Late Time Cosmological Phase Transitions I: Particle Physics Models and Cosmic Evolution

Joshua A. Frieman¹
Christopher T. Hill¹
Richard Watkins¹,²

¹Fermi National Accelerator Laboratory
P.O. Box 500, Batavia, Illinois, 60510

²Enrico Fermi Institute
Department of Astronomy and Astrophysics
The University of Chicago
5640 South Ellis Avenue, Chicago, IL, 60637

Abstract

We describe a natural particle physics basis for late-time phase transitions in the Universe. Such a transition can seed the formation of large scale structure while leaving a minimal imprint upon the microwave background anisotropy. The key ingredient is an ultra-light pseudo-Nambu-Goldstone boson with an astronomically large ($\mathcal{O}(\text{kpc-Mpc})$) Compton wavelength. We analyze the cosmological signatures of and constraints upon a wide class of scenarios which do not involve domain walls. In addition to seeding structure, coherent ultra-light bosons may also provide unclustered dark matter in a spatially flat universe, $\Omega_\phi \simeq 1$. 
I. Introduction

In a few remarkable instances, modern particle physics theory has predicted the existence of new phenomena on macroscopic distance scales. One well-known example is the axion, a hypothetical pseudo-Nambu Goldstone boson (pNGB) associated with the Peccei-Quinn symmetry, introduced to solve the strong CP problem. Axions arise when a global $U(1)_{PQ}$ symmetry is spontaneously broken by the vacuum expectation value of a complex scalar at the scale $f_a$, $\langle \Phi \rangle = f_a e^{i \alpha}$. At this scale, the axion, the angular field $\alpha$ around the infinitely degenerate minimum of the potential, is a massless Nambu-Goldstone boson. QCD instantons explicitly break the global symmetry at the scale $f_\pi \sim 100$ MeV, generating the axion mass, $m_a \sim O(m_\pi f_\pi / f_a)$. Since its couplings and mass are suppressed by inverse powers of $f_a$, the axion is very light and very weakly interacting. Nevertheless, it can play an important role in astrophysics and cosmology; indeed, astrophysical and cosmological arguments constrain the global symmetry breaking scale to lie in a narrow window around $f_a \sim 10^{10} - 10^{12}$ GeV. Thus, the axion mass $m_a \sim 10^{-5} eV (10^{12} GeV / f_a)$, and its Compton wavelength is macroscopic, $\lambda_a \sim (f_a / 10^{12} GeV)$ cm. Although motivated by the strong CP problem, the axion is a particular instance of a more general phenomenon, and it portends an important lesson: the physical world may contain many new phenomena in the far infrared which are not directly accessible in the laboratory, but which may play an important role in the development of the early Universe.

Axions may be generalized to include familons and majorons, as well as more exotic objects. Recently, a class of pNGBs closely related to familons (called 'schizons'), with masses of order $m_\phi \sim m_{\text{fermion}}^2 / f$, has been analyzed in some detail. If one associates $m_{\text{fermion}}$ with a hypothetical neutrino mass, $m_\nu \sim 0.01 - 1$ eV, and $f \sim M_{\text{GUT}} - M_{\text{pl}} \sim 10^{15} - 10^{19}$ GeV, one arrives at a cosmologically interesting scale for the boson Compton wavelength, $\lambda_\phi \sim f / m_\nu^2 \sim$ kiloparsecs – Megaparsecs. This naturally leads to the idea of a "late time phase transition (LTPT)," i.e., a vacuum rearrangement occuring at the
very low temperature $T_c \sim m_\nu$, which may generate structure on a correspondingly large scale.\textsuperscript{5} If it occurs after decoupling of the cosmic background radiation at $T_{\text{dec}} \simeq 0.3$ eV, such a late transition opens the possibility of forming large-scale structure without imprinting an excessive angular anisotropy $\delta T/T$ on the $3^\circ K$ microwave background. It was originally proposed that "soft" domain walls, with thickness of order $m_\phi^{-1} \sim f/m_\nu^2$, would form in the discrete symmetry breaking at $T_c \sim m_\nu$ and generate non-linear density fluctuations, while (it was hoped) avoiding direct imprinting of the walls upon the background radiation.\textsuperscript{5}

The idea of a late time phase transition is, we believe, more general than the particular realizations that have been suggested. However, it is in need of specific detailed models and further theoretical refinement before it can be assessed and tested. Preliminary numerical analysis has suggested that soft domain walls of ultra-light pNGB's may remain potentially problematic, since it has been argued that at least one large wall ultimately extends across our Hubble volume\textsuperscript{6}, leading to unacceptably large microwave background fluctuations\textsuperscript{7}. It is unclear, however, whether these simulations are ultimately definitive; they do not necessarily have the resolution to see the formation of small structures, such as "vacuum bags," which Widrow and others have argued may provide an attractive mechanism for the formation of accretion centers.\textsuperscript{8} Moreover, it is conceivable that one might still arrive at a viable soft-domain wall scenario with suitable modifications, e.g., if the walls are subject to strong friction due to the surrounding medium, or if they divide regions of differing neutrino mass\textsuperscript{9}. In this paper, we will not consider domain wall scenarios further.

Recently, Press, Ryden, and Spergel (PRS) have considered an alternative possibility of a second-order late time phase transition and its consequences for large-scale structure,\textsuperscript{10} without invoking domain walls. In their model, non-linear fluctuations in an evolving soft boson field directly generate large-scale structure (and voids) on scales $\sim 30h^{-1}$ Mpc. In addition, the coherent oscillations of the field provide dark matter
with critical density, \( \Omega_\phi = 1 \). The soft bosons are non-relativistic, but their Compton wavelength is so large that they do not cluster on galaxy or cluster scales, in agreement with the fact that the inferred density of matter clustered on these scales is only \( \Omega \simeq 0.1 - 0.2 \). In this model, the dark matter in galaxies and clusters is thus purely baryonic; this is marginally consistent with limits on the baryon density \( \Omega_b \) from big bang nucleosynthesis. If viable, this scenario is attractive, since it brings together a variety of cosmological problems and solves them in one model. Yet, an important issue is whether such a model is reasonable from the viewpoint of particle physics.

Our only rational guideline in thinking about ultra-low mass particles is the principle of "Naturalness." This is a well-defined operational principle in theoretical elementary particle physics, first stated by 't Hooft. In this regard, small mass scales must be "protected" by symmetries, such that when the small masses are set to zero they cannot be generated in any order of perturbation theory, owing to the restrictive symmetry. We will not enter here into a general or complete discussion of this mechanism, and how to insure its implementation, since the literature of particle theory is infused with this principle (see eg., refs. (4, 12) for a more lengthy discussion of naturalness in the schizon models and in thermal physics of soft-bosons). In its strongest form, the principle of naturalness requires that small mass scales (or large hierarchies) must appear as a consequence of some plausible mechanism (in addition to being protected once they appear). While the cosmological implications depend upon dynamics and are generally insensitive to whether or not a given model Lagrangian has been fine-tuned, the form of any given low-energy effective Lagrangian, and its finite temperature corrections, will be strongly influenced by the symmetries of the interactions of the full theory. Let us summarize the constraint of naturalness in the present context.

As mentioned above, there can exist a general class of pNGBs ("schizons"), with masses \( m_s \sim \frac{m_{\text{fermion}}^2}{f_s} \), where the decay constant \( f_s \) is naturally associated with
the GUT or Planck scale, $f_* \sim 10^{15} - 10^{19}$ GeV. Here, the small schizon mass is protected by fermionic chiral symmetries or additional discrete symmetries and is therefore technically natural. That is, when certain fermion mass terms are set to zero in the Lagrangian, the schizon mass goes to zero; the fermion mass terms will not be generated in any order of perturbation theory. Familons or schizons are really not significantly less compelling than axions; indeed, if one accepts the existence of axions, then it would seem undemocratic of nature not to supply pNGB brethren such as familons or schizons.

From this perspective, the "soft boson" model of PRS, while cosmologically interesting, is highly unnatural. The Lagrangian is that of a self-interacting complex scalar field, with potential $V(\phi^\dagger \phi) = \lambda(\phi^\dagger \phi - v^2)^2$ and the scale $v \simeq 10^{17}$ GeV. The mass term for the scalar field, $m_\phi = 2\lambda^{1/2} v$, is fine-tuned to be of order $m_\phi \simeq 2 \times 10^{-28}$ eV $\simeq (30 \text{ kpc})^{-1}$, by fixing the self-coupling constant to be $\lambda \simeq 10^{-108}$. Moreover, the field is assumed to have normal strength interactions with other particles. In any quantum field-theoretic version of the model, these interactions would lead to an uncontrollable quadratic divergence of the mass term (and a logarithmic divergence of the self-coupling). Thus, to maintain the small mass term one must fine-tune the theory in each order of perturbation theory. Also, even if one is willing to accept the unnatural fine-tuning of the Lagrangian, the initial condition imposed on the field, $\phi_0/v \lesssim 10^{-30}$ is ad hoc and fine-tuned. Nevertheless, given the interesting cosmological consequences of the PRS scenario, we feel it is worth pursuing more plausible particle physics models which can incorporate its attractive features.

Thus, in this paper we attempt to "naturalize" and amplify the proposal of a late time phase transition. We will largely dispense with domain walls and instead follow the route of PRS without, however, invoking unnatural fine-tunings of parameters or of initial conditions. In Sec. II, we exploit the aforementioned properties of pNGBs to construct a class of well-defined LTPT models which are acceptable from the point of view of naturalness. In Sec. III, we abstract the general features of these models and
discuss the resulting cosmological scenarios in detail. We follow with the conclusion, in which we also speculate on other possible consequences of an ultra-light pNGB field which became dynamical at recent epochs. In the second paper of this series (hereafter, paper II), we study structure formation arising from the dynamics of a light pNGB field under a wider variety of initial conditions, and we explore concomitant constraints from the microwave background anisotropy.

II. Models

Above, we argued that an ultralight boson mass scale naturally arises only in models which implement symmetries in an appropriate way. In this section, we review two classes of toy models with the desired properties: neutrino schizon models and a hidden axion model.

Consider first the low energy effective Lagrangian which contains a neutrino field $\nu$:

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi + \bar{\nu}_L \gamma^\mu \partial_\mu \nu_L + \bar{\nu}_R \gamma^\mu \partial_\mu \nu_R + m_0 \bar{\nu}_L \nu_R e^{i \phi^f} + h.c. \quad (2.1)$$

where $\nu_{(R,L)}$ are respectively right- and left-handed projections, $\nu_{(R,L)} = (1 \pm \gamma^5)/2$. The term proportional to $m_0$ can arise from a Yukawa coupling $g \bar{\nu}_L \nu_R \Phi + h.c.$, where the complex scalar field $\Phi$ has a non-zero vacuum expectation value, $(\Phi) = f e^{i \phi^f} / \sqrt{2}$, and $m_0 \equiv g f / \sqrt{2}$. This is a familiar chiral Lagrangian, possessing the continuous chiral symmetry:

$$\nu_L \rightarrow e^{i \alpha} \nu_L; \quad \nu_R \rightarrow e^{-i \alpha} \nu_R; \quad \phi \rightarrow \phi + 2 \alpha f \quad (2.2)$$

Such a theory has several important and well-known properties: (1) it can be embedded in a fully renormalizable theory in which a $U(1)$-invariant complex field develops a vacuum expectation value, $(\Phi) = f$, and $\phi$ is then the residual Nambu–Goldstone boson; (2) $\mathcal{L}$ is itself renormalizable for a small cut-off $\Lambda$, up to suppressed counterterms of order $\Lambda / f$; (3) $\phi$ will be identically massless unless terms are introduced which explicitly
break the chiral symmetry; (4) $\phi$ satisfies "Adler decoupling," i.e., we may replace $\nu$ everywhere by $\nu'$:

$$\nu'_L = \nu_L e^{i\phi/2f}; \quad \nu'_R = \nu_R e^{-i\phi/2f};$$

(2.3)

and we thus see that $\phi$ disappears in the mass term but couples derivatively to the neutrino as $\partial^\mu \phi \bar{\nu}' \gamma^5 \gamma_\mu \nu' / f$. Therefore, for small $\phi$ momentum $q_\mu$, $\phi$ emission or absorption amplitudes will tend to zero. As a consequence of this decoupling theorem, $\phi$ will not mediate a long range $1/r^2$ force (we note, however, that the decoupling theorem can be violated when the symmetry is broken by a chirally non-invariant mass term).

Let us now consider explicitly breaking the symmetry. To the Lagrangian of (2.1) we may add a small mass term of unknown origin. Usually this comes from some deeper symmetry breaking in the theory which breaks the continuous $U(1)$ down to a discrete subgroup $Z_N$. For example, let us break $U(1) \rightarrow Z_2$. This implies that $\phi \rightarrow \phi + n\pi f$ remains an invariance. So we now have:

$$\mathcal{L}' = \mathcal{L} + \kappa^4 \cos \left( \frac{2\phi}{f} + \theta \right)$$

(2.4)

The physical mass of $\phi$ is determined by shifting $\phi$ to a local minimum and then expanding the cosine to quadratic order in $\phi$. We obtain:

$$m_{\phi} = 2\kappa^2 / f$$

(2.5)

and the quartic interaction term $\lambda \phi^4 / 4!$, where

$$\lambda = 16\kappa^4 / f^4.$$  

(2.6)

The limit $\kappa \rightarrow 0$ is the symmetry limit of the theory in which we recover the full continuous $U(1)$ chiral symmetry. Therefore, $\kappa$ can be naturally small in the technical sense; radiative corrections from the full theory, or even for the effective theory, will only multiplicatively renormalize $\kappa$, since when $\kappa = 0$ the symmetry prevents these effects from generating a nonzero $\kappa$ (hence, small $\kappa_{\text{bare}}$ produces a small $\kappa_{\text{renormalized}} \propto \kappa_{\text{bare}}$).
in perturbation theory). Of course, as we stated above, the mass term comes from some deeper structure in the theory, and when we take the symmetry limit we know that this deeper symmetry breaking structure has also gone to its symmetry limit. However, just adding the mass term for $\phi$ has one special property: it breaks the symmetry only in the mass term of $\phi$ and the decoupling theorem will still hold.

What kind of deeper structure might give rise to such a mass term for $\phi$? In the case of QCD, the proton and neutron are analogues of the $N$ field and the pion is the analogue of $\phi$. The deeper structure that breaks the chiral symmetry there is the presence of light quark masses, which are not chirally invariant. This leads to the nonzero pion mass, and a small non-chirally invariant contribution to the nucleon mass (the $\sigma$ term). We can make an analogy to this situation in the present case by adding an explicit neutrino mass term to the Lagrangian that breaks the chiral symmetry. The low energy Lagrangian then becomes:

$$L = \frac{1}{2} \partial_{\mu} \phi \partial^{\mu} \phi + \bar{\nu}_{L} i \gamma^{\mu} \partial_{\mu} \nu_{L} + \bar{\nu}_{R} i \gamma^{\mu} \partial_{\mu} \nu_{R}$$

$$+ (m_{0} \bar{\nu}_{L} \nu_{R} e^{i \phi/\ell} + e \bar{\nu}_{L} \nu_{R} + h.c.) + \kappa^{4} \cos \left( \frac{\phi}{f} + \theta \right)$$

(2.7)

Notice that the $U(1)$ symmetry is now broken here to the trivial center by the neutrino mass terms as well as the cosine term, and therefore only the residual discrete symmetry $\phi \rightarrow \phi + 2\pi f$ remains as an invariance. (As a consequence, the argument of the pNGB potential is now $\phi/f$ instead of $2\phi/f$.)

Now, if $\kappa \rightarrow 0$ we must also set $e \rightarrow 0$ to recover the true symmetry limit. However, a nonzero $\kappa$ will always be induced at one loop by the presence of a nonzero $e$ and $m_{0}$. With a cutoff $\Lambda < f$, we find the induced term:

$$L_{1-loop} = \frac{m_{0} e \Lambda^{2}}{8 \pi^{2}} \cos(\phi/f).$$

(2.8)

We can freely view this as the origin of the scale $\kappa^{2} \sim \sqrt{m_{0} e \Lambda}$. Thus far, this is a neutrino version of the schizon model of Hill and Ross. We note, however, that we
have ignored the possible effects of CP-violation which can lead to induced Yukawa
couplings of $\phi$ to $\bar{\nu}\nu$, and thus to new long-range forces between neutrinos of quasi-
gravitational strength\textsuperscript{4,13}; a more general model should allow for a CP-phase on the $\epsilon$
term as discussed below.

In this theory we see that the induced scalar mass will be of order:

$$m_\phi^2 \sim m_0 \epsilon (\Lambda^2/f^2) \sim m_0 \epsilon.$$  \hfill (2.9)

The mass for $\phi$ can now be technically naturally small since we can tune the symme-
try breaking parameter $\epsilon$ to be arbitrarily tiny for large $m_0$; e.g., the observed neutrino
mass is $m_\nu \sim m_0 \sim 1$ eV while $m_\phi \sim (100 \text{ Mpc})^{-1}$ if $\epsilon \sim 10^{-60}$ eV. The symmetry
guarantees that radiative corrections will not change this result.

This theory will generally have a late-time phase transition at a temperature of order
the neutrino mass, $T \sim m_0 \sim 1$ eV.\textsuperscript{12} However, having to input the small parameter $\epsilon$
is still not completely satisfactory. Preferable is a scheme in which the scale of $\sim 100$
Megaparsec is generated by a strongly natural mechanism that involves only the putative
scales of particle physics, such as $m_\nu \sim 1$ eV and $M_{\text{GUT}}$ or $M_{\text{Planck}}$, in concert with
some additional symmetry principle.

In the Lagrangian of eqn.(2.7), we observed the appearance of a ("large") quadrati-
cally divergent contribution to the induced mass of $\phi$, eq.(2.8). Can we somehow reduce
the degree of divergence of this induced term? Yes, in fact residual symmetries can
readily control this. Consider first the following schizon Lagrangian invariant under a
residual $Z_2$ discrete symmetry:

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi + \bar{\nu} i \gamma^\mu \partial_\mu \nu + \bar{\chi} i \gamma^\mu \partial_\mu \chi$$

$$+ (m_0 + \epsilon e^{i\phi/f+\beta}) \bar{\nu} L \nu_R + (m_0 - \epsilon e^{i\phi/f+\beta}) \bar{\chi} L \chi_R + h.c.$$  \hfill (2.10)

where the allowed CP-violating phase, $\beta$, is arbitrary. The continuous $U(1)$ chiral
symmetry is broken down to a residual $Z_2$ discrete symmetry:

$$\nu \rightarrow \chi; \quad \chi \rightarrow \nu; \quad \phi \rightarrow \phi + \pi f$$  \hfill (2.11)
If one now computes the induced $\phi$ mass term, one observes that the $\Lambda^2$ term of eqn.(2.8) cancels. The leading contribution is now only log-divergent:

$$L_{1\text{-loop}} \sim \frac{m_0^2 \epsilon^2}{8\pi^2} \ln(\Lambda^2/m_0^2) \cos(2\phi/f + 2\beta).$$

(2.12)

The $\phi$ mass scale that is now induced is of order

$$m_\phi \sim m_0 \epsilon/f.$$

(2.13)

In this case, if $\epsilon \sim m_0 \sim m_\nu$, we retrieve the desired result that $m_\phi \sim m_\nu^2/f$, so that an ultralight schizon emerges without inputting a tiny mass scale.

These models can be easily generalized to further soften the contribution of fermion loops and eliminate the cut-off dependence in the induced $\phi$ mass altogether. Consider the $Z_N$-invariant chiral theory of $N$ neutrinos,

$$L = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi + \sum^{N-1}_{j=0} \bar{\nu}_j i \gamma^\mu \partial_\mu \nu_j + (m_0 + \epsilon e^{i(\phi/f + 2\pi j/N)}) \bar{\nu}_L \nu_{jR} + \text{h.c.}$$

(2.14)

(Hereafter, we suppress possible CP-violating phases $\beta$.) The $U(1)$ chiral symmetry is broken to a residual $Z_N$ discrete symmetry:

$$\nu_j \rightarrow \nu_{j+1}; \quad \nu_{N-1} \rightarrow \nu_0; \quad \phi \rightarrow \phi + 2\pi j f/N.$$  

(2.15)

The 1-loop correction is now

$$\sum^{N-1}_{j=0} \frac{M_j^4}{16\pi^2} \ln \left( \frac{\Lambda^2}{M_j^2} \right),$$

(2.16)

where

$$M_j^2 = m_0^2 + \epsilon^2 + 2m_0 \epsilon \cos \left( \frac{\phi}{f} + \frac{2\pi j}{N} \right).$$

(2.17)

This respects the discrete symmetry. For $N > 2$, the sum $\Sigma_j M_j^4$ is independent of $\phi$; thus, the $\phi$-dependent term is independent of the cutoff $\Lambda$, and for $N > 2$ we can write

$$V(\phi) = - \sum_j \frac{M_j^4}{16\pi^2} \ln M_j^2.$$  

(2.18)
In this case, the $\phi$-potential is explicitly calculable.

In addition to neutrino schizons, one can readily envisage other candidates for "natural" low mass elementary particles. For example, consider a hidden, unbroken $SU(2)$ gauge theory ("quasicolor") containing $n'_f$ massless fermion flavors ("quasi-quarks"), which unifies with $SU(3)_c$ at a scale $M_{GUT}$. By the usual renormalization group, this new force will become confining at an energy scale

$$\Lambda_{SU(2)} \simeq M_{GUT} \left( \frac{\Lambda_{QCD}}{M_{GUT}} \right)^{(33-2n_f)/(22-2n'_f)}$$

Taking $M_{GUT} \sim 10^{17}$ GeV, $n_f = 6$, $n'_f = 4$, and using $\Lambda_{QCD} \simeq 0.1$ GeV, we obtain $\Lambda_{SU(2)} \approx 0.1$ eV. If the "quasi-quarks" are massless, then this theory will contain massless quasi-pions, some diquark "quasi-baryons" with masses of order $2\Lambda_{SU(2)}$, and an analogue of the $\eta'$ with a mass of order $\Lambda_{SU(2)}$ (due to macroscopic instanton effects). These are all "quasi-hadron" boundstates on the scale of $\Lambda_{SU(2)}$, and all of the phenomena will be natural in the strong sense. A quasi-hadronic phase transition would occur when the Universe has a temperature $T \simeq \Lambda_{SU(2)}$, i.e., at a redshift $z \simeq 400$. If such a phase transition occurred within the hidden sector of a unified theory, e.g., if the $SU(2)$ is the low energy remnant of the hidden $E_8$ of superstring models, we would only know of its existence through the gravitational effects that it produces. If the $\theta$ parameter of the $SU(2)$ theory becomes associated with a dynamical field through a Peccei-Quinn-like mechanism, i.e., via an anomalous global symmetry with spontaneous symmetry breaking scale $f_{qa}$, the resulting pNGB, the 'quaxion', will obtain a mass $m_{qa} \simeq \Lambda_{SU(2)}^2/f_{qa}$. For $f_{qa} \sim M_{GUT} \sim 10^{17}$ GeV, we again find $m_{qa} \sim 10^{-28}$ eV, with a Compton wavelength of order 30 kpc.

It is also possible that such cosine potentials can be generated nonperturbatively in a theory if the associated symmetry has an anomalous current; this happens in QCD where instantons nonperturbatively generate a large $\eta'$ mass, or a nonvanishing axion mass, through the axial $U(1)$ anomaly. In the present context the prefactor of the
cosine potential might be the large scale $\Lambda^4 \sim f^4$ multiplied by an extra "tunneling suppression" factor of order $\exp(-8\pi/\alpha)$, yielding a very small mass for the pNGB. The application of ideas like this to soft-boson models has been suggested by Ovrut and Thomas$^{15}$; this may ultimately be a very appealing way of generating soft-boson mass scales.

To study phase transitions in these models, we need to know their behavior at finite temperature $T$. For the neutrino schizon models, this involves integrating over a thermal Fermi-Dirac distribution of neutrinos to get the one-loop finite temperature effective potential $V_T(\phi)$. There are two subtleties here: first, one is interested in studying effects at $T \sim m_\nu \sim eV$, yet the light neutrinos freeze out of thermal equilibrium at $T_F \sim 1$ MeV. Although the neutrinos are no longer in thermal equilibrium, their distribution function retains the form of a redshifted thermal distribution so long as they are semi-relativistic. (This distribution is characterized by an effective temperature $T_{\text{eff}} = T_F a(t_F)/a(t)$, where $a(t)$ is the cosmic scale factor; one thus computes $V_{T,\text{eff}}(\phi)$. ) Second, the reactions that keep $\phi$ in thermal equilibrium have rates of order $\Gamma \sim T^3/f^2$. The field will be in thermal equilibrium if this rate is larger than the Hubble expansion rate $H$, where, for a radiation-dominated universe, $H \sim T^2/M_{\text{Planck}}$. Thus, the condition that $\phi$ be in equilibrium, $\Gamma \gg H$, corresponds to $T \gg f^2/M_{\text{Planck}}$. Hence, for $f \sim M_{\text{GUT}}$, $\phi$ decouples thermally very early. However, we emphasize that this does not invalidate use of the effective potential. In the neutrino-schizon models, one computes the temperature corrections by calculating the value of operators to which $\phi$ couples, such as $\bar{\nu}\nu$, in the appropriate density matrix for the neutrinos. We emphasize that this is a coherent or classical field treatment of $\phi$ and $\bar{\nu}\nu$, corresponding to a classical limit of quantum mechanics, and it is not invalidated by the small reaction rates for $\phi$ particles to scatter incoherently off of a given $\nu$ excitation$^{16}$. 

12
The relevant temperature corrections have been calculated for the $Z_N$ neutrino schizon models in ref. 12. In the high temperature limit, $T \gg m_0$, the result is

$$\Delta V_T(\phi) = \sum_j \frac{M_j^4(\phi)}{16\pi^2} \ln \left( \frac{M_j^2(\phi)}{T^2} \right)$$

(2.20)

Comparison with eqn.(2.18) shows that, for the $Z_2$ model, the coefficient of the cosine potential changes sign at a critical temperature $T \sim (m_0 \pm \epsilon)$: the schizon undergoes a second order phase transition at a temperature comparable to the neutrino mass. If the zero-temperature minimum of the potential is at $\phi/f = \pi$, with maxima at $\phi/f = 0, 2\pi$, then at temperatures $T \gtrless m_\nu$, $\phi/f = \pi$ becomes a local maximum of the potential, and the zero-temperature maxima become minima. This is illustrated in Fig. 1. On the other hand, for the $N > 2$ models, because of the $\phi$-independence of the expression $\sum_j M_j^4$, in the sum of the vacuum contribution of eq.(2.18) and the finite temperature contributions of eq.(2.20) the $\phi$ dependence cancels at $T \gg m_0$. In this case, the $\phi$ potential becomes asymptotically flat as the temperature is raised, as shown in Fig.2. As a result, the transition at $T \sim m_\nu$ in this case is analogous to the transition for an axion potential arising from QCD instantons at $T \sim \Lambda_{QCD}$ in which the axion mass turns on. Note that this is also the expected finite temperature behavior for the hidden $SU(2)$ quaxion model outlined above. There is, however, a qualitative difference between the $Z_{N>2}$ and quaxion models: for the schizon models, the potential turns off only logarithmically as the temperature is raised, while for quaxions the instanton-induced potential is suppressed roughly as an inverse power of the temperature. For axion models, it is well known that the temperature dependence of the mass plays an important role in determining the cosmological bound on $f$ due to the density of coherent axion oscillations. For the schizon models, however, this temperature dependence is small: when the schizon starts oscillating, its mass differs only logarithmically from its zero temperature value. The cosmological implications of these different behaviors will be discussed in the next section.
In the model of Ovrut and Thomas, it appears that there is no temperature dependence of the potential for \( T \ll f \); in this case, the "phase transition" will simply involve the 'unfreezing' of the boson field when the Hubble expansion parameter becomes of order the \( \phi \)-mass. If, however, the tunneling effects which generate the potential are associated with a mass scale \( M \), then we would generally expect the potential to turn on as a power of temperature at \( T \sim M \), as in the (qu)axion model.

III. Cosmology and Late Time Phase Transitions

Having discussed a variety of underlying particle physics models for late phase transitions, we now turn to their cosmological implications. We will examine consequences that are relatively insensitive to the details of particular models. Thus, we assume the existence of a generic pNGB with a phenomenological Lagrangian given as a function of temperature:

\[
L_{pNGB} = \frac{1}{2} \partial_{\mu} \phi \partial^{\mu} \phi + V_T(\phi),
\]

where the finite-temperature potential is

\[
V_T(\phi) = \mu^4 (2 + c(T) [\cos(\phi/f) - 1]).
\]

Here, \( \phi \) is a real pNGB scalar field, \( f \) is the scale of global spontaneous symmetry breaking, assumed to be of order the grand unification or Planck scale, \( f \sim 10^{15} - 10^{19} \) GeV, and \( \mu \) is an explicit symmetry breaking scale associated, e.g., with a light neutrino mass or a hidden strong interaction scale, taken to be of order \( 10^{-2} - 1.0 \) eV. By suitable rescaling of \( \mu \) we can set \( c(T = 0) = 1 \), and we have inserted a constant term in the potential to insure a vanishing cosmological constant in the usual way.

From the discussion of particle physics models in Section II, we extract three broad classes of models for the temperature-dependence of the potential:
I. Models in which $c(T) \sim 0$ at high temperatures, $T \gg T_c \sim \mu$, and the potential turns on approximately logarithmically at a critical temperature $T \sim T_c \sim \mu$; for these models, $c(T) \sim 1$ for $T \leq T_c$. The $Z_N$ schizon models with $N > 2$ are in this category.

II. Models in which $c(T)$ is a slowly varying function of temperature that changes sign at a critical temperature $T_c \sim \mu$, defined by $c(T_c) = 0$; for example, in the $Z_2$ schizon model, $c(T) \sim \ln(T_c/T)$, where $T_c \sim \mu \sim m_\nu$. For models of this type, the Universe goes through a second-order phase transition at the critical temperature $T_c$. For $T \neq T_c$, we will generally assume $|c(T)| \sim 1$.

III. Models in which $c(T) \sim 0$ for high temperatures, $T \gg T_c$, and the potential turns on roughly as a power law in temperature near a critical temperature, $c(T) \sim (T_c/T)^n$, $n \sim \mu$ for $T \gtrsim T_c$. For these models, $c(T) \sim 1$ for $T \lesssim T_c$. For example, the quaxion model, where the instanton-induced potential is suppressed at high temperature, is of this type.

We will not explicitly consider a fourth logical possibility, that $c(T)$ is constant for all $T$, as in the Ovrut–Thomas model with a high energy tunneling scale. This possibility is implicitly included in some of the cases discussed below.

Abstracting from the models of the preceding section, we generally expect the critical temperature to be of order the explicit breaking scale, $T_c \sim \mu$; we will define

$$\xi = T_c/\mu = O(1)$$

as a third parameter of the model, which is naturally of order unity. We note that, in some cases, the potential term in eqn.(3.1) may actually be absent until the Universe cools to the electroweak symmetry breaking scale, $T \sim 100$ GeV. Indeed, in the schizon models this is necessarily so since the potential term arises from quark or lepton masses, which do not appear until electroweak breaking.
With the generic temperature-dependence of the potential in hand, we now discuss the cosmological evolution of the scalar field in broad outline. We focus here upon the spatially homogeneous, zero-momentum mode of the field, \( \phi(t) = \langle \phi(\vec{x}, t) \rangle \), where the brackets here denote spatial averaging. In paper II, we study the evolution of spatial fluctuations in the field; here, we are implicitly assuming that these fluctuations do not strongly perturb the zero mode. This is certainly the case if the fluctuation amplitude is small compared to \( \phi \) as would be expected, e.g., after inflation if the reheat temperature \( T_{RH} < f \): in this case, aside from inflation-induced quantum fluctuations, the field will be homogeneous over many present Hubble volumes. On the other hand, if inflation did not take place, or if \( T_{RH} > f \) so that the global symmetry is broken again after inflation, we generically expect large spatial gradients ("Kibble gradients") in the field due to the fact that \( \phi/f \) is uncorrelated on scales larger than the Hubble radius when the transition at \( T \sim f \) occurs. In this case, it may not make sense to talk of a zero-momentum mode\(^{17}\), but, at least in the absence of light boson production by topological defects, the zero-mode treatment does lead to an estimate of the scalar energy density which should be accurate in order of magnitude, since the (neglected) gradient energy term is comparable to the potential energy term\(^{18}\).

We assume that at temperatures \( f \gg T \gg T_c \), \( \phi \) is a classical field expectation value, randomly placed on its potential. The scalar equation of motion is

\[
\ddot{\phi} + 3H \dot{\phi} + \frac{dV(\phi)}{d\phi} = 0, \tag{3.5}
\]

where the Hubble parameter is given by \( H^2 = (8\pi / 3m_p^2) \rho \) for a spatially flat universe. At these high temperatures, the Compton wavelength of the field is much larger than the Hubble radius, so the potential term is negligible compared to the Hubble damping term (\( \phi \) is effectively massless). In this limit, the solution is time-independent, \( \phi(t) = \phi_1 \), i.e., the field is frozen to its initial value.
As the temperature drops below $T_r$, defined by the point where

$$|m_\phi^2(T_r)| \equiv |d^2 V_{T_r}/d\phi^2|_{\phi=\phi_1} \simeq 9H^2(T_r), \quad (3.6)$$

the $\phi$ field becomes free to roll down the potential, modulo Hubble damping. There will then result spatially coherent oscillations about the potential minimum; once the amplitude is sufficiently small that the oscillations are approximately harmonic, the $\phi$ stress tensor is that of a nonrelativistic particle (pressureless dust). The coherent $\phi$ oscillations may currently dominate the mass density of the universe, providing (unclustered) dark matter, but they must satisfy $\Omega_\phi \leq 1$. This condition provides a constraint on the parameter space of $f, \mu, \text{and } \xi$ (see below).

We now discuss the cosmic evolution of the scalar field for the three classes of models in more depth.

**III.1 MODEL I**

For these models, the potential is flat at high temperatures and turns on rapidly at the critical temperature $T_c$ (see Fig. 2). We shall approximate their behavior by letting

$$c(T) = \begin{cases} 0, & T > T_c \\ 1, & T < T_c \end{cases} \quad (3.7)$$

The model is described by 3 parameters, $f, \mu, \text{and } \xi$ (or $T_c$), and it is useful to define 3 important characteristic temperatures: $T_c$ is the critical temperature as defined above (i.e., when the potential goes from being flat to curved); $T_r$ is the temperature, defined above, when the Hubble damping ceases and the field begins to oscillate in the potential; and $T_d$ is the temperature when the energy density of the scalar field begins to dominate over baryons (if ever). By definition, $T_r \leq T_c$.

We will assume that the initial value of the field, $\phi_1$, is not extremely close to the maximum or minimum of the zero-temperature potential. For a typical initial
value of the field, the potential energy density at the critical temperature is of order
$$V(T_c(\phi_1)) \approx \mu^4.$$  It is convenient to define a parameter \(\alpha\) which characterizes the ratio of
the scalar energy density to the baryon density,

$$\alpha(T) \equiv \left( \frac{\rho(\phi)}{\rho_B} \right)_T.$$  (3.8)

Thus, the temperature \(T_d\) is defined by \(\alpha(T_d) = 1\). Since \(\rho_B(T_c) = \Omega_B \rho_{\text{crit}}(T_c/T_0)^3\),
and using \(T_c = \xi \mu\), we have

$$\alpha(T_c) = \frac{\mu T_0^3}{\Omega_B \rho_{\text{crit}} \xi^3} = 8.5 \xi^{-3} \left( \frac{\mu}{\text{eV}} \right) \left( \frac{0.02}{\Omega_B h^2} \right)$$  (3.9)

where, in the last equality, we have used the present microwave background temperature,
\(T_0 = 2.4 \times 10^{-4}\) eV, and the critical density \(\rho_{\text{crit}} = 8.1 \times 10^{-11}h^2\) eV. We also define
the ratio

$$\beta(T) \equiv \left( \frac{|m_{\phi|^2}}{g H^2} \right)_T.$$  (3.10)

Thus, the temperature \(T_r\), at which the field starts oscillating, is defined by \(\beta(T_r) = 1\).

It is convenient to divide the discussion of this model into two sub-cases: 1) \(\alpha(T_c) < 1\), i.e., \(\mu < (\xi^3 / 8.5)(\Omega_B h^2 / 0.02)\) eV: baryons dominate at \(T = T_c\) and \(T_c > T_d\); 2) \(\alpha(T_c) > 1\): \(\phi\) dominates at \(T_c\), and \(T_d > T_c\). We consider these two cases separately.

III.1.1 Case 1: \(\alpha(T_c) < 1\)

Since baryons dominate at \(T_c\), we must use the baryon density to determine the
expansion rate in the denominator of \(\beta\) (eqn.3.10); for a typical field value on the
potential, we have \(|m_{\phi|^2}^2| \approx \mu^4/f^2\), and so

$$\beta(T_c) = \alpha(T_c) \left( \frac{m_{\phi|^2}}{24\pi f^2} \right).$$  (3.11)

For this case, since \(\alpha(T_c) < 1\), we have \(\beta(T_c) < m_{\phi|^2}/24\pi f^2\). Again, we subdivide the possibilities:
la) \( f > m_{Pl}/\sqrt{24\pi} = 1.4 \times 10^{18}\) GeV: in this case, we are guaranteed to have \( \beta(T_c) < 1 \), so that \( T_r < T_c \): once the potential 'turns on', the field remains frozen to its initial value for a while, because of Hubble damping. As a result, the scalar energy density remains approximately constant (instead of redshifting), acting as a cosmological constant during this phase. Since the baryon density redshifts, the ratio \( \alpha(T) \) grows as \( \alpha \sim T^{-3} \), giving us the more general relation \( \beta(T) = \alpha(T)m_{Pl}^2/24\pi f^2 \), so that \( \beta(T) < \alpha(T) \). Thus \( \alpha(T) \) reaches unity before \( \beta(T) \) does: \( \phi \) is frozen to its initial value until after it dominates the energy density, i.e., we have the relation \( T_c > T_d > T_r \). This can be seen from evaluating

\[
T_d = T_0 \left( \frac{\mu^4}{\Omega_B \rho_{\text{crit}}} \right)^{1/3} \\
= 2\,\text{eV} \left( \frac{\mu}{\text{eV}} \right)^{4/3} \left( \frac{0.02}{\Omega_B h^2} \right)^{1/3} 
\]

while the critical temperature is given by

\[
\frac{T_c}{T_0} = \left( \frac{\mu^4}{\alpha(T_c) \Omega_B \rho_{\text{crit}}} \right)^{1/3}
\]

Thus \( T_d/T_c = \alpha^{1/3}(T_c) < 1 \). The evolution of the scalar and baryon densities is shown schematically in Fig. 3.

Once the temperature drops below \( T_d \), the scalar energy density dominates the expansion rate, and the expression for \( \beta(T) \) changes. Initially, at \( T \) just below \( T_d \), the field is still frozen at its initial value, so the density \( V(\phi) \simeq \mu^4 \simeq \text{constant} \). In this case, we find \( \beta(T \lesssim T_d) = \frac{m_{Pl}^2}{24\pi f^2} \). Since \( f > m_{Pl}/\sqrt{24\pi} \), this implies \( \beta(T \lesssim T_d) < 1 \), so the field remains approximately frozen. In actuality, however, the field does evolve toward the minimum, but on a timescale slower than the expansion time \( H^{-1} \). When \( \beta(T) \lesssim 1 \), the field slowly rolls down the potential with a characteristic growth rate \( \dot{\phi} \sim \exp(m_{Pl}^2 t/3H) \sim \exp(3Ht/\beta) \). One can think of this as a brief period of slow rollover quasi-inflation beginning at \( T = T_d \) and ending when the field begins to roll rapidly
down the potential at $T = T_r$. As we shall see, however, this epoch must be very brief, so that the universe never enters a full-blown exponential de Sitter inflation.

With a simple approximation, we can analyze the slow rollover epoch: we assume that the baryon density is negligible compared to the scalar density at $T < T_d$ and that the scalar energy density remains constant from $T = T_d$ down to $T = T_r$, below which it subsequently scales as non-relativistic matter. As a result, we have, at the present epoch,

$$\frac{\Omega_B}{\Omega_\phi} = \left(\frac{\rho_B}{\rho_\phi}\right)_0 \simeq \left(\frac{T_r}{T_d}\right)^3 = e^{-3\mathcal{N}_e} \tag{3.14}$$

where $\mathcal{N}_e = \ln(T_d/T_r)$ is the number of e-folds of the scale factor during the quasi-inflationary epoch between $T_d$ and $T_r$ (note that we are not assuming exponential expansion during this phase). Since big bang nucleosynthesis suggests $\Omega_B \simeq 0.01 - 0.21$, we require

$$\mathcal{N}_e \simeq (1 \pm 0.5) + \frac{1}{3} \ln(\Omega_\phi) \tag{3.15}$$
e-folds of quasi-inflation, i.e., the scalar field should slow-roll only for about 1 expansion time.

The slow roll-over of a scalar field in a potential of the form (3.1) has been analyzed recently in the context of a model of inflation\textsuperscript{19}, and we can apply those results here. The slow rollover phase ends, and the field begins oscillating, when the field reaches a value $\phi = \phi_r$, given implicitly by $|V'(\phi_r)m_{Pl}/V(\phi_r)| = \sqrt{48\pi}$, or

$$\frac{\sin(\phi_r/f)}{1 + \cos(\phi_r/f)} = \frac{\sqrt{48\pi}}{m_{Pl}}. \tag{3.16}$$

Note that this generally happens while $|V''(\phi)| \ll 9H^2$, i.e., slightly before the field begins oscillating according to our criterion ($\beta(T) = 1$) above. For example, for $f = m_{Pl}$, we have $\phi_r/f = 2.98$, while for $f = m_{Pl}/\sqrt{24\pi}$, $\phi_r/f = 1.9$. If the field begins (at $T \simeq T_d$) at a value $\phi = \phi_1$, the number of slow roll e-folds before it begins to oscillate is

$$\mathcal{N}_e(\phi_1, \phi_r, f) = \frac{16\pi f^2}{m_{Pl}^2} \ln \left[\frac{\sin(\phi_r/2f)}{\sin(\phi_1/2f)}\right] \tag{3.17}$$

20
To achieve \( N_e = 0.5 - 1.5 \) requires, for example, \( \phi_1/f = 2.63 - 2.82 \) if \( f = m_{Pl} \) and \( \phi_1/f = 0.17 - 0.79 \) if \( f = m_{Pl}/\sqrt{24\pi} \). Note that these values are only indicative: for \( N_e \approx 1 \), the approximation we have adopted that \( \rho(\phi) > \rho_B \) is not well justified.

1b) \( f < m_{Pl}/\sqrt{24\pi} \): In this case, from eqn.(3.11), we have \( \beta(T) > \alpha(T) \), so that \( \phi \) starts to oscillate before it dominates the energy density. Once it begins oscillating, \( \rho_\phi \) scales like non-relativistic matter, i.e., like \( \rho_B \), so it never dominates the energy density, \( \rho_\phi < \rho_B \) for all \( T \). (See Fig. 4). In this case, \( \phi \) cannot be the dark matter, but it can still play an important role in structure formation (see paper II).

For this case, the parameter space again naturally splits into two regimes: 1bi) either the oscillations occur after the phase transition, \( T_r < T_c \), which happens if \( \beta(T_c) < 1 \), or (1bii) the field starts oscillating as soon as the potential turns on, \( T_r = T_c \), which corresponds to \( \beta(T_c) \geq 1 \). From eqns. (3.9) and (3.11), in the first regime (1bi),

\[
\mu < \frac{\xi^3 \text{eV} \ 24\pi f^2 \Omega_B h^2}{8.5 \ m_{Pl}^2 \ 0.02}, \quad T_r < T_c
\]  

(3.18)

and the oscillations begin at the temperature

\[
T_r = T_0 \left( \frac{\mu^4 m_{Pl}^2}{24\pi f^2 \Omega_B \beta_{\text{crit}}} \right)^{1/3}
\]  

(3.19)

The second regime (1bii) corresponds to

\[
\frac{24\pi f^2}{m_{Pl}^2} < 8.5 \xi^{-3} \left( \frac{\mu}{\text{eV}} \right) \left( \frac{0.02}{\Omega_B h^2} \right) < 1, \quad T_r = T_c
\]  

(3.20)

In this case, the ratio of the Hubble radius to the \( \phi \) Compton wavelength when it begins oscillating is approximately \( N_r \approx \beta^{1/2}(T_c) \), or

\[
N_r = \frac{3 m_{Pl}}{\sqrt{24\pi \xi^3 f}} \left( \frac{\mu}{\text{eV}} \right)^{1/2} \left( \frac{0.02}{\Omega_B h^2} \right)^{1/2}
\]  

(3.21)

which may be substantially larger than one. This is in contrast to the usual case, e.g., for the axion, which starts oscillating when \( N_r = m_\phi/3H = H^{-1}/3\lambda_\phi \approx 1 \). This difference has important ramifications when we consider structure formation in these models.
Briefly, in the scenario of ref.(10), the scalar field generates non-linear perturbations on the scale of its Compton wavelength; unless \( N_r \gg 1 \) or the transition occurs before recombination, this may lead to unacceptably large microwave background anisotropies (see paper II).

III.1.2 Case 2: \( \alpha(T_c) > 1 \)

In this case, the scalar field dominates the energy density at \( T_c \), and \( T_d > T_c \geq T_r \). From the discussion below eqn.(3.13), before the field starts redshifting we have \( \beta(T) \simeq m_{Pl}^2/24\pi f^2 \). We therefore split the discussion again into two subcases.

2a) \( f > m_{Pl}/\sqrt{24\pi} \) : in this case, \( \beta(T_c) < 1 \), so the field is frozen at \( T_c \) (i.e., \( T_c > T_r \)) and may undergo a brief quasi-inflationary slow-rollover. This is similar to case (1a). Inflation actually begins at \( T \simeq T_d \) given by eqn.(3.12). However, in the range \( T_d < T < T_c \), the potential is flat, and the field does not roll classically. The number of e-folds of inflation between \( T_c \) and \( T_r \) is constrained by the requirement, analogous to eqns.(3.14) and (3.15), that

\[
\alpha(T_r) = \alpha(T_c)e^{3N_e(T_c,T_r)} = \frac{\Omega_\phi}{\Omega_B} \leq \frac{1}{\Omega_B}, \tag{3.22}
\]

where \( \alpha(T_c) \) is given by eqn.(3.9). Since, for case 2, \( \alpha(T_c) > 1 \), this yields the requirement that \( N_e(T_c,T_r) < (1/3)\ln(\Omega_B^{-1}) < 0.5 - 1.5 \); again, the number of inflationary e-folds must be small. For fixed \( \mu \) and \( \xi \), eqn.(3.22) fixes the number of e-folds required to have the proper ratio of scalar to baryon density today; from eqns.(3.16) and (3.17), this fixes the required initial value of the scalar field \( \phi_1 \). Since the number of e-folds is small, the initial value of the field must in general be reasonably large, i.e., not fine-tuned very close to the origin.

2b) \( f < m_{Pl}/\sqrt{24\pi} \) : here, \( \beta(T_c) > 1 \), so the field starts oscillating at the critical temperature, \( T_r = T_c \). In this case, the ratio of the scalar to baryon densities is preserved, \( \alpha(T_0) = \alpha(T_c) \). For \( \Omega_\phi \leq 1 \), we thus require \( \alpha(T_c) \leq 1/\Omega_B \); using eqn.(3.9), this
yields the constraint

$$\mu \leq \frac{\xi^3 \rho_{\text{crit}}}{T_0^3} = 1.4 h_{50}^2 \xi^3 \text{eV} \ , \ \Omega_\phi \leq 1$$ (3.23)

where the present value of the Hubble parameter is $H_0 = 50 h_{50} \text{ km/sec/Mpc}$, and observations indicate $1 < h_{50} < 2$. (We use the parameter $h_{50}$ because its default value, $h_{50} = 1$, yields marginal agreement with globular cluster ages if $\Omega = 1$.) For fixed physical scale $\mu$, we may think of this as determining the required value of $\xi$. For $\mu = O(\text{eV})$, the scalar field provides closure density if $\xi = O(1)$. Since we expect $\xi$ to be of order unity, this shows that the natural explicit symmetry breaking scale for a coherently oscillating field that dominates the present energy density is in the eV range. That this is far below the QCD scale, $\Lambda_{\text{QCD}} \sim 100 \text{ MeV}$, that sets the scale for axion oscillations, does not contradict the fact that axions may also dominate the density (if $f \sim 10^{12} \text{ GeV}$), because the temperature dependence of the axion potential is different (see model III below).

For closure density bosons ($\Omega_\phi = 1$), using $\xi(\mu)$ from eqn.(3.23) we find the critical temperature

$$T_c = \mu \xi(\mu) = 0.9 h_{50}^{-2/3} \text{eV} \left(\frac{\mu}{\text{eV}}\right)^{4/3} \ , \ \Omega_\phi = 1$$ (3.24)

and the redshift of the transition,

$$1 + z_c = \frac{T_c}{T_0} = 3.7 \times 10^3 h_{50}^{-2/3} \left(\frac{\mu}{\text{eV}}\right)^{4/3} \ , \ \Omega_\phi = 1$$ (3.25)

If we require the transition to occur after recombination, i.e., at $z_c < 1000$, and before the first known quasars, i.e., $z_c > 5$, eqn.(3.25) gives

$$8 \times 10^{-3} \text{ eV} < \mu h_{50}^{-1/2} < 0.4 \text{ eV}$$ (3.26)

This also corresponds to the range

$$0.2 < \xi h_{50}^{1/2} < 0.7$$ (3.27)
(For completeness, for case 2a above, the corresponding constraint is $5 < z_r < 10^3$, which yields a range for $\mu$ identical to eqn.(3.26).) For reference, we note that the ratio of the Hubble radius to the scalar Compton wavelength at the onset of oscillations is given in this case by

$$ N_r \simeq \beta^{1/2}(T_c) = \frac{m_p}{\sqrt{24\pi f}} \quad (3.28) $$

For case 2b, this is larger than one.

The parameter space for the different cases of Model I is displayed in Fig.5.

III.2 MODEL II

For these models, the high temperature potential does not vanish, but is inverted from the low temperature potential, as in Fig. 1. We approximate the behavior of this class of models by choosing the temperature-dependent coefficient in eqn.(3.1) to have the form

$$ c(T) = \begin{cases} 
-1, & T > T_c \\
0, & T = T_c \\
1, & T < T_c
\end{cases} \quad (3.29) $$

The analysis is similar to that of Model I, the primary difference being that in some cases the field may now oscillate before the critical temperature is reached. We must now define two temperatures: $T_r^{(1)} \ (T_r^{(2)})$ is the temperature when the field starts oscillating in the high-temperature (low-temperature) potential. Clearly, $T_r^{(1)} \geq T_c \geq T_r^{(2)}$.

We divide the analysis into the same subcases as for Model I. Then, for cases 1a, 1bi, and 2a, the behavior of Model II is identical to Model I, because the field is always frozen in the high temperature potential ($\beta(T > T_c) < 1$); for these cases, the constraints are summarized in Fig. 5. We now turn to the cases where the field does oscillate in the high temperature potential.
1bii) For \( \mu \) in the range given by eqn.(3.20), the field begins oscillating in the high temperature potential, corresponding to \( T_r^{(1)} > T_c = T_r^{(2)} \). The onset of these oscillations occurs when \( \beta(T_r^{(1)}) = 1 \), which implies

\[
\frac{T_r^{(1)}}{T_0} = \left( \frac{\mu^4 m_{Pl}^2}{\Omega_B \rho_{crit} 24 \pi f^2} \right)^{1/3}
\]  

(3.30)

As in Model I, for this case the scalar oscillations never dominate the energy density of the universe. Nevertheless, it is of interest to study the behavior of the field.

When the field starts oscillating at \( T_r^{(1)} \), typically there will be rough equipartition between the kinetic and potential energy of the field, and the energy density decreases with the expansion. After a short time, the kinetic energy has redshifted to become negligible compared to the potential energy, and the energy momentum tensor of the field approximates that of a vacuum state, with vacuum energy \( \rho_{vac} \sim 2 \mu^4 \). This does not lead to inflation, however, because by construction the scalar field never dominates the total energy density for this case. Defining \( \tilde{\rho}_\phi = \rho_\phi - \rho_{vac} \), it is easy to show that \( \tilde{\rho}_\phi \) redshifts like non-relativistic matter, \( \tilde{\rho}_\phi \sim R^{-3} \). This implies that the oscillation amplitude decays as

\[
\bar{\phi}(T) = \bar{\phi}_1 \left( \frac{T}{T_r^{(1)}} \right)^{3/2}
\]  

(3.31)

during the high temperature phase. Using eqns.(3.13) and (3.30), the amplitude at the critical temperature is then

\[
\frac{\bar{\phi}(T_c)}{\bar{\phi}_1} = \beta^{-1/2}(T_c) = \sqrt{24 \pi f} \frac{\xi^{3/2}}{m_{Pl}} \left( \frac{e V \Omega_B h^2}{\mu} \right)^{1/2}
\]  

(3.32)

In this case, \( T_r^{(2)} \approx T_c \), and the ratio of the Hubble radius to the Compton wavelength at the onset of the low temperature oscillations is just the factor by which the amplitude has been damped in the high temperature phase,

\[
N_r = \beta^{1/2}(T_c) = \left( \frac{\bar{\phi}(T_c)}{\bar{\phi}_1} \right)^{-1}
\]  

(3.33)
2b) As for Model I, this is the most interesting case, since the scalar field can dominate the energy density. As above, we have \( T_{r(1)} > T_d > T_c \simeq T_r^{(2)} \); the situation is qualitatively similar to Model I (2b), except that in this case due to the high temperature damped oscillations, the field may be localized near the origin at the critical temperature, \( \bar{\phi}(T_c) \ll f \). This factor is straightforwardly estimated:

\[
\frac{\bar{\phi}(T_c)}{\phi_1} = \left( \frac{24\pi f^2}{m_{pl}^2 \alpha(T_c)} \right)^{1/2} = \sqrt{\frac{24\pi}{m_{pl}}} \left( \frac{\Omega_B}{\Omega_\phi} \right)^{1/2}
\]

where in the second equality we have used \( \alpha(T_c) = \Omega_\phi/\Omega_B \). In this case, the ratio of the Hubble radius to the Compton wavelength at \( T_r = T_r^{(2)} \) is

\[
N_r = \beta^{1/2}(T_c) = \frac{m_{pl}}{2\pi f} = \frac{\bar{\phi}_1}{\bar{\phi}(T_c)} \left( \frac{\Omega_B}{\Omega_\phi} \right)^{1/2}
\]

For some applications in paper II, we will be interested in the case where the high temperature (in addition to the low temperature) oscillations begin after recombination, i.e., \( z_r^{(1)} = T^{(1)}/T_0 < 10^3 \). Using eqn.(3.30), this happens if

\[
\mu < 0.2 \text{ eV} \left( \frac{\sqrt{24\pi h_{50}}}{m_{pl}} \right)^{1/2} \left( \frac{\Omega_B}{0.06} \right)^{1/4}
\]

III.3 MODEL III: QUAXIONS

For this class of models, we take the coefficient in eqn.(3.1) to be

\[
c(T) = \left( \frac{T_c}{T} \right)^n , \quad T > T_c
\]

\[
= 1 , \quad T < T_c
\]

From the effective Lagrangian viewpoint, the parameter \( n > 0 \) corresponds to a fourth parameter of the model (compared to three for models I and II); however, it is determined by microphysics. In this case, the potential is flat at \( T \gg T_c \) and gradually turns on, reaching full strength (and its zero-temperature value) at \( T_c \). For field values away
from the maxima of the potential, it is convenient to expand $V_T(\phi)$ around the potential minimum, $\tilde{\phi} = \phi - \pi f$,

$$V_T(\tilde{\phi}) = f(T) + \frac{1}{2}m^2_\phi(T)\tilde{\phi}^2$$  \hspace{1cm} (3.38a)

where

$$f(T) = 2\mu^4 \left[1 - \left(\frac{T_c}{T}\right)^n\right], \quad T > T_c$$ \hspace{1cm} (3.38b)

$$m_\phi(T) = \frac{\mu^2}{f} \left(\frac{\xi \mu}{T}\right)^{n/2}$$

We define the parameters $\alpha(T)$ and $\beta(T)$ as before. In this case, like the axion, the field can start oscillating at $T_r > T_c$, and it is convenient to divide the cases slightly differently than above\textsuperscript{20}: 1) $\alpha(T_r) < 1$, i.e., a baryon-dominated universe when the oscillations begin, and (2) $\alpha(T_r) > 1$, in which case the scalar field dominates the energy density when it begins to oscillate. For a typical initial field value, $\rho_\phi(T_r) \approx \mu^4$, so that

$$\alpha(T_r) = \left(\frac{\rho_\phi}{\rho_B}\right)_{T_r} = \frac{\mu^4}{\Omega_B \rho_{\text{crit}}(T_r/T_b)^3}$$  \hspace{1cm} (3.39)

### III.3.1 Case 1: $\alpha(T_r) < 1$

In this case, $\rho_\phi(T_r) < \rho_B(T_r)$. The temperature $T_r$ when the field starts oscillating is determined by $\beta(T_r) = m^2_\phi(T_r)/9H^2(T_r) = 1$, where the temperature-dependent mass is given by eqn.(3.38b) and the expansion rate at $T_r$ is fixed by the baryon energy density. Straightforward manipulation gives

$$\frac{T_r}{T_c} = \left[\frac{m^2_\phi 8.5}{24\pi f^2 \xi^3} \left(\frac{\mu}{eV}\right) \left(\frac{0.02}{\Omega_B h^2}\right)\right]^{1/(3+n)}, \quad T_r > T_c$$ \hspace{1cm} (3.40)

$$= \left[\frac{m^2_\phi 8.5}{24\pi f^2 \xi^3} \left(\frac{\mu}{eV}\right) \left(\frac{0.02}{\Omega_B h^2}\right)\right]^{1/3}, \quad T_r < T_c$$

From (3.39) and (3.40), we have

$$\alpha(T_r) = \frac{24\pi f^2}{m^2_\phi} \left(\frac{T_r}{T_c}\right)^n$$  \hspace{1cm} (3.41)
so the condition $\alpha(T_r) < 1$ requires

$$\frac{T_r}{T_c} < \left(\frac{m_{pl}^2}{24\pi f^2}\right)^{1/n}$$

(3.42)

As before, we subdivide this case into two subclasses depending on the value of $f$:

1a) $f > m_{pl}/\sqrt{24\pi}$: In this case, from (3.42) we immediately have $T_r < T_c$. Since the field does not begin oscillating until the potential reaches its zero-temperature form, the high-temperature behavior of the potential is irrelevant in this instance. This is therefore identical to Model I, case 1a.

1b) $f < m_{pl}/\sqrt{24\pi}$: in this case, the ratio $T_r/T_c$ may be smaller or larger than one, depending on the value of $\mu$. From eqn.(3.40), the condition $T_r = T_c$ clearly corresponds to

$$\mu = \mu_{\text{crit}} = \frac{\xi^2 eV}{8.5} \left(\frac{\Omega_B h^2}{0.02}\right) \left(\frac{24\pi f^2}{m_{pl}^2}\right), \quad T_r = T_c$$

(3.43)

(Compare eqns.(3.18) and (3.20).) We thus subdivide this case into two subclasses:(1bi) $\mu < \mu_{\text{crit}}$, which implies $T_r < T_c$, and (1bii) $\mu > \mu_{\text{crit}}$, i.e., $T_r > T_c$.

1bi) $\mu < \mu_{\text{crit}}$: this is identical to case (1bi) of Model I. Again, the high-temperature nature of the potential plays no role, since $T_r < T_c$.

1bii) $\mu > \mu_{\text{crit}}$: in this case, the field starts oscillating while the potential is still evolving. The condition $\alpha(T_r) < 1$ restricts $\mu$ to the range

$$\left(\frac{m_{pl}^2}{24\pi f^2}\right)^{(n+1)/n} > \frac{\mu}{\mu_{\text{crit}}} > 1, \quad T_r > T_c$$

(3.44)

Although the scalar field never dominates the energy density in this case, it is of interest to study the damping of its oscillations in the evolving potential. As before, if we subtract the instantaneous zero-point energy by defining $\bar{\rho}_\phi = \rho_\phi - f(T)$, then for the harmonically oscillating field we find

$$\dot{\bar{\rho}}_\phi + 3H\bar{\rho}_\phi \left(1 - \frac{m_{\phi}}{m}\right) = 0$$

(3.45)
which implies
\[ \ddot{\rho}_\phi \propto \frac{m_\phi(T)}{R^3} \]

(3.46)

Since \( \ddot{\rho}_\phi \sim m_\phi^2(T)\dot{\phi}^2 \), the oscillation amplitude decays as \( \dot{\phi} \sim m_\phi^{-1/2}(T)R^{-3/2} \). Using eqn.(3.38b), this gives the amplitude at \( T_c \),
\[ \frac{\ddot{\phi}(T_c)}{\dot{\phi}_1} = \left( \frac{T_c}{T_r} \right)^{(n+6)/4} \]

(3.47)

and the relative energy density
\[ \frac{\alpha(T_c)}{\alpha(T_r)} = \left( \frac{T_c}{T_r} \right)^{n/2} \]

(3.48)

where the ratio in parenthesis is given by the first eqn.(3.40) and we have assumed \( \dot{\phi}_1 \sim f \). As in the axion case, the commencement of \( \phi \) oscillations in the evolving potential reduces the relative scalar energy density.

III.3.2 Case 2: \( \alpha(T_r) > 1 \)

In this case, the scalar field already dominates the energy density when it begins to oscillate, \( T_d > T_r \). From eqns.(3.10) and (3.38), for \( T_d > T \geq T_r \), we find
\[ \beta(T) = \frac{m_\phi^2}{24\pi f^2} \left( \frac{T_c}{T} \right)^n, \quad T > T_c \]
\[ = \frac{m_\phi^2}{24\pi f^2}, \quad T < T_c \]

(3.49)

where we have used the fact that \( \rho_\phi \simeq \mu^4 \) for \( T \geq T_r \).

As usual, this motivates us to consider two subcases:

2a) \( f > m_{Pl}/\sqrt{24\pi} \): This case is similar to Model I, case (2a). That is, in order to end up with an acceptable value of \( \Omega_\phi/\Omega_B \), the scalar field must undergo a very brief period of slow rolling quasi-inflation. The only difference is that, in the present case, the field can in principle start evolving classically at \( T > T_c \). However, from eqn.(3.49), the rollover rate in this high-temperature regime is exponentially suppressed,
since $\phi \sim \exp(3Ht\beta)$ and $\beta \sim T^{-n}$. As a result, in practice, the initial value(s) of $\phi_1$ required to give $\Omega_\phi/\Omega_B \lesssim 1/(0.01 - 0.2)$ in this case will be similar to that in Model I.

2b) $f < m_{Pl}/\sqrt{24\pi}$: in this case, from eqn.(3.49), we have

$$\frac{T_r}{T_c} = \left( \frac{m_{Pl}^2}{24\pi f^2} \right)^{1/n} > 1 \quad (3.50)$$

and the field starts oscillating above the critical temperature. In the temperature range $T_r > T > T_c$, the field oscillates in the evolving potential, and its energy density is approximately (Cf. eqn.(3.38))

$$\rho_\phi = f(T) + m_\phi^2(T)\phi^2 \quad (3.51)$$

The oscillation amplitude $\phi$ decays according to eqn.(3.47); again assuming a typical initial amplitude $\phi_1 \simeq f$ and using eqn.(3.50), we find

$$\rho_\phi(T_c) \simeq \mu^4 \left( \frac{24\pi f^2}{m_{Pl}^2} \right)^{(6+n)/2n} \quad (3.52)$$

Now the analysis follows along the lines leading to eqn.(3.23): we form the ratio $\alpha(T_c) = \rho_\phi(T_c)/\rho_B(T_c)$ and use the fact that this ratio is conserved for $T < T_c$, i.e., $\alpha(T_0) = \alpha(T_c)$. Then setting $\Omega_\phi \leq 1$, i.e., $\alpha(T) \leq 1/\Omega_B$, yields the constraint

$$\mu \leq \frac{\xi^3 \rho_{\text{crit}}}{T_0^3} \left( \frac{24\pi f^2}{m_{Pl}^2} \right)^{(6+n)/2n} = 1.4h_{50}^2\xi^3 \text{eV} \left( \frac{m_{Pl}^2}{24\pi f^2} \right)^{(8+n)/2n} \quad (\Omega_\phi \leq 1) \quad (3.53)$$

Unlike eqn.(3.23), the cosmological constraint on the scale $\mu$ here depends on $f$, as in the usual axion case; in particular, as $f$ is decreased below the Planck scale, the upper bound on $\mu$ is relaxed\textsuperscript{21}, depending on the index $n$.

Assuming the scalar field saturates closure density determines the parameter $\xi$ and, analogously to eqns.(3.24) and (3.25), we have the critical temperature

$$T_c = \mu\xi(\mu) = 0.9h_{50}^{-2/3} \text{eV} \left( \frac{\mu}{\text{eV}} \right)^{4/3} \left( \frac{24\pi f^2}{m_{Pl}^2} \right)^\frac{1}{n} \quad , \quad (\Omega_\phi = 1) \quad (3.54)$$
and the redshift of the transition,

\[ 1 + z_c = \frac{T_c}{T_0} = 3.7 \times 10^3 h_{50}^{-2/3} \left( \frac{\mu}{\text{eV}} \right)^{4/3} \left( \frac{24\pi f^2}{m_{\text{Pl}}^2} \right)^{\frac{1}{3} + \frac{1}{6}}, \quad (\Omega_\phi = 1) \]  

(3.55)

IV. Conclusions

We have studied a variety of particle physics models which can potentially generate the observed large-scale structure in the context of late time phase transitions involving ultra-low mass bosons. Some of these models can simultaneously explain the dark-matter in the Universe. These models are distinguished in containing a fundamental large distance scale, \( d \), which emerges as a hierarchical ratio of microscopic scales, e.g., \( d \sim M_{\text{GUT}}/m_\phi^2 \). Symmetry principles, abstracted and borrowed from experience with QCD, allow us to build underlying theories of such a phase transition at a very low energy scale, \( T_c \lesssim \text{eV} \), that are technically and even strongly natural. This low energy scale corresponds to a moderately recent cosmological epoch, \( z_c \lesssim 1000 \), and such models may therefore have other potentially striking observational signatures, as in the scenarios for large-scale periodic structure discussed in ref.(21).

The common element of the models considered here is the existence of an ultra-light pseudo-Nambu-Goldstone boson, a lighter cousin of the conventional QCD axion, which becomes dynamical at this late epoch. We have analyzed the cosmic evolution of this field in three broad classes of models and have identified the regions of parameter space for which it can make a significant contribution to the energy density of the Universe. Unlike the conventional axion, an ultra-light boson associated with global symmetry breaking at the GUT scale, \( f \sim 10^{16} \text{ GeV} \), can provide \( \Omega_\phi \sim 1 \) without invoking special initial conditions. We note that such a field is generically very weakly coupled and thus difficult to detect.
For completeness, we close with several comments about other possible implications of, and constraints upon, these models. The first have to do with topological defects. If inflation does not take place, or if it does occur but with reheat temperature $T_{RH} > f$, the boson field is not expected to be homogeneous over scales larger than the Hubble radius at the critical temperature $T_c$. In this case, the symmetry breaking at $T_c \sim \mu$ can lead to the formation of domain walls, which may endanger the isotropy of the microwave background. As in axion models, there are a number of ways to avoid this domain wall problem. For example, one can have additional terms in the scalar potential which explicitly break the residual discrete symmetry (bias the potential), lifting the vacuum degeneracy; this effectively destroys the topological stability of the walls, driving them into the false vacuum regions. As long as the vacuum asymmetry is large enough, the walls can disappear before they do any damage. Alternatively, if the broken global symmetry is $U(1)$, then cosmic strings form in the transition at $T \sim f$; these become the boundaries of the walls which form later at $T \sim \mu$. In models with $N = 1$ minimum of the boson potential, there is one wall per string, and the wall-string system destroys itself before it dominates the energy density. We note that in this case the strings radiate bosons until the boson mass turns on at $T_r$, and the resulting cosmic energy density of string-radiated bosons may be larger than that calculated from ‘misalignment’ production in section III by a factor $\sim \ln(ft_r) \sim 100$. In order for the gravitational effects of the strings not to perturb the microwave background, the symmetry breaking scale is then restricted to roughly $f \lesssim 10^{17}$ GeV.

A final speculative possibility concerns the interesting cosmological role of an ultra-light field which has not yet become dynamical (is still frozen) or only became dynamical at moderate redshift (say, $z \sim 8 - 10$). The origin of the smallness of the cosmological constant is still shrouded in mystery, but let us suppose that there is some mechanism which sets the ultimate true vacuum energy density of the universe to zero. Then, at late times, classically we expect the vacuum energy to be dominated by the lightest field
which has not yet evolved to its ground state. If there is an ultralight pseudo-Goldstone boson which has yet to start oscillating, i.e., if it satisfies $T_r < T_0 = 2.4 \times 10^{-4}$ eV, then it will act as a (temporary) cosmological constant, with present vacuum energy of order $\rho_{\text{vac}} \sim \mu^4$. If $\mu \simeq 3 \times 10^{-3}$ eV, then $\Omega_{\text{vac}} \simeq 1$. This may have potential benefits such as increasing the age of the universe (for fixed value of the Hubble parameter) and boosting the large-scale transfer function for density fluctuations. The simplest way to implement this is to have $T_c < T_0$, which requires $\xi = T_c/\mu < 0.08$; one could relax this bound somewhat if the global symmetry breaking scale $f > m_{\text{Pl}}/\sqrt{24\pi}$.

Alternatively, if the field became dynamical at moderate redshift, it could lead to an epoch of cosmological loitering, in which the scale factor of the universe passes through a moderate Eddington-Lemaître-like coasting phase, allowing extra time for the growth of large-scale structure.

Acknowledgements

This work was supported in part by the DOE at Chicago and Fermilab and by NASA grant NAGW-1340 at Fermilab.
REFERENCES


14. This idea has also been considered by G. Starkman, private communication (unpublished).


16. To argue that the time scale for \( \phi \) to 'feel' the neutrinos is determined by the reaction rate for a \( \phi \) particle to scatter off of \( \nu \) would be tantamount to claiming that the time scale for gravitation to 'feel' the presence of a massive object is given by the reaction rate for single gravitons on single particles.


20. For models I and II, we divided cases 1 and 2 according to the value of \( \alpha(T_c) \).

For model III, it is more convenient to choose the value of \( \alpha(T_P) \), because the field amplitude may be Hubble damped at \( T_c \), so that \( \rho_{\phi}(T_c) \) may be substantially below \( \sim \mu^4 \).


FIGURE CAPTIONS

Figure 1: The temperature-dependent potential $V_T(\phi) = V(\phi) + \Delta V_T(\phi)$ for the pseudo-Nambu-Goldstone field in the $Z_2$ schizon model. The field exhibits a second order phase transition at a temperature $T \sim m_0$; this is characteristic of models we classify as Type II (from ref.(12)).

Figure 2: The temperature-dependent potential for the scalar field in a $Z_3$ schizon model; this is typical of $Z_N$ models with $N > 2$: the potential is flat at high temperature, developing a minimum at $T \sim m_0$. We classify these as Type I models (from ref.(12)).

Figure 3: The evolution of the scalar and baryon energy densities $\rho_\phi$ and $\rho_B$ as a function of the scale factor $a(t)$ for model I, case 1a. In this case, the scalar field presently dominates the energy density.

Figure 4: The evolution of $\rho_\phi$ and $\rho_B$ for model I, case 1b. Here, the scalar field never dominates the energy density of the universe.

Figure 5: The parameter space for models I and II, showing the different cases in the plane $f$ (GeV) vs. $\mu/\xi^3$ (eV).
\[
\frac{V(\phi) + \Delta V_\pi(0) - \Delta V_\tau(0)}{m_-^2} \]

\[
f^2 m_0^2 = -m_-
\]

\[
m_+ = 3m_-
\]

\[
T = 10m_-
\]

\[
T = 5m_-
\]

\[
T = 3m_-
\]

\[
T = 0
\]

\[
\phi / f \to \pi / 2
\]
\( \rho_\phi > \rho_B \), \( T_c > T_d > T_r \)

\[ f = \frac{M_{Pl}}{(24\pi)^{1/2}} \]

\( T_c > T_r, \rho_\phi < \rho_B \)

\( \Omega_\phi h^2 = 1 \)

\( T_r = T_c, \rho_\phi < \rho_B \)

\( \rho_c > \rho_\phi > \rho_B \)

\( \rho_\phi > \rho_c \), \( T_d > T_c = T_r \)