Journal of Physics A: Mathematical and General

Topology of cosmic domains and strings

J. Phys. A: Math. Gen., Vol. 9, No. 8, 1976. Printed in Great Britain. © 1976

Topology of cosmic domains and strings

T W B Kibble
Blackett Laboratory, Imperial College, Prince Consort Road, London SW7 2BZ, UK

Received 11 March 1976

Abstract. The possible domain structures which can arise in the universe in a spontaneously broken gauge theory are studied. It is shown that the formation of domain walls, strings or monopoles depends on the homotopy groups of the manifold of degenerate vacua. The subsequent evolution of these structures is investigated. It is argued that while theories generating domain walls can probably be eliminated (because of their unacceptable gravitational effects), a cosmic network of strings may well have been formed and may have had important cosmological effects.

1. Introduction

Gauge theories with spontaneous symmetry breaking have come to play a central role in elementary particle theory. Kirzhnits (1972), and Kirzhnits and Linde (1972, 1974) suggested that as in ferromagnets and superconductors the full symmetry may be restored above some critical temperature. That this actually happens in a class of theories where the symmetry breaking occurs through the acquisition of a vacuum expectation value by an elementary scalar field has been demonstrated by Weinberg (1974) while Jacobs (1974) and Harrington and Yildiz (1975) have examined models of dynamical symmetry breaking in which the role of the order parameter is played by a composite field operator. (See also Bernard 1974, Dolan and Jackiw 1974, Dashen *et al* 1975, and Linde 1975.)

In the hot big-bang model, the universe must at one time have exceeded the critical temperature so that initially the symmetry was unbroken. It is then natural to enquire whether as it expands and cools it might acquire a domain structure, as in a ferromagnet cooled through its Curie point. Zel'dovich *et al* (1974; see also Kobzarev *et al* 1974) have discussed this question, and in particular pointed out the important gravitational effects to be expected of domain walls. Everett (1974) has studied the propagation of waves across a domain boundary.

The aim of this paper is to discuss the topology and scale of the possible cosmic structures that might arise. After reviewing the results of Weinberg and others on phase transitions in a simple class of models in § 2, we discuss in § 3 the initial formation of 'protodomains' as the universe cools. The possible topological configurations are examined in § 4. These include domain walls, strings and monopoles. We show that their occurrence is largely determined by the topology of the manifold M of degenerate vacuum states (specifically by its homotopy groups). (Coleman (1976) has stated the same result in a different context. In the case of monopoles it has been proved by Krive and Chudnovskiĭ 1975.) In § 5 we examine the later evolution of these structures. We show that domain walls can be of two main types with very different transmissivity, and that highly reflecting walls may behave very differently from the essentially transparent ones considered by Zel'dovich *et al* (1974). In all cases however the typical scale of the domain structure will grow with time until it is comparable with the radius of the universe. Hence the argument of Zel'dovich *et al*, to the effect that domain walls cannot have persisted beyond the recombination era because their gravitational effect would have destroyed the isotropy of the 3 K background radiation, applies. If domain walls existed they must have disappeared by then. This in turn is possible only if the universe has a small built-in asymmetry. The exclusion of theories generating domain walls is an interesting example of a restriction on elementary particle theories derived from cosmology.

The general conclusion is that there is a rich variety of possible topological structures which might have appeared in the early history of the universe. Few of these (monopoles excepted) are likely to be stable enough to have survived to the present, but they may nevertheless be of importance in understanding the history of the universe, for example the evolution of galaxies. The conclusions are summarized in more detail in § 6.

2. The phase transition

Although our discussion will be quite general, for illustrative purposes it is convenient to have a specific example in mind. Let us consider an N-component real scalar field ϕ with a Lagrangian invariant under the orthogonal group O(N), and coupled in the usual way to $\frac{1}{2}N(N-1)$ vector fields represented by an antisymmetric matrix B_μ. We can take

$$L=\tfrac{1}{2}(D_\mu\phi)^2-\tfrac{1}{8}g^2(\phi^2-\eta^2)^2+\tfrac{1}{8}\mathrm{Tr}(B_{\mu\nu}B^{\mu\nu}) \tag{1}$$

with

$$D_\mu\phi=\partial_\mu\phi-eB_\mu\phi$$
$$B_{\mu\nu}=\partial_\nu B_\mu-\partial_\mu B_\nu+e[B_\mu, B_\nu].$$

The coupling constants g and e are not necessarily related, but we shall assume that they are of a similar order of magnitude (and both small).

At zero temperature the O(N) symmetry here is spontaneously broken to O($N-1$), with ϕ acquiring a vacuum expectation of order η. In the tree approximation,

$$\langle\phi\rangle^2=\eta^2 \tag{2}$$

so that the manifold of degenerate vacua is an $(N-1)$ sphere S^{N-1}.

Let us recall the more general situation. In a model with symmetry group G, the vacuum expectation value $\langle\phi\rangle$ will be restricted to lie on some orbit of G. If H is the isotropy subgroup of G at one point $\langle\phi\rangle$, i.e. the subgroup of transformations leaving $\langle\phi\rangle$ unaltered, then the orbit may be identified with the coset space $M=G/H$. Physically H is the subgroup of unbroken symmetries, and M is the manifold of degenerate vacua. As we shall see, the topological properties of M (specifically its homotopy groups) largely determine the geometry of possible domain structures.

At a finite temperature T the expectation value of ϕ in a thermal equilibrium state must be found by minimizing the free energy, or equivalently the temperature-dependent effective potential. The leading temperature dependence at high T and small coupling constant comes from the one-loop diagrams (Weinberg 1974, Dolan and Jackiw 1974, Bernard 1974). Including these terms, we have

$$V(\phi)=\tfrac{1}{8}g^2(\phi^2-\eta^2)^2+\tfrac{1}{48}[(N+2)g^2+6(N-1)e^2]T^2\phi^2, \tag{3}$$

as in the Landau–Ginsberg theory of superconductivity. (See for example Schrieffer 1964.) The minimum occurs at $\phi=0$ and so the symmetry is unbroken for T larger than the transition temperature

$$T_c=\eta\left(\frac{N+2}{12}+\frac{N-1}{2}\frac{e^2}{g^2}\right)^{-1/2}. \tag{4}$$

This is the normal phase. Below T_c, we have an ordered phase: ϕ acquires a vacuum expectation value, which plays the role of the order parameter, and whose magnitude is determined by

$$\langle\phi\rangle^2=\eta^2[1-(T^2/T_c^2)]. \tag{5}$$

Thus the manifold of degenerate equilibrium states for all $T<T_c$ is an $(N-1)$ sphere, $M=O(N)/O(N-1)=S^{N-1}$.

In more complicated models there may be several transition temperatures, and as Weinberg (1974) has shown the symmetry may even increase as the temperature drops. However we shall not consider such cases, but assume a single transition temperature above which the symmetry is unbroken.

3. Formation of protodomains

Let us consider a 'hot big-bang' universe and examine what happens as it expands and cools through the transition temperature T_c. In unified models of weak and electromagnetic interactions T_c is of the order of the square root of the Fermi coupling constant, $G_F^{1/2}$, i.e. a few hundred GeV. Thus the transition occurs when the universe is aged between 10^{-10} and 10^{-12} seconds and far above nuclear densities. In other models, however, T_c might be considerably smaller and the transition would occur correspondingly later.

For T near T_c there will be large fluctuations in ϕ. Once T has fallen well below T_c, we may expect ϕ to have settled down with a non-zero expectation value corresponding to some point on M. No point is preferred over any other. As in an isotropic ferromagnet cooled below its Curie point the choice will be determined by whatever small fields happen to be present, arising from random fluctuations. Moreover this choice will be made independently in different regions of space, provided they are far enough apart. (What is far enough we shall discuss shortly.) Thus we can anticipate the formation of an initial domain structure with the expectation value of ϕ, the order parameter, varying from region to region in a more or less random way. Of course for energetic reasons a constant or slowly varying $\langle\phi\rangle$ is preferred and so much of this initially chaotic variation will quickly die away. The interesting question is whether any residue remains—in particular whether normal regions can be 'trapped' like flux tubes in a superconductor.

Because domains are most familar in the context of ferromagnetism it may be well to point out at the outset a crucial difference between that case and ours. The long-range dipole–dipole interaction between spins ensures that it is energetically favourable for a large ferromagnet to break up into domains with different magnetization directions.

However in the models we are considering, the source of the gauge field always involves the derivative of the order parameter $\langle\phi\rangle$, and vanishes for constant $\langle\phi\rangle$. Thus there is no long-range force between differently oriented domains, and the true ground state necessarily has a spatially uniform $\langle\phi\rangle$. The phenomena we are concerned with are non-equilibrium effects.

It is perhaps conceivable that in rather different theories domains might not be neutral with respect to some of the charges associated with gauge fields. However it is hard to see how domain formation in such cases could be energetically allowed. It is important to realize that even though the gauge field acquires a mass (Kislinger and Morley 1975) the effective interaction between non-neutral domains is still long-range. Physically, the situation is the same as for photons in a plasma. Because of the plasmon mass, charge *fluctuations* are shielded, but an unbalanced *net* charge will nevertheless yield a long-range Coulomb force. Net charge separation is energetically prohibitive.

Another possibility is that the universe as a whole may be non-neutral, if it is open and expanding. But then its mean charge would define a particular direction in the group space so that the direction of symmetry breaking would be fixed by the initial conditions rather than by random choice, and would be everywhere the same. Thus no domain structure would arise. However, a small degree of non-neutrality would be compatible with domain structure. Initially at least, we shall assume the overall neutrality of the universe, but we shall return to this question later, in § 5.

Let us now consider the initial scale of the 'protodomains' formed as the universe cools below T_c. At a temperature $T < T_c$, the difference in free-energy density between states with $\langle\phi\rangle$ equal to zero and to its equilibrium value (i.e. between 'normal' and 'ordered' phases) is easily found from (3)–(5) to be

$$\Delta f = \tfrac{1}{8}g^2(\langle\phi\rangle^2)^2. \tag{6}$$

It rises rapidly towards its zero-temperature value

$$\Delta f_0 = \tfrac{1}{8}g^2\eta^4$$

as T falls below T_c. The correlation length ξ which determines the scale of fluctuations in ϕ is the inverse of the scalar-meson mass (the 'Higgs scalar'), given approximately by

$$\xi^{-1} = m_S = g|\langle\phi\rangle|. \tag{7}$$

It is of course infinite at the phase transition. As in a superconductor there is a second correlation length, the penetration depth λ given by

$$\lambda^{-1} = m_V = e|\langle\phi\rangle|$$

which is of importance in discussing 'angular' oscillations, i.e. oscillations in the orientation of ϕ. However 'radial' oscillations, in the magnitude of ϕ, are controlled by ξ.

If T is only a little less than T_c, a fluctuation back to $\phi = 0$ remains quite probable. The free energy associated with such a fluctuation with scale ξ is, ignoring factors of order unity,

$$(2\xi)^3\Delta f \simeq |\langle\phi\rangle|/g.$$

This fluctuation will have high probability so long as the free energy required is substantially less than the thermal energy T. (We use units in which Boltzmann's constant, as well as $\hbar$ and c, is equal to unity.) The two are equal when

$$\frac{1}{T^2}=\frac{1}{T_c^2}+\frac{g^2}{\eta^2} \tag{8}$$

at which time the correlation length ξ is given by

$$\xi^{-1}\simeq g^2 T. \tag{9}$$

For weak coupling this is not too different from $g^2 T_c$, or roughly $g^2\eta$.

Thereafter fluctuations back to $\phi=0$ rapidly become less likely, so that the distinction between normal and ordered phases is well established as is that between ordered phases corresponding to well separated points on M (i.e. very different orientations of $\langle\phi\rangle$). The correlation length at this time thus determines the initial scale of the protodomains. Beyond this point it continues to fall, but the fluctuations are no longer large enough to disturb the established long-range order. In fact the scale of the domains over which $\langle\phi\rangle$ is nearly constant will tend to grow with time because of the $(\nabla\phi)^2$ term in the energy.

4. Topology of domain structures

Our next task is to examine the possible geometric configurations of the initial domain structure. It will be convenient to discuss a somewhat idealized model. Let us suppose that space is divided into cells of variable shape and size, with a mean dimension of the order of the correlation length obtained above, and that within each cell $\langle\phi\rangle$ is initially nearly constant, but with random variation from one cell to another. Although in reality $\langle\phi\rangle$ may be expected to vary continuously from point to point, this model seems to be a reasonable idealization of the kind of structure we might expect.

The interesting question is what happens on the boundaries where two or more cells meet. Consider two cells characterized by $\langle\phi\rangle=\phi_1$ and $\langle\phi\rangle=\phi_2$ respectively, with a common boundary surface. If there is a continuous path in the manifold M joining the points ϕ_1 and ϕ_2, then because of the $(\nabla\phi)^2$ term in the energy it will be energetically favourable for the width of the region over which $\langle\phi\rangle$ varies from ϕ_1 to ϕ_2 to expand, leaving a smoothly varying $\langle\phi\rangle$ in place of a sharp boundary. However, if M is disconnected, and ϕ_1 and ϕ_2 belong to different connected components, then no smooth transition is possible. One can pass from one ordered phase to the other only by going through a normal region. Hence a wall of normal phase will be formed between the two cells.

The simplest example of a model with disconnected M is the case $N=1$ of the model described in § 2. In that case there is only a discrete reflection symmetry, and no gauge fields. The manifold M is the 0-sphere S^0 comprising two points corresponding to the two possible values of $\langle\phi\rangle$, for example at $T=0$, $\langle\phi\rangle=\pm\eta$. In this case there will be two distinct, though internally indistinguishable, ordered phases. Any large enough volume of space will be divided more or less equally between the two with thin walls of normal phase, where $\langle\phi\rangle$ passes through zero, separating them. The width of the domain wall is determined by a balance between Δf and the $(\nabla\phi)^2$ term in the energy, and is in fact the correlation length ξ. (A solution with a plane boundary separating two domains is effectively a 'soliton'. See for example, Scott *et al* (1973).)

The subsequent development of this complex geometric structure is governed largely by the surface tension σ of the wall which is of order $\xi\Delta f$ or, by (6) and (7),

$$\sigma \simeq \xi\Delta f \simeq g|\langle\phi\rangle|^3. \tag{10}$$

(We again ignore factors of order unity.) Because of surface tension, the area of wall will tend to decrease with time. Thus in any fixed finite volume one of the two ordered phases will eventually tend to predominate. The rate at which this happens and the scale of the structure after a given time we shall discuss later.

Domain walls can form only if M is disconnected, i.e. if as in the $N = 1$ model there is a discrete broken symmetry. There may of course be continuous symmetries that are broken too. If M has two connected components, the structure produced is a kind of emulsion of two (similar) ordered phases. The walls can never terminate and must either form closed surfaces or extend to infinity. If the number of connected components is three or more we have a more complicated structure, an emulsion of several phases.

Let us now consider what happens when three cells meet along an edge. We suppose that the corresponding expectation values ϕ_1, ϕ_2, ϕ_3 belong to the same connected component of M, so that no pair of cells is separated by a domain wall. Then as we go around the edge we will find the expectation value $\langle\phi\rangle$ following a smooth path in M from ϕ_1 to ϕ_2, to ϕ_3, and back to ϕ_1. Now if this closed path can be continuously deformed in M to a point, i.e. if it is null-homotopic, then the ordered phase can be extended in a continuous manner into the region of the edge. All three domains can then fuse, leaving only a smoothly varying $\langle\phi\rangle$. But if it cannot be so deformed then a region of normal phase will be trapped along the edge, exactly like a vortex filament in a superfluid (see for example Putterman 1974) or a flux tube in a type II superconductor (see Schrieffer 1964). The Nielsen–Olesen (1973) string is an example of a solution of this kind, essentially for the model of § 2 with $N = 2$, in which case the manifold M is a circle, S^1. Note that although the gauge field does not seem to play an important topological role, it is vital in making the energy per unit length of the string finite.

We note that the strings have a tension whose order of magnitude is

$$\mu \simeq \xi^2\Delta f \simeq |\langle\phi\rangle|^2. \tag{11}$$

What is relevant in deciding whether strings can exist is the first homotopy group $\pi_1(M)$, comprising the homotopy classes of maps from S^1 into M, i.e. of closed paths in M (see Hu 1959). If M is simply connected, $\pi_1(M)$ is trivial and strings are impossible. But whenever $\pi_1(M)$ is nontrivial they will occur. Not all edges will of course form strings but we can estimate how many do so if we make two simple assumptions: that the initial choices of $\langle\phi\rangle$ in different cells are uncorrelated, and that where two cells meet and fuse $\langle\phi\rangle$ varies along the shortest path in M between the two original values ϕ_1 and ϕ_2. Then we find for the case where M is a circle that one in four of the original edges will form into strings.

The number of independent generators of $\pi_1(M)$ influences the type of structure that the strings can form. If it is one, as when M is a circle, there is only one basic type of string. In principle n strings might coalesce to form a multiple string around which the phase of $\langle\phi\rangle$ varies by $2n\pi$, though whether this is energetically favoured depends on the details of the model (essentially on the ratio e/g, as in the distinction between type I and type II superconductors—see Schrieffer 1964). If $\pi_1(M)$ has more than one generator, then several types of strings can exist. We may then expect to find vertices where three strings join, and space will be filled with a three-dimensional network.

We can go on to discuss what happens at a vertex where four cells meet with order parameters ϕ_1, ϕ_2, ϕ_3, ϕ_4. These points may be regarded as the vertices of a (curvilinear) tetrahedron in M, with edges representing the paths along which $\langle\phi\rangle$ varies across each bounding surface. We assume that each closed circuit, such as ϕ_1–ϕ_2–ϕ_3–ϕ_1, is null-homotopic, so that the faces of the tetrahedron can be filled in without leaving M. If the closed surface so formed can be shrunk to a point in M, then $\langle\phi\rangle$ can be extended continuously into the region around the vertex. But if not, then a normal region will be trapped, as in the monopole solution of 't Hooft (1974). For monopoles to exist we require that the second homotopy group $\pi_2(M)$ be nontrivial, as it is for the $N=3$ model. (Compare Krive and Chudnovskiĭ (1975).)

Such objects if they do exist will be localized structures with a scale typical of elementary particles, and regarded as such. They will not be significant on a cosmic scale, and we shall not discuss them further. (The considerations of this paper might perhaps be relevant in estimating the probable density of monopoles. If M is a sphere S^2 then under the same assumptions as before one in eight of the initial vertices should form a monopole. However to obtain an estimate of present density would require a careful study of annihilation mechanisms, which are its principal determinant.)

We have seen that to trap a k-dimensional normal region within the ordered phase requires, for $k=0$ or 1, that the homotopy group $\pi_{2-k}(M)$ be nontrivial, i.e. that there exist non-contractible maps of S^{2-k} into M. This criterion applies also for $k=2$, although $\pi_0(M)$ is not strictly speaking a homotopy group: it has no group structure but represents merely the number of connected components of M (see for example Hilton and Wylie 1960, p 276).

It is perhaps amusing to note that if $\pi_0(M)$, $\pi_1(M)$ and $\pi_2(M)$ are all trivial, we can still obtain a topological characteristic of our particular universe (if it is finite and homeomorphic to S^3) by considering $\pi_3(M)$: namely, the homotopy class of the function $\boldsymbol{x}\to\langle\phi(\boldsymbol{x})\rangle$, which is with high probability an invariant once we are well past the transition temperature. However it is hard to envisage any possible physical relevance for it.

If we are interested in cosmic structures, we have then two candidates: domain walls when $\pi_0(M)$ is nontrivial and strings when $\pi_1(M)$ is. The next question we must ask is how these structures would evolve, once formed.

5. Evolution of domain walls and strings

The motion of a domain wall is controlled primarily by its surface tension (equation (10)). The corresponding surface energy is also the inertial mass of the wall. In other words the wall is 'massless' in the sense usually implied in talking of massless strings: its action integral is simply proportional to the three-volume it sweeps out in space–time. As a result of the equality between areal mass density and surface tension the velocity of waves of short wavelength along the wall is the velocity of light.

The surface is in local equilibrium when its mean curvature is zero, i.e. when the two principal radii of curvature are equal and opposite, $\pm R$ say. (The simplest nontrivial example is the catenoid $(x^2+y^2)^{1/2}=\cosh z$.) Any deviation from local equilibrium will lead to waves spreading out rapidly from their point of origin. Among the possible damping mechanisms the most obvious is the interaction of the walls with surrounding matter. A wall moving with velocity v through matter of density ρ will experience a retarding force per unit area

$$\alpha v = a\rho\bar{u}v \tag{12}$$

where a is a numerical constant depending on the reflectivity of the wall and $\bar{u}$ a suitably defined mean velocity of the matter particles. The characteristic time for damping out short-wavelength oscillations is then

$$t_{\rm d} = 2\sigma/\alpha. \tag{13}$$

Clearly it is important to estimate the degree of transparency of the walls. If it were possible for the domain structure to survive until near the present, this question would also have a more direct relevance to the possible observability of domains. It was from this point of view that it was tackled by Everett (1974). However, domain walls surviving so long can be ruled out by the argument of Zel'dovich *et al* (1974).

In the very early stages the wall thickness ξ is comparable with the (relativistic) thermal wavelength T^{-1}. In fact at the time of formation of protodomains, identified in § 3, we have $\xi T \simeq g^{-2} > 1$. Thereafter, however, ξ falls rapidly to its limiting value of order η^{-1}, and T decreases steadily though more slowly. Once it has fallen to a small fraction of $T_{\rm c}$—say $T \simeq 1$ GeV which occurs at $t \simeq 10^{-6}$ s—we can legitimately treat the domain walls as very thin, and replace them by sharp discontinuities across which the fields are related by boundary conditions of the usual type.

Most of the significant evolution occurs during the radiation era, so we are mainly interested in the transparency of the walls to *light.* Let us consider an electromagnetic wave striking a domain wall. We suppose, as is now conventional, that the photon belongs to a multiplet of gauge bosons whose other members are very massive. Two very different situations can arise, depending on the precise model of symmetry breaking. If a *different* component of the field is massless on the far side of the wall then, as Everett (1974) showed, the probability of reflection is large. In that case the constant a in equation (12) will be of order unity and the motion of the walls will be heavily damped by radiation pressure. On the other hand if the electromagnetic field corresponds to the *same* component on both sides, as in the models considered by Zel'dovich *et al* (1974), it is clear that the wall will be almost transparent and so little affected by radiation pressure. It is not hard to write down models of both types.

In the latter case, one must examine other damping mechanisms—for example the absorption and emission of radiation—but none seems likely to be very effective. Thus the walls will continue to move with relativistic speeds, and the subsequent evolution will be more or less along the lines described by Zel'dovich *et al.*

For the moment, however, let us assume that the walls are highly reflecting, so that a is of order unity, and

$$\alpha \simeq \rho \simeq T^4.$$

Once T is sufficiently below $T_{\rm c}$ for σ to have reached its asymptotic value, we then have

$$t_{\rm d} \simeq g\eta^3/T^4 \tag{14}$$

which grows like t^2.

Initially $t_{\rm d}$ is roughly of order g/η, and therefore small compared to the length scale $1/g^2\eta$ of the protodomains. Consequently the domain walls will evolve rather rapidly towards local equilibrium. Moreover isolated pockets of one phase surrounded by another will quickly disappear. In the case where M has two components, we are left with an emulsion of two phases each occupying a single connected region (cf Broadbent and Hammersley 1957) and separated by a single, highly convoluted wall whose mean curvature is everywhere near zero. (The structure is in some ways rather similar to the matter–antimatter emulsion considered by Omnès (1971).)

Such a configuration is stable against short-wavelength perturbations, but it cannot be completely stable. It is well known that the only stable infinite-area solution of Plateau's problem (to find minimum-area surfaces of given perimeter) is the plane (see Thompson 1942). In general, waves shorter than a typical radius of curvature will be damped but disturbances of longer wavelength will grow exponentially in time. The equation of motion for small deviations from equilibrium is

$$\frac{\partial^2 \zeta}{\partial t^2}+\frac{2}{t_{\rm d}}\frac{\partial \zeta}{\partial t}-\Delta\zeta-\frac{2}{R^2}\zeta=0, \qquad (15)$$

where ζ is the normal displacement and Δ the two-dimensional Laplace–Beltrami operator. Thus when $R \gg t_{\rm d}$, the characteristic time for growth of long-wavelength perturbations is $R^2/t_{\rm d}$.

Eventually the surface will jump to a new equilibrium configuration of smaller area which in turn will be unstable to yet longer wavelengths. The length scale l of the emulsion—the volume to wall-area ratio for any large enough volume—will therefore grow with time. Roughly, we may identify the characteristic time for growth of l with that for the long-wavelength perturbations, so that

$$\frac{1}{l}\frac{{\rm d}l}{{\rm d}t}\simeq\frac{t_{\rm d}}{l^2}.$$

Hence l grows like the square root of the elapsed time, as one might expect from the diffusion-like character of the process. Since this is much faster than the slow growth of $t_{\rm d}$, the approximation $l \gg t_{\rm d}$ remains valid. However over a long period of time the growth of $t_{\rm d} \propto t^2$ leads to $l \propto t^{3/2}$. Clearly therefore $t_{\rm d}$ will overtake l. It is easy to see that this happens when both are of the same order as the age of the universe; in fact when

$$l \simeq t_{\rm d} \simeq t = t_1 \simeq 1/G\sigma, \qquad (16)$$

where G is the Newtonian gravitational constant (which enters via the relation $tT^2 \sim 1/G^{1/2}$). It is interesting to note that the very different mechanism envisaged by Zel'dovich *et al* yields a similar result for the time at which the scale of the emulsion becomes comparable with the radius of the universe. With the parameters chosen earlier we have $t_1 \simeq 10^7$ s, but because of its strong η dependence, proportional to η^{-3}, t_1 is very model-dependent.

The degree of inhomogeneity introduced into the universe by the gravitational effect of the domain walls (as shown by Zel'dovich *et al* 1974) is of the order $G\sigma l \simeq l/t_1$. It is clearly unacceptably large unless either there are many domain walls remaining ($l/t_1 < 10^{-3}$) or none at all. The first alternative is implausible. Domain walls with the required spacing can be ruled out on observational grounds, at any rate if they are substantially reflecting. The second is impossible if we start from completely random initial choices of the order parameter. The typical scale of the emulsion can never much exceed the distance over which causal corrections are possible, i.e. the present radius of the universe.

There seems to be only one way of accommodating models that generate domain walls, namely, as Zel'dovich *et al* suggest, to allow a small initial bias. It would require only a minute non-neutrality in the initial state of the universe with respect to some 'charge'—related perhaps to the preference for matter over antimatter—to ensure a slight but consistent excess in the number of protodomains of one type. This excess would tend to become more pronounced with time and eventually lead to complete dominance of the preferred phase. It would not be hard to realize this case by suitable choice of model parameters. But such initial asymmetry is unaesthetic, and not at all in the spirit of spontaneous symmetry breaking.

In all probability therefore we should rule out models with discrete symmetry breaking, which lead to domain walls—an interesting example of a restriction on elementary particle theories deriving from cosmology.

There is no such objection to cosmic strings. A single domain wall in the universe (with the parameters chosen here) would have a mass far in excess of the accepted upper limit for the mass of the universe. By contrast, the mass of a string might be of order 1 mg m^{-1}, so that the total mass of a string of length equal to the current radius of the universe would be less than that of a minor planet.

The evolution of the network of cosmic strings would in many respects be similar to that of walls, although of course there is only one local equilibrium configuration for a string—a straight line—in contrast to the infinite variety of surfaces of zero mean curvature.

Consider a section of string with radius of curvature R. Because of its 'masslessness', it experiences an initial acceleration $1/R$. However there is a damping force arising as before from the interaction with other matter. If the effective (linear) cross section of the string is simply its width ξ then the ratio of retarding force to mass is the same as that found earlier for a wall. Thus we may assume that the damping time t_d is similar. It follows that the string will acquire a limiting velocity t_d/R and the kink will be straightened out in a time of order R^2/t_d.

As the strings move they will sometimes cross. When they do, they may be apt to change partners, leaving a pair of sharply kinked strings which straighten out in time, and so on. It is not entirely clear how far the existence of the topological invariants of strings studied by Khan (1975) might inhibit this exchange process. Assuming however that the exchange can occur, it is clear that occasionally a small closed loop may form which then shrinks to a point and vanishes. Thus the overall length of string will decrease, or equivalently the length scale l will increase, on a time scale l^2/t_d, exactly as for domain walls. With most choices of parameters, this means that the scale now is comparable with the radius of the universe. We can expect only one string within the visible universe, so that looking for cosmic strings directly would be pointless.

If the manifold of equilibrium states is more complex, so that $\pi_1(M)$ has more than one generator, then as we have seen a network of strings can be formed. Although superficially such a network might seem to have more stability its evolution would probably not in fact be much slower. It seems most unlikely that any structure could have survived until the present on a small enough scale to be directly visible.

Nonetheless the existence of such a network of cosmic strings may have had profound effects on the earlier history of the universe, at a stage where the number of strings was large. For example, strings can produce significantly local inhomogeneities by their interaction with matter. The details of this mechanism and its effectiveness require further study.

6. Conclusions and discussion

It may be well to begin by recalling our basic assumptions. We have assumed that the universe is correctly described by a spontaneously broken gauge theory exhibiting a phase transition at a critical temperature T_c, above which the symmetry is restored. We have taken for granted the hot big-bang model of the universe, with no maximum temperature, whence it follows that in its very early history the universe was above T_c, in the 'normal' phase. Finally we have generally assumed the overall neutrality of the universe with respect not only to electric charge but also to all the other charges associated with gauge fields.

On this basis we showed that a domain structure can be expected to arise. The topological character of this structure depends on the homotopy groups $\pi_k(M)$ of the manifold M of degenerate vacua. Domain walls can form if $\pi_0(M)$ is nontrivial, i.e. if M is non-connected. If it has n connected components we find an n-phase emulsion. The formation of cosmic strings requires that $\pi_1(M)$ be nontrivial, i.e. that M is not formed of simply connected components. Finally, 'monopoles' can form if $\pi_2(M)$ is nontrivial.

The later evolution of domain walls is governed by their surface tension and their interaction with matter. Different types of domain walls can occur with very different transparency, but in all cases the overall scale of the structure will grow with time. In general we may expect it now to be comparable with the radius of the universe. Domain walls on anything like this scale can be ruled out (Zel'dovich *et al* 1974) because their gravitational effect would lead to unacceptable anisotropy in the black-body background radiation. The only way of accommodating theories with spontaneously broken discrete symmetries (and hence domain walls) is to relax the requirement of complete neutrality of the universe, so that one of the ordered phases is slightly preferred and eventually comes to occupy all of space. However this is not a very attractive solution, and it may be better to regard this as an argument against such theories.

Networks of strings will evolve in a similar way under the combined effects of tension and interaction with matter. Once again, the scale of the structure will grow with time, probably at a similar rate. One cannot expect to find significant numbers of cosmic strings in the visible universe now, but their presence may have had an important effect on the earlier evolution of the universe.

References

Bernard C 1974 *Phys. Rev.* D **9** 3312
Broadbent S R and Hammersley J M 1957 *Proc. Camb. Phil. Soc.* **53** 629
Coleman S 1976 *Int. School on Subnuclear Physics 'Ettore Majorana' Erice 1975* to be published
Dashen R F, Ma S K and Rajaraman R 1975 *Phys. Rev.* **11** 1499
Dolan L and Jackiw R 1974 *Phys. Rev.* D **9** 3320
Everett A E 1974 *Phys. Rev.* D **10** 3161
Harrington B J and Yildiz A 1975 *Phys. Rev.* D **11** 779, 1705
Hilton P J and Wylie S 1960 *Homology Theory* (Cambridge: Cambridge University Press) p 276
't Hooft G 1974 *Nucl. Phys.* B **79** 276
Hu S-T 1959 *Homotopy Theory* (New York: Academic Press)
Jacobs L 1974 *Phys. Rev.* D **10** 3956
Khan I 1975 *Trieste preprint* IC/75/148
Kirzhnits D A 1972 *Zh. Eksp. Teor. Fiz. Pis'ma Red.* **15** 745 (1972 *Sov. Phys.-JETP Lett.* **165** 529)
Kirzhnits D A and Linde A D 1972 *Phys. Lett.* **42B** 471
—— 1974 *Zh. Eksp. Teor. Fiz.* **67** 1263 (1974 *Sov. Phys.-JETP* **40** 628)

Kislinger M B and Morley P D 1975 *Chicago preprint EFI* 75/8
Kobzarev I Yu, Okun L B and Zel'dovich Ya B 1974 *Phys. Lett.* **50B** 340
Krive I V and Chudnovskiĭ E M 1975 *Zh. Eksp. Teor. Fiz. Pis'ma Red.* **21** 271 (1975 *Sov. Phys.–JETP Lett.* **21** 124)
Linde A D 1975 *P N Lebedev Physical Institute preprint no.* 123
Nielsen H B and Olesen P 1973 *Nucl. Phys.* B **61** 45
Omnès R 1971 *Astron. Astrophys.* **10** 228
Putterman S J 1974 *Superfluid Hydrodynamics* (Amsterdam: North Holland)
Schrieffer J R 1964 *Theory of Superconductivity* (New York: Benjamin)
Scott A, Chu, F and McLaughlin D 1973 *Proc. IEEE* **61** 1443
Thompson d'A W 1942 *On Growth and Form* (Cambridge: Cambridge University Press) 2nd edn
Weinberg S 1974 *Phys. Rev.* D **9** 3357
Zel'dovich Ya B, Kobzarev I Yu and Okun L B 1974 *Zh. Eksp. Teor. Fiz.* **67** 3 (1975 *Sov. Phys.–JETP* **40** 1)