Nuclear Physics B172 (1980) 224–284

BARYON NUMBER GENERATION IN THE EARLY UNIVERSE*

Edward W. KOLB[1]

W.K. Kellogg Radiation Laboratory, California Institute of Technology, Pasadena, California 91125, USA

Stephen WOLFRAM[2]

Theoretical Physics Laboratory, California Institute of Technology, Pasadena, California 91125, USA

Received 29 November 1979
(Final version received 10 March 1980)

The generation of an excess of baryons over antibaryons in the very early universe due to *CP*- and *B*-violating interactions is described. The Boltzmann equation is used to perform detailed calculations of the time development of such an excess in several simple illustrative models. Complications encountered in applications of the results to specific models are discussed.

1. Introduction

Theories in which quarks and leptons are treated in a unified manner often lead naturally to the speculation that there should exist interactions which violate baryon and lepton number. The rather stringent limit ($\gtrsim 10^{30}$ yr) on the proton lifetime suggests that such interactions, if present, must be mediated by very massive particles (with $m \gtrsim 10^{14}$ GeV). A null result in the forthcoming generation of searches for proton decay could rule out the detailed predictions of the present generation of models, but could never provide an ultimate proof for the absence of baryon number (B) violation. For further information on B-violating interactions, one must forsake terrestrial experiments, and rely on indirect evidence from the early universe. According to the standard hot big bang model for the early universe, temperatures at sufficiently early times should have been high enough to overwhelm suppressions from large intermediate masses, rendering the rates for any B-violating reactions comparable to those for B-conserving ones.

If baryon number and the various (μ, e, ...) lepton numbers were absolutely conserved by all possible interactions occurring in the early universe, then the total baryon and lepton numbers of the present universe must simply reflect their apparently arbitrarily imposed initial values. A plausible guess would be that the initial total baryon and lepton numbers were exactly zero (as the total electric charge appears to be). However, if this is to be viable some mechanism must exist which serves either to separate baryons and antibaryons or to hide antibaryons in the

* A summary of parts of this work is given in ref. [1].

[1] Work supported in part by the National Science Foundation [PHY76-83685].

[2] Work supported in part by the US DOE [DE-AC-03-79ER0068] and a Feynman fellowship.

universe, since otherwise nearly all baryons should have annihilated away at temperatures above about 50 MeV in the early universe, and the present observed mean matter density, corresponding to $n_B/n_\gamma \simeq 10^{-9}$, would not be present. The fundamental prediction of models in which baryons and antibaryons are separated in the universe is the existence of antimatter galaxies. While there is quite convincing evidence against the presence of antimatter within our own galaxy (mostly based on the absence of obvious annihilation products), the constitution of other galaxies cannot definitely be ascertained. There is, however, an apparently insuperable theoretical difficulty with baryon-antibaryon symmetric models for the universe: the separation must have acted at times $\lesssim 0.1$ sec ($T \gtrsim 50$ MeV) in order to prevent complete annihilation, but at these times the processes of separation cannot yet have acted over regions containing more than about 10^{56} particles ($\sim 0.1\ M_\odot$), since a light signal could not by that time have traversed a larger distance.

If B-violating interactions do occur at very high energies, the present baryon number of the universe could no longer be specified simply as an initial condition. In fact, the presence of such interactions should serve to destroy all but perhaps a small fraction of any initial baryon number (see sect. 4), and insufficient baryons would have survived to explain the observed $n_B/n_\gamma \simeq 10^{-9}$. If no further effects occurred, then models involving appreciable B violation at very high energies would presumably be in disagreement with the standard cosmological model. However, if C and CP invariances are also violated in B-violating reactions, it is possible that a calculable baryon excess may be generated after any initial baryon excess has been erased, thus allowing large B violation at high temperatures without inconsistency with the standard hot big bang model of the early universe. Nevertheless, since CPT invariance must remain intact, CP violation can have no effect unless a definite arrow of time is defined (see subsect. 2.1). In thermal equilibrium no preferred time direction exists. However, the expansion of the universe may result in small deviations from thermal equilibrium, which allow an excess of baryons over antibaryons to be generated by the action of B-, CP-violating interactions. The relaxation time necessary to regain true equilibrium in which the baryon asymmetry has been destroyed again may increase faster than the age of the universe, thus freezing the asymmetry. A model along these basic lines was considered by Sakharov in 1966 [2], and since the development of grand unified gauge theories in which B violation is rampant, the generation of a baryon excess in the early universe has been discussed extensively [3–8]. In this paper we perform a detailed calculation of the development of a baryon excess in several simple illustrative models. Sect. 2 introduces the models, and derives Boltzmann equations for the time evolution of the number densities of the various particles involved. In sect. 3 we discuss the solution of these equations in the early universe, and find that for plausible choices of parameters, numerical solutions are obligatory. The final results depend sensitively on the parameters; the observed n_B/n_γ should therefore place interesting constraints on models for the very early universe and the interactions occurring in it. The final

n_B/n_γ produced in any given model is always proportional to an unknown CP violation in super-high-energy interactions. The origin and magnitude of this CP violation is probably unconnected to that observed in the K^0 system. Nevertheless, since there is always an upper limit to CP-violating phases, any model involving B violation which cannot generate sufficient n_B/n_γ even with maximal CP violation must presumably be considered in disagreement with the standard cosmology. The methods developed in this paper are easily generalized to an arbitrary model (see subsect. 2.4); in a forthcoming work we shall describe the constraints which result [29].

In grand unified models where the μ, τ "families" are treated as simple replications of the lowest e family, it is inevitable that cosmological mechanisms which yield a net baryon number (and hence e number) in the universe should also generate net μ, τ asymmetries of the same magnitude. Such an asymmetry in massless $\bar{\nu}_\tau$, $\bar{\nu}_\mu$ number densities would, however, be quite negligible ($O(10^{-9})$) and presumably unobservable. However, the same asymmetry should also exist for possible more massive absolutely stable replications (their masses are irrelevant if they are much smaller than the temperature $\sim 10^{15}$ GeV at which the asymmetries must be generated). The observed deceleration parameter for the universe suggests that the mean energy density does not exceed that observed in nucleons by more than about an order of magnitude. Thus there cannot exist absolutely stable particles much heavier than the proton in the concentrations suggested by grand unified models with the mechanism for baryon asymmetry generation described below. This constraint strengthens existing limits on neutral and charged heavy leptons and hadrons derived previously without grand unified models [13].

2. Basic formalism

2.1. INTRODUCTION

Let $\mathcal{M}(i\to j)$ be the amplitude for a transition from a state i to state j, and let $\bar{i}$ be the state obtained by applying a CP transformation to i. Then the CPT theorem (the validity of which is necessary to justify use of quantum field theory) implies that

$$\mathcal{M}(i\to j)=\mathcal{M}(\bar{j}\to\bar{i})\,,\qquad (CPT\text{ invariance})\,. \tag{2.1.1}$$

CP invariance (and hence, by CPT, T invariance), when valid, demands

$$\mathcal{M}(i\to j)=\mathcal{M}(\bar{i}\to\bar{j})=\mathcal{M}(j\to i)\,,\qquad (CP\text{ invariance})\,. \tag{2.1.2}$$

The requirement of unitarity (that the probabilities for all possible transitions to and from a state i should sum to one) yields*

$$\sum_j|\mathcal{M}(i\to j)|^2=\sum_j|\mathcal{M}(j\to i)|^2\,,\qquad (\text{unitarity})\,. \tag{2.1.3}$$

But from (2.1.1)(the sum over j includes all states and their antistates)*

$$\sum_j |\mathcal{M}(i \to j)|^2 = \sum_j |\mathcal{M}(j \to \bar{i})|^2 = \sum_j |\mathcal{M}(j \to i)|^2 , \qquad (CPT + \text{unitarity}) . \quad (2.1.4)$$

In thermal equilibrium (and in the absence of chemical potentials corresponding to non-zero conserved quantum numbers) all states j of a system with a given energy are equally populated**. Eq. (2.1.4) then shows that transitions from these states (interactions) must produce states i and their CP conjugates $\bar{i}$ in equal numbers. Thus no excess of particles over antiparticles (and hence, for example, a net baryon number) may develop in a system in thermal equilibrium, even if CP invariance is violated. (A restricted form of this result was given in ref. [5].)

From eqs. (2.1.1) and (2.1.3) one finds

$$\sum_j |\mathcal{M}(i \to j)|^2 = \sum_j |\mathcal{M}(\bar{i} \to j)|^2 , \qquad (CPT + \text{unitarity}) , \qquad (2.1.5)$$

implying that the total cross section for interactions between a set of particles and their CP conjugates are equal, and that the total decay rate of a particle and its antiparticle must be equal. (CPT invariance alone implies the equality of the elastic scattering cross sections $|\mathcal{M}(i \to i)|^2 = |\mathcal{M}(\bar{i} \to \bar{i})|^2$.) However, the corresponding result

$$|\mathcal{M}(i \to j)|^2 = |\mathcal{M}(\bar{i} \to \bar{j})|^2 \equiv |\mathcal{M}(j \to i)|^2 , \qquad (CP \text{ invariance}) , \qquad (2.1.6)$$

for *specific* final states requires CP invariance (e.g., [10]). Thus if the interaction inducing the decay of a particle (say X) violates CP invariance, then the decay of a system containing an equal number of X and $\bar{\text{X}}$ can result in unequal numbers of, say, b and $\bar{\text{b}}$. Note that for a system with only two states, unitarity gives $|\mathcal{M}(1 \to 1)|^2 + |\mathcal{M}(1 \to 2)|^2 = |\mathcal{M}(1 \to 1)|^2 + |\mathcal{M}(2 \to 1)|^2$ so that the result (2.1.6) always holds.

Above, we argued that in thermal equilibrium, no excess of particles over antiparticles may develop. In addition, any pre-existing excess tends to be diminished by interactions: eq. (2.1.5) shows that the total cross sections for the destruction of the states i and $\bar{i}$ are equal. Hence, if there exist, say, more i than $\bar{i}$, then the rate of i destruction is larger than the rate for $\bar{i}$ destruction. Moreover, eq. (2.1.5) implies that in thermal equilibrium i and $\bar{i}$ are produced in equal numbers. Thus, interactions tend to destroy any excess of, say, i over $\bar{i}$ in thermal equilibrium. According to Boltzmann's H theorem (which holds regardless of T invariance, as discussed in appendix A) any closed system will evolve on average in the absence of external influences to a state in which all particles not carrying absolutely conserved quantum numbers are distributed equally in phase space. No difference between the densities of any species of particles and their antiparticles may survive (unless they

* This relation is modified if some of the final-state particles are indistinguishable from those in an ambient gas. The necessary quantum statistics corrections are given in appendix A.

** This result holds if all particles obey Maxwell–Boltzmann statistics. The modified form with quantum statistics is given in Eq. (A.23).

are distinguished by absolutely conserved quantum numbers). However, as discussed in appendix A, the expansion of the universe adds extra terms to the Boltzmann transport equation which invalidate the H theorem if some participating particles are massive. The expansion of the universe may therefore prevent the achievement of complete thermal equilibrium and allow a baryon excess to be generated: the relaxation time necessary to destroy the excess often increases faster than the age of the universe (see subsect. 3.2) and hence a net baryon number may persist.

CP invariance requires hermiticity of the transition matrix $\underset{\sim}{T} = i(\underset{\sim}{1} - \underset{\sim}{S})$, and in terms of the $\underset{\sim}{T}$ matrix the CP-invariance constraint (2.1.6) becomes

$$|\underset{\sim}{T}_{ij}|^2 = |\underset{\sim}{T}_{\bar{i}\bar{j}}|^2 = |\underset{\sim}{T}_{ji}|^2\,, \qquad (CP \text{ invariance})\,. \tag{2.1.7}$$

The unitarity requirement $\underset{\sim}{S}^\dagger \underset{\sim}{S} = \underset{\sim}{S}\underset{\sim}{S}^\dagger = 1$ [which gave eq. (2.1.3)] is

$$\underset{\sim}{T}_{ij} - \underset{\sim}{T}^*_{ji} = i \sum_n \underset{\sim}{T}_{in}(\underset{\sim}{T}^\dagger)_{nj}\,, \qquad (\text{unitarity})\,. \tag{2.1.8}$$

Thus unitarity constrains possible violations of CP invariance, and from eq. (2.1.8) one finds that deviations from (2.1.7) must obey

$$|\underset{\sim}{T}_{ij}|^2 - |\underset{\sim}{T}_{ji}|^2 = \left| i\Big(\sum_n \underset{\sim}{T}\underset{\sim}{T}^+\Big)_{ij} + \underset{\sim}{T}^*_{ji}\right|^2 - |\underset{\sim}{T}_{ji}|^2$$

$$= -2\,\mathrm{Im}\left[\Big(\sum_n \underset{\sim}{T}\underset{\sim}{T}^+\Big)_{ij} \underset{\sim}{T}^*_{ji}\right] + \left|\Big(\sum_n \underset{\sim}{T}\underset{\sim}{T}^+\Big)_{ij}\right|^2\,. \tag{2.1.9}$$

If the rates of transitions $i \to j$ are governed by some small parameter, say α, so that $|\mathcal{M}(i \to j)|^2 = \mathrm{O}(\alpha^k)$, then eq. (2.1.9) shows that any CP-violating difference $|\mathcal{M}(i \to j)|^2 - |\mathcal{M}(j \to i)|^2$ must be at least of order α^{k+1}. (Regardless of perturbation theory, one may show that CP-violating effects in any scattering process with high c.m. energy $\sqrt{s}$ are suppressed by $\mathrm{O}(1/\log(s/s_0))$ [for decays $s \to m^2$].) Hence, CP-violating effects must arise from loop diagram corrections to the processes $i \to j$. In appendix B, it is shown that these corrections must also involve B-violating interactions [11]. In addition, the intermediate states in the loops must correspond to physical systems n (so that the Feynman amplitudes have absorptive parts due to s-channel discontinuities), in order to contribute to (2.1.9). Thus even if the intermediate particles have CP violating complex couplings, they can produce a violation of (2.1.6) only when their masses are sufficiently small to allow them to propagate on their mass-shells in intermediate states. (Note that, as discussed in subsect. 2.4, when absolute conservation laws allow a particle to mix with its own antiparticle [as for K^0, $\bar{K}^0$], CP-violating mixing may occur without physical intermediate states.)

We assume that the early universe consists primarily of an effective number ξ (see appendix C) of massless particle species (none forming highly degenerate Fermi gases) and we usually take it to be homogeneous and isotropic. Then the Robertson–Walker scale parameter R for the early universe (which corresponds to its radius of curvature) should expand with time (t) according to ($\hbar = c = k = 1$) (e.g., [9])*

$$\frac{1}{R}\frac{\mathrm{d}R}{\mathrm{d}t} \simeq \left(\frac{8\pi\rho(t)}{3}\right)^{1/2}\frac{1}{m_{\mathscr{P}}}, \tag{2.1.10}$$

where $\rho(t)$ is the energy density of the universe, and $m_{\mathscr{P}}$ is the Planck mass

$$m_{\mathscr{P}} = G^{-1/2} \simeq 1.2\times 10^{19}\ \mathrm{GeV} = 1.2\times 10^{4}\ \Pi\mathrm{eV}\,. \tag{2.1.11}$$

In keeping with SI conventions, we take 1 ΠeV ≡ 1 TTeV ≡ 10^{24} eV ≡ 10^{15} GeV.

Let $f_i(\boldsymbol{p}, \boldsymbol{r}, t)$ be the density of a particle species i in phase space (i.e., the number of i per volume element $\mathrm{d}^3\boldsymbol{p}\ \mathrm{d}^3\boldsymbol{r}$). The assumptions of homogeneity and isotropy imply $f_i(\boldsymbol{p}, \boldsymbol{r}, t) = f_i(\boldsymbol{p}, t) = f_i(p, t)$; we usually do not display explicitly the dependence of $f_i(p)$ on time. We denote the number of i particles per unit volume (of configuration space) by

$$n_i = g_i \int \frac{\mathrm{d}^3\boldsymbol{p}}{(2\pi)^3} f_i(p)\,, \tag{2.1.12}$$

where g_i is the number of accessible spin states for the species $i (g_i = 2s_i + 1$ for $m_i > 0$; $g_i = 2$ for $m_i = 0$, $s_i > 0$; $g_i = 1$ for $m_i = 0$, $s_i = 0$; for particles with small mass, some spin states may be decoupled from interactions). We shall usually make the simplifying approximation that all particles obey Maxwell–Boltzmann statistics and have only one spin state; the small corrections resulting the indistinguishability of the particles are discussed in subsect. 2.4. Then the massless neutral (at least in baryon number) particles (denoted generically as γ) which comprise a large fraction of the contents of the early universe, should be Maxwell–Boltzmann distributed at all times due to their interactions, so that

$$f_\gamma(p) = \mathrm{e}^{-E/T} = \mathrm{e}^{-|\boldsymbol{p}|/T}\,,$$

$$n_\gamma = \frac{T^3}{\pi^2}\,,$$

$$\rho_\gamma = \frac{3}{\pi^2}T^4\,, \tag{2.1.13}$$

* This result may essentially be obtained by Newtonian considerations. Taking the universe to contain a homogeneous gas with density ρ, the energy equation for a unit mass shell of the gas with radius R becomes $\frac{1}{2}\dot{R}^2 - GM/R = E_{\mathrm{tot}} = -\frac{1}{2}k$, where $M = \frac{4}{3}\pi\rho R^3$ is the total mass contained within the shell (relativistic considerations show that pressure terms do not contribute). At least in the early universe the total energy may be neglected with respect to the separate kinetic and potential energies and

$$(\dot{R}/R)^2 \simeq \frac{8\pi}{3}G\rho = \frac{8\pi}{3}\frac{\rho}{m_{\mathscr{P}}^2}\,.$$

where T is their common temperature (referred to as "the temperature" of the universe). The expansion of the universe redshifts all $\boldsymbol{p}$ like $1/R$, so that $T \sim 1/R$, and (dots denote time derivatives)*

$$\frac{\dot{R}}{R} = -\frac{\dot{T}}{T} = (\tfrac{8}{3}\pi G\rho)^{1/2} = \frac{T^2}{m_{\mathrm{P}}},$$

$$m_{\mathrm{P}} = (\tfrac{1}{8}\pi)^{1/2} m_{\mathscr{P}}/\sqrt{\xi(T)} \simeq 0.63\, m_{\mathscr{P}}/\sqrt{\xi(T)}$$

$$\simeq 7.5 \times 10^3/\sqrt{\xi(T)}\,\mathrm{TTeV}\,. \tag{2.1.14}$$

Only for massless particles are the equilibrium $f_i(p)$ self-similar under rescalings of $\boldsymbol{p}$; for massive particles the mass provides an intrinsic scale and the $f_i(p)$ change their form as the universe expands. The expansion of the universe dilutes the number densities of all types of particles, even in the absence of interactions, at a rate

$$\frac{\mathrm{d}n_i}{\mathrm{d}t} = V\frac{\mathrm{d}V}{\mathrm{d}t}n_i = -\frac{3\dot{R}}{R}n_i\,. \tag{2.1.15}$$

In keeping with the simple big bang cosmology we shall assume** that all species of particles in the universe were initially in thermal equilibrium and spread homogeneously (the gravitational field opposing expansion must, however, remain far from equilibrium). Two effects modify this "equilibrium" state. First, regardless of expansion, long-range gravitational forces render a homogeneous state unstable, and lead to clumping. (This formally tends to increase the entropy of the universe, but in fact produces a more ordered state, eventually containing stars, etc.) Second, the expansion of a homogeneous universe can give rise to deviations from equilibrium, some of which may never have time to relax away. The expansion of the universe causes the momenta of all particles to redshift so that $\boldsymbol{p} \sim 1/R$. So long as the energy density of the universe is dominated by ultrarelativistic particle species, the temperature of the universe will likewise redshift according to $T \sim 1/R$. The equilibrium $f_i(p) \sim \exp(-|\boldsymbol{p}|/T)$ for massless particles remain, by construction, unchanged by this expansion (so long as homogeneity is preserved). Since the expansion of the universe is taken to be adiabatic, leading to the energy equation $\mathrm{d}(\rho R^3) + p\,\mathrm{d}(R^3) = 0$, the retention of equilibrium distributions by massless particles shows that the entropy of a universe initially in thermal equilibrium and composed solely of such massless particles (assumed homogeneous and with zero equilibrium chemical potentials) should remain unchanged with time. However, as mentioned above, the equilibrium distributions $f_i(p) \sim \exp(-\sqrt{|\boldsymbol{p}|^2 + m^2}/T$ for massive particles change their form when $|\boldsymbol{p}| \sim T \sim 1/R$ becomes smaller than $\sim m$. Several collision

* The temperature dependence of the expansion rate in eq. (2.1.14) obtains only in a homogeneous universe.

** We comment on the consistency of this assumption in sect. 4.

times are then necessary for the actual number distributions of the massive particles to relax into their equilibrium forms; if the rate of expansion is much larger than the rate of interactions, then significant deviations from equilibrium may result. When massive particles are present, therefore, the expansion of the universe is no longer reversible; deviations from thermal equilibrium may occur, and in their relaxation, the entropy of the universe may increase slightly*. (When equilibrium is destroyed by expansion, the gravitational field opposing the expansion becomes slightly closer to equilibrium; the increase in its entropy compensates the slight decrease in the entropy of the massive particle species.) If the expansion of the universe was slow enough, then any deviations from thermal equilibrium would eventually relax to zero. However, in several cases, the relaxation is always prevented by expansion. One example of this effect is in the survival of massive stable particles from the early universe [13]. At some time after the equilibrium number density of the massive particles starts to decrease rapidly, the rate for annihilation reactions becomes slower than the expansion rate of the universe, and the number of the particles is permanently frozen: the particles are so separated by the expansion that the probability for them to interact becomes negligible. In this way, the expansion of the universe causes the number density to deviate from its equilibrium form, and prevents its relaxation in a finite time. The generation of a baryon asymmetry is an effect of a similar character. That B, but not CP, violating reactions should occur only too slowly to destroy any baryon asymmetry (deviation from equilibrium) produced placed stringent constraints on models.

2.2. A VERY SIMPLE MODEL

In this subsection, we describe baryon number generation in a very simple model; subsect. 2.3 treats a more complicated and more realistic model, while in subsect. 2.4 we discuss the complications of the general case.

Let b be a nearly massless particle carrying baryon number $B=\frac{1}{2}$ and $\bar{\text{b}}$ its antiparticle, with $B=-\frac{1}{2}$, and let φ be a massive particle with $\bar{\varphi}\equiv\varphi$. We allow a small violation of CP invariance in the rates of $2\to 2$ scattering processes among the b and φ, and consider the generation of a net baryon number when a system initially symmetric in b and $\bar{\text{b}}$ cools, as in the early universe. We take the scattering amplitudes in the simple model to be ($|\mathcal{M}_0|^2=\mathrm{O}(\alpha^2)$, where α is a small coupling constant)

$$|\mathcal{M}(\text{bb}\to\bar{\text{b}}\bar{\text{b}})|^2=(1+\zeta)\tfrac{1}{2}|\mathcal{M}_0|^2\,,$$

$$|\mathcal{M}(\text{bb}\to\varphi\varphi)|^2=|\mathcal{M}(\varphi\varphi\to\bar{\text{b}}\bar{\text{b}})|^2=(1-\zeta)\tfrac{1}{2}|\mathcal{M}_0|^2\,,$$

* This increase may be considered to result from the appearance of a bulk (or second) viscosity [24] $\zeta>0$, which opposes the expansion, ζ vanishes for ideal gases in the limits $m\gg T$ and $m\ll T$, but is positive at intermediate temperatures $m\sim T$ (and is typically proportional to the relaxation time).

$$|\mathcal{M}(\bar{b}\bar{b}\to bb)|^2=(1+\bar{\zeta})\tfrac{1}{2}|\mathcal{M}_0|^2\,,$$

$$|\mathcal{M}(\bar{b}\bar{b}\to\varphi\varphi)|^2=|\mathcal{M}(\varphi\varphi\to bb)|^2=(1-\bar{\zeta})\tfrac{1}{2}|\mathcal{M}_0|^2\,, \tag{2.2.1}$$

where $-1<\zeta,\bar{\zeta}<1$. This parametrization ensures the *CPT*-invariance constraint (2.1.5) that the total cross sections for bb and $\bar{b}\bar{b}$ interactions should be equal. *CP* invariance would require $\zeta=\bar{\zeta}$; we shall, however, consider the *CP*-violating case $\zeta-\bar{\zeta}=O(\alpha)\neq 0$, so that, for example $|\mathcal{M}(bb\to\bar{b}\bar{b})|^2, |\mathcal{M}(\bar{b}\bar{b}\to bb)|^2=O(\alpha^2)$, $|\mathcal{M}(bb\to\bar{b}\bar{b})|^2-|\mathcal{M}(\bar{b}\bar{b}\to bb)|^2=O(\alpha^3)$. We assume here that the φ can interact only through the processes of eq. (2.2.1), but that the b, $\bar{b}$ also undergo baryon-number conserving interactions (such as $\gamma b\to\gamma b$ or $\gamma\gamma\to b\bar{b}$) with the other particles in the universe. Such reactions should typically have rates $O(\alpha^2)$ and serve to distribute the b and $\bar{b}$ in phase space in a Maxwell–Boltzmann manner. The time necessary to attain this state of kinetic equilibrium should be much shorter than the time $O(1/\alpha^3)$ on which the net baryon number $n_b-n_{\bar{b}}$ changes through the processes of eq. (2.2.1). Hence, at all times

$$f_b(p)\simeq e^{-(E-\mu)/T}\,,$$

$$f_{\bar{b}}(p)\simeq e^{-(E+\mu)/T}\,,$$

$$\frac{n_B}{n_\gamma}\equiv\left(\frac{n_b-n_{\bar{b}}}{n_\gamma}\right)\simeq 2\sinh\left(\frac{\mu}{T}\right)\,, \tag{2.2.2}$$

where μ is a baryon number chemical potential, which is changed only by *B*-violating processes (in the present model, these occur on a time-scale at least $O(1/\alpha)$ longer than the reactions which thermalize the b, $\bar{b}$ into Maxwell–Boltzmann distributions). The fact that the chemical potentials in f_b and $f_{\bar{b}}$ are exactly opposite is a consequence of processes such as $\gamma\gamma\to b\bar{b}$, which maintain $b\bar{b}$ in chemical equilibrium with $\gamma\gamma$.

The time evolution of the φ number density and of the total baryon number $n_B=n_b-n_{\bar{b}}$ due to the processes (2.2.1) is described by the Boltzmann transport equations*

$$\begin{aligned}\frac{dn_\varphi}{dt}+\frac{3\dot{R}}{R}n_\varphi=2\Lambda_{12}^{34}\,[&f_b(p_1)f_b(p_2)|\mathcal{M}(bb\to\varphi\varphi)|^2\\&+f_{\bar{b}}(p_1)f_{\bar{b}}(p_2)|\mathcal{M}(\bar{b}\bar{b}\to\varphi\varphi)|^2\\&-f_\varphi(p_1)f_\varphi(p_2)|\mathcal{M}(\varphi\varphi\to bb)|^2-f_\varphi(p_1)f_\varphi(p_2)|\mathcal{M}(\varphi\varphi\to\bar{b}\bar{b})|^2]\,,\end{aligned} \tag{2.2.3a}$$

* These equations are entirely classical. Identical particle corrections are discussed in subsect. 2.4. Genuine quantum-mechanical interference effects should be important only when the mean distance between successive collisions is shorter than the wavelengths of the participating particles. This circumstance may well occur at high temperatures, and a discussion of its consequences is given in sect. 4. The rough number of particles in causal contact at a temperature T is given (see subsect. 2.4) roughly by $10^{-2}(m_P/T)^3$ and is therefore sufficiently large at the temperatures we consider for statistical methods to be applicable.

$$\frac{\mathrm{d}n_{\mathrm{B}}}{\mathrm{d}t}+\frac{3\dot{R}}{R}n_{\mathrm{B}}=\Lambda_{12}^{34}\,[2f_{\bar{\mathrm{b}}}(p_1)f_{\bar{\mathrm{b}}}(p_2)|\mathcal{M}(\bar{\mathrm{b}}\bar{\mathrm{b}}\to\mathrm{bb})|^2-2f_{\mathrm{b}}(p_1)f_{\mathrm{b}}(p_2)|\mathcal{M}(\mathrm{bb}\to\bar{\mathrm{b}}\bar{\mathrm{b}})|^2$$
$$+f_\varphi(p_1)f_\varphi(p_2)|\mathcal{M}(\varphi\varphi\to\mathrm{bb})|^2-f_{\mathrm{b}}(p_1)f_{\mathrm{b}}(p_2)|\mathcal{M}(\mathrm{bb}\to\varphi\varphi)|^2$$
$$-f_\varphi(p_1)f_\varphi(p_2)|\mathcal{M}(\varphi\varphi\to\bar{\mathrm{b}}\bar{\mathrm{b}})|^2+f_{\bar{\mathrm{b}}}(p_1)f_{\bar{\mathrm{b}}}(p_2)|\mathcal{M}(\bar{\mathrm{b}}\bar{\mathrm{b}}\to\varphi\varphi)|^2]\,, \tag{2.2.3b}$$

where the integral operator

$$\Lambda_{\mathrm{a}_1,\mathrm{a}_2\ldots}^{\mathrm{b}_1,\mathrm{b}_2\ldots}\equiv\int\frac{\mathrm{d}^4p_{\mathrm{a}_1}}{(2\pi)^3}\int\frac{\mathrm{d}^4p_{\mathrm{a}_2}}{(2\pi)^3}\cdots\int\frac{\mathrm{d}^4p_{\mathrm{b}_1}}{(2\pi)^3}\int\frac{\mathrm{d}^4p_{\mathrm{b}_2}}{(2\pi)^3}\cdots$$
$$\times\delta(p_{\mathrm{a}_1}^2-m_{\mathrm{a}_1}^2)\delta(p_{\mathrm{a}_2}^2-m_{\mathrm{a}_2}^2)\ldots\delta(p_{\mathrm{b}_1}^2-m_{\mathrm{b}_1}^2)\delta(p_{\mathrm{b}_2}^2-m_{\mathrm{b}_2}^2)\ldots$$
$$\times(2\pi)^4\delta^4\Big(\sum_i p_{\mathrm{b}_i}-\sum_j p_{\mathrm{a}_j}\Big)$$
$$=\int\frac{\mathrm{d}^3\boldsymbol{p}_{\mathrm{a}_1}}{2E_{\mathrm{a}_1}(2\pi)^3}\int\frac{\mathrm{d}^3\boldsymbol{p}_{\mathrm{a}_2}}{2E_{\mathrm{a}_2}(2\pi)^3}\cdots\int\frac{\mathrm{d}^3\boldsymbol{p}_{\mathrm{b}_1}}{2E_{\mathrm{b}_1}(2\pi)^3}\int\frac{\mathrm{d}^3\boldsymbol{p}_{\mathrm{b}_2}}{2E_{\mathrm{b}_2}(2\pi)^3}\cdots$$
$$\times(2\pi)^4\delta^4(\sum p_{\mathrm{b}_i}-\sum p_{\mathrm{a}_j}) \tag{2.2.4}$$

represents appropriate integration over initial- and final-state phase space in the scattering processes. (When Λ has only upper or only lower indices, no momentum conservation δ function is included.) The second term on the left-hand side of eq. (2.2.3) accounts for the dilution of the number densities due to the expansion of the universe, as in eq. (2.1.15). (A proof of its form for Robertson–Walker metrics is given e.g., in ref. [5].) By considering

$$Y_{\mathrm{A}}\equiv n_{\mathrm{A}}/n_\gamma\,, \tag{2.2.5}$$

the explicit expansion term is removed ($\dot{Y}_{\mathrm{A}}=\dot{n}_{\mathrm{A}}/n_\gamma-n_{\mathrm{A}}\dot{n}_\gamma/n_\gamma^2=\dot{n}_{\mathrm{A}}/n_\gamma+(3\dot{R}/R)n_{\mathrm{A}}/n_\gamma$). The various terms on the right-hand side of (2.2.3) represent the effects of the processes (2.2.1) (for example, the first term on the right-hand side of (2.2.3a) accounts for the increase in the number of φ due to $\mathrm{bb}\to\varphi\varphi$).

To simplify (2.2.3), we first substitute the parametrizations (2.2.1):

$$\frac{\mathrm{d}Y_\varphi}{\mathrm{d}t}=\frac{1}{n_\gamma}\Lambda_{12}^{34}\,[\{(1-\zeta)f_{\mathrm{b}}(p_1)f_{\mathrm{b}}(p_2)+(1-\bar{\zeta})f_{\bar{\mathrm{b}}}(p_1)f_{\bar{\mathrm{b}}}(p_2)$$
$$-[(1-\bar{\zeta})+(1-\zeta)]f_\varphi(p_1)f_\varphi(p_2)\}|\mathcal{M}_0(p_1,p_2,p_3,p_4)|^2]\,, \tag{2.2.6a}$$

$$\frac{\mathrm{d}Y_{\mathrm{B}}}{\mathrm{d}t}=\frac{1}{2n_\gamma}\Lambda_{12}^{34}[\{2f_{\bar{\mathrm{b}}}(p_1)f_{\bar{\mathrm{b}}}(p_2)(1+\bar{\zeta})-2f_{\mathrm{b}}(p_1)f_{\mathrm{b}}(p_2)(1+\zeta)$$
$$+[(1-\bar{\zeta})-(1-\zeta)]f_\varphi(p_1)f_\varphi(p_2)-(1-\zeta)f_{\mathrm{b}}(p_1)f_{\mathrm{b}}(p_2)$$
$$+(1-\bar{\zeta})f_{\bar{\mathrm{b}}}(p_1)f_{\bar{\mathrm{b}}}(p_2)\}|\mathcal{M}_0(p_1,p_2,p_3,p_4)|^2]\,. \tag{2.2.6b}$$

Using the 4-momentum conservation δ function in Λ, one may write here

$$f_{\overset{(-)}{b}}(p_1)f_{\overset{(-)}{b}}(p_2) = e^{-(E_1+E_2)/T}\, e^{\overset{(+)}{-}2\mu/T}$$
$$= f_{\varphi}^{eq}(p_3)f_{\varphi}^{eq}(p_4)\, e^{\overset{(+)}{-}2\mu/T} \tag{2.2.7}$$

where we have defined

$$f_{A}^{eq}(p_A) = e^{-E_A/T}\,, \tag{2.2.8}$$

which is the phase-space distribution for a species of particles (perhaps massive) in thermal equilibrium at temperature T, and with zero chemical potential. The total equilibrium number density is given by (see appendix C)

$$\begin{aligned} n_A^{eq} &= \int \frac{d^3p_A}{(2\pi)^3} f_A^{eq}(p_A) \\ &= \int \frac{d^3p_A}{(2\pi)^3} e^{-E_A/T} = \int \frac{d^3p_A}{(2\pi)^3} \exp\left(-\sqrt{p_A^2+m_A^2}/T\right) \\ &= \frac{T^3}{2\pi^2}\int_{x_A}^{\infty} z\sqrt{z^2-x_A^2}\, e^{-z}\, dz \\ &= \frac{T^3}{2\pi^2} x_A^2 K_2(x_A)\,, \end{aligned}$$
$$x_A \equiv m_A/T\,, \tag{2.2.9}$$

where K_2 is a modified Bessel function (see appendix C) [as $x\to 0$, $x^2K_2(x)\to 2$, and (2.2.9) reverts to the massless result (2.1.13)]. We assume here that $n_B/n_\gamma \ll 1$, and may thus write (2.2.7) in the form

$$f_{\overset{(-)}{b}}(p_1)f_{\overset{(-)}{b}}(p_2) \simeq f_{\varphi}^{eq}(p_3)f_{\varphi}^{eq}(p_4)(1 \overset{+}{\underset{(-)}{}} Y_B)\,. \tag{2.2.10}$$

(Non-linear terms in Y_B for the model of subsect. 2.3 are discussed in subsect. 2.4.3.) Hence (2.2.6) becomes

$$\begin{aligned} \frac{dY_\varphi}{dt} \simeq \frac{1}{n_\gamma} \Lambda_{12}^{34}[\{ &-[(1-\bar\zeta)+(1-\zeta)]f_\varphi(p_1)f_\varphi(p_2) \\ &+[(1-\zeta)(1+Y_B)+(1-\bar\zeta)(1-Y_B)]f_\varphi^{eq}(p_1)f_\varphi^{eq}(p_2)\} \\ &\times |\mathcal{M}_0(p_1,p_2,p_3,p_4)|^2]\,, \end{aligned} \tag{2.2.10a}$$

$$\begin{aligned} \frac{dY_B}{dt} \simeq \frac{1}{2n_\gamma} \Lambda_{12}^{34}[\{ &[(3+\bar\zeta)(1-Y_B)-(3+\zeta)(1+Y_B)]f_\varphi^{eq}(p_1)f_\varphi^{eq}(p_2) \\ &+(\zeta-\bar\zeta)f_\varphi(p_1)f_\varphi(p_2)\}|\mathcal{M}_0(p_1,p_2,p_3,p_4)|^2]\,. \end{aligned} \tag{2.2.10b}$$

The integration over the phase space available for the final momenta p_3 and p_4

introduces the total $2\to 2$ scattering cross sections

$$v\sigma(a_1a_2\to a_3a_4)=\frac{1}{(2E_{a_1})(2E_{a_2})}\int\frac{d^3p_{a_3}}{2E_{a_3}}\int\frac{d^3p_{a_4}}{2E_{a_4}}\times\frac{\delta^4(p_{a_1}+p_{a_2}-p_{a_3}-p_{a_4})}{(2\pi)^2}|\mathcal{M}(a_1a_2\to a_3a_4)|^2\,, \tag{2.2.11}$$

where v is the relative velocity of the incoming particles a_1 and a_2, which depends only on $E_{a_1}+E_{a_2}$. Eq. (2.2.10) then becomes

$$\frac{dY_\varphi}{dt}\simeq n_\gamma\langle\sigma_0 v\rangle\left\{2\left[1-\left(\frac{\zeta+\bar{\zeta}}{2}\right)\right][(Y_\varphi^{eq})^2-Y_\varphi^2]-(\zeta-\bar{\zeta})(Y_\varphi^{eq})^2Y_B\right\}, \tag{2.2.12a}$$

$$\frac{dY_B}{dt}\simeq n_\gamma\langle\sigma_0 v\rangle\left\{\left(\frac{\zeta-\bar{\zeta}}{2}\right)[Y_\varphi^2-(Y_\varphi^{eq})^2]-\left[3+\left(\frac{\zeta+\bar{\zeta}}{2}\right)\right](Y_\varphi^{eq})^2Y_B\right\}, \tag{2.2.12b}$$

where $\langle\sigma_0 v\rangle$ denotes the average of the cross section $\sigma_0 v$ over the incoming energy distribution. The first term in eq. (2.2.12a) is simply $n_\gamma\langle\sigma_{tot}(\varphi\varphi)v\rangle[(Y_\varphi^{eq})^2-Y_\varphi^2]$ and is familiar from studies of the survival of stable heavy particles produced in the early universe [13]. The second term in eq. (2.2.12a) contains the two small parameters $\zeta-\bar{\zeta}$ and Y_B, and may usually be ignored. The first term in eq. (2.2.12b) is approximately $-n_\gamma\langle(\sigma(\varphi\varphi\to bb)-\sigma(\varphi\varphi\to\bar{b}\bar{b}))/\sigma_{tot}(\varphi\varphi))v\rangle\, dY_\varphi/dt$ and accounts for the small disparity between b and $\bar{b}$ production in $\varphi\varphi$ annihilation. The second term in eq. (2.2.12b) is proportional to the total cross section for baryon-number violating interactions, and causes Y_B to relax towards zero when the system is in thermal equilibrium ($Y_\varphi=Y_\varphi^{eq}$). Note that the rate of baryon number generation (2.2.12b) is proportional to the deviation of the φ number density from its equilibrium value; if $m_\varphi=0$, then in the present model, the expansion of the universe cannot alter $f_\varphi(p)=e^{-p/T}$, and no net baryon number results [5]. (In subsect. 2.4 we discuss inhomogenities and differential heating effects which may produce non-zero baryon number even if all particles participating directly in B-violating processes are massless.)

2.3. A SIMPLE MODEL*

As in the previous model, let b and $\bar{b}$ be nearly massless particles carrying baryon numbers $B=\frac{1}{2}$ and $B=-\frac{1}{2}$ respectively. Let $\overset{(-)}{X}$ be some massive boson which mediates baryon-number violating interactions. We take the decay amplitudes for

* A model similar to this was proposed in ref. [6].

the $\overset{(-)}{\mathrm{X}}$ to be

$$
\begin{aligned}
|\mathcal{M}(\mathrm{X}\to \mathrm{bb})|^2 &= (1+\eta)\tfrac{1}{2}|\mathcal{M}_0|^2\,,\\
|\mathcal{M}(\mathrm{X}\to \bar{\mathrm{b}}\bar{\mathrm{b}})|^2 &= (1-\eta)\tfrac{1}{2}|\mathcal{M}_0|^2\,,\\
|\mathcal{M}(\bar{\mathrm{X}}\to \bar{\mathrm{b}}\bar{\mathrm{b}})|^2 &= (1+\bar{\eta})\tfrac{1}{2}|\mathcal{M}_0|^2\,,\\
|\mathcal{M}(\bar{\mathrm{X}}\to \mathrm{bb})|^2 &= (1-\bar{\eta})\tfrac{1}{2}|\mathcal{M}_0|^2\,,
\end{aligned}
\tag{2.3.1}
$$

where now $|\mathcal{M}_0|^2$ is of order α. The parametrization (2.3.1) respects the *CPT*-invariance constraint (2.1.5). *CP* invariance would imply $\eta=\bar{\eta}$, but we take $\eta-\bar{\eta}=\mathrm{O}(\alpha)\neq 0$. Hence a state initially containing an equal number of X and $\bar{\mathrm{X}}$ will decay, in the absence of back reactions, to a system with a net baryon number $n_{\mathrm{B}}\simeq(\eta-\bar{\eta})\frac{1}{2}(n^0_{\mathrm{X}}+n^0_{\bar{\mathrm{X}}})$. (Back reactions can be ignored if the $\overset{(-)}{\mathrm{X}}$ are emitted as thermal radiation into an infinite vacuum, or are concentrated into a beam.) *CPT* conjugation gives the rates for the inverse decay processes

$$
\begin{aligned}
|\mathcal{M}(\mathrm{bb}\to \mathrm{X})|^2 &= |\mathcal{M}(\bar{\mathrm{X}}\to \bar{\mathrm{b}}\bar{\mathrm{b}})|^2 = (1+\bar{\eta})\tfrac{1}{2}|\mathcal{M}_0|^2\\
|\mathcal{M}(\mathrm{bb}\to \bar{\mathrm{X}})|^2 &= |\mathcal{M}(\mathrm{X}\to \bar{\mathrm{b}}\bar{\mathrm{b}})|^2 = (1-\eta)\tfrac{1}{2}|\mathcal{M}_0|\\
|\mathcal{M}(\bar{\mathrm{b}}\bar{\mathrm{b}}\to \mathrm{X})|^2 &= |\mathcal{M}(\bar{\mathrm{X}}\to \mathrm{bb})|^2 = (1-\bar{\eta})\tfrac{1}{2}|\mathcal{M}_0|^2\\
|\mathcal{M}(\bar{\mathrm{b}}\bar{\mathrm{b}}\to \bar{\mathrm{X}})|^2 &= |\mathcal{M}(\mathrm{X}\to \mathrm{bb})|^2 = (1+\eta)\tfrac{1}{2}|\mathcal{M}_0|^2\,.
\end{aligned}
\tag{2.3.2}
$$

Note that if X and $\bar{\mathrm{X}}$ decay preferentially produce b (i.e., $\eta>\bar{\eta}$), then according to *CPT* invariance, inverse decay processes must preferentially destroy $\bar{\mathrm{b}}$. Thus, if only decay and inverse decay are considered, a system even in thermal equilibrium cannot fail to generate a net baryon number. (This rather relevant point has also been noted in ref. [14], but appears to have been neglected elsewhere.) However, according to eq. (2.1.4), which follows purely from *CPT* invariance and unitarity, no excess of b over $\bar{\mathrm{b}}$ can develop in thermal equilibrium. We take the total rate for $\overset{(-)}{\mathrm{X}}$ decay to be $\sim\alpha m_{\mathrm{X}}$. Then eq. (2.1.9) shows that *CP*-violating effects in these decays must be at least $\mathrm{O}(\alpha^2)$ (hence $\eta-\bar{\eta}=\mathrm{O}(\alpha)$). Eq. (2.1.4) applies only when summed over all possible initial states j which can produce $\overset{(-)}{i}$ to a given order in α. Decay and inverse decay are, however, not the only possible interactions between $\overset{(-)}{\mathrm{X}}$ and $\overset{(-)}{\mathrm{b}}$ to $\mathrm{O}(\alpha^2)$: $2\to 2$ scattering processes, such as $\mathrm{bb}\to\bar{\mathrm{b}}\bar{\mathrm{b}}$ mediated by s-channel $\overset{(-)}{\mathrm{X}}$ exchange, also occur. We shall below that after including these processes, eq. (2.1.4) is respected, and no baryon excess develops in thermal equilibrium.

Although the X and $\bar{\mathrm{X}}$ in (2.3.1) are taken to have identical decay modes, we shall, for simplicity ignore any mixing between them (until subsect. 2.4.4). This may be enforced by considering two species of b (each with $B=\frac{1}{2}$), and taking $\mathrm{X}\to \mathrm{b}_1\mathrm{b}_1$, $\mathrm{X}\to\bar{\mathrm{b}}_2\bar{\mathrm{b}}_2$, $\bar{\mathrm{X}}\to\bar{\mathrm{b}}_1\bar{\mathrm{b}}_1$, $\bar{\mathrm{X}}\to \mathrm{b}_2\mathrm{b}_2$. The formulae below are unaffected by these distinctions. Note that, as discussed in subsect. 2.4, the model of this section may be slightly simplified by taking X and $\bar{\mathrm{X}}$ to be indistinguishable, so that $\eta=-\bar{\eta}$, and X is an eigenstate of CP. Then CP invariance requires $|\mathcal{M}(\mathrm{X}\to\mathrm{bb})|^2 = |\mathcal{M}(\mathrm{X}\to\bar{\mathrm{b}}\bar{\mathrm{b}})|^2$, or $\eta = 0$. This case is exemplified by semileptonic K_L^0 decay, where CP violation is revealed in $\Gamma(\mathrm{K}_\mathrm{L}^0\to\pi^+\mathrm{e}^-\bar{\nu}) \neq \Gamma(\mathrm{K}_\mathrm{L}^0\to\pi^-\mathrm{e}^+\nu)$.

2.3.1. The $\overset{(-)}{\mathrm{X}}$ number density. In calculating the time evolution of the $\overset{(-)}{\mathrm{X}}$ number density, we work to $\mathrm{O}(\alpha)$, at which only the decay and inverse decay processes of equations (2.3.1) and (2.3.2) contribute. In addition to these baryon number violating interactions, the $\overset{(-)}{\mathrm{X}}$ may also undergo baryon number conserving interactions (such as $\gamma\mathrm{X}\to\gamma\mathrm{X}$ or $\gamma\gamma\to\mathrm{X}\bar{\mathrm{X}}$) with other particles in the universe. Typically, these processes will be $\mathrm{O}(\alpha^2)$, and therefore occur on a longer time-scale than the X decays we consider. However, it is possible that the typical coupling constants involved are larger than for the B-violating decays, or that the number of light particle species $\gtrsim \mathrm{O}(1/\alpha)$, so that $\overset{(-)}{\mathrm{X}}$ may undergo several B-conserving scatterings before decay. In this case, as with the $\overset{(-)}{\mathrm{b}}$ in the model of subsect. 2.2, the $\overset{(-)}{\mathrm{X}}$ will be brought into kinetic equilibrium before they decay, and assume a Maxwell–Boltzmann distribution in phase space, so that

$$f_{\overset{(-)}{\mathrm{X}}}(px) \simeq \exp\left[-(E_{\overset{(-)}{\mathrm{X}}} - \mu_{\overset{(-)}{\mathrm{X}}})/T\right]. \tag{2.3.3}$$

Processes such as $\gamma\gamma\to\mathrm{X}\bar{\mathrm{X}}$ would lead to $\mu_\mathrm{X} = \mu_{\bar{\mathrm{X}}}$. The chemical potentials $\mu_{\overset{(-)}{\mathrm{X}}}$ in eq. (2.3.3) are in any case determined by the processes (2.3.1) and (2.3.2) in which single $\overset{(-)}{\mathrm{X}}$ are created or destroyed.

To $\mathrm{O}(\alpha)$, the X number density evolves with time according to the equation [analogous to (2.2.3a)]

$$\frac{\mathrm{d}n_\mathrm{X}}{\mathrm{d}t} + \frac{3\dot{R}}{R} n_\mathrm{X} = \Lambda_{12}^{\mathrm{X}}\left[-f_\mathrm{X}(p_\mathrm{X})|\mathcal{M}(\mathrm{X}\to\mathrm{bb})|^2 - f_\mathrm{X}(p_\mathrm{X})|\mathcal{M}(\mathrm{X}\to\bar{\mathrm{b}}\bar{\mathrm{b}})|^2 + f_\mathrm{b}(p_1)f_\mathrm{b}(p_2)|\mathcal{M}(\mathrm{bb}\to\mathrm{X})|^2 + f_{\bar{\mathrm{b}}}(p_1)f_{\bar{\mathrm{b}}}(p_2)|\mathcal{M}(\bar{\mathrm{b}}\bar{\mathrm{b}}\to\mathrm{X})|^2\right]. \tag{2.3.4}$$

The first two terms on the right-hand side account for X decays, while the second two represent inverse decay processes. In addition to the reactions (2.3.1) and (2.3.2) in which a net baryon number is created or destroyed the $\overset{(-)}{\mathrm{b}}$ also undergo baryon-conserving interactions with other particles in the universe. These B-conserving processes (such as $\gamma\mathrm{b}\to\gamma\mathrm{b}$) occur at a rate $\mathrm{O}(\alpha^2)$; at temperatures $T \ll m_\mathrm{X}$ they are much faster than any B-violating interactions mediated by $\overset{(-)}{\mathrm{X}}$ exchange. The relative rates of the various processes at high temperatures are discussed in sect. 4, and it seems likely that in most cases, B-conserving reactions occur with larger rates than do B-violating ones. Hence the $\overset{(-)}{\mathrm{b}}$ should be Maxwell–Boltzmann distributed in phase space as in eq. (2.2.2), with their chemical potential determined by B-violating processes. Using the momentum conservation δ function in Λ, one may then write (assuming $Y_\mathrm{B} \ll 1$)

$$f_{\overset{(-)}{\mathrm{b}}}(p_1)f_{\overset{(-)}{\mathrm{b}}}(p_2) \simeq f_\mathrm{X}^{\mathrm{eq}}(p_\mathrm{X})(1 \overset{(-)}{+} Y_\mathrm{B}), \tag{2.3.5}$$

where, as above, $f_{\rm X}^{\rm eq}(p_{\rm X}) = {\rm e}^{-E_{\rm X}/T}$ is the distribution of X in thermal equilibrium at temperature T and with zero chemical potential. On inserting eq. (2.3.5) into eq. (2.3.4), the p_1 and p_2 integrations are weighted only by the matrix element and the available phase space; they therefore yield simply the total decay rates*

$$\Gamma({\rm A}\to {\rm a}_1{\rm a}_2) = \frac{1}{2E_{\rm A}}\int\frac{{\rm d}^3p_{{\rm a}_1}}{2E_{{\rm a}_1}}\int\frac{{\rm d}^3p_{{\rm a}_2}}{2E_{{\rm a}_2}}\frac{\delta^4(p_{\rm A}-p_{{\rm a}_1}-p_{{\rm a}_2})}{(2\pi)^2}|\mathcal{M}({\rm A}\to {\rm a}_1{\rm a}_2)|^2$$

$$= \frac{m_{\rm A}}{E_{\rm A}}\Gamma_{\rm R}({\rm A}\to {\rm a}_1{\rm a}_2) = \frac{|\mathcal{M}({\rm A}\to {\rm a}_1{\rm a}_2)|^2}{16\pi E_{\rm A}}, \tag{2.3.6}$$

where $\Gamma_{\rm R}$ is the rate measured in the rest frame of the decaying A. Eq. (2.3.4) may then be written in the form

$$\frac{{\rm d}n_{\rm X}}{{\rm d}t}+\frac{3\dot R}{R}n_{\rm X} = -\int\frac{{\rm d}^3p_{\rm X}}{(2\pi)^3}(f_{\rm X}(p_{\rm X})-f_{\rm X}^{\rm eq}(p_{\rm X}))\{\Gamma({\rm X}\to {\rm bb})+\Gamma({\rm X}\to \bar{\rm b}\bar{\rm b})\}$$

$$- Y_{\rm B}\int\frac{{\rm d}^3p_{\rm X}}{(2\pi)^3}f_{\rm X}^{\rm eq}(p_{\rm X})\{\Gamma(\bar{\rm X}\to {\rm bb})-\Gamma(\bar{\rm X}\to\bar{\rm b}\bar{\rm b})\}. \tag{2.3.7}$$

Performing the final $p_{\rm X}$ integration, and using the parametrization (2.3.2) then gives

$$\frac{{\rm d}Y_{\rm X}}{{\rm d}t} \simeq -\langle\Gamma_{\rm X}\rangle\{(Y_{\rm X}-Y_{\rm X}^{\rm eq})-\bar\eta Y_{\rm B}Y_{\rm X}^{\rm eq}\}, \tag{2.3.8}$$

where $\langle\Gamma_{\rm X}\rangle$ denotes the total X decay rate averaged over the time-dilation factors for the decaying particles. [In writing eq. (2.3.8) we have made the approximation that the actual X momentum distribution does not differ from the $\mu = 0$ equilibrium form sufficiently to affect the time-dilation factor. This is certainly the case if thermalizing reactions occur sufficiently fast to produce an $\overset{(-)}{\rm X}$ distribution of the form (2.3.3). Note that we assume all decaying $\overset{(-)}{\rm X}$ to be exactly on-shell; in practice they should have a distribution of invariant masses peaked at $m_{\rm X}$ and with a width $\sim\Gamma_{\rm X}\sim\alpha m_{\rm X}$**: we make the narrow resonance approximation $\Gamma_{\rm X} \ll m_{\rm X}$.] Charge conjugation ($n_{\rm X}\to n_{\bar{\rm X}}$, $n_{\rm B}\to -n_{\rm B}$, $\eta\to\bar\eta$; $\Gamma_{\rm X}=\Gamma_{\bar{\rm X}}$ and $m_{\rm X}=m_{\bar{\rm X}}$ so that $n_{\rm X}^{\rm eq}=n_{\bar{\rm X}}^{\rm eq}$ by CPT invariance) gives the corresponding equation for the $\bar{\rm X}$ number density

$$\frac{{\rm d}Y_{\bar{\rm X}}}{{\rm d}t} \simeq -\langle\Gamma_{\rm X}\rangle\{(Y_{\bar{\rm X}}-Y_{\rm X}^{\rm eq})+\eta Y_{\rm B}Y_{\rm X}^{\rm eq}\}. \tag{2.3.9}$$

* We may take the two-body decay matrix element to be independent of any momenta. Even $\varepsilon\cdot p$ terms must vanish on averaging over initial and final spins.

** In addition to the finite intrinsic width of the X resonance, there should be additional collision broadening at high temperatures; the resonance width should become of order the inverse time between collisions $\sim\alpha^2 T$ or $\alpha^2(T^3/m_{\rm X}^2)$ (see sect. 4). However, most X decay at temperatures $\sim\alpha m_{\rm X}$, where such effects are probably negligible. (Doppler broadening, familiar from the spectra of hot gases, is irrelevant here; it serves only to smear the energies of $\overset{(-)}{\rm b}$ emitted in $\overset{(-)}{\rm X}$ decay.)

It is convenient to write eqs. (2.3.8) and (2.3.9) in terms of

$$Y_+ = \tfrac{1}{2}(Y_X + Y_{\bar{X}}) \tag{2.3.10a}$$

and

$$Y_- = \tfrac{1}{2}(Y_X - Y_{\bar{X}}): \tag{2.3.10b}$$

$$\frac{dY_+}{dt} \simeq -\langle\Gamma_X\rangle\left\{(Y_+ - Y_+^{eq}) + \left(\frac{\eta - \bar{\eta}}{2}\right) Y_B Y_+^{eq}\right\},$$

$$\frac{dY_-}{dt} \simeq -\langle\Gamma_X\rangle\left\{Y_- - \left(\frac{\eta + \bar{\eta}}{2}\right) Y_B Y_+^{eq}\right\}. \tag{2.3.10c}$$

2.3.2. The $\overset{(-)}{b}$ number density. The Boltzmann equation for the b number density in the model of equations (2.3.1) and (2.3.2) is

$$\begin{aligned}\frac{dn_b}{dt} + 3\frac{\dot{R}}{R} n_b = \Lambda_{12}^{X}\, &[+f_X(p_X)|\mathcal{M}(X\to bb)|^2 + f_{\bar{X}}(p_X)|\mathcal{M}(\bar{X}\to bb)|^2 \\ &- f_b(p_1)f_b(p_2)|\mathcal{M}(bb\to X)|^2 - f_b(p_1)f_b(p_2)|\mathcal{M}(bb\to \bar{X})|^2] \\ &+ \Lambda_{12}^{34}[+f_{\bar{b}}(p_1)f_{\bar{b}}(p_2)|\mathcal{M}'(\bar{b}\bar{b}\to bb)|^2 - f_b(p_1)f_b(p_2)|\mathcal{M}'(bb\to \bar{b}\bar{b})|^2]\,.\end{aligned} \tag{2.3.11}$$

The first term in eq. (2.3.11) accounts for the decay and inverse decay processes $X\to bb$, $\bar{X}\to bb$, $bb\to X$ and $bb\to\bar{X}$. The second term in (2.3.11) accounts for those $2\to 2$ scattering processes that are not already included as successive inverse decay and decay processes, (as would $bb\to X\to \bar{b}\bar{b}$, with a *real* intermediate X.) The amplitude for $bb\to\bar{b}\bar{b}$ due to s-channel exchange of a single X contains two terms: a part corresponding to the propagation of an on-shell intermediate X (which is important only when the incoming energies lie within the X resonance curve), and, as usual, a part accounting for off-shell X exchange. [The t- and u-channel exchange diagrams at lowest order receive no contributions from physical intermediate states. Note that processes such as $b\to bbb$ are energetically forbidden in (2.3.11).] We write

$$|\mathcal{M}'(a_1a_2\to a_3a_4)|^2 = |\mathcal{M}(a_1a_2\to a_3a_4)|^2 - |\mathcal{M}_{RIS}(a_1a_2\to a_3a_4)|^2\,, \tag{2.3.12}$$

where $\mathcal{M}_{RIS}$ denotes the contribution from physical intermediate states, already included in the first term of the Boltzmann equation (2.3.11) as successive lower-order ($2\to 1$ and $1\to 2$) processes.

Subtracting from eq. (2.3.11) the charge-conjugated equation for the $\bar{b}$ number density, we obtain an equation for the evolution of the total baryon number density $Y_B = Y_b - Y_{\bar{b}} = (n_b - n_{\bar{b}})/n_\gamma$:

$$\begin{aligned}\frac{dY_B}{dt} &\simeq \langle\Gamma_X\rangle\{\eta Y_X - \bar{\eta} Y_{\bar{X}} + (\eta - \bar{\eta})Y_X^{eq} - 2Y_B Y_X^{eq}\} \\ &\quad - \frac{2}{n_\gamma}\Lambda_{12}^{34}\{e^{-(E_1+E_2)/T}(|\mathcal{M}'(b_1b_2\to\bar{b}\bar{b})|^2 - |\mathcal{M}'(\bar{b}_1\bar{b}_2\to bb)|^2)\}\end{aligned}$$

$$-\frac{2Y_B}{n_\gamma}\Lambda_{12}^{34}\{e^{-(E_1+E_2)/T}(|\mathcal{M}'(b_1b_2\to\bar{b}\bar{b})|^2+|\mathcal{M}'(b_1b_2\to bb)|^2)\}\,. \tag{2.3.13}$$

Notice that, as mentioned above, even when the $\overset{(-)}{X}$ are in thermal equilibrium, so that $Y_X = Y_{\bar{X}} = Y_X^{eq}$, the two terms on the right-hand side of this equation do not individually vanish even when $Y_B = 0$: the $2\to 2$ scattering processes must conspire with decay and inverse decay processes to maintain thermal equilibrium.

In the model of this section, the only B-violating $2\to 2$ reactions which may occur to $O(\alpha^2)$ are $bb\to\bar{b}\bar{b}$ and $\bar{b}\bar{b}\to bb$. But the unitarity requirement (2.1.3) then demands $|\mathcal{M}(bb\to\bar{b}\bar{b})|^2 = |\mathcal{M}(\bar{b}\bar{b}\to bb)|^2$ for the total matrix elements of these processes. However, in the $|\mathcal{M}'|^2$ which actually enter the Boltzmann equation (2.3.13), the part $|\mathcal{M}_{RIS}|^2$, which arises from real intermediate X exchanges already accounted for by the first term of eq. (2.3.13), has been subtracted out. Unlike the total $|\mathcal{M}|^2$, $|\mathcal{M}_{RIS}|^2$ (and hence $|\mathcal{M}'|^2$) may differ at $O(\alpha^2)$ between $bb\to\bar{b}\bar{b}$ and $\bar{b}\bar{b}\to bb$. In the narrow-width approximation (which has already been made in (2.3.13) by assigning the decaying X and definite mass m_X) the contributions of a real intermediate X to the process $bb\to\bar{b}\bar{b}$ and $\bar{b}\bar{b}\to bb$ become

$$\begin{aligned}|\mathcal{M}_{RIS}(b_1b_2\to\bar{b}\bar{b})|^2 &\simeq \frac{\pi}{m_X\Gamma_X}\,\delta((p_{b_1}+p_{b_2})^2-m_X^2)\\ &\quad\times\{|\mathcal{M}(bb\to X)|^2|\mathcal{M}(X\to\bar{b}\bar{b})|^2+|\mathcal{M}(bb\to\bar{X})|^2|\mathcal{M}(\bar{X}\to\bar{b}\bar{b})|^2\}\\ &= \frac{\pi|\mathcal{M}_0|^4}{2m_X\Gamma_X}\,\delta((p_{b_1}+p_{b_2})^2-m_X^2)(1-\eta)(1+\bar{\eta})\\ &= -|\mathcal{M}'(b_1b_2\to\bar{b}\bar{b})|^2+|\mathcal{M}(b_1b_2\to\bar{b}\bar{b})|^2\,,\end{aligned} \tag{2.3.14a}$$

$$\begin{aligned}|\mathcal{M}_{RIS}(\bar{b}_1\bar{b}_2\to bb)|^2 &\simeq \frac{\pi|\mathcal{M}_0|^4}{2m_X\Gamma_X}\,\delta((p_{b_1}+p_{b_2})^2-m_X^2)(1-\bar{\eta})(1+\eta)\\ &= -|\mathcal{M}'(\bar{b}_1\bar{b}_2\to bb)|^2+|\mathcal{M}(\bar{b}_1\bar{b}_2\to bb)|^2\,.\end{aligned} \tag{2.3.14b}$$

Because of the $1/\Gamma_X \sim 1/\alpha m_X$ factor arising from the integral under the X resonance curve, these terms are of order α, rather than $O(\alpha^2)$ as expected for $2\to 2$ scatterings. Using the fact that in our model $|\mathcal{M}(bb\to\bar{b}\bar{b})|^2 = |\mathcal{M}(\bar{b}\bar{b}\to bb)|^2$ for the total amplitudes [in general these terms may have a *CP*-violating difference $O(\alpha^3)$], we may write the difference appearing in the second term on the right-hand side of eq. (2.3.13) as

$$\begin{aligned}|\mathcal{M}'(b_1b_2\to\bar{b}\bar{b})|^2-|\mathcal{M}'(\bar{b}_1\bar{b}_2\to bb)|^2 &= |\mathcal{M}_{RIS}(\bar{b}_1\bar{b}_2\to bb)|^2-|\mathcal{M}_{RIS}(b_1b_2\to\bar{b}\bar{b})|^2\\ &\simeq \frac{\pi(\eta-\bar{\eta})|\mathcal{M}_0|^4}{m_X\Gamma_X}\,\delta((p_{b_1}+p_{b_2})^2-m_X^2)\,.\end{aligned} \tag{2.3.14c}$$

Then the complete second term in eq. (2.3.13) may be written in the form

$$-\frac{2}{n_\gamma}\frac{1}{(2\pi)^8}\int\frac{d^3p_1}{2E_1}\int\frac{d^3p_2}{2E_2}\int\frac{d^3p_3}{2E_3}\int\frac{d^3p_4}{2E_4}\delta^4(p_1+p_2-p_3-p_4)$$

$$\times f_X^{eq}(p_1+p_2)\frac{\pi|\mathcal{M}_0|^4}{m_X\Gamma_X}\delta((p_1+p_2)^2-m_X^2)(\eta-\bar{\eta})\,, \tag{2.3.15}$$

where the X mass-shell delta function has allowed us to replace $\exp(-[E_1+E_2]/T^2)$ by $f_X^{eq}(p_1+p_2)$. [The extra $O(\alpha^3)$ terms which may in general appear in (2.3.15) from *CP*-violating loop corrections to genuine $2\to 2$ scattering processes may contain parts not proportional to f_X^{eq}, but these cannot be retained consistently in view of other approximations.] To simplify this we apply the results

$$\Gamma_X=\frac{1}{2m_X}\int\frac{d^3p_{b_1}}{2E_{b_1}}\int\frac{d^3p_{b_2}}{2E_{b_2}}\frac{\delta^4(p_X-p_{b_1}-p_{b_2})}{(2\pi)^2}|\mathcal{M}_0|^2=\frac{|\mathcal{M}_0|^2}{16\pi m_X}\,, \tag{2.3.16}$$

and (from appendix C)

$$\int\frac{d^3p_1}{2E_1}\int\frac{d^3p_2}{2E_2}f_X^{eq}(p_1+p_2)\delta((p_1+p_2)^2-m_X^2)=\left(\frac{\langle\Gamma_X\rangle}{\Gamma_X}\right)n_X^{eq}\frac{2\pi^4}{m_X}\,, \tag{2.3.17}$$

so that eq. (2.3.15) becomes just

$$-2Y_X^{eq}(\eta-\bar{\eta})\langle\Gamma_X\rangle\,, \tag{2.3.18}$$

and thus elegantly cancels the first term of (2.3.13) in thermal equilibrium, as required by eq. (2.1.4). (This seemingly miraculous result may formally be obtained by considering the sum of double ($t=\pm\infty$) cuts in vacuum diagrams with finite temperature propagators, without treating separately one- and two-body initial states as is done here.)

The last term of eq. (2.3.13) may be written using eq. (2.2.11) in the form

$$-2Y_B n_\gamma\langle v\sigma'(bb\to\bar{b}\bar{b})+v\sigma'(\bar{b}\bar{b}\to bb)\rangle\,, \tag{2.3.19}$$

where in σ' the contribution from real intermediate X exchange in the s-channel has been subtracted out by replacing the full X propogator by its principal part. In addition, since $Y_B\ll 1$, $n_{(\bar{b})}$ is approximated by n_γ. [No Y_X^{eq} factor, as in eq. (2.3.18), appears here, since the incoming c.m.s. energy is no longer constrained to be m_X, and at low temperatures will be much smaller.]

Finally, therefore, eq. (2.3.13) for the time development of the baryon number density may be written in the simple form

$$\frac{dY_B}{dt}\simeq\langle\Gamma_X\rangle\{(\eta-\bar{\eta})(Y_+-Y_+^{eq})+(\eta+\bar{\eta})Y_-\}$$

$$-2Y_B\{\langle\Gamma_X\rangle Y_+^{eq}+n_\gamma\langle v\{\sigma'(bb\to\bar{b}\bar{b})+\sigma'(\bar{b}\bar{b}\to bb)\}\rangle\}\,, \tag{2.3.20}$$

$$Y_\pm=\tfrac{1}{2}(Y_X\pm Y_{\bar{X}})\,;$$

as expected from the discussion of subsect. 2.1, the rate of baryon generation vanishes when the system is in thermal equilibrium ($Y_X = Y_{\bar{X}} = Y_X^{eq}$), while any pre-existing baryon number is destroyed at a rate governed by the total rate for B-violating processes. Note that any possible *CP* violation in the $2 \to 2$ reactions would be ineffective at producing a baryon excess, since the masslessness of the $\bar{b}$ prevents deviations from thermal equilibrium [5].

In sect. 3 we discuss the solution of eq. (2.3.10) and (2.3.20). First, however, we consider some possible complications.

2.4. COMPLICATIONS

2.4.1. More particles and more decay modes. The evolution of the number density $n_\chi = Y_\chi n_\gamma$ of a massive particle species χ due to decay and inverse decay processes is given in direct analogy with eq. (2.3.4) by

$$\frac{\mathrm{d}Y_\chi}{\mathrm{d}t} = \Lambda_{\mathrm{k}}^{\chi}\left[-\sum_{\mathrm{k}} f_\chi(p_\chi)|\mathcal{M}(\chi \to \mathrm{k})|^2 + \sum_{\mathrm{k}}\left(\prod_{\beta\in\mathrm{k}} f_\beta(p_\beta)\right)|\mathcal{M}(\mathrm{k} \to \chi)|^2\right]. \tag{2.4.1}$$

But, so long as all excesses of particles over antiparticles are small,

$$\prod_{\substack{\beta\in\mathrm{k} \\ \sum p_\beta = p_\chi}} f_\beta(p_\beta) = f_\chi^{eq}(p_\chi)\left(1 + \sum_\beta \frac{\mu^\beta}{T}\right)$$

$$= f_\chi^{eq}(p_\chi)(1 + \tfrac{1}{2}\sum (Y_\beta - Y_{\bar{\beta}}))\,, \tag{2.4.2}$$

and hence

$$\frac{\mathrm{d}Y_\chi}{\mathrm{d}t} = \langle\Gamma_\chi\rangle(-Y_\chi + Y_\chi^{eq}) - Y_\chi^{eq}\sum_{\mathrm{k}}\left\{\left(\sum_\beta (N_\beta - N_{\bar{\beta}})_{\mathrm{k}}\left(\frac{Y_\beta - Y_{\bar{\beta}}}{2}\right)\right)\langle\Gamma(\bar{\chi} \to \mathrm{k})\rangle\right\}, \tag{2.4.3}$$

where $(N_\beta)_{\mathrm{k}}$ denotes the number of β particles in the state k. This equation holds even if several massive species χ are present. (If some χ may mix, further complications may occur, as discussed below.) By charge conjugation and subtraction, we obtain from (2.4.3) equations analogous to (2.3.10) (and with corresponding approximations):

$$\begin{aligned}\frac{\mathrm{d}Y_+}{\mathrm{d}t} &\simeq -\Gamma_\chi(Y_+ - Y_+^{eq}) + \tfrac{1}{2}Y_+^{eq}\sum_{\mathrm{k}}\left\{\left(\sum_\beta (N_\beta - N_{\bar{\beta}})_{\mathrm{k}}\left(\frac{Y_\beta - Y_{\bar{\beta}}}{2}\right)\right)\right. \\ &\qquad\qquad \left.\times\langle\Gamma(\bar{\chi}\to\mathrm{k}) - \Gamma(\chi\to\bar{\mathrm{k}})\rangle\right\} \\ &\simeq -\langle\Gamma_\chi\rangle(Y_+ - Y_+^{eq})\,,\end{aligned} \tag{2.4.4a}$$

$$\frac{\mathrm{d}Y_-}{\mathrm{d}t} \simeq -\langle\Gamma_\chi\rangle Y_- + \tfrac{1}{2}Y_+^{eq}\sum_k\left\{\left(\sum_\beta (N_\beta - N_{\bar{\beta}})_{\mathrm{k}}\left(\frac{Y_\beta - Y_{\bar{\beta}}}{2}\right)\right)\langle\Gamma(\bar{\chi}\to\mathrm{k}) + \Gamma(\chi\to\bar{\mathrm{k}})\rangle\right\}, \tag{2.4.4b}$$

$$Y_{\pm}=\tfrac{1}{2}(Y_{\chi}\pm Y_{\bar{\chi}})\,. \tag{2.4.4c}$$

In analogy with eq. (2.3.13), the density of a quantum number B violated in decays and scatterings involving particles χ evolves according to [sums on χ run over both particles and antiparticles ($\bar{\chi}$), and for simplicity we assume that any particle β for which $n_\beta \neq n_{\bar{\beta}}$ carries baryon number]

$$\begin{aligned}
\frac{\mathrm{d}Y_\mathrm{B}}{\mathrm{d}t} &= \sum_{\chi}\left\{\sum_{\mathrm{f}}(N_\mathrm{B})_\mathrm{f}Y_\chi \Lambda[|\mathcal{M}(\chi\to\mathrm{f})|^2-|\mathcal{M}(\chi\to\bar{\mathrm{f}})|^2]\right\}\\
&\quad-\sum_{\chi}\left\{\sum_{\mathrm{f}}(N_\mathrm{B})_\mathrm{f}Y_\chi^{\mathrm{eq}}\Lambda[|\mathcal{M}(\mathrm{f}\to\chi)|^2-|\mathcal{M}(\bar{\mathrm{f}}\to\chi)|^2]\right\}\\
&\quad-Y_\mathrm{B}\sum_{\chi}\left\{\sum_{\mathrm{f}}[(N_\mathrm{B})_\mathrm{f}]^2Y_\chi^{\mathrm{eq}}\Lambda[|\mathcal{M}(\mathrm{f}\to\chi)|^2+|\mathcal{M}(\bar{\mathrm{f}}\to\chi)|^2]\right\}\\
&\quad+\sum_{\mathrm{f},\mathrm{f}'}\left\{[(N_\mathrm{B})_\mathrm{f}-(N_\mathrm{B})_{\mathrm{f}'}]\sum_{\chi}\left\{\frac{Y_\chi^{\mathrm{eq}}}{\Gamma_\chi}\Lambda[|\mathcal{M}(\mathrm{f}\to\chi)|^2|\mathcal{M}(\chi\to\mathrm{f}')|^2\right.\right.\\
&\qquad\qquad\left.\left.-|\mathcal{M}(\bar{\mathrm{f}}\to\chi)|^2|\mathcal{M}(\chi\to\bar{\mathrm{f}}')|^2]\right\}\right\}\\
&\quad-\tfrac{1}{2}Y_\mathrm{B}n_\gamma\sum_{\mathrm{f},\mathrm{f}'}\{[(N_\mathrm{B})_\mathrm{f}-(N_\mathrm{B})_{\mathrm{f}'}](N_\mathrm{B})_\mathrm{f}\Lambda[|\mathcal{M}'(\mathrm{f}\to\mathrm{f}')|^2]\}\,.
\end{aligned} \tag{2.4.5}$$

Using CPT invariance, the second term may immediately be rewritten as

$$+\sum_{\chi}\left\{\sum_{\mathrm{f}}(N_\mathrm{B})_\mathrm{f}Y_\chi^{\mathrm{eq}}\Lambda[|\mathcal{M}(\chi\to\mathrm{f})|^2-|\mathcal{M}(\chi\to\bar{\mathrm{f}})|^2]\right\}, \tag{2.4.6}$$

which is equal to the first term when $Y_\chi = Y_\chi^{\mathrm{eq}}$, while the fourth term may be simplified to

$$\begin{aligned}
&-\sum_{\mathrm{f}}\left\{(N_\mathrm{B})_\mathrm{f}\sum_{\chi}\Lambda\,[|\mathcal{M}(\chi\to\mathrm{f})|^2-|\mathcal{M}(\chi\to\bar{\mathrm{f}})|^2-|\mathcal{M}(\mathrm{f}\to\chi)|^2+|\mathcal{M}(\bar{\mathrm{f}}\to\chi)|^2]\right\}\\
&\qquad=-2\sum_{\mathrm{f}}\left\{(N_\mathrm{B})_\mathrm{f}\sum_{\chi}\Lambda[|\mathcal{M}(\chi\to\mathrm{f})|^2-|\mathcal{M}(\chi\to\bar{\mathrm{f}})|^2]\right\},
\end{aligned} \tag{2.4.7}$$

which has exactly the form necessary to negate the second term, as required by the theorem (2.1.4). The second and fifth terms in eq. (2.4.5) may be written in the manifestly negative forms

$$-Y_\mathrm{B}\sum_{\chi}\left\{\sum_{\mathrm{f}}[(N_\mathrm{B})_\mathrm{f}]^2Y_\chi^{\mathrm{eq}}\,\langle\Gamma(\chi\to\mathrm{f})+\Gamma(\chi\to\bar{\mathrm{f}})\rangle\right\}, \tag{2.4.8a}$$

and

$$\begin{aligned}
&-\tfrac{1}{4}Y_\mathrm{B}n_\gamma\sum_{\mathrm{f},\mathrm{f}'}\{[(N_\mathrm{B})_\mathrm{f}-(N_\mathrm{B})_{\mathrm{f}'}]^2\Lambda[|\mathcal{M}'(\mathrm{f}\to\mathrm{f}')|^2+|\mathcal{M}'(\mathrm{f}'\to\mathrm{f})|^2]\}\\
&\qquad\simeq-\frac{Y_\mathrm{B}n_\gamma}{2}\sum_{\mathrm{f},\mathrm{f}'}\{[(N_\mathrm{B})_\mathrm{f}-(N_\mathrm{B})_{\mathrm{f}'}]^2\langle v\sigma'(\mathrm{f}\to\mathrm{f}')\rangle\}
\end{aligned} \tag{2.4.8b}$$

respectively, exhibiting the fact that B-violating processes in thermal equilibrium must always act to destroy any initial net baryon number. Performing the remaining phase space integrations, eq. (2.4.5) may finally be written as

$$\frac{\mathrm{d}Y_{\mathrm{B}}}{\mathrm{d}t}=\sum_{\chi}\left\{(Y_{\chi}-Y_{\chi}^{\mathrm{eq}})\sum_{\mathrm{f}}\{(N_{\mathrm{B}})_{\mathrm{f}}\langle\Gamma(\chi\to\mathrm{f})-\Gamma(\chi\to\bar{\mathrm{f}})\rangle\}\right\}$$

$$-Y_{\mathrm{B}}\left\{\sum_{\chi}Y_{\chi}^{\mathrm{eq}}\sum_{\mathrm{f}}\{[(N_{\mathrm{B}})_{\mathrm{f}}]^{2}\langle\Gamma(\chi\to\mathrm{f})+\Gamma(\chi\to\bar{\mathrm{f}})\rangle\}\right\}$$

$$-\tfrac{1}{2}Y_{\mathrm{B}}n_{\gamma}\sum_{\mathrm{f,f'}}\left\{[(N_{\mathrm{B}})_{\mathrm{f}}-(N_{\mathrm{B}})_{\mathrm{f'}}]^{2}\langle v\sigma'(\mathrm{f}\to\mathrm{f'})\rangle\right\}. \qquad (2.4.9)$$

In many supposedly more realistic models, it is necessary to generalize these equations to describe the development of several approximately-conserved quantum numbers (e.g., B, e, μ, . . .). Often some combination of the quantum numbers (e.g., $B-L$) may be absolutely conserved (typically in order to conserve fermion number). Note that if heavy unstable fermions are present, they may be treated in this analysis as χ.

2.4.2. Spin and statistics. In the discussion above, we have assumed that all particles have only one spin state, and obey Maxwell–Boltzmann statistics. Accounting for more spin states changes no formulae: inclusion of appropriate Fermi–Dirac or Bose–Einstein statistics complicates the proof of the theorems discussed in subsect. 2.1 (see appendix A) and yields some small corrections.

For a particle A with g_{A} accessible spin states, we define $f_{\mathrm{A}}(p_{\mathrm{A}})$ to be the phase-space density for each single spin state, but take n_{A} to be the total number density of A summed over spins, and thus write $n_{\mathrm{A}}=g_{\mathrm{A}}\Lambda_{\mathrm{A}}[f_{\mathrm{A}}(p_{\mathrm{A}})]$. The matrix elements $|\mathcal{M}|^2$ are taken to be summed over the possible spin states of initial and final particles, so that the total rates for reactions are given by products of the form $\Lambda[f_{\mathrm{A}}f_{\mathrm{B}}\ldots|\mathcal{M}|^2]$, without further spin factors. To write these rates in terms of the total initial particle number densities n_{i}, rather than f_{i} would require division by the requisite g_{i} factors. However, since we define (as usual) cross sections and widths to be averaged, rather than summed, over initial spin states, the rates written in terms of n_{i} and these cross-sections require no explicit spin multiplicity factors.

When the density $f_{\mathrm{A}}(p_{\mathrm{A}})$ of a species A of particles in phase space becomes close to one, so that cells at least in some region of phase space have a high probability to be occupied, the rates for reactions in which A are produced must be modified to account for quantum statistics effects. If A is a fermion, then these rates contain a factor $(1-f_{\mathrm{A}}(p))$ for each A produced, thereby implementing the exclusion principle that no more than one A (with a given spin direction) may occupy a single cell in phase space. If A instead obeys Bose–Einstein statistics, then each produced A introduces a factor $(1+f_{\mathrm{A}}(p))$ to account for stimulated emission, as discussed in appendix A. These factors appear not only for the final particles produced in a process, but also for each intermediate virtual particle and hence modify the unitarity relation (2.1.3) [to (A.22)]. This will guarantee the cancellation between the two-body processes, and decay-inverse decay as discussed in subsect. 2.3. Taking b, $\bar{\text{b}}$ to be fermions, and X a boson, eq. (2.3.4) becomes

$$\begin{aligned}\frac{\mathrm{d}n_{\mathrm{X}}}{\mathrm{d}t}+3\frac{\dot{R}}{R}n_{\mathrm{X}}=\Lambda_{12}\{&-f_{\mathrm{X}}(p_{\mathrm{X}})[1-f_{\mathrm{b}}(p_1)][1-f_{\mathrm{b}}(p_2)]|\mathcal{M}(\mathrm{X}\to\mathrm{bb})|^2\\&-f_{\mathrm{X}}(p_{\mathrm{X}})[1-f_{\bar{\mathrm{b}}}(p_1)][1-f_{\bar{\mathrm{b}}}(p_2)]|\mathcal{M}(\mathrm{X}\to\bar{\mathrm{b}}\bar{\mathrm{b}})|^2\\&+f_{\mathrm{b}}(p_1)f_{\mathrm{b}}(p_2)[1+f_{\mathrm{X}}(p_{\mathrm{X}})]|\mathcal{M}(\mathrm{bb}\to\mathrm{X})|^2\\&+f_{\bar{\mathrm{b}}}(p_1)f_{\bar{\mathrm{b}}}(p_2)[1+f_{\mathrm{X}}(p_{\mathrm{X}})]|\mathcal{M}(\bar{\mathrm{b}}\bar{\mathrm{b}}\to\mathrm{X})|^2\}\,.\end{aligned}\tag{2.4.10}$$

We again assume that b, $\bar{\text{b}}$ are in kinetic equilibrium. If $f^{\mathrm{eq}}(\mathrm{p})$ is Fermi-Dirac (Bose–Einstein) distributed, then

$$f^{\mathrm{eq}}(p)=[\mathrm{e}^{(E-\mu)/T}\underset{(-)}{+}1]^{-1}=\mathrm{e}^{-(E-\mu)/T}(1\underset{(+)}{-}f^{\mathrm{eq}}(p))\,,\tag{2.4.11}$$

so that the product of $f_{\overset{(-)}{\mathrm{b}}}(p_1)f_{\overset{(-)}{\mathrm{b}}}(p_2)$ in eq. (2.4.10) may be written as

$$\begin{aligned}f_{\overset{(-)}{\mathrm{b}}}(p_1)f_{\overset{(-)}{\mathrm{b}}}(p_2)&=\mathrm{e}^{\overset{+}{(-)}2\mu/T}\,\mathrm{e}^{-(E_1+E_2)/T}[1-f_{\overset{(-)}{\mathrm{b}}}(p_1)][1-f_{\overset{(-)}{\mathrm{b}}}(p_2)]\\&=\mathrm{e}^{\overset{+}{(-)}2\mu/T}\frac{f_{\mathrm{X}}^{\mathrm{eq}}(p_{\mathrm{X}})}{1+f_{\mathrm{X}}^{\mathrm{eq}}(p_{\mathrm{X}})}[1-f_{\overset{(-)}{\mathrm{b}}}(p_1)][1-f_{\overset{(-)}{\mathrm{b}}}(p_2)],\end{aligned}\tag{2.4.12}$$

where the second equality follows from the energy conservation δ function $\Lambda_{12}^{\mathrm{X}}$. We now assume that the fermion chemical potential is small, and take $[1-f_{\mathrm{b}}^{\mathrm{eq}}(p_1)]\approx[1-f_{\bar{\mathrm{b}}}^{\mathrm{eq}}(p_2)]$, so that (2.4.10) becomes

$$\begin{aligned}\frac{\mathrm{d}n_{\mathrm{X}}}{\mathrm{d}t}+3\frac{\dot{R}}{R}n_{\mathrm{X}}=\Lambda_{12}^{\mathrm{X}}\Big\{&-f_{\mathrm{X}}(p_{\mathrm{X}})(|\tilde{\mathcal{M}}(\mathrm{X}\to\mathrm{bb})|^2+|\tilde{\mathcal{M}}(\mathrm{X}\to\bar{\mathrm{b}}\bar{\mathrm{b}})|^2)\\&+f_{\mathrm{X}}^{\mathrm{eq}}(p_{\mathrm{X}})\left(\frac{1+f_{\mathrm{X}}(p_{\mathrm{X}})}{1+f_{\mathrm{X}}^{\mathrm{eq}}(p_{\mathrm{X}})}\right)|\tilde{\mathcal{M}}(\bar{\mathrm{X}}\to\bar{\mathrm{b}}\bar{\mathrm{b}}|^2\,\mathrm{e}^{2\mu/T}+|\tilde{\mathcal{M}}(\bar{\mathrm{X}}\to\mathrm{bb})|^2\,\mathrm{e}^{-2\mu/T})\Big\}\,,\end{aligned}$$

where we have defined

$$|\tilde{\mathcal{M}}(\mathrm{a}\to\mathrm{cc})|^2=|\mathcal{M}(\mathrm{a}\to\mathrm{cc})|^2[1-f_{\mathrm{b}}(p_{\mathrm{c}})][1-f_{\mathrm{b}}(p_{\mathrm{c}})]\,.\tag{2.4.14}$$

Quantum statistics corrections should be small so long as

$$|\tilde{\mathcal{M}}|\approx|\mathcal{M}|\,,\tag{2.4.15a}$$

$$\left(\frac{1+f_{\mathrm{X}}(p_{\mathrm{X}})}{1+f_{\mathrm{X}}^{\mathrm{eq}}(p_{\mathrm{X}})}\right)\approx 1\,.\tag{2.4.15b}$$

The correction to the decay rate of particle A with mass m_{A} decaying at rest to two massless fermions c, in the presence of a gas of c in thermal equilibrium at a temperature T is given simply by

$$\frac{\tilde{\Gamma}}{\Gamma}=\frac{|\tilde{\mathcal{M}}|^2}{|\mathcal{M}|^2}\int\frac{d^3p_1}{2E_1}\int\frac{d^3p_2}{2E_2}$$

$$\times[[1\underset{(+)}{-}f_c^{eq}(p_1)][1\underset{(+)}{-}f_c^{eq}(p_2)]\delta^4(p_1+p_2-p_A)]\Big/\int\frac{d^3p_1}{2E_1}\int\frac{d^3p_2}{2E_2}\delta^4(p_1+p_2-p_A)$$

$$=\frac{1}{[1\underset{(-)}{+}e^{-m_A/2T}]^2}\,. \tag{2.4.16}$$

As $T\to 0$, the density of the c gas goes to zero, and $\tilde{\Gamma}/\Gamma\to 1$. If, as in (2.4.13), the c are fermions, then when $T\to\infty$, $\tilde{\Gamma}/\Gamma\to\frac{1}{4}:\frac{3}{4}$ of the final-state phase space is excluded by Pauli's principle ($\tilde{\Gamma}/\Gamma$ climbs slowly up to 1 as T decreases; for $T=m_A$, $\tilde{\Gamma}/\Gamma\simeq 0.4$). When the c are bosons, the presence of an ambient c gas causes stimulated emission, and increases the decay rate. In fact, as $T\to\infty$, $\tilde{\Gamma}/\Gamma\sim(2T/m_A)^2$, reflecting the approach to Bose condensation in the f_c^{eq}. (As T decreases, $\tilde{\Gamma}/\Gamma$ falls steadily to 1; for $T=m_A$, $\tilde{\Gamma}/\Gamma\simeq 6.5$.) The correction (2.4.16) is for decay at rest: for high-energy A, $\tilde{\Gamma}/\Gamma$ tends to one. In the region $T\leqslant m_X$ which dominates baryon production $\langle\tilde{\Gamma}/\Gamma\rangle$ is close to one. Suppression of X decays by Pauli exclusion merely causes Y_B to be generated at slightly lower temperatures: its final value is unaffected, except in as far as the processes which destroy Y_B are less effective. The correction is reversed if the X decay products are bosons rather than fermions.

If we assume that the actual X distribution is not far from its $\mu=0$ equilibrium form (but perhaps at a different temperature) in most regions of phase space, then the approximation (2.4.15b) is good. However, for a Bose gas, when $E/T\ll 1$, the phase-space density may become large. The approximation (2.4.15b) should nevertheless remain valid for two reasons. First, the equilibrium distribution $f_X^{eq}(p_X)$ will also become large, tending to cancel the growth in $f_X(p_X)$. Also the region or phase space where $f_X(p_X)$ is expected to be large is for p_X/T small and m/T small. Since most baryon production occurs at $T>m$, only in the small p_X region is $f_X(p_X)\gtrsim 1$. However, in this region the distribution function is multiplied by p_X in calculating the number density, which lessens the contribution of the region where $f_X(p_X)\gtrsim 1$.

2.4.3. Large baryon excesses. In 2.3, we always assumed that $Y_B\equiv n_B/n_\gamma\ll 1$, and made the linear approximation

$$f_b(p)=e^{-(E-\mu)/T}=e^{-E/T}\left(1+\frac{\mu}{T}\right)$$

$$=e^{-E/T}(1+\tfrac{1}{2}Y_B)\,. \tag{2.4.17}$$

However, in most cases, the formalism of subsect. 2.3 does not require this approximation. Retaining the full non-linear form

$$Y_B = e^{\mu/T} - e^{-\mu/T} = 2\sinh\left(\frac{\mu}{T}\right),$$

$$f_b(p) = e^{-E/T}\left(\sqrt{1+\tfrac{1}{4}Y_B^2}+\tfrac{1}{2}Y_B\right), \tag{2.4.18}$$

the final equations (2.3.10) and (2.3.20) become

$$\frac{dY_+}{dt} \simeq -\langle\Gamma_X\rangle\{(Y_+ - Y_+^{eq}) + \tfrac{1}{2}Y_B Y_+^{eq}\{(\eta-\bar\eta)\sqrt{1+\tfrac{1}{4}Y_B^2} - Y_B\}\}, \tag{2.4.19a}$$

$$\frac{dY_-}{dt} \simeq -\langle\Gamma_X\rangle\left\{Y_- - \left(\frac{\eta+\bar\eta}{2}\right)Y_B Y_+^{eq}\sqrt{1+\tfrac{1}{4}Y_B^2}\right\}, \tag{2.4.19b}$$

$$\begin{aligned}\frac{dY_B}{dt} \simeq{}& \langle\Gamma_X\rangle\{(\eta-\bar\eta)(Y_+ - Y_+^{eq}) + (\eta+\bar\eta)Y_-\}\\ &-2Y_B\langle\Gamma_X\rangle Y_+^{eq}\{\sqrt{1+\tfrac{1}{4}Y_B^2} - \tfrac{1}{4}Y_B(\eta-\bar\eta)\}\\ &-2Y_B n_\gamma\langle v\{\bar\sigma'(bb\to\bar b\bar b)+\bar\sigma'(\bar b\bar b\to bb)\}\rangle\sqrt{1+\tfrac{1}{4}Y_B^2}\\ &-Y_B^2 n_\gamma\langle v\{\bar\sigma'(bb\to\bar b\bar b)-\bar\sigma'(\bar b\bar b\to bb)\}\rangle,\end{aligned} \tag{2.4.19c}$$

where we have used the fact that

$$e^{\pm 2\mu/T} = 1 + \tfrac{1}{2}Y_B^2 \pm Y_B\sqrt{1+\tfrac{1}{4}Y_B^2}. \tag{2.4.19d}$$

In the limit $|Y_B| \gg 1$, but ignoring Pauli exclusion effects, these reduce to

$$\begin{aligned}\frac{dY_+}{dt} &\simeq \tfrac{1}{2}Y_B^2 Y_+^{eq}\langle\Gamma_X\rangle\left\{1 - \frac{Y_B}{|Y_B|}\left(\frac{\eta-\bar\eta}{2}\right)\right\},\\ \frac{dY_-}{dt} &\simeq -\left(\frac{\eta+\bar\eta}{2}\right)\tfrac{1}{2}Y_B^2 Y_+^{eq}\langle\Gamma_X\rangle,\\ \frac{dY_B}{dt} &\simeq -Y_B^2 Y_+^{eq}\langle\Gamma_X\rangle\left\{1-\frac{Y_B}{|Y_B|}\left(\frac{\eta-\bar\eta}{2}\right)\right\}\\ &\quad - Y_B^2 n_\gamma \begin{cases}\langle v\bar\sigma'(bb\to\bar b\bar b)\rangle, & Y_B>0\\ \langle v\bar\sigma'(\bar b\bar b\to bb)\rangle, & Y_B<0\end{cases}.\end{aligned} \tag{2.4.20}$$

For very large $|Y_B|$ the b or $\bar b$ should form a degenerate Fermi gas; while exclusion effects render the bb or $\bar b\bar b$ elastic scattering cross section much suppressed, they do not affect the baryon number descruction processes of eq. (2.4.20) since the phase space available to the products of these reactions is unrestricted by the presence of the Fermi gas. As shown in eq. (4.8), even if degenerate massless particles dominate the energy density and hence expansion rate of the universe, the expansion term absorbed on the left-hand side of eqs. (2.4.20) remains unchanged. For the usual hot universes considered in sect. 3, Y_B is sufficiently small that the non-linear terms in eq. (2.4.19) are entirely irrelevant. [Note that, as shown in sect. 3, if the initial $Y_B \ll (\eta-\bar\eta)$, then the final Y_B generated is always less than $\eta-\bar\eta$; hence, for example, the last term in eq. (2.4.19) cannot dominate even if $\eta-\bar{\eta}$ is very small.] Notice that chemical potentials associated with quantum numbers which are genuinely conserved in the processes considered exactly cancel out in all equations [e.g., (2.2.7)].

2.4.4. Mixing. As mentioned in subsect. 2.3, when it is not forbidden by absolute conservation laws (as exhibited for example by their ability to decay into the same final state), X and $\bar{\mathrm{X}}$ should mix*; the mixed states X_1 and X_2 diagonalize the hamiltonian and have definite masses and decay widths (typically, the X_1 and X_2 will be split by an amount $m_{\mathrm{X}_1}-m_{\mathrm{X}_2}\sim\Gamma_{\mathrm{X}}$). *CP*-violating effects in X decays may then arise in two ways: either because the eigenstates $\mathrm{X}_{1,2}$ consist of a combination of X, $\bar{\mathrm{X}}$ which is not a *CP* eigenstate, or because the final decays of the $\mathrm{X}_{1,2}$ exhibit *CP* violation (in the manner described in subsect. 2.3). The observed *CP* violation in K^0 system appears to be dominantly of the former type.

We first consider the case in which $\mathrm{X}_{1,2}$ are the *CP* eigenstates:

$$\begin{aligned}|\mathrm{X}_1\rangle&=\sqrt{\tfrac{1}{2}}\{|\mathrm{X}\rangle+|\bar{\mathrm{X}}\rangle\}\,,\\ |\mathrm{X}_2\rangle&=\sqrt{\tfrac{1}{2}}\{|\mathrm{X}\rangle-|\bar{\mathrm{X}}\rangle\}\,,\\ CP|\mathrm{X}_1\rangle&=+|\mathrm{X}_1\rangle\,,\\ CP|\mathrm{X}_2\rangle&=-|\mathrm{X}_2\rangle\,.\end{aligned}\tag{2.4.21}$$

Then *CP* invariance requires

$$\mathcal{M}(\mathrm{X}_k\to i)=\mathcal{M}(\mathrm{X}_k\to\bar{i})\,,\qquad(k=1,2)\,.\tag{2.4.22}$$

The unitarity and *CPT*-invariance constraint (2.1.5) is impotent in this case; the result (2.1.9) still applies, however, so that *CP*-violating differences $(|\mathcal{M}(\mathrm{X}_k\to i)|^2-|\mathcal{M}(\mathrm{X}_k\to\bar{i})|^2)/\sum_i|\mathcal{M}(\mathrm{X}_k\to i)|^2$ must again be $\mathrm{O}(\alpha)$ in perturbation theory. For example, if for some state s,

$$\begin{aligned}\mathcal{M}(\mathrm{X}\to\mathrm{s})&=\mathcal{M}_0\,,&\qquad\mathcal{M}(\bar{\mathrm{X}}\to\bar{\mathrm{s}})&=\mathcal{M}_1\,,\\ \mathcal{M}(\mathrm{X}\to\bar{\mathrm{s}})&=0\,,&\qquad\mathcal{M}(\bar{\mathrm{X}}\to\mathrm{s})&=0\,,\end{aligned}\tag{2.4.23}$$

then (we take the X^0-$\bar{\mathrm{X}}^0$ mixing to occur through an intermediate state with $m>m_{\mathrm{X}}$; in these cases none of the X^0, $\bar{\mathrm{X}}^0$ decay final states need be identical)

$$\begin{aligned}|\mathcal{M}(\mathrm{X}_1\to\mathrm{s})|^2&=|\mathcal{M}(\mathrm{X}_2\to\mathrm{s})|^2=\tfrac{1}{2}|\mathcal{M}_0|^2,\\ |\mathcal{M}(\mathrm{X}_1\to\bar{\mathrm{s}})|^2&=|\mathcal{M}(\mathrm{X}_2\to\bar{\mathrm{s}})|^2=\tfrac{1}{2}|\mathcal{M}_1|^2\,,\end{aligned}\tag{2.4.24}$$

which may differ by $\mathrm{O}(\alpha^2)$. (If s was a *CP* eigenstate, so that $\mathrm{s}\equiv\bar{\mathrm{s}}$, then $\mathcal{M}(\overset{(-)}{\mathrm{X}}\to\mathrm{s})=$

* Since quarks have third-integer electric charges, but leptons integer ones, B-violating bosons in grand unified gauge models typically carry electric charge, and therefore cannot mix with their antiparticles.

$\pm\mathcal{M}(\overset{(-)}{\mathrm{X}}\to\bar{\mathrm{s}})$.) When $\mathrm{X}_{1,2}$ are *CP* eigenstates, their number densities individually satisfy equations analogous to those derived in subsect. 2.3 (on setting $\eta=-\bar{\eta}=$ O(α) and identifying X and $\bar{\mathrm{X}}$ there). Notice that with the choice (2.4.23) of matrix elements, the state X may yield only s and $\bar{\mathrm{X}}$ only $\bar{\mathrm{s}}$. Because of the X-$\bar{\mathrm{X}}$ mixing, it is the $\mathrm{X}_{1,2}$ rather than $\bar{\mathrm{X}}$ states which exhibit the characteristic $\exp[(iE-\Gamma)t]$ time dependence; the number of X or $\bar{\mathrm{X}}$ oscillates at the beat frequency $\sim(E_{\mathrm{X}_1}-E_{\mathrm{X}_2})$ $(=m_{\mathrm{X}_1}-m_{\mathrm{X}_2}$, when the X are at rest). [The case of the K^0 system is slightly more complicated than that treated here. In practical experiments, the $\overset{(-)}{\mathrm{K}}{}^0$ are produced by strong interaction $2\to2$ processes for which the matrix elements are arranged somewhat analogously to (2.4.23), in that strangeness +1 initial states give only K^0 and $S=-1$ only $\bar{\mathrm{K}}^0$ (assuming the recoil final particle has $S=0$). The weak interactions responsible for $\overset{(-)}{\mathrm{K}}{}^0$ mixing and decay are not involved in their production.]

The second type of *CP* violation arises when X_1 and X_2 are no longer *CP* eigenstates as in (2.4.21), but are rather of the form (in the K^0 system, these combinations are conventionally denoted K_S and K_L (e.g., [15]))

$$
\begin{aligned}
|\mathrm{X}_1\rangle &= \frac{1}{\sqrt{1+|\varepsilon|^2}}\{|\underline{\mathrm{X}}_1\rangle+\varepsilon|\underline{\mathrm{X}}_2\rangle\} \\
&= \frac{1}{\sqrt{2(1+|\varepsilon|^2)}}\{(1+\varepsilon)|\mathrm{X}\rangle+(1-\varepsilon)|\bar{\mathrm{X}}\rangle\}\,, \\
|\mathrm{X}_2\rangle &= \frac{1}{\sqrt{1+|\varepsilon|^2}}\{\varepsilon|\underline{\mathrm{X}}_1\rangle+|\underline{\mathrm{X}}_2\rangle\} \\
&= \frac{1}{\sqrt{2(1+|\varepsilon|^2)}}\{(1+\varepsilon)|\mathrm{X}\rangle-(1-\varepsilon)\bar{\mathrm{X}}\rangle\}
\end{aligned}
\tag{2.4.25}
$$

where ε measures the *CP* impurity of the states. $\mathrm{X}-\bar{\mathrm{X}}$ mixing occurs when the matrix elements $\mathcal{M}(\mathrm{X}\to\bar{\mathrm{X}})$, $\mathcal{M}(\bar{\mathrm{X}}\to\mathrm{X})$ are non-zero (if they involve as intermediate states shared X, $\bar{\mathrm{X}}$ decay modes, they are typically $\sim\Gamma_\mathrm{X}$). In the eigenvectors (2.4.25) of the X-$\bar{\mathrm{X}}$ propagation matrix, the parameter ε measures the *CP*-violating difference of the ratio $\mathcal{M}(\mathrm{X}\to\bar{\mathrm{X}})/\mathcal{M}(\bar{\mathrm{X}}\to\mathrm{X})$ from one. Whereas unitarity places severe constraints (2.1.9) on *CP* violation in $|\mathcal{M}(\mathrm{X}\to i)|^2\neq|\mathcal{M}(\bar{\mathrm{X}}\to\bar{i})|^2$, these constraints do not apply to the unsquared amplitudes: to obtain $\varepsilon\neq0$, it is sufficient for interactions generating $\mathrm{X}\to\bar{\mathrm{X}}$ to have complex coupling constants (which appear conjugated in $\bar{\mathrm{X}}\to\mathrm{X}$); no further restrictions apply. (In the K^0 system, $\varepsilon\neq0$ may well arise from relatively complex couplings of t, c and u quarks in box diagrams with virtual WW intermediate states (e.g., [12]). We shall assume here that the decays of the states $\mathrm{X}_{1,2}$ are *CP* conserving (as appears to be the case for $\mathrm{K}^0_{\mathrm{L,S}}$); all *CP*-violating effects arise from the X-$\bar{\mathrm{X}}$ mixing which causes the eigenstates $\mathrm{X}_{1,2}$ of the hamiltonian not to be *CP* eigenstates. We take the X, $\bar{\mathrm{X}}$ decay amplitudes

$$\mathcal{M}(\mathrm{X}\to i)=\mathcal{M}(\bar{\mathrm{X}}\to \bar{i})=(1+\lambda)\tfrac{1}{2}\mathcal{M}_0\,,$$
$$\mathcal{M}(\mathrm{X}\to \bar{i})=\mathcal{M}(\bar{\mathrm{X}}\to i)=(1-\lambda)\tfrac{1}{2}\mathcal{M}_0\,, \tag{2.4.26}$$

so that the decay rates for the states (2.4.25) become (for simplicity taking λ real)

$$|\mathcal{M}(\mathrm{X}_1\to i)|^2=\frac{|\mathcal{M}_0|^2}{2(1+|\varepsilon|^2)}\{1+2\lambda\ \mathrm{Re}\ \varepsilon+\lambda^2|\varepsilon|^2\}\,,$$
$$|\mathcal{M}(\mathrm{X}_1\to \bar{i})|^2=\frac{|\mathcal{M}_0|^2}{2(1+|\varepsilon|^2)}\{1-2\lambda\ \mathrm{Re}\ \varepsilon+\lambda^2|\varepsilon|^2\}\,,$$
$$|\mathcal{M}(\mathrm{X}_2\to i)|^2=\frac{|\mathcal{M}_0|^2}{2(1+|\varepsilon|^2)}\{|\varepsilon|^2+2\lambda\ \mathrm{Re}\ \varepsilon+\lambda^2\}\,,$$
$$|\mathcal{M}(\mathrm{X}_2\to \bar{i})|^2=\frac{|\mathcal{M}_0|^2}{2(1+|\varepsilon|^2)}\{|\varepsilon|^2-2\lambda\ \mathrm{Re}\ \varepsilon+\lambda^2\}\,. \tag{2.4.27}$$

(If i is a CP eigenstate with $CP|i\rangle=+|i\rangle$, then in (2.4.26), $\lambda=0$: this is the case for the $\pi\pi$ final state in K^0 decays. In semileptonic ($\pi\ell\nu$) K^0 decays, the $\Delta S=\Delta Q$ rule implies $\lambda=1$.) (The value of $|\varepsilon|^2$ in the simple model (2.4.27) is determined from the total X_1 and X_2 decay rates by a quadratic equation. In general, $|\varepsilon|^2$ satisfies the unitarity constraint $|\varepsilon|^2\leq\sum_i\Gamma(\mathrm{X}_1\to i)\Gamma(\mathrm{X}_2\to i)/[(m_1-m_2)^2+\frac{1}{4}(\Gamma_1+\Gamma_2)^2]$ [15].)

The CP violation in the rates (2.4.27) can generate an excess of i over $\bar{i}$ in a system which is initially symmetrical in X and $\bar{\mathrm{X}}$. For example, the free decay (without back reactions) of X and $\bar{\mathrm{X}}$ produced in equal numbers, and with wave functions of random phase, as when they are emitted in thermal radiation, generates an excess of i over $\bar{i}$ given by

$$\frac{n_i-n_{\bar{i}}}{n_i+n_{\bar{i}}}\simeq\frac{8\lambda\ \mathrm{Re}\ \varepsilon}{(1+|\varepsilon|^2)(1+\lambda^2)}\,. \tag{2.4.28}$$

2.4.5. Multiple temperatures and annihilation heating. Eq. (2.1.4) shows that if all particles are in thermal equilibrium with zero chemical potential, then no baryon number can be generated. Once a massless particle species has been brought into kinetic equilibrium, equation (2.1.13) shows that the expansion of the universe alone cannot destroy its equilibrium distribution in phase space. Thus other influences are necessary to modify the distributions of massless particles so as to allow reactions between them to generate baryon number. One possible such influence might be the presence or growth of inhomogeneities in the universe. Another possible mechanism would be differential heating of various massless particle species by the annihilation products of other, massive, particles decoupling from thermal equilibrium. As mentioned in subsect. 2.1, the expansion of a homogeneous universe (containing weakly interacting particles) should approximately conserve the entropy

$$S=R^3\left(\frac{\rho+p-\mu n}{T}\right)\simeq\left(\frac{4}{\pi^2}\right)\xi T^3R^3\,, \tag{2.4.29}$$

where ξ is the effective number of particle species at temperature $T(\xi=\frac{1}{2}$ for each boson spin state with $m \ll T$ and $\xi=\frac{7}{16}$ for non-degenerate ultrarelativistic fermion spin states), as described in appendix C. p is the pressure*: for an ultrarelativistic ideal gas in equilibrium $p=\frac{1}{3}\rho$ while for dust $p=0$. (For any ideal equilibrium gas, $p=(\gamma-1)\rho$, where $\gamma=c_p/c_v$. Systems whose components interact strongly may have pressures up to $p=\rho$. Such pressure should probably occur if the universe undergoes a phase transition.) As T falls below the mass of a particular species, the contribution of that species to the energy density of the universe (and hence to ξ) drops rapidly to zero (unless the particles carry a non-vanishing chemical potential). The energy density originally carried by the disappearing species is transferred to its lighter annihilation products; their interactions with the rest of the universe raise the tempertures of other particle species in such a way as to conserve (2.4.29), so that $\xi T^3=$ constant. However, the rate at which the energy is shared among the species depends on their cross sections for interaction with the annihilation products. If the cross section for a particular species is too small, it may never receive its full share of the energy, and remain at a lower temperature. (This behavior is exhibited by light neutrinos in the present universe; below $T \sim 1$ MeV, the rate for $e^+e^- \to \nu\bar{\nu}$ reactions becomes very small, and when the $e^\pm$ annihilate at $T \sim 0.1$ MeV, all their energy goes into photons. Consequently, the temperature of photons in the present universe should be about $[(\xi_\gamma+\xi_{e^+}+\xi_{e^-})/\xi_\gamma]^{1/3}=(\frac{11}{4})^{1/3}\simeq 1.4$ times higher than that of the neutrinos. Similarly, heavy quark species should dominantly annihilate into gluons, which heat the lighter quarks, but not leptons, in the universe.) In eq. (2.2.12), for example, the rate of baryon number generation is proportional to the difference of the actual φ number density from its equilibrium value with $\mu=0$ and at the temperature of the b, $\bar{\text{b}}$. Even if $m_\varphi=0$, so that the φ phase-space distribution remains of the form (2.2.13), its temperature may differ from that of the $\overset{(-)}{\text{b}}$ (because of weaker interactions with annihilation products), and baryon number generation may occur until the φ and $\overset{(-)}{\text{b}}$ are brought to the same temperature. Thus baryon-number conserving annihilation of massive species can indirectly result in baryon number generation by B-violating reactions between massless particles.

In models such as those of subsects 2.2 and 2.3, the baryon number of the universe remains constant after the temperature has fallen below the masses of the particles mediating B-violating interactions. When lighter species of particles annihilate, they increase the temperature, and thus number, of photons, but leave the baryon number

* In general,

$$\{n,\rho,p\}=\int_0^\infty \frac{d^3p}{(2\pi)^3}\left\{1, E, \frac{|\boldsymbol{p}|^2}{3E}\right\} f_{eq}(p)\,,$$

while

$$v_{sound}^2=\partial p/\partial\rho\,.$$

unchanged. These effects reduce the original Y_B produced by a factor

$$\frac{Y_B(T \simeq 0)}{Y_B(T \simeq m_X)} \simeq \frac{\xi(T \simeq 0)}{\xi(T \simeq m_X)}\,; \tag{2.4.30}$$

In typical grand unified models, this factor is $O(\frac{1}{100})$*.

The approximate conservation of entropy for the universe may of course be drastically violated if its contents undergo a first order phase transition (with specific latent heat Δ). Such a transition would reduce Y_B by a factor $\sim \xi T^3/(\xi T^3 + \Delta/T)$. As mentioned above, most phase transitions, regardless of order, would result in a temporary increase in the pressure of the universe. The phase transitions associated with the spontaneous breaking of gauge symmetries (mentioned below) are expected to be second-order, or first-order with very small latent heats. The phase transition at $T \sim \Lambda$ in asymptotically free theories (eiter QCD or a higher Λ "technicolor" group) to confinement and chiral symmetry breaking is also probably second order.

2.4.6. Phase transitions and spontaneous symmetry restoration. The only known method for breaking local gauge symmetries and providing masses for gauge bosons without destroying renormalizability is by the introduction of Higgs fields ϕ with non-zero vacuum expectation values $\langle\phi\rangle$. With this mechanism, the mass of a gauge boson V is given by

$$m_V^2 = \tfrac{1}{2}g^2\langle\phi\rangle^2\,, \tag{2.4.31}$$

where g is the gauge coupling constant, and $\langle\phi\rangle$ is determined by minimizing the effective potential

$$V(\phi) = -\tfrac{1}{2}\mu^2\phi^2 + \tfrac{1}{4}\lambda\phi^2 + O(g^4\phi^4 \log(\phi^2/m^2))\,, \tag{2.4.32}$$

which gives the energy density of the "vacuum" as a function of the strength of a uniform classical Higgs field ϕ. The masses of the quantized fluctuations in this condensate (Higgs particles H) are given by

$$m_H^2 = \left.\frac{\delta^2 V}{\delta\phi^2}\right|_{\phi=\langle\phi\rangle} \simeq 2\mu^2\,, \tag{2.4.33a}$$

$$\langle\phi\rangle \simeq \mu/\sqrt{\lambda}\,. \tag{2.4.33b}$$

The vacuum energy density implied by minimizing $V(\phi)$ then has the absurdly large value

$$\rho = V(\langle\phi\rangle) \simeq -\frac{m_H^2 m_V^2}{2g^2} \simeq -2\times 10^{21}(m_H[\mathrm{GeV}])^2\ \mathrm{g\,cm^{-3}}\,. \tag{2.4.34}$$

If this energy density is present, it presumably has no gravitational effects; to accord with observation it must otherwise be delicately cancelled by a cosmological term in the Einstein field equations. (When the symmetry is restored, the Higgs condensate energy density disappears, leaving uncancelled the cosmological term; however, at

* If a massive species survives for a long time between decoupling from thermal equilibrium and decaying, the large deviations from equilibrium can occur, and large amounts of entropy may be generated [31].

the relevant temperatures, its effects on the expansion rate of the universe are typically overwhelmed by the radiation energy density present.) The energy density (2.4.34) perhaps discredits any cosmological considerations of the Higgs mechanism.

At zero temperature, $\langle\phi\rangle$ will always be given by minimizing the "vacuum" energy (2.4.32). However, at high temperatures, thermal fluctuations in ϕ become much larger than $\langle\phi\rangle$, and the "vacuum" state is no longer concentrated at the minimum of $V(\phi)$. The mean Higgs condensate strength at high temperatures is typically given by*

$$\begin{aligned} \langle\phi(T)\rangle^2 &\simeq \langle\phi_0^2\rangle\left(\frac{T_c^2 - T^2}{T_c^2}\right), \qquad && T < T_c\,, \\ &= 0\,, && T > T_c\,, \\ T_c &\simeq m_H\,. && \end{aligned} \tag{2.4.35}$$

(In the region close to the phase transition $[(T^2 - T_c^2)/T_c^2 \lesssim g^2]$, the perturbative methods used to derive (2.4.35) fail, so that the precise nature of the transition cannot be determined.) For $T > T_c$, therefore, the gauge bosons and Higgs particles will be effectively massless; when the universe cools below $T = T_c$, their masses grow slowly to the $T = 0$ values. Note that the constraint $V(\langle\phi\rangle) < V(0)$ necessary for spontaneous symmetry breakdown to occur at low temperatures implies

$$T_c \simeq m_H \gtrsim g^2 m_V\,. \tag{2.4.36}$$

Above this critical temperature, all particles (X) should be effectively massless, and therefore exist in their equilibrium number densities, so that

$$Y_X \simeq 1\,, \qquad (T \gtrsim T_c)\,. \tag{2.4.37}$$

When the universe cools below T_c, the particles become massive and may decay. The smaller the ratio T_c/m_X is, the smaller will be the back reactions to these decays, and thus the larger the final baryon number generated. However, the bound (2.3.36) suggests that the values of m_X for which back reactions will not destroy the baryon number produced are not in fact much extended by these considerations of symmetry restoration.

In addition to the masses of the decaying X particles, spontaneous symmetry breakdown may also determine the strength of *CP* violation (e.g., [11]). In this case, *CP*-violating effects should disappear above a critical temperature T_c^{CP}. If $T_c^{CP}/(\alpha m_X) \ll 1$, then most X decays will be *CP* conserving, and so no baryon asymmetry may be generated.

An intriguing (but probably irrelevant) possibility is that domains with different "order parameters", (typically $\langle\phi\rangle$) signalling different symmetry breaking, may

* Much more complicated behavior, perhaps with $\langle\phi(\infty)\rangle^2 > 0$, can be obtained by introducing several coupled Higgs fields.

have formed in the early universe just below the critical temperature. If, as in the phase transition leading to $m_X \neq 0$, different values of $\langle\phi\rangle$ imply different vacuum energies, then the "true vacuum" in which $V(\langle\phi\rangle)$ is at its global minimum should quickly overwhelm the regions of false vacuum. However, if there are many possible $\langle\phi\rangle$, as characterized, for example, by a phase angle, which give the same vacuum energy, then domains may survive. Thus it is possible that the sign and perhaps magnitude of *CP* violation may initially have differed from one region (domain) in the universe to another. However, it is probable that insufficient surface tension would exist to prevent the domains from mixing freely. Moreover, the maximum size of a domain is presumably governed by the distance over which a light signal could have propogated by the time of the phase transition: larger regions could not yet be in "causal contact" and therefore could not act collectively. At the temperatures $T \sim 10^{15}$ GeV probably relevant for B-violating processes, the maximum number of particles in a domain is $\sim 10^5$. The possibility that domains in which baryons were generated should have collected together and repelled antibaryons seems extremely implausible. (Note that the "Leidenfrost effect" by which radiation pressure from $N\bar{N}$ annihilation may hold matter and antimatter regions apart is entirely impotent at $T \gg m_q, m_N$, since it relies on the conversion of N rest energy into photon momentum in annihilation.)

3. Results in a simple model

3.1. INTRODUCTION

In this subsection, we present solutions to eqs. (2.3.10) and (2.3.20) which describe baryon number generation in the simple model of subsect. 2.3. In terms of the dimensionless variables

$$x \equiv \frac{m_X}{T}, \qquad x_P = \frac{m_X}{m_P},$$

$$\frac{dY}{dx} = \frac{x}{m_X x_P}\frac{dY}{dt}, \tag{3.1.1}$$

where the effective Planck mass m_P was defined in (2.1.14), these equations become

$$\frac{dY_+}{dx} \simeq -A(x)\left\{(Y_+ - Y_+^{eq}) + \left(\frac{\eta - \bar{\eta}}{2}\right) Y_B Y_+^{eq}\right\}, \tag{3.1.2a}$$

$$\frac{dY_-}{dx} \simeq -A(x)\left\{Y_- - \left(\frac{\eta + \bar{\eta}}{2}\right) Y_B Y_+^{eq}\right\}, \tag{3.1.2b}$$

$$\frac{dY_B}{dx} \simeq A(x)\left\{\{(\eta - \bar{\eta})(Y_+ - Y_+^{eq}) + (\eta + \bar{\eta}) Y_-\}\right. \tag{3.1.2c}$$

$$-2Y_{\mathrm{B}}\left\{Y_{+}^{\mathrm{eq}}+\frac{n_{\gamma}}{\langle\Gamma_{\mathrm{X}}\rangle}\langle v\{\sigma'(\mathrm{bb}\to\bar{\mathrm{b}}\bar{\mathrm{b}})+\sigma'(\bar{\mathrm{b}}\bar{\mathrm{b}}\to\mathrm{bb})\}\rangle\right\}\Bigg\},$$

$$A(x)=\frac{x}{x_{\mathrm{P}}}\frac{\langle\Gamma_{\mathrm{X}}\rangle}{m_{\mathrm{X}}}, \tag{3.1.2d}$$

$$Y_{\pm}=\tfrac{1}{2}(Y_{\mathrm{X}}\pm Y_{\bar{\mathrm{X}}}). \tag{3.1.2e}$$

For simplicity we shall henceforth take

$$\frac{\eta+\bar{\eta}}{2}=1$$

although none of our results are sensitve to this choice. The *CP* violation parameter $\eta-\bar{\eta}$ will be denoted by

$$\varepsilon\equiv(\eta-\bar{\eta}).$$

Unitarity of the decay rates (2.3.1) requires $|\varepsilon|\leq 2$, and according to eq. (2.1.9), ε is formally of order α. We shall write Y_{B}^{0} for the final baryon number density (at zero temperature); we do not include in Y_{B}^{0} the final factor ξ discussed in subsect. 2.4 to account for increase in the photon number density.

If the contents of the universe were in thermal equilibrium at sufficiently early times, the solutions to eqs. (3.1.2) must satisfy the initial conditions [we assume as in sect. 2 that all particles (including γ) have only one spin state and obey Maxwell–Boltzmann statistics: if X has g_{X} spin states (and γ are assigned Bose–Einstein statistics) then $Y_{+}^{\infty}=\frac{1}{2}g_{\mathrm{X}}$ if it is a boson, and $Y_{+}^{\infty}=\frac{3}{8}g_{\mathrm{X}}$ if a fermion]:

$$\begin{aligned}&Y_{+}(x=0)=Y_{+}^{\infty}=Y_{+}^{\mathrm{eq}}(x=0)=1,\\&Y_{-}(x=0)=Y_{-}^{\infty}=0\\&Y_{\mathrm{B}}(x=0)=Y_{\mathrm{B}}^{\infty},\end{aligned} \tag{3.1.3}$$

where Y_{B}^{∞} is a possible initial baryon number. At lower temperatures, the equilibrium X number density is given by (cf., eq. (2.2.9) and appendix C)

$$Y_{+}^{\mathrm{eq}}(x)=Y_{\mathrm{X}}^{\mathrm{eq}}(x)=\tfrac{1}{2}x^{2}K_{2}(x)=\frac{m_{\mathrm{X}}^{3}}{2\pi^{2}x}K_{2}(x), \tag{3.1.4}$$

where K_2 is a modified Bessel function (see appendix C). We take the time-dilation factor in the effective X width to be averaged over an equilibrium X energy distribution, so that

$$\langle\Gamma_{\mathrm{X}}\rangle=\left\langle\frac{m_{\mathrm{X}}}{E_{\mathrm{X}}}\right\rangle\Gamma_{\mathrm{X}}=\frac{K_{1}(x)}{K_{2}(x)}\Gamma_{\mathrm{X}}. \tag{3.1.5}$$

For the numerical solutions of subsect. 3.3, we use

$$\Gamma_{\mathrm{X}}=\tfrac{1}{4}m_{\mathrm{X}}\alpha \qquad (g_{\mathrm{X}}=1)$$

$$=\frac{m_X\alpha}{2g_X} \qquad (g_X>1)\,, \tag{3.1.6}$$

corresponding to the decay of an X with g_X spin states to two identical (spin-$\frac{1}{2}$) fermions, coupled with strength $e=\sqrt{4\pi\alpha}$. (The factor of 2 between $g_X=1$ and $g_X>1$ in (3.1.7) arises because only half the possible final fermion spin states are accessible from a spin-0 X.) We usually take $g_X=1$. The cross sections $\sigma'(\mathrm{bb}\to\bar{\mathrm{b}}\bar{\mathrm{b}})$ and $\sigma'(\bar{\mathrm{b}}\bar{\mathrm{b}}\to\mathrm{bb})$ (in which the contribution of real intermediate X already included in previous terms has been subtracted off, by removing the pole part of the exchanged X propagator) are equal to $O(\alpha^2)$. The high-energy behavior of the σ' crucial for the destruction of any initial baryon number will be discussed in sect. 4. For baryon number generation, the form of σ' at c.m.s. energies $\sqrt{s}\lesssim m_X$ is important. In the low-energy limit, it is of the usual Fermi form

$$v\sigma'(\mathrm{bb}\to\bar{\mathrm{b}}\bar{\mathrm{b}})\simeq\frac{c\alpha^2 s}{m_X^2}\,, \qquad (s\ll m_X^2)\,, \tag{3.1.7}$$

where typically $c\simeq\pi g_X$ (see below for specific cases). Averaging this cross section over thermal energy distributions for the incoming b gives (see appendix C)

$$\langle v\sigma'(\mathrm{bb}\to\bar{\mathrm{b}}\bar{\mathrm{b}})\rangle\simeq\frac{18c\alpha^2T^2}{m_X^4}\,, \qquad (T\ll m_X)\,. \tag{3.1.8}$$

The detailed from of σ' as a function of s depends on the couplings and spin of the exchanged X. For scalar X exchange, one finds [using the same couplings as in (3.1.6)]*

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t}=\frac{\pi\alpha^2}{s^2}\left[\frac{t^2}{t'^2}+\frac{s^2}{s'^2}+2\frac{st}{s't'}\right],$$

$$\sigma=\int_{-s}^{0}\frac{\mathrm{d}\sigma}{\mathrm{d}t}\mathrm{d}t\,,$$

$$v'\equiv v-m_X^2\,, \qquad (v=s,t)\,. \tag{3.1.9}$$

In the high-energy limit, this cross section becomes

$$\sigma\simeq\frac{4\pi\alpha^2}{3s}\left[1+\frac{m_X^2}{s}\left(\tfrac{5}{4}-\log\frac{s}{m_X^2}\right)+\cdots\right], \qquad (s\gg m_X^2)\,, \tag{3.1.10}$$

while at low energies it reduces to

$$\sigma\simeq\frac{\pi\alpha^2 s}{3m_X^4}\left[1+\tfrac{7}{2}\frac{s}{m_X^2}-\cdots\right], \qquad (s\ll m_X^2)\,. \tag{3.1.11}$$

* In this equation, we have approximated the intermediate X propagator by its zero-temperature form. As discussed in sect. 4, the effective X mass at high temperatures is probably given by an inverse Debye screening length.

Note that in (3.1.9) the contribution of t-channel, as well as s-channel, X-exchange has been included. To obtain σ', one must sutbract from σ the cross section obtained by keeping only the pole part of the X propagator. For a vector X (with coupling $e\gamma_\mu$), the total cross section becomes

$$\frac{d\sigma}{dt}=\frac{2\pi\alpha^2}{s^2}\left[\frac{s^2+u^2}{t'^2}+\frac{t^2+u^2}{s'^2}+\frac{2u^2}{s't'}\right],$$

$$\sigma=\int_{-s}^{0}\frac{d\sigma}{dt}\,dt,$$

$$v'=v-m_X^2,\qquad (v=s,t,u),$$

$$s+t+u=0. \qquad (3.1.12)$$

In the high-energy limit, this yields

$$\sigma\simeq\frac{4\pi\alpha^2}{m_X^2}\left[1+\frac{m_X^2}{s}\left(\tfrac{7}{3}-6\log\frac{s}{m_X^2}\right)+\cdots\right],\qquad (s\gg m_X^2), \qquad (3.1.13)$$

and in the low-energy limit

$$\sigma\simeq\frac{16\pi\alpha^2 s}{3m_X^4}\left[1+\frac{s}{4m_X^2}-\cdots\right],\qquad (s\ll m_X^2). \qquad (3.1.14)$$

For the numerical calculations of subsect. 3.3, these cross sections are averaged over the relevant initial energy distributions; for the purposes of analytical approximation, one may estimate the complete $\langle v\sigma'\rangle$ by replacing s in $v\sigma'$ by $\langle s\rangle=18T^2$.

3.2. APPROXIMATE ANALYTICAL SOLUTIONS

If Y_B and Y_- always remain small, eqs. (3.1.2) reduce simply to

$$\frac{dY_B}{dx}=-\varepsilon\frac{dY_+}{dx},\qquad (Y_B, Y_-\ll 1), \qquad (3.2.1)$$

corresponding to baryon-number generation by free X decays, with no back reactions. In this approximation, the baryon number generated Y_B^0 is trivially given by

$$Y_B^0\simeq Y_B^\infty+\varepsilon\{Y_+^\infty-Y_+^0\}$$

$$=Y_B^\infty+\varepsilon,\qquad (Y_B, Y_B^\infty, Y_-\ll 1), \qquad (3.2.2)$$

where, as usual, Y_B^∞ is a possible initial baryon number. Numerical solutions in subsect. 3.3 suggest that this approximation is typically accurate for $Y_B^\infty\leq\varepsilon$ and $\alpha\leq 10^{-3}$ or $x_P=m_X/m_P\gtrsim 10^{-4}$. Note that (3.2.2) provides an upper bound on $|Y_B^0|$; back reactions always tend to diminish the baryon density.

At high temperatures, and taking for simplicity $Y_B^\infty=0$, eqs. (3.1.2) become (the necessary small x expansions of Y_+^{eq}, etc., are given in appendix C):

$$\frac{\mathrm{d}Y_+}{\mathrm{d}x} \simeq -ax^2\{Y_+ - 1 + \tfrac{1}{4}x^2 + \mathrm{O}(x^4 \log x) + \mathrm{O}(x^2(Y_+ - 1)) + \cdots\}, \quad (3.2.3a)$$

$$\frac{\mathrm{d}Y_-}{\mathrm{d}x} \simeq -ax^2\{Y_- - Y_\mathrm{B} + \cdots\}, \quad (3.2.3b)$$

$$\frac{\mathrm{d}Y_\mathrm{B}}{\mathrm{d}x} \simeq -a\varepsilon x^2\{Y_+ - 1 + \tfrac{1}{4}x^2 + \cdots\}, \qquad (x \ll 1), \quad (3.2.3c)$$

where

$$a = \frac{\Gamma_\mathrm{X}}{x_\mathrm{P} m_\mathrm{X}} = \frac{m_\mathrm{P}\Gamma_\mathrm{X}}{m_\mathrm{X}^2}. \quad (3.2.3d)$$

For small x, the solutions to these equations are

$$Y_+ \simeq 1 - \tfrac{1}{20}ax^5 + \mathrm{O}(x^7 \log x), \quad (3.2.4a)$$

$$Y_- \simeq \tfrac{1}{160}\varepsilon a^2 x^8 - \mathrm{O}(x^{11}), \quad (3.2.4b)$$

$$Y_\mathrm{B} \simeq \varepsilon\{1 - Y_+\} \simeq \tfrac{1}{20}\varepsilon a x^5 + \mathrm{O}(x^7 \log x), \qquad (x \ll 1). \quad (3.2.4c)$$

In subsect. 3.3, we shall find that these forms are often adequate until $x \simeq 1$ (the largest discrepancies are usually in Y_-).

At low temperatures (large x) the X undergo exponential decay, and their number is typically negligible for $x \gg 1$. Only the last term in eq. (3.1.2c) for Y_B is thus important at large x. Using the low-energy point form (3.1.7) for the $2 \to 2$ cross sections, and taking $s = \langle s \rangle = 18T^2 = 18(m_\mathrm{X}/x)^2$, eq. (3.1.2c) becomes

$$\frac{\mathrm{d}Y_\mathrm{B}}{\mathrm{d}x} \simeq -\frac{72c\alpha^2}{\pi^2 x_\mathrm{P}}\frac{Y_\mathrm{B}}{x^4} \equiv -\lambda\frac{\alpha^2 m_\mathrm{P}}{m_\mathrm{X}}\frac{Y_\mathrm{B}}{x^4}, \qquad (x \gg 1). \quad (3.2.5)$$

Any baryon excess generated by decay and inverse decay at high temperatures is therefore depeleted at low temperatures through baryon-number violating $2 \to 2$ reactions, falling roughly like

$$Y_\mathrm{B}(x) \sim \exp\left[\frac{\lambda\alpha^2}{3x_\mathrm{P}x^3}\right], \qquad (x \gg 1), \quad (3.2.6)$$

and eventually tending to a constant non-zero value. (For fixed temperature, the exponent here $\sim\alpha^2/m_\mathrm{X}^4$, which arises simply from the Fermi low-energy form for the $\mathrm{bb} \to \bar{\mathrm{b}}\bar{\mathrm{b}}$ cross section.) Numerical results in subsect. 3.3 suggest that in practice, when $\lambda\alpha^2/3x_\mathrm{P} \gtrsim 1$, this behavior typically sets in when x rises above about 2.

If the temperature T of the universe falls with time t according to

$$T \simeq \frac{m_\mathrm{P}}{(tm_\mathrm{P})^\gamma}, \quad (3.2.7)$$

(where $\gamma=\frac{1}{2}$ for a radiation-dominated universe at small t, $\gamma=\frac{3}{2}$ for a matter-dominated universe with deceleration parameter $q_0=\frac{1}{2}$ and $\gamma=1$ for a closed universe ($q_0<\frac{1}{2}$) [9]), then the relaxation (3.2.6) of baryon density with time due to low-energy $2\rightarrow 2$ interactions is roughly

$$Y_B(t)\sim\exp\left[\frac{\lambda\alpha^2}{x_P^4(m_P t)^{5\gamma-1}(5\gamma-1)}\right]. \tag{3.2.8}$$

Hence, if $\gamma>\frac{1}{5}$ [as in eq. (3.2.6), for which $\gamma=\frac{1}{2}$], $Y_B(t)$ cannot relax to zero even when $t\rightarrow\infty$: the age of the universe then grows faster than the time necessary to establish chemical equilibrium; the fluctuation in baryon number has been frozen by the expansion of the universe, and survives forever, albeit perhaps somewhat diminished from its high temperature value. As discussed in appendix A, this failure to destroy baryon number even after an infinite time is a consequence of the extra expansion terms in the Boltzmann equation, which invalidate Boltzmann's H theorem $dH/dt\leq 0$*. On the other hand, in a universe with $\gamma<\frac{1}{5}$, $2\rightarrow 2$ processes occur with a sufficient rate to combat expansion, and any baryon number generated at high temperatures eventually relaxes exponentially to zero. To attain $\gamma<\frac{1}{5}$ would require the introduction of a cosmological term into the Einstein field equation, which can serve even to halt expansion (as in the Lemaître universe) and allow chemical equilibrium to be established. (These results are not specific to the model of subsect. 2.3 considered. In practice, however, gravitational or other clumping will drastically change the rate for B-violating interactions at large t: for example, two quarks confined within a proton have a much higher amplitude to come sufficiently close together to annihilate than would two free quarks in an ideal homogeneous gas with the same density as the proton gas**. Note that even in the presumably physical case $\gamma>\frac{1}{5}$, baryon number generated at high temperatures would be diminished to an unacceptably low level if $\lambda\alpha^2 m_P/m_X$ were too large. The final baryon number usually depends, however, on the behavior of eqs. (3.1.2) in the region $x\sim 1$, where simple analytical approximations fail; a numerical solution to (3.1.2) is therefore necessary.

3.3. NUMERICAL RESULTS

In this subsection we give numerical solutions to eqs. (3.1.2) as a function of the three dimensionless parameters ε, x_P and α. Except in considerations of the destruction of an initial non-zero Y_B^∞ at very high temperatures, for which eqs. (3.1.2) are no longer accurate (see sect. 4) the precise form for the widths and cross sections assumed is largely irrelevant; only the very model independent low-energy form (3.1.7) for the $2\rightarrow 2$ cross sections is important (these cross sections are essentially just those which should induce proton decay).

* When applied solely to the matter in the universe, but not to the gravitational field generated by it.

** For two quarks within a proton, $|\Psi(0)|^2\sim 1/r_P^3\sim(0.7\text{ GeV})^3$, while if the quarks were free in an ideal gas of number density n, $|\Psi(0)|^2\sim n$.

Baryon number violating interactions such as those in the simple model of subsect. 2.3 treated here should lead to proton decay, with a lifetime given by the very low-energy limit of (3.1.7) as roughly $\tau_P \sim m_X^4/(\alpha^2 m_N^5)$. The experimental $\tau_P \gtrsim 10^{30}$ yr then implies $m_X \gtrsim 10^{14}$ GeV; in the SU(5) grand unified model, estimates suggest that $m_X \simeq 10^{15}$ GeV $= 1\ \Pi$ eV*. We use $m_X = 1\ \Pi$ eV as a standard value for our numerical results. The relevant coupling constant α depends on the precise nature of the X in our model. If X is a gauge (vector) boson, then α should presumably be the corresponding effective gauge coupling constant at an invariant mass $\sim\sqrt{s}$ for $2 \to 2$ scatterings and $\sim m_X$ for X decays. A typical value obtained for this coupling constant in the SU(5) model is $\alpha \sim \frac{1}{40}$. On the other hand, if X is a scalar (presumably Higgs) boson, as is probably obligatory in generating baryon number from an SU(5) model, the relevant coupling constant is largely unknown, but it is probably rather small ($\lesssim 10^{-3}$). The value of the CP-violation parameter ε is even more uncertain. Nevertheless, all our numerical results for Y_B (and Y_-) in fact depend linearly on ε to within a few percent, even when $\varepsilon \simeq 1$. As a standard, we take the quite unmotivated value $\varepsilon = 10^{-6}$**. Finally, we must specify the effective Planck mass defined by eq. (2.1.14), which depends on the number of species ξ contributing to the energy density of the universe at the temperatures $\sim m_X$ considered. If no new species of particles (except t, $W^{\pm}$, Z^0) beyond those already detected exist with masses $\lesssim m_X$, then $\xi \simeq 100$. With this, and the choice $m_X = 1\ \Pi$ eV, the dimensionless parameter $x_P \simeq 2 \times 10^{-3}$. m_P determines the rate of expansion in the early universe; inhomogeneities or perturbations in the metric could lead to different expansion rates for different regions of the universe. Such effects may be parametrized by different values for x_P.

Fig. 1 shows the development of the $\overset{(-)}{X}$ and baryon densities as a function of the inverse temperature $x = m_X/T$, with $\alpha = \frac{1}{40}$, $m_X = 1\Pi$ eV, $\xi = 100$, and $\varepsilon = 10^{-6}$. The dashed lines in fig. 1 are the analytical approximations for small x discussed in subsect. 3.2. Note that the changes in the actual Y_+ lag behind those in Y_+^{eq}.

In fig. 2 we show the relative sizes of the terms contributing to dY_B/dt with the parameters used in fig. 1. As expected, for $x > 1$ (in this case $x \gtrsim 10$) all terms proportional to Y_X decrease exponentially so that the only remaining contribution is from the two-body scattering.

Fig. 3 illustrates the sensitivity of Y_B to the parameters of the model. Unless otherwise indicated, the parameters are the same as for fig. 1. Fig. 3a shows that the final Y_B is independent of an initial Y_B^∞, so long as Y_B^∞ is small. As discussed in subsect. 2.4, the destruction of a very large Y_B^∞ cannot be treated using eq. (2.3.20). Fig. 3a also exhibits the linear proportionality of Y_B^0 on ε, Figs. 3b–d illustrate the dependence of Y_B on m_X, α, and ξ respectively.

* Note that even within this specific model, only leading log contributions are included in deducing m_X [16]; the subleading log terms have not yet been calculated using consistent prescriptions.

** Note, of course, that by construction $\varepsilon \leq 1$, so that models giving Y_B too small even when $\varepsilon = 1$ should be considered in disagreement with the standard cosmology.

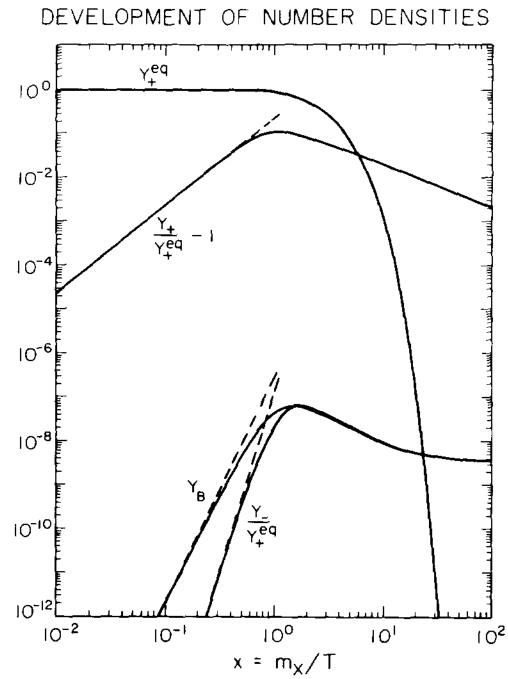

Fig. 1. Numerical solutions (solid lines) and analytical approximations (dashed lines) for the number densities in the model of subsect. 2.3. The differential equations for Y_+, Y_-, and Y_B are given in (3.12). The standard choices of parameters used in this and later figures are $m_X = 1\Pi$ eV, $\alpha = \frac{1}{40}$, $\xi = 100$, $\varepsilon = 10^{-6}$, $\frac{1}{2}(\eta - \bar{\eta}) = 1$, and $Y_B^\infty = 0$, ($1\Pi\,\mathrm{eV} \equiv 10^{15}$ GeV).

Finally in fig. 4, we give the final value of Y_B/ε as a function of x_P (i.e., of m_X) for various values of α.

4. The destruction of initial baryon number

In this section, we discuss the behavior of the baryon density in the universe at temperatures $T \gg m_X$, where the X mass is irrelevant, and any B-violating interactions should occur as fast as B-conserving ones. At times $t \lesssim 1/m_{\mathscr{P}}$ (corresponding to temperatures $T \gtrsim m_{\mathscr{P}}$), it is undoubtedly not permissible to consider only background gravitational effects; quantized fluctuations in the expansion rate, and direct gravitational contributions to particle interactions (which determine the equation of state) presumably become overwhelming. Nevertheless, the perhaps overly naive estimates given below indicate that such effects should rapidly become unimportant when T falls below $\simeq m_{\mathscr{P}}$.

As mentioned in sect. 1, the likely excess of baryons over antibaryons in the present universe probably cannot be explained solely by postulating a small initial

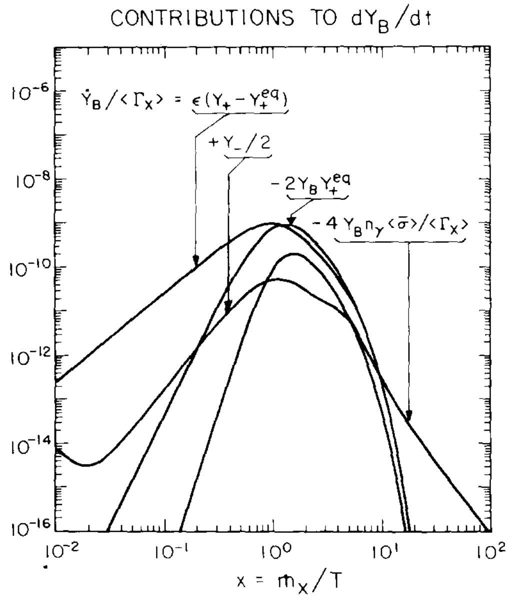

Fig. 2. The relative magnitudes of the terms in eq. (3.1.2) contributing to the time development of the baryon density $\dot{Y}_B$. Notice that for $x = m_X/T \gg 1$, all terms proportional to Y_+ or Y_- become exponentially unimportant, and the largest contribution to $\dot{Y}_B$ is from $2 \to 2$ scattering processes.

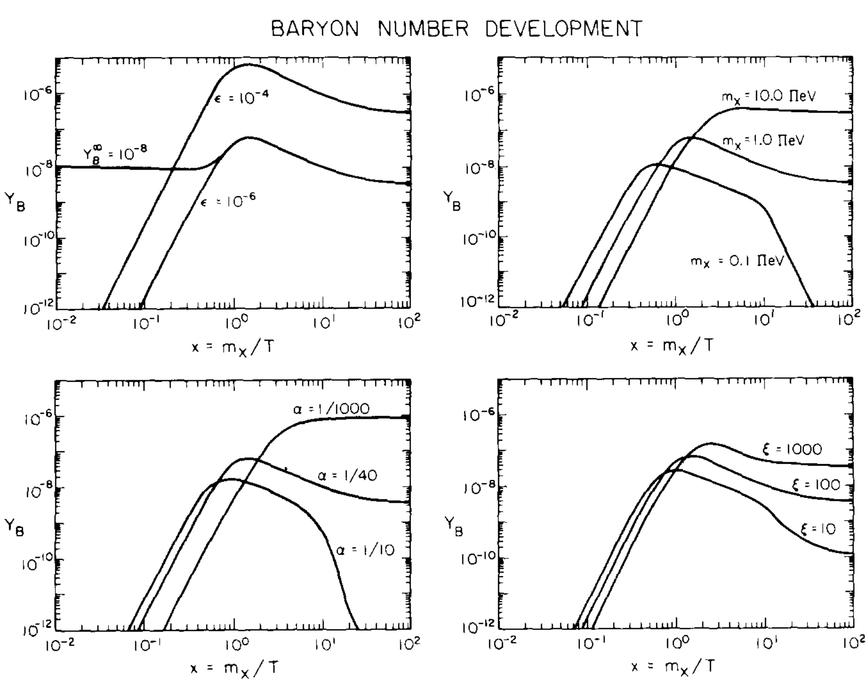

Fig. 3. The sensitivity of the baryon number development to the input parameters. Unless otherwise indicated, the parameters are the same as those used for fig. 1.

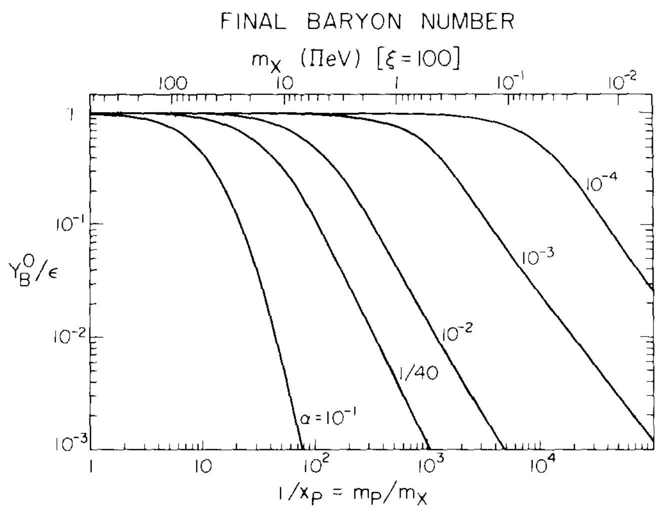

Fig. 4. The final baryon to photon ratio (divided by the CP-violation parameter ε) as a function of the ratio x_P of the effective Planck mass to m_X, for several values of the coupling constant α. The upper scale shows the values of m_X with the choice $\xi = 100$ in the definition of the effective Planck mass (2.1.14).

net baryon number if B-violating processes occur rapidly at high temperatures, since these would tend to eradicate the initial baryon number. The purpose of this section is to estimate the rate of relaxation to a $B = 0$ state, and to determine whether it would be complete before when the temperature fell below $\sim m_X$, and B-violating processes become rare. We consider the fate of a baryon excess existing at $T \simeq m_{\mathcal{P}}$ (the possible genesis of such an initial condition is, of course, quite unknown, despite Polkinghorne's program [17]). If particles carrying baryon number also carry electric charges (but the sum of the charges of all baryon species is non-zero), then to maintain the overall charge neutrality of the universe any baryon excess must be compensated by a suitable antilepton excess. Then, if for example, $B \pm L$ is absolutely conserved, destruction of an initial baryon excess would be accompanied by destruction of the corresponding (anti) lepton excess. We assume as above, purely for simplicity, that only one light neutral fermion species carrying baryon number exists, and that there is only one species X of B-violating boson, with $m_X \ll m_{\mathcal{P}}$.

According to eq. (2.3.20), at $T \gg m_X$ (so that, by assumption, $Y_X = Y_X^{eq}$) a small baryon excess $Y_B = (n_b - n_{\bar{b}})/n_\gamma \lesssim 1$ in the model of subsect. 2.3 should be relaxed according to

$$\frac{dY_B}{dt} \simeq -2\,Y_B\{\langle \Gamma_X\rangle + n_\gamma \langle v\sigma(bb \to \bar{b}\bar{b})\rangle\}\,. \tag{4.1}$$

The first term in eq. (4.1) represents the absorption of baryons by inverse X decays (e.g., $bb \to X$); the X produced eventually decay approximately symmetrically to baryons and antibaryons, and hence baryon asymmetry is diminished. The second term describes the direct destruction of baryon number by $2 \to 2$ scattering processes.

(As discussed in sect. 3, these act at temperatures $T \leqslant m_X$ to diminish baryon number generated in X decays). Ignoring temporarily the first term in (4.1), an initial baryon concentration Y_B^∞ should relax roughly according to

$$Y_B(t) \sim Y_B^\infty \exp\left[-m_P \int_{m_{\mathcal{P}}}^{T(t)} \langle v\sigma \rangle \,\mathrm{d}T\right], \tag{4.2}$$

assuming as above that the universe expands according to

$$T \sim \sqrt{\frac{m_P}{t}}. \tag{4.3}$$

At high c.m. energies $\sqrt{s}$, the total cross section for two-body scattering through t-channel exchange of a spin-j particle behaves like

$$\sigma \sim (s/m^2)^j/s. \tag{4.4}$$

As evidenced by eq. (3.1.13), $\sigma(\mathrm{bb} \to \bar{\mathrm{b}}\bar{\mathrm{b}})$ due to exchange of a vector X goes to a constant value $\sim\alpha^2/m_X^2$ at high s. (On the other hand, a scalar X exchange gives a cross section asymptotically falling like α^2/s.) Note that, as is typical in non-abelian gauge theories, the change in quantum numbers $(\mathrm{b} \to \bar{\mathrm{b}})$ may be effected without any exchange of transverse momentum; the range of the B-violating interaction is limited only by the mass of the exchanged X. If $\sigma \sim \alpha^2/m_X^2$, then eq. (4.2) implies that any initial baryon excess would be diminished, dominantly near the Planck time, by a factor $\sim\exp[-\alpha^2 m_P^2/m_X^2]$, which would probably be sufficient to destroy any initial baryon excess $Y_B^\infty \leqslant 1$. (If $\sigma \sim \alpha^2/s$, then the destruction factor would only be $\sim\exp[-\alpha^2 m_P/m_X]$.) A cross section $\sigma \sim \alpha^2/m_X^2$ results essentially from interactions with an X cloud (of opacity α) surrounding a baryon, with area $\sim 1/m_X^2$. At high temperatures, there should be $\sim n/m_X^3 \sim \xi T^3/(\pi^2 m_X^3)$ particles within the range of the X interaction from a single baryon. These particles should contribute to scattering from the baryon, giving rise to processes with many-body initial states. As usual, the simple Boltzmann equation is unable to account for such effects of a "long-range" X interaction*. However, as in electron-ion plasmas [18], it is presumably permissible, at least in the near-equilibrium state considered, to account for higher-body processes simply by introduction of an effective screened two-body cross section. The antibaryons surrounding a baryon typically screen its "X charge" at distances beyond the Debye length (note that all particle species carrying any X charge contribute with the same sign to the screening length)

$$\lambda_D \simeq \left(\frac{T}{32\pi^2 \alpha n}\right)^{1/2} \simeq \frac{1}{(\sqrt{32\alpha\xi})T}; \tag{4.5}$$

* In fact, in the presence of "long-range forces" (acting over times longer than the collision time), the Boltzmann equation ceases to be applicable. The equation assumes that the momenta of particles are uncorrelated before each scattering, but ignores processes involving more than two initial particles (e.g., two sequential two-body interactions). When long-range forces exist, the effects neglected in this way become inportant, and one must formally resort to more complicated equations [19]. So long as kinetic equilibrium prevails, however, consideration of an effective screened cross section should be adequate.

this screening may be described by assigning the X an effective mass $\sim 1/\lambda_D$. The mean time $\sim\lambda_D$ between successive collisions of the X becomes smaller than its Compton time $\sim 1/m_X$ at temperatures $T \gg m_X$. The effective baryon destruction cross section at high temperatures should therefore be* $\sim\alpha^2\lambda_D^2 \sim \alpha/(\xi T^2)$, due to each species of B-violating exchange. If, as in many complicated and "realistic" models, the number of particle species mediating B-violating processes at very high temperatures is much larger than those conserving baryon number, then the total B destruction effective cross section should be $\sim\xi\alpha/(\xi T^2) = \alpha/T^2$. With this form, an initial baryon excess would be diminished by a factor $\sim\exp[-\alpha m_P/m_X]$, which is probably $\lesssim 10^{-6}$, so that the present Y_B could not be explained in an initially hot universe. (Recall that $Y_B(T=0) \simeq Y_B(T \simeq m_X)/\xi(T \simeq m_X)$). The use of an effective screened cross section will tend, if anything, to underestimate the rate of B destruction. Notice that the modification $\sim 1/\lambda_D \sim \sqrt{\alpha}T$ to the effective X mass is largely irrelevant at the temperatures $\lesssim m_X$ considered in previous sections.

The X width which governs the first term in eq. (4.1) is given at temperatures not too far above m_X by $\langle\Gamma_X\rangle \simeq (m_X/T)\Gamma_X \sim m_X\alpha\sqrt{t/m_P}$. This form assumes that the produced X is on its mass-shell, and therefore that the incoming total c.m.s. energy must lie within the X resonance curve; this restricts the angle of the incoming particles to be $\lesssim m_X/E_X$, and thus introduces the m_X/E_X "time-dilation factor" in $\langle\Gamma_X\rangle$. However, at temperatures $T \gg \Gamma_X$, the mean free path of X for scattering will typically be much shorter than the mean decay length. Hence the X resonance should be collision broadened, and the m_X/E_X factor resulting from the impossibility of producing X with invariant masses $\gg m_X$, should disappear. Then, according to eq. (4.1) inverse decay processes should diminish an initial baryon excess by a factor $\gtrsim\exp[-\alpha m_P/m_X]$. (If many B-violating bosons exist, dominating ξ at high temperatures, then the factor becomes $\gtrsim\exp[-\xi\alpha m_P/m_X]$.)

Most grand unified models based on simple gauge groups [e.g., SU(5)] are asymptotically free, so that the effective gauge coupling constant falls logarithmically with increasing energy or temperature, and presumably always remains small. Nevertheless, the effective coupling constant for Higgs interactions often increases with temperature**, and could diverge (reach its Landau singularity $\Lambda^2 \sim \mu^2\exp(\beta_0/g^2(\mu^2))$*** beyond which perturbation theory is useless at temperatures

* In electron-ion plasmas, the effective Coulomb cross section involves a log, rather than a power of λ_D; in that case, the $f(p)$ change only if scatterings deflect particles into a different momentum state, so that for the relevant effective total cross section, the differential cross section is weighted by the change in momentum. Here, the $f(p)$ may change due to changes in quantum numbers alone, with no change in momentum.

** Bounds on Higgs couplings [20] based on the structure of the effective potential (vacuum energy density as a function of Higgs field strength) responsible for spontaneous symmetry breakdown need only apply at low temperatures $T \ll m_X$; at higher temperatures, large thermal fluctuations in the fields restores the broken symmetry (see subsect. 2.4) and render the effective potential irrelevant.

***Recall that asymptotically free coupling constants behave as $g^2(s) \sim 1/\log(s/\Lambda^2)$, and diverge in the small s infrared region, while asymptotically strong couplings behave according to $g^2(s) \sim 1/\log(\Lambda^2/s)$. For QED, $\Lambda^2 \sim m_e^2\exp(1/\alpha)$.

below the Planck mass. (If this occurred at $T \lesssim m_X$, then the calculations of baryon number generation given above must be modified considerably; in the standard SU(5) model this case does not occur, except in the presence of massive fermions forbidden by other considerations [20]. Nevertheless, divergence below the Planck mass could easily be achieved.) The presence of such strong interactions at very high temperatures should lead to phenomena analogous to those encountered in the Hagedorn–Frautschi model for hadronic matter (now believed to be inappropriate because of the quark composite nature of the hadron states considered). When energy is added to a strongly interacting system, it may not simply increase the kinetic energies (and hence temperature) of its constituent particles, but rather serve only to generate more massive particles with small momenta (these particles might be bound by strong Higgs interactions★). Hence there may exist a maximum temperature for the universe, governed by the point at which Higgs couplings become strong. The behavior of the universe at earlier times may then be shielded: only the decay products of the massive Higgs bound states initially present will be visible when the universe has cooled below the maximum temperature; their nature should presumably be determined solely from the dynamics of the strong Higgs interactions.

Next, we consider the autolysis of a cold universe, initially consisting of a zero temperature Fermi gas of (ξ types of) baryons with large chemical potential $\mu\,(\mu/T \gg 1)$. The expansion rate of such a universe is given by

$$\frac{\dot{R}}{R} = -\frac{\dot{T}}{T} = \xi\frac{\mu^2}{m_{\mathscr{P}}} = -\frac{\dot{\mu}}{\mu}, \tag{4.6}$$

where $\xi = (2n_f/3\pi^2)^{1/2}$, and n_f is the number of degenerate fermion species. According to eq. (2.4.19), the baryon density in a universe with $|Y_B| \gg 1$ should be destroyed at a rate (Pauli exclusion effects are negligible, because the phase space for the annihilation products is unrestricted by the presence of the degenerate baryon sea)

$$\begin{aligned}\dot{Y}_B &= -\frac{n_b^2}{n_\gamma}\langle v\sigma\rangle \\ &= -Y_B n_b \langle v\sigma\rangle\,.\end{aligned} \tag{4.7}$$

The number density is [see (C.36)]

$$n_{FD}\left(T, m=0, \frac{\mu}{T}\right) = \tfrac{1}{2}\xi^2\mu^3\,, \tag{4.8}$$

where according to (4.6)

$$\mu^2 = \frac{2m_{\mathscr{P}}}{\xi t}\,. \tag{4.9}$$

★ Which represent (in ladder approximation) only spin-0 exchanges, yielding universally attractive forces.

Just as in the high temperature case discussed above, B-violating exchange at high densities should also be screened. (Although the initial baryon number of a cold universe is taken to be large, the initial SU(3) color charges, etc., were presumably zero, so that the total charge to which the X couples was zero, allowing screening.) Hence the effective B destruction cross section in a cold universe should be* $\sim\alpha/\mu^2$, so that eq. (4.7) becomes

$$Y_{\mathrm{B}}(t)=\frac{2\xi^2}{\alpha}\left(\frac{\xi}{2m_{\mathscr{P}}}\right)^{1/2}t^{-1/2}\,. \tag{4.10}$$

Thus in the presence of B-violating interactions, an initially cold universe with $Y_{\mathrm{B}}^{\infty}\gg 1$ should relax to a hot universe with $Y_{\mathrm{B}}\lesssim m_{\mathrm{X}}/(\alpha m_{\mathrm{P}})\ll 1$.

Finally, we discuss the structure of the universe very close to the Planck time, and comment on the consistency of our assumption of initial equilibrium. The cross section for production or annihilation of scalar particle pairs into photon pairs at high energies is given by $\sigma\sim 8\pi\alpha^2/s$. The corresponding cross section for production or annihilation into spin-2 graviton pairs is given at lowest order by [21] $\sigma\sim 19\pi s/6m_{\mathscr{P}}^4$. When $T\gtrsim\sqrt{\alpha}m_{\mathscr{P}}$, therefore, graviton-induced interactions should be important, but at lower temperatures, they should rapidly become irrelevant. We have assumed above that the contents of the early universe behave as an ideal ultrarelavistic gas. In fact, Coulomb interactions alone should affect the properties of the gas, giving, for example, [22]

$$\rho\simeq\frac{3T^4}{\pi^2}\left[1-\alpha\frac{\sqrt{\alpha\xi}}{3\sqrt{\pi}}+\cdots\right], \tag{4.11}$$

an evidently irrelevant correction. Even assuming homogeneity, gravitational interactions would provide a correction perhaps $\mathrm{O}((T/m_{\mathscr{P}})^3)$; once again, the effect becomes overwhelming at $T\simeq m_{\mathscr{P}}$, but quickly becomes negligible below it. Of course, even excepting quantum corrections, the treatment of gravitational effects in the early universe is made difficult by the genuinely long range nature of gravity: since all masses and gravity is universally attractive, no screening occurs, and the Boltzmann equation becomes entirely inadequate.

In keeping with the simplest big bang cosmology, we have assumed that the universe is initially in a state of kinetic equlibrium (so that all particles follow equilibrium distributions in phase space, albeit perhaps with non-zero chemical potential). Nevertheless, of course, the gravitational field must be far from equilibrium, otherwise no expansion would occur. One might therefore be led to modify the usual initial assumptions, and postulate instead that not all particle species were in equilibrium at the Planck time [23]: they would come to kinetic equilibrium only after a few collision times. (Nevertheless, despite the fact that a finite time should be

* At high temperatures, the Boltzmann distribution of the screening particles is reflected in the exponential form $V_{\mathrm{X}}(r)\sim\exp(r/\lambda_{\mathrm{D}})/r$ for the screened potential. In a degenerate Fermi gas, the potential again reflects the particle distribution, and $V_{\mathrm{X}}(r)\sim\theta(\lambda_{\mathrm{D}}-r)/r$.

required for a signal to propagate from one part of the universe to another, it appears that the expansion of the different parts began at "times" closer than would have allowed a light signal to be exchanged between them. This perhaps surprising phenomenon may indicate that equilibration was more rapid in the very early universe.) In the presence of long-range interactions, one may consider only an effective collision time $\simeq 1/(n\sigma_{\mathrm{eff}}) \simeq (\alpha T)^{-1} \log(1/\sqrt{\alpha})$. (Assuming not too many conservation laws, the α here should probably be multiplied in practice by $\sim\xi^2$.) Typically, therefore, kinetic equilibrium should be achieved quite rapidly. Two effects could be thought to modify this result. First, the large effective masses of particles at high temperatures could affect their equilibrium distributions. In fact, this modification is already included in consideration of suitable screened cross sections. Second, at a time t, only particles within a Jeans volume $\simeq(v_{\mathrm{sound}}t)^3$ could apparently be in causal contact. However, this does not necessarily provide an infrared cutoff on particle momenta (density fluctuations with larger wavelengths can exist).

We are grateful to many people for discussions, including A.D. Dolgov, E. Dwek, S. Frautschi, William A. Fowler, G.C. Fox, T.J. Goldman, S.E. Koonin, D.L. Tubbs, and R.V. Wagoner.

Appendix A

BOLTZMANN'S H THEOREM

A.1. MAXWELL-BOLTZMANN STATISTICS

In this appendix, we discuss the H theorem, which shows that any closed system obeying Boltzmann's equation will evolve with time, in the absence of external influences, to an equilibrium state in which Boltzmann's H function is minimal (entropy is maximal). If interactions in the system violate baryon number, then the final equilibrium state can contain no excess of baryons over antibaryons. In an expanding universe, however, the H theorem is modified, and the baryon asymmetries discussed above may be produced and remain while expansion persists.

To investigate the approach to equilibrium, we consider the development of the quantity

$$H = \sum_{\alpha} f_\alpha(\log f_\alpha - 1)\,, \tag{A.1}$$

where the index α labels both the momentum state and the particle type (in the notation used above $H = \sum_i \int \mathrm{d}^3 p_i f_i(p_i)\,(\log f_i(p) - 1)$, where i runs only over particle types). [Note that (A.1) differs slightly from usual definitions of H in which the $-f_\alpha$ term is absent. When no particles are created or destroyed, $\sum_\alpha f_\alpha$ gives the total number of particles and is constant and irrelevant. The definition (A.1) is more convenient when particle creation is included.] H changes with time according to

$$\frac{dH}{dt} = \sum_{\alpha} \frac{df_\alpha}{dt} \log f_\alpha . \tag{A.2}$$

H is stationary when $df_\alpha/dt = 0$, which may occur either if no collisions take place, or by detailed balancing in equilibrium. In the presence of interactions, the f_α are taken to evolve with time according to the Boltzmann equation (the sum over states α_i, β_i accounts for usual phase-space integration)

$$\frac{df_{\alpha_1}}{dt} = \sum_{\substack{\alpha_2,\ldots,\alpha_m \\ \beta_1,\ldots,\beta_n}} \{f_{\beta_1} \cdots f_{\beta_n} |\mathcal{M}(\beta_1 \cdots \beta_n \to \alpha_1 \cdots \alpha_m)|^2 - f_{\alpha_1} \cdots f_{\alpha_m} |\mathcal{M}(\alpha_1 \cdots \alpha_m \to \beta_1 \cdots \beta_n)|^2\} . \tag{A.3}$$

(This may represent a sum of terms with different numbers m and n of initial or final particles in collisions.) Then Boltzmann's H theorem states that any system (in which the momenta of particles are initially uncorrelated) will evolve towards equilibrium so that

$$\frac{dH}{dt} = \sum_{\alpha} \frac{df_\alpha}{dt} \log f_\alpha \leq 0 , \tag{A.4}$$

which is a microscopic statement of the second law of thermodynamics. Adding the forms obtained by permuting the dummy indices α_i, (A.4) may be written (dropping irrelevant constant factors) as

$$\frac{dH}{dt} = \sum_{\alpha,\beta} \{F_\beta |\mathcal{M}(\beta \to \alpha)|^2 - F_\alpha |\mathcal{M}(\alpha \to \beta)|^2\} \log (F_\alpha) , \tag{A.5a}$$

where

$$F_\alpha \equiv f_{\alpha_1} \cdots f_{\alpha_n} , \tag{A.5b}$$

and the sums on α, β run over all α_i and β_i.

We first discuss the proof of (A.4) when all interactions respect time-reversal (CP) invariance, so that all pairs of T-conjugated matrix elements in (A.5) are equal. Then

$$\begin{aligned} \frac{dH}{dt} &= \sum_{\alpha,\beta} |\mathcal{M}(\alpha \to \beta)|^2 \{F_\beta - F_\alpha\} \log (F_\alpha) \\ &= \sum_{\alpha,\beta} |\mathcal{M}(\alpha \to \beta)|^2 \{F_\beta - F_\alpha\} \log (F_\alpha / F_\beta) , \end{aligned} \tag{A.6}$$

where the second form is obtained by adding forms in which the α_i and β_i are permuted in all possible ways. The terms in the sum (A.6) now each have the form

$$(b-a) \log (a/b) \leq 0 \tag{A.7}$$

(where equality holds only when $a = b$), and are thus separately not positive, so that

$dH/dt \leq 0$. To achieve the equilibrium state $dH/dt = 0$, each term must vanish, so that

$$F_\alpha = F_\beta = f_{\alpha_1} \cdots f_{\alpha_m} = f_{\beta_1} \cdots f_{\beta_n} \quad \text{(equilibrium)} \tag{A.8}$$

for all sets of α_i and β_i corresponding to initial and final states of possible collisions as represented in (A.3)*. If particles carry no absolutely conserved internal quantum numbers, then the f_i in (A.8) may depend only on the energies E_i and must follow the Maxwell-Boltzmann distribution

$$f_\alpha = e^{-E_\alpha/T}, \tag{A.9}$$

where, as usual, $-1/T$ may be considered as a Lagrange multiplier enforcing energy conservation. The phase-space distributions for particles carrying absolutely conserved quantum numbers may differ from (A.9) by factors $\exp(\lambda_i\mu/T)$ where λ_i gives the value of the quantum number carried by species i, while μ is a chemical potential which parametrizes the total concentration of the quantum number. (For non-abelian quantum numbers, the relevant factors are the corresponding group elements.) Nevertheless, in the absence of absolutely conserved quantum numbers, the equilibrium distributions must follow (A.9), and in particular if the particles and antiparticles of a species are not absolutely conserved, their phase-space distributions must be identical in equilibrium; no particle–antiparticle asymmetries may exist.

This result also holds when the fundamental interactions exhibit CP violation; it is based solely on unitarity [25]. The unitarity constraint (2.1.3) implies

$$\sum_{\alpha,\beta} F_\alpha\{|\mathcal{M}(\alpha\to\beta)|^2 - |\mathcal{M}(\beta\to\alpha)|^2\} = \sum_{\alpha,\beta} |\mathcal{M}(\beta\to\alpha)|^2\{F_\beta - F_\alpha\} = 0\,. \tag{A.10}$$

Permuting dummy indices in the second term of (A.5a), but without assuming equality of the T-conjugated matrix elements yields

$$\frac{dH}{dt} = \sum_{\alpha,\beta} |\mathcal{M}(\beta\to\alpha)|^2\{F_\beta \log F_\alpha - F_\beta \log F_\beta\}\,. \tag{A.11}$$

Inserting the result (A.10) then gives

$$\frac{dH}{dt} = \sum_{\alpha,\beta} |\mathcal{M}(\beta\to\alpha)|^2\{F_\beta \log F_\alpha - F_\beta \log F_\beta + F_\beta - F_\alpha\} \leq 0\,, \tag{A.12}$$

* One might expect that the equilibrium condition (A.8) could be deduced from $df_\alpha/dt = 0$ for a single state α. However, this implies only

$$\sum_\beta (F_\alpha - F_\beta)|\mathcal{M}(\alpha\to\beta)|^2 = 0\,;$$

individual terms could be non-zero but cancel in the complete sum. Consideration of $\sum_\alpha d(f_\alpha)^2/dt = 0$
yields $\sum_{\alpha,\beta}(F_\alpha - F_\beta)^2|\mathcal{M}(\alpha\to\beta)|^2 = 0$, thus providing an alternative derivation of (A.8).

since

$$b \log (a/b) + (b-a) = \int_b^a \log (x/a)\, \mathrm{d}x \leq 0\,. \tag{A.13}$$

Once again, distributions must tend to the equilibrium form (A.8). That this conclusion is independent of T-violation in the collisions was to be expected, since it also holds for systems in static external (e.g., magnetic) fields, and when internal degrees of freedom are excited in molecules by collisions. Of course, as usual, the validity of (A.4) relies on the assumption of molecular chaos, according to which the momenta of particles are uncorrelated before each collision (but clearly not after). This is presumably true only at the separated maxima of H, where evolution in t or $-t$ would decrease H. Strictly, the time dependence of a single-particle phase-space distribution should depend on the joint two-particle distribution (as in the BBGKY hierarchy); this may not always factorize as required for molecular chaos. Nevertheless, eq. (A.3) and hence the H theorem (A.4) is presumably adequate in an average sense, except when collision rates become very large, or long-range forces are present, as discussed in sect. 4.

Note that the result $\mathrm{d}H/\mathrm{d}t \leq 0$ relies on the form (A.3) of the Boltzmann equation. In an expanding universe, the extra term of eq. (2.1.15) must be added to account for the expansion. This term typically gives a positive contribution to $\mathrm{d}H/\mathrm{d}t$, which may overwhelm the negative contributions from collisions and give $\mathrm{d}H/\mathrm{d}t > 0$, so that baryon asymmetries can be generated.

A.2. QUANTUM STATISTICS

In eqs. (A.1) and (A.3) we have assumed that all particles are classically distinguishable, and therefore obey Maxwell-Boltzmann statistics. As mentioned in subsect. 2.4, when the particles are identical, (A.3) becomes (known as the Uehling-Uhlenbeck equation)

$$\begin{aligned}\frac{\mathrm{d}f_{\alpha_1}}{\mathrm{d}t} = \sum_{\substack{\alpha_2,\ldots,\alpha_m \\ \beta_1,\ldots,\beta_n}} \{ & f_{\beta_1} \cdots f_{\beta_n} (1+\theta_{\alpha_1} f_{\alpha_1}) \cdots (1+\theta_{\alpha_m} f_{\alpha_m}) |\mathcal{M}(\beta_1 \cdots \beta_n \to \alpha_1 \cdots \alpha_m)|^2 \\ & - f_{\alpha_1} \cdots f_{\alpha_m} (1+\theta_{\beta_1} f_{\beta_1}) \cdots (1+\theta_{\beta_n} f_{\beta_n}) |\mathcal{M}(\alpha_1 \cdots \alpha_m \to \beta_1 \cdots \beta_n)|^2 \}\,,\end{aligned} \tag{A.14}$$

where $\theta_{\alpha_i} = +1$ if the particle α_i is a boson and $= -1$ if it is a fermion. We write the factor $(1+\theta_{\alpha_1} f_{\alpha_1}) \cdots (1+\theta_{\alpha_n} f_{\alpha_n}) = s_{\alpha_1} \cdots s_{\alpha_n} = S_\alpha$ and denote $\Theta_\alpha = \theta_{\alpha_1} \cdots \theta_{\alpha_n}$. For final fermions, the extra $(1-f)$ factor implements the Pauli exclusion principle which forbids any fermion from being emitted into a cell in phase space which is already occupied (in a finite quantization volume, this is achieved by a $1-N_\alpha$ factor for each finite cell, which becomes $1-f$ in the continuum limit). For final bosons, the two terms in each $(1+f)$ factor represent spontaneous and stimulated emission respectively. The presence of this correction may formally be considered to result from the

$\sqrt{N+1}$ factor when a creation operator acts on an N boson state ($a^{\dagger}|N\rangle = \sqrt{N+1}|N+1\rangle$), which follows from the commutation relations of the boson field operators. A slightly more direct derivation is based on the fact that the total amplitude for any process is the sum of the amplitudes associated with each possible permutation of identical bosons. (For fermions, the exchange of each pair introduces a minus sign.) Consider a state $|N_{\rm i}\rangle$ consisting of N indistinguishable non-interacting bosons. For combinational purposes, assign each boson to one of N separate "substates" of $|N_{\rm i}\rangle$. The number of possible assignments is $N!$ First, let the bosons propagate undisturbed to a final state $|N_{\rm f}\rangle$, which is again divided into N "substates", labelled in $N!$ ways. The amplitude for the propagation of each boson is exp (iEt). The total amplitude for all N bosons to propagate them $|N_{\rm i}\rangle$ to $|N_{\rm f}\rangle$ is the sum of the $N!$ amplitudes corresponding to each possible assignment of substates for the bosons in the initial and final state. Dividing by the number of possible relabellings of the substates in $|N_{\rm i}\rangle$ and $|N_{\rm f}\rangle$ gives for the total propagation probability $|N!\,{\rm e}^{iEt}|^2/(N!)^2 = 1$. Now, however, consider the introduction of an extra boson during the propagation, with amplitude A. The final state $|(N+1)_{\rm f}\rangle$ is now divided into $(N+1)$ substates, which may be labelled in $(N+1)!$ ways. The total probability for the propagation with the addition of one extra boson is then $|(N+1)!A\,{\rm e}^{iEt}|^2/(N+1)!) = (N+1)|A|^2$: it is enhanced by a factor $(N+1)$ relative to the probability $|A|^2$ for the introduction of the boson into an initially empty state. Taking the continuum limit then gives the $(1+f)$ correction factor.

A suitable quantity by which to measure the approach to thermal equilibrium of a system of identical particles is

$$H = \sum_{\alpha} \{f_\alpha(\log f_\alpha - 1) - \theta_\alpha s_\alpha(\log s_\alpha - 1)\}\,. \tag{A.15}$$

This H evolves with time according to

$$\frac{{\rm d}H}{{\rm d}t} = \sum_\alpha \frac{{\rm d}f_\alpha}{{\rm d}t}\log(f_\alpha/s_\alpha)\,, \tag{A.16}$$

which may be written, using the modified Boltzmann equation (A.14), as

$$\frac{{\rm d}H}{{\rm d}t} \sum_{\alpha,\beta} \{F_\beta S_\alpha|\mathcal{M}(\beta\to\alpha)|^2 - F_\alpha S_\beta|\mathcal{M}(\alpha\to\beta)|^2\}\log(F_\alpha/S_\alpha)\,. \tag{A.17}$$

We first assume T invariance, so that $\mathcal{M}(\beta\to\alpha) = \mathcal{M}(\alpha\to\beta)$. In this case, eq. (A.17) becomes [analogous to eq. (A.6)]

$$\begin{aligned}\frac{{\rm d}H}{{\rm d}t} &= \sum_{\alpha,\beta} |\mathcal{M}(\alpha\to\beta)|^2\{F_\beta S_\alpha - F_\alpha S_\beta\}\log(F_\alpha/S_\alpha)\\ &= \sum_{\alpha,\beta} |\mathcal{M}(\alpha\to\beta)|^2\{F_\beta S_\alpha - F_\alpha S_\beta\}\log\frac{F_\alpha S_\beta}{F_\beta S_\alpha}\\ &\leq 0\,,\end{aligned} \tag{A.18}$$

where the final inequality follows from (A.7). Hence the system will evolve on average to an equilibrium state in which H is minimal, and

$$\frac{F_\alpha}{S_\alpha}=\frac{F_\beta}{S_\beta}=\frac{f_{\alpha_1}}{1+\theta_{\alpha_1}f_{\alpha_1}}\cdots\frac{f_{\alpha_m}}{1+\theta_{\alpha_m}f_{\beta_m}}$$
$$=\frac{f_{\beta_1}}{1+\theta_{\beta_1}f_{\beta_1}}\cdots\frac{f_{\beta_n}}{1+\theta_{\beta_n}f_{\beta_n}} \tag{A.19}$$

for all set of α_i and β_i corresponding to initial and final states of possible scattering processes. In analogy with (A.9), energy conservation then implies

$$\frac{f_\alpha}{1+\theta_\alpha f_\alpha}=\mathrm{e}^{-(E_\alpha-\mu_\alpha)/T}, \tag{A.20}$$

where μ_α represents possible absolutely conserved quantum numbers. Solving (A.20) for f_α gives [cf., eq. (2.4.11)]

$$f_\alpha=[\mathrm{e}^{(E_\alpha-\mu_\alpha)/T}-\theta_\alpha]^{-1}, \tag{A.21}$$

which is the usual equilibrium Bose–Einstein or Fermi–Dirac distribution ($\theta=+1$ for bosons, $\theta=-1$ for fermions).

When CP invariance is violated, one must use unitarity to prove the H theorem $\mathrm{d}H/\mathrm{d}t\leq 0$. In the presence of indistinguishable particles the unitarity relation (2.1.3) is modified, and becomes [26]

$$\sum_j S_j|\mathcal{M}(i\to j)|^2=\sum_j S_j|\mathcal{M}(j\to i)|^2, \tag{A.22}$$

where S_j is the product of quantum statistics correction factors defined above. The CPT-invariance constraint (2.1.1) yields the results [analogous to eqs (2.1.4) and (2.1.5)]

$$\sum_j S_j|\mathcal{M}(i\to j)|^2=\sum_j S_j|\mathcal{M}(j\to i)|^2=\sum_j S_j|\mathcal{M}(j\to i)|^2$$
$$=\sum_j S_j|\mathcal{M}(\bar{i}\to j)|^2. \tag{A.23}$$

To show that no asymmetry between particles and antiparticles may be generated in thermal equilibrium, we must prove that

$$\sum_j F_j|\mathcal{M}(j\to\bar{i})|^2=\sum_j F_j|\mathcal{M}(j\to i)|^2, \tag{A.24}$$

where $F_j=f_{j_1}\cdots f_{j_n}$ is the product of the relevant incoming particle Bose-Einstein or Fermi-Dirac equilibrium phase-space densities. According to eq. (A.19), in thermal equilibrium

$$\frac{F_\alpha}{S_\alpha}=\frac{F_\beta}{S_\beta}, \tag{A.25}$$

for all sets of particles α, β. (This is the generalization of the result for distinguishable Maxwell–Boltzmann particles used in subsect. 2.1 that all states of a given energy are equally populated.) Inserting the relation (A.25) in the unitarity equation (A.23) gives directly the desired result (A.24).

Using the unitarity result (A.22), the proof of the H theorem for indistinguishable particles undergoing CP-violating interactions proceeds quite analogously to the distinguishable particle case treated above. On exchanging dummy indices, eq. (A.17) may be written [cf., (A.11)]

$$\frac{\mathrm{d}H}{\mathrm{d}t}=\sum_{\alpha,\beta}|\mathcal{M}(\beta\to\alpha)|^2\{F_\beta S_\alpha\log(F_\alpha/S_\alpha)-F_\beta S_\alpha\log(F_\beta/S_\beta)\}\,. \tag{A.26}$$

The unitarity relation (A.22) implies [cf., (A.10)]

$$\sum_{\alpha,\beta}|\mathcal{M}(\beta\to\alpha)|^2\{F_\beta S_\alpha-F_\alpha S_\beta\}=0\,, \tag{A.27}$$

and inserting this in eq. (A.26) gives

$$\frac{\mathrm{d}H}{\mathrm{d}t}=\sum_{\alpha,\beta}|\mathcal{M}(\beta\to\alpha)|^2\{F_\beta S_\alpha\log\frac{F_\alpha S_\beta}{F_\beta S_\alpha}+F_\beta S_\alpha-F_\alpha S_\beta\}\leqslant 0\,, \tag{A.28}$$

by eq. (A.13). Thus the validity of the H theorem is unaffected by indistinguishable particle effects, even when CP violation is present*.

Appendix B

CONSTRAINTS ON *CP* VIOLATION

In eq. (2.1.9) we showed that to obtain a CP-violating difference between partial cross sections of the form

$$|\mathcal{M}(i\to j)|^2-|\mathcal{M}(\bar{i}\to\bar{j})|^2\neq 0 \tag{B.1}$$

cannot (except when $i\bar{i}$ mixing is possible, as discussed in subsect. 2.4) occur in the Born approximation, and requires loop corrections with absorptive (imaginary) parts. Here, we first show that a difference between cross sections summed over all final states $\overset{(-)}{j}$ with baryon number $\overset{(-)}{B}$

$$\sum_{j\in B}|\mathcal{M}(i\to j)|^2-\sum_{\bar{j}\in\bar{B}}|\mathcal{M}(\bar{i}\to\bar{j})|^2\neq 0\,, \tag{B.2}$$

cannot occur unless the loop corrections are also B-violating. We write

$$\mathcal{M}(i\to j)=\sum_n\mathcal{M}_1(i\to n)\mathcal{M}_2(n\to j)\,, \tag{B.3}$$

* The proof given here finally dispels doubts raised e.g., in refs. [25, 30].

where $\mathcal{M}_1$ is a B-violating Born amplitude (which is necessarily CP conserving) and $\mathcal{M}_2$ represents a correction (analogous to rescattering) which introduces CP violation, but is taken to be B conserving. Then

$$\sum_{j\in B} |\mathcal{M}(i\to j)|^2 = \sum_{j\in B}\left[\sum_n \mathcal{M}_1(i\to n)\mathcal{M}_2(n\to j)\right] \times\left[\sum_{n'} \mathcal{M}_1(i\to n')\mathcal{M}_2(n'\to j)\right]^{\dagger} . \tag{B.4}$$

Since $\mathcal{M}_2$ is baryon conserving, it must obey the unitarity constraint

$$\sum_{j\in B} \mathcal{M}_2(n\to j)\mathcal{M}_2^*(j\to n') = \delta_{nn'} , \tag{B.5}$$

when summed only over the accessible states j of a given baryon number. Using this result, eq (B.4) becomes

$$\begin{aligned}\sum_{j\in B} |\mathcal{M}(i\to j)|^2 &= \left|\sum_{n\in B} \mathcal{M}_1(i\to n)\right|^2 \\ &= \left|\sum_{\bar{n}\in \bar{B}} \mathcal{M}_1(\bar{i}\to \bar{n})\right|^2 \\ &= \sum_{\bar{j}\in\bar{B}} |\mathcal{M}_1(\bar{i}\to\bar{j})|^2 ,\end{aligned} \tag{B.6}$$

since $\mathcal{M}_1$ is CP conserving. Hence to obtain a CP-violating difference of the form (B.2) after summing over states $j\in B$, the hamiltonian responsible for the loop correction (hence $\mathcal{M}_2$) must violate baryon number. This constraint severely restricts baryon generation in gauge models, as discussed in ref. [11, 29].

Appendix C

NUMBER DENSITY INTEGRALS

In this appendix, we give some integrals of equilibrium phase-space distributions used in sect. 2. As elsewhere, we take units such that (k is Boltzmann's constant)

$$\hbar = c = k = 1 . \tag{C.1}$$

We consider a uniform ideal gas of particles with mass m in thermal equilibrium at temperature T, and write

$$x \equiv m/T . \tag{C.2}$$

The number density of such particles in phase space is given by (see subsect. 2.1)*

* The factor $1/(2\pi)^3$ arises because a phase-space cell containing a single mode of the field has volume h^3 rather than $\hbar^3$.

$$\frac{dN}{d^3\boldsymbol{p}\, d^3x} \equiv f(\boldsymbol{p}) = f(p) = \frac{g}{(2\pi)^3} \frac{1}{e^{(E-\mu)/T} + \theta}\,,$$

$$E = \sqrt{\boldsymbol{p}^2 + m^2}\,, \tag{C.3}$$

where $\theta = \underset{(-)}{+}1$ for particles obeying Fermi–Dirac (FD) (Bose–Einstein (BE)) statistics, and $\theta = 0$ in the classical (distinguishable particles) approximation of Maxwell–Boltzmann (MB) statistics, used extensively above. g gives the number of spin states for each particle. [Usually $g = 2s + 1$ for massive particles and $g = 2$, $s > 0$ for $m = 0$. However, because the interaction cross sections for the various spin states of a weakly interacting particle may differ, its full complement of spin states may not be in thermal equilibrium at a particular temperature. Thus, for example, the second spin state of a neutrino carrying a small but non-zero mass may not be in thermal equilibrium if only one of its spin states may participate directly in weak interactions.] The μ appearing in eq. (C.3) is a possible chemical potential which serves to constrain the total number of particles. When a species of particles are distributed in phase space according to eq. (C.3), they are said to be in *kinetic* equilibrium. They are in *chemical* equilibrium only if $\mu = 0$ (unless they carry an absolutely conserved quantum number, such as electric charge, with respect to which the complete system is not neutral). Baryons in the early universe should quickly be brought into kinetic equilibrium by collisions; only much slower B- and CP-violating scatterings can produce chemical equilibrium, with $\mu = 0$.

The total particle number density may be obtained by integrating (C.3) over the available momentum states:

$$n\left(T, \frac{m}{T}, \frac{\mu}{T}\right) = \int_0^\infty \frac{d^3\boldsymbol{p}}{(2\pi)^3} \frac{g}{e^{(E-\mu)/T} + \theta}$$

$$= \frac{gT^3}{2\pi^3} \int_x^\infty z\sqrt{z^2 - x^2}\,[e^{(z-\mu/T)} + \theta]^{-1}\, dz\,. \tag{C.4}$$

C.1. NUMBER DENSITIES WITH MAXWELL–BOLTZMANN STATISTICS

In the approximation of Maxwell–Boltzmann statistics, the number density integral becomes

$$n_{\mathrm{MB}}\left(T, \frac{m}{T}, \frac{\mu}{T}\right) = \frac{gT^3}{2\pi^2} \int_0^\infty z\sqrt{z^2 - x^2}\, e^{-(z-\mu/T)}\, dz\,, \qquad z = E/T\,. \tag{C.5}$$

This integral may be expressed in terms of a modified Bessel function (using the notation of ref. [27])

$$n_{\mathrm{MB}}\left(T, \frac{m}{T}, \frac{\mu}{T}\right) = \frac{gT^3\, e^{-\mu/T}}{2\pi^2} x^2 K_2(x)\,. \tag{C.6}$$

For large x (corresponding to low temperatures) the asymptotic expansion of the

Bessel function

$$K_\nu(x) \simeq \sqrt{\frac{\pi}{2x}}\,\mathrm{e}^{-x}\left[1+\frac{(\nu^2-1/4)}{2x}+\frac{(\nu^2-1/4)(\nu^2-9/4)}{8x^2}+\cdots\right], \qquad (x \gg 1), \tag{C.7}$$

gives the usual Boltzmann factor

$$n_{\mathrm{MB}}\left(T, \frac{m}{T}, \frac{\mu}{T}\right) \simeq g\left(\frac{mT}{2\pi}\right)^{3/2} \mathrm{e}^{-(m-\mu)/T}\left[1+\frac{15}{8x}+\frac{105}{128x^2}-\cdots\right], \qquad (m \gg T). \tag{C.8}$$

At small x (high temperatures) expansion of the Bessel function

$$K_\nu(x) = \frac{1}{2}\sum_{k=0}^{\nu-1}(-1)^k\frac{(\nu-k-1)!}{k!}(\tfrac{1}{2}x)^{2k-\nu} + (-1)^{\nu+1}\sum_{k=0}^{\infty}\frac{(\frac{1}{2}x)^{2k+\nu}}{(k!(\nu+k)!)}\left[\log(\tfrac{1}{2}x)+\gamma+\sum_{i=1}^{k}\frac{1}{2i}+\sum_{i=1}^{k+\nu}\frac{1}{2i}\right], \tag{C.9}$$

gives

$$n_{\mathrm{MB}}\left(T, \frac{m}{T}, \frac{\mu}{T}\right) \simeq \frac{gT^3\,\mathrm{e}^{\mu/T}}{\pi^2}\left[1-\tfrac{1}{4}x^2-\tfrac{1}{64}x^4(4\log(\tfrac{1}{2}x)+4\gamma-3)+\cdots\right], \qquad (m \ll T), \tag{C.10}$$

where $\gamma(\simeq 0.5772)$ is Euler's constant. Hence the ratio of the number density of a massive species of particles to the number density of photons at high temperatures is roughly

$$\begin{aligned} n_{\mathrm{MB}}\left(T, \frac{m}{T}, \frac{\mu}{T}\right)/n^{\gamma}_{\mathrm{MB}}(T,0,0) &\simeq \tfrac{1}{2}g\,\mathrm{e}^{\mu/T}\left[1-\left(\frac{m}{2T}\right)^2-\cdots\right] \qquad (m \ll T) \\ &\simeq \tfrac{1}{2}g\left[1+\frac{\mu}{T}-\left(\frac{m}{2T}\right)^2+\tfrac{1}{2}\left(\frac{\mu}{T}\right)^2-\cdots\right] \qquad (\mu, m \ll T), \end{aligned} \tag{C.11}$$

where, for consistency, we have approximated the photon number distribution by Maxwell–Boltzmann statistics.

C.2. MOMENTS OF THE MAXWELL–BOLTZMANN DISTRIBUTION

The mean time dilation in the decay lifetime of Maxwell–Boltzmann particles in thermal equilibrium at a temperature $T = m/x$ is given by

$$\begin{aligned} \left\langle\frac{m}{E}\right\rangle_{\mathrm{MB}} &= \frac{\int_x^\infty x\sqrt{z^2-x^2}\,\mathrm{e}^{-z}\,\mathrm{d}z}{\int_x^\infty z\sqrt{z^2-x^2}\,\mathrm{e}^{-z}\,\mathrm{d}z} \\ &= \frac{K_1(x)}{K_2(x)}. \end{aligned} \tag{C.12}$$

At high temperatures (small x) the particles are relativistic, with energies $\sim T$, and

$$\left\langle \frac{m}{E} \right\rangle_{\mathrm{MB}} \simeq \frac{m}{T}[1+\tfrac{1}{2}x^2(\log(\tfrac{1}{2}x)+\gamma)+\cdots], \qquad (m \ll T), \tag{C.13}$$

while in the non-relativistic limit (high x):

$$\left\langle \frac{m}{E} \right\rangle_{\mathrm{MB}} \simeq 1 - \frac{3}{2x} + \frac{15}{8x^2} - \cdots, \qquad (m \gg T). \tag{C.14}$$

Next, we consider the mean energy of the particles in a Maxwell–Boltzmann gas. This may be found directly by integrating (C.1) with weight $E = Tz$. However, we shall here use a less direct method, since it introduces several useful results. In non-relativistic classical statistical mechanics the equipartition theorem states that each (quadratic) degree of freedom of a particle which is Boltzmann distributed has a mean thermal energy of $\frac{1}{2}T$. To obtain the relativistic generalization of the equipartition theorem one must find a quantity $Q(p)$ whose mean value

$$\langle Q \rangle_{\mathrm{MB}} = \frac{\int \mathrm{d}^3p \, \mathrm{e}^{-E/T} Q(p)}{\int \mathrm{d}^3p \, \mathrm{e}^{-E/T}}, \tag{C.15}$$

depends only on T and not on m/T or μ/T. A suitable such quantity is

$$Q = \boldsymbol{p}_i \frac{\partial E}{\partial \boldsymbol{p}_i}, \tag{C.16}$$

where i labels some component of the three-momentum $\boldsymbol{p}$. Inserting the choice (C.16) into (C.15) one finds on integrating by parts:

$$\langle Q \rangle_{\mathrm{MB}} \equiv \left\langle \boldsymbol{p}_i \frac{\partial E}{\partial \boldsymbol{p}_i} \right\rangle_{\mathrm{MB}} = T, \tag{C.17}$$

for all m and μ. In the non-relativistic limit, $E = m + \Sigma\, p_i^2/2m$, so that $Q = mv_i^2$, and summing over i one regains the standard result that the mean kinetic energy of a particle in thermal equilibrium at a temperature T is $\frac{3}{2}T$. For relativistic particles, $Q = \boldsymbol{p}_i^2/E$, so that

$$\left\langle \frac{\boldsymbol{p}^2}{E} \right\rangle_{\mathrm{MB}} = \left\langle E - \frac{m^2}{E} \right\rangle_{\mathrm{MB}} = 3T, \tag{C.18}$$

regardless of the value of m (or μ). This result is therefore a relativistic generalization of the equipartition theorem. One may now use eqs. (C.12) and (C.18) to find

$$\begin{aligned} \langle E \rangle_{\mathrm{MB}} &= 3T + m\frac{K_1(x)}{K_2(x)} \\ &\simeq 3T + \frac{m^2}{T} + \cdots, \qquad (T \gg m), \end{aligned}$$

$$\simeq m+\frac{3T}{2}+\frac{15T^2}{8m}-\cdots \qquad (T \ll m)\,. \tag{C.19}$$

Eq. (C.6) then gives the energy density of the gas:

$$\begin{aligned}\rho_{\mathrm{MB}}\left(T,\frac{m}{T},\frac{\mu}{T}\right) &= n_{\mathrm{MB}}\left(T,\frac{m}{T},\frac{\mu}{T}\right)\langle E\rangle_{\mathrm{MB}} \\ &= \frac{3gT^4\,\mathrm{e}^{\mu/T}}{2\pi^2}[x^2K_2(x)+\tfrac{1}{3}x^3K_1(x)] \\ &\simeq gm\left(\frac{mT}{2\pi}\right)^{3/2}\mathrm{e}^{-(m-\mu)/T}\left[1+\frac{27}{8x}+\cdots\right] \qquad (m \ll T) \\ &\simeq \frac{3gT^4\,\mathrm{e}^{\mu/T}}{\pi^2}[1-\tfrac{1}{12}x^2+\cdots] \qquad (T \gg m)\,. \end{aligned} \tag{C.20}$$

C.3. NUMBER DENSITIES WITH QUANTUM STATISTICS

For fermions (bosons) the number density integral (C.4) becomes

$$n\left(T,x=\frac{m}{T},\frac{\mu}{T}\right)=\frac{gT^3}{2\pi^2}\int_x^\infty z\sqrt{z^2-x^2}\,\mathrm{e}^{-z}/[1\underset{(-)}{+}\mathrm{e}^{-(z-\mu/T)}]\,\mathrm{d}z\,, \tag{C.21}$$

which cannot be expressed in terms of the usual special functions, but may formally be written as

$$n_{\substack{\mathrm{FD}\\(\mathrm{BE})}}\left(T,x=\frac{m}{T},\frac{\mu}{T}\right)=\frac{gT^3}{2\pi^2}\sum_{k=0}^{\infty}\frac{[(\mp)\,\mathrm{e}^{\mu/T}]^k}{(k+1)}x^2K_2((k+1)x)\,. \tag{C.22}$$

Since $K_2(z)$ falls off exponentially for large z, this series converges rapidly, and is convenient for numerical evaluation. However, for small x the series is not uniformly convergent. Nevertheless it is easy to find the high temperature behavior of eq. (C.21) when $\mu=0$:

$$\begin{aligned} n_{\mathrm{BE}}\left(T,\frac{m}{T},0\right) &\simeq \frac{gT^3}{\pi^2}\left[\zeta(3)+\frac{1}{4}\left(\frac{m}{T}\right)^2\log\frac{m}{T}+\cdots\right], \\ n_{\mathrm{FD}}\left(T,\frac{m}{T},0\right) &\simeq \frac{gT^3}{\pi^2}\left[\tfrac{3}{4}\zeta(3)-\left(\frac{m}{T}\right)^2\frac{\log 2}{2}+\cdots\right], \qquad (T\gg m)\,, \end{aligned} \tag{C.23}$$

where $\zeta(3)=\sum_{i=1}^{\infty}1/i^3\simeq 1.202$, while for Maxwell–Boltzmann statistics, eq. (C.10) gives

$$n_{\mathrm{MB}}\left(T,\frac{m}{T},0\right)\simeq\frac{gT^3}{\pi^2}\left[1-\frac{1}{4}\left(\frac{m}{T}\right)^2+\cdots\right], \qquad (T\gg m)\,. \tag{C.24}$$

At low temperatures, of course, the effects of quantum statistics (represented by the terms in (C.22) with $k>0$) are exponentially unimportant, and thus [from eq. (C.8)]

$$n_{\mathrm{BE}}\left(T,\frac{m}{T},0\right)\simeq n_{\mathrm{FD}}\left(T,\frac{m}{T},0\right)\simeq n_{\mathrm{MB}}\left(T,\frac{m}{T},0\right)$$
$$\simeq g\left(\frac{mT}{2\pi}\right)^{3/2}\mathrm{e}^{-m/T}\,,\qquad (T\ll m)\,. \tag{C.25}$$

[Further terms in the asymptotic series are given in eq. (C.37)]. The ratio $n_{\mathrm{FD}}(T,m/T,0)/n_{\mathrm{MB}}(T,m/T,0)$ goes monotonically from $\frac{3}{4}\zeta(3)(\zeta(3))$ to 1 as m/T goes from 0 to ∞; the largest changes occur around $m=T$.
(BE)

C.4. MASSLESS PARTICLE NUMBER DENSITIES

For massless particles, it is simple to perform the integrals (C.4) retaining the chemical potential μ, in terms of the polylogarithm functions [28]

$$\mathrm{Li}_n(x)=\int_0^x\frac{\mathrm{Li}_{n-1}(t)}{t}\,\mathrm{d}t=\sum_{k=1}^{\infty}\frac{x^k}{k^n}$$
$$=\frac{(-1)^n}{\Gamma(n-1)}\int_0^1\frac{\log^{n-2}(t)\log(1-xt)}{t}\,\mathrm{d}t\,,\qquad (n\geqslant 2)\,,$$
$$\mathrm{Li}_1(x)=-\log(1-x)\,,\qquad \mathrm{Li}_n(1)=\zeta(n)\,,$$
$$\mathrm{Li}_n(-1)=(2^{1-n}-1)\zeta(n)\,, \tag{C.26}$$

using the formula

$$\int_0^\infty\frac{z^{n-1}\,\mathrm{d}z}{\mathrm{e}^z-Y}=\frac{\mathrm{Li}_n(Y)\Gamma(n)}{Y}\,,$$
$$=\Gamma(n)\,,\qquad (Y=0)\,. \tag{C.27}$$

One finds* ($\zeta(2)=\frac{1}{6}\pi^2\simeq 1.645$, $\zeta(3)\simeq 1.202$, $\zeta(4)=\frac{1}{90}\pi^4\simeq 1.082$)

* For these expansions, we used the relations (valid for $0<x\ll 1$):

$$\mathrm{Li}_2(1-x)=\zeta(2)-\mathrm{Li}_2(x)-\log x\log(1-x)$$
$$\simeq\zeta(2)-x(-\log(x)+1)-\cdots\,,$$
$$\mathrm{Li}_3(1-x)\simeq\zeta(3)-x\zeta(2)-\tfrac{1}{4}x^2(2\log x+2\zeta(2)-3)+\cdots\,,$$
$$\mathrm{Li}_n(1-x)\simeq\zeta(n)-x\zeta(n-1)+\cdots\,,\qquad (n>2)\,.$$

The last result may trivially be derived from the series expansion in (C.26). It is also convenient to apply the relation

$$\mathrm{Li}_n(x^2)=2^{n-1}[\mathrm{Li}_n(x)+\mathrm{Li}_n(-x)]\,,$$

to obtain (for $|x|\ll 1$)

$$\mathrm{Li}_n(-1+x)\simeq(2^{1-n}-1)\zeta(n)+x(1-2^{2-n})\zeta(n-1)+\cdots\,,\qquad (n>2)\,.$$

$$n_{\mathrm{BE}}\left(T, 0, \frac{\mu}{T}\right) = \frac{gT^3}{\pi^2} \mathrm{Li}_3(\mathrm{e}^{\mu/T}) \simeq \frac{gT^3}{\pi^2}\left[\zeta(3) + \frac{\mu}{T}\zeta(2) + \cdots\right],$$

$$n_{\mathrm{FD}}\left(T, 0, \frac{\mu}{T}\right) = -\frac{gT^3}{\pi^2} \mathrm{Li}_3(-\mathrm{e}^{\mu/T}) \simeq \frac{gT^3}{\pi^2}\left[\tfrac{3}{4}\zeta(3) + \frac{\mu}{2T}\zeta(2) + \cdots\right],$$

$$n_{\mathrm{MB}}\left(T, 0, \frac{\mu}{T}\right) = \frac{gT^3}{\pi^2} \mathrm{e}^{\mu/T}, \tag{C.28}$$

where $\mu \leq 0$ for the Bose–Einstein case. [If the number density of bosons in a system increases above its value given above when $\mu = 0$, then Bose condensation must occur: many particles collect in the ground state and the approximation of a sum over discrete energy levels by the integral (C.4) is no longer satisfactory (so that finite volume effects become important).]

C.5. ENERGY DENSITIES

The expansion rate of the early universe is determined by its energy density, which is conveniently parametrized in terms of the "effective number of species in thermal equilibrium", defined by

$$\rho = \xi\left(\frac{m}{T}, \frac{\mu}{T}\right)\rho_\gamma, \tag{C.29}$$

where ρ_γ is the energy density of a (genuine Bose–Einstein) photon gas

$$\rho_\gamma = \frac{6\zeta(4)}{\pi^2} T^4 = \tfrac{1}{15}\pi^2 T^4 . \tag{C.30}$$

The contribution of a massive particle with $\mu = 0$ to ξ is given by eq. (C.25) as

$$\xi\left(\frac{m}{T}, 0\right) \simeq \frac{15g}{\pi^2} \frac{m}{T}\left(\frac{m}{2\pi T}\right)^{3/2} \mathrm{e}^{-m/T}, \qquad (m \gg T), \tag{C.31}$$

which is usually negligibly small. Eq. (C.20) gives us complete form of ξ for particles obeying Maxwell–Boltzmann statistics. At high temperatures, one finds

$$\xi_{\mathrm{BE}}\left(\frac{m}{T}, 0\right) \simeq \zeta(4)\left[1 - \frac{5}{2\pi^2}\left(\frac{m}{T}\right)^2 + \cdots\right] \simeq 1.082g + \cdots,$$

$$\xi_{\mathrm{FD}}\left(\frac{m}{T}, 0\right) \simeq \tfrac{7}{8}\zeta(4)g\left[1 - \frac{5}{4\pi^2}\left(\frac{m}{T}\right)^2 + \cdots\right] \simeq 0.947g + \cdots,$$

$$\xi_{\mathrm{MB}}\left(\frac{m}{T}, 0\right) \simeq g\left[1 - \tfrac{1}{12}\left(\frac{m}{T}\right)^2 + \cdots\right]. \tag{C.32}$$

For zero-mass particles with non-zero chemical potential*

$$\xi_{\rm BE}\left(0,\frac{\mu}{T}\right)=\frac{g}{2}\,{\rm Li}_4\,({\rm e}^{\mu/T})\simeq\zeta(4)g\left[1+\frac{\mu}{T}\frac{\zeta(3)}{\zeta(4)}+\cdots\right],$$

$$\xi_{\rm FD}\left(0,\frac{\mu}{T}\right)=\frac{-g}{2}\,{\rm Li}_4\,(-{\rm e}^{\mu/T})\simeq\tfrac{7}{8}\zeta(4)g\left[1+\frac{\mu}{T}\frac{6\zeta(3)}{7\zeta(4)}+\cdots\right],$$

$$\xi_{\rm MB}\left(0,\frac{\mu}{T}\right)=g\,{\rm e}^{\mu/T}, \tag{C.33}$$

where the expansions are for $\mu/T \ll 1$. The mean energies $(=\rho/n)$ for zero-mass particles in thermal equilibrium with $\mu=0$ are given by

$$\langle E\rangle_{\rm BE}\simeq 3\frac{\zeta(4)}{\zeta(3)}T\simeq 2.6T,$$

$$\langle E\rangle_{\rm FD}\simeq\frac{7}{2}\frac{\zeta(4)}{\zeta(3)}T\simeq 3.2T,$$

$$\langle E\rangle_{\rm MB}\simeq 3T,\qquad (m=\mu=0), \tag{C.34}$$

while the dispersions of the energy distributions about these means are

$$\sigma=[\langle E^2\rangle-\langle E\rangle^2]^{1/2},$$

$$\sigma_{\rm BE}\simeq 1.77T,$$

$$\sigma_{\rm FD}\simeq 1.75T,$$

$$\sigma_{\rm MB}\simeq\sqrt{3}{\rm T}\simeq 1.73T,\qquad (m=\mu=0). \tag{C.35}$$

Since no known bosons carry absolutely conserved quantum numbers (other than electric charge) it seems unlikely that a high chemical potential for a boson species, leading to Bose condensation, could be enforced in the early universe. However, according to "cold" models for the early universe (discussed in sect. 4), degenerate Fermi gases existed at early times, having

$$n_{\rm FD}\left(T,0,\frac{\mu}{T}\right)\simeq\frac{g\mu^3}{6\pi^2}\left[1+6\zeta(2)\left(\frac{T}{\mu}\right)^2+\cdots\right],$$

$$\rho_{\rm FD}\left(T,0,\frac{\mu}{T}\right)\simeq\frac{g\mu^4}{8\pi^2}\left[1+12\zeta(2)\left(\frac{T}{\mu}\right)^2+\cdots\right],\qquad (\mu\gg T). \tag{C.36}$$

* For these expansions we used the result $(x \gg 1)$

$$\mathrm{Li}_n\,(-x)\simeq-\frac{\log^n(x)}{n!}-\frac{\log^{n-2}(x)}{(n-2)!}\zeta(2)-\cdots,$$

where the second term is absent for $n \leqslant 2$.

C.6. NON-RELATIVISTIC LIMIT OF QUANTUM STATISTICS

In the non-relativistic limit $m \gg T$, eq. (C.4) becomes

$$n_{\substack{\mathrm{FC} \\ (\mathrm{BE})}}\left(T, \frac{m}{T}, \frac{\mu}{T}\right) \approx \frac{g}{2\pi^2} \int_0^{\infty} \frac{p^2 \, \mathrm{d}p}{\mathrm{e}^{(m-\mu)/T} \mathrm{e}^{p^2/(2mT)} \underset{(-)}{+} 1}, \qquad (m \gg T), \tag{C.37}$$

which may be written in terms of polylogarithm functions

$$n_{\mathrm{BE}}\left(T, \frac{m}{T}, \frac{\mu}{T}\right) \approx g\left(\frac{mT}{2\pi}\right)^{3/2} \mathrm{Li}_{3/2}\left(\mathrm{e}^{-(m-\mu)/T}\right)$$

$$n_{\mathrm{FD}}\left(T, \frac{m}{T}, \frac{\mu}{T}\right) \approx -g\left(\frac{mT}{2\pi}\right)^{3/2} \mathrm{Li}_{3/2}\left(-\mathrm{e}^{-(m-\mu)/T}\right), \qquad (m \gg T), \tag{C.38}$$

where from eq. (C.26)

$$\mathrm{Li}_{3/2}(x) = \sum_{k=1}^{\infty} \frac{x^k}{k^{3/2}}, \tag{C.38}$$

while

$$n_{\mathrm{MB}}\left(T, \frac{m}{T}, \frac{\mu}{T}\right) \simeq g\left(\frac{mT}{2\pi}\right)^{3/2} \mathrm{e}^{-(m-\mu)/T}, \qquad (m \gg T). \tag{C.40}$$

(Note that for $n \leqslant 2$, $\mathrm{Li}_n(x)$ has an infinite derivative at $x = 1$). The energy densities of massive fermions and bosons in thermal equilibrium at low temperatures are given by

$$\rho_{\mathrm{BE}}\left(T, \frac{m}{T}, \frac{\mu}{T}\right) \approx n_{\mathrm{BE}} m + \frac{3}{4m}\left(\frac{mT}{2\pi}\right)^{5/2} \mathrm{Li}_{5/2}\left(\mathrm{e}^{-(m-\mu)/T}\right),$$

$$\rho_{\mathrm{FD}}\left(T, \frac{m}{T}, \frac{\mu}{T}\right) \approx n_{\mathrm{FD}} m - \frac{3}{4m}\left(\frac{mT}{2\pi}\right)^{5/2} \mathrm{Li}_{5/2}\left(-\mathrm{e}^{-(m-\mu)/T}\right). \tag{C.41}$$

C.7. A TWO-PARTICLE INTEGRAL

In subsect. 2.3, the two-particle integral

$$I \equiv \int \frac{\mathrm{d}^3 p_1}{2E_1} \int \frac{\mathrm{d}^3 p_2}{2E_2} f^{\mathrm{eq}}(p_1) f^{\mathrm{eq}}(p_2) \delta(s - m_{\mathrm{X}}^2), \tag{C.42}$$

where $s = (p_1 + p_2)^2$, appeared in connection with the rate for $2 \to 2$ scattering of massless b via nearly on-shell X exchange. Performing the integral over c.m. angles gives

$$\begin{aligned} I &= \pi^2 \int \mathrm{d}E_1 \int \mathrm{d}E_2 \, \mathrm{e}^{-E_1/T} \mathrm{e}^{-E_2/T} \int_0^{4E_1E_2} \mathrm{d}s \, \delta(s - m_{\mathrm{X}}^2) \\ &= \pi^2 T m_{\mathrm{X}} K_1\left(\frac{m_{\mathrm{X}}}{T}\right), \end{aligned} \tag{C.43}$$

where K_1 is a modified Bessel function. Making use of (C.6) and (C.12), I may be written as [cf., (2.3.17)]

$$I=\left\langle\frac{m_X}{E}\right\rangle\frac{2\pi^4}{m}n_X^{eq}$$

$$=\left(\frac{\langle\Gamma_X\rangle}{\Gamma_X}\right)n_X^{eq}\frac{2\pi^4}{m_X}. \tag{C.44}$$

References

[1] E.W. Kolb and S. Wolfram, Phys. Lett. 91B (1980) 217
[2] A.D. Sakharov, ZhETF Pis'ma 5 (1967) 32
[3] M. Yoshimura, Phys. Rev. Lett. 41 (1978) 281; (E: 42 (1979) 746)
[4] S. Dimopoulos and L. Susskind, Phys. Rev. D18 (1978) 4500; Phys. Lett. 81B (1979) 416
[5] D. Toussaint, S.B. Treiman, F. Wilczek and A. Zee, Phys. Rev. D19 (1979) 1036
[6] S. Weinberg, Phys. Rev. Lett. 42 (1979) 850
[7] J. Ellis, M.K. Gaillard and D.V. Nanopoulos, Phys. Lett. 80B (1979) 360; (E:82B (1979) 464)
[8] A. Yu. Ignatev et al., Phys. Lett. 87B (1979) 114
[9] S. Weinberg, Gravitation and cosmology (Wiley, New York, 1972), ch. 15;
C.W. Misner, K.S. Thorne and J.A. Wheeler, Gravitation (Freeman, San Francisco, 1973), ch. 27, 28;
L.D. Landau and E.M. Lifshitz, The classical theory of fields (Pergamon, Oxford, 1975), 4th ed., ch. 14
[10] V.B. Berestetskii, E.M. Lifshitz and L.P. Pitaevskii, Relativistic quantum theory (Pergamon, Oxford, 1971), p. 239
[11] D.V. Nanopoulos and S. Weinberg, Phys. Rev. D20 (1979) 2484
[12] J. Ellis, M.K. Gaillard, and D.V. Nanopoulos, Nucl. Phys. B109 (1976) 213
[13] B. Lee and S. Weinberg, Phys. Rev. Lett. 39 (1977) 169;
S. Wolfram, Phys. Lett. 82B (1979) 65
[14] A.D. Dolgov and Y.B. Zeldovich, Cosmology and elementary particles, preprint (1979)
[15] R.E. Marshak, Riazuddin and C.P. Ryan, Theory of weak interactions in particle physics (Interscience, New York, 1969)
[16] A.T. Buras, J. Ellis, M.K. Gaillard, and D.V. Nanopoulos, Nucl. Phys. B135 (1978) 66
[17] Nature 272 (1978) 662
[18] D.C. Montgomery and D.A. Tidman, Plasma kinetic theory (McGraw-Hill, 1964)
[19] T-Y Wu, Kinetic equations of gases and plasmas (Addison-Wesley, 1966)
[20] H.D. Politzer and S. Wolfram, Phys. Lett. 82B (1979) 242;
P.Q. Hung, Phys. Rev. Lett. 42 (1979) 873
[21] B.S. DeWitt, Phys. Rev. 162 (1967) 1239
[22] L.D. Landau and E.M. Lifshitz, Statistical physics, 2nd ed. (Pergamon Press, 1969)
[23] J. Ellis and G. Steigman, Non-equilibrium in the early universe, CERN preprint TH-2745
[24] L.D. Landau and E.M. Lifshitz, Fluid mechanics (Pergamon, 1959) p. 304;
J.M. Stewart, Non-equilibrium relativistic kinetic theory (Springer-Verlag, 1971)
[25] A. Aharony, *in* Modern developments in thermodynamics (Wiley, New York, 1973), pp. 95–114
[26] R.V. Wagoner, The early universe, Stanford preprint (1979)
[27] I.S. Gradestyn and I.M. Ryzhik, Table of integrals, series, and products (Academic Press, New York, 1965), p. 951
[28] K.S. Köblig, J.A. Mignaco and E. Remiddi, BIT 10 (1970) 38
[29] J. Harvey, E. Kolb, D. Reiss, and S. Wolfram, Cosmological baryon number generation in gauge theories, Caltech preprint, in preparation
[30] A. Aharony, Phys. Lett. 37A (1971) 45
[31] J.A. Harvey, E.W. Kolb, D.B. Reiss and S. Wolfram, Cosmological constraints on heavy weakly interacting fermions, Caltech preprint OAP-592, CALT-68-784