The Development of Baryon Asymmetry in the Early Universe

Edward W. Kolb*

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

and

Stephen Wolfram†

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

*Work supported in part by the National Science Foundation [PHY76-83685].
†Work supported in part by the Department of Energy [EY76-C-03-0068] and a Feynman fellowship.
ABSTRACT

The development of an excess of baryons over antibaryons due to CP and baryon number violating reactions during the very early stages of the big bang is calculated in simple models using the Boltzmann equation.
There are observational and theoretical indications that the local preponderance of baryons over antibaryons extends throughout the universe (at least since the time when the temperature $T \sim 100$ MeV) with an average ratio of baryon to photon densities [1] $n_B/n_\gamma = Y_B \sim 10^{-9}$. If baryon number ($B$) were absolutely conserved in all processes, this small baryon excess must have been present since the beginning of the universe. However, many grand unified gauge models [2] require superheavy particles (typically with masses $m_X \sim 10^{15}$ GeV $\equiv 1$ TeV) which mediate baryon- and lepton-number ($L$) violating interactions. Any direct evidence for these must presumably come from an observation of proton decay. In the standard hot big bang model [1], the temperature $T$ (of light particle species) in the early universe fell with time $t$ according to (taking units such that $\hbar = c = k = 1$) $T \propto \sqrt{m_p/2t}$ where $m_p = (45/8\pi^3)^{1/2} \frac{m_p}{\sqrt{\xi(T)}} = 5 \times 10^3/\sqrt{\xi}$ MeV, and $m_p = G^{-1/2} \sim 10^{19}$ GeV is the Planck mass, while $\xi$ gives the effective number of particle species in equilibrium ($\xi = \frac{1}{2} \frac{7}{16}$ for each ultrarelativistic boson (nondegenerate fermion) spin state). At temperatures $T \geq m_X$, B-violating interactions should have been important, and they should probably have destroyed or at least much diminished any initial baryon excess. (This occurs even when, for example, $B-L$ is absolutely conserved, since then an initial baryon excess would presumably be accompanied by a lepton excess, so as to maintain the accurate charge neutrality of the universe.) It is interesting (and in some models necessary) to postulate that B-violating interactions in the very early universe could give rise to a calculable baryon excess even from an initially symmetrical state. For this to be possible, the rates for reactions producing baryons and antibaryons must differ, and hence the interactions responsible must violate CP invariance. We describe here a simple but general method for calculating $B$ generation in any specific model. We clarify and extend previous estimates [3].
Let $M(i \rightarrow j)$ be the amplitude for transitions from the state $i$ to $j$, and let $\bar{i}$ be the CP conjugate of $i$ (particles $\rightarrow$ antiparticles and spins reversed). Then CPT invariance demands $M(i \rightarrow j) = M(j \rightarrow \bar{i})$, while CP invariance would require $M(i \rightarrow j) = M(\bar{j} \rightarrow i) = M(j \rightarrow i)$. Unitarity [4] (transitions to and from $i$ must occur with total probability 1) demands \[
abla \sum_j |M(i \rightarrow j)|^2 = \sum_j |M(j \rightarrow i)|^2; \] combining this with the constraint of CPT invariance yields (the sum over $j$ includes all states and their antistates) \[
abla \sum_j |M(i \rightarrow j)|^2 = \sum_j |M(\bar{i} \rightarrow j)|^2 = \sum_j |M(j \rightarrow i)|^2. \quad (1) \]

In thermal equilibrium (and in the absence of chemical potentials representing nonzero conserved quantum numbers) all states $j$ of a system with a given energy are equally populated. Then the last equality in (1) shows that transitions from these states (interactions) must produce $i$ and $\bar{i}$ in equal numbers; thus no excess of particles over antiparticles may develop in a system in thermal equilibrium, even if CP is violated. In addition, the first equality in (1) shows that the total cross-sections for destroying particles and antiparticles must be equal. Since in thermal equilibrium no excess of $i$ over $\bar{i}$ may develop, this implies that any initial excess must be destroyed.

The phase space distribution $f_i(p)$ (number per unit cell $d^3p \, d^3x$ [5]) for a species $i$ develops with time (on average) according to a Boltzmann transport equation. A closed system with no external influences obeys Boltzmann's H theorem (which holds regardless of T (i.e., CP) invariance [6]), so that from any initial state the $f_i(p)$ evolve (on average) to their equilibrium forms for which $f_\bar{i}(p) = f_i(p)$, and no baryon excess may survive.
However, in an expanding universe, extra terms must be added to the Boltzmann equations, and if some participating particles are massive [7], a baryon excess may be generated; the relaxation time necessary to destroy the excess often increases faster than the age of the universe [8].

Equation (1) requires that the total rates for processes with particle and antiparticle initial states be equal. CP violation allows the rates for specific conjugate reactions to differ; unitarity nevertheless requires $(T = i(1-S), SS^+ = S^+S = 1)$ [9]:

$$|M(2 \rightarrow 1)|^2 = |M(\bar{2} \rightarrow \bar{1})|^2 = |T_{ij}|^2 - |T_{ji}|^2$$

$$= 2 \text{Im}[\sum_{n} T_{ij}^{\dagger} T_{ji} - \sum_{n} T_{ij}^{\dagger} T_{ji}]^2.$$  \hspace{1cm} (2)

Hence the fractional difference between conjugate rates must be at least $O(\alpha)$ where $\alpha$ is some coupling constant [10]. Moreover, the loop diagrams giving CP violation must allow physical intermediate states $n$. (These loop corrections must usually also be B-violating to give a difference in rates when summed over all final states $\frac{(-)}{3}$ with a given $(-)B$ [11].)

Let $\frac{(-)}{b}$ be an "(anti)baryon" with $B = \frac{\epsilon}{(-) \frac{1}{2}}$. For simplicity we assume here that all particles (including photons) obey Maxwell-Boltzmann statistics and have only one spin state. In our first (very simple) model, we consider CP, B violating $2\leftrightarrow 2$ reactions involving $\frac{(-)}{b}$ and a heavy neutral particle $\phi$; we take their rates to be (this parametrization ensures unitarity and CPT invariance)

$$|M(b \rightarrow \bar{b} \bar{b})|^2 = (1 + \zeta)|M_0|^2/2$$

$$|M(b \rightarrow \phi \phi)|^2 = |M(\phi \phi \rightarrow \bar{b} \bar{b})|^2 = (1 - \zeta)|M_0|^2/2$$

$$|M(\bar{b} \bar{b} \rightarrow b b)|^2 = (1 + \zeta)|M_0|^2/2$$
\[ |M(\bar{b}b + \phi\phi)|^2 = |M(\phi\phi + b\bar{b})|^2 = (1 - \zeta^2) |M_0|^2/2 \]  \hspace{1cm} (3)

where \( \zeta = O(\alpha) \) measures the magnitude of CP violation. The number of a species \( i \) per unit volume \( n_i \equiv \int d^3p/(2\pi)^3 f_i(p) \) decreases with time even without collisions in an expanding universe according to \((R\text{ is the Robertson-Walker scale factor; dots denote time derivatives})\)

\[ \frac{dn_i}{dt} = n_i \frac{d(1/V)}{dt} = - \frac{3\dot{R}}{R} n_i = \frac{3\dot{T}}{T} n_i = - \frac{3T^2}{R^2 n_i}. \]  \hspace{1cm} (4)

The \( n_i \) are also changed by collisions; the (average) time development of the \( \phi \) and baryon number \((n_B' \equiv n_b - n_\gamma)\) densities is given by the Boltzmann equations \((\gamma \equiv n_\gamma / n_\gamma, \text{where } \gamma \text{ is a massless particle; } |M_0|^2 = O(\alpha^2))\)

\[ n_\gamma \frac{\dot{\phi}}{\dot{\phi}} = n_\gamma \frac{\dot{f}_b(p_1) f_b(p_2)}{f_b(p_1) f_b(p_2) |M(b\bar{b} + \phi\phi)|^2} + f_{\bar{b}}(p_1) f_{\bar{b}}(p_2) |M(b\bar{b} + \phi\phi)|^2 - f_\phi(p_1) f_\phi(p_2) (|M(\phi\phi + b\bar{b})|^2)
\]

\[ + \left[ |M(\phi\phi + b\bar{b})|^2 \right] \]  \hspace{1cm} (5a)

\[ n_\gamma \frac{\dot{b}}{\dot{b}} = A_{12} |M(b\bar{b} + \phi\phi)|^2 + |M(\phi\phi + b\bar{b})|^2)
\]

\[ + f_{\bar{b}}(p_1) f_{\bar{b}}(p_2) (2|M(b\bar{b} + \phi\phi)|^2 + |M(b\bar{b} + \phi\phi)|^2)
\]

\[ + f_\phi(p_1) f_\phi(p_2) (|M(\phi\phi + b\bar{b})|^2 - |M(\phi\phi + b\bar{b})|^2) \]  \hspace{1cm} (5b)

where the operator \( A \) represents suitable integration over initial and final state momenta. We assume that the \( b \) undergo baryon-conserving collisions with a frequency much higher than the \( O(\alpha^3) \) rate on which \( n_B \) changes (as is presumably the case in realistic models). They are therefore always in kinetic
equilibrium with the rest of the universe, and hence Maxwell-Boltzmann distributed in phase space:

\[ f_b^{-}(p) \approx \exp[-(E/T)\mu]/T \]

\[ Y_B \equiv (n_b - n_{\bar{b}})/n_\gamma \approx 2 \sinh(\mu/T) \quad . \tag{6} \]

\( \mu \) is a baryon number chemical potential, which is changed only by \( \bar{B} \)-violating processes, and would vanish if chemical equilibrium prevailed. Assuming \( Y_B \ll 1 \), one may use momentum conservation in (5) to write

\[ f_b^{-}(p_1) f_b^{-}(p_2) \approx \exp[-(E_3 + E_4)/T](1 + \frac{Y_B}{\mu^2}) f_\phi^\text{eq}(p_3) f_\phi^\text{eq}(p_4)(1 + \frac{Y_B}{\mu^2}) , \]

where \( f_\phi^\text{eq}(p) = \exp(-E/T) \) is the equilibrium distribution of \( \phi \) at temperature \( T \). The equilibrium \( \phi \) number density \( n_\phi^\text{eq} = T^3/(2\pi^2)(m_\phi/T)^2 K_2(m_\phi/T) \), where \( K_2 \) is a modified Bessel function [12] (as \( m_\phi \to 0 \), \( n_\phi^\text{eq} \to T^3/\pi^2 \); as \( T \to 0 \), \( n_\phi^\text{eq} \to (m_\phi T^2/2\pi)^{3/2} \exp(-m_\phi/T) \). Then substituting the parametrization (3) and performing phase space integrations, (5) becomes

\[ \dot{Y}_\phi = n_\gamma \langle \sigma_o v \rangle \left\{ 2[1 - \frac{\xi + \bar{\xi}}{2}]\langle \gamma_\phi^\text{eq} \rangle^2 - \frac{\gamma_\phi^2}{\mu^2} \right\} , \tag{7a} \]

\[ \dot{Y}_B = n_\gamma \langle \sigma_o v \rangle \left\{ \frac{\xi - \bar{\xi}}{2}\left[ \gamma_\phi^2 - \frac{\gamma_\phi^\text{eq}^2}{\mu^2} \right] - \left[ 3 + \frac{\xi + \bar{\xi}}{2} \right]\langle \gamma_\phi^\text{eq} \rangle^2 Y_B \right\} , \tag{7b} \]

where \( \langle \sigma_o v \rangle \) is the cross-section corresponding to \( |\mu|^2 \) averaged over a flux of incoming particles in equilibrium energy distributions. Equation (7b) exhibits the necessity of deviation from equilibrium for \( \bar{B} \) generation, and the destruction of \( Y_B \) in equilibrium.

We now turn to a slightly more realistic but more complicated model in
which massive particles \( X \) decay to \( b \) with rates \( \gamma_X = 0(a) \)

\[
\begin{align*}
|M(X \rightarrow bb)|^2 &= |M(b\bar{b} + \overline{X})|^2 = (1 + \eta)\gamma_X/2 \\
|M(X \rightarrow \overline{b}\bar{b})|^2 &= |M(b\bar{b} + X)|^2 = (1 - \eta)\gamma_X/2 \\
|M(\overline{X} \rightarrow bb)|^2 &= |M(b\bar{b} + X)|^2 = (1 - \overline{\eta})\gamma_X/2 \\
|M(\overline{X} \rightarrow b\bar{b})|^2 &= |M(b\bar{b} + X)|^2 = (1 + \overline{\eta})\gamma_X/2 
\end{align*}
\]

(8)

Note that if \( X \) decays preferentially to produce \( b \), then CPT invariance implies that \( \overline{b} \) are preferentially destroyed in inverse processes; thus \( X \) decays and inverse decays (DID) alone would generate a net \( B \) even if all particles were in thermal equilibrium, in contravention of the theorem (1) [13]. However, the CP violation parameter \( (\eta - \overline{\eta}) \) is \( 0(a) \), and hence changes in \( n_B \) from DID are of the same order as \( 2+2 \) scattering processes, such as \( bb + \overline{b}\bar{b} \). It will turn out that s-channel exchange of nearly on-shell \( X \) in \( bb \leftrightarrow b\bar{b} \) cancels the DID contribution to \( \dot{\gamma}_B \) so as to recover \( \dot{\gamma}_B = 0 \) in thermal equilibrium. In direct analogy with eqs. (5) and (7), and using the assumption (6), the equation for the evolution of the \( X \) number density \( n^{(-)} \) becomes

\[ \dot{\gamma}_X = -\langle \Gamma_X \rangle \left[ (\gamma_X - \gamma_X^{eq}) - \overline{\eta} \gamma_B \gamma_X^{eq} \right] ; \]

(9)

the corresponding equation for \( \dot{\gamma}_X \) is obtained by charge conjugation \( (Y_X \leftrightarrow Y_{\overline{X}}, Y_B \leftrightarrow -Y_B, \eta \leftrightarrow \overline{\eta}) \). The \( \langle \Gamma_X \rangle \) in (9) is the total \( X \) decay width multiplied by the time dilation factor \( n_X/E_X \) and averaged over the equilibrium \( X \) energy distribution [14]. The baryon concentration evolves according to
\[ \dot{Y}_B = \langle \Gamma_X \rangle \left[ (\eta - \bar{\eta}) Y_{X-} + (\eta + \bar{\eta}) Y_{X+} - 2 Y_B^\text{eq} \right] \]

\[ - \frac{2}{n_Y} \lambda_1 \left\{ f_X^\text{eq}(p_1 + p_2) \left[ |M'(bb + \bar{b}\bar{b})|^2 - |M'(\bar{b}\bar{b} \to bb)|^2 \right] \right\} \]

\[ - 2 Y_B^\text{eq} \left\{ f_X^\text{eq}(p_1 + p_2) \left[ |M'(bb \to \bar{b}\bar{b})|^2 + |M'(\bar{b}\bar{b} \to bb)|^2 \right] \right\} \]

(10)

where the first term is from DID (and does not separately vanish when \( Y(-) = Y_{X-}^\text{eq} \)), while the second two terms arise from \( 2 \to 2 \) scatterings. The DID term accounts for sequential inverse decay and decay processes involving real \( X \): these are therefore subtracted from the true \( 2 \to 2 \) scattering terms by writing \( |M'(i+j)|^2 = |M(i+j)|^2 - |M_{\text{RIX}}(i+j)|^2 \), where \( M_{\text{RIX}}(i+j) \) is the amplitude for \( i \to j \) due to on-shell s channel \( X \) exchange. In the narrow \( X \) width approximation, \( |M_{\text{RIX}}(i+j)|^2 \approx |M(i \to X)|^2 |M(X \to j)|^2 / \Gamma_X \); the presence of the \( \Gamma_X \) denominator renders it \( O(\alpha) \). According to the theorem (1), the CP violating difference of total rates \( |M(bb + \bar{b}\bar{b})|^2 - |M(\bar{b}\bar{b} + bb)|^2 = O(\alpha^3) \). Hence \( |M'(bb \to \bar{b}\bar{b})|^2 - |M'(\bar{b}\bar{b} \to bb)|^2 = |M_{\text{RIX}}(\bar{b}\bar{b} \to bb)|^2 - |M_{\text{RIX}}(bb \to \bar{b}\bar{b})|^2 + O(\alpha^3) \) = \( O(\alpha^2) \), and the second term in eq. (10) becomes \( -2 \langle \Gamma_X \rangle (\eta - \bar{\eta}) Y_{X-}^\text{eq} \), thereby elegantly cancelling the first term in thermal equilibrium. Finally, therefore,

\[ \dot{Y}_B = \langle \Gamma_X \rangle \left\{ (\eta - \bar{\eta}) \left[ (Y_{X-} + Y_{X+})/2 \right] - Y_{X-}^\text{eq} \right\} \]

\[ + (\eta + \bar{\eta}) (Y_{X-} - Y_{X+}/2) \]

\[ - 2 Y_B \left\{ \langle \Gamma_X \rangle Y_{X-}^\text{eq} + \eta \langle \omega' (bb \to \bar{b}\bar{b}) + \omega' (\bar{b}\bar{b} \to bb) \rangle \right\} \]

(11)

The differential equations (10) and (11) must now be solved with the initial condition \( Y_X(t=0) = Y_{X-}^\text{eq}(0) \), and possibly an initial baryon density.
Figure 1 shows the solutions with guesses for parameters based on the SU(5) model [2] ($m_X = 10^{15}$ GeV and $10^{14}$ GeV; $\alpha = 1/40$ (vector decays), or $10^{-3}$ (scalar decays)). If all $X$ initially in thermal equilibrium decayed with no back reactions, the $Y_B$ generated would be simply $\eta \bar{\eta}$. For small $\alpha$ or large $m_X/m_p$, this upper limit is approached. (At small $x \equiv m_X/T$, series solution of equations (10) and (11) gives $(Y_X + Y_\bar{X})/2 = 1 - ax^5/20$; $(Y_X - Y_\bar{X})/2 = (\eta - \bar{\eta})ax^8/160$; $Y_B = (\eta - \bar{\eta})ax^5/20$, where $a = m_p\Gamma_X/m_X^2$.) For $T \ll m_X$, baryon number is destroyed by $2 \to 2$ reactions with $\sigma \sim a^2 T^2/m_X^4$ roughly like $Y_B(T) \sim \exp[\alpha^2 m_p T^3/m_X^4]$ [15], so that $Y_B \rightarrow$ constant as $T \rightarrow 0$, but if $m_X$ is small, the final $Y_B$ is much diminished from its value at higher $T$. The $Y_B$ generated is always roughly linearly proportional to $\eta \bar{\eta}$, but is a sensitive function of $m_X/m_p$ and $\alpha$; for realistic values of these parameters, a numerical solution is probably essential.

According to equation (11), any baryon excess existing at the Planck time $t_p = 1/m_p$ should be diminished by inverse decays at $T \gg m_X$ so that $Y_B(t)/Y_B(t_p) \sim \exp[-\alpha m_p T^2]$; any initial $Y_B$ should be reduced by a factor $\sim \exp[-m_p/m_X]$ before CP violating processes can generate $Y_B$ at $T \lesssim m_X$. B-violating $2 \to 2$ scatterings at temperatures $m_p > T > m_X$ should reduce an initial $Y_B$ by a factor $\sim \exp[-m_p \int^{m_X} T \langle \nu \nu \rangle dT]$. One might expect that $\langle \nu \nu \rangle \sim \alpha^2/m_X^2$ at high energies due to t-channel vector $X$ exchange; however, the effective $\langle \nu \nu \rangle$ presumably relevant for the Boltzmann equation is rather $\langle \nu \nu_{\text{eff}} \rangle \sim \alpha^2/\lambda_D^2$ where the Debye screening length $\lambda_D \sim [\sqrt{32\alpha} T]^{-1}$. In this approximation $2 \to 2$ and higher multiplicity collisions are probably no more effective at destroying an initial $Y_B$ than are inverse decays.

We conclude therefore that B-violating reactions in the very early universe might well destroy any initial baryon number existing around the Planck time ($1/m_p$), requiring subsequent B and CP-violating interactions.
to generate the observed baryon asymmetry. The methods described here [16]
allow a calculation of the resulting baryon excess in any specific model;
the simple examples considered suggest that the observed $Y_B$ should place
stringent constraints on parameters of the model.

We are grateful to many people for discussions, including A. D. Dolgov,
S. Frautschi, William A. Fowler, G. C. Fox, T. J. Goldman, S. E. Koonin,
and D. L. Tubbs. The work of E.W.K. was supported in part by the National
Science Foundation [PHY76-83685], and of S.W. by the Department of Energy
[EY76-C-03-0068] and a Feynman fellowship.
References


5. We assume homogeneity and isotropy, so that \( f(p, x) = f(p) = f(p) \).


7. It is not necessary that these participate directly in \( B \) violating reactions.

8. In the simple models discussed below, this phenomenon occurs if the universe is homogeneous and always cools faster than \( T \sim m_p/(t m_p)^{1/5} \); in practice the quark excess will always be contained in baryons where their probability for collisions remains constant rather than falling as in a homogeneous expanding universe.
9. This constraint applies only if no initial or final particles may mix with their antiparticles (as in the $K^0$ system). CP-violating mixing requires a difference $M(i \rightarrow j) - M(i \rightarrow \bar{j}) \neq 0$ in amplitudes rather than rates.

10. Regardless of perturbation theory, CP violation is asymptotically suppressed by powers of $\log(s)$, where $s$ is the invariant mass of the initial state.


13. This rather relevant point has also been noticed by A. D. Dolgov and Ya. B. Zeldovich (ref. 3), but was apparently neglected elsewhere.

14. Strictly, $m_\chi/E_\chi$ should be averaged separately for the various terms of (9); if $\chi$ is in kinetic equilibrium, however, the averages are equal. Note that we have implicitly assumed all produced and decaying $\chi$ to be exactly on their mass shells. However, particularly at high $T$, the mean $\chi$ collision time $\ll 1/\Gamma_\chi$, so that the $\chi$ resonance is collision broadened, and produced or decaying $\chi$ may be far off shell. The $m_\chi/E_\chi$ factor for inverse decays essentially arises from the fact that the incoming particles must subtend a sufficiently small angle to have invariant mass $m_\chi$; if produced $\chi$ are far off shell, the $m_\chi/E_\chi$ in DID should disappear.

15. If $T \sim m_p/(\Gamma_\chi m_\chi)^{1/5}$ as in footnote 8, then $Y_B \sim \exp(-1/T)$; the universe expands sufficiently slowly for $Y_B$ to relax to zero.

17. Results depend only on $m_X$ through the dimensionless combination $m_X/m_p$; here we take $\xi = 100$ in the definition of $m_p$. Note that inhomogeneities in the early universe may be manifest in different expansion rates and hence different effective $\xi$ for different regions. The final $Y_B$ produced could vary considerably between the regions.
The development of baryon number density (solid curves) as a function of inverse temperature in the model of eq. (11) for various choices of parameters (unless otherwise indicated, $\alpha = 1/40$ and $m_\chi = 1 \text{ MeV} \pm 10^{15} \text{ GeV}$ [17]). The dashed and dotted curves give $(Y_\chi + Y_\nu)/2$ and $(Y_\chi - Y_\nu)/2$, respectively. In all cases we have taken the CP violation parameter, $\eta - \bar{\eta} = 10^{-6}$. 

Figure Caption