Evolution of cosmological baryon asymmetries. I. The role of gauge bosons

J. N. Fry, Keith A. Olive, and Michael S. Turner
Astronomy and Astrophysics Center, Enrico Fermi Institute, The University of Chicago, Chicago, Illinois 60637
(Received 11 March 1980)

The time evolution of the baryon asymmetry (kn_B/s) due to the interactions of a superheavy gauge boson (mass $M_X \sim 10^{15}$ GeV, coupling strength $\alpha \sim 1/45$) is obtained by numerically integrating the Boltzmann equations. Particle interactions in the very early universe ($t \lesssim 10^{-35}$ sec) are assumed to be described by the SU(5) grand unification theory. To a good approximation the results depend upon one parameter, $K \equiv 2.9\times 10^{17}\,\alpha$ GeV$/M_X$. If C and CP are not violated in the decays of the superheavy boson no asymmetry develops, and any initial baryon asymmetry is reduced by a factor of $\cong \exp(-5.5K)$. If both C and CP are violated then an initially symmetrical universe evolves a baryon asymmetry which today corresponds to $kn_B/s \cong 7.8\times 10^{-3}\epsilon/[1+(16K)^{1.3}]$, where $\epsilon/2$ is the baryon excess produced when an X-$\bar{X}$ pair decays. Decays and inverse decays of superheavy bosons are primarily responsible for these results (as Weinberg and Wilczek suggested); however for $K \gg 1$ baryon production falls off much less rapidly than they had expected. A gauge boson of mass 3×10^{14} GeV could have generated the observed asymmetry $kn_B/s \cong 10^{-9.8\pm 1.6}$ if $\epsilon \cong 10^{-4.3\pm 1.6}$. In a companion paper the role of Higgs bosons is considered.

I. INTRODUCTION

A. Background

In the late 1940's, Gamow and others[1] suggested that all elements heavier than hydrogen could have been produced by nuclear reactions involving primordial protons and neutrons during the first few minutes of the evolution of the universe. Although this was always an attractive idea, it was not until nearly 20 years later that the extensive nucleosynthesis calculations of Wagoner, Fowler, and Hoyle[2] tested all the details of this scenario. Their calculations showed that although few nuclei heavier than mass 4 were formed because of the lack of stable nuclei with $A=5$ or $A=8$, Gamow's idea could quite nicely explain the observed 25% mass fraction of ^{4}He. This was an accomplishment, for while the production of heavier elements could be accounted for by reactions in stars and by explosive nucleosynthesis events in supernovas, there did not seem to be a natural way to produce this large amount of helium. The big-bang helium-synthesis calculation is now one of the strongest pieces of evidence in support of the big-bang hypothesis.

There is an analogous problem regarding the net amount of matter itself, specifically, say, the baryon to photon ratio n_B/n_γ which is between 3×10^{-11} and 5×10^{-8} for Ω between 0.005 and 2, h between 0.5 and 1.0 (see Sec. II). Also in need of explanation is the discrepancy in the amounts of matter and antimatter. Although microphysics is very nearly particle-antiparticle symmetric, the universe appears to contain negligible amounts of antimatter, the ratio of antimatter to matter being less than 10^{-4} on scales up to clusters of galaxies.[3] Since the number of photons changes with temperature (because of reactions like e^+e^- annihilations at $T \cong 0.5$ MeV), a more useful quantity to consider is the baryon number to specific entropy ratio kn_B/s. The specific entropy is related to the numbers and temperatures of relativistic particles, presently, photons and neutrinos (see Sec. II). The observations give kn_B/s between 4×10^{-12} and 8×10^{-9}, or $kn_B/s = 10^{-9.8\pm 1.6}$. If the expansion of the universe is isentropic (adiabatic), and if baryon number is conserved, kn_B/s is a constant.

These observations deal a death blow to conventional symmetrical big-bang cosmologies. With no baryon-violating forces, a universe that is initially symmetrical remains so, and in the simplest scenario incomplete annihilation leaves a residual baryon and antibaryon to photon ratio, $n_b/n_\gamma = n_{\bar{b}}/n_\gamma \sim 10^{-18}$ (Ref. 3). To avoid this, one must put in the asymmetry as an *ad hoc* initial condition, or invoke some mechanism to separate matter and antimatter before annihilation on scales that are then much larger than the horizon.[4]

When the age of the universe was $t \lesssim 10^{-7}$ sec, the temperature was $kT \gtrsim 1$ GeV, and baryons and antibaryons were both about as abundant as photons. The net baryon number we observe today was present in a small baryon asymmetry of 1 part in 10^{10}. Many authors, including Sakharov, Zeldovich, Hawking, Weinberg, and Parker[5] have suggested that an initially symmetrical universe could have dynamically evolved this asymmetry if there were some mechanism to violate baryon conservation. A recent resurgence of interest in this idea was spurred by the suggestion of Yoshimura[6] and Ignatiev *et al.*[7] that if particle interactions in the early universe were described by grand unified theories (GUT's) such an asymmetry would evolve and could be calculated.

Grand unified theories unify the weak, electromagnetic, and strong interactions by spontaneously broken gauge theories[8,9] with full symmetry restored at energies $E \gtrsim 10^{14}$ GeV, where the extrapolated coupling constants become equal. These theories are largely motivated by the success of the SU(2)×U(1) (Weinberg-Salam) theory of weak and electromagnetic interactions and the SU(3) color gauge theory (quantum chromodynamics) of strong interactions. A useful consequence of grand unification is that baryon number and lepton number are not individually conserved, as baryons and leptons are placed in the same multiplets and are thus mixed under gauge transformations. However, baryon- and lepton-nonconserving interactions are mediated by superheavy gauge and Higgs particles with masses $M \gtrsim 10^{14}$ GeV, and at ordinary energies these conservation laws are nearly obeyed. A measure of this is the stability of the proton, whole lifetime is $\tau \gtrsim 10^{30}$ yr.[10] This limit also requires that unification take place above 10^{14} GeV.

Three ingredients are necessary to generate a baryon asymmetry: (1) baryon-nonconserving forces, (2) C and CP violations to provide an arrow for the direction of the violation, and (3) a departure from thermal equilibrium. GUT's provide (1) and they can naturally contain (2) through loop processes,[11] while (3) arises in the early universe, when particle reaction rates Γ lag behind the rate of cosmological expansion H. Ingredient (3) is not obvious; in fact, Yoshimura missed this point and it was subsequently pointed out by Dimopoulos and Susskind,[12] Toussaint *et al.*,[13] and Weinberg.[14] Qualitatively, though, the need for a departure from equilibrium is easy to understand. If a symmetric universe is in thermal equilibrium, particle-number densities behave as $\exp(-m/kT)$. CPT invariance guarantees that a particle and its antiparticle have the same mass, and unitarity requires that the total production rates of a particle and its antiparticle be equal, so their densities remain equal during expansion and no asymmetry arises, regardless of B-, C-, and CP-violating interactions.

B. Goals and assumptions

Of the many scenarios for producing a cosmological baryon asymmetry, the most promising appears to be the out-of-equilibrium decay scenario suggested by Weinberg[14] and Wilczek.[13] The goal of this paper is to, as much as possible, perform the same careful numerical calculation for the time evolution of the baryon asymmetry as Wagoner, Fowler, and Hoyle[2] did for big-bang nucleosynthesis. Thereby, we hope to provide quantitative verification of what is at present an appealing idea. However, we face a number of problems that force us to be less precise than we would like. First, which (if any) GUT is the correct description of nature? What are the details of the CP violation? Are the contents of the universe in kinetic equilibrium? If not, what are the particle distributions? Are there non-GUT processes which also generate a baryon asymmetry? We address these difficulties by making some working assumptions.

(1) Weinberg and Nanopoulos[11] have shown that in the most general GUT which contains only the usual fermions (arbitrary number of generations, each with "up"-, "down"-, e- and ν_e-like particles) and which breaks down to $SU(3)_c \times SU(2)_L \times U(1)$ there are three generic types of superheavy bosons whose interactions violate baryon number and hence can generate a baryon excess. The first two are gauge bosons: (i) XY—an isodoublet, color triplet (charge $\pm\frac{4}{3}$, $\pm\frac{1}{3}$) and (ii) $X'Y'$—an isodoublet, color triplet (charge $\pm\frac{2}{3}$, $\pm\frac{1}{3}$). The third type is a Higgs boson: (iii) H—isosinglet, color triplet (charge $\pm\frac{1}{3}$). The superheavy gauge bosons of SU(5) are of the XY type; in this paper we will consider baryon generation by the XY gauge bosons and will use the SU(5) theory to calculate the interactions of these bosons. In a more general GUT the couplings of the XY bosons are of the same form with different coupling strength, so that by adjusting the coupling strength our results can be made more general than the SU(5) theory. Since the couplings of the $X'Y'$ bosons are similar to those of the XY bosons we take our results to be valid for these bosons also. In a companion paper[15] we will consider the role of the H boson in baryon generation. In both of these papers we shall consider the effect of just one superheavy species; of course, in an arbitrary GUT there may be many species of superheavy bosons and in order to understand the complete evolution of kn_B/s one must take into account the effects of all the superheavy species. This issue is addressed in a third paper.[16]

(2) The only processes which we include in our computations which involve the superheavy bosons are of order α (their decays and inverse decays) and of order α^2 (baryon-nonconserving fermion collisions, Compton-type and annihilationlike reactions). We find that the second-order processes are not important and expect higher-order processes to be negligible.

(3) The largest uncertainty is the CP violation. At present there is no detailed model of CP violation in SU(5) or other GUT's. Therefore, we parametrize the CP violation by $\epsilon/2 \equiv$ (net baryon number produced in the decay of an X-$\bar{X}$ pair). Because the lowest-order diagram for decay (the

tree graph) conserves CP, ϵ must be of order α or smaller.

(4) Whether the constituents are in partial equilibrium with B- and L-conserving interactions occurring rapidly on the expansion time scale is an important question. If they are not, the Boltzmann equations for the evolution of particle distributions are coupled integrodifferential equations in both time and energy, and further, we do not know what initial distributions to require. If we have partial equilibrium then the system reduces to ordinary differential equations for the numbers of each species per comoving volume. At the energies of interest, $kT \gtrsim 10^{14}$ GeV, it is not clear whether or not B- and L-conserving reactions are occurring rapidly enough ($\Gamma \gtrsim H$) to maintain partial equilibrium. However, thermal distributions are "most probable," and noninteracting massless particles in a thermal distribution remain in such a distribution with a temperature which is red-shifted, $T \propto R(t)^{-1}$, as the universe expands, as with photons and neutrinos today. To allow for our ignorance, we use a variety of particle energy distributions which allow us to solve ordinary differential equations. We find our results, except in one instance which we point out, are relatively insensitive to the form of the distribution.

(5) Finally, there is the question whether or not non-GUT processes, such as those operating at the quantum gravity epoch or the evaporation of primordial black holes, also contribute baryon asymmetries. We address this by considering cases in which the initial asymmetry may not be zero. This question is discussed in more detail by Carr and Turner.[17]

C. Out-of-equilibrium decay scenario

Since we find the basic idea of Weinberg and Wilczek to be qualitatively correct, we will briefly outline it as a framework against which to discuss our results. In this scenario, at the Planck time (5×10^{-44} sec) the universe is a hot soup of all the fundamental particles in thermal equilibrium. These particles include quarks, leptons, gauge and Higgs particles; in particular, the numbers of baryons and antibaryons are equal. The key particles are the superheavy bosons which mediate baryon nonconservation (either gauge or Higgs bosons) and are designated S. They have coupling strength α; for a gauge boson $\alpha \cong 10^{-2}$; for a Higgs boson, $\alpha \cong 10^{-4}-10^{-6}$. These bosons have decay channels of different total baryon number. We will take $\hbar = c = k_B = 1$.

Nothing of importance happens until the temperature of the universe drops to $T \lesssim M_S$. At this point, S's and $\bar{S}$'s, which were as abundant as photons and other relativistic particles, must diminish rapidly in number to maintain a thermal distribution, since for highly relativistic particles the number density is $n \sim T^3$ while for nonrelativistic particles $n \sim (MT)^{3/2} e^{-M/T}$. This can happen only if the decay rate $\Gamma_D \sim \alpha M_S$ is faster than the expansion rate $H \sim (G\rho)^{1/2} \sim g^{1/2} T^2/m_P$ [g is the effective number of particles $g \cong 10^2$; $m_P =$ Planck mass $\cong 10^{19}$ GeV (see Sec. II)]. (Decays are the dominant means of reducing the number of S, S's since the annihilation rate is of order α^2.) If $\Gamma_D > H$ when $T \cong M_S$, then $S, \bar{S}$'s can decay fast enough to maintain thermal equilibrium. This is the case if $M_S \lesssim 10^{18}\alpha$ GeV. No departure from equilibrium occurs and no baryon asymmetry is generated.

On the other hand, if $\Gamma_D < H$ when $T \cong M_S$ (that is, $M_S \gtrsim 10^{18}\alpha\,\text{GeV}$), the lifetime of the S is greater than the age of the universe, and decays are not yet happening on the cosmological time scale. S and $\bar{S}$ bosons remain as abundant as photons until the universe becomes as old as the S lifetime. At this point $S, \bar{S}$'s decay freely (inverse decays are blocked since typical particles have energies $\sim T \ll M_S$). The mean net baryon number of their products need not be equal and opposite, and $\epsilon \sim$ (size of CP violation) will parametrize the net baryon number generated when an S-$\bar{S}$ pair decays.

After these decays baryon number is effectively conserved until today, as rates of baryon nonconserving processes are very small, so this baryon excess remains constant. Its value is

$$n_B = n_b - n_{\bar{b}} \sim \epsilon n_X (\text{before decay}) \sim \epsilon n_\gamma \,. \qquad (1.1)$$

In an isentropic expansion, the total entropy also remains constant; the specific entropy density is the number of relativistic species times the density of a relativistic species $s/k \sim g n_\gamma$. The baryon-to-specific-entropy ratio remains constant, $kn_B/s \sim \epsilon/g \sim 10^{-2}\epsilon$. After antibaryons and most of the baryons have annihilated, this asymmetry is the total number of residual baryons, and this should be the baryon to entropy ratio measured today $kn_B/s(obs) = 10^{-9.8\pm1.6}$.

D. Summary of results

Both as a preamble and for those who wish to skip the detailed discussion in later sections, we offer here a summary of our results and an outline of the rest of the paper. In our numerical calculations we have considered cases both without and with C and CP violations in the decays and inverse decays of the superheavy gauge bosons. In this first case (C or CP conserved) no asymmetry can evolve and we follow the evolution of a universe which has an initial baryon asymmetry at the Planck time. In the second case (C and CP violated), an asymmetry can evolve and we examine situations with both zero and nonzero initial asymmetries.

As mentioned earlier, we have included decays (D) and inverse decays (ID) of the superheavy gauge bosons as well as baryon-nonconserving (BNC) collisions mediated by the superheavy gauge bosons. In general our results depend on the two parameters α and M_X, the coupling and mass of the gauge particles. Since we find that the "effectiveness" (interaction rate/expansion rate) is proportional to $K \equiv 2.9\times10^{17}$ α GeV$/M_X$ for D and ID and to αK for BNC processes, we use the equivalent and more useful parameters K and α.

1. C or CP conserved

If C or CP is not violated, no baryon asymmetry can be generated, but any initial asymmetry can be damped. In SU(5), $B-L=$ (baryon number $-$ lepton number) is conserved. Therefore, unless the universe starts with $B-L=0$ it can never evolve to a state with $B=L=0$ by GUT processes. In this way, other processes such as quantum gravity can leave a residual asymmetry which cannot be erased.

If the universe begins with $B-L=0$, an initial baryon asymmetry (hence also a lepton asymmetry) can be damped by D and ID, and also BNC. The damping for fixed K is insensitive to α: BNC processes are not important. The amount of damping is approximately $\exp(-5.5K)$ when we assume partial equilibrium. If this assumption is relaxed, the damping becomes approximately $\exp(-8.3K)$ because the presence of more quarks and leptons (due to X decays) increases the rate of inverse decays. This is the only result which depends at all significantly on the assumption of partial equilibrium. D and ID neutralize baryon asymmetries by a two-step process: $q+q\rightarrow\bar{X};\bar{X}\rightarrow\bar{q}+\bar{l}$ ($\Delta B=-1$, $\Delta L=-1$; $q=$quark, $l=$lepton). If the universe begins with $B-L\neq0$, then there is always an asymmetry which cannot be damped out. In this case, all that can happen is a redistribution of the asymmetry among all species. When we start the universe with $B=0$ but $B-L\neq0$ (only an initial lepton asymmetry) for $M_X \ll 2\times10^{18}$ α GeV the initial asymmetry is redistributed to make $|B|\cong|L|$. For $M_X\gg 2\times10^{18}$ α GeV, little redistribution takes place.

2. C and CP violated

In this case an initially symmetrical universe can dynamically evolve a baryon asymmetry, as outlined in (C) above. Here we observe a very gradual transition between the two limiting cases described there. For $M_X \gg M_C = 4.6\times10^{18}$ α GeV, an initially symmetric universe develops a maximum baryon-to-entropy ratio of $kn_B/s=7.8\times10^{-3}\epsilon$. We note that the size of the asymmetry and the critical mass are nearly but not exactly those estimated above. For $M_X\ll M_C$ the baryon-to-entropy ratio is smaller by a factor $(M_X/M_C)^{1.3}$, with the whole regime described reasonably well by

$$kn_B/s \cong 7.8\times10^{-3}\epsilon/[1+(M_C/M_X)^{1.3}] \cong 7.8\times10^{-3}\epsilon/[1+(16K)^{1.3}]\,. \tag{1.2}$$

This is insensitive to α for fixed K, and is also insensitive (varies by $\lesssim 50\%$) to whether or not the light particles (quarks and leptons) are in partial equilibrium.

If the universe starts with a large asymmetry, much larger than that which would be generated, for $M_X\gg M_C$ it is slightly damped as in the case of no CP violation, while for $M_X\ll M_C$ the final asymmetry approaches that which would have developed for zero initial asymmetry; here, unless $B-L\neq0$ initially, we are unable to tell what happened before.

Popular values for the superheavy masses are around 3×10^{14} GeV.[18] If a gauge boson of this mass is to produce the observed $kn_B/s\cong10^{-9.8\pm1.6}$, then we must have $\epsilon\cong10^{-4.3\pm1.6}$. If the CP violation in the X-X system is due to single-loop exchanges of superheavy Higgs bosons, then we would expect $\epsilon\cong\alpha_{\rm Higgs}=10^{-4}-10^{-6}$, which agrees with the above requirement.

In the rest of this paper we will develop in full glory the calculations which lead to those results. In Sec. II we will discuss kinetic theory and the analog of the Boltzmann equation in the expanding universe. In Sec. III we will review our assumptions and derive the equations governing the development of each species. In Sec. IV we present in detail our results when C or CP is conserved and in Sec. V when they are violated. In Sec. VI we assess our results and indicate further work that remains to be done. A discussion of SU(5), the various matrix elements needed for this calculation, and the full set of Boltzmann equations are contained in the Appendices.

II. KINETIC THEORY IN THE EXPANDING UNIVERSE

A. The Friedmann-Robertson-Walker universe

We perform our calculations against the background of the simplest big-bang model, the Friedmann-Robertson-Walker (FRW) universe.

The metric for an isotropic and homogeneous universe can be written in the Robertson-Walker form

$$ds^2 = dt^2 - R^2(t)[dr^2/(1-kr^2) + r^2(d\theta^2 + \sin^2\theta\, d\psi^2)], \tag{2.1}$$

where t is proper time measured by a comoving observer, r, θ, and ψ are comoving spatial coordinates, $k=0$ or ± 1 is a measure of curvature, and $R(t)$ is the Robertson-Walker scale factor. We use units such that $\hbar = k_B = c = 1$.

The dynamical equations for this metric are[19]

$$(\dot{R}/R)^2 = 8\pi G\rho/3 - k/R^2 + \Lambda/3, \tag{2.2}$$

$$d(\rho R^3)/dt + p\,d(R^3)/dt = 0, \tag{2.3}$$

where Λ is the cosmological constant, ρ the total energy density, and p the (isotropic) pressure.

The fundamental constituents of the universe, quarks, leptons, and gauge and Higgs bosons, are pointlike and at very high energies should behave like ideal relativistic gases. In thermal equilibrium at temperature T the energy densities of relativistic species are

$$\rho_b = (g\pi^2/30)T^4, \quad \rho_f = \tfrac{7}{8}\rho_b, \tag{2.4}$$

where g is the degeneracy factor; b refers to bosons and f to fermions. Thus, the total energy density ρ in relativistic particles is given by $(g_*\pi^2/30)T^4$, where g_* is the effective total degeneracy,

$$g_* = \sum_b g_b + \tfrac{7}{8}\sum_f g_f, \tag{2.5}$$

where the sum $\sum_b$ is over all the relativistic bosons and $\sum_f$ is over all relativistic fermions. Note that g_* is temperature dependent, since as T falls below M_i, species i drops out of the sum. For the minimal SU(5) model, at very high energies ($E \gg 10^{15}$ GeV) when all species are relativistic $g_* \cong 160$.

At very early times the terms k/R^2 and $\Lambda/3$ in (2.2) are negligible, so that the age and temperature of the universe are related by

$$t = (45/16\pi^3)^{1/2} g_*^{-1/2} m_P/T^2 = 0.154(160/g_*)^{1/2} T_P/T^2. \tag{2.6}$$

The quantity m_P is the Planck mass ($\equiv G^{-1/2} = 1.22 \times 10^{19}$ GeV). $T_P = 1.88\times 10^{18}$ GeV is the temperature of the universe at the Planck time (0.538×10^{-43} sec) for $g_* = 160$.

The specific entropy $s/k_B = s$ is also dominated by relativistic species, for which $s = (\rho + p)/T = \frac{4}{3}\rho/T$. Thus we find that at early times the entropy density is

$$s = g_*(2\pi^2/45)T^3. \tag{2.7}$$

When baryon-violating interactions cease, the net density of baryons $n_B = n_b - n_{\bar{b}}$ will also behave as $n_B \propto R^{-3} \propto T^3$ simply because of expansion, so that the ratio n_B/s remains constant.

Today the known relativistic particle species present in the universe are γ, ν_e, ν_μ, and (presumably) ν_τ. The temperature of the photons is $T_\gamma \cong 2.9$ K. The neutrino species are at a lower temperature, $T_\nu = (\frac{4}{11})^{1/3} T_\gamma$, since photons were heated by e^+e^- annihilations ($T \sim \frac{1}{2}$ MeV) after neutrinos decoupled ($T \sim 1$ MeV). Therefore, the entropy density today is

$$s = (2\pi^2/45)[2 + 6(\tfrac{7}{8})(\tfrac{4}{11})]T_\gamma^3, \tag{2.8}$$

where the first term in brackets comes from photons ($g_\gamma = 2$) and the second from neutrinos [$g_\nu = 6$, $\frac{4}{11} = (T_\nu/T_\gamma)^3$, and the factor $\frac{7}{8}$ for fermions].

The present baryon mass density ρ_b is conveniently expressed in terms of the parameters h [$H_0 \equiv 100h$ (km/sec)/Mpc] and $\Omega = \rho_b/\rho_c$ ($\rho_c = 3H_0^2/8\pi G$ = density necessary to close the universe for $\Lambda = 0$) (provided this mass is in baryons),

$$\rho_b = \Omega h^2 1.88\times 10^{-29}\ \mathrm{g/cm^3}, \tag{2.9}$$

or the baryon number density as

$$n_b = \rho_b/m = \Omega h^2 1.13\times 10^{-5}\ \mathrm{cm^{-3}} \tag{2.10}$$

(m = proton mass). This is also the net baryon number density n_B. Our knowledge of h and Ω is poor; we take as extreme limits $\frac{1}{2} < h < 1$ and $0.005 < \Omega < 2$. (The upper limit $\Omega = 2$ corresponds to a deceleration parameter $q_0 = 1$, while galactic masses determined from rotation curves indicate $\Omega > 0.005$.) With these ranges, we find that the baryon-to-entropy ratio today is

$$kn_B/s = 3.25\times 10^{-9}\Omega h^2(2.9/T_\gamma)^3 = 10^{-9.8\pm 1.6}. \tag{2.11}$$

This ratio has remained constant since baryon-nonconserving interactions became ineffective ($T \sim 10^{14}$ GeV).

Today it is almost equivalent to speak of the baryon-per-photon ratio. Essentially all the photons in the universe are in the 3-K background; the number density of photons is

$$n_\gamma = 2\zeta(3)/\pi^2 T_\gamma^3 \tag{2.12}$$

[$\zeta(3) \cong 1.20206$]. Comparing (2.8) and (2.12), we find $s \cong 7.02 n_\gamma$, so today

$$n_B/n_\gamma = 7.02\ kn_B/s = 10^{-8.9\pm 1.6}. \tag{2.13}$$

If one uses big-bang nucleosynthesis to place an upper limit on Ω_b (the contribution of baryons alone) more restrictive limits are obtained: $kn_B/s = 10^{-10.8\pm 0.9}$ and $n_B/n_\gamma = 10^{-9.9\pm 0.9}$ (Ref. 20).

B. Kinetic theory

The fundamental equations needed for the evolution of phase-space densities of the various species in the universe are the general-relativistic Boltzmann equations. Ehlers[21] has written an excellent review on this subject, and Wagoner[22] has also discussed these equations. We shall follow Wagoner's notation.

In order for the results of relativistic kinetic theory to be valid, two conditions must be met: first, the contituents must behave similar to ideal gases, and second, the spacetime must be locally flat. The first condition is satisfied if the average potential energy between particles is much less than their average kinetic energy $\sim T$. In GUT's, the fundamental particles are pointlike, and at high energies their interactions are weak enough so that they can be treated as perturbations to free-particle states. At times much later than the Planck time ($\sim 10^{-43}$ sec), the second condition will also be satisfied, since local curvature fluctuations will be negligible.

The fundamental dynamical quantities are the phase-space densities $N_i(p^\mu, x^\mu)$, where i runs over particle species. The number of particles in phase-space element $dV\,d\Pi_i$ is

$$dN_i = u_\alpha p_i^\alpha N_i(p_i^\mu, x^\mu)\,dV\,d\Pi_i\,, \tag{2.14}$$

where u^α is the four-velocity of the observer, dV his element of proper volume, and $d\Pi_i$ is the frame-invariant momentum element

$$d\Pi_i = (2\pi)^{-3} g_i d^3p_i/p_i^0\,. \tag{2.15}$$

Here, g_i is the degeneracy factor for species i, and the p_i^μ are components of physical momentum. The number density of species i per proper volume as seen by an observer with four-velocity $u^\alpha = (1, 0, 0, 0)$ (comoving) is

$$n_i = \int u_\alpha p_i^\alpha N_i d\Pi_i = g_i/(2\pi)^3 \int N_i d^3p_i\,. \tag{2.16}$$

If it were not for interactions, the phase-space density N_i of each species would remain constant along the phase-space trajectory of the system (Liouville's theorem). The change in N_i along the phase-space path is due solely to interactions, and is given by

$$L_i N_i(p_i^\mu, x^\mu) = \sum_{\substack{j,\ldots\\ l,m,\ldots}} \int d\Pi_j \cdots d\Pi_l d\Pi_m \cdots [N_l N_m \cdots (1 \pm N_i)(1 \pm N_j) \cdots W(p_l p_m \cdots \to p_i p_j \cdots) - N_i N_j \cdots (1 \pm N_l)(1 \pm N_m) \cdots W(p_i p_j \cdots \to p_l p_m \cdots)]\,, \tag{2.17}$$

where L_i is the Liouville operator for species i ($L_i = m_i D/D\tau$ for a massive particle). The factors $(1+N)$ are stimulated emission factors for bosons and $(1-N)$ exclusion factors for fermions. The invariant transition rate W is

$$W = (s/2^n)\,|\mathcal{M}|^2 (2\pi)^4 \delta^{(4)}\Big(\sum_{\text{in}} p_i^\mu - \sum_{\text{out}} p_j^\mu\Big)\,, \tag{2.18}$$

where s is a statistical term which contains a factor $(m!)^{-1}$ for each set of m identical incoming or outgoing particles, and n is the total number of incoming and outgoing particles. The invariant amplitude $\mathcal{M}$ is calculated according to the rules of Bjorken and Drell,[23] averaged over initial and final spins, colors, and generations, all of which are treated as degrees of freedom contributing to the total degeneracy g of each species.

The Liouville operator can be written as $L_i = p_i^\mu \partial/\partial x^\mu - \Gamma^\mu{}_{\alpha\beta} p_i^\alpha p_i^\beta \partial/\partial p_i^\mu$, so Eq. (2.17) becomes

$$\partial N_i/\partial t = (p_i^0)^{-1}[\Gamma^\mu{}_{\alpha\beta} p_i^\alpha p_i^\beta \partial N_i/\partial p_i^\mu - p_i^j \partial N_i/\partial x^j + \text{right-hand side of (2.17)}]\,, \tag{2.19}$$

where the index j is summed from 1 to 3.

In the case of an isotropic, homogeneous universe, $\partial/\partial x^j \equiv 0$ and $N(p^\mu) = N(|\vec{p}|)$, so that [recall $(p^0)^2 - |\vec{p}|^2 = m^2$; we write $p = |\vec{p}|$]

$$\partial N_i/\partial_t = (\dot{R}/R)p\,\partial N_i/\partial p + (1/p_i^0) \sum_{\substack{j,\ldots\\ l,m,\ldots}} \int d\Pi_j \cdots \int d\Pi_l d\Pi_m \cdots [N_l N_m \cdots (1 \pm N_i)(1 \pm N_j) \cdots W(p_l p_m \cdots \to p_i p_j \cdots) - N_i N_j \cdots (1 \pm N_l)(1 \pm N_m) \cdots W(p_i p_j \cdots \to p_l p_m \cdots)]\,. \tag{2.20}$$

The first term on the right-hand side of (2.20) represents the effects of the expanding universe, which will appear as red-shifts and dilutions ($n \propto R^{-3}$) of noninteracting particles, while the second term includes the effect of interactions. Equation (2.20) is the generalization of the Boltzmann equation for the expanding FRW universe.

In the absence of interactions, the distribution

$N=[\exp(p/T)\pm 1]^{-1}$, with $T\propto R^{-1}$ is a solution to (2.20) which corresponds to a gas of massless fermions (+) or bosons (−), which remains a thermal distribution with a red-shifted temperature. This is the condition of photons and neutrinos today.

Again in the absence of interactions, the distribution $N=\exp(-p^2/2mT)$, with $T\propto R^{-2}$, is a solution. It corresponds to a gas of massive, nonrelativistic particles whose kinetic energies have a Maxwell-Boltzmann distribution and whose temperature is red-shifted $T\propto R^{-2}$.

Integrating these distributions over momenta, one finds in both cases that the spatial number densities behave as $n\propto R^{-3}$, as expected since the total volume occupied by particles grows as $V \propto R^3$.

C. Equilibrium

In complete thermal equilibrium, a system appears stationary: all macroscopic properties, such as particle distributions, assume their most probable values. Since the universe is expanding, thermal equilibrium in the usual sense cannot exist. However, if particle-interaction rates are rapid compared to the expansion [i.e., the cosmological term in (2.20) is small compared to the other terms on the right-hand side], then equilibrium-particle distributions will be established on the expansion time scale.

Therefore, on time scales short compared to the expansion, but long compared to reaction rates, the universe will take on the appearance of a system in thermal equilibrium with particle distributions being given by their usual thermal forms—$\{\exp[(E+\mu)/T]\pm 1\}^{-1}$, etc.—with the cosmological temperature T decreasing $T\propto R^{-1}$. As long as the reaction rates which govern the distribution of a given species are rapid, that species will pass through a succession of equilibriumlike states. We shall take this as our definition of thermal equilibrium in the expanding universe.

III. EQUATIONS FOR THE EVOLUTION OF BARYON NUMBER

A. Assumptions

In principle, to follow the evolution of the baryon asymmetry we need to solve the coupled set of Boltzmann integrodifferential equations (2.20) for the distributions of all species of quarks, leptons, and all gauge and Higgs bosons, including all their interactions. This is clearly a formidable project. Therefore, we have made two key assumptions and a number of approximations along the way. The major assumptions involve which particle species are in fact important and the nature of the energy distributions; they allow us to restrict our attention to a manageable set of ordinary differential equations.

1. Species of interest

As a starting point we will assume that the world is described by the minimal SU(5) model: three generations of quarks and leptons, 24 gauge particles, and 34 Higgs particles (a complex $\underline{5}$ and a $\underline{24}$). Although in this minimal model it appears that the size of the CP violation is far too small due to cancellations at lower orders of perturbation theory,[11] the effect of enlarging the Higgs sector to avoid these cancellations is only to increase the total number of particle species, which in turn affects the expansion rate $H\propto g_*^{1/2}$, and the entropy $s\propto g_*$. In the minimal model, $g_*\cong 160$, and we display the dependence of our results on $(g_*/160)$ where appropriate.

In SU(5) there are two types of bosons that mediate B- and L-violating interactions—an isospin doublet, color triplet of gauge particles (X, Y) and a isospin singlet, color triplet of superheavy Higgs scalars (H). The X has charge $\pm\frac{4}{3}$, the Y has charge $\pm\frac{1}{3}$, and the H also has charge $\pm\frac{1}{3}$. In the most general GUT which breaks down to $SU(3)_C\times SU(2)_L\times U(1)$ there is only one additional type of boson that mediates B and L violations—an isospin doublet, color triplet with charge $\pm\frac{2}{3}$, $\mp\frac{1}{3}$ $(X'Y')$.[11] These bosons are also vector particles.

In this paper we will only consider the effect of the X and Y gauge bosons on the evolution of the baryon asymmetry. Because the X' and Y' bosons couple similarly we will assume that these results also describe the evolution of kn_B/s due to a single set of $X'Y'$ bosons. In a companion paper[15] we treat the effect of a single species of Higgs bosons (H), and in a third paper[16] in this series we consider the effect of many superheavy species. Therefore, of all the species present, we will follow only quarks, leptons, and the superheavy gauge particles X and Y. The other species will be assumed to be present in thermal numbers, boson species contributing $\rho_i=(g_i\pi^2/30)T^4$ to the total density of the universe and fermion species $\frac{7}{8}$ of this amount. All particles except X and Y will be treated as massless, which is a very good approximation at the temperatures of interest $T \gtrsim 10^{14}$ GeV.

We include the three generations of quarks and leptons that have been at least partially observed. Since SU(5) is symmetrical with respect to generation and colors, we treat generation and color as additional, spinlike degeneracies. Specifically, all up-like quarks (u, c, and t) are referred to as U, with an effective degeneracy factor $g_U=(2$ spin states) × (3 generations) × (3 colors) = 18; all downlike quarks (d, s, and b) are designated D, and D also has an effective degeneracy $g_D = 18$. All electronlike leptons (e, μ, and τ) are labeled by L, where L has an effective degeneracy $g_L = (2$ spin states) × (3 generations) = 6, and finally, all neutrinos (ν_e, ν_μ, and ν_τ) are labeled by ν, where ν has an effective degeneracy $g_\nu = (1$ helicity state) × (3 generations) = 3.

The X and Y superheavy gauge bosons, labeled X and Y, each have an effective degeneracy $g_X = g_Y = (3 \text{ spins}) \times (3 \text{ colors}) = 9$. The X and Y masses are assumed to be equal and each is denoted M_X. Any difference is due to the SU(2) × U(1) weak symmetry breaking, and is of order 100 GeV, negligible at temperatures near 10^{14} GeV. Strictly speaking, the X and Y bosons have masses and three polarizations only when the SU(5) symmetry is broken. When the full symmetry is realized, $T \gtrsim 100 M_X$ we should have instead massless gauge bosons, with two spin states, and additional Higgs particles. However, for $T \gtrsim 100 M_X$ all particles are effectively massless anyway and the processes of interest for baryon generation are occurring at $T \lesssim M_X$. For these reasons, we treat X and Y always as massive particles.

2. *Particle energy distributions*

In order to reduce the Boltzmann equations from integrodifferential equations in phase space to ordinary differential equations in time only, we take the following model for the distributions:

$$N_F(p,t) = F(t)/[\exp(p/T)+1],$$
$$N_B(p,t) = B(t) f(p/T), \tag{3.1}$$

where $F = U$, D, L and ν, and $B = X$ and Y. T is the universal temperature and $p = |\vec{p}|$; $f(p, T)$ is the momentum distribution of the X's and Y's, and will be discussed later. From (3.1) we can find the number density of one of the fermions, say U, $n_U = \int d\Pi_U p^0 N_U = g_U A U(t)$, where $A = [3\zeta(3)/4\pi^2]T^3$.

Besides reducing our problem to a system of ordinary differential equations, this model also automatically cancels the $\partial N/\partial p$ terms which represent the effect of cosmological expansion;

$$[\partial/\partial t - (\dot{R}/R) p \partial/\partial p] N_F(p,t) = \dot{F}(t)/[\exp(p/T)+1]. \tag{3.2}$$

When each of Eqs. (2.20) for species i is integrated over $\int d\Pi_i p_i^0$, the result is a set of 12 coupled ordinary differential equations for the functions $U(t)$, $D(t)$, . . . , $\bar{U}(t)$, $\bar{D}(t)$, . . . ; for example,

$$\dot{U}(t) = (g_U A)^{-1} \sum_{\substack{j,\ldots \\ l,m,\ldots}} \int d\Pi_U d\Pi_j \cdots d\Pi_l d\Pi_m \cdots [N_l N_m \cdots W(p_l p_m \cdots \to p_U p_j \cdots) - N_U N_j \cdots W(p_U p_j \cdots \to p_l p_m \cdots)], \tag{3.3}$$

where in addition, we have neglected the degeneracy factors (stimulated emission and exclusion) in (2.20). None of the species should be highly degenerate, and the errors caused by this approximation are probably about 30%, certainly consistent with the degree of approximation throughout. The integrals on the right-hand side can be performed so that $\dot{U}$ can be expressed in terms of $U(t)$, $D(t)$, . . . , and corresponding barred quantities, with coefficients having known temperature dependences. Equations (2.6) relate time and temperature.

If all processes were happening rapidly on the expansion time scale, B and L violating included, thermal equilibrium in the sense described in Sec. II would be established, and we would have $U = D = L = \nu = \bar{U} = \bar{D} = \bar{L} = \bar{\nu} \equiv 1$. Because of the B-violating processes, all species would have zero chemical potential (if we assume total $B - L = 0$ for the universe). If all processes except those violating B and L were occurring rapidly, "partial" thermal equilibrium would be established, and quark and lepton distributions should be given by $\{\exp[(\mu+p)/T]+1\}^{-1}$. Particle and antiparticle chemical potentials would be related by $\mu + \bar{\mu} = 0$, since by assumption reactions such as $u + \bar{u} \rightleftharpoons 2\gamma$ would be occurring rapidly. If the asymmetries are small, $|\mu/T| \ll 1$, then $\{\exp[(\pm\mu+p)/T]+1\}^{-1} \cong (1 \mp \delta)[\exp(p/T)+1)^{-1}$; in our parametrization, $U \cong 1+\delta$; $\bar{U} \cong 1-\delta$; $U + \bar{U} = 2$, etc. For small asymmetries our parametrization is equivalent to using a small chemical potential, and greatly simplifies the Boltzmann equations.

3. *Partial equilibrium*

It is not clear whether the B- and L-conserving reactions occur rapidly enough to ensure $U + \bar{U} = 2$, etc. To determine this explicitly all B- and L-conserving processes should be included, however, this would greatly complicate the problem. We can estimate the rates of these reactions: at very high T (for us $T \gtrsim 10^{14}$ GeV), any two-body process $ij \to lm$, which transfers momentum $\Delta q \sim T$, should have a cross section $\sigma \sim \alpha^2/T^2$. Since number densities are $n_i \sim (g_i/\pi^2)T^3$, the total rate of such scatterings ought to be $\Gamma = n\sigma v \sim (g_*/\pi^2)\alpha^2 T$. These reactions will be effective at maintaining equilibrium if $\Gamma \gtrsim H \sim 6T^2/T_P$, or $T \lesssim 3\times 10^{15}$ GeV. More careful consideration of this question by Ellis and Steigman[24] leads to the same result. This temperature, above which even partial equilibrium cannot be maintained, is very close to the temperatures of interest for baryon generation, $T \cong M_X \cong 10^{15}$ GeV.

If these B- and L-conserving interactions are not effective, there is no restriction $\mu + \overline{\mu} = 0$, or equivalently $U + \overline{U} = 2$, and in fact the quantities $U+\overline{U}$, $D+\overline{D}$, $L+\overline{L}$, $\nu+\overline{\nu}$ become greater than 2 from the decays of X and Y bosons into quarks and leptons which are not subsequently redistributed by B-conserving reactions such as $u + \overline{u} \rightleftharpoons 2\gamma$, etc.

Rather than including all these interactions we consider two limiting cases. As we shall discuss later, our results vary little between these two cases, with one exception. In one extreme, we suppose that partial equilibrium is maintained exactly and we set $U+\overline{U} = D+\overline{D} = L+\overline{L} = \nu+\overline{\nu} = 2$. In the other, we suppose that B- and L-conserving interactions are completely ineffective, and we allow $U+\overline{U}$, etc., to evolve as X and Y bosons decay according to the B-violating processes we include, with no B-conserving redistribution. In this case, after all decays, the quantities $U+\overline{U}$, etc., have increased from 2 to about 4. Here, we implicitly assume that the X and Y decays do not affect the shape of the U, D, L, ν distributions but just roughly double the total number of quarks and leptons.

4. The X and Y distributions

For a baryon asymmetry to develop, the superheavy gauge bosons must at some point be out of equilibrium, i.e., when T falls below M_X, X and Y bosons are present in larger numbers than they would have if they were in thermal equilibrium. Thus, we do not want to extend partial equilibrium to include X's and Y's. However, in this case they are taken to be in kinetic equilibrium, with energies, but not numbers, distributed thermally. This can be modeled as

$$N_B(p,t) = B(t)a(T)/\{\exp[(p^2+m^2)^{1/2}/T]-1\}, \qquad (3.4)$$

where $a(T)$ is chosen so that $n_X = X(t)[g_X\zeta(3)/\pi^2]T^3$; that is, $X(t)$ and $Y(t)$ reflect the abundances of these particles relative to the number that would be present if they were highly relativistic (massless).

In the other case, if Compton scattering and other thermalizing reactions are not occurring rapidly, kinetic equilibrium with the rest of the universe will not be maintained. Recall from Sec. II that the temperature of a nonrelativistic noninteracting gas falls more rapidly, $T \propto R^{-2}$ instead of $T \propto R^{-1}$. To model this, we choose

$$N_B(p,t) = B(t)/[\exp(|\vec{p}|/T)-1]. \qquad (3.5)$$

This distribution, $[\exp(p/T)-1]^{-1}$, is also a solution of the Boltzmann equation without interactions. At very high temperatures, $T \gg M$, it corresponds to a relativistic gas with $T \propto R^{-1}$, while at low temperatures $T \sim \mathrm{KE} = p^2/2M \propto R^{-2}$.

These two distributions (3.4) and (3.5) serve to represent the two limiting cases of kinetic equilibrium and free expansion of the X and Y bosons. It will turn out that our results are insensitive to the choice of the X and Y distributions.

B. The interactions

With the simplifying assumptions discussed above we are ready to discuss the interactions we include. These are decays (D) and inverse decays (ID) of X and Y bosons, baryon-nonconserving (BNC) collisions which involve two incoming and two outgoing fermions, and annihilationlike and Compton-type BNC processes involving two bosons and two fermions. The decay channels for the X are $X \to \overline{U}\overline{U}$, $X \to DL$; the decay channels for the Y are $Y \to \overline{U}\overline{D}$, $Y \to UL$, and $Y \to D\nu$. All these processes are of order α. All the BNC fermion collisions are contained in the processes $UU \to \overline{D}\overline{L}$, $UD \to \overline{U}\overline{L}$, $UD \to \overline{D}\overline{\nu}$, and $U\nu \to \overline{D}\overline{D}$, along with their CP-conjugated and time-reversed counterparts. These processes, mediated by exchange of X and Y bosons, are $O(\alpha^2)$.

The annihilationlike and Compton-type BNC processes of order α^2 for $XY \to \overline{U}L$, $YY \to \overline{D}L$ or $\overline{U}\nu$, and $UX \to \overline{Y}L$, $UY \to \overline{X}L$ or $\overline{Y}\nu$, $DY \to \overline{Y}L$ together with their CP-conjugated and time-reversed counterparts. These are less important because of the smaller abundances of X's and Y's at times of interest; when included, their effects are less than 10% of the effect of BNC fermion collisions, which are themselves not very important. Since they are not important, are in principle no different from the BNC processes we do include, and lead to a proliferation of terms in the equations, they are omitted for clarity in this discussion (although they have been included in our computations).

1. Decays and inverse decays

The details of calculating the various matrix elements we need are contained in Appendix A. The total decay rate for an X or Y gauge boson at rest is just $\Gamma_D = \tau^{-1} = \alpha M_X$. The rate when not at rest has a time dilation factor, $\Gamma = \alpha M_X{}^2/E$, suppressing decays and inverse decays at high temperatures, $T \gg M_X$. The rates which enter the Boltzmann equation are integrals over the distributions of X and Y and their decay products.

2. Baryon- (lepton-) nonconserving fermion collisions

Again the matrix elements are calculated in Appendix A. At low temperatures, $T \lesssim M_X$, a typical thermal averaged cross section would be expected to be $\sigma \sim \alpha^2 T^2/M_X{}^4$, since at these energies the interactions mediated by X and Y bosons are effectively point interactions with a Fermi constant $\sim\alpha/M_X{}^2$. This is the behavior we find.

At high temperatures, $T \gg M_X$, one expects on dimensional grounds that a typical cross section should be $\sigma \sim \alpha^2/T^2$. Instead, we find at high temperatures $\sigma \sim \alpha^2/M_X{}^2$; rather than decreasing with temperature the cross section remains constant. The reason for this is simple: the total cross section becomes dominated by very soft t-channel exchanges.

There are at least two effects that should prevent this. The first and probably more important effect is Debye screening. In a hot plasma, the electromagnetic interaction has a finite range $R_D = (\alpha n/T)^{-1/2}/4\pi$ due to shielding effects,[25] where n is the density of screening electrons. Equivalently, the photon acquires an effective mass, $m \propto R_D{}^{-1}$. A similar screening effect should occur for interactions mediated by X and Y bosons, generating an effective mass $M_{\rm eff} \sim R_D{}^{-1} \cong 36\alpha^{1/2} T$ for these particles.

The second effect involves causality. In the early universe the horizon distance, or the distance over which a signal could have propagated since the big bang ($t=0$), is $R_H = 2t$. An interaction with $\sigma > \pi(2t)^2$ would appear to be noncausal. Stated differently, the range of any interaction should be $R \lesssim 2t$, or the effective mass should be $M_{\rm eff} \gtrsim (2t)^{-1} = 3.3T^2/T_p$.

To allow for these effects, we use at high temperatures an effective mass for the X and Y, $M_{\rm eff} = \max\{M_X, M_{\rm Debye}, M_{\rm horizon}\} = \max\{M_X, 36.1\,\alpha^{1/2}T, 3.3T^2/T_p\}$, so that at high temperatures cross sections $\sigma \sim \alpha^2/M_{\rm eff}{}^2$. Thus, at the highest temperatures, $T \gtrsim (11\alpha^{1/2})T_p$ (and only for small coupling constants) the range is limited by the horizon; for the intermediate range $M_X \lesssim T \lesssim (11\alpha^{1/2})T_p$, Debye screening limits the range of the interaction and gives the form $\sigma \sim T^{-2}$ expected on dimensional grounds.

In order to obtain the rates for BNC processes at all temperatures, we interpolate smoothly between high- and low-temperature limiting forms. The results are not sensitive to our assumptions about the high-temperature regime, since baryon generation does not occur until $T \lesssim M_X$, when horizon effects and Debye screening are not important.

3. The CP violation[26]

Even in the absence of C and CP invariance, CPT invariance and unitarity put strong constraints on the matrix elements. Unitarity of the S matrix, or conservation of probability, implies that $\underline{S}\underline{S}^\dagger = \underline{S}^\dagger \underline{S} = 1$. $\underline{S}$ is defined by writing the transition probability from incoming state i to outgoing state f as $|S_{fi}|^2$; CPT invariance requires $S_{fi} = S_{\bar{i}\bar{f}}$, where the bar indicates CP-conjugate states. From unitarity and CPT invariance we have

$$\sum_n |S_{in}|^2 = \sum_n |S_{ni}|^2 = \sum_n |S_{n\bar{i}}|^2 = \sum_n |S_{\bar{i}n}|^2 , \qquad (3.6)$$

where n runs over all physical states of the system. This will be discussed more later.

The T matrix is defined by $\underline{T} = i(\underline{1} - \underline{S})$, and the matrix elements and transition rates $W(j \to i)$ are proportional to $|T_{ij}|^2$. The relations in (3.6) also hold for $|T_{ij}|^2$ and thus for $W(j \to i)$; specifically, we have[27]

$$\sum_n W(i \to n) = \sum_n W(\bar{i} \to n) = \sum_n W(n \to i) = \sum_n W(n \to \bar{i}) . \qquad (3.7)$$

For simplicity we assume that CP is violated only for processes involving the X bosons. In the absence of definite models of CP violation involving these particles, we parametrize the CP violation as follows. When C or CP is conserved, $W \equiv W(X \to \overline{U}\overline{U}) \sim O(\alpha)$ describes all decay and inverse decay rates:

$$\begin{aligned} W(X \to \overline{U}\overline{U}) &= W(\overline{U}\overline{U} \to X) \\ &= W(\overline{X} \to UU) = W(UU \to \overline{X}) = W , \\ W(X \to DL) &= W(DL \to X) \\ &= W(\overline{X} \to \overline{D}\overline{L}) = W(\overline{D}\overline{L} \to \overline{X}) = 3W . \end{aligned} \qquad (3.8)$$

Without loss of generality when CP is violated we can choose $W(X \to \overline{U}\overline{U}) = W$, $W(X \to DL) = 3W$, $W(\overline{X} \to UU) = (1+\epsilon)W$, and $W(\overline{X} \to \overline{D}\overline{L}) = 3(1+\epsilon')W$. The unitarity constraint (3.7) then gives

$$g_U{}^2 W + 3g_D g_L W = g_U{}^2(1+\epsilon)W + 3g_D g_L(1+\epsilon')W . \tag{3.9}$$

Recall that matrix elements are averaged over all initial- and final-state degeneracies, so that factors for final-state degeneracies must be included in (3.9) when doing the sum over n. Since $g_U{}^2 = 3g_D g_L$, Eq. (3.9) requires $\epsilon = -\epsilon'$.

Combining this with CPT invariance, $W(i \to j) = W(\bar{j} \to \bar{i})$, we find

$$\begin{aligned} &W(X \to \overline{U}\overline{U}) = W(UU \to \overline{X}) = W ,\\ &W(\overline{X} \to UU) = W(\overline{U}\overline{U} \to X) = W(1+\epsilon) ,\\ &W(X \to DL) = W(\overline{D}\overline{L} \to \overline{X}) = 3W ,\\ &W(\overline{X} \to \overline{D}\overline{L}) = W(DL \to X) = 3W(1-\epsilon) . \end{aligned} \tag{3.10}$$

From these relations it follows that the mean net baryon number produced in X and $\overline{X}$ decays is

$$\begin{aligned} B_X &= [W(-\tfrac{2}{3}) + W(\tfrac{1}{3})]/(2W) = -\tfrac{1}{6} ,\\ B_{\overline{X}} &= [W(1+\epsilon)(\tfrac{2}{3}) + W(1-\epsilon)(-\tfrac{1}{3})]/(2W)\\ &= +\tfrac{1}{6} + \epsilon/2 \end{aligned} \tag{3.11}$$

so that the net baryon excess produced by an X-$\overline{X}$ pair is $\epsilon/2$. CPT invariance and unitarity require (Ref. 11) $\epsilon \lesssim O(\alpha)$. For the following discussion we assume $\epsilon \sim O(\alpha)$, although the reasoning can be generalized for $\epsilon < O(\alpha)$.

In addition to CP violations in the decays and inverse decays of X bosons, the CPT invariance and unitarity constraints require CP violations of the same order in BNC processes. For example, consider transitions from initial states UU or $\overline{U}\overline{U}$. Equation (3.7) requires

$$\sum_n W(UU \to n) = \sum_n W(\overline{U}\overline{U} \to n) . \tag{3.12}$$

If we include all processes up to $O(\alpha^2)$, then the equality is valid in the sense that any discrepancy is $O(\alpha^3)$. The states UU and $\overline{D}\overline{L}$ are certainly included in the sum over n on the left-hand side, but there is ambiguity concerning the state $\overline{X}$. In some sense, on time scales less than its lifetime, the $\overline{X}$ is a stable particle and represents a real physical state, but in other circumstances, on time scales greater than its lifetime, the $\overline{X}$ is merely a resonance and not a real physical state. Thus, there is an ambiguity in the treatment of the destruction of the state UU by inverse decay into an $\overline{X}$. This is resolved by self-consistency. If the $\overline{X}$ is considered to be a resonance, then UU destruction by inverse decay is included in the matrix elements for $UU \to UU$ and $UU \to \overline{D}\overline{L}$ when these s-channel processes are nearly on resonance, the virtual X nearly on mass shell. In this case the initial UU state "resonates" (forms an $\overline{X}$) as either a UU or $\overline{D}\overline{L}$ final state.

On the other hand, if the $\overline{X}$ is treated as a stable particle then UU destruction by inverse decay is included explicitly in the sum on the left-hand side of (3.12) in the rate $W(UU \to \overline{X})$. In this case, to avoid double counting, nearly on-shell processes $UU \to UU$ and $UU \to \overline{D}\overline{L}$ must not be included in the sum, as they also represent inverse decay which has already been taken into account. Similar reasoning holds for the right-hand side of (3.12).

For our purposes it is more convenient to think of X bosons as nearly stable physical particles; in writing the Boltzmann equations we treat them as a distinct particle species. Therefore, the unitarity relationship (3.12) becomes

$$\begin{aligned} &\int\int d\Pi_{\overline{D}} d\Pi_{\overline{L}} W_T(UU \to \overline{D}\overline{L}) + \int d\Pi_{\overline{X}} W(UU \to \overline{X})\\ &\quad = \int\int d\Pi_D d\Pi_L W_T(\overline{U}\overline{U} \to DL) + \int d\Pi_X W(\overline{U}\overline{U} \to X)\\ &\qquad + O(\alpha^3) . \end{aligned} \tag{3.13}$$

The terms $UU \to UU$ and $\overline{U}\overline{U} \to \overline{U}\overline{U}$ cancel by CPT invariance, and $\int d\Pi$ are implicit in the sum over states $\sum_n$. The W_T represent "true" scatterings, where processes with nearly on-shell virtual X's or $\overline{X}$'s are removed. Since $W(\overline{X} \to UU)$ and $W(X \to \overline{U}\overline{U})$ differ by a term of order α^2 for $\epsilon \sim O(\alpha)$, the true scattering rates also differ by $O(\alpha^2)$, which implies a CP violation of $O(\alpha^2)$ in scattering. In general, the violation is $O(\alpha\epsilon)$, with (3.13) valid up to terms $O(\alpha^2\epsilon)$.

In the interaction terms of the Boltzmann equations we need both the CP-conserving and CP-violating parts of BNC processes such as $UU \to \overline{D}\overline{L}$. The CP-conserving part can be calculated without regard to CP violations as detailed in Appendix A, being careful to exclude the contribution from s-channel processes with a nearly on-shell X. The CP-violating rates that are necessary can all be found from equations similar to (3.13). In terms of the transition rate W, the CP-violating part of $UU \to \overline{D}\overline{L}$ is

$$\begin{aligned} &\int\int d\Pi_D d\Pi_L [W_T(UU \to \overline{D}\overline{L}) - W_T(\overline{U}\overline{U} \to DL)]\\ &\quad = \int d\Pi_X [W(\overline{U}\overline{U} \to X) - W(UU \to \overline{X})] = \epsilon \int d\Pi_X W . \end{aligned} \tag{3.14}$$

C. The master equations—discussion

Rather than the 12 quantities $U, D, \ldots, \overline{U}, \overline{D}, \ldots$, it is more convenient to introduce $I_\pm = I \pm \overline{I}$ where $I = U, D, L, \nu, X$, and Y. A set of 12 coupled ordinary differential equations for these quantities in time (or equivalently, temperature) is obtained from (3.3) when the integrals over the appropriate rates W (calculated in Appendix A) have been performed. The equations are collected in Appendix B. We discuss here some of their properties.

1. The + equations

The six equations for + quantities, $U_+, D_+, \ldots,$ decouple from the other equations when terms second order in − quantities are ignored. The evolution of + quantities depends on (1) D and ID, and (2) whether or not we impose partial equilibrium for quarks and leptons and kinetic equilibrium for X and Y bosons. Although BNC processes can also redistribute the + quantities, their effect is minor compared to what all the B- and L-conserving processes can do, so we have not included BNC processes in the + equations. With regard to (2) when we do impose partial equilibrium we fix $U_+ = D_+ = L_+ = \nu_+ = 2$. The X and Y momentum distributions are $f_X(p, T) = a(T)\{\exp[(p^2 + m^2)^{1/2}/T] - 1\}^{-1}$, and only X_+ and Y_+ evolve with time due to D and ID. When we assume B- and L-conserving thermalizing reactions to be completely ineffective, all six + quantities evolve by D and ID. In this case, $f_X = [\exp(p/T) - 1]^{-1}$.

In both cases, the + equations are formally accurate through $O(\alpha)$, since higher-order processes have not been included. Given the uncertainties about partial and kinetic equilibrium, it seems reasonable not to worry about higher-order processes, but since the crucial feature for baryon generation is that X and Y bosons cannot diminish rapidly enough to remain in equilibrium, we did estimate the effect of annihilations in the establishment (or not) of equilibrium numbers of X and Y bosons.

There are numerous annihilationlike processes which can consume X bosons: $X\overline{X} \to U\overline{U}, D\overline{D}, L\overline{L}, \nu\overline{\nu}, \gamma\gamma, ZZ, gg, W^+W^-, \gamma Z, \gamma g, Zg$; $X\overline{Y} \to L\overline{\nu}, \overline{U}D$; $XY \to \overline{U}L$; $X\gamma \to DL, \overline{U}\overline{U}$; $Xg \to DL, \overline{U}\overline{U}$; $XZ \to DL, \overline{U}\overline{U}$; $XW \to D\nu, UL, \overline{U}\overline{D}$; $XU \to \overline{D}W, \overline{U}\gamma, \overline{U}g, \overline{U}Z$; $XD \to \overline{U}W$; $X\overline{U} \to WL$; $X\overline{D} \to \nu W, \gamma L, gL, ZL$; $X\overline{\nu} \to DW$ (g = gluon). These processes are all $O(\alpha^2)$, but their sheer number is alarming, in light of our inclusion of only D and ID.

We have carefully estimated the total effect of annihilations, and we find that the ratio of the annihilation rate Γ_A to the decay rate Γ_D is

$$\Gamma_A/\Gamma_D \sim (T/M_X)^3 r\alpha + (T/M_X)^5 r\,\alpha \tag{3.15}$$

for $T \lesssim M_X$, where $r \cong 30$. When $\alpha \lesssim 10^{-2}$, neglecting these processes should be a reasonable approximation. However, in SU(5), $\alpha \cong \frac{1}{45}$, and this omission is not as secure. There is also the temperature dependence, though. The equilibrium number density of a massive particle compared to that of a massless particle does not fall by a factor of 2 until $(T/M) \cong 0.5$. Therefore, by the time the X bosons must really begin to decrease rapidly in number, $T \lesssim 0.5 M_X$, Γ_A/Γ_D is already smaller than 5α, and continues to decrease as T falls. At most, annihilations might allow X's to remain in equilibrium a little longer than D and ID alone, but if a departure from equilibrium will occur when only D and ID are considered, it will also occur when annihilations are included, and the important effects occur later, when annihilations are negligible (see Sec. V).

2. The - equations

Explicitly, these equations will contain only B- and L-nonconserving processes, but since their rates depend on + quantities they implicitly depend on B- and L-conserving interactions as well. We assume that all − quantities are small and only terms linear in − quantities are retained. Since the observed asymmetry is $kn_B/s \cong 10^{-9.8}$, this is probably a good assumption.

As a typical − equation we display the equation for the evolution of U_-,

$$\begin{aligned} U_-'/(zK) = & -\gamma_D[2X_- + Y_-/2] - \gamma_{\rm ID}[(UL)_-/4 + (UD)_-/2 + 2U_+U_-] \\ & - [2s_1U_+U_- + (s_2 + s_3/4)(UD)_- + s_4D_+D_-/2 + s_2(UL)_- + s_1(DL)_- + s_4(U\nu)_-/4 + s_3(D\nu)_-/4] \\ & + \epsilon[\gamma_D X_+ - \gamma_{\rm ID} D_+L_+/2]\,. \end{aligned} \tag{3.16}$$

We use as the independent variable $z = M_X/T$; the prime denotes d/dz. The notation $(AB)_-$ stands for $(A_+B_- + B_+A_-)$. Here $K = (160/g_*)^{1/2} \times (2.9 \times 10^{17}\ \alpha\ \text{GeV}/M_X)$; $K \simeq$ (decay rate)/(expansion rate) when $T \simeq M_X$.

The first group of terms in brackets on the right-hand side of (3.16) represents the baryon generation by decays when there is an asymmetry between X-$\overline{X}$ or Y-$\overline{Y}$. The second group represents baryon production or damping by inverse decays. The L_+U_-, D_+U_-, and U_+U_- terms cause the U asymmetry to be damped, while the U_+L_- and U_+D_- terms can result in generation or damping, depending on the signs of L_- and D_-.

The third group of terms in brackets represents the effect of BNC collisions involving two incoming and outgoing fermions. The s_i are all $O(\alpha)$. The terms containing U_- all result in U damping, while the terms proportional to U_+ can damp or generate asymmetry depending on the signs of other asymmetries.

Finally, the fourth set of terms represents net U_- production due to CP violations in processes involving X bosons. The first term comes directly from X-$\overline{X}$ decays, which produce a baryon asymmetry of $\epsilon/2$ for each pair which decays. The second term comes from the CP-violating part of BNC processes which is required by unitarity. Because of the relation (3.14) dictated by unitarity, the coefficient of D_+L_+ is $\gamma_{\rm ID}$ rather than one of s_i.

When equilibrium is maintained, all reaction rates large compared to the expansion rate, we have $U_+=D_+=L_+=\nu_+=2$ and $X_+=Y_+=2\gamma_{\rm ID}/\gamma_{\rm D}$ (obtained from the X_+, Y_+ equations by setting derivatives equal to zero). Then all the CP-violating terms exactly cancel. In equilibrium if the asymmetries are zero they remain so; this cancellation occurs in all $-$ equations. This is a manifestation of the theorem mentioned earlier; even in the presence of B, C, and CP violations, a baryon symmetrical universe will remain so as long as thermal equilibrium is maintained. The importance of unitarity is clear, since it is unitarity which dictates the form of the CP-violating terms.

D. Notes on the numerical integration

1. The integration scheme

These equations require small time steps for a few ranges of $z=M_X/T$ where rates are large, while for efficiency the steps should be larger where rates are small. For choosing time steps essentially independently, Runge-Kutta schemes are appropriate [Ref. 28, Eqs. (25.5.7) and (25.5.8)]. The third-order scheme was used, with a fixed maximum time step which was reduced if necessary to ensure that either $|\Delta y/y|=(y_{n+1}-y_n)/y_n<10^{-1.5}$ or $|\Delta y^{(2)}-\Delta y^{(3)}|/\Delta y^{(2)}=(k_1-2k_2+k_3)/6k_2<10^{-3}$.

2. Initial conditions

All numerical integrations begin at the Planck time $t=0.538\times10^{-43}$ sec, when the temperature of the universe was $T_P=1.88\times10^{18}$ GeV. The quantities U_+, D_+, L_+, and ν_+ are given the values that they would have in thermal equilibrium, 2. The quantities X_+ and Y_+ are also set to their equilibrium values $X_+=Y_+=2\gamma_{\rm ID}/\gamma_{\rm D}\simeq2$; for $T\gg M_X$, X and Y bosons are as abundant as any other ultrarelativistic species. Although no interactions are occurring rapidly enough to establish equilibrium and the number of particles within the horizon is a few, we assume that somehow these conditions are created. It may be that quantum-gravitational processes occurring before the Planck time established thermal distributions. Hartle and Hu[29] have shown that the damping of anisotropies during the pre-Planck epoch through particle creation by the gravitational field leads to nearly thermal distributions of particles. Even in the absence of anisotropies they find that when the so-called trace anomalies are taken into account, the resulting metric is "horizonless," so that there is no problem of causality, and of all distributions thermal distributions are in some sense most probable.

The quantities X_- and Y_- are always taken to be initially zero. In some instances we begin our calculations with initial baryon and/or lepton asymmetries. The baryon and lepton asymmetries are given by

$$\begin{aligned} n_B(T)&=\tfrac{1}{3}\int p^0 d\Pi(N_U+N_D-N_{\overline{U}}-N_{\overline{D}})\\ &=[g_U\zeta(3)/4\pi^2](U_-+D_-)T^3\\ &=0.55(U_-+D_-)T^3\,,\\ n_L(T)&=\int p^0 d\Pi(N_L+N_\nu-N_{\overline{L}}-N_{\overline{\nu}})\\ &=[3\zeta(3)/4\pi^2](g_LL_-+g_\nu\nu_-)T^3\\ &=0.55(L_-+\nu_-/2)\,. \end{aligned}\tag{3.17}$$

The specific entropy is $s/k=(2\pi^2g_*/45)T^3$ [Eq. (2.7)], so the baryon- and lepton-to-entropy ratios are

$$\begin{aligned} kn_B/s&=(7.8\times10^{-3})(160/g_*)(U_-+D_-)\,,\\ kn_L/s&=(7.8\times10^{-3})(160/g_*)(L_-+\nu_-/2)\,. \end{aligned}\tag{3.18}$$

IV. C OR CP NOT VIOLATED ($\epsilon=0$)

If the decays of the superheavy bosons do not violate both C and CP, in our model $\epsilon=0$ and no asymmetries develop in an initially particle-antiparticle symmetrical universe. This is easy to see: all the $-$ equations are homogeneous except for sources which are proportional to ϵ. However, if the universe is initially asymmetrical, the size of the asymmetry may be damped by the B- and L-violating interactions. Decays and inverse decays can reduce an asymmetry by a two step process: An excess of U quarks, say, can produce $\overline{X}$'s by inverse decay, $UU\to\overline{X}$, and the subsequent decays will be by both channels, $\overline{X}\to UU$ and $\overline{X}\to\overline{D}\overline{L}$. The BNC reactions can accomplish this directly, $UU\to\overline{D}\overline{L}$ by virtual exchanges.

The case $\epsilon=0$ has interest even though no baryon asymmetry develops for two reasons. First, it allows us to develop some intuition for the system without the complication introduced by CP violation, and second, if ϵ is very small, $\epsilon < 10^{-9}$, then the observed asymmetry must be the result of non-GUT processes; the most likely at this time appear to be either an arbitrary initial condition or quantum gravitational effects. The damping of processes which occur prior to the GUT epoch must be calculated in order to determine the initial value of the asymmetry necessary to explain the observed $kn_B/s = 10^{-9.8\pm1.6}$.

Since SU(5) and some other theories conserve $B-L$, we distinguish two classes of initial asymmetry, with and without initial $B-L=0$. In the former case it is possible for the initial asymmetry to be reduced to zero by the B- and L-violating interactions. In the latter this is not possible; since $B-L$ is exactly conserved, the most that can happen is a redistribution of B and L asymmetries among all species.

Since for $\epsilon = 0$ the $-$ equations are homogeneous, the solutions scale with the value of the initial asymmetry. The relevant quantity is the ratio of final to initial asymmetry. When $B-L=0$, we use the size of the initial baryon (or lepton) excess as the scale.

A. Initial $B-L=0$

In this case the initial baryon and lepton asymmetries are equal $(kn_B/s)_i = (kn_L/s)_i$. The time development of kn_B/s is shown in Fig. 1. As can be seen from this figure, any damping that does occur happens for $T \cong M_X$. This is easy to understand. The expansion rate of the universe is $H = (8\pi G\rho/3)^{1/2} = 3.3T^2/T_P$. At high temperatures, rates of D and ID are $\Gamma_D \simeq \Gamma_{ID} \simeq \alpha M_X^2/T$ because of time dilation, while BNC scattering processes occur at a rate $\Gamma = n\sigma v \sim T^3(\alpha/T^2) \sim \alpha T$ using the Debye-screened cross sections. For $T \gg M_X$, both of these rates are small compared to H.

For $T \ll M_X$, BNC processes have rates $\Gamma_{BNC} \sim T^3(\alpha^2 T^2/M_X^4) \sim \alpha^2 T^5/M_X^4$ and again $\Gamma_{BNC} \ll H$. At low temperatures inverse decays are suppressed by $\exp(-M_X/T)$, as typical fermion pairs are not energetic enough to produce superheavy bosons. The decay rate $\Gamma_D \sim \alpha M_X$ will now surpass the expansion rate, however, it is the two-step process described above which is required to damp initial asymmetries, and the inverse decays are not happening.

The amount of damping that occurs depends on the effectiveness of the B- and L-violating reactions for $T \cong M_X$. This effectiveness is proportional to $K = \alpha T_p/(6.6M_X) = 2.9\times10^{17}\alpha$ GeV$/M_X$; for $T \cong M_X$, $\Gamma_D/H \simeq \Gamma_{ID}/H \simeq K$ and $\Gamma_{BNC}/H \simeq \alpha K$. In Fig. 1 curves are plotted for $K=0.2$, $K=1.0$,

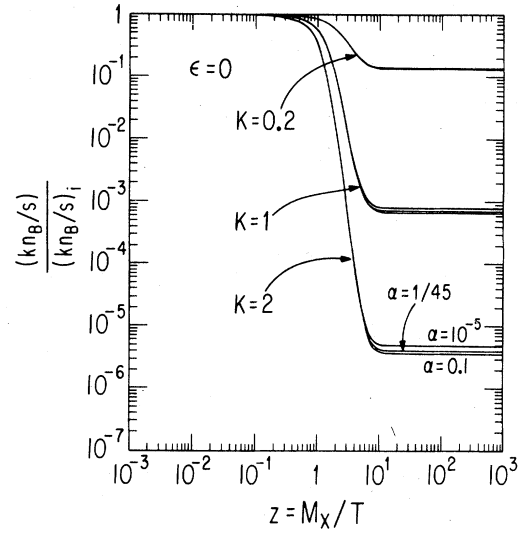

FIG. 1. The time evolution ($t \sim z^2$) of an initial asymmetry with $B-L=0$ is shown in the case where C or CP is conserved ($\epsilon=0$). Results are presented for $K=0.2$, 1.0, 2.0 and $\alpha=10^{-5}$, $\frac{1}{45}$, 0.1 ($K=2.9\times10^{17}\alpha\,\mathrm{GeV}/M_X \sim \Gamma_D/H$ at $z=1$). Most of the damping occurs for $z\sim1$ and is due to D and ID since the results are insensitive to α (for fixed K, $\alpha=0$ corresponds to neglecting all processes except D and ID). Partial equilibrium was assumed ($U_+ = D_+ = L_+ = \nu_+ = 2$).

and $K=2.0$; as K increases, the amount by which the initial asymmetry is damped increases markedly.

Another feature which is apparent in Fig. 1 is that the final asymmetry is insensitive to α for fixed K. Since for fixed K setting $\alpha=0$ is equivalent to including only decays and inverse decays of the X and Y bosons, most of the damping must be due to these processes.

Since the BNC processes die out as a power of the temperature while inverse decays are cut off exponentially, it might be expected that they could be significant for later times, after inverse decays are cut off. For late times, including only the scattering terms gives approximately

$$Q'_- \cong -KzsQ_-, \qquad (4.1)$$

where Q_- stands for either U_- or D_- and $s \simeq 1000\alpha z^{-5}$, $z = M_X/T \gg 1$. This is easily solved: $Q_-(\infty) = Q_-(z_f)\exp(-1000\alpha K/z_f^3)$. The value z_f for which inverse decays stop is found from setting $\Gamma_{ID}/H = 1$; z_f is a solution of $Kz_f^{7/2}\exp(-z_f) = 1$. For $K=3.2$ we obtain $z_f = 8.8$; for $K=10$, $z_f = 10.5$. For both cases, the extra damping is about $\exp(-8\alpha)$. For $\alpha = \frac{1}{45}$, this is insignificant, and even for α as large as 0.1, this is less than a factor of 2.

In Fig. 2 we show the asymmetry damping factor as a function of K for the values $\alpha = 0.1$, $\alpha = \frac{1}{45}$,

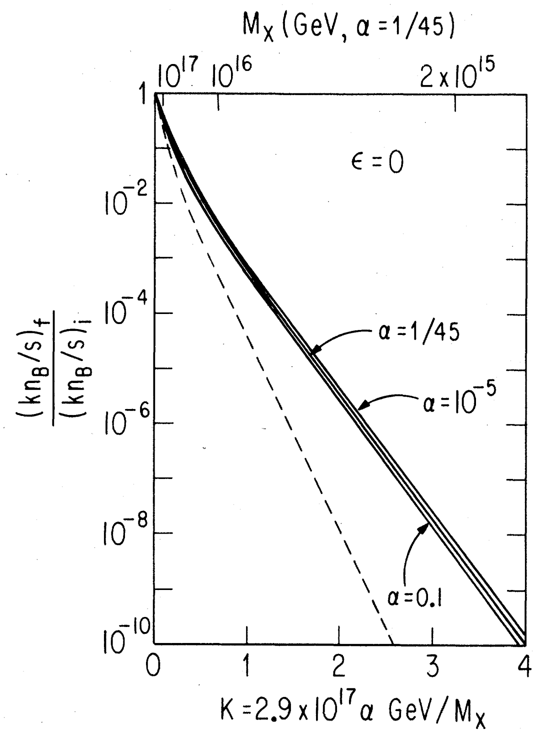

FIG. 2. The ratio of final to initial asymmetry (damping) is shown as a function of K and α when $B-L=0$ and C or CP is conserved ($\epsilon=0$). A mass scale is given at the top for $\alpha=\frac{1}{45}$. The solid lines correspond to the limiting case of partial equilibrium; the broken line corresponds to the limiting case where B- and L-conserving interactions are ineffective (and $\alpha=\frac{1}{45}$). The change in slope of these curves near $K\sim 1$ is discussed in Sec. IV. For large K the solid lines $\sim\exp(-5.5K)$ and the broken line $\sim\exp(-8.3K)$.

and $\alpha=10^{-5}$. The results are relatively insensitive to α, and the damping for large K, $\alpha=10^{-5}$ is approximately given by $\exp(-5.5K)$ when we assume partial equilibrium, $U_+=D_+=L_+=\nu_+=2$ and kinetic equilibrium among X and Y bosons. In our other case, where B- and L-conserving interactions are completely ineffective, the damping is more efficient, as shown by the dotted line in Fig. 2; for large K, $\alpha=10^{-5}$, it is approximated by $\exp(-8.3K)$. This is simple to understand. Here, U_+, D_+, L_+, and ν_+ all become greater than 2 and the rate of inverse decays is correspondingly greater. This is the only result which depends significantly on the assumption regarding equilibrium.

The shape of the damping curve in Fig. 2 is a bit perplexing. For $\alpha=10^{-5}$ and $K\lesssim 1$ the damping is given by $\exp(-10K)$, and for $K\gg 1$ the damping is given by $0.13\exp(-5.5K)$. This is straightforward to explain. If one assumes partial equilibrium ($U_+=D_+=L_+=\nu_+=2$), that X_- and Y_- take on their equilibrium values (found by solving $X_-'=Y_-'=0$; this should be a good approximation for $K\gtrsim 1$), and ignores BNC processes ($\alpha\equiv 0$), then the equations for U_-, D_-, L_-, and ν_- are a set of coupled, linear differential equations.

This set can be written most conveniently in terms of its eigenvectors and eigenvalues,

$$A_j'=-(zK)\gamma_{\rm ID}\lambda_j A_j\,, \qquad (4.2)$$

where the eigenvectors A_j are linear combinations of U_-, D_-, L_-, and ν_- and j runs from 1 to 4. The solution to (4.2) is given by

$$A_j(z)=A_j(0)\exp\left[-\lambda_j K\int_0^z z'\gamma_{\rm ID}(z')dz'\right]. \qquad (4.3)$$

Our system has two zero eigenvalues ($\lambda_1=\lambda_2=0$) corresponding to the two conserved quantities: (i) $B-L$, and (ii) charge Q. These zero eigenvalues guarantee that any initial asymmetry [specified by $U_-(0)$, $D_-(0)$, $L_-(0)$, and $\nu_-(0)$] with Q and/or $B-L\neq 0$ always has Q and/or $B-L\neq 0$. The other two eigenvalues are $\lambda_3=3.29$ and $\lambda_4=6.83$. The time evolution of any initial asymmetry can be given as a linear combination of the solutions in (4.3). If $B-L=Q=0$, then the solution is given in terms of $A_3(z)$ and $A_4(z)$ alone. In particular, if we choose $U_-(0)=D_-(0)=L_-(0)=\frac{1}{2}\nu_-(0)$ (as we did to produce Figs. 1 and 2), then kn_B/s is given by

$$(kn_B/s)=(kn_B/s)_i\times\left\{0.13\exp\left[-3.29K\int_0^z z'\gamma_{\rm ID}(z')dz'\right]+0.87\exp\left[-6.83K\int_0^z z'\gamma_{\rm ID}(z')dz'\right]\right\}, \qquad (4.4)$$

$$(kn_B/s)_f/(kn_B/s)_i=[0.13\exp(-5.5K)+0.87\exp(-10K)]\,, \qquad (4.5)$$

where $\int_0^\infty z\gamma_{\rm ID}dz$ has been evaluated numerically. For $K\gg 1$ ($M_X\ll 2.9\times10^{17}\ \alpha$ GeV), the damping should be given by $\exp(-5.5K)$—which agrees well with the numerical results shown in Fig. 2. For $K\lesssim 1$ ($M_X\gtrsim 2.9\times10^{17}\ \alpha$ GeV), the damping should be given by $\exp(-10K)$—which also agrees with the numerical results in Fig. 2. Unless the initial asymmetry is "pure" A_4, then for large K it will be damped by $\exp(-5.5K)$.

B. Initial $B-L\neq 0$

Because SU(5) exactly conserves $B-L$, if $B-L\neq 0$ initially no interactions can reduce both B and L to zero. At late times, $z\gg 1$, asymmetries of order the size of $B-L$ must exist in B and/or L. When B and L violations are effective, $K\gg 1$, regardless of which species had the initial asymmetries, all species U, D, L, and ν have final asymmetries of about the same order. On the other hand, when $K \ll 1$ only species with initial asymmetries have significant final asymmetries.

The situation in which the universe initially has only a neutrino asymmetry is illustrated in Fig. 3, where $(kn_B/s)_f/(kn_L/s)_i$ is shown as a function of K. We conclude that if some process in the early universe produced only a neutrino asymmetry $kn_L/s \cong 10^{-9}$, then even in the absence of CP violation X's and Y's could have translated this into a baryon asymmetry of the right size if $M_X \lesssim 10^{17}$ GeV.

C. 5-ness

Because we averaged all quantities over spin we missed an additional zero eigenmode corresponding to another conserved quantity, 5-ness. 5-ness is

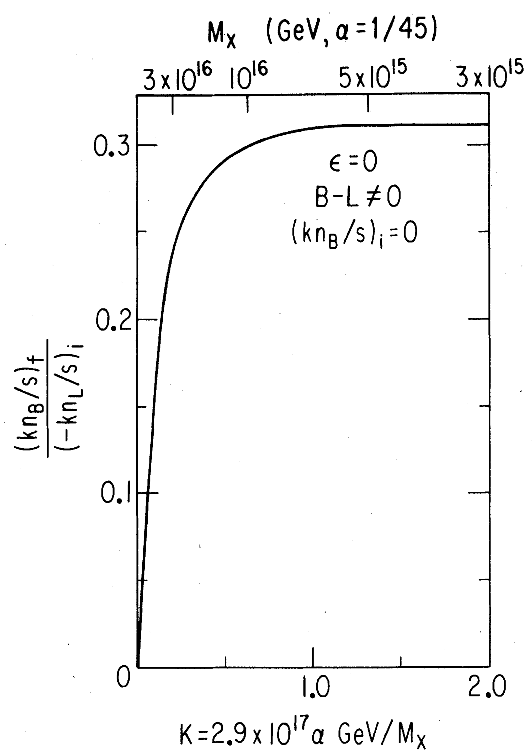

FIG. 3. The ratio of final baryon asymmetry to initial lepton asymmetry is shown as a function of K when $B-L \neq 0$ and C or CP is conserved ($\epsilon = 0$). A mass scale is given at the top for $\alpha = \frac{1}{45}$. The initial baryon asymmetry is zero; because SU(5) conserves $B-L$ and $B-L \neq 0$, B- and L-violating interactions cannot damp the initial asymmetry significantly, they can only redistribute it. Here an initial lepton asymmetry results in both final lepton and baryon asymmetries. The curve shown is for $\alpha = \frac{1}{45}$; since our results are insensitive to α for fixed K, it is approximately correct for all $\alpha \lesssim 0.1$. Partial equilibrium was assumed.

the property of being in the $\underline{5}$ representation of SU(5); all the particles in the $\underline{5}$ have 5-ness +1, those in the $\overline{\underline{5}}$ have 5-ness − 1, and all others 5-ness zero. The interactions of the X and Y conserve 5-ness, so like an initial asymmetry with $B - L \neq 0$, one with net 5-ness will not be damped.[30]

However, Higgs bosons *do not* conserve 5-ness. Therefore, the interactions of a Higgs boson (whose interactions may or may not conserve B and L) together with X and Y gauge bosons could erase asymmetries with net 5-ness. Whether the interactions of such Higgs bosons are rapid enough ($\Gamma > H$) at the appropriate time ($T \sim M_X$) depends upon the Higgs structure and is thus very model dependent. This is an issue we will not address further in this paper.

V. C AND CP VIOLATED ($\epsilon \neq 0$)

Certainly the more interesting situation is when $\epsilon \neq 0$ and interactions in the early universe when $T \cong M_X$ can allow a baryon asymmetry to arise dynamically. Once again we consider two classes of initial conditions: first, the case in which the universe is initially completely symmetrical, $B = L = 0$, and second, the case in which an initial asymmetry is already present with $B \neq 0$, $L \neq 0$, but $B - L = 0$. The asymmetry with which we start our computations might be due to pre-GUT processes or might truly represent an initial condition.

In the first situation ($B = L = 0$ initially) since the asymmetry equations are linear in all the − quantities with sources proportional to ϵ, the results scale simply as ϵ. In the second situation, $B = L \neq 0$ initially, there is an additional scale in the problem which can be taken to be the ratio of the initial asymmetry to ϵ. For a given value of this ratio the results again scale as ϵ. All our results are for an initial asymmetry which is 100 times the maximum that could be produced by the GUT interactions alone for the given value of ϵ.

A. Initial $B = L = 0$

The time development of the baryon asymmetry is shown in Fig. 4 as a function of $z \equiv M_X/T$. Once again we observe that most of the interesting processes occur near $T \cong M_X$. As discussed in Sec. IV, this is because of the reactions—D, ID, and BNC—are effective, $\Gamma/H \gtrsim 1$, only for $T \cong M_X$. Also, for the most part the results for fixed K are insensitive to α. The exception to this is the curve for $M_X = 2.9 \times 10^{16}$ α GeV ($K = 10$) and $\alpha = 0.1$, where one can see additional damping due to BNC processes by a factor of roughly 2. As in Sec. IV we expect damping by a factor of $\exp(-1000\alpha K/z_f^3)$ by these processes after inverse decays stop.

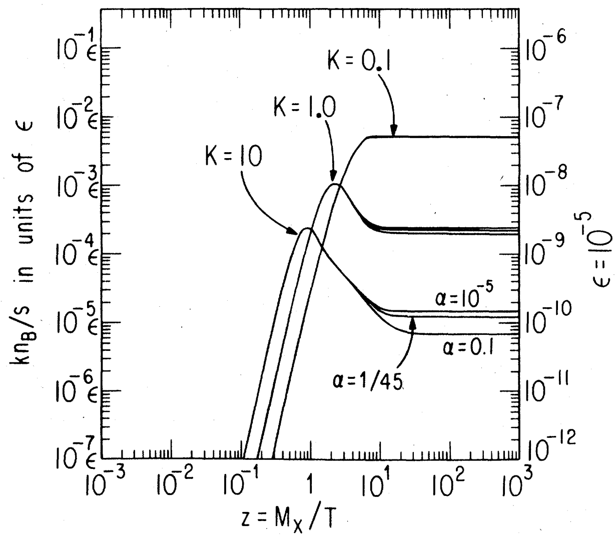

FIG. 4. The time evolution ($t \sim z^2$) of the baryon asymmetry (kn_B/s) in units of ϵ is shown in the case when C and CP are violated ($\epsilon \neq 0$) and the initial asymmetry is zero. Results are presented for $K = 0.1$, 1.0, 10 and $\alpha = 10^{-5}$, $\frac{1}{45}$, 0.1 ($K = 2.9\times10^{17}\alpha$ GeV$/M_X$). For $K \ll 1$, the qualitative picture of Weinberg and Wilczek is borne out—the asymmetry is produced by delayed ($z \gg 1$) free decays. For $K \gtrsim 1$ the scenario is more complex: initially there is a period of nonequilibrium and growth ($z \ll 1$), $kn_B/s \sim z^5$, followed by quasiequilibrium, $kn_B/s \sim z^{-1}K^{-1}$, and finally when ID freeze out [$z \gtrsim z_f$, $z_f^{7/2}\exp(-z_f)K = 1$]kn_B/s freezes out and remains constant. D and ID are primarily responsible for these results; BNC processes are only important for large K and α. Partial equilibrium was assumed, although these results are not sensitive to this assumption.

This is about a factor of 2 for $K = 10$ and $\alpha = 0.1$, and $\alpha = 0.1$ is larger than what is usually expected for the grand unification coupling strength ($\alpha = 0.1$ was chosen to maximize the effects of BNC processes). For the rest of our discussion we will assume that the effects of BNC processes are not important.

When M_X is very large, $M_X \gg 3\times10^{17}$ α GeV, $K \ll 1$, decays are ineffective when $T \cong M_X$, or, equivalently, the lifetime of the X is long compared to the age of the universe at that point. Here, eventually the X bosons decay freely, producing an asymmetry of $kn_B/s = 7.8\times10^{-3}\epsilon \simeq \epsilon/g_*$. We find the qualitative picture of Weinberg and Wilczek to be basically correct for $M_X \gtrsim 3\times10^{18}$ α GeV, $K \lesssim 0.1$ (see Fig. 4).

This behavior can be understood simply from the Boltzmann equations. The relevant equations from from Appendix B are (B5), (B7), and (B8), for X_+, U_-, and D_-. Ignoring BNC terms and couplings, we have

$$X_+' = -3\gamma_D K z (X_+ - 2\gamma_{ID}/\gamma_D)\,, \quad (5.1a)$$

$$U_-' = 3\gamma_D K z[(-\tfrac{11}{6})(\gamma_{ID}/\gamma_D)U_- + (\epsilon/3)(X_+ - 2\gamma_{ID}/\gamma_D)]\,, \quad (5.1b)$$

$$D_-' = 3\gamma_D K z[(-\tfrac{5}{6})(\gamma_{ID}/\gamma_D)D_- + (\epsilon/6)(X_+ - 2\gamma_{ID}/\gamma_D)]\,, \quad (5.1c)$$

where we assume $U_+ = D_+ = L_+ = \nu_+ = 2$, partial equilibrium established by B- and L-conserving interactions. From (5.1a) we see that the equilibrium value of X_+ is $2\gamma_{ID}/\gamma_D$, and when X_+ assumes this value the source terms in (5.1b) and (5.1c) vanish. For small z, $\gamma_{ID}/\gamma_D \simeq 1 - z^2/4$; at high temperatures the equilibrium value $X_+ \cong 2$ means X bosons are as abundant as other relativistic particles. For large z, $\gamma_{ID}/\gamma_D \simeq (\pi/8)^{1/2}z^{3/2}e^{-z}$; the equilibrium abundance of X's is smaller than that of relativistic particles by $\sim\exp(-M/T)$.

The rates for decays and inverse decays are $\Gamma_D \sim \alpha M_X \gamma_D$ and $\Gamma_{ID} \sim \alpha M_X \gamma_{ID}$ with effectiveness $\Gamma_D/H \simeq K\gamma_D z^2$ and $\Gamma_{ID}/H \simeq K\gamma_{ID} z^2$. Since $\gamma_D \lesssim 1$ (for small z, $\gamma_D \sim z$, because of time dilation, while for large z, $\gamma_D \cong \frac{2}{3}$, all rates are small for $K \ll 1$, until $Kz^2 \gtrsim 1$ ($z \gtrsim K^{-1/2}$). At this point, $\gamma_{ID}/\gamma_D \sim e^{-K^{-1/2}} \ll 1$, so the equations can be integrated neglecting the terms proportional to γ_{ID}/γ_D. The solutions are

$$\begin{aligned} X_+(z) &\cong X_+(0)\exp(-Kz^2) \cong 2\exp(-Kz^2)\,, \\ U_-(z) &\cong (2\epsilon/3)[1 - \exp(-Kz^2)]\,, \\ D_-(z) &\cong (\epsilon/3)[1 - \exp(-Kz^2)]\,. \end{aligned} \quad (5.2)$$

The final asymmetry is $kn_B/s = 7.8\times10^{-3}(U_- + D_-) = 7.8\times10^{-3}\epsilon$. Recall $K = \alpha M_X/(6.6T_p)$, $z = M_X/T$, and $t = 0.154\, T_p/T^2$ so that $Kz^2 = t/\tau_X$ ($\tau_X^{-1} = \alpha M_X$) and this corresponds exactly to late free decay.

The behavior in the other limit, $K \gg 1$, can also be understood easily. Here there are three regions: First, a period of growth for $z \lesssim 1$, followed by a region of slow decline $kn_B/s \sim z^{-1}$, and a final, asymptotically constant region.

1. $z \ll 1$

Initially, at $t = t_p$, X_+ is set equal to $2\gamma_{ID}/\gamma_D \simeq 2$. As z increases, the equilibrium value of X_+, $2\gamma_{ID}/\gamma_D$, begins to decrease because the X is massive, but at early times decays are suppressed by time dilation and X_+ remains equal to 2. Thus the source terms in (5.1b) and (5.1c) are nonzero and

$$D_-' \sim U_-' \sim \epsilon K z \gamma_D (1 - \gamma_{ID}/\gamma_D)\,. \quad (5.3)$$

We expect $1 - \gamma_{ID}/\gamma_D \simeq z^2/4$, and since $\gamma_D \simeq z/3$ the growth of U_-, D_-, and kn_B/s should be as z^5; in particular, the behavior $kn_B/s \sim (7.8\times10^{-3}\epsilon)Kz^5$ is seen in Fig. 4 for $z \ll 1$.

2. $K^{-1/3} < z < z_f$

The source term for U_- and D_- is $\propto \Delta \equiv X_+ - 2\gamma_{ID}/\gamma_D$ which measures the degree to which X_+ is out of equilibrium. Early ($z \ll 1$) $\Delta \sim z^2$, but decays will begin to damp out Δ when $zX_+' \simeq \Delta$, which occurs when $z^3 K \cong 1$. At about this same time, the damping terms in the U_- and D_- equations become comparable to the source terms, since

$$\frac{(\gamma_{ID}/\gamma_D)U_-}{\epsilon(X_+ - 2\gamma_{ID}/\gamma_D)} \sim \frac{K\epsilon z^5}{\epsilon z^2} \sim Kz^3, \tag{5.4}$$

so that growth of U_- and D_- halts here. This behavior can be seen in Fig. 4; for $M_X = 2.9\times10^{16}\ \alpha$ GeV ($K=10$), kn_B/s stops growing at $z \cong 0.6$, while $10^{-1/3} \cong 0.5$.

For $z \gtrsim 1$, $\gamma_D \simeq \frac{2}{3}$ and $\gamma_{ID}/\gamma_D \simeq (\pi/8)^{1/2} z^{3/2} e^{-z}$. The effectiveness of decays and inverse decays are $\Gamma_D/H \sim Kz^2$ and $\Gamma_{ID}/H \sim Kz^{7/2}e^{-z}$. Decays will always be effective for $z \gtrsim 1$; however, inverse decays effectively cease when $Kz_f^{7/2}e^{-z}f = 1$. For $1 < z < z_f$ D and ID are occurring rapidly on the expansion time scale and a quasiequilibrium is established; in particular, with large positive and negative contributions to the rates, during this period U_- and D_- should assume values such that $U_-' \simeq D_-' \simeq 0$ or

$$U_- \approx D_- \approx \tfrac{1}{5}\epsilon\Delta/(\gamma_{ID}/\gamma_D)\,. \tag{5.5}$$

The deviation from equilibrium Δ, and hence the asymmetries U_- and D_-, are determined from (5.1a) (for $z \gg \frac{3}{2}$):

$$\Delta' = (X_+ - 2\gamma_{ID}/\gamma_D)' \cong (\pi/2)^{1/2} z^{3/2} e^{-z} - 2Kz\Delta\,. \tag{5.6}$$

For $K \gg 1$, the term $-2Kz\Delta$ will be large enough to establish $\Delta' \cong 0$, so that

$$\Delta \cong K^{-1}(\pi/8)^{1/2} z^{1/2} e^{-z}\,. \tag{5.7}$$

Combining (5.5) and (5.7) we obtain that, for $1 \le z \le z_f$, U_- and D_- are given by

$$U_- \approx D_- \cong \tfrac{1}{5}\epsilon z^{-1} K^{-1}\,. \tag{5.8}$$

Therefore, when $K \gtrsim 1$, $kn_B/s = 7.8\times10^{-3}(U_- + D_-)$ should decline as z^{-1} for $1 \lesssim z \lesssim z_f$. This is the behavior seen in Fig. 4.

3. $z > z_f$

The two competing rates $(\gamma_{ID}/\gamma_D)U_-$ and $\epsilon(X_+ - 2\gamma_{ID}/\gamma_D)$ in (5.1b) both remain large as long as both D and ID are effective. When this is no longer true, we expect that U_- will no longer change significantly and will be frozen at the value it has then. Since the slowest processes in (5.1) are inverse decays, we expect this freeze out to occur at z_f (when $\Gamma_{ID} = H$). For $z > z_f$ we expect U_-, D_-, and kn_B/s to be given by their values at z_f,

$$U_-(z > z_f) \approx D_-(z > z_f) \cong \tfrac{1}{5}\epsilon z_f^{-1} K^{-1}\,,$$
$$kn_B/s \cong 3.0\times10^{-3}\epsilon z_f^{-1}K^{-1}\,. \tag{5.9}$$

Asymptotically $z_f \to \ln(K)$ but for the values of K of interest here this is a poor approximation. For $K=10$, we find $z_f \cong 11$, which is very close to where the curve flattens in Fig. 4. From (5.9) we then expect $kn_B/s \simeq 2.7\times10^{-5}\epsilon$, which agrees rather well with the final value for $K=10$ and $\alpha = 10^{-5}$; recall in the discussion above we neglected all BNC processes.

B. Initial $B = L \neq 0$

In this case we start with an initial asymmetry $(kn_B/s)_{\text{initial}} = 0.78\epsilon$, which is 100 times larger than $7.8\times10^{-3}\epsilon$, the maximum that could be produced by GUT processes alone. In light of the discussion above, the results here are easy to describe. The evolution of the baryon asymmetry with time in this case is shown in Fig. 5.

For $K \lesssim 1$, the initial asymmetry is damped by D and ID when $T \cong M_X$. However, the damping is not enough to establish the quasiequilibrium situation where $U_- \simeq (\frac{1}{5})\epsilon\Delta/(\gamma_{ID}/\gamma_D)$. The final value of

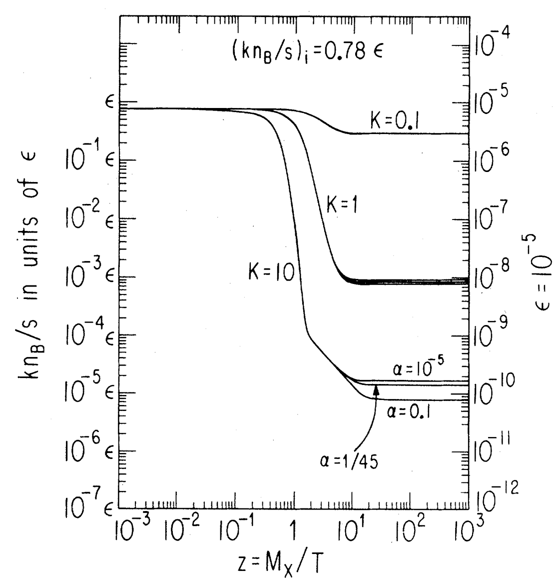

FIG. 5. The time evolution ($t \sim z^2$) of the Baryon asymmetry in units of ϵ is shown when $B-L=0$, $\epsilon \neq 0$, and there is an initial asymmetry $kn_B/s = 0.78\epsilon$. Results are presented for $K=0.1$, 1.0, 10 and $\alpha = 10^{-5}$, $\frac{1}{45}$, 0.1. For $K=0.1$ the initial asymmetry is damped slightly by D and ID. For $K=10$ D and ID rapidly damp the initial asymmetry and a period of quasiequilibrium is established, $kn_B/s \sim K^{-1}z^{-1}$, and the final asymmetry is essentially independent of the initial asymmetry (cf. Fig. 4).

the asymmetry is the same as it would have been if $\epsilon = 0$; the residue left from the initial asymmetry overwhelms that generated by the CP-violating effects.

For $K \gtrsim 3$, the rates of D and ID for $T \cong M_X$ are great enough to bring U_- and D_- to their quasiequilibrium values (5.8). Quasiequilibrium is maintained until $z \sim z_f$, and finally $(z > z_f)$ kn_B/s takes on the value it would have had there been no initial asymmetry. The division between these two cases will depend on the ratio of the initial asymmetry to that which could be generated $(7.8\times10^{-3}\epsilon)$.

C. Summary

When $M_X \gtrsim 4.6\times10^{18}$ α GeV ($K \lesssim 1$), an asymmetry $kn_B/s \cong 7.8\times10^{-3}\epsilon$ evolves, just as in the qualitative picture of Weinberg and Wilczek. When $M_X \lesssim 4.6\times10^{18}$ α GeV ($K \gtrsim 1$) there is an initial period of nonquilibrium growth ($T \gtrsim M_X$), followed by a period of quasiequilibrium where the symmetry falls as T ($M_X \gtrsim T \gtrsim M_X/z_f$). Then, when $T \cong M_X/z_f$, ID freeze out and the asymmetry remains constant at $kn_B/s \cong 3.0\times10^{-3}\epsilon z_f^{-1}K^{-1}$.

Figure 6 shows the results of the numerical integration, the final asymmetry kn_B/s as a function of α and M_X. For αK not too large, the results are relatively insensitive to α, indicating D and ID are primarily responsible for these results, with the exception of the CP-violating parts of the BNC processes required by unitarity, which are expressible in terms of ID rates.

Although we expect for large K to find $kn_B/s \propto K^{-1}$ with logarithmic corrections, these corrections are large for $K \lesssim 10$, and the results for $\alpha = \frac{1}{45}$ are well represented by

$$kn_B/s = 7.8\times10^{-3}\epsilon/[1 + (16K)^{1.3}]\,. \qquad (5.10)$$

The value $\epsilon = 10^{-5}$ was chosen for the right-hand ordinate scale since Weinberg and Nanopoulos have shown that in SU(5) $\epsilon \lesssim \alpha \simeq 10^{-4} - 10^{-6}$. For $M_X = 3\times10^{14}$ GeV, we need $\epsilon \simeq 10^{-4.3}$ to obtain $kn_B/s = 10^{-9.8}$.

The final asymmetry as a function of K is shown in Fig. 7 when there is an initial asymmetry, $kn_B/s = 0.78\epsilon$. For an arbitrary initial asymmetry (with $B - L = 0$) the results are well represented by

$$kn_B/s \cong (kn_B/s)_{\text{initial}} \exp(-5.5K) + 7.8\times10^{-3}\epsilon/[1 + (16K)^{1.3}]\,. \qquad (5.11)$$

All of the results in this section were obtained

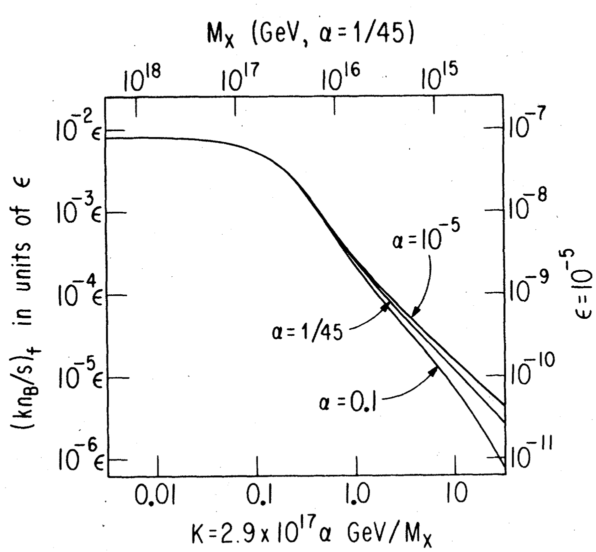

FIG. 6. The final baryon asymmetry (kn_B/s) in units of ϵ is shown as a function of K and α when the initial asymmetries are zero ($B_i = L_i = 0$), and C and CP are violated ($\epsilon \neq 0$). A mass scale is given for $\alpha = \frac{1}{45}$. These results are relatively insensitive to α for fixed K and do not depend significantly on the assumptions about the effectiveness of B- and L-conserving interactions. For $K \ll 1$, the X bosons decay late and freely, producing an asymmetry $(kn_B/s) \sim 7.8\times10^{-3}$. For $K \gtrsim 1$ the final asymmetry has a rather weak dependence upon K, decreasing as $K^{-1.3}$ (see discussion in Sec. V).

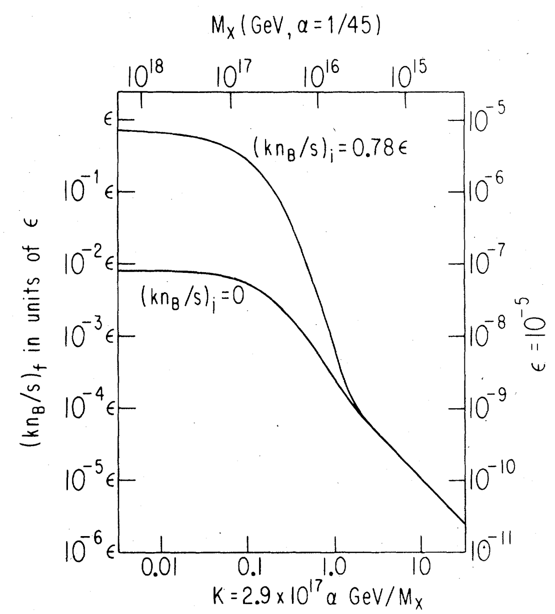

FIG. 7. The final baryon asymmetry (kn_B/s) in units of ϵ is shown as a function of K and $\alpha = \frac{1}{45}$ when $B - L = 0$, $\epsilon \neq 0$, and $(kn_B/s)_i = 0.78\epsilon$. A mass scale is given for $\alpha = \frac{1}{45}$. For reference the curve for $(kn_B/s)_i = 0$ and $\alpha = \frac{1}{45}$ is also shown. For $K \ll 1$, the initial asymmetry is damped only slightly, and the final asymmetry is equal to the initial asymmetry times a damping factor. For $K \gg 1$, the initial asymmetry is damped sufficiently so that the final asymmetry is independent of the initial asymmetry (the two curves merge).

assuming partial equilibrium for the fermions, $U_+=2$ etc., and kinetic equilibrium for the X and Y bosons. In the other limiting case, B- and L-conserving interacting ineffective, the final results only differ by $\lesssim 50\%$.

VI. CONCLUDING REMARKS

A. Summary

If C or CP is a symmetry of the GUT ($\epsilon=0$), then B- and L-violating interactions can only reduce initial asymmetries. The damping is primarily due to the D and ID which occur when $T \sim M_X$. If the initial asymmetry has $B-L=0$, then that asymmetry is reduced by a factor of $\exp(-5.5K - -8.3K)$ depending upon whether or not B- and L-conserving interactions are effective ($5.5K$ if they are, $8.3K$ if they are not). On the other hand, if $B-L$ is initially nonzero, then initial asymmetries can only be redistributed.

There is one interesting consequence of this damping. Suppose the proton is seen to decay with a lifetime of $\lesssim 10^{33}$ yr, the level of sensitivity expected to be achieved this decade. Then this implies the existence of a superheavy boson which mediates B and L violations with a mass $\lesssim 10^{15}$ GeV (for $\alpha=\frac{1}{45}$). Then if one wants to explain the observed asymmetry today, $kn_B/s=10^{-9.8}$, as the result of an initial condition rather than being cosmologically generated, that initial asymmetry must be larger than $10^{-9.8}$ since it will be reduced by the D and ID of the superheavy boson. If $B-L=0$ initially, then $10^{-9.8} \lesssim \exp(-1.6\times 10^{18}\ \alpha\ \text{GeV}/10^{15}\ \text{GeV})\ (kn_B/s)_i \simeq 10^{-15}(kn_B/s)_i$—making it impossible to explain the observed asymmetry in terms of an initial condition. Cosmological baryon generation may be a necessity rather than a luxury.

If C and CP are violated in the GUT ($\epsilon \neq 0$), then a baryon asymmetry can evolve from an initially symmetrical universe. If the mass of the superheavy boson is greater than $M_C \cong 4.6\times 10^{18}\ \alpha$ GeV, then just as in the Weinberg-Wilczek picture an asymmetry $kn_B/s \cong 7.8\times 10^{-3}\epsilon$ evolves due to delayed decays. If M_X is less than M_C, then the situation is a bit more complex—there is a period of quasiequilibrium and a finally asymmetry, $kn_B/s \cong 7.8\times 10^{-3}\epsilon(M_X/M_C)^{1.3}$ emerges. These results are primarily due to decays and inverse decays occurring when $T\sim M_X$. If the superheavy masses are $\sim 3\times 10^{14}$ GeV, then $\epsilon \sim 10^{-4.3}$ is needed for gauge bosons to account for $kn_B/s=10^{-9.8}$.

Our results for $\epsilon \neq 0$ are in qualitative agreement with those obtained by Kolb and Wolfram[26] for their "simple model." We do differ in that they found BNC processes have a significant effect for $\alpha \gtrsim 10^{-2}$. (Their simple model was designed to simulate a typical GUT and has two particle species—b, baryon number $\frac{1}{2}$, and X, a superheavy boson.)

In light of the number of simplifying assumptions and approximations made, it is appropriate to discuss which of them in retrospect seem reasonable and which of them require further study—in short, what are the strengths and weaknesses of this paper?

B. Strengths

Recall that the observed asymmetry $kn_B/s \approx 10^{-9.8\pm 1.6}$ is known to within an order of magnitude or so. To that level of accuracy the following assumptions seem justifiable: (i) The neglect of degeneracy factors, and in the future possibly the use of Maxwell-Boltzmann distributions for particle number densities to further simplify the calculations. (ii) The neglect of annihilations (at least for gauge bosons), and other higher-order processes. For $\alpha \lesssim \frac{1}{45}$ the effects of BNC processes of $O(\alpha^2)$ were small, justifying the neglect of $O(\alpha^3)$ and higher processes. (iii) Except with regard to the damping of initial asymmetries, the question of whether or not B- and L-conserving interactions can maintain a partial equilibrium seems to be unimportant. Taking the two extreme views (completely effective, completely ineffective) changes the final results only by $\lesssim 50\%$. (iv) Finally, the question of which GUT should be used in these calculations. As was discussed earlier, as far as baryon generation goes, the important features of a GUT are the number of different species of superheavy boson which mediate B and L violations. Since there are only three basic types of superheavy bosons which mediate B and L violations, our results in this paper and the companion paper[15] should accurately reflect what any species might do by itself. The issue of what happens when they all operate at the same time (species of similar mass) or in sequence (species of different mass) is a more complex one and is addressed in a third paper.[16]

C. Weakness—work to be done

By far, the weakest link in this calculation is the CP violation. We have parametrized it by ϵ, and the only definite theory-independent constraint one has on ϵ is that it is less than $\alpha_{\text{gauge}} \sim 10^{-2}$. Progress here is sorely needed. There are also two more difficult issues to address. (i) Were equilibrium distributions established at $t\approx 10^{-43}$ sec, and, if so, by what method? (ii) What are the effects of non-GUT processes (quantum gravity, primordial black holes, particle creation by curvature, etc.) on the baryon asymmetry?

Note added in proof. Segre and Turner have recently shown that if a fourth generation of heavy quarks ($M \sim 30-200$ GeV) exists then superheavy Higgs bosons can produce kn_B/s of the right magnitude in the minimal SU(5) model [University of Chicago report (unpublished)].

ACKNOWLEDGMENTS

This work grew out of a series of discussions one of us (M.S.T.) had with Gordon Baym at the Aspen Center for Physics during the summer of 1979. In addition we wish to acknowledge the helpful conversations we have had with many people including A. Buras, M. Crawford, T. Curtright, E. Kolb, I. Lerche, S. Love, and D. Schramm. This work was supported in part by J.N.F.'s McCormick Fellowship, K.A.O.'s Hertz Fellowship, M.S.T.'s Fermi Fellowship, and NSF Grant No. AST 78-20402.

APPENDIX A

Spurred by the success of both the $SU(2)\times U(1)$ and SU(3) color-gauge theories, there have been a number of attempts to obtain a unified theory of the strong, weak, and electromagnetic interactions. The simplest grand unified theory is the SU(5) model of Georgi and Glashow.[9] Baryon nonconservation is a direct consequence of this theory and others, for the simple reason that quarks and leptons are grouped together in the same multiplets. SU(5) and other GUT's can incorporate *CP* violation, however, the details remain to be worked out. For the rates computed in this appendix, *CP* (and hence *T*) are assumed to be symmetries of the theory. The *CP* violation necessary for baryon generation is added in an *ad hoc* but self-consistent manner in Sec. III. All the matrix elements needed for our computations are contained in this appendix or can be obtained from rates in this appendix as described in Sec. III.

In SU(5) the fermions (quarks and leptons) are placed in $\underline{\bar{5}}$ and $\underline{10}$ representations,

$$\underline{\bar{5}} = \begin{pmatrix} \bar{d}_R \\ \bar{d}_Y \\ \bar{d}_B \\ e^- \\ \nu_e \end{pmatrix}_L , \quad \underline{10} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & \bar{u}_B & -\bar{u}_Y & -u_R & -d_R \\ -\bar{u}_B & 0 & \bar{u}_R & -u_Y & -d_Y \\ \bar{u}_Y & -\bar{u}_R & 0 & -u_B & -d_B \\ u_R & u_Y & u_B & 0 & -e^+ \\ d_R & d_Y & d_B & e^+ & 0 \end{pmatrix}_L , \tag{A1}$$

where *R*, *Y*, and *B* denote the three colors, and this pattern is followed for each generation. Here we shall assume that there are three generations. The theory includes 24 gauge bosons: the photon, $W^\pm$, Z^0, 8 gluons, a color triplet with charge $\pm\frac{4}{3}(X)$, and a color triplet with charge $\pm\frac{1}{3}(Y)$. The superheavy X and Y gauge bosons [mass M_X—X and Y masses are split only by the breaking of the $SU(2)\times U(1)$ symmetry] mediate baryon nonconservation. The minimal Higgs structure necessary to break SU(5) to $SU(3)\times SU(2)\times U(1)$ and then finally to $SU(3)\times U(1)$ is an adjoint $\underline{24}$ and a complex $\underline{5}$. At low energies ($E < 100$ GeV) there are 19 physical degrees of freedom in the Higgs sector. Among the Higgs bosons is a superheavy color triplet with charge $\pm\frac{1}{3}(H)$ that also mediates B and L violations (it couples to the same fermion channels as the Y). However, in this paper we will not specifically be interested in this triplet.

The parts of the SU(5) Lagrangian which specify the B- and L-violating interactions mediated by the superheavy gauge bosons X and Y are[31]

$$(g/\sqrt{2})X^-_{i\mu}(\epsilon_{ijk}\overline{U}^c_{\alpha kL}\gamma^\mu U_{\alpha jL} + \overline{D}_{\alpha i}\gamma^5\gamma^\mu L^c_\alpha)\,,$$

$$(g/\sqrt{2})Y^-_{i\mu}(\epsilon_{ijk}\overline{U}^c_{\alpha kL}\gamma^\mu D_{\alpha jL} + \overline{D}_{\alpha iR}\gamma^\mu \nu^c_{\alpha R} - \overline{U}_{\alpha iL}\gamma^\mu L^c_{\alpha L})\,, \tag{A2}$$

where $\alpha = (1, 2, 3)$, $U_\alpha = (u, c, t)$, $D_\alpha = (d, s, b)$, $L_\alpha = (e, \mu, \tau)$, and $\nu_\alpha = (\nu_e, \nu_\mu, \nu_\tau)$. The subscripts i, j, k are color indices, μ is the Lorentz index, and the superscript c denotes charge conjugation. All other notation is that of Bjorken and Drell.[23] Note that $g^2 = 4\pi\alpha$.

This Lagrangian leads to the following fermion couplings for the X and Y bosons: $X^- \to \overline{U}_\alpha \overline{U}_\alpha, D_\alpha L_\alpha$ and $Y^- \to \overline{U}_\alpha \overline{D}_\alpha, D_\alpha \nu_\alpha, U_\alpha L_\alpha$. The X and Y mediated BNC fermion scatterings which we require are $UU \to \overline{D}\overline{L}$, $UD \to \overline{U}\overline{L}$, $UD \to \overline{D}\overline{\nu}$, $U\nu \to \overline{D}\overline{D}$, and their CP-conjugated and T-reversed reactions. They can all be obtained from the amplitudes for the following Feynman diagrams: $L_\alpha + D_\alpha \to \overline{U}_{1\beta} + \overline{U}_{2\beta}$ (virtual X), $L_\alpha + U_{1\alpha} \to \overline{D}_\beta + \overline{U}_{2\beta}$ (virtual Y), $D_{1\alpha} + \nu_\alpha \to \overline{U}_\beta + \overline{D}_{2\beta}$ (virtual Y), by appropriate transformations.

Following Wagoner's notation,[22] the elementary transition rate W is given by

$$W = (s/2^n)\,|\mathfrak{M}|^2(2\pi)^4\delta^{(4)}\left(\sum_{\rm in} p^\mu - \sum_{\rm out} p^\mu\right), \tag{2.18$'$}$$

where $|\mathfrak{M}|^2$ (the invariant matrix element) is averaged over initial and final spins, colors, and generations.

1. Decays and inverse decays

For the following processes the matrix elements and transition rates are given by

(i) $X^- \rightleftarrows \bar{U}_{1\alpha} + \bar{U}_{2\alpha}$:

$$|\mathfrak{M}|^2 = (2\pi\alpha/81)M_X{}^2,\quad W = (\pi\alpha/324)M_X{}^2(2\pi)^4\delta^{(4)}(\textstyle\sum p); \tag{A3}$$

(ii) $X^- \rightleftarrows L_\alpha + D_\alpha$:

$$|\mathfrak{M}|^2 = (2\pi\alpha/27)M_X{}^2,\quad W = (\pi\alpha/108)M_X{}^2(2\pi)^4\delta^{(4)}(\textstyle\sum p); \tag{A4}$$

(iii) $Y^- \rightleftarrows \bar{U}_\alpha + \bar{D}_\alpha$:

$$|\mathfrak{M}|^2 = (2\pi\alpha/81)M_X{}^2,\quad W = (\pi\alpha/324)M_X{}^2(2\pi)^4\delta^{(4)}(\textstyle\sum p); \tag{A5}$$

(iv) $Y^- \rightleftarrows D_\alpha + \nu_\alpha$:

$$|\mathfrak{M}|^2 = (2\pi\alpha/27)M_X{}^2,\quad W = (\pi\alpha/108)M_X{}^2(2\pi)^4\delta^{(4)}(\textstyle\sum p); \tag{A6}$$

(v) $Y^- \rightleftarrows U_\alpha + L_\alpha$:

$$|\mathfrak{M}|^2 = (\pi\alpha/27)M_X{}^2,\quad W = (\pi\alpha/216)M_X{}^2(2\pi)^4\delta^{(4)}(\textstyle\sum p). \tag{A7}$$

2. Baryon-nonconserving fermion collisions

The matrix elements for these four processes are given by

(i) $U_1U_2 \rightleftarrows \bar{D}\bar{L}$, $\bar{U}_1\bar{U}_2 \rightleftarrows DL$:

$$\begin{aligned}|\mathfrak{M}|^2 = (8\pi^2\alpha^2/81)\{&2[(p_1\cdot p_D)^2 + (p_2\cdot p_D)^2](2p_1\cdot p_2 - M_X{}^2)^{-2} + (p_2\cdot p_D)^2(2p_1\cdot p_D + M_X{}^2)^{-2}\\ &+ (p_1\cdot p_D)^2(2p_2\cdot p_D + M_X{}^2)^{-2} + \tfrac{2}{3}(p_2\cdot p_D)^2(2p_1\cdot p_D + M_X{}^2)^{-1}(M_X{}^2 - 2p_1\cdot p_2)^{-1}\\ &+ \tfrac{2}{3}(p_1\cdot p_D)^2(2p_2 p_D + M_X{}^2)^{-1}(M_X{}^2 - 2p_1\cdot p_2)^{-1}\};\end{aligned} \tag{A8}$$

(ii) $U_1D \rightleftarrows \bar{U}_2\bar{L}$, $\bar{U}_1\bar{D} \rightleftarrows U_2L$.

The matrix element is obtained from (A8) by the substitutions $p_2 \to -p_2$ and $p_D \to -p_D$

(iii) $UD_1 \rightleftarrows \bar{D}_2\bar{\nu}$, $\bar{U}\bar{D}_1 \rightleftarrows D_2\nu$:

$$|\mathfrak{M}|^2 = (16\pi^2\alpha^2/81)\{[(p_2\cdot p_U)(p_1\cdot p_\nu)]/[2(p_1\cdot p_U) - M_X{}^2]^2 + [(p_1\cdot p_U)(p_2\cdot p_\nu)]/[2(p_2\cdot p_U) + M_X{}^2]^2\}. \tag{A9}$$

(iv) $U\nu \rightleftarrows \bar{D}_1\bar{D}_2$, $\bar{U}\bar{\nu} \rightleftarrows D_1D_2$.

This matrix element is obtained from (A9) by the substitutions $p_1 \to -p_1$ and $p_\nu \to -p_\nu$.

The elementary transition rates W for (i), (ii), (iii), and (iv) are obtained from (A8) and (A9) by

$$W = (s/16)\,|\mathfrak{M}|^2(2\pi)^4\delta^{(4)}(\textstyle\sum p), \tag{A10}$$

where $s = \frac{1}{2}$ for (i) and (iv), and $s = 1$ for (ii) and (iii).

APPENDIX B

After all the assumptions that are discussed in Sec. III have been made, the Boltzmann equations reduce to 12 coupled, ordinary differential equations. As the independent variable we use $z \equiv M_X/T$. The comoving time coordinate and temperature of the universe are related by (2.6), so that $d/dz = (5.8\times10^{17}\ \text{GeV}/M_X)\times(160/g_*)^{1/2}M_X{}^{-1}z(d/dt)$.

Although Compton-type and annihilationlike BNC processes were included in our calculations, we have omitted the terms corresponding to these interactions here. The effects of these interactions were small ($\lesssim$10% of the effects due to BNC processes involving two incoming and two outgoing fermions, which themselves were small) and the number of additional terms was large. The following set of 12 equations are the master equations:

$$U_+'/(zK) = \gamma_D(2X_+ + 3Y_+/2) - \gamma_{ID}(U_+^2 + U_+D_+/2 + U_+L_+/4), \tag{B1}$$

$$D_+'/(zK) = \gamma_D(X_+ + 3Y_+/2) - \gamma_{ID}(U_+D_+/2 + D_+L_+/2 + D_+\nu_+/4), \tag{B2}$$

$$L_+'/(zK) = \gamma_D(3X_+ + 3Y_+/2) - \gamma_{ID}(3U_+L_+/4 + 3D_+L_+/2), \tag{B3}$$

$$\nu_+'/(zK) = 3\gamma_D Y_+ - \gamma_{ID}(3D_+\nu_+/2), \tag{B4}$$

$$X_+'/(zK) = -3\gamma_D X_+ + \gamma_{ID}(3U_+^2/4 + 3D_+L_+/4), \tag{B5}$$

$$Y_+'/(zK) = -3\gamma_D Y_+ + \gamma_{ID}(3U_+D_+/4 + 3U_+L_+/8 + 3D_+\nu_+/8), \tag{B6}$$

$$U'_-/(zK)=-\gamma_D[2X_-+Y_-/2]-\gamma_{ID}[2U_+U_-+(UD)_-/2+(UL)_-/4]-s_1[2U_+U_-+(DL)_-]-s_2[(UD)_-+(UL)_-]$$
$$-s_3[(UD)_-/4+(D\nu)_-/4]-s_4[D_+D_-/2+(U\nu)_-/4]+\epsilon[\gamma_D X_+-\gamma_{ID}D_+L_+/2]\,, \tag{B7}$$

$$D'_-/(zK)=\gamma_D[X_--Y_-/2]-\gamma_{ID}[(UD)_-/2+(DL)_-/2+(D\nu)_-/4]-s_1[U_+U_-+(DL)_-/2]-s_2[(UD)_-/2+(UL)_-/2]$$
$$-s_3[(UD)_-/2+(D\nu)_-/2]-s_4[D_+D_-+(U\nu)_-/2]+\epsilon[\gamma_D X_+/2-\gamma_{ID}U_+^2/4]\,, \tag{B8}$$

$$L'_-/(zK)=\gamma_D[3X_-+3Y_-/2]-\gamma_{ID}[3(UL)_-/4+3(DL)_-/2]-s_1[3U_+U_-+3(DL)_-/2]-s_2[3(UD)_-/2+3(UL)_-/2]$$
$$+\epsilon[3\gamma_D X_+/2-3\gamma_{ID}U_+^2/4]\,, \tag{B9}$$

$$\nu'_-/(zK)=3\gamma_D Y_--\gamma_{ID}[3(D\nu)_-/2]-s_3[3(UD)_-/2+3(D\nu)_-/2]-s_4[3D_+D_-+3(U\nu)_-/2]\,, \tag{B10}$$

$$X'_-/(zK)=-3\gamma_D X_--\gamma_{ID}[3U_+U_-/2-3(DL)_-/4]+\epsilon[3\gamma_{ID}U_+^2/8-3\gamma_{ID}D_+L_+/8]\,, \tag{B11}$$

$$Y'_-/(zK)=-3\gamma_D Y_--\gamma_{ID}[3(UD)_-/4-3(UL)_-/8-3(D\nu)_-/8]\,, \tag{B12}$$

where

$$(Q_1Q_2)_-\equiv Q_{1+}Q_{2-}+Q_{2+}Q_{1-}$$

and

$$K\equiv(2.9\times10^{17}\ \alpha\ \text{GeV}/M_X)(160/g_*)^{1/2}\,.$$

The dimensionless quantities γ_D, γ_{ID} and the s_i are given in terms of the elementary transition rates W (see Appendix A) by

$$\gamma_D=(\alpha M_X)^{-1}\frac{4}{3}\frac{\int\int\int W(X\to\bar U_1\bar U_2)f\,d\Pi_X\,d\Pi_{\bar U_1}d\Pi_{\bar U_2}}{g_X\zeta(3)T^3/\pi^2}\,, \tag{B13}$$

$$\gamma_{ID}=(\alpha M_X)^{-1}\frac{4}{3}\frac{\int\int\int W(U_1U_2\to\bar X)[\exp(p_1/T)+1]^{-1}[\exp(p_2/T)+1]^{-1}d\Pi_{\bar X}d\Pi_{U_1}d\Pi_{U_2}}{g_X\zeta(3)T^3/\pi^2}\,, \tag{B14}$$

$$s_i=(\alpha M_X)^{-1}\frac{8}{3}\frac{\int\int\int\int W_i(1,2\to3,4)[\exp(p_1/T)+1]^{-1}[\exp(p_2/T)+1]^{-1}d\Pi_1d\Pi_2d\Pi_3d\Pi_4}{g_U\zeta(3)T^3/\pi^2}\,, \tag{B15}$$

where the phase-space density of X bosons is $N_X(p,t)=X(t)f(p,T)$ and the reaction i is $UU\to\bar D\bar L(i=1)$, $UD\to\bar U\bar L(i=2)$, $UD\to\bar D\bar\nu(i=3)$, $U\nu\to\bar D\bar D(i=4)$. The rates for decays, inverse decays, and BNC fermion collisions are $\Gamma_D\approx\gamma_D\alpha M_X$, $\Gamma_{ID}\approx\gamma_{ID}\alpha M_X$, and $\Gamma_{BNC}\approx s_i\alpha M_X$.

The quantities γ_D and γ_{ID} can be evaluated,

$$\gamma_D=\frac{2}{3}\frac{M_X\int f\,d\Pi_X}{g_X\zeta(3)T^3/\pi^2}\cong\frac{2}{3}\quad(z\gg1)\,, \tag{B16}$$

$$\gamma_{ID}=\frac{2}{3}\frac{z^2}{\zeta(3)}\int_1^\infty\ln\left(\frac{\cosh\{[u+(u^2-1)^{1/2}]z/4\}}{\cosh\{[u-(u^2-1)^{1/2}]z/4\}}\right)\times\frac{du}{\exp(uz)-1}$$
$$\cong\tfrac{2}{3}(\pi/8)^{1/2}z^{3/2}e^{-z}\quad(z\gg1)\,. \tag{B17}$$

For $z\gg1$ the decay rate is temperature independent since all X bosons are very nonrelativistic, and inverse decays are suppressed since typical fermions are not energetic enough to create an X boson.

The s_i are calculated in the low- ($z\gg1$) and high-temperature ($z\ll1$) limits; the intermediate behavior is obtained by smooth interpolation. Specifically, we found that

$$s_1\cong1.72\alpha z^{-1}(32.3z^{-4})/[1+32.3(M_{\text{eff}}/T)^2z^{-4}]\,, \tag{B18}$$

$$s_2=8.6\alpha z^{-1}(26z^{-4})/[1+26(M_{\text{eff}}/T)^2z^{-4}]\,, \tag{B19}$$

$$s_3\cong3.43\alpha z^{-1}(35.6z^{-4})/[1+35.6(M_{\text{eff}}/T)^2z^{-4}]\,, \tag{B20}$$

$$s_4\cong3.44\alpha z^{-1}(8.84z^{-4})/[1+8.84(M_{\text{eff}}/T)^2z^{-4}]\,. \tag{B21}$$

The quantity M_{eff} is the effective X (or Y) mass when Debye screening and horizon effects are taken into account (see Sec. III). We have smoothly connected the different regimes by using

$$M_{\text{eff}}^2=M_X^2[1+1300\alpha z^{-2}+11z^{-4}(M_X/T_p)^2]\,. \tag{B22}$$

For $z\gg1$, $M_{\text{eff}}^2=M_X^2$; for $36.1\alpha^{1/2}>z>0.1\alpha^{-1/2}\times(M_X/T_p)^2$, $M_{\text{eff}}^2\approx(1300\alpha z^{-2})M_X^2$; and for $z<0.1\alpha^{-1/2}(M_X/T_p)^2$, $M_{\text{eff}}^2\approx(11z^{-4}M_X^2/T_p^2)M_X^2$. The screening and horizon effects are only important at high temperatures ($z\lesssim1$).

[1]G. Gamow, Phys. Rev. 70, 572 (1946); R. A. Alpher, H. A. Bethe, and G. Gamow, *ibid.* 73, 803 (1948).
[2]R. V. Wagoner, W. A. Fowler, and F. Hoyle, Astrophys. J. 148, 3 (1967).
[3]G. Steigman, Annu. Rev. Astron. Astrophys. 14, 339 (1976).
[4]R. L. Omnes, Phys. Rev. Lett. 23, 38 (1969); Astron. Astrophys. 10, 228 (1971).
[5]L. Parker, in *Asymptotic Structure of Space-Time*, edited by F. P. Esposito and L. Witten (Plenum, New York, 1977); A. Sakharov, Zh. Eksp. Teor. Fiz. Pis'ma Red. 5, 32 (1967) [JETP Lett. 5, 24 (1967)]; S. Weinberg, in *Lectures on Particles and Field Theory*, edited by S. Deser and K. Ford (Prentice-Hall, Englewood Cliffs, 1964).
[6]M. Yoshimura, Phys. Rev. Lett. 41, 281 (1978).
[7]A. Ignatiev, N. Krasnikov, V. Kuzmin, and A. Tavkhelidze, Phys. Lett. 76B, 436 (1978).
[8]J. C. Pati and A. Salam, Phys. Rev. D 8, 1240 (1973).
[9]H. Georgi and S. L. Glashow, Phys. Rev. Lett. 32, 438 (1974).
[10]J. Learned, F. Reines, and A. Soni, Phys. Rev. Lett. 43, 907 (1979).
[11]See, e.g., D. V. Nanopoulos and S. Weinberg, Phys. Rev. D 20, 2484 (1979).
[12]S. Dimopoulos and L. Susskind, Phys. Rev. D 18, 4500 (1976).
[13]D. Toussaint, S. B. Treiman, F. Wilczek, and A. Zee, Phys. Rev. D 19, 1036 (1979).
[14]S. Weinberg, Phys. Rev. Lett. 42, 850 (1979).
[15]J. N. Fry, K. A. Olive, and M. S. Turner, following paper, Phys. Rev. D 22, 2977 (1980).
[16]J. N. Fry, K. A. Olive, and M. S. Turner, Enrico Fermi Institute Report No. 80-25 (unpublished).
[17]B. Carr and M. Turner, Enrico Fermi Institute Report No. 80-09 (unpublished).
[18]T. J. Goldman and D. A. Ross, Phys. Lett. 84B, 208 (1979).
[19]S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972).
[20]K. A. Olive, D. N. Schramm, G. Steigman, M. S. Turner, and J. Yang, Enrico Fermi Institute Report No. 80-42, 1980 (unpublished).
[21]J. Ehlers, in *General Relativity and Cosmology, Proceedings of the International School of Physics*, edited by R. K. Sachs (Academic, New York, 1971).
[22]R. V. Wagoner, in *Physical Cosmology*, proceedings of Session XXIX of the Les Houches Summer School, 1979, edited by R. Balian, J. Audouze, and D. N. Schramm (North-Holland, Amsterdam, 1980).
[23]J. D. Bjorken and S. D. Drell, *Relativistic Quantum Fields* (McGraw-Hill, New York, 1965).
[24]J. Ellis and G. Steigman, Phys. Lett. 89B, 186 (1980).
[25]See, e.g., J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1962).
[26]Our understanding of the subtleties involved in the *CP* violation benefited greatly from conversations with E. Kolb and S. Love, and a report given to us by Kolb and Wolfram [E. W. Kolb and S. Wolfram, Phys. Lett. 91B, 217 (1980); Nucl. Phys. B172, 224 (1980)].
[27]These relationships are only valid in the absence of ambient particles or when degeneracy factors are neglected. When these effects are included these relationships must be modified (see, for example, Wagoner, Ref. 22). Since we ignore degeneracy factors these equations are applicable for our purposes.
[28]*Handbook of Mathematical Functions*, edited by M. Abramowitz and I. Stegun (Dover, New York, 1964).
[29]J. B. Hartle and B. L. Hu, Phys. Rev. D 20, 1772 (1979); 21, 2756 (1980); 22, 2091 (1980).
[30]S. B. Treiman and F. Wilczek, Phys. Lett. 95B, 222 (1980).
[31]A. Buras, J. Ellis, M. K. Gaillard, and D. V. Nanopoulos, Nucl. Phys. B135, 66 (1978).