GLOBAL SPACE–TIME EFFECTS ON FIRST-ORDER PHASE TRANSITIONS FROM GRAND UNIFICATION

P. HUT
Astronomical Institute, 1018 WB Amsterdam, The Netherlands

and

F.R. KLINKHAMER
Astronomical Institute, 2300 RA Leiden, The Netherlands

Received 7 May 1981

We argue that the supercooling of a first-order phase transition proceeds only to $T \sim 10^{11}$ GeV (calculated for a Coleman–Weinberg potential). Then the barrier width between real and false vacuum as calculated in flat space–time becomes comparable to the scale set by the event horizon, and mode mixing might induce the transition.

We consider phase transitions at grand unification energies, which might have taken place in the early history of the universe. First we discuss the local physics and its implications on the expansion of the universe and secondly we turn to possible effects of the global space–time structure.

Gauge theories for the unification of the strong, electromagnetic and weak interactions have at least two transitions towards a larger symmetry at high energies:

$$G \xrightarrow[M_U \sim 10^{15}\ \mathrm{GeV}]{} SU(3) \times SU(2) \times U(1) \xrightarrow[M_{WS} \sim 10^{2}\ \mathrm{GeV}]{} SU(3) \times U(1)\,,$$

where the group G describes the unified interaction with a single coupling constant g and with quarks and leptons in common representations (for reviews see refs. [1,2]). In order to apply these ideas to the early universe [3], finite-temperature field theory is used and one finds symmetry restoration for high enough temperatures, resembling phase transitions [4]. Spontaneous breaking of symmetries is due to non-zero expectation values [‡1] of the Higgs scalars φ introduced in the theory. This mechanism preserves renormalisation while providing masses to some gauge bosons and fermions [5]. In a first-order phase transition (1 PT) the shift towards non-zero $\langle\varphi\rangle$ is discontinuous, in contrast to the smooth change for a second-order phase transition. Which type of transition occurs depends on the parameters in the effective potential of the scalars $V(\varphi_c, T)$ [6]. To have breaking at very different energies some very special fine-tuning in V is required [7], which may hint [2] to symmetry breaking by radiative terms only [6]. This in turn might explain the hierarchy of hierarchies $M_{WS}/M_U \ll M_U/M_{Pl} \ll 1$, if the quartic coupling constants (λ) are of order g^2 at the Planck energy $M_{Pl} = G^{-1/2} = 1.2 \times 10^{19}$ GeV ($\hbar = c = k = 1$) [8]. Perhaps superunification [9] leads to radiatively broken $G = SU(5)$. The important point here is that Coleman–Weinberg (CW) breaking leads to strongly first-order phase transitions [10,11]. In the following we will use this potential for numerical estimates.

We now consider the scenario for a 1 PT in the cooling universe. Initially the vacuum is symmetric because of a positive temperature-dependent mass^2 term in the effective potential V [12]. For temperatures $T < T_c$ [where for the two minima $V(\varphi_c = 0, T_c) = V(\varphi_c \neq 0, T_c)$] the transition to the energetically favourable broken state is blocked and the

‡1 For $T = 0$ vacuum expectation values, for $T \neq 0$ relative to a Gibbs ensemble of temperature T.

0 031–9163/81/0000–0000/$ 02.50 © North-Holland Publishing Company

universe cools far below the critical temperature T_c to T_{end} when the transition takes place and the latent heat reheats the universe to fT_c, where $f = O(1)$ follows from entropy conservation and depends on the available particle states before and after reheating [13]. For CW breaking of SU(5) the potential is ($\varphi \ll T \ll \sigma$)

$$V = \tfrac{5}{8} g^2 T^2 \varphi_c^2 + B\varphi_c^4 (\ln \varphi_c^2/\sigma^2 - \tfrac{1}{2}) + \tfrac{1}{2} B\sigma^4 , \quad (1)$$

with φ_c a classical scalar field, $\langle\varphi\rangle_{T=0} = \sigma$ and numerical constant $B = 8 \times 10^{-4}$ [11,14]. The zero level is such that $V(\sigma, T=0) = 0$, as required by the presently observed zero cosmological constant [15]. The part of the universe still in the symmetric state has a constant vacuum energy density $\rho_v \sim T_c^4$ which for $T < T_c = 0.3\ \sigma$ [11] leads to exponential expansion $a \propto \exp(t/\tau)$ as follows from the Friedmann equation [3]:

$$(\dot{a}/a)^2 = (8\pi/3M_{Pl}^2)(\rho_v + \tfrac{1}{30}\pi N T^4) , \quad (2)$$

where $a(t)$ is the scale factor and N the effective number of degrees of freedom of relativistic particles. The transition to the broken vacuum takes place through nucleation with a minimal bubble radius; either through thermal excitation or through tunneling [16].

How precisely the transition to the broken state for the whole universe takes place is of crucial importance and in general we can distinguish three cases:

(1) Thermal nucleation rates have a maximum just below T_c and either the bubble density gets high enough and they quickly fill the universe or else the bubbles cannot catch up with the continuously accelerated expansion of the rest of the universe [17,18].

(2) Nucleation through tunneling has a constant rate per space volume. This leads to a large supercooling, which originally was the motivation to consider 1 PT's in order to prevent high monopole densities [19] ‡2. The nucleation rate for the CW SU(5) model equals the expansion rate only at $T = O(1\text{GeV})$ and the transition is directly to SU(3) × SU(2) × U(1), not through an intermediate SU(4) × U(1) [11]. We note that the barrier vanishes here at $T = 0$, which need not be general (e.g. if $\lambda \ll g^4$ in the abelian Higgs model [4]).

‡2 The suppression mechanism of ref. [20] requires large Higgs masses $m_H \gtrsim m_X$, which is not the case for CW breaking.

(3) Nucleation will be immediate if at T_1 the metastable symmetric vacuum becomes unstable (abelian Higgs model: $3g^4/16\pi^2 < \lambda < g^4$; in SU(5) region d of ref. [18]). The false vacuum shifts to the broken state while releasing its latent heat. If there typically is one monopole in a horizon volume their densities would be tolerable for $T_1 \lesssim 10^{12}$ GeV, although correlations starting to form after T_1 suggest the creation of more monopoles, and require $T_1 \lesssim 10^8$ GeV [13,21].

In all cases it is important to reheat locally at least to the Higgs mass so that a new baryon asymmetry can be created, since previously developed asymmetries are strongly diluted by the supercooling.

In first-order phase transitions the energy density of the metastable vacuum drives the expansion of the universe. However, we must also try to incorporate effects of the *global* structure of this accelerated expansion. In particular we will point out where the flat-space equilibrium theory used so far breaks down. Without a consistent theory of quantum gravity and one of non-equilibrium corrections to finite-temperature particle interactions, we can only attempt to give a qualitative picture, as suggested by a semiclassical approximation (see below) analogous to the Hawking effect ‡3.

The major difference between cosmology and flat space–time (laboratory) physics is the existence of horizons:

(1) A particle horizon limits the region of possible causal contact with a given observer before a given time, and thus includes all comoving particles which have intersected the observers past lightcone.

(2) An event horizon limits the region which will in future have the possibility for contact with a geodetic observer and thus is the boundary of the past lightcone of the observer for $t \to \infty$.

In the standard Big Bang model there are only particle horizons. With relativistic particles, adiabatic expansion $T \propto a^{-1}$, we have from (2) $a \propto t^{1/2}$ and hence the maximum proper distance travelled by a light signal up to time t is [3]

$$d_H = a(t) \int_0^t dt'/a(t') = 2t \qquad (a \propto t^{1/2}) , \quad (3a)$$

$$\sim \tau e^{t/\tau} \qquad (a \propto e^{t/\tau}), \quad (3b)$$

‡3 We thank Eardly and Press for reminding us of the Hawking radiation in de Sitter space.

The exponential expansion during a 1 PT introduces also event horizons: two observers initially separated by a distance significantly greater than the exponential timescale $\tau \sim M_{Pl}\,\rho_v^{-1/2}$ will never be able to communicate, because the intermediate region expands with a constant acceleration so that light signals never catch up. The distance to the event horizon is [22]

$$D_H = (3/\Lambda)^{1/2} = (3/8\pi)^{1/2}\,M_{Pl}\,\rho_v^{-1/2}\,. \qquad (4)$$

Here lies the origin of all evil: the flat space–time approximation is expected to be valid for distances much smaller than that of the event horizon, for which a locally inertial reference frame is a good approximation. But for length scales comparable to or larger than the event horizon distance D_H, or, equivalently, for energies $\lesssim D_H^{-1}$ the flat space–time approximation clearly breaks down.

How does this affect our previous discussion of the 1 PT? We expect changes to be small if we discuss effects at values of the classical field φ_c (or $\langle\varphi\rangle$) much larger than D_H^{-1}. But the description of the stability or calculations of tunneling rates of the symmetric vacuum will fail for small φ_c. An (educated) guess of the temperature T^* below which the flat space–time approximation breaks down is made by equating the barrier width $\Delta\varphi_c$ and $D_H^{-1} \sim T_c^2/T_{Pl}$. For the CW potential we find roughly

$$\Delta\varphi_c^2 \sim (g^2T^2/2B)\ln(g\sigma/T) \sim 10^{\,2}g^2T^2,$$

where we replaced the second term in the RHS of (1) by $-2B\,\varphi^4\ln(g\sigma/T)$ [14] and hence the flat space–time treatment of the barrier will not be valid for

$$T < (T_c/T_{Pl})10^{-1}g^{-1}T_c \sim 10^{11}\ \text{GeV}. \qquad (5)$$

Because confinement of the Higgs expectation value in the metastable symmetric state requires local effects on scales which are globally distorted by the background metric for $T < T^*$, we expect that the universe cools to $\sim T^*$ and then shifts to the broken state, thereby ending the supercooling prematurely. Although we cannot give a rigorous proof for our assertion, we will now discuss several analogies in its favour.

Since $T^* \ll T_{Pl}$ one may use a semi-classical approach where gravity is treated classically through general relativity (GR) and the particles as quantum fields. Note the contrast between the local approach in GR and the global treatment for the particle fields. In particular the definition of the Hilbert space of particle states requires a priori causality relations, whereas only the solution of GR equations provides time- or space-like separations between events (cf. ref. [22]). Unambiguous particle states can only be defined for space–time backgrounds somehow related to Minkowski space [24], to which Robertson–Walker and de Sitter spaces are linked by a conformal transformation. The problem is to treat the backreaction of the particles on the metric and as a first guess one uses a somehow regulated energy–momentum tensor of the particle fields as a classical source term in the Einstein equations.

A major result of the semi-classical approach is the prediction of thermal radiation from an isolated black hole with temperature [25]:

$$T_H = (8\pi)^{-1}\,M_{Pl}^2/M_{BH}, \qquad (6)$$

which can be derived in several ways:

(1) A mapping of a complete set of particle states from the asymptotic past into one of the asymptotic future shows that an incoming vacuum state leads to the emergence of a thermal particle spectrum [25].

(2) Thermodynamic considerations suggest (6), up to a factor of order unity, because information on the quantum states is lost by the existence of an event horizon. The entropy of the black hole can be estimated, from which $T_H = (\partial S/\partial M)^{-1}$ follows [26].

(3) Path integrals on a complexified Schwarzschild metric give a propagator of the form of a thermal Green's function with temperature T_H [27].

These derivations are consistent, because the asymptotic region of space–time is flat, where particle states can be defined unambiguously. All distant observers agree on the Hawking radiation (6), but an infalling observer near the horizon will hardly see any radiation [28]. This observer dependency always occurs locally in globally curved space–times for wavelengths of the order of the curvature radius (cf. Schwarzschild radius $2GM \sim T_H^{-1}$). Even in a Minkowski vacuum a constantly accelerated observer will detect thermal radiation, because his detector measures positive frequencies with respect to his own proper time [29]. For the accelerated observer there also is an event horizon. Both the inertial and accelerated observer agree that the detector will be excited, but they differ on the interpretation, namely bremsstrahlung and absorption, respectively.

Also in our cosmological context similar phenomena occur. In a Friedmann universe the particles determine a preferred restframe. If the vacuum energy density, which is locally Lorentz invariant, dominates over that of the particles, accelerated expansion takes place (2), and the universe asymptotically approaches de Sitter space, where geodetic observers are equivalent. Gibbons and Hawking [22] showed with the same path-integral technique as for the black-hole case that every geodetic detector will see radiation with a temperature

$$T_{GH} = (12)^{-1/2}\pi^{-1}\Lambda^{1/2} \sim M_{Pl}^{-1}\rho_V^{1/2} \sim T_c^2/T_{Pl}. \quad (7)$$

Two differences with the black hole case are in order: (1) absorption of the thermal particles does not destabilize the event horizon [22], in contrast to the increasing rate of black-hole evaporation; (2) no observer independent definition of this radiation is possible ‡4; indeed *if* this could be done, local Lorentz invariance would give an infinite total energy density from the superposition of the finite contributions (7) of all equivalent observers.

To make the link with our assertion that the false vacuum indeed decays at a temperature T^* (5), we now give a general physical picture for the above results. Parker [28] notes that whenever a physical system is externally disturbed on a time-scale τ, modes with frequencies $\omega \lesssim \omega_{cr} \sim \tau^{-1}$ are excited. Both in the black-hole and the de Sitter case, the particle production results from a mixing of positive and negative frequencies of the particle fields, caused by the time dependence or curvature of the background metric. Modes are excited with energies

$$\omega \lesssim \omega_{cr}, \quad \omega_{cr} \sim (GM)^{-1} \quad \text{(Schwarzschild)}, \quad (8a)$$

$$\sim \Lambda^{1/2} \quad \text{(de Sitter)}, \quad (8b)$$

which agrees with the exponential drop in a Planck spectrum for $\omega > T$, with T given by (6) or (7). Parker also shows that relations (8) imply particle creation near enough to the event horizons so that the Heisenberg uncertainty for detection would be large enough to compensate for the negative energy of one of the particles, thus providing an energy reservoir for detection of Hawking radiation. As mentioned above, the backreaction on the event horizon is different in the black-hole and the de Sitter case.

Similarly, mode mixing will occur for energies given by (8b) during the exponential expansion in a 1 PT in the early universe. As soon as the potential barrier around the false vacuum becomes narrower than this range, decay is no longer prohibited rigorously and we expect the transition to occur.

Finally we compare our discussion with a recent article by Shore [30], who considers CW-breaking with a given strong curvature in de Sitter space ‡5. He finds symmetry restoration for a curvature R with a Hawking temperature (7) larger than the Higgs or gauge boson mass for conformally ($\frac{1}{6}R\varphi^2$) or minimally (no $R\varphi^2$) coupled scalars, respectively. But in a 1 PT the curvature due to the false vacuum is less than this critical value by a factor $\sim T_c^2/T_{Pl}^2$. Hence this curvature will not affect the existence of an asymmetric true vacuum. In our view, the important effect of space–time curvature is not a qualitative change in the effective potential, such as the disappearance of a minimum, but the irrelevance of a stabilizing narrow barrier around the false vacuum when mode mixing occurs with respect to a flat space–time approximation.

Some implications of our suggestion of a globally induced transition to the broken vacuum at $T^* \sim 10^{11}$ GeV: (1) perhaps low enough monopole densities [13] even if there are no other suppression mechanisms [31]; (2) no unnatural ($< M_{WS} \sim 10^2$ GeV) supercooling; (3) a smooth transition and reheating, thereby saving baryon number and helium synthesis. All's well that ends well.

‡4 But perhaps two geodetic observers, passing each other, might agree, because added to the expected Doppler shift is an Unruh-type radiation from their relative acceleration [23].

‡5 Abbot [14] considers barrier penetration for an ad hoc $R\varphi^2$ term in the flat-space CW potential. But for the numerical value of the curvature R determined by ρ_V his temperature-independent barrier is of the order of our uncertainty range D_H^{-1}.

References

[1] D.V. Nanopoulos, XVième Rencontre de Moriond, CERN-TH 2896 (1980).
[2] J. Ellis, 21st Scottisch Universities Summer School in Physics, CERN-TH 2942 (1980).
[3] S. Weinberg, Gravitation and cosmology (Wiley, 1972).

[4] D.A. Kirzhnits and A.D. Linde, A.. Phys. 101 (1976) 195; and references therein.
[5] J.C. Taylor, Gauge theories of weak interactions (Cambridge U.P., 1976).
[6] S. Coleman and E. Weinberg, Phys. Rev. D7 (1973) 1888.
[7] A.J. Buras, J. Ellis, M.K. Gaillard and D.V. Nanopoulos, Nucl. Phys. B135 (1978) 66.
[8] J. Ellis, M.K. Gaillard, A. Peterman and C.T. Sachrajda, Nucl. Phys. B164 (1980) 253.
[9] J. Ellis, M.K. Gaillard and B. Zumino, Phys. Lett. 94B (1980) 343.
[10] M. Daniel, Phys. Lett. 98B (1981) 371.
[11] A. Billoire and K. Tamvakis, preprint CERN-TH 3019 (1981).
[12] S. Weinberg, Phys. Rev. D9 (1974) 3357.
[13] M. Einhorn and K. Sato, NORDITA preprint (July 1980).
[14] L.F. Abbot, preprint CERN-TH 3018 (1981).
[15] E.W. Kolb and S. Wolfram, Astrophys. J. 239 (1980) 428.
[16] S. Coleman, Phys. Rev. D15 (1977) 2929;
C.G. Callan and S. Coleman, Phys. Rev. D16 (1977) 1762.
[17] K. Sato, NORDITA preprint-80/29 (January 1980).
[18] A.H. Guth and E.J. Weinberg, Phys. Rev. D23 (1981) 876.
[19] M.B. Einhorn, D.L. Stein and D. Toussaint, Phys. Rev. D21 (1980) 3295;
A.H. Guth and S.H. Tye, Phys. Rev. Lett. 44 (1980) 631, 963.
[20] F.A. Bais and S. Rudaz, preprint CERN-TH 2885 (1980).
[21] M. Einhorn, in: Unification of fundamental particle interactions, eds. J. Ellis and P. van Nieuwenhuizen (Plenum, 1980).
[22] G.W. Gibbons and S.W. Hawking, Phys. Rev. D15 (1977) 2738.
[23] L. Smolin, preprint (September 1979).
[24] B.L. Hu, in: Recent developments of general relativity, ed. R. Ruffini (1980).
[25] S.W. Hawking, Commun. Math. Phys. 43 (1975) 199.
[26] J.D. Bekenstein, Phys. Rev. D9 (1974) 3292;
S.W. Hawking, Phys. Rev. D13 (1976) 191.
[27] J.B. Hartle and S.W. Hawking, Phys. Rev. D13 (1976) 2188.
[28] L. Parker, in: Asymptotic structure of spacetime,eds. F. Esposito and L. Witten (Plenum, 1977).
[29] W.G. Unruh, Phys. Rev. D14 (1976) 870.
[30] G.M. Shore, Ann. Phys. 128 (1979) 376.
[31] A.D. Linde, Phys. Lett. 96B (1980) 293.