19 Particle astrophysics and inflation

In Chapter 18 we put forward a history of the universe. The picture is extremely simple. Its inputs were Einstein's equations and the assumptions of homogeneity and isotropy. We also used our knowledge of the laws of atomic, nuclear and particle physics. We saw a number of striking confirmations of this basic picture, but there are many puzzles.

1. The most fundamental problem is that we do not know the laws of physics relevant to temperatures greater than about 100 GeV. If there is only a single Higgs doublet at the weak scale, it is possible that we can extend this picture back to far earlier times. If there is, say, supersymmetry or large extra dimensions, the story could change drastically. Even if things are simple at the weak scale, we will not be able to extend the picture all the way back to $t = 0$. We have already seen that the classical gravity analysis breaks down.
2. There are a number of features of the *present* picture we cannot account for within the Standard Model. Specifically, what is dark matter? There is no candidate among the particles of the Standard Model. Is it some new kind of particle? As we will see, there are plausible candidates from the theoretical structures we have proposed, and they are the subject of intense experimental searches.
3. Dark energy is very mysterious. Assuming that it is a cosmological constant, it can be thought of as the vacuum energy of the underlying microphysical theory. As a number, it is totally bizarre. Its natural value should be set by the largest relevant scale, perhaps the Planck or unification scale, or the scale of supersymmetry breaking. Other proposals have been put forward to model dark energy. One which has been extensively investigated is known as quintessence, the possibility that the energy is that of a slowly varying scalar field. Such models typically do not predict $w \neq -1$ (see Section 18.1), and many are already ruled out by observation. But it should be stressed that these models are, if anything, less plausible than the possibility of a cosmological constant. First, one needs to explain why the underlying microphysical theory produces an essentially zero cosmological constant and a potential whose curvature is smaller than the present value of H. Then one needs to understand why the slowly varying field produces the energy density observed today, without disturbing the successes of the cosmological picture for earlier times. It is probably fair to say that no convincing explanation of either aspect of the problem has been forthcoming.
4. The value of the present baryon to photon ratio is puzzling:

$$\frac{n_B}{n_\gamma} = \left(6.1\,^{+0.3}_{-0.2}\right) \times 10^{-10}. \tag{19.1}$$

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

As we will see, the question can be phrased as why is this so small, or why is it so large? If the universe was always in thermal equilibrium, this number is a constant. So at very early times, there was a very tiny excess of particles over antiparticles. One might imagine that this is simply an initial condition but, as A. Sakharov first pointed out, this is a number that one might hope to explain through cosmology combined with microphysical theory. As we will discuss in detail later, it is necessary that the underlying microphysics violates baryon number and CP and that there is a significant departure from thermal equilibrium. The Standard Model, as we have seen, violates both and can generate a baryon number but, as we will see, it is far too small. So, *modifications of the known physical laws are required to account for the observed density of baryons.*

5. Homogeneity, flatness and topological objects such as monopoles pose puzzles which suggest a phenomenon known as *inflation*. Consider, first, homogeneity. This certainly made the equations simple to solve, but it is puzzling. If we look at the cosmic microwave background, the temperature variation in different directions in the sky is equal to about a part in 10^5. But, as we look out at distances as much as 13.8 billion light years away, points separated by a tiny fraction of a degree were separated, at 100 000 years after the big bang, by an enormous distance compared with the horizon at that time. The problem is that, as we look back, the horizon decreases in size as $\sqrt{t}$. So points separated by a degree were, at that time, separated by about 10^7 light years. But signals could not have traveled more than 10^5 light years by this time. So if these points had not been in causal contact by recombination, why should they have identical temperatures? For nucleosynthesis, which occurs much earlier, the question is even more dramatic.
6. Flatness ($\Omega_{\rm tot} = 1$) may not seem puzzling at first, but consider again the structure of the FRW metric. We have seen that the Friedmann equation can be recast as

$$\frac{8\pi G_{\rm N}\rho}{H^2} = \Omega - 1. \tag{19.2}$$

Suppose, for example, that $\Omega = 0.999$ today. Then, at recombination, the left-hand side of this equation was more than eight orders of magnitude smaller. So the energy density was equal to the critical density with extraordinary precision. This apparent fine tuning gets more and more extreme as we look further back in time.
7. Monopoles: we have seen that simple grand unified theories predict the existence of magnetic monopoles. Their masses are typically of order the grand unification scale. So unless their density were many orders of magnitude (perhaps 14!) smaller than the density of baryons, their total energy density would be far greater than the observed energy density of the universe. Astrophysical limits turn out to be even smaller; passing through the galaxy, monopoles would deplete the magnetic field. This sets a limit, known as the Parker bound, on the monopole flux in the galaxy:

$$\mathcal{F} < 10^{-16}\left(\frac{M_{\rm mon}}{10^{17}\,{\rm GeV}}\right)\,{\rm cm}^{-2}\,{\rm s}^{-1}\,{\rm sr}^{-1}. \tag{19.3}$$

However, we might expect, in a grand unified theory, quite extensive monopole production. We have seen that monopoles are topological objects. If there is a phase

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

transition between phases of broken and unbroken $SU(5)$, we would expect twists in the fields on scales of order the Hubble radius at this time, and a density of monopoles of order one per horizon volume. If the transition occurs at $T_0 = 10^{16}$ GeV, the Hubble radius is of order T^2/M_p so the density, in units of the photon density T^3, is of order

$$\frac{n_{\rm mon}}{n_\gamma} = \frac{T^3}{M_p^3} \tag{19.4}$$

and can be *larger* than the baryon density.

In the following sections we discuss these issues. We will study a possible solution to the homogeneity, flatness and monopole problems: inflation, the hypothesis that the universe underwent a period of extremely rapid expansion. We will see that there is some evidence that this phenomenon really occurred. Certainly there is nothing within the Standard Model itself which can give rise to inflation, so this points to the presence of some new phenomena, perhaps fields or perhaps more complicated entities, which are crucial to understanding the universe we see around us. We will describe some simple models of inflation, especially slow-roll inflation and chaotic and hybrid inflation, and some of their successes and the puzzles which they raise. We will discuss inflationary theory's biggest success, that quantum mechanical fluctuations during inflation give rise to the perturbations which grow to give the structure we see around us in the universe. This introduction is not comprehensive but should give the reader some tools to approach the vast literature which exists on this subject.

We next turn to the problem of dark matter. We focus on two candidates: the lightest supersymmetric particle of the MSSM, and the axion. We explain how these particles might rather naturally be produced with the observed energy density and discuss briefly the prospects for their direct detection. Then we turn to baryogenesis. We explain why the Standard Model has all the ingredients to produce an excess of baryons over antibaryons but, given the value of the Higgs mass, this baryon number cannot be nearly as large as is observed. We then turn to baryon production in some of our proposals for physics beyond the Standard Model, focusing on three possibilities: heavy particle decay in grand unified theories, leptogenesis and coherent production by scalar fields.

19.1 Inflation

The underlying idea behind inflation is that the universe behaved for a time as if (or nearly as if) the energy density was dominated by a positive cosmological constant, Λ. During this era the Friedmann equation was that for de Sitter space,

$$H_{\rm i}^2 = \left(\frac{\dot a}{a}\right)^2 = \frac{8\pi G_{\rm N}}{3}\Lambda, \tag{19.5}$$

with solution

$$a(t) = e^{H_{\rm i} t}. \tag{19.6}$$

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

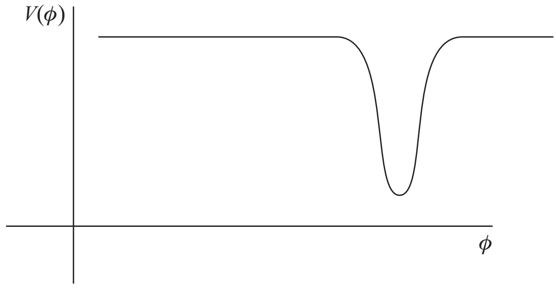

Fig. 19.1 A typical inflationary potential has a region in which $V(\phi)$ varies slowly and then settles into a minimum.

If this situation had held for a time interval such that, say, $\Delta t H_{\rm i} = 60$ then the universe would have expanded by an enormous factor. Suppose, for example, that Λ was 10^{16} GeV; correspondingly $H_{\rm i} \approx 10^{14}$ GeV. Then a patch of size H^{-1} would have grown to be almost a centimeter in size. If, at the end of this period of inflation, the temperature of the universe had been 10^{16} GeV, this patch would have grown, up to the present time, by a factor 10^{29}. This is about the size of our present horizon!

One possibility for how this might have come about is called *slow-roll inflation*. Here one has a scalar field ϕ with potential $V(\phi)$; $V(\phi)$, for some range of ϕ, is slowly varying (Fig. 19.1). What we have called $H_{\rm i}$ is determined by the average value of the potential in the plateau region, V_0. If we assume that we have a patch of initial size a bit larger than $H_{\rm i}^{-1}$, then we can write down an equation of motion for the zero-momentum mode of the field ϕ in this region:

$$g^{\mu\nu} D_\mu \partial_\nu \phi + V'(\phi) = 0. \tag{19.7}$$

Because of our assumptions of homogeneity and isotropy, we can take the metric to have the Robertson–Walker form:

$$\ddot{\phi} + 3h\dot{\phi} + V'(\phi) = 0. \tag{19.8}$$

We assume that the field is moving slowly, so that we can neglect the $\ddot{\phi}$ term. Shortly, we will check whether this assumption is self-consistent. In this limit the equation of motion is first order:

$$\dot{\phi} = -\frac{V'}{3H}. \tag{19.9}$$

We can integrate this equation to get Δt, the time it takes the field to traverse the plateau of the potential; $\Delta t H$ is roughly the number of e-folds of inflation, N, thus

$$N = \frac{1}{M_{\rm p}^2} \int d\phi \frac{V(\phi)}{V'(\phi)}. \tag{19.10}$$

The requirement for obtaining adequate inflation is: that N should be larger than about 60.

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

Now we can determine the conditions for the validity of the slow-roll approximation. We simply want to check, from our solution, that $\ddot{\phi} \ll 3H\dot{\phi}$ and $V'(\phi)$. Differentiating Eq. (19.9) leads to the conditions

$$\epsilon = \frac{1}{2}M_p^2\left(\frac{V'}{V}\right)^2 \ll 1 \tag{19.11}$$

and

$$\eta = M_p^2\frac{V''}{V}, \quad |\eta| \ll 1. \tag{19.12}$$

How did inflation end? Near the minimum of the potential we can approximate it as quadratic. So we might try to study an equation of the form

$$\ddot{\phi} + 3H\phi + m^2\phi = 0. \tag{19.13}$$

Were it not for the expansion of the universe, this equation would have a solution

$$\phi = \phi_0 \cos mt. \tag{19.14}$$

In quantum mechanical language this would describe a coherent state of particles, with energy density

$$\rho = \frac{1}{2}m^2\phi_0^2. \tag{19.15}$$

These particles have zero momentum; the pressure, $T_{ij} = p\delta_{ij} = 0$. So, if this field dominates the energy density of the universe, we know that

$$a \sim t^{2/3}, \quad H = \frac{2}{3}t. \tag{19.16}$$

In our toy model we might imagine that $m \sim 10^{16}$ GeV $\gg H$, so we could solve the equation by assuming

$$\phi(t) = f(t)\cos mt \tag{19.17}$$

for a slowly varying function f. Substituting Eqs. (9.1) and (19.16) into Eq. (19.13), one finds

$$f(t) = \frac{1}{t}. \tag{19.18}$$

Note that this means that

$$\rho = \rho_0\left(\frac{t}{t_0}\right)^2 = \rho_0\left(\frac{a}{a_0}\right)^3. \tag{19.19}$$

To summarize, we are describing a system which behaves like pressureless dust – zero-momentum particles – and which is diluted by the expansion of the universe.

This description also gives us a clue as to the fate of the field ϕ. Supposing that the ϕ particles have a finite width Γ, they will decay in time $1/\Gamma$. We can include this in our equation of motion, writing

$$\ddot{\phi} + (3H + \Gamma)\dot{\phi} + V'(\phi) = 0. \tag{19.20}$$

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

When the particles decay, if their decay products include, for example, ordinary quarks, leptons and gauge fields then their interactions will bring them quickly to equilibrium. We can be at least somewhat quantitative about this. The condensate disappears at a time set by $H \approx \Gamma$. If the universe quickly comes to equilibrium, the temperature must satisfy

$$\frac{\pi^2}{30} g T^4 = H^2 \frac{3}{8\pi G_{\rm N}}. \tag{19.21}$$

At this temperature we can estimate the rate of interaction. Since the typical particle energy will be of order T, the cross sections will be of order

$$\sigma = \frac{\alpha_i^2}{T^2} \tag{19.22}$$

We can multiply this by the density, $n = (\pi^2/30) g^* T^3$, to obtain a reaction rate. For inflation at the scales we are discussing, this is enormous compared with H. The details by which equilibrium is established have been studied with some care. We can imagine that when a ϕ particle first decays, it produces two very high energy particles. These will have rather small cross sections for scattering with other high-energy decay products, but these interactions degrade the energy, and so the cross sections for subsequent interactions – and for interactions with previously produced particles – are larger. A more careful study leads to a behavior with time where the temperature rises to a maximum and then falls. This maximum temperature is

$$T_{\rm max} \approx 0.8 g_*^{-1/4} m^{1/2} (\Gamma M_{\rm p})^{1/4}, \tag{19.23}$$

where m is the mass of the *inflaton*.

19.1.1 Fluctuations: the formation of structure

One of the most exciting features of inflation is that it predicts that the universe is not exactly homogeneous and isotropic. We cannot do justice to this subject in this short section, but we can at least give the flavor of the analysis and collect the crucial formulae. In order to have inflation we need the metric and fields to be reasonably uniform over a region of size $H_{\rm i}^{-1}$. But, because of quantum fluctuations, the fields and in particular the scalar field ϕ cannot be completely uniform. We can estimate the size of these quantum fluctuations without great difficulty. In order that inflation occurs at all, we need $m_\phi \ll H$. So we will treat ϕ as a massless free field in de Sitter space. As in flat space, we can expand the field ϕ in Fourier modes:

$$\phi(\vec{x}, t) = \int \frac{d^3k}{(2\pi)^3} \left[e^{i\vec{k}\cdot\vec{x}} h(\vec{k}, t) + \text{c.c.} \right]. \tag{19.24}$$

The expansion coefficients h obey the equation $D_\mu \partial_\nu h(\vec{k}, t) e^{i\vec{k}\cdot\vec{x}} = 0$, yielding, in the FRW background,

$$\ddot{h} + 3H\dot{h} + \frac{k^2}{a^2} h = 0. \tag{19.25}$$

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

Here k/a is the red-shifted momentum. In the case of de Sitter space, a grows exponentially rapidly. As soon as $k/a \sim H$ the system becomes overdamped, and the value of h is essentially frozen. We will see this in a moment when we write down an explicit solution of the equation.

It is convenient to change our choice of time variable. Rather than take the FRW form for the metric, we take a metric more symmetric between space and time:

$$ds^2 = a^2(t)(-d\eta^2 + d\vec{x}^2). \tag{19.26}$$

Here, in terms of our original variables,

$$\frac{d\eta}{dt} = \pm\frac{1}{a}. \tag{19.27}$$

So, choosing the + sign,

$$\eta = \int \frac{da}{(\dot{a}/a)a^2} = \int \frac{da}{Ha^2} \tag{19.28}$$

and

$$\eta = \frac{1}{H_a}. \tag{19.29}$$

In these coordinates, the equation of motion for $h(k, \eta)$ reads:

$$\delta\ddot{\phi} + 2aH\delta\dot{\phi} + k^2\delta\phi = 0. \tag{19.30}$$

This equation is straightforward to solve. The solution can be written in terms of Bessel functions, but more transparently as:

$$\delta\phi_k = \frac{e^{-ik\eta}}{\sqrt{ik}}\left(1 - \frac{i}{k\eta}\right). \tag{19.31}$$

Note that for large times $\eta \to 0$.

Further analysis is required to convert this expression into a fluctuation spectrum. The result is that the fluctuations in the energy density are roughly scale invariant, and

$$\frac{\delta\rho}{\rho} \approx \frac{H^2}{5\pi\dot{\phi}}. \tag{19.32}$$

Using the slow-roll equation

$$3H\dot{\phi} = V' \tag{19.33}$$

gives

$$\frac{\delta\rho}{\rho} = \frac{3H^*}{5\pi V'} = \frac{3V^{3/2}}{5\pi V' M_p^3}. \tag{19.34}$$

Much more detailed discussions of these formulas can be found in the suggested reading. These fluctuations quickly pass out of the horizon during inflation. While outside of the

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

horizon, they are frozen. Subsequently, however, they reenter the horizon and begin to grow. Measurements of the CMBR combined with Eq. (19.34) yield

$$\frac{V^{3/2}}{V'} = 5.15 \times 10^{-4} M_{\mathrm{p}}^{3} \tag{19.35}$$

on horizon scales. Fluctuations which were within the horizon at the time of matter–radiation equality have grown linearly with time since then. At about 1 billion years after the big bang they became non-linear, and this appears to account adequately for the observed structure in the universe. Precise studies of the CMBR, of the formation of structure and of Type Ia supernovas, as well as of the missing mass in structures on a wide range of scales, has allowed the determination of the composition of the universe.

Other observables of inflation are the spectral index n_s, which measures the deviation of the power spectrum from perfect scale invariance, and r, the ratio of the tensor and scalar fluctuations. The tensor modes arise due to gravitational radiation and are only observable if the scale of inflation is sufficiently high. These quantities, in slow-roll models, are given by

$$n_s = 1 + 2\eta - y\epsilon, \quad r = 16\epsilon. \tag{19.36}$$

The spectral index has been measured by the Planck collaboration as

$$n_s = 0.9624, \tag{19.37}$$

where the error is about 1%; r is not yet well known. If and when it is measured, it will determine the scale of inflation,

$$\Lambda_{\mathrm{inf}} = \left(\frac{r}{0.7}\right)^{1/4} \times (1.8 \times 10^{16})\ \mathrm{GeV}. \tag{19.38}$$

But, while the inflationary scenario is compelling and has significant observational support, we lack a persuasive microphysical understanding of these phenomena. This is undoubtedly one of the great challenges of theoretical physics. In the next section we describe various classes of models.

19.1.2 Models of inflation

Experiments of the last two decades, and especially WMAP and Planck, have provided strong support for the phenomenon of inflation. This is likely to be information about physics at extraordinarily high energy scales, well above those likely to be accessible to foreseeable accelerators. However, translating the data into a microscopic model is extremely challenging. There are almost as many models of inflation as there are physicists who have thought about the problem, and we cannot possibly sample them all; in this section we survey a few. No existing model is terribly compelling. Essentially, all must be tuned in order to obtain enough e-foldings of inflation and small enough $\delta\rho/\rho$. First, it is known from observations that the Hubble constant during inflation cannot have been much larger than 10^{16} GeV. This means that the scalar mass cannot be comparable to the Planck mass, so we face the usual problem of light scalars. In fact, the difficulties are more severe since the scalar mass must be much lighter than the Hubble scale of inflation, in order to ensure slow roll; as we will see, even with supersymmetry this requires percent-level tunings. Further tunings are typically required to obtain the required fluctuation spectrum.

19.1.2.1 Chaotic inflation

A particularly simple class of models yields what is known as *chaotic inflation*. These models illustrate both possible predictions and the issues of naturalness and tuning. An example is a theory with a single scalar field, with a simple potential

$$V = \frac{1}{2}m^2\phi^2. \tag{19.39}$$

Requiring 60 e-foldings of inflation gives

$$N = \frac{1}{4}\frac{\phi^2}{M_p^2} \tag{19.40}$$

Correspondingly, $\epsilon = 8.3 \times 10^{-3} = \eta$. One predicts then that the spectral index n_s is approximately 0.967, close to the value measured by Planck, and $r = 0.133$.

While the predictions are interesting, the model is hard to take seriously as a microscopic theory. In particular, solving Eq. (19.34) for m, one obtains $m^2 = 4.6 \times 10^{-12}M_p^2$. There is no symmetry which accounts for this; moreover, we require that the coefficients of all other powers of ϕ be extremely small as well. More generally, the fact that $\phi \gg M_p$ means that we have no control of the physics. Despite these concerns this model has proven useful for considering many aspects of inflation, and it has been argued that some of its features may characterize a larger class of models. Later, we will consider a possible setting for this idea without these problems.

19.1.2.2 Inflation with supersymmetry: hybrid inflation

Given that supersymmetry naturally produces light scalars, supersymmetry would seem a natural context in which to construct models of inflation. We have mentioned that in supersymmetric field theories and in string theory one often encounters moduli, i.e. scalar fields whose potentials vanish in the some limit. Banks suggested that, for such fields, a potential of the form

$$V = \mu^4 f(\phi/M_p) \tag{19.41}$$

will often arise. Here μ is an energy scale determined by some dynamical phenomenon such as the scale of supersymmetry breaking. For such a potential, assuming that typical field values are of order M_p we have from Eq. (19.34),

$$\frac{\delta\rho}{\rho} \approx \frac{\mu^2}{M_p^2}. \tag{19.42}$$

From this we have $\mu \approx 10^{15.5}$ GeV. The number of e-foldings is generically of order one; the potential must be tuned to the level of 1%, for example, if one is to obtain sufficient inflation. Still, this may seem less troubling than having many couplings less than 10^{-12}.

Note that μ, the energy scale in a supersymmetric model of this kind with a single field, is far larger than those we have considered previously for low-energy supersymmetry breaking. Banks proposed that, at the minimum of the potential, supersymmetry is unbroken with vanishing $\langle W \rangle$ as a result of an R symmetry.

Another class of models of some interest are known as hybrid models. These involve at least two fields. They are particularly interesting in the context of supersymmetry, where such models have been dubbed "supernatural" by Guth and Randall since the presence of light scalars is again natural.

Hybrid inflation is often described in terms of fields and potentials with rather detailed, special, features but it can be characterized in a conceptual way. Inflation occurs in all such models on a pseudomoduli space, in a region where supersymmetry is badly broken (possibly by a larger amount than in the present universe) and on which the potential is slowly varying. We have seen that moduli are common in supersymmetric theories. We will find that they are ubiquitous in string theory. The simplest (supersymmetric) hybrid model has two fields, I and ϕ:

$$W = I(\kappa\phi^2 - \mu^2). \tag{19.43}$$

The field ϕ is known as the waterfall field. Classically, for large I the potential is independent of I,

$$V_{\rm cl} = \mu^4 \quad (\phi = 0).$$

This is the regime of inflation. Quantum mechanical effects generate a potential for I such that it rolls slowly from larger to smaller values. Inflation ends either when the slow-roll conditions are not satisfied or when I is small enough that the ϕ curvature is negative. In any case, at this point ϕ moves quickly towards its minimum. As it oscillates about the minimum, reheating occurs.

The quantum mechanical corrections control the dynamics of the inflaton. These involve a Coleman–Weinberg calculation of a type with which we are now familiar:

$$V(I) = \mu^4 \left(1 + \frac{\kappa^2}{16\pi^2} \log \frac{|I|^2}{\mu^2}\right). \tag{19.44}$$

Here κ is constrained to be extremely small in order that the fluctuation spectrum be of the correct size; κ is proportional, in fact, to V_I, the energy during inflation. The quantum corrections determine the slow-roll parameters. We have

$$\kappa = 0.17 \times \left(\frac{\mu}{10^{15}\,\text{GeV}}\right)^2 = 7.1 \times 10^5 \times \left(\frac{\mu}{M_{\rm p}}\right)^2. \tag{19.45}$$

The problem of fine tuning in these models can be readily characterized. For example, Planck-suppressed terms in the Kahler potential K can spoil slow roll;

$$K = \frac{\alpha}{M_{\rm p}^2} I^\dagger I I^\dagger I \tag{19.46}$$

gives too large an η value unless $\alpha \sim 10^{-2}$. This is an irreducible tuning of (supersymmetric) hybrid models. However, the very small value required of κ is arguably a more

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

severe tuning issue. In any case the model as it stands predicts $n_s > 1$ and is ruled out by the results from the Planck satellite, Eq. (19.3). Modifications are possible which avoid this prediction. Indeed, the moduli space of the simplest model does not closely resemble those we will encounter in string theory; broadening these considerations leads to different possibilities.

19.1.3 Constraints on reheating: the gravitino problem

In the context of supersymmetric theories, it is thought that there may be an upper bound on the reheating temperature. This is the problem of producing too many gravitinos. The gravitino lifetime is quite long,

$$\Gamma_{3/2} \approx \frac{m_{3/2}^3}{M_{\rm p}^2}; \tag{19.47}$$

gravitinos might even be stable. As a minimal requirement we need to suppose that gravitinos did not dominate the energy density at the time of nucleosynthesis. Otherwise the expansion rate at the time of nucleosynthesis is not consistent with the observed abundances of the light elements but, even more dramatically, their decay products would break up He^4 and other nuclei. Even though gravitinos are very weakly interacting, there is a danger that they would be overproduced during the period of reheating that follows inflation. A natural estimate for their production rate per unit volume is obtained by assuming that they are produced in two-body scattering, by light particles with densities of order T^3, and that their cross sections behave as $1/M_{\rm p}^2$:

$$n^2\langle\sigma v\rangle \approx T^6\langle\sigma v\rangle \approx \frac{T^6}{M_{\rm p}^2}. \tag{19.48}$$

Integrating this over a Hubble time $M_{\rm p}/T^2$ and dividing by the photon density, of order T^3, gives a rough estimate:

$$\frac{n_{3/2}}{s} \sim \frac{T}{M_{\rm p}}. \tag{19.49}$$

Assuming 1 TeV for the gravitino mass, the requirement that gravitinos do not dominate before nucleosynthesis gives $T < 10^{12}$ GeV. But this is too crude. Considering the destruction of deuterium and lithium gives $T < 10^9$ GeV or possibly a much smaller value. This is a strong constraint on the nature of reheating after inflation, but it is not a problem for the low-scale hybrid models we discussed in the previous subsection.

19.2 The axion as the dark matter

Within the set of ideas we have discussed for physics beyond the Standard Model, there are two promising candidates for dark matter. One is the axion, which we discussed in Chapter 5 as a possible solution to the strong CP problem. A second is the lightest supersymmetric

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

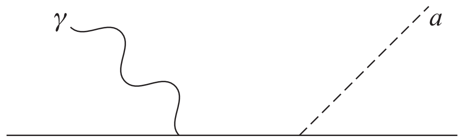

Fig. 19.2 In a Bremsstrahlung-like process, a lepton or nucleon can emit an axion when struck by a photon.

particle in models with an unbroken R-parity. We first discuss the axion, mainly because the theory is particularly simple. To begin, we need to consider the astrophysics of the axion a little further. There is a lower bound on the axion decay constant or, equivalently, an upper bound on its mass, arising from processes in stars. Axions can be produced by collisions deep within a star. Then, because of their small cross section, most axions will escape, carrying off energy. This has the potential to disrupt the star. We can set a limit by requiring that the flux of energy from the stars is not more than a modest fraction of the total energy flux.

To estimate these effects, we can first ask what sorts of processes might be problematic. A pair of photons can collide and produce an axion (using the $aF\tilde{F}$ coupling of the axion to the photon). Axions can be produced from nuclei or electrons in Compton-like and Bremsstrahlung processes (Fig. 19.2). The typical energies will be of order T.

For the Compton-like process of Fig. 19.2, the cross section is of order given by

$$\sigma_a \approx \frac{\alpha}{f_a^2}. \tag{19.50}$$

The total rate per unit time for a given electron to scatter off a photon in this way will be proportional to the photon density, which we will simply approximate as T^3. To obtain the total emission per unit volume we need to multiply, as well, by the electron density in the star. In the Sun this number is of order the total number of protons or electrons, 1.16×10^{57}, divided by the cube of the solar radius (in particle physics units, 3.5×10^{25} GeV^{-1}). This corresponds to

$$n_e \approx 3 \times 10^{-16}\ (\text{GeV})^3 \text{ electrons.} \tag{19.51}$$

Rather than calculate the absolute rate, we will compare it with the rate for neutrino production. We would expect that if axions carry off far more energy than neutrinos, this would be problematic. For neutrino production we might take n_e^2 and multiply by a typical weak cross section:

$$\sigma_\nu = G_\text{F}^2 E_\nu^2. \tag{19.52}$$

where E_ν is a typical neutrino energy. Finally, we take the temperature in the core of the star to be of order 1 MeV. Taking $f_a = 10^9$ gives, for the axion production rate,

$$R_\text{a} = 10^{-47}\ \text{GeV}^{-4} \tag{19.53}$$

while

$$R_\nu = 10^{-47}\ \text{GeV}^{-4}. \tag{19.54}$$

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

Clearly this analysis is crude; much more care is required in enumerating the different processes, evaluating their cross sections and integrating over particle momentum distributions. But this rough calculation indicates that 10^9 GeV is a plausible lower limit on the axion decay constant.

So, we have a lower bound on the axion decay constant. An upper bound arises from cosmology. Suppose that the Peccei–Quinn symmetry breaks before inflation. Then, throughout what will become the observable universe, the axion field is essentially constant. But, at early times the axion potential is negligible. To be more precise, consider the equation of motion of the axion field:

$$\ddot{a} + 3H\dot{a} + V'(a) = 0. \tag{19.55}$$

At very early times $H \gg m$ and the system is overdamped. The axion simply does not move. If the universe were very hot, the axion mass would actually have been much smaller than its current value. This is explained in Appendix C, but is easy to understand: at very high temperatures, the leading contribution in QCD to the axion potential comes from instantons. Instanton corrections are suppressed by $\exp\left[\frac{-8\pi^2}{g^2(T)}\right] = (\Lambda/T)^{b_0}$. They are also suppressed by powers of the quark masses. In other words, they behave as

$$V(a) = \prod_f m_f \Lambda^{b_0} T^{-b_0+n_f-4} \cos\theta, \tag{19.56}$$

where $\theta = a/f_a$ and n_f is the number of flavors with mass $\ll T$. This goes very rapidly to zero at temperatures above the QCD scale.

So the value of the axion field – the θ-angle – at early times, is most likely to be simply a random variable. Let us consider, then, the subsequent evolution of the system. The equation of motion for such a scalar field in an FRW background is by now quite familiar:

$$\ddot{a} + H\dot{a} + V'(a) = 0. \tag{19.57}$$

The potential $V(a)$ also depends on $T(t)$, which complicates the solution slightly, so let us first solve the problem with just the zero-temperature axion potential. In this case, the axion will start to oscillate when $H \sim m_a$. After this, the axions on the one hand dilute like matter, i.e. as $1/a^3$. The energy in radiation, on the other hand, dilutes like $a^4 \propto T^{-4}$. Assuming radiation domination when the axion starts to oscillate, we can determine the temperature at that time. Using our standard formula for the energy density,

$$\rho = \frac{\pi^2}{30} g^* T^4, \tag{19.58}$$

we have, just above the QCD phase transition, $g^* \approx 48$ (with the gluons, three quark flavors, three light neutrinos and the photon). Just below it we do not have the quarks or gluons but we should include the pions, so $g^* \approx 30$. Taking the larger value,

$$T_a = 10^2 \text{ GeV} \left(\frac{10^{11}\text{GeV}}{f_a}\right)^{1/2}. \tag{19.59}$$

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

At this time, the fraction of the energy density in axions is approximately

$$\frac{\rho_a}{\rho} = \frac{\frac{1}{2} f_a^2 m_a^2}{\rho} \approx \frac{1}{6} \frac{f_a^2}{M_p^2}. \tag{19.60}$$

So, if $f_a = 10^{11}$ GeV, axions come to dominate the energy density quite late, at $T \approx 10^{-3}$ eV. The temperature of axion domination scales with f_a, so 10^{16} GeV axions would dominate the energy density at 100 eV, which would be problematic.

However, the axion potential, as we have seen, is highly suppressed at temperatures above a few hundred MeV. So oscillation, sets in much later, in fact. We can make another crude estimate by simply supposing that the axion potential turns on at $T = 100$ TeV. In this case the axion fraction is large, of order $1/g^*$. So, if the axion density is to be compatible with the observed dark matter density for any value of f_a, we need to allow for the detailed temperature dependence of the axion mass. Using our formula for the axion potential as a function of temperature we can ask when the associated mass becomes of order the Hubble constant. After that time, axion oscillations are more rapid than the Hubble expansion so we might expect that the axion density will damp, subsequently, like matter. Let us take, specifically, $f_a = 10^{11}$ GeV. For the axion mass we can take

$$m_a(T) \approx 0.1 m_a(T = 0) \left(\frac{\Lambda_{\rm QCD}}{T} \right)^{3.7}. \tag{19.61}$$

The axion then starts to oscillate when $T \approx 1.5$ GeV. At this time, axions represent about 10^{-9} of the energy density. One needs to do a bit more work to show that, in the subsequent evolution, the energy in axions relative to the energy in radiation falls roughly as $1/T$ but, for this decay constant, the axion and radiation energies become equal at roughly 1 eV. If the decay constant is significantly higher than 10^{11} GeV then the axions start to oscillate too late and dominate the energy density too early. If the decay constant is significantly smaller then the axions cannot constitute the presently observed dark matter.

So, on the one hand it is remarkable that there is a rather narrow range of axion decay constants that are consistent with observation. On the other hand, some assumptions that we have made in this section are open to question. In particular, as we will see when we discuss the problem of moduli in cosmology, there are reasons to suspect that the universe may never have been hotter than tens of MeV. In this case the upper limit on the axion decay constant, as we will discuss further below, can be much weaker.

19.3 The LSP as the dark matter

Stability is one criterion for a dark matter candidate; a suitable production rate is another. We can make a crude calculation, which indicates that with susy breaking in the TeV range the LSP density is in a suitable range for the LSP to be the dark matter. Consider particles X with mass of order 100 GeV and interacting with weak interaction strength.

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

Their annihilation and production cross sections go as $G_F^2 E^2$. So, in the early universe, the corresponding interaction rate is of order

$$\Gamma \approx \rho_X G_F^2 E^2 \approx \rho_X G_F^2 T^2. \tag{19.62}$$

These interactions will drop out of equilibrium when the mass of the particle X is small compared with the temperature, so that there is a large Boltzmann suppression of their production. This will occur when this rate is of order the expansion rate, or

$$T^3 e^{-M_X/T} \langle v\sigma \rangle \sim \frac{T^2}{M_p}. \tag{19.63}$$

Since the exponent is very small, once $T \sim 10 M_X$ we can get a rough estimate of the density by saying that

$$e^{-M_X/T} T^3 \sim \frac{G_F^{-2}}{M_p}. \tag{19.64}$$

The ratio of the X particle density and the total entropy, n_X/s, is then given by

$$\frac{n_X}{s} \approx \frac{G_F^{-2}}{(M_p T^3)}. \tag{19.65}$$

Assuming that $M_X \sim 100$ GeV and $T \sim 10$ GeV, this gives about 10^{-9} for the right-hand side. Since the energy density in radiation damps as T^{-4} while that for matter damps as T^{-3}, this gives matter–radiation equality at temperatures of order an electronvolt, as in the standard big bang cosmology.

Needless to say, this calculation is quite crude. Extensive, and far more sophisticated calculations have been done to find the regions of parameter space in different supersymmetric models which are compatible with the observed dark matter density. The basic starting point for these analyses is the Boltzmann equation. If the basic process is of the form $1 + 2 \leftrightarrow 3 + 4$ then

$$\begin{aligned} a^{-3} \frac{d}{dt}(n_1 a^3) = & \int \frac{d^3 p_1}{(2\pi)^3 2E_1} \int \frac{d^3 p_2}{(2\pi)^3 2E_2} \int \frac{d^3 p_3}{(2\pi)^3 2E_3} \int \frac{d^3 p_4}{(2\pi)^3 2E_4} \\ & \times [f_3 f_4 (1 \pm f_1)(1 \pm f_2) - f_1 f_2 (1 \pm f_3)(1 \pm f_4)] \\ & \times (2\pi)^4 \delta^4 (p_1 + p_2 - p_3 - p_4) |\mathcal{M}|^2, \end{aligned} \tag{19.66}$$

where $\mathcal{M}$ is the invariant matrix element for the scattering process under consideration. The functions $f_1, \ldots, f_4$ are the distribution functions for the different species. These equations can be simplified in the high-temperature limit using Boltzmann statistics:

$$f(E) \rightarrow e^{\mu/T} e^{-E/T}. \tag{19.67}$$

Interactions are still fast enough at this time to maintain the equilibrium of the X momentum distributions (i.e. kinetic equilibrium) but not that of the X number. So it is the limiting

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

value of the X chemical potential, μ_X, which we seek. In this limit, we have

$$f_3 f_4(1 \pm f_1)(1 \pm f_2) - f_1 f_2(1 \pm f_3)(1 \pm f_4)$$
$$\to e^{-(E_1+E_2)/T}\left(e^{(\mu_1+\mu_2)/T} - e^{(\mu_3+\mu_4)/T}\right). \quad (19.68)$$

Here we have used $E_1 + E_2 = E_3 + E_4$.

Things simplify further since all but the X particle (particle 1) are light and are nearly in equilibrium. Defining $n_i^{(0)}$ as the distributions in the absence of a chemical potential, and defining the thermally averaged cross section

$$\langle \sigma v \rangle = \frac{1}{n_1^{(0)} n_2^{(0)}} \int \frac{d^3p_1}{(2\pi)^3 2E_1} \int \frac{d^3p_2}{(2\pi)^3 2E_2} \int \frac{d^3p_3}{(2\pi)^3 2E_3} \int \frac{d^3p_4}{(2\pi)^3 2E_4} |\mathcal{M}|^2, \quad (19.69)$$

we have

$$a^{-3}\frac{d(n_x a^3)}{dt} = n_X^{(0)} n_2 \langle \sigma v \rangle \left(1 - \frac{n_X}{n_X^{(0)}}\right). \quad (19.70)$$

Detailed solutions of these equations (often without some of these simplifications) reveal, as one would expect, a range of parameters in the MSSM that are compatible with the observed dark matter density (Fig. 19.3).

So, while it is disturbing that we need to impose additional symmetries in the MSSM in order to avoid proton decay, it is also exciting that this leads to a possible solution of one of the most critical problems of cosmology: the identity of the dark matter.

Having contemplated stable, weakly interacting particles as the dark matter, it is clear that this is a possibility that one can consider without invoking supersymmetry. One can

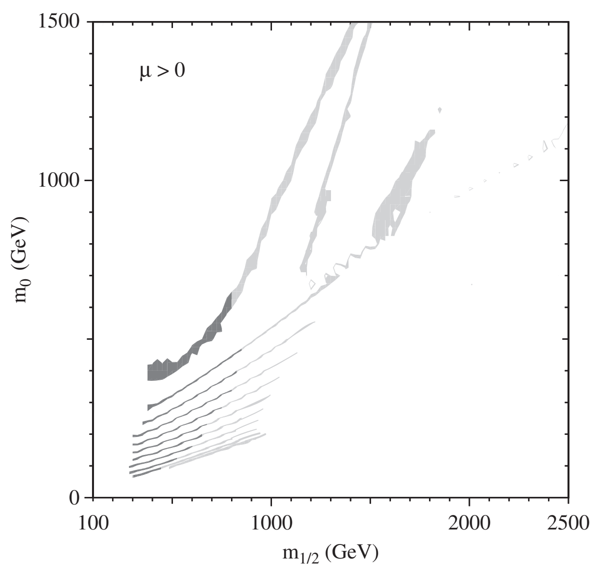

Fig. 19.3 Reprinted from J. R. Ellis *et al.*, Supersymmetric dark matter, *Phys. Lett. B*, **565**, 176 (2003), Figure 9. Copyright 2003, with permission from Elsevier.

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

simply postulate the existence of a massive stable particle with weak interactions with Standard Model fields. The fact that such a particle automatically leads to more or less the observed dark matter density is referred to as the "wimp miracle". One can also suppose that this particle has interactions with other particles, possibly lighter than it. Indeed, partly in response to potential signals, physicists have explored a broad range of possibilities.

19.3.1 The search for wimp dark matter

There is a variety of strategies that one can contemplate to search for weakly interacting masssive particle (*wimp*) dark matter, and this has become a major area of experimental and theoretical activity. There is no space here for an extensive review, so we will just mention the main strategies.

1. *Direct detection of dark matter* Here one searches for the scattering of dark matter particles off a target. Typically the targets are heavy nuclei, and one searches for the energy transferred to the recoiling nucleus. Such experiments must be conducted deep underground. Detectors must be sensitive to tiny energy depositions.
2. *Indirect detection of dark matter* Here one looks for the annihilation of dark matter particles against each other with the production of pairs of photons or neutrinos, for example. The galactic center, which is believed to contain a high concentration of dark matter particles, is a particularly interesting potential source for such events. A variety of experiments, particularly involving satellites such as PAMELA, Fermi and AMS, have been engaged in such searches.
3. *Accelerator searches* Many models of dark matter predict observable accelerator signals. Clearly direct observation in accelerators, complemented by discovery in either direct or indirect searches, would have the potential to provide a convincing discovery.

Among direct and indirect searches, significant exclusions have been achieved, There have also been tantalizing hints of possible signals.

19.4 The moduli problem

We have seen that in supersymmetric theories there are frequently light moduli (we have invoked this idea in our discussion of hybrid inflation). In string models we will find that such fields are ubiquitous. Such moduli, if they exist, pose a cosmological problem with some resemblance to the problems of axion cosmology.

In this section we will formulate the problem as it arises in gravity-mediated supersymmetry breaking. The potential for a modulus ϕ would be expected to take the form

$$V(\phi) = m_{3/2}^2 M_{\rm p}^2 f(\phi/M_{\rm p}). \tag{19.71}$$

By assumption f has a minimum at some value ϕ of order $M_{\rm p}$. In the early universe, when the Hubble parameter is much greater than $m_{3/2}$, this potential is effectively quite small and there is in general no obvious reason why the field should sit at its minimum. So, when

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

$H \sim m_{3/2}$, the field is likely to lie at a distance M_{p} in field space from the minimum and to store an energy of order $m_{3/2}^2 M_{\text{p}}^2$. Like the axion, after this time, assuming it is within the domain of attraction of the minimum, it will oscillate, behaving like pressureless dust. Almost immediately, given our assumptions about scales, it comes to dominate the energy density of the universe and continues to do so until it decays. The problem is that the decay occurs quite late and the temperature after the decay is likely to be quite low.

We can estimate this temperature after ϕ decay, T_{r}, by first considering the lifetime of the ϕ particle. We might expect this to be

$$\Gamma_\phi = \frac{m_{3/2}^3}{M_{\text{p}}^2}, \tag{19.72}$$

assuming that the couplings of the ϕ field to other light fields are suppressed by a single power of M_{p}. Assuming that the decay products quickly thermalize, and noting that Γ_ϕ is the Hubble constant at the time of ϕ decay, gives

$$T_{\text{r}}^4 \approx \frac{m_{3/2}^6}{M_{\text{p}}^2} \approx (10 \text{ keV})^4. \tag{19.73}$$

Here we are assuming that $m_{3/2} \approx 1$ TeV. This is a temperature well below the temperature at which nucleosynthesis occurs. So, in such a picture, the universe is matter dominated during nucleosynthesis. But the situation is actually far worse: the decay products almost certainly destroy deuterium and the other light nuclei.

Two plausible resolutions for this puzzle have been put forward. The first is the obvious one, that there may simply be no moduli. Related to this, it is possible that, at the minimum of their potential, all the moduli may be charged under unbroken symmetries; these might be new discrete symmetries, for example beyond those of the MSSM. Furthermore, they may be much more strongly interacting than suggested above. In models with some degree of low-energy supersymmetry, there is a problem with this proposal. Assuming that the strong CP problem is solved by an axion, this field is accompanied by another scalar. This scalar must acquire mass in a supersymmetry-violating fashion, otherwise it would be quite heavy. Conceivably, of course, either there is no axion or supersymmetry is broken at an extremely high energy scale.

Alternatively, the moduli might be significantly more massive than 1 TeV. Note that T_{r} scales, like the moduli masses, to the $3/2$ power, so if the moduli masses are of order 30 - TeV or more, this temperature could be sufficiently high (10 MeV) that nucleosynthesis occurs (again).

Such a scenario raises interesting questions. First, one could well imagine that one or more of these moduli play the role of the inflaton. In this case the reheating temperature would be much lower than usually contemplated. Indeed, in effect the universe was never very hot. The conventional picture of the thermal production of dark matter cannot be operative. Even if the late-decaying moduli are not connected with inflation, these decays will dilute whatever dark matter might have been produced earlier. This dilution factor can easily be a factor of $10^9 - 10^{12}$. Any baryon number produced before these decays is also diluted by this factor. One can hope that the baryons are produced in the decays of these moduli, but this requires one to understand why such low-energy baryon-number violation does not cause difficulties for proton decay. In the rest of this section, we will consider non-thermal mechanisms to produce the dark matter; in the next section we will discuss possible mechanisms to produce the baryon asymmetry, and we will see that there is one mechanism which is capable of producing a large enough asymmetry to survive moduli decays.

19.4.1 The axion as dark matter again

If moduli dominated the energy density of the universe for some period, then the cosmological constraints on the axion mass and decay constant are appreciably modified. These can be formulated quite simply. If the axion initially has amplitude f_a then, when the axion begins to oscillate and decay, at $H \approx m_a$, the fraction of the energy density stored in axions is of order f_a^2/M_{p}^2. If, when the moduli decay, they reheat the universe to 10 MeV, the ratio of axions (dark matter) to radiation is

$$r_a = \frac{f_a^2}{M_{\mathrm{p}}^2} \frac{10\,\mathrm{MeV}}{T}. \tag{19.74}$$

In order that this fraction be of order unity only when the temperature is of order 1 eV, we require $f_a < 10^{14.5}$ GeV. This is close to, say, the unification scale.

19.4.2 Moduli and wimp dark matter

As we have noted, moduli domination followed by reheating to nucleosynthesis temperatures does not permit the usual thermal production of wimps. One possibility which has been widely considered is that dark matter might be produced in moduli decays. The problem with this is that typically dark matter is then overproduced. In an approximately supersymmetric limit, moduli decays to particles and their superpartners have equal branching ratios. This means that, when the moduli decay, an order-one fraction goes into each accessible state. If the LSP is one of these decay products, it will likely be overproduced (typically, subsequent annihilations are not strong enough to avoid this difficulty). There may be special ranges of parameters where dark matter production in this way is possible; alternatively, one might argue that a picture with moduli favors axions, or some other coherently produced particle, as the dark matter.

19.5 Baryogenesis

The baryon to photon ratio, n_{B}/n_γ, is quite small. At early times, when QCD was nearly a free theory, this slight excess would have been extremely unimportant. But, for the structure of our present universe, it is terribly important. One might imagine that n_{B}/n_γ is simply an initial condition, but it would be more satisfying if we could have some microphysical understanding of this asymmetry between matter and antimatter. Andrei

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

Sakharov, after the experimental discovery of CP violation, was the first to state precisely the conditions under which the laws of physics could lead to a prediction for the asymmetry.

1. *The underlying laws must violate baryon number* This condition is obvious; if there is, for example, no net baryon number initially, and if baryon number is conserved, the baryon number will always be zero.
2. *The laws of nature must violate CP* Otherwise, for every particle produced in interactions, an antiparticle will be produced as well.
3. *The universe, in its history, must have experienced a departure from thermal equilibrium* Otherwise, the CPT theorem ensures that the numbers of baryons and of antibaryons at equilibrium are zero. This can be proven with various levels of rigor, but one way to understand it is to observe that CPT ensures that the masses of the baryons and antibaryons are identical, so at equilibrium their distributions should be the same.

Subsequently, there have been many proposals for how the asymmetry might arise. In the next sections, we will describe several. *Leptogenesis* relies on lepton-number violation, something we know is true in nature but of whose underlying microphysics we are ignorant. *Baryogenesis through coherent scalar fields* (Affleck–Dine baryogenesis) also seems plausible. It is only operative if supersymmetry is unbroken up to comparatively low energies, but it can operate quite late in the evolution of the universe and can be extremely efficient. This could be important in situations like moduli decay or hybrid inflation where the entropy of the universe is produced very late, after the baryon number.

19.5.1 Baryogenesis through heavy particle decays

One well-motivated framework in which to consider baryogenesis is grand unification. Here one can satisfy all the requirements for baryogenesis. Baryon-number violation is one of the hallmarks of GUTs, and these models possess various sources of CP violation. As far as departure from equilibrium is concerned, the decays of massive gauge bosons X provide good candidates for a mechanism. To understand in a little more detail how the asymmetry can come about, note that CPT requires that the total decay rate of X is the same as that of its antiparticle $\bar{X}$. But it does not require equality of the number of decays to particular final states (partial widths). So, starting with equal numbers of X and $\bar{X}$ particles, there can be a slight asymmetry between processes such as

$$X \to d + e^-, \quad X \to d + u^c \tag{19.75}$$

and

$$\bar{X} \to d^c + e^+, \quad \bar{X} \to d^c + u, \tag{19.76}$$

where the superscript c denotes an antiparticle. The tree graphs for these processes are necessarily equal; any CP-violating phase simply cancels out when we take the absolute square of the amplitude (see Fig. 19.4). This is not true in higher order, where additional phases associated with real intermediate states can appear. Actually computing the baryon asymmetry requires an analysis of the Boltzmann equations, of the kind that we have encountered in our discussion of dark matter.

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

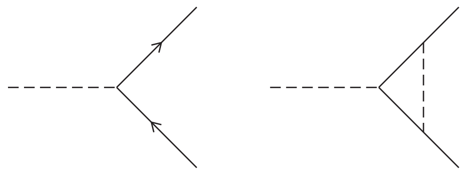

Fig. 19.4 Tree and loop diagrams whose interference can lead to an asymmetry in heavy particle decay.

There are reasons to believe, however, that GUT baryogenesis is not the origin of the observed baryon asymmetry. Perhaps the most compelling of these has to do with inflation. Assuming that there was a period of inflation, any pre-existing baryon number was greatly diluted by this. So, in order that one produces baryons through X boson decay, it is necessary that the reheating temperature after inflation be at least comparable with the X boson mass; but, we have seen that the scale of inflation is constrained to be less than 10^{16} GeV so we would require very efficient conversion of the energy density during inflation to radiation for this mechanism to be operative. Also, as we have explained, in supersymmetric theories a reheating temperature greater than 10^9 GeV leads to the overproduction of gravitinos.

19.5.2 Electroweak baryogenesis

The Standard Model, for some range of parameters, can satisfy all the conditions for baryogenesis. We saw in our discussion of instantons that the Standard Model violates baryon number. This violation is extremely small at low temperatures, so small that it is unlikely that in the history of the universe a single baryon has decayed in this way. The rate is so small because baryon-number violation is a tunneling process. If one could excite the system to high energies, one might expect that the rate would be enhanced. At high enough energies the system might simply be above the barrier. One can find the configuration which corresponds to sitting on top of the barrier by looking for static but unstable solutions of the equations of motion. Such a solution is known. It is called a *sphaleron* (from the Greek, meaning "ready to fall"). The barrier is quite high – from familiar scaling arguments, the sphaleron energy is of order $E_{\rm sp} = 1/(\alpha M_W)$. But this configuration is large compared with its energy; its size is of order M_W. As a result, it is difficult to produce in high-energy scattering. Two particles with enough energy to produce the sphaleron need to have momenta much higher than M_W. As a result, their overlap with the sphaleron configuration is exponentially suppressed.

At high temperatures one might expect that the sphaleron rate would be controlled by a Boltzmann factor, $e^{-E_{\rm sp}/T}$. Then, as the temperature increases, the rate would grow significantly. This turns out to be the case. In fact the rate is even larger than one might expect from this estimate, because $E_{\rm sp}$ itself is a function of T. At very high temperatures the rate has no exponential suppression at all and behaves as:

$$\Gamma = (\alpha_w T)^4. \tag{19.77}$$

These phenomena are discussed in Appendix C.

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

If the Higgs mass is not too large, the Standard Model can produce a significant departure from equilibrium. As the temperature rises, a simple calculation, described in Appendix C, shows that the Higgs mass increases (the mass-squared value becomes less negative) with temperature. At very high temperatures, the $SU(2) \times U(1)$ symmetry is restored. The phase transition between these two phases, for a sufficiently light Higgs, is first order. It proceeds by the formation of bubbles of the unbroken phase. The surfaces of these bubbles can be sites for baryon number production. These phenomena are also discussed in Appendix C. So, the third of Sakharov's conditions can be satisfied.

Finally, we know that the Standard Model violates CP. We also know, however, that it is crucial that there are three generations and that this CP violation vanishes if any quark masses are zero. As a result, even if the Higgs mass is small enough that the transition is strongly first order, any baryon number produced is suppressed by several powers of Yukawa couplings and is far too small to account for the observed matter–antimatter asymmetry.

In the MSSM the situation is somewhat better. There is a larger region of the parameter space in which the transition is first order and, as we have seen, there are many new sources of CP violation. As a result there is, as of the time of writing, a small range of parameters where the observed asymmetry could be produced in this way.

19.5.3 Leptogenesis

There is compelling evidence that neutrinos have mass. The most economical explanation of these masses is that they arise from a seesaw, involving gauge singlet fermions N_a. These couplings violate lepton number. So, according to Sakharov's principles, we might hope to produce a lepton asymmetry in their decays. Because the electroweak interactions violate baryon and lepton number at high temperatures, the production of a lepton number leads to the production of baryon number.

In general, there may be several N_a fields, with couplings of the form

$$\mathcal{L}_N = M_{ab} N_a N_b + h_{ai} H L_i N_a + \text{c.c.} \tag{19.78}$$

In a model with three Ns, there are CP-violating phases in the Yukawa couplings of the Ns to the light Higgs. The heaviest of the right-handed neutrinos, say N_1, can decay to ℓ and a Higgs, or to $\bar{\ell}$ and a Higgs. At tree level, as in the case of GUT baryogenesis, the rates for production of leptons and antileptons are equal, even though there are CP-violating phases in the couplings. It is necessary, again, to look at quantum corrections, since in these dynamical phases can appear in the amplitudes. At one loop the decay amplitude for N has a discontinuity associated with the fact that the intermediate N_1 and N_2 can be on-shell (a similar situation to that in Fig. 19.4). So, one obtains an asymmetry ϵ proportional to the imaginary parts of the Yukawa couplings of the Ns to the Higgs:

$$\epsilon = \frac{\Gamma(N_1 \to \ell H_2) - \Gamma(N_1 \to \bar{\ell}\bar{H}_2)}{\Gamma(N_1 \to \ell H_2) + \Gamma(N_1 \to \bar{\ell}\bar{H}_2)} = \frac{1}{8\pi} \frac{1}{hh^\dagger} \sum_{i=2,3} \text{Im}[(h_\nu h_\nu^\dagger)_{1i}]^2 f\left(\frac{M_i^2}{M_1^2}\right), \tag{19.79}$$

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

where f is a function that represents radiative corrections. For example, in the Standard Model, $f = \sqrt{x}[(x-2)/(x-1) + (x+1)\ln(1+1/x)]$ while in the MSSM, $f = \sqrt{x}[2/(x-1) + \ln(1+1/x)]$. Here we have allowed for the possibility of multiple Higgs fields, with H_2 coupling to the leptons. The rough order of magnitude here is readily understood by simply counting loop factors. It need not be very small.

Now, as the universe cools through temperatures of order the masses of the Ns, they drop out of equilibrium and their decays can lead to an excess of neutrinos over antineutrinos. Detailed predictions can be obtained by integrating a suitable set of Boltzmann equations. However, a rough estimate can be obtained by noting that the N_as drop out of equilibrium when their production rate becomes comparable with the expansion rate of the universe. If α represents a typical coupling, this occurs roughly when

$$\pi\alpha^2 T e^{-M_N/T} \approx \frac{T^2}{M_{\mathrm{p}}}. \tag{19.80}$$

Assuming that, in the polynomial terms, $T \sim M_N/10$ gives a density at this time of order

$$\frac{\rho_N}{\rho_{\mathrm{tot}}} \sim \frac{\pi T}{M_{\mathrm{p}}\alpha^2}. \tag{19.81}$$

Multiplying by ϵ, the average asymmetry in N decays, this estimate suggests a lepton number – and hence a baryon number – of order

$$\frac{\rho_{\mathrm{B}}}{\rho_{\mathrm{tot}}} \approx \epsilon \frac{M_N}{10\pi\alpha^2 M_{\mathrm{p}}}. \tag{19.82}$$

We have seen that ϵ is suppressed by a loop factor and by Yukawa couplings. So the above number can easily be compatible with observations, or even somewhat larger, depending on a variety of unknown parameters.

These decays, then, produce a net lepton number but not a net baryon number (hence they produce a net $B-L$). The resulting lepton number will be further processed by sphaleron interactions, yielding a net lepton and baryon number (recall that sphaleron interactions preserve $B-L$ but violate B and L separately). One can determine the resulting asymmetry by an elementary thermodynamics exercise: one introduces chemical potentials for each neutrino, quark and charged lepton species and then considers the various interactions between the species at equilibrium. For any allowed chemical reaction, the sum of the chemical potentials on each side of the reaction must be equal. For neutrinos, the relations come from:

1. the sphaleron interactions themselves,

$$\sum_i \left(3\mu_{q_i} + \mu_{\ell_i}\right) = 0; \tag{19.83}$$

2. a similar relation for QCD sphalerons,

$$\sum_i \left(2\mu_{q_i} - \mu_{u_i} - \mu_{d_i}\right) = 0; \tag{19.84}$$

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

3. the vanishing of the total hypercharge of the universe,

$$\sum_i \left(\mu_{q_i} - 2\mu_{\bar{u}_i} + \mu_{\bar{d}_i} - \mu_{\ell_i} + \mu_{\bar{e}_i}\right) + \frac{2}{N}\mu_{\rm H} = 0; \tag{19.85}$$

4. the quark and lepton Yukawa coupling relations

$$\mu_{q_i} - \mu_\phi - \mu_{d_j} = 0, \quad \mu_{q_i} - \mu_\phi - \mu_{u_j} = 0, \quad \mu_{\ell_i} - \mu_\phi - \mu_{e_j} = 0. \tag{19.86}$$

The number of equations here is the same as the number of unknowns. Combining these, one can solve for the chemical potentials in terms of the lepton chemical potential and, finally, in terms of the initial $B - L$. With N generations we obtain,

$$B = \frac{8N + 4}{22N + 13}(B - L). \tag{19.87}$$

Reasonable values of the neutrino parameters give asymmetries of the order we seek to explain. Note the sources of small numbers:

1. the phases in the couplings;
2. the loop factor;
3. the small density of the N particles when they drop out of equilibrium; parametrically, one has, e.g., for production,

$$\Gamma \sim e^{(-M/T)} g^2 T, \tag{19.88}$$

which is much less than $H \sim T^2/M_{\rm p}$ once the density is suppressed by $T/M_{\rm p}$, i.e. Γ is of order 10^{-6} for a 10^{13} GeV particle.

It should be noted that implementing this mechanism requires a high reheating temperature after inflation, of order the mass of the right-handed neutrinos. It is conceivable, as we have seen, that the reheating temperature is this high. It is also possible that the right-handed neutrinos are light. If the reheating temperatures (after inflation or moduli decay) are *very* low, some other mechanism to produce the dark matter is required.

It is interesting to ask, assuming that these processes are the source of the observed asymmetry, how many parameters which enter into the computation can be measured? It is likely that, over time, many parameters of the light neutrino mass matrices, including possible CP-violating phases, will be measured. But, while these measurements determine some N_i couplings and masses, they are not in general enough. In order to give a precise calculation, analogous to nucleosynthesis calculations, of the baryon number density one needs additional information about the masses of the fields N_i. One either requires some other (currently unforeseen) experimental access to this higher-scale physics or a compelling theory of neutrino mass in which symmetries, perhaps, reduce the number of parameters.

19.5.4 Baryogenesis through coherent scalar fields

In supersymmetric theories the ordinary quarks and leptons are accompanied by scalar fields. These scalar fields carry baryon and lepton number. A coherent field, i.e. a large classical value of such a field, can in principle carry a large amount of baryon number. As

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

we will see, it is quite plausible that such fields were excited in the early universe, and this could have led to a baryon asymmetry.

To understand the basics of the mechanism, consider first a model with a single complex scalar field. Take the Lagrangian to be

$$\mathcal{L} = |\partial_\mu \phi|^2 - m^2|\phi|^2. \tag{19.89}$$

This Lagrangian has a symmetry, $\phi \to e^{i\alpha}\phi$, and a corresponding conserved current, which we will refer to as the baryon number:

$$j^\mu_{\rm B} = i(\phi^* \partial^\mu \phi - \phi \partial^\mu \phi^*). \tag{19.90}$$

It also possesses a CP symmetry:

$$\phi \leftrightarrow \phi^*. \tag{19.91}$$

With supersymmetry in mind we will think of m as of order M_W.

If we focus on the behavior of spatially constant fields, $\phi(\vec{x}, t) = \phi(t)$, this system is equivalent to an isotropic harmonic oscillator in two dimensions. In field theory, however, we expect that higher-dimensional terms will break the symmetry. In the isotropic oscillator analogy, this corresponds to anharmonic terms which break the rotational invariance. With a general initial condition the system will develop some non-zero angular momentum. If the motion is damped, so that the amplitude of the oscillations decreases, these rotationally non-invariant terms will become less important with time.

In the supersymmetric field theories of interest, supersymmetry will be broken by small quartic and higher-order couplings as well as by soft masses for the scalars. So, as a simple model, take:

$$\mathcal{L}_I = \lambda|\phi|^4 + \epsilon\phi^3\phi^* + \sigma\phi^4 + \text{c.c.} \tag{19.92}$$

These interactions clearly violate the conservation of B. For general complex ϵ and σ, they also violate CP. As we will shortly see, once supersymmetry is broken, quartic and higher-order couplings can be generated but these couplings $\lambda, \epsilon, \sigma \ldots$ will be extremely small, $\mathcal{O}(m_{3/2}^2/M_{\rm p}^2)$ or $\mathcal{O}(m_{3/2}^2/M_{\rm GUT}^2)$.

In order that these tiny couplings could have led to an appreciable baryon number, it is necessary that the fields, at some stage, were very large. To see how the cosmic evolution of this system can lead to a non-zero baryon number, first note that at very early times, when the Hubble constant $H \gg m$ (see Eq. (19.89)), the mass of the field is irrelevant. It is thus reasonable to suppose that at this early time $\phi = \phi_0 \gg 0$; later we will describe some specific suggestions as to how this might come about. This system then evolves like the axion or moduli. In the radiation- and matter-dominated eras, respectively, one has that

$$\phi = \frac{\phi_0}{(mt)^{3/2}} \sin mt \qquad \text{(radiation)} \tag{19.93}$$

$$\phi = \frac{\phi_0}{mt} \sin mt \qquad \text{(matter).} \tag{19.94}$$

In either case the energy behaves, in terms of the scale factor $R(t)$, as

$$E \approx m^2 \phi_0^2 \left(\frac{R_0}{R}\right)^3, \tag{19.95}$$

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

i.e. it decreases as R^3, as would the energy of pressureless dust. One can think of this oscillating field as a coherent state of ϕ particles with $\vec{p} = 0$.

Now let us consider the effects of the quartic couplings. Since the field amplitude damps with time, their significance will decrease with time. Suppose, initially, that $\phi = \phi_0$ is real. Then the real and imaginary parts of ϕ satisfy, in the approximation that ϵ and δ are small,

$$\ddot{\phi}_i + 3H\dot{\phi}_i + m^2\phi_i \approx \text{Im}(\epsilon + \delta)\phi_r^3. \tag{19.96}$$

For large times, the right-hand side falls off as $t^{-9/2}$ whereas the left-hand side falls off only as $t^{-3/2}$. As a result, just as in our mechanical analogy, baryon number (angular momentum) violation becomes negligible. Equation (19.96) goes over to the free equation, with a solution of the form

$$\phi_i = a_\text{r} \frac{\text{Im}(\epsilon + \delta)\phi_0^3}{m^2(mt)^{3/4}} \sin(mt + \delta_\text{r}) \qquad \text{(radiation)}, \tag{19.97}$$

$$\phi_i = a_\text{m} \frac{\text{Im}(\epsilon + \delta)\phi_0^3}{m^3 t} \sin(mt + \delta_\text{m}) \qquad \text{(matter)}, \tag{19.98}$$

in the radiation- and matter-dominated cases, respectively. The constants δ_m, δ_r, a_m and a_r can easily be obtained numerically, and are of order unity:

$$a_\text{r} = 0.85, \quad a_\text{m} = 0.85, \quad \delta_\text{r} = -0.91, \quad \delta_\text{m} = 1.54. \tag{19.99}$$

However, now we have a non-zero baryon number; substituting in the expression for the current,

$$n_\text{B} = 2a_\text{r}\,\text{Im}(\epsilon + \delta)\frac{\phi_0^2}{m(mt)^2} \sin(\delta_\text{r} + \pi/8) \qquad \text{(radiation)} \tag{19.100}$$

$$n_\text{B} = 2a_\text{m} \quad \text{Im}(\epsilon + \delta)\frac{\phi_0^2}{m(mt)^2} \sin\delta_\text{m} \qquad \text{(matter)}. \tag{19.101}$$

Note that CP violation can be provided here by phases in the couplings and/or the initial fields. Note also that as expected, n_B is conserved at late times, in the sense that the baryon number per comoving volume is constant.

This mechanism for generating baryon number could be considered without supersymmetry. In that case, several questions would be begged.

- What are the scalar fields carrying baryon number?
- Why are the ϕ^4 terms so small?
- How are the scalars in the condensate (see Section 19.8) converted to more familiar particles?

In the context of supersymmetry there is a natural answer to each question. First, as we have stressed, there are scalar fields carrying baryon and lepton number. As we will see, in the limit in which supersymmetry is unbroken, there are typically approximate flat directions in the field space in which the quartic terms in the potential vanish. Finally, the scalar quarks and leptons can decay (in a baryon- and lepton-number-conserving fashion) to ordinary quarks.

19.6 Flat directions and baryogenesis

To discuss the problem of baryon number generation, we first want to examine the theory in a limit in which we ignore the soft susy-breaking terms. After all, at very early times, $H \gg M_W$ and these terms were irrelevant. We are now quite familiar with the fact that supersymmetric theories often exhibit flat directions. At the renormalizable level the MSSM possesses many flat directions. A simple example is

$$H_u = \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad L_1 = \begin{pmatrix} v \\ 0 \end{pmatrix}, \tag{19.102}$$

where L_1 denotes the first-generation lepton doublet and v is an (at this point arbitrary) expectation value. This direction is characterized by a modulus which carries lepton number. Written in a gauge-invariant fashion, $\Phi = H_u L$. As we have seen, producing a lepton number is for all intents and purposes like producing a baryon number. Non-renormalizable, higher-dimensional terms, with more fields, can lift the flat direction. For example, the quartic term in the superpotential,

$$\mathcal{L}_4 = \frac{1}{M}(H_u L)^2 \tag{19.103}$$

respects all the gauge symmetries and is invariant under R-parity. Here M denotes some very large scale, perhaps the planck mass M_p. The term (19.103) gives rise to a potential

$$V_\mathrm{lift} = \frac{|v|^6}{M^2}. \tag{19.104}$$

There are many more flat directions, and many of them do carry baryon or lepton number. A flat direction with both baryon and lepton numbers excited is the following:

$$\text{first generation, } Q_1^1 = b, \quad \bar{u}_2 = a, \quad L_2 = b; \tag{19.105}$$

$$\text{second generation, } \bar{d}_1 = \sqrt{|b|^2 + |a|^2}, \tag{19.106}$$

$$\text{third generation, } \bar{d}_3 = a. \tag{19.107}$$

(On Q the upper index is a color index and the lower index is an $SU(2)$ index; we have suppressed the generation indices.)

Higher-dimensional operators can again lift this flat direction. In this case the leading term is

$$\mathcal{L}_7 = \frac{1}{M^3}[Q^1 \bar{d}^2 L^1][\bar{u}^1 \bar{d}^2 \bar{d}^3]. \tag{19.108}$$

Here the superscripts denote flavor. We have suppressed the color and $SU(2)$ indices, but the brackets indicate sets of fields which are contracted in $SU(3)$- and $SU(2)$-invariant ways. In addition to being completely gauge invariant, this operator is invariant

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

under ordinary R-parity. (There are lower-dimensional operators, including operators of dimension four, which violate R-parity.) It gives rise to a term in the potential

$$V_{\text{lift}} = \frac{\Phi^{10}}{M^6}. \tag{19.109}$$

Here Φ refers in a generic way to the fields whose vevs parameterize the flat directions (a, b).

19.7 Supersymmetry breaking in the early universe

We have indicated that higher-dimensional, supersymmetric operators give rise to potentials in the flat directions. To fully understand the behavior of the fields in the early universe, we need to consider supersymmetry breaking, which gives rise to additional potential terms.

In the early universe, we expect that supersymmetry was much more badly broken than it is in the present era. For example, during inflation, the non-zero energy density (the cosmological constant) breaks supersymmetry. Suppose that I is the inflaton field and that the inflaton potential arises because of a non-zero value of the auxiliary field for I, $F_I = \partial W/\partial I$. So, during inflation, supersymmetry is broken by a large amount. Not surprisingly, as a result there can be an appreciable supersymmetry-breaking potential for the field Φ. These contributions to the potential have the form

$$V_H = H^2\Phi^2 f\left(\Phi^2/M_{\rm p}^2\right). \tag{19.110}$$

It is perfectly possible for the second derivative of the potential near the origin to be negative. In this case, writing our higher-dimensional term as

$$W_n = \frac{1}{M^n}\Phi^{n+3}, \tag{19.111}$$

the potential takes the form

$$V = -H^2|\Phi|^2 + \frac{1}{M^{2n}}|\Phi|^{2n+4}. \tag{19.112}$$

The minimum of the potential then lies at

$$\Phi_0 \approx M\left(\frac{H}{M}\right)^{1/(n+1)}. \tag{19.113}$$

More generally, one can see that the higher the dimension of the operator that raises the flat direction, the larger the starting value of the field and the larger the ultimate value of the baryon number. Typically, there is plenty of time for the field to find its minimum during inflation. After inflation, H decreases and the field Φ evolves adiabatically, oscillating slowly about the local minimum for some time.

Our examples illustrate that, in models with R-parity, the value of n and hence the size of the initial field can vary appreciably. Which flat direction is most important depends on

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

the form of the mass matrix (i.e. it depends on in which directions the curvature of the potential is negative). With further symmetries, it is possible that n is larger and even that all operators which might lift the flat direction are forbidden. For the rest of this section, however, we will continue to assume that the flat directions are lifted by terms in the superpotential. If they are not, the required analysis is different since the lifting of a flat direction is entirely associated with supersymmetry breaking.

19.7.1 Appearance of the baryon number

The term $|\partial W/\partial\Phi|^2$ in the potential does not break either baryon number or CP. In most models it turns out that the leading sources of B and CP violation come from supersymmetry-breaking terms associated with F_I. These have the form

$$am_{3/2}W + bHW. \tag{19.114}$$

Here a and b are complex dimensionless constants. The relative phase in these two terms, δ, violates CP. This is crucial; if the two terms carry the same phase then this phase can be eliminated by a field redefinition, and we have to look elsewhere for possible CP-violating effects. Examining Eqs. (19.103) and (19.108), one sees that the term proportional to W violates B and/or L. In following the evolution of the field Φ, the important era occurs when $H \sim m_{3/2}$. At this point the phase misalignment of the two terms, along with the B-violating coupling, leads to the appearance of a baryon number. From the equations of motion, the equation for the time rate of change of the baryon number is

$$\frac{dn_{\rm B}}{dt} = \frac{\sin\delta\, m_{3/2}}{M^n}\phi^{n+3}. \tag{19.115}$$

Assuming that the relevant time is H^{-1}, one is led to the estimate (supported by numerical studies)

$$n_{\rm B} = \frac{1}{M^n}\sin\delta\,\Phi_0^{n+3}. \tag{19.116}$$

Here, Φ_0 is determined by $H \approx m_{3/2}$, i.e. $\Phi_0^{2n+2} = m_{3/2}^2 M^{2n}$.

19.8 The fate of the condensate

Of course, we do not live in a universe dominated by a coherent scalar field. In this section we consider the fate of a homogeneous condensate in the early universe, ignoring possible inhomogeneities. The following section will deal with the inhomogeneities and the interesting array of phenomena to which they might give rise.

We will adopt the following model for inflation. The features of this picture are true of many models of inflation but by no means all. We will suppose that the energy scale of inflation is $E \sim 10^{15}$ GeV and that inflation is due to a field, the inflaton I. We will take the amplitude of the inflaton, just after inflation, to be of order $M \approx 10^{18}$ GeV (the usual reduced Planck mass). Correspondingly, we will take the mass of the inflaton to

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

be $m_I = 10^{12}$ GeV (so that $m_I^2 M_p^2 \approx E^4$). Correspondingly, the Hubble constant during inflation is of order $H_I \approx E^2/M_p \approx 10^{12}$ GeV.

After inflation ends the inflaton oscillates about the minimum of its potential, much like the field Φ, until it decays. We will suppose that the inflaton couples to ordinary particles with a rate suppressed by a single power of the Planck mass. Dimensional analysis then gives, for a rough value of the inflaton lifetime,

$$\Gamma_I = \frac{m_I^3}{M^2} \sim 1 \text{ GeV}. \tag{19.117}$$

The reheating temperature can then be obtained by equating the energy density at time, $H \approx \Gamma(\rho = 3H^2M^2)$, to the energy density of the final plasma:

$$T_R = T\left(t = \Gamma_I^{-1}\right) \sim (\Gamma_I M_p)^{1/2} \sim 10^9 \text{ GeV}. \tag{19.118}$$

The decay of the inflaton is not sudden, however, but leads to a gradual reheating of the universe, as described, for example, in the book by Kolb and Turner (1990). As a function of time,

$$T \approx \left(T_R^2 H(t) M_p\right)^{1/4}. \tag{19.119}$$

where $H(t)$ is the Hubble parameter as a function of time. For the field Φ our basic assumption is that during inflation it obtains a large value, in accord with Eq. (19.113). When inflation ends the inflaton, by assumption, still dominates the energy density for a time, oscillating about its minimum; the universe is matter dominated during this period. The field Φ now oscillates about a time-dependent minimum, given by Eq. (19.113). The minimum decreases in value with time, dropping to zero when $H \sim m_{3/2}$. During this evolution, a baryon number develops classically. This number is frozen once $H \sim m_{3/2}$.

Eventually the condensate will decay, through a variety of processes. As we have stressed, the condensate can be thought of as a coherent state of Φ particles. These particles – linear combinations of the squark and slepton fields – are unstable and will decay. However, for $H \leq m_{3/2}$ these lifetimes are much longer than those in the absence of the condensate. The reason is that the fields to which Φ couples have mass of order Φ, and Φ is large. Particles which are light in the presence of large ϕ form an ambient thermal bath. In most cases, the most important process which destroys the condensate is what we might call evaporation: particles in the ambient thermal bath can scatter off the particles in the condensate.

We can make a crude estimate for the reaction rate as follows. Because the particles which couple directly to Φ are heavy, interactions of Φ particles with light particles must involve loops. So we include a loop factor in the amplitude, of order α_2^2, the weak coupling squared. Because of the large masses, the amplitude is suppressed by Φ. Squaring and multiplying by the thermal density of the scattered particles gives a crude estimate for the reaction rate.

$$\Gamma_p \sim \alpha_2^2 \pi \frac{1}{\Phi^2} \left(T_R^2 H M\right)^{3/4}. \tag{19.120}$$

The condensate will evaporate when this quantity is of order H. Since we know the time dependence of Φ, this allows us to solve for this time. One finds that equality occurs, in

https://doi.org/10.1017/9781009290883.025 Published online by Cambridge University Press

the case $n = 1$, for $H_I \sim 10^2$–10^3 GeV. For $n > 1$ it occurs significantly later; for $n < 4$ it occurs before the decay of the inflaton and for $n \geq 4$ a slightly different analysis is required from that which follows. In other words, for the case $n = 1$, the condensate evaporates shortly after the baryon number is created but for larger n, it evaporates significantly later.

The expansion of the universe is unaffected by the condensate as long as the energy density in the condensate, $\rho_\Phi \sim m_\Phi^2 \Phi^2$, is much smaller than that of the inflaton, $\rho_I \sim H^2 M^2$. Assuming that $m_\Phi \sim m_{3/2} \sim 0.1$–$1$ TeV, a typical supersymmetry-breaking scale, one can estimate the ratio of the two densities at the time when $H \sim m_{3/2}$ as

$$\frac{\rho_\Phi}{\rho_I} \sim \left(\frac{m_{3/2}}{M_{\rm p}}\right)^{2/(n+1)}. \tag{19.121}$$

We are now in a position to calculate the baryon to photon ratio in this model. Given our estimate of the inflaton lifetime, the coherent motion of the inflaton still dominates the energy density when the condensate evaporates. The baryon number equals the Φ density just before evaporation divided by the Φ mass (assumed to be of order $m_{3/2}$), while the inflaton number is ρ_I/M_I. So the baryon to inflaton ratio follows from Eq. (19.121). With the assumption that the inflaton energy density is converted to radiation at the reheating temperature, $T_{\rm R}$, we obtain

$$\frac{n_{\rm B}}{n_\gamma} \sim \frac{n_{\rm B}}{\rho_I/T_{\rm R}} \sim \frac{n_{\rm B}}{n_\Phi}\frac{T_{\rm R}}{m_\Phi}\frac{\rho_\Phi}{\rho_I} \sim 10^{-10}\left(\frac{T_{\rm R}}{10^9\ {\rm GeV}}\right)\left(\frac{M_{\rm p}}{m_{3/2}}\right)^{(n-1)/(n+1)}. \tag{19.122}$$

Clearly the precise result depends on factors beyond those indicated here explicitly, such as the precise mass of the Φ particle(s). But as a rough estimate it is rather robust. For $n = 1$, it is in *precisely the right range* to explain the observed baryon asymmetry. For larger n, it can be significantly larger. In light of our discussion of the late decays of moduli this is potentially quite interesting. These decays produce a huge amount of entropy, typically increasing it by a factor 10^7 or so. The baryon density is diluted by a corresponding factor. But we see that coherent production can readily yield, prior to moduli decay, baryon to photon densities of the needed size.

There are many other issues which can be studied, both in leptogenesis and in Affleck–Dine baryogenesis, but it appears that both types of process might well account for the observed baryon asymmetry. The discovery (or not) of low-energy supersymmetry, and further studies of neutrino masses, might make one or the other picture more persuasive. Both pose challenges, as they involve couplings which we are not likely to be able to measure directly.

19.9 Dark energy

It has long been recognized that any cosmological constant in nature is far smaller than the scales of particle physics. Before the discovery of dark energy, many physicists conjectured that for some reason of principle this energy was zero. However, as we have seen, we now know that the dark energy is non-zero and in fact that it is the largest component of the energy density of the universe. Present data is compatible with the idea that this energy density represents a cosmological constant (for a cosmological constant, in the equation of state, we have $p = w\rho$ and $w = -1$; the Planck satellite, for example, gives $w = -1.10^{+0.08}_{-0.07}$) but other suggestions, typically involving time-varying scalar fields, have been offered and future surveys will improve the measurement of w.

Apart from its smallness, another puzzle surrounding the cosmological constant is simply one of coincidence: why is the dark energy density today comparable to the dark matter density? Weinberg has argued that it could not be much different from this in a universe containing stars and galaxies, provided that all the other laws of nature are as we observe. The basic point is that if the dark energy were, say, 10^3 times more dense than we observe, it would have come to dominate the energy density when the universe was much younger than it is today, at a time prior to the formation of galaxies and stars. The rapid acceleration after that time would have prevented the formation of structure. More refined versions of the argument give estimates for the dark energy within a factor ten of the measured value.

Weinberg speculated that perhaps the universe is much larger than we see (i.e. than our current horizon) and that in other regions it has different values of the cosmological constant. Only in those regions where Λ is very small would stars – and hence observers – form. Weinberg called this possible explanation (actually a prediction) of Λ the *weak anthropic argument*. We will return to this question in our studies of string theory, where we will see that such a *landscape* of ground states may exist.

Suggested reading

Seminal papers on inflation include that of Guth (1981), which proposed a version of inflation now often referred to as “old inflation,” and those of Linde (1982) and Albrecht and Steinhardt (1982), which contain the germ of the slow-roll inflation idea stressed in this work. The ideas of hybrid inflation were developed by Linde (1994); those specifically discussed here were introduced by Randall *et al*. (1996) and Berkooz *et al*. (2004). There are a number of good texts on inflation and related issues, some of which we have mentioned in the previous chapter. These texts include those of Dodelson (2004), Kolb and Turner (1990) and Linde (1990). Dodelson provide a particularly up-to-date discussion of dark matter, including more detailed calculations than those presented here, and dark energy, including surveys of observational results. For a review of axions and their cosmology and astrophysics, see Turner (1990). For more recent papers which raise questions about the cosmological axion limits see, for example, Banks *et al*. (2003). The cosmological moduli problem, and possible solutions, were first discussed by Banks *et al*. (1994) and de Carlos *et al*. (1993). A general review of electroweak baryogenesis, including detailed discussions of phenomena at the bubble walls, appears in Cohen *et al*. (1993). A discussion of electroweak baryogenesis within the MSSM is given in Carena *et al*. (2003). A detailed review of baryogenesis is to be found in Buchmuller *et al*. (2005), while Enqvist and Mazumdar (2003) focuses on Affleck–Dine baryogenesis. A more comprehensive review of baryogenesis mechanisms appears in Dine and Kusenko (2003). Aspects of the cosmological constant, and especially Weinberg's anthropic prediction of Λ, are explained clearly in Weinberg (1989), with more recent additions in Vilenkin (1995) and Weinberg (2000).

Exercises

(1) Verify the slow-roll conditions, Eqs. (19.11) and (19.12). Determine the number of e-foldings and the size of $\delta\rho/\rho$ as a function of N.
(2) Work through the limits on the axion in more detail. Attempt to analyze the behavior of the axion energy in the high-temperature regime.
(3) Construct a discrete R symmetry which guarantees that the $H_U L$ flat direction is exactly flat. Assuming that the universe reheats to 100 MeV when a modulus decays, estimate the final baryon number of the universe in this case.
(4) Suppose that the characteristic scale of supersymmetry breaking is much higher than 1 TeV, say 10^9 GeV. Discuss baryogenesis by coherent scalar fields in such a situation.