Cosmological Production of Superheavy Magnetic Monopoles

John P. Preskill
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts 02138
(Received 21 June 1979)

Grand unified models of elementary particle interactions contain stable superheavy magnetic monopoles. The density of such monopoles in the early universe is estimated to be unacceptably large. Cosmological monopole production may be suppressed if the phase transition at the grand unification mass scale is strongly first order.

There has been much interest recently in grand unified models of elementary particle interactions,[1] in which a simple gauge group breaks down at a very large mass scale to a group containing the $SU(3)\otimes SU(2)\otimes U(1)$ of the strong, weak, and electromagnetic interactions. Because the unbroken symmetry group contains a U(1) factor, general arguments[2] imply that such models contain topologically stable solitons which carry U(1) magnetic charges and have masses of the order of the scale of the symmetry breakdown. In this note, I will argue that an unacceptably large number of such superheavy magnetic monopoles (M) and antimonopoles ($\overline{M}$) might have been produced in the early universe, indicating a possible discrepancy between standard big-bang cosmology and grand unified models.

The masses of the superheavy particles in grand unified models are characterized by a zero-temperature scalar-field[3] expectation value, v_0, which is expected to be of the order of 10^{15} GeV.[4] The M mass is approximately given by $m \approx hv_0$, where h is the M U(1) magnetic charge.[2,5] The smallest allowed magnetic charge is $h = 2\pi/q$, where q is the minimal U(1) charge of a particle which transforms as a singlet under the unbroken subgroup.[2] If symmetry breakdown occurs at many different mass scales, then the M mass is determined by the largest scale at which a U(1) factor appears in the unbroken subgroup, but its classical size at zero temperature is of the order of $(v_{\rm min})^{-1}$, where $v_{\rm min}$ is the smallest mass scale at which the U(1) factor is altered.

If the temperature T is greater than a critical temperature T_c, which is of the order of v_0,[6] the full gauge symmetry is restored, and no M's are present. Suppose that very early in the history of the universe, T exceeds T_c.[7] When the universe cools below T_c, M's can be produced. Because M's, unlike the other superheavy particles in these models, are absolutely stable, their density per comoving volume can be reduced only by annihilation of M-$\overline{M}$ pairs. We will see that the expansion of the universe halts this annihilation process. If M's are produced copiously when $T \approx T_c$, then many M's remain when the temperature is much lower—enough to dominate the mass density of the universe by many orders of magnitude.

In a recent, paper, Zel'dovich and Khlopov[8] have considered the cosmological production of M's with a mass of order 10^4 GeV. However, these authors make the implausible assumption that collisions produce a thermal density of M's when $T \lesssim T_c$. I will argue that, if the phase transition at $T = T_c$ is second order (or weakly first order), the production of an appreciable density of M's is a consequence of the large fluctuations near the critical point. If the phase transition is strongly first order, the situation is less clear, and the density of M's may be tolerably small.

We need to estimate both the M density produced initially and the rate at which the density per comoving volume decreases. Since the latter question can be answered more precisely, I consider it first.[9] Below a temperature T_i at which the M production rate is negligible compared with the expansion rate of the universe, the M density is governed by the rate of M-$\overline{M}$ annihilation. If we may ignore M-$\overline{M}$ correlations, we have

$$dn/dt = -Dn^2 - (3\dot{R}/R)n, \tag{1}$$

where n is the M density per unit volume (M and $\overline{M}$ densities are assumed equal), R is the scale factor of the universe, and D characterizes the annihilation process. If the expansion is adiabatic ($RT \approx$ const) and the universe is radiation-dominated, the expansion rate is[10]

$$\dot{R}/R = -\dot{T}/T = T^2/Cm_{\rm P}. \tag{2}$$

Here, T is the temperature of the universe, $m_{\rm P} = G^{-1/2} = 1.2\times10^{19}$ GeV is the Planck mass, and $C = (45/4\pi^3 N)^{1/2} = 0.60N^{-1/2}$, where N is the effective number of spin degrees of freedom due to particles light compared with the temperature.[11] If D depends on the temperature like a simple power, $D = (A/m^2)(m/T)^p$, Eqs. (1) and (2) can be integrated:

$$r(T)=\left\{\frac{1}{r(T_i)}+\frac{A}{p-1}\frac{Cm_{\rm P}}{m}\left[\left(\frac{m}{T}\right)^{p-1}-\left(\frac{m}{T_i}\right)^{p-1}\right]\right\}^{-1}, \tag{3}$$

where $r=n/T^3$. If $p<1$, the expansion of the universe cuts off the annihilation when $T\ll T_i$, and $r(T)$ approaches a constant value. If $p>1$, the annihilation persists at low temperatures; for $T\ll T_i$, the density becomes independent of its initial value, and is given by

$$r(T)\approx\frac{p-1}{A}\frac{m}{Cm_{\rm P}}\left(\frac{T}{m}\right)^{p-1}, \tag{4}$$

if $r(T)\ll r(T_i)$.

For T less than the inverse size of the M, the interactions between M and light charged particles, and between M and $\overline{M}$, are dominated by the long-range magnetic coupling. We can estimate scattering cross sections by calculating them classically. M-$\overline{M}$ can annihilate by capturing each other in magnetic Coulomb bound states, and then cascading down. As long as the M mean free path λ is shorter than the capture distance $a_c=h^2/4\pi T$, the annihilation rate is given by the flux of M's diffusing through the dense plasma of charged particles toward an $\overline{M}$. This flux is $\varphi=h^2(\tau/m)n$, where τ is the mean time between collisions in which the M is scattered by a large angle. The cross section for large-angle scattering of a thermal relativistic particle with charge q is $\sigma\approx(hq/4\pi)^2T^{-2}$, and the M is itself scattered by a large angle after m/T encounters. Therefore the collision time is $\tau\approx m(BT^2)^{-1}$ where $B=(3/4\pi^2)\zeta(3)\sum_i(hq_i/4\pi)^2$. (The sum is over all spin states of relativistic charged particles). We see that D in Eq. (1) is $D=\varphi n^{-1}=h^2(BT^2)^{-1}$. At the temperature $T_f=(4\pi/h^2)^2mB^{-2}$ at which $\lambda\approx a_c$, Eq. (4) gives

$$r(T_f)\approx\frac{1}{Bh^2}\left(\frac{4\pi}{h^2}\right)^2\frac{m}{Cm_{\rm P}}, \tag{5}$$

if $r(T_f)\ll r(T_i)$.

When $T<T_f$, M and $\overline{M}$ can capture each other only by emitting radiation, and D in Eq. (3) is $\langle\sigma v\rangle$, the average value of the product of the M-$\overline{M}$ capture cross section and the relative velocity. A classical calculation yields $D\approx(h^2/4\pi)^2m^{-2}(m/T)^{9/10}$. From Eqs. (3) and (5), we see that capture by emission of radiation does not reduce $r(T)$ below $r(T_f)$. If $r(T_i)$ is smaller than the right-hand side of Eq. (5) (about 10^{-10} for $m\approx10^{16}$ GeV, $h^2/4\pi\approx75$, and $B\approx10$) annihilation does not reduce r at all for $T<T_i$, as long as M-$\overline{M}$ correlations are negligible.

I have neglected the size of the M. However, when $T\lesssim v_{\rm min}\approx250$ GeV, the classical size b of the M grows to $b\approx v_{\rm min}^{-1}$, which is large compared with $(h^2/4\pi T)(T/m)^{3/10}$, the typical impact parameter for which capture can occur by emission of radiation. Hence the capture cross section may be dominated by nonelectromagnetic interactions. If one makes the very optimistic assumption that an M-$\overline{M}$ pair with an impact parameter of order b can capture by emitting light scalar particles, then D in Eq. (1) is $D\approx\langle\sigma v\rangle\approx b^2(T/m)^{1/2}$. For $T_i\approx v_{\rm min}$, Eq. (3) gives us $r(T)\approx(v_{\rm min}/Cm_{\rm P})(m/v_{\rm min})^{1/2}$, if $r(T)\ll r(T_i)$. Comparing with Eq. (5), we see that capture by scalar emission cannot reduce r significantly unless $m\gtrsim10^{18}$ GeV. For smaller values of m, the size of the M can be safely ignored.

I have also neglected correlations between M's which may be produced by the gravitational effects of inhomogeneities in the energy density of the universe. When electrons and positrons freeze out at $T<1$ MeV, the mean collision time of the M becomes large enough so that such effects are potentially important. When galaxies form, M's and $\overline{M}$'s accumulate rapidly in the cores of galaxies and stars, where the annihilation rate is greatly enhanced by the relatively large number densities.

Experimental limits[12] on the M flux in cosmic rays do not apply if $m\approx10^{16}$ GeV. Such massive M's are not accelerated to relativistic velocities by galactic magnetic fields, and so do not ionize strongly. Also, they do not bind to matter in Earth's crust, because of the strong pull of Earth's gravitational field. We can, however, obtain a bound on the present value of r from the simple observation that the mass density due to M's must not exceed the limit on the mass density of the universe imposed by the observed Hubble constant and deceleration parameter.[13] This constraint is $r(2.7\,{}^\circ{\rm K})\lesssim(10\text{ eV})/m$, or $r\lesssim10^{-24}$ for $m\approx10^{16}$ GeV. To find a bound on $r(T)$ which applies when Eq. (1) is still valid, we note that the standard scenario of helium synthesis[14] requires that the mass of M's does not dominate the universe when $T\approx1$ MeV. We demand that

$$r(T=1\text{ MeV})\lesssim(1\text{ MeV})/m, \tag{6}$$

or $r \lesssim 10^{-19}$ for $m \approx 10^{16}$ GeV. Since M-$\overline{M}$ annihilation cannot reduce a large initial density below $r \approx 10^{-10}$, we conclude that $r(T_i) \lesssim 10^{-19}$.

Now we must estimate the initial density $r(T_i)$. As the universe cools below T_c, one does not expect M-$\overline{M}$ pairs to be brought into statistical equilibrium by collisions, because M-$\overline{M}$ pair production is expected to be suppressed by a small factor of order $\exp(-a/g^2)$.[15] Nevertheless, when $T \lesssim T_c$, the scalar field φ undergoes large random fluctuations; the zeros of φ can be identified as M's and $\overline{M}$'s. Positions of M's and $\overline{M}$'s are strongly correlated on a scale of the order of the correlation length,[16] which is the inverse of the largest scalar or vector mass. As T decreases, M's and $\overline{M}$'s pair up and annihilate, but pairs which are widely separated can survive.

When T is far enough below T_c, large fluctuations have no significant effect on M-$\overline{M}$ annihilation. The criterion for large fluctuations to be unimportant is $T \lesssim T_G$, where[17]

$$\lambda^{1/2} T_G / [4\pi v(T_G)] \approx 1. \tag{7}$$

Here, $v(T) = v_0(1 - T^2/T_c^2)^{1/2}$ is a scalar field expectation value and λ is the largest scalar self-coupling in the theory, normalized so that $m_s(T) = \lambda^{1/2} v(T)$ is the largest scalar mass. The most optimistic assumption we can make is that M's and $\overline{M}$'s annihilate rapidly enough to remain in statistical equilibrium until $T \approx T_G$. In that case, when $T \approx T_G$, the density of widely separated M-$\overline{M}$ pairs is suppressed by a Boltzmann factor:

$$r(T_G) \approx \exp[-m(T_G)/T_G] \approx \exp(-\lambda^{1/2}/g), \tag{8}$$

where g is the gauge coupling. If $\lambda^{1/2} g^{-1}$ is not large, there are many unpaired monopoles. After closely paired M's and $\overline{M}$'s annihilate, the widely separated M's which remain feel only the long-range magnetic coupling. When typical M-$\overline{M}$ separations are of the order of $a_c \approx h^2/4\pi T$, M-$\overline{M}$ correlations may be ignored, and Eq. (1) applies. Hence the initial M density in Eq. (3) should be of the order of $r(T_i) \approx (4\pi/h^2)^3 \approx 10^{-6}$.

It appears that M's and $\overline{M}$'s are copiously produced when $T \approx T_c$. The rapid expansion of the universe prevents the complete annihilation of M's, so that M's dominate the energy density of the universe before the time of helium synthesis. This conclusion, which is incompatible with standard cosmology, might be avoided if there is a scalar self-coupling λ which is sufficiently large that $\exp(-\lambda^{1/2} g^{-1}) \lesssim 10^{-19}$. If $g^2/4\pi \approx 1/45$ at the grand unification scale, the required value of λ is so large that it interferes with the integration of the renormalization-group equations down to ordinary energies. One should note, though, that the suppression factor in Eq. (8) is very sensitive to numerical factors which may occur in the exponent, but cannot be calculated accurately.

The above discussion applies if the phase transition is second order. If the phase transition is strongly first order, $r(T_i)$ is more difficult to estimate; conceivably, M production is severely suppressed. In a first-order phase transition, expanding bubbles of the stable asymmetric vacuum form in the metastable symmetric vacuum when T is of the order of a nucleation temperature T_N.[18] M's can be produced by expanding bubbles or collisions between bubbles, but one might hope that, if the discontinuity in the M mass is large compared with T_N, $r(T_i)$ can be as small as 10^{-19}. A strongly first-order phase transition can be generated by an explicit cubic term in the scalar potential, or by higher-order corrections to the finite-temperature effective potential.[19] Such higher-order corrections are important if the scalar mass matrix has eigenvalues which are sufficiently small. In particular, an interesting possibility is that M production is strongly suppressed when bare scalar masses vanish.[20]

This possible conflict between big-bang cosmology and grand unified models might be resolved in several less attractive ways. Perhaps there is no grand unification. If the gauge group contains a U(1) factor at arbitrarily large mass scales, there need never be any M's. Perhaps the universe was never so hot that symmetry restoration occurred. Then, however, the baryon excess in the universe is no longer explained.[7]

Perhaps M production does not occur if the unification scale v_0 is quite close to the Planck mass $m_{\rm P}$. If the M mass exceeds $m_{\rm P}$, quantum gravity corrections might invalidate the analysis in this note. The standard calculations, however, suggest that v_0 is comfortably below $m_{\rm P}$.[1,4]

Finally, consider the consequences of a lower unification scale. The constraint in Eq. (6) can be satisfied by r as given in Eq. (5) for $m \lesssim 10^{11}$ GeV. However, a grand unification scale as low as 10^{10} GeV seems to be ruled out by the observed bound on the lifetime of the proton.[21]

I am grateful to M. Peskin and B. Halperin for many helpful comments and suggestions. I have also benefitted from conversations with P. Ginsparg, A. Guth, E. Purcell, H. Tye, S. Weinberg, and E. Witten. This research was supported in part by the National Science Foundation under Grant No. PHY77-22864.

[1]H. Georgi, H. R. Quinn, and S. Weinberg, Phys. Rev. Lett. 33, 451 (1974); H. Georgi and S. L. Glashow, Phys. Rev. Lett. 32, 438 (1974).

[2]G. 't Hooft, Nucl. Phys. B79, 276 (1974), and B105, 538 (1976); A. Polyakov, Pis'ma Eksp. Teor. Fiz. 20, 430 (1974) [JETP Lett. 20, 194 (1974)].

[3]My arguments should apply even if the symmetry breakdown is dynamical; a composite order parameter then replaces the elementary scalar field.

[4]T. J. Goldman and D. A. Ross, Phys. Lett. 84B, 208 (1979).

[5]E. B. Bogomol'nyi, Yad. Fiz. 24, 861 (1976) [Sov. J. Nucl. Phys. 24, 449 (1976)].

[6]D. A. Kirzhnits and A. D. Linde, Phys. Lett. 42B, 471 (1972); S. Weinberg, Phys. Rev. D 9, 3357 (1974); L. Dolan and R. Jackiw, Phys. Rev. D 9, 3320 (1974).

[7]This assumption is required in calculations of the baryon-to-entropy ratio. S. Weinberg, Phys. Rev. Lett. 42, 850 (1979), and references therein.

[8]Ya. B. Zel'dovich and M. Y. Khlopov, Phys. Lett. 79B, 239 (1979).

[9]This discussion of M-$\bar{M}$ annihilation follows closely that of Ref. 8.

[10]Units are chosen such that $h = c = k = 1$.

[11]In a minimal grand unified model, $C \approx 1/20$ near the critical temperature. I take C to be a constant, although it varies slowly as particle species freeze out.

[12]R. L. Fleischer *et al.*, Phys. Rev. 184, 1393, 1398 (1969); H. H. Kolm *et al.*, Phys. Rev. D 4, 1285 (1971); L. W. Alvarez *et al.*, Phys. Rev. D 4, 3260 (1971).

[13]J. Kristian, A. Sandage, and J. Westphal, Astrophys. J. 221, 383 (1978).

[14]P. J. E. Peebles, Astrophys. J. 146, 542 (1966).

[15]E. Witten, Harvard University Report No. HUTP-79/A007, 1979 (to be published).

[16]B. Halperin, private communication.

[17]V. L. Ginzburg, Yad. Fiz. 2, 2031 (1960) [Sov. Phys. Solid State 2, 1824 (1960)].

[18]S. Coleman, Phys. Rev. D 15, 2929 (1977); A. D. Linde, Phys. Lett. 70B, 306 (1977).

[19]D. A. Kirzhnits and A. D. Linde, Ann. Phys. (N.Y.) 101, 195 (1976).

[20]S. Coleman and E. Weinberg, Phys. Rev. D 7, 1888 (1973); E. Gildener and S. Weinberg, Phys. Rev. D 13, 3333 (1976).

[21]A. Buras, J. Ellis, M. K. Gaillard, and D. V. Nanopoulos, Nucl. Phys. B135, 66 (1978).