FERMILAB-CONF-87/130-A

The Origin of Baryons in the Universe [1]

EDWARD W. KOLB

NASA/Fermilab Astrophysics Center
Fermi National Accelerator Laboratory
Batavia, Illinois 60510
USA

ABSTRACT

In this talk I will review some recent work on the origin of the baryon asymmetry in the Universe. An asymmetry in the number of baryons compared to the number of antibaryons is crucial for nucleons to survive annihilation in the early moments of the big bang to later form nuclei, stars, planets, people, etc. The standard picture of the origin of the baryon asymmetry in the decay of a supermassive boson will be presented in detail. Some recent developments regarding effects at the electroweak phase transition, as well as possible low-temperature scenarios will be reviewed.

This symposium *The Origin and Distribution of Elements* concerns one of the most fundamental questions in science, namely "what is the Universe made of?" There are many possible answers to the question. An astronomer might answer that the Universe is made of stars or galaxies, some space dirt, and a small insignificant fraction in the form of planets. If pressed for a more fundamental answer, the astronomer might admit that stars are made of three types of elements, hydrogen, helium and metals. The abbreviated periodic table of the astronomer is often adequate for most purposes in astronomy. If pressed still further for a more fundamental answer, the astronomer might say "go see a chemist." A chemist might point to the actual periodic table and say that the Universe is made of atoms. A Universe of atoms as the fundamental building blocks might be the Universe of

[1]Talk presented at the American Chemical Society meeting in New Orleans, Louisiana, August 31 – September 4, 1987

the chemist, but there is a more fundamental answer. A physicist might say that the universe consists of electrons and nuclei. When asked the inevitable question the physicist will answer that electrons aren't made of anything, they are fundamental, but the nuclei are made of nucleons (neutrons and protons), which in turn are made of three quarks and some gluons added for good measure to keep the nucleons and nuclei bound.

Of course the answers of the astronomer, the chemist, and the physicist are all correct in some sense, but no single answer is the complete picture. In order to understand the Universe it is necessary to understand the structure of stars and galaxies, which requires an understanding of the fundamental particles down the entire chain from atoms to quarks. Often the best answer depends upon the scale of the problem.

Now that we have established what the Universe is made of, it is necessary to understand why it has the structure it does. This symposium is concerned with the origin of the elements. In order to understand the origin of the elements it is necessary to understand the origin of the building blocks of the elements, the nucleons. In order to understand the origin of the nucleons, it is necessary to consult that noble and rare breed of scientist, the cosmologist. The question I will address is why are there nucleons at all. A possible answer to the question would be to invoke the anthropic principle and say that if there weren't nucleons in the Universe, people wouldn't be around to hold a symposium on the subject. In my opinion, even if said with a British accent, the anthropic principle explains nothing in this instance. However, it is possible to understand the existence of nucleons on the basis of physical law. Before describing such efforts, I will quantify the question in terms of a number known as the "Baryon Number of the Universe."

It is possible to assign to quarks and antiquarks a conserved quantum number known as baryon number. The baryon number of a quark is $1/3$ and the baryon number of an antiquark is $-1/3$. Nucleons, which consist of three quarks, have a baryon number of $+1$. The baryon number is conserved in all laboratory experiments. For experiments at low energy, $E \lesssim m_N \simeq 1$ GeV, where m_N is the mass of the nucleon, the conservation of baryon number is equivalent to the conservation of mass number.

It is possible to estimate the average number density of baryons (plus antibaryons) in the Universe from the observed total mass density of the universe. It is convenient to express present energy densities in terms of a critical density, $\rho_C = 3H_0^2/8\pi G$, where H_0 is the Hubble constant (the present expansion rate of the Universe) and G is Newton's constant. In terms of $h = H_0/100$ km^{-1}sec^{-1}Mpc^{-1}, $\rho_C = 1.88 \times 10^{-29}h^2$g cm^{-3}. In terms of $\Omega_+ = (\rho_b + \rho_{\bar{b}})/\rho_C$, $n_b + n_{\bar{b}} = 1.12 \times 10^{-5}\Omega_+ \; h^2$cm^{-3}. If there are no antibaryons in the Universe (as argued below), then $n_b + n_{\bar{b}} = n_b$, and from the limit $\Omega_+ \leq 1$, it is possible to estimate the baryon charge density, $n_B = n_b - n_{\bar{b}}$. It is possible to create baryons and antibaryons so long as the difference in baryons and antibaryons is conserved. n_B should scale in expansion as inverse of the volume of the Universe. The entropy density $s = (\rho + p)/T$ also scales in the same way, so the baryon number of the Universe, $B = n_B/s$ should be constant so long as the baryon number is locally conserved and the entropy remains constant. The entropy density today is $s \simeq 2800$ cm^{-3}, so

$$B = \frac{n_B}{s} = 4 \times 10^{-9}\Omega_+ h^2. \tag{1}$$

If there is no antimatter in the Universe, $\Omega_+ = \Omega_N$, the present contribution to the total mass density due to nucleons. Limits on the TOTAL mass density of the Universe give $0.01 < \Omega_N < 1$, while limits from primordial nucleosynthesis give the much more restrictive range $0.01 < \Omega_N h^2 < 0.04$.

As mentioned above, our Universe seems to have a baryon asymmetry, more baryons than antibaryons, i.e., more matter than antimatter. Antimatter is rare on earth. It exists in "large" quantities only in the antiproton accumulators at Fermilab and CERN. Antimatter is also rare in the solar system. The fact that Neil Armstrong survived his "one small step" is evidence that the moon is matter. Planetary probes have visited seven of the nine (ten?) planets, and have shown that the solar system is matter.

Cosmic rays provide a sample of the entire galaxy (at least). Antiprotons are seen in cosmic rays at about the 10^{-4} level compared to protons. These cosmic rays are usually assumed to be secondary particles, and not primary particles from an antimatter source. The flux of antimatter nuclei is also below the 10^{-4} level compared to nuclei, and there is no clean event that signals detection of an antinucleus. Cosmic rays are solid evidence that there is a galactic asymmetry between baryons and antibaryons, and that this asymmetry is maximal, i.e., everything is baryons, and there aren't any antibaryons in the galaxy.

Our evidence on larger scales is somewhat more uncertain. There are clusters of galaxies containing intercluster gas. If both matter galaxies and antimatter galaxies exist in the same cluster there would result a large amount of annihilation which would contribute a large γ-ray flux. The absence of such a large γ-ray flux is evidence that clusters of galaxies are either all baryons or all antibaryons. There is basically no information on scales larger than clusters of galaxies ($\simeq 10^{14}\ M_\odot$).

If antimatter exists in appreciable quantities, it must be separated on scales larger than about $10^{14}\ M_\odot$. However this separation must be done before $T \simeq 1$ GeV, or the nucleons and antinucleons would have annihilated below acceptable limits. If there are equal numbers of baryons and antibaryons the evolution of the baryon number density, n_B, is described by the equation

$$\dot{n}_B + 3Hn_B = ((n_B^{eq})^2 - n_B^2)\langle\sigma_A|v|\rangle \tag{2}$$

where H is the expansion rate of the Universe, n_B^{eq} is the equilibrium number density of baryons, and $\langle\sigma_A v\rangle$ is the thermal average of the annihilation rate times relative velocity. The baryon number density tracks its equilibrium value, $n_B^{eq} \simeq (m_N T)^{3/2}\exp(-m_N/T)$, until the baryon density is so low that annihilation effectively ceases. For baryons, this happens when the density of baryons relative to the entropy density is about 10^{-19}.

There seems to be convincing arguments for a universal baryon asymmetry. Details of the arguments for a baryon asymmetry and for a symmetric Universe can be found in Ref. 1. The most reasonable conclusion is that at $T \geq 1$ GeV, there was an asymmetry between the number of baryons and the number of antibaryons. This asymmetry is characterized by the baryon number discussed above, $B = 4 \times 10^{-9}\Omega_N h^2$. Although the baryon asymmetry is maximal today (i.e., no antimatter) at $T \geq 1$ GeV, the temperature was high enough to create $N\bar{N}$ pairs, and $B \simeq 10^{-9}$ means that for every billion antibaryons there were a billion and one baryons.

The answer to the question why are there are baryons in the Universe has been replaced with another question, why is there a baryon asymmetry. The goal of this talk is to explain how such a curious (but crucial) number could arise in a Universe with symmetric, $B = 0$, or even better, random, initial conditions.

There are three basic ingredients necessary to generate a non-zero B from an initial symmetric state.[2] • *Baryon Number Violation:* There must obviously be a violation of baryon number. If baryon number is conserved in all interactions, the present baryon asymmetry must simply reflect the initial conditions. • *C and CP Violation:* Since B is odd under C and CP, they both must be violated to generate a non-zero B. • *Non-Equilibrium Conditions:* In chemical equilibrium

particle		final state	branching ratio	B
X	$\rightarrow$	qq	r	2/3
X	$\rightarrow$	$\bar{q}\bar{l}$	$1-r$	-1/3
$\bar{X}$	$\rightarrow$	$\bar{q}\bar{q}$	$\bar{r}$	-2/3
$\bar{X}$	$\rightarrow$	$q\bar{l}$	$1-\bar{r}$	1/3

Table 1: Final states and branching ratios

the entropy is maximal when the chemical potentials associated with all non-conserved quantum numbers vanish. If baryon number is not conserved, and the Universe ever obtains chemical equilibrium, B must be zero.

To illustrate the simple model, consider a particle X which decays to final states with different baryon number (hence violating baryon number) with branching ratios given in Table 1.[3,4,5,6] Note that CP is violated if $r \neq \bar{r}$, but CPT is conserved, since the total rate for X decay is the same as the total rate for $\bar{X}$ decay.

Imagine a box containing equal amounts of X and $\bar{X}$. The baryon number produced by the decay of the X's is proportional to $B_X = r(2/3) + (1-r)(-1/3)$, and the baryon number produced by the $\bar{X}$'s is proportional to $B_{\bar{X}} = \bar{r}(-2/3) + (1-\bar{r})(1/3)$. The net asymmetry produced by an equal amount of X and $\bar{X}$'s is proportional to $\epsilon = B_X + B_{\bar{X}} = r - \bar{r}$. The baryon number vanishes, of course, if CP is conserved ($r = \bar{r}$). If there are no further baryon number violating reactions (equilibrium not maintained) a net baryon asymmetry will result. This extremely simple picture illustrates the basic idea.

Now, consider the possibility of B, C, and CP violation, and non-equilibrium conditions in the early Universe.

• *B Violation*: The existence of baryon number violation seems to be a generic feature of Grand Unified Theories (GUTs). If strong and electroweak interactions are unified, quarks and leptons typically appear as elements of a common irreducible representation of the gauge group, and gauge bosons can mediate interactions that violate B. The limit on the stability of the proton, $\tau \geq 10^{31}$ years, implies that these bosons should have masses in excess of 10^{14}GeV or so. The weaker coupling allows Higgs bosons that mediate baryon-number violation to have a somewhat smaller mass. In both cases the large mass of the intermediate bosons is responsible for the feebleness of baryon-number violation today. This suppression should have been overcome at the extremely high temperatures

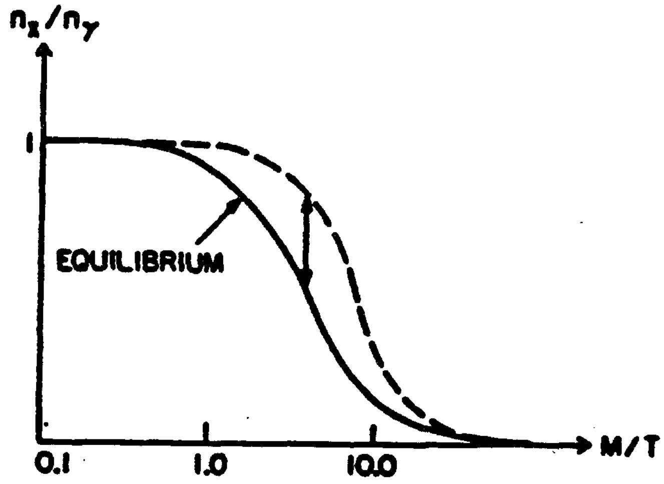

Figure 1: Departure from equilibrium denoted as the arrow between equilibrium (solid) and actual (dashed) ratio of n_X/n_γ

With a similar calculation for the other decay mode

$$\epsilon_X = \frac{4}{\Gamma_X} \mathrm{Im} I_{XY} \mathrm{Im}(g_1^* g_2 g_3^* g_4) \left[(B_{i_4} - B_{i_3}) - (B_{i_2} - B_{i_1}) \right]. \tag{7}$$

For Y decay $\epsilon_Y = -\epsilon_X$, with $X \leftrightarrow Y$ everywhere.

There are several lessons to be learned from this tedious exercise. There must be *two* baryon number violating bosons, with masses greater than the sum of the masses in the internal loops. CP violation is manifest as complex coupling constants. The X and Y particles in the above example must not be degenerate, or the baryon number produced in X decay will cancel the baryon number produced in Y decay.

• *Non-Equilibrium Conditions*: The non-equilibrium conditions are provided by the expansion of the Universe. If the expansion rate is faster than particle interaction rates, non-equilibrium can result. Assume that at the Planck time X and $\bar{X}$ are present in equilibrium, $n_X = n_{\bar{X}} = n_\gamma$. In LTE $n_X = n_{\bar{X}} = n_\gamma$ for $T \geq m_X$, and $\ll n_\gamma$ for $T \leq m_X$. However LTE will be obtained only if the X interaction rates are greater than H. If LTE is not maintained, n_X will remain equal to n_γ and there will be an excess of X relative to the equilibrium abundance. This is illustrated in Fig. 1.

The conditions necessary for a departure from equilibrium can be quantified. The decay rate of the X, denoted by Γ_D, the inverse decay rate (i.e., X produc-

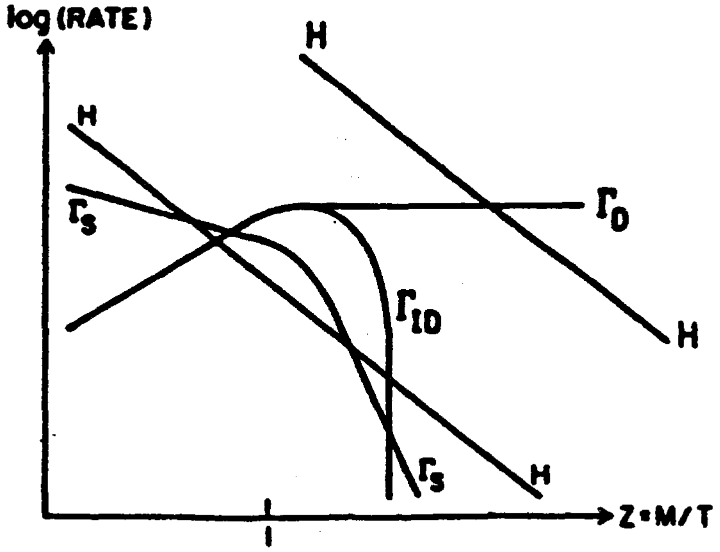

Figure 2: Rates in the early Universe

tion), denoted as Γ_{ID}, and the two-body scattering rate, denoted as Γ_S, and the expansion rate, denoted as H, are given by

$$\begin{aligned}
\Gamma_D &= \alpha m_X \begin{cases} m_X/T & m_X \leq T \\ 1 & m_X \geq T \end{cases} \\
\Gamma_{ID} &= \Gamma_D \exp(-m_X/T) \\
\Gamma_S &= n\sigma \simeq T^3 \alpha^2 \frac{T^2}{(T^2+m_X^2)^2} \\
H &\sim g_*^{1/2} T^2/m_{Pl}.
\end{aligned} \tag{8}$$

A comparison of the rates are given in Fig. 2. Note that all the reaction rates depend upon m_X, while H is independent of m_X. It is necessary to determine what sets the scale of H relative to the other rates. As shown in Fig. 2, H can be either large or small compared to the rates at $T = m_X$.

The relevant quantity to determine if X will be over abundant is the ratio of the decay rate to the expansion rate at $T = m_X$, given by

$$K \equiv (\Gamma_D/H)_{T=m_X} = \frac{\alpha m_{Pl}}{g_*^{1/2} m_X}. \tag{9}$$

If $K \ll 1$, equilibrium will not be maintained and the X will become over abundant. In this limit the X just drifts along and then decays. In the limit of pure drift and decay (the British limit) $B = g_*^{-1}\epsilon$. If $K \gg 1$, the X will be in good causal contact, there will be no departure from equilibrium, and baryogenesis will be thwarted. Exactly what is meant by "$\gg$" and "$\ll$" can be found by a simple model. If X is a gauge boson, then $\alpha = g^2_{\rm GUT}/4\pi \simeq 1/45$, and $K = 7\times 10^{15}{\rm GeV}/M$. If X is not a gauge boson, the effective α may be smaller, and the corresponding K may be smaller.

Detailed calculations of the complete rate equations in Grand Unified Theories based upon the gauge groups SU(5) and SO(10) have been done.[7] The scenario for the generation of a baryon asymmetry indeed can work in GUTs. The GUT scenario has become the standard model for Baryogenesis.

There have been two recent developments that might change the standard GUT scenario. The first has to do with non-perturbative effects at the electroweak

The θ-vacuum structure in electroweak theories leads to the anomalous non-conservation of baryon number.[8] There are equivalent vacua associated with different fermion numbers. Transitions between the different vacua are accompanied by a change in the baryon number. This process is usually associated with instantons that describe the tunneling between different θ-vacua. Since the process is inherently non-perturbative, the rate for baryon number non-conservation is proportional to $\exp(-4\pi/\alpha_W)$. This quantum tunneling is unimportant today, and was certainly unimportant in the early Universe.

At finite temperature the transitions between different vacua can be driven by finite temperature effects. Kuzmin, Rubakov, and Shaposhnikov (KRS) have recently shown that the tunneling at finite temperature may be appreciable.[9] In their analysis they considered an SU_2 gauge theory, which is a good approximation to the Weinberg-Salam model in the limit $\sin^2\theta_W \to 0$, α_W fixed. The Lagrangian for the SU_2 model with gauge coupling g_W is given by

$$\mathcal{L} = \frac{1}{2}D_\mu\phi^a D^\mu\phi^a - \frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} - \frac{1}{4}\lambda(\phi^a\phi^a - \sigma^2/2)^2 + \mathcal{L}_{\rm fermions}. \tag{10}$$

The basic point is that fluctuations in A_μ and ϕ caused by finite temperature effects cause transitions between different θ-vacua, with associated baryon number violation. The calculation of the transition rate is related to the calculation of the fate of the false vacuum at finite temperature. However, in this case the calculation is much more complicated. The first approximation made is to ignore the effect of the fermion fields, and to only consider A_μ and ϕ as the dynamical fields. The validity of this approximation will be discussed later. The second approximation is to replace the finite temperature "bounce" action by the free energy at the maximum of the barrier between adjacent θ-vacua.

The maximum free energy in the transition is dominated by a static configuration, given in the $A_0 = 0$ gauge by $(r^2 = \vec{x}^2)$

$$A^i_{cl} = \frac{i}{g}\frac{\varepsilon_{ijk}x_j\tau_k}{r^2}f(\xi) \qquad \phi = \frac{\sigma}{\sqrt{2}}\frac{i\vec{\tau}\cdot\vec{x}}{r}\begin{bmatrix} 0 \\ 1 \end{bmatrix}h(\xi), \tag{11}$$

where $\vec{\tau}$ are the Pauli spin matrices, and $f(0) = h(0) = 0$, $f(\infty) = h(\infty) = 1$ are functions of the dimensionless parameter $\xi = r/r_0 = rg_W\sigma$. The free energy of this configuration is given by $F \simeq 2M_W/\alpha_W$, where $M_W = g_W\sigma/2$.

With the KRS assumption that the action for the tunneling rate at finite temperature is given by F, the change in the baryon number is $\dot{B} = -CBT\exp(-F/T)$, where C is a dimensionless constant of order unity, and the overall factor of T appears on dimensional grounds. Since CP is conserved, B cannot be generated by the transitions. In the early Universe g, λ, and σ are a function of temperature. The temperature dependence of g and λ is only proportional to $\ln T$ and will be ignored. The temperature dependence of σ around the critical temperature is much more important. At $T \geq T_c$, $\sigma = 0$, and as the temperature passes through the critical temperature, $\sigma \to \sigma_0 = 250$ GeV.

The rate for baryon number non-conservation is given by $\Gamma_{\Delta B} = CT\exp(-F/T)$, which is greater than the expansion rate for $T \geq T^* \simeq 200$ GeV. This would be a long enough period to eradicate any baryon number. Since there is no anomaly in $B - L$,the transitions conserve $B - L$, so more precisely, it would eradicate any baryon number with zero projection on $B - L$. The two possible solutions would be either to generate an asymmetry in $B - L$ in the same manner an asymmetry in B is generated, [2] or to generate the baryon asymmetry after the electroweak transition.

The potential impact of the KRS calculation certainly warrants more detailed calculations of the finite-temperature induced tunneling between θ-vacua. In particular, one approximation made by KRS that might be questioned is the neglect of the plasma effects on the transition. The classical field configuration are spatially large. The characteristic size of the configurations are several r_0. Within this size are $10^6 - 10^4$ particles with weak charge. Since the configuration of A_μ is pure magnetic and the plasma is a magnetic conductor, there are no plasma screening

[2] In many GUTs (in particular SO_{10}) this is easy to accomplish.

effects in the *static* configuration. However in the evolution to and from the static configuration there are time-dependent gauge fields and A_μ must have an electric component which *will* be affected by screening. The problem of screening can be addressed in a U_1 theory by calculating how long it takes to establish a U_1 field configuration with a characteristic size of several r_0. By detailed balance the time it takes to establish this unstable field configuration is the same as the time it takes for the field configuration to decay. In the limit that the conductivity of the Universe is infinite, the time is infinite, i.e., in a perfect conductor currents in the plasma are induced to oppose the establishment of the field. Of course, the Universe is not a perfect conductor, and the time for the establishment of the coherent field, Γ_A, will be proportional to the conductivity.

It is possible to map the problem at hand into one treated before. Consider the field configuration as a particle (called the Sphaleron) with mass $m = F$ and decay width $\Gamma = \Gamma_A$. Since baryon number is violated in the decay (and CP is conserved), the change in the baryon number is $\dot{n_B} + 3Hn_B = -n_B n_S^{eq}\Gamma_S$, where as usual $n_S^{eq} \simeq (mT)^{3/2}\exp(-m/T)$. Determination of Γ_S will allow a determination of the efficiency of the process in damping the baryon asymmetry.

A second development that may have implications for baryogenesis is the search for low-temperature scenarios. In the conventional GUT scenario, baryogenesis occurs at a temperature $T \simeq 10^{14}$GeV. New Inflation is a possible explanation to the origin of the large entropy in the Universe.[10] If inflation produces a large amount of entropy, it would dilute any baryon number that might exist before inflation and baryogenesis must occur after (or at least in the latter stages of) inflation. Exactly how inflation is implemented is unclear at present. However it seems that the temperature of the Universe after inflation is low, i.e. less than the GUT scale. The standard decaying particle model for baryogenesis is impossible if one starts at a temperature much less than the mass of the decaying particle. If the reheat temperature is low another mechanism for baryogenesis must be found.

One example of such a mechanism has recently been proposed by Dimopoulos and Hall.[11] They point out that in the supersymmetric extension of the standard SU(3)$\times$SU(2)$\times$U(1) model with explicit R-parity violation it is possible to have relatively light baryon number violating particles. In this model the proton is stable in the presence of light baryon-number violating bosons because the proton is lighter than all fermionic superpartners.

In conclusion, it is possible to understand the presence of baryons in the Universe as the result of the operation of physical laws. Whether the standard GUT scenario proves correct, or perhaps is modified by non-perturbative effects at the electroweak phase transition, or perhaps replaced by a low-temperature scenario, baryogenesis is an example of the importance of microphysics in understanding the large-scale structure of the Universe.

References

1. The evidence for the asymmetry is reviewed by G. Steigman, *Ann. Rev. Ast. Astrophys.* **14**, 339 (1976). The view that the Universe is symmetric is reviewed by F. Stecker, *Ann. NY Acad. Sci.* **375**, 69 (1981).

2. A. D. Sakharov, *JETP Lett.* **5**, 24 (1967).

3. E. W. Kolb and S. Wolfram, *Nucl. Phys.* **B172**, 224 (1980).

4. E. W. Kolb and M. S. Turner, *Ann. Rev. Nucl. Part. Sci.* **33**, 645 (1983).

5. J. N. Fry, K. A. Olive, and M. S. Turner, *Phys. Rev. D.* **22**, 2953 (1980).

6. M. Yoshimura, *Phys. Rev. Lett.* **41**, 281 (1978); D. Toussaint, S. B. Trieman, F. Wilczek, and A. Zee, *Phys. Rev. D* **19**, 1036 (1979); S. Dimopoulos and L. Susskind, *Phys. Rev. D* **18**, 4500 (1978); J. Ellis, M. K. Gaillard, and D. V. Nanopoulos, *Phys. Lett.* **B80**, 360 (1979); S. Weinberg, *Phys. Rev. Lett.* **42**, 850 (1979).

7. J. A. Harvey, E. W. Kolb, D. B. Reiss, and S. Wolfram, *Nucl. Phys.* **B201**, 16 (1982).

8. G. 't Hooft, *Phys. Rev. Lett.* **37**, 8 (1976).

9. V. A. Kuzmin, V. A. Rubakov, and M. E. Shaposhnikov, *Phys. Lett.* **155B**, 36 (1985).

10. For a review of inflation, see M. S. Turner, in *Proceedings of the 1984 Cargese School*, ed. J. Audouze and J. Tran Thanh Van.

11. S. Dimopoulos and L. Hall, LBL preprint LBL-23516 (1987).