Gravitational Baryogenesis

Hooman Davoudiasl, Ryuichiro Kitano, Graham D. Kribs, Hitoshi Murayama,* and Paul J. Steinhardt†
School of Natural Sciences, Institute for Advanced Study, Princeton, New Jersey 08540, USA
(Received 4 May 2004; published 10 November 2004)

We show that a gravitational interaction between the derivative of the Ricci scalar curvature and the baryon-number current dynamically breaks CPT in an expanding Universe and, combined with baryon-number-violating interactions, can drive the Universe towards an equilibrium baryon asymmetry that is observationally acceptable.

DOI: 10.1103/PhysRevLett.93.201301 PACS numbers: 98.80.Cq, 04.50.+h, 11.30.Er

The successful predictions of big-bang nucleosynthesis (BBN) [1], highly precise measurements of the cosmic microwave background [2], and the absence of intense radiation from matter-antimatter annihilation [3] all indicate that the Universe contains an excess of matter over antimatter. Numerically, the baryon-to-entropy ratio is $n_B/s = 9.2^{+0.6}_{-0.4} \times 10^{-11}$. What remains a mystery is how the baryon asymmetry was generated.

The contemporary view is that the baryon asymmetry is generated dynamically as the Universe expands and cools. Sakharov [4] argued that three conditions are necessary: (1) baryon number (B) nonconserving interactions; (2) C and CP violation; and (3) a departure from thermal equilibrium. To satisfy the latter two conditions, the conventional approach has been to introduce interactions that violate C and CP *in vacuo* and a period when the Universe is out of thermal equilibrium.

In this Letter, we propose a mechanism that generates an observationally acceptable B-asymmetry while maintaining thermal equilibrium. The key ingredient is a CP-violating interaction between the derivative of the Ricci scalar curvature $\mathcal{R}$ and the B-current J^μ:

$$\frac{1}{M_*^2}\int d^4x\sqrt{-g}(\partial_\mu \mathcal{R})J^\mu, \tag{1}$$

where M_* is the cutoff scale of the effective theory. It is not necessary that J^μ be the B-current; any current that leads to net $B - L$ charge in equilibrium (L is the lepton number) so that the asymmetry will not be wiped out by the electroweak anomaly [5] is sufficient for our purpose. It is natural to expect such an operator in the low-energy effective field theory of quantum gravity if M_* is of order the reduced Planck scale $M_P = (8\pi G_N)^{-1/2} \simeq 2.4 \times 10^{18}$ GeV. We also note that it can be obtained in supergravity theories from a higher dimensional operator $[\mathrm{Im}\mathcal{D}^2\overline{\mathcal{D}}^2 Q^* e^V Q]_D$ in the notation of Refs. [6,7].

The interaction in Eq. (1) violates CP and is CPT conserving *in vacuo*. However, this interaction dynamically breaks CPT in an expanding Universe and biases the energetics in favor of causing an asymmetry between particles and antiparticles.

To generate a B-asymmetry using the interaction in (1), we also require that there be B-violating processes in thermal equilibrium. We denote the temperature at which B-violation decouples by T_D. Given these ingredients, the B-asymmetry in our setup is generated as follows. In an expanding Universe, where $\mathcal{R} \sim H^2$ and $\dot{\mathcal{R}}$ are nonzero (where a dot means time derivative), the interaction in Eq. (1) gives opposite sign energy contributions that differ for particle versus antiparticle, and thereby dynamically violates CPT. This modifies thermal equilibrium distributions in a similar fashion to a chemical potential $\mu \sim \pm\dot{\mathcal{R}}/M_*^2$, driving the Universe towards nonzero *equilibrium* B-asymmetry via the B-violating interactions. Once the temperature drops below T_D, as the Universe expands and cools, the asymmetry can no longer change and is frozen. Then a net asymmetry remains:

$$\frac{n_B}{s} \approx \frac{\dot{\mathcal{R}}}{M_*^2 T}\bigg|_{T_D}. \tag{2}$$

Our approach is closely related to "spontaneous baryogenesis" [8], which relies on the derivative coupling between a spatially uniform scalar field and the B-current, $(\partial_\mu\varphi)J^\mu$. With φ, though, the construction is considerably more complicated. The required scalar has to be added by hand, whereas the term in Eq. (1) is expected to be present in an effective theory of gravity. The initial conditions for φ must be separately specified and justified: φ must be forced to evolve homogeneously in one direction versus the other to produce an asymmetry and must be spatially uniform. In contrast, the time evolution of $\mathcal{R} \propto H^2$ is required in a cosmological background and it is highly spatially uniform because the Universe is highly homogeneous. The oscillation of φ around its minimum is also a complication because the average $\dot{\varphi}$ is zero, tending to cancel the asymmetry [9], whereas the mean value of $\dot{\mathcal{R}} \sim H^3$ does not vanish.

To produce B-asymmetry by the gravitational interaction in Eq. (1), several factors have to be considered. For a constant equation of state w, where w is the ratio of the pressure p to the energy density ρ, $\mathcal{R}$ is proportional to $(1 - 3w)$, and its time derivative is given by

$$\dot{\mathcal{R}} = -(1-3w)\frac{\dot{\rho}}{M_P^2} = \sqrt{3}(1-3w)(1+w)\frac{\rho^{3/2}}{M_P^3}. \tag{3}$$

We will examine B generation for some cosmologically important values of w.

We start from the radiation-dominated era following inflation when $w \approx 1/3$. If w were equal to precisely 1/3, then the right-hand side of Eq. (3) would vanish and there would be no effect. However, $w = 1/3$ only applies in the limit of exact conformal invariance, $T^\mu_\mu = 0$. In practice, interactions among massless particles lead to running coupling constants, and hence, the trace anomaly that makes $T^\mu_\mu \propto \beta(g) F^{\mu\nu} F_{\mu\nu} \neq 0$. The thermodynamic potential of a plasma of an $SU(N_c)$ gauge theory, with coupling g and N_f flavors, leads to [10]

$$1 - 3w = \frac{5}{6\pi^2} \frac{g^4}{(4\pi)^2} \frac{(N_c + \frac{5}{4}N_f)(\frac{11}{3}N_c - \frac{2}{3}N_f)}{2 + \frac{7}{2}[N_c N_f/(N_c^2 - 1)]} \tag{4}$$

up to $O(g^5)$ corrections, where the last factor in the numerator is the beta function coefficient. Typical gauge groups and matter content at very high energies can easily yield $1 - 3w \sim 10^{-2}$–10^{-1}. There may also be mass thresholds that lead to conformal violation. Then, Eq. (2) gives

$$\frac{n_B}{s} \approx (1 - 3w) \frac{T_D^5}{M_*^2 M_P^3}. \tag{5}$$

The upper bound on tensor mode fluctuations constrains the inflationary scale to be $M_I \leq 3.3 \times 10^{16}$ GeV [11], and obviously for this scenario $T_D < T_{\rm RD} < M_I$, where $T_{\rm RD}$ is the temperature at which the Universe becomes radiation-dominated (*i.e.*, the reheat temperature in this case). It is remarkable that the asymmetry can be sufficiently large even for $M_* \simeq M_P$ if $T_D \simeq M_I$, implying that tensor mode fluctuations should soon be observed.

A second case of cosmic relevance is $w = 0$. This corresponds to the matter domination epoch which characterizes, for example, conventional perturbative reheating via a scalar field $\phi_{\rm osc}$, as it oscillates around the minimum of a quadratic potential. During the oscillation phase, $\rho \propto a^{-3}$, $a \propto t^{2/3}$, and $\phi_{\rm osc}$ decays at a rate Γ into radiation, whose energy density becomes equal to that of the scalar field when $H \simeq \Gamma \simeq T_{\rm RD}^2/M_P$. Therefore,

$$\rho_{osc} \simeq T_{\rm RD}^4 \left(\frac{a_{RD}}{a}\right)^3, \tag{6}$$

$$\rho_R \simeq T_{\rm RD}^4 \left(\frac{a_{RD}}{a}\right)^{3/2}. \tag{7}$$

The latter equation suggests $T \simeq T_{\rm RD}(a_{\rm RD}/a)^{3/8}$. At the time of decoupling $T_D > T_{\rm RD}$ and Eq. (2) gives

$$\frac{n_B}{s} \simeq \frac{T_D^{11}}{M_*^2 M_P^3 T_{\rm RD}^6}. \tag{8}$$

This asymmetry, however, is diluted by a continuous production of entropy. The dilution factor is given by $(T_{\rm RD}/T_D)^5$ and hence, the final asymmetry is

$$\frac{n_B}{s} \simeq \frac{T_D^6}{M_*^2 M_P^3 T_{\rm RD}}. \tag{9}$$

Within the linear approximation made in Eq. (2), the initial asymmetry in Eq. (8) cannot be larger than $O(1)$. Therefore, $T_{\rm RD}$ cannot be smaller than about $10^{-2} T_D$ to obtain the correct B-asymmetry, giving an upper limit

$$\frac{n_B}{s} \lesssim 10^2 \frac{T_D^5}{M_*^2 M_P^3}. \tag{10}$$

This result is 3–4 orders of magnitude enhanced relative to Eq. (5) and allows for $T_{\rm RD} \simeq 10^{14}$ GeV.

A third possibility is to generate the B-asymmetry while a nonthermal component with $w > 1/3$ dominates the Universe. The nonthermal energy component decreases more rapidly than radiation, so there is no need for this component to decay into additional radiation and produce more entropy in order to enter the radiation-dominated epoch. Hence, once n_B/s is set at T_D, it remains constant. The absence of further dilution opens up the range of allowed parameters. One example where this can occur is in the ekpyrotic [12] or cyclic [13] Universe where the kinetic-energy density of a scalar field ϕ dominates immediately after the bang over a smaller, subdominant radiation component. The scalar field is the modulus that describes the interbrane separation. Another possibility is that the inflaton ϕ falls down a steep potential at the end of inflation and shoots out as a massless scalar field [14]. More general w can be realized, for example, by a coherent oscillation of ϕ about the minimum of the potential $V(\phi) = \lambda \phi^{2N}/M_P^{2N-4}$ with a coupling constant λ. This form of $V(\phi)$ yields $w = (N - 1)/(N + 1)$, where $1/3 < w \leq 1$ for $N > 2$ [15]. Such a potential is natural in supersymmetry using a discrete symmetry, and it is easy to verify that soft supersymmetry breaking effects do not spoil its desired behavior.

The Universe is initially dominated by $\rho_\phi \sim a^{-3(1+w)}$, which decreases faster than a subdominant radiation component $\rho_R \sim a^{-4}$. The ϕ-dominated Universe expands as $a \propto t^{2/[3(1+w)]}$, while the temperature drops as $T(t) = T_{\rm RD}[a_{\rm RD}/a(t)] = T_{\rm RD}(t_{\rm RD}/t)^{2/[3(1+w)]}$. Therefore,

$$\rho_\phi \simeq T_{\rm RD}^4 \left(\frac{a_{\rm RD}}{a}\right)^{3(1+w)}, \tag{11}$$

$$\rho_R \simeq T_{\rm RD}^4 \left(\frac{a_{\rm RD}}{a}\right)^4. \tag{12}$$

Combining these relations, we find the asymmetry

$$\frac{n_B}{s} \sim \frac{T_D^8}{M_*^2 M_P^3 T_{\rm RD}^3} \left(\frac{T_{\rm RD}}{T_D}\right)^{9(1-w)/2}. \tag{13}$$

As in this scenario we can make T_D significantly greater than $T_{\rm RD}$ and since $w > 1/3$, the B-asymmetry in Eq. (13) can be significantly enhanced relative to that in Eq. (5) by a factor $(T_D/T_{\rm RD})^{3(3w-1)/2}$. Henceforth we focus on the case $w > 1/3$.

Next, we discuss the origin of the B-violating interaction that is necessary for any of the baryogenesis scenarios considered here. To keep the discussion general, we assume that B-violating interactions are generated by an operator $\mathcal{O}_B$ of mass dimension $D = 4 + n$. The rate of such interactions is given by $\Gamma_B = T^{2n+1}/M_B^{2n}$, where M_B is the mass scale associated with $\mathcal{O}_B$. Decoupling of B-violating processes occurs at $T \sim T_D$, when Γ_B falls below $H \sim (T_{\rm RD}^2/M_P)(T/T_{\rm RD})^{3(1+w)/2}$. The decoupling temperature T_D is then estimated to be

$$T_D \sim T_{\rm RD}\left(\frac{M_B^{2n}}{M_P T_{\rm RD}^{2n-1}}\right)^{2/(4n-3w-1)}. \qquad (14)$$

When $w = 1$ and $n = 1$, Eq. (14) is not applicable since the interactions with dimension five operators, such as those responsible for the neutrino mass seesaw mechanism [16], are either in thermal equilibrium all the time or never. Using Eqs. (13) and (14), we can identify the values of $(T_{\rm RD}, M_B)$ that generate $n_B/s \sim 10^{-10}$. Assuming $n = 1$, Fig. 1 shows the result for various values of w. The plots for $n > 1$ are similar.

In supergravity theories, gravitino production places severe bounds on the highest temperature $T_{\rm max}$ attained in the early Universe. Apart from order unity coefficients, the gravitino abundance $Y_{3/2}$ is given by

$$Y_{3/2} = \frac{n_{3/2}}{s} \sim 10^{-4}\frac{T_{\rm RD}}{M_P}\left(\frac{T_{\rm max}}{T_{\rm RD}}\right)^{3(1-w)/2}, \qquad (15)$$

where we have used Boltzmann's equation $dn_{3/2}/dt + 3Hn_{3/2} = \sigma_{eff} n_R^2$, with $\sigma_{\rm eff} \sim 1/M_P^2$ and $n_R \sim T^3$. We set $T_{\rm max} = T_D$ for numerical estimates below. The bounds come from two constraints: (i) ensuring that the products of BBN will not be dissociated by late gravitino decays and (ii) avoiding overclosure of the Universe by gravitinos, if they are the lightest supersymmetric particles (LSP's), or by the LSP's that would be produced in gravitino decays. Here, we assume that the gravitino has a mass $m_{3/2} \gtrsim 100$ TeV and decays rapidly—before BBN—as expected in anomaly mediated supersymmetry breaking scenarios [17]. Then, there is only constraint (ii) that the LSP's produced in gravitino decay not overclose the Universe, which requires $Y_{3/2} < 4 \times 10^{-12}(100\ {\rm GeV}/{\rm m_{LSP}})$. This range is indicated in Fig. 1 in dark gray shading, assuming ${\rm m_{LSP}} = 100$ GeV. On the other hand, the entire range shown in Fig. 1 is allowed if R parity is violated so that the LSP decays before BBN, if the LSP is much lighter than 100 GeV, if the gravitino is lighter than keV, or if there is no supersymmetry at all.

It is fascinating that the correct B-asymmetry can be obtained without overproducing gravitinos. The typical energy scale for the B violation in this scenario is about 10^{14} GeV as seen in Fig. 1. For instance, this energy scale is consistent with what is expected for the Majorana mass of a right-handed neutrino in the seesaw mechanism [16] which violates B in conjunction with electroweak anomaly effects [5]. This would predict nearly degenerate Majorana neutrinos with masses above 0.1 eV and neutrinoless double-beta decay ($0\nu\beta\beta$) at a rate observable in near-future experiments [18]. In comparison, thermal

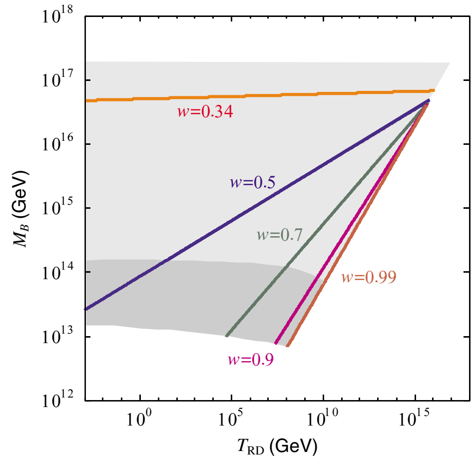

FIG. 1 (color online). The range of $T_{\rm RD}$, M_B, and w that can generate an observationally acceptable B-asymmetry, assuming a dimension five ($n = 1$) B-violating operator *and no significant entropy production after decoupling*, is shown in dark gray and light gray. In supersymmetric theories, where the gravitino decays rapidly to an LSP with ${\rm m_{LSP}} \approx 100$ GeV, the allowed region is restricted to the dark gray region only to avoid overclosure of the Universe by LSP's. Both light and dark gray regions are allowed if the ${\rm m_{LSP}} \ll 100$ GeV, if the LSP decays, if the gravitino is lighter than keV, or if there is no supersymmetry.

leptogenesis places a tight upper bound on the neutrino mass $m_\nu \leq 0.11$ eV [19], and observation of $0\nu\beta\beta$ at larger m_ν would exclude that scenario. Another possibility for B-violation is the $D = 7$ operator $W = (UDD) \times (UDD)/M_B^3$, which satisfies all experimental constraints provided $M_B \gtrsim 100$ TeV.

Higher M_B can be accommodated if there is significant entropy production below $T_{\rm RD}$. The regions above the curves in Fig. 1 produce a larger B-asymmetry and we can afford such an entropy production, which in turn also dilutes the gravitinos. Then, a much wider region of the parameter space becomes available. The predictions for a $D = 5$ operator, with entropy production below T_D, are presented in Fig. 2.

Another concern in the case of inflation is that gravitinos may be produced by quantum fluctuations in the de Sitter phase. If the gravitino mass can be ignored during the inflation, its coupling to the background is conformal, and no gravitinos are produced [20]. Depending on the details of the model of inflation, the gravitino mass may be enhanced and hence its production [21] for helicity $\pm 3/2$ states. We have checked that for certain models, *e.g.*, the supersymmetric hybrid inflation model [22], the gravitino constraint is easy to satisfy [21]. Note that the helicity $\pm 1/2$ states are actually "eaten" inflatino which decays quickly and hence are harmless [23,24]. Yet another concern is that gravitinos may be overproduced by brane collisions in an ekpyrotic or cyclic model, but this issue lies beyond the scope of this Letter.

Finally, we point out that M_* does not have to be as high as the Planck scale. In fact, if the B-violation is soft,

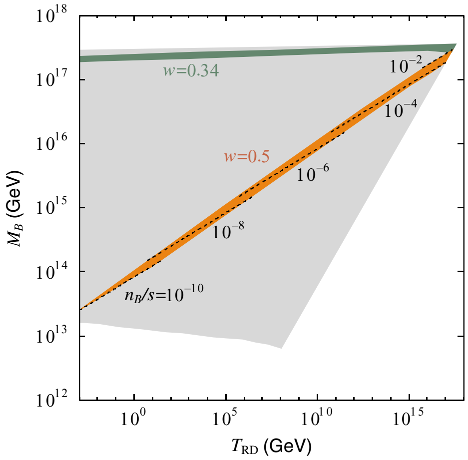

FIG. 2 (color online). The range of $T_{\rm RD}$, M_B, and w that can generate an observationally acceptable B-asymmetry and avoid overclosure of the Universe by stable LSP's expands to cover essentially both the light and dark gray regions in Fig. 1 *if we allow for entropy to be produced after the B-violating interactions and gravitinos decouple.* For each w, the allowed line in Fig. 1 becomes an allowed strip (shown for two cases in the Figure) because we now include the possibility that the B-asymmetry may be overproduced at decoupling to some degree and brought to its proper value by entropy production after decoupling. For the case $w = 0.5$, we indicate along the strip the values of n_B/s before dilution. The additional entropy reduces the gravitino/LSP density, thereby opening up the allowed range of $T_{\rm RD}$ and M_B.

for example, by the Majorana mass M_R of the right-handed neutrino, the operator in Eq. (1) does not cause any unitarity violation up to the Planck scale even if $M_*^2 \simeq M_R M_P$. This is an interesting possibility that we will not pursue further in this Letter.

In summary, we have presented a new framework for baryogenesis where CP violation lies in a gravitational interaction. The expansion of the Universe promotes the microscopic CP violation to a dynamical violation of CPT that shifts the relative energies of particles and antiparticles. A CP-conserving B-violating interaction in equilibrium can then create the asymmetry that gets frozen when the B-violating interaction decouples. We have shown that it is possible to obtain the correct magnitude of the B-asymmetry in many different cosmological scenarios: radiation-dominated ($w \approx 1/3$), matter-dominated ($w = 0$), and kinetic-energy dominated ($1/3 < w \leq 1$) Universes. In the last case, in particular, one can obtain the correct B-asymmetry while keeping the gravitino abundance low enough to avoid overclosure by its decay products. We envision that particle physics beyond the standard model can provide the required B violation at an energy scale above 10^{13} GeV, giving a whole new class of realistic models of baryogenesis.

We thank D. Gross, P. Langacker, J. Lykken, J. Maldacena, L. Susskind, and E. Witten for discussions. R. K. is the Marvin L. Goldberger Member and G. D. K. is a Frank and Peggy Taplin Member at the Institute for Advanced Study (IAS). P. J. S. is Keck Distinguished Visiting Professor at the IAS with support from the Wm.-Keck Foundation and the Monell Foundation. H. M. was supported by the IAS, funds for Natural Sciences. This work was also supported in part by the DOE under Contracts No. DE-FG02-90ER40542 and No. DE-AC03-76SF00098, and in part by NSF Grant No. PHY-0098840.

*On leave of absence from Department of Physics, University of CA, Berkeley, CA 94720
†On leave of absence from Department of Physics, Princeton University, Princeton, NJ 08544

[1] S. Burles, K. M. Nollett, and M. S. Turner, Phys. Rev. D **63**, 063512 (2001).
[2] C. Bennett *et al.*, Astrophys. J. Suppl. Ser. **148**, 1 (2003); *ibid.* **148**, 175 (2003).
[3] A. G. Cohen, A. De Rujula, and S. L. Glashow, Astrophys. J. **495**, 539 (1998).
[4] A. D. Sakharov, JETP Lett. **5**, 24 (1967).
[5] V. A. Kuzmin, V. A. Rubakov, and M. E. Shaposhnikov, Phys. Lett. B **155**, 36 (1985).
[6] T. Kugo and S. Uehara, Nucl. Phys. B **222**, 125 (1983).
[7] T. Kugo and S. Uehara, Prog. Theor. Phys. **73**, 235 (1985).
[8] A. G. Cohen and D. B. Kaplan, Phys. Lett. B **199**, 251 (1987).
[9] A. Dolgov, K. Freese, R. Rangarajan, and M. Srednicki, Phys. Rev. D **56**, 6155 (1997).
[10] K. Kajantie, M. Laine, K. Rummukainen, and Y. Schroder, Phys. Rev. D **67**, 105008 (2003).
[11] H. V. Peiris *et al.*, Astrophys. J. Suppl. Ser. **148**, 213 (2003).
[12] J. Khoury, B. A. Ovrut, P. J. Steinhardt, and N. Turok, Phys. Rev. D **64**, 123522 (2001).
[13] P. J. Steinhardt and N. Turok, Science **296**, 1436 (2002); Phys. Rev. D **65**, 126003 (2002).
[14] P. J. E. Peebles and A. Vilenkin, Phys. Rev. D **59**, 063505 (1999).
[15] M. S. Turner, Phys. Rev. D **28**, 1243 (1983).
[16] T. Yanagida, in *Proceedings of the Workshop on the Unified Theory and the Baryon Number in the Universe*, edited by O. Sawada and A. Sugamoto, p. 95 (KEK Report No. 79-18, 1979); M. Gell-Mann, P. Ramond, and R. Slansky, in *Supergravity*, edited by P. van Nieuwenhuizen and D. Z. Freedman (North-Holland, Amsterdam, 1979), p. 315.
[17] L. Randall and R. Sundrum, Nucl. Phys. B **557**, 79 (1999); G. F. Giudice, M. A. Luty, H. Murayama, and R. Rattazzi, J. High Energy Phys. 12 (1998) 027.
[18] G. Gratta, Int. J. Mod. Phys. A **19**, 1155 (2004).
[19] W. Buchmuller, P. Di Bari, and M. Plumacher, Nucl. Phys. B **665**, 445 (2003).
[20] L. Parker, Phys. Rev. **183**, 1057 (1969).
[21] M. Lemoine, Phys. Rev. D **60**, 103522 (1999).
[22] G. R. Dvali, Q. Shafi, and R. K. Schaefer, Phys. Rev. Lett. **73**, 1886 (1994).
[23] H. P. Nilles, M. Peloso, and L. Sorbo, J. High Energy Phys. 04 (2001) 004.
[24] P. B. Greene, K. Kadota, and H. Murayama, Phys. Rev. D **68**, 043502 (2003).