COSMOLOGY OF THE INVISIBLE AXION

John PRESKILL [1], Mark B. WISE [2]
Lyman Laboratory of Physics, Harvard University, Cambridge, MA 02138, USA

and

Frank WILCZEK
Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106, USA

We identify a new cosmological problem for models which solve the strong *CP* puzzle with an invisible axion, unrelated to the domain wall problem. Because the axion is very weakly coupled, the energy density stored in the oscillations of the classical axion field does not dissipate rapidly; it exceeds the critical density needed to close the universe unless $f_a \lesssim 10^{12}$ GeV, where f_a is the axion decay constant. If this bound is saturated, axions may comprise the dark matter of the universe.

Ever since the discovery [1] of the *CP*-violating parameter $\bar{\theta}$ of the strong interactions, the small value of $\bar{\theta}$ has posed a serious puzzle. Current experimental limits [2] on the electric dipole moment of the neutron require [3] $\lesssim 10^{-9}$. The small value of $\bar{\theta}$ is naturally explained in models with a spontaneously broken Peccei–Quinn symmetry [4], but the resulting axion [5] has not been seen. In a recently proposed variant of the Peccei–Quinn scheme [6,7], the axion mass and its couplings to ordinary matter are inversely proportional to the large vacuum expectation value f_a of an $[SU(2) \times U(1)]_W$-singlet scalar field. For $f_a \gtrsim 10$ TeV, this axion is consistent with all known laboratory experiments, but a more stringent constraint on f_a comes from astrophysical considerations [8]. To suppress the power radiated in axions by the helium core of a red giant star to an acceptable level, it is required that $f_a \gtrsim 10^9$ GeV. Since f_a must be very large anyway, one is tempted to suppose [7] that $f_a \sim 10^{15}$ GeV, the mass scale of grand unification [9].

In this paper, we point out a new cosmological constraint on the "invisible" axion satisfying $f_a \gtrsim 10^9$ GeV. Because the couplings of this axion are so weak, the energy density stored in the oscillations of the classical axion field does not dissipate rapidly. To prevent this energy density from exceeding the critical density we must require $f_a \lesssim 10^{12}$ GeV; the invisible axion must not be too visible. The most interesting possibility is $f_a \sim 10^{12}$ GeV, for in that case a nondissipative, pressureless axion gas could constitute the dark matter of the universe.

The axion is a pseudo-Goldstone boson; all of its couplings are suppressed by inverse powers of the axion decay constant f_a. Its only nonderivative coupling, to $\mathrm{tr}(F\tilde{F})$, arises from the color anomaly of the Peccei–Quinn current. In the effective field theory describing physics at energies below f_a, the action of the nonlinearly realized $U(1)_{PQ}$ Peccei–Quinn symmetry on the axion field a is

$$a/f_a \rightarrow a/f_a + \theta. \tag{1}$$

We have chosen to normalize f_a so that a global rotation of a/f_a is equivalent to a rotation of the QCD angle θ; a/f_a must therefore be regarded as a periodic variable with period $2\pi N$, where N is the number of elements in the discrete subgroup of the $U(1)_{PQ}$ symmetry group which is not explicitly broken by nonperturbative QCD effects.

In the standard big-bang cosmology, the axion is in thermal equilibrium at temperatures $T > f_a$. At temperatures satisfying $f_a \gg T$ and $H \gg m_a$, where H is the Hubble parameter and m_a is the axion mass, the axion is decoupled, and behaves like a free massless scalar field. Fluctuations in the angle a/f_a averaged over a volume λ^3 are expected to be of order one if λ exceeds the horizon size t. The fluctuations begin to red-shift away when the wavelength enters the horizon [10]; thereafter, they decay like $\delta a/f_a \sim \lambda/t$. Hence a/f_a is approximately constant over length scales which are smaller than the horizon size, and assumes an arbitrarily chosen value which is typically of order one. In the new inflationary cosmology [11], the size of a causally connected region is enormous, and the background value of a/f_a can be effectively constant throughout the universe, provided that the inflation occurs after the spontaneous breakdown of the Peccei–Quinn symmetry. (There *are* fluctuations in the background value, due to the Hawking temperature.) The fluctuations about this background value, generated by reheating, are small. [+1]

The nonderivative couplings of the axion can be described by an effective lagrangian

$$\mathcal{L} = \tfrac{1}{2}(\partial_\mu a)^2 - f_a^2 m_a^2 F(a/f_a), \qquad (2)$$

where F is a dimensionless function with period 2π and the Taylor expansion $F(x) = \tfrac{1}{2}x^2 + \dots$. ($F$ is also a function of the temperature T.) Because a/f_a is a periodic variable with period $2\pi N$, the potential in eq. (2) has N distinct minima, and if the axion field is at any one of these minima, $\bar{\theta} \simeq 0$ and the CP violation of the strong interactions is acceptably small. However, as noted by Sikivie [12], if $N > 1$ we expect the axion field to choose different minima in causally disconnected regions of space, giving rise to axion domains separated by walls, in conflict with standard cosmology [13].

[+1] If the axion fluctuations on all scales are in thermal equilibrium at the reheating temperature T_{RH}, so that $\delta a \sim (T_{RH}/\lambda)^{1/2}$ initially, then fluctuations entering the "horizon" H^{-1} at a lower temperature T are larger than thermal fluctuations by the factor T_{RH}/T; thermal equilibrium is not maintained by the expansion for $\lambda > H^{-1}$, because the axion is not conformally coupled. Averaging over a "horizon" volume, we have $\delta a^2/f_a^2 \sim (T_{RH}/f_a)^2 \times (H/T)$. These fluctuations are very small, if $T_{RG} \lesssim f_a$ and H^{-1} is much larger than a thermal wavelength.

The problem pointed out by Sikivie can be avoided in various ways. Several authors have shown that it is possible to construct models in which N is effectively 1, so that there are no domains [14]. Furthermore, in the new inflationary cosmology, it is expected that a/f_a is nearly constant throughout the visible universe, if the inflation occurs after the Peccei–Quinn symmetry is spontaneously broken; it is therefore possible for the entire visible universe to lie within a single domain.

But there is another possible cosmological problem with the axion which has not been discussed previously. Whether or not the axion potential in eq. (2) has multiple vacua, and whether or not the universe has undergone a period of exponential inflation, the expectation value $\langle a \rangle$ of the axion field has no reason to be at the minimum of the potential when the universe cools to a temperature such that $H \sim m_a$. Typically, we expect $\langle a \rangle \sim f_a$. Thus, the axion field has a potential energy density of order $f_a^2 m_a^2$, and begins to oscillate coherently.

There are two mechanisms by which the energy stored in these oscillations can be dissipated — Hubble expansion and particle production. We will show that, for $f_a > 10^{13}$ GeV, neither mechanism can prevent the axion energy density from exceeding the critical density needed to close the universe, assuming that $\langle a \rangle \sim f_a$ initially.

We first neglect particle production and analyze the effect of the Hubble expansion; it will be shown below that particle production really is negligible. Because the other modes red-shift away rapidly, we consider only the zero-momentum mode of the axion field. For values of $\langle a \rangle/f_a$ which are somewhat smaller than one, it is reasonable to approximate the potential in eq. (2) by the quadratic term, and the equation satisfied by the zero-momentum mode is

$$d^2\langle a \rangle/dt^2 + 3H(t)d\langle a \rangle/dt + m_a^2(t)\langle a \rangle = 0. \qquad (3)$$

As will be described in more detail below, the axion mass m_a varies as the temperature changes, so we have allowed it to be time-dependent in eq. (3). If H and $\dot{m}_a/m_a$ are small compared to m_a, eq. (3) has an approximate solution of the form

$$\langle a \rangle = A(t)\cos(m_a t), \qquad (4)$$

where $A(t)$ satisfies

$$d(m_a A^2)/dt = -3H(m_a A^2). \qquad (5)$$

Eq. (5) can be integrated, giving

$$(m_a A^2)_f/(m_a A^2)_i = (R_f/R_i)^{-3}, \tag{6}$$

where R is the Robertson–Walker scale factor. This result has a simple interpretation. The zero-momentum mode of the axion field has energy density $\rho_a = \frac{1}{2}m_a^2 A^2$, and may be regarded as a coherent state of axion particles at rest with number density $n_a = \frac{1}{2}m_a A^2$. Eq. (6) says that the number of zero-momentum axions per comoving volume in the expanding universe is conserved. This statement is true because we have neglected axion interactions and have assumed that $m_a(t)$ is so slowly varying that no additional axions can be produced. Eq. (6) begins to apply when the axion Compton wavelength comes within the "horizon" H^{-1}.

Using eq. (6), we can derive an expression for the axion energy density ρ_a as a function of the photon temperature T_f,

$$\rho_a = \frac{1}{2}(m_a)_i(m_a)_f f_a^2 (N_f/N_i)^{3/4}(T_f/T_i)^3(A_i/f_a)^2. \tag{7}$$

Here N is the effective number of spin degrees of freedom, and the factor $(N_f/N_i)^{3/4}$ quantifies the effect of photon heating as the degrees of freedom freeze out. The amplitude A begins to decay according to eq. (6) when $m_a(t) \sim 3H(t)$. Thus we expect that A_i/f_a is of order one when T_i is chosen to satisfy

$$m_a(T_i) \sim 3H = (24\pi G\rho)^{1/2} = 39 T_i^2/m_p, \tag{8}$$

where $m_p = 1.2 \times 10^{19}$ GeV is the Planck mass. [The radiation species contributing to ρ in eq. (8) are assumed to be photons, gluons, three neutrinos, two charged leptons, and three quark flavors.] If the temperature T_f is much smaller than the electron mass, we have $(N_f/N_i)^{3/4} \sim 0.13$, recalling that neutrinos decouple before electron–positron annihilation. Therefore, combining (7) and (8) yields

$$\rho_a \simeq 2.5(f_a/m_p)(m_a f_a/T_i)T_f^3(A_i/f_a)^2, \tag{9}$$

where m_a is now the zero temperature value of the axion mass, and A_i/f_a is expected to be of order one.

To proceed with the evaluation of ρ_a using (9), we must solve eq. (8) for T_i; first we need to derive an expression for $m_a(T)$. A global rotation of a/f_a is equivalent to a rotation of the QCD angle θ. Therefore, the axion mass can be calculated in terms of the θ-dependent free energy density as

$$m_a^2(T) = (1/f_a^2)\mathrm{d}^2 F(T,\theta)/\mathrm{d}\theta^2|_{\bar\theta=0}. \tag{10}$$

At high temperature, the θ-dependent part of the free energy density can be computed in the dilute-instanton-gas approximation. Using the methods of ref. [15], we find, for QCD with three quark flavors,

$$m_a(T) \sim (2 \times 10^{-2})(\Lambda^2/f_a)(m_u m_d m_s/\Lambda^3)^{1/2}$$

$$\times (\Lambda/\pi T)^4 [9 \ln(\pi T/\Lambda)]^3, \tag{11}$$

where $m_{u,d,s}$ are the quark masses, and $\Lambda \sim 200$ MeV is the QCD scale factor.

At low temperature, the chiral symmetries of QCD are spontaneously broken, and the axion mass must be computed using the methods of current algebra; rotating θ causes the quark condensate to rotate also. Generalizing the calculation of ref. [16], the result is

$$m_a = (1/f_a)[m_u m_d]^{1/2}/(m_u + m_d)]f_\pi m_\pi, \tag{12}$$

assuming $m_{u,d} \ll m_s$ and choosing $f_\pi = 93$ MeV. To find the axion energy density today, we combine (9) and (12), making the substitutions $T_f = 2.7$ K and $(m_u m_d)^{1/2}/(m_u + m_d) \sim 0.5$. After dividing by the critical density, $\rho_c \sim 2 \times 10^{-29} h^2$ g cm^{-3}, we obtain

$$\Omega_a = \rho_a/\rho_c$$

$$= 10^7 (f_a/m_p)(200 \text{ MeV}/T_i)(A_i/f_a)^2 h^{-2}, \tag{13}$$

where h is the Hubble constant in units of 100 km sec^{-1} Mpc^{-1}. By solving (8) and (11), we find that T_i is a slowly decreasing function of f_a/m_p which takes the value $T_i \sim 4\Lambda \sim 800$ MeV for $f_a/m_p \sim 4 \times 10^{-7}$. Therefore, the axion energy density exceeds the critical density unless

$$f_a \lesssim 4 \times 10^{12} \text{ GeV}, \tag{14}$$

if A_i/f_a is of order one. This bound is our main result.

It has been implicit in the above discussion that the change in m_a during the chiral phase transition can be regarded as adiabatic. The assumption is valid if the chiral phase transition is second-order in the chiral symmetry limit, for it is then very weakly first-order in the presence of small non-zero quark masses. If the phase transition actually occurs suddenly, however, Ω_a is larger by about a factor of four, and the limit on f_a is correspondingly more stringent, $f_a \lesssim 10^{12}$ GeV. Other sources of uncertainty are the Hubble constant h and the QCD parameter $\Lambda = (200$

MeV)λ. Roughly, our result is $f_a \lesssim (4 \times 10^{12}\,\text{GeV})$
$\times h^2 \lambda^{3/4}$.

It should also be noted that (14) was derived by
assuming the initial value of $\langle a \rangle/f_a$ to be of order one.
In a conventional cosmological scenario, there are
large fluctuations in a/f_a over distances exceeding the
horizon size. The axion energy density begins to be-
have like nonrelativistic matter when the axion
Compton wavelength comes within the horizon, and
it is unavoidably true that at this time the energy den-
sity stored in the nonrelativistic axion modes is of
order $f_a^2 m_a^2$ [+2]. In an inflationary cosmology in which
the inflation occurs after the Peccei—Quinn symmetry
is spontaneously broken, however, a/f_a can be ap-
proximately constant throughout the visible universe.
If its arbitrarily chosen value, which we expect to be
of order one, happens to be small, the bound on f_a
can be relaxed. For example, if $f_a \sim 10^{15}$ GeV, the
axion energy density is acceptably small for
$(A_i/f_a)^2 \lesssim 10^{-3}$ [+3]. Thus, for the value of f_a sug-
gested by grand unified theories, we are faced with a
naturalness problem which is not quite so severe as
the strong CP problem that the invisible axion is in-
tended to solve. [+4]

Let us now examine whether the axion energy
density can be dissipated by particle production [18].
Naively, particle production is insignificant, because
the invisible axion with $f_a \gtrsim 10^9$ GeV has a lifetime

---

[+2] Another means of dissipating the axion energy density
has been discussed by Vilenkin and Everett [17]. They
propose that nearly all of the energy density is carried by
oscillating pieces of domain wall bounded by strings, and
that these oscillations are damped by the emission of grav-
itational radiation. We expect that their mechanism can
actually dissipate only a small fraction of the axion energy
density, and have ignored it in our discussion.

[+3] Reheating does not much affect A_i/f_a. If T_i is the temper-
ature at which the axion Compton wavelength becomes com-
parable to the "horizon" size H^{-1}, the fluctuations aver-
aged over a "horizon" volume are $\delta a^2/f_a^2 \sim (T_{RH}/f_a)^2$
$\times (40 T_i/m_p)$, which is quite small for $T_{RH} \lesssim f_a$.

[+4] The grand unified axion may be made acceptable in yet
another way, if inflation is used to avoid axion domain
walls. In a model in which the PQ symmetry is spontane-
ously broken at the GUT scale, we actually expect
$f_a \sim (10^{15}\,\text{GeV})/N$, where Z_N is the subgroup of $U(1)_{PQ}$
which is not explicitly broken by nonperturbative QCD
effects. We can make N large by incorporating into the
model many superheavy fermions with PQ charges; for N
sufficiently large, the constraint (14) can be satisfied.

---

which far exceeds the age of the universe. However,
the oscillating axion field represents a coherent
space consisting of a large density of zero-momentum
axions, and one wonders whether coherence effects
can greatly amplify the rate at which this state decays.
We will show that in fact the cosmological red shift
prevents such coherence effects from substantially re-
ducing the axion energy density.

The two-photon decay mode of the axion, which
is the dominant one at zero temperature, is forbidden
for an axion at rest when $T \gtrsim 6$ keV$(10^9\,\text{GeV}/f_a)^{2/3}$,
because the electron plasma frequency exceeds the
photon energy $\frac{1}{2}m_a$. At higher temperatures, the dom-
inant process that destroys zero-momentum axions
is the $4a \to 2a$ process, in which the energy of four
zero-momentum axions is transferred to two axions
with $E = 2m_a$; this energy can subsequently red-shift
away harmlessly. The $4a \to 2a$ process causes the
axion zero-momentum mode to decay at the net rate

$$\Gamma_{4a \to 2a} \sim (m_a^3/f_a^2)(A/f_a)^6 n_{2m_a}, \tag{15}$$

where n_{2m_a} is the occupation number for the $E =
2m_a$ modes. (This is the difference of the $4a \to 2a$
rate and the rate for the inverse process.) In standard
cosmology we expect $n_{2m_a} \sim f_a^2/3m_a^2$ while $H \gtrsim 2m_a$.
(In an inflationary scenario, n_{2m_a} is much smaller.)
Subsequently, n_{2m_a} and A begin to decrease rapidly
because of the Hubble expansion. Therefore, $\Gamma_{4a \to 2a}$
never becomes as large as H, and the $4a \to 2a$ process
cannot significantly deplete the axion zero-momen-
tum mode.

Eventually, the universe cools to the point where
the two-photon decay mode of the axion is allowed.
At this point, were it not for the cosmological red
shift, induced emission would cause the $E = \frac{1}{2}m_a$
photon modes to grow exponentially like $e^{\Gamma t}$ where
[19]

$$\Gamma \sim \alpha(A/f_a)m_a. \tag{16}$$

But this exponential growth is blocked by the red-
shift if $\Gamma\tau \lesssim 1$, where $\tau \sim \Gamma(Hm_a)^{-1}$ is the time re-
quired for the band of width Γ of growing modes to
be red-shifted by an amount of order Γ. The condition

$$\Gamma\tau \sim \alpha^2(A/f_a)^2 m_a H^{-1} \lesssim 1 \tag{17}$$

is well satisfied at all times. Thus, the photon modes
do not grow exponentially, and the net $a \to \gamma\gamma$ de-
cay rate of the axion is

$$\Gamma_{a\to\gamma\gamma} \sim \alpha^2(m_a^3/f_a^2)n_\gamma$$

$$\sim 10^{-24} \text{ yr}^{-1}(10^9 \text{ GeV}/f_a)^5 n_\gamma, \tag{18}$$

where $n_\gamma = [\exp(m_a/2T) - 1]^{-1}$ is a thermal occupation number. The axions produced in the early universe are still around.

The axion field is still oscillating coherently today, but with a very small amplitude. From eq. (9) we can deduce that, when the photon temperature is 2.7 K, the amplitude is

$$A/f_a \sim 10^{-22}(f_a/10^9 \text{ GeV})^{1/2}(100 \text{ MeV}/T_i)^{1/2}(A_i/f_a). \tag{19}$$

Therefore, although $\bar\theta$ is still oscillating today with frequency m_a, the *CP* violation in the strong interactions is extremely small; the Peccei–Quinn mechanism works. (Actually, the *CP* violation in the weak interactions shifts the minimum of the axion potential to $\bar\theta \sim 10^{-14}$, in the Kobayashi–Maskawa [20] model.)

The astrophysical and cosmological constraints on the axion together require $10^9 \text{ GeV} \lesssim f_a \lesssim 10^{12} \text{ GeV}$; if we wish to solve the strong *CP* problem by the Peccei–Quinn mechanism, there must be a symmetry-breaking scale in that range. Could the new physics associated with this symmetry-breaking scale be experimentally accessible? One possibility is that there are other flavor symmetries, besides the Peccei–Quinn U(1) symmetry, which are spontaneously broken at the scale f_a. Thus there may be other very weakly coupled Goldstone bosons besides the axion, perhaps including ones with generation nondiagonal couplings. For example, if $f_a \sim 10^{11}$ GeV, the branching ratios for $\mu \to e\phi$ and $K^+ \to \pi^+\phi$, where ϕ is a Goldstone boson, may not be far below current experimental limits.

If f_a is of order 10^{12} GeV, we are led to entertain the possibility that axions make up a significant part of the dark matter of the universe. An axion gas with $\Omega_a \sim 1$ behaves like dust; it is nondissipative and pressureless. [+5] Fluctuations in the axion energy

---

[+5] The energy–momentum tensor of a *coherent* axion gas includes an oscillating pressure term, proportional to $\cos 2m_a t$, which *averages* to zero. But it can be shown by explicit calculation that small fluctuations in the axion energy density grow like $t^{2/3}$ when the universe is dominated by axions; this is the same behavior as for nonrelativistic matter.

density can begin to grow when the universe becomes matter dominated by axions at $T \sim 10$ eV, and these fluctuations can act as a seed for galaxy formation when ordinary matter decouples from radiation at $T \sim 0.3$ eV.

Notes added. The astrophysical bound on f_a derived in ref. [8] has been reexamined by Fukugita et al. [22] who claim that the actual limit is $f_a \gtrsim 4 \times 10^7$ GeV.

The result eq. (12) was previously derived by Dine et al. in ref. [6].

References

[1] G. 't Hooft, Phys. Rev. Lett. 37 (1976) 8; Phys. Rev. D14 (1976) 3432;
C.G. Callan, R.F. Dashen and D.J. Gross, Phys. Lett. 63B (1976) 334;
R. Jackiw and C. Rebbi, Phys. Rev. Lett. 37 (1976) 177.
[2] W.B. Dress et al., Phys. Rev. D15 (1977) 9.
[3] V. Baluni, Phys. Rev. D19 (1979) 2227;
R.J. Crewther et al., Phys. Lett. 88B 1979) 123.
[4] R. Peccei and H. Quinn, Phys. Rev. Lett. 38 (1977) 1440.
[5] S. Weinberg, Phys. Rev. Lett. 40 (1978) 223;
F. Wilczek, Phys. Rev. Lett. 46 (1978) 279.
[6] J.E. Kim, Phys. Rev. Lett. 43 (1979) 103;
M. Dine, W. Fischler and M. Srednicki, Phys. Lett. 104B (1981) 199.
[7] M.B. Wise, H. Georgi and S.L. Glashow, Phys. Rev. Lett. 47 (1981) 402.
[8] D. Dicus et al., Phys. Rev. D18 (1978) 1829.
[9] H. Georgi and S.L. Glashow, Phys. Rev. Lett. 32 (1974) 438;
H. Georgi, H. Quinn and S. Weinberg, Phys. Rev. Lett. 33 (1974) 451.
[10] A. Vilenkin, Phys. Rev. Lett. 48 (1982) 59.
[11] A. Guth, Phys. Rev. D23 (1981) 347;
A.D. Linde, Phys. Lett. 106B (1982) 389;
A. Albrecht and P.J. Steinhardt, Phys. Rev. Lett. 48 (1982) 1220.
[12] P. Sikivie, Phys. Rev. Lett. 48 (1982) 1156.
[13] Ya.B. Zeldovich, I.Yu. Kobzarev and L.B. Okun, Zh. Eksp. Teor. Fiz. 67 (1974) 3.

[14] G. Lazarides and Q. Shafi, Phys. Lett. 115B (1982) 21;
H. Georgi and M.B. Wise, Phys. Lett. 116B (1982) 123;
S. Dimopoulos et al., Phys. Lett., 117B (1982) 185.

[15] D.J. Gross, R.D. Pisarski and L.G. Yaffe, Rev. Mod.
Phys. 53 (1981) 43.

[16] M. Peskin, unpublished;
S. Weinberg, Phys. Rev. Lett. 40 (1978) 223;
W. Bardeen and S.-H.H. Tye, Phys. Lett. 74B (1978)
229.

[17] A. Vilenkin and A.E. Everett, Phys. Rev. Lett. 48
(1982) 1867.

[18] L.F. Abbott, E. Farhi and M.B. Wise, Phys. Lett. 117B
(1982) 29.

[19] N.W. McLachlan, Theory and Application of Mathieu
Functions (Oxford, U.P. London, 1947);
B.R. Mollow and R.J. Glauber, Phys. Rev. 160 (1967)
1076.

[20] M. Kobayashi and K. Maskawa, Prog. Theor. Phys. 49
(1973) 652.

[21] L. Abbott and P. Sikivie, to appear.

[22] M. Fukugita et al., Phys. Rev. Lett. 48 (1982) 1522.