International Journal of Modern Physics A, Vol. 9, No. 13 (1994) 2117–2189

TOPOLOGICAL DEFECTS AND STRUCTURE FORMATION*

ROBERT H. BRANDENBERGER
Department of Physics, Brown University, Providence, RI 02912, USA

Received 6 December 1993

Topological defects are produced during phase transitions in the very early Universe. They arise in most unified theories of strong, weak and electromagnetic interactions. This article focuses on the role of topological defects in cosmology, with particular emphasis on the models of structure formation based on defects. The role of topological defects in baryogenesis is also reviewed.

1. Introduction

Topological defects[1] are inevitably produced during phase transitions in the very early Universe. One or several types of defects arise in most unified particle physics models of strong, weak and electromagnetic interactions. Unless they are inflated away[2] (i.e. unless there is a period of inflation after the phase transition during which the defects are produced), topological defects will play an important role in cosmology (for recent reviews see, for example, Refs. 3–6). The main aim of this article is to give a comprehensive pedagogical review of the role of topological defects in cosmology.

The reason why topological defects can play a role in structure formation in the early Universe is very simple. Topological defects carry energy. This energy leads to an extra attractive gravitational force, and hence the defects can act as seeds for cosmic structures. In particular, cosmic strings (for a recent review see, for example, Ref. 7) and global textures can lead to attractive scenarios for the formation of galaxies and large scale structure.

In cosmology there are many open questions. What is the origin of the observed large scale structure in the Universe?[9] Why are galaxies clustered? What explains the specific amplitudes and slopes of the galaxy and galaxy cluster correlation functions?[10] Do the models of structure formation predict cosmic microwave anisotropies in agreement with the recent COBE results?[11] There is a wealth of quantitative data in need of a consistent theoretical explanation.

Before 1980 there was no causal explanation starting from first principles for any of the above questions. Around 1980 there was a crucial realization which has

*Based on a Troisième Cycle lecture series given at the EPFL, Lausanne, Switzerland, 4/22-5/13 1993.

now led to the development of several self-consistent theories. The key point was to realize that in the very early Universe, matter cannot be described in terms of an ideal gas. At very high temperatures and energies, the matter content of the Universe must be described in terms of quantum fields. In "old cosmology" (prior to 1980), space–time was described by general relativity and matter as an ideal gas; in "new cosmology" space–time is still described by general relativity, but matter in terms of fields. Evidently, from a conceptual point of view the present theories are not satisfactory since they describe one sector in terms of classical physics, and the other quantum mechanically.

Since in "new cosmology" matter is described in terms of fields, the possibility of having phase transitions in the early Universe arises (for current particle physics models of matter this is a rather generic prediction). Hence, topological defects are formed, and they may play a role in structure formation.

Concrete issues which must be addressed and which are discussed in this article are:

- Why do defects form?
- What kind of defects arise?
- How to they evolve after formation?
- What are the mechanisms by which defects give rise to cosmic structures?
- Are there direct observational signatures for topological defect models?

Phase transitions and therefore topological defects arise in many areas of physics. In metals and other solids, defects arise during the fluid to solid phase transition.[12] The resulting crystal defects can be point, line or planar defects. In liquid crystals, defects arise during the transition from the disordered phase to the ordered phase.[13] There have recently been some very ingenious studies of the dynamics of defect formation in liquid crystals.[14] Vortex lines also appear in the transition from the normal to the superfluid phase of ^{3}He,[15] and in the normal to superconducting transition in superconducting materials.[16]

The above are all examples of topological defects in condensed matter systems as opposed to relativistic field theory. There are important similarities and differences between the two classes of systems. In condensed matter systems the dynamics is generally friction-dominated and nonrelativistic, whereas in flat space–time the dynamics of topological defects in field theory and cosmology is relativistic. Hence, condensed matter systems give a good analogy for the topology and formation of defects, but not for the dynamics.

This article begins with a survey of the successes and problems of standard "big bang" cosmology. In Sec. 3, phase transitions and the origin and nature of topological defects are reviewed. In Sec. 4, the formation and evolution of topological defects (particularly cosmic strings) in the early Universe are treated. Sections 5–8 deal with topological defect models for the formation of structure. First, an introduction to the basic mechanisms and a survey of the most relevant data are given. Then, the cosmic string and global texture models of structure formation are

developed, and in Sec. 8 the predictions of the two models are compared with those of inflationary Universe scenarios. Sections 9 and 10 deal with the microphysics of cosmic strings. Baryon-number-violating scattering processes based on strings are discussed in Sec. 9, and the role of topological defects in baryogenesis is analyzed in the final section. This article is based in part on lectures[17] given at the 7th Swieca Summer School in Particles and Fields (see also Ref. 18).

A word concerning notation. Unless mentioned otherwise, units in which $c = \hbar = k_B = 1$ are employed. The metric $g_{\mu\nu}$ is taken to have signature $(+,-,-,-)$. Greek indices run over space and time; Latin ones over spatial indices only. The Hubble expansion rate is $H(t) = \dot{a}(t)/a(t)$, where $a(t)$ is the scale factor of a Friedmann–Robertson–Walker (FRW) Universe. The present value of H is $h \times 100$ km s^{-1} Mpc^{-1}, where $0.4 < h < 1$. Unless stated otherwise, the value of h is taken to be 0.5. As usual, G and $m_{\rm Pl}$ stand for Newton's constant and Planck mass, respectively, and $z(t)$ is the redshift at time t. Distances are measured in Mpc, where 1 Mpc is $3.26 \cdot 10^6$ light years. t_0 usually denotes the present time, and T_0 stands for the present temperature.

2. Standard Cosmology: Successes and Problems

The standard big bang cosmology rests on three theoretical pillars: the cosmological principle, Einstein's general theory of relativity and a perfect fluid description of matter.

The cosmological principle[19] states that on large distance scales the Universe is homogeneous. From an observational point of view this is an extremely non-trivial statement. On small scales the Universe looks extremely inhomogeneous. The inhomogeneities of the solar system are obvious to everyone, and even to the naked eye it is apparent that stars are not randomly distributed. They are bound into galaxies, dynamical entities whose visible radius is about 10^4 pc. Telescopic observations show that galaxies are not randomly distributed either. Dense clumps of galaxies can be identified as Abell clusters. In turn, Abell cluster positions are correlated to produce the large scale structure dominated by sheets (or filaments) with the typical scale 100 Mpc observed in recent redshift surveys.[9] Until recently, every new survey probing the Universe to greater depth revealed new structures on the scale of the sample volume. In terms of the visible distribution of matter there was no evidence for large scale homogeneity. This situation changed in 1992 with the announcement[20] that a new redshift survey complete to a depth of about 800 Mpc had discovered no prominent structures on scales larger than 100 Mpc. This is the first observational evidence from optical measurements in favor of the cosmological principle. However, to put this result in perspective we must keep in mind that the observed isotropy of the CMB temperature to better than 10^{-5} on large angular scales has been excellent evidence for the validity of the cosmological principle.

The second theoretical pillar is general relativity, the theory which determines the dynamics of the Universe. According to general relativity, space–time is a

smooth manifold. Together with the cosmological principle this tells us that it is possible to choose a family of hypersurfaces with maximal symmetry. These are the homogeneous constant time hypersurfaces. The metric of these surfaces is[21]

$$ds^2 = a(t)^2 \left[\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2\theta \, d\varphi^2) \right] \tag{2.1}$$

when using spherical polar coordinates. The constant k is $+1$, 0 or -1 for closed, flat or open surfaces respectively. The function $a(t)$ is the scale factor of the Universe. By a coordinate choice, it could be set equal to 1 at any given time. However, the time dependence of $a(t)$ indicates how the spatial sections evolve as a function of time. The full space–time metric is

$$ds^2 = dt^2 - a(t)^2 \left[\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2\theta \, d\varphi^2) \right] . \tag{2.2}$$

According to the Einstein equivalence principle, test particles move on geodesics with respect to the metric given by ds^2. This implies that the peculiar velocity $\underline{v}_p$ obeys the equation

$$\dot{\underline{v}}_p + \frac{\dot{a}}{a} \underline{v}_p = 0 . \tag{2.3}$$

Here,

$$\underline{v}_p \equiv a(t) \frac{d\underline{x}^c}{dt} , \tag{2.4}$$

$\underline{x}^c$ being the comoving coordinates r, θ and φ. Equation (2.3) implies that in an expanding Universe, the peculiar velocity $\underline{v}_p$ decreases:

$$\underline{v}_p(t) \sim a^{-1}(t) . \tag{2.5}$$

Therefore, trajectories with constant $\underline{x}^c$ are geodesics and correspond to particles at rest. The velocity $\underline{v}_p$ is the physical velocity relative to the expansion of the Universe (see Fig. 1).

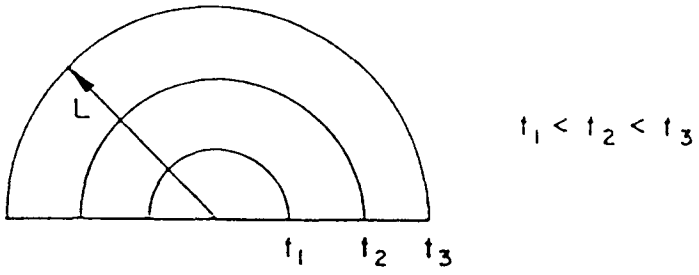

Fig. 1. Sketch of the expanding Universe. Concentric circles indicate space at fixed time, time increasing as the radius gets larger. Points at rest have constant comoving coordinates. Their world lines are straight lines through the origin (e.g. L).

From observations[22] it is known that the Universe is at present expanding. Looking at distant galaxies, we detect a redshift z of light which increases linearly with the distance d of the object:

$$z(d) \sim d\,, \tag{2.6}$$

where the redshift z is defined by

$$z = \frac{\lambda_0}{\lambda_e} - 1\,, \tag{2.7}$$

with λ_0 and λ_e being the wavelengths measured by us and by the source. The relationship (2.6) is explained by taking galaxies to be at rest in comoving coordinates, and $a(t)$ to be increasing. In this case, for $z \ll 1$

$$z \simeq H(t_0)d \quad \left[H(t) = \frac{\dot{a}}{a}(t)\right]\,, \tag{2.8}$$

t_0 being the present time.

The most important consequence of general relativity for the history of the Universe is that it relates the expansion rate to the matter content. The Einstein field equations follow from the action

$$S = \int d^4x\,\sqrt{-g}(R - 16\pi G\mathcal{L}_M)\,, \tag{2.9}$$

where R is the Ricci scalar curvature, g is the determinant of the metric, and $\mathcal{L}_M$ is the Lagrange density for matter. Evaluating the equations of motion obtained by varying (2.9) with respect to $g_{\mu\nu}$ for a metric of the form (2.2) leads to the famous FRW equations

$$\left(\frac{\dot{a}}{a}\right)^2 - \frac{k}{a^2} = \frac{8\pi G}{3}\rho\,, \tag{2.10}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p)\,. \tag{2.11}$$

Note that for fixed energy density ρ, the evolution of $a(t)$ depends in an important way on the pressure p.

Besides explaining Hubble's redshift–distance relation (2.6), standard big bang cosmology makes two key quantitative predictions: the existence of a black body cosmic microwave background[23] and nucleosynthesis, the generation of light elements.[23–25]

Consider ordinary matter made up of atoms in an expanding Universe. The energy density in matter scales as $a(t)^{-3}$, and the temperature $T(t)$ as $a^{-1}(t)$. Thus, standard cosmology predicts that as we go back in time, the Universe was warmer. In particular, at a critical temperature $T_{\rm rec}$, matter becomes ionized. Before $t_{\rm rec}$ (the time corresponding to $T_{\rm rec}$) the Universe was opaque to photons; after $t_{\rm rec}$ it was transparent. To be more precise, $T_{\rm rec}$ is the temperature when photons fall

out of thermal equilibrium. Thereafter they propagate without scattering. The black body nature of the spectrum of photons is maintained, but the temperature redshifts. Hence, the big bang model predicts[23,26] a black body spectrum of photons of temperature

$$T_0 = T_{\rm rec} z_{\rm rec}^{-1} \tag{2.12}$$

where the cosmological redshift $z(t)$ is given by

$$1 + z(t) = \frac{a(t_0)}{a(t)} \tag{2.13}$$

and $z_{\rm rec} = z(t_{\rm rec}) \sim 10^3$.

In 1965, Penzias and Wilson[27] discovered this remnant black body radiation at a temperature of about 3° K. Since the spectrum peaks in the microwave region it is now called the CMB (cosmic microwave background). Recent satellite (COBE)[28] and rocket[29] experiments have confirmed the black body nature of the CMB to very high accuracy. The temperature is $2.73°\ {\rm K} = T_0$.

Given the existence of the CMB, we know that matter has two components: dust [with energy density $\rho_m(t)$] and radiation [with density $\rho_r(t)$]. At the present time t_0, $\rho_m(t) \gg \rho_r(t)$. The radiation energy density is determined by T_0, and the matter energy density can be estimated by analyzing the dynamics of galaxies and clusters and using the virial theorem. However, since

$$\rho_m(t) \sim a(t)^{-3}\,, \quad \rho_r(t) \sim a(t)^{-4}\,, \tag{2.14}$$

as we go back in time the fraction of energy density in radiation increases, and the two components become equal at $t_{\rm eq}$, the time of equal matter and radiation. The corresponding redshift is

$$z_{\rm eq} = \Omega h_{50}^{-2} 10^4\,, \tag{2.15}$$

where

$$\Omega = \frac{\rho}{\rho_{\rm cr}}(t_0)\,, \tag{2.16}$$

$\rho_{\rm cr}$ being the density for a spatially flat Universe (the critical density), and h_{50} in the value of H in units of 50 km s^{-1} Mpc^{-1}.

The time $t_{\rm eq}$ is important for structure formation.[30] It is only after $t_{\rm eq}$ that perturbations on scales smaller than the Hubble radius $H^{-1}(t)$ can grow. Before then, the radiation pressure prevents growth. A temperature–time plot of the early Universe is sketched in Fig. 2. Note that $t_{\rm eq} < t_{\rm rec}$.

The second quantitative prediction of standard big bang cosmology concerns nucleosynthesis. Above a temperature of about $10^{9°}$ K, the nuclear interactions are sufficiently fast to prevent neutrons and protons from fusing. However, below that temperature, it is thermodynamically favorable[23–25] for neutrons and protons to fuse and form deuterium, helium 3, helium 4 and lithium 7 through a long and interconnected chain of reactions. The resulting light element abundances depend

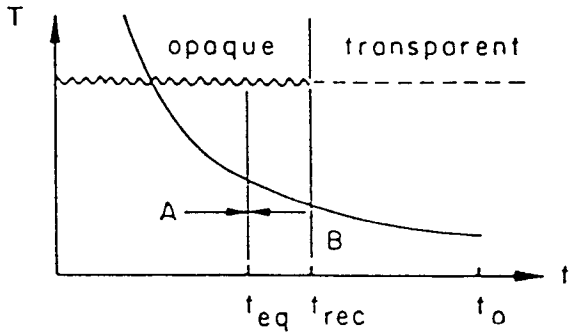

Fig. 2. Temperature–time diagram of standard big bang cosmology. The present time, time of decoupling and time of equal matter and radiation are t_0, t_{rec} and t_{eq}, respectively. The Universe is radiation-dominated before t_{eq} (Region A), and matter-dominated in Region B. Before and after t_{rec} the Universe is opaque and transparent, respectively, to microwave photons.

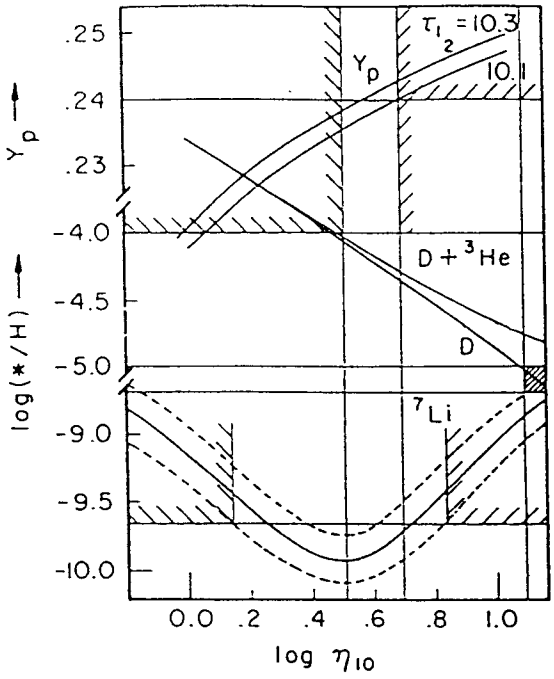

Fig. 3. Light element abundances as a function of the baryon to entropy ratio η. Y_p is the ^{4}He mass fraction, calculated assuming three light neutrino species, for two different values $\tau_{1/2}$ of the neutrino half life (in minutes). The abundances of $D+^3$He and of ^{7}Li are shown as ratios of their number density relative to the number density of H. The horizontal lines indicate limits from observations: upper limit of 10^{-4} for the $D+^3$He abundance, and upper limits on ^{7}Li from observations of dwarf stars. The ^{7}Li curve is shown with $\pm 2\sigma$ errors. Combining ^{4}He and ^{3}He $+D$ limits leaves only a small window for η which is allowed. A major success of primordial nucleosynthesis is that the ^{7}Li abundance matches well. Note that η_{10} is η in units of 10^{-10}.

sensitively on the expansion rate of the Universe and on Ω_B, the fraction of energy density ρ_B at present in baryons relative to the critical density ρ_c. In Fig. 3, recent[31] theoretical calculations of the abundances are shown and compared with observations. Demanding agreement with all abundances leaves only a narrow window:

$$3 \times 10^{-10} < \eta < 10^{-9}, \tag{2.17}$$

where η is the ratio of baryon number density n_B to entropy density s:

$$\eta = \frac{n_B}{s}. \tag{2.18}$$

From (2.17), it follows that Ω_B is constrained:

$$0.01 < \Omega_B h^2 < 0.035. \tag{2.19}$$

In particular, if the Universe is spatially flat, there must be nonbaryonic dark matter. We will return to the dark matter issue shortly.

In summary: the three observational pillars of standard big bang cosmology are Hubble's redshift–distance relation, the existence and black body spectrum of the CMB, and the concordance between observed and theoretically determined light element abundances.

Standard big bang cosmology is faced with several important problems: the age, dark matter, homogeneity, flatness, and formation of structure problems (this list is not exhaustive,[32] but contains what I consider to be the key problems). In addition, standard big bang cosmology does not explain the small value of the cosmological constant — a problem which no present cosmological model addresses (see, however, Ref. 33 for some recent progress on this issue). Of the above five problems, only the first is a possible conflict of the theory with observations. The other four are questions which are left unanswered by the theory. Extensions of the standard model are needed to address these issues.

Globular cluster ages have been estimated to lie in the range $(13-18) \times 10^9$ yr.[34] Nuclear cosmochronology gives an age of structures in the range $(10-20) \times 10^9$ yr. However, big bang cosmology (in the absence of a cosmological constant) predicts the age

$$\tau = \frac{7}{h} \cdot 10^2 \text{ yr}. \tag{2.20}$$

Thus, theory and observations are compatible only if $h < 0.55$.

The dark matter problem[35] problem has various aspects. There is more matter in galaxies than is visible in stars. This follows if one studies galactic velocity rotation curves and observes that the velocity remains constant beyond the visible radius of the galaxy. Whereas[34] the contribution of stars to Ω is less than 0.01, galaxies as a whole contribute > 0.02 to Ω. The second level of the dark matter problem lies in clusters. By studying cluster dynamics it can be inferred that the contribution to Ω exceeds 0.1. Finally, observations on the largest scales give an even larger contribution to Ω. From Virgo infall it follows that $\Omega > 0.3$, and from

large scale velocity measurements (POTENT)[36] or from infrared galaxy (IRAS) surveys[37] it follows that

$$0.5 < \Omega < 3\,. \tag{2.21}$$

Inflation predicts that $\Omega \simeq 1$, unless fine-tuned initial conditions are chosen.[38]

The bottom line of the dark matter problem is that more mass is observed than can be in light (hence the name "dark matter"), and that most of this missing matter must be nonbaryonic [since by (2.19) $\Omega_B < 0.14$]. Note that there must be some baryonic dark matter, and that all of the galactic dark matter could be baryonic. This comment will be relevant when we discuss cosmic strings and galaxy formation in Sec. 6. If $h = 0.5$, then most of the cluster dark matter must be nonbaryonic. Standard cosmology does not address the issue of what the dark matter is. New cosmology can provide a solution. Quantum field theory models of matter in the Universe give rise to several candidate particles which could constitute the dark matter.[39]

The last three problems are the classic problems which the inflationary Universe scenario addresses.[2] In Fig. 4, the homogeneity (or horizon) problem is illustrated. As is sketched, the region $\ell_p(t_{\rm rec})$ over which the CMB is observed to be homogeneous to better than one part in 10^5 is much larger than the forward light cone $\ell_f(t_{\rm rec})$ at $t_{\rm rec}$, which is the maximal distance over which microphysical forces could have caused the homogeneity:

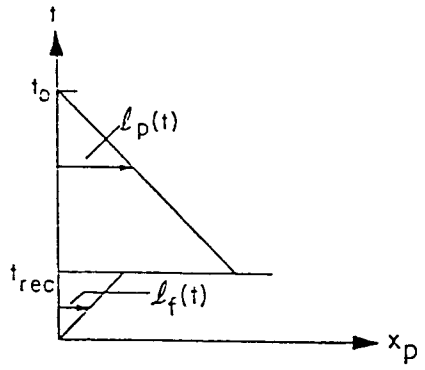

Fig. 4. A sketch of the homogeneity problem: the past light cone $\ell_p(t)$ at the time $t_{\rm rec}$ of last scattering is much larger than the forward light cone $\ell_f(t)$ at $t_{\rm rec}$.

$$\ell_p(t_{\rm rec}) = a(t_{\rm rec}) \int_{t_{\rm rec}}^{t_0} dt\, a^{-1}(t) \simeq 3t_0 \left[1 - \left(\frac{t_{\rm rec}}{t_0}\right)^{1/3}\right] a(t_{\rm rec})\,, \tag{2.22}$$

$$\ell_f(t_{\rm rec}) = a(t_{\rm rec}) \int_0^{t_{\rm rec}} dt\, a^{-1}(t) \simeq 3t_0^{2/3}\, t_{\rm rec}^{1/3} a(t_{\rm rec})\,. \tag{2.23}$$

From the above equations it is obvious that $\ell_p(t_{\rm eq}) \gg \ell_f(t_{\rm eq})$. Hence, standard cosmology cannot explain the observed isotropy of the CMB.

In standard cosmology and in an expanding Universe, $\Omega = 1$ is an unstable fixed point. This can be seen as follows. For a spatially flat Universe ($\Omega = 1$)

$$H^2 = \frac{8\pi G}{3}\rho_c\,, \tag{2.24}$$

whereas for a nonflat Universe

$$H^2 + \varepsilon T^2 = \frac{8\pi G}{3}\rho\,, \tag{2.25}$$

with

$$\varepsilon = \frac{k}{(aT)^2}\,. \tag{2.26}$$

The quantity ε is proportional to $s^{-2/3}$, where s is the entropy density. Hence, in standard cosmology, ε is constant. Combining (2.24) and (2.25) gives

$$\frac{\rho - \rho_c}{\rho_c} = \frac{3}{8\pi G}\frac{\varepsilon T^2}{\rho_c} \sim T^{-2}\,. \tag{2.27}$$

Thus, as the temperature decreases, $\Omega - 1$ increases. In fact, in order to explain the present small value of $\Omega - 1 \sim \mathcal{O}(1)$, the initial energy density had to be extremely close to critical density. For example, at $T = 10^{15}$ GeV, (2.27) implies that

$$\frac{\rho - \rho_c}{\rho_c} \sim 10^{-50}\,. \tag{2.28}$$

What is the origin of these fine-tuned initial conditions? This is the flatness problem of standard cosmology.

The last problem of the standard cosmological model I will mention is the "formation of structure problem." Observations indicate that galaxies and even clusters of galaxies have nonrandom correlations on scales larger than 50 Mpc.[9,40,41] This scale is comparable to the comoving horizon at $t_{\rm eq}$. Thus, if the initial density perturbations were produced much before $t_{\rm eq}$, the correlations could not be explained by a causal mechanism. Gravity alone is, in general, too weak to build up correlations on the scale of clusters after $t_{\rm eq}$ (see, however, the explosion scenario of Ref. 42). Hence, the two questions of what generates the primordial density perturbations and what causes the observed correlations, do not have an answer in the context of standard cosmology. This problem is illustrated in Fig. 5.

In 1981, based on previous work by many other people (see, for example, Refs. 43–45 for a detailed bibliography of early works), Guth realized[2] that having a sufficiently long phase in the very early Universe during which the scale factor expands exponentially,

$$a(t) \sim e^{Ht}\,, \tag{2.29}$$

can potentially solve the last three problems listed above. This phase of exponential expansion is called the de Sitter or inflationary phase.

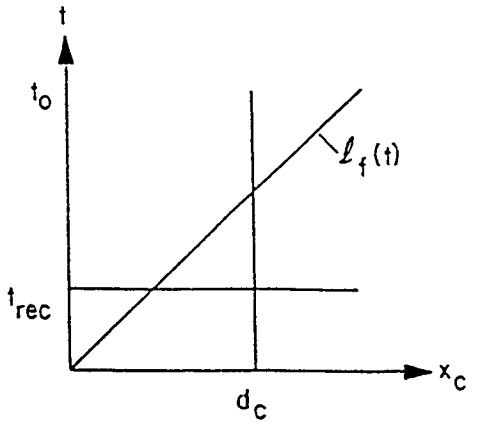

Fig. 5. A sketch (conformal separation vs. time) of the formation of structure problem: the comoving separation d_c between two clusters is larger than the forward light cone at time t_{rec}.

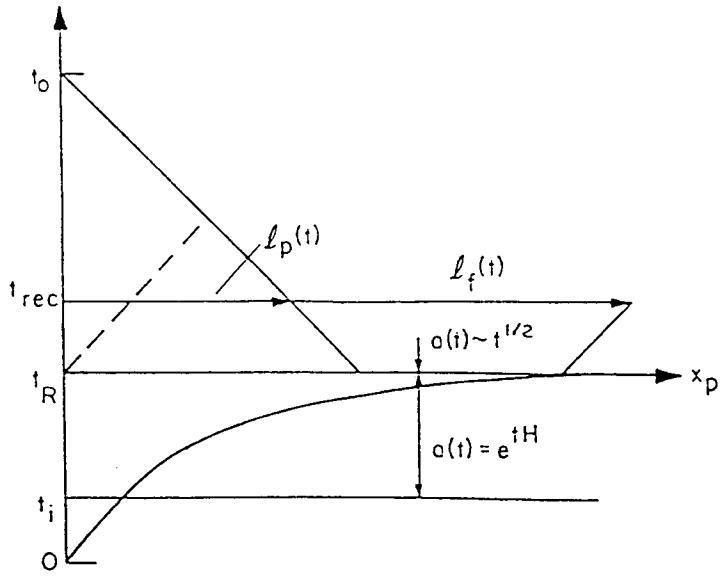

Fig. 6. A sketch (physical coordinates vs. time) of the solution of the homogeneity problem. Due to exponential expansion the forward light cone is larger than the past light cone at t_{rec}. The dashed line is the apparent horizon or Hubble radius.

Figure 6 sketches how a period of inflation can solve the homogeneity problem. t_i denotes the onset of inflation, t_R the end. $\Delta t = t_R - t_i$ is the period of inflation. During inflation, the forward light cone increases exponentially compared to a model without inflation, whereas the past light cone is not affected for $t \geq t_R$. Hence,

provided Δt is sufficiently large, $\ell_f(t_{\rm rec})$ will be greater than $\ell_p(t_{\rm rec})$. The condition on Δt depends on the temperature T_R corresponding to time t_R, the temperature of reheating. Demanding that $\ell_f(t_R) > \ell_p(t_R)$ we find, using (2.22) and (2.23), the criterion

$$e^{\Delta t H} \geq \frac{\ell_p(t_R)}{\ell_f(t_R)} \simeq \left(\frac{t_0}{t_R}\right)^{1/2} = \left(\frac{T_R}{T_0}\right) \sim 10^{27} \tag{2.30}$$

for $T_R \sim 10^{14}$ GeV and $T_0 \sim 10^{-13}$ GeV (the present microwave background temperature). Thus, in order to solve the homogeneity problem, a period of inflation with

$$\Delta t > 50\, H^{-1} \tag{2.31}$$

is required.

Inflation also can solve the flatness problem.[2,46] The key point is that the entropy density s is no longer constant. As will be explained later, the temperatures at t_i and t_R are essentially equal. Hence, the entropy increases during inflation by a factor of $\exp(3H\Delta t)$. Thus, ϵ decreases by $\exp(-2H\Delta t)$. With the numbers used in (2.30),

$$\epsilon_{\rm after} \sim 10^{-54} \epsilon_{\rm before}\,. \tag{2.32}$$

Hence, $(\rho - \rho_c)/\rho$ can be of order 1 both at t_i and at the present time. In fact, if inflation occurs at all, the theory then predicts that at the present time $\Omega = 1$ to a high accuracy (now $\Omega < 1$ would require special initial conditions).

What was said above can be rephrased geometrically: during inflation, the curvature radius of the Universe — measured on a fixed physical scale — increases exponentially. Thus, a piece of space looks essentially flat after inflation even if it had measurable curvature before.

Most important, inflation provides a mechanism which in a causal way generates the primordial perturbations required for galaxies, clusters and even larger objects.[47–50] In inflationary Universe models, the "apparent" horizon $3t$ and the "actual" horizon (the forward light cone) do not coincide at late times. Provided (2.30) is satisfied, then (as sketched in Fig. 7) all scales within our apparent horizon were inside the actual horizon since t_i. Thus, it is in principle possible to have a causal generation mechanism for perturbations.

As will be shown at the beginning of the following section, in order to obtain inflation it is necessary to describe matter in terms of field theory. In particular, there must be a phase transition in the early Universe during which the equation of state of matter changes, leading to different expansion rates of the Universe before and after the transition.

In order to obtain sufficient inflation,[2] the phase transition must be slow. For particle-physics-motivated matter theories this is hard to achieve, as will again be shown in the next section. However, matter theories generically predict phase transitions, and the latter lead to the formation of topological defects.

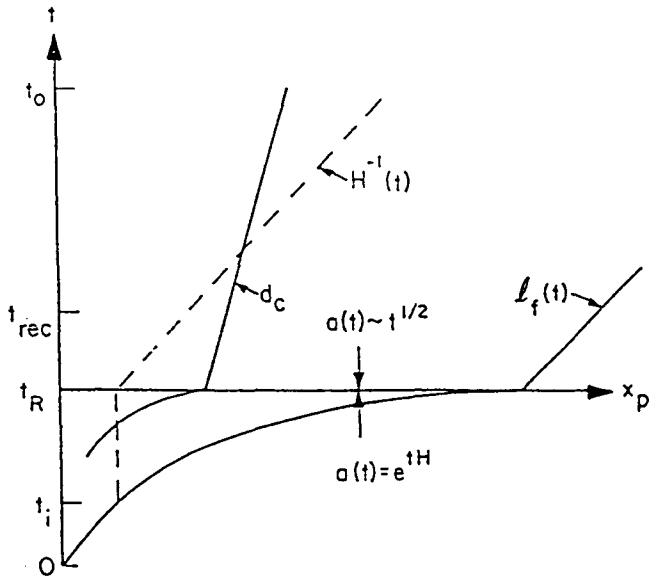

Fig. 7. A sketch (physical coordinates vs. time) of the solution of the formation of structure problem. The separation d_c between two clusters is always smaller than the forward light cone. The dashed line is the Hubble radius $H^{-1}(t)$.

Topological defect models provide a second mechanism for producing structures in the Universe on large scales. To be specific, consider a model giving rise to cosmic strings (see Sec. 4). The strings produced during the phase transition will form a network of random walks. This network has an infinite extent and gives rise to nonrandom correlations on all scales for point objects seeded by the strings (see Sec. 6). This and similar topological defect models for structure formation will be the focus of a substantial part of this article.

3. Phase Transitions and Topological Defects

3.1. *Scalar fields and cosmology*

Let us for a moment return to the inflationary Universe. As stated in (2.29), the requirement for inflation is to have a time interval during which the scale factor $a(t)$ expands exponentially. Recall from (2.10) and (2.11) that for a spatially flat FRW Universe the Einstein equations reduce to

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho, \tag{3.1}$$

$$2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 = -8\pi G p. \tag{3.2}$$

However, by the use of $a(t) = e^{tH}$, the left hand side of (3.2) can be evaluated using (3.1) to give

$$2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 = 8\pi G\rho . \tag{3.3}$$

In order for (3.2) and (3.3) to agree, the equation of state of matter must be

$$p = -\rho . \tag{3.4}$$

Therefore, "new cosmology" (i.e. description of matter in terms of fields) is required in order to obtain inflation.

It is not hard to show that an equation of state like (3.4) with negative pressure can be obtained if matter is described in terms of scalar fields. Consider the Lagrangian $L(\varphi)$ for a theory of a scalar field $\varphi(\underline{x}, t)$:

$$\mathcal{L}(\varphi) = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - V(\varphi) . \tag{3.5}$$

Given the Lagrangian, the energy–momentum tensor $T_{\mu\nu}$ can be determined as in any classical field theory (see Ref. 51). In a Universe with FRW metric

$$g_{\mu\nu} = \mathrm{diag}\,(1, -a^2(t), -a^2(t), -a^2(t)) \tag{3.6}$$

we obtain for $\rho = T_{00}$ and $p = \frac{1}{3}\sum_{i=1}^{3} T_{ii}$

$$\begin{aligned} \rho(\underline{x}, t) &= \frac{1}{2}\dot{\varphi}^2(\underline{x}, t) + \frac{1}{2}a^{-2}(\nabla\varphi)^2 + V(\varphi) , \\ p(\underline{x}, t) &= \frac{1}{3}\sum_{i=1}^{3} T_{ii} = \frac{1}{2}\dot{\varphi}^2(\underline{x}, t) - \frac{1}{6}a^{-2}(\nabla\varphi)^2 - V(\varphi) . \end{aligned} \tag{3.7}$$

Thus, if $\varphi(\underline{x}, t_i) = \text{const}$ and $\dot{\varphi}(\underline{x}, t_i) = 0$ at some initial time t_i and $V(\varphi(\underline{x}, t_i)) > 0$, then the equation of state becomes $p = -\rho$ and leads to inflation.

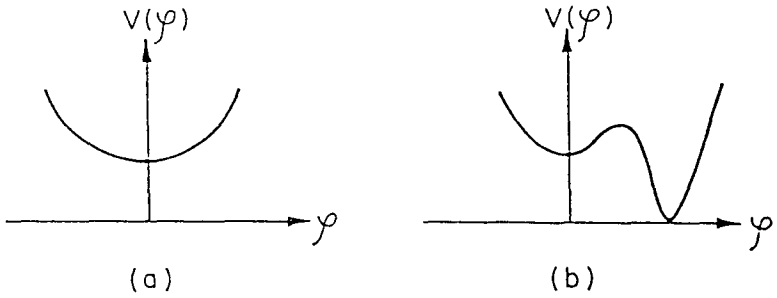

Fig. 8. A sketch of two potentials which can give rise to inflation.

Two examples which give inflation are shown in Fig. 8. In (a), inflation occurs at the stable fixed point $\varphi(\underline{x}, t_i) = 0 = \dot{\varphi}(\underline{x}, t_i)$. However, this model is ruled out by observation: the inflationary phase has no ending. $V(0)$ acts as a permanent nonvanishing cosmological constant. In (b), a finite period of inflation can arise if $\varphi(\underline{x})$ is trapped at the local minimum $\varphi = 0$ with $\dot{\varphi}(\underline{x}) = 0$. However, in this case $\varphi(\underline{x})$ can make a sudden transition at some time $t_R > t_i$ through the potential barrier and move to $\varphi(\underline{x}) = a$. Thus, for $t_i < t < t_R$ the Universe expands exponentially, whereas for $t > t_R$ the contribution of φ to the expansion of the Universe vanishes and we get the usual FRW cosmology. There are two obvious questions: How does the transition occur and why should the scalar field have $V(\varphi) = 0$ at the global minimum? In the following section the first question will be addressed. The second question is part of the cosmological constant problem, for which there is as yet no convincing explanation. Before studying the dynamics of the phase transition, we need to digress and discuss finite temperature effects.

3.2. *Finite temperature field theory*

The evolution of particles in vacuum and in a thermal bath are very different. Similarly, the evolution of fields changes when coupled to a thermal bath. Under certain conditions, the changes may be absorbed in a temperature-dependent potential, the finite temperature effective potential.[52] Here, a heuristic derivation of this potential will be given. (See Ref. 53 or the original articles[52] for the actual derivation.)

We assume that the scalar field $\varphi(\underline{x}, t)$ is coupled to a thermal bath which is represented by a second scalar field $\psi(\underline{x}, t)$ in thermal equilibrium. The Lagrangian for φ is

$$\mathcal{L} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - V(\varphi) - \frac{1}{2}\hat{\lambda}\varphi^2\psi^2\,, \tag{3.8}$$

where $\hat{\lambda}$ is a coupling constant. The action from which the equations of motion are derived is

$$S = \int d^4x\,\sqrt{-g}\mathcal{L}\,, \tag{3.9}$$

where g is the determinant of the metric (3.6). The resulting equation of motion for $\varphi(\underline{x}, t)$ is

$$\ddot{\varphi} + 3H\dot{\varphi} - a^{-2}\nabla^2\varphi = -V'(\varphi) - \hat{\lambda}\psi^2\varphi\,. \tag{3.10}$$

If ψ is in thermal equilibrium, we may replace ψ^2 by its thermal expectation value $\langle\psi^2\rangle_T$. Now,

$$\langle\psi^2\rangle_T \sim T^2\,, \tag{3.11}$$

which can be seen as follows: in thermal equilibrium, the energy density of ψ equals that of one degree of freedom in the thermal bath. In particular, the potential energy density $V(\psi)$ of ψ is of that order of magnitude. Let

$$V(\psi) = \lambda_\psi\psi^4 \tag{3.12}$$

with a coupling constant λ_ψ, which we take to be of order 1 (if λ_ψ is too small, ψ will not be in thermal equilibrium). Since the thermal energy density is proportional to T^4, (3.11) follows. (3.10) can be rewritten as

$$\ddot{\psi} + 3H\dot{\varphi} - a^{-2}\nabla^2\varphi = -V_T'(\varphi)\,, \tag{3.13}$$

where

$$V_T(\varphi) = V(\varphi) + \frac{1}{2}\hat{\lambda}T^2\varphi^2 \tag{3.14}$$

is called the finite temperature effective potential. Note that in (3.14), $\hat{\lambda}$ has been rescaled to absorb the constant of proportionality in (3.11).

These considerations will now be applied to Example A, a scalar field model with potential

$$V(\varphi) = \frac{1}{4}\lambda(\varphi^2 - \eta^2)^2 \tag{3.15}$$

(η is called the scale of symmetry breaking). The finite temperature effective potential becomes (see Fig. 9)

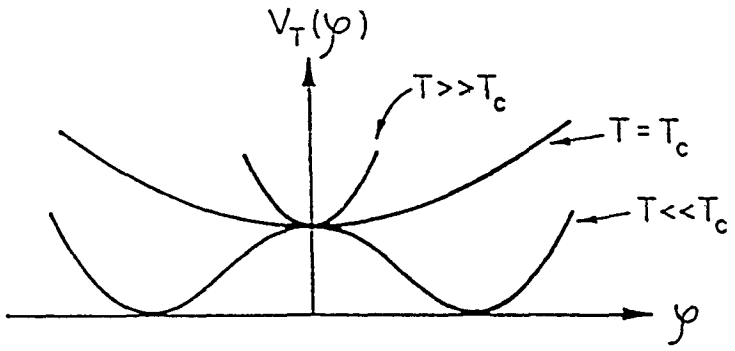

Fig. 9. The finite temperature effective potential for Example A.

$$V_T(\varphi) = \frac{1}{4}\lambda\varphi^4 - \frac{1}{2}\left(\lambda\eta^2 - \hat{\lambda}T^2\right)\varphi^2 + \frac{1}{4}\lambda\eta^4\,. \tag{3.16}$$

For very high temperatures, the effective mass term is positive and hence the energetically favorable state is $\langle\varphi\rangle = 0$. For very low temperatures, on the other hand, the mass term has a negative sign which leads to spontaneous symmetry breaking. The temperature at which the mass term vanishes defines the critical temperature T_c:

$$T_c = \hat{\lambda}^{-1/2}\lambda^{1/2}\eta\,. \tag{3.17}$$

As Example B, consider a theory with potential

$$V(\varphi) = \frac{1}{4}\varphi^4 - \frac{1}{3}(a+b)\varphi^3 + \frac{1}{2}ab\varphi^2 \tag{3.18}$$

with $\frac{1}{2}a > b > 0$. The finite temperature effective potential is obtained by adding $\frac{1}{2}\hat{\lambda}T^2\varphi^2$ to the right hand side of (3.18). $V_T(\varphi)$ is sketched in Fig. 10 for various values of T. The critical temperature T_c is defined as the temperature when the two minima of $V_T(\varphi)$ become degenerate.

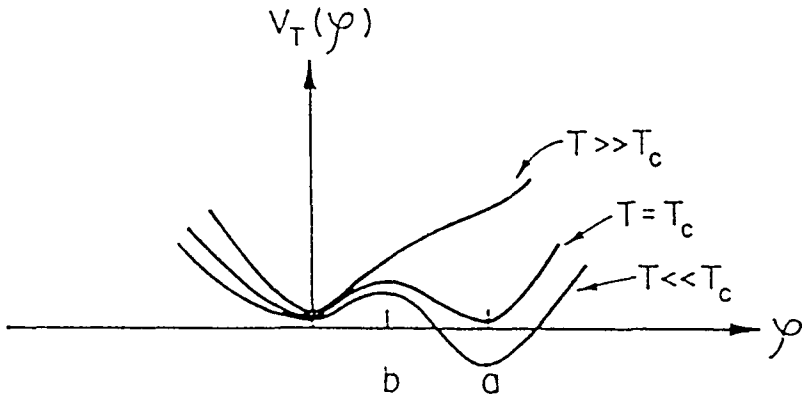

Fig. 10. The finite temperature effective potential for Example B.

It is important to note that the use of finite temperature effective potential methods is legitimate only if the system is in thermal equilibrium. This point was stressed in Refs. 54 and 55, although the fact should be obvious from the derivation given above. To be more precise, we require the ψ field to be in thermal equilibrium and the coupling constant $\hat{\lambda}$ of (3.8), which mediates the energy exchange between the φ and ψ fields, to be large. However (see, for example, Ref. 18), observational constraints stemming from the amplitude of the primordial energy density fluctuation spectrum force the self-coupling constant λ of φ to be extremely small. Since at one-loop order the interaction term $\frac{1}{2}\hat{\lambda}\varphi^2\psi^2$ induces contributions to λ, it is unnatural to have λ very small and $\hat{\lambda}$ unsuppressed. Hence, in many inflationary Universe models — particularly in new inflation[56] and in chaotic inflation[54] — finite temperature effective potential methods are inapplicable.

3.3. *Phase transitions*

The temperature dependence of the finite temperature effective potential in quantum field theory leads to phase transitions in the very early Universe. These transitions are either first or second order.

Example A of the previous section provides a model in which the transition is second order (see Fig. 9). For $T \gg T_c$, the expectation value of the scalar field φ vanishes at all points $\underline{x}$ in space:

$$\langle\varphi(\underline{x})\rangle = 0\,. \tag{3.19}$$

For $T < T_c$, this value of $\langle\varphi(\underline{x})\rangle$ becomes unstable and $\langle\varphi(\underline{x})\rangle$ evolves smoothly in time to a new value $\pm\eta$. The direction is determined by thermal and quantum fluctuations and is therefore not uniform in space. There will be domains of average radius $\xi(t)$ in which $\langle\varphi(\underline{x})\rangle$ is coherent. By causality, the coherence length is bounded from above by the horizon. However, typical values of $\xi(t)$ are proportional to $\lambda^{-1}\eta^{-1}$ if φ was in thermal equilibrium before the phase transition.[1]

In condensed matter physics, a transition of the above type is said to proceed by spinodal decomposition,[57] triggered by a rapid quench.

In Example B of the previous section (see Fig. 10), the phase transition is first order. For $T > T_c$, the expectation value $\langle\varphi(x)\rangle$ is approximately 0, the minimum of the high temperature effective potential. Provided the zero temperature potential has a sufficiently high barrier separating the metastable state $\varphi = 0$ from the global minimum (compared to the energy density in thermal fluctuations at $T = T_c$), $\varphi(\underline{x})$ will remain trapped at $\varphi = 0$ also for $T < T_c$. In the notation of Ref. 58, the field φ is trapped in the false vacuum. After some time (determined again by the potential barrier), the false vacuum will decay by quantum tunneling.

Tunneling in quantum field theory was discussed in Refs. 58–61 (for reviews see, for example, Refs. 62 and 53). The transition proceeds by bubble nucleation. There is a probability per unit time and volume that at a point $\underline{x}$ in space a bubble of "true vacuum" $\varphi(\underline{x}) = a$ will nucleate. The nucleation radius is microscopical. As long as the potential barrier is large, the bubble radius will increase with the speed of light after nucleation. Thus, a bubble of $\varphi = a$ expands in a surrounding "sea" of false vacuum $\varphi = 0$.

To conclude, let us stress the most important differences between the two types of phase transitions discussed above. In a second order transition, the dynamics is determined mainly by classical physics. The transition occurs homogeneously in space (apart from the phase boundaries which — as discussed below — become topological defects), and $\langle\varphi(x)\rangle$ evolves continuously in time. In first order transitions, quantum mechanics is essential. The process is extremely inhomogeneous, and $\langle\varphi(x)\rangle$ is discontinuous as a function of time.

3.4. *Particle physics connection*

Up to this point, we have discussed only scalar field toy models. The actual case of interest is a unified gauge theory of strong, weak and electromagnetic interactions. Such a theory (see, for example, Ref. 63) contains massless fermion fields ψ, gauge fields A_μ and scalar fields φ. The Lagrangian $L(\psi, A_\mu, \varphi)$ is invariant under the action of some internal symmetry group G. This grand unified symmetry is broken spontaneously (see, for example, Ref. 64) in one or several stages for the "standard model" symmetry group $\mathrm{SU(3)} \times \mathrm{SU(2)} \times \mathrm{U(1)}$:

$$G \to H \to \cdots \to \mathrm{SU(3)} \times \mathrm{SU(2)} \times \mathrm{U(1)}\,. \tag{3.20}$$

Spontaneous symmetry breaking is achieved by postulating that a scalar (Higgs) field takes on an expectation value which is not invariant under the full group G,

but only under the subgroup H. In order for this to be possible, the potential $V(\varphi)$ for this Higgs field φ must have a global minimum at $\varphi \neq 0$ but a high temperature minimum at $\varphi = 0$. Therefore, such a model will lead to a symmetry-breaking phase transition in the early Universe. Hence, the possibility of topological defect formation arises.

As a toy model, consider a theory of a two-component real scalar field

$$\varphi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}, \tag{3.21}$$

with Lagrangian

$$\mathcal{L}(\varphi) = \frac{1}{2}\partial_\mu \varphi \partial^\mu \varphi - V(\varphi), \tag{3.22}$$

and symmetry-breaking potential

$$V(\varphi) = \frac{1}{4}\lambda(\varphi^2 - \eta^2)^2. \tag{3.23}$$

Summation over the internal index $i = 1,2$ is implicit in (3.22) and (3.23). Obviously, the Lagrangian is invariant under a U(1) rotation U(α) in an internal field space $\Re^2$. However, a choice of vacuum state

$$\varphi_{\text{vac}} = \eta \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{3.24}$$

breaks this symmetry completely.

The concept of vacuum manifold $\mathcal{M}$ will be crucial in the following sections. In general, $\mathcal{M}$ is defined as the set of field configurations which minimize the free energy modulo gauge transformations. For a gauge theory with a symmetry-breaking scheme as in the first stage of (3.20),

$$\mathcal{M} \simeq \frac{G}{H}. \tag{3.25}$$

In the above toy model,

$$\mathcal{M} = \{\varphi : V(\varphi) = V_{\text{min}} = 0\}, \tag{3.26}$$

I will now argue that in "realistic" particle physics models, a symmetry-breaking phase transition proceeds too fast to allow for inflation. Hence, the topological defects which form in the phase transition are not inflated away [i.e. moved beyond the present Hubble radius $H^{-1}(t)$ by the exponential expansion of the forward light cone], but remain inside the Hubble radius and are therefore important for cosmology.

The argument is based on the equation of motion for the Higgs field. If one neglects inhomogeneities in φ, interactions with other fields, and the expansion of the Universe, this equation becomes

$$\ddot{\varphi} \simeq -\frac{\partial V}{\partial \varphi} \simeq \lambda \eta^2 \varphi. \tag{3.27}$$

The typical time scale τ for the solution is

$$\tau = \lambda^{-1/2}\eta^{-1}. \tag{3.28}$$

Sufficient inflation requires that

$$\tau \gg H^{-1}. \tag{3.29}$$

However, from the FRW equation (3.1) it follows that

$$H^2 \simeq \frac{8\pi G}{3}V(0) = \frac{2\pi}{3}\lambda G\eta^4, \tag{3.30}$$

and hence

$$\tau \ll H^{-1} \tag{3.31}$$

unless $\eta > m_{\mathrm{Pl}}$.

In order to obtain inflation, it is in general [54,65] necessary to introduce a scalar field φ which is *not* the Higgs field of a unified field theory. We shall not further consider such models and shall restrict our attention to matter theories motivated by particle physics.

3.5. *Classification of Topological Defects*

Different particle physics models admit different types of topological defects. The topology of the vacuum manifold $\mathcal{M}$ determines the type of defect.

The classification of defects is based on homotopy classes of $\mathcal{M}$. To introduce the ideas (see, for example, Ref. 66), consider maps ψ from S^n to $\mathcal{M}$, n being an integer. Two maps ψ_1 and ψ_2 are called homotopically equivalent ($\psi_1 \sim \psi_2$) if there exists a continuous one-parameter family of maps

$$\psi(t): S^n \to \mathcal{M} \quad t \in [0,1] \tag{3.32}$$

with

$$\psi(0) = \psi_1, \quad \psi(1) = \psi_2. \tag{3.33}$$

The nth homotopy "group" of $\mathcal{M}$, $\Pi_n(\mathcal{M})$, is the set of all homotopy classes of maps $S^n \to \mathcal{M}$. Except for $n = 0, \Pi_n(\mathcal{M})$ is a group.

Let us consider a symmetry-breaking pattern

$$G \to H \tag{3.34}$$

of a group G with

$$\Pi_0(G) = \Pi_1(G) = 1 \tag{3.35}$$

to a subgroup H. As stated in (3.25), the vacuum manifold $\mathcal{M}$ is isomorphic to G/H. By the use of exact sequences[66] it can be shown[1,4] that

$$\Pi_1(\mathcal{M}) = \Pi_0(H), \tag{3.36}$$

$$\Pi_2(\mathcal{M}) = \Pi_1(H), \tag{3.37}$$

assuming $\Pi_2(G) = 1$ for the second relation to hold.

Let us consider two relevant examples. The smallest grand unified group is SU(5). If we consider the symmetry-breaking pattern

$$\mathrm{SU}(5) \to \mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)\,, \tag{3.38}$$

we can use (3.36) and (3.37) to conclude that

$$\Pi_1(\mathcal{M}) = 1\,, \tag{3.39}$$

$$\Pi_2(\mathcal{M}) = \Pi_1(\mathrm{U}(1)) = \mathcal{Z}\,. \tag{3.40}$$

As will be shown in the following section, (3.39) and (3.40) imply that no cosmic strings but monopoles form during the phase transition associated with (3.38).

As a second example, consider $G = \mathrm{SO}(10)$ and the symmetry-breaking pattern

$$\mathrm{SO}(10) \to \mathrm{SU}(5) \times \mathcal{Z}_2\,. \tag{3.41}$$

In this example

$$\Pi_1(\mathcal{M}) = \mathcal{Z}_2\,, \tag{3.42}$$

and hence cosmic strings form during the symmetry-breaking phase transition.

Let us give a brief overview of the classification of topological defects. To be specific, consider a theory with an n-component real scalar field φ with symmetry-breaking potential (3.23).

There are various types of local and global topological defects[3] (regions of trapped energy density), depending on the number of components of φ. The words "local" and "global" refer to whether the symmetry which is broken is a gauge or global symmetry. In the case of local symmetries, the topological defects have a well-defined core outside of which φ contains no energy density in spite of non-vanishing gradients $\nabla\varphi$: the gauge fields A_μ can absorb the gradient, i.e. $D_\mu\varphi = 0$ when $\partial_\mu\varphi \neq 0$, where the covariant derivative D_μ is defined as

$$D_\mu = \partial_\mu - igA_\mu\,, \tag{3.43}$$

g being the gauge coupling constant. Global topological defects, however, have long range density fields and forces.

Table 1. Contains a list of topological defects with their topological characteristics. A "v" marks acceptable theories, an "x" theories which are in conflict with observations (for $\eta \sim 10^{16}$ GeV).

	n	Topology	Local defect	Global defect
Domain wall	1	$\Pi_0(\mathcal{M}) \neq 1$	x	x
Cosmic string	2	$\Pi_1(\mathcal{M}) \neq 1$	v	v
Monopole	3	$\Pi_2(\mathcal{M}) \neq 1$	x	v
Texture	4	$\Pi_3(\mathcal{M}) \neq 1$	—	v

In the following subsections we will construct the various types of defects. The cosmological implications will be discussed in later sections.

The easiest defect to construct is the domain wall. Consider $n = 1$ [or, more generally, a theory with $\Pi_0(\mathcal{M}) \neq 1$]. In this case, the vacuum manifold consists of two points:

$$\varphi_{\text{vac}} = \pm\eta . \tag{3.44}$$

During the symmetry-breaking phase transition, regions in physical space $\mathcal{R}^3$ with $\varphi = \pm\eta$ will form. These regions are separated by two-dimensional surfaces (walls) with $\varphi \notin \mathcal{M}$ (see Fig. 11). These are the domain walls. Since $\varphi \notin \mathcal{M}$ in the walls, $V(\varphi) > 0$ and hence the walls carry energy per unit area. Via the usual gravitational force, this energy can act as a seed for structures in the Universe.

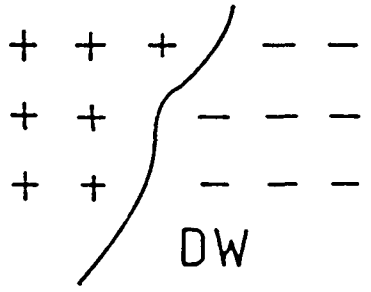

Fig. 11. A two-dimensional cross section through space showing a domain wall (DW) separating a region with $\varphi = \eta$ (+) from a neighboring region with $\varphi = -\eta$ (−).

3.6. *Cosmic strings*

Consider a theory in which matter consists of a gauge field A_μ and a complex scalar field ϕ whose dynamics is given by the Lagrangian

$$\mathcal{L} = \frac{1}{2} D_\mu \phi D^\mu \phi - V(\phi) + \frac{1}{4} F_{\mu\nu} F^{\mu\nu} , \tag{3.45}$$

where $F_{\mu\nu}$ is the field strength tensor. The potential $V(\phi)$ has the symmetry-breaking "Mexican hat" shape (see Fig. 12):

$$V(\phi) = \frac{1}{4}\lambda(|\phi|^2 - \eta^2)^2 . \tag{3.46}$$

Hence, the vacuum manifold $\mathcal{M}$, the space of minimum energy density configurations, is a circle S^1.

The theory described by (3.45) and (3.46) admits one-dimensional topological defects, cosmic strings. In the Abelian Higgs model of this example the string solutions were first found by Nielsen and Olesen.[67] It is possible to construct

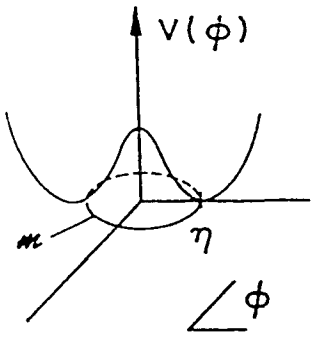

Fig. 12. The zero temperature potential energy of the complex scalar field used in the cosmic string model.

string configurations which are translationally invariant along the z axis. On a circle C in the $x-y$ plane with radius r (see Fig. 13), the boundary conditions for ϕ are

$$\phi(r,\theta) = \eta e^{i\theta}\,, \tag{3.47}$$

where θ is the polar angle along C.

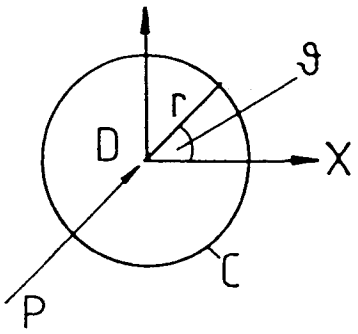

Fig. 13. Sketch of the cosmic string construction of Subsec. 3.6. (See text for notation.)

The configuration (3.47) has winding number 1: at all points of the circle, ϕ takes on values in $\mathcal{M}$, and as φ varies from 0 to 2π, ϕ winds once round $\mathcal{M}$. By continuity it follows that there must be a point p on the disk D bounded by C where $\phi = 0$. By translational symmetry there is a line of points with $\phi = 0$. This line is the center of the cosmic string. The cosmic string is a line of trapped potential energy. In order to minimize the total energy given the prescribed topology (i.e. winding number), the thickness of the string [i.e. radius over which $V(\phi)$ deviates significantly from 0] must be finite. As first shown in Ref. 67, the width w of a string is

$$w \simeq \lambda^{-1/2}\eta^{-1}, \tag{3.48}$$

from which it follows that the mass per unit length μ is

$$\mu \simeq \eta^2, \tag{3.49}$$

i.e. independent of the self-coupling constant λ.

Cosmic strings arise in any model in which the vacuum manifold satisfies the topological criterion

$$\Pi_1(\mathcal{M}) \neq 1. \tag{3.50}$$

Any field configuration $\phi(\underline{x})$ is characterized by an integer n, the element of $\Pi_1(\mathcal{M})$ corresponding to $\phi(\underline{x})$. (Roughly speaking, n can be viewed as the number of times the map φ from $\mathcal{C}$ to $\mathcal{M}$ covers $\mathcal{M}$.)

A cosmic string is an example of a *topological defect* has a well-defined core, a region in space where $\phi \notin \mathcal{M}$ and hence $V(\phi) > 0$. There is an associated winding number, and it is quantized. Hence, a topological defect is stable. Furthermore, topological defects exist for theories with global and local symmetry groups.

3.7. *Monopoles*

If the theory contains three real scalar fields ϕ_i with potential (3.46) (if $|\phi|^2 = \sum_{i=1}^{3} \phi_i^2$), then $\Pi_2(\mathcal{M}) \neq 1$ and monopoles result. The construction of a monopole configuration is illustrated in Fig. 14. As the origin in physical space we select a point which is to become the center of the monopole. Consider a sphere S_r of radius r surrounding this point. A spherically symmetric monopole configuration is obtained by the identity map

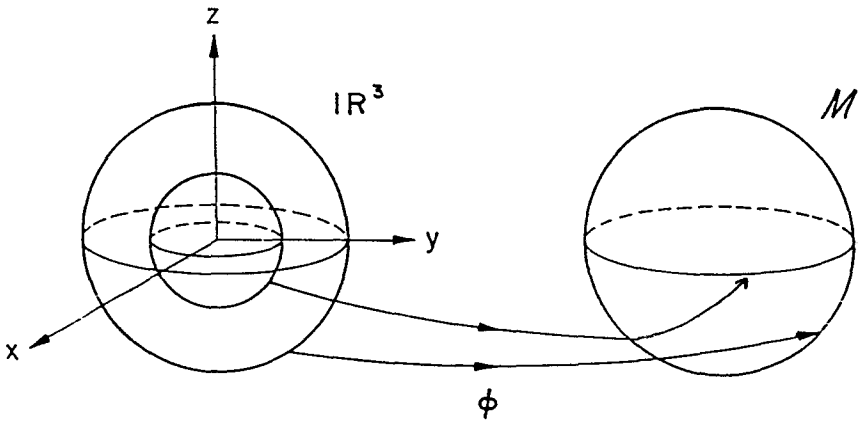

Fig. 14. Construction of a monopole: on the left is physical space, on the right the vacuum manifold. The field configuration ϕ maps spheres in space onto $\mathcal{M}$. However, a core region of space near the origin is mapped onto field values not in $\mathcal{M}$.

$$S_r \to \mathcal{M} = S^2\,, \quad (r,\theta,\varphi) \underset{\varphi}{\to} (\theta,\varphi)\,. \tag{3.51}$$

This configuration has winding number 1. Since the winding number of maps $S^2 \to S^2$ is quantized, it cannot change as r varies. Thus, the only way to obtain a single-valued field configuration at $r = 0$ is for $\varphi(r,\theta,\varphi)$ to leave $\mathcal{M}$ as $r \to 0$. In particular, there is a point (e.g. $r = 0$) for which $\varphi = 0$. This is the center of the monopole. We see that monopoles are topological defects: they contain a core, have quantized winding number and are stable.

3.8. *Global textures*

Next, consider a theory of four real scalar fields given by the Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - V(\phi) \tag{3.52}$$

with

$$V(\phi) = \frac{1}{4}\lambda\left(\sum_{i=1}^{4}\phi_i^2 - \eta^2\right)^2\,. \tag{3.53}$$

In this case, the vacuum manifold is $\mathcal{M} = S^3$ with topology

$$\Pi_3(\mathcal{M}) \neq 1\,, \tag{3.54}$$

and the corresponding defects are the global textures.[1,68,69]

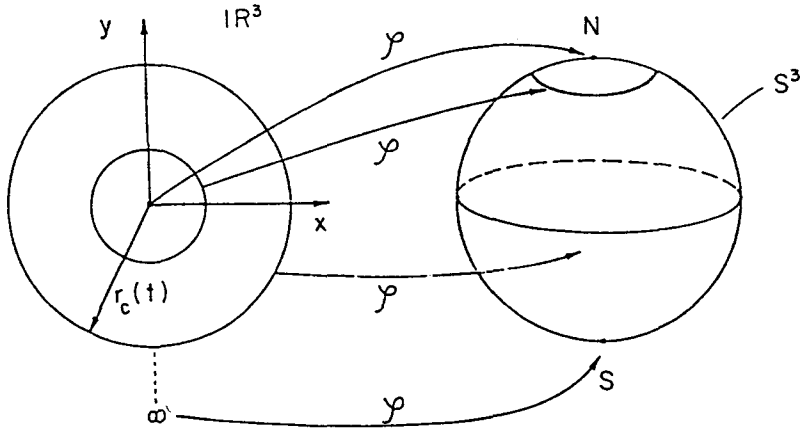

Fig. 15. Construction of a global texture: on the left is physical space, on the right the vacuum manifold. The field configuration ϕ is a map from space to the vacuum manifold (see text).

Textures, however, are quite different than the previous topological defects. The texture construction will render this manifest (see Fig. 15). To construct a radially symmetric texture, we give a field configuration $\phi(x)$ which maps physical space

onto $\mathcal{M}$. The origin 0 in space (an arbitrary point which will be the center of the texture) is mapped onto the north pole N of $\mathcal{M}$. Spheres surrounding 0 are mapped onto spheres surrounding N. In particular, some sphere with radius $r_c(t)$ is mapped onto the equator sphere of $\mathcal{M}$. The distance $r_c(t)$ can be defined as the radius of the texture. Inside this sphere, $\phi(x)$ covers half the vacuum manifold. Finally, the sphere at infinity is mapped onto the south pole of $\mathcal{M}$. The configuration $\phi(\underline{x})$ can be parametrized by[69]

$$\phi(x,y,z) = \left(\cos\chi(r), \sin\chi(r)\frac{x}{r}, \sin\chi(r)\frac{y}{r}, \sin\chi(r)\frac{z}{r}\right) \tag{3.55}$$

in terms of a function $\chi(r)$ with $\chi(0) = 0$ and $\chi(\infty) = \pi$. Note that at all points in space, $\phi(\underline{x})$ lies in $\mathcal{M}$. There is no defect core. All the energy is spatial gradient (and possibly kinetic) energy.

In a cosmological context, there is infinite energy available in an infinite space. Hence, it is not necessary that $\chi(r) \to \pi$ as $r \to \infty$. We can have

$$\chi(r) \to \chi_{\max} < \pi \quad \text{as} \quad r \to \infty . \tag{3.56}$$

In this case, only a fraction,

$$n = \frac{\chi_{\max}}{\pi} - \frac{\sin 2\chi_{\max}}{2\pi} , \tag{3.57}$$

of the vacuum manifold is covered: the winding number n is not quantized. This is a reflection of the fact that whereas topologically nontrivial maps from S^3 to S^3 exist, all maps from R^3 to S^3 can be deformed to the trivial map.

Textures in R^3 are unstable. For the configuration described above, the instability means that $r_c(t) \to 0$ as t increases: the texture collapses. When $r_c(t)$ is microscopical, there will be sufficient energy inside the core to cause $\phi(0)$ to leave $\mathcal{M}$, pass through 0 and equilibrate at $\chi(0) = \pi$: the texture unwinds.

A further difference compared to topological defects: textures are relevant only for theories with global symmetry. Since all the energy is in spatial gradients, for a local theory the gauge fields can reorient themselves such as to cancel the energy:

$$D_\mu \phi = 0 . \tag{3.58}$$

Therefore, it is reasonable to regard textures as an example of a new class of defects, *semitopological defects*. In contrast to topological defects, there is no core, and $\phi(\underline{x}) \epsilon \mathcal{M}$ for all $\underline{x}$. In particular, there is no potential energy. Second, the winding number is not quantized, and hence the defects are unstable. Finally, they exist only in theories with a global internal symmetry.

4. Formation and Evolution of Topological Defects

4.1. *Kibble mechanism*

The Kibble mechanism[1] ensures that in theories which admit topological or semitopological defects, such defects will be produced during a phase transition in the very early Universe.

Consider a mechanical toy model, first introduced by Mazenko, Unruh and Wald[55] in the context of inflationary Universe models, which is useful in understanding the scalar field evolution. Consider (see Fig. 16) a lattice of points on a flat table. At each point, a pencil is pivoted. It is free to rotate and oscillate. The tips of nearest neighbor pencils are connected with springs (to mimic the spatial gradient terms in the scalar field Lagrangian). Newtonian gravity creates a potential energy $V(\varphi)$ for each pencil (φ is the angle relative to the vertical direction). $V(\varphi)$ is minimized for $|\varphi| = \eta$ (in our toy model $\eta = \pi/2$). Hence, the Lagrangian of this pencil model is analogous to that of a scalar field with symmetry-breaking potential (3.46).

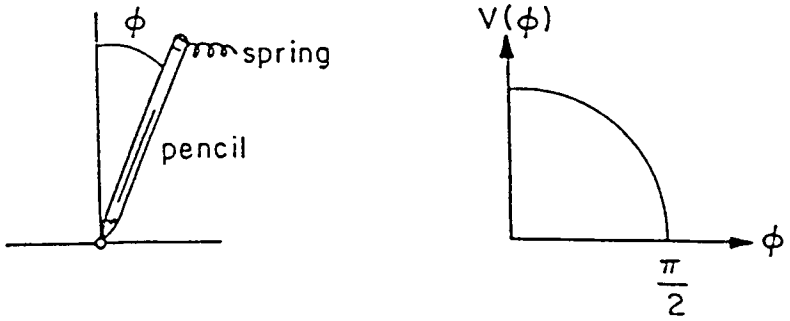

Fig. 16. The pencil model: the potential energy of a simple pencil has the same form as that of scalar fields used for spontaneous symmetry breaking. The springs connecting nearest neighbor pencils give rise to contributions to the energy which mimic spatial gradient terms in field theory.

At high temperatures $T \gg T_c$, all pencils undergo large amplitude high frequency oscillations. However, by causality, the phases of oscillation of pencils with large separation s are uncorrelated. For a system in thermal equilibrium, the length s beyond which phases are random is the correlation length $\xi(t)$. By causality there is an *a priori* causality bound on ξ:

$$\xi(t_c) < t_c\,, \tag{4.1}$$

where t_c is the causal horizon, at temperature T_c.

The critical temperature T_c is the temperature at which the thermal energy is equal to the energy a pencil needs to jump from a horizontal to a vertical position. For $T < T_c$, all pencils want to lie flat on the table. However, their orientations are random beyond a distance of $\xi(t_c)$.

The boundaries between the domains of correlated orientation become topological defects. Hence, it follows from the above causality argument that during the phase transition a network of defects with mean separation $\xi(t) \leq t$ will form.

For models of structure formation and for defects formed in grand unified phase transitions, we are interested in models with a scale of symmetry breaking $\eta \sim 10^{16}$ GeV corresponding to a time of formation $t_c \sim 10^{-35}$ sec.

The evolution of the field configuration (and thus of the network of defects) may be very complicated. However, the same causality argument as used above tells us that no correlations on scales $> t$ can be established (see, however, the caveats of Refs. 70 and 71). Hence, even at times $t \gg t_c$, a network of defects will persist with

$$\xi(t) \leq t \,. \tag{4.2}$$

For some applications, it is important to have a realistic estimate for the initial separation of defects rather than the general causality bound (4.1). In order to improve on (4.1) we must assume that φ is in thermal equilibrium. Our goal is to estimate $\xi(t)$ once the defects have stabilized after the phase transition.

We follow the methods of Refs. 1, 3 and 72. The first step of the analysis is to calculate the location $\varphi_m(T)$ of the minimum of the finite temperature effective potential $V_T(\varphi)$. From (3.16) it follows that for $T \leq T_c$

$$\varphi_m(T) = \eta \left[1 - \left(\frac{T}{T_c}\right)^2\right]^{1/2} . \tag{4.3}$$

The difference in free energy between $\varphi = \varphi_m(T)$ and $\varphi = 0$ is

$$\Delta V(T) = V(0) - V_T(\varphi_m(T)) = \frac{1}{4}\lambda\eta^4 \left[1 - \left(\frac{T}{T_c}\right)^2\right]^2 . \tag{4.4}$$

The protodomain size is given by equating spatial gradient and potential energy density,

$$\left[\frac{\varphi_m(T)}{\xi}\right]^2 \simeq \lambda\eta^4 \left[1 - \left(\frac{T}{T_c}\right)^2\right]^2 , \tag{4.5}$$

giving

$$\xi(T) \simeq \lambda^{-1/2}\eta^{-1} \left[1 - \left(\frac{T}{T_c}\right)^2\right]^{-1/2} . \tag{4.6}$$

The Ginsburg temperature T_G is defined as the temperature below which a thermal fluctuation has insufficient energy to take a correlation volume of "true vacuum" $\varphi = \varphi_m(T)$ into the symmetric configuration $\varphi = 0$. Thus, the criterion for T_G is

$$\xi^3(T_G)\Delta V(T_G) = T_G \,. \tag{4.7}$$

Inserting (4.4) and (4.6) we obtain

$$T_G = \lambda^{-1/2}\eta \left[1 - \left(\frac{T_G}{T_c}\right)^2\right]^{1/2} \tag{4.8}$$

and therefore [from (4.6)]

$$\xi(T_G) = \lambda^{-1} T_G^{-1} \sim \lambda^{-1} \eta^{-1} . \tag{4.9}$$

As a result of the above analysis, it follows that the network of topological defects "freezes out" (i.e. is not continuously destroyed and recreated by thermal fluctuations) at time t_G corresponding to the temperature T_G. The correlation length at the Ginsburg temperature is microscopic, i.e. proportional to η^{-1}, and thus much smaller than the naive causality argument (4.1) would indicate. This fact will be important for cosmic-string-driven baryogenesis.

Note that the Kibble mechanism was discussed above in the context of a global symmetry-breaking scenario. As pointed out in Ref. 73, the situation is more complicated in local theories in which the gauge field can cancel spatial gradients in φ in the energy functional, and in which spatial gradients in φ can be gauged away. Nevertheless, as demonstrated numerically (in $2+1$ dimensions) in Ref. 74 and shown analytically in Ref. 75, the Kibble mechanism also applies to local symmetries.

4.2. *Domain wall and monopole problems*

As stated in Table 1, models with domain walls and local monopoles are ruled out on cosmological grounds. In both cases, the problem is that the energy density in the defects dominates over the energy density in radiation already at the time of nucleosynthesis; in other words, the defects would overclose the Universe. The reasons why this problem occurs, however, are different for domain walls and local monopoles.

Let us demonstrate explicitly why stable domain walls are a cosmological disaster.[76] If domain walls form during a phase transition in the early Universe, it follows by causality (see, however, the caveats of Refs. 70 and 71) that even today there will be at least one wall per Hubble volume. Assuming one wall per Hubble volume, the energy density $\rho_{\rm DW}$ of matter in domain walls is

$$\rho_{\rm DW}(t) \sim \eta^3 t^{-1} , \tag{4.10}$$

whereas the critical density ρ_c is

$$\rho_c = H^2 \frac{3}{8\pi G} \sim m_{\rm Pl}^2 t^{-2} . \tag{4.11}$$

Hence, for $\eta \sim 10^{16}$ GeV the ratio of (4.10) and (4.11) is

$$\frac{\rho_{\rm DW}}{\rho_c}(t) \sim \left(\frac{\eta}{m_{\rm Pl}}\right)^2 (\eta t) \sim 10^{52} . \tag{4.12}$$

The above argument depends in an essential way on the dimension of the defect. One cosmic string per Hubble volume leads to an energy density $\rho_{\rm cs}$ in string,

$$\rho_{\rm cs} \sim \eta^2 t^{-2} , \tag{4.13}$$

which scales like the background radiation density. Later in this section we shall see that the scaling (4.13) does indeed hold in the cosmic string model. Hence, cosmic strings do not lead to cosmological problems. On the contrary, since for GUT models with $\eta \sim 10^{16}$ GeV

$$\frac{\rho_{\rm cs}}{\rho_{\rm c}} \sim \left(\frac{\eta}{m_{\rm Pl}}\right)^2 \sim 10^{-6}, \tag{4.14}$$

cosmic strings in these models could provide the seed perturbations responsible for structure formation.

Theories with local monopoles are ruled out on cosmological grounds[77] (see again the caveats of Refs. 70 and 71) for rather different reasons. Since there are no long range forces between local monopoles, their number density in comoving coordinates does not decrease. Since their contribution to the energy density scales as $a^{-3}(t)$, they will come to dominate the mass of the Universe, provided η is sufficiently large.

Theories with global monopoles[78] are not ruled out, since there are long range forces between monopoles which lead to a "scaling solution" with a fixed number of monopoles per Hubble volume.

4.3. *Cosmic string evolution*

Applied to cosmic strings, the Kibble mechanism implies that at the time of the phase transition, a network of cosmic strings with typical step length $\xi(t_G)$ will form. According to numerical simulations,[79] about 80% of the initial energy is in infinite strings and 20% in closed loops.

The evolution of the cosmic string network for $t > t_G$ is complicated. The key processes are loop production by intersections of infinite strings (see Fig. 17) and loop shrinking by gravitational radiation. These two processes combine to create a mechanism by which the infinite string network loses energy (and length as measured in comoving coordinates). It will be shown that, as a consequence, the correlation length of the string network is always proportional to its causality limit:

$$\xi(t) \sim t. \tag{4.15}$$

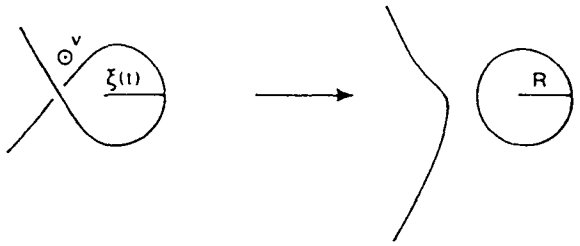

Fig. 17. Formation of loops by self-intersection of infinite strings. According to the original cosmic string scenario, loops form with radius R determined by the instantaneous correlation length of the infinite string network.

Hence, the energy density $\rho_\infty(t)$ in long strings is a fixed fraction of the background energy density $\rho_c(t)$:

$$\rho_\infty(t) \sim \mu\xi(t)^{-2} \sim \mu t^{-2} \tag{4.16}$$

or

$$\frac{\rho_\infty(t)}{\rho_c(t)} \sim G\mu\,. \tag{4.17}$$

We conclude that the cosmic string network approaches a "scaling solution"[80] in which the statistical properties of the network are time-independent if all distances are scaled to the horizon distance.

The origin of the scaling solution for the infinite string network can be understood heuristically, as follows. If the curvature radius $\xi(t)$ of this network is much larger than the Hubble radius t, the network will be frozen in comoving coordinates (since the Hubble damping term dominates in the equations of motion). Hence in the radiation-dominated FRW epoch

$$\xi(t) \sim a(t) \sim t^{1/2} \tag{4.18}$$

and the Hubble radius will catch up to $\xi(t)$. Conversely, if $\xi(t) \ll t$ then the tension term in the equations of motion for the string will dominate, the strings will oscillate relativistically and there will be frequent self-intersections of the strings, leading to rapid loop production and to increasing $\xi(t)/t$. Combining these two arguments, we conclude that there must be a "dynamical fixed point" with $\xi(t) \sim t$.

A first step in a more rigorous analysis of cosmic string evolution is the derivation of the effective equation of motion for the strings. Note that this equation must follow from the field equations since the string is merely a particular topologically stable field configuration.

The equations of motion of a string can be derived from the Nambu action

$$S = -\mu \int d\sigma\, d\tau \left(-\det g^{(2)}_{ab}\right)^{1/2}, \quad a,b = 0,1\,, \tag{4.19}$$

$g^{(2)}_{ab}$ is the world-sheet metric and σ and τ are the world-sheet coordinates. In flat space–time, τ can be taken to be coordinate time, and σ is an affine parameter along the string. In terms of the string coordinates $X^\mu(\sigma,\tau)$ and the metric $g^{(4)}_{\mu\nu}$ of the background space–time,

$$g^{(2)}_{ab} = X^\mu_{,a} X^\nu_{,b} g^{(4)}_{\mu\nu}\,. \tag{4.20}$$

From general symmetry considerations, it is possible to argue that the Nambu action is the correct action. However, I shall follow Foerster[81] and Turok[82] and give a direct heuristic derivation. We start from a general quantum field theory Lagrangian, $\mathcal{L}_{\rm QFT}$. The action is

$$S = \int d^4y\, \mathcal{L}_{\rm QFT}\left(\phi(y)\right)\,. \tag{4.21}$$

We assume the existence of a linear topological defect at $X^\mu(\sigma,\tau)$. The idea is to change variables so that σ and τ are two of the new coordinates, and to expand S to lowest order in w/R, where w is the width of the string and R its curvature radius. As the other new coordinates we take coordinates ρ^2 and ρ^3 in the normal plane to $X^\mu(\sigma,\tau)$. Thus the coordinate transformation takes the coordinates y^μ ($\mu = 0,1,2,3$) to new ones, $\sigma^\nu = (\tau,\sigma,\rho^2,\rho^3)$:

$$y^\mu(\sigma^a) = X^\mu(\sigma,\tau) + \rho^i n_i^\mu(\sigma,\tau)\,, \tag{4.22}$$

where $i = 2,3$ and n_i^μ are the basis vectors in the normal plane to the string world-sheet. The measure transforms as

$$\int d^4y = \int d\sigma\, d\tau\, d\rho^2\, d\rho^3 (\det M_a^\mu) \tag{4.23}$$

with

$$M_a^\mu = \frac{\partial y^\mu}{\partial \sigma^a} = \begin{pmatrix} \partial X^\mu/\partial(\sigma,\tau) \\ n_i^\mu \end{pmatrix} + O(\rho)\,. \tag{4.24}$$

The determinant can easily be evaluated using the following trick:

$$\det M_a^\mu = (-\det \eta_{\mu\nu} M_a^\mu M_b^\nu)^{1/2} \equiv \sqrt{-\det D_{ab}}\,, \tag{4.25}$$

$$\begin{aligned} D &= \begin{pmatrix} \dfrac{\partial x^\mu}{\partial(\sigma,\tau)} \dfrac{\partial X^\nu}{\partial(\sigma,\tau)}\eta_{\mu\nu} & \dfrac{\partial X^\mu}{\partial(\sigma,\tau)} n_b^\nu \eta_{\mu\nu} \\ \dfrac{\partial X^\mu}{\partial(\sigma,\tau)} n_a^\nu \eta_{\mu\nu} & n_a^\mu n_b^\nu \eta_{\mu\nu} \end{pmatrix} \\ &= \begin{pmatrix} X^\mu_{,a} X^\nu_{,b} \eta_{\mu\nu} & 0 \\ 0 & \delta_{ab} \end{pmatrix} + 0\left(\frac{w}{R}\right)\,. \end{aligned} \tag{4.26}$$

Hence

$$\begin{aligned} S &= \int d\sigma\, d\tau \left(-\det g_{ab}^{(2)}\right)^{1/2} \int d\rho^2\, d\rho^3\, \mathcal{L}(y(\sigma,\tau,\rho^2,\rho^3)) + O\left(\frac{w}{R}\right) \\ &= -\mu \int d\sigma\, d\tau \left(-\det g_{ab}^{(2)}\right)^{1/2} + O\left(\frac{w}{R}\right)\,. \end{aligned} \tag{4.27}$$

Here, $-\mu$ is the integral of $\mathcal{L}$ in the normal plane of X. To first order in w/R, it equals the integral of $-\mathcal{H}$; hence it is the mass per unit length.

This derivation of the Nambu action is instructive as it indicates a method for calculating corrections to the equations of motion of the string when extra fields are present, such as for superconducting cosmic strings. It also gives a way of calculating the finite thickness corrections to the equations of motion which will be important at cusps (see below).

In flat space–time we can consistently choose $\tau = t$, $\dot{x} \cdot x' = 0$ and $\dot{x}^2 + x'^2 = 0$. The equations of motion derived from the Nambu action then become

$$\ddot{\underline{x}} - \underline{x}'' = 0\,, \tag{4.28}$$

where $'$ indicates the derivative with respect to σ. As expected, we obtain the relativistic wave equation. The general solution can be decomposed into a left-moving and a right-moving mode:

$$\underline{x}(t,\sigma) = \frac{1}{2}\left[\underline{a}(\sigma - t) + \underline{b}(\sigma + t)\right]. \tag{4.29}$$

The gauge conditions imply that

$$\dot{\underline{a}}^2 = \dot{\underline{b}}^2 = 1\,. \tag{4.30}$$

For a loop, $\underline{x}(\sigma,t)$ is periodic and hence the time average of $\dot{\underline{a}}$ and $\dot{\underline{b}}$ vanishes. Thus, $\dot{\underline{a}}$ and $\dot{\underline{b}}$ are closed curves on the unit sphere with vanishing average. Two such curves generically intersect if they are continuous.[83] An intersection corresponds to a point with $\underline{x}' = 0$ and $|\dot{\underline{x}}| = 1$. Such a point moving at the speed of light is called a cusp. Note that $\dot{\underline{x}}(\sigma,t)$ need not be continuous. Points of discontinuity are called kinks. Both cusps and kinks will be smoothed out by finite thickness effects.[84]

The Nambu action does not describe what happens when two strings hit. This process has been studied numerically for both global[85] and local[86] strings. The authors of those papers set up scalar field configurations corresponding to two strings approaching one another and evolve the complete classical scalar field equations. The result of the analysis is that strings do not cross but exchange ends, provided the relative velocity is smaller than 0.9. Thus, by self-intersecting, an infinite string will split off a loop (Fig. 17). An important open problem is to understand this process analytically. For a special value of the coupling constant Ruback has given a mathematical explanation[87] (see also Shellard and Ruback in Ref. 86).

There are two parts to the nontrivial evolution of the cosmic string network. Firstly, loops are produced by self-intersections of infinite strings. Loops oscillate due to the tension and slowly decay by emitting gravitational radiation.[88] Combining the two steps we have a process by which energy is transferred from the cosmic string network to radiation.

There are analytical indications that a stable "scaling solution" for the cosmic string network exists.[4] According to it, there are always of the order 1 infinite string segments crossing every Hubble volume. The correlation length $\xi(t)$ of an infinite string is of the order t. Hence, at time t loops of radius $R \sim t$ are produced, of the order 1 loop per Hubble volume per expansion time. A heuristic argument for the scaling solution is due to Vilenkin[4]. Take $\tilde{\nu}(t)$ to be the mean number of infinite string segments per Hubble volume. Then the energy density in infinite strings is

$$\rho_\infty(t) = \mu\tilde{\nu}(t)t^{-2}\,. \tag{4.31}$$

The number of loops $n(t)$ produced per unit volume is given by

$$\frac{dn(t)}{dt} = c\tilde{\nu}^2 t^{-4}, \tag{4.32}$$

where c is a constant of order 1. Conservation of energy in strings gives

$$\frac{d\rho_\infty(t)}{dt} + \frac{3}{2t}\rho_\infty(t) = -c'\mu t \frac{dn}{dt} = -cc'\mu\tilde{\nu}^2 t^{-3} \tag{4.33}$$

or, written as an equation for $\tilde{\nu}(t)$,

$$\dot{\tilde{\nu}} - \frac{\tilde{\nu}}{2t} = -cc'\tilde{\nu}^2 t^{-1}. \tag{4.34}$$

Thus if $\tilde{\nu} \gg 1$ then $\dot{\tilde{\nu}} < 0$, while if $\tilde{\nu} \ll 1$ then $\dot{\tilde{\nu}} > 0$. Hence there will be a stable solution with $\tilde{\nu} \sim 1$.

The precise value of $\tilde{\nu}$ must be determined in numerical simulations. These simulations are rather difficult because of the large dynamic range required and due to singularities which arise in the evolution equations near cusps. In the radiation-dominated epoch, $\tilde{\nu}$ is still uncertain by a factor of about 10. The first results were reported in Ref. 89. More recent results are due three groups. Bennett and Bouchet[90] and Allen and Shellard[91] are converging on a value $10 < \tilde{\nu} < 20$, whereas Albrecht and Turok[92] obtain a value which is about 100.

4.4. *Scaling solution for strings*

The scaling solution for the infinite strings implies that the network of strings looks the same at all times when scaled to the Hubble radius. This should also imply that the distribution of cosmic string loops is scale-invariant in the same sense. At present, however, there is no convincing evidence from numerical simulations that this is really the case.

A scaling solution for loops implies that the distribution of $R_i(t)$, the radius of loops at the time of formation, is time-independent after division by t. To simplify the discussion, I shall assume that the distribution in monochromatic, i.e.

$$\frac{R_i(t)}{t} = \alpha. \tag{4.35}$$

From Fig. 17, we expect that $\alpha \sim 1$. The numerical simulations, however, now give $\alpha < 10^{-2}$.[90,91]

From the scaling solution (4.15) for the infinite strings we can derive the scaling solution for loops. We assume that the energy density in long strings — inasmuch as it is not redshifted — must go into loops. β shall be a measure for the mean length ℓ in a loop of "radius" R:

$$\ell = \beta R. \tag{4.36}$$

Since per expansion time and Hubble volume about one loop of radius $R_i(t)$ is produced, we know that the number density in physical coordinates of loops of radius $R_i(t)$ is

$$n(R_i(t),t) = ct^{-4}\,, \tag{4.37}$$

with a constant c which can be calculated from (4.31), (4.35) and (4.36). If we neglect gravitational radiation, this number density simply redshifts:

$$n(R,t) = \left[\frac{z(t)}{z(t_f(R))}\right]^3 n(R,t_f(R))\,, \tag{4.38}$$

where $t_f(R)$ is the time when loops of radius R are formed. Isolating the R dependence, we obtain

$$n(R,t) \sim R^{-4} z(R)^{-3}\,, \tag{4.39}$$

where $z(R)$ is the redshift at time $t = R$. We have the following special cases:

$$\begin{aligned} &n(R,t) \sim R^{-5/2} t^{-3/2}\,, && t < t_{\rm eq}\,, \\ &n(R,t) \sim R^{-5/2} t_{\rm eq}^{1/2} t^{-2} && t > t_{\rm eq}\,, \quad t_f(R) < t_{\rm eq}\,, \\ &n(R,t) \sim R^{-2} t^{-2} && t > t_{\rm eq}\,, \quad t_f(R) > t_{\rm eq}\,. \end{aligned} \tag{4.40}$$

The proportionality constant c is

$$c = \frac{1}{2}\beta^{-1}\alpha^{-2}\tilde{\nu} \tag{4.41}$$

(see, for example, Ref. 93). In deriving (4.41) it is important to note that $n(R_i(t),t)\,dR_i$ is the number density of loops in the radius interval $[R_i, R_i + dR_i]$. Hence, in the radiation-dominated epoch

$$n(R,t) = \nu R^{-5/2} t^{-3/2} \tag{4.42}$$

with

$$\nu = \frac{1}{2}\beta^{-1}\alpha^{1/2}\tilde{\nu}\,. \tag{4.43}$$

From (4.43) we can read off the uncertainties in ν based on the uncertainties in the numerical results. Both $\alpha^{1/2}$ and $\tilde{\nu}$ are determined only up to one order of magnitude. Hence, any quantitative results which depend on the exact value of ν must be treated with a grain of salt.

Gravitational radiation leads to a lower cutoff in $n(R,t)$. Loops with radius smaller than this cutoff were all formed at essentially the same time and hence have the same number density. Thus, $n(R)$ becomes flat. The power in gravitational radiation P_G can be estimated using the quadrupole formula.[94] For a loop of radius R and mass M

$$P_G = \frac{1}{5} G \langle \dddot{Q} \dddot{Q} \rangle\,, \tag{4.44}$$

where Q is the quadrupole moment, $Q \sim MR^2$, and since the frequency of oscillation is $\omega = R^{-1}$

$$P_G \sim G(MR^2)^2\omega^6 \sim (G\mu)\mu\,. \tag{4.45}$$

Even though the quadrupole approximation breaks down since the loops move relativistically, (4.45) gives a good order of magnitude of the power of gravitational radiation. Improved calculations[88] give

$$P_G = \gamma(G\mu)\mu \tag{4.46}$$

with $\gamma \sim 50$. (4.36) and (4.46) imply that

$$\dot{R} = \tilde{\gamma}G\mu \tag{4.47}$$

with $\tilde{\gamma} \equiv \gamma/\beta \sim 5$ (using $\beta \simeq 10^{89}$). Note that the rate of decrease is constant. Hence,

$$R(t) = R_i - (t - t_i)\tilde{\gamma}G\mu \tag{4.48}$$

and the cutoff loop radius is

$$R_c \sim \tilde{\gamma}G\mu t_i\,. \tag{4.49}$$

Let us summarize the scaling solution:

(1) At all times the network of infinite strings looks the same when scaled by the Hubble radius. A small number of infinite string segments cross each Hubble volume and $\rho_\infty(t)$ is given by (4.31).

(2) There is a distribution of loops of all sizes, $0 \leq R < t$. Assuming scaling for loops,

$$n(R,t) = \nu R^{-4}\left[\frac{z(t)}{z(R)}\right]^3\,, \quad R \in [\tilde{\gamma}G\mu t, \alpha t] \tag{4.50}$$

where $\alpha^{-1}R$ is the time of formation of a loop of radius R. Also,

$$n(R,t) = n(\tilde{\gamma}G\mu t, t)\,, \quad R < \tilde{\gamma}G\mu t\,. \tag{4.51}$$

Although the qualitative characteristics of the cosmic string scaling solution are well established, the quantitative details are not. The main reason for this is the fact that the Nambu action breaks down at kinks and cusps. However, kinks and cusps inevitably form and are responsible for the small scale structure on strings. In fact, coarse-graining by integrating out the small scale structure may give an equation of state for strings which deviates from that of a Nambu string.[95] Attempts at understanding the small scale structure on strings are under way.[96]

4.5. *Scaling solution for textures*

Applied to textures,[69] the Kibble mechanism implies that on all scales $r \geq \xi(t_G)$, field configurations with nonvanishing winding number n_w are frozen in after the phase transition at time t_G.

The general causality argument implies that at all times $t > t_G$, on scales greater than t, the field configuration will be random, and hence configurations with winding number $n_w \neq 0$ will persist.

The texture dynamics is determined by the field equation

$$\ddot{\varphi} + 3H\dot{\varphi} - a^{-2}\nabla^2\varphi = -V'(\varphi)\,. \tag{4.52}$$

The second term on the left hand side is proportional to $H^2\varphi$, and the third term scales as $k^2a^{-2}\varphi$, where k^{-1} is the wavelength of the inhomogeneity of the field configuration. Comparing these two terms, we see that for

$$ka^{-1} < H\,, \tag{4.53}$$

i.e. on scale larger than the Hubble radius, a texture configuration will be frozen in, whereas for

$$ka^{-1} > H\,, \tag{4.54}$$

on scales smaller that the Hubble radius, the motion of φ will be relativistic, and the texture configuration will become homogeneous on a time scale of t.

Hence, in the texture scenario, at time $t \gg t_G$, the field configuration $\varphi(\underline{x})$ is predicted to be homogeneous on scales smaller than t, but inhomogeneous on larger scales. There is a finite (and time-independent) probability $p(n_c)$ that per horizon volume there is a field configuration with winding n_w greater than n_c. The probability $p(n_c)$ can be determined using combinatorial arguments.[97]

The evolution of texture configurations has been studied numerically in Refs. 98–100 and analytically in Ref. 101 (see also Ref. 102). The qualitative features of the evolution depend on whether n_w is greater or smaller than the critical winding n_c. If $n_w > n_c$, then the texture will contract, the winding will increase, and eventually the texture will unwind by φ jumping over the potential barrier at the texture center. On the other hand, if $n_w < n_c$, the field configuration will dissipate. It will expand with n_w decreasing to zero continuously.

The physics of texture evolution can be most readily analyzed by studying the spherically symmetric configuration (3.55) with χ a function of r and t. For such a configuration, the action in an expanding FRW background metric is

$$\mathcal{S} = \int d^4x\, r^2 a^3 \left(\dot{\chi}^2 - a^{-2}\chi'^2 - 2a^{-2}r^{-2}\sin^2\chi\right)\,. \tag{4.55}$$

The second term corresponds to the radial gradient energy, the third to angular gradient energy.

Consider a field configuration as sketched in Fig. 18(a) with $\chi_{\max} > n_c > \pi/2$. In this case, the forces F_1 and F_2 act as indicated in the figure. In order to minimize the angular gradient energy, $\chi_{\max}$ wants to increase (see F_1). However, for $r < x$ there is an additional (and in part opposing) force F_2: in order to reduce the angular gradient energy, $r(\chi)$ will tend to decrease, i.e. the field configuration will contract and its total winding will increase.

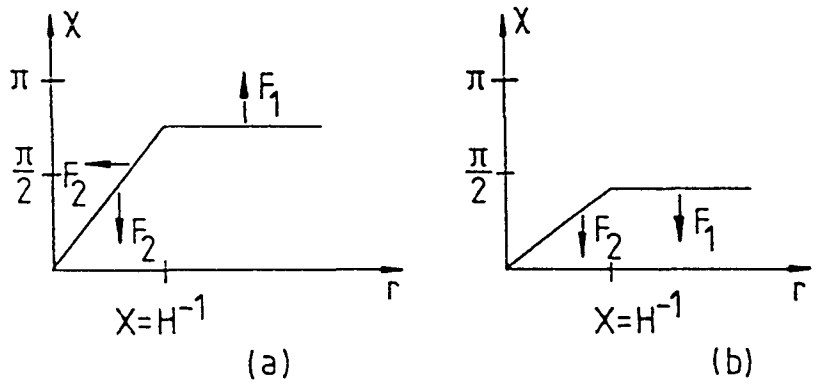

Fig. 18. A sketch of the forces acting on a radially symmetric texture configuration and which cause unwinding in Fig. 18(a) if $n_w > n_c$, and dissipation if $n_w < n_c$ [Fig. 18(b)].

However, for $\chi_{\max} < \pi/2$, the angular gradient force tends to reduce $\chi_{\max}$, and to reduce the radial gradients, $\chi(r)$ will tend to decrease even for $\chi < \chi_{\max}$. Hence, the field configuration will dissipate [see Fig. 18(b)].

If n_w is only slightly larger than $\pi/2$, the force F_1 of Fig. 15(a) is too weak to offset the force F_2 described above. The critical winding n_c is hence larger than 0.5. The precise value has been determined in Refs. 99 and 101. It depends on the length of the "plateau region" of $\chi(r)$ (which is the inter-texture separation) and on the expansion rate of the Universe. For realistic parameters[101]

$$0.65 < n_c < 0.75\,. \tag{4.56}$$

As will be shown in a later section, only textures with $n_w > n_c$ generate localized density perturbations which can act as seeds for cosmic structure formation.

5. Introduction to Structure Formation

5.1. *Power spectrum*

In Sec. 2, the cosmology of a homogeneous and isotropic Universe was reviewed. In order to understand structure formation, it is essential to study the evolution of inhomogeneities at a linearized level. This will be adequate for understanding the early evolution of density perturbations in the Universe.

The starting point of the relativistic theory of cosmological perturbations[103] is the linearized Einstein equations. If we take the general Einstein equations

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}\,, \tag{5.1}$$

where $G_{\mu\nu}(g_{\alpha\beta})$ is the Einstein tensor and $T_{\mu\nu}$ is the energy–momentum tensor of matter, and expand about a cosmological background solution

$$g_{\mu\nu}{}^{(0)} = \mathrm{diag}\left(1, -a^2(t), -a^2(t), -a^2(t)\right) \tag{5.2}$$

and

$$T_\nu^{\mu(0)} = \mathrm{diag}\,(\rho, -p, -p, -p) \tag{5.3}$$

(see Sec. 2), then we obtain the linearized equations

$$\delta G_{\mu\nu}\left(g_{\alpha\beta}{}^{(0)} + h_{\alpha\beta}\right) = 8\pi G \delta T_{\mu\nu}\,. \tag{5.4}$$

Equation (5.4) relates the perturbation $h_{\mu\nu}$ of the metric, i.e.

$$h_{\mu\nu} = g_{\mu\nu} - g_{\mu\nu}{}^{(0)}\,, \tag{5.5}$$

to the matter perturbations.

To gain a heuristic understanding of how the perturbations evolve, recall that gravity is a purely attractive force. Given an initial mass perturbation δm, the force on surrounding particles will be

$$F \sim \delta m\,. \tag{5.6}$$

Since (neglecting for a moment the expansion of the Universe)

$$\delta \ddot{m} \sim F\,, \tag{5.7}$$

we see that in a nonexpanding background the growth of perturbations is exponential. In an expanding background, there will be a damping term depending on H. Hence, perturbations will increase only as a power of time.

The details of the analysis are rather complicated (see Refs. 30 and 104 for recent reviews). The result is that the density contrast $\delta\rho$ grows as follows:

$$\delta\rho(t) \sim \begin{cases} t^{2/3} & t > t_{\mathrm{eq}} \\ t & t < t_{\mathrm{eq}}\,, \quad \lambda > t \\ \mathrm{const} & t < t_{\mathrm{eq}}\,, \quad \lambda < t\,. \end{cases} \tag{5.8}$$

Note that on length scales λ greater than the Hubble radius t, the quantity $\delta\rho$ is not gauge-invariant, i.e. it depends on the slicing of space–time.[104] The quantity which is gauge-invariant is the relativistic potential Φ, which is time-independent if the equation of state of the background cosmology is constant.[104] In a gauge in which $g_{\mu\nu}$ is diagonal and for models of matter in which δT_{ij} is diagonal at linearized level (a condition satisfied by most interesting models of matter), Φ can be identified as follows:

$$g_{\mu\nu} = (1 + 2\Phi)dt^2 - a^2(t)(1 - 2\Phi)d\underline{x}^2\,. \tag{5.9}$$

We will use the results of (5.8) when describing the evolution of the power spectrum.

The main quantity of interest is the r.m.s. mass excess $(\delta M/M)(k,t)$ at time t in a sphere of radius k^{-1}. Given a smooth density distribution

$$\rho(\underline{x}, t) = \rho_0(t) + \delta\rho(\underline{x}, t)\,, \tag{5.10}$$

the r.m.s. mass excess can be related to the Fourier mode $\delta\rho(k)$ in a straightforward manner.[18] The result is

$$\left(\frac{\delta M}{M}\right)^2(k,t) \simeq k^3\left|\frac{\delta\rho}{\rho_0}\right|^2(k,t)\,. \tag{5.11}$$

The adopted convention for Fourier transformation is

$$\delta\rho(\underline{x}) = (2\pi)^{-3/2}V^{-1/2}\int d^3k\, e^{i\underline{k}\,\underline{x}}\delta\rho(\underline{k})\,. \tag{5.12}$$

The result (5.11) holds provided $|\delta\rho(k)|^2$ is proportional to k^n with $n > -3$. An intuitive way to understand the result is as follows: perturbations with wave number larger than k average to zero in a volume k^{-3}, perturbations with wave number smaller then k are phase-space-suppressed such that $(\delta M/M)(k)$ receives its major contribution from Fourier modes of wave number k. Their phase space volume is k^3.

The most commonly used function describing the ensemble of perturbations is the spectrum $P(k)$. By definition, the power spectrum is the square of the modulus of the Fourier space density contrast:

$$P(k) = \left|\frac{\delta\rho}{\rho_0}(k)\right|^2\,. \tag{5.13}$$

Hence [from (5.11)], $P(k)$ is related to the r.m.s. mass fluctuations $(\delta M/M)(k,t)$ on physical length scale

$$\lambda_k = a(t)\frac{2\pi}{k} \tag{5.14}$$

at time t via

$$\left(\frac{\delta M}{M}\right)^2(k,t) \simeq k^3P(k)\,. \tag{5.15}$$

The scaling solution for topological defect models implies that when measured at the time $t_H(k)$ at which the wavelength λ_k equals the Hubble radius, the r.m.s. mass perturbation $(\delta M/M)(k)$ is independent of k (scale invariant), i.e.

$$\frac{\delta M}{M}(k, t_H(k)) = \text{const}\,. \tag{5.16}$$

This is because at any time t, a constant fraction of the mass M inside the Hubble radius is contained in the topological defects. For example, one cosmic string of length t contains mass $\delta M = \mu t$ compared to the total mass $M \sim t^3\rho(t) \sim t$ inside the Hubble radius, thus leaving the ratio $\delta M/M$ time-independent.

Equation (5.16) is the same result as is obtained for inflationary Universe models. Hence, we conclude that all three main models of structure formation — adiabatic random phase perturbations from inflation, cosmic strings, and global textures — to a first approximation produce a scale-invariant spectrum.

To convert (5.16) into an expression for the power spectrum $P(k)$, we use the fact that $\delta M/M$ grows as the scale factor $a(t)$ during the matter-dominated epoch on scales smaller than the Hubble radius [see (5.8)]:

$$\frac{\delta M}{M}(k,t) = \left[\frac{t}{t_H(k)}\right]^{2/3} \frac{\delta M}{M}(k, t_H(k)) . \tag{5.17}$$

On scales larger than the Hubble radius at $t_{\rm eq}$

$$t_H(k) = 2\pi k^{-1} a(t_H(k)) \sim t_H^{2/3}(k) k^{-1} , \tag{5.18}$$

and hence

$$t_H(k) \sim k^{-3} . \tag{5.19}$$

Therefore, combining (5.16), (5.17) and (5.19) we obtain

$$\frac{\delta M}{M}(k,t) \sim k^2 . \tag{5.20}$$

From (5.15) it follows that

$$P(k) \sim k^n \tag{5.21}$$

with $n = 1$.

Recently, there has been some interest in deviations from scale invariance. In models of inflation, a deviation comes about[105] because H decreases slowly during inflation. In topological defect models, numerical[106,107] and semianalytical[108] studies have also shown small deviations from scale invariance. These deviations, however, are all small and quite model-dependent.

5.2. *CMB anisotropies*

Density perturbations give rise to anisotropies in the temperature of the CMB. There are three main contributions (see Fig. 19):

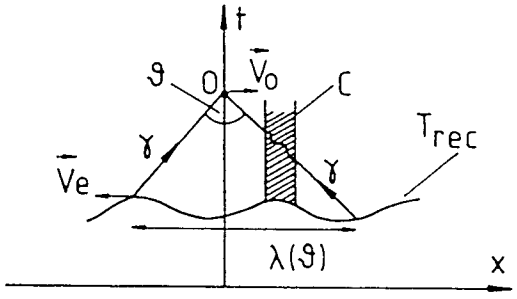

Fig. 19. Space–time plot sketching the origin of CMB temperature anisotropies. The surface labeled $T_{\rm rec}$ is the last scattering surface. O is the observer at the present time measuring photons γ impinging from directions in the sky separated by angle ϑ. The shaded area labeled C stands for a local overdensity, leading to distortions of geodesics. Possible velocities of observer and emitter are indicated as $\vec{v_o}$ and $\vec{v_e}$.

(i) Density fluctuations at the time of last scattering $t_{\rm rec}$ imply that the surface of constant temperature $T_{\rm rec}$ was not flat. Hence, photons arriving at the observer O from different directions (separated by angle θ) have traveled a different amount of time. Since they were emitted with the same frequency, they arrive with differing frequencies, thus leading to temperature anisotropies. This is the Sachs–Wolfe effect.[109]

(ii) A localized clump of energy between the last scattering and the present time will distort the geodesics passing through it, and hence lead to temperature differences (the Rees–Sciama effect[110]).

(iii) If either the last scattering surface or the observer is moving with respect to the frame given by the background cosmology, Doppler anisotropies will result.

The observed[111] dipole anisotropy of the CMB is used to determine the Earth's velocity relative to the rest frame of the CMB. After subtraction of the dipole, the remaining anisotropy on large angular scales is dominated by the Sachs–Wolfe effect (see, for example, Ref. 30). The specific relationship between $\delta T/T$ and density perturbations in the case of adiabatic fluctuations is given by[30,104]

$$\frac{\delta T}{T}(t_0, \vartheta) = \frac{1}{3}\Phi\left(t_{\rm rec}, \lambda(\vartheta)\right) . \tag{5.22}$$

According to relativistic perturbation theory, the relativistic potential Φ of (5.9) is independent of time in the absence of changes in the equation of state of the background.[104] Hence

$$\frac{\delta T}{T}(t_0, \vartheta) = \frac{1}{3}\Phi\left(t_H(\lambda), \lambda(\vartheta)\right) \simeq \frac{1}{3}\frac{\delta\rho}{\rho}\left(t_H(\lambda), \lambda\right) , \tag{5.23}$$

where $\lambda(\vartheta)$ is the comoving length scale corresponding to angular separation θ at $t_{\rm rec}$, and $t_H(\lambda)$ is the time when this scale equals the Hubble radius t. We have used the fact that at Hubble radius crossing Φ and $\delta\rho/\rho$ coincide up to a numerical factor of order 1.[104]

Recently, both the COBE satellite[11] and the MIT balloon[112] experiment have detected CMB anisotropies on angular scales larger than 7°. The quadrupole anisotropy reported is

$$\frac{\delta T}{T}(90°) \simeq 5 \cdot 10^{-6} , \tag{5.24}$$

and the low l harmonies of $\delta T/T$ are consistent with a power law spectrum of density perturbations with

$$n = 1.1 \pm 0.5 . \tag{5.25}$$

It is reassuring that the observed value of n is so close to the value $n = 1$ favored in both inflationary and topological defect models of structure formation. The amplitude of $\delta T/T$ is also of the right order of magnitude to agree with theoretical

predictions based on structure formation arguments. Since by (5.8) density perturbations grow as $a(t)$ and since today $\delta\rho/\rho$ on scales of clusters is about 1, we predict that on these scales

$$\frac{\delta T}{T}(\vartheta) \sim \frac{1}{3} z(t_{\rm eq})^{-1} \sim 3 \cdot 10^{-5} . \tag{5.26}$$

This result agrees with the extrapolation of the COBE results to cluster scales [11] to better than a factor of 2.

The results (5.24) and (5.25) cannot at the present time be used to differentiate between inflationary Universe and topological defect theories, since the models all predict a similar slope of $P(k)$. However, the amplitude of $\delta T/T$ can be used to normalize the power spectrum in any given model. Large scale structure observations provide an independent normalization. These two normalizations must be consistent in order for a theoretical model to work. For cosmic strings, the normalizations of $P(k)$ agree[106,108] well; for an unbiased texture model, the normalization factors differ by about 3.[107]

In the near future, maps of CMB anisotropies will be obtained[113] which are signal-dominated in every pixel (the COBE maps are dominated by noise). At that point, statistical[114,115] measures of CMB maps can be evaluated which can pick out the non-Gaussian signatures of topological defect models. Indeed, nonrandom phases of the Fourier modes of $\delta\rho$ are the key feature of such models (see Fig. 20).

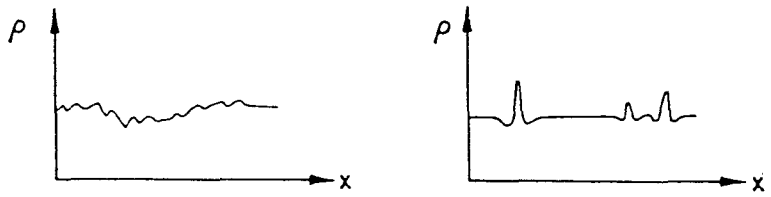

Fig. 20. A comparison of density perturbations along a line in space between (a) random phase models and (b) theories based on nonadiabatic seeds.

Note that CMB anisotropies form a more severe testing ground for cosmological models than large scale structure data obtained from optical telescopes because they are (at least on angular scales larger than the horizon at the last scattering) not contaminated by hydrodynamical and nonlinear effects.

5.3. *Large scale structure data*

Here, I shall mention some of the data obtained from optical telescopes which at the present time seems most relevant to cosmology.

First, however, it is important to note a discrepancy between theory and observations. Observations become increasingly uncertain on increasing length scales, whereas theoretical predictions decrease in accuracy as the length scale

decreases, the reason being that on large scales linear theory is applicable and gravity is the only force which needs to be considered, whereas on smaller scales nonlinear and messy nongravitational effects become important. However, this state of affairs also implies that progress on either the theoretical or the observational front will yield a double benefit. For example, new data on larger scales not only increases the amount of data available, but also leads into a regime with smaller theoretical uncertainties. Since we know that observers will be providing lots of new data in the next few years, we can be sure that cosmology will remain an extremely exciting field.

There is a lot of new data on the large scale structure of the Universe. Redshift surveys have provided three-dimensional maps of the distribution of galaxies. An example is the recent Center for Astrophysics (CFA) survey[9] of slices on the northern celestial sphere which show overdense sheets of galaxies with dimensions $(50 \times 50 \times 5)$ Mpc^3 (with $h = 1$) separated by large voids. There is evidence for superclusters,[116] filaments longer than $50h^{-1}$ Mpc,[117] voids of diameter $60h^{-1}$ Mpc[118] and overdense sheets of galaxies[119] in regions of the sky different from that covered in the CFA survey.

A second window on the large scale structure of the Universe comes from measuring the peculiar velocities of galaxies averaged over large regions[120] from which one can infer information about the magnitude of density perturbations on large scales.[121] Preliminary observations indicate surprisingly large velocities on scales of $60h^{-1}$ Mpc, although consensus on the interpretation of the data is still lacking.[120,122]

On the scale of clusters the data is still rather uncertain. The quantities one would like to focus on are the overall mass scale (the mean mass of an object which satisfies a fixed operational definition of a cluster), the multiplicity function and the two-point (and in the future higher-point) correlation functions.

A cluster of galaxies (Abell[123] cluster) is defined as a region in the sky with more than 50 bright galaxies in a sphere of radius $1.5h^{-1}$ Mpc (to be compared to the mean separation of bright galaxies,[124] which is about $5h^{-1}$ Mpc). The mean separation of clusters is about $50h^{-1}$ Mpc, and their masses are of the order $10^{14} M_\odot$[125] (where $M_\odot$ is the solar mass).

The multiplicity function $n(M)$ of clusters gives the number density of clusters of mass M per unit physical volume. $n(M)dM$ is the number density of objects in the mass interval $[M, M + dM]$.

The two point correlation function $\xi(r)$ of clusters measures the nonrandomness of the distribution of clusters and is defined by

$$\xi(r) = \left\langle \frac{n(r) - n_0}{n_0} \right\rangle , \tag{5.27}$$

where $n(r)$ is the number density at a fixed distance r from a given object and n_0 is the average density, and the angular brackets indicate averaging over the objects.

The initial observational results[10] (Fig. 21) gave

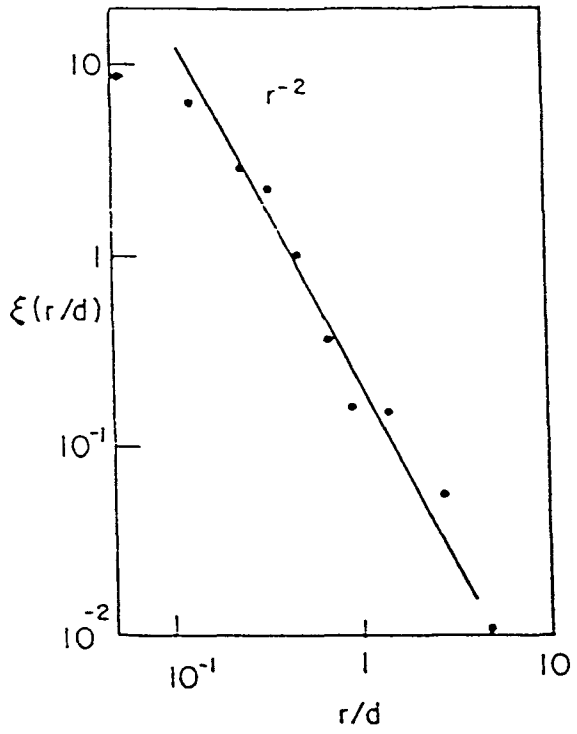

Fig. 21. The observed cluster correlation function plotted against relative separation (from Ref. 10). The solid circles are the data points. Shown is the fit to a r^{-2} slope.

$$\xi(r) \simeq \left(\frac{r}{r_0}\right)^{-1.8} \tag{5.28}$$

with correlation length $r_0 \sim 25h^{-1}$ Mpc. Recently,[126] some criticism of these results has been raised. However, it is unlikely that the entire effect is fictitious.

For galaxies, similar quantities can be measured and compared with observations. Large spiral galaxies have masses $M \sim 10^{11} M_\odot$. Their mean separation is about $5h^{-1}$ Mpc. The mass function[127] (Fig. 22) gives the distribution of masses, and the two-point correlation function describes the nonrandomness of the distribution. The two-point correlation function of galaxies has the same power law form as (5.28), with correlation length $r_0 \sim 8$ Mpc. For galaxies, one can in addition probe the interior structure by measuring the velocity rotation curve $v(r)$, which is related to the density profile (baryonic plus dark matter) by

$$v(r) \sim r^2 \rho(r) . \tag{5.29}$$

Observations[128] indicate that far beyond the disk radius, $v(r) \sim$ const (Fig. 23), which indicates the existence of dark matter with a density profile

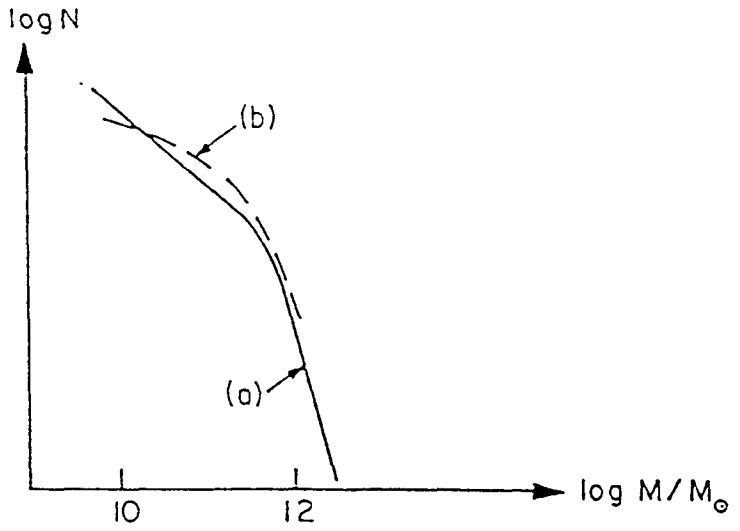

Fig. 22. The mass function of galaxies (determined from the luminosity function assuming constant mass to light ratio) from Ref. 127 [(a) from Bahcall, (b) from Binggeli].

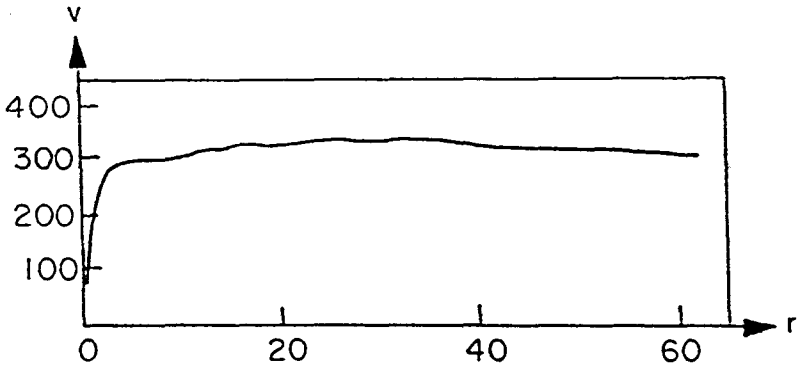

Fig. 23. A typical velocity rotation curve (for NGC488 HI data, taken from Ref. 128). The radius is in kpc, the velocity in $\mathrm{km\,s^{-1}}$.

$$\rho(r) \sim r^{-2}. \tag{5.30}$$

We can also measure the angular momentum of galaxies. Typical numbers for large spirals are in the range[129] 10^{73}–10^{75} $\mathrm{cm^2 g\ s^{-1}}$.

6. Cosmic Strings and Structure Formation

The starting point of the structure formation scenario in the cosmic string theory is the scaling solution for the cosmic string network, according to which at all times t (in particular at t_{eq}, the time when perturbations can start to grow) there will be a

few long strings crossing each Hubble volume, plus a distribution of loops of radius $R \ll t$ (see Fig. 24).

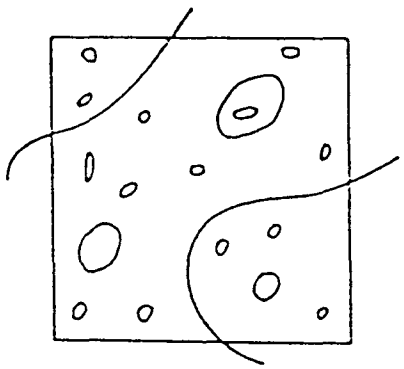

Fig. 24. Sketch of the scaling solution for the cosmic string network. The box corresponds to one Hubble volume at arbitrary time t.

The cosmic string model admits three mechanisms for structure formation: loops, filaments and wakes. Cosmic string loops have the same time-averaged field as a point source with mass[130]

$$M(R) = \beta R \mu , \tag{6.1}$$

R being the loop radius and $\beta \sim 2\pi$. Hence, loops will be seeds for spherical accretion of dust and radiation.[80]

For loops with $R \leq t_{\rm eq}$, growth of perturbations in a model dominated by cold dark matter starts at $t_{\rm eq}$. Hence, the mass at the present time will be

$$M(R, t_0) = z(t_{\rm eq}) \beta R \mu . \tag{6.2}$$

In the original cosmic string model[80,131] it was assumed that loops dominate over wakes. In this case, the theory could be normalized (i.e. μ could be determined) by demanding that loops with the mean separation of clusters $d_{\rm cl}$ [from the discussion in Subsec. 4.4 it follows that the loop radius $R(d_{\rm cl})$ is determined by the mean separation] accrete the correct mass, i.e. that

$$M(R(d_{\rm cl}), t_0) = 10^{14} M_\odot . \tag{6.3}$$

This condition yields[131]

$$\mu \simeq 10^{36} \ {\rm GeV}^2 . \tag{6.4}$$

Thus, if cosmic strings are to be relevant for structure formation, they must arise due to symmetry breaking at energy scale $\eta \simeq 10^{16}$ GeV. This scale happens to be the scale of unification of weak, strong and electromagnetic interactions. It is tantalizing to speculate that cosmology is telling us that there indeed is new physics at the GUT scale.

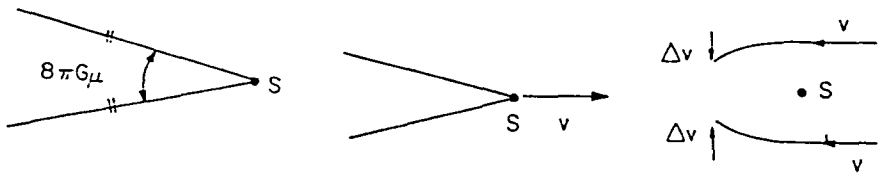

Fig. 25. Sketch of the mechanism by which a long straight cosmic string moving with velocity v in the transverse direction through a plasma induces a velocity perturbations Δv towards the wake. Shown on the left is the deficit angle, in the center is a sketch of the string moving in the plasma, and on the right is a sketch of how the plasma moves in the frame in which the string is at rest.

The second mechanism involves long strings moving with relativistic speed in their normal plane which give rise to velocity perturbations in their wake. The mechanism is illustrated in Fig. 25: space normal to the string is a cone with deficit angle[132]

$$\alpha = 8\pi G\mu . \tag{6.5}$$

If the string is moving with normal velocity v through a bath of dark matter, a velocity perturbation

$$\delta v = 4\pi G\mu v\gamma \tag{6.6}$$

[with $\gamma = (1 - v^2)^{-1/2}$] towards the plane behind the string results.[133] At times after t_{eq}, this induces planar overdensities, the most prominent (i.e. thickest at the present time) and numerous of which were created at t_{eq}, the time of equal matter and radiation.[134,135] The corresponding planar dimensions are (in comoving coordinates)

$$t_{eq}z(t_{eq}) \times t_{eq}z(t_{eq})v \sim (40 \times 40v)\ \mathrm{Mpc}^2 . \tag{6.7}$$

An intuitive understanding of the origin of the above distinguished scale can be obtained as follows. Viewed from a distance, the density perturbation grows as in a linear theory, i.e.

$$\frac{\delta\rho}{\rho}(t) = \left(\frac{t}{t_i}\right)^{2/3} \frac{\delta\rho}{\rho}(t_i) , \tag{6.8}$$

for a perturbation set up at a time $t_i > t_{eq}$. Since the initial $\delta\rho/\rho(t_i)$ is independent of t_i, the largest density contrast comes from the earliest t_i, namely $t_i = t_{eq}$. Fluctuations created at $t_i < t_{eq}$ are erased by the large thermal velocities. Thus, strings at t_{eq} create the most prominent wakes. They are also the must numerous, since the comoving separation of strings decreases as t_i decreases.

A more rigorous way to obtain the above result is to consider the evolution of the velocity perturbations induced by a wake in the Zel'dovich approximation.[30] The height of a dark matter particle above the wake can be written as

$$h(q,t) = a(t)[q - \psi(q,t)] , \tag{6.9}$$

where q is the initial comoving distance, and $\psi(q,t)$ is the comoving displacement caused by the presence of the wake. The thickness of the wake at time t is determined by the value of q for which

$$\dot{h}(q,t) = 0\,. \tag{6.10}$$

Obviously, the value of q for which $\dot{h} = 0$ increases as the time when the perturbation ψ begins to grow is moved back in time. Hence, the earliest wakes will be the thickest.

The details of the calculation depend on whether the dark matter is hot or cold (see Sec. 8). For hot dark matter, the large thermal velocities delay the beginning of the growth of ψ on small scales. A detailed analysis[136,137] shows that for hot dark matter no perturbations become nonlinear unless

$$G\mu > 5 \cdot 10^{-7}\,. \tag{6.11}$$

In this case, the value of the redshift $z(q)$ at which $\dot{h} = 0$ is maximal for the value of q (the thickness) given by

$$q \sim G\mu v\gamma(v) z(t_{\rm eq})^2 t_{\rm eq} \sim 4v \text{ Mpc} \tag{6.12}$$

for wakes created at $t_i = t_{\rm eq}$. Note that the scales of cosmic string wakes [see (6.7) and (6.12)] compare favorably with the measures of the observed sheets of galaxies.[9]

Wakes arise if there is little small scale structure on the string. In this case, the string tension equals the mass density, the string moves at relativistic speeds, and there is no local gravitational attraction towards the string.

In contrast, if there is small scale structure on strings, then[95] the string tension T is smaller than the mass per unit length μ and the metric of a string in the z direction becomes

$$ds^2 = (1 + h_{00})[dt^2 - dz^2 - dr^2 - (1 - 8G\mu)r^2 d\theta^2] \tag{6.13}$$

with

$$h_{00} = 4G(\mu - T)\ln\frac{r}{r_0}\,, \tag{6.14}$$

r_0 being the string width. Since h_{00} does not vanish, there is a gravitational force towards the string which gives rise to cylindrical accretion, thus producing filaments.

As is evident from the last term in the metric (6.13), space perpendicular to the string remains conical, with the deficit angle given by (6.5). However, since the string is no longer relativistic, the transverse velocities v of the string network are expected to be smaller, and hence the induced wakes will be shorter.

Which of the mechanisms — filaments or wakes — dominates is determined by the competition between the velocity induced by h_{00} and the velocity perturbation of the wake. The total velocity is[138,139]

$$u = -\frac{2\pi G(\mu - T)}{v\gamma(v)} - 4\pi G\mu v\gamma(v)\,, \tag{6.15}$$

the first term giving filaments, the second producing wakes. Hence, for small v the former will dominate, for large v the latter.

By the same argument as for wakes, the most numerous and prominent filaments will have the distinguished scale

$$t_{\rm eq} z(t_{\rm eq}) \times d_f \times d_f \,, \tag{6.16}$$

where d_f can be calculated using the Zel'dovich approximation.

7. Textures and Structure Formation

The starting point of the texture scenario of structure formation is the scaling solution for textures: at any time t, there is a fixed probability p that a texture configuration is entering the Hubble radius and starting to collapse.

In the texture model it is the contraction of the field configuration which leads to density perturbations. At the time when the texture enters the horizon, an isocurvature perturbation is established: the energy density in the scalar field is compensated by a deficit in radiation. However, the contraction of the scalar field configuration leads to a clumping of gradient and kinetic energy at the center of the texture[140] (Fig. 26). This, in turn, provides the seed perturbations which cause dark matter and radiation to collapse in a spherical manner.[140–143]

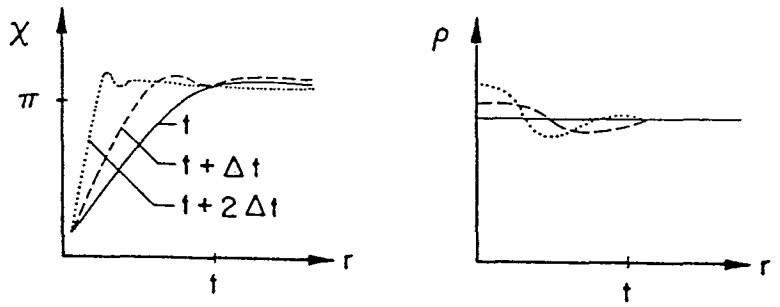

Fig. 26. A sketch of the density perturbation produced by a collapsing texture. The left graph shows the time evolution of the field $\chi(r)$ as a function of radius r and time [see (3.55)]. The contraction of $\chi(r)$ leads to a spatial gradient energy perturbation at the center of the texture, as illustrated on the right. The energy is denoted by ρ. Solid lines denote the initial time, dashed lines are at time $t + \Delta t$, and dotted lines correspond to time $t + 2\Delta t$, where Δt is a fraction of the Hubble expansion time (the typical time scale for the dynamics).

The texture model has a significant advantage over the cosmic string theory: it is much more amenable to analytical and numerical calculations. The reason is simple: whereas the evolution of the network of cosmic stings is very complicated, and the detailed statistics of the scaling solution are not yet known with good accuracy, the evolution of textures is essentially trivial: they collapse in less than a Hubble expansion time and set up well-defined cosmological perturbations whose subsequent evolution can be analyzed without great problems.

In the scaling solution, the comoving number density of textures n obeys.[8,69]

$$\frac{dn}{d\tau} = \frac{c}{\tau^4}\,, \tag{7.1}$$

where $c \simeq 0.04$ (see also Refs. 97 and 100), and τ is conformal time determined by

$$dt = a(t)d\tau\,. \tag{7.2}$$

A collapsing texture with center at $\underline{x} = 0$ induces a velocity perturbation[8,140]

$$\Delta \underline{v} = -F(\underline{x},\tau)\varepsilon\,\hat{\underline{x}}\,, \tag{7.3}$$

where $\hat{\underline{x}}$ is the unit vector in the direction $\underline{x}$,

$$\varepsilon = 8\pi^2 G\eta^2 \tag{7.4}$$

is the measure of the strength of the velocity kick (recall that η is the symmetry-breaking scale, the single free parameter in the texture model) and

$$F(\underline{x},\tau) \simeq \begin{cases} \left(1 - \dfrac{x}{\tau}\right) & x < \tau \\ 0 & x > \tau \end{cases} \tag{7.5}$$

gives the profile of the velocity field. In the above, $\underline{x}$ is a comoving coordinate. In agreement with the causality conditions, F vanishes on scales larger than the horizon ($|\underline{x}| \equiv x > \tau$).

The density fluctuation $\delta(\underline{x},\tau,\tau_i)$ at time τ induced by a texture collapsing at $x = 0$ at time τ_i is determined by the initial velocity perturbation. The result is[8,141]

$$\delta(\underline{x},\tau,\tau_i) = \frac{2F(\underline{x},\tau_i)\varepsilon\tau_*}{x}a(\eta_i)[\delta_2(\eta_i)\delta_1(\eta) - \delta_1(\eta_i)\delta_2(\eta)]\,, \tag{7.6}$$

where $\delta_1(\eta)$ and $\delta_2(\eta)$ are the growing and decaying mode solutions of the linear perturbation equations, and

$$\frac{1}{\tau_*} = \left[\frac{8\pi G\rho(t_{\rm eq})}{3}\right]^{1/2}. \tag{7.7}$$

In the above, the scale factor has been normalized such that

$$a(t_{\rm eq}) = 1\,. \tag{7.8}$$

In this case the growing mode is

$$\delta_1(a(\eta)) = 1 + \frac{3a}{2}\,. \tag{7.9}$$

Based on (7.1) and (7.6), many interesting observables can be computed rather easily. As an example, in Fig. 27 the resulting power spectrum of density perturbations is shown and compared to what is obtained in the "standard CDM model," a theory based on quantum fluctuations from inflation in a Universe dominated by cold dark matter.

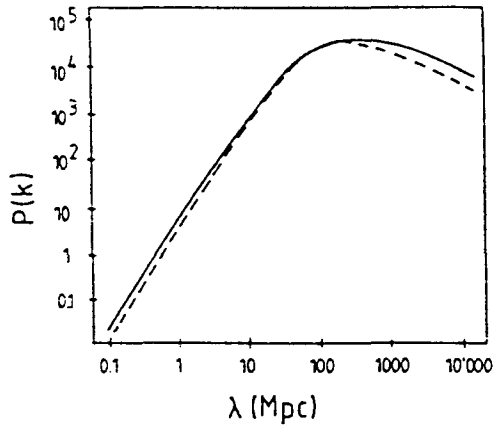

Fig. 27. A comparison of the power spectrum of density perturbations in a global texture model (solid line) and in the standard CDM model (dashed curve). Note the larger power in the texture model on large scales (from Ref. 143). The units of $P(k)$ are arbitrary.

As in the cosmic string model, in the global texture scenario the length scale of the dominant structures is the comoving Hubble radius at $t_{\rm eq}$. Textures generated at $t_{\rm eq}$ are the most numerous, and the perturbations induced by them have the most time to grow.

In both cosmic string and texture models, the fluctuations are non-Gaussian, which means that the Fourier modes of the density perturbation $\delta\rho$ have nonrandom phases. Most inflationary Universe models, in contrast, predict (in linear theory) random phase fluctuations which can be viewed as a superposition of small amplitude plane wave perturbations with uncorrelated phases (for some subtle issues see Refs. 144 and 145).

8. Comparison

8.1. *Role of hot and cold dark matter*

Before discussing some key observations which will allow us to distinguish between the different models, I will consider the role of dark matter. The key issue is free streaming. Recall that cold dark matter consists of particles which have negligible velocity v at $t_{\rm eq}$, the time when subhorizon scale perturbations can start growing:

$$v(t_{\rm eq}) \ll 1 \quad \text{(CDM)}\,. \tag{8.1}$$

For hot dark matter, on the other hand,

$$v(t_{\rm eq}) \sim 1 \quad (\text{HDM}). \tag{8.2}$$

Due to their large thermal velocities, it is not possible to establish HDM perturbations at early times on small scales. Fluctuations are erased by free streaming on all scales smaller than the free streaming length

$$\lambda_J^c(t) = v(t)z(t)t \tag{8.3}$$

(in comoving coordinates). For $t > t_{\rm eq}$, the free streaming length decreases as $t^{-1/3}$. The maximal streaming length is

$$\lambda_J^{\max} = \lambda_J^c(t_{\rm eq}), \tag{8.4}$$

which for $v(t_{\rm eq}) \sim 0.1$ (appropriate for 25 eV neutrinos) exceeds the scale of galaxies.

In inflationary Universe models and in the texture theory, the density perturbations are essentially dark matter fluctuations. The above free streaming analysis then shows that, if the dark matter is hot, no perturbations on the scale of galaxies will survive independently of larger scale structures. Hence, these theories are acceptable only if most of the dark matter is cold.

Cosmic string theories, in contrast, work well — if not even better — with hot dark matter.[146,147,7,136,137] The cosmic string seeds survive free streaming. The growth of perturbations on small scales λ is delayed [it starts once $\lambda = \lambda_J(t)$] but not prevented.

Let us summarize the main characteristics of the cosmic string, global texture, and inflationary Universe theories of structure formation. Inflation predicts random phase perturbations. The density distribution is "Swiss-cheese-like," and the model is consistent with basic observations only for CDM. The global texture and cosmic string models both give nonrandom phase perturbations. The topology is dominated by spherical peaks for textures, whereas it is planar or filamentary for cosmic strings (depending on the small scale structure on the strings). Textures require CDM, whereas cosmic strings work better with HDM.

8.2. *Large scale structure signatures for cosmic strings and textures*

The *genus curve*[148] is a statistical measure for the topology of large scale structure. Given a smooth density field $\rho(\underline{x})$, we pick a density ρ_0 and consider the surface S_{ρ_0} where $\rho(\underline{x}) = \rho_0$ and calculate the genus $\nu(S)$ of this surface:

$$\nu = \#\text{ of holes of } S - \#\text{ of disconnected components of } S + 1. \tag{8.5}$$

The genus curve is the graph of ν as a function of ρ_0.

For perturbations with Poisson statistics, the genus curve can be calculated analytically (Fig. 28). The inflationary CDM model in the linear regime falls in

this category. The genus curve is symmetric about the mean density $\bar{\rho}$. In the texture model, the symmetry about $\bar{\rho}$ is broken and the genus curve is shifted to the left.[143] In the cosmic string model, there is a pronounced asymmetry between $\rho > \bar{\rho}$ and $\rho < \bar{\rho}$. At small densities, the genus curve measures the (small number of) large voids, whereas for $\rho > \bar{\rho}$ the curve picks out[149] the large number of high density peaks which result from the fragmentation of the wakes (Fig. 28).

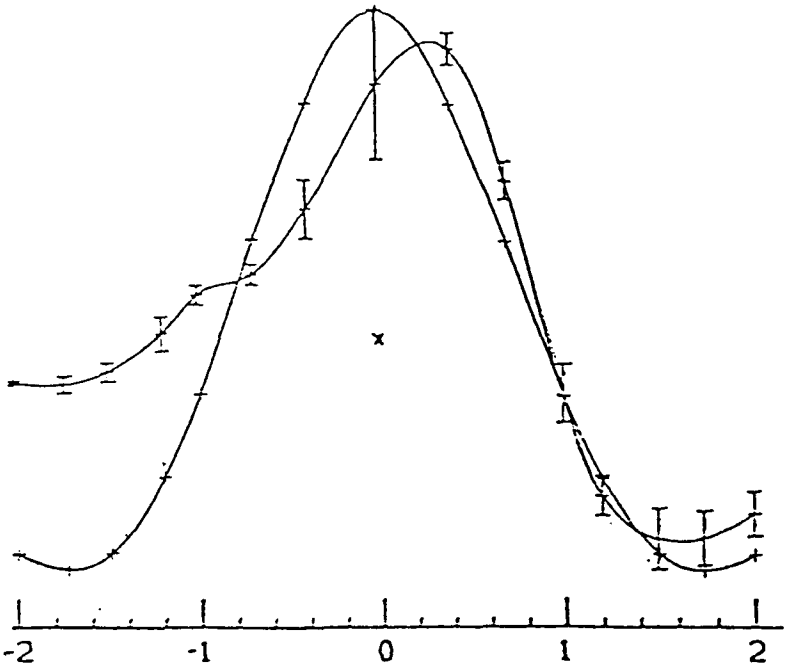

Fig. 28. The genus curve of the smoothed mass density field in a cosmic string wake toy model compared to the symmetric curve which results in the case of a model with a random distribution of mass points. The vertical axis is the genus (with genus 0 at the height of the "x"); the horizontal axis is a measure of density ("0" denotes average density).

The *counts in cell statistic*[150] can be successfully applied to distinguish between distributions of galaxies with the same power spectrum but different phases. The statistic is obtained by dividing the sample volume into equal size cells, counting the number $f(n)$ of cells containing n galaxies, and plotting $f(n)$ as a function of n.

We[151] have applied this statistic to a set of toy models of large scale structure. In each case, the sample volume was $(150\ \text{Mpc})^3$, the cell size $(3.75\ \text{Mpc})^3$, and the samples contained 90,000 galaxies. We compared a texture model (galaxies distributed in spherical clumps separated by 30 Mpc with a Gaussian radial density field of width 9 Mpc), a cosmic string model dominated by filaments [all galaxies

randomly distributed in filaments of dimensions $(60 \times 4 \times 4)$ Mpc^3 with mean separation 30 Mpc], a cosmic string wake model [same separation and wake dimensions $(40 \times 40 \times 2)$ Mpc^3], a cold dark matter model (obtained by Fourier-transforming the CDM power spectrum and assigning random phases), and a Poisson distribution of galaxies.

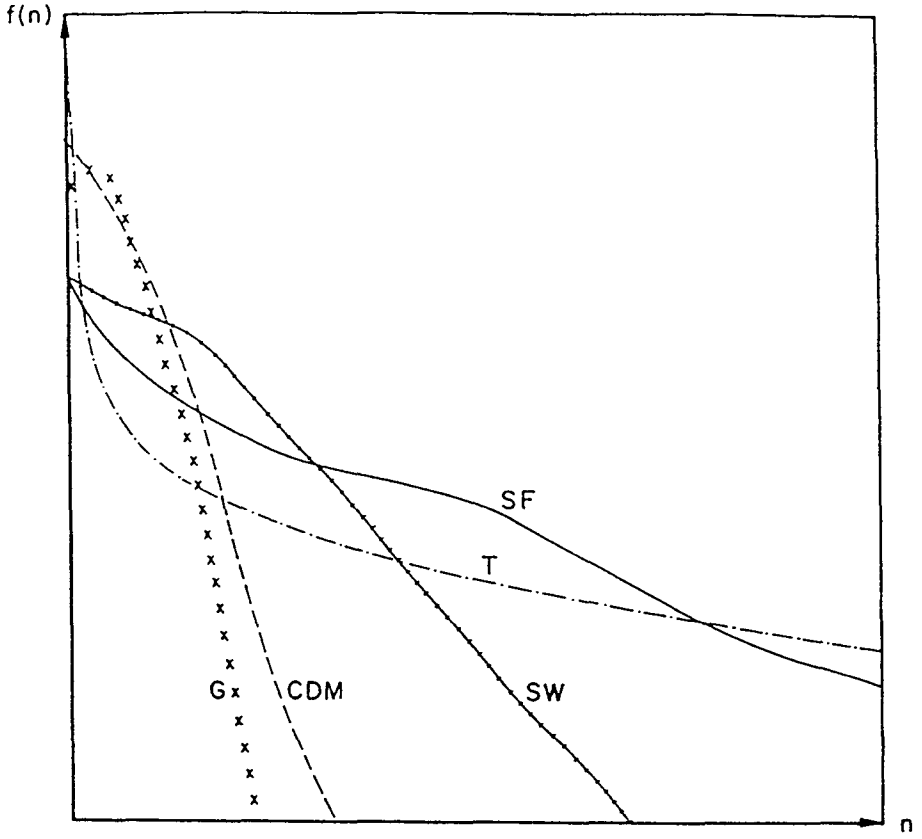

Fig. 29. The three-dimensional counts in cell statistics for a Poisson model (G), a cold dark matter model (CDM), cosmic string wakes (SW), string filaments (SF) and textures (T).

As shown in Fig. 29, the predicted curves differ significantly, demonstrating that this statistic is an excellent one at distinguishing different theories with the same power spectrum. The counts in cell statistic can also be applied to effectively two-dimensional surveys such as single slices of the CFA redshift survey.[9] The predictions of our theoretical toy models are shown in Fig. 30.

A third statistic which has proved useful[152] for distinguishing models with Gaussian and non-Gaussian phases is the void probability function $p(R)$, the probability that a sphere of radius R contains no galaxies.

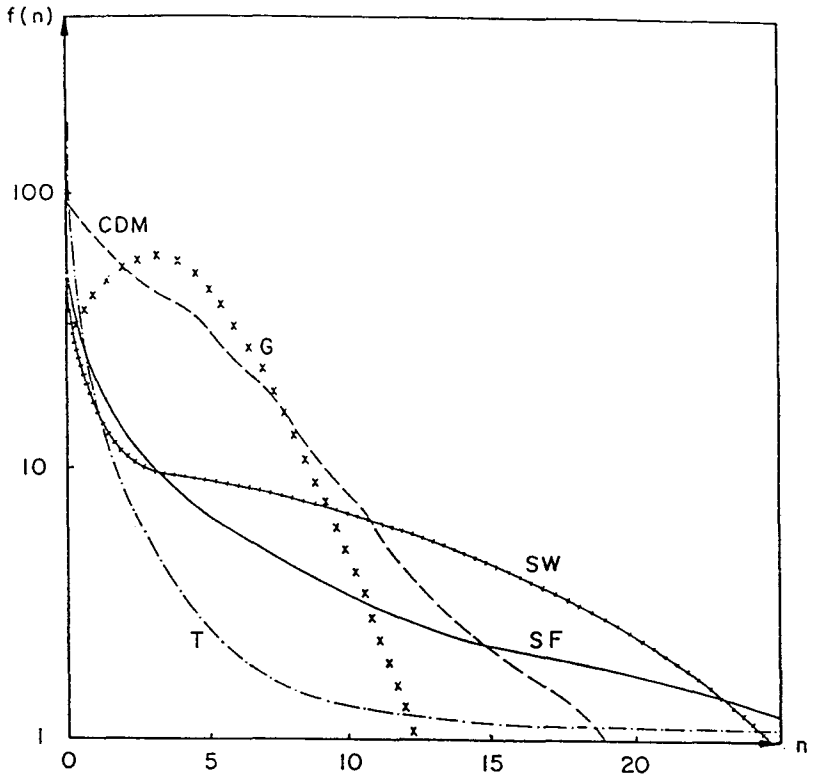

Fig. 30. The two-dimensional counts in cell statistic for a slice of the Universe of the dimensions of a CFA slice, evaluated for the same models as in Fig. 32.

8.3. *Signatures in the microwave background*

Inflationary Universe models predict essentially random phase fluctuations in the microwave background with a scale-invariant spectrum ($n = 1$). Small deviations from scale invariance are model-dependent and have recently been discussed in detail in Ref. 105. In all models, the amplitude must be consistent with structure formation. The COBE discovery[11] of anisotropies in the CMB has provided severe constraints on inflationary models. They are consistent with the present data only if the bias parameter b is about 1, which must be compared to the value $b \simeq 2.5$, which is the best value for galaxy formation in this model.[153] Note that full sky coverage is not essential for testing inflationary models since, in any set of local observations of $\delta T/T$, the results will form a Gaussian distribution about the r.m.s. value. Mixed dark matter models do slightly better and have recently been studied vigorously.

Cosmic string models predict non-Gaussian temperature anisotropies. One mechanism gives rise to localized linear temperature discontinuities;[154] its origin is illustrated in Fig. 31. Photons passing on different sides of a long straight string moving with velocity v reach the observer with a Doppler shift:

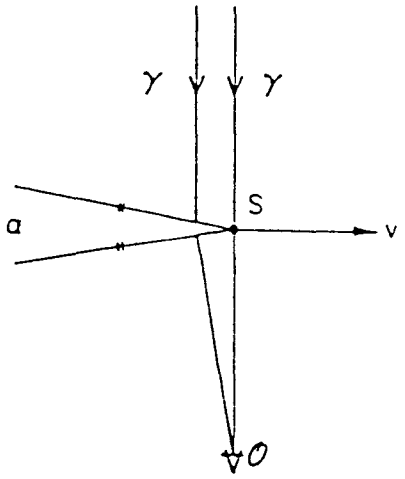

Fig. 31. Sketch of the mechanism producing linear discontinuities in the microwave temperature for photons γ passing on different sides of a moving string S (velocity v). O is the observer. Space perpendicular to the string is conical (deficit angle α).

$$\frac{\delta T}{T} \sim 8\pi G\mu v\gamma(v)\,. \tag{8.6}$$

To detect such discontinuities, an appropriate survey strategy (for example, a full sky survey) with small angular resolution is crucial. The distribution of strings also gives rise to Sachs–Wolfe type anisotropies.[155]

The theoretical error bars in the normalization of CMB anisotropies from strings are rather large — a direct consequence of the fact that the precise form of the scaling solution for the string network is not well determined. Nevertheless, we can consider a fixed set of cosmic string parameters and ask whether the normalizations of $G\mu$ from large scale structure data and from COBE are consistent. This has been done numerically in Ref. 106, and using an analytical toy model in Ref. 108.

The analytical model[108] is based on adding up as a random walk the individual Doppler shifts from strings which the microwave photons separated by angular scale v pass on different sides, and using this method to compute $\Delta T/T(\theta)$. By the use of the Bennett–Bouchet[156] string parameters, the result for $G\mu$ becomes

$$G\mu = (1.3 \pm 0.5)10^{-6}\,, \tag{8.7}$$

in good agreement with the requirements from large scale structure formation.[7]

To detect the predicted anisotropies from textures, it is again essential to have a full sky survey. However, larger angular resolution is adequate this time, since the specific signature for textures is a small number ($\sim$ 10) of hot and cold disks with amplitude[157]

$$\frac{\delta T}{T} \sim 0.06 \times 16\pi G\eta^2 \sim 3 \cdot 10^{-5} \tag{8.8}$$

and angular size of about 10°. The hot and cold spots are due to photons falling into the expanding Goldstone boson radiation field which results after texture collapse or due to photons climbing out of the potential well of the collapsing texture[158] (see Fig. 32).

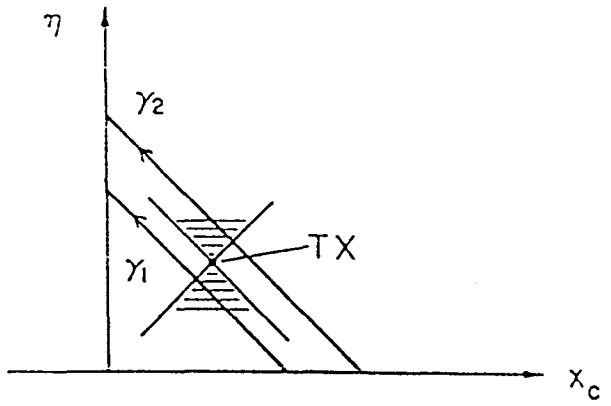

Fig. 32. Space–time diagram of a collapsing texture (backward light cone) and the resulting expanding Goldstone boson radiation (forward light cone). Unwinding of the texture occurs at point "TX." The light ray γ_2 falls into the potential well and is blueshifted, the ray γ_1 is redshifted.

Note that the texture model is not ruled out by the recent COBE results. The amplitude (8.8) is lower than the pixel sensitivity of the COBE maps. However, the predicted quadrupole CMB anisotropy (normalizing η by the large scale structure data) exceeds the COBE data point by a factor of between 2.5 and 4.[107,159] Hence, biasing must be invoked in order to try to explain the large scale structure data given the reduced value of η mandated by the discovery of CMB anisotropies.

8.4. *Conclusions*

Topological defect models of structure formation generically give rise to a scale-invariant power spectrum and are hence in good agreement with the recent COBE results on anisotropies of the CMB. The amplitude of the quadrupole temperature fluctuation can be used to normalize the models. For cosmic strings, the resulting normalization agrees well with the normalization from large scale structure data. For textures, there is a mismatch which requires the introduction of biasing. For textures, biasing is required, as is the case in the CDM model, where COBE demands a bias parameter $b \simeq 1$, whereas galaxy formation is said to demand[153] $b \simeq 2.5$.

It was emphasized that r.m.s. data is intrinsically unable to differentiate between topological defect models (with nonrandom phases) and inflationary modes (with random phases). We need statistics which are sensitive to nonrandom phases.

The most economical model for structure formation may be the model based on cosmic strings and hot dark matter. It requires no new particles (although it does require a finite neutrino mass), it agrees well with COBE and with the CFA redshift data, and it has clear signatures both for large scale structure and for CMB statistics.

9. Microphysics of Topological Defects

In Secs. 5–8, applications of topological defects were considered for which only the gravitational effects of such defects were important. However, there is much more structure in defects than is visible through gravity only. Topological defects are localized coherent excitations of gauge and Higgs fields. Matter interactions with the defects can probe this structure, give rise to effects of potential importance in cosmology, and possibly even to new direct signatures for defect models. In this and the following section, I will touch on two aspects involving microphysics. The discussion is not intended to be complete, but rather to whet the curiosity of the reader. First, we consider scattering of fermions by strings and monopoles.

Consider cosmic strings arising in a GUT model. The gauge and Higgs fields excited in the string mediate baryon- and lepton-number-violating interactions, for example (in the case of the Higgs field ϕ) via a term $\mathcal{L}_I$ in the interaction Lagrangian,

$$\mathcal{L}_I = g\bar{\psi}\phi\psi\,, \tag{9.1}$$

where g is the Yukawa coupling constant and ψ is a G multiplet of fermion fields, G being the gauge group.

Due to the term (9.1) in the interaction Lagrangian (and similar terms due to the GUT gauge fields coupling to ψ), there is a nonvanishing cross section σ for a quark to be scattered into a lepton by the defect. The conserved charges which are different in the initial and the final state are absorbed by the defect (which is here treated as a static background field).

In the case of fermion–monopole scattering, we would naively expect the cross section to be the geometrical one, $\sigma_{\rm geom}$:

$$\sigma_{\rm geom} \sim \eta^{-2}\,, \tag{9.2}$$

η^{-1} being the monopole radius. However, as first discovered by Callan[160] and Rubakov,[161] there is in fact an enhancement of the cross section and

$$\sigma \sim m^{-2}\,, \tag{9.3}$$

m being the fermion mass. Note that (9.3) represents a strong interaction cross section. Is there a similar enhancement for cosmic strings?

Classical physics considerations indicate that there will be no enhancement of the baryon-number-violating inelastic cross section for fermion–cosmic string

scattering.[162] For monopoles, the enhancement of the cross section can be viewed[162] as the consequence of an attractive magnetic moment–magnetic field force for s wave scattering (see Fig. 33). The potential energy is

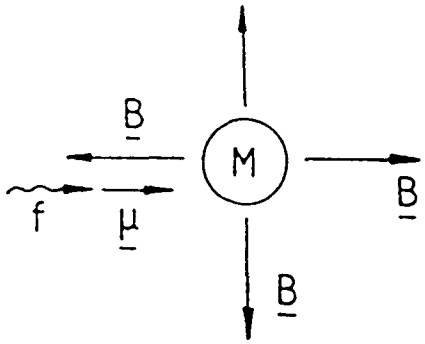

Fig. 33. The magnetic field $\underline{B}$ of a magnetic monopole M is parallel to the magnetic moment μ of an incident s wave fermion f.

$$V(r) \sim -(\underline{\mu} \cdot \underline{B}) \sim \frac{1}{r^2}, \tag{9.4}$$

where $\underline{B}$ is the U(1) magnetic field of the monopole, and $\underline{\mu}$ is the magnetic moment of the fermion. This results in an attractive force

$$F(r) \sim \frac{1}{r^3}. \tag{9.5}$$

In the case of ordinary (i.e. nonsuperconducting) cosmic strings, there is no long range magnetic field and hence no enhancement of the inelastic scattering cross section due to classical physics effects.[162] For superconducting cosmic strings, $\underline{B}$ is circular in the plane perpendicular to the string, and hence again $V(r)$ vanishes for s wave scattering.[163]

However, this is not the full story. Quantum-mechanical effects due to the non-vanishing gauge field at long distances from the string may induce an enhancement of the cross section, an effect reminiscent of but not identical to the Aharonov–Bohm[164] effect. (The AB effect appears in elastic scattering whereas our effect is due to inelastic scattering.) Let us therefore consider the quantum-mechanical scattering of fermions by cosmic strings.

Fermion–cosmic string scattering was first considered using a first-quantized approach, in Refs. 165 and 166; and then extended to inelastic scattering, in Ref. 167. The same results can be derived[168] using a perturbative second-quantized approach discussed in Ref. 162.

We consider fermion–defect scattering to first order in perturbation theory, treating the defect fields as a fixed background. The transition matrix element is

$$\mathcal{A} = {}_{\infty}\langle \psi' D | \psi D \rangle_{-\infty} = \langle \psi' D | S | \psi D \rangle \,, \tag{9.6}$$

where D stands for the defect and S is the scattering matrix. To first order in g

$$\mathcal{A} = \langle \psi' D | \int d^4x \, \mathcal{L}_I | \psi D \rangle = g \int dt \, d^3\underline{x} \langle \psi' | \bar{\psi}\psi | \psi \rangle \langle D | \phi | D \rangle \,, \tag{9.7}$$

where we have split the Hilbert space expectation value into its two tensor components.

From $\mathcal{A}$ we obtain the differential cross section through the usual relation,

$$\frac{d\sigma}{d\Omega} \sim \frac{1}{T} V \int d^3k' |\mathcal{A}|^2 \,, \tag{9.8}$$

where k' is the momentum of the final particle and V, T are cutoff volume and time respectively.

The procedure[162] now is to first evaluate $\mathcal{A}$ using free fermion wave functions, obtaining the unamplified cross section $(d\sigma/d\Omega)_0$. In a second step, we solve the Dirac equation in the presence of the defect and determine the amplification factor

$$A \equiv \frac{\psi(\eta^{-1})}{\psi_{\text{free}}(\eta^{-1})} \tag{9.9}$$

as the ratio of the actual and free wave functions ψ and ψ_{free}, respectively, evaluated at the core radius. Care must be taken to normalize both wave functions in the same way at infinity. Note that $r = 0$ (or $\rho = 0$) is taken to be the defect center for monopoles (or strings). The actual cross section becomes

$$\frac{d\sigma}{d\Omega} = A^4 \left(\frac{d\sigma}{d\Omega} \right)_0 \tag{9.10}$$

(four powers of A because in $|\mathcal{A}|^2$ two fermion wave functions appear and both are squared).

As an example,[168] consider the defect to be a cosmic string. The "geometric" amplitude $\mathcal{A}_0$ for a string along the z axis is

$$\mathcal{A}_0 \sim g\eta\eta^{-2} \int dt \, dz \frac{m}{(E_k E_{k'})^{1/2} V^{1/2}} e^{i(E_k - E_{k'})t} e^{i(k'_z - k_z)z} \left(\frac{m}{\eta} \right) . \tag{9.11}$$

The first factor of η comes from the amplitude of ϕ inside the core, integration over the core gives η^{-2}, the factor m comes from fermion spin sums (strictly speaking, the summation occurs only when one is computing the cross section), and the other factor, m/η, is due to the suppression of the free s wave amplitude at the core radius. Taking the square and performing one z and t integral gives

$$(\mathcal{A}_0)^2 \sim g^2 \left(\frac{m}{\eta} \right)^2 \frac{1}{E_k^2 V} LT \delta(E_k - E_{k'}) \delta(k_z - k_{z'}) \left(\frac{m}{\eta} \right)^2 . \tag{9.12}$$

The two delta functions express conservation of fermion energy and momentum along the string. L is a cutoff length. Integrating over $\underline{k}'$ and taking $E_{k'} \sim m$ gives the following differential cross section per unit length:

$$\frac{d\sigma}{d\Omega dL} \sim g^2 \left(\frac{m}{\eta}\right)^4 \frac{1}{m}. \tag{9.13}$$

The second step of the analysis is to calculate the amplification of the fermion wave function at the core radius by solving the Dirac equation in the background field of the string:

$$i\not{D}\psi - m\psi = 0. \tag{9.14}$$

The covariant derivative D_μ is

$$D_\mu = \partial_\mu + ieA_\mu. \tag{9.15}$$

Here, e is the charge of the fermions with respect to the subgroup of G for which A_μ is the gauge field. We assume that in units of $2\pi/e$ the flux of the string is α.

The solution of the Dirac equation makes use of the techniques pioneered by de Vega.[169] Since we are interested in field configurations independent of z, we choose a basis of γ matrices in which only γ_z mixes the upper and lower two-spinors of the four-spinor ψ:

$$\psi = \begin{pmatrix} \psi_u \\ \psi_d \end{pmatrix} \tag{9.16}$$

(see Refs. 162 and 168 for details). The Dirac equation (9.14) then decouples into two separate equations for ψ_u and ψ_d, respectively. Each is a system of two coupled first order differential equations.

Next, we expand ψ_u (and similarly ψ_d) in partial waves:

$$\psi_\mu = \sum_{n=-\infty}^{\infty} \begin{pmatrix} F_n(\rho) \\ G_n(\rho)e^{i\vartheta} \end{pmatrix} e^{in\vartheta - i\omega t}, \tag{9.17}$$

where ρ and ϑ are the polar coordinates in the plane perpendicular to the string. The radial function $F_n(\rho)$ satisfies a Bessel equation with index depending on α. It is necessary to solve the equation inside and outside the core, and to match the solutions at the core radius. For simplicity, we take the flux of the string to be concentrated on the cylinder $\rho = \eta^{-1}$.

Finally, for given α we search for the mode $n(\alpha)$ for which $F_n(\eta^{-1})$ is largest and use this coefficient to determine the amplification factor A. We find nontrivial amplification:

$$A = \left(\frac{\eta}{m}\right)^{p(\alpha)}; \tag{9.18}$$

$p(\alpha)$ is plotted in Fig. 34.

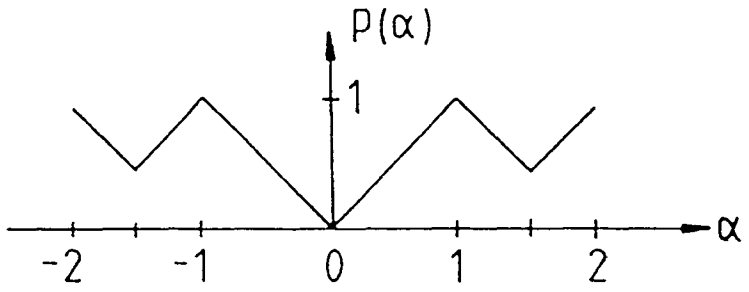

Fig. 34. The amplification exponent p plotted as a function of the fractional flux α on the string, in a model in which the fermions do not couple to the Higgs field.

We conclude that, in general, there is enhancement of the inelastic baryon-number-violating cross section due to intrinsically quantum-mechanical effects, namely the presence of a nonvanishing gauge potential in the far field limit. The maximal cross section is

$$\left.\frac{d\sigma}{d\Omega dL}\right|_{\max} \sim \frac{1}{m}, \tag{9.19}$$

i.e. a strong interaction cross section. However, $d\sigma/d\Omega\, dL$ is in general suppressed, the suppression rate depending on the flux/charge ratio α.

10. Topological Defects and Baryogenesis

Even through baryon number violation processes can be mediated by scattering of fermions from topological defects, no net baryon asymmetry can be created because both initial and final particles are in thermal equilibrium. As realized already by Sakharov,[170] in order to generate a nonvanishing baryon to entropy ratio n_B/s, it is necessary to have — in addition to the presence of baryon number violation processes — CP violation, and the processes which violate baryon number must occur out of thermal equilibrium. Two ways in which topological defects can contribute to baryogenesis are sketched in this section. The first occurs at the unification scale, the second at the electroweak scale.

10.1. *Collapsing cosmic strings and baryogenesis*

The standard mechanism[171] of GUT scale baryogenesis is based on the out-of-equilibrium decay of superheavy Higgs and gauge particles which freeze out of the initial thermal bath of particles at a temperature T_f of the order of the critical temperature T_c of the symmetry-breaking phase transition.

The above is a viable mechanism. Baryon-number-violating interactions arise in the Lagrangian [see (9.1)], CP violation already exists in the standard model, and the out-of-thermal-equilibrium condition is achieved as described in the previous paragraph.

However, if T_f is significantly lower than the mass m_X of the superheavy fields, then the net baryon asymmetry from the above process will be exponentially suppressed since[172] for $T_f < T < m_X$ the number density n_X of X particles decays exponentially:

$$n_X \sim e^{-\frac{m_X}{T}} . \tag{10.1}$$

As first pointed out in Refs. 173 and 174, cosmic strings can contribute to baryogenesis in the subset of GUT models which admit cosmic string solutions. The topological defects form out-of-equilibrium concentrations of GUT Higgs and gauge fields. Any emission of quanta from the strings proceeds through n_B-violating interactions, and hence a net baryon number may be generated. In Ref. 173, baryogenesis from cosmic strings collapsing onto a line was considered, and in Ref. 174 the contribution of cusp annihilation[175] to n_B/s was calculated. A third source [176] is the final collapse of a cosmic string loop.

As discussed in Sec. 4, cosmic string loops slowly shrink by emitting gravitational radiation. Once the radius R becomes comparable to the thickness w, there is no topological barrier for the remnant ball of energy

$$E \sim 2\pi\mu w \sim 2\pi\lambda^{-1/2}\eta \tag{10.2}$$

to decay into N_Q quanta of superheavy particles which in turn decay into low mass particles producing a net n_B. From (10.2) and $m_X \sim \eta$ it follows that

$$N_Q \sim 2\pi\lambda^{-1} . \tag{10.3}$$

If ΔB is the average net baryon number per decay, then the net baryon asymmetry produced per loop is

$$\Delta N_B \sim 2\pi\lambda^{-1}\Delta B . \tag{10.4}$$

Integrating over the distribution of loops after the time of the phase transition and dividing by the entropy density, we obtain[176]

$$\frac{n_B}{s} \sim \lambda^3 N^{-1} N_Q \Delta B , \tag{10.5}$$

where λ is the self-coupling constant of the Higgs field potential and N is the number of spin degrees of freedom in thermal equilibrium. In order to obtain this result we took $T_f = T_c$ and the mean separation of strings to be given by (4.9).

Since the number density of string loops decays only as a power law (redshift effect), the contribution of collapsing cosmic string loops to n_B/s decays only as T_F^{+3} and hence may dominate if $T_F \ll m_X$.

The result (10.5) must be compared to the maximal n_B/s which can be produced by the usual GUT mechanism:[177]

$$\frac{n_B}{s} \sim N^{-1} N_X \Delta B , \tag{10.6}$$

where N_X is the number of helicity states of all superheavy particles contributing to ΔB. Note that ΔB depends on the particular GUT model chosen. It enters both the standard and the topological defect baryogenesis mechanism in the same way. The value of ΔB has been estimated[177] to lie in the range $10^{-13} < \Delta B < 10^{-2}$. Thus, it is in principle possible to generate the observed n_B/s ratio, which lies in the range

$$\frac{n_B}{s} \sim 10^{-10}. \tag{10.7}$$

10.2. *Electroweak baryogenesis with electroweak strings*

Recently, it has been realized that nonperturbative processes associated with the SU(2) anomaly[178] of the standard model can erase a primordial $B + L$ asymmetry above the critical temperature of the electroweak symmetry-breaking transition. Thus, unless a primordial $B - L$ asymmetry is generated at the GUT scale (or extra symmetries such as those arising in supersymmetry are introduced,[179]) the observed n_B/s must be produced below the electroweak scale.

Initial ideas on how to generate a nonvanishing n_B/s were discussed in Ref. 180, and recently several concrete mechanisms were suggested.[181–183] All of these mechanisms work only if the electroweak phase transition is first order and therefore proceeds by bubble nucleation.

In electroweak baryogenesis, n_B violation occurs nonperturbatively through the anomaly, and CP violation is already contained in the standard model (although stronger CP violation is introduced in Refs. 181–183 by enlarging the Higgs sector of the model). The role of the first order phase transition in Refs. 181–183 is to generate bubble walls in which the n_B-violating processes can take place out of thermal equilibrium.

However, the order of the electroweak phase transition is not yet known. Therefore it is important to point out that electroweak baryogenesis mechanisms also exist if the phase transition is second order.[184] Topological defects may play the role of the bubble walls in a first order transition.

However, in the electroweak model and minimal extensions thereof, such as the two-Higgs-doublet model, there are no stable topological defects. But, as realized recently by Vachaspati[185] and considered much earlier by Nambu,[186] energetically stable stringlike defects can be constructed in the standard model by embedding the Nielsen–Olesen U(1) string into the standard model in the following way:

$$\Phi = f_{\mathrm{NO}}(\rho) e^{i\vartheta} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{10.8}$$

$$Z = A_{\mathrm{NO}}, \tag{10.9}$$

$$W' = W^2 = A = 0, \tag{10.10}$$

where $f_{\mathrm{NO}}(\rho)e^{i\vartheta}$ and A_{NO} are the Nielsen–Olesen string configurations.

It can be shown[187] that for large Weinberg angles, the above electroweak string configuration is energetically stable. Generally, electroweak strings have finite length.[185] Their ends can be viewed as a monopole–antimonopole pair.

Electroweak baryogenesis works as follows.[181,182] In the outer shell of the bubble wall where

$$|\Phi| < g\eta\,, \tag{10.11}$$

g being the gauge coupling constant, sphaleron[182] and local texture unwinding[181] events take place out of thermal equilibrium at the rate

$$\Gamma_B \sim \alpha_W^4 T^4\,. \tag{10.12}$$

Since there is CP violation in this region, a net baryon asymmetry will be generated. It is important that the region of n_B violation is in motion with a preferred direction (otherwise there is no net CP violation).

In a model with a second order phase transition and electroweak strings, the role of the bubble wall is played by the tips of the string which are moving towards each other due to the tension of the string.[184]

The typical length and separation of strings immediately after the phase transition is the Ginsburg length [see (4.9)]

$$\xi(t_G) \sim \lambda^{-1}\eta^{-1}\,. \tag{10.13}$$

The net baryon asymmetry produced by one electroweak string is

$$N_B = \int_{t_G}^{t_G+\Delta t_S} dt \frac{dV}{dt} \Delta t_c \varepsilon \Gamma_B\,, \tag{10.14}$$

where dV/dt is the rate of change of the volume in which n_B-violating processes are taking place out of equilibrium, Δt_c is the length of time any point in space remains in the above volume, and ε is the CP violation parameter defined such that $\varepsilon\Gamma_B$ is the net rate of baryon-number-producing processes per unit volume. Δt_S is the contraction time of the string.

Using (10.12) and the geometrical values for dV/dt, Δt_c and Δt_S (see Ref. 184 for details), we obtain

$$\frac{n_B}{s} \sim \lambda g^2 \varepsilon \alpha_W{}^4\,, \tag{10.15}$$

which is smaller than what is obtained with a first order phase transition only by a geometrical factor (the region in which n_B violation takes place out of equilibrium is much smaller).

Note, however, that the above model of electroweak baryogenesis using topological defects arising in a second order phase transition works only if extreme conditions on the parameters of the theory are satisfied: we require stable electroweak strings, a second order transition, sufficient CP violation ($\varepsilon \sim 1$), constraints on the coupling constants to make sure that sphalerons or textures fit into the region

$|\phi| < g\eta$ and that n_B-violating effects in the true vacuum are suppressed at $t_G + \Delta t_s$. Nevertheless, it is an existence proof that electroweak baryogenesis with a second order phase transition is possible. A more robust scenario of electroweak baryogenesis in a second order phase transition makes use of contracting topological defects left over from a phase transition at a scale $\eta > \eta_{\rm EW}$, $\eta_{\rm EW}$ being the electroweak symmetry-breaking scale.[188]

Acknowledgments

This article is based on a series of lectures given as part of the Troisième Cycle de la Suisse Romande at the EPFL, Lausanne, Switzerland, from 22 April to 13 May 1993. I wish to thank the organizers and in particular Prof. Jean-Pierre Derendinger for inviting me to give those lectures. I am grateful to all members of the particle theory group of the Université de Neuchâtel for their hospitality. Special thanks to Nicole Derendinger and Peter Bamert for typing the manuscript. Financial support from the US Department of Energy under grant FG02-91ER40688 is also acknowledged.

References

1. T. W. B. Kibble, *J. Phys.* **A9**, 1387 (1976).
2. A. Guth, *Phys. Rev.* **D23**, 347 (1981).
3. T. W. G. Kibble, *Phys. Rep.* **67**, 183 (1980).
4. A. Vilenkin, *Phys. Rep.* **121**, 263 (1985).
5. N. Turok, "Phase transitions as the origin of large-scale structure," in *Particles, Strings and Supernovae* (TASI-88), eds. A. Jevicki and C.-I. Tan (World Scientific, Singapore, 1989).
6. A. Vilenkin and E. P. S. Shellard, *Topological Defects and Cosmology* (Cambridge Univ. Press, Cambridge, 1994).
7. R. Brandenberger, *Phys. Scripta* **T36**, 114 (1991).
8. N. Turok, *Phys. Scripta* **T36**, 135 (1991).
9. V. de Lapparent, M. Geller and J. Huchra, *Ap. J. (Lett.)* **302**, L1 (1986).
10. N. Bahcall and R. Soneira, *Ap. J.* **270**, 20 (1983); A. Klypin and A. Kopylov, *Sov. Astron. Lett.* **9**, 41 (1983).
11. G. Smoot *et al.*, *Ap. J. (Lett.)* **396**, L1 (1992).
12. M. Mermin, *Rev. Mod. Phys.* **51**, 591 (1979).
13. P. de Gennes, *The Physics of Liquid Crystals* (Clarendon, Oxford, 1974).
14. I. Chuang, R. Durrer, N. Turok and B. Yurke, *Science* **251**, 1336 (1991); M. Bowick, L. Chandar, E. Schiff and A. Srivastava, *Science*, 18 Feb. 1994.
15. M. Salomaa and G. Volovik, *Rev. Mod. Phys.* **59**, 533 (1987).
16. A. Abrikosov, *JETP* **5**, 1174 (1957).
17. R. Brandenberger, "Lectures on modern cosmology and structure formation," Brown preprint BROWN-HET-893 (1993), to be publ. in the proc. of the 7th Swieca Summer School in Particles and Fields (10–23 Jan. 1993), eds. O. Eboli and V. Ribelles (World Scientific, Singapore, 1993).
18. R. Brandenberger, in *Physics of the Early Universe*, proc. of the 1989 Scottish Univ. Summer School in Physics, eds. J. Peacock, A. Heavens and A. Davies (SUSSP, Edinburgh, 1990).
19. E. Milne, *Z. Astrophys.* **6**, 1 (1933).

20. S. Shechtman, P. Schechter, A. Oemler, D. Tucker, R. Kirshner and H. Lin, "Strip-mining the sourthern sky: scratching the surface," in *Clusters and Superclusters of Galaxies*, ed. A. Fabian (Kluwer, Dordrecht, 1993).
21. See e.g.: S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972); Ya. B. Zel'dovich and I. Novikov, *The Structure and Evolution of the Universe* (Univ. of Chicago Press, Chicago, 1983).
22. E. Hubble, *Proc. Nat. Acad. Sci.* **15**, 168 (1927).
23. R. Alpher and R. Herman, *Rev. Mod. Phys.* **22**, 153 (1950); G. Gamov, *Phys. Rev.* **70**, 572 (1946).
24. R. Alpher, H. Bethe and G. Gamov, *Phys. Rev.* **73**, 803 (1948); R. Alpher and R. Herman, *Nature* **162**, 774 (1948).
25. For an excellent introduction see: S. Weinberg, *The First Three Minutes* (Basic Books, New York, 1988).
26. R. Dicke, P. J. E. Peebles, P. Roll and D. Wilkinson, *Ap. J.* **142**, 414 (1965).
27. A. Penzias and R. Wilson, *Ap. J.* **142**, 419 (1965).
28. J. Mather *et al.*, *Ap. J. (Lett.)* **354**, L37 (1990).
29. H. Gush, M. Halpern and E. Wishnow, *Phys. Rev. Lett.* **65**, 937 (1990).
30. See e.g.: G. Efstathiou, in *Physics of the Early Universe*, proc. of the 1989 Scottish Univ. Summer School in Physics, eds. J. Peacock, A. Heavens and A. Davies (SUSSP, Edinburgh, 1990).
31. B. Pagel, *Ann. N.Y. Acad. Sci.* **647**, 131 (1991).
32. H. Arp, G. Burbidge, F. Hoyle, J. Narlikar and N. Vickramasinghe, *Nature* **346**, 807 (1990).
33. N. Tsamis and R. Woodard, *Phys. Lett.* **B301**, 351 (1993).
34. J. Tonry, "Relativistic astrophysics and particle cosmology" (TEXAS-PASCOS 92, Berkeley, 12–18 Dec. 1992), publ. in *Ann. N.Y. Acad. Sci.* **688**, 114 (1993).
35. V. Trimble, *Ann. Rev. Astr. Astrophys.* **25**, 423 (1987).
36. E. Bertschinger and A. Dekel, *Ap. J.* **336**, L5 (1989).
37. M. Strauss, M. Davis, A. Yahil and J. Huchra, *Ap. J.* **385**, 421 (1992).
38. G. Ellis, *Class. Quantum Grav.* **5**, 207 (1988); G. Ellis, D. Lyth and M. Mijic, *Phys. Lett.* **B271**, 52 (1991); M. Madsen, G. Ellis, J. Mimoso and J. Butcher, *Phys. Rev.* **D46**, 1399 (1992).
39. J. Primack, D. Seckel and B. Sadoulet, *Ann. Rev. Nucl. Part. Sci.* **38**, 751 (1988).
40. L. da Costa, in *The Distribution of Matter in the Universe*, eds. D. Gerbal and G. Mamon, in press (1991).
41. T. Broadhurst, R. Ellis, D. Koo and A. Szalay, *Nature* **343**, 726 (1990).
42. J. Ostriker and L. Cowie, *Ap. J. (Lett.)* **243**, L127 (1981).
43. A. Linde, *Particle Physics and Inflationary Cosmology* (Harwood, Chur, 1990).
44. S. Blau and A. Guth, "Inflationary cosmology," in *300 Years of Gravitation*, eds. S. Hawking and W. Israel (Cambridge Univ. Press, Cambridge, 1987).
45. K. Olive, *Phys. Rep.* **190**, 307 (1990).
46. D. Kazanas, *Ap. J.* **241**, L59 (1980).
47. W. Press, *Phys. Scr.* **21**, 702 (1980).
48. G. Chibisov and V. Mukhanov, "Galaxy formation and phonons," Lebedev Physical Institute Preprint No. 162 (1980); G. Chibisov and V. Mukhanov, *Mon. Not. R. Astron. Soc.* **200**, 535 (1982).
49. V. Lukash, *Pis'ma Zh. Eksp. Teor. Fiz.* **31**, 631 (1980).
50. K. Sato, *Mon. Not. R. Astron. Soc.* **195**, 467 (1981).
51. L. Landau and E. Lifshitz, *Theoretical Physics, Vol. II: Classical Fields* (Pergamon, London, 1958).

52. D. Kirzhnits and A. Linde, *Pis'ma Zh. Eksp. Teor. Fiz.* **15**, 745 (1972); D. Kirzhnits and A. Linde, *Zh. Eksp. Teor. Fiz.* **67**, 1263 (1974); C. Bernard, *Phys. Rev.* **D9**, 3313 (1974); L. Dolan and R. Jackiw, *Phys. Rev.* **D9**, 3320 (1974); S. Weinberg, *Phys. Rev.* **D9**, 3357 (1974).
53. R. Brandenberger, *Rev. Mod. Phys.* **57**, 1 (1985).
54. A. Linde, *Phys. Lett.* **B129**, 177 (1983).
55. G. Mazenko, W. Unruh and R. Wald, *Phys. Rev.* **D31**, 273 (1985).
56. A. Linde, *Phys. Lett.* **B108**, 389 (1982); A. Albrecht and P. Steinhardt, *Phys. Rev. Lett.* **48**, 1220 (1982).
57. J. Langer, *Physica* **73**, 61 (1974).
58. S. Coleman, *Phys. Rev.* **D15**, 2929 (1977); C. Callan and S. Coleman, *Phys. Rev.* **D16**, 1762 (1977).
59. M. Voloshin, Yu. Kobzarev and L. Okun, *Sov. J. Nucl. Phys.* **20**, 644 (1975).
60. M. Stone, *Phys. Rev.* **D14**, 3568 (1976); M. Stone, *Phys. Lett.* **B67**, 186 (1977).
61. P. Frampton, *Phys. Rev. Lett.* **37**, 1380 (1976).
62. S. Coleman, in *The Whys of Subnuclear Physics* (Erice, 1977), ed. A. Zichichi (Plenum, New York, 1979).
63. G. Ross, *Grand Unified Theories* (Benjamin, Reading, 1985).
64. S. Coleman, "Secret symmetry," in *New Phenomena in Subnuclear Physics* (Erice, 1975), ed. A. Zichichi (Plenum, New York, 1977).
65. A. Starobinsky, *Phys. Lett.* **B91**, 99 (1980).
66. N. Steenrod, *Topology of Fibre Bundles* (Princeton Univ. Press, Princeton, 1951).
67. H. Nielsen and P. Olesen, *Nucl. Phys.* **B61**, 45 (1973).
68. R. Davis, *Phys. Rev.* **D35**, 3705 (1987).
69. N. Turok, *Phys. Rev. Lett.* **63**, 2625 (1989).
70. P. Langacker and S.-Y. Pi, *Phys. Rev. Lett.* **45**, 1 (1980).
71. T. W. B. Kibble and E. Weinberg, *Phys. Rev.* **D43**, 3188 (1991).
72. T. W. B. Kibble, *Acta Phys. Pol.* **B13**, 723 (1982).
73. S. Rudaz and A. Srivastava, *Mod. Phys. Lett.* **A8**, 1443 (1993).
74. J. Ye and R. Brandenberger, *Nucl. Phys.* **B346**, 149 (1990).
75. M. Hindmarsh, A.-C. Davis and R. Brandenberger, "Formation of topological defects in first order phase transitions," *Phys. Rev. D*, in press (1994).
76. Ya. B. Zel'dovich, I. Kobzarev and L. Okun, *Zh. Eksp. Teor. Fiz.* **67**, 3 (1974).
77. Ya. B. Zel'dovich and M. Khlopov, *Phys. Lett.* **B79**, 239 (1978); J. Preskill, *Phys. Rev. Lett.* **43**, 1365 (1979).
78. M. Barriola and A. Vilenkin, *Phys. Rev. Lett.* **63**, 341 (1989); S. Rhie and D. Bennett, *Phys. Rev. Lett.* **65**, 1709 (1990).
79. T. Vachaspati and A. Vilenkin, *Phys. Rev.* **D30**, 2036 (1984).
80. Ya. B. Zel'dovich, *Mon. Not. R. Astron. Soc.* **192**, 663 (1980); A. Vilenkin, *Phys. Rev. Lett.* **46**, 1169 (1981).
81. D. Foerster, *Nucl. Phys.* **B81**, 84 (1974).
82. N. Turok, in *Proceedings of the 1987 CERN/ESO Winter School on Cosmology and Particle Physics* (World Scientific, Singapore, 1988).
83. T. W. B. Kibble and N. Turok, *Phys. Lett.* **B116**, 141 (1982).
84. R. Brandenberger, *Nucl. Phys.* **B293**, 812 (1987).
85. E. P. S. Shellard, *Nucl. Phys.* **B283**, 624 (1987).
86. R. Matzner, *Computers in Physics* **1**, 51 (1988); K. Moriarty, E. Myers and C. Rebbi, *Phys. Lett.* **B207**, 411 (1988); E. P. S. Shellard and P. Ruback, *Phys. Lett.* **B209**, 262 (1988).
87. P. Ruback, *Nucl. Phys.* **B296**, 669 (1988).

88. T. Vachaspati and A. Vilenkin, *Phys. Rev.* **D31**, 3052 (1985); N. Turok, *Nucl. Phys.* **B242**, 520 (1984); C. Burden, *Phys. Lett.* **B164**, 277 (1985).
89. A. Albrecht and N. Turok, *Phys. Rev. Lett.* **54**, 1868 (1985).
90. D. Bennett and F. Bouchet, *Phys. Rev. Lett.* **60**, 257 (1988).
91. B. Allen and E. P. S. Shellard, *Phys. Rev. Lett.* **64**, 119 (1990).
92. A. Albrecht and N. Turok, *Phys. Rev.* **D40**, 973 (1989).
93. R. Brandenberger and J. Kung, in *The Formation and Evolution of Cosmic Strings*, eds. G. Gibbons, S. Hawking and T. Vachaspati (Cambridge Univ. Press, Cambridge, 1990).
94. See e.g.: C. Misner, K. Thorne and J. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
95. B. Carter, *Phys. Rev.* **D41**, 3869 (1990).
96. E. Copeland, T. W. B. Kibble and D. Austin, *Phys. Rev.* **D45**, 1000 (1992).
97. T. Prokopec, *Phys. Lett.* **B262**, 215 (1991); R. Leese and T. Prokopec, *Phys. Rev.* **D44**, 3749 (1991).
98. D. Spergel, N. Turok, W. Press and B. Ryden, *Phys. Rev.* **D43**, 1038 (1991).
99. T. Prokopec, A. Sornborger and R. Brandenberger, *Phys. Rev.* **D45**, 1971 (1992).
100. J. Borrill, E. Copeland and A. Liddle, *Phys. Lett.* **B258**, 310 (1991).
101. A. Sornborger, *Phys. Rev.* **D48**, 3517 (1993).
102. L. Perivolaropoulos, *Phys. Rev.* **D46**, 1858 (1992).
103. E. Lifschitz, *Zh. Eksp. Teor. Fiz.* **16**, 587 (1946).
104. V. Mukhanov, H. Feldman and R. Brandenberger, *Phys. Rep.* **215**, 203 (1992).
105. L. Krauss and M. White, *Phys. Rev. Lett.* **69**, 869 (1992); D. Salopek, *Phys. Rev. Lett.* **69**, 3602 (1992); R. Davis, H. Hodges, G. Smoot, P. Steinhardt and M. Turner, *Phys. Rev. Lett.* **69**, 1856 (1992); A. Liddle and D. Lyth, *Phys. Lett.* **B291**, 391 (1992); F. Lucchin, S. Matarrese and S. Mollerach, *Ap. J. (Lett.)* **401**, L49 (1992); T. Souradeep and V. Sahni, *Mod. Phys. Lett.* **A7**, 3541 (1992).
106. D. Bennett, A. Stebbins and F. Bouchet, *Ap. J. (Lett.)* **399**, L5 (1992).
107. D. Bennett and S. Rhie, *Ap. J. (Lett.)* **406**, L7 (1993).
108. L. Perivolaropoulos, *Phys. Lett.* **B298**, 305 (1993).
109. R. Sachs and A. Wolfe, *Ap. J.* **147**, 73 (1967).
110. M. Rees and D. Sciama, *Nature* **217**, 511 (1968).
111. D. Fixsen, E. Cheng and D. Wilkinson, *Phys. Rev. Lett* **50**, 620 (1983).
112. K. Ganga, E. Cheng, S. Meyer and L. Page, *Ap. J. (Lett.)* **410**, L54 (1993).
113. A. Lasenby, "Ground-based observations of the cosmic microwave background," *Current Topics in Astrofundamental Physics*, eds. N. Sanchez and A. Zichichi (World Scientific, Singapore, 1993).
114. P. Coles and J. Barrow, *Mon. Not. R. Astron. Soc.* **228**, 407 (1987); P. Coles, *Mon. Not. R. Astron. Soc.* **234**, 501 (1988); J. Gott *et al.*, *Ap. J.* **340**, 625 (1989).
115. L. Perivolaropoulos, *Phys. Rev.* **D48**, 1530 (1993); R. Moessner, L. Perivolaropoulos and R. Brandenberger, "A cosmic string specific signature on the cosmic microwave background," Brown Univ. preprint BROWN-HET-911 (1993), *Ap. J.* (in press); P. Graham, N. Turok, P. Lubin and J. Schuster, "A simple test for non-Gaussianity in CMBR measurements," Princeton preprint PUP-TH-1408 (1993).
116. Ya. B. Zel'dovich, J. Einasto and S. Shandarin, *Nature* **300**, 407 (1982); J. Oort, *Ann. Rev. Astron. Astrophys.* **21**, 373 (1983); R. B. Tully, *Ap. J.* **257**, 389 (1982); S. Gregory, L. Thomson and W. Tifft, *Ap. J.* **243**, 411 (1980).
117. G. Chincarini and H. Rood, *Nature* **257**, 294 (1975); J. Einasto, M. Joeveer and E. Saar, *Mon. Not. R. Astron. Soc.* **193**, 353 (1980); R. Giovanelli and M. Haynes, *Astron. J.* **87**, 1355 (1982); D. Batuski and J. Burns, *Ap. J.* **299**, 5 (1985).

118. R. Kirshner, A. Oemler, P. Schechter and S. Shechtman, *Ap. J. (Lett.)* **248**, L57 (1981).
119. M. Joeveer, J. Einasto and E. Tago, *Mon. Not. R. Astron. Soc.* **185**, 357 (1978); L. da Costa *et al.*, *Ap. J.* **327**, 544 (1988).
120. A. Dressler, D. Lynden-Bell, D. Burstein, R. Davies, S. Faber, R. Terlevich and G. Wegner, *Ap. J.* **313**, 42 (1987); C. Collins, R. Joseph and N. Robertson, *Nature* **320**, 506 (1986).
121. E. Bertschinger and A. Dekel, *Ap. J. (Lett.)* **336**, L5 (1989).
122. M. Aaronson, G. Bothun, J. Mould, R. Schommer and M. Cornell, *Ap. J.* **302**, 536 (1986); J. Lucey and D. Carter, *Mon. Not. R. Astron. Soc.* **235**, 1177 (1988).
123. G. Abell, *Ap. J. Suppl.* **3**, 211 (1958).
124. M. Davis and J. Huchra, *Ap. J.* **254**, 437 (1982).
125. P. J. E. Peebles, in *Physical Cosmology*, 1979 Les Houches Lectures, eds. R. Balian, J. Audouze and D. Schramm (North-Holland, Amsterdam, 1980).
126. A. Dekel, G. Blumenthal, J. Primack and S. Olivier, *Ap. J. (Lett.)* **338**, L5 (1989); W. Sutherland, *Mon. Not. R. Astron. Soc.* **234**, 159 (1988).
127. N. Bahcall, *Ann. Rev. Astron. Astrophys.* **15**, 505 (1977); B. Binggeli, in *Nearly Normal Galaxies*, ed. S. Faber (Springer, New York, 1986).
128. See e.g.: E. Athanassoula and A. Bosma, in *Large Scale Structures of the Universe*, IAU Symposium No. 130, eds. J. Audouze *et al.* (Kluwer, Dordrecht, 1988).
129. See e.g.: S. M. Fall, in *Internal Kinetics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel, Dordrecht, 1983).
130. N. Turok, *Nucl. Phys.* **B242**, 520 (1984).
131. N. Turok and R. Brandenberger, *Phys. Rev.* **D33**, 2175 (1986); A. Stebbins, *Ap. J. (Lett.)* **303**, L21 (1986); H. Sato, *Prog. Theor. Phys.* **75**, 1342 (1986).
132. A. Vilenkin, *Phys. Rev.* **D23**, 852 (1981); J. Gott, *Ap. J.* **288**, 422 (1985); W. Hiscock, *Phys. Rev.* **D31**, 3288 (1985); B. Linet, *Gen. Relativ. Gravit.* **17**, 1109 (1985); D. Garfinkle, *Phys. Rev.* **D32**, 1323 (1985); R. Gregory, *Phys. Rev. Lett.* **59**, 740 (1987).
133. J. Silk and A. Vilenkin, *Phys. Rev. Lett.* **53**, 1700 (1984).
134. T. Vachaspati, *Phys. Rev. Lett.* **57**, 1655 (1986).
135. A. Stebbins, S. Veeraraghavan, R. Brandenberger, J. Silk and N. Turok, *Ap. J.* **322**, 1 (1987).
136. R. Brandenberger, L. Perivolaropoulos and A. Stebbins, *Int. J. Mod. Phys.* **A5**, 1633 (1990); L. Perivolarapoulos, R. Brandenberger and A. Stebbins, *Phys. Rev.* **D41**, 1764 (1990).
137. E. Bertschinger and P. Watts, *Ap. J.* **328**, 23 (1988).
138. D. Vollick, *Phys. Rev.* **D45**, 1884 (1992).
139. T. Vachaspati and A. Vilenkin, *Phys. Rev. Lett.* **67**, 1057 (1991).
140. A. Gooding, D. Spergel and N. Turok, *Ap. J. (Lett.)* **372**, L5 (1991).
141. C. Park, D. Spergel and N. Turok, *Ap. J. (Lett.)* **373**, L53 (1991).
142. R. Cen, J. Ostriker, D. Spergel and N. Turok, *Ap. J.* **383**, 1 (1991).
143. A. Gooding, C. Park, D. Spergel, N. Turok and J. Gott, *Ap. J.* **393**, 42 (1992).
144. L. Grishchuk and Y. Sidorov, *Class. Quantum Grav.* **6**, L161 (1989); L. Grishchuk and Y. Sidorov, *Phys. Rev.* **D42**, 3413 (1990); L. Grishchuk, *Phys. Rev. Lett.* **70**, 2371 (1993).
145. T. Prokopec, *Class. Quantum Grav.* **10**, 2295 (1993); A. Albrecht, P. Ferreira, M. Joyce and T. Prokopec, "Inflation and squeezed quantum states," Imperial College preprint (1993).
146. R. Brandenberger, N. Kaiser, D. Schramm and N. Turok, *Phys. Rev. Lett.* **59**, 2371 (1987).

147. R. Brandenberger, N. Kaiser and N. Turok, *Phys. Rev.* **D36**, 2242 (1987).
148. J. Gott, A. Melott and M. Dickinson, *Ap. J.* **306**, 341 (1986).
149. J. Gerber and R. Brandenberger, "Topology of large-scale structure in a cosmic string wake model," Brown Univ. preprint BROWN-HET-829 (1991).
150. W. Saslaw, *Ap. J.* **297**, 49 (1985).
151. S. Ramsey, senior thesis, Brown Univ. (1992); D. M. Kaplan, senior thesis, Brown Univ. (1993); R. Brandenberger, D. M. Kaplan and S. Ramsey, "Some statistics for measuring large-scale structure," Brown preprint BROWN-HET-922 (1993).
152. E. Valentini and R. Brandenberger, in preparation (1993); D. Weinberg and S. Cole, *Mon. Not. R. Astron. Soc.* **259**, 652 (1992).
153. S. White, C. Frenk, M. Davis and G. Efstathiou, *Ap. J.* **313**, 505 (1987).
154. N. Kaiser and A. Stebbins, *Nature* **310**, 391 (1984).
155. J. Traschen, N. Turok and R. Brandenberger, *Phys. Rev.* **D34**, 919 (1986); S. Veeraraghavan and A. Stebbins, *Ap. J. (Lett.)* **395**, L55 (1992).
156. D. Bennett and F. Bouchet, *Phys. Rev.* **D41**, 2408 (1990).
157. R. Durrer and D. Spergel, "Microwave anisotropies from texture seeded structure formation," publ. in *Santa Monica Astropart. 1990*, 288 (1991); R. Durrer, A. Howard and Z.-H. Zhou, *Phys. Rev.* **D49**, 681 (1994).
158. N. Turok and D. Spergel, *Phys. Rev. Lett.* **64**, 2736 (1990).
159. U.-L. Pen, D. Spergel and N. Turok, *Phys. Rev.* **D49**, 692 (1994).
160. C. Callan, *Phys. Rev.* **D25**, 2141 (1982); *Phys. Rev.* **D26**, 2058 (1982).
161. V. Rubakov, *Pis'ma Zh. Eksp. Fiz.* **33**, 658 (1981) [*JETP Lett.* **33**, 644 (1981)]; *Nucl. Phys.* **B203**, 311 (1982).
162. R. Brandenberger, A.-C. Davis and A. Matheson, *Nucl. Phys.* **B307**, 909 (1988).
163. R. Brandenberger and L. Perivolaropoulos, *Phys. Lett.* **B208**, 396 (1988).
164. Y. Aharanov and D. Bohm, *Phys. Rev.* **119**, 485 (1959).
165. M. Alford and F. Wilczek, *Phys. Rev. Lett.* **62**, 1071 (1989).
166. P. de Sousa Gerbert, *Phys. Rev.* **D40**, 1346 (1989); P. de Sousa Gerbert and R. Jackiw, *Commun. Math. Phys.* **124**, 229 (1989).
167. M. Alford, J. March-Russell and F. Wilczek, *Nucl. Phys.* **B328**, 140 (1989).
168. W. Perkins, L. Perivolaropoulos, A.-C. Davis, R. Brandenberger and A. Matheson, *Nucl. Phys.* **B353**, 237 (1991).
169. H. de Vega, *Phys. Rev.* **D18**, 2932 (1978).
170. A. Sakharov, *Pis'ma Zh. Eksp. Teor. Fiz.* **5**, 32 (1967).
171. M. Yoshimura, *Phys. Rev. Lett.* **41**, 281 (1978); A. Ignatiev, N. Krasnikov, V. Kuzmin and A. Tavkhelidze, *Phys. Lett.* **B76**, 436 (1978); S. Dimopoulos and L. Susskind, *Phys. Rev.* **D18**, 4500 (1978); S. Weinberg, *Phys. Rev. Lett.* **42**, 850 (1979); D. Toussaint, S. Trieman, F. Wilczek and A. Zee, *Phys. Rev.* **D19**, 1036 (1979).
172. See e.g.: S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972).
173. P. Bhattacharjee, T. W. B. Kibble and N. Turok, *Phys. Lett.* **B119**, 95 (1982).
174. M. Kawasaki and K. Maeda, *Phys. Lett.* **B209**, 271 (1988).
175. R. Brandenberger, *Nucl. Phys.* **B293**, 812 (1987).
176. R. Brandenberger, A.-C. Davis and M. Hindmarsh, *Phys. Lett.* **263**, 239 (1991); A.-C. Davis and M. Earnshaw, *Nucl. Phys.* **B394**, 21 (1993).
177. D. Nanopoulos and S. Weinberg, *Phys. Rev.* **D20**, 2484 (1979).
178. G. 't Hooft, *Phys. Rev. Lett.* **37**, 8 (1976).
179. L. Ibáñez and F. Quevedo, *Phys. Lett.* **B283**, 261 (1992).
180. M. Shaposhnikov, *JETP Lett.* **44**, 465 (1986); M. Shaposhnikov, *Nucl. Phys.* **B287**, 757 (1987); L. McLerran, *Phys. Rev. Lett.* **62**, 1075 (1989).

181. N. Turok and J. Zadrozny, *Phys. Rev. Lett.* **65**, 2331 (1990); N. Turok and J. Zadrozny, *Nucl. Phys.* **B358**, 471 (1991); L. McLerran, M. Shaposhnikov, N. Turok and M. Voloshin, *Phys. Lett.* **B256**, 451 (1991).
182. A. Cohen, D. Kaplan and A. Nelson, *Phys. Lett.* **B263**, 86 (1991).
183. A. Nelson, D. Kaplan and A. Cohen, *Nucl. Phys.* **B373**, 453 (1992).
184. R. Brandenberger and A.-C. Davis, *Phys. Lett.* **B308**, 79 (1993).
185. T. Vachaspati, *Phys. Rev. Lett.* **68**, 1977 (1992).
186. Y. Nambu, *Nucl. Phys.* **B130**, 505 (1977).
187. M. James, L. Perivolaropoulos and T. Vachaspati, *Phys. Rev* **D46**, R5232 (1992).
188. R. Brandenberger, A.-C. Davis and M. Trodden, "Cosmics strings and electroweak baryogenesis," Brown preprint BROWN-HET-935 (1994).