Evolution of cosmic string networks

Andreas Albrecht

*National Aeronautics and Space Administration/Fermilab Astrophysics Center, Fermi National Accelerator Laboratory,
Batavia, Illinois 60510*

Neil Turok

Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544

(Received 17 March 1989)

A discussion of the evolution and observable consequences of a network of cosmic strings is given. A simple model for the evolution of the string network is presented and related to the statistical mechanics of string networks. The model predicts the long-string density throughout the history of the Universe from a single parameter, which we calculate in radiation-era simulations. The statistical mechanics arguments indicate a particular "thermal" form for the spectrum of loops chopped off the network. Detailed numerical simulations of string networks in expanding backgrounds are performed to test the model. Consequences for large-scale structure, the microwave and gravity-wave backgrounds, nucleosynthesis, and gravitational lensing are calculated.

I. INTRODUCTION

The cosmic-string theory of the formation of structure in the Universe is both simple and, in principle, highly predictive.[1,2]

The existence of cosmic strings could provide one of the few possible ways to test unification physics. They would also provide an explanation for the origin of galaxies and clusters of galaxies, linking the spontaneous breaking of the underlying gauge symmetry in particle interactions with the breaking of spatial symmetry in the Universe.

The presence of cosmic strings in a unified gauge theory is purely a question of topology. The simplest SO(10) model, for example, predicts strings.[3] Many "superstring-inspired" models also predict cosmic strings. These strings would be formed at a symmetry-breaking phase transition, generally occurring at the grand unified scale. The distribution of such strings is predicted by the "Kibble mechanism," in which one takes the distribution of Higgs-field phases to be random on scales larger than the correlation length at the Ginzburg temperature, when thermal fluctuations are no longer string enough to erase the strings. This distribution has been calculated by Vachaspati and Vilenkin,[4] and others, and we shall take it as our starting point. Recently Hodges has performed dynamical simulations of string formation,[5] which support this picture. The crucial feature of the initial network is that most of the string appears in the form of "infinite" strings which wander like random walks clear across the Universe. It is this fact that guarantees that some cosmic strings will still be present at any time after they are produced.

One of the advantages of the cosmic-string theory of structure formation is that given a hot homogeneous big bang, the distribution of strings and the perturbations they induce in the surrounding matter is fully specified. They are independent of the precise initial conditions or parameters in the full field theory. This is because the *motion* of the strings is governed by a purely geometrical

action, the Nambu action, and a network of strings quickly enters a *scaling solution* which is independent of the precise initial distribution of string on small scales.[6-8] The magnitude of the perturbations produced by the strings is set by a single parameter, the string tension μ. For a recent review and references, see Ref. 9.

Of course cosmic strings (at least in the simplest scenario we consider here) do not address many of the other important cosmological questions. They do not explain the homogeneity and isotropy of the Universe or provide a significant contribution to the dark matter. It is also awkward (but not impossible) to construct models which provide cosmic strings, inflation, and dark matter (often with much unwanted besides). However, besides it being plausible on general grounds that a phase transition early in the Universe was responsible for the formation of large-scale structure, cosmic strings are a reasonably generic prediction of grand unification, and have the great virtue of being quite testable independent of the details of the underlying theory. We shall confine ourselves in this paper to the simplest cosmic-string scenario, which is highly predictive, and as we will see is already close to being ruled out by observation.

Our first calculations of the evolution of cosmic-string networks, in 1985, provided indications that a cosmic-string network might be just what is needed to explain the observed large-scale structure.[7] Indeed the correlation function of Abell clusters, one of the few systematic statistics on the largest-scale structure in the Universe, was miraculously fitted with no adjustable parameters by the correlation function of string loops chopped off the scaling string network.[10]

Since then, there has justifiably been a lot of interest in pursuing the predictions of cosmic strings further. However for the past few years the numerical problems posed by cosmic-string evolution have been a stumbling block. Perhaps more importantly, there has been little progress in understanding the evolution of string networks analytically. There have even been serious questions raised about whether the scaling solution exists at all.[11]

In this paper we hope to resolve many of these problems.

We shall describe the evolution of strings using a new set of nonsingular, "gauge-fixed" variables, and an improved method of detecting and enacting string interactions. These have led to numerical results significantly different from our first results, with the long-string density an order of magnitude larger than we originally thought. We shall describe some numerical tests of the new code which lead us to realistic estimates of our systematic errors, which we believe to be of the order of 50% (see Note added in proof). Our new string scaling density is approximately twice that reported in independent work by Bennett and Bouchet last year,[12] so the difference is within our estimated errors. There are still significant differences in our more detailed results, however, which we are in the process of trying to resolve.

More importantly, we shall describe conceptual improvements in our understanding of the problem. In particular we shall describe a simple model for the evolution of the long-string density inspired by earlier work of Kibble[8] and Bennett,[11] but representing an advance on that work. In particular we present a simple model for the velocities of the long strings, and argue from string statistical mechanics that the *sign* of the term governing long-string-loop energy exchange is fixed. This reduces the problem of calculating the string scaling density to the calculation of a single dimensionless number, the chopping efficiency. This may well be calculable in flat-spacetime simulations. The model allows one to calculate the scaling density for the string network in the radiation, transition or matter eras from a single parameter measured in string simulations. So, for example, given only the radiation-era scaling density we can *predict* the matter-era scaling density and even follow the string density right through the transition between the two eras.

We shall relate the model to the statistical mechanics of string networks in flat spacetime, and argue that the loops chopped off a string network may be viewed as "thermal" radiation from a hot body, the network of long strings. We also show how a string-dominated universe would be inconsistent with flat-spacetime statistical mechanics of strings, and from this argue that the scaling solution is inevitable. The statistical arguments lead to a form of the "loop production function," which we check in our simulations. As a consequence of this discussion, we argue that it should be possible, at least approximately, to calculate the scaling density of the string network in *flat-spacetime* simulations, which present few numerical problems.

Finally we discuss the observable consequences of our new numerical results. We show that the masses of galaxy clusters are reasonable if the string tensions $\approx (10^{16} \text{ GeV})^2$ (i.e., $G\mu = 10^{-6}$), and calculate the mass function in the one-loop–one-object picture for cold- or hot-dark matter. We then discuss distortions in the microwave background produced by strings, the gravity wave background caused by radiation from strings, and lensing of galaxies by strings. Any one of these observations could constrain the string tension to be too low for any appreciable structure formation, at least by gravita-

tional accretion.

We shall only deal briefly with the predicted large-scale structure from our new results. As a consequence of the higher density of a long string, the correlations produced in the distribution of galaxies and clusters of galaxies may not be as clear-cut as we originally thought. Our new simulations do find loop-loop correlations quite similar to the original results (as do the simulations of Bennett and Bouchet[13]), but the translation of these into a cluster-cluster correlation function is a nontrivial matter and we shall defer a full treatment of the problem to a future publication.

The paper is organized as follows. In Sec. II we give a detailed discussion of the equations of motion of cosmic strings in expanding backgrounds. We introduce a new set of "gauge-fixed" nonsingular variables which are useful in numerical evolution, and discuss the motion of small loops. The later parts of this section are technical, and the reader interested in our more important results may skip to Sec. III. There we present and solve a simple analytic model for the evolution of the string network. In Sec. IV we discuss the statistical mechanics of string networks in flat spacetime, and what insight they give us into the expanding Universe case. This discussion makes it clear that the scaling solution for the string network is inevitable, and provides a qualitative picture for the distribution of strings on all scales in the scaling solution. We suggest how the scaling density might be calculated from flat-spacetime string simulations. In Sec. V we present our numerical results. The simple model presented in Sec. III is shown to fit the results remarkably well. The next three sections are devoted to calculation of the observable consequences of a cosmic-string network. Section VI discusses the general issue of how to attribute "typical" properties to loops chopped off the network, and Sec. VII discusses the spectrum of massive objects accreted by the strings in cold- or hot-dark-matter-dominated universes. In Sec. VIII we discuss the other observational signatures of cosmic strings—fluctuations in the microwave background, the gravity-wave background, and lensing of galaxies.

Throughout this article we use units where $\hbar = c = k_B = 1$.

II. EQUATIONS OF MOTION

Immediately after cosmic strings form, they are heavily damped by collisions with particles in the surrounding medium.[14] This causes the strings to straighten out, so that the typical radius of curvature on the string rapidly becomes much larger than its width. The damping ceases to be important at a temperature of order $(G\mu)^{1/2}\mu^{1/2}$, and thereafter it is a very good approximation to treat them as infinitely thin relativistic strings[15,16] described by the Nambu action[17]

$$S = -\mu \int dA , \qquad (2.1)$$

which is simply proportional to the area traced out by the string world sheet in spacetime. In this section we will derive some properties of Nambu strings which will be used in our general discussion. Then we will discuss

some more technical issues which pertain to our numerical techniques.

The area element

$$dA = \sqrt{-q} \tag{2.2}$$

is obtained from the induced metric on the world sheet

$$q_{\alpha\beta} = g_{\mu\nu}(x)\partial_\alpha x^\mu \partial_\beta x^\nu . \tag{2.3}$$

Here $g_{\mu\nu}$ is the background spacetime metric and $x^\mu(\sigma)$ are the spacetime coordinates of the world sheet, parametrized by $\sigma^\alpha = (\tau, \sigma)$. A nice feature (2.1) is that it is purely geometrical: μ clearly drops out of the equations of motion, which depends solely on the background spacetime metric. But the early Universe was, as far as we can tell from present observations of the microwave background and matter density, a very nearly flat Friedmann-Robertson-Walker radiation- and/or matter-dominated spacetime.

Furthermore, according to the numerical calculations of Matzner[18] and Shellard,[19] (local) cosmic strings have very simple interactions—two colliding strings *always* reconnect the other way (Fig. 1). Thus the evolution of the string network is completely specified and has no adjustable parameters.

The equations of motion derived from (2.1) are

$$\partial_\alpha(\sqrt{-q}\,q^{\alpha\beta}\partial_\beta x^\mu) + \Gamma^\mu_{\nu\lambda}\partial_\alpha x^\nu \partial_\beta x^\lambda \sqrt{-q}\,q^{\alpha\beta} = 0 , \tag{2.4}$$

where $\Gamma^\mu_{\nu\lambda}$ are the Christoffel symbols for the background. To solve (2.4) we need to pick coordinates for the background and for the world sheet [corresponding to fixing the gauge invariance of (2.1) under general coordinate transformations and world-sheet reparametrizations]. It is convenient to pick coordinates in which the metric is conformally flat:

$$ds^2 = dt^2 - a(t)^2 d\mathbf{x}^2 = a(\eta)^2(d\eta^2 - d\mathbf{x}^2) . \tag{2.5}$$

We also choose $\tau = \eta(\sigma)$, simply slicing the world sheet at constant conformal time. We can use the remaining freedom in redefining σ to choose the velocity of each point labeled by σ to be always perpendicular to the string, i.e., $q_{\tau\sigma} = \dot{\mathbf{x}} \cdot \mathbf{x}' = 0$ (our notation is $\dot{x} = \partial_\tau x$, $x' = \partial_\sigma x$, etc.). With these choices,[6] (2.4) reduce to

$$\ddot{\mathbf{x}} + 2h\dot{\mathbf{x}}(1 - \dot{\mathbf{x}}^2) = \frac{1}{\epsilon}\partial_\sigma\left[\frac{\mathbf{x}'}{\epsilon}\right] ,$$

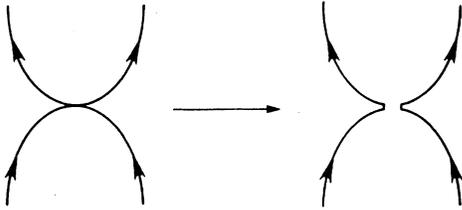$$\dot{\epsilon} = -2h\epsilon\dot{\mathbf{x}}^2 , \tag{2.6}$$

FIG. 1. Two colliding strings always reconnect the other way. This is the case for the simplest, directional strings. We shall not consider nondirectional strings in this paper.

with $h = \dot{a}/a$ and

$$\epsilon = \left[\frac{\mathbf{x}'^2}{1 - \dot{\mathbf{x}}^2}\right]^{1/2} , \quad \dot{\mathbf{x}} \cdot \mathbf{x}' = 0 . \tag{2.7}$$

The stress-energy tensor calculated from (2.1) is

$$\begin{aligned}
T^{\mu\nu}(\mathbf{x},\eta) &= \frac{-2}{\sqrt{-g}}\frac{\delta S}{\delta g_{\mu\nu}} \\
&= \frac{\mu}{\sqrt{-g}}\int d^2\sigma\sqrt{-q}\,q^{\alpha\beta}\partial_\alpha x^\mu \partial_\beta x^\nu \delta^4(x - x(\sigma)) \\
&= \frac{\mu}{a^4}\int d\sigma(\epsilon\dot{x}^\mu \dot{x}^\nu - \epsilon^{-1}x'^\mu x'^\nu)\delta^3(\mathbf{x} - \mathbf{x}(\sigma,\eta)) .
\end{aligned} \tag{2.8}$$

In particular, the total energy and momentum in the string are given by

$$\begin{aligned}
E &\equiv \int d^3\mathbf{x}\,a^3 T^{tt} = \int d^3\mathbf{x}\,a^3 T^\eta_\eta = \mu a(\eta)\int d\sigma\,\epsilon , \\
P^i &\equiv \int d^3\mathbf{x}\,a^3 T^{ti} = \mu a(\eta)\int d\sigma\,\epsilon\dot{\mathbf{x}} ,
\end{aligned} \tag{2.9}$$

where t is defined in (2.5). Thus in this gauge $\mu a\epsilon$ is the energy per unit parameter length along the string. From (2.6) it follows that

$$\dot{E} = h(1 - 2V^2)E , \quad \dot{\mathbf{P}} = -h\mathbf{P} , \tag{2.10}$$

where the average velocity squared on the string

$$V^2 \equiv \frac{\int d\sigma\,\epsilon\dot{\mathbf{x}}^2}{\int d\sigma\,\epsilon} . \tag{2.11}$$

Equations (2.10) and (2.11) will prove very useful in Sec. III, where we develop an intuitive physical picture of the network evolution. We now turn to more technical aspects of the equations which will come into play when discussing numerical issues.

Returning to (2.6), it is well known that in Minkowski spacetime ($h = 0$) we can choose σ so that $\epsilon = 1$ initially, and (2.6) will preserve this for all time.[20] We then have as the full system of equations

$$\ddot{\mathbf{x}} = \mathbf{x}'' , \quad \dot{\mathbf{x}} \cdot \mathbf{x}' = 0 , \quad \dot{\mathbf{x}}^2 + \mathbf{x}'^2 = 1 , \tag{2.12}$$

which are solved in terms of "left-movers" **a** and "right-movers" **b** as follows:

$$\mathbf{x} = \tfrac{1}{2}[\mathbf{a}(\sigma + \tau) + \mathbf{b}(\sigma - \tau)] ,$$

$$\mathbf{a}'^2 = \mathbf{b}'^2 = 1 , \tag{2.13}$$

so $\mathbf{a}'$ and $\mathbf{b}'$ are constrained to be unit vectors, but **a** and **b** are otherwise arbitrary functions. For a closed loop, $\mathbf{a}'$ and $\mathbf{b}'$ describe closed trajectories on the surface of a unit sphere.[21] Furthermore, it is easily seen that in its center-of-mass frame a loop's motion is periodic with period $L/2$ where L is its length, defined by energy over μ. Where the curves $\mathbf{a}'$ and $-\mathbf{b}'$ cross on the sphere, "*cusps*" occur—the string instantaneously reaches the speed of light and doubles back on itself.[22]

This suggests defining approximate "left-movers" and "right-movers" in the expanding Universe case as well: we set

$$\mathbf{l} = \dot{\mathbf{x}} + \mathbf{x}'/\epsilon , \quad \mathbf{r} = \dot{\mathbf{x}} - \mathbf{x}'/\epsilon , \tag{2.14}$$

which automatically obey

$$l^2 = r^2 = 1 \ . \tag{2.15}$$

Substituting into (2.6) we find

$$\dot{l} = l'/\epsilon - h r + h (l \cdot r) l \ ,$$
$$\dot{r} = -r'/\epsilon - h l + h (l \cdot r) r \ , \tag{2.16}$$
$$\dot{\epsilon} = -h \epsilon (1 + l \cdot r) \ .$$

Equations (2.15) and (2.16) provide a closed set of equations governing the evolution of $\dot{x}$ and x'. They are completely "gauge fixed" but do not have any "singular" variables. Furthermore, we can obviously evolve l and r exactly in flat spacetime; for constant ϵ, l is constant along constant $\epsilon \sigma + \tau$ and r is constant along constant $\epsilon \sigma - \tau$. Our numerical scheme is designed to include the effects of expansion as a small perturbation about the exact flat-spacetime solution (as we explain in the Appendix).

By contrast, using (2.6) directly, with ϵ defined from (2.7) presents difficult numerical problems—generically some points on the string reach the speed of light instantaneously and at these points x'^2 must go to zero as well to keep ϵ finite. In our first paper[23] (and in the paper of Bennett and Bouchet[12]) the gauge conditions (2.7) were not imposed numerically—the variables $\dot{x}, x'$, and ϵ were evolved according to their own equations. In our first numerical code this led to the problem that ϵ actually drifted away from its definition over time, which meant that our simulations did not conserve energy properly.

Our new numerical scheme preserves (2.15) automatically and is much simpler to implement—no "fudges" are needed. Bennett and Bouchet have recently also written a new code using the nonsingular variables (2.16). Our detailed results still appear to be significantly different, and we are at present engaged in trying to locate the reasons for this.

In addition to $x'(\sigma)$ (which may be obtained from l, r, and ϵ), we need a single position in order to reconstruct any loop of string. It is obviously better not to single out any particular point but to use the center of mass of the loop as our one extra variable for each loop. This is kept and evolved (using the known and updated center-of-mass velocity) for every loop.

It shall be very useful to us later to show that loops whose size is much smaller than the Hubble radius evolve to a good approximation as if they were in flat spacetime. For short times one can see this as follows. If h is small, and we choose $\epsilon = 1$, then (2.12) is a first approximation to (2.6). Taking a solution of (2.10) one finds that the curvature term on the right-hand side of (2.6) is of order the inverse (comoving) curvature radius of the string. The damping term is of order h, the inverse comoving Hubble radius. Thus for small loops the damping term is a small perturbation.

What is the effect of the damping term over long times? First, consider the evolution of the center of mass of the loop. Defining the velocity of the loop to be $\mathbf{v}_c = \mathbf{P}/E$ we find from (2.10) that

$$\dot{\mathbf{v}}_c + 2h (1 - V^2) \mathbf{v}_c = 0 \ . \tag{2.17}$$

We can solve this in the approximation that we treat the V^2 term by averaging over an oscillation of the loop, taken to be a zero-order (i.e., flat-spacetime) solution. This is valid to order h as explained above.

The flat-spacetime result is very simple[24] (we can set $\epsilon = 1$ here)

$$V^2 \equiv \frac{\int \dot{\mathbf{x}}^2 d\sigma \, d\eta}{TL} = \frac{-\int \ddot{\mathbf{x}} \cdot \mathbf{x}}{TL} + \frac{\left[\int d\sigma \, \dot{\mathbf{x}} \cdot \mathbf{x} \right]_0^T}{TL}$$
$$= -\frac{\int \mathbf{x}'' \cdot \mathbf{x}}{TL} \mathbf{v}_c^2$$
$$= \frac{\int (1 - \dot{\mathbf{x}}^2)}{TL} + \mathbf{v}_c^2$$
$$\rightarrow V^2 = \tfrac{1}{2}(1 + \mathbf{v}_c^2) \ . \tag{2.18}$$

Here T, L, and v_c are the period, length, and center-of-mass velocity of the loop. Note that for a loop at rest, the average velocity squared $V^2 = \frac{1}{2}$. Using (2.18) in (2.17) we find that, to order h^2,

$$\dot{\mathbf{v}}_c + h \mathbf{v}_c (1 - \mathbf{v}_c^2) = 0 \ , \tag{2.19}$$

which is exactly the equation for a point particle in an expanding background.

Furthermore, substituting (2.18) into (2.10) we find

$$\dot{E} = -h \mathbf{v}_c^2 E \ . \tag{2.20}$$

This makes sense; a very fast moving loop $(\mathbf{v}_c^2 \approx 1)$ is like a photon—its energy is red-shifted as a^{-1}. However a nonrelativistic loop has nearly constant energy. Equation (2.19) is easily solved, and yields

$$v_c = \frac{v_{ci} a^{-1}}{(1 - v_{ci}^2 + v_{ci}^2 a^{-2})^{1/2}} \ , \tag{2.21}$$

where the initial velocity is v_{ci} and the initial scale factor is chosen to be 1. Now (2.20) can be integrated for E:

$$\frac{E}{E_i} = (1 - v_{ci}^2)^{1/2} \left[\frac{v_{ci}^2 + a^2 (1 - v_{ci}^2)}{a^2 (1 - v_{ci}^2)} \right]^{1/2} \tag{2.22}$$

so that as $a \rightarrow \infty$ the energy remaining in the loop is simply the rest-mass energy.

What about the internal oscillations of the loop? It is possible to choose coordinates so that the Christoffel symbols vanish, and the metric is the Minkowski metric along any *world line*.[25] (It is well known that this is possible at a point, and this is demonstrated in most textbooks. It is actually crucial for the equivalence principle, which is all we are really using here, that one can do so along a geodesic world line, since observers live on world lines, not at points.) In particular we may do so along a line chosen to run through the center of the "world tube" swept out by a closed loop as it moves through spacetime. The time coordinate in the case where the line is geodesic is just the proper time for the particle traveling along the geodesic. Returning to (2.4) it is now convenient to pick "orthogonal gauge," where $q_{\alpha\beta} = \Omega \eta_{\alpha\beta}$, $\eta_{\alpha\beta} = \text{diag}(1, -1)$, in which the string equations become

$$\ddot{x}^\mu - x''^\mu + \Gamma^\mu_{\nu\lambda}(\dot{x}^\nu\dot{x}^\lambda - x'^\nu x'^\lambda) = 0 \ ,$$

$$g_{\nu\lambda}(\dot{x}^\nu\dot{x}^\lambda + x'^\nu x'^\lambda) = g_{\nu\lambda}(\dot{x}^\nu x'^\lambda) = 0 \ . \tag{2.23}$$

The zeroth-order solution is simply $x^0 = \tau$, with $\mathbf{x}(\sigma,\tau)$ obeying the flat-space equations (2.12). Recalling that the Christoffel symbols are to be evaluated on the world tube, we see that the second term in the evolution equation is of order r/R_H^2 where r is the spatial size of the loop and R_H the Hubble radius (in these coordinate), compared to the first two terms which are of the order of $1/r$. In the cases of interest, $R_H \propto t$, it is easily seen by using a Green's function that the second term causes negligible disturbance to the evolution of the loop in the long time limit. Thus the loop evolves, to a better and better approximation, as if it were in flat spacetime in its own local inertial frame.

III. THE SCALING SOLUTION

The notion of scaling plays a central role in understanding the evolution of a network of cosmic strings. The idea is that statistically the properties of the string network are the same at two different times, once all linear dimensions are rescaled by the ratio of the two Hubble lengths. The scaling picture says that not only does the scaling solution exist, but that any initial string network satisfying "randomness" on large scales will evolve towards the scaling solution with time. This means that observable predictions based on cosmic strings are quite insensitive to the initial string configuration. It also makes parameter-independent predictions possible. As we emphasized above, the spatial distribution of the strings does not depend on μ, the string tension, since the equations of motion are independent of μ. Thus correlation properties of the network are independent of μ.

First, let us define some useful terms. We define Hubble's constant $H \equiv (da/dt)/a$ and the Hubble radius $R_H \equiv H^{-1}$. The "length" l of a string loop is defined as e/μ where e is its energy. "Long strings" are strings whose length is longer than the Hubble radius R_H. "Loops" are strings shorter than R_H, although technically of course, much of the "long string" may be in the form of finite loops as well.

The main idea of the "scaling solution"[23,26] is that there is a single scale in the problem: the Hubble radius R_H. All other scales are determined in terms of R_H. Thus there is a total length in long strings of the order of R_H per volume R_H^3, so

$$\rho_L \propto \frac{\mu}{R_H^2} \ . \tag{3.1}$$

Einstein's equations tell us that the total energy density is also proportional to R_H^{-2}, so the ratio of the two energy densities is constant in time.

In this section we will set up a simple model for string evolution which exhibits scaling behavior and which we will show (in Sec. V) agrees well with the numerical simulations. This model is based on a "one-scale" principle which allows a simple understanding of the network evolution, even far away from the scaling solution.

The initial-string network formed at the phase transition is composed largely of long strings which wander as random walks right across the Universe. The precise details of the phase transition are irrelevant—the distribution of long strings is a result of the phases of the Higgs field being uncorrelated on large scales.[1,4]

Let us consider the evolution of these long strings. We define a length scale ξ on the string

$$\rho_L = \frac{\mu}{\xi^2} \ , \tag{3.2}$$

where ρ_L is the density in long strings.

As long as reconnection is frequent between the strings, which should be the case if $\xi \ll R_H$, it will keep the network "random" so that ξ should be related, by a constant factor of order unity, to the typical radius of curvature on the string, and the typical distance between strings. This is the "one-scale" principle. It results in a simple model of the string network, where the string distribution is characterized by ξ alone, even when scaling ($\xi \propto R_H$) has not yet been reached.

The existence of the scaling solution may then be argued in the following way. If ξ becomes much smaller than R_H, the long strings rapidly chop off loops and the long-string density falls, so ξ grows faster than R_H. If ξ grows larger than R_H, chopping off becomes infrequent, the string density rises, and ξ falls relative to R_H.

To see this more quantitatively, let us assume ξ is $\ll R_H$, so that from the discussion of the previous section the strings average velocity squared V^2 should, as explained above, be close to $\frac{1}{2}$. Thus neglecting interactions, the energy in string should remain constant [Eq. (2.10)] so the string density should evolve as matter. In the radiation era this means ρ_L is growing compared with scaling, which would have ρ_L evolving as radiation.

Now let us include the effect of interactions. As mentioned above, these simply cause two colliding strings to reconnect the other way. Statistical mechanical calculations of the density of states for free Nambu strings which we will discuss further in the next section show that at low density $\rho \ll \mu^2$ there are many more states available for the long string to chop itself up into loops than there are for it to remain in long string. Even though we are far from equilibrium in the present discussion, it means that we can expect chopping off of loops from long string to be favored by phase space over reconnection of loops. In fact the time scale for the string to chop a given fraction of its length off into loops must be related to ξ by a constant factor, the chopping "efficiency" c, which one would expect to be a fairly small fraction. We shall have more to say about the value of c in Sec. V.

Putting this together, we have

$$\frac{d\rho_L}{dt} = -3H\rho_L + (1 - 2V^2)H\rho_L - c\frac{\rho_L}{\xi} \ . \tag{3.3}$$

The first term alone would give ρ_L evolving as matter. The second term gives the correction to $\dot{E} = 0$ from (2.10). The third term represents the loss of energy into loops. Note that it is the *physical* time we use here.

We shall adopt a simple model for predicting V^2. The quantity $(1-2V^2)H$ just measures growth of energy (or length) of the string due to stretching. If a long string is a random walk with a correlation length ξ then it takes a total length $L = R_H^2/\xi$ to cross each volume R_H^3 (assuming $R_H \gg \xi$). On scales larger than R_H the string is stretched by the expansion of the Universe. If we imagine "pinning" a long string to the background at two points separated by R_H, the string will be stretched at a rate $\dot{L} = HR_H = 1$. Dividing by L we obtain $\dot{L}/L = H\xi/R_H$, so we take [using (2.10)]

$$(1-2V^2)H = H\xi/R_H \tag{3.4}$$

and (3.3) is just

$$\frac{d\rho_L}{dt} = -3H\rho_L + H\frac{\xi}{R_H}\rho_L - c\frac{\rho_L}{\xi} . \tag{3.5}$$

Note that although the physical picture behind Eq. (3.4) assumed $\xi \ll R_H$, the formula also makes sense in the other extreme. A string which is straight on the scale of R_H (that is, $\xi = R_H$) should have $L \propto a$, and $V^2 = 0$.

Now we define the number of correlation lengths ξ per Hubble radius:

$$\gamma \equiv \frac{R_H}{\xi} = (\rho_L R_H^2/\mu)^{1/2} \tag{3.6}$$

and (3.5) becomes

$$\frac{d\gamma}{dt} = -\frac{H}{2}[c\gamma^2 - (2\dot{R}_H - 3)\gamma - 1] . \tag{3.7}$$

In the radiation era $\dot{R}_H = 2$; setting the right-hand side equal to zero we find the fixed point

$$\gamma_r = \frac{1}{c}\frac{1+\sqrt{1+4c}}{2} \approx \frac{1}{c}, \quad c \ll 1 . \tag{3.8}$$

Similarly in the matter era $\dot{R}_H = \frac{3}{2}$ and the fixed point is given by

$$\gamma_m = \frac{1}{\sqrt{c}} . \tag{3.9}$$

In the radiation era the chopping term must make up an extra $-H\rho_L$ [in (3.5)] as well as counteract the stretching term in order to keep the string scaling as radiation $(\dot{\rho} = -4H\rho)$. In the matter era however (where $\dot{\rho} = -3H\rho$) the chopping term need only balance against the stretching term. Thus one expects ξ to be larger in the matter era since less chopping is required. The result, for $c \ll 1$, that $\gamma_m \approx \sqrt{\gamma_r}$ is remarkably simple and as we will see is verified to good accuracy in our simulations. Likewise we predict the velocities of the long strings in both eras:

$$V_r^2 = \tfrac{1}{2}(1-c), \quad V_m^2 = \tfrac{1}{2}(1-\sqrt{c}) . \tag{3.10}$$

It is also clear that the scaling solutions are *stable* — the right-hand sides of (3.7) are inverted parabolas with the scaling solution at the positive root. If γ is greater than the scaling solution it falls, if it is less than the scaling solution it rises. In fact (3.7) may be solved in the radiation era to give

$$\delta\gamma \equiv \gamma - \gamma_r = \frac{\epsilon\delta\gamma_i}{a^{\epsilon/2}(\epsilon + c\delta\gamma_i) - c\delta\gamma_i}$$

$$\approx \delta\gamma_i a^{-\epsilon/2}, \quad c\delta\gamma_i \ll 1 ,$$

$$\epsilon \equiv \sqrt{1+4c} , \tag{3.11}$$

where $\delta\gamma_i$ is the initial deviation from scaling. Thus γ approaches its scaling value rather slowly, as $a^{-\epsilon/2}$.

Likewise in the matter era one finds

$$\delta\gamma \equiv \gamma - \gamma_m = \frac{2\delta\gamma_i}{a^{\sqrt{c}}(2+\sqrt{c}\,\delta\gamma_i) - c\delta\gamma_i}$$

$$\approx \delta\gamma_i a^{-\sqrt{c}}, \quad \sqrt{c}\,\delta\gamma_i \ll 1 . \tag{3.12}$$

For small c this approaches scaling even more slowly.

At this point we should also mention a small correction to the model due to the V dependence of the chopping efficiency c. This is a small effect because both in radiation and matter eras V^2 is not very far from $\frac{1}{2}$. To a first approximation we should simply have c proportional to V because this determines the rate of interactions of the long string. This we can include by writing $c = c_r V/V_r$.

The model is easily generalized to the matter-radiation transition. In this case it is more convenient to change variables from t to $a(t)$ which is given by

$$\left(\frac{da}{dt}\right)^2 = \frac{8\pi G\rho_{m,\mathrm{eq}}}{3}(a^{-1} + a^{-2}) , \tag{3.13}$$

where $\rho_{m,\mathrm{eq}}$ is the matter density at t_{eq}, the time of equal radiation and matter density and we take $a(t_{\mathrm{eq}}) = 1$.

Now we find (3.3) becomes

$$\frac{d\gamma}{da} = \frac{1}{2}\left[\frac{\gamma}{a(1+a)} - c\frac{\gamma^2}{a} + \frac{1}{a}\right] . \tag{3.14}$$

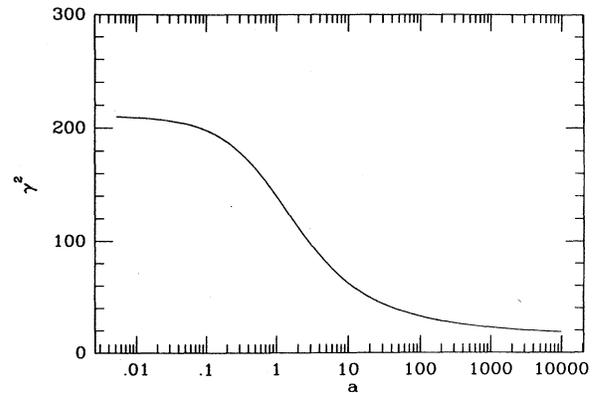This equation is simple to solve numerically, and the result is plotted in Fig. 2, for $c_r = 0.074$, the result of our

FIG. 2. The solution to our "one-scale" model for the long-string density ρ_L. The figure shows $\gamma^2 \equiv \rho_L R_H^2/\mu$ (the long-string density in Hubble radius units) as a function of scale factor a. The units are chosen so that $a = 1$ at equal matter and radiation density, and the single parameter in the model, the chopping efficiency c, is taken from our numerical results in the radiation era.

numerical simulations explained in Sec. V.

Let us now work out the consequences of the scaling solution for the distribution of loops being chopped off the long strings. The chopping term in (3.3) represents the loss of energy into loops: if we define the dimensionless *energy production function* $f(x)$ to be the energy loss from the long string into loops of length l to $l+dl$ per correlation volume per unit time we have

$$c\frac{\rho_L}{\xi}=\frac{\mu}{\xi^3}\int_0^{l_c}\frac{dl}{\xi}f\left[\frac{l}{\xi}\right],$$

$$\partial_t\rho_l(l)=-3H\rho_l(l)+\frac{\mu}{\xi^4}f\left[\frac{l}{\xi}\right],$$

(3.15)

where $\rho_l(l)dl$ is the energy density in loops of length l to $l+dl$. The first term just represents dilution due to the expansion of the Universe. The cutoff l_c is simply a result of our definition of long string—in our numerical results, for example, we shall typically *define* $f(x)$ to be zero for $x>x_c=2(R_H/\xi)$. We shall see that the precise value of the cutoff is irrelevant— for x of order 10 or greater chopping off and reconnection balance very closely. Note that f as we define it here includes both chopping off and reconnection. In the next section it will be useful to differentiate between these explicitly.

Here we make the important assumption that f is only a function of l/ξ. The one-scale principle has come into play again. It says that no matter what the scale ξ of the network, the loop production process looks the same when scaled with ξ. In particular, if one understands loop production in the radiation era, then an appropriate rescaling will describe loop production in the matter era, and even in the radiation-matter transition. We shall check this in Sec. V. Note that f is the *net* amount of energy lost per Hubble time per unit volume—it includes both chopping off and reconnection. As we explained above, we expect chopping off to be greater than reconnection from the statistical mechanical results, and we shall explore how this works in detail in the next section. We have ignored effects of expansion on the loops, where the main (small) effect is to redshift kinetic energy away as explained in Sec. II.

What about the loop distribution? Of course loops fragment after being chopped off the network, so it is convenient to divide loops into two types; those destined to self-intersect (variables corresponding to these will have a subscript I) and those which are not (these will have a subscript NI). In particular we have $\rho_I(l)$ $=\rho_I(l)+\rho_{NI}(l)$ and $f(x)=f_{L\to I}(x)+f_{L\to NI}(x)$ as the production function from long string and (3.15) becomes

$$\partial_t\rho_I(l)=-3H\rho_I(l)+\frac{1}{(\xi)^4}\left[f_{L\to I}\left[\frac{l}{\xi}\right]-f_{I\to NI}\left[\frac{l}{\xi}\right]\right],$$

(3.16)

$$\partial_t\rho_{NI}(l)=-3H\rho_{NI}(l)$$
$$+\frac{1}{(\xi)^4}\left[f_{L\to NI}\left[\frac{l}{\xi}\right]+f_{I\to NI}\left[\frac{l}{\xi}\right]\right].$$

In fact $f_{NI}\equiv f_{L\to NI}+f_{I\to NI}$, the non-self-intersecting

energy production function, determines the final density in non-self-intersecting loops—integrating (3.16) and using $\xi=R_H/\gamma$ we find in the scaling solution in the radiation era that

$$\rho_{NI}(l)dl=\mu\lambda_r\left[\frac{R_H}{l}\right]^{3/2}\frac{dl}{R_H^3},$$

$$\lambda_r=\frac{\gamma_r^{5/2}}{2}\int dx\,x^{1/2}f_{NI}(x)$$

(3.17)

and in the matter era that

$$\rho_{NI}(l)dl=\mu\lambda_m\left[\frac{R_H}{l}\right]\frac{dl}{R_H^3},$$

$$\lambda_m=\frac{2\gamma_m^3}{3}\int dx\,f_{NI}(x)$$

(3.18)

assuming of course that the integrals converge. We shall present strong evidence for this and determine λ_m and λ_r from our numerical simulations. The number density of loops of length l to $l+dl$ is given by $n(l)\equiv\rho_l(l)dl/(\mu l)$.

We have so far ignored the slow decay of loops into gravitational radiation, which is crucial in the radiation era since the loop density (3.17) scales as matter, and without this process the loop density would come to dominate the Universe. A loop of initial length l_i produced at a time t_i loses energy at a rate $\dot{l}=-\Gamma G\mu$ with Γ a constant which depends on the loop trajectory but typically $\Gamma\approx50$ for simple trajectories.[22,27] Thus the length of a loop varies with time as $l=l_i-\Gamma G\mu(t-t_i)$. Equations (3.17) and (3.18) are the densities as a function of the initial length l_i—substituting for l_i we obtain, in the radiation era for $t\gg t_i$,

$$\rho_{NI}(l)dl=\mu\lambda_r\frac{dl}{R_H^{3/2}(l+\Gamma G\mu t)^{3/2}},$$

(3.19)

and in the matter era that

$$\rho_{NI}(l)dl=\mu\lambda_m\frac{dl}{R_H^2(l+\Gamma G\mu t)}.$$

(3.20)

Thus loop decay determines a cutoff in the size of "typical" loops of the order of $\Gamma G\mu t$.

As we discussed above, in the scaling solution we have $\int dx\,f(x)=c$. This led to $\rho_L\approx\mu/(c^2R_H^2)$ in the radiation era and $\approx\mu/(cR_H^3)$ in the matter era. Likewise we should have for the density of long string plus intersecting loops ρ_{LI} similar relations in terms of the integral of the non-self-intersecting energy production function, $\int dx\,f_{NI}(x)\equiv c_{NI}$. This provides a useful consistency check which we will apply to our string simulations in Sec. V.

IV. STATISTICAL MECHANICS AND THE ENERGY PRODUCTION FUNCTION

In this section we shall discuss the statistical mechanics of string networks. This will be useful in several ways. The most important fact we will learn is that phase space favors the chopping up of long string into loops over the

reconnection of loops onto long string. This fixes the *sign* of our constant c of the previous section, and guarantees that a scaling solution must exist, as we have discussed. We shall argue that it is inconsistent with flat-spacetime statistical mechanics to have a string-dominated universe where chopping off is balanced by reconnection.[28] This conclusion is different from that reached by Bennett[11] and Kibble.[26]

We also suggest that the long strings may be viewed as a radiating hot body. Loop production can be then thought of as the "blackbody radiation" coming off the network. This suggests a simple form for the shape of the loop production function. In the next section we show that the predicted shape is fit very well by the simulations.

Let us begin by reviewing flat-spacetime string statistical mechanics. To discuss this, one needs to be able to count states in the space of all possible string configurations. The measure on the space of states may be obtained by quantizing the strings,[28] or from a simpler argument we present below.

It is useful to think in terms of the "left-movers" and "right-movers" of (2.13). Let us introduce a fundamental energy scale Δ and construct the set of all loops of energy $e = N\Delta$ by choosing $\mathbf{a}'$ and $\mathbf{b}'$ to be randomly chosen unit vectors at each step Δ along the string (energy is proportional to parameter length in this gauge). This scheme has been used by Hawking[29] and by York.[30,31] We must furthermore restrict $\mathbf{a}'$ and $\mathbf{b}'$ to only take a discrete set of directions D at each step. This procedure guarantees that the displacement $\Delta \mathbf{x}'$ and momentum $\Delta \dot{\mathbf{x}}$ are themselves taken from a discrete set (quantized) for every step along the string. The number of such sets $\mathbf{a}'(\sigma)$ and $\mathbf{b}'(\sigma)$ is just $D^{2e/\Delta} \equiv \exp(be)$. Thus b^{-1} ($\propto \Delta$) emerges as fundamental energy scale (the precise value of b depends on exactly how one defines the set D). For any one of these sets, $\mathbf{a}(\sigma)$ and $\mathbf{b}(\sigma)$ execute random walks given by adding up all the $\mathbf{a}'$ and $\mathbf{b}'$ vectors. (Copeland, Haws, and Rivers[32] give another "classical" treatment of this problem where they count random walks in *position* space rather than in terms of $\mathbf{a}$ and $\mathbf{b}$. However, they neglect the momentum degrees of freedom of the string segments. See also Ref. 33 for a similar discussion of the string initial conditions.)

We still must further restrict ourselves to closed loops in their center-of-mass frame. This is only the case when we impose the two additional constraints

$$\int d\sigma \, \mathbf{a}'(\sigma) = \int d\sigma \, \mathbf{b}'(\sigma) = 0 \ . \qquad (4.1)$$

Thus we need to select out the class of random walks which return to the origin. As is well known, the probability of a given walk returning to the origin after N steps is proportional to $N^{-3/2}$ in three space dimensions. Taking this into account for both $\mathbf{a}'$ and $\mathbf{b}'$ and using $e = m$ in the center-of-mass frame of the loop we obtain $m^{-3}e^{bm}$ for the number of closed strings of mass m. However this is still not quite right because any given loop could be obtained by starting at any of the m/Δ steps on it, and so we have overcounted by this factor in our construction. So the number of different configurations of mass m to $m + dm$ is just

$$n(m)dm \propto m^{-4}e^{bm}dm, \quad bm \gg 1 \qquad (4.2)$$

with a and b constants depending on D and Δ. Note that the measure requires some fundamental scale of allowed "wiggles" on the strings. The argument given is easily generalized to any number of dimensions and to open strings. In each case it agrees with the "Hardy-Ramanujan" formula which one obtains when the theory is quantized.

It is also clear with the above measure what typical loops look like: $\mathbf{x}(\sigma) = \int^\sigma \frac{1}{2}(\mathbf{a}' + \mathbf{b}')$ clearly executes a random walk in space as we track along the string.

Now let us discuss equilibrium distribution of loops in flat spacetime. From (4.2) one calculates the equilibrium number density of loops in the energy range e to $e + de$ at finite temperature $T = \beta^{-1}$:

$$n_{\text{eq}}(e)de \propto \int d^3\mathbf{p} \int dm \, n(m)\delta(e - \sqrt{\mathbf{p}^2 + m^2})\exp(-\beta e)$$

$$\approx ae^{-5/2}\exp[(b - \beta)e], \quad be \gg 1 \ , \qquad (4.3)$$

where one uses $m \approx e - \mathbf{p}^2/2e$ to perform the mass integration. The total energy density is given by

$$\rho = \int_{e_c} de \, n_{\text{eq}}(e)e \approx \int_{e_c} ae^{-3/2}\exp[(b - \beta)e] \ , \qquad (4.4)$$

where e_c is a lower-energy cutoff, usually taken to be of order b^{-1}.

There are several important points about (4.4). First, the expression makes no sense for $\beta < b$, i.e., $T > T_H \equiv b^{-1}$, the Hagedorn temperature. This is a reflection of the fact that the canonical ensemble is not defined above T_H. More importantly, (4.4) tends to a *finite* limit ρ_H at T_H. Densities above ρ_H cannot be described with the canonical ensemble—a microcanonical description is necessary. In fact the initial-string configuration is described by this $\rho > \rho_H$ configuration with most of the string in "infinite" string and the remainder in a distribution of loops, given by (4.4) with $\beta = b$.

Densities below ρ_H are well described by the canonical ensemble.[28] From (4.3) and (4.4) one can see that the loop distribution and energy density are dominated by the smallest loops allowed, with energies of e_c. Smith and Vilenkin, and Sakellariadou and Vilenkin have numerically evolved boxes of string in flat spacetime[34] and have verified these results. A box of long string (with $\rho < \rho_H$) grinds itself up into small loops which settle into the distribution given by (4.3). In their case the smallest allowed loop is set by their numerical scheme, and the scale of "wiggles" by the lattice spacing. For cosmic strings, the "smallest allowed" loop would be a loop which is not much longer than a string thickness, which also sets the scale of allowed wiggles. Loops smaller than this can disintegrate into their constituent fields. In *thermal equilibrium*, at any temperature below the Hagedorn temperature large strings would be exponentially rare. Long strings survive only to the extent that the network *does not* equilibrate.

Boxes of string in which there is a net string winding number across the box were also considered in Ref. 28. These impose a topological constraint which forces there

to be a minimum number of long strings in the box. The energy in wiggles on the long strings is suppressed by the Boltzmann factor at low-string density, just as large loops are. Even though the long strings carry kinks whose separation can be as small as the scale of allowed wiggles, they remain very straight. Not surprisingly, the equilibration process does its best to put energy into the statistically favored small loops.

In the context of cosmology, equilibrium statistical mechanics cannot be applied directly. However, lessons learned in the above discussion have important implications.

The main difference of course is that as the Universe expands, the mean separation between strings grows and the loop velocities redshift away. This has the effect of "turning off" the interactions between different loops. As we shall discuss, this interrupts the fragmentation process. In the expanding Universe, we expect a given loop to fragment only down to the scale set by the smallest wiggles it had acquired before it stopped interacting with other strings. (We neglect the effects of gravitational radiation in this analysis since they only become important over much longer time scales than we are concerned with here.)

It is important to emphasize here that loops broken off a string characterized by the scale ξ do *not* fragment indefinitely — the scale of the typical fragments is set by ξ. This has been shown convincingly by York[31] in flat-spacetime simulations of loop self-intersection (see also Ref. 35). He also gives the following simple argument. Defining the space of loops as we have above, all loops consist of straight sections connected at kinks. Take a loop with N left-moving kinks and N right-moving kinks. If this breaks off a child loop, there have to be at least five kinks on the child loop. Two are created at the crossing, and at least three are necessary to "bend the string around" back to the crossing point. This is so because kinks are actually *planar* (velocity of the string perpendicular to the plane of a kink has to be equal on both sides of it). Thus the case with only two kinks "bending the string around" is degenerate. Now only two new kinks are created on the "parent" loop, while three were lost from it. Thus eventually the parent loop runs out of kinks and can no longer chop off loops. The maximum number of child loops allowed is just $2N - 4$. This argument ignores reconnection of fragments onto the parent loop, but York has verified that for an isolated initial-loop reconnection has very little effect on the final-energy distribution produced for $x = l/\xi > 1$.

York also found that the probability for a loop of N straight segments to be non-self-intersecting was exponentially small for large N. This can also be understood by considering the measure we discussed above. As a simple case, consider constructing $\mathbf{a}$ and $\mathbf{b}$ on a simple cubic lattice. Since there are six directions for $\mathbf{a}'$ and $\mathbf{b}'$ at each step, the number of closed loops on this lattice scales with N as $6^{2N}N^{-4}$, as discussed above. Now an intersection occurs on a loop if and only if $\mathbf{a}(x+L)+\mathbf{b}(x+L) = \mathbf{a}(x)+\mathbf{b}(x)$ for some L and x. In other words, if $\mathbf{a}$ and $-\mathbf{b}$ trace out identical vectors along stretches of the same length L. In the case of our lattice,

an intersection of length 1 occurs if $\mathbf{a}$ and $\mathbf{b}$ travel in the same direction anywhere (since both are closed curves, they must traverse both forwards and backwards along any direction they take). The only way to avoid such "intersections" is if $\mathbf{a}$ lies in a two-dimensional plane and $\mathbf{b}$ lies along a line, and vice versa. Similarly intersections of length 2 can occur if $\mathbf{a}$ and $\mathbf{b}$ go forward and back along any link in two adjacent steps. $\mathbf{b}$, the one-dimensional closed walk, is guaranteed to do this at some point, but $\mathbf{a}$ can avoid it by never reversing on itself. Now it is easy to see that all intersections up to length 4, and all odd length intersections, are avoided if the conditions imposed so far are met. Assuming that higher L intersections are rare, we estimate the total number of non-self-intersecting loops as $(3^N/N)(2^N/N^{1/2})(1/N)$, where the denominators come about because one curve is two dimensional and the other one dimensional, and we include the overcounting factor N as before. Thus the fraction of all loops which are non-self-intersecting is proportional to $N^{3/2}6^{-N}$. Of course the details of the result depend strongly on the lattice, and our decision to treat intersections of length 1 (which are really "cusps") as intersections, but the conclusion that the probability of a loop being non-self-intersecting decreasing exponentially with N should be independent of these details.

The finite fragmentation result means that if a string network is *smooth* on the fundamental scale given by e_c, then interactions between different loops are crucial to the equilibration process. It is only through reconnections that extra wiggles can be introduced in order to allow fragmentation all the way down to the scale e_c. The exponential result for the probability of non-self-intersecting loops means that we can expect the vast majority of final non-self-intersecting loops to be in "simple" trajectories, of only a few steps on the relevant scale.

One can estimate this scale, which determines when loop fragmentation stops, by determining when the probability per unit time for any loop to hit a long string falls below the expansion rate. The former is essentially $\approx kl/\xi^2$, with l the loop length, and $k < 1$ a factor determining the geometrical cross section for a loop to hit a long string. (Remember, ξ^{-2} gives the length density of the long strings.) Thus loops with $l \ll k^{-1}\gamma^{-1}\xi$ are very unlikely to interact with the long string. One can check that they are even less likely to interact with similarly sized loops. So the long string-loop system cannot "thermalize" further than this scale, smaller than but proportional to $\xi \propto R_H$ in the scaling solution.

What about the long-string distribution? As the Universe expands it proceeds to chop itself up into loops. As it does so, the scale ξ on the long string grows. Provided the long string remains in a random configuration it will be characterized by the scale ξ, and there will be of order 1 string of length ξ per volume ξ^3. In fact the distribution of long string on scales much larger than ξ should look very much like the high-density string phase (above the Hagedorn density), where the smallest allowed scale of wiggles on the string is taken to be ξ. This is so not because the string network has equilibrated, but simply because this is the most probable configuration. It follows that most of the energy density in long strings

should be in strings much larger than the Hubble radius. We emphasize that this configuration occurs because the long string is random, and this is the most likely configuration. The often used "causality" bound is quite misleading on this point.[36]

The actual density of long string is set by the chopping efficiency (the parameter c described in Sec. III), and is considerably higher than one per Hubble volume (the naive "causality" bound). The long string cannot be straight on scales much larger than the mean separation of long-string segments, since interactions will tend to randomize it on that scale. This is exactly the picture we presented in Sec. III.

Now we are ready to discuss the form of the energy production function.

In flat spacetime and in thermal equilibrium we can calculate the spectrum of loops chopped off a long string by detailed balance. The process of chopping off a loop is the time reverse of the process where a loop collides with a long string. In equilibrium the two processes must therefore proceed at the same rate. The rate at which our long string collides with loops of length l to $l+dl$ is given by $Lk\,ln_{eq}(l)dl$ where L is its length and k is a constant. We assume that the cross section for a string of length l to hit other strings is proportional to l and that the characteristic velocity involved is a constant of order unity. Thus in equilibrium the rate of loss of energy to loops of energy e to $e+de$ is given by

$$\frac{\dot{E}(e)}{E}=ke\frac{e}{\mu}n_{eq}(e)\equiv\frac{1}{\xi}f_{off}\left[\frac{l}{\xi}\right],\qquad(4.5)$$

which defines the energy production function for loops coming off the long string: $f_{off}(x)$. The production function $f(x)$ of the previous section includes both this and the reconnection function. Using n_{eq} from (4.3) we find

$$f_{off}(x)=Ax^{-1/2}e^{-Bx},\quad Bx>1\;.\qquad(4.6)$$

A useful analogy is with a hot body, connected via a conducting rod to a radiating surface. The temperature is greatest at the hot body, and least at the surface. The long-string network may be viewed as very hot, at the Hagedorn temperature in fact. This is so even though it has never had time to equilibrate—it is just because it is a "typical" configuration of strings which are constrained to be straight (because smaller wiggles get chopped off) on a scale similar to the scale defined by their density ξ. Because the long strings are at the Hagedorn temperature the production function from the long strings should have $B=0$, and be a pure $x^{-1/2}$ at large x. This may also be understood more directly—the probability for a random walk to return to the origin after a length l scales as $l^{-3/2}$, as we mentioned before. Thus if the long strings remain random walks, the number of fragments breaking off with length l to $l+dl$ should scale the same way. The energy production function should therefore scale as $l^{-1/2}$, in agreement with the argument above.

Going down in scale, the network explores smaller scales as chopping up occurs and is described by a configuration at a density lower than the Hagedorn density, and with $B>0$. Energy flows continuously from the

"hot" long strings to the "colder" loops. Finally, after all chopping is done with, a distribution of non-self-intersecting loops will leave the network. The exact form of this distribution is determined by the nature of the "cooling" process, and by how the loops which leave the network break into nonintersecting loops. We have yet to model all these effects in a detailed way, but we continue our discussion under the simple assumption that the distribution of nonintersecting loops is well described by a production function of the form (4.6). We shall see in Sec. V that this assumed form gives a remarkably good fit to the numerical results.

Now we turn to the reconnection of loops onto long strings. This is the main difference between the flat spacetime and expanding cases—in the former case all loops eventually reconnect, so the scale of allowed "wiggles" can go to zero. As we have argued however, in the expanding case the allowed scale grows in proportion to $\xi\propto t$, and a substantial fraction of loops never reconnect. For simplicity we shall deal only with non-self-intersecting loops. We shall consider the rest of the loops as part of the long string network.

The rate of loss of energy in loops of length l to $l+dl$ due to reconnection onto long string is given by $k\rho_l(l)l/\xi^2$. Using this, we have, for the energy density in loops from l to $l+dl$,

$$\dot{\rho}_l(l)=-3H\rho_l+\frac{\mu}{\xi^4}f_{off}\left[\frac{l}{\xi}\right]-k\frac{l}{\xi^2}\rho_l(l)\;,\qquad(4.7)$$

which is easy to integrate using the form (4.6) [it is helpful to integrate with respect to $x=l/\xi$ instead of t; we assume for simplicity that (4.4) holds down to $x=0$]. For the final density in loops of length l to $l+dl$ in the radiation era one finds

$$\rho_l(l)=\mu\frac{1}{2}\frac{\gamma^{5/2}}{R_H^{3/2}l^{3/2}}e^{(kl\gamma)/2\xi}\int_{1/\xi}^{\infty}dx\,f_{off}(x)e^{(-k\gamma x)/2}$$

$$=\mu\frac{1}{2}\frac{\gamma^{5/2}}{l^{3/2}R_H^{3/2}}Ae^{-(Bl)/\xi}\frac{1}{B+\gamma k/2}\qquad(4.8)$$

and the net energy production function $f(l/\xi)$ $=f_{off}(l/\xi)-k\xi^2l\rho_l/\mu$ is found to be a fraction $B/(B+\gamma k/2)$ of the chopping off function f_{off}. The reconnection function $r(l/\xi)=-k\xi^2l\rho_l/\mu$, defined as the negative contribution to the net production function, is a fraction $\gamma k/2B$ of the net production function. As we shall see later this prediction, that the reconnection function is an l-independent fraction of the production function in the radiation era, is well borne out in our simulations.

Note that B plays a crucial role above. If we set $B=0$ then we would find that production exactly canceled reconnection and the string density would scale as matter. It is the fact that a substantial fraction of the energy lost from long strings goes into loops whose size is set by ξ that guarantees the scaling solution. We know that B has to be nonzero to ensure that Eq. (4.7) produces the correct Boltzmann distribution in flat spacetime.

In fact (4.7) and our previous argument put a much stronger constraint on the loop production function. The

segments of long string should interact with the "sea" of small loops that surround them just as the topologically constrained long strings mentioned before. In flat spacetime, at low-string density, chopping off and reconnection of small loops onto a long string balance, with very little energy in the "wiggles" on the long string. In an expanding Universe, where many loops coming off never get the chance to reconnect, it must be that chopping off is at least as efficient as reconnection on small scales. Demanding this, (4.7) *requires* that $f_{off}(x)$ diverge at least as fast as $x^{-1/2}$ at small-x. In our numerical results we shall see evidence of this $x^{-1/2}$ small-x behavior building up as the simulation proceeds.

This in fact justifies our approximation above in assuming the form (4.6) down to $x = 0$. If we impose the condition that chopping off is greater than reconnection for all l in (4.7) then our conclusions above are not qualitatively altered by deviations at small x from (4.6). Bennett[11] considered Eq. (4.7) for general $f(x)$ in much the same way as we have. However he allowed a form for f for which small loops reconnected onto long strings at a faster rate than they were chopped off. This then allowed a string-dominated universe. Our statistical discussion makes it clear that this cannot happen.

As we explained above, as well as "infinite" strings, there should be a distribution of large loops given by (4.8) with $B = 0$. One of the striking results we shall show in the next section is how little density resides in loops with length greater than ξ. From (4.8) the density in large loops is suppressed by $1/k$. We assumed that the cross section for all loops to reconnect was proportional to kl above, with k the *same* constant. However large loops look more like a collection of random segments intertwined with the long strings, so k should be of order unity. In contrast, small loops are "curled up" by some geometrical factor—while the cross section should still be proportional to l, because they are mostly in the "simplest" trajectories, k should be smaller. So the constant k should really also depend on l, and have the effect of suppressing the density in large loops relative to small ones.

In the matter era the integrals are more complicated. However for small l one finds instead

$$\rho_l(l) = \mu \frac{2}{3} \frac{\gamma^3}{l R_H^2} A e^{-(Bl)/\xi} \left[\frac{\pi}{B + \frac{2}{3} k\gamma} \right]^{1/2} . \tag{4.9}$$

In this case the reconnection function is seen to be a decreasing fraction of the net production function at small l. For small l the reconnections should be unimportant and Eq. (4.6) should give a good approximation to the net loop production function in the matter era as well.

This picture of the string network as a radiating hot body characterized by the scale ξ is quite different from previous analytic approaches to string networks which allowed for a string-dominated universe. Our discussion makes it clear that this cannot occur—if the scale on the string network ξ became much smaller than the horizon, then in a time of the order of a few ξ the string network would approach the equilibrium solution, with the long string chopping itself up into loops. Therefore ξ/R_H

would fall. Reconnection does indeed reduce the final loop density [Eq. (4.7)] but as long as B and k are finite we are inevitably driven into the scaling solution.

It may well be possible to calculate the loop production function from flat-spacetime simulations. Certainly the production function in equilibrium may be calculated as a function of string density and the cutoff scale for wiggles. With knowledge of exactly how the interaction rate for the network depends on string density one should be able to calculate the effective ξ as well. Using our formalism in the previous section one could then calculate the scaling density. This would be worth doing—simulations in flat spacetime are much easier to do.

V. NUMERICAL RESULTS

In the previous sections we have developed an analytic model for the evolution of the string network. In this section we shall compare its predictions with our numerical results. The reader interested in the details of our numerical methods should consult the Appendix.

We shall also discuss some tests of our code in order to give estimates of the possible errors in our results. In the Appendix we discuss some extra checks.

Let us begin by discussing the density in long strings. In Sec. III we developed a simple model for the time evolution of the long-string density. It contains a single free parameter, the chopping parameter c. If we calculate c from the scaling density in the radiation era, we can use the model to predict the scaling density in the matter era, and the rate of approach to scaling in both eras. The model also predicts the string density throughout the matter-radiation transition, which is important in calculating the growth of density perturbations produced by strings.

Figure 3 shows the model compared to our numerical simulations. In these simulations there is one parameter which determines the initial conditions, the ratio of the Hubble radius R_H to the correlation length ξ, which we have called γ. This determines the long-string density: $\rho_L = \gamma^2 \mu / R_H^2$. We define the long strings to be loops whose length is larger than $2R_H$. As we shall see, there is little energy density in large loops, so the results are very insensitive to exactly where the dividing line between "long string" and "loops" is drawn. For each run γ^2 is plotted against scale factor.

In Fig. 3(a) the results are shown for the radiation era. Three different runs are shown, in solid lines. In the first γ falls, in the second it is steady, and in the third it rises. We deduce from this that $\gamma^2 \approx 210$ is the scaling solution. From (2.8) we see that this corresponds to $c_r = (1 + \gamma)/\gamma^2 \approx 0.074$.

With this one parameter fixed, we can now calculate the predicted density evolution for each run from our scaling model. The predictions are marked in dashed lines—they clearly fit the numerical results very well.

In the figures the units are chosen such that the abscissa is also the Hubble radius in units of a comoving initial correlation length. For example, the longest run was in a box 26 initial correlation lengths across which contains a

volume $(26/45)^3 R_H^3 \approx 0.2 R_H^3$ by the end of the run. The statistical fluctuations in these graphs are small—they are quite smooth. This is not too surprising since even by the end of the longest run the box contained $N = 0.2 R_H^3 / \xi^3 \approx 600$ correlation volumes. As we shall see, the long-string segments are uncorrelated on scales larger than ξ so an estimate of the statistical error in the long-string density is simply $1/\sqrt{N} \approx 4\%$. The possible systematic errors, due to numerical errors in the string evolution code and finite cutoff effects, are more important, and we will discuss these later.

Figure 3(b) shows the results for the matter-radiation transition. Again the model fits well. The abscissa here is the scale factor in units where it is unity at equal matter and radiation density.

Figure 3(c) shows the results for the matter era. We calculated c_r above. Including the velocity correction we have $c_m \approx c_r (1 - \frac{1}{2}\sqrt{c_r}) \approx 0.064$ so the predicted matter scaling density is given by $\gamma_m^2 = 1/c_m \approx 16$. Using the scaling density from the radiation era, we have *predicted* the matter scaling density, and again the model fits the simulations very well.

Judging from this we believe that the solution to our model shown in Fig. 3 should give an accurate representation of the string density throughout the history of the Universe.

Our scaling model relies on an estimate of the amount of stretching for long strings, which is particularly important near the matter scaling solution. Figure 4 shows the average V^2 on the long strings for each of the runs shown in Fig. 3, as a function of conformal time. The predicted velocity is calculated from $V^2 = \frac{1}{2}(1 - \xi/R_H) = \frac{1}{2}(1 - \gamma^{-1})$, with γ as plotted in Fig. 3.

The velocity model gives slightly low predicted V^2 during the radiation or transition eras [Figs. 4(a) and 4(b)] although always by less than 10%. It fits the numerical results remarkably well during the matter era [Fig. 4(c)]. In these runs V^2 approaches the predicted value and settles on it where the run continues long enough. Since this very simple model works well in the matter era, where stretching is most important, we have not tried to improve on it.

Further support for both the scaling density model and the velocity model comes from calculating the shape of the long strings. In Fig. 5 we have measured the average straight-line distance squared d^2 between two points separated by a length $l = \text{(energy}/\mu)l$ along the string for the longest radiation run. We only did this for $l < L/10$ where L is the length of the loop involved, since otherwise the simple random-walk formula breaks down.[28] As can be seen, $d^2 \approx l\xi$ to good accuracy throughout the simulation (during which ξ grows by almost 8 in physical units). We implicitly assumed this in constructing the velocity model.

Another test of the long-string configuration is shown in Fig. 6, where the correlation function of the tangent vector along the string $\langle \mathbf{x}'(l) \cdot \mathbf{x}'(0) \rangle$ is shown as a function of the length (energy/μ) along the string. In Fig. 6(a) the correlation function is plotted for various times during the long radiation run, to check for scaling. In Fig. 6(b) the radiation and matter runs are compared with the
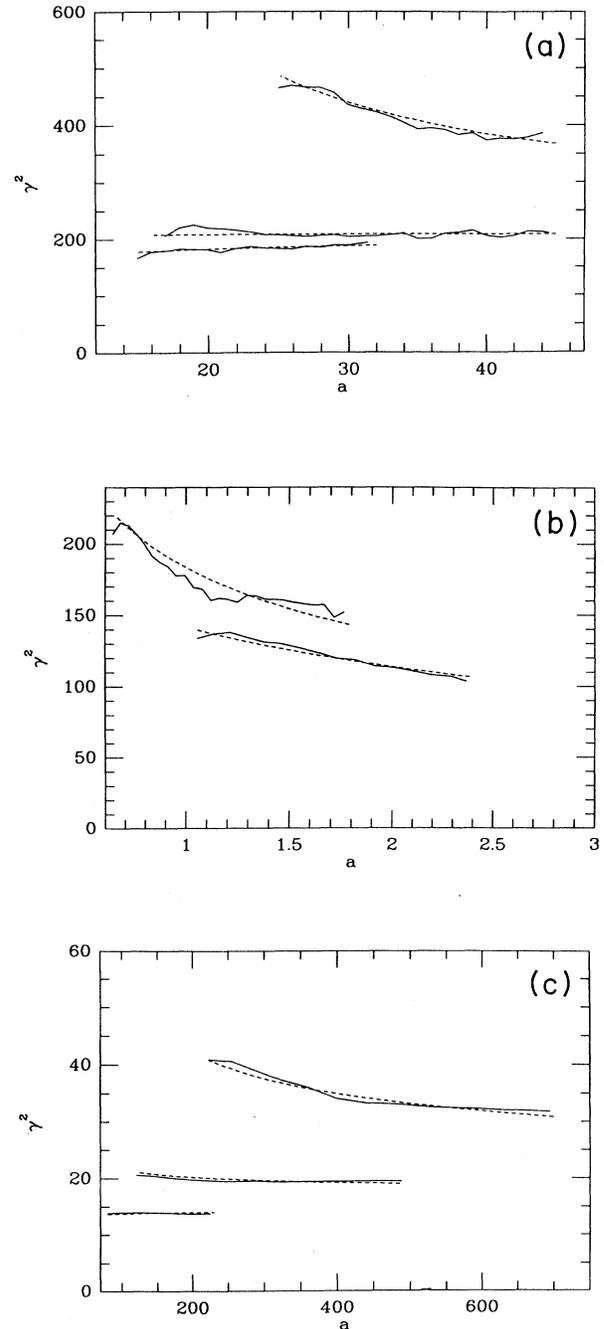

FIG. 3. The model for the density in long string presented in Sec. II compared to numerical calculations, for different initial string densities. (a) shows the results in the radiation era. The numerical result is the solid line and the model prediction the dashed line in each case. The vertical axis is the long-string density in Hubble radius units, $\gamma^2 = \rho_L R_H^2 / \mu$, and the horizontal axis the scale factor. The single parameter in the model, the chopping efficiency c, is determined from the run for which γ^2 is constant, and approximately equal to 210. (b) shows the model versus simulations in the matter-radiation transition. The scale factor is in units where it equals unity at equal matter and radiation density. The model predictions here are deduced from the value of c measured in the radiation era. (c) shows the matter-era results similarly.

initial conditions. The correlation functions are nonzero for $l > \xi$ in the scaling network, but not in the initial conditions. We interpret this as evidence of the "stretching" we have assumed in our velocity model. The matter era network is somewhat more highly correlated, which corresponds to the greater degree of stretching we expect in the matter era.

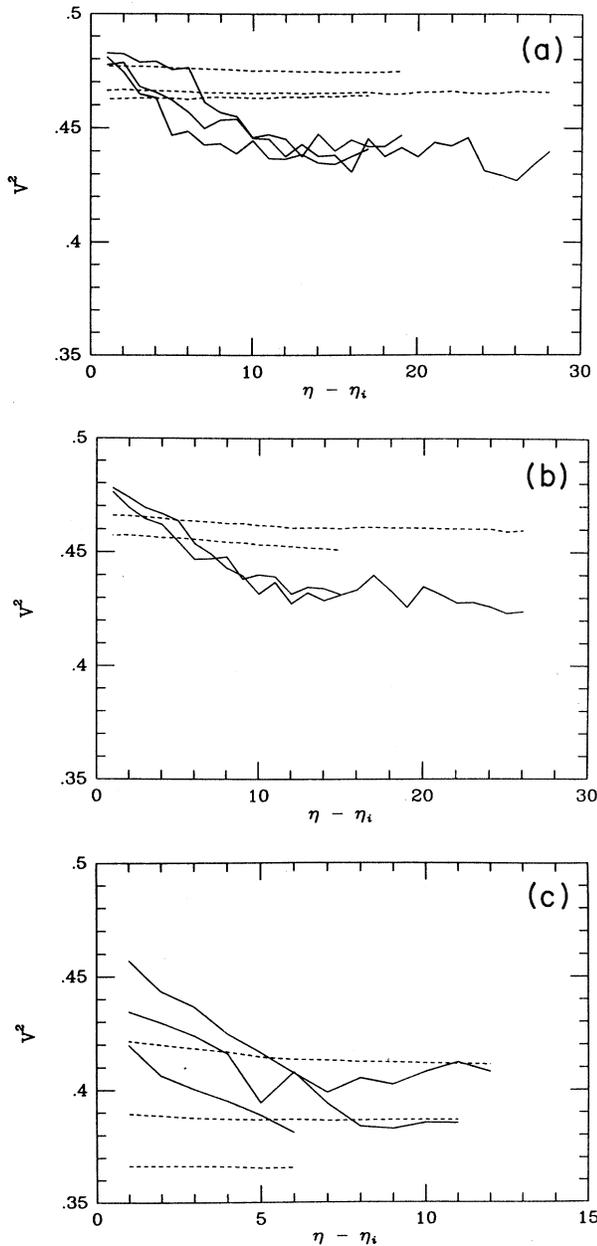Thus, aside from small stretching effects, the long strings look like random walks with step length ξ. Thus

FIG. 5. The shape of the long strings as measured by d, the straight-line distance between two points on the string, compared to l, the length (energy/μ) along the string between them. d^2 is plotted against l at different times in the longest radiation-era simulation.

FIG. 4. The model for the average velocity squared of the long strings is compared to the numerical results for each run in the previous figure. Here the horizontal axis is the conformal time minus the initial conformal time. The model prediction is shown in dashed lines, the numerical results in solid lines. (a)–(c) show radiation, transition, and matter runs, respectively.
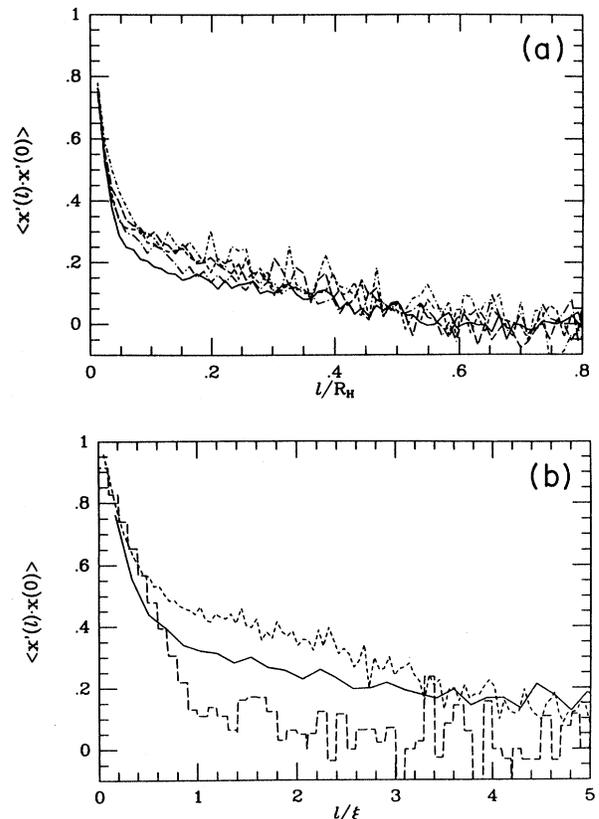
FIG. 6. The shape of the long strings as measured by the correlation function for the tangent vector to the string, as a function of l. (a) shows the radiation-era run at different times in the simulation (the solid line corresponds to an early time when scaling has not really set in). (b) shows the radiation (solid line) and matter (dashed line) runs (when they have reached scaling) compared to the initial conditions (broad dashes)— some evidence of "stretching" of the long strings is seen.

the initial conditions of the string network provide a surprisingly good approximation to the scaling configuration of the long strings (with the appropriate choice for ξ, of course). This is very reassuring—it means that the initial conditions for our simulations are quite close to the final scaling solution, at least for the long-string distribution.

The energy lost from long strings goes into loops. It is an important consistency check on our simulations to see that enough energy is lost from the long strings to keep

the network scaling. Figure 7 shows the energy production function $f(l/\xi)$ [Eq. (3.16)] during the course of the longest radiation run. The function f (the solid line) is shown for five different ranges of the scale factor—the run started at 16, and the average value of f is shown for the ranges 20–25, 25–30, 30–35, 35–40, 40–45. While there are obviously quite large fluctuations in f, it settles down after 25 and is reasonably similar during the different time intervals thereafter. Thus it shows good evidence of scaling. As we discussed in Sec. III, the in-
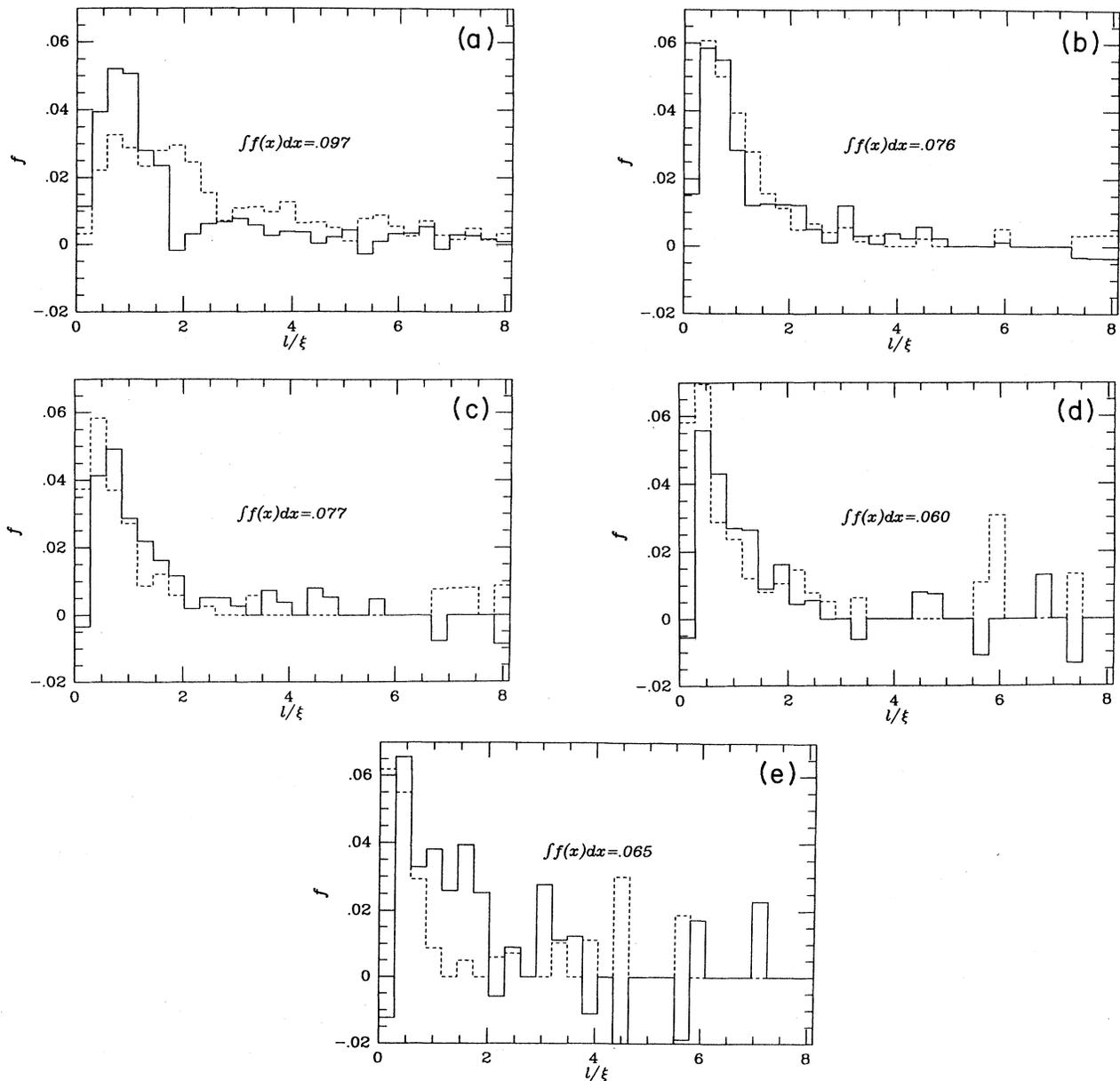
FIG. 7. The net energy production function $f(l/\xi)$ defined in Eq. (3.16) is plotted (the solid line) against l/ξ for different time intervals during the course of the longest radiation run. It settles down and appears to be roughly constant towards the end of the run. The reconnection function is similarly shown in dashed lines. Note that the total energy coming off the strings equals the sum of these two curves, so reconnection is approximately one-half of chopping off. The integral of it, $\int f(x)dx$, is also shown on each figure. See the text for further details. The production function averaged over the whole scaling part of the run appears in Fig. 9.

tegral of f is equal to c_r in the scaling solution—if the integral is larger than c_r, then γ^2 will fall, if it is smaller then γ^2 rises. Above we calculated $c_r \approx 0.074$. Comparing the value of the integral of f, shown with each figure, with 0.074, one can see the value of γ^2 rise or fall over the corresponding period in Fig. 3(a). This is a good consistency check of our results.

Also shown (the dashed lines) in each case is the reconnection function $r(l/\xi)$, we briefly discussed in Sec. IV. This is defined as the (negative) contribution to f coming from loops reconnecting onto long strings. The other component of f is due to chopping off, and equals the sum of the solid and dashed lines in the figures. It is very noticeable that the reconnection function has the same shape, and approximately the same magnitude as f. This is exactly as we predicted in Sec. IV.

Our results for f show it going to zero at $l/R_H \approx 0.02$. Note that $f = f_{\text{off}} + r$, and the reconnection function does grow at small x, which implies that f_{off} does as well. Our discussion in Sec. IV indicated that f_{off} should diverge at least as fast as r at small x, which should go like $x^{-1/2}$. Our simulations show $f_{\text{off}} \approx r$ at small x. However, the corresponding loops are of order $l/11 \approx 0.08$ in rms radius, in units of an initial comoving correlation length, in which $R_H = 45$ at the end of the run. This is certainly comparable to the resolution of our crossing detection boxes, $\frac{1}{15} \approx 0.07$ in these units, and so could well be a cutoff effect. We checked this in a much higher resolution run, with 5 times the usual number of points per initial correlation length on the string, and 6 times smaller effective crossing detection (see the Appendix). We could only do a small simulation—the box was 12^3. In Figs. 8(a) and 8(b) the average production and reconnection functions f and r over the entire course of this run are compared with the results for our long radiation run over the same period. There is indeed some enhancement of the loop production function at small scales, but contributing less than 10% to the integral. Similarly, the reconnection function is slightly smaller in the higher resolution run. Since the eventual scaling density $\gamma^2 \propto 1/c^2$ and c the chopping efficiency, is directly related to the integral of f, we make a rough estimate of the error in c by doubling this, to obtain 20%. Consequently our scaling density could be too high by 40%. A really systematic estimate of our errors would require a higher resolution run in a large box, which we have not yet performed (see Note added in proof).

In fact our production functions look quite similar in form to those found by York[31] in exact simulations of the breakup of loops in flat spacetime. His results suggest that for isolated loops breaking up in flat spacetime f actually turns flat below the scale $\approx \xi$ on the string.

Now we turn to the matter-era simulations. Figure 9 shows the matter-era energy production function f_m (solid line) and reconnection function r_m (dotted-dashed line). The radiation-era production function (the average of the last four graphs in Fig. 7) is also shown for comparison (dashed line). f_m is obviously quite similar in form to f_r when expressed in units of ξ, the scale on the strings. f_m does appear to be somewhat steeper however.

As in the radiation case we calculate the integral $c_m = \int dx\, f_m(x) \approx 0.051$, which is a little lower than the predicted value 0.064 calculated from the radiation-era run.

The difference could be a statistical fluctuation—the matter run was smaller (21^3) and may not have reached scaling. However, even if such a difference were real we would regard the simple one-scale model as quite successful. Most of the variation in the size of the chopping term ($c\xi^{-3}$) between the matter and radiation eras is given by the ξ dependence. The ξ dependence is set by the one-scale model and correctly accounts for a factor of $(\gamma_r/\gamma_m)^3 \approx 3000$ change in the chopping rate (in Hubble units).

The final distribution of loops produced by the network is determined by $f_{\text{NI}}(x)$, the non-self-intersecting energy production function. For each loop we store the "birthday" (the time it was produced) and a loop is defined as non-self-intersecting if its "age" is longer than one period (its energy divided by 2μ). In order not to bias the result towards small loops, we require that all loops accepted as non-self-intersecting have lived as long as a period of the largest accepted loop. Every loop gives a contribution to the function f_{NI} when it is born. If one is interested in
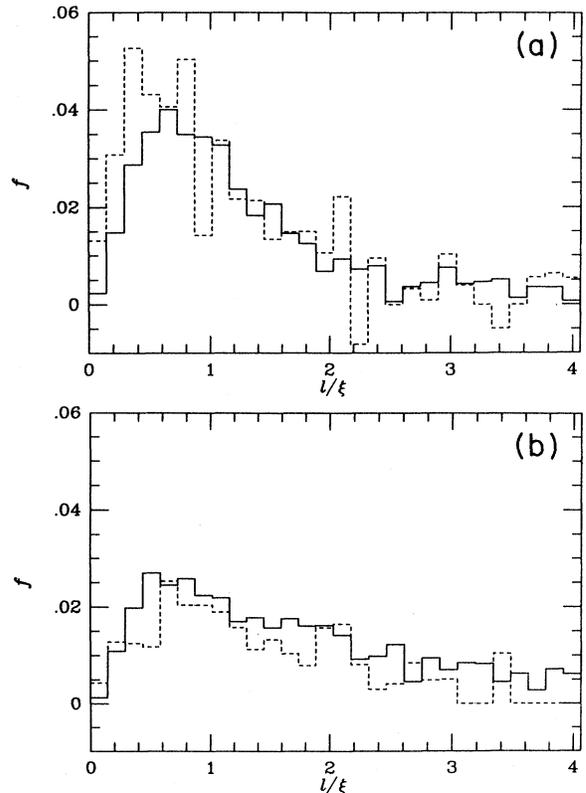

FIG. 8. Production and reconnection functions in a run with 5 times the usual resolution are compared with the same from a "normal" run over the same period. (a) compares the production functions $f(x)$, with the high resolution run result being the dashed line and the "normal" run the solid line. (b) compares the reconnection functions similarly.

the final-loop distribution, it is of course important to measure the contribution to f_{NI} over some time interval by looking at the simulation at some much later time, so that the loops counted are those that do not eventually reconnect. We shall show graphs of f_{NI} calculated from the surviving loops at the end of the simulation.

First, in Fig. 10(a) we show $f_{NI}(1/\xi)$ in the radiation era, for different times in the simulation. The solid line is for scale factor 26–32, the dashed line for 32–38, and the dotted-dashed line for 38–44. The most noticeable deviation from scaling is at small x, where one can see more and more smaller loops being produced. This accords with our discussion in Sec. IV, where we showed that $f_{off}(x)$ must diverge at least as fast as $x^{-1/2}$ for small x. Over most of its range f_{NI} does appear to be scaling, however. In Fig. 10(b) we plot the integral $\int dx\, f_{NI}(x) \approx 0.11$ which does appear quite constant over the course of the simulation. This is higher than the integral of the total energy production function from long string discussed above, $\int dx\, f_r(x) \approx 0.071$. As we discussed at the end of Sec. III, this implies that $(\rho_L + \rho_I)/\rho_L = \int f_{NI}/\int f \approx 1.5$. Thus approximately half the long-string density again is contained in eventually reconnecting loops shorter than $2R_H$. Figure 10(c) shows the integral $\int dx\, f_{NI}(x)$ in the matter era run for two time intervals. As can be seen, it falls during the run, to about 0.06. This is not very different from $\int dx\, f_m(x) \approx 0.051$ mentioned above. In our model for the scaling network we expect these two quantities to be the same.

Now for many purposes it is important to have the parameters λ_m and λ_r in the final-loop distribution (3.17) and (3.18). In fact for loops produced in the radiation era it is better to define this in terms of rest mass rather than energy since some of the initial energy in a loop is redshifted away, and the most important effects (mass accretion, gravitational radiation) occur late in its lifetime. Redefining all energies as rest masses, we have calculated $I = \int dx\, x^{1/2} f_r(x)$ from the radiation run, and this is

shown in Fig. 11. It seems to scale quite well. The average value over the course of the run is $I \approx 0.033$, which results in a value $\lambda_r = \frac{1}{2}\gamma_r^{5/2} I \approx 11$. As a check on this, we have calculated the loop distribution at fixed times in the simulation to compare with (3.17). In Fig. 12(a) the number of loops with length greater than l in a Hubble volume is plotted against $(1/R_H)^{-3/2}$. The slope of this graph, according to (3.17), should be $\frac{2}{3}\lambda_r$. As can be seen, the slope stabilizes during the course of the run, at a value corresponding to $\lambda_r \approx 10.5$, in good agreement with
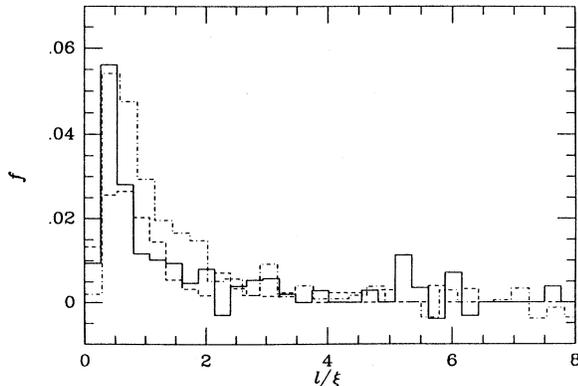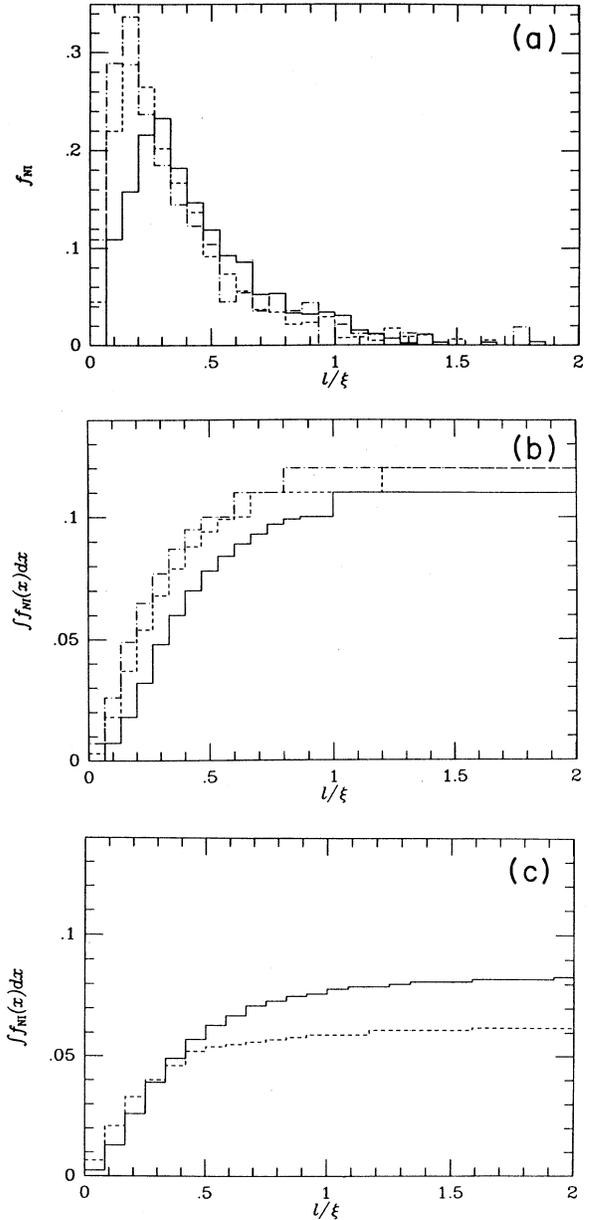

FIG. 9. Matter-era energy production function f_m (solid line) and reconnection function r_m (dashed line) as functions of l/ξ. Also shown is the average radiation-era production function (dotted-dashed line).

FIG. 10. Non-self-intersecting loop-production function at three different times in the radiation era. (a) shows the function $f_{NI}(x)$, and (b) shows $\int_0^x f_{NI}(x)dx$ vs x through the course of the simulation. (c) shows $\int_0^x f_{NI}(x)$ for the matter-era run at two different times.
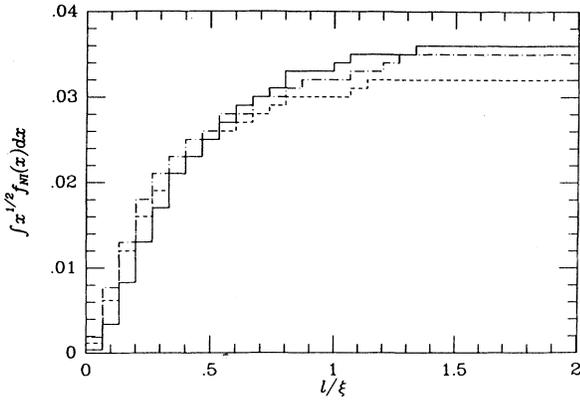
FIG. 11. The integral $\int_0^x x^{1/2} f_{NI}(x)$ (relevant in determining the final-loop distribution) is plotted against x. It is shown over several ranges of scale factor in the long radiation run.
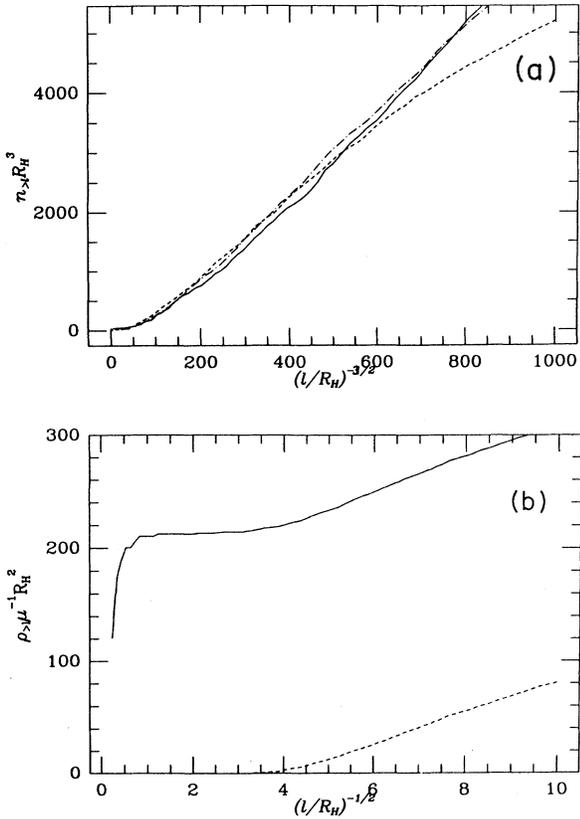

FIG. 12. The distribution of strings at fixed time in the radiation-era simulation. (a) shows the number of loops in a volume R_H^3 with length greater than l versus $(l/R_H)^{-3/2}$ at different scale factors in the long radiation run (conformal times 25, 35, and 45 go with dashed, dotted-dashed, and solid, respectively). (b) shows the energy density (in units of μR_H^{-2}) in loops of length greater than l vs $(l/R_H)^{-1/2}$. Note the small contribution of loops greater than $\xi \approx 0.07 R_H [(l/R_H)^{-1/2} \approx 3.8]$. The dashed line represents nonintersecting loops only.

the value obtained from f_{NI}. Figure 12(b) shows the energy density (in scaling units) in loops greater than l vs $(l/R_H)^{-1/2}$. It is very noticeable that there is not much energy around in loops whose length is a few ξ—the long strings lose their energy primarily into smaller loops. The explanation for this was given in Sec. IV.

The matter-era loop-density parameter λ_m is given directly from the values of γ_m and $\int dx\, f_{NI}(x)$ quoted above, and Eq. (3.18) as approximately 2.1.

In our previous paper on string simulations we worked in terms of radius r rather than rest energy (m) of the loops. For comparison with our original results, we use $\beta = m/(\mu r)$ to convert between the two. In Sec. VI we find $\beta \approx 11$ on average. Originally we defined a parameter ν giving the differential number density in loops of radius r; this is related to λ_r and β by $\nu = \lambda_r/(2\beta)^{3/2} \approx 0.1$, 10 times higher than our first calculations.[23]

In Sec. IV we presented a statistical model for $f_{NI}(x)$. [Eq. (4.6)]. We check this model in Fig. 13 where $\ln(f_{NI})$ is graphed against e/ξ. In the radiation era it is well fitted by the predicted form, as shown with the dotted line. In the matter era the statistics are poorer, and as we discussed, reconnection is expected to alter the form somewhat [recall that the reconnection term in (4.7) is
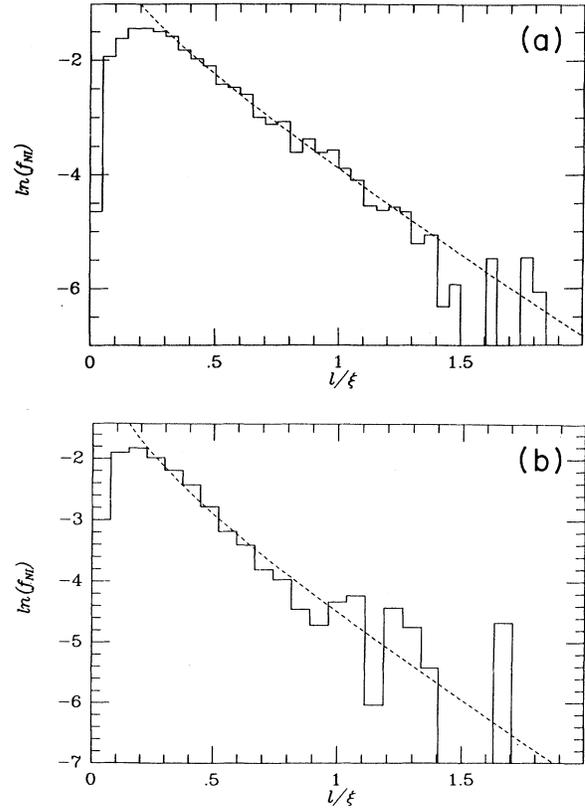

FIG. 13. The natural log of the nonintersecting loop production function (solid line) is plotted against l/ξ, and compared with a fit to the form [Eq. (4.6)] given by our statistical model (dashed line). (a) is for the radiation era, and (b) is for the matter era.

constant for small l]. For simplicity we have nevertheless fitted the result to the simple form (4.6). The fit in the radiation run is $\ln(f_{NI}) = -1.3 - 2.6x - \frac{1}{2}\ln(x)$ and in the matter run is $\ln(f_{NI}) = -2.0 - 2.5x - \frac{1}{2}\ln(x)$, so the matter-era function is smaller. Of course the predicted form only applies for $Bx \gg 1$ — in fact it fits well down to $Bx \approx 1$. Below that point, the results of York suggest f_{NI} should turn flat.

In Fig. 14 we compare the matter and radiation non-self-intersecting loop production functions directly. Their similarity in form is evident but the matter f_{NI} is definitely smaller. Rescaling the matter f_{NI} by a factor 1.55, the two functions look very similar as is shown in Fig. 14(b). The factor 1.55 can be accounted for almost entirely by the fact that the overall chopping efficiency is down by a factor $c_m/c_r = 1/1.45$ in the matter era. This should affect the overall rate of producing nonintersecting loops, and thus the overall scale of f_{NI}.

As we mentioned above, we also did a much higher resolution run in a smaller box to check for small-scale cutoff effects, mainly due to our crossing detection method. In Fig. 15 we compare the non-self-intersecting loop production function f_{NI} for this run (the average

over the entire course of the run) with that for our long radiation run calculated over a similar period. The main effect is obviously an enhancement on small scales, but as the integral of f_{NI} shows, apart from this finite correction, the functions are very similar in shape (see Note added in proof).

Finally the "thermal" model we have used to describe the energy production function can also be used to predict the form of the velocity distribution expected for loops being chopped off the network. This is given by a relativistic Boltzmann distribution

$$\bar{v}(m) = \frac{\int d^3\mathbf{p}\, e^{-\beta\sqrt{\mathbf{p}^2+m^2}}(p/\sqrt{p^2+m^2})}{\int d^3\mathbf{p}\, e^{-\beta\sqrt{\mathbf{p}^2+m^2}}}$$

$$\equiv I(\beta m) , \tag{5.1}$$

where β is the inverse effective temperature of the network. Unfortunately this cannot be deduced directly from f_{NI} because $B = (\beta - b)\xi$, as we discussed in Sec. IV. However, if we fit (5.1) to the measured velocity distribution of loops in our simulations [Figs. 16(a) and 16(b)] we find a reasonable fit for $\beta\xi = 8$ in the radiation era and 10 in the matter era. These values are indeed
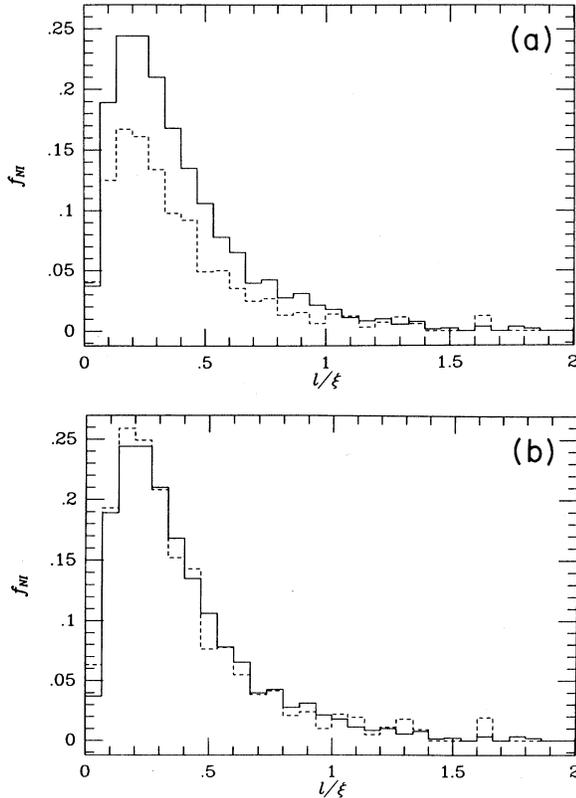

FIG. 14. Comparison of the matter (dashed) and radiation era (solid) non-self-intersecting loop production functions. Both are plotted as a function of l/ξ, and represent the average over the latter part of each simulation, when they appear to be scaling (the matter function only approximately so). (a) compares them directly, and in (b) the matter function is multiplied by 1.55, the ratio of the chopping efficiencies. After this rescaling, the functions are very similar.
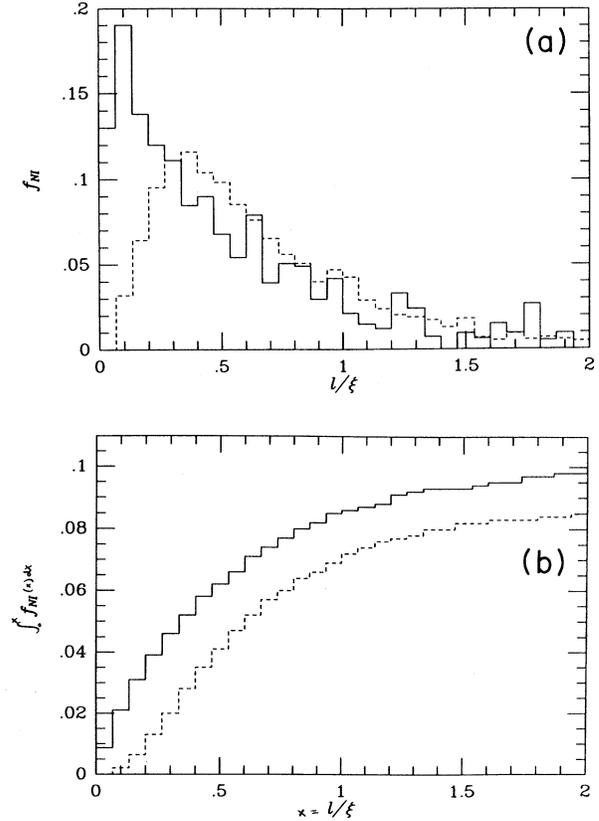

FIG. 15. The non-self-intersecting loop production function for the high resolution test run (solid) averaged over the entire length of the run (scale factor 16 to 30) compared with the radiation-era function (dashed) over the same period. (a) compares the functions directly, and (b) compares $\int_0^x f_{NI}(x)dx$ vs x.
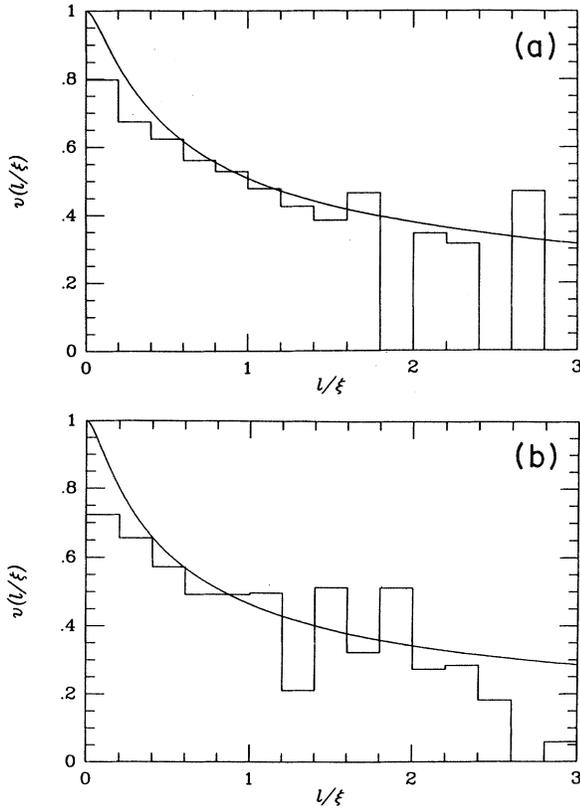
FIG. 16. The velocity distribution for loops chopped off the scaling network. Center-of-mass velocity is plotted against l/ξ. (a) shows the radiation run and (b) the matter run. The curve shown on each figure is a fit using the "thermal" model explained in Sec. V.

greater than B. For large loops, the predicted velocity scales as $m^{-1/2}$, as one would deduce naively by arguing that large loops have $N = m/\xi$ segments and the segments have uncorrelated velocities.

We now note one way in which the distribution of nonintersecting loops is nonthermal. A standard characterization of the wiggliness of a loop is β_r, which is just the ratio of l to the rms fluctuation in position of the loop (the "radius," r). We use the subscript r here to distinguish from the inverse temperature. Our "random walk" measure developed in this section would predict that β_r should increase linearly with r. What we find, however, is $\beta_r \approx 11$ *independent* of r. It appears that the requirement that loops are nonintersecting selects only the relatively "simple" ones. This result also indicates that the velocity distribution cannot be explained by the simple "uncorrelated segments" argument. The fact that β_r is independent of r means that the larger loops cannot be thought of as having more uncorrelated segments than smaller ones.

We conclude this section with some pictures from our simulations. In Fig. 17 we show boxes of string with size $(3\xi)^3$ taken from our radiation run at different times. The boxes are rescaled to be the same size, so the predic-

tion of scaling is that the boxes should all look similar to one another. Of course, as time evolves and ξ grows our simulation populates the loop distribution down to smaller values of $1/\xi$.

VI. MEASURES

The observable effects of cosmic strings depend in large part on distribution of nonintersecting loops which come off the network of longer strings. We will want to discuss such things as average sizes and velocities of the loops, but there are many different ways to take averages. In this section we discuss different ways one can put a measure on the distribution of loops, so we can then make physically appropriate choices.

Equation (3.16) defined the growth of energy density in nonintersecting loops $\rho_{NI}(l)$ as a function of time in terms of the nonintersecting energy production function, $f_{NI}(l/\xi)$. As our attention shifts from energetics to actual numbers of loops it is useful to study

$$N(l) \equiv \frac{a^3 \rho_{NI}(l)}{\mu l} . \tag{6.1}$$

Dividing by μl converts an energy density to a number density, and multiplying by a^3 turns a number density into an actual number per comoving volume. Using Eq. (3.18) one gets

$$\partial_t N(l) = \frac{a^3}{l} \frac{l}{\xi^4} f_{NI}(l/\xi) . \tag{6.2}$$

In this discussion we are labeling each loop with two labels, its length l, and the time t at which it is produced. This labeling can be made more explicit by considering $N(l, t_i)$, where $N(l,t)de\, dt$ gives the number of loops in volume a^3 with l between l and $l + dl$ and t between t and $t + dt$. Then $N(l,t)$ is simply $\partial_t N(l)$ as given above. Using $\gamma \equiv R_H/\xi$ as in Sec. III we get

$$N(l,t) = \frac{a^3}{l} \frac{\gamma^4}{R_H^4} f_{NI}\left[\frac{\gamma l}{R_H}\right] . \tag{6.3}$$

It will be useful to focus on the loops produced in the matter era, during which $a^3/R^2 = \text{const} = a_0^3/R_0^2$ and one gets

$$N(l,t) = \frac{a_0^3}{R_0^2} \frac{\gamma^4}{R_H^2} \frac{f_{NI}\left[\frac{\gamma l}{R_H}\right]}{l} . \tag{6.4}$$

One can change to any new set of variables $y(l,t)$ and $z(l,t)$ as long as the transformation is nonsingular. One then gets

$$N(y,z) \equiv N(l(y,z), t(y,z)) \left| \frac{\partial l}{\partial y} \frac{\partial t}{\partial z} - \frac{\partial t}{\partial y} \frac{\partial l}{\partial z} \right| . \tag{6.5}$$

For example, using $x \equiv l/\xi$ gives

$$N(x, R_H) = \frac{a_0^3}{R_0^2} \frac{\gamma^4}{R_H^2} \frac{f_{NI}(x)}{x \dot{R}_H} \tag{6.6}$$

and

$$N(x,l) = \frac{a_0^3}{R_0^2} \frac{\gamma^3}{l^2} \frac{f_{NI}(x)}{\dot{R}_H} . \qquad (6.7)$$

As an illustration of the averaging procedure, let us calculate the average value of x for loops of a particular length. This average is given by

$$\bar{x}(l) \equiv \frac{\int_0^\infty x N(x,l) dx}{\int_0^\infty N(x,l) dx} , \qquad (6.8)$$

which reduces to

$$\bar{x}(l) = \frac{\int_0^\infty x f_{NI}(x) dx}{\int_0^\infty f_{NI}(x) dx} = 0.4 , \qquad (6.9)$$

whereas the average value of x for loops produced at a given time (or value of R_H) is

$$\bar{x}(R_H) \equiv \frac{\int_0^\infty x N(x,R) dx}{\int_0^\infty N(x,R) dx}$$

$$= \frac{\int_0^\infty f_{NI}(x) dx}{\int_0^\infty x^{-1} f_{NI}(x) dx} = 0.2 . \qquad (6.10)$$

For both averages the numerical values come from evaluating integrals over f_{NI} from our string simulations.

Note that both $\bar{x}(R_H)$ and $\bar{x}(l)$ are independent of their respective arguments. This fact is ensured by scaling. However, the two averages are different because a
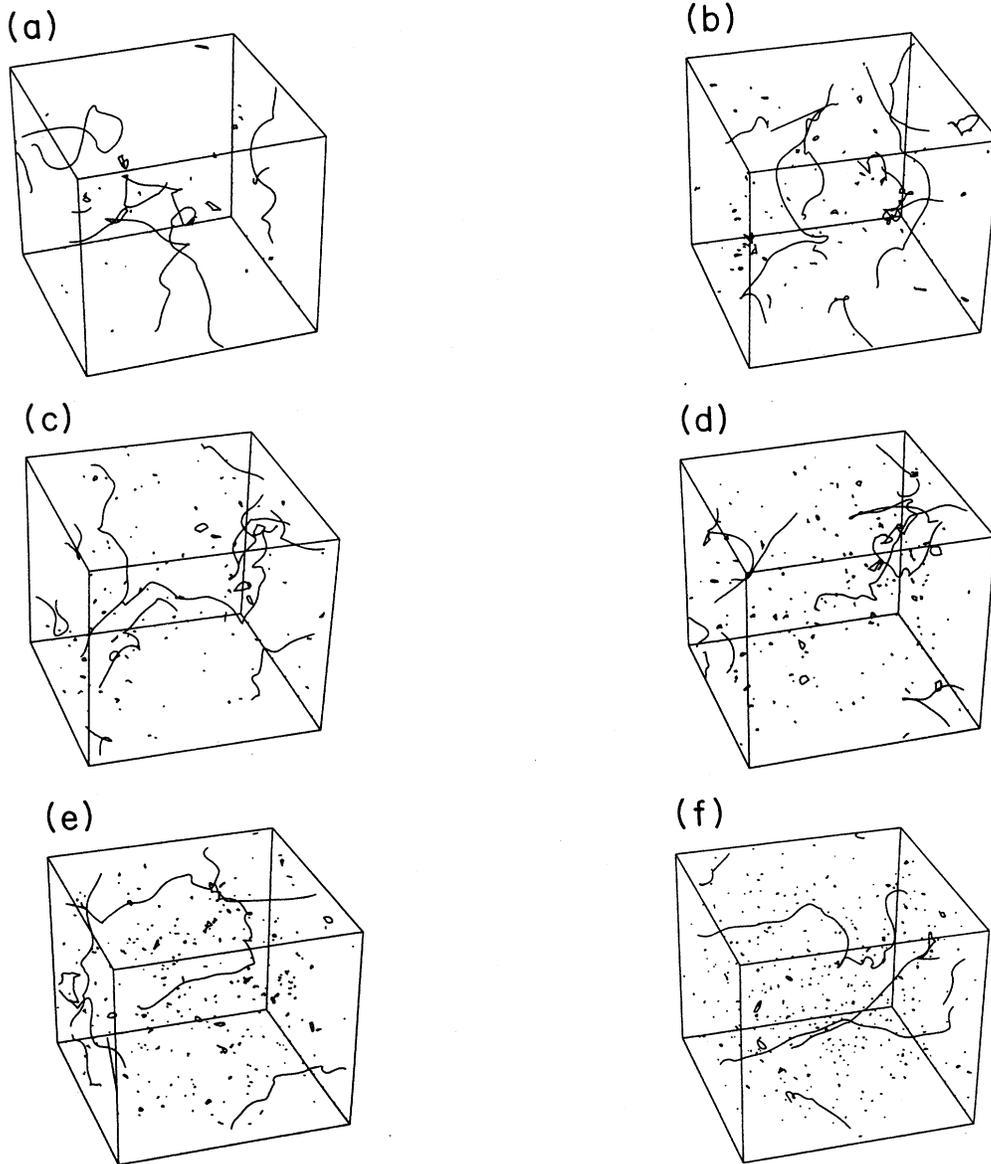

FIG. 17. A set of $(3\xi)^3$ boxes of string from our radiation-era simulations. The physical Hubble length (in units of the Hubble length at the beginning of the run) is 2.3 for (a), 3.1 for (b), 3.5 for (c), 5.3 for (d), and 6.9 for (e), and 7.6 for (f).

given loop is being averaged in with a different set of other loops in each case. The value of $\bar{x}(l)$ depends more on the behavior of $f_{\mathrm{NI}}(x)$ for larger values of x than does $\bar{x}(R_H)$. Although at a given time there are fewer loops produced at large x, the loops produced at fixed length with larger x must be produced at earlier times, when the overall densities are higher. Thus larger x loops are more important at fixed lengths than at fixed times.

We conclude this section by calculating averages of two other interesting quantities. The parameter β_r (defined at the end of Sec. IV) is the rest mass of the loop divided by its radius r, defined as the rms fluctuation in the position of the string. It is a measure of how wiggly the loop is. We find

$$\bar{\beta}(l) \equiv \frac{\int_0^\infty \beta(x)f_{\mathrm{NI}}(x)dx}{\int_0^\infty f_{\mathrm{NI}}(x)dx} = 11.4 \tag{6.11}$$

and

$$\bar{\beta}(R_H) \equiv \frac{\int_0^\infty x^{-1}\beta(x)f_{\mathrm{NI}}(x)dx}{\int_0^\infty x^{-1}f_{\mathrm{NI}}(x)dx} = 10.7 \ . \tag{6.12}$$

Similarly we calculate the average speed of the loop when it was created:

$$\bar{v}(l) \equiv \frac{\int_0^\infty v(x)f_{\mathrm{NI}}(x)dx}{\int_0^\infty f_{\mathrm{NI}}(x)dx} = 0.72 \tag{6.13}$$

and

$$\bar{v}(R_H) \equiv \frac{\int_0^\infty x^{-1}v(x)f_{\mathrm{NI}}(x)dx}{\int_0^\infty x^{-1}f_{\mathrm{NI}}(x)dx} = 0.75 \ . \tag{6.14}$$

Note that there do not seem to be significant differences between the two types of averages in these cases. A simple explanation is that most loops with $l \approx \xi$ are very similar up to an overall scale. They just are the simplest possible nonintersecting loops there can be. It seems reasonable however, to expect the "effective ξ" for a given small loop to vary, as a manifestation of fluctuations around the average ξ for the network. Quantities such as β and v would not depend on an overall scale, but only on the wiggliness of the loop, which determines the relative size of the radius, and how coherently the velocities of individual bits of string add up into net motion of the loop. Thus our two averages, one centered on loops with $l \approx 0.4\xi$ and the other on loops with $l \approx 0.2\xi$ give similar results.

VII. ACCRETION OF MATTER ONTO STRINGS

Cosmic strings can initiate gravitational collapse in an initially homogeneous distribution of matter. It has been suggested that a scaling distribution of a cosmic string might provide sufficient perturbation on an otherwise homogeneous distribution of matter to account for all the structure we observe in the Universe today.[2,24,37] Preliminary calculations were encouraging, but hitherto the lack of a solid understanding of the scaling network has made precise calculations difficult. While there are still

uncertainties, we feel they are now small enough to justify the detailed calculation of large-scale structure based on our string simulations.

We shall not attempt this here. Instead we shall make a few preliminary estimates to indicate how suitable the real scaling network might be for the formation of observed large-scale structure.

Much initial work on this subject was based on the one-loop–one-object hypothesis.[24,28] One assumes that individual nonintersecting loops seed the collapse of objects which are still distinct today. Larger loops form larger objects—there would be complete correspondence between loops of string and observed objects. There are several ways in which the one loop-one object hypothesis may fall short, but as we shall see, it is still a useful starting point for the discussion.

Let us first consider very-large-scale structure, formed by loops produced after the matter-radiation transition. It is well established that the mass M accreted on a loop with length l (rest mass divided by μ) laid down in the matter era is given by

$$M = \alpha \mu l (1+z) \ , \tag{7.1}$$

where z is the red-shift at which the loop is laid down, and α is a factor which is generally of order unity. For example, in the spherical collapse model around a stationary point mass, if M is defined as the mass which has reached its final virial radius, $\alpha = 0.38$. If M is defined as the mass "turned around" then $\alpha = 0.57$ (Ref. 39). This formula is true in the case where the Universe is flat and the dark matter is cold. For hot-dark matter it is also true for masses $M \gg M_J \approx 1.5 \times 10^{14} h_{50}^{-4} M_\odot$, the neutrino Jeans mass.[40] Equation (7.1) can be written as

$$M = \mathcal{M} \frac{x}{\gamma_m}(1+z)^{-1/2} = \mathcal{M}\frac{x}{\gamma_m}\left[\frac{R_H}{R_0}\right]^{1/3} , \tag{7.2}$$

$$\mathcal{M} \equiv \alpha \mu R_0 \approx \alpha \mu_6 h_{50}^{-1} 10^{17} M_\odot \ , \tag{7.3}$$

and R_0 is the current Hubble length ($R_0 = h_{50}^{-1} \times 6000$ Mpc), and μ_6 (typically ~ 1 for grand unified strings) measures μ in units of $10^{-6}G^{-1} \approx 2.1 \times 10^7$ $M_\odot \text{pc}^{-1}$. In what follows we shall set $\mu_6 = \alpha = h_{50} = \Omega = 1$.

We can use Eq. (7.2) and the techniques of Sec. VI to label loops according to the mass they accrete, and to arrive at

$$N(M,x) = \frac{3a_0^3}{\dot{R}_H R_0^3}\frac{1}{M}\left[\frac{\mathcal{M}}{M}\right]^3 \gamma_m x^2 f_{\mathrm{NI}}(x) \ . \tag{7.4}$$

Furthermore, one can construct

$$N(M) = \int_0^\infty N(M,x)dx \ . \tag{7.5}$$

Dividing out by a_0^3 (a_0 is the scale factor today) given the number density of objects of mass M today:

$$n(M) = \frac{1}{R_0^3}\frac{1}{M}\left[\frac{\mathcal{M}}{M}\right]^3 2\gamma_m \mathscr{F}_2 \ , \tag{7.6}$$

where

$$\mathscr{F}_n \equiv \int_0^\infty x^n f(x)dx \ . \tag{7.7}$$

The dependence of $n(M)$ on $\mathscr{F}_2$ means that it depends more on the behavior of the energy production function at larger values of x. For example, one can calculate $\bar{x}(M)$ in the same manner as $\bar{x}(l)$ and $\bar{x}(R_H)$ were calculated in Sec. VI:

$$\bar{x}(M) = \frac{\mathscr{F}_3}{\mathscr{F}_2} = 0.9 \approx 2\bar{x}(l) \approx 4\bar{x}(R_H) \ . \tag{7.8}$$

Even though at any given time most loops are produced with $l < \xi$ (that is, $x < 1$), as far as M is concerned these smaller loops are grouped together with loops that were produced earlier and at larger x. Because the overall density of string is higher at earlier times the large x loops dominate for any fixed M. This behavior is good for string simulations, since it will always be the low-x regime which is most susceptible to finite resolution effects. This emphasis on larger x is most relevant on large scales, where the translation from perturbations to observable structure is least ambiguous anyway, and means that quite precise calculations of the predicted structure should be possible.

Now we turn to smaller scales. A simple estimate of the virialized mass accreted on loops which appeared in the radiation era with cold-dark matter is

$$M = \alpha \mu l (1 + z_{\text{eq}}) \ , \tag{7.9}$$

where $z_{\text{eq}} \approx 6000$ is the redshift at equal matter and radiation density. In the spherical collapse model, $\alpha \approx 0.95$.

For these loops $n(M)$ can be calculated from (3.19) by a simple change of variables [we ignore the small extra numerical factors obtained by evolving (3.19) through the transition era precisely]. The result, evaluated for today is

$$n(M) = \frac{1}{R_0^3} \frac{\lambda_r}{M} \left[\frac{\mathscr{M}}{M} \right]^{3/2} (1 + z_{\text{eq}})^{3/4} \ . \tag{7.10}$$

The mean separation of objects of mass greater than M is

$$d_{>M} = \left[\int_M^\infty n(M) dM \right]^{-1/3} \ . \tag{7.11}$$

Using $n(M)$ from Eq. (7.6) for loops laid down in the matter era, Eq. (7.11) gives

$$d_{>M} = R_0 \frac{M}{\mathscr{M}} (\gamma_m \mathscr{F}_2)^{-1/3} \tag{7.12}$$

and plugging in Eq. (7.10) for loops put down in the radiation era gives

$$d_{>M} = R_0 \left[\frac{M}{\mathscr{M}} \right]^{1/2} (\tfrac{2}{3}\lambda_r)^{-1/3}(1 + z_{\text{eq}})^{-1/4} \ . \tag{7.13}$$

The smallest loops around at t_{eq} accrete masses of about

$$M_{\text{min}} \approx 10^{11} M_\odot \ . \tag{7.14}$$

Of course we have ignored the important transition era. It would be straightforward to use our numerical solution for γ through the transition, the form of the energy production function, and the exact growth formula for seeds laid down during the transition, to solve for $d_{>M}$ for all M. In fact this is necessary for scales corre-

sponding to galaxies. The mean separation of bright galaxies is $d \approx 10 h_{50}^{-1}$ Mpc. If we assume the corresponding loops were formed in the radiation era, using (3.19), these have lengths $l \approx 0.1 R_H \approx \xi$ where ξ is the scale on the network at equal matter and radiation density. This contradicts our calculation $\bar{x}(l) \approx 0.4$, and implies that these loops were produced in the transition era. In fact according to our new results, both galaxy- and cluster-forming loops were produced *after* equal density.

We plot $d_{>M}$ in Fig. 18. The plot corresponds to $\mu_6 = h_{50} = \alpha = 1$, and the solid line corresponds to cold-dark matter. We have simply extrapolated the large M and small M results to where they meet. The kink in the curve corresponds to the radiation-matter transition, and we expect that corrections to our crude analysis will smooth it out.

From the figure, the mean separation of objects with a mass of $10^{15} M_\odot$ (the typical mass of richness class 1 clusters) is around 100 Mpc, just what is observed for these clusters. On smaller scales the story gets more complicated—the loops have caused all the matter in the Universe to go nonlinear. One can estimate this critical scale M_c by solving

$$M_c n_>(M_c) = \rho_{\text{tot}} = \frac{3}{8\pi} G^{-1} R_H^{-2} \ , \tag{7.15}$$

which gives

$$M_c \approx 10^{14} M_\odot \ . \tag{7.16}$$

One would expect that for masses below M_c the correspondence between loops and individual objects would be lost. Figure 18 may also indicate that there may be too many objects around with masses of order $10^{14} M_\odot$.

If the dark matter is hot then the accretion is suppressed on scales $M < M_J \approx 1.5 \times 10^{14} h_{50}^{-4} M_\odot$, the
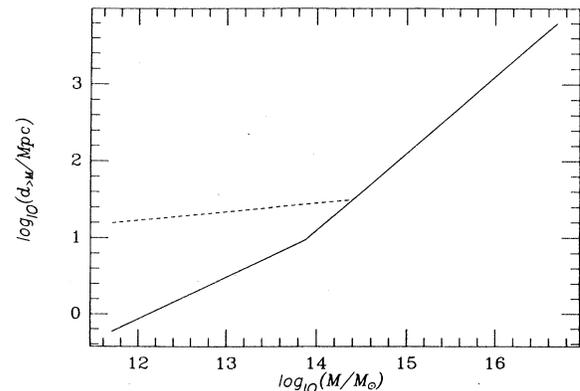

FIG. 18. The mean separation $d_{>M}$ of objects of mass greater than M in the one-loop–one-object picture. The solid line is the cold-dark-matter calculation, and the dashed line shows the correction on small scales if the dark matter is hot (a single massive neutrino making up $\Omega = 1$). d is defined as the inverse cube root of the number density. The cold-dark-matter curve cannot be taken literally for small M because merging is important, as explained in the text.

neutrino Jeans mass at equal density. Instead of (7.1) we have found the virialized mass M:

$$M = \frac{\alpha(\mu l)^3(1+Z_{eq})^3}{M_J^2} , \qquad (7.17)$$

where $\alpha = 0.47$ in the spherical collapse model. In this case, accretion only starts on a scale M when the Jeans mass has fallen to that scale. The final distribution of low-mass objects is now determined by the loop density in the matter era. We find

$$d_{>M} = R_0 \lambda_m^{-1/3} \left[\frac{M}{M_J} \right]^{1/9} \left[\frac{M_J}{\mathcal{M}} \right]^{1/3} (1+Z_{eq})^{-1/3} .$$

$$(7.18)$$

The result is plotted in the dashed line in Fig. 18. Now objects with the mean separation of galaxies have much lower masses than with cold-dark matter. There is far less merging in the hot-dark matter model, and the one-loop–one-object picture should be more reasonable.

Because of the initial relativistic velocities of the loops, the matter they initially accrete forms in a pencil-like wake behind them.[41] We have ignored this so far. It is interesting to know the length of these pencils compared to their mean separation. If this number is large, one expects a large amount of interference between different pencils, which would reduce the time over which the one-loop–one-object correspondence is maintained. When a nonintersecting loop breaks off the network it has some speed v_i which redshifts with time. The loops produced in the radiation era have had ample opportunity to slow down before gravitational collapse sets in. Loops produced in the matter era when the Hubble length is R_i travel a distance $d_t = 2R_i v_i$ (in the Newtonian approximation) before stopping. To gauge the effect of this motion on the mass distribution we compare the comoving distance (d_t) the loop travel with the mean separation of objects with mass $> M$:

$$\frac{d_t}{d_{>M}} = 2v \mathcal{F}^{-1/3} x^{-1} \qquad (7.19)$$

so

$$\left\langle \frac{d_t}{d_{>M}} \right\rangle \Bigg|_M = 2\gamma_m^{4/3} \mathcal{F}_2^{1/3} \frac{\mathcal{F}_1}{\mathcal{F}_2} \bar{v} \approx 5\bar{v} . \qquad (7.20)$$

where

$$\bar{v} \equiv \frac{1}{\mathcal{F}_1} \int_0^\infty v(x) x f_{NI}(x) dx = 0.7 . \qquad (7.21)$$

The upshot is that individual pencils are not a whole lot longer than their mean separation. One should only get serious mixing of the pencils of a given mass when their corresponding mass density approaches the total mass density of the Universe.

It is also possible that these pencils fragment to form several objects rather than one. This would only alter our calculations by an overall factor—the scaling with M would be unaffected.

Another very important factor which we have also neglected is the wakes of the long string. Because the

scaling network has more long string inside a Hubble volume than was originally expected, wakes of long string will play a more important part in perturbing the surrounding matter. A very crude estimate of the magnitude of the perturbations induced by the long strings is obtained by multiplying the fractional density perturbation provided by the long strings in the radiation era, $\rho_L/\rho_{rad} = \gamma^2(8\pi/3)G\mu \approx 1.6 \times 10^{-3}$, by the linear growth factor for perturbations produced in the radiation era, $\frac{3}{2}(1+Z_{eq})$. We find $\delta\rho/\rho \approx 10$ today, on a length scale of $\xi \approx R_H/10$ at the matter-radiation transition. This corresponds to a comoving scale of about $5h_{50}^{-2}$ Mpc today. The wakes produced by long strings could well be a significant feature in the distribution of galaxies today. We shall leave this issue for later investigation.

We conclude that at first glance the prospects for cosmic strings to play an important role in structure formation remain very good, despite changes in our understanding of the scaling network. A more thorough analysis, using the string from our simulation as sources for gravitational instability is underway.[42]

VIII. OTHER OBSERVABLE IMPLICATIONS

One of the most promising aspects of the cosmic-string theory is its unique set of observable predictions. We have already briefly discussed the consequences for large-scale structure, but in this section we will focus on some of the even more specific signatures of cosmic strings—the gravity-wave background, lensing, and microwave-background anisotropies.

The microwave-background anisotropies produced by strings are potentially their clearest signature.[43,44] Stebbins has performed detailed calculations of the distortions produced by the direct effects of the strings themselves,[45] and with Bouchet and Bennett has calculated the expected level of anisotropy from realistic string simulations. We have little to add to this here, but note that our *matter* scaling density is only a factor of 2 above the density they used, and this difference is small compared to the uncertainties in the calculation—in particular the standard redshift of last scattering may be significantly altered in the presence of strings.[46] The limits these authors obtain from detailed comparison with observations is $G\mu < 5 \times 10^{-6}$, so as yet the constraint is quite weak. The precise nature of the pattern, the magnitude of the Sachs-Wolfe, and other indirect effects on the microwave background are all in principle calculable from our simulation results. The expected constraints from these effects are probably similar however.

We have already discussed the importance of loop decay into gravity waves—this sets the scale of the smallest typical loops around at any moment in the Universe. The gravity-wave background in observable periods today is a result of many loops, so it is expected to be described by Gaussian statistics with a characteristic power spectrum. The observationally relevant gravity waves were produced in the radiation era, and for these the power spectrum has a very simple form, which follows directly from the scaling solution—there is equal power in each logarithmic frequency interval.[47] The amplitude is set by the

parameter $G\mu$, and the bounds provided by the millisecond pulsar timing observations[48] constrain $G\mu$ strongly as we shall see (Ref. 49).

A string loop radiates energy into gravity waves at a rate $\dot{E} = -\Gamma G\mu^2$ with Γ a constant which depends on the precise loop trajectory but for typical trajectories $\Gamma \approx 50$ (Refs. 22 and 27). The frequencies of the waves emitted are integer multiples N of the loop's inverse period $2\mu/E$, where E is its energy. Most of the energy is emitted in waves whose period is less than $\frac{1}{100}$ or so of the loop's period. Waves observed today with a period of $T \approx 1$ yr were therefore emitted at a time t_e given by 1 yr $Z_{eq}^{-1}(t_e/t_{eq})^{1/2} \approx \Gamma G\mu t_e/(2N)$ obtained by red-shifting the wave back and equating its period to that of the waves emitted by the smallest typical loops at that time. Using $Z_{eq} \approx 10^4$ and $t_{eq} \approx 10^{12}$ s we find $t_e \approx 10^4 N^2$ s, well before matter-radiation equality. So we need to focus on the gravity waves emitted during the radiation-era scaling solution.

The spectrum of gravity waves is easy to calculate in the scaling solution.[27] Using the number density in loops (3.19) we find that the total number density of loops in the scaling solution is $n = \lambda_r/[3\sqrt{2}(\Gamma G_\mu)^{3/2}t^3]$. Each loop radiates at a rate $\Gamma G\mu^2$. The fact that the whole loop distribution is scaling tells us the spectrum must be scaling too: the energy density emitted in frequencies ω_e to $\omega_e + d\omega_e$ in the time interval t_e to $t_e + dt_e$ is $n\Gamma G\mu^2 dt_e g(\omega_e t_e)d\omega_e/\omega_e$ where $g(x)$ is a dimensionless function whose integral $\int dx\, g(x)/x$ must be unity. At some time t later this energy density is redshifted by $(t_e/t)^2$. Now, just as in our discussion of measures in Sec. VI, we change variables to the current frequency $\omega = \omega_e(t_e/t)^{1/2}$ (assuming t is still in the radiation era) and integrate to find the energy density in waves from ω to $\omega + d\omega$ at some much later time:

$$\rho(\omega)d\omega = \frac{\lambda_r}{3\sqrt{2}(\gamma G\mu)^{1/2}} \int \frac{dt_e}{t_e^3} \frac{t_e^2}{t^2} g(\omega(tt_e)^{1/2}) \frac{d\omega}{\omega} \ . \quad (8.1)$$

Changing variables to $x = \omega(tt_e)^{1/2}$, integrating, and dividing by the density in radiation $\rho_{rad} = 3/(32\pi Gt^2)$ we find

$$\frac{\omega\rho(\omega)}{\rho_{rad}} = \frac{64\pi}{9\sqrt{2}} \frac{(G\mu)^{1/2}\lambda_r}{\Gamma^{1/2}}$$

$$= 2.2 \times 10^{-2}\lambda_{r10} \frac{(G\mu_6)^{1/2}}{\Gamma_{50}^{1/2}} \ , \quad (8.2)$$

where $\lambda_r = 10\lambda_{r10}$, $G\mu = 10^{-6}G\mu_6$, and $\Gamma = 50\Gamma_{50}$. This derivation makes it clear that the ω^{-1} dependence is a consequence of scaling. Because the waves of interest were emitted well before t_{eq} it is completely unnecessary to discuss the precise frequency distribution emitted by each loop. Present tentative limits from millisecond pulsar timing [D. Steinbring and J. Taylor (private communication)] indicate that (8.2) is constrained to be less than 0.18 for frequencies of 0.68 yr^{-1}, 0.006 for frequencies of 0.25 yr^{-1}, and 0.018 for frequencies of 0.16 yr^{-1}. (We take the number of light neutrinos $N_\gamma = 3$.)

At first sight, with our new parameters the cosmic-string scenario with $G\mu = 10^{-6}$ is clearly ruled out. How-

ever this conclusion is not yet fully justified. There are several small effects that could decrease (8.2) by a few: the result is dominated by the smallest, longest surviving loops, if loops split up after they have lost a fraction of their energy this could easily decrease (8.2) by a factor of 2 or so, and there could be additional energy-loss mechanisms operating—annihilation of regions of the string near cusps, for example. Note that we have already included the loss of energy due to the redshifting of loops' velocities by defining γ_r using the rest mass of loops.

The greatest uncertainty in the limit comes from the accuracy of our simulations—the value for λ_r from our simulations is sensitive to the small-scale resolution of the simulations and is probably too high (see Sec. V). However there really is very little room to maneuver, and a modest improvement in the limits could convincingly rule out $G\mu > 10^{-7}$, which is the minimum required for the simplest gravitational accretion cosmic-string scenario (see Sec. VII).

The gravity-wave background from strings leads to another limit on $G\mu$, from nucleosynthesis.[11,50] For the standard nucleosynthesis scenario to work, the total density in gravity waves is constrained to be less than 18% of the total density at that time. A minor complication is that as the temperature falls through low-mass particle thresholds, the photons are reheated while the gravity waves are not. This effectively dilutes the density in gravity waves, by a factor of ≈ 2 if we only include the known quarks, leptons, and gauge particles (see, e.g., Ref. 11). Integrating $\rho(\omega)$ in (8.2) from frequencies emitted at the time when strings started moving freely $\omega \approx t_{GUT}^{-1}$ down to those emitted at the time of nucleosynthesis $\omega \approx t_{nuc}^{-1}$ and using $\ln(t_{nuc}/t_{GUT}) \approx 60$ we find that the density in gravity waves at nucleosynthesis is given by

$$\frac{\rho_{gw}}{\rho_{rad}} = 0.6\lambda_{r10} \frac{(G\mu_6)^{1/2}}{\Gamma_{50}^{1/2}} \ . \quad (8.3)$$

We see that the string scenario with our new parameters $G\mu = 10^{-6}$ is again past the edge of the nucleosynthesis limit. The uncertainties, which all go in the direction of weakening the bound, could produce the factor of 3 required to make the theory acceptable. Unfortunately this constraint is not likely to become any tighter.

The second direct observational test is gravitational lensing by strings.[51] Recently Cowie and Hu have reported a candidate event for lensing by a string loop[52]—four double galaxies with angular separations of approximately 2.5 arcseconds in a region approximately 40 arcseconds square. Our numerical results enable us to ask whether such events should occur with a reasonable probability.

The most numerous loops around today are those about to disappear into gravity waves, of length $l_c = \Gamma G\mu t_0 \approx 200$ kpc. Such loops at a redshift z would have an angular circumference of $\approx 5z^{-1}$ arcseconds. From our matter-era loop-production function, we calculate that they were produced at a red-shift of ≈ 200, in the matter scaling regime. Thus the number density is given to a good approximation by (3.20).

Strings produce a double image with separation[51]

$$\delta\phi = \frac{r-d}{r} 8\pi G\mu \sin\theta R \ , \tag{8.4}$$

where d is the distance from observer to string, and r the distance from observer to object. R is a relativistic factor equal to $(1-v^2)^{1/2}/(1-\mathbf{n}\cdot\mathbf{v})$ where $\mathbf{n}$ is the unit vector along the line of sight and $\mathbf{v}$ the string velocity. Expressed in angular units, the typical lensing angle is $\delta = 4\pi G\mu = 2.6(G\mu)_6$ arcseconds. A string loop of length l lenses everything behind it within an angle $\delta\phi$. Thus the fraction of the whole sky which is lensed by string as we go out a distance r is given by

$$F = \frac{1}{4\pi} 8\pi G\mu \frac{\pi}{4} \int_0^r dr' 4\pi r'^2$$

$$\times \int_{l_c}^{R_H/\gamma} dl \ n(l)$$

$$\times \int_0^{r'} dd \frac{l}{\sqrt{2}d} \frac{3d^2}{r^3} \frac{r-d}{r} \ , \tag{8.5}$$

where we integrate out in r' and average the loop distribution over d. The loop distribution is cut off at $l\approx\xi\approx R_H/\gamma_m$. The factor of $\sqrt{2}$ arises in translating loop energy into spatial length and the factor of $\pi/4$ in averaging over $\sin\theta$. We ignore the relativistic factor, which is of order unity. The result is

$$F \approx 4\pi G\mu \ln(1/8\pi G\mu)\frac{\lambda_m}{\sqrt{2}} \left[\frac{r}{R_H}\right]^2$$

$$\approx 2\times 10^{-4} Z^2 \lambda_{m2} G\mu_6 \ , \tag{8.6}$$

where Z is the redshift depth of the sample, and for $Z<1$ we use $r\approx ZR_H$.

Note that since $n(l)\propto l^{-2}$ in the matter era, loops in each logarithmic interval of l contribute equally to F. Taking $G\mu = 10^{-6}$, the *number* of loops with length greater than l contributing to (8.6) is similarly calculated to be $\approx 2.10^5 Z^3 (l_c/l)\lambda_{m2}$.

Long strings would lens an additional fraction

$$F \approx \frac{1}{4\pi} 4\pi G\mu \frac{\pi}{4} \frac{\rho_L}{\sqrt{2}\mu} \int_0^r 4\pi r'^2 dr' \frac{1}{r'}$$

$$\approx 4\pi G\mu \frac{\gamma_m^2}{\sqrt{2}} \left[\frac{r}{R_H}\right]^2$$

$$\approx 10^{-4} Z^2 \gamma_{m4}^2 G\mu_6 \tag{8.7}$$

a slightly smaller fraction of the sky than that lensed by loops. For a sample extending out to $Z=0.5$, the spatial length in long string would be approximately $Z^3\gamma_m^3\xi/\sqrt{2}\approx 6\xi$, and the angular scale subtended by a segment ξ would be approximately $1/(\gamma_m Z)\approx 30°$. So the long strings would appear quite straight.

Thus out to a redshift of 0.5 almost 10^{-4} of the sky is lensed. There are approximately 5000 galaxies per square degree with R magnitudes greater than 22.5 (which Cowie and Hu estimate to be a reasonable requirement for galaxies to be recognizable as twinned). We therefore expect one lensed galaxy per 2 square degrees on the sky. A typical plate covers approximately 10^{-2} square de-

grees, so several hundred are needed to observe a single lensed galaxy.

What would the distribution of lensed galaxies look like? *A priori* a lensed galaxy is equally likely to be associated with loops in each logarithmic interval of the range $l_c < l < R_H/\gamma$, a range of 5000 in length. The small loops, however, will typically lens a single object—using the fact that the typical configuration is where the loop's redshift is half of the object's, in a survey of depth $Z=0.5$ a loop of length l lenses on average $0.02l/l_c$ galaxies. Thus most galaxy pairs lensed by loops in the interval $l_c < l < 100l_c$ would be isolated. However most lensed galaxies would obviously come in groups lensed by the same string. If galaxies were randomly distributed then one would have to go an angular distance of approximately 1000 arcseconds along the string to see the next galaxy pair (given by multiplying 2.5 arcseconds by θ and requiring that this solid angle contain a galaxy). There is significant galaxy clustering at small separations, but for deep redshifts this is quite small (being washed out by random galaxies along the line of sight). Cowie and Hu's event seems rather fortuitous from this point of view— Typically the lensed galaxies would be far more widely separated. So the configuration they find is rather unexpected, but a more detailed analysis is required to put a precise figure on this.

Cowie and Hu have seen a single event involving four "lensed" galaxies in "several hundred" plates, which is also on the fortunate side, since one expects one per 200 plates from the above, but not unreasonably so. As they note, these can be explained by a string loop at $Z\approx 0.07$ with $G\mu = 10^{-6}$. Cowie and Hu, and Bennett and Koo, are undertaking a deeper redshift survey in the same region to look for further double images along the "lines" between their lensed galaxies. It remains to be seen whether further searches will reveal more events.

IX. CONCLUSION

In the course of this paper we hope to have convinced the reader that the evolution of a cosmic-string network is both an interesting and solvable problem in nonequilibrium statistical mechanics.

It is also important—the existence (or proven absence) of cosmic strings would have far-reaching implications for our understanding of grand unification physics, and the nature of the Universe at very early times. And cosmic strings could also provide a firm foundation for our understanding of large-scale structure in the Universe.

Our model for the scaling density, and our statistical discussion of the form of the loop production function, provide the necessary tools with which to follow the string network through the transition from radiation to matter domination. This will be essential for any precise calculation of the final effects of the string network. The heuristic picture of the string network as a hot radiating body provides a good description of the network on all length scales.

We have briefly discussed the consequences of our new results for large-scale structure, and found that as far as

predicting the observed mass spectrum of objects, the cosmic-string scenario with $G\mu = 10^{-6}$ is still in good shape. The cold-dark-matter scenario is now complicated, however, by merging of low-mass objects. The hot-dark-matter scenario seems to fit the mass function from galaxies up to clusters quite well in the one-loop–one-object picture. The long strings are likely to be quite important, however, and we defer a discussion of their effects to future work.[42]

The correlation properties of string-induced perturbations are now within reach of precise calculation. While our simulations confirm the existence and approximate magnitude of the r^{-2} loop-loop correlations originally found by one of us,[10] there is clearly more work ahead in translating them into a precise observational statement. We are optimistic that the result will soon be available.

We have discussed in some detail the other observational predictions and constraints on the cosmic-string theory. In particular we discussed the microwave background, the millisecond pulsar timing constraint, the nucleosynthesis constraint, and gravitational lensing. The $G\mu = 10^{-6}$ scenario is very close to being convincingly ruled out by pulsar timing, and is only marginally consistent with the nucleosynthesis constraint. The millisecond pulsar constraint is the most likely to rule out (or confirm) the theory in the near future. As we have emphasized, precise calculations of the string parameters are crucial in this.

Note added in proof. Since this paper was submitted we have performed exact string simulations for strings in flat spacetime to check the scaling model we present here. The model works very well, but the value of the chopping efficiency c is higher (by less than a factor of 2) than in the simulations reported here. From this, and considering the effect of extrapolating the form of our loop production function to smaller scales (below our resolution), we estimate that our value for c could be too low by as much as a factor of 2. If we make the simplistic assumption that all moments of the energy production functions $f(x)$ and $f_{NI}(x)$ have the same uncertainties, then we find that γ_r could be too high by a factor of 2, λ_r too high by a factor of $2^{3/2} \approx 3$, γ_m could be too high by $2^{1/2}$, and λ_m by a factor $2^{1/2}$. In fact higher moments are less dependent on small x and so should have lower uncertainties. We caution also that there could be systematic uncertainties on the scale of the resolution which would alter these estimates.

ACKNOWLEDGMENTS

APPENDIX: NUMERICAL METHODS

In this section we give some details of the methods we have used to form and evolve the strings and detect their crossings. We have run our program on a VAX 8600. A typical 26^3 run takes around 50 hours of CPU time.

We choose initial conditions by throwing down phases for the string-forming field at random on a cubic lattice and checking edges for the presence of strings.[4] These initial conditions are slightly artificial—the string is static and has 90 degree kinks on it. We believe this is not important—the string starts moving at relativistic velocities very soon, and all indications are that within a time of the order of a few correlation lengths the network approaches the scaling solution.

To evolve the strings, as we emphasized in Sec. II, we use our new nonsingular variables with the gauge fixed throughout. We discretize the string in σ but the positions and velocities of the points on the string are continuous. We actually evolve the "left-movers" l, "right-movers" r, and ϵ of Eq. (2.14). Our large runs are done in cubes of 26 initial correlation lengths on a side with 10 points on the string per initial correlation length. We use periodic boundary conditions throughout.

The program actually stores the positions $\mathbf{x}(\sigma_k)$ of the points σ_k marked with crosses in Fig. 19 and the momenta $\mathbf{u} = \epsilon \dot{\mathbf{x}}$ at the points half-way between, marked with circles. $\mathbf{x}'$ is calculated from the difference between the positions at neighboring crossed points divided by $d\sigma$, which is fixed (typically at $\frac{1}{10}$ of an initial correlation length) from the beginning. Thus left and right movers, and $\epsilon = \sqrt{\mathbf{u}^2 + \mathbf{x}'^2}$ are defined at the circled points.

In flat spacetime the evolution is trivial. If ϵ is uniform along the string initially, it remains so. The left mover at $k' - 1$ simply equals the left mover at k, and the right mover at $k' + 1$ equals the right mover at k.

In the expanding Universe case this is no longer quite true. Nevertheless, the corrections are small, and we wish to remain close to the exact flat-space solution.
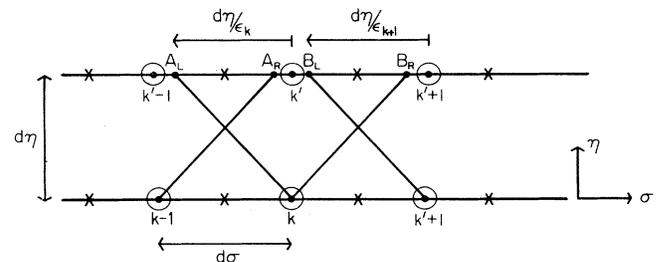

FIG. 19. The numerical string world sheet. At a given value of the conformal time (η) the positions in space $\mathbf{x}(\sigma_k)$ of a discrete set of points on the string labeled by parameter values σ_k, $d\sigma$ apart, are stored. These points are labeled with crosses in the figure. The velocities $\dot{\mathbf{x}}$ and tangent vectors $\mathbf{x}'$ are stored at the points half-way between, labeled with circles. To update the left mover at k one first calculates the left mover at A_L and B_L and then interpolates on a unit sphere to find the left mover at k'. Similarly the right movers at A_R and B_R are calculated and then interpolated between. Thus one recovers the velocity and tangent vector at k'.

First, ϵ is no longer uniform. To respect the Courant condition $d\eta < \epsilon \, d\sigma$ required for stability (essentially the information used should be sufficient to fill the backward light cone of any new point) we must choose our time step according to the minimum value of ϵ on the entire network. This does not include small loops, which are evolved separately as we will explain. One could choose different time steps for different loops, but we have not found this necessary. We find that the minimum value of ϵ evolves approximately as $a^{-1.5}$, where a is the scale factor.

We proceed by first calculating the left mover at the points A_L and B_L in Fig. 3. This is sensible because in the Taylor expansion of $l(\sigma - d\eta/\epsilon, \eta + d\eta)$, an infinite series of terms in $l', l'', l''', \ldots$ cancel; ignoring derivatives of ϵ- and h-dependent terms we have

$$l(\sigma - d\eta/\epsilon, \eta + d\eta)$$

$$= l + d\eta \, \dot{l} + \frac{1}{2}\frac{d\eta^2}{\epsilon^2}l'' + \frac{d\eta^2}{2}\ddot{l} - \frac{d\eta^2}{\epsilon}\dot{l}' + \cdots$$

$$= l(\sigma, \eta) \qquad (A1)$$

using $\dot{l} = l'/\epsilon$ to reexpress all η derivatives as σ derivatives. Of course this is just a reflection of the flat-space solution.

Including all terms up to $d\eta^3$ from the full equation for $\dot{l}$ we find

$$l(\sigma - d\eta/\epsilon/\eta + d\eta)$$

$$= l + d\eta[-h\mathbf{r} + h(l \cdot \mathbf{r})l]$$

$$+ \frac{d\eta^2}{2}\left[Bl + Cl' + D\mathbf{r} - 2h\frac{\mathbf{r}' \cdot l}{\epsilon}l + 2h\frac{\mathbf{r}'}{\epsilon}\right],$$

$$B = -h^2 + 3h^2(l \cdot \mathbf{r})^2 + \dot{h}(l \cdot \mathbf{r}), \qquad (A2)$$

$$C = \frac{1}{\epsilon}\left[\frac{\epsilon'}{\epsilon^2} + h(1 + l \cdot \mathbf{r})\right], \quad D = -\dot{h} - 2h^2 l \cdot \mathbf{r}.$$

Now we notice that to this order we may remove the l' and $\mathbf{r}'$ terms by defining

$$\tilde{l} = l\left[\sigma + C\frac{d\eta^2}{2}, \eta\right], \quad \tilde{\mathbf{r}} = \mathbf{r}\left[\sigma - \frac{d\eta}{\epsilon}, \eta\right] \qquad (A3)$$

so that (A2) becomes

$$l(\sigma - d\eta/\epsilon, \eta + d\eta) = \tilde{l} + d\eta[-h\tilde{\mathbf{r}} + h(l \cdot \tilde{\mathbf{r}})l]$$

$$+ \frac{d\eta^2}{2}(Bl + D\mathbf{r}). \qquad (A4)$$

How do we calculate (A3)? Recalling that in our scheme l and $\mathbf{r}$ are unit vectors, we interpolate along a great circle on the unit sphere between l_k and l_{k+1} a distance $C d\eta^2/2$ (for $C > 0$) or between l_k and l_{k-1} for $C < 0$. We find $\tilde{\mathbf{r}}$ similarly. Now we have l at A_L, we calculate l at B_L similarly. We then interpolate between l_{A_L} and l_{B_L} to find $l_{k'}$. Note that (A4) preserves the constraint $l^2 = 1$ up to terms involving $d\eta^3$, but not exactly. We therefore rescale l at A_L and B_L to make its length unity before interpolating on the sphere. This introduces

errors of order $d\eta^3$, no longer than those we have included anyway.

We calculate $\mathbf{r}_{k'}$ similarly. To update ϵ we note that $\alpha = \ln(a\epsilon)$ obeys

$$\dot{\alpha} = -hl \cdot \mathbf{r}$$

$$\rightarrow \alpha(\sigma, \eta + d\eta)$$

$$= \alpha(\sigma, \eta) - d\eta \, hl \cdot \mathbf{r} - \frac{d\eta^2}{2}[\dot{h}l \cdot \mathbf{r} + h(\dot{l} \cdot \mathbf{r})]$$

$$= \alpha(\sigma, \eta) = \frac{d\eta}{2}h(l \cdot \mathbf{r} + l_u \cdot \mathbf{r}_u) - \frac{d\eta^2}{2}\dot{h}l \cdot \mathbf{r} \qquad (A5)$$

up to order $d\eta^3$. Here l_u and $\mathbf{r}_u$ are the new values of l and $\mathbf{r}$ at σ, i.e., $l_{k'}$ and $\mathbf{r}_{k'}$.

Now from the updated l, $\mathbf{r}$, and ϵ along a loop we can reconstruct $\mathbf{x}'$ and $\mathbf{u} = \epsilon\dot{\mathbf{x}}$. However reconstructing the position of the whole loop is a bit more complicated. In particular the constraint $\int d\sigma \, \mathbf{x}' = 0$ required to close the loop may be violated by numerical errors. We deal with this by making sure that such errors are distributed evenly over the whole loop, so no "gaps" appear. We use the "center of mass" of the loop $\mathbf{x}_c = (1/N)\sum \mathbf{x}_k$ which is evolved separately as each point $\mathbf{x}_k$ on a loop of length N is evolved. This is exactly the center of mass if ϵ is not uniform, but is always perfectly well defined. Having updated $\mathbf{x}_c$ we reconstruct the positions of points on the loop by

$$\mathbf{x}_1 = \mathbf{x}_c - \frac{ds}{N}[(N-1)\mathbf{x}_1' + (N-2)\mathbf{x}_2' + \cdots + \mathbf{x}_{N-1}'],$$

$$(A6)$$

$$\mathbf{x}_2 = \mathbf{x}_c - \frac{ds}{N}[(N-1)\mathbf{x}_2' + (N-2)\mathbf{x}_3' + \cdots + \mathbf{x}_N'],$$

and so on. Here $\mathbf{x}_k' \equiv (\mathbf{x}_{k+1} - \mathbf{x}_k)/ds$. These formulas, being cyclically related, obviously treat all $\mathbf{x}_k$ on the same footing, and so errors in $\sum \mathbf{x}_k' = 0$ are distributed evenly.

Loops whose energy is less than μR_H are evolved separately as follows. ϵ is treated as uniform along the string, and the left and right movers are evolved as if they were in flat spacetime, with the *global* time step instead of the time step appropriate to the loop. The center of mass and ϵ are evolved using the analytic approximation explained in Sec. I for loops well inside the Hubble radius. We do not believe that these two approximations can lead to serious error—the first readjusts each loop *once* as it is chopped off by a small amount and the second means that the loop evolves slightly faster than it actually should. Since the oscillatory motion is periodic anyway, this results in self-intersections occurring slightly earlier than they should. There are no *cumulative* numerical errors introduced in the internal oscillation of the loops with this procedure (it is periodic to machine accuracy), which is very important since small loops have to undergo many internal oscillations over the course of the simulation, and cumulative errors would lead to spurious self-intersections. This is particularly important for loops with small numbers of points where evolving anything but the flat-spacetime solution would quickly produce nonperiodic motion.

Now we turn to the method for detecting and enacting string interactions.

This part of the program is the most time consuming, and it is very important to use a method which is efficient. The most naive method, checking each string segment with every other for crossing each time step obviously scales as N^2 where N is the total number of points on the string (typically 250 000 or so in our large simulations) and is prohibitively slow. We use a method which scales as N. The procedure is first to divide all of space into small (comoving) boxes (typically $\frac{1}{5}$ of an initial correlation length). Each box corresponds to an element of a large array. Then we look for self-intersections of all loops in turn. We do this by tracking along a loop, calculating which box we are in, and recording the label of the present point in (the array element for) that box. We update the boxes as we go, with the result that each box contains the label for the last point on the string in it or is empty. If we come to a new box and find a point on the current string in it, we have a candidate crossing event. Now we track forward from the current point, and backward on the detected point up to and including the first point that leaves the box in both cases. We now have two stretches of string which may intersect in the current time step. Every pair of segments, one segment from each stretch of string, is checked for crossing in detail.

The detailed crossing check works as follows. We use the fact that in our chosen evolution method the velocity of a segment is always perpendicular to the segment. We assume that each segment has this velocity for the whole of the following time step, so the world sheet we assume for the segment during the next time step is a rectangular blade of length $d\sigma$ and width $V\,d\eta$, where V is the velocity of the segment. Imagine going to the rest frame of one segment by a Galilean transformation (i.e., simply adding the negative of its velocity to the velocity of the other segment). Now the problem of whether the two segments intersect becomes simply whether the first string segment (stationary in this frame) pierces the parallelogram swept out by the other segment. This is a simple geometrical problem which may be solved exactly using two cross products. The details are explained in Fig. 20. One drawback with our method is that the world sheet assumed, with rectangular blades swept out by each segments, has "gaps" between neighboring segments so that it is possible for us to miss intersections where the rectangular blades do not intersect. We have checked for the significance of this by increasing the number of points along the string, which should have the effect of narrowing the "gaps" for the same curvature along the string, and as we shall discuss this did not have much effect.

If a pair of segments do intersect, then the values of x at the ends of each new segment are determined but the segments are not. We choose the magnitude of the velocity for each segment so that each segment carries off half the total energy of the initial two segments. The direction of the velocity is chosen as follows. First, the center of mass of the four points involved in the intersecting segments is found. Then for each of the two new segments, the vector joining this point to the center of the segment is constructed. Finally, this is projected onto the plane through the center of each segment to make it perpendicular to the segment. Since the velocities of the segments
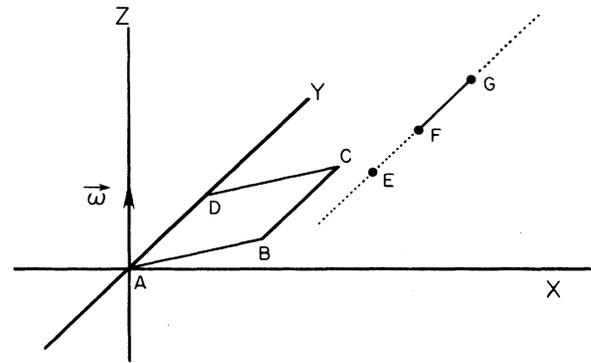

FIG. 20. Detailed crossing detection for two segments on the string. In the Galilean rest frame of segment FG, the other segment traces out a parallelogram $ABCD$ (which is in the x-y plane in the figure). Now the line obtained by continuing FG, given by $\vec{x}(t) = \vec{AF} + t\vec{FG}$, pierces the $z = 0$ plane at the point E. This is found by calculating $\vec{w} \equiv \vec{AB} \wedge \vec{AD}$ and then solving $\vec{x}(t)\cdot\vec{w} = 0$ for t. The first test is that we must have $0 < t < 1$. Then we must have $0 < (\vec{AE} \wedge \vec{AD})\cdot\vec{w} < \vec{w}^2$, so E lies between the parallel lines through AD and BC, and finally $0 < (\vec{AB} \wedge \vec{AE})\cdot\vec{w} < \vec{w}^2$, so E lies between AB and CD. If all three conditions are met, the segments intersect.

before projection are away from each other, and projection changes each by at most 90°, the velocities of the two new segments are guaranteed to be away from each other, so they will not intersect in the same time step. Our procedure is rather artificial, and undoubtedly could be improved on, by solving for the motion of the intersecting segments in flat spacetime, for example. One must however avoid creating low-energy new segments, since these would slow down the time step of the simulation.

Once a self-intersection is found, the new loop is broken off, the boxes it passes through (including the box where the intersection was found) reinitialized (i.e., labeled as empty), and the tracking process continued. This could in principle miss multiple self-intersections in the same box (which are rather unlikely). All self-intersections are found this way. We also impose the condition that any loop have at least three points on it, a fairly minimal requirement.

Then loops are checked for intersections with other loops. One again tracks along each loop, leaving point labels behind in the detection boxes. If one enters a new box where a point from a different loop is located, then one checks all pairs of segments on the two stretches of string running through the box, just as above. The same procedure is followed for defining the new segments.

Now the biggest problem with this method is that one may miss (a) segments which cross over a face of two adjoining detection boxes or (b) loops smaller than the size of the detection boxes. We remedy these problems by repeating the whole procedure outlined above not just once but 3 times per time step, but choosing the lattice for the detection boxes to be displaced by 0, $\frac{1}{3}$, and $\frac{2}{3}$ of a detection box side in each of the x, y, z directions each of the 3 times. This makes our resolution for crossing detection

(the maximal size of the largest loop which could escape detection) one-third of a detection box side. This is certainly a cautious estimate for the breakup of small loops, since these are moving and if a crossing is not detected in one period it may be the next time around.

We found that repeating the detection 3 times did indeed increase the number of detected crossings significantly, but repeating it 6 times did not then produce a noticeable difference. We discuss a test run where we used 50 instead of 10 points per initial correlation length on the string, and used detection boxes of $\frac{2}{25}$ instead of $\frac{1}{5}$ of an initial correlation length. In addition we repeated the checking process in offset boxes 6 times instead of 3; thus providing an overall factor of 5 increase in detection resolution.

[1]T. W. B. Kibble, J. Phys. A **9**, 1387 (1976); Phys. Rep. **67**, 183 (1980).

[2]Ya Zel'dovich, Mon. Not. R. Astron. Soc. **192**, 663 (1980); A. Vilenkin, Phys. Rev. Lett. **46**, 1169 (1980); **46**, 1496(E) (1980); Phys. Rep. **121**, 263 (1985).

[3]T. W. B. Kibble, G. Lazarides, and Q. Shafi, Phys. Rev. D **26**, 435 (1982); D. Olive and N. Turok, Phys. Lett. **117B**, 193 (1982).

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

[5]H. Hodges, Sant Cruz report, 1988 (unpublished).

[6]N. Turok and P. Bhattacharjee, Phys. Rev. D **33**, 1557 (1984).

[7]A. Albrecht and N. Turok, Phys. Rev. Lett. **54**, 1868 (1985).

[8]T. W. B. Kibble, Nucl. Phys. **B252**, 227 (1985).

[9]N. Turok, in *Particle Physics and Astrophysics: Current Viewpoints*, edited by H. Mitter and F. Widder (Springer, New York, 1989).

[10]N. Turok, Phys. Rev. Lett. **55**, 1801 (1985).

[11]D. Bennett, Phys. Rev. D **33**, 872 (1986); **34**, 3592 (1986).

[12]D. Bennett and F. Bouchet, Phys. Rev. Lett. **60**, 257 (1988).

[13]F. Bouchet, LBL report, 1988 (unpublished); D. Bennett, Fermilab report, 1988 (unpublished); F. Bouchet and D. Bennett, presentation at Yale Workshop on Cosmic String, 1988 (unpublished).

[14]T. W. B. Kibble, J. Phys. A **9**, 1387 (1976); Phys. Rep. **67**, 183 (1980).

[15]K. Maeda and N. Turok, Phys. Lett. B **202**, 376 (1988).

[16]R. Gregory, Phys. Lett. B **206**, 199 (1988).

[17]Y. Nambu, in *Symmetries and Quark Models,* proceedings of the International Conference, Detroit, Michigan, 1969, edited by R. Chand (Gordon and Breach, New York, 1970).

[18]R. Matzner, Comput. Phys. **2**, 51 (1988).

[19]E. P. S. Shellard, Nucl. Phys. **B282**, 624 (1987); E. P. S. Shellard and P. Ruback, Phys. Lett. B **209**, 262 (1988).

[20]P. Goddard, G. Goldstone, C. Rebbi, and C. Thorn, Nucl. Phys. **B56**, 109 (1973).

[21]T. W. B. Kibble and N. Turok, Phys. Lett. **116B**, 141 (1982); A. Albrecht and T. York, Phys. Rev. D **38**, 295 (1988).

[22]N. Turok, Nucl. Phys. **B242**, 520 (1984).

[23]A. Albrecht and N. Turok, Phys. Rev. Lett. **54**, 1868 (1985).

[24]N. Turok, Phys. Lett. **123B**, 287 (1983).

[25]P. K. Raschewski, *Riemannsche Geometrie und Tensoranalysis* (VEB Deutscher der Wissenschaften, Berlin, 1959).

[26]T. W. B. Kibble, Nucl. Phys. **B252**, 227 (1985).

[27]T. Vachaspati and A. Vilenkin, Phys. Rev. D **31**, 3052 (1985); D. Garfinkle and T. Vachaspati, *ibid.* **36**, 2229 (1987).

[28]D. Mitchell and N. Turok, Phys. Rev. Lett. **58**, 1577 (1987); Nucl. Phys. **B294**, 1138 (1987).

[29]S. W. Hawking, DAMTP report, 1987 (unpublished).

[30]T. York, presentation at the Yale Workshop on Cosmic String (Ref. 13).

[31]T. York, Phys. Rev. D **40**, 277 (1989).

[32]E. Copeland, D. Haws, and R. Rivers, Nucl. Phys. **B298**, 445 (1989).

[33]J. Frieman and R. Scherrer, Phys. Rev. D **33**, 3556 (1986).

[34]G. Smith and A. Vilenkin, Phys. Rev. D **36**, 987 (1987); M. Sakellariadou and A. Vilenkin, *ibid.* **37**, 885 (1988).

[35]R. Scherrer and W. Press, Smithsonian report, 1988 (unpublished).

[36]A. Everett, T. Vachaspati, and A. Vilenkin, Phys. Rev. D **31**, 1925 (1985); E. Copeland, D. Haws, T. Kibble, D. Mitchell, and N. Turok, Nucl. Phys. **B298**, 445 (1988).

[37]A. Vilenkin and Q. Shafi, Phys. Rev. Lett. **51**, 1716 (1983).

[38]N. Turok and R. Brandenberger, Phys. Rev. D **33**, 2175 (1986).

[39]P. J. E. Peebles, *The Large Scale Structure of the Universe* (Princeton University Press, Princeton, NJ, 1980).

[40]R. Brandenberger, N. Kaiser, D. Schramm, and N. Turok, Phys. Rev. Lett. **59**, 2371 (1987); R. Brandenberger, N. Kaiser, and N. Turok, Phys. Rev. D **36**, 335 (1987); E. Bertschinger and P. Watts, Astrophys. J. **328**, 23 (1988).

[41]E. Bertschinger, Astrophys. J. **316**, 489 (1987).

[42]A. Albrecht, A. Stebbins, and N. Turok, 1989 (in preparation).

[43]N. Kaiser and A. Stebbins, Nature (London) **310**, 391 (1984).

[44]J. R. Gott, Astrophys. J. **288**, 422 (1985).

[45]A. Stebbins, Astrophys. J. **327**, 584 (1988).

[46]F. Bouchet, D. Bennett, and A. Stebbins, Nature (London) **335**, 410 (1988).

[47]A. Vilenkin, Phys. Lett. **107B**, 47 (1981).

[48]L. A. Rawley, J. H. Taylor, M. M. Davis, and D. W. Allan, Science **238**, 761 (1987).

[49]C. Hogan and M. Rees, Nature (London) **311**, 109 (1984).

[50]R. L. Davis, Phys. Rev. D **32**, 3172 (1985).

[51]A. Vilenkin, Phys. Rev. D **23**, 852 (1981); Nature (London) **322**, 612 (1986); see also C. Hogan, Steward Observatory Report No. 698, 1987 (unpublished).

[52]L. Cowie and E. Hu, Astrophys. J. **318**, L33 (1987).