

2. The Hot Universe

As we wind the clock back towards the Big Bang, the energy density in the universe increases. As this happens, particles interact more and more frequently, and the state of the universe is well approximated by a hot fluid in equilibrium. This is sometimes referred to as the *primeval fireball* of the Big Bang. The purpose of this section is to introduce a few basic properties of this fireball.

It is worth sketching the big picture. First we play the movie in reverse. As we go back in time, the Universe becomes hotter and hotter and things fall apart. Running the movie forward, the Universe cools and various objects form.

For example, there is an important event, roughly 300,000 years after the Big Bang, when atoms form for the first time. Prior to this, the temperature was higher than the 13.6 eV binding energy of hydrogen, and the electrons were stripped from the protons. This moment in time is known as *recombination* and will be described in Section 2.3. (Obviously a better name would simply be “combination” since the electrons and protons combined for the first time, but we don’t get to decide these things). This is a key moment in the history of the universe. Prior to this time, space was filled with a charged plasma through which light is unable to propagate. But when the electrons and protons form to make (mostly) neutral hydrogen, the universe becomes transparent. The cosmic microwave background, which will be discussed in Section 2.2, dates from this time.

At yet earlier times, the universe was so hot that nuclei fail to cling together and they fall apart into their constituent protons and neutrons. This process – which, running forwards in time is known as nucleosynthesis – happens around 3 minutes after the Big Bang and is understood in exquisite detail. We will describe some of the basic reactions in Section 2.5.3.

As we continue to trace the clock further back, the universe is heated to extraordinary temperatures, corresponding to the energies probed in particle accelerators and beyond. Taking knowledge from particle physics, even here we have a good idea of what happens. At some point, known as the QCD phase transition, protons and neutrons melt, dissolving into a soup of their constituents known as the quark-gluon plasma. Earlier still, at the electroweak phase transition, the condensate of the Higgs boson melts. Beyond this, we have little clear knowledge but there are still other events that we know must occur. The purpose of this chapter is to tell this story.

2.1 Some Statistical Mechanics

Our first task is to build a language that allows us to describe stuff that is hot. We will cherry pick a few key results that we need. A much fuller discussion of the subject can be found in the lectures on [Statistical Physics](#) and the lectures on [Kinetic Theory](#).

Ideas such as heat and temperature are not part of the fundamental laws of physics. There is no such thing, for example, as the temperature of a single electron. Instead, these are examples of *emergent phenomena*, concepts which arise only when a sufficiently large number of particles are thrown together. In domestic situations, where we usually apply these ideas, large means $N \sim 10^{23}$ particles. As we will see, in the cosmological setting N can be substantially larger.

When dealing with such a large number of particles, we need to shift our point of view. The kinds of things that we usually discuss in classical physics, such as the position and momentum of each individual particle, no longer hold any interest. Instead, we want to know coarse-grained properties of the system. For example, we might like to know the probability that a particle chosen at random has a momentum \mathbf{p} . In what follows, we call this probability distribution $f(\mathbf{p}; t)$.

Equilibrium

In general, the distribution $f(\mathbf{p}, t)$ will be very complicated. But patience brings rewards. If we wait a suitably long time, the individual particles will collide with each other, transferring energy and momentum among themselves until, eventually, any knowledge about the initial conditions is effectively lost. The resulting state is known as *equilibrium* and is described by a time-independent probability distribution $f(\mathbf{p})$. In equilibrium, the constituent particles are flying around in random directions. But, if you focus only on the coarse-grained probability distribution, everything appears calm.

Equilibrium states are characterised by a number of macroscopic quantities. These will be dealt with in detail in the [Statistical Physics](#) course, but here we summarise some key facts.

The most important characteristic of an equilibrium state is *temperature*. This is related to the average energy of the state in a way that we will make precise below. The reason that temperature plays such an important role is due to the following property: suppose that we have two different systems, each individually in equilibrium, one at temperature T_1 and the other at temperature T_2 . We then bring the two systems together and allow them to exchange energy. If $T_1 = T_2$, then the two systems remain unaffected by this, and the combined system is in equilibrium. In contrast, if $T_1 \neq$

T_2 , then a net energy will flow from the hotter system to the colder system, and the combined system will eventually settle down to a new equilibrium state at some intermediate temperature. Two systems which have the same temperature are said to be in *thermal equilibrium*.

Other kinds of equilibria are also possible. One that we will meet later in this section arises when two systems are able to exchange particles. Often we will be interested in this when one type of particle can transmute into another. In this case, we characterise the system by another quantity known as the *chemical potential*. (The name comes from chemical reactions although, in this course, will be more interested in processes in atomic or particle physics.) The chemical potential has the property that if two systems have the same value then, when brought together, there will not be a net transfer of particles from one system to the other. In this case, the systems are said to be in *chemical equilibrium*.

2.1.1 The Boltzmann Distribution

For now we will focus on states in thermal equilibrium. The thermal properties of a state are closely related to its energy which, in turn, is related to the momentum of the constituent particles. This means that understanding thermal equilibrium is akin to understanding the momentum distribution $f(\mathbf{p})$ of particles. We will see a number of examples of this in what follows.

A microscopic understanding of thermal equilibrium was first provided by Boltzmann. It turns out that the result is somewhat easier to state in the language of quantum mechanics, although it also applies to the classical world. Consider a system with discrete energy eigenstates $|n\rangle$, each with energy E_n . In *thermal equilibrium* at temperature T , the probability that the system sits in the state $|n\rangle$ is given by the *Boltzmann distribution*,

$$p(n) = \frac{e^{-E_n/k_B T}}{Z} \quad (2.1)$$

Here k_B is the *Boltzmann constant*, defined to be

$$k_B \approx 1.381 \times 10^{-23} \text{ J K}^{-1}$$

This fundamental constant provides a translation between temperatures and energies. Meanwhile Z is simply a normalisation constant designed to ensure that

$$\sum_n p(n) = 1 \quad \Rightarrow \quad Z = \sum_n e^{-E_n/k_B T}$$

This normalisation factor Z has its own name: it is called the *partition function* and it plays a starring role in most treatments of statistical mechanics. For our purposes, it will suffice to keep Z firmly in the background.

It is possible to derive the Boltzmann distribution from more elementary principles. (Such a derivation can be found in the lectures on [Statistical Physics](#).) Here, we will simply take the distribution (2.1) to be the definition of both thermal equilibrium and the temperature.

The Boltzmann distribution gives us some simple intuition for the meaning of thermal equilibrium. We see that the any state with $E_n \ll k_B T$ has a more or less equal chance of being occupied, while any state with $E_n \gg k_B T$ has a vanishingly small chance of being occupied. In this way $k_B T$ sets the characteristic energy scale of the system.

We'll see many variations of the Boltzmann distribution in what follows. It gets tedious to keep writing $1/k_B T$. For this reason we define

$$\beta = \frac{1}{k_B T}$$

We will be careless in what follows and also refer to β as “temperature”: obviously it is actually (proportional to) the inverse temperature. The Boltzmann distribution then reads

$$p(n) = \frac{e^{-\beta E_n}}{Z}$$

Above, we mentioned the key property of temperature: it determines whether two systems sit in thermal equilibrium. We should check that this is indeed obeyed by the Boltzmann distribution. Suppose that we have two systems, A and B , both at the same temperature β , but with different microscopic constituents, meaning that their energy levels differ. If we bring the two systems together, we expect that the combined system also sits in a Boltzmann distribution at temperature β . Happily, this is indeed the case. To see this note that we have independent probability distributions for A and B , so the combined probability distribution is given by

$$p(n, m) = \frac{e^{-\beta E_n^A}}{Z_A} \frac{e^{-\beta E_m^B}}{Z_B} = \frac{e^{-\beta(E_n^A + E_m^B)}}{Z_A Z_B}$$

But this is again of the Boltzmann form. The denominator $Z_A Z_B$ can be written as

$$Z_A Z_B = \left(\sum_n e^{-\beta E_n^A} \right) \left(\sum_m e^{-\beta E_m^B} \right) = \sum_{n,m} e^{-\beta E_n^A} e^{-\beta E_m^B} = \sum_{n,m} e^{-\beta(E_n^A + E_m^B)}$$

where we recognise this final expression as Z_{A+B} , the partition function of the combined system $A + B$. This had to be the case to ensure that the joint probability distribution $p(n, m)$ is correctly normalised.

It's worth re-iterating what we have learned. You might think that if we combined two systems, separately in equilibrium, then there would be no energy transfer from one to the other if the average energies coincide, i.e. $\langle E^A \rangle = \langle E^B \rangle$, with

$$\langle E \rangle = \frac{1}{Z} \sum_n E_n e^{-\beta E_n}$$

However, this is not the right criterion. As we have seen above, the average energies of the two systems can be very different. It is the temperatures that must coincide.

2.1.2 The Ideal Gas

As our first application of the Boltzmann distribution, consider a gas of non-relativistic particles, each of mass m . We will assume that there are no interactions between these particles, so the energy of each is given by

$$E = \frac{1}{2}mv^2 \tag{2.2}$$

Before we proceed, I should mention a subtlety. We've turned off interactions in order to make our life simpler. Yet, from our earlier discussion, it should be clear that interactions are crucial if we are ever going to reach equilibrium, since this requires a large number of collisions to share energy and momentum between particles! This is one of many annoying and fiddly issues that plague the fundamentals of statistical mechanics. We will argue this subtlety away by pretending that the interactions are strong enough to drive the system to equilibrium, but small enough to ignore when describing equilibrium. Obviously this is unsatisfactory. We can do better, but it is more work. (See, for example, the discussion of the interacting gas in the lectures on [Statistical Physics](#) or the derivation of the approach to equilibrium in the lectures on [Kinetic Theory](#).) We will also see this issue rear its head in a physical context in Section 2.3.4 when we discuss the phenomenon of decoupling in the early universe.

We consider a gas of particles. We'll assume that each particle is independent of the others, and focus on the state of a just single particle, specified by the momentum \mathbf{p} or, equivalently, the velocity $\mathbf{v} = \mathbf{p}/m$. If the momentum is continuous (or finely spaced) we should talk about the probability that the velocity lies in some some volume d^3v centred around \mathbf{v} . We denote the probability distribution as $f(\mathbf{v}) d^3v$. The Boltzmann

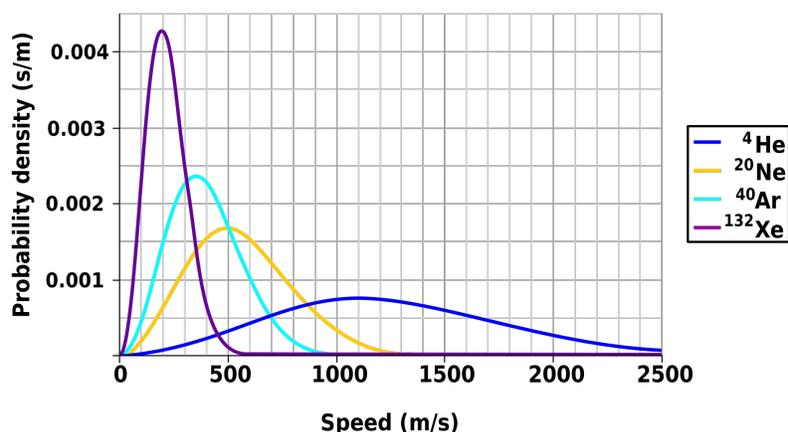


Figure 27: The distribution of the speeds of various molecules at $T = 25$ C. (Image taken from Wikipedia.)

distribution (2.1) tells us that this is

$$f(\mathbf{v}) d^3v = \frac{e^{-\beta m v^2/2}}{Z} d^3v \quad (2.3)$$

where Z is a normalisation factor that we will determine shortly.

Our real interest lies in the speed $v = |\mathbf{v}|$. The corresponding speed distribution $f(v) dv = f(\mathbf{v}) d^3v$ is

$$f(v) dv = \frac{4\pi v^2}{Z} e^{-\beta m v^2/2} dv \quad (2.4)$$

Note that we have an extra factor of $4\pi v^2$ when considering the probability distribution over speeds v , as opposed to velocities \mathbf{v} . This reflects the fact that there’s “more ways” to have a high velocity than a low velocity: the factor of $4\pi v^2$ is the area of the sphere swept out by a velocity vector \mathbf{v} .

We require that

$$\int_0^\infty dv f(v) = 1 \quad \Rightarrow \quad Z = \left(\frac{2\pi k_B T}{m} \right)^{3/2}$$

Finally, we find the probability that the particle has speed between v and $v + dv$ to be

$$f(v) dv = 4\pi v^2 \left(\frac{m}{2\pi k_B T} \right)^{3/2} e^{-m v^2/2k_B T} dv \quad (2.5)$$

This is known as the *Maxwell-Boltzmann distribution*. It tells us the distribution of the speeds of gas molecules in this room.

Pressure and the Equation of State

We can use the Maxwell-Boltzmann distribution to compute the pressure of a gas. The pressure arises from the constant bombardment by the underlying atoms and can be calculated with some basic physics. Consider a wall of area A that lies in the (y, z) -plane. Let n denote the density of particles (i.e. $n = N/V$ where N is the number of particles and V the volume). In some short time interval Δt , the following happens:

- A particle with velocity \mathbf{v} will hit the wall if it lies within a distance $\Delta L = |v_x|\Delta t$ of the wall *and* if it's travelling towards the wall, rather than away. The number of such particles with velocity centred around \mathbf{v} is

$$\frac{1}{2}nA|v_x|\Delta t d^3v$$

with a factor of $1/2$ picking out only those particles that travel in the right direction.

- After each such collision, the momentum of the particle changes from p_x to $-p_x$, with p_y and p_z left unchanged. As before, this holds only for the initial $p_x > 0$. We therefore write the impulse imparted by each particle as $2|p_x|$.
- This impulse is equated with $F_x\Delta t$ where F_x is the force on the wall. The force arising from particles with velocity in the region d^3v about \mathbf{v} is

$$F_x\Delta t = \left(\frac{1}{2}nA|v_x|\Delta t d^3v\right) \times 2|p_x| \quad \Rightarrow \quad F_x = nAv_x p_x d^3v$$

where we dropped the modulus signs on the grounds that the sign of the momentum p_x is the same as the sign of the velocity v_x .

- The pressure on the wall is the force per unit area, $P = F_x/A$. We learn that the pressure from those particles with velocity in the region of \mathbf{v} is

$$P = nv_x p_x d^3v$$

At this stage we invoke isotropy of the gas, which means that $\mathbf{v} \cdot \mathbf{p} = v_x p_x + v_y p_y + v_z p_z = 3v_x p_x$. We therefore have

$$P = \frac{n}{3} \mathbf{v} \cdot \mathbf{p} d^3v \tag{2.6}$$

The last stage is to integrate over all velocities, weighted with the probability distribution. In the final form (2.6), the pressure is related to the speed v rather than the

(component of the) velocity v_x . This means that we can use the Maxwell-Boltzmann distribution over speeds (2.5) and write

$$P = \frac{1}{3} \int dv n \mathbf{v} \cdot \mathbf{p} f(v) \quad (2.7)$$

This coincides with our earlier result (1.33) (albeit using slightly different notation for the probability distributions).

The expression (2.7) holds for both relativistic and non-relativistic systems, a fact that we will make use of later. For now, we care only for the non-relativistic case with $\mathbf{p} = m\mathbf{v}$. Here we have

$$P = \frac{4\pi n}{3} \left(\frac{m}{2\pi k_B T} \right)^{3/2} \int dv mv^4 e^{-mv^2/2k_B T}$$

The integral is straightforward: it is given by

$$\int_0^\infty dx x^4 e^{-ax^2} = \frac{3}{8} \sqrt{\frac{\pi}{a^5}}$$

Using this, we find a familiar friend

$$P = nk_B T$$

This is the equation of state for an ideal gas.

We can also calculate the average kinetic energy. If the gas contains N particles, the total energy is

$$\langle E \rangle = \frac{N}{2} m \langle v^2 \rangle = N \int_0^\infty dv \frac{1}{2} m v^2 f(v) = \frac{3}{2} N k_B T \quad (2.8)$$

This confirms the result (1.37) that we met when we first introduced non-relativistic fluids.

2.2 The Cosmic Microwave Background

The universe is bathed in a sea of thermal radiation, known as the *cosmic microwave background*, or the CMB. This was the first piece of evidence for the hot Big Bang – the idea that the early universe was filled with a fireball – and remains one of the most compelling. In this section, we describe some of the basic properties of this radiation.

2.2.1 Blackbody Radiation

To start, we want to derive the properties of a thermal gas of photons. Such a gas is known, unhelpfully, as *blackbody radiation*.

The state of a single photon is specified by its momentum $\mathbf{p} = \hbar\mathbf{k}$, with \mathbf{k} the *wavevector*. The energy of the photon is given by

$$E = pc = \hbar\omega$$

where $\omega = ck$ is the (angular) frequency of the photon.

Blackbody radiation comes with a new conceptual ingredient, because the number of photons is not a conserved quantity. This means that when considering the possible states of the gas, we should include states with an arbitrary number of photons. We do this by stating how many photons $N(\mathbf{p})$ sit in the state \mathbf{p} .

In thermal equilibrium, we will not have a definite number of photons $N(\mathbf{p})$, but rather some probability distribution over the number of photons, Focussing on a fixed state $\mathbf{p} = \hbar\mathbf{k}$, the average number of particles is dictated by the Boltzmann distribution

$$\langle N(\mathbf{p}) \rangle = \frac{1}{Z} \sum_{n=0}^{\infty} n e^{-\beta n \hbar \omega} \quad \text{with } Z = \sum_{n=0}^{\infty} e^{-\beta n \hbar \omega}$$

We can easily do both of these sums. Defining $x = e^{-\beta \hbar \omega}$, the partition function is given by

$$Z = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

Meanwhile the numerator of $\langle N(\mathbf{p}) \rangle$ takes the form

$$\sum_{n=0}^{\infty} n x^n = x \sum_{n=0}^{\infty} n x^{n-1} = x \frac{dZ}{dx} = \frac{x}{(1-x)^2}$$

We learn that the average number of particles with momentum \mathbf{p} is

$$\langle N(\mathbf{p}) \rangle = \frac{1}{e^{\beta \hbar \omega} - 1} \tag{2.9}$$

For $k_B T \ll \hbar \omega$, the number of photons is exponentially small. In contrast, when $k_B T \gg \hbar \omega$, the number of photons grows linearly as $\langle N(\mathbf{p}) \rangle \approx k_B T / \hbar \omega$.

Density of States

Our next task is to determine the average number of photons $\langle N(\omega) \rangle$ with given energy $\hbar \omega$. To do this, we must count the number of states \mathbf{p} which have energy $\hbar \omega$.

It's easier to count objects that are discrete rather than continuous. For this reason, we'll put our system in a square box with sides of length L . At the end of the calculation, we can happily send $L \rightarrow \infty$. In such a box, the wavevector is quantised: it takes values

$$k_i = \frac{2\pi q_i}{L} \quad q_i \in \mathbf{Z}$$

This is true for both a classical wave or a quantum particle; in both cases, an integer number of wavelengths must fit in the box.

Different states are labelled by the integers q_i . When counting, or summing over such states, we should therefore sum over the q_i . However, for very large boxes, so that L is much bigger than any other length scale in the game, we can approximate this sum by an integral,

$$\sum_q \approx \frac{L^3}{(2\pi)^3} \int d^3k = \frac{4\pi V}{(2\pi)^3} \int_0^\infty dk k^2 \quad (2.10)$$

where $V = L^3$ is the volume of the box. The formula above counts all states. But the final form has a simple interpretation: the number of states with the magnitude of the wavevector between k and $k + dk$ is $4\pi V k^2 / (2\pi)^3$. Note that the $4\pi k^2$ term is reminiscent of the $4\pi v^2$ term that appeared in the Maxwell-Boltzmann distribution; both have the same origin.

We would like to compute the number of states with frequency between ω and $\omega + d\omega$. For this, we simply use

$$\omega = ck \quad \Rightarrow \quad \frac{4\pi V}{(2\pi)^3} \int dk k^2 = \frac{4\pi V}{(2\pi c)^3} \int d\omega \omega^2$$

This tells us that the number of states with frequency between ω and $\omega + d\omega$ is $4\pi V \omega^2 / (2\pi c)^3$.

There is one final fact that we need. Photons come with two polarisation states. This means that the total number of states is twice the number above. We can now combine this with our earlier result (2.9). In thermal equilibrium, the average number of photons with frequency between ω and $\omega + d\omega$ is

$$\langle N(\omega) \rangle d\omega = 2 \times \frac{4\pi V}{(2\pi c)^3} \frac{\omega^2}{e^{\beta\hbar\omega} - 1} d\omega$$

We usually write this in terms of the number density $n = N/V$. Moreover, we will be a little lazy and drop the expectation value $\langle n \rangle$ signs. The distribution of photons in a

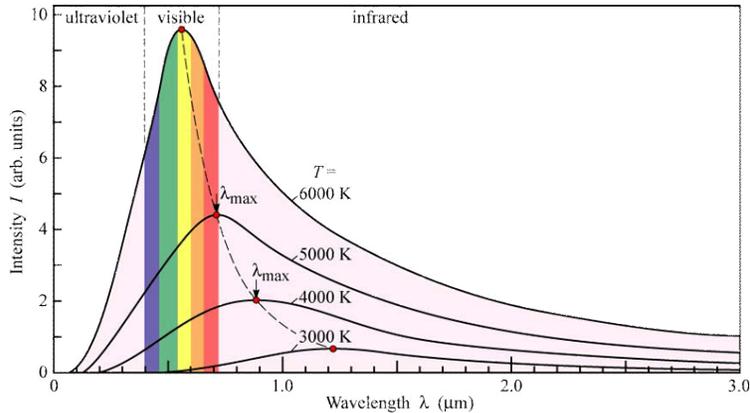


Figure 28: The distribution of colours at various temperatures.

thermal bath is then written as

$$n(\omega) = \frac{1}{\pi^2 c^3} \frac{\omega^2}{e^{\beta \hbar \omega} - 1} \quad (2.11)$$

This is the *Planck blackbody distribution*. For a fixed temperature, $\beta = 1/k_B T$, the distribution tells us how many photons of a given frequency – and hence, of a given colour – are present. The distribution peaks in visible light for temperatures around 6000 K, which is the temperature of the surface of the Sun. (Presumably the Sun evolved to be at exactly the right temperature so that our eyes can see it. Or something.)

The Equation of State

We now have all the information that we need to compute the equation of state. First the energy density. This is straightforward: we just need to integrate

$$\rho = \int_0^\infty d\omega \hbar \omega n(\omega) \quad (2.12)$$

Next the pressure. We can import our previous formula (2.7), now with $\mathbf{v} \cdot \mathbf{p} = \hbar c k = \hbar \omega$. But this gives precisely the same integral as the energy density; it differs only by the overall factor of 1/3,

$$P = \frac{1}{3} \rho$$

This, of course, is the relativistic equation of state that we used when describing the expanding universe.

Finally, we can actually do the integral (2.12). In fact, there's a couple of quantities of interest. The energy density is

$$\rho = \frac{\hbar}{\pi^2 c^3} \int_0^\infty d\omega \frac{\omega^3}{e^{\beta\hbar\omega} - 1} = \frac{(k_B T)^4}{\pi^2 \hbar^3 c^3} \int_0^\infty dy \frac{y^3}{e^y - 1}$$

Meanwhile, the total number density is

$$n = \int_0^\infty d\omega n(\omega) = \frac{1}{\pi^2 c^3} \int_0^\infty d\omega \frac{\omega^2}{e^{\beta\hbar\omega} - 1} = \frac{(k_B T)^3}{\pi^2 \hbar^3 c^3} \int_0^\infty dy \frac{y^2}{e^y - 1}$$

Both of these integrals take a similar form. Here we just quote the general result without proof:

$$I_n = \int_0^\infty dy \frac{y^n}{e^y - 1} = \Gamma(n+1)\zeta(n+1) \quad (2.13)$$

The Gamma function is the analytic continuation of the factorial function to the real numbers; when evaluated on the integers it gives $\Gamma(n+1) = n!$. Meanwhile, the Riemann zeta function is defined, for $\text{Re}(s) > 1$, as $\zeta(s) = \sum_{q=1}^\infty q^{-s}$. It turns out that $\zeta(4) = \pi^4/90$, giving us $I_3 = \pi^4/15$. In contrast, there is no such simple expression for $\zeta(3) \approx 1.20$. It is sometimes referred to as *Apéry's constant*. A derivation of (2.13) can be found in Section 3.5.3 of the lectures on [Statistical Physics](#).

We learn that the energy density is

$$\rho = \frac{\pi^2}{15\hbar^3 c^3} (k_B T)^4 \quad (2.14)$$

Meanwhile, the total number density is

$$n = \frac{2\zeta(3)}{\pi^2 \hbar^3 c^3} (k_B T)^3 \quad (2.15)$$

Notice, in particular, that the number density of photons varies with the temperature. This will be important in what follows.

2.2.2 The CMB Today

The universe today is filled with a sea of photons, the cosmic microwave background. This is the afterglow of the fireball that filled the universe in its earliest moments. The frequency spectrum of the photons is a perfect fit to the blackbody spectrum, with at a temperature

$$T_{\text{CMB}} = 2.726 \pm 0.0006 \text{ K} \quad (2.16)$$

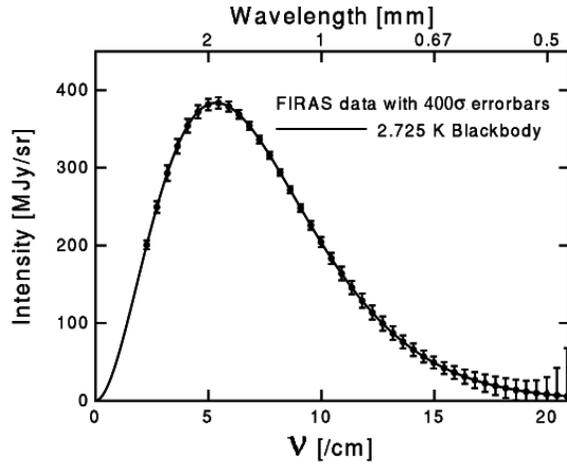


Figure 29: The blackbody spectrum of the CMB, measured in 1990 by the FIRAS detector on the COBE satellite. The error bars have been enlarged by a factor of 400 just to help you see them.

This spectrum is shown in Figure 29. There are small, local deviations in this temperature at the level of

$$\frac{\Delta T}{T_{\text{CMB}}} \sim 10^{-5}$$

These fluctuations will be discussed further in Section 3.4.

From the temperature (2.16), we can determine the energy density and number density in photons. From (2.14), the energy density is given by

$$\rho_\gamma \approx 4.3 \times 10^{-14} \text{ kg m}^{-3}$$

We can compare this to the critical energy density (1.71), $\rho_{\text{crit},0} = 8.5 \times 10^{-10} \text{ kg m}^{-3}$ to find

$$\Omega_\gamma = \frac{\rho_\gamma}{\rho_{\text{crit},0}} \approx 5 \times 10^{-5}$$

This is the value (1.69) that we quoted previously. There are, of course, further photons in starlight, but they are dwarfed in both energy and number by the CMB.

From (2.15), the number density of CMB photons is

$$n_\gamma = 4 \times 10^8 \text{ m}^{-3} = 400 \text{ cm}^{-3}$$

We can compare this to the number of baryons (i.e. protons and neutrons). The density of baryons is (1.74) $\Omega_B \approx 0.05$, so the total mass in baryons is

$$\rho_B \approx \Omega_B \rho_{\text{crit},0} \approx 4 \times 10^{-11} \text{ kg m}^{-3}$$

The mass of the proton and neutron are roughly the same, at $m_p \approx 1.7 \times 10^{-27} \text{ kg}$. This places the number density of baryons as

$$n_B = \frac{\rho_B}{m_p} \approx 0.3 \text{ m}^{-3}$$

We see that there are many more photons in the universe than baryons: the ratio is

$$\eta \equiv \frac{n_B}{n_\gamma} \approx 10^{-9} \quad (2.17)$$

This is one of the fundamental numbers in cosmology. As we will see, this ratio has been pretty much constant since the first second or so after the Big Bang and plays a crucial role in both nucleosynthesis (the formation of heavier nuclei) and in recombination (the formation of atoms). We do not, currently, have a good theoretical understanding of where this number fundamentally comes from: it is something that we can only derive from observation.

The CMB is a Relic

There is an important twist to the story above. We have computed the expected distribution of photons in thermal equilibrium, and found that it matches perfectly with the spectrum of the cosmic microwave background. The twist is that the CMB is *not* in equilibrium!

Recall that equilibrium is a property that arises when particles are constantly interacting. Yet the CMB photons have barely spoken to anyone for the past 13 billion years. The occasional photon may bump into a planet, or an infra-red detector fitted to a satellite, but most just wend their merry way through the universe, uninterrupted.

How then did the CMB photons come to form a perfect equilibrium spectrum? The answer is that this dates from a time when the photons were interacting frequently with matter. Fluids like this, that have long since fallen out of thermal equilibrium, but nonetheless retain their thermal character, are called *relics*.

There are a couple of questions that we would like to address. The first is: when were the photons last interacting and, hence, last genuinely in equilibrium? This is called the time of *last scattering*, t_{last} and we will compute it in Section 2.3 below. The second question is: what happened to the distribution of photons subsequently?

We start by answering the second of these questions. Once the photons no longer interact, they are essentially free particles. As the universe expands, each photon is redshifted as explained in Section 1.1.3. This means that the wavelength is stretched and, correspondingly, the frequency is decreased as the universe expands.

$$\lambda(t) = \lambda_{\text{last}} \frac{a(t)}{a(t_{\text{last}})} \quad \Rightarrow \quad \omega(t) = \omega_{\text{last}} \frac{a(t_{\text{last}})}{a(t)} \quad (2.18)$$

At the same time, the number of photons is diluted by a factor of $(a(t_{\text{last}})/a(t))^3$ as the universe expands. Putting these two effects together, an initial blackbody distribution (2.11) will, if left alone, evolve as

$$n(\omega_{\text{last}}; T_{\text{last}}, t) d\omega_{\text{last}} = \frac{1}{\pi^2 c^3} \left(\frac{a(t_{\text{last}})}{a(t)} \right)^3 \frac{\omega_{\text{last}}^2}{e^{\beta \hbar \omega_{\text{last}}} - 1} d\omega_{\text{last}}$$

The $1/a^3$ dilution factor is absorbed into the frequency in the ω^2 and $d\omega$ terms. But not in the exponent. However, the resulting distribution can be put back into blackbody form if we think of the temperature as time dependent

$$n(\omega; T, t) d\omega = \frac{1}{\pi^2 c^3} \frac{\omega(t)^2}{e^{\beta(t) \hbar \omega(t)} - 1} d\omega(t)$$

where the $\beta(t) = 1/k_B T(t)$, with the time varying temperature

$$T(t) = T_{\text{last}} \frac{a(t_{\text{last}})}{a(t)} \quad (2.19)$$

We see that, left alone, a blackbody distribution will keep the same overall form, but with the temperature scaling as $T \sim 1/a$.

This means that, if we can figure out the temperature T_{last} when the photons were last in equilibrium, then we can immediately determine the redshift at which this occurred $1 + z_{\text{last}} = a(t_{\text{last}})^{-1}$. We'll compute both of these in Section 2.3.

2.2.3 The Discovery of the CMB

In 1964, two radio astronomers, Arno Penzias and Robert Wilson, got a new toy. The microwave horn antenna was originally used by their employers, the Bell telephone company, for satellite communication. Now Penzias and Wilson hoped to do some science with it, measuring the radio noise emitted in the direction away from the plane of the galaxy.

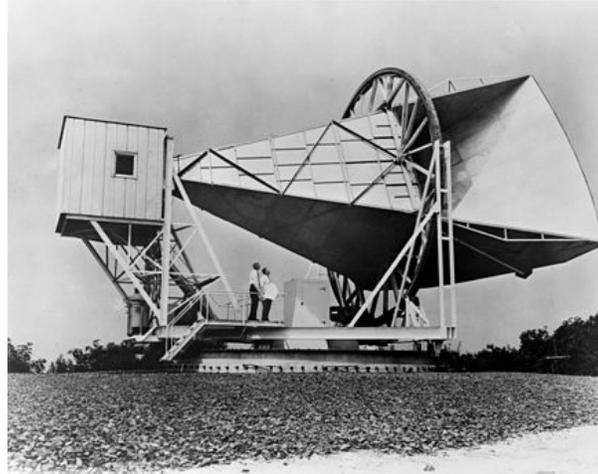


Figure 30: The Holmdel radio antenna at Bell Telephone Laboratories.

To their surprise, they found a background noise which did not depend on the direction in which they pointed their antenna. Nor did it depend on the time of day or the time of a year. Taken seriously, this suggested that the noise was a message from the wider universe.

There was, however, an alternative, more mundane explanation. Maybe the noise was coming from the antenna itself, some undiscovered systematic effect that they had failed to take into account. Indeed, they soon found a putative source of the noise: a pair of pigeons had taken roost and deposited what Penzias called “a white dielectric material” over much of the antenna. They removed this material (and shot the pigeons), but the noise remained. What Penzias and Wilson had on their hands was not pigeon shit, but one of the great discoveries of the twentieth century: the afterglow of the Big Bang itself, with a temperature that they measured to lie between 2.5 K and 4.5 K. In 1965 they published their result with the attention-grabbing title: “A Measurement of Excess Antenna Temperature at 4080 Mc/s”.

Penzias and Wilson were not unaware of the significance of their finding. In the year since they first found the noise, they had done what good scientists should always do: they talked to their friends. They were soon put in touch with the group in nearby Princeton where Jim Peebles, a theoretical cosmologist, had recently predicted a background radiation with a temperature of a few degrees, based on the idea of nucleosynthesis in the very early universe (an idea we will describe in Section 2.5.3). Meanwhile, three experimental colleagues, Dicke, Roll and Wilkinson had cobbled together a small antenna in the hope of searching for this radiation. These four scientists

wrote a companion paper, outlining the importance of the discovery. In 1978, Penzias and Wilson were awarded the Nobel prize. It took another 39 years before Peebles gained the same recognition.

In fact there had been earlier predictions of the CMB. In the 1940s, Gammow together with Alpher and Herman suggested that the early universe began only with neutrons and, through somewhat dodgy calculations, concluded that there should be a background radiation at 5 K. Later other scientists, including Zel'dovich in the Soviet Union, and Hoyle and Taylor in England, used nucleosynthesis to predict the existence of the CMB at a few degrees. Yet none of these results were taken sufficiently seriously to search for the signal before Penzias and Wilson made their serendipitous discovery.

Detecting the CMB was just the beginning of the story. The radiation is not, it turns out, perfectly uniform but contains small anisotropies. These contain precious information about the make-up of the universe when it was much younger. A number of theorists, including Harrison, Zel'dovich, and Peebles and Yu, predicted that these anisotropies could be observed at a level of 10^{-4} to 10^{-5} . These were finally detected by the NASA COBE satellite in the early 1990s. Since then a number of ground based telescopes, including BOOMERanG and MAXIMA, and a two full sky maps from the satellites WMAP and Planck, have mapped out the CMB in exquisite detail. We will describe these anisotropies in Section 3.

2.3 Recombination

We've learned that the CMB is a relic, with its perfect blackbody spectrum a remnant of an earlier, more intense time in the universe, when the photons were in equilibrium with matter. We would like to gain a better understanding of this time.

Photons interact with electric charge. Nowadays, the vast majority of matter in the universe is in the form of neutral atoms, and electrons interact only with the charged constituents of the atoms. Such interactions are relatively weak. However, there was a time in the early universe when the temperature was so great that electrons and protons could no longer bind into neutral atoms. Instead, the universe was filled with a plasma. In this era, the matter and photons interacted strongly and were in equilibrium.

The CMB that we see today dates from this time. Or, more precisely, from the time when electrons and protons first bound themselves into neutral hydrogen, emitting a photon in the process



The moment at which this occurs is called *recombination*. As the arrows illustrate, this process can happen in both directions.

Interactions like (2.20) involve one particle type transmuted into a different type. This means that the number of, say, hydrogen atoms is not fixed but fluctuating. We need to introduce a new concept that allows us to deal with such situations. This concept is the *chemical potential*.

2.3.1 The Chemical Potential

The chemical potential offers a slight generalisation of the Boltzmann distribution which is useful in situations where the number of particles in a system is not fixed. It was, as the name suggests, originally introduced to describe chemical reactions but we will re-purpose it to describe atomic reactions like (2.20) (and, later, nuclear reactions).

Although our ultimate goal is to describe atomic reactions, we can first introduce the chemical potential in a more mundane setting. Suppose that we have a fixed number of atoms N in a box of size V . If we focus attention on some large, fixed sub-volume $V' \subset V$, then we would expect the gas in V' to share the same macroscopic properties, such as temperature and pressure, as the whole gas in V . But particles can happily fly in and out of V' and the total number in this region is not fixed. Instead, there is some probability distribution which has the property that the average number density coincides with N/V .

In this situation, it's clear that we should consider states of all possible particle number in V' . There is a possibility, albeit a very small one, that V' contains no particles at all. There is also a small possibility that it contains all the particles.

If we work in the language of quantum mechanics, each state $|n\rangle$ in the system can be assigned both an energy E_n and a particle number N_n . Correspondingly, equilibrium states are characterised by two macroscopic properties: the temperature T and the chemical potential μ . These are defined through the generalised Boltzmann distribution

$$p(n) = \frac{e^{-\beta(E_n - \mu N_n)}}{\mathcal{Z}} \quad (2.21)$$

where $\mathcal{Z} = \sum_n e^{-\beta(E_n - \mu N_n)}$ is again the appropriate normalisation factor. In the language of statistical mechanics, this is referred to as the *grand canonical ensemble*.

Clearly, the distribution has the same exponential form as the Boltzmann distribution. This is important. We learned in Section 2.1.1 that two isolated systems which sit at the same temperature will remain in thermal equilibrium when brought together,

meaning that there will be no transfer of energy from one system to the other. Exactly the same argument tells us that if two isolated systems have the same chemical potential then, when brought together, there will be no net flux of particles from one system to the other. In this case, we say that the systems are in *chemical equilibrium*.

Notice that the requirement for equilibrium is *not* that the number densities of the systems are equal: it is the chemical potentials that must be equal. This is entirely analogous to the statement that it is temperature, rather than energy density, that determines whether systems are in thermal equilibrium.

We'll see examples of how to wield the chemical below, but before we do it's worth mentioning a few issues.

- In general, we can introduce a different chemical potential for every conserved quantity in the system. This is because conserved quantities commute with the Hamiltonian, and so it makes sense to label microscopic states by both the energy and a further quantum number. One familiar example is electric charge Q . Here, the corresponding chemical potential is voltage.

This leads to an almost-contradictory pair of statements. First, we can only introduce a chemical potential for any conserved quantity. Second, the purpose of the chemical potential is to allow this conserved quantity to fluctuate! If you're confused about this, then think back to the volume $V' \subset V$, or to the meaning of voltage in electromagnetism, both of which give examples where these statements hold.

- The story above is very similar to our derivation of the Planck blackbody distribution for photons. There too we labeled states by both energy and particle number, but we didn't introduce a chemical potential. What's different now? This is actually a rather subtle issue. Ultimately it is related to the fact that we ignore interactions while simultaneously pretending that they are crucial to reach equilibrium. As soon as we take these interactions into account, the number of photons is not conserved so we can't label states by both energy and photon number. This is what prohibits us from introducing a chemical potential for photons. In contrast, we can introduce a chemical potential in situations where particle number (or some other quantity) is conserved even in the presence of interactions.

2.3.2 Non-Relativistic Gases Revisited

For our first application of the chemical potential, we're going to re-derive the ideal gas equation. At first sight, this will appear to be only a more complicated derivation

of something we've seen already. The pay-off will come only in Section 2.3.3 where we will understand recombination and the atomic reaction (2.20).

We consider non-relativistic particles, with energy

$$E_{\mathbf{p}} = \frac{p^2}{2m}$$

As with our calculation of photons, we now consider states that have arbitrary numbers of particles. We choose to specify these states by stating how many particles $n_{\mathbf{p}}$ have momentum \mathbf{p} . For each choice of momentum, the number of particles⁷ can be $n_{\mathbf{p}} = 0, 1, 2, \dots$. The generalised Boltzmann distribution (2.21) then tells us that the average number of particles with momentum \mathbf{p} is

$$\langle N(\mathbf{p}) \rangle = \frac{1}{\mathcal{Z}_{\mathbf{p}}} \sum_{n_{\mathbf{p}}=0}^{\infty} n_{\mathbf{p}} e^{-\beta(n_{\mathbf{p}}E_{\mathbf{p}} - \mu n_{\mathbf{p}})}$$

where the normalisation factor (or, in fancy language, the grand canonical partition function) is given by the geometric series

$$\mathcal{Z}_{\mathbf{p}} = \sum_{n=0}^{\infty} e^{-\beta n_{\mathbf{p}}(E_{\mathbf{p}} - \mu)} = \frac{e^{\beta(E_{\mathbf{p}} - \mu)}}{e^{\beta(E_{\mathbf{p}} - \mu)} - 1}$$

This is exactly the same calculation as we saw for photons in Section 2.2.1, but with the additional minor complication of a chemical potential. Note that computing $\mathcal{Z}_{\mathbf{p}}$ allows us to immediately determine the expected number of particles since we can write

$$\langle N(\mathbf{p}) \rangle = \frac{1}{\beta} \frac{\partial}{\partial \mu} \log \mathcal{Z}_{\mathbf{p}} = \frac{1}{e^{\beta(E_{\mathbf{p}} - \mu)} - 1} \quad (2.22)$$

This is known as the *Bose-Einstein distribution* and will be discussed further in Section 2.4.

To compute the average total number of particles, we simply need to integrate over all momenta \mathbf{p} . We must include the density of states, but this is identical to the calculation we did for photons, with the result (2.10). The total average number of particles is then

$$N = \frac{V}{(2\pi\hbar)^3} \int d^3p N(\mathbf{p})$$

⁷Actually, there is a subtlety here: I am implicitly assuming that the particles are bosons. We'll look at this more closely in Section 2.4.

where we've been a little lazy and dropped the $\langle \cdot \rangle$ brackets on $N(\mathbf{p})$. We usually write this in terms of the particle density $n = N/V$,

$$n = \frac{1}{(2\pi\hbar)^3} \int d^3p N(\mathbf{p}) = \frac{4\pi}{(2\pi\hbar)^3} \int_0^\infty dp \frac{p^2}{e^{-\beta\mu} e^{\beta p^2/2m} - 1} \quad (2.23)$$

where, in the second equality, we have chosen to integrate using spherical polar coordinates, picking up a factor of 4π from the angular integrals and a factor of p^2 in the Jacobian for our troubles. We have also used the explicit expression $E_{\mathbf{p}} = p^2/2m$ for the energy in the distribution.

At this stage, we have an annoying looking integral to do. To proceed, let's pick a value of the chemical potential μ such that $e^{-\beta\mu} \gg 1$. (We'll see what this means physically below.) We can then drop the -1 in the denominator and approximate the integral as

$$n \approx \frac{4\pi}{(2\pi\hbar)^3} e^{\beta\mu} \int_0^\infty dp p^2 e^{-\beta p^2/2m} = \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2} e^{\beta\mu} \quad (2.24)$$

Let's try to interpret this. Read naively, it seems to tell us that the number density of particles depends on the temperature. But that's certainly *not* what happens for the gas in this room, where ρ and P depend on temperature but the number density $n = N/V$ is fixed. We can achieve this by taking the chemical potential μ to also depend on temperature. Specifically, we wish to describe a gas with fixed n , then we simply invert the equation above to get an expression for the chemical potential

$$e^{\beta\mu} = \left(\frac{2\pi\hbar^2}{mk_B T} \right)^{3/2} n \quad (2.25)$$

Before we proceed, we can use this result to understand what the condition $e^{-\beta\mu} \gg 1$, that we used to do the integral, is forcing upon us. Comparing to the expression above, it says that the number density is bounded above by

$$n \ll \left(\frac{mk_B T}{2\pi\hbar^2} \right)^{3/2}$$

This is sensible. It's telling us that the ideal gas can't be too dense. In particular, the average distance between particles should be much larger than the length scale set by $\lambda = \sqrt{2\pi\hbar^2/mk_B T}$. This is the average de Broglie wavelength of particles at temperature T . If n is increased so that the separation between particles is comparable to λ then quantum effects kick in and we have to return to our original integral (2.23) and make a different approximation to do the integral and understand the physics. (This path will lead to the beautiful phenomenon of Bose-Einstein condensation, but it is a subject for a [different course](#).)

We can now calculate the energy density and pressure. Once again, taking the limit $e^{-\beta\mu} \gg 1$, the energy density is given by

$$\begin{aligned}\rho &= \frac{1}{(2\pi\hbar)^3} \int d^3p E_{\mathbf{p}} N(\mathbf{p}) \\ &\approx \frac{4\pi}{(2\pi\hbar)^3} e^{\beta\mu} \int_0^\infty dp \frac{p^4}{2m} e^{-\beta p^2/2m} = \frac{3}{2} n k_B T\end{aligned}$$

This is a result that we have met before (2.8). Meanwhile, we can use our expression (2.7) to compute the pressure,

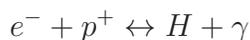
$$\begin{aligned}P &= \frac{1}{(2\pi\hbar)^3} \int d^3p \frac{\mathbf{v} \cdot \mathbf{p}}{3} N(\mathbf{p}) \\ &= \frac{4\pi}{(2\pi\hbar)^3} e^{\beta\mu} \int_0^\infty dp \frac{p^4}{3m} e^{-\beta p^2/2m} = n k_B T\end{aligned}$$

Again, this recovers the familiar ideal gas equation.

So far, the chemical potential has not bought us anything new. We have simply recovered old results in a slightly more convoluted framework in which the number of particles can fluctuate. But, as we will now see, this is exactly what we need to deal with atomic reactions.

2.3.3 The Saha Equation

We would like to consider a gas of electrons and protons in equilibrium at some temperature. They have the possibility to combine and form hydrogen, which we will think of as an atomic reaction, akin to the chemical reactions that we met in school. It is



The question we would like to ask is: what proportion of the particles are hydrogen, and what proportion are electron-proton pairs?

To simplify life, we will assume that the hydrogen atom forms in its ground state, with a binding energy

$$E_{\text{bind}} \approx 13.6 \text{ eV}$$

In fact, this turn out to be a bad assumption! We explain why at the end of this section.

Naively, we would expect hydrogen to ionize when we reach temperatures of $k_B T \approx E_{\text{bind}}$. It's certainly true that for temperature $k_B T \gg E_{\text{bind}}$, the electrons can no longer cling on to the protons, and any hydrogen atom is surely ripped apart. However, it will ultimately turn out that hydrogen only forms at temperatures significantly lower than E_{bind} .

We'll treat each of the massive particles – the electron, proton and hydrogen atom – in a similar way to the non-relativistic gas that we met in Section 2.3.2. There will, however, be two differences. First, we include the rest mass energy of the atoms, so each particle has energy

$$E_{\mathbf{p}} = mc^2 + \frac{p^2}{2m}$$

This will be useful as we can think of the binding energy E_{bind} as the mass difference

$$(m_e + m_p - m_H)c^2 = E_{\text{bind}} \approx 13.6 \text{ eV} \quad (2.26)$$

Secondly, each of our particles comes with a number g of internal states. The electron and proton each have $g_e = g_p = 2$ corresponding to the two spin states, referred to as “spin up” and “spin down”. (These are analogous to the two polarisation states of the photon that we included when discussing blackbody radiation.) For hydrogen, we have $g_H = 4$; the electron and proton spin can either be aligned, to give a spin 0 particle, or anti-aligned to give 3 different spin 1 states.

With these two amendments, our expression for the number density (2.24) of the different species of particles is given by

$$n_i = g_i \left(\frac{m_i k_B T}{2\pi \hbar^2} \right)^{3/2} e^{-\beta(m_i c^2 - \mu_i)} \quad (2.27)$$

Note that the rest mass energy mc^2 in the energy can be absorbed by a constant shift of the chemical potential.

Now we can use the chemical potential for something new. We require that these particles are in chemical equilibrium. This means that there is no rapid change from $e^- + p^+$ pairs into hydrogen, or vice versa: the numbers of electrons, protons and hydrogen are balanced. This is ensured if the chemical potentials are related by

$$\mu_e + \mu_p = \mu_H \quad (2.28)$$

This follows from our original discussion of what it means to be in chemical equilibrium. Recall that if two isolated systems have the same chemical potential then, when brought together, there will be no net flux of particles from one system to the other. This mimics the statement about thermal equilibrium, where if two isolated systems have the same temperature then, when brought together, there will be no net flux of energy from one to the other.

There is no chemical potential for photons because they're not conserved. In particular, in addition to the reaction $e^- + p^+ \leftrightarrow H + \gamma$ there can also be reactions in which the binding results in two photons, $e^- + p^+ \leftrightarrow H + \gamma + \gamma$, which is ultimately why it makes no sense to talk about a chemical potential for photons. (Some authors write this, misleadingly, as $\mu_\gamma = 0$.)

We can use the condition for chemical equilibrium (2.28) to eliminate the chemical potentials in (2.27) to find

$$\frac{n_H}{n_e n_p} = \frac{g_H}{g_e g_p} \left(\frac{m_H}{m_e m_p} \frac{2\pi\hbar^2}{k_B T} \right)^{3/2} e^{-\beta(m_H - m_e - m_p)c^2} \quad (2.29)$$

In the pre-factor, it makes sense to approximate $m_H \approx m_p$. However, in the exponent, the difference between these masses is crucial; it is the binding energy of hydrogen (2.26). Finally, we use the observed fact that the universe is electrically neutral, so

$$n_e = n_p$$

We then have

$$\frac{n_H}{n_e^2} = \left(\frac{2\pi\hbar^2}{m_e k_B T} \right)^{3/2} e^{\beta E_{\text{bind}}} \quad (2.30)$$

This is the *Saha equation*.

Our goal is to understand the fraction of electron-proton pairs that have combined into hydrogen. To this end, we define the *ionisation fraction*

$$X_e = \frac{n_e}{n_B} \approx \frac{n_e}{n_p + n_H}$$

where, in the second equality, we're ignoring neutrons and higher elements. (We'll see in Section 2.5.3 that this is a fairly good approximation.) Since $n_e = n_p$, if $X_e = 1$ it means that all the electrons are free. If $X_e = 0.1$, it means that only 10% of the electrons are free, the remainder bound inside hydrogen.

Using $n_e = n_p$, we have $1 - X_e = n_H/n_B$ and so

$$\frac{1 - X_e}{X_e^2} = \frac{n_H}{n_e^2} n_B$$

The Saha equation gives us an expression for n_H/n_e^2 . But to translate this into the fraction X_e , we also need to know the number of baryons. This we take from observation. First, we convert the number of baryons into the number of photons, using (2.17),

$$\eta = \frac{n_B}{n_\gamma} \approx 10^{-9}$$

Here we need to use the fact that $\eta \approx 10^{-9}$ has remained constant since recombination. Next, we use the fact that photons sit at the same temperature as the electrons, protons and hydrogen because they are all in equilibrium. This means that we can then use our earlier expression (2.15) for the number of photons

$$n_\gamma = \frac{2\zeta(3)}{\pi^2 \hbar^3 c^3} (k_B T)^3$$

Combining these gives our final answer

$$\frac{1 - X_e}{X_e^2} = \eta \frac{2\zeta(3)}{\pi^2} \left(\frac{2\pi k_B T}{m_e c^2} \right)^{3/2} e^{\beta E_{\text{bind}}} \quad (2.31)$$

Suppose that we look at temperature $k_B T \sim E_{\text{bind}}$, which is when we might naively have thought recombination takes place. We see that there are two very small numbers in the game: the factor of $\eta \sim 10^{-9}$ and $k_B T / m_e c^2$, where the electron mass is $m_e c^2 \approx 0.5 \text{ MeV} = 5 \times 10^5 \text{ eV}$. These ensure that at $k_B T \sim E_{\text{bind}}$, the ionisation fraction X_e is very close to unity. In other words, nearly all the electrons remain free and unbound. In large part this is of the enormous number of photons, which mean that whenever a proton and electron bind, one can still find sufficient high energy photons in the tail of the blackbody distribution to knock them apart.

Recombination only takes place when the $e^{\beta E_{\text{bind}}}$ factor is sufficient to compensate both the η and $k_B T / m_e c^2$ factors. Clearly recombination isn't a one-off process; it happens continuously as the temperature varies. As a benchmark, we'll calculate the temperature when $X_e = 0.1$, so 90% of the electrons are sitting happily in their hydrogen homes. From (2.31), we learn that this occurs when $\beta E_{\text{bind}} \approx 45$, or

$$k_B T_{\text{rec}} \approx 0.3 \text{ eV} \quad \Rightarrow \quad T_{\text{rec}} \approx 3600 \text{ K}$$

This corresponds to a redshift of

$$z_{\text{rec}} = \frac{T_{\text{rec}}}{T_0} \approx 1300$$

This is significantly later than matter-radiation equality which, as we saw in (1.73), occurs at $z_{\text{eq}} \approx 3400$. This means that, during recombination, the universe is matter dominated, with $a(t) \sim (t/t_0)^{2/3}$. We can therefore date the time of recombination to,

$$t_{\text{rec}} \approx \frac{t_0}{(1 + z_{\text{rec}})^{3/2}} \approx 300,000 \text{ years}$$

After recombination, the constituents of the universe have been mostly neutral atoms. Roughly speaking this means that the universe is transparent and photons can propagate freely. We will look more closely at this statement a little more closely below.

Mea Culpa

The full story is significantly more complicated than the one told above. As we have seen, at the time of recombination the temperature is much lower than the 13.6 eV binding energy of the 1s state of hydrogen. This means that whenever a 1s state forms, it emits a photon which has significantly higher energy than the photons in thermal bath. The most likely outcome is that this high energy photon hits a different hydrogen atom, splitting it into its constituent proton and electron, resulting in no net change in the number of atoms! Instead, recombination must proceed through a rather more tortuous route.

The hydrogen atom doesn't just have a ground state: there are a whole tower of excited states. These can form without emitting a high energy photon and, indeed, at these low temperatures the thermal bath of photons is in equilibrium with the tower of excited states of hydrogen. There are then two, rather inefficient processes, which populate the 1s state. The 2s state decays down to 1s by emitting two photons (to preserve angular momentum), neither of which have enough energy to re-ionize other atoms. Alternatively, the 2p state can decay to 1s, emitting a photon whose energy is barely enough to excite another hydrogen atom out of the ground state. If this photon experiences redshift, then it can no longer do the job and we increase the number of atoms in the ground state. More details can be found in the book by Weinberg. These issues do not greatly change the values of T_{rec} and z_{rec} that we computed above.

2.3.4 Freeze Out and Last Scattering

Photons interact with electric charge. After electrons and protons combine to form neutral hydrogen, the photons scatter much less frequently and the universe becomes transparent. After this time, the photons are essentially decoupled.

Similar scenarios play out a number of times in the early universe: particles, which once interacted frequently, stop talking to their neighbours and subsequently evolve without care for what's going on around them. This process is common enough that it is worth exploring in a little detail. As we will see, at heart it hinges on what it means for particle to be in “equilibrium”.

Strictly speaking, an expanding universe is a time dependent background in which the concept of equilibrium does not apply. In most situations, such a comment would be rightly dismissed as the height of pedantry. The expansion of the universe does not, for example, stop me applying the laws of thermodynamics to my morning cup of tea. However, in the very early universe this can become an issue.

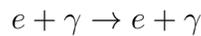
For a system to be in equilibrium, the constituent particles must frequently interact, exchanging energy and momentum. For any species of particle (or pair of species) we can define the *interaction rate* Γ . A particle will, on average, interact with another particle in a time $t_{\text{int}} = 1/\Gamma$. It makes sense to talk about equilibrium provided that the universe hasn't significantly changed in the time t_{int} . The expansion of the universe is governed by the Hubble parameter, so we can sensibly talk about equilibrium provided

$$\Gamma \gg H$$

In contrast, if $\Gamma \ll H$ then by the time particles interact the universe has undergone significant expansion. In this case, thermal equilibrium cannot be maintained.

For many processes, both the interaction rate and temperature scale with time, but in different ways. The result is that particles retain equilibrium at early times, but decouple from the thermal bath at late time. This decoupling occurs when $\Gamma \approx H$ and is known as *freeze out*.

We now apply these ideas to photons, where freeze out also goes by the name of *last scattering*. In the early universe, the photons are scattered primarily by the electrons (because they are much lighter than the protons) in a process known as *Thomson scattering*



The scattering is elastic, meaning that the energy, and therefore the frequency, of the photon is unchanged in the process. For Thomson scattering, the interaction rate is given by

$$\Gamma = n_e \sigma_T c$$

where σ_T is the cross-section, a quantity which characterises the strength of the scattering. We computed the cross-section for Thomson scattering in the lectures on [Electromagnetism](#) (see Section 6.3.1 of these lectures) where we showed it was given by

$$\sigma_T = \frac{\mu_0^2 e^4}{6\pi m_e^2 c^2} \approx 6 \times 10^{-30} \text{ m}^2$$

Note the dependence on the electron mass m_e ; the corresponding cross-section for scattering off protons is more than a million times smaller.

Last scattering occurs at the temperature T_{last} such that $\Gamma(T_{\text{last}}) \approx H(t_{\text{last}})$. We can express the interaction rate by replacing the number density of electrons with the number density of photons,

$$\Gamma(T_{\text{last}}) = n_B X_e(T_{\text{last}}) \sigma_{Tc} = \eta \sigma_T \frac{2\zeta(3)}{\pi^2 \hbar^3 c^2} (k_B T_{\text{last}})^3 X_e(T_{\text{last}}) \quad (2.32)$$

Meanwhile, we can trace back the current value of the Hubble constant, through the matter dominated era, to last scattering. Meanwhile, to compute $H(T_{\text{last}})$, we use the formula (1.67)

$$\left(\frac{H}{H_0}\right)^2 = \frac{\Omega_r}{a^4} + \frac{\Omega_m}{a^3} + \frac{\Omega_k}{a^2} + \Omega_\Lambda$$

Evaluated at recombination, radiation, curvature and the cosmological constant are all irrelevant, and this formula becomes

$$\left(\frac{H}{H_0}\right)^2 \approx \frac{\Omega_m}{a^3}$$

Using the fact that temperature scales as $T \sim 1/a$, we then have

$$H(T_{\text{last}}) = H_0 \sqrt{\Omega_m} \left(\frac{T_{\text{last}}}{T_0}\right)^{3/2}$$

Equating this with (2.32) gives

$$X_e(T_{\text{last}}) (k_B T_{\text{last}})^{3/2} = \frac{\pi^2 \hbar^3 c^2}{2\zeta(3)} \frac{H_0 \sqrt{\Omega_m}}{\eta \sigma_T (k_B T_0)^{3/2}}$$

Using (2.31) to solve for $X_e(T_{\text{last}})$ (which is a little fiddly) we find that photons stop interacting with matter only when

$$X_e(T_{\text{last}}) \approx 0.01$$

We learn that the vast majority of electrons must be housed in neutral hydrogen, with only 1% of the original electrons remaining free, before light can happily travel unimpeded. This corresponds to a temperature

$$k_B T_{\text{last}} \approx 0.27 \text{ eV} \quad \Rightarrow \quad T_{\text{last}} \approx 3100 \text{ K}$$

and, correspondingly, a time somewhat after recombination,

$$z_{\text{last}} = \frac{T_{\text{last}}}{T_0} \approx 1100 \quad \Rightarrow \quad t_{\text{last}} = \frac{t_0}{(1 + z_{\text{last}})^{3/2}} \approx 350,000 \text{ years}$$

After this time, the universe becomes transparent. The cosmic microwave background is a snapshot of the universe from this time.

2.4 Bosons and Fermions

To better understand the physics of the Big Bang, there is one last topic from statistical physics that we will need to understand. This follows from a simple statement: quantum particles are indistinguishable. It's not just that the particles look the same: there is a very real sense in which there is no way to tell them apart.

Consider a state with two identical particles. Now swap the positions of the particles. This doesn't give us a new state: it is exactly the same state as before (at least up to a minus sign). This subtle effect plays a key role in thermal systems where we're taking averages over different states. The possibility of a minus sign is important, and means that quantum particles come in two different types, called *bosons* and *fermions*.

Consider a state with two identical particles. These particles are called *bosons* if the wavefunction is symmetric under exchange of the particles.

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = \psi(\mathbf{x}_2, \mathbf{x}_1)$$

The particles are *fermions* if the wavefunction is anti-symmetric

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = -\psi(\mathbf{x}_2, \mathbf{x}_1)$$

Importantly, if you try to put two fermions on top of each other then the wavefunction vanishes: $\psi(\mathbf{x}, \mathbf{x}) = 0$. This is a reflection of the *Pauli exclusion principle* which states that two or more fermions cannot sit in the same state. For both bosons and fermions, if you do the exchange twice then you get back to the original state.

There is a deep theorem – known as the spin-statistics theorem – which states that the type of particle is determined by its spin (an intrinsic angular momentum carried by elementary particles). Particles that have integer spin are bosons; particles that have half-integer spin are fermions.

Examples of spin 1/2 particles, all of which are fermions, include the electron, the various quarks, and neutrinos. Furthermore, protons and neutrons (which, roughly speaking, consist of three quarks) also have spin 1/2 and so are fermions.

The most familiar example of a boson is the photon. It has spin 1. Other spin 1 particles include the W and Z-bosons (responsible for the weak nuclear force) and gluons (responsible for the strong nuclear force). The only elementary spin 0 particle is the Higgs boson. Finally, the graviton has spin 2 and is also a boson.

While this exhausts the elementary particles, the ideas that we develop here also apply to composite objects like atoms. These too are either bosons or fermions. Since the number of electrons is always equal to the number of protons, it is left to the neutrons to determine the nature of the atom: an odd number of neutrons and it's a fermion; an even number and it's a boson.

2.4.1 Bose-Einstein and Fermi-Dirac Distributions

The generalised Boltzmann distribution (2.21) specifies the probability that we sit in a state $|n\rangle$ with some fixed energy E_n and particle number N_n .

In what follows, we will restrict attention to non-interacting particles. In this case, there is a simple way to construct the full set of states $|n\rangle$ starting from the single-particle Hilbert space. The state of a single particle is specified by its momentum $\mathbf{p} = \hbar\mathbf{k}$. (There may also be some extra, discrete internal degrees of freedom like polarisation or spin; we'll account for these later.) We'll denote this single particle state as $|\mathbf{p}\rangle$. For a relativistic particle, the energy is

$$E_{\mathbf{p}} = \sqrt{m^2c^4 + p^2c^2} \quad (2.33)$$

To specify the full multi-particle state $|n\rangle$, we need to say how many particles $n_{\mathbf{p}}$ occupy the state $|\mathbf{p}\rangle$. The possible values of $n_{\mathbf{p}}$ depend on whether the underlying particle is a boson or fermion:

$$\begin{aligned} \text{Bosons :} & \quad n_{\mathbf{p}} = 0, 1, 2, \dots \\ \text{Fermions :} & \quad n_{\mathbf{p}} = 0, 1 \end{aligned}$$

In our previous discussions of blackbody radiation in Section 2.2.1 and the non-relativistic gas in Section 2.3.2, we did the counting appropriate for bosons. This is fine for blackbody radiation, since photons are bosons, but was an implicit assumption in the case of a non-relativistic gas.

The other alternative is a fermion. For these particles, the Pauli exclusion principle says that a given single-particle state $|\mathbf{p}\rangle$ is either empty or occupied. But you can't put more than one fermion there. This is entirely analogous to the way the periodic table is constructed in chemistry, by filling successive shells, except now the states are in momentum space. (A better analogy is the way a band is filled in solid state physics as described in the lectures on [Quantum Mechanics](#).) For bosonic particles, there is no such restriction: you can pile up as many as you like.

Now we can compute some quantities, like the average particle number and average energy. We deal with bosons and fermions in turn

For bosons, the calculation is exactly the same as we saw in Section 2.3.2. For a given momentum \mathbf{p} , the average number of photons is

$$\langle N(\mathbf{p}) \rangle = \frac{1}{\mathcal{Z}_{\mathbf{p}}} \sum_{n_{\mathbf{p}}=0}^{\infty} n_{\mathbf{p}} e^{-\beta(n_{\mathbf{p}}E_{\mathbf{p}} - \mu n_{\mathbf{p}})} = \frac{1}{\beta} \frac{\partial}{\partial \mu} \log \mathcal{Z}_{\mathbf{p}}$$

where the normalisation factor is given by the geometric series

$$\mathcal{Z}_{\mathbf{p}} = \sum_{n=0}^{\infty} e^{-\beta n_{\mathbf{p}}(E_{\mathbf{p}} - \mu)} = \frac{e^{\beta(E_{\mathbf{p}} - \mu)}}{e^{\beta(E_{\mathbf{p}} - \mu)} - 1}$$

As in the previous section, we will be a little lazy and drop the expectation value, so $\langle N(\mathbf{p}) \rangle \equiv N(\mathbf{p})$. Then we have

$$N(\mathbf{p}) = \frac{1}{e^{\beta(E_{\mathbf{p}} - \mu)} - 1} \quad (2.34)$$

This is known as the *Bose-Einstein distribution*.

For fermions, the calculation is easier still. We can have only $n_{\mathbf{p}} = 0$ or 1 particles in a given state $|\mathbf{p}\rangle$ so the average occupation number is

$$N(\mathbf{p}) = \frac{1}{\mathcal{Z}_{\mathbf{p}}} \sum_{n_{\mathbf{p}}=0,1} n_{\mathbf{p}} e^{-\beta(n_{\mathbf{p}}E_{\mathbf{p}} - \mu n_{\mathbf{p}})} \quad \text{with} \quad \mathcal{Z}_{\mathbf{p}} = \sum_{n_{\mathbf{p}}=0,1} e^{-\beta(n_{\mathbf{p}}E_{\mathbf{p}} - \mu n_{\mathbf{p}})}$$

Again, keeping the $\langle \cdot \rangle$ expectation value signs implicit, we have

$$N(\mathbf{p}) = \frac{1}{e^{\beta(E_{\mathbf{p}} - \mu)} + 1} \quad (2.35)$$

This is the *Fermi-Dirac distribution*.

For both bosons and fermions, the calculation of the density of states (2.10) proceeds as before, so that if we integrate over all possible momenta, it should be weighted by

$$\frac{V}{(2\pi\hbar)^3} \int d^3p$$

with the pre-factor telling us how quantum states are in a small region d^3p .

If we include the degeneracy factor g , which tells us the number of internal states of the particle, the number density $n = N/V$ is given by

$$n = \frac{g}{(2\pi\hbar)^3} \int d^3p N(\mathbf{p}) \quad (2.36)$$

Similarly, the energy density is

$$\rho = \frac{g}{(2\pi\hbar)^3} \int d^3p E_{\mathbf{p}} N(\mathbf{p}) \quad (2.37)$$

and the pressure (2.7) is

$$P = \frac{g}{(2\pi\hbar)^3} \int d^3p \frac{\mathbf{v} \cdot \mathbf{p}}{3} N(\mathbf{p}) \quad (2.38)$$

We'll now apply these in various examples.

The Non-Relativistic Gas Yet Again

In Section 2.3.2, we computed various quantities of a non-relativistic gas, so that the energy of each particle is

$$E_{\mathbf{p}} = \frac{p^2}{2m}$$

When we evaluated various quantities using the chemical potential approach, we implicitly assumed that the constituent atoms of the gas were bosons so, for example, our expression for the expression for the number density (2.23),

$$n_{\text{boson}} = \frac{g}{(2\pi\hbar)^3} \int d^3p N(\mathbf{p}) = \frac{4\pi g}{(2\pi\hbar)^3} \int_0^\infty dp \frac{p^2}{e^{-\beta\mu} e^{\beta p^2/2m} - 1}$$

If, instead, we have a gas comprising of fermions then we should replace this expression with

$$n_{\text{fermion}} = \frac{g}{(2\pi\hbar)^3} \int d^3p N(\mathbf{p}) = \frac{4\pi g}{(2\pi\hbar)^3} \int_0^\infty dp \frac{p^2}{e^{-\beta\mu} e^{\beta p^2/2m} + 1}$$

We can then ask: how does the physics change?

If we focus on the high temperature regime of non-relativistic gases, the answer to this question is: very little! This is because we evaluate these integrals using the approximation $e^{-\beta\mu} \gg 1$, and we can immediately drop the ± 1 in the denominator. This means that both bosons and fermions give rise to the same ideal gas equation.

We do start to see small differences in the behaviour of the gases if we expand the integrals to the next order in $e^{\beta\mu}$. We see much larger differences if we instead study the integrals in a very low-temperature limit. These stories are told in the lectures on [Statistical Physics](#) but they hold little cosmological interest.

Instead, the difference between bosons and fermions in cosmology is really only important when we turn to very high temperatures, where the gas becomes relativistic.

2.4.2 Ultra-Relativistic Gases

As we will see in the next section, as we go further back in time, the universe gets hot. Really hot. For any particle, there will be a time such that

$$k_B T \gg 2mc^2$$

In this regime, particle-anti-particle pairs can be created in the fireball. When this happens, both the mass and the chemical potential are negligible. We say that the particles are *ultra-relativistic*, with their energy given approximately as

$$E_{\mathbf{p}} \approx pc$$

just as for a massless particle. We can use our techniques to study the behaviour of gases in this regime.

We start with ultra-relativistic bosons. We work with vanishing chemical potential, $\mu = 0$. (This will ensure that we have equal numbers of particles and anti-particles. The presence of a chemical potential results in a preference for one over the other, and will be explored in Examples Sheet 3.) The integral (2.36) for the number density gives

$$n_{\text{boson}} = \frac{4\pi g}{(2\pi\hbar)^3} \int dp \frac{p^2}{e^{\beta pc} - 1} = \frac{gI_2}{2\pi^2\hbar^3 c^3} (k_B T)^3$$

while the energy density is

$$\rho_{\text{boson}} = \frac{4\pi g}{(2\pi\hbar)^3} \int dp \frac{p^3 c}{e^{\beta pc} - 1} = \frac{gI_3}{2\pi^2\hbar^3 c^3} (k_B T)^4$$

where we've used the definition (2.13) of the integral

$$I_n = \int_0^\infty dy \frac{y^n}{e^y - 1} = \Gamma(n+1)\zeta(n+1)$$

In both cases, the integrals coincide with those that we met for blackbody radiation

Meanwhile, for fermions we have

$$n_{\text{fermion}} = \frac{4\pi g}{(2\pi\hbar)^3} \int dp \frac{p^2}{e^{\beta pc} + 1} = \frac{gJ_2}{2\pi^2\hbar^3 c^3} (k_B T)^3$$

and

$$\rho_{\text{fermion}} = \frac{4\pi g}{(2\pi\hbar)^3} \int dp \frac{p^3 c}{e^{\beta pc} + 1} = \frac{g J_3}{2\pi^2 \hbar^3 c^3} (k_B T)^4$$

where, this time, we get the integral

$$J_n = \int_0^\infty dy \frac{y^n}{e^y + 1} = \int_0^\infty dy \left[\frac{y^n}{e^y - 1} - \frac{2y^n}{e^{2y} - 1} \right] = \left(1 - \frac{1}{2^n}\right) I_n$$

The upshot of these calculations is that the number density is

$$n = \frac{g \zeta(3)}{\pi^2 \hbar^3 c^3} (k_B T)^3 \times \begin{cases} 1 & \text{for bosons} \\ \frac{3}{4} & \text{for fermions} \end{cases}$$

and the energy density is

$$\rho = \frac{g \pi^2}{30 \hbar^3 c^3} (k_B T)^4 \times \begin{cases} 1 & \text{for bosons} \\ \frac{7}{8} & \text{for fermions} \end{cases}$$

The differences are just small numerical factors but, as we will see, these become important in cosmology.

Ultimately, we will be interested in gases that contain many different species of particles. In this case, it is conventional to define the effective number of relativistic species in thermal equilibrium as

$$g_\star(T) = \sum_{\text{bosons}} g_i + \frac{7}{8} \sum_{\text{fermions}} g_i \quad (2.39)$$

As the temperature drops below a particle's mass threshold, $k_B T < m_i c^2$, this particle is removed from the sum. In this way, the number of relativistic species is both time and temperature dependent. The energy density from all relativistic species is then written as

$$\rho = g_\star \frac{\pi^2}{30 \hbar^3 c^3} (k_B T)^4 \quad (2.40)$$

To calculate g_\star in different epochs, we need to know the matter content of the Standard Model and, eventually, the identity of dark matter. We'll make a start on this in the next section.

2.5 The Hot Big Bang

We have seen that for the first 300,000 years or so, the universe was filled with a fireball in which photons were in thermal equilibrium with matter. We would like to understand what happens to this fireball as we dial the clock back further. This collection of ideas goes by the name of the hot Big Bang theory.

2.5.1 Temperature vs Time

It turns out, unsurprisingly, that the fireball is hotter at earlier times. This is simplest to describe if we go back to when the universe is radiation dominated, at $z > 3400$ or $t < 50,000$ years. Here, the energy density scales as (1.41),

$$\rho \sim \frac{1}{a^4}$$

We can compare this to the thermal energy density of photons, given by (2.14)

$$\rho = \frac{\pi^2}{15\hbar^3 c^3} (k_B T)^4$$

To see that the temperature scales inversely with the scale factor

$$T \sim \frac{1}{a} \tag{2.41}$$

This is the same temperature scaling that we saw for the CMB after recombination (2.19). Indeed, the underlying arguments are also the same: the energy of each photon is blue-shifted as we go back in time, while their number density increases, resulting in the $\rho \sim 1/a^4$ behaviour. The difference is that now the photons are in equilibrium. If they are disturbed in some way, they will return to their equilibrium state. In contrast, if the photons are disturbed after recombination they will retain a memory of this.

What happens during the time $1100 < z < 3400$, before recombination but when matter was the dominant energy component? First consider a universe with only non-relativistic matter, with number density n . The energy density is

$$\rho_m = nmc^2 + \frac{1}{2}nmv^2$$

The first term drives the expansion of the universe and is independent of temperature. The second term, which we completely ignored in Section 1 on the grounds that it is negligible, depends on temperature. This was computed in (2.8) and is given by $\frac{1}{2}nmv^2 = \frac{3}{2}nk_B T$.

As the universe expands, the velocity of non-relativistic particles is red-shifted as $v \sim 1/a$. (This is hopefully intuitive, but we have not actually demonstrated this previously. We will derive this redshift in Section 3.1.3.) This means that, in a universe with *only* non-relativistic matter, we would have

$$T \sim \frac{1}{a^2}$$

So what happens when we have both matter and radiation? We would expect that the temperature scaling sits somewhere between $T \sim 1/a$ and $T \sim 1/a^2$. In fact, it is entirely dominated by the radiation contribution. This can be traced to the fact that there are many more photons than baryons; $\eta = n_B/n_\gamma \approx 10^{-9}$. A comparable ratio is expected to hold for dark matter. This means that the photons, rather than matter, dictate the heat capacity of the thermal bath. The upshot is that the temperature scales as $T \sim 1/a$ throughout the period of the fireball. Moreover, as we saw in Section 2.2, the temperature of the photons continues to scale as $T \sim 1/a$ even after they decouple.

Doing a Better Job

The formula $T \sim 1/a$ gives us an approximate scaling. But we can do better.

We start with the continuity equation (1.39) for relativistic matter, with $P = \rho/3$, is

$$\dot{\rho} = -3H(\rho + P) = -4H\rho \quad (2.42)$$

But for ultra-relativistic gases, we know that the energy density is given by (2.43), have

$$\rho = g_\star \frac{\pi^2}{30 \hbar^3 c^3} (k_B T)^4 \quad (2.43)$$

where g_\star is the effective number of relativistic degrees of freedom (2.39). Differentiating this with respect to time, and assuming that g_\star is constant, we have

$$\dot{\rho} = \frac{4\dot{T}}{T}\rho \quad \Rightarrow \quad \dot{T} = -HT$$

where the second expression comes from (2.42). This is just re-deriving the fact that $T \sim 1/a$. However, now we have use the Friedmann equation to determine the Hubble parameter in the radiation dominated universe,

$$H^2 = \frac{8\pi G}{3c^2}\rho = A(k_B T)^4 \quad \text{with } A = \frac{8\pi^3 G}{90 \hbar^3 c^5} g_\star$$

This leaves us with a straightforward differential equation for the temperature,

$$k_B \dot{T} = -\sqrt{A}(k_B T)^3 \quad \Rightarrow \quad t = \frac{1}{2\sqrt{A}} \frac{1}{(k_B T)^2} + \text{constant} \quad (2.44)$$

We choose to set the integration constant to zero. This means that the temperature diverges as we approach the Big Bang singularity at $t = 0$. All times will be measured from this singularity.

To turn this into something physical, we need to make sense of the morass of fundamental constants in A . The presence of Newton's constant is associated with a very high energy scale known as the *Planck mass* with the corresponding *Planck energy*,

$$M_{\text{pl}} c^2 = \sqrt{\frac{\hbar c^5}{8\pi G}} \approx 2.4 \times 10^{21} \text{ MeV}$$

Meanwhile, the value of Planck's constant is

$$\hbar \approx 6.6 \times 10^{-16} \text{ eV s} = 6.6 \times 10^{-22} \text{ MeV s}$$

These combine to give

$$\hbar M_{\text{pl}} c^2 \approx 1.6 \text{ MeV}^2 \text{ s}$$

Putting these numbers into (2.44) gives us an expression that tells us the temperature T at a given time t ,

$$\left(\frac{t}{1 \text{ second}} \right) \approx \frac{2.4}{g_\star^{1/2}} \left(\frac{1 \text{ MeV}}{k_B T} \right)^2 \quad (2.45)$$

Ignoring the constants of order 1, we say that the universe was at a temperature of $k_B T = 1 \text{ MeV}$ approximately 1 second after the Big Bang.

As an aside: most textbooks derive the relationship (2.45) by assuming conservation of entropy (which, it turns out, ensures that $g_\star T^3 a^3$ is constant). The derivation given above is entirely equivalent to this.

To finish, we need to get a handle on the effective number of relativistic degrees of freedom g_\star . In the very early universe many particles were relativistic and g_\star is bigger. As the universe cools, it goes through a number of stages where g_\star drops discontinuously as the heavier particles become non-relativistic.

For example, when temperatures are around $k_B T \sim 10^6 \text{ eV} \equiv 1 \text{ MeV}$, the relativistic species are the photon (with $g_\gamma = 2$), three neutrinos and their anti-neutrinos (each with $g_\nu = 1$) and the electron and positron (each with $g_e = 2$). The effective number of relativistic species is then

$$g_\star = 2 + \frac{7}{8} (3 \times 1 + 3 \times 1 + 2 + 2) = 10.75 \quad (2.46)$$

As we go back in time, more and more species contribute. By the time we get to $k_B T \sim 100 \text{ GeV}$, all the particles of the Standard Model are relativistic and contribute $g_\star = 106.75$.

In contrast, as we move forward in time, g_\star decreases. Considering only the masses of Standard Model particles, one might naively think that, as electrons and positrons annihilate and become non-relativistic, we're left only with photons, neutrinos and anti-neutrinos. This would give

$$g_\star = 2 + \frac{7}{8} (3 + 3) = 7.25$$

Unfortunately, at this point one of many subtleties arises. It turns out that the neutrinos are very weakly interacting and have already decoupled from thermal equilibrium by the time electrons and protons annihilate. When the annihilation finally happens, the bath of photons is heated while the neutrinos are unaffected. We can still use the formula (2.43), but we need an amended definition of g_\star to include the fact that neutrinos and electrons are both relativistic, but sitting at different temperatures. For now, I will simply give the answer:

$$g_\star \approx 3.4 \quad (2.47)$$

I will very briefly explain where this comes from in Section 2.5.4.

A Longish Aside on Neutrinos

Why do neutrinos only contribute 1 degree of freedom to (2.46) while the electron has 2? After all, they are both spin- $\frac{1}{2}$ particles. To explain this, we need to get a little dirty with some particle physics.

First, for many decades we thought that neutrinos are massless. In this case, the right characterisation is not spin, but something called *helicity*. Massless particles necessarily travel at the speed of light; their spin is aligned with their direction of travel. If the spin points in the same direction as the momentum, then it is said to be right-handed; if it points in the opposite direction then it is said to be left-handed. It

is a fact that we've only ever observed neutrinos with left-handed helicity and it was long believed that the right-handed neutrinos simply do not exist. Similarly, we've only observed anti-neutrinos with right-handed helicity; there appear to be no left-handed anti-neutrinos. If this were true, we would indeed get the $g = 1$ count that we saw above.

However, we now know that neutrinos do, in fact, have a very small mass. Here is where things get a little complicated. Roughly speaking, there are two different kinds of masses that neutrinos could have: they are called the *Majorana mass* and the *Dirac mass*. Unfortunately, we don't yet know which of these masses (or combination of masses) the neutrino actually has, although we very much hope to find out in the near future.

The Majorana mass is the simplest to understand. In this scenario, the neutrino is its own anti-particle. If the neutrino has a Majorana mass then what we think of as the right-handed anti-neutrino is really the same thing as the right-handed neutrino. In this case, the counting goes through in the same way, but we drape different words around the numbers: instead of getting $1 + 1$ from each neutrino + anti-neutrino, we instead get 2 spin states for each neutrino, and no separate contribution from the anti-neutrino.

Alternatively, the neutrino may have a Dirac mass. In this case, it looks much more similar to the electron, and the correct counting is 2 spin states for each neutrino, and another 2 for each anti-neutrino. Here is where things get interesting because, as we will explain in Section 2.5.3, we know from Big Bang nucleosynthesis that the count (2.46) of $g_* = 10.75$ was correct a few minutes after the Big Bang. For this reason, it must be the case that 2 of the 4 degrees of freedom interact very weakly with the thermal bath, and drop out of equilibrium in the very early universe. Their energy must then be diluted relative to everything else, so that it's negligible by the time we get to nucleosynthesis. (For example, there are various phase transitions in the early universe that could dump significant amounts of energy into half of the neutrino degrees of freedom, leaving the other half unaffected.)

2.5.2 The Thermal History of our Universe

The essence of the hot Big Bang theory is simply to take the temperature scaling $T \sim 1/a$ and push it as far back as we can, telling the story of what happens along the way.

As we go further back in time, more matter joins the fray. For some species of particles, this is because the interaction rate is sufficiently large at early times that it couples to the thermal bath. For example, there was a time when both neutrinos and (we think) dark matter were in equilibrium with the thermal bath, before both underwent freeze out.

For other species of particle, the temperatures are so great (roughly $k_B T \approx 2mc^2$) that particle-anti-particle pairs can emerge from the vacuum. For example, for the first six seconds after the Big Bang, both electrons and positrons filled the fireball in almost equal numbers.

The goal of the Big Bang theory is to combine knowledge of particle physics with our understanding of thermal physics to paint an accurate picture for what happened at various stages of the fireball. A summary of some of the key events in the early history of the universe is given in the following table. In the remainder of this section, we will tell some of these stories.

What	When (t)	When (z)	When (T)
Inflation	10^{-36} s ?	10^{28} ?	?
Baryogenesis	?	?	?
Electroweak phase transition	10^{-12} s	10^{15}	10^{22} K
QCD phase transition	10^{-6} s	10^{12}	10^{16} K
Dark Matter Freeze-Out	?	?	?
Neutrino Decoupling	1 second	6×10^9	10^{10} K
$e^- e^+$ Annihilation	6 second	2×10^9	5×10^9 K
Nucleosynthesis	3 minutes	4×10^8	10^9 K
Matter-Radiation Equality	50,000 years	3400	8700 K
Recombination	$\sim 300,000$ years	1300	3600 K
Last Scattering	350,000 years	1100	3100 K
Matter- Λ Equality	10^{10} years	0.4	3.8 K
Today	1.4×10^{10} years	0	2.7 K

2.5.3 Nucleosynthesis

One of the best understood processes in the Big Bang fireball is the formation of deuterium, helium and heavier nuclei from the thermal bath of protons and neutrons. This is known as *Big Bang nucleosynthesis*. It is a wonderfully delicate calculation, that involves input from many different parts of physics. The agreement with observation could fail in a myriad of ways, yet the end result agrees perfectly with the observed abundance of light elements. This is one of the great triumphs of the Big Bang theory.

Full calculations of nucleosynthesis are challenging. Here we simply offer a crude sketch of the formation of deuterium and helium nuclei.

Neutrons and Protons

Our story starts at early times, $t \ll 1$ second, when the temperature reached $k_B T \gg 1$ MeV. The mass of the electron is

$$m_e c^2 \approx 0.5 \text{ MeV}$$

so at this time the thermal bath contains many relativistic electron-positron pairs. These are in equilibrium with photons and neutrinos, both of which are relativistic, together with non-relativistic protons and neutrons. Equilibrium is maintained through interactions mediated by the weak nuclear force

$$n + \nu_e \leftrightarrow p + e^- \quad , \quad n + e^+ \leftrightarrow p + \bar{\nu}_e$$

These reactions arise from the same kind of process as beta decay, $n \rightarrow p + e^- + \bar{\nu}_e$.

The chemical potentials for electrons and neutrinos are vanishingly small. Chemical equilibrium then requires $\mu_n = \mu_p$, and the ratio of neutron to proton densities can be calculated using the equation (2.24) for a non-relativistic gas,

$$\frac{n_n}{n_p} = \left(\frac{m_n}{m_p} \right)^{3/2} e^{-\beta(m_n - m_p)c^2}$$

The proton and neutron have a very small mass difference,

$$\begin{aligned} m_n c^2 &\approx 939.6 \text{ MeV} \\ m_p c^2 &\approx 938.3 \text{ MeV} \end{aligned}$$

This mass difference can be neglected in the prefactor, but is crucial in the exponent. This gives the ratio of protons to neutrons while equilibrium is maintained

$$\frac{n_n}{n_p} \approx e^{-\beta\Delta mc^2} \quad \text{with} \quad \Delta mc^2 \approx 1.3 \text{ MeV}$$

For $k_B T \gg \Delta mc^2$, there are more or less equal numbers of protons and neutrons. But as the temperature falls, so too does the number of neutrons.

However, the exponential decay in neutron number does not continue indefinitely. At some point, the weak interaction rate will drop to $\Gamma \sim H$, at which point the neutrons freeze out, and their number then remains constant. (Actually, this last point isn't quite true as we will see below but let's run with it for now!)

The interaction rate can be written as $\Gamma = n\sigma v$. where σ is the cross-section. At this point, I need to pull some facts about the weak force out of the hat. The cross-section varies as temperature as $\sigma v \sim G_F T^2$ with $G_F \approx 1.2 \times 10^{-5} \text{ GeV}^{-2}$ a constant that characterises the strength of the weak force. Meanwhile, the number density scales as $n \sim T^3$. This means that $\Gamma \sim T^5$.

The Hubble parameter scales as $H \sim 1/a^2 \sim T^2$ in the radiation dominated epoch. So we do indeed expect to find $\Gamma \gg H$ at early times and $\Gamma \ll H$ at later times. It turns out that neutrons decouple at the temperature

$$k_B T_{\text{dec}} \approx 0.8 \text{ MeV}$$

Putting this into (2.45), and using $g_* \approx 3.4$, we find that neutrons decouple around

$$t_{\text{dec}} \approx 2 \text{ seconds}$$

after the Big Bang.

At freeze out, we are then left with a neutron-to-proton ratio of

$$\frac{n_n}{n_p} \approx \exp\left(-\frac{1.3}{0.8}\right) \approx \frac{1}{5}$$

In fact, this isn't the end of the story. Left alone, neutrons are unstable to beta decay with a half life of a little over 10 minutes. This means that, after freeze out, the number density of neutrons decays as

$$n_n(t) \approx \frac{1}{5} n_p(t_{\text{dec}}) e^{-t/\tau_n} \tag{2.48}$$

where $\tau_n \approx 880$ second. If we want to do something with those neutrons (like use them to form heavier nuclei) then we need to hurry up: the clock is ticking.

Deuterium

Ultimately, we want to make elements heavier than hydrogen. But these heavier nuclei contain more than two nucleons. For example, the lightest is ${}^3\text{He}$ which contains two protons and a neutron. But the chance of three particles colliding at the same time to form such a nuclei is way too small. Instead, we must take baby steps, building up by colliding two particles at a time.

The first such step is, it turns out, the most difficult. This is the step to *deuterium*, or heavy hydrogen, consisting of a bound state of a proton and neutron that forms through the reaction



The binding energy is

$$E_{\text{bind}} = m_n + m_p - m_D \approx 2.2 \text{ MeV}$$

Both the proton and neutron have spin $1/2$, and so have $g_n = g_p = 2$. In deuterium, the spins are aligned to form a spin 1 particle, with $g_D = 3$. The fraction of deuterium is then determined by the Saha equation (2.29), using the same arguments that we saw in recombination

$$\frac{n_D}{n_n n_p} = \frac{3}{4} \left(\frac{m_D}{m_n m_p} \frac{2\pi\hbar^2}{k_B T} \right)^{3/2} e^{\beta E_{\text{bind}}}$$

Approximating $m_n \approx m_p \approx \frac{1}{2}m_D$ in the pre-factor, the ratio of deuterium to protons can be written as

$$\frac{n_D}{n_p} \approx \frac{3}{4} n_n \left(\frac{4\pi\hbar^2}{m_p k_B T} \right)^{3/2} e^{\beta E_{\text{bind}}}$$

We calculated the time-dependent neutron density n_n in (2.48). We will need this time-dependent expression soon, but for now it's sufficient to get a ballpark figure and, in this vein, we will simply approximate the number of neutrons as

$$n_n \approx n_p \approx \eta n_\gamma$$

The baryon-to-photon ratio has not had the opportunity to significantly change between nucleosynthesis and the present day, so we have $\eta \approx 10^{-9}$. (The last time it changed was when electrons and positrons annihilated, with $e^- + e^+ \rightarrow \gamma + \gamma$.) Using the expression $n_\gamma \approx (k_B T/c)^3$ from (2.15) for the number of photons, we then have

$$\frac{n_D}{n_p} \approx \eta \left(\frac{k_B T}{m_p c^2} \right)^{3/2} e^{\beta E_{\text{bind}}} \quad (2.49)$$

We see that we only get an appreciable number of deuterium atoms when the temperature drops to a suitably small value. This delay in deuterium formation is mostly due to the large number of photons as seen in the factor η . These same photons are responsible for the delay in hydrogen formation 300,000 years later: in both cases, any putative bound state is quickly broken apart as it is bombarded by high-energy photons at the tail end of the blackbody distribution.

Solving (2.49), we find that $n_D/n_p \sim 1$ only when $\beta E_{\text{bind}} \approx 35$, or

$$k_B T \lesssim 0.06 \text{ MeV}$$

Importantly, this is after the neutrinos have decoupled. Using (2.45), again with $g_\star \approx 3.4$, we find that deuterium begins to form at

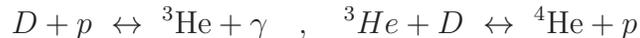
$$t \approx 360 \text{ seconds}$$

This is around six minutes after the Big Bang. Fortunately (for all of us), six minutes is not yet the 10.5 minutes that it takes neutrons to decay. But it's getting tight. Had the details been different so that, say, it took 12 minutes rather than 6 for deuterium to form, then we would not be around today to tell the tale. Building a universe is, it turns out, a delicate business.

Helium and Heavier Nuclei

Heavier nuclei have significantly larger binding energies. For example, the binding energy for ${}^3\text{He}$ is 7.7 MeV, while for ${}^4\text{He}$ it is 28 MeV. In perfect thermal equilibrium, these would be present in much larger abundancies. However, the densities are too low, and time too short, for these nuclei to form in reactions involving three or more nucleons coming together. Instead, they can only form in any significant levels after deuterium has formed. And, as we saw above, this takes some time. This is known as the *deuterium bottleneck*.

Once deuterium is present, however, there is no obstacle to forming helium. This happens almost instantaneously through



Because the binding energy is so much higher, all remaining neutrons rapidly bind into ${}^4\text{He}$ nuclei. At this point, we use the time-dependent form for the neutron density (2.48) which tells us that the number of remaining neutrons at this time is

$$\frac{n_n}{n_p} = \frac{1}{5} e^{-360/880} \approx 0.13$$

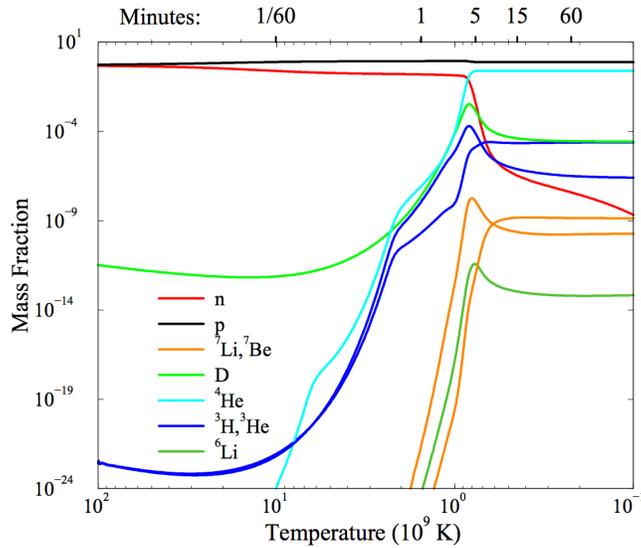


Figure 31: The abundance of light nuclei in the early universe.

Since each ${}^4\text{He}$ atom contains two neutrons, the ratio of helium to hydrogen is given by

$$\frac{n_{\text{He}}}{n_{\text{H}}} = \frac{n_n/2}{n_p - n_n} \approx 0.07$$

A helium atom is four times heavier than a hydrogen atom, which means that roughly 25% of the baryonic mass sits in helium, with the rest in hydrogen. This is close to the observed abundance.

Only trace amounts of heavier elements are created during Big Bang nucleosynthesis. For each proton, approximately 10^{-5} deuterium nuclei and 10^{-5} ${}^3\text{He}$ nuclei survive. Astrophysical calculations show that this is a million times greater than the amount that can be created in stars. There are even smaller amounts of ${}^7\text{Li}$ and ${}^7\text{Be}$, all in good agreement with observation.

The time dependence of the abundance of various elements is shown⁸ in Figure 31. You can see the red neutron curve start to drop off as the neutrons decay, and the abundance of the other elements rising as finally the deuterium bottleneck is overcome.

⁸This figure is taken from Burles, Nollett and Turner, *Big-Bang Nucleosynthesis: Linking Inner Space and Outer Space*, [astro-ph/99033](https://arxiv.org/abs/astro-ph/99033).

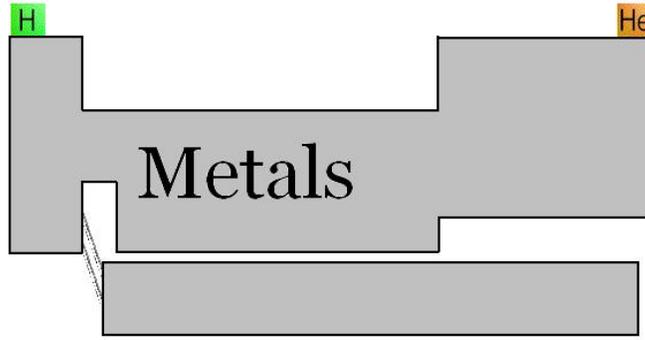


Figure 32: The elements, according to cosmologists.

Any heavier elements arise only much later in the evolution of the universe when they are forged in stars. Because of this, cosmologists have developed their own version of the periodic table, shown in the Figure 32. It is, in many ways, a significant improvement over the one adopted by atomic and condensed matter physicists.

Dependence on Cosmological Parameters

The agreement between the calculated and observed abundancies provides strong support for the seemingly outlandish idea that we know what we're talking about when the universe was only a few minutes old. The results depend in detail on a number of specific facts from both particle physics and nuclear physics.

One input into the calculation is particularly striking. The time at which deuterium finally forms is determined by the equation (2.45) which, in turn, depends on the number of relativistic species g_* . If there are more relativistic species in thermal equilibrium with the heat bath then the deuterium bottleneck is overcome sooner, resulting in a larger fraction of helium. Yet, the contribution from the light Standard Model degrees of freedom (i.e. the photon and neutrinos) gives excellent agreement with observation.

This puts strong constraints on the role of dark matter in the early universe. Given its current prominence, we might naively have thought that the relativistic energy density in the early universe would receive a significant contribution from dark matter. The success of Big Bang nucleosynthesis tells us that this is not the case. Either there are no light particles in the dark sector (so the dark sector is dark even if you live there) or hot dark particles fell out of equilibrium long before nucleosynthesis took place and so sit at a much lower temperature when the all action is happening.

2.5.4 Further Topics

There are many more stories to tell about the early universe. These lie beyond the scope of this course, but here is a taster. Going back in time, we have...

Electron-Positron Annihilation

Prior to nucleosynthesis, the fireball included both electrons and positrons. At around $k_B T \sim 1$ MeV these annihilate, injecting energy into the thermal bath of photons and slightly raising their temperature.

We can give an estimate for this. Prior to annihilation, the photons and electron-positron pairs were in equilibrium, giving

$$g_\star = 2 + \frac{7}{8}(2 + 2) = \frac{11}{2}$$

Afterwards, there are only photons with

$$g_\star = 2$$

So far, we haven't looked closely at what happens when g_\star changes. This is because we need an important concept that we haven't yet introduced: entropy. This is discussed in detail in the lectures on [Statistical Physics](#) where we show that the entropy of an ultra-relativistic gas is proportional to $g_\star T^3 a^3$.

The annihilation of electron-positron pairs is an *adiabatic process*, which means that entropy is conserved. Since g_\star decreases by a factor of 11/4, this means that $T^3 a^3$ increases by 11/4. Or

$$T_{\text{after}} = \left(\frac{11}{4}\right)^{1/3} T_{\text{before}}$$

There is one last twist to the story. The electrons and positrons do not all annihilate. There must be a very slight excess of electrons that are left over at the end. This, of course, is the stuff we're made of.

Before annihilation, the number of electron-positron pairs was the same order of magnitude as the number of photons, and these have persisted to the present day. Meanwhile, electric neutrality of the universe ensures that the number of left over electrons is comparable to the number of baryons currently in the universe. This means that the slight excess of electrons in the early universe must be roughly equal to the famous ratio $\eta \sim 10^{-9}$ of baryons to photons in the present day. In other words, in the early universe there was one extra electron for every billion electron-positron pairs. Understanding the origin of this imbalance is the goal of baryogenesis and is briefly described below.

Neutrino Decoupling

Neutrinos are very weakly interacting. They decouple from the thermal bath at temperatures of $T \sim 1$ MeV. Neutrinos have masses $m_\nu c^2$ between an meV and an eV (the exact masses are not well known) so they are very relativistic when they decouple. Like the photons after recombination, neutrinos preserve their relativistic distribution even after they decouple.

However, in contrast to the photons, neutrinos do not get the energy boost from electron-positron annihilation. This means that their temperature after this event lags behind the photon temperature, with

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

This relation persists to the present day. It is expected that there is a cosmic neutrino background filling the universe, with a temperature $T \approx (4/11)^{1/3} T_{\text{CMB}} \approx 1.9$ K. This has not yet been observed although there is an experiment, currently in the design phase, which aims to detect it.

When we have relativistic species that sit at different temperatures, we need to revisit our formula for the effective number of degrees of freedom g_\star . We can continue to write the total energy density as

$$\rho = g_\star \frac{\pi^2}{30 \hbar^3 c^3} (k_B T)^4$$

if we now define g_\star to be the sum over all relativistic species, whether or not they are in equilibrium,

$$g_\star(T) = \sum_{\text{bosons}} g_i \left(\frac{T_i}{T}\right)^4 + \frac{7}{8} \sum_{\text{fermions}} g_i \left(\frac{T_i}{T}\right)^4$$

where T_i is the temperature of each species. In particular, after e^-e^+ annihilation, when nucleosynthesis occurs, the relativistic species are photons and electrons with

$$g_\star = 2 + \frac{7}{8} \left[2 \times 3 \times \left(\frac{4}{11}\right)^4 \right] \approx 3.4$$

This is the value quoted in (2.47) and used when discussing nucleosynthesis.

QCD Phase Transition

At a temperature of $k_B T \approx 150$ MeV, protons and neutrons melt. They disassociate into a soup of quarks and gluons, known as the quark-gluon plasma. This state of matter has been created at particle accelerators here on Earth, freeing the quarks from their nucleon prison for the first time in 13.8 billion years.

Electroweak Phase Transition

In the Standard Model, fundamental particles such as the electron and quarks get their mass from the Higgs mechanism. Above $k_B T \approx 100$ GeV, this mechanism ceases to work. At this point all particles in the Standard Model are massless and in thermal equilibrium.

Dark Matter Freeze Out

Clearly there are many things we don't know about dark matter. We don't, for example, know if it has any interactions with the stuff we're made of. We can, however, make some assumptions and see where it leads us.

One of the most popular candidates for dark matter is a stable, massive particle that interacts only weakly with itself and with the Standard Model. These are known as weakly interacting massive particles, or WIMPs. Nearly all theories that go beyond the Standard Model predict such objects.

If the particle interacts even weakly with the Standard Model then there will be a time when dark matter is in equilibrium with the thermal bath. As the temperature lowers, the dark matter will freeze out. With very little input — just the mass and cross-section of the dark matter — we can then compute the expected abundance of dark matter seen today.

Here something nice happens. If we take the mass to be around $M_X \sim 100$ GeV, which is the energies probed by the LHC, and the cross-section to be $\sigma v \sim G_F$, which is the strength of the weak nuclear force, then we do indeed find the observed abundance of dark matter. With an overblown rhetorical flourish, this coincidence is known as the *WIMP miracle*. It was one of the reasons for optimism that it might be possible to create dark matter at the LHC. Needless to say, this was not borne out. Furthermore, a slew of impressive experiments, designed to directly detect passing dark matter, have so-far, offered only null results. While WIMPs remain a possible candidate for dark matter, there is no compelling observation beyond the coincidence above to suggest they are intimately tied to the weak force at the 100 GeV scale.

Baryogenesis

The universe contains lots of matter but very little anti-matter. How did this asymmetry come to be?

One possibility is that it is an initial condition on the universe. Another is that the universe started with equal amounts of matter and anti-matter, but somehow a small dynamical shift took place that preferred one over the other. This latter process is known as *baryogenesis*.

We don't have an established theory of baryogenesis; whatever caused it must lie beyond the Standard Model. Nonetheless, there are criteria, known as the *Sakharov conditions* that must be obeyed for baryogenesis to occur:

- The first criterion is the most obvious: baryon number cannot be a conserved quantity. Here “baryon number” refers to baryons minus anti-baryons. In a symmetric universe, this starts off as zero. We want it to end up non-zero.

In the Standard Model, baryon number is conserved. (In fact, strictly speaking $B - L$ is conserved where B is baryon number and L is lepton number, but this is a story for another day.) But it is straightforward to cook up interactions at higher energy scales which violate baryon number.

- There is a symmetry known as CP which, roughly speaking, says that particles and anti-particles behave the same. This too must be violated for baryogenesis to occur, since particles should be favoured over anti-particles.

In fact, CP is violated in the Standard Model. It's not clear if this is sufficient, or if further CP violation is needed in the interaction beyond the Standard Model.

- The final criterion is the least obvious: the early universe must deviate from thermal equilibrium. This is needed so that the interactions in one direction differ from the interactions running in reverse.

A deviation from thermal equilibrium occurs when the universe undergoes a first order phase transition. (You can read more about phase transitions in the lectures on [Statistical Physics](#) and [Statistical Field Theory](#).) The electroweak phase transition appears to be a fairly smooth crossover, which is not violent enough to do the job. For baryogenesis to occur, we most likely need a different phase transition early in the universe.

There are many models of baryogenesis, but currently no smoking gun experiment or observation to determine which, if any, is correct.