Monopole annihilation and causality

Allen E. Everett, Tanmay Vachaspati, and Alexander Vilenkin
Department of Physics and Astronomy, Tufts University, Medford, Massachusetts 02155
(Received 29 October 1984)

It is shown that, contrary to a recent claim, causality alone does not impose any interesting constraints on the rate of monopole annihilation. We study a two-dimensional analog of the models in which monopoles get connected by strings and use a Monte Carlo simulation of the phase transitions to determine the length distribution of strings. The result is that long strings are exponentially suppressed. We then argue that for any initial length distribution of strings in two or three dimensions, an exponential distribution is eventually established by intercommutings or by spontaneous breaking of strings.

I. INTRODUCTION

Recently there has been discussion[1–3] as to whether general arguments based only on considerations of causality can be used to limit the rate of annihilation of particles and antiparticles carrying a conserved charge Q. Although there are interesting questions of physical principles involved, the main motivation for considering the problem has been its possible applicability to the case where the particles in question are very heavy magnetic monopoles that may have been produced in the early universe.[4] In particular, the arguments are relevant to models in which the U(1) symmetry, whose appearance as a factor in the unbroken symmetry group of the Hamiltonian gives rise to monopoles, is broken at a latter stage of symmetry breaking. In such models, the monopoles and antimonopoles (M's and $\overline{M}$'s) become connected by strings, i.e., tubes of false vacuum carrying a magnetic flux, and the strings contract, pulling the $M\overline{M}$ pair together. An example of such a model is that of Langacker and Pi.[5] The lifetime τ of an $M\overline{M}$ pair connected by a string is determined by the time it takes to dissipate the string energy.[6] It depends on particular dissipation mechanisms operating and can be different for different models. In any case, causality requires that $\tau > l$, where l is the length of the string. The efficiency of this monopole annihilation mechanism can thus be related to the length distribution of strings.

If all monopoles and antimonopoles are connected by the shortest possible strings of length $\sim d$, where d is the typical monopole separation, then causality requires only that the lifetime of the system is $\tau > d$. Obviously, this is not a very interesting constraint. Weinberg[1,2] has argued that there exist much more stringent constraints on the rate of monopole annihilation. He has conjectured that magnetic-charge-density fluctuations cannot be erased on scales greater than the causal horizon. (We shall call this Assumption A.) Then it follows that the monopole density cannot decrease faster than a power law, $n_{\min} \propto t^{-5/2}$ [in a universe expanding like $a(t) \propto t^{1/2}$]. One can also argue that since the directions of the Higgs field are not correlated on scales greater than the horizon, there should be at least of order one monopole per horizon volume at any time. (We shall call this Assumption B.) Indeed, the magnetic charge can be written as a surface integral,

$$Q = \frac{1}{4\pi} \int \mathbf{B} \cdot d\mathbf{S} , \tag{1}$$

where $\mathbf{B}$ can be expressed in terms of the Higgs field. The disappearance of Q on scales greater than the horizon implies that the Higgs field has developed certain correlations on such scales. Assumption B gives a somewhat weaker constraint on the monopole density: $n_{\min} \sim t^{-3}$.

In this paper it will be shown that both Assumptions A and B are in fact wrong. We shall prove this by giving counterexamples in which these assumptions fail. In the next section we shall introduce a two-dimensional model that has all essential features of the system of monopoles and strings. The length distribution of strings will be found by a Monte Carlo simulation of the phase transitions in this model (Sec. III). The result is that strings much longer than the typical monopole separation are exponentially suppressed, and thus causality gives no interesting constraints on the rate of monopole annihilation. In Sec. IV it will be argued that even if one prepares the system with another length distributon of strings, an exponential distribution will eventually be established by various physical processes in the system.

Before we get to the proof that Assumptions A and B are false, in the case of monopoles connected by strings, it is worth explaining why they are not necessarily true in general. Any annihilation mechanism is consistent with causality as long as it does not require superluminal velocities. In general this implies neither Assumption A nor B; their validity in a particular case depends on the details of the dynamics. If, for example, annihilation is due to random diffusive motion of particles and antiparticles, then the charge fluctuations remain on all scales $> (Dt)^{1/2}$, where D is the diffusion constant and can vanish even more slowly than demanded by Assumption A. In contrast, as noted in Ref. 3, an initial charge density in an ohmic medium vanishes exponentially on all scales. Physically, this is because the neutralization of the initial charge density is accomplished by small displacements of free conduction charges in the medium. The same conclusion applies to magnetic charge fluctuations. The dynamics of the monopole motion in this case is governed by a long-range magnetic field **B** obeying

$$\nabla\cdot\mathbf{B}=4\pi\rho\ , \tag{2}$$

where ρ is the magnetic charge density. Lee and Weinberg[2] have pointed out that in the case of electric charge, the impossibility of the spontaneous creation of isolated charges makes it difficult to see how one can establish fluctuations in a conductor on a scale larger than the horizon, and hence it is not clear that this constitutes a valid counterexample to Assumption A. However, this objection does not hold in the case of magnetic monopoles produced at a phase transition in the early universe. There, the charge fluctuations do appear spontaneously, and the topological properties of the scalar and vector fields automatically assure that Eqs. (1) and (2) are satisfied. In particular, Eq. (1) holds for all surfaces, including those larger than the horizon. The long-range magnetic field **B** is established instantaneously and does not require any superluminal propagation.

We now turn to Assumption B, which has a different philosophy behind it. The assumption is basically that no correlations can be established between the values of the Higgs field on scales greater than the horizon. In our case, the quantity of interest is

$$I=\int_S \mathbf{B}\cdot d\mathbf{S}\ . \tag{3}$$

The disappearance of this integral for all or most of the surfaces bigger than the horizon certainly implies some kind of correlations in the Higgs field on such scales. To see how such correlations may arise, let us divide the volume enclosed by the surface S into a large number N of regions much smaller than the horizon. Then

$$I=\sum_{k=1}^{N} I_k\ , \tag{4}$$

where I_k is the value of integral (3) for the kth region. Local physics imposes correlations on the Higgs field on scales smaller than the horizon, and it is conceivable that all I_k rapidly decrease with time. It is clear that the large-scale integral I will exhibit the same behavior. This is simply due to the fact the I is an additive quantity and its vanishing on small scales implies that it vanishes on any scale. In contrast, the two-point function does not have this additive property and will show no correlations on large scales.

II. THE MODEL

For ease of numerical simulation, we have replaced the three-dimensional system of monopoles connected by strings with a two-dimensional system having most of the same properties. To do this we consider a system possessing a U(1) symmetry that undergoes spontaneous symmetry breaking to a discrete Z_2 symmetry. In this process vacuum strings are formed.[7] When the Z_2 symmetry is then spontaneously broken at a second stage of symmetry breaking, the strings become connected by domain walls that contract, pulling the strings together.[8] When projected on two dimensions, the strings play the role of monopoles, and the domain walls play the role of strings, so that in two space dimensions this model describes a system of two-dimensional monopoles connected by strings; we will, however, continue to speak in the three-dimensional language, referring to strings and walls rather than monopoles and strings.

We can exhibit an explicit model with these properties. Consider a model with two complex scalar fields ϕ_1 and ϕ_2 and take the potential term in the Lagrangian to be given by

$$V(\phi_i)=g_1(\,|\phi_1{}^2\,|-\eta^2)^2+g_2\mathrm{Re}(\phi_2{}^2\phi_1^*)+g_3(\phi_2{}^2\phi_2^{*2})/2\ , \tag{5}$$

where the g_i are coupling constants and, for simplicity, an explicit mass term for ϕ_2 has been omitted. V is invariant under the U(1) transformations $\phi_i\to\exp(iq_i\theta)\phi_i$, where the U(1) charges q_i are 1 and $\frac{1}{2}$ for ϕ_1 and ϕ_2, respectively. Assuming $g_1\gg g_2/\eta, g_3$, the first term in V can be treated separately to a good approximation and implies that ϕ_1 will develop a vacuum expectation value

$$\langle\phi_1\rangle=\eta e^{i\alpha}\ , \tag{6}$$

where α is an arbitrary phase, at a temperature of order η. This breaks the continuous U(1) symmetry spontaneously, but leaves unbroken a symmetry under the discrete transformation

$$\phi_i\to\exp(2\pi i q_i)\phi_i \tag{7}$$

under which $\phi_1\to\phi_1$ and $\phi_2\to-\phi_2$. Minimizing the potential with respect to ϕ_2 at fixed ϕ_1, one finds that the minimum of the potential occurs at

$$\langle\phi_2\rangle^2=-(g_2/g_3)\langle\phi_1\rangle \tag{8}$$

and hence ϕ_2 will develop a vacuum expectation value in a second phase transition that will break the discrete symmetry of Eq. (7). Letting θ_2 be the phase of $\langle\phi_2\rangle$, one sees from Eqs. (6) and (8) that

$$\theta_2=\alpha/2\ \text{ or }\ \theta_2=\alpha/2+\pi\ , \tag{9}$$

where the ambiguity reflects the sign ambiguity in ϕ_2 from Eq. (8).

In the first phase transition, strings are formed. The value of α changes by $\pm 2\pi$ along a closed curve in space enclosing a string. (Strings around which α changes by $2n\pi$ are presumably unstable[9] at this stage and decay into n simple strings.) From Eq. (9) one sees that if θ_2 varied smoothly around a string, it would change by only π. Since $\langle\phi_2\rangle$ must be single valued, θ_2 must, however, change by an integral multiple of 2π around any closed curve, and hence around a curve enclosing a string θ_2 must change at some point from one to the other of the two solutions allowed by Eqs. (8) and (9); that is, there will be two nearby points on the curve, at one of which $\theta_2=\alpha/2$ and at the other $\theta_2=\alpha/2+\pi$, with the two points being separated by a domain wall. Thus each string formed in the first phase transition becomes attached to a domain wall; since the wall is a region of false vacuum, it will tend to contract to minimize its energy, so that the two strings connected to its edges will be pulled together. (There will also be closed domain walls formed, whose projections onto two dimensions are closed curves unattached to strings. These are not relevant for our discussion.)

Note that α may increase in the same direction or in opposite directions around the two strings joined by a domain wall. In the first case when the two strings are pulled together they will annihilate. In the second case they form a double string around which α changes by 4π and θ_2 by 2π. Such strings are topologically stable, i.e., have no domain walls attached to them, after the second phase transition. Thus the physical vacuum state of the system after the second phase transition will contain a number of stable strings. Hence two simple strings can annihilate into the physical vacuum whether they have opposite directions and annihilate entirely, or have the same direction and "annihilate" to form a stable double string. The projection of this model onto two dimensions, therefore, does not correspond to two-dimensional monopoles and antimonopoles carrying a conserved charge, but rather to monopoles carrying a charge that is conserved modulo 2. Thus in our model, I in Eq. (3) is only defined mod2, and the analog of Eq. (1) in our model is an equality mod2, namely,

$$N_w(C)=N_s(C)\ (\mathrm{mod}2)\ , \tag{3'}$$

where $N_s(C)$ is the total number of strings enclosed by a closed curve C, and $N_w(C)$ the total number of domain walls passing through C. Despite this difference the model appears to contain all of the essential features of models in which $M\overline{M}$ pairs become connected by strings.

It is perhaps worth pointing out that the U(1) symmetry, which is spontaneously broken in this model, is exact. This is to be contrasted with models involving a broken approximate global U(1) symmetry;[8] such models are relevant in axion theories.[10] Numerical simulations of the formation of strings and domain walls in such a model have been done by Vachaspati and Vilenkin and by Sikivie.[11,12] The distinction between such models is relevant in the present context since Lee and Weinberg[2] have argued that in models involving broken approximate U(1) symmetries Assumption A is not expected to be valid, and so such models cannot be used as counterexamples.

III. MONTE CARLO SIMULATION

The procedure used to simulate the formation of strings and walls in the model of Sec. II is very similar to that used to study the formation of strings in Ref. 11. We consider an $n\times n$ planar lattice, and assign the value of the phase of the Higgs field ϕ_1, i.e., the value of α, randomly at each point on the lattice. We make the simple approximation of allowing the phase of ϕ_1 to take on only one of the three possible values 0, $2\pi/3$, or $4\pi/3$. We take the lattice spacing to correspond to the correlation length ξ of the field ϕ_1 so that α varies smoothly over a lattice spacing. We mimic this smooth variation by taking the sense of rotation of α in going from one vertex of the lattice to an adjoining one to be that which involves the smaller magnitude of the change in angle; e.g., α is taken to increase in going from a vertex with $\alpha=0$ to one with $\alpha=2\pi/3$, but to decrease in going from a vertex with $\alpha=0$ to one with $\alpha=4\pi/3$. A string passes through a face of the lattice if α increases by 2π as one goes once completely around the face; examples of configurations that do and do not correspond to strings are shown in Fig. 1. From Eq. (9) the three possible values of α correspond to three possible pairs of values of θ_2, $(0,\pi)$, $(\pi/3,4\pi/3)$, and $(2\pi/3,5\pi/3)$ for $\alpha=0$, $2\pi/3$, and $4\pi/3$, respectively. Values of θ_2 for each vertex are determined consistently with the values of α by using Eq. (9) and determining at random whether or not to add π. If no domain wall separates two adjacent vertices, then from Eq. (9) and our rules for assigning α, the magnitude of the difference between the values of θ_2 at those vertices cannot exceed $\pi/3$. (As before, of course, we take the sense of rotation of θ_2 in going from one vertex to the next to be that which minimizes $|\Delta\theta_2|$; if $|\Delta\theta_2|=\pi$, the sense of rotation is not defined, but our result will not depend on the sense of rotation chosen in that case.) If the magnitude of the difference in the values of θ_2 assigned to two adjacent vertices exceeds $\pi/3$, then a domain wall must cut the line joining them. With this rule for determining the position of domain walls, it is easy to convince oneself that, if a face of the lattice contains a string, there will always be an odd number of domain walls passing through the edges of that face, consistent with the fact that a domain wall terminates on the string. For a lattice face containing a string, there are two possibilities in our simulation; there may be a single segment of domain wall entering it, or three segments, the latter case representing a wall entering the face through one edge and leaving through a second, in addition to the wall segment terminating on the string. These two possibilities are illustrated in Fig. 2, where we also show possible examples of assignments of values of θ_2 at the corners of the face that are consistent with the values of α illustrated in Fig. 1(a), and that result in each of the two possible types of wall configurations. Similarly, it is easy to see that for a face of the lattice containing no string our algorithm always gives an even number of wall segments, 0, 2, or 4, entering it. These possibilities are shown in Fig. 3, where for each of the possible types of configuration we show a possible set of values of θ_2, consistent with the values of α in the example of Fig. 1(b), which give rise to it. It follows from the preceding remarks that the lines representing the intersections of domain walls with the lattice plain either form closed

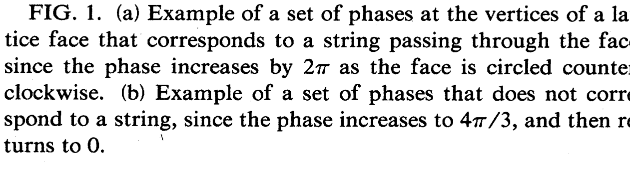

FIG. 1. (a) Example of a set of phases at the vertices of a lattice face that corresponds to a string passing through the face, since the phase increases by 2π as the face is circled counterclockwise. (b) Example of a set of phases that does not correspond to a string, since the phase increases to $4\pi/3$, and then returns to 0.

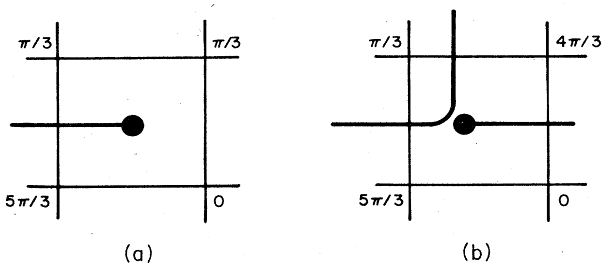

FIG. 2. (a) Example of a set of phases, obtainable from those in Fig. 1(a) by the prescription of Eq. (8), that corresponds to a wall indicated by the heavy line attached to the string. (b) Same as (a) but with the phases corresponding to a second wall entering and leaving the lattice face in addition to the wall attached to the string. A random choice is made of which wall segment is attached to the string.

curves or open curves ending on strings. (We identify opposite edges of the lattice with one another by imposing periodic boundary conditions, so that the possibility of curves leaving the lattice does not occur.)

Having determined the positions of the lines representing the intersections of domain walls with the plain of the lattice, the computer follows each such line to determine its length, and also whether it is closed or terminates on a string. In the case of a configuration like that in Fig. 2(b), the computer decides at random which of the three wall segments terminates on the string; similarly in configurations like those of Fig. 3(c), the computer decides at random which pairs of wall segments are connected with each other, subject to the condition that the walls do not cross.

The largest lattice that we used was 300×300 in size. In one typical example of a run with such a lattice, we found a total of 13 282 open wall segments; that is, 26 584 strings passed through the area of the lattice. This number is consistent with the fact that our algorithm can easily be shown to give a probability of $\frac{8}{27}$ that any particular lattice face contains a string; thus the average number of strings for a lattice of size 300 is $300^2\times(8/27)=26\,666$. There were also 5564 closed curves formed in the lattice by wall segments that did not terminate on strings; these are not directly relevant to the present discussion. We adopt units in which the lattice spacing $\xi=1$. The longest open wall segment has a length of 47, and there are only five segments with length $\geq$ 40. Let $n(L)$ be the density of open wall segments with length L, i.e., the number of such segments per unit area of the lattice. In Fig. 4 we show a plot of $\ln n(L)$ vs L for $1\leq L<40$. It will be seen that the plot is consistent with being linear, implying an exponential falloff of $n(L)$ with L of the form

$$n(L)=n(1)\exp[(L-1)\sigma]\ . \tag{10}$$

A fit to the curve of Fig. 4 yields $\sigma=0.193$, so that $n(L)$ decreases exponentially on a scale of a few correlation lengths ξ. Note that ξ is of the order of the typical separation between two nearest-neighbor strings, and thus our results indicate that, statistically, strings (and, presumably, monopoles in the three-dimensional case) almost always become attached to rather near neighbors.

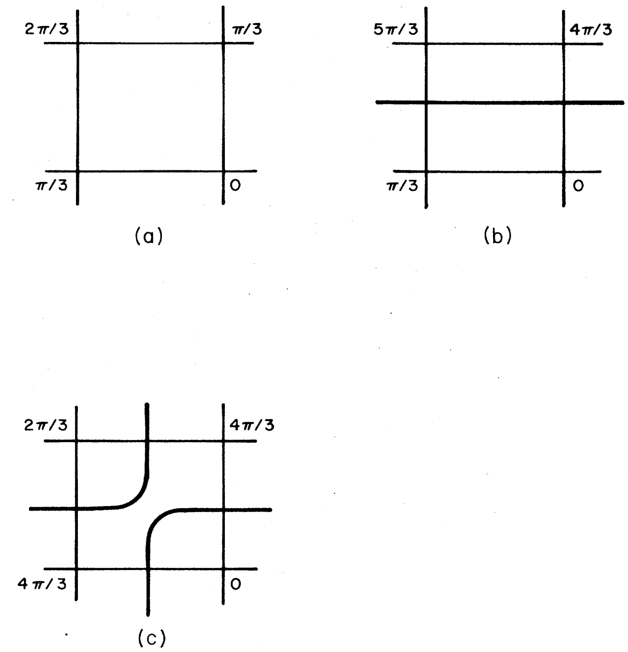

FIG. 3. (a) Example of a set of phases, obtainable from those of Fig. 1(b) by the prescription of Eq. (8), for which no wall enters or leaves the lattice face. (b) Same as (a) but with the phases corresponding to a single wall entering and leaving the lattice face. (c) Same as (a) and (b) but with two walls entering and leaving the lattice face. A random choice is made as to how the wall segments are connected.

The number of wall segments that are long compared to the typical distance between strings is exponentially small.

Note that, by imposing periodic boundary conditions so that the lattice becomes a closed space, we are forcing any wall segment that is attached to one string to terminate on a second string within the lattice. However, since $1/\sigma \ll 300$, the size of the lattice, and since even the

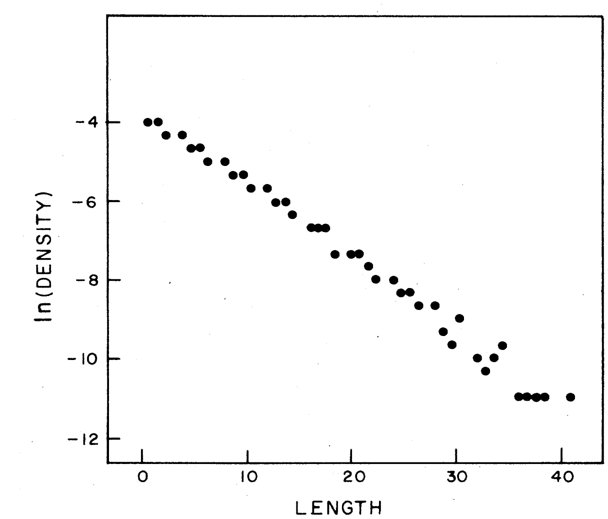

FIG. 4. A plot of $\ln n(L)$ vs L for one of the runs with a 300×300 lattice, where $n(L)$ is the density of open wall segments of length L.

lengths of the longest strings found are small compared to 300, the finite size of the lattice is not affecting our results.

To check the statistical significance of the results, four runs were made with a 300×300 lattice and different sets of random phases. The results of all four runs were qualitatively similar to the run shown in Fig. 3. Taking the average of the four runs yields the result $\sigma=0.196\pm 0.003$. As a further check on the effect of lattice size, runs were made with 100×100 and 200×200 lattices. Again the results were qualitatively similar to Fig. 3. The 100×100 and 200×200 lattices yielded values $\sigma=0.175$ and 0.183, and the longest strings found on the two lattices had lengths of 37 and 40, respectively.

In the simulation described above we used the same lattice spacing to simulate both phase transitions. This amounts to an assumption that the two correlation lengths correspond to the same co-moving scale. Such an assumption may be justified if both phase transitions are second order, in which case the correlation length is $\xi\sim T^{-1}$, where T is the transition temperature. If one or both transitions are first order, or if substantial annihilation occurs between the two phase transitions, the correlation lengths will be different. While this would certainly change the numerical details, i.e., the value of σ, there seems to be no reason why it should change the quantitative conclusion that the length distribution falls off exponentially. We made some effort to explore this directly in our numerical simulation by taking a larger correlation length for the second phase transition. However, with the lattice sizes available, the results were not very significant statistically, though certainly consistent with the conclusion that one was still seeing an exponential decrease in $n(L)$.

If Fig. 4 is not convincing enough, we have plotted in Fig. 5 the natural logarithm of the density of wall segment versus the natural logarithm of the segment length. It is evident from the graph that $n(L)$ decreases faster than a power law, in contradiction to Assumptions A and B. If one arbitarily fits the curve in Fig. 5 over its whole length to a straight line and translates the results into a fit to the density of the form $n(L)=aL^{-q}$, one obtains as a best fit $q=2.37$. This already gives a faster falloff than allowed by either Assumptions A or B applied to the case of two dimensions. Moreover, the curve in Fig. 5 is clearly not a straight line, and obviously still larger values of q would be obtained by fitting only a region of the curve corresponding to large values of L in an attempt to find the asymptotic L dependence. Moreover, the value of q depends strongly on the lattice size. For 100×100 lattice, one finds $q=1.46$, and for 200×200, $q=1.67$. As expected, as one goes to larger lattices, so that the length distribution is probed out to larger values of L, the value of q required to attempt to mimic the exponential dependence increases.

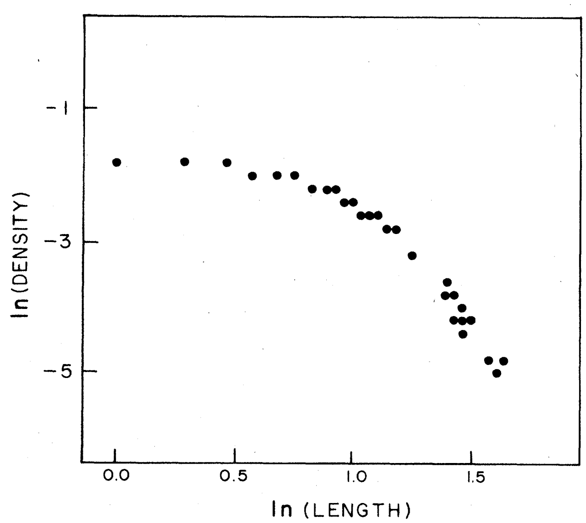

FIG. 5. The same data as in Fig. 4, plotted as $\ln n(L)$ vs $\ln L$.

IV. DISCUSSION

In the previous two sections, we have studied a two-dimensional analog of the system of monopoles connected by strings. We have shown that the density of $M\overline{M}$ pairs connected by strings of length L decreases exponentially with L. With such a length distribution of strings, it is clear that the monopole density will also decrease exponentially with time, and the Assumptions A and B will be violated. This counterexample is already sufficient to prove that the causality constraints in Assumptions A and B do not have general validity. Now it will be argued that these constraints also fail in the most interesting case of monopoles and strings in three dimensions.

Although it is natural to assume that the length distribution of strings in three dimensions will also be exponential, let us suppose that we start out with some other distribution, say $n(L)\propto L^{-\alpha}$. [We must require that $\alpha>2$, since otherwise the length of strings per unit volume, $\int n(L)L\,dL$, would diverge.] We expect long strings to have the shape of random walks; tension in curved strings will cause them to move at high speeds, and thus the strings will frequently intersect. Intersecting strings can intercommute (or change partners); as a result long strings will be cut in small pieces by intercommuting with much more numerous shorter strings. It is easily understood that the resulting length distribution of strings will be exponential: the probability for a string of length L to avoid intercommuting decreases exponentially with L. The kinetics of this process have been discussed in Ref. 6.

To take the worst possible case, suppose that the intercommuting probability is negligible, so that the strings can pass freely through one another. Still, there is a physical process that will eventually establish an exponential length distribution of strings. The strings connecting the monopoles are not topologically stable: they can break, producing monopoles and antimonopoles at the free ends. The semiclassical tunneling probability for this process per unit length of string per unit time is[6]

$$\kappa\propto\exp(-\pi m^2/\mu)\ ,$$

where μ is the string tension and m is the monopole mass. The probability for a string of length L to break in a time interval t is $\sim\kappa Lt$. Hence all strings longer than $(\kappa t)^{-1}$ will break into smaller pieces with a high probability.

The probability for a string with $L \gg (\kappa t)^{-1}$ to avoid breaking exponentially decreases with L, and thus an exponential length distribution of strings will be established. If the symmetry breaking scale of strings is much smaller than that of monopoles, the probability κ can be vanishingly small. However, as a matter of principle, both Assumptions A and B will be eventually violated.

To summarize, our conclusion is that causality alone does not impose any interesting constraints on the monopole annihilation. The actual rate of annihilation in each particular model can be determined by studying various dissipation processes, as done in Refs. 2, 6, and 13.

ACKNOWLEDGMENTS

One of us (A.V.) is grateful to Alan Guth for discussions and to the National Science Foundation for support.

[1]E. Weinberg, Phys. Lett. **126B**, 441 (1983).

[2]K. Lee and E. Weinberg, Nucl. Phys. **B246**, 354 (1984).

[3]A. Vilenkin, Phys. Lett. **136B**, 47 (1984).

[4]J. Preskill, Phys. Rev. Lett. **43**, 1365 (1979).

[5]P. Langacker and S.-Y. Pi, Phys. Rev. Lett. **45**, 1 (1980).

[6]A. Vilenkin, Nucl. Phys. **B196**, 240 (1982).

[7]T. W. B. Kibble, G. Lazarides, and Q. Shafi, Phys. Rev. D **26**, 435 (1982).

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

[9]P. Shellard (private communication).

[10]R. Peccei and H. Quinn, Phys. Rev. Lett. **38**, 199 (1981).

[11]T. Vachaspati and A. Vilenkin, Phys. Rev. D **30**, 2036 (1984).

[12]P. Sikivie, Proceeding of the Summer School on Particle Physics, Gif-sur-Yvette, 1983 (unpublished).

[13]A. Vilenkin, Phys. Rep. (to be published).