The origin of matter in the universe: A brief review

Marcelo Gleiser

Citation: AIP Conference Proceedings **400**, 497 (1997); doi: 10.1063/1.53231
View online: https://doi.org/10.1063/1.53231
View Table of Contents: http://aip.scitation.org/toc/apc/400/1
Published by the American Institute of Physics

The Origin of Matter in the Universe: A Brief Review

Marcelo Gleiser

Department of Physics and Astronomy, Dartmouth College[1]
Hanover, NH 03755

Abstract. In this talk I briefly review the main ideas and challenges involved in the computation of the observed baryonic excess in the Universe.

I THE SAKHAROV CONDITIONS AND GUT BARYOGENESIS

Given that the observational evidence is for a Universe with a primordial baryon asymmetry [1,2], we have two choices; either this asymmetry is the result of an initial condition, or it was attained through dynamical processes that took place in the early Universe. In 1967, just a couple of years after the discovery of the microwave background radiation, Sakharov wrote a ground-breaking work in which he appealed to the drastic environment of the early stages of the hot big-bang model to spell out the 3 conditions for dynamically generating the baryon asymmetry of the Universe [3]. Here they are, with some modifications:
i) Baryon number violating interactions: Clearly, if we are to generate any excess baryons, our model must have interactions which violate baryon number. However, the same interactions also produce antibaryons at the same rate. We need a second condition;
ii) C and CP violating interactions: Combined violation of charge conjugation (C) and charge conjugation combined with parity (CP) can provide a bias to enhance the production of baryons over antibaryons. However, in thermal equilibrium $n_{\rm b} = n_{\bar{\rm b}}$, and any asymmetry would be wiped out. We need a third condition;
iii) Departure from thermal equilibrium: Nonequilibrium conditions guarantee that the phase-space density of baryons and antibaryons will not be the same.

[1] NSF Presidential Faculty Fellow.

CP400, *First Latin American Symposium on High Energy Physics/ VII Mexican School of Particles and Fields*, edited by D'Olivo/Klein-Kreisler/Mendez

Hence, provided there is no entropy production later on, the net ratio n_B/s will remain constant.

Given the above conditions, we have to search for the particle physics models that both satisfy them and are capable of generating the correct asymmetry. The first models that attempted to compute the baryon asymmetry dynamically were Grand Unified Theory (GUT) models [4]. A typical mechanism of GUT baryogenesis is known as the "out-of-equilibrium decay scenario"; one insures that the heavy X bosons have a long enough lifetime so that their inverse decays go out of equilibrium as they are still abundant. Baryon number is produced by the free decay of the heavy Xs, as the inverse rate is shut off.

Interesting as they are, GUT models of baryogenesis have serious obstacles to overcome. Here I mention only the obstacle related to electroweak scale phenomena. The vacuum manifold of the electroweak model exhibits a very rich structure, with degenerate minima separated by energy barriers (in field configuration space). Different minima have different baryon (and lepton) number, with the net difference between two minima being given by the number of families. Thus, for the standard model, each jump between two adjacent minima leads to the creation of 3 baryons and 3 leptons, with net $B-L$ conservation and $B+L$ violation. At $T=0$, tunneling between adjacent minima is mediated by instantons, and, as shown by 't Hooft [5], the tunneling rate is suppressed by the weak coupling constant ($\Gamma \sim e^{-4\pi/\alpha_W} \sim 10^{-170}$). That is why the proton is stable. However, as pointed out by Kuzmin, Rubakov, and Shaposhnikov, at finite temperatures ($T \sim 100$ GeV), one could hop over the barrier, tremendously enhancing the rate of baryon number violation [6]. The height of the barrier is given by the action of an unstable static solution of the field equations known as the sphaleron [7].

Being a thermal process, the rate of baryon number violation is controlled by the energy of the sphaleron configuration, $\Gamma \sim \exp[-\beta E_S]$, with $E_S \simeq M_W/\alpha_W$, where M_W is the W-boson mass. Note that $M_W/\alpha_W = \langle\phi\rangle/g$, where $\langle\phi\rangle$ is the vacuum expectation value of the Higgs field. For temperatures above the critical temperature for electroweak symmetry restoration, it has been shown that sphaleron processes are not exponentially suppressed, with the rate being roughly $\Gamma \sim (\alpha_W T)^4$ [8]. Even though this opens the possibility of generating the baryonic asymmetry at the electroweak scale, it is bad news for GUT baryogenesis. Unless the original GUT model was $B-L$ conserving, any net baryon number generated then would be brought to zero by the efficient anomalous electroweak processes. There are several alternative models for baryogenesis invoking more or less exotic physics. The interested reader is directed to the review by Olive, listed in Ref. 1. I now move on to discuss the promises and challenges of electroweak baryogenesis.

II ELECTROWEAK BARYOGENESIS

As pointed out above, temperature effects can lead to efficient baryon number violation at the electroweak scale. Can the other two Sakharov conditions be satisfied in the early Universe so that the observed baryon number could be generated during the electroweak phase transition? The short answer is that in principle yes, but probably not in the context of the minimal standard model. Let us first see why it is possible to satisfy all conditions for baryogenesis in the context of the standard model.

Departure from thermal equilibrium is obtained by invoking a first order phase transition. After summing over matter and gauge fields, one obtains a temperature corrected effective potential for the magnitude of the Higgs field, ϕ. The potential describes two phases, the symmetric phase with $\langle\phi\rangle = 0$ and massless gauge and matter fields, and the broken-symmetric phase with $\langle\phi\rangle = \phi_+(T)$, with massive gauge and matter fields. The loop contributions from the gauge fields generate a cubic term in the effective potential, which creates a barrier separating the two phases. This result depends on a perturbative evaluation of the effective potential, which presents problems for large Higgs masses as I will discuss later. At 1-loop, the potential can be written as [9]

$$V_{\rm EW}(\phi,T) = D\left(T^2 - T_2^2\right)\phi^2 - ET\phi^3 + \frac{1}{4}\lambda_T\phi^4, \tag{1}$$

where the constants D and E are given by $D = \left[6(M_W/\sigma)^2 + 3(M_Z/\sigma)^2 + 6(M_T/\sigma)^2\right]/24 \sim 0.17$, and $E = \left[6(M_W/\sigma)^3 + 3(M_Z/\sigma)^3\right]/12\pi \sim 0.01$, where I used, $M_W = 80.6$ GeV, $M_Z = 91.2$ GeV, $M_T = 174$ GeV [10], and $\sigma = 246$ GeV. The (lengthy) expression for λ_T, the temperature corrected Higgs self-coupling, can be found in Ref. 9. At the critical temperature, $T_C = T_2/\sqrt{1 - E^2/\lambda_T D}$, the minima have the same free energy, $V_{\rm EW}(\phi_+, T_C) = V_{\rm EW}(0, T_C)$. As $E \to 0$, $T_C \to T_2$ and the transition is second order. Since E and D are fixed, the strength of the transition is controlled by the value of the Higgs mass, or λ.

Assuming that the above potential (or something close to it) correctly describes the two phases, as the Universe cools belows T_C the symmetric phase becomes metastable and will decay by nucleation of bubbles of the broken-symmetric phase which will grow and percolate completing the transition. Departure from equilibrium will occur in the expanding bubble walls. This scenario relies on the assumption that the transition is strong enough so that the usual homogeneous nucleation mechanism correctly describes the approach to equilibrium. As I will discuss later, this may not be the case for "weak" transitions. For now, we forget this problem and move on to briefly examine how to generate the baryonic asymmetry with expanding bubbles.

The last condition for generating baryon number is C and CP violation. It is known that C and CP violation are present in the standard model. However, the CP violation from the Kobayashi-Maskawa (KM) phase is too small to generate the required baryon asymmetry. Even though the debate is still going on, efficient baryogenesis within the standard model is a remote possibility.

For many, this is enough motivation to go beyond the standard model in search of extensions which have an enhanced CP violation built in. Several models have been proposed so far, although the simplest invoke either more generations of massive fermions, or multiple massive Higgs doublets with additional CP violation in this sector of the theory. Instead of looking into all models in detail, I will just briefly describe the essential ingredients common to most models.

The transition is assumed to proceed by bubble nucleation. Outside the bubbles the Universe is in the symmetric phase, and baryon number violation is occurring at the rate $\Gamma \sim (\alpha_{\mathrm{W}} T)^4$. Inside the bubble the Universe is in the broken symmetric phase and the rate of baryon number violation is $\Gamma \sim \exp[-\beta E_{\mathrm{S}}]$. Since we want any net excess baryon number to be preserved in the broken phase, we must shut off the sphaleron rate inside the bubble. This imposes a constraint on the strength of the phase transition, as $E_{\mathrm{S}} \simeq \langle\phi(T)\rangle/g$; that is, we must have a large "jump" in the vacuum expectation value of ϕ during the transition, $\langle\phi(T)\rangle/T \geq 1$, as shown by Shaposhnikov [11].

Inside the bubble wall the fields are far from equilibrium and there is CP violation, and thus a net asymmetry can be induced by the moving wall. In practice, computations are complicated by several factors, such as the dependence on the net asymmetry on the bubble velocity and on its thickness [12]. Different charge transport mechanisms based on leptons as opposed to quarks have been proposed, which enhance the net baryonic asymmetry produced [13]. However, the basic picture is that as matter traverses the moving wall an asymmetry is produced. And since baryon number violation is suppressed inside the bubble, a net asymmetry survives in the broken phase. Even though no compelling model exists at present, and several open questions related to the complicated nonequilibrium dynamics remain, it is fair to say that the correct baryon asymmetry may have been generated during the electroweak phase transition, possibly in some extension of the standard model. However, I would like to stress that this conclusion has two crucial assumptions built in it; that we know how to compute the effective potential reliably, and that the transition is strong enough to proceed by bubble nucleation. In the next Section I briefly discuss some of the issues involved and how they may be concealing interesting new physics.

III CHALLENGES TO ELECTROWEAK BARYOGENESIS

A The Effective Potential

A crucial ingredient in the computation of the net baryon number generated during the electroweak phase transition is the effective potential. In order to trust our predictions, we must be able to compute it reliably. However, it is well known that perturbation theory is bound to fail due to severe infrared problems. It is easy to see why this happens. At finite temperatures, the loop expansion parameter involving gauge fields is $g^2T/M_{\mathbf{gauge}}$. Since $M_{\mathbf{gauge}} = g\langle\phi\rangle$, in the neighborhood of $\langle\phi\rangle = 0$ the expansion diverges. This behavior can be improved by summing over ring, or daisy, diagrams [14].

Another problem that appears in the evaluation of the effective potential is due to loop corrections involving the Higgs boson. For second order phase transitions, the vanishing of the effective potential's curvature at the critical temperature leads to the existence of critical phenomena characterized by diverging correlation lengths. Even though there is no infrared-stable fixed point for first order transitions, for large Higgs masses the transition is weak enough to induce large fluctuations about equilibrium; the mean-field estimate for the correlation length $\xi(T) = M^{-1}(T)$ is certainly innacurate. This behavior has led some authors [15,16] to invoke ε-expansion methods to deal with the infrared divergences. Another alternative is to go to the computer and study the equilibrium properties of the standard model on the lattice [17]. Recent results are encouraging inasmuch as they seem to be consistent with perturbative results in the broken phase for fairly small Higgs masses. Furthermore, they indicate how the transition becomes weaker for large values of the Higgs mass, $M_{\mathrm{H}} \geq 60$ GeV.

B Weak vs. Strong First Order Transitions

In order to avoid the erasure of the produced net baryon number inside the broken-symmetric phase, the sphaleron rate must be suppressed within the bubble. As mentioned earlier, this amounts to imposing a large enough "jump" on the vacuum expectation value of ϕ during the transition. In other words, the transition cannot be too weakly first order. But what does it mean, really, to be "weakly" or "strongly" first order?

This is a very important point which must not be overlooked (although it often is!); the vacuum decay formalism used for the computation of nucleation rates relies on a semi-classical expansion of the effective action. That is, we assume we start at a *homogeneous* phase of false vacuum, and evaluate the rate by summing over small amplitude fluctuations about the metastable state [18]. This approximation must break down for weak enough transitions, when we expect large fluctuations to be present within the metastable phase. An explicit example of this breakdown was recently discussed, where the extra free energy available due to the presence of large-amplitude fluctuations was incorporated into the computation of the decay rate [19].

In Ref. 15, it was suggested that weak transitions may evolve by a different mechanism, characterized by substantial mixing of the two phases as the critical temperature is approached from above (*i.e.* as the Universe cools to T_C). They estimated the fraction of the total volume occupied by the broken-symmetric phase by assuming that the dominant fluctuations about equilibrium are subcritical bubbles of roughly a correlation volume which interpolate between the two phases. Their approach was later refined by the authors of Ref. [21] who found, within their approximations, that the 1-loop electroweak potential shows considerable mixing for $M_H \geq 55$ GeV. Clearly, the presence of large-amplitude, nonperturbative thermal fluctuations compromises the validity of the effective potential, since it does not incorporate such corrections.

In order to understand the shortcomings of the mean-field approximation in this context, numerical simulations in 2d [23] and 3d [24] were performed, which focused on the amount of "phase mixing" promoted by thermal fluctuations.

The results show that the problem boils down to how well localized the system is about the symmetric phase as it approaches the critical temperature. If the system is well localized about the symmetric phase, it will become metastable as the temperature drops below T_C and the transition can be called "strong". In this case, the mean-field approximation is reliable. Otherwise, large-amplitude fluctuations away from the symmetric phase rapidly grow, causing substantial mixing between the two phases. This will be a "weak" transition, which will not evolve by bubble nucleation. Defining $\langle\phi\rangle_V$ as the volume averaged field and $\phi_{\rm inf}$ as the inflection point nearest to the $\phi = 0$ minimum, the criterion for a strong transition can be written as [23]

$$\langle\phi\rangle_V < \phi_{\rm inf} \quad . \tag{2}$$

Recently, an analytical model, based on the subcritical bubbles method, was shown to qualitatively and *quantitatively* describe the results obtained by the 3d simulation [25]. The fact that subcritical bubbles successfully model the effects of thermal fluctuations promoting phase mixing and the breakdown of the mean-field approximation with subsequent symmetry restoration, supports previous estimates which showed that the assumption of homogeneous nucleation is incompatible with standard model baryogenesis for $M_H \leq 55$ GeV [21,22]. It is straightforward to adapt these computations to extensions of the standard model. Thus, the requirement that the transition proceeds by bubble nucleation can be used, together with the subcritical bubbles method, to constrain the parameters of the potential.

ACKNOWLEDGMENTS

I am grateful to my collaborators Rocky Kolb, Andrew Heckler, Graciela Gelmini, Mark Alford, Julian Borrill, and Rudnei Ramos for the many long discussions on bubbles and phase transitions. I am also grateful to Juan Carlos D'Olivo and the local organizing committee for their warm hospitality during this Conference. This work was partially supported by the National Science Foundation through a Presidential Faculty Fellows Award no. PHY-9453431 and by a NASA grant no. NAGW-4270.

REFERENCES

1. For reviews see, L. Yaffe, hep-ph/9512265, A. G. Cohen, D. B. Kaplan, and A. E. Nelson, Annu. Rev. Nucl. Part. Sci. **43**, 27 (1993); A. Dolgov, Phys. Rep. **222**, 311 (1992); K. A. Olive, in "Matter under extreme conditions", eds. H. Latal and W. Schweiger (Springer-Verlag, Berlin, 1994).
2. G. Steigman, Ann. Rev. Astron. Astrophys. **14**, 339 (1976).
3. A. D. Sakharov, JETP Lett. **5**, 24 (1967).
4. For a review see E. W. Kolb and M. S. Turner, Ann. Rev. Nucl. Part. Sci. **33**, 645 (1983); *ibid.* The Early Universe, (Addison-Wesley, Redwood, CA, 1990).
5. G. t'Hooft, Phys. Rev. Lett. **37**, 8 (1976); Phys. Rev. D**14**, 3432 (1976).
6. V. A. Kuzmin, V. A. Rubakov, and M. E. Shaposhnikov, Phys. Lett. **B155**, 36 (1985).
7. N. S. Manton, Phys. Rev. D**28**, 2019 (1983); F. R. Klinkhammer and N. S. Manton, Phys. Rev. D**30**, 2212 (1984).
8. P. Arnold and L. McLerran, Phys. Rev. D**36**, 581 (1987); *ibid.* D**37**, 1020 (1988); J. Ambjorn and A. Krasnitz, Phys. Lett. **B362**, 97 (1995).
9. G. W. Anderson and L. J. Hall, Phys. Rev. D **45**, 2685 (1992); M. Dine, P. Huet, and R. Singleton, Nucl. Phys. **B375**, 625 (1992).
10. F. Abe et al. (CDF Collaboration), Phys. Rev. Lett. **73**, 225 (1994).
11. C. Jarlskog, Z. Physik C**29**, 491 (1985); Phys. Rev. Lett. **55**, 1039 (1985); M. E. Shaposhnikov, Nucl. Phys. B**287**, 757 (1987); B**299**, 797 (1988).
12. B. Liu, L. McLerran, and N. Turok, Phys. Rev. **D46**, 2668 (1992); M. Dine and S. Thomas, Phys. Lett. **B328**, 73 (1994); G.D. Moore and T. Prokopec, Phys. Rev. Lett. **75**, 777 (1995).
13. A. G. Cohen, D. B. Kaplan, and A. E. Nelson, Phys. Lett. B**245**, 561 (1990); Nucl. Phys. B**349**, 727 (1991); B**373**, 453 (1992); Phys. Lett. **B336**, 41 (1994); M. Joyce, T. Prokopec, and N. Turok, Phys. Rev. Lett. **75**, 1695 (1995).
14. P. Arnold and O. Espinosa, Phys. Rev. **D47**, 3546 (1993); M. Dine, P. Huet, R.G. Leigh, A. Linde, and D. Linde, Phys. Rev. **D46**, 550 (1992); C.G. Boyd, D.E. Brahm, and S. Hsu, Phys. Rev. **D48**, 4963 (1993); M. Quiros, J.R. Spinosa, and F. Zwirner, Phys. Lett. **B314**, 206 (1993); W. Buchmüller, T. Helbig, and D. Walliser, Nucl. Phys. **B407**, 387 (1993); M. Carrington, Phys. Rev. **D45**, 2933 (1992).

15. M. Gleiser and E.W. Kolb, Phys. Rev. **D48**, 1560 (1993).
16. P. Arnold and L. Yaffe, Phys. Rev. **D49**, 3003 (1994).
17. K. Farakos, K. Kajantie, K. Rummukainen, and M. Shaposhnikov, hep-lat/9510020; Nucl. Phys. **B407**, 356 (1993); *ibid.* **B425**, 67 (1994); *ibid.***B442**, 317 (1995); B. Bunk, E.-M. Ilgenfritz, J. Kripfganz, and A. Schiller, Phys. Lett. **B284**, 372 (1992); Nucl. Phys. **B403**, 453 (1993); Z. Fodor et al. Nucl. Phys. **B439**, 147 (1995).
18. J. S. Langer, Ann. Phys. (NY) **41**, 108 (1967); *ibid.* **54**, 258 (1969); M. B. Voloshin, I. Yu. Kobzarev, and L. B. Okun', Yad. Fiz. **20**, 1229 (1974) [Sov. J. Nucl. Phys. **20**, 644 (1975); S. Coleman, Phys. Rev. **D15**, 2929 (1977); C. Callan and S. Coleman, Phys. Rev. **D16**, 1762 (1977); A. D. Linde, Nucl. Phys. **B216**, 421 (1983); [Erratum: **B223**, 544 (1983)]; M. Gleiser, G. Marques, and R. Ramos, Phys. Rev. **D48**, 1571 (1993); D. Brahm and C. Lee, Phys. Rev. **D49**, 4094 (1994).
19. M. Gleiser and A. Heckler, Phys. Rev. Lett. **76**, 180 (1996).
20. M. Gleiser and E. W. Kolb, Phys. Rev. Lett. **69**, 1304 (1992); M. Gleiser, E. W. Kolb, and R. Watkins, Nucl. Phys. **B364**, 411 (1991); N. Tetradis, Z. Phys. **C57**, 331 (1993).
21. G. Gelmini and M. Gleiser, Nucl. Phys. B**419**, 129 (1994).
22. M. Gleiser and R. Ramos, Phys. Lett. B**300**, 271 (1993).
23. M. Gleiser, Phys. Rev. Lett. **73**, 3495 (1994).
24. J. Borrill and M. Gleiser, Phys. Rev. **D51**, 4111 (1995).
25. M. Gleiser, A. Heckler, and E.W. Kolb, cond-mat/9512032, submitted to Physical Review Letters.