Fermi National Accelerator Laboratory

FERMILAB-Conf-88/116-A
August 1988

Phase Transitions as the Origin of Large Scale Structure in the Universe

Neil Turok[1]

NASA/Fermilab Astrophysics Center
Fermi National Accelerator Laboratory, Batavia, Illinois 60510

Abstract

A review of the formation of large scale structure through gravitational growth of primordial perturbations is given. This is followed by a discussion of how symmetry breaking phase transitions in the early universe might have produced the required perturbations, in particular through the formation and evolution of a network of cosmic strings. Finally the statistical mechanics of string networks, for both cosmic and fundamental strings is discussed, leading to some more speculative ideas on the possible role of fundamental strings (superstrings or heterotic strings) in the very early universe.

August, 1988

Lectures presented at the XXVII Internationale Universitatswochen fur Kernphysik, Schladming, Austria 22 February - 3 March 1988; Theoretical Advanced Studies Institute, Brown University, June 1988; CCAST Symposium on Particle Physics and Cosmology, Nanjing, China 30 June - 13 July 1988.

---

[1]Address after September 1st 1988: Joseph Henry Laboratory, Princeton University, Princeton NJ08544

1 Very Large Scale Structure and Gravity

The standard big bang model of the early universe is both simple and successful. It assumes that at early times the universe was a homogeneous and isotropic thermal bath of all the particle species see today. From this and Einstein's equations it correctly predicts the Hubble expansion, microwave background radiation and the abundances of the light elements (for a review, see [1]).

Nevertheless it is obviously only a first approximation - the universe is very clearly not homogeneous and isotropic today and so there must have been some inhomogeneity at early times. In fact, as we shall see in this section, a number of recent observations point to the existence of structure in the universe on scales much larger than the distances matter has moved (under gravity) since the big bang. This is exciting - the new very large scale observations are giving us a more or less direct 'window' on the primordial fluctuations.

Where could the large scale structure in the universe have come from?

A beautiful idea has emerged from fundamental particle physics in the last few years. It is based on the notion of spontaneous symmetry breaking, which has emerged as fundamental in unified theories of particle interactions. What could be simpler than that the mechanism which broke the symmetry of particle interactions also broke the spatial symmetry of the universe?

This happens through the dynamics of the symmetry breaking process, as I shall discuss in section 2. At very high temperatures the symmetry is unbroken but as the universe cools and expands a phase transition occurs,

forming defects as the universe goes out of thermal equlibrium which survive to much later epochs and seed the formation of structure.

One particular class of defects shows particularly interesting behavior in an expanding universe. Cosmic strings[2] are linelike defects formed in symmetry breaking phase transitions, and have the property that they evolve into a 'scaling solution', where the strings remain a fixed fraction of the total density throughout the radiation dominated era. This results in a highly predictive theory of galaxy formation, an alternative to the more popular theory based on quantum fluctuations produced during a period of inflation. I will concentrate on cosmic strings here but will also compare and contrast the two scenarios in section 2.

Of course we still do not know what the right fundamental theory is. Some grand unified theories predict strings and others do not. Over the last few years there has been increased hope that superstring theories may offer an 'ultimate' unified theory, including gravity. Some 'superstring-inspired' models also predict cosmic strings, but so far these are so contrived that it is hard to take them seriously.

In the final section I will describe an interesting connection between cosmic strings and fundamental strings. The latter may be used as an approximate description of the former, and for fundamental strings the statistical mechanics is relatively straightforward. I hope to give a clear and simple account of this in section 3, and discuss what insight this gives us into both cosmic and fundamental strings in the early universe.

1.1 Observations and Motivations

Cosmology is one of the most rapidly growing areas in physics for two very good reasons.

First the observational situation is qualitatively better than just a few years ago and improving all the time. Advances in telescope technology have made large scale complete surveys of the three dimensional galaxy distribution feasible. There are also now very stringent and improving limits on the microwave background anisotropy and gravity wave background, which put strong constraints on the magnitude of the primordial inhomogeneities - typically the magnitude of the density fluctuation must have been less than a part in ten thousand.

What makes the large scale observations interesting is that the distribution of galaxies is very clearly not random on very large scales. Early three dimensional surveys revealed giant 'filaments' (roughly linear overdense regions in the galaxy distribution) about $200h_{50}^{-1}$ Megaparsecs long and $10h_{50}^{-1}$ Mpc across [3], large 'voids' $120h_{50}^{-1}$ Mpc in diameter [4]. A complete 'slice of the universe' out to $300h_{50}^{-1}$ Mpc shows the galaxy distribution to be 'frothy' i.e. many galaxies lying on the surfaces of almost empty 'bubbles' $40-60h_{50}^{-1}$ Mpc across [5], while deep 'needle' surveys out to further than $2000h_{50}^{-1}$ Mpc have shown a shell-like distribution of galaxies on a scale of as much as $200h_{50}^{-1}$ Mpc [11]. The rich clusters of galaxies, called Abell clusters [12], are (roughly) defined observationally to be regions $3h_{50}^{-1}$ Mpc in radius containing more than 50 bright galaxies. These seem to be significantly clustered on scales of at least $200h_{50}^{-1}$ Mpc [13]. For comparison the present Hubble radius H_0^{-1} is $6000h_{50}^{-1}$ Mpc, the mean galaxy separation (i.e. the inverse cube root of the number density) is $10h_{50}^{-1}$ Mpc [14] and the mean Abell cluster separation is $110h_{50}^{-1}$ Mpc [13]. h_{50} is of course Hubbles constant H_0 in units of $50kms^{-1}$, observationally $1 < h_{50} < 2$.

Statistics to describe these structures are still being developed - the simplest is the 2-point correlation function $\xi(r)$, defined to be the excess probability over random of finding a galaxy at a distance r from another galaxy [7]. There are different ideas relating the primordial fluctuations to the final $\xi(r)$ we see today.

We now believe that most of the mass in the universe is dark. There is no strong reason why the shining matter distribution should exactly trace the dark matter distribution. What processes could end up making the matter in one part of the universe shine and leave another part dark? One popular idea is that of *threshold biasing* - that shining matter might only be formed at the *peaks* of the underlying matter distribution [21,22]. With Gaussian noise, the peaks are more correlated than the underlying distribution, and this is used crucially to get the standard inflation plus CDM model (which I shall call the ICDM model) to agree with the galaxy-galaxy correlation function and pair velocity distribution. However this idea, at least as applied in the ICDM model, does not appear to work particularly well as far as fitting large scale structure[79].

Another idea, and one that is at least partly realised in the cosmic string scenario, is that the distribution of galaxies and clusters may reflect a perturbation distribution with a *scale invariant* correlation function i.e. objects of mean separation d have a correlation function $\xi(r/d)$ with ξ the same function for all objects. This is consistent with the observational fact that clusters poorer than Abell clusters [16], the richer Abell clusters [13] and even superclusters[15] have clustering apparently consistent with a single scale invariant correlation function. In fact ξ is approximately equal to $0.2(r/d)^{-2}$. Written the same way galaxies appear slightly *more* correlated: $\xi_{gg} \approx (r/d)^{-2}$[18], consistent with the view that gravity has enhanced the correlations on small scales.

Statistics on these observations are still not good and the ideas for explaining them clearly preliminary, but the observations are improving all the time and becoming increasingly constraining. Clearly data like this are going to play a crucial role in deciding which theories of the origin of large scale structure are viable.

The fact that the distribution of galaxies does not necessarily reflect the distribution of the mass (and in particular the dark matter) in a simple way represents a profound problem for matching cosmological theories to observation. Which is why recent determinations of the actual peculiar velocity field (the velocities of galaxies relative to the overall Hubble flow) are of such great importance. These seem to indicate a very large scale flow field centred on a region ('the great attractor') $80h_{50}^{-1}$ Mpc from us with magnitude about $500kms^{-1}$ at our galaxy [17]. I will discuss below how this velocity field is related, in the linear regime, to the mass perturbation. If these observations are correct, they practically rule out the ICDM model.

So what if there is large scale structure in the universe? The most striking thing about it is that matter could not have moved these distances under gravity since the big bang. In the linear regime (and on these scales the

structure is not yet nonlinear) there is a precise relation between the distance a particle has moved relative to the Hubble flow since the big bang δr and its present peculiar velocity δv: $\delta r = H_0^{-1}\delta v$ (see below). Now galaxies are very rarely observed to have velocities of greater than $600kms^{-1} = 2.10^{-3}c$ relative to the observed structures. This means they can have moved no more than $12h_{50}^{-1}$Mpc since the big bang. So when we see much larger structures we are seeing the primordial density perturbation field itself! This is an encouraging thought.

The second good reason for a growing interest in cosmology is that only in the last few years have any well-defined ideas as to the *origin* of large scale structure come forward. In particular cosmic strings and quantum fluctuations in inflationary models are well defined theories with clear predictions. For example, the quantum fluctuation scenario with hot dark matter is close to being ruled out by its excessively large predicted anisotropy in the microwave background [19].

1.2 Dark Matter

The greatest single obstacle to the interpretation of the observations of large scale structure is the by now firmly established fact that most of the matter in the universe is dark (For a nice review, see [9]). The shining matter - gas and stars - is only the 'tip of the iceberg'. The best evidence for dark matter comes from observations of the rotation curves of spiral galaxies[20]. To keep a particle in orbit around a galaxy requires an acceleration of v^2/r where v and r are the orbital velocity and radius. Assuming that this is provided by the gravitational pull of the mass $M_{<r}$ inside the orbit one finds

$M_{<r} = \frac{v^2 r}{G}$. Observations indicate that v^2 remains constant with increasing r out to distances several times the optical radius of a galaxy. Thus most of the mass is dark. Estimates of the mass of apparently virialised clusters of galaxies confirm this - and extrapolating the observed mass to light ratios in clusters to the whole universe leads to an estimate of Ω typically between 0.1 and 0.3. This may well be an underestimate, particularly if galaxies form preferentially in denser regions like clusters, making the mass to light ratio larger outside clusters.

Nucleosynthesis calculations[8] in fact indicate that even the *baryons* are mostly dark! They yield the bounds $.04 < \Omega_{baryons} h_{50}^2 < 0.14$ whereas the observed $\Omega_{luminous}$ i.e. the density in stars and gas is less than 0.02. So we can't even see the baryons!

However things are worse than this, according to most theorists. There is a strong prejudice that $\Omega = 1$. This is because in a matter or radiation dominated universe $\Omega = 1$ is an unstable fixed point of Einstein's equation. If Ω deviates from 1 then the deviation grows rapidly in time. Given that Ω is today fairly close to unity it would seem crazy that we should apparently be here precisely around the time when it started to deviate from unity. Inflation 'solves' this problem (called the flatness problem) - during a period of inflation $\Omega = 1$ becomes an *attractor* in Einstein's equation so that in a typical region of the universe which underwent inflation Ω would still be very close to 1.

1.3 Newtonian Approximation

For many purposes the Newtonian approximation to Einstein's equation is entirely adequate. In particular it is a good approximation when dealing with scales (a) far inside the horizon so velocities are small and (b) where gravity is fairly weak so the dimensionless gravitational potential is much less than 1. It is also a lot simpler than full relativistic perturbation theory and brings out the main points about the growth of density perturbations under gravity.

So let us begin with Newton's equations. A particle at 'physical' coordinate $\vec{r}(t)$ obeys

$$\ddot{\vec{r}} = -\vec{\nabla}_r \Phi \tag{1}$$

where the Newtonian potential Φ is given by

$$\vec{\nabla}_r^2 \Phi = 4\pi G\rho \tag{2}$$

with ρ the density. Before discussing perturbations, let us describe the homogeneous matter dominated background solution to (1) and (2). We set $\rho = \rho_b$, independent of $\vec{r}$, and find in spherical coordinates about some point chosen as the origin (where $r = \Phi = 0$),

$$\begin{aligned} \Phi &= \frac{2\pi G\rho_b r^2}{3} \\ \ddot{\vec{r}} &= -\frac{4\pi G\rho_b \vec{r}}{3} \\ \rho_b r^3 &= const \end{aligned} \tag{3}$$

the last equation being matter conservation. Integrating we find

$$\frac{\dot{r}^2}{2} - \frac{4\pi G\rho_b r^2}{3} = const \tag{4}$$

and choosing the constant to be zero (i.e. Ω to be 1) we find

$$
\begin{aligned}
r &\propto t^{\frac{2}{3}} \\
\rho_b &= \frac{1}{6\pi G t^2}
\end{aligned}
\tag{5}
$$

the standard Friedmann matter dominated universe. This solution is true for every particle, and we write it as $\vec{r} = a(t)\vec{x}$ where $a(t)$ is the scale factor. We can choose $a(t)$ equal to 1 initially in which case $\vec{x}$ is the initial coordinate of the particle (the 'comoving coordinate'). The Newtonian approximation gives the correct result as a consequence of homogeneity and Birkhoff's theorem[6] - it is only the contribution of the matter *inside* a spherical shell which is important, and this shell may be chosen small enough so that fields are weak and velocities small so Newtons equations apply. By homogeneity the same rate of expansion must apply on all scales.

1.4 Linear Perturbations in Cold Dark Matter

For cold dark matter we also have mass conservation in a comoving volume

$$\int \rho \, dV_{comoving} = const \tag{6}$$

Now imagine we give each particle a small perturbation at some early time and follow it thereafter

$$\vec{r} = a(t)(\vec{x} + \vec{\psi}(\vec{x}, t)) \tag{7}$$

$\vec{\psi}$ is called the comoving displacement. Mass conservation tells us the resulting density perturbation: if the initial unperturbed density is $\rho_i = \rho_b a^3$ then

$$
\begin{aligned}
\rho_i d^3\vec{x} &= \rho_b a^3 d^3\vec{x} \\
&= \rho(\vec{r}) d^3\vec{r}
\end{aligned}
\tag{8}
$$

so the total density

$$
\begin{aligned}
\rho(\vec{r}) &= \frac{a^3\rho_b}{Det(\frac{\partial \vec{r}}{\partial \vec{x}})} \\
&\approx \rho_b(1 - \vec{\nabla}_x . \vec{\psi})
\end{aligned}
\tag{9}
$$

where we expand the determinant to first order in the perturbation using $Det(1 + A) = 1 + Tr(A) + o(A^2)$. Thus the fractional density perturbation is given by

$$
\begin{aligned}
\frac{\delta\rho}{\rho}(\vec{r}) &\equiv \frac{\rho(\vec{r}) - \rho_b}{\rho_b} \\
&= -\vec{\nabla}_x . \vec{\psi}
\end{aligned}
\tag{10}
$$

At this point it is convenient to Fourier transform $\vec{\psi}(\vec{x})$ and decompose it into two pieces;

$$
\begin{aligned}
\vec{\psi}(\vec{k}) &\equiv \int d^3\vec{x} e^{i\vec{k}.\vec{x}} \vec{\psi}(\vec{x}) \\
&\equiv \psi_{\parallel}(\vec{k})\frac{\vec{k}}{k} + \psi_{\perp}(\vec{k})
\end{aligned}
\tag{11}
$$

$\psi_{\parallel}$ (parallel to $\vec{k}$) is the *compressional* piece and $\psi_{\perp}$ (perpendicular to $\vec{k}$) is the *rotational* piece. Now from (2) and (9) only the *compressional* piece causes a perturbation in the potential $\delta\Phi$. Fourier transforming and solving (2) we find

$$\vec{\nabla}_r \delta\Phi = -4\pi G \rho_b a \vec{\psi}_{\parallel} \tag{12}$$

where to lowest order we use $\vec{\nabla}_r = a^{-1}\vec{\nabla}_x$. Now (1) decomposes into

$$
\begin{aligned}
\ddot{\vec{\psi}}_{\parallel} + 2\frac{\dot{a}}{a}\dot{\vec{\psi}}_{\parallel} &= 4\pi G \rho_b \vec{\psi}_{\parallel} \\
\ddot{\vec{\psi}}_{\perp} + 2\frac{\dot{a}}{a}\dot{\vec{\psi}}_{\perp} &= 0
\end{aligned}
\tag{13}
$$

These equations have a simple interpretation. The left hand side describes a test particle moving in comoving coordinates - setting it to zero results in $\dot{\vec\psi} \propto a^{-2}$ so the peculiar velocity (the physical velocity $\dot{\vec r}$ minus the Hubble flow $\frac{\dot a}{a}\vec r$) $\vec v_p = a\dot{\vec\psi} \propto a^{-1}$. In comoving coordinates a test particle slows down. This is all that happens to the rotational piece. The right hand side of (13) describes the feedback of the compressional perturbation on itself - the gravitational instability first discovered by Jeans. If the background were not expanding, $\vec\psi$ would grow exponentially but in a matter dominated universe with $a \propto t^{\frac{2}{3}}$ we find

$$\begin{aligned} \ddot{\vec\psi}_\parallel + \frac{4}{3t}\dot{\vec\psi}_\parallel &= 4\pi G\rho_b\psi_\parallel = \frac{2}{3t^2}\psi_\parallel \\ \ddot{\vec\psi}_\perp + \frac{4}{3t}\dot{\vec\psi}_\perp &= 0 \\ \rightarrow \vec\psi_\parallel &= \vec A(\vec x)t^{\frac{2}{3}} + \vec B(\vec x)t^{-1} \\ \vec\psi_\perp &= \vec C(\vec x)t^{-\frac{1}{3}} + \vec D(\vec x) \end{aligned} \tag{14}$$

with $\vec A, \vec B, \vec C$, and $\vec D$ arbitrary functions of x. From (10) $\delta\rho/\rho$ behaves just as $\psi_\parallel$, so in the growing mode it grows like $a(t)$.

More generally the right hand side of (13) has an extra term $-a^{-1}\vec\nabla_r\Phi_s(\vec r)$ if there is a source driving the perturbation producing a potential Φ_s. For example for a point mass it reads $-Gm\vec x/(a^2x^3)$. Likewise in the equation for $\delta\rho/\rho$ the source term is $4\pi G\rho_s(\vec x)$, where $\rho_s(\vec x)$ is the density of the source. Notice that with cold dark matter the evolution is *purely local* in comoving coordinates.

In general if one has a theory for the origin of the perturbations one wants to evolve the perturbations through the radiation era into the matter era. Now the density ρ_b on the right hand side of (13) is only the *matter* density,

at early times much less than the total density. Setting it to zero and using $a \propto t^{\frac{1}{2}}$, the radiation dominated result, one finds the solution $\psi = E + F\ln(t)$ with E and F constants i.e. slow logarithmic growth. In fact solving (13) through the transition one finds at late times ($a >> a_{eq}$, the scale factor at equal matter and radiation density)

$$\vec\psi = \vec A(\vec x)(1 + \frac{3}{2}\frac{a}{a_{eq}}) \tag{15}$$

where $\vec A(\vec x)$ is the perturbation at early times [7]. Thus there is very little growth in the radiation era and one can think of the perturbation growth as beginning at a_{eq}.

1.5 Local Nonlinearity 'doesn't matter'

Whilst these equations are correct only in the linear regime they can be used perfectly well on a surface surrounding a nonlinear mass concentration as long as the density perturbation is always small *on the surface*. If one has a comoving surface S surrounding a volume V then the change in the volume enclosed by S is

$$\delta V = a^2\int_S \vec\psi.d\vec S \tag{16}$$

and the change in the mean density ρ inside S defined by $\rho_{av}(V + \delta V) \equiv \rho_i V_i = \rho_b V$ where $V = V_i a^3$. Now we find

$$\begin{aligned} \frac{\delta\rho}{\rho} &\equiv \frac{(\rho_{av} - \rho_b)}{\rho_b} \\ &= -\frac{\delta V}{V} \\ &= -\frac{\int_S \vec\psi.d\vec S}{V_i} \end{aligned} \tag{17}$$

which of course grows just as (14). Thus linear theory has a wide range of applicability.

1.6 Peculiar velocities

Linear theory also gives information about peculiar velocities in a very direct way - we see from $\vec{v}_p = a\dot{\vec{\psi}}$ and (14) that peculiar velocities grow as $t^{\frac{1}{3}}$ in the compressional mode. We also see using $\vec{\psi} \propto a$ in the growing mode that $\vec{v}_p = \dot{a}\vec{\psi} = H\delta\vec{r}$, the relation used in the introduction.

In particular geometries v_p is related to $\delta\rho/\rho$ - using (17) we find that defining the Hubble flow $v_H = Hr$ we get

$$\begin{aligned} \frac{v_p}{v_H} &= \frac{\delta\rho}{\rho} \text{ for planar collapse} \\ &= \frac{1}{2}\frac{\delta\rho}{\rho} \text{ for cylindrical collapse} \\ &= \frac{1}{3}\frac{\delta\rho}{\rho} \text{ for spherical collapse} \end{aligned} \quad (18)$$

Observations of peculiar velocities such as those recently reported[17] thus give a unique window on $\delta\rho/\rho$. If they are correct, we are moving in part of a very roughly spherical collapse pattern centred on the 'great attractor', 80 h_{50}^{-1} Mpc away from us. Our velocity towards it is approximately $500kms^{-1}$. This corresponds to a value of $\delta\rho/\rho$ of approximately 0.4 on this scale.

1.7 Sources

As discussed above gravitational sources (like loops of cosmic string) add an additional driving term to the right hand side of (13). It may then be solved fairly simply with a Green's function[65],[70]. Since the equation is

completely local $\delta\rho/\rho(x)$ simply evolves in proportion to $\delta\rho_S(x)$, the source density. In linear theory therefore a point mass always remains a point mass - its mass simply grows as $a(t)$. In the linear regime (13) with a source which 'turns on' in the matter era at t_i is easily solved with a Green's function

$$\begin{aligned} \delta\rho/\rho(x,t) &= \int_{t_i}^{t} dt' G(t,t') 4\pi G\rho_s(\vec{x},t') \\ G(t,t') &= \frac{3}{5}(t^{\frac{2}{3}}t'^{-\frac{2}{3}} - t^{-1}t') \end{aligned} \quad . \quad (19)$$

Extended sources like large cosmic string loops trace out an extended 'trumpet - shaped' structure in comoving coordinates leading to a density contrast of the same pattern[70], quite similar to the giant 'filament' mentioned above[3]!

1.8 Hot Dark Matter

If the dark matter particles have appreciable thermal velocities at t_{eq}, when perturbations start to grow, then the above approach must be modified - it is hard to follow the particles! Instead one must follow the particle trajectories in phase space and I will discuss this in the next section.

However the main effect is simple to understand. Roughly speaking, particles move a distance $d \approx vt$ in an expansion time. where v is their velocity. Structure on smaller scales than this is washed out by particle diffusion ('free streaming') while structure on larger scales evolves unaffected, just as with cold dark matter. This length scale corresponds to a mass scale, often called the 'neutrino Jeans mass'. Since peculiar velocities are damped as we discussed above, the thermal velocity v of the particles decays as a^{-1}. The comoving scale $d_c = d/a \propto t^{-\frac{1}{3}} \propto a^{-\frac{1}{2}}$ corresponds to a mass scale

$M_J \propto a^{-\frac{1}{2}}$. With neutrinos for example, of mass 25 h_{50}^2 eV such as would be required to make $\Omega = 1$ while respecting the nucleosynthesis bound, M_J is of order 10^{15} $M_\odot$ at matter-radiation equality and about 10^9 $M_\odot$ today.

1.9 Linearising Liouville

Clustering of hot dark matter can be treated quite precisely by linearising the equation for the phase space density, the Liouville equation. As before I will remain strictly Newtonian.

The phase space density f is defined so that the number of particles in the phase space volume $d^3\vec{r}d^3\vec{p}$ at time t is $f(\vec{r},\vec{p},t)d^3\vec{r}d^3\vec{p}$. Thus the ordinary density is given by $\rho = m\int d^3\vec{p}\ f$, where m is the mass of the particles. Conservation of particle number yields the continuity equation

$$\frac{\partial f}{\partial t} + \vec{\nabla}.(f\vec{v}) = 0 \tag{20}$$

where the divergence and velocity are in the full phase space: $\vec{\nabla} = (\vec{\nabla}_r, \vec{\nabla}_p)$, $\vec{v} = (\dot{\vec{r}}, \dot{\vec{p}})$. Recalling Liouvilles theorem that $\vec{\nabla}.\vec{v} = \vec{\nabla}_r.\dot{\vec{r}} + \vec{\nabla}_p.\dot{\vec{p}} = \vec{\nabla}_r.\vec{\nabla}_p H - \vec{\nabla}_p.\vec{\nabla}_r H = 0$ where the total Hamiltonian is H, we get

$$\frac{\partial f}{\partial t} + \vec{v}.\vec{\nabla} f = 0 \tag{21}$$

which is the statement that the function f is constant along particle trajectories. We will return to this in the next section. Next we use the fact that $(\dot{\vec{r}}, \dot{\vec{p}}) = (\vec{p}/m, -m\vec{\nabla}_r\Phi)$ with Φ the Newtonian potential. It is convenient to change coordinates to $\vec{x} = \vec{r}/a, \vec{q} = a\vec{p} - ma\dot{\vec{r}}$, comoving position and peculiar momentum, and $d\tau = dt/a$, conformal time. This yields

$$\frac{1}{a}\frac{\partial f}{\partial \tau} + \frac{\vec{q}}{ma^2}.\vec{\nabla}_x f - m\ddot{a}\vec{r}.\vec{\nabla}_q f - m\vec{\nabla}_x\Phi.\vec{\nabla}_q f = 0 \tag{22}$$

(Φ is independent of $\vec{q}$). Here $\ddot{a}\vec{r} = \frac{d^2a}{dt^2}\vec{r} = -\vec{\nabla}_x\Phi_0$ from Newton's equations for the background as discussed above, where Φ_0 is the zeroth order part of the potential. The homogeneous, zeroth order solution to this is

$$\begin{aligned} f_0 &= \frac{2}{e^{\frac{q}{T_\nu a}} + 1} \\ \rightarrow \rho_0 &= \frac{m}{a^3}\int d^3\vec{q}\frac{2}{e^{\frac{q}{T_\nu a}} + 1} = \frac{2mT_\nu^3}{\pi^2}\eta_3 \end{aligned} \tag{23}$$

where $\eta_3 \approx 1.803$ and $T_\nu \propto a^{-1}$ is the neutrino temperature, equal to $(4/11)^{\frac{1}{3}}$ the temperature of the microwave background radiation. We use the relativistic phase space density of one species of neutrino (and antineutrino) when they decoupled as initial conditions, and this is preserved as the universe expands even as the neutrinos go out of thermal equilibrium and go nonrelativistic, as we learn from (22). As an excercise the reader can calculate the required mass of the neutrino needed to make $\Omega = 1$, stated earlier.

It is fairly straightforward to linearise and solve (23) through the matter-radiation transition including sources just as for cold dark matter[24,25]. The interested reader may refer to [26] for a pedagogical account.

1.10 Phase Space Constraints for Hot Dark Matter

The fact that for neutrino (and more generally fermionic) dark matter the initial phase space density (23) has a maximum, namely $\frac{1}{2}$ per neutrino or antineutrino, has a nice consequence first pointed out by Tremaine and Gunn [27]. Because it is constant along particle trajectories, it cannot increase beyond this maximum. This puts a strong constraint, valid even in the nonlinear regime, on the clustering of neutrinos, relevant to the issue of whether neutrinos can form the dark matter in galactic halos.

Crudely, the phase space density corresponding to a virialised lump of N neutrinos with volume r^3 and momentum p_ν is just $Nr^{-3}p_\nu^{-3}$ and by the constraint, this cannot be larger one. However for a virialised lump we have $v^2 \approx GM/r \approx GNm/r$. Solving for N and substituting back into the previous formula we obtain $m_\nu^4 > 1/(vr^2G)$ (I use units where $\hbar$ is 1). A more detailed treatment[27] yields the bound

$$m_\nu > 30eV\left(\frac{100\ km\ s^{-1}}{\sigma}\right)^{\frac{1}{4}}\left(\frac{15\ kpc}{r_c}\right)^{\frac{1}{2}} \tag{24}$$

assuming the collapsed object is an isothermal sphere with core radius r_c and one dimensional velocity dispersion σ. Observations of dwarf galaxies with dark matter halos can therefore put strong constraints on the neutrino scenario.

1.11 Summary

In this section I have given a flavour of the new observations and what is encouraging about them. I have also given a brief introduction to the kind of calculation necessary to follow the gravitational clustering of matter around primordial perturbations. In the next section I will give an account of one theory as to where these perturbations came from, the theory of cosmic strings.

2 Phase Transitions, Defects and Fluctuations

In this section I will discuss the physics of symmetry breaking phase transitions predicted to occur in grand unified theories in the very early universe[28,29,2]. The discussion will necessarily be qualitative and crude - it is difficult to do precise calculations in this area! Nevertheless according to our present view the precise details are not crucial in most cases. This is because the phase transitions concerned happened very early in the life of the universe, typically at times of 10^{-34} seconds. Transitions producing domain walls or monopoles would have been disastrous fairly quickly - the defects quickly coming to dominate the universe. Those producing string however, which I shall concentrate on as the most interesting case, lead to a 'scaling solution' for the string network, in which the strings remain a fixed fraction of the total energy density in the universe. As I shall discuss, we have good reason to believe that the late time evolution of the scaling solution is rather insensitive to the precise details of the initial conditions.

For the moment I will ignore the possibility that the phase transitions were inflationary - indeed it is very difficult to arrange for inflation in a gauge symmetry-breaking transition - but will return to the question of how cosmic strings relate to inflation later on.

2.1 Physical picture of a phase transition

According to our current and successful understanding of high energy physics, the underlying field theory describing nature has a high degree of symmetry but most of this symmetry is broken at low energy.

As the simplest example of how this works, consider the theory of a scalar field ϕ invariant under $\phi \to -\phi$. If we give the field a potential $V(\phi)$ as in Figure 2.1 then the potential shares this symmetry, so the classical ground state of the theory is degenerate - in vacuo ϕ can sit at + or at −. The chosen configuration breaks the symmetry[1]

At high temperature however the symmetry is 'restored' simply because thermal excitations are energetic enough to 'kick' the field ϕ over the potential barrier. Any masses become irrelevant at high temperatures and so by dimensions one must have $< \phi >^2 \approx T^2$. Thus ϕ explores a wide range of the potential. As the system cools the field explores a smaller and smaller range until at low temperatures it flops down into one of the two minima. In some parts of space it chooses the + minimum and in other parts the − minimum. Because of this defects are inevitably formed - in going from a + region to a − region ϕ must cross over the potential barrier in $V(\phi)$. Where it does so there is a localised lump of potential and gradient energy - a *domain wall* separating the + and − regions. It is important to realise that these defects are, at low temperatures, a *nonequilibrium* phenomenon - they are Boltzmann suppressed. To understand them one needs to understand how the fields go out of equilibrium as the universe expands.

According to the most naive picture, at high T the system is characterised by thermal oscillations with $k \approx T$. Shorter wavelength modes are

[1]This idea is *classical* -quantum mechanically of course tunnelling can occur between the + and − states. The true ground state is a symmetric combination of + and −. In field theory however, in macroscopic volumes the tunnelling rate is very tiny so the classical picture is reasonable - there is simply a *classical* probability of being in + or −.

suppressed. Thus the crudest picture of the ϕ field is that it is smooth on a scale of the thermal wavelength. On might therefore expect that the size of the + and − domains would, as long as one stayed in equilibrium, be just this. However what is really relevant is not the shortest wavelength modes - these oscillate vigorously and simply average out to zero as the system cools. Rather it is the *long* wavelength modes, with their more sluggish oscillations, which determine how the system goes out of equilibrium.

2.2 Gaussian Approximation

Let us try to understand this in a little more detail. To begin with assume that Φ has a long wavelength component $\bar{\Phi}$ which we shall initially treat as constant. We write $\Phi = \bar{\Phi} + \delta\Phi$ where $\delta\Phi$ is the short wavelength component. I will be more precise about the division into short and long wavelengths later. We then calculate the partition function for Φ by first summing over the fluctuations $\delta\Phi$ about $\bar{\Phi}$ in the Gaussian approximation (i.e. keeping only quadratic terms in $\delta\Phi$ in the energy).

The small fluctuations are just massive fields whose mass is determined by the value of $\bar{\Phi}$. For $V(\Phi) = \frac{\lambda}{8}(\Phi^2 - \eta^2)^2$ (shown in Figure 2.1) the mass squared of $\delta\Phi$ is $m^2 = V''(\bar{\Phi}) = \frac{\lambda}{2}(3\bar{\Phi}^2 - \eta^2)$. In general if Φ is a Higgs field it gives a mass to other fields as well - in particular fermion and gauge fields. Their contribution may be included in exactly the same way.

Now recall that a massive boson field has a partition function

$$\begin{aligned} Z_b &= Tr(e^{-\beta H}) \\ &= \sum_{\{n_{\zeta}\}} e^{-\beta \sum_{\zeta} E_b} \end{aligned}$$

$$\begin{aligned} &= \prod_{\vec{k}} \sum_{n_{\vec{k}}=0}^{\infty} e^{-\beta n_{\vec{k}}\omega_k} \\ &= \prod_{\vec{k}} \frac{1}{(1-e^{-\beta\omega_k})} \end{aligned} \tag{25}$$

where $\vec{k}$ labels each momentum mode and $n_{\vec{k}}$ are the occupation numbers for each mode. ω_k is just $\sqrt{\vec{k}^2 + m^2}$, and $\beta = 1/T$. For fermions one has instead

$$Z_f = \prod_{\vec{k}} (1 + e^{-\beta\omega_k}) \tag{26}$$

The free energy contributed by a massive field is (recall $Z \equiv e^{-\beta F}$)

$$\begin{aligned} F_b &= \frac{V}{\beta} \int d^3k\, ln(1 - e^{-\beta\omega_k}) \\ &\approx V(-\frac{\pi^2 T^4}{90} + \frac{m^2T^2}{24} + ...) \quad \frac{m}{T} << 1 \end{aligned} \tag{27}$$

at high temperature. A Dirac fermion contributes

$$\begin{aligned} F_f &= 4\frac{V}{\beta} \int d^3k\, ln(1 + e^{-\beta\omega_k}) \\ &\approx V(-\frac{7\pi^2 T^4}{180} + \frac{m^2T^2}{12} + ...) \quad \frac{m}{T} << 1 \end{aligned} \tag{28}$$

similarly. Here V is the volume : we replace $\sum_{\vec{k}}$ by $V \int d^3k \equiv V \int d^3k/(2\pi)^3$. The high T expansion is obtained by scaling out the T dependence and Taylor expanding the integrand in m/T about $m/T = 0$.

It is easily seen that (27) and (28) are dominated by modes with $k \approx T$, just as we expected above. Thus for $m/T << 1$ we can regard (27) and (28) as the contribution to the full partition function of only summing over the high k modes, leaving the low k modes (i.e. $k << T$) not summed over. We then have an effective theory for the low k modes with a potential energy

given by adding the contributions of (27) and (28) to the zero temperature potential for Φ

$$\begin{aligned} V_{eff}(\Phi) &\approx V(\Phi) + m^2(\Phi)\frac{T^2}{24} \\ &\approx \frac{\lambda}{8}\Phi^4 - \frac{\lambda}{4}\eta^2\Phi^2 + \frac{\lambda}{16}\Phi^2T^2 \\ &\approx \frac{\lambda}{8}(\Phi^2 - \eta_{eff}^2)^2 \end{aligned} \tag{29}$$

ignoring Φ independent terms. Here

$$\begin{aligned} \eta_{eff}^2 &= \eta^2(1 - \frac{T^2}{T_c^2}) \\ T_c &= 2\eta \end{aligned} \tag{30}$$

To repeat, the extra temperature dependent terms come from summing over the short wavelength fluctuations about the long wavelength 'background'. This has a very important effect - at high temperature it changes the sign of the Φ^2 term and pushes the minimum to $\Phi = 0$, the 'unbroken' phase. The transition occurs when the Φ field becomes effectively massless, at $T = T_c = 2\eta$, called the critical temperature. This is shown in Figure 2.2. If the Φ field couples to other fields then it also gives them a mass and their contribution may be included in V_{eff} similarly. For example Φ generally gives gauge fields a mass squared $\approx g^2\Phi^2$ giving a correction term in (29) of $\approx g^2\Phi^2T^2$ and we find $T_c \approx \sqrt{\lambda/g^2}\eta$.

What is the cause of this 'restoration' of symmetry? In the partition function we are really finding the most likely configuration for the system plus heat bath (see Lecture 3). In this context this amounts to minimising the free energy with respect to Φ. The most important effect of Φ being nonzero and giving the other modes a mass is to cut down the number of states available (at given T) to that mode. This is an *entropy* effect, and at high T as we have seen it tends to favour $\bar{\Phi} = 0$.

One other point must be made. In the Gaussian approximation the exact expressions in (27) and (28) do not make sense for all $\bar{\Phi}$: for $\bar{\Phi}^2 < \eta^2/3$ we have $m^2 = \frac{\lambda}{2}(3\bar{\Phi}^2 - \eta^2) < 0$. This means that ω_k and hence F are not even real! This is clearly a problem with the approximation. Nevertheless the qualitative behaviour exhibited by the small m/T expansion is quite believable.

The function $V_{eff}(\bar{\Phi})$ is often called the finite temperature effective potential. In quantum mechanics, by considerations similar to those I explained in the context of symmetry breaking, this should never have two minima - $\bar{\Phi}$ can tunnel through from one minimum to another. However we are interested in an evolving macroscopic system, and we are approximating it as a *classical* $\bar{\Phi}$ plus a quantum fluctuation. Ignoring quantum tunnelling should be reasonable and so $V_{eff}(\bar{\Phi})$ is the relevant quantity. However as far as I am aware none of this has been made rigourous and I would certainly encourage those of you feeling queasy with it to try and do better yourselves[30]!

2.3 Spatial Distribution of the Fields near T_c

The interpretation of $V_{eff}(\bar{\Phi})$ as an effective potential for the long wavelength modes means that we can use it to discuss their distribution at high T. Clearly $\bar{\Phi}$ will not be spatially constant at finite T. To understand its spatial distribution consider a massive field in equilibrium. As we approach T_c from above $\bar{\Phi}$ looks like a massive field with m decreasing to zero. For T below

T_c, in the neighbourhood of the minima $\bar{\Phi} = \pm\eta_{eff}$ of $V_{eff}(\bar{\Phi})$, $\bar{\Phi}$ looks like a massive field with mass squared $m^2_{eff} = V''_{eff}(\bar{\Phi}) = \lambda\eta^2_{eff} = \lambda\eta^2(1-(T/T_c)^2)$. Now in thermal equilibrium a massive field has spatial correlation function

$$< \Phi(r)\Phi(0) >= \frac{Tr(\Phi(r)\Phi(0)e^{-\beta H})}{Tr(e^{-\beta H})} \tag{31}$$

The trace over all states is conveniently calculated in an occupation number basis just as the partition function in (25) was. With the mode expansion for Φ one finds

$$< n_{\vec{k}} \mid \Phi(r)\Phi(0) \mid n_{\vec{k}} > \;=\; \int \frac{d^3k}{2\omega_k} e^{i\vec{k}.\vec{r}}(2n_{\vec{k}} + 1) \tag{32}$$

and so summing over all states

$$< \Phi(r)\Phi(0) > \;=\; \int \frac{d^3k}{2\omega_k} e^{i\vec{k}.\vec{r}}\left(\frac{2}{(e^{\beta\omega_k} - 1)} + 1\right) \tag{33}$$

the second piece being the usual vacuum piece. For $r >> m^{-1} >> T^{-1}$ the integral in (33) is dominated by long wavelength modes and gives

$$< \Phi(r)\Phi(0) > \;\approx\; \frac{T}{4\pi r}e^{-mr} \tag{34}$$

This tells us that $\bar{\Phi}$ is uncorrelated on scales larger than $\xi \equiv m^{-1}_{eff}$, which is called the correlation length. If we defined $\bar{\Phi}_R$ as the average of Φ over a region of size $R >> \xi$ then it is straightforward to check that (34) implies the same result for $< \bar{\Phi}^2_R >$ as that obtained if $\bar{\Phi}_R$ were the result of averaging over $(R/\xi)^3$ volumes within which $\bar{\Phi}$ took on a randomly chosen values $\pm\sqrt{m_{eff}T}$. This serves as a rough picture of the distribution of Φ. Near the transition point the magnitude of the fluctuation, $\sqrt{m_{eff}T}$, is much less than T but the range, m^{-1}_{eff}, is much greater than T^{-1}. The energy in a typical fluctuation $\approx$ *potential* $\times$ *volume* $\approx m^2_{eff}\Phi^2\xi^3 \approx T$, as expected.

In fact if the system remains in equilibrium the correlation length goes to infinity at T_c, when $m_{eff} = 0$. In any real system with a finite rate of cooling, this never occurs. Instead the maximum correlation length is determined by the rate of cooling. I leave it as an exercise to find the maximum value of ξ by using the equilibrium value, with $T \propto a^{-1}$ and demanding that its rate of increase $\dot{\xi}/\xi$ not exceed the Hubble expansion rate $\dot{a}/a$.

2.4 Out of Equilibrium: Defect Formation

As the universe cools below a symmetry breaking phase transition, even though V_{eff} has the symmetry breaking form thermal fluctuations are initially still large enough to erase whole domains - in the example above a whole + domain may be 'flipped' to a − domain. Domains are really 'frozen in' very soon after the temperature falls below the energy required to kick them over the potential barrier. This is because below this temperature 'flipping' becomes exponentially (Boltzmann) suppressed, and the finite expansion rate of the universe means that it rapidly becomes negligible. The temperature at which domains are still frequently 'flipped' is given by

$$T \approx \xi^3 \Delta V_{eff} \approx \xi^3 \frac{\lambda}{8} \eta_{eff}^4 \tag{35}$$

being the energy required to take a domain ξ^3 over the potential barrier separating the two minima of V_{eff}. Using the equilibrium expressions right down to this temperature we find

$$\begin{aligned} T &= \frac{\eta}{8\sqrt{\lambda}}\left(1 - \left(\frac{T}{T_c}\right)^2\right)^{\frac{1}{2}} \\ \rightarrow T^{-1} &= \sqrt{\frac{64\lambda}{\eta^2} + \frac{1}{T_c^2}} \end{aligned}$$

$$\approx T_c \tag{36}$$

for small λ. This temperature, the temperature of domains freezing in, is called the *Ginzburg* temperature T_G. Below this temperature the domains and defects rapidly go out of thermal equilibrium[2].

At T_G we have $\xi = m_{eff}^{-1} = \eta_{eff}^{-1}/\sqrt{\lambda} \approx 1/(8T_c\lambda)$ which for λ not too small and $T_c \approx m_{Gut} << m_{Planck}$ is smaller than the horison scale at this time. For very small λ the assumption of thermal equilibrium would be dubious. However λ must be quite small for the small m/T expansion to be reasonable.

I have given a very simplistic account of what is a very complex subject. For one recent attempt to do better see [31].

2.5 Defects and Topology

In the previous section we discussed the simplest example of a symmetry, the discrete Z_2 symmetry $\Phi \rightarrow -\Phi$. In unified theories we are interested in more complicated symmetries including continuous symmetries. Nevertheless defect formation can be understood in a simple and general way.

Say we have a theory invariant under a symmetry group G which is spontaneously broken by a Higgs field Φ to a subgroup H, defined as $H = \{g \in G : g\Phi = \Phi\}$. Now the whole theory and in particular the Higgs potential $V(\Phi)$ is invariant under $\Phi \rightarrow g\Phi$. Thus if we have one point η on the minimum V_{min} of V we can obtain a whole set of values of Φ degenerate in energy by applying elements of G to η. In fact if there is no 'accidental' degeneracy in the theory the resulting set will be the entire space V_{min}. V_{min} is not however equal to G because if we apply two different grroup elements g and g' we may obtain the same point on V_{min}. This happens if $g\Phi = g'\Phi$ which is true iff $g^{-1}g' \in H$ so that g and g' are related by right multiplication in H . V_{min} is therefore obtained by ***identifying*** elements in G related by right multiplication in H - it is a ***coset space***

$$V_{min} = \frac{G}{H} \tag{37}$$

Thus the topology of V_{min} is simply a group theoretic entity, not depending on details of the Higgs potential or its coupling constants. In our simple example, $G = Z_2$ was broken to $H = 1$ by Φ and so from (37) $V_{min} = Z_2$, the two points + and −.

As another simple example, consider a theory with a continuous $G = U(1)$ broken by a complex scalar field Φ transforming under $g = e^{i\theta}$ as $\Phi \to e^{iQ\theta}\Phi$. Here $H = 1$ so from (37) $V_{min} = U(1)$ which is just a circle. The corresponding quartic potential, $V = \lambda(|\Phi|^2 - \eta^2)^2$ is the 'Mexican hat' potential. As a more complicated example if $G = SO(3)$ is broken by a 3 - component vector $\hat{\Phi}$ to a subgroup $H = SO(2)$ (rotations about the axis defined by $\hat{\Phi}$) then we find $V_{min} = SO(3)/SO(2) = S_2$, a two sphere.

The topology of V_{min} has very important physical consequences. In the 'domain' picture explained above, $\hat{\Phi}$ falls onto different points of V_{min} in different domains. Moving through space, Φ will wind around V_{min} and this will in general result in defects just as in Section 2.1. In the simplest case, V_{min} has two points + and −. In going along any path in space from a domain where $\Phi = +$ to one where $\Phi = -$, Φ has to climb over the potential barrier in between. This means that somewher there must be a localised wall of energy density separating the two domains, a 'domain wall'. In the second example above, as we traverse a loop in space Φ will in general wind

around the circle V_{min}. This means that somewhere inside the loop Φ must vanish. In fact Φ must vanish somewhere on every surface bounded by the loop. Where it vanishes, there is a nonzero energy density. The lowest energy configuration compatible with Φ winding around V_{min} on the loop is a *line* of energy density - a string - passing through the loop. Neither walls nor strings can have edges or ends - they must either be closed or infinite. In the third example above, traversing a two- sphere in space one would find in general that Φ would wind around the S^2 of V_{min}. This results in a localised *point* defect - a magnetic monopole (see Figure 2.3).

2.6 A Useful Theorem

A useful theorem in the case of strings is that the group of loops in $V_{min} - G/H$ not continuously deformable into one another (the homotopy group $\pi_1(G/H)$) is given by

$$\pi_1(\frac{G}{H}) = \pi_0(H) \tag{38}$$

in the case where G is simply connected. $\pi_0(H)$ is just the number of disconnected components of H. A pictorial 'proof' of this statement is given in Figure 2.4. This formula is still useful when G is simple but not simply connected because in this case one has $G/H = \tilde{G}/\tilde{H}$ where $\tilde{G}$ is the simply connected covering group of G and $\tilde{H}$ is the subgroup of $\tilde{G}$ leaving Φ invariant. For example $SO(3)$ can be broken by *two* 3-vectors to 1. This produces strings because the simply connected covering group of $SO(3)$ is $\tilde{G} = SU(2)$ and it would be broken to a subgroup $\tilde{H} = Z_2$, because the 3-vectors are in the adjoint representation of SU(2), and are left invariant by the two el-

perconducting strings form in theories where Φ couples to another electrically charged field χ in such a way as to force χ to be nonzero on the string. These strings may carry a net electric charge or current[41]. With the assumption of a primordial magnetic field (which is hard to justify) these strings are the basis for the OTW scenario of explosive galaxy formation[42]. I will not discuss this scenario in these lectures.

2.8 Simulating String Formation

Cosmic string formation was originally understood numerically by simply throwing down values for the Higgs field on V_{min} at random on a lattice of domains of size ξ, with a prescription for smoothly varying the Higgs field phase in going from one domain to the next. This is of course a crude model and ignores any interactions. The scale ξ is the correlation length at the Ginzburg temperature as discussed in section 2.4.

In more detail, for U(1) strings V_{min} is approximated by three points assigned the values 0,1 or 2. For a cubic lattice of domains, any edge is surrounded by four domains. The rule for going from one domain to the next is that Φ takes the shortest path along V_{min}. This is illustrated in Figure 2.6. The reader can easily check that at any vertex on the lattice the same number of strings leave as enter. With periodic or reflecting boundary conditions this means that the strings have no ends - they are all in the form of closed loops.

The numerical simulations first done by Vachaspati and Vilenkin [56] have been checked by many people on a variety of lattices and are in good agreement [49,46,52]. Most (about 80%) of the string is in strings as large as the

box in which the simulation is performed ('infinite' strings). The remainder is in the form of closed loops. These have a *scale invariant* distribution - the number of loops of size (a *course* measure like r.m.s. radius) greater than R per unit volume, $n_>(R) \propto R^{-3}$. The term scale invariant simply means that the result depends only on R and not on the domain size ξ. Dimensional analysis then determines the power law. Both the 'infinite' strings and the loops are in the form of Brownian random walks. Recently David Mitchell and myself have shown how these results may be obtained by counting states in the quantised closed bosonic string, an intriguing connection I shall explain in Lecture 3.

A picture of the string network at formation is shown in Figure 2.7.

2.9 Evolution of Strings

After the strings form we have to evolve them. At first sight this appears horribly complicated - they are inherently nonlinear entities. However two effects straighten out the strings very rapidly. In the early stages (until the temperature falls to $\approx (G\mu)^{\frac{1}{2}} m_{GUT}$) the strings are heavily damped by *collisions with particles*[35]. Secondly the *expansion of the universe stretches* the string out. The typical radius of curvature of the string increases rapidly while the width of the string remains fixed.

Quite quickly it becomes a very good approximation to treat the strings as infinitely thin relativistic lines or 'Nambu' strings, the action for which is simply the area of the two dimensional worldsheet they sweep out in spaceements of its centre. Thus we would have $\pi_1(V_{min}) = Z_2$. These strings are more complicated than the $U(1)$ strings discussed above - in this case a string running in one direction is topologically equivalent to one running in the opposite direction. Strings of this type form in many simple grand unified theories, for example in the symmetry breaking transitions

$$\begin{aligned} SO(10) &\rightarrow \text{`}SU(5)XZ_2\text{'} \\ E(6) &\rightarrow \text{`}SO(10)XZ_2\text{'} \end{aligned} \qquad (39)$$

the first being mediated by a Higgs **126** and the second by a **351**. The quotation marks indicate that the full global structure of H is more complicated, as explained in [32].

In low energy models 'inspired' by superstrings cosmic strings of the simplest U(1) type as well as those of the more complicated Z_n type occur[33].

As an excercise I leave the reader to construct the 'proof' analogous to that in Figure 2.4 that $\pi_2(\frac{G}{H}) = \pi_1(H)$ assuming $\pi_2(G)$ and $\pi_1(G)$ are trivial. The latter are always true for simple Lie groups. Since H must ultimately contain the $U(1)$ of electromagnetism, if the grand unified theory is based on a simple group then magnetic monopoles are inevitable. This is the basis of the famous 'magnetic monopole problem' [34].

2.7 Cosmic Strings

Strings are the most interesting topological defects because their existence is compatible with the observed universe and may even explain its large scale structure. Their possible existence in relativistic field theories was first pointed out by Nielsen and Olesen [45]. They are simple generalisations of

the flux lines found in superconductors. For a review, see [35].

The microscopic structure of a string is illustrated in Figure 2.5 for the simplest case, a U(1) string. Outside the string, Φ winds around V_{min} as we go around the string. At the centre of the string Φ goes to zero. The 'magnetic' gauge field of the $U(1)$ group that is broken in forming the string is nonzero and points along the string in one direction in the minimum energy configuration. The energy per unit length μ of a static straight string is obtained from the energy functional

$$\begin{aligned} \mu \equiv E/L &= \int d^2x[(\vec{D}\Phi)^2 + \frac{1}{2}\vec{B}^2 + V(\Phi)] \\ &= 2\pi\eta^2 f(\frac{\lambda}{g^2}) \approx m_{Gut}^2 \\ V(\Phi) &= \frac{\lambda}{2}(|\Phi|^2 - \eta^2)^2 \\ \vec{D}\Phi &= (\vec{\nabla} - ig\vec{A})\Phi \end{aligned} \qquad (40)$$

where f is a slowly varying function equal to unity when $\lambda = g^2$ [38,39]. These strings generally carry no electric charges and their strongest coupling to other matter is through their gravitational field. The gravitational coupling is thus the only parameter relevant to the (ordinary) cosmic string theory of galaxy formation. It is given by the dimensionless number $G\mu \approx (m_{Gut}/m_{Planck})^2 \approx 10^{-6}$ typically in those grand unified theories predicting strings.

In some theories the string can couple to long range fields other than gravity. *Global* strings occur where the $U(1)$ symmetry broken by Φ is not gauged. Here there is a massless Goldstone boson which couples to the string and is radiated from it. This kind of string occurs in axion models[40]. *Su-*

time [43]

$$\begin{aligned} S &= -\mu \int dA = -\mu \int d^2\sigma \sqrt{-\gamma} \\ \gamma &= det(\gamma_{\mu\nu}) \quad \gamma_{\mu\nu} = \partial_\mu x^a(\sigma)\partial_\nu x^b(\sigma) g_{ab}(x(\sigma)) \end{aligned} \tag{41}$$

where $g_{ab}(x)$ is the metric of the background spacetime. $x^a(\sigma)$ are the four spacetime coordinates of the string worldsheet which is parametrised by the two worldsheet coordinates $\sigma^\mu = (\tau, \sigma)$. The string worldsheet has one spacelike coordinate σ and one timelike coordinate τ.

Let me explain how (41) arises. For a $U(1)$ theory producing strings the action is

$$\begin{aligned} S &= \int d^4y \sqrt{-g}((D\Phi)^2 - V(\Phi) - \frac{1}{4}F^2) \\ V(\Phi) &= \frac{\lambda}{2}(\mid \Phi \mid^2 - \eta^2)^2 \\ D_\mu \Phi &= (\partial_\mu + igA_\mu)\Phi \\ F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu \end{aligned} \tag{42}$$

We seek an approximate classical solution to the full field theory equations in the form of a curved and moving string. Locally it should look like a Lorentz boosted version of the static cylindrically symmetric solutions obtained by minimising the energy functional (40). It should also be built around a curved worldsheet $x^a(\sigma)$.

First we imagine such a worldsheet is given and construct two normal vectors to it, $n^a_A(\vec{\sigma}), A = 1, 2$. These obey $n^a_A x^b_{,\mu} = 0$ and are orthonormal $n^a_A n^b_B g_{ab} = -\delta_{ab}$. Now for any point in space y^a closer to the string than its radius of curvature R we can associate two worldsheet coordinates σ^μ and

two radial coordinates ρ^A uniquely

$$y^a = x^a(\sigma) + \rho^A n^a_A(\sigma) \tag{43}$$

This is illustrated in Figure 2.8. We will use this relation to change variables from the four y^a coordinates to the four $\xi^\alpha = (\sigma^\mu, \rho^A)$ coordinates. Our ansatz for the field configuration corresponding to this worldsheet is the $\Phi(y) = \Phi_s(\rho)$ and $A^a(y) = n^a_A(\sigma)A^A_s(\rho)$ where the subscript s refers to the static cylindrically symmetric solution. The reader can check that this ansatz gives the exact solution for a straight moving string (i.e. just the Lorentz boosted version of the static straight solution). We expect that it will provide us with an approximate solution if the radius of curvature of the string R is much greater than its width W. Now we simply change coordinates in the action (42) from y^a to ξ^α. The Jacobian is most easily calculated by noting that

$$\sqrt{-g} det(\frac{\partial y}{\partial \xi}) = \sqrt{-det(g_{ab}\frac{\partial y^a}{\partial \xi^\alpha}\frac{\partial y^b}{\partial \xi^\beta})} \tag{44}$$

and

$$\begin{aligned} M_{\alpha\beta} &= g_{ab}\frac{\partial y^a}{\partial \xi^\alpha}\frac{\partial y^b}{\partial \xi^\beta} \\ &= diag(\gamma_{\mu\nu}, -\delta_{AB}) + o(\rho/R) \end{aligned} \tag{45}$$

where worldsheet derivatives of x^a are down by order ρ/R. Thus $det M = det\gamma$ to order ρ/R. We also have to the same order $(D\Phi)^2 = M^{\alpha\beta}D_\alpha\Phi D_\beta\Phi = -(D_A\Phi_s)^2$ and $F^2 = F^2_s$ so when the $d^2\rho$ integral is performed the Lagrangian reduces to minus the energy per unit length, μ. The terms of order ρ are down by W/R after integrating ρ across the string profile, where W is the string width. Thus to lowest order we find the action (41). For loops seeding galaxies R is of the order of 10 parsecs and W is of order $m_{gut}^{-1} \approx 10^{-33} cm$ so W/R is indeed pretty small! However in the very small regions of the string worldsheet like 'cusps' (see below) where the worldsheet becomes highly curved we expect the corrections to be important.

In fact the $o(W/R)$ corrections to (41) vanish by symmetry - the first nontrivial corrections arise at order $(W/R)^2$. They are

$$\Delta S = \mu \int d^2\sigma(\alpha K^2 - \beta\omega^2) \tag{46}$$

where K is the extrinsic curvature of the worldsheet and w the 'twist' along the string [44]. α and β are positive constants of order W^2. The sign of the extrinsic curvature term corresponds to a *negative* rigidity - cosmic string likes to be crinkled rather than straight. This might lead to interesting effects in the late time evolution of loops, where presumably it will be counteracted by the gravitational radiation acting to reduce high frequency modes on the string.

The nicest feature of the Nambu action is that it is completely geometrical. The parameter μ does not enter into the equations of motion at all. The motion of the string just depends on the background spacetime metric. Waves propagate along the string with the speed of light.

Of course our discussion breaks down where two strings collide. In this case one has to solve the full nonlinear field equations. However it really is a very local process (the string fields approach the vacuum exponentially rapidly outside the string) and can be done perfectly well in flat spacetime. Numerical calculations by Shellard [36](for global strings) and Matzner [37]

(for local strings) show that when two strings collide they always reconnect the other way (see Figure 2.9) for velocities less than .95 in the centre of mass frame, i.e. essentially always. This is a very nice result because the string interactions are therefore also fixed and cannot be adjusted. For strings with more complicated internal structure like superconducting strings this may not be the case.

2.10 Free Strings in an Expanding Universe

We know from present observation that at early times Ω was very close to unity and that the universe was radiation dominated and quite homogeneous. Thus the background metric in which we have to evolve the strings is completely specified. The metric is

$$ds^2 = dt^2 - a^2(t)d\vec{x}^2 = a^2(\eta)(d\eta^2 - d\vec{x}^2) \tag{47}$$

with $a(t) \propto t^{\frac{1}{2}} \propto \eta$ and η is the *conformal* time, in which the metric is conformally related to the Minkowski metric. The action (41) is invariant under *reparametrisations* of the string worldsheet. This means that to obtain a well defined initial value problem one has to choose a gauge. A convenient gauge is

$$\begin{aligned} x^0(\tau,\sigma) &= \tau \\ \partial_\tau \vec{x}.\partial_\sigma \vec{x} &= 0 \end{aligned} \tag{48}$$

(i.e. choosing τ to be the time and choosing σ at each instant so that the motion of the point labelled by σ is perpendicular to the string). Upon varying (41) one finds in this gauge the equations[47]

$$\begin{aligned}
\ddot{\vec{x}} + 2\frac{\dot{a}}{a}\dot{\vec{x}}(1-\dot{\vec{x}}^2) &= \frac{1}{\epsilon}\partial_\sigma\left(\frac{\vec{x}'}{\epsilon}\right) \\
\dot{\epsilon} &= -2\frac{\dot{a}}{a}\epsilon\dot{\vec{x}}^2 \\
\epsilon &= \sqrt{\frac{\vec{x}'^2}{1-\dot{\vec{x}}^2}}
\end{aligned} \tag{49}$$

where dots refer to τ derivatives and primes to σ derivatives. The quantity ϵ has a simple meaning - the string energy per parameter σ length (see Section 2.14 below) is given by $\mu a(\eta)\epsilon$. The numerator in ϵ is just the comoving length of the string per unit σ, corresponding to a potential energy μ times (physical length). The denominator contributes the kinetic energy of the string.

In (49) there are two competing forces. On the left hand side is a damping term due to the expansion of the universe. On the right hand side is the curvature term which accelerates the string. The result of the competition between these terms is that for curvature scales R on the string much *smaller* than the Hubble radius $R_H = H^{-1} = a/\frac{da}{dt}$ the curvature wins and the string evolves as in flat spacetime. One finds in the case [47]

$$E \approx const \rightarrow \rho \propto a^{-3} \quad R << R_H \tag{50}$$

Highly curved strings behave like *matter*. For curvature scales much *greater* than the Hubble radius the damping term wins so $\dot{\vec{x}} \approx 0$ and the string is *conformally* stretched, its shape remaining fixed but its length growing like a. Thus

$$E \propto a \rightarrow \rho \propto a^{-2} \quad R >> R_H \tag{51}$$

Either of these two behaviours would be cosmologically disastrous - the string would quickly come to dominate the radiation background (for which $\rho \propto a^{-4}$). However this is not the end of the story. Loops smaller than the Hubble radius are not a problem because they slowly radiate themselves away into gravity waves for which $\rho \propto a^{-4}$. As we shall see the long strings which would cause problems if there were no string-string interactions gradually chop themselves up into loops which then radiate themselves away.

2.11 Inside the horizon: Free Strings in Flat Spacetime

In flat spacetime the string equations are particularly simple and can be solved analytically as follows [105,48]. From (49) we have $\dot{\epsilon} = 0$. If we initially parametrise the string so that $\epsilon(\sigma)$ is equal to 1 along the string then it will remain so for all time. We then have

$$\begin{aligned}
\ddot{\vec{x}} &= \vec{x}'' \\
\dot{\vec{x}}.\vec{x}' &= 0 \quad \dot{\vec{x}}^2 + \vec{x}'^2 = 1
\end{aligned} \tag{52}$$

The general solution is given by

$$\begin{aligned}
\vec{x} &= \frac{1}{2}(\vec{a}(\sigma+\tau) + \vec{b}(\sigma-\tau)) \\
\vec{a}'^2 &= \vec{b}'^2 = 1
\end{aligned} \tag{53}$$

Here $\vec{a}'$ and $\vec{b}'$ are arbitrary curves on the suface of a unit sphere. For a closed loop in its centre of mass frame $\vec{a}'$ and $\vec{b}'$ are in addition constrained to have $\int d\sigma\vec{a}' = \int d\sigma\vec{b}' = 0$. Some simple nonintersecting Fourier mode solutions [55] are shown on Figure 2.10 (courtesy A.Stebbins).

It is easy to check that a closed loop has period $L/2$ where L is the parameter length in this gauge. Some of the trajectories in Figure 2.10 have 'cusps' - points where $|\vec{x}'| = 0$ and the string reverses back on itself. From (52) these points have $|\dot{\vec{x}}| = 1$ - they are instantaneously moving at the speed of light![55]. Gravitational radiation is 'beamed' out from these points[55,58].

2.12 Interacting String networks in an Expanding Universe: the Scaling Solution

Now we proceed to the full problem of the evolution of an interacting string network in an expanding background. We simply take the initial configuration as shown in Figure 2.7 and evolve it with the equations (49) and the rule for interactions that whenever two strings cross we reconnect them the other way as in Figure 2.9.

A. Albrecht and myself have been developing a numerical code to do this for several years [49]. We have recently made qualitative improvements over our early code and brought the results to what (for us!) is a rather convincing state. Bennett and Bouchet [52] have also developed a code and our latest code agrees at least in part with theirs.

A picture of part of one of our simulations is shown in Figure 2.11 . We begin with initial conditions as in Figure 2.7 in a 24^3 or 30^3 box, with 10 points on the string for every initial 'correlation length' ξ. We then evolve (49) numerically. The only parameter to vary in the simulation is the ratio of ξ to the Hubble radius $R_H \equiv a/\frac{da}{dt} = 2t$. The initial string density is approximately $\rho_s = \mu/\xi^2$, so starting at large or small R_H/ξ corresponds to starting at high or low string density relative to the background radiation

density ($\rho_b = \rho_{radiation} = 3/(32\pi G t^2) = 3/(8\pi G R_H^2)$).

It is useful to separate the strings into 'long string' , strings longer than the Hubble radius, and 'loops', strings shorter than the Hubble radius. The precise dividing line is not very important.

Figure 2.12 shows how the density in long string (defined as loops for whom E/μ is greater than $2R_H$) behaves for differing initial string densities. A good way to think of this is that if there is a fixed number of long strings of length $\approx R_H$ per Hubble volume R_H^3 then the long string density scales with time as R_H^{-2} so the ratio ρ_s/ρ_b is constant. As can be seen this ratio does indeed appear to approach a constant and $\rho_s \approx 200/R_H^2$.

Why is this? The main idea of the 'scaling solution' [49,50,51] is that one should define a length scale ξ by

$$\rho_L = \frac{\mu}{\xi^2} \tag{54}$$

where ρ_L is the density in long strings. As long as reconnection keeps the network 'random' then ξ should also be the typical curvature scale on the string. Now as long as $\xi << R_H$, the strings should evolve as in flat space so that the energy in the long strings (ignoring chopping off loops) scales like matter. Also, the timescale for the long strings to chop off a fraction of their length into loops is simply ξ (recall the characteristic velocity of waves on the string is just the speed of light which we set equal to 1). Thus ignoring reconnections of loops we find

$$\dot{\rho}_L = -3\frac{\dot{a}}{a}\rho_L - c\frac{\rho_L}{\xi} \tag{55}$$

with c a constant (here dot refers to $\frac{d}{dt}$ not $\frac{d}{d\tau}$). The first term is simply a result of the expansion, the second the loss of energy into loops. Ignoring reconnections is reasonable in the light of the results of the next lecture, where I will show that in flat space chopping off is heavily favoured by phase space over reconnection .

Equation (55) has a simple solution: for a radiation dominated background $a \propto t^{\frac{1}{2}}$ it is given by

$$\rho_{sc} = \frac{\mu}{c^2 R_H^2} = \frac{\mu}{c^2 (2t)^2} \tag{56}$$

This is called the ***scaling solution.***

In fact the scaling solution is a stable fixed point of (55) as may be seen by defining $y = (\rho_L / \rho_{sc})$ and calculating

$$\dot{y} = \frac{y}{R_H}(1 - \sqrt{y}) \tag{57}$$

If y is greater than unity then it decreases, but if it is less than unity it increases. Thus $y = 1$ is a stable fixed point. Taking the initial scale factor a_i to be unity (57) may be solved for an initial value y_i to yield

$$y = \frac{y_i \sqrt{a}}{(1 - y_i(1 - \sqrt{a}))} \tag{58}$$

Thus y approaches unity rather slowly.

The dashed lines on Figure 2.12 show the predicted approach to scaling from (58) with y_i chosen soon after the initial time (there is a short initial transient due to the fact that we start our strings off stationary in comoving coordinates). As can be seen the results fit (58) rather well. Figure 2.11 shows the string configuration in a carved out Hubble volume (actually $(R_H/2)^3$) in the scaling solution.

2.13 Loops

The distribution of loops produced in the scaling solution plays a very important role in the cosmic string theory of large scale structure formation. The importance of loops was first emphasised in [53] by Vilenkin, and the simple identification of one loop with one object today proposed in [54,64,66].

The energy lost from long strings in (55) goes into loops. Defining $\rho_l(e)$ to be the density contributed by loops of energy e to $e + de$ we have an equation

$$\dot{\rho}_l(e) = -3\frac{\dot{a}}{a}\rho_l + \frac{1}{R_H^2} f(e, t) \tag{59}$$

which defines the dimensionless *loop production function* f. In the scaling solution there is only one length and time scale in the problem, the Hubble radius $R_H = 2t$ in the radiation era. Thus f being dimensionless can only be a function of $e/(\mu R_H)$. Some of the loops produced will further selfintersect. I will ignore this complication here by regarding f as the loop production function for non-self-intersecting loops. Integrating (59) we find

$$\begin{aligned} \rho_l(e) de &= \left(\frac{\mu R_H}{e}\right)^{\frac{1}{2}} \frac{\lambda de}{R_H^2} \\ \lambda &= \frac{1}{2} \int_0^{x_C} dx x^{\frac{1}{2}} f(x) \end{aligned} \tag{60}$$

for loops with $e/(\mu R_H) << 1$ and we define $f(x)$ to be zero for $x > x_C \approx 1$ · such loops are simply part of the long string. Similarly the number of loops per unit volume with energy e to $e + de$ is

$$n_l(e) de = \left(\frac{\mu R_H}{e}\right)^{\frac{1}{2}} \frac{\lambda de}{e R_H^2} \tag{61}$$

Our original simulations found $\lambda \equiv \nu (2\beta)^{\frac{1}{2}}$ (in the old notation; see [66]) approximately equal to 0.8. The results of our latest simulations will shortly be revealed [49]. Equation (60) shows that the loop density scales like matter - the small loops have been around longest so they dominate the total density in strings. This would be a disaster if the loops were totally stable. However they do slowly decay by emitting gravity waves, at a rate which is exactly calculable for simple loop trajectories [55,57,59]. The rate of energy loss is given by $\dot{\epsilon} = -\Gamma G\mu^2$ independent of the size of the loops, with Γ generically of order 50-100 for simple loops. The energy of a loop thus decreases slowly with time: $\epsilon(t) = \epsilon_i - \Gamma G\mu^2(t - t_i)$ where ϵ_i and t_i are the initial values. For $t >> t_i$ this yields the final loop distribution

$$n_l(\epsilon,t)d\epsilon = (\frac{\mu}{R_H})^{\frac{3}{2}} \frac{\lambda d\epsilon}{(\epsilon + \Gamma G\mu t)^{\frac{5}{2}}} \tag{62}$$

These loops act as gravitating 'seeds' which start accreting matter at around t_{eq} in the manner described in Lecture 1. The mass a loop accretes is proportional to it's mass (with cold dark matter and as long as it was produced before t_{eq}, see below). Smaller loops are more numerous and give rise to smaller objects like galaxies. Larger loops are rarer and give rise to larger objects like clusters of galaxies. Thus (62) gives a simple prediction for number density versus mass for objects in todays universe (though this is in general altered by merging).

2.14 Strings as Density Perturbations

Cosmic strings of the simplest kind interact with matter mainly through gravity. As discussed in Lecture 1 density perturbations only start growing appreciably when the universe becomes matter dominated. To see how strings cause matter perturbations to grow in their vicinity we need to cal-

culate their gravitational field.

For simplicity I will only discuss strings in a flat background in the weak field limit (the gravitational potential is of order $G\mu << 1$ so this is a very good approximation most of the time for cosmic strings. It may fail at 'cusps'). The string stress tensor is given from the Nambu action as

$$\begin{aligned} T^{ab}(t,\vec{x}) &= -\frac{2}{\sqrt{-g}}\frac{\delta S}{\delta g_{ab}}\Big|_{g_{ab}=\eta_{ab}} \\ &= \mu\int d\sigma \begin{pmatrix} 1 & \dot{x}^i \\ \dot{x}^j & \dot{x}^i\dot{x}^j - x'^i x'^j \end{pmatrix} \delta(\vec{x} - \vec{x}(t,\sigma)) \end{aligned} \tag{63}$$

in the gauge (48) where η_{ab} is the Minkowski metric and $\vec{x}(t,\sigma)$ the string trajectory in space. For a straight static string along the z axis , $\vec{x}(t,\sigma) = (0,0,\sigma)$, we find $T_{00} = -T_{33} = \mu\delta^2(\vec{x})$ with other components zero. The sign of the T_{33} term corresponds to a *negative* pressure i.e. a tension along the z axis.

To calculate the metric perturbation produced by the string we write the metric as $g_{ab} = \eta_{ab} + h_{ab}$ with $| h |\approx G\mu << 1$. In the harmonic gauge $g^{ab}\Gamma^c_{ab} = 0$ the linearised Einstein equations become [6]

$$\partial^2 h_{ab} = -16\pi G(T_{ab} - \frac{1}{2}\eta_{ab}T^c_c) \tag{64}$$

The usual Newtonian potential is just $h_{00}/2$: for our straight string it is easily seen that the source for h_{00} vanishes exactly! The only regular solution is $h_{00} = 0$ everywhere, so there is no Newtonian potential outside the string. h_{33} also vanishes but there is a nonzero h_{11} and h_{22};

$$\begin{aligned} \nabla^2 h_{11} &= \nabla^2 h_{22} = 16\pi G\mu\delta(\vec{x}) \\ \Rightarrow h_{11} &= h_{22} = 8\pi G\mu ln(r/W) \end{aligned} \tag{65}$$

smoothed out and the object become more and more spherical. The shape traced out by a loop is in fact like a 'trumpet' [70]. This is seen as follows. The peculiar velocity of a loop $v_{pec} = a(t)\dot{\vec{x}}$ decays [1] as a^{-1}. This means that, in the matter dominated era, the comoving distance a loop moves is given by

$$\Delta x = \int_{t_i}^{t} v_i(\frac{a_i}{a})^2 dt = 3v_i t_i (1 - (\frac{a_i}{a})^{\frac{1}{2}}) \tag{70}$$

where v_i is the initial velocity, at scale factor a_i and time t_i. Thus a loop slows down and stops in comoving coordinates in the matter era (in the radiation era Δx increases logarithmically with time). However the radius of the loop in comoving coordinates r_c shrinks (the *physical* radius is approximately constant, ignoring oscillations) so

$$r_c = r_i \frac{a_i}{a} \tag{71}$$

So defining the distance from where the loop ends up as $z = 3v_i t_i - \Delta x$ we see that the shape traced out by the loop is

$$r_c \propto z^2 \tag{72}$$

the surface of revolution of a parabola - a 'trumpet' shape. As stated above, nonlinearity smears this out, making the accreted lump in the end rather spherical [65]. For galaxies this would be the case today. However larger clusters should still exhibit this shape - in fact the Perseus- Pisces filament [3] does look rather 'trumpet' like! Bingelli has made an attempt to measure the eccentricity for Abell clusters - this will serve as an interesting test in the future[10].

[1] More precisely the momentum $\int d\sigma \epsilon \dot{\vec{x}}$ decays as a^{-1} as can be seen from (49), but if its velocity is substantially less than 1 and it is well inside the horizon then the energy is roughly constant.

2.16 Loops as seeds for galaxies and clusters

The string plus cold dark matter (SCDM) scenario is the easiest to understand. It is clear from Section 1 that the *total mass* accreted by a loop around at t_{eq} is today roughly $M_{loop}(1 + Z_{eq})$. In fact a precise calculation in the spherical collapse model yields [67]

$$m_{loop}(1 + Z_{eq}) = \frac{\xi}{5} M(\frac{\rho}{\rho_b})^{\frac{1}{2}} \tag{73}$$

where M is the mass of the accreted object, ρ/ρ_b is the overdensity today and ξ is a 'loss of growth' factor, equal to 1 for loops laid down well before t_{eq}, equal to 4 for loops laid down at t_{eq} and greater for loops laid down still later.

It is now simple to calculate the required string tension. First we pick a class of objects, for example Abell clusters. Their number density is $1/(110h_{50}^{-1}Mpc)^3$. Let me assume that the loops seeding them are all produced in the radiation era for simplicity (they are actually produced at or just after the matter-radiation transition according to our latest simulations). Now we take the loop distribution, given in equation (62), and evolve it through the matter - radiation transition to the present day. We then ask for what cutoff in energy will the integrated number density be equal to $1/(110h_{50}^{-1}Mpc)^3$. Doing this[66,67], one finds that the length of loops producing clusters is $L = \epsilon/\mu \approx 20kpc$. We can then demand that loops of this size are massive enough to accrete clusters by today, from (73) and the observed mass and overdensities of clusters [67]. One finds

$$G\mu = 2.10^{-6} h_{50}^{-1} \sigma_{700}^{\frac{3}{2}} d_{110}^{-2} \lambda^{-\frac{3}{2}} \xi_4 \tag{74}$$

choosing $h_{11} = h_{22} = 0$ at the string width W. Inside the string the true stress energy is not singular like the 'Nambu' approximation (63) and this removes the apparent singularity at $r = 0$ [63]. Clearly the weak field approximation breaks down at very large radii as well. Ignoring this for the moment we have in cylindrical coordinates

$$ds^2 = dt^2 - dz^2 - (1 - 8G\mu ln(r/W))(dr^2 + r^2 d\phi^2) \tag{66}$$

We can remove the prefactor to the dr^2 term by defining a new radial variable $dr' = dr(1 - 4G\mu ln(r/W))$ or $r' = r(1 + 4G\mu(1 - ln(r/W))$. Neglecting terms of order $(G\mu)^2$, consistent with our approximation, we then find

$$\begin{aligned} ds^2 &= dt^2 - dz^2 - dr'^2 - r'^2(1 - 8G\mu)d\phi^2 \\ &= dt^2 - dz^2 - dr'^2 - r'^2 d\phi'^2 \end{aligned} \tag{67}$$

upon defining $\phi' = (1 - 4G\mu)\phi$. This is not only regular everywhere, it is flat space! However the new angular variable ϕ' only runs from 0 to $2\pi - 8\pi G\mu$ - the geometry is that of 'flat space minus a wedge', or a cone [61,60]. This is illustrated in Figure 2.13. The only difference is at radii of order W, the string width, where the apex of the cone is smoothly rounded out[62,63].

What about loops? These *do* produce a Newtonian potential - one finds

$$\partial^2 h_{00} = -16\pi G\mu \int d\sigma \dot{\vec{x}}^2 \delta^3(\vec{x} - \vec{x}(\sigma, t)) \tag{68}$$

For an oscillating loop the right hand side oscillates peiodically in the centre of mass frame of the loop. This leads to a peculiar pulsating gravitational field[54]. It's time average however is simple, because of the following

identity. If the length of the loop is L and its period T we have

$$\begin{aligned} \int \frac{d\sigma}{L} \int \frac{dt}{T} \dot{\vec{x}}^2 &= \int\int -\vec{x}.\ddot{\vec{x}} = \int\int -\vec{x}.\vec{x}'' \\ &= \int\int \vec{x}'^2 = \int\int (1 - \dot{\vec{x}}^2) \\ &\Rightarrow \int\int \dot{\vec{x}}^2 = \frac{1}{2} \end{aligned} \tag{69}$$

where we used integration by parts and the equations (52). The average velocity squared for a closed loop is thus $\frac{1}{2}$. Now if we time average (68) we find the time averaged long range field of a loop is that of a shell of Newtonian dust with suface density proportional to $\dot{\vec{x}}^2$ and with total mass equal to μL, the mass of the loop. At distances from the loop much greater than its size we can treat it to a good approximation as a point mass.

Vachaspati [58] has tried to use the time variation of h_{ab} and the *beaming* effect mentioned earlier [55] to argue that an accreting loop will eject matter into an interesting shape. However the actual amount of mass which is in the region where h is so large that it moves the matter substantially during the existence of a cusp is tiny, so most of the matter around the loop simply sees the time averaged field, which shows no beaming effect [55].

2.15 Accretion by Loops

Accretion by moving loops has been looked at in some detail in [65] and [70]. In linear theory, the equation for the density perturbation $\delta\rho/\rho(\vec{x})$ is purely local in comoving coordinates, as noted in Lecture 1. This means that the surface 'traced out' by the loop as it moves through space will appear in the density perturbation it produces. This is only true while the perturbation is linear - as soon as collapse and virialisation occur the structure will be

Here σ_{700} is the velocity dispersion in a cluster in units of 700 kms^{-1}, d their mean separation in units of 110 h_{50}^{-1}Mpc, λ the constant in (62) and ξ the growth factor mentioned before. There are clearly large uncertainties in the normalisation of $G\mu$, probably of an order of magnitude either way. In particular there are several effects which might reduce $G\mu$. If clusters form by mergers of several loops, or directly from curved sections of the long strings, this would decrease the required value. This is one of the main motivations for the more detailed calculations now being performed in which we are directly evolving the matter distribution around a string network produced in the simulations[71].

Independently of $G\mu$ however, the ***spatial distribution*** of the loops and long strings should have testable consequences for the spatial distribution of galaxies and clusters, particularly clear cut on scales larger than those where gravitational clustering has been efficient (i.e. greater than 12 h_{50}^{-1} Mpc or so, see Lecture 1). Unlike theories like the ICDM model, there are no parameters to adjust here - the evolution of the strings is independent of the string tension. An early calculation [64] showed that remarkably enough the two-point correlation function of loops produced in a cosmic string simulation (in the radiation era) matched the observed cluster-cluster correlation function very closely. Subsequent simulations by Bennett and Bouchet, and those by Albrecht and myself using our improved code, have confirmed this correlation of loops when they are produced. The more detailed calculations using improved codes and including the radiation-matter transition and the effect of the long strings will show whether this apparently miraculous agreement

with observation is confirmed. [2]

2.17 Cosmic String and Hot Dark Matter

It has recently been shown that cosmic string could solve the usual problems of the standard Hot Dark Matter (IHDM) scenario [25,83] (see also the earlier work of [112]). I will call this the SHDM scenario. This is nice - the massive neutrino is still the best motivated dark matter candidate. Of course exactly why $m_\nu = 30eV$ as is required is still a mystery (as are the other fermion masses!).

Without going into details, the reason for the improvement is that loops survive the free streaming era and begin accreting baryons slowly at around recombination. The baryon/loop lumps then start accreting neutrinos efficiently when the neutrino Jeans mass falls to galaxy scales. Flat rotation curves are obtained automatically (once heralded as a success of the standard ICDM scenario). The loss of growth on galaxy scales means that a higher value of $G\mu$ is needed - this brings the theory more into line with the observed large scale velocity fields [78]. Clusters and their correlations are similar to SCDM[25,26,83].

[2] It has been argued, in particular by Bennett and Bouchet[68], that if one extrapolates the loop positions for several expansion times then the correlations are washed out by loop velocities. I believe this is a red herring - loops forming clusters are produced at or just after t_{eq} - accretion begins as soon as the loops are produced. Thus the *initial* positions of the loops is most relevant. Second, some (e.g. Scherrer and Press[69]) have argued that loop fragmentation produces fragments with fairly high average velocities $\approx 0.6c$, so one can use randomly distributed loops to calculate structure formation. However it is the larger loops around (and even those that will subsequently self-intersect) at or after t_{eq} that will form clusters, and these have much lower velocities [49]. String plus matter simulations will hopefully resolve the issue.

2.18 Cosmic String and Inflation?

How do cosmic strings fit in with inflation? Many cosmologists today take the idea of inflation very seriously (some even refer to it as a 'paradigm') because it 'explains' why the universe is so large, so homogeneous and so flat today.

However inflation is not a gospel. Many people including myself feel rather unhappy with the 'explanations' it offers. At a general level, it is certainly true that inflation will turn a wider class of initial conditions into a large homogeneous universe like our own. However it certainly will not turn *arbitrary* initial conditions into universes like ours and until it is possible to say precisely 'how much better' inflation makes things, one really cannot judge the idea.

It also seems very likely that it will be quite some time before we can say anything very definite about the initial conditions of the universe that is in any way testable or provable. The whole discussion may well simply remain in the realm of philosophy rather than physics.

In cosmology today there is a strong case for being conservative in our ambitions. It has always been easy in cosmology to get into the realms of untestable speculation. What is new today is that there is a rapidly increasing data base upon which theories can be tested. And there are now at least two rather precise theories to test. This is the most hopeful aspect of the subject. Now the aspect of inflation that is certainly a major advance from this point of view is that the fluctuation spectrum it 'predicts' (at least in the simplest version [73]) has definite consequences for the galaxy distribution,

the microwave background fluctuations and so on. Cosmic strings are from this point of view also a very good theory, with clear testable predictions.

In fact one can make the case that cosmic strings are a much better justified extrapolation of low energy physics. The underlying successful idea in high energy physics has been that of gauge symmetry and spontaneous symmetry breaking. Cosmic strings are fairly generically (though certainly not always) produced in symmetry breaking phase transitions, and the precise details of the symmetry breaking (i.e. whether it occurs dynamically via a condensate or through a scalar 'Higgs field') are not crucial. On the other hand it is very difficult to make inflation work in a gauge symmetry breaking phase transition. Instead one has to use a gauge *singlet* inflaton field which 'allows' one to tune the coupling constant to very small values. This raises real questions about the initial distribution of the inflaton field - it was almost certainly not in thermal equilibrium. It also makes the subject very ad hoc.

I am often faced with the question 'well how do cosmic strings explain the flatness and homogeneity of the universe?' to which I reply 'they don't'. It is perfectly legitimate in seeking to explain the large scale structure in the universe to use the fact that we *know* the universe was homogeneous and flat at early times, from the microwave background. This is exactly what we do in nucleosynthesis calculations for example.

Inflation does not go well with cosmic strings (as you might have guessed from my critical comments!). During inflation the string density would fall exponentially so that after inflation there would be far fewer than one string per horizon volme. Indeed that is how inflation solves the monopole problem (to which there are other solutions [74,75,76]). For the reasons I have given, this does not bother me too much.

If you want to use inflation to solve the monopole problem, but have cosmic strings provide the fluctuations later on, you have to 'stand on your head'. You must first form the monopoles, inflate them away and then reheat to a temperature high enough to form strings but too low to form monopoles. This is tricky - in most inflationary models the universe heats up to temperatures well below the 10^{16} GeV scale typically needed to form strings. However in cosmic strings the mass per unit length $\mu \approx \eta^2$ is only weakly dependent on couplings whereas the phase transition temperature $T_c \approx \sqrt{\lambda/g^2}\eta$ can be made much lower by decreasing λ or imposing discrete symmetries which only allow high powers of Φ in the Higgs potential which effectively does the same thing [77].

Just as none of the models for inflation alone is compelling, none of the models for inflation plus strings is particularly inspiring. In fact they are quite ugly [81,80].

2.19 Cosmic String versus Inflation

Let me now turn to comparing the predictions of cosmic strings with those of 'inflation'. By inflation I mean the simplest prediction of quantum fluctuations from inflation - the scale-free Harrison- Zel'dovich spectrum[84,85,86].

The theories are really very different. In the case of inflation it makes sense to think of the density perturbation field as a linear superposition of

plane waves with random phases

$$\frac{\delta\rho(\vec{x})}{\rho} = \sum_{\vec{k}} \delta_{\vec{k}} e^{i\vec{k}.\vec{x}} \tag{75}$$

with the probability distribution for each $\delta_{\vec{k}}$ being a Gaussian : if $\delta_{\vec{k}} = r_{\vec{k}} e^{i\theta_{\vec{k}}}$ then each mode is independent (except for the reality condition which forces $\delta_{-\vec{k}} = \delta^*_{\vec{k}}$) with probability measure

$$\prod_{\vec{k}} \frac{d\theta_{\vec{k}}}{2\pi} \frac{dr_{\vec{k}} r_{\vec{k}}}{P_k} e^{-\frac{r_{\vec{k}}^2}{2P_k}} \tag{76}$$

This is a result of the inflaton field being essentially a free massless field during inflation. The power spectrum $P_k = <|\delta_{\vec{k}}|^2>$ completely defines the theory. The fact that the different $\vec{k}$ modes are decoupled means that any quantity like the excess mass in a ball of radius R

$$\delta M = \int_0^R d^3r \delta\rho(\vec{r}) \tag{77}$$

is also Gaussian distributed, by the central limit theorem. This makes large fluctuations away from the 'typical' fluctuations very rare.

With strings the situation is quite different, and

$$\delta\rho(\vec{x}) = \sum_{loops, long strings} \delta\rho(\vec{x}) \tag{78}$$

is often the simplest way to look at the density perturbations. Many different $\vec{k}$ modes are correlated in a loop for example. The probability distribution is nonGaussian as can be seen from (62) - in a volume V (ignoring correlations) the number of loops expected with mass M to $M + dM$ is

$$N(M)dM \propto M^{-\frac{5}{2}} dM \tag{79}$$

When this number is much less than one, it is just the probability of finding a loop of mass M in the volume. Thus $P(\delta M) \propto (\delta M)^{-\frac{5}{2}}$ at large δM. Any one realisation of cosmic string looks like a collection of spikes and lines in the density distribution - small loops and long strings cause very high density perturbations locally. Fluctuations from inflation however look like a superposition of plane waves that is everywhere linear. This should make it easier in principle to identify objects with initial density perturbations in the cosmic string theory - the distinction between a lump and the surrounding matter is more clear-cut. In inflation there are mass fluctuations on every scale right down to the smallest scales. This makes the identification of 'galaxies' quite hazardous. Of course merging of lumps, which is important in the cosmic string plus CDM scenario[87] tends to destroy the identification of one loop with one object.

The 'typical' mass fluctuation on a given scale δ_M is actually quite similar in both theories - $\delta M/M \propto R^{-1.35}$ approximately for scales R corresponding to the mean separation of galaxies [72] for inflation, which leads to the successful prediction of the galaxy-galaxy correlation function. For strings, the mass m of the largest 'typical' loop in a volume R^3 is given from the loop distribution (62) by $n_>(m)R^3 \approx 1$ so $\delta M \propto R^2$ and $\delta M/M \propto R^{-1}$ instead. So the gross mass fluctuation spectrum on galaxy scales is not very different.

One of the virtues of the prediction from inflation is that it is much easier to calculate with. The linear phase of the evolution is relatively trivial. With strings however it is necessary to do a computer simulation of the strings even before including the growth of perturbations. This should not of course mistakenly be seen as a reason why one theory is 'better'.

The successes of the ICDM theory are mostly on small scales where clustering has gone nonlinear. With the correct *biassing* factor (initially put in by hand but now argued by some to emerge naturally [23]) the ICDM model matches the galaxy correlation function, pair velocity distribution, and gives flat rotation curves in good agreement with observation[21]. This is impressive but it must be remembered that the nonlinear dissipative baryonic physics has not yet been put in, and this is actually in determining whether the postulated bias actually happens.

In contrast, the ICDM model has fared rather poorly in matching the large scale observations. These are, as I argued in Section 1 the most unambiguous testing ground for theories of the origin of structure. The basic problem of the ICDM model is that it has very little power on large scales - the predicted cluster cluster correlation function for example goes through zero at only 80 h_{50}^{-2} Mpc[22]. With suitable *biassing* it is claimed that it can produce 'bubbles' such as those seen in the 'slice of the universe'[21]. However with this biassing it fails by a long way to match the observed velocity field [17,79]. While the theory is not yet clearly ruled out it certainly has enough problems to make the investigation of alternatives such as cosmic strings worthwhile. In particular the SHDM model offers particular promise.

Cosmic strings are as I have said harder to calculate with. However there are already indications that cosmic strings have a lot of power on large scales. Early calculations of the correlation function for loops showed remarkable agreement with that for Abell clusters[64]. The later simulations have confirmed these correlations are present and have the correct r^{-2} form. However they have also found a much higher scaling density and this makes the long strings relatively more important. The detailed calculation of large scale structure produced by long strings is underway. This will also be important in accurately normalising the theory.

2.20 Other Observational Tests for Cosmic Strings

Cosmic strings have several fairly direct observational consequences which give us some confidence that if they exist, and were massive enough to seed the observed large scale structure, they should be observable fairly soon.

The most clear-cut effect, as with other theories, is the predicted anisotropy of the microwave background. Observations of intrinsic anisotropies would be a major advance in cosmology - the statistics of the fluctuations would tell us a lot about the kind of processes that could have caused them.

With cosmic strings the effect is quite dramatic. Moving strings lead to linear discontinuity across the sky in the temperature of the microwave background[88]. This is a result of the conical metric discussed in Section 2.14. In the string's restframe the there is a difference of $8\pi G\mu$ in the angle at which light from the same point on the last scattering surface of the microwave radiation is recieved by the observer. The observer thus appears to have a small component $8\pi G\mu v_s$ *towards* the last scattering surface behind the string, where v_s is the string's velocity perpendicular to the line of sight. This results in a Doppler shift of the radiation temperature observed of

$$\frac{\delta T}{T} = 8\pi G\mu v_s \qquad (80)$$

of the order 10^{-5} if $G\mu = 10^{-6}$. This is not far below present sensitivities. The difficulty is that strings are actually quite rare on the last scattering

surface, so one has to cover quite a large angle in order to be likely to see one at all. Recently Bouchet et al. have calculated the expected pattern from a string simulation [89] - they deduce the limit

$$G\mu < 5 \times 10^{-6} \qquad (81)$$

from the experiment of Melchiorri et. al. [90].

Another unique signature of cosmic strings is their gravitational lensing effect. The conical metric leads, in the case of a straight string, to two identical images for objects behind the string with angular separation

$$\delta\theta = 8\pi G\mu \frac{d_s}{d_s + d_o} \qquad (82)$$

where d_s and $d_o > d_s$ are the distances from us to the string and the object. This is of the order of a few arc seconds if $G\mu = 10^{-6}$. What is unique is that a string would produce a line of double images. In fact a candidate event involving four double galaxies was reported recently by Cowie and Hu [91]. Deeper observations in the same area of the sky to see whether galaxies behind the putative string connecting the objects are also lensed are currently underway.

A less direct but still powerful test of the cosmic string theory comes from the gravitational radiation produced by the strings during the radiation dominated era. The fact that the millisecond pulsar is so 'quiet' imposes strong limits on the level of gravitational background that can be around today. The expected contribution to Ω_{GW} today from cosmic strings with periods of the order of a year or so (the millisecond pulsar observations put the stringest constaint on waves with this period) is[92,93,106,95]

$$\Omega_{GW} \approx 2 \times 10^{-7} \lambda (\frac{\Gamma}{50})^{-\frac{1}{2}} (\frac{G\mu}{10^{-6}})^{\frac{1}{2}} h_{50}^{-2} \Omega^{-1} \tag{83}$$

whereas the observations already constrain $\Omega_{GW} < 4 \times 10^{-7}$ [96]. Things are already getting tight! This limit is improving, and will rapidly improve with time should another quiet millisecond pulsar be discovered, and should be able to convincingly rule out the scenario with $G\mu > 10^{-6}$ in 10 years time.

The other major effect of the gravity waves produced by strings in the radiation era is the disruption of standard nucleosynthesis. This gives the bound [97,51,95,98]

$$G\mu < 10^{-6} (\frac{\Gamma}{50}) \lambda^{-2} \tag{84}$$

The strength of these bounds can only be assessed when we have an accurate calculation of the normalisation of $G\mu$, as discussed in Section 2.16. One cautionary remark about the gravitational radiation bounds must be made. The calculations of the rate of loss of radiation were made using the Nambu action and equations. If there is more complicated dynamics, particularly in the late history of a loop, as I discussed in connection with corrections to the Nambu action, or if loops continue self-intersect late in their lifetime Γ could increase, and there could be other types of radiation emitted. Both of these effects would weaken the bounds (83) and (84). Finally there are many other large scale structure predictions waiting to be made and tested - higher correlation functions, void statistics, velocity field predictions and so on.

3 Fundamental Strings and Cosmology

In this lecture I will run the risk of confusing you by discussing both cosmic strings (i.e. topological defects in spontaneously broken gauge theories) and fundamental strings which I have not so far mentioned at all. The latter are *quite different* in many important ways - they are the elementary excitations in string theories which attempt to unify gravity with the other forces.

However fundamental strings are similar in some important respects to cosmic strings. In particular the 'Nambu-Goto' action which provides an *approximate* description of the motion of cosmic string as long as the radius of curvature is much greater than the string width is the *exact* action for free fundamental strings. Provided we keep in mind the circumstances where the 'Nambu-Goto' action breaks down for cosmic strings, we can use fundamental strings as an approximation to cosmic strings. We shall also see that our numerical results for cosmic strings in turn provide considerable insight into the behaviour of fundamental strings at high density (at least in the weak coupling limit).

The more immediate goal of this lecture will be to understand and explain some of the numerical results on string networks mentioned above. In particular we would like to understand the distribution of strings at formation (Section 2.8) where

a) most (roughly 80%) of the string is in 'infinite' string

b) the remainder is in a scale invariant distribution of loops.

c) both types of string are in the form of 'Brownian' walks.

We would also like to understand why the chopping off of loops from long string is favoured over reconnection, which was very important in leading to the scaling solution (55). Our understanding of this below will lead to a simple argument that the string network cannot come to dominate the universe and that the scaling solution is, barring very long time transients, more or less inevitable.

Of course we will only obtain precise results for strings in equilibrium in Minkowski space. These will nevertheless be relevant to strings in an expanding background in the limiting regime where ξ is much smaller than the Hubble radius.

3.1 Counting String States : the Quantised String

In studying string statistical mechanics we are simply interested in finding the configurations which maximise the number of states available to a certain length[3] of string in a box. Assuming ergodicity, the system will end up in these, the most probable configurations. This is the idea behind the *micro-canonical ensemble* and is completely independent of notions of temperature, heat baths and so on.

But how do we count states for the strings? Classically, a short string has just as many configurations as a long string. Thus one might imagine that the greatest number of states would be obtained by putting the string into the smallest possible loops. However this argument is incorrect - in counting states one has to put a *measure* on configuration space, and this brings in a length scale, roughly the size of the smallest wiggles allowed on the string. A good way to do this is actually to *quantise* the strings - then the states form a well defined discrete set and the scale of the wiggles is given in terms of the string tension. In fact the number of states for a large string is exponentially greater than that for a small string as we shall see.

[3] I shall use length L and energy E interchangeably for strings, since $E = \mu \int d\sigma$ in the gauge $\epsilon = 1$. $\int d\sigma$ can be taken as the definition of the length, related to the naive one $L_N = \int d\sigma \mid \vec{x}' \mid$ by the factor $\mid \vec{x}' \mid$. Since $<\mid \vec{x}' \mid^2>= \frac{1}{2}$ for a closed loop by an argument analogous to (69) the two definitions are quite similar.

The free quantised string is a beautiful example of a nonlinear but exactly solvable quantum theory. For a thorough review see [102]. Here I will only summarise the main formulae we need.

We only wish to count the physical degrees of freedom of the strings so we use the 'light-cone' gauge of [105]. In this gauge the transverse coordinates of the quantised string (i= 2,3..d-1 in d spacetime dimensions) obey

$$x^i(\sigma,\tau) = q^i + \frac{p^i}{2\pi\mu} + \frac{i}{2\sqrt{\pi\mu}} \sum_{n\neq 0} \frac{\alpha^i_n}{n} e^{-2in(\tau-\sigma)} + \frac{\tilde{\alpha}^i_n}{n} e^{-2in(\tau-\sigma)} \quad (85)$$

$$\left[\alpha^i_n, \alpha^j_m\right] = n\delta_{n+m,0}\delta^{ij} \quad (86)$$

and similarly for the $\tilde{\alpha}$ oscillators. The string has been parametrised so that σ runs from 0 to π. In addition we have the rest mass formula and constraints

$$\begin{aligned} m^2 &= 4\pi\mu(N + \tilde{N} - 2) \\ N &= \tilde{N} \\ N &= \sum_{n>0} \alpha^i_{-n}\alpha^i_n \end{aligned} \quad (87)$$

and similarly for $\tilde{N}$. Here N and $\tilde{N}$ are the level number operators for the left and right moving oscillators respectively.

I shall ignore the tachyon altogether in what follows. This is the state with $N = \tilde{N} = 0$ and has negative mass squared. For cosmic strings, we shall not trust the string spectrum right down to the lowest mass states since the Nambu approximation is not valid for small loops. For heterotic strings or superstrings there are no tachyons anyway. I will also for simplicity ignore fermionic or winding number modes - including these is straightforward and does not affect the main conclusions.

A complete set of states for a single string is obtained by applying the α^i_{-n} creation operators to the vacuum, which is labelled by the centre of mass momentum [102]. For example at level number 0 we have the tachyon, with $m^2 = -8\pi\mu$. At level 1 we have the tensor states $\alpha^i_{-1}\tilde{\alpha}^j_{-1} \mid 0 >$, with $m^2 = 0$ and so on. In general the level number is obtained by adding the subscripts of the α or $\tilde{\alpha}$ oscillators (which from (87) must give the same result). Thus the number of states at level n involves (apart from the left-right degeneracy and the transverse indices) just the number of different ways of writing n as a sum of smaller positive integers. In mathematics this is called the partition function, and it was first realised by the mathematicians Hardy and Ramanujan that there is a simple way to calculate it for large n[106].

The method is as follows. One considers the expression

$$\begin{aligned} P(z) &= Tr(z^N) \\ &= (1 + z + z^2 + ..)^{d-2}(1 + z^2 + z^4 + ..)^{d-2}(1 + z^3 ..)^{d-2} \quad (88) \end{aligned}$$

where the trace is over the entire Fock space of one set of oscillators (say the left moving oscillators) and factorises into the product of traces for each oscillator separately. In fact

$$\begin{aligned} P(z) &= (\prod_{n=1}^{\infty} \frac{1}{(1 - z^n)})^{d-2} \\ &= \sum_{n=0}^{\infty} d_n z^n \quad (89) \end{aligned}$$

where d_n is the degeneracy for the left moving oscillators at level n because the trace in (88) adds one term z^n for every state at level n.

Now we invert (89) for d_n

$$d_n = \oint \frac{dz}{2\pi i z} z^{-n} P(z) \quad (90)$$

where the contour runs around the origin in the complex z plane. Now the integrand is easily seen to have a saddle point for z just less than 1 on the real axis - for real z the function $P(z)$ blows up as z approaches 1 and conversely z^{-n} blows up as z decreases away from 1. We therefore distort the contour in (90) to run over the saddle point. To find the saddle point we calculate $P(z)$ near $z = 1$;

$$\begin{aligned} P(z) &= e^{-(d-2)\sum_{n=1}^{\infty} ln(1-z^n)} = e^{(d-2)\sum_{n,m=1}^{\infty} \frac{z^{mn}}{m}} \\ &= e^{(d-2)\sum_{m=1}^{\infty} \frac{z^m}{m(1-z^m)}} \sim e^{\frac{d-2}{1-z}\sum_{m=1}^{\infty} \frac{1}{m^2}} \\ &= e^{\frac{\pi^2(d-2)}{6(1-z)}} \quad (91) \end{aligned}$$

We then perform the Gaussian saddle point integral to obtain

$$d_n \sim e^{\pi\sqrt{\frac{2n(d-2)}{3}}} \quad (92)$$

as the leading behaviour. A more detailed calculation gives the prefactor [103,99]

$$d_n \approx C n^{-\frac{d+1}{4}} e^{\pi\sqrt{\frac{2n(d-2)}{3}}} \quad (93)$$

with C a constant. The total degeneracy at level n is simply the square of (93) - d_n ways for the left moving oscillators times d_n ways for the right moving oscillators.

Using the mass formula $m^2 \approx 8\pi\mu n$ we find for the number of states for a string of mass m to $m + dm$

$$\begin{aligned} \rho(m)dm &\approx cm^{-d}e^{bm}dm \\ b &= \sqrt{\frac{\pi(d-2)}{3\mu}} \end{aligned} \tag{94}$$

Thus as stated above the number of states available to a string grows exponentially with its mass. c is a constant depending on μ with dimensions of inverse volume.

3.2 The Fractal Dimension of Big Strings

As we have seen there is an exponentially large number of states available to a string of mass m. What do these states look like? This is surprisingly easy to answer[99]. Consider the position operator (85) at $\tau = 0$. Using it we can construct the operator which measures the average squared radius of the string

$$\begin{aligned} \Delta r^2 &= (d-1)\int \frac{d\sigma}{\pi}(x^i - q^i)^2 = \frac{d-1}{2\pi\mu}(R + \tilde{R} + 2\sum_{n=1}^{\infty}\frac{1}{n}) \\ R &= \sum_{n=1}^{\infty}\frac{\alpha^i_{-n}\alpha^i_n}{n^2} \\ \tilde{R} &= \sum_{n=1}^{\infty}\frac{\tilde{\alpha}^i_{-n}\tilde{\alpha}^i_n}{n^2} \end{aligned} \tag{95}$$

where i is not summed over - a single transverse index is used and the answer obtained by isotropy. The last term is a logarithmic divergence obtained by 'normal ordering' the oscillators. For cosmic string there is a cutoff corresponding to modes whose wavelength is of order the string width - for fundamental string this is a delicate issue but I will take the view that there

should also always be an ultraviolet cutoff in any 'measurement' of the size of the string (short wavelength modes diffract light more for example). With any reasonable cutoff the logarithm will turn out to be negligible compared to the first two terms.

Ignoring it, we find for example for the two extreme states at level n

$$\begin{aligned} \Delta r^2(\alpha_{-1})^n \,|\, 0> &\propto n(\alpha_{-1})^n \,|\, 0> \\ \Delta r^2\alpha_{-n} \,|\, 0> &\propto n^{-1}\alpha_{-n} \,|\, 0> \end{aligned} \tag{96}$$

The first set of states may be thought of as a long and 'straight' loops with a size Δr proportional to their length m/μ. The second set of states are increasingly tightly wound loops whose size actually decreases with increasing m.

Now we wish to calculate the average of Δr^2 for *all* states at level n - this will tell us what typical string configurations look like. It is easily seen that this is obtained by

$$\begin{aligned} \sum_{|\psi>\ \mathrm{at\ level}\ n} <\psi \,|\, \Delta r^2 \,|\, \psi> &= \frac{d-1}{\mu\pi}\frac{d}{dy}Tr(x^N y^R)\,|_{y=1}|_{\mathrm{coeff\ of}\ x^n} \\ &= \frac{d-1}{\mu\pi}\frac{1}{d-2}P(x)ln(P(x))\,|_{\mathrm{coeff\ of}\ x^n} \end{aligned} \tag{97}$$

Evaluating the right hand side by a contour integral just as we did for d_n in the previous section, and dividing the answer by d_n, the degeneracy of level n, we find

$$\Delta r^2 = \frac{d-1}{\mu\pi}\sqrt{\frac{n}{6(d-1)}} \propto m \tag{98}$$

Since the mass m is proportional to the length, this means that the typical string trajectory has a fractal dimension of 2, the same as a Brownian random walk. This also applies to the point-to-point separation on a string [99].

3.3 Failure of Canonical Ensemble

We are now ready to discuss the statistical mechanics of strings in general. Since string interactions involve the splitting and joining of strings, the number of strings is not conserved. The problem we wish to address is simple - given a box of volume V containing string with energy E to $E + dE$, which configurations are most probable? We shall make the usual assumption that all states are *a priori* equally likely so the probability distribution is just given by counting states.

Historically this problem was first addressed by Hagedorn, Frautschi and Carlitz [107,108,109] who were motivated by the desire to understand hadronic matter at high densities. Nowadays we believe that the string model of hadrons is only good at low energies, and at high densities hadronic matter behaves more like free quarks and gluons. Nevertheless many of the important conclusions were reached by these early works. For some recent papers see [104]. Here I will present a simplified account of the work of D. Mitchell and myself on the subject [99](see also [100]).

Let me begin by reviewing the basics of canonical statistical mechanics. One starts from the partition function

$$Z(\beta) = \int_0^\infty dE\, \Omega(E) e^{-\beta E} \qquad (99)$$

where $\Omega(E)$ is the number of states for the system to have energy in the range E to $E + dE$. β is $1/T$ as usual. The Boltzmann factor $e^{-\beta E}$ may be thought of as the decrease in the number of states for a heat bath if we take energy E from it into the system: the integrand of $Z(\beta)$ is the E dependent part of the total number of states for system plus heat bath. For 'normal'

macroscopic systems, $\Omega(E)$ grows rapidly with E, typically as E^n where n is the number of particles in the system. In fact this is where the $e^{-\beta E}$ for the heat bath comes from. Assume that for the heat bath the number of states $\Omega_{hb}(E) \propto E^n$. From the entropy, defined as $S = ln(\Omega)$ we calculate the temperature via $\frac{1}{T} \equiv \frac{dS}{dE} = \frac{n}{E}$. The definition of a heat bath is that we hold T fixed and let n and E go to infinity.

Now if our system is in contact with the heat bath but the total energy E_{tot} conserved, the energy of the heat bath is given by $E_{tot} - E$ where E is the energy in the system. Thus the number of states for the heat bath is given up to constants by $ln(\Omega_{hb}(E_{tot} - E)) = n ln(E_{tot} - E) \propto -E/T$ expanding to first order in E/E_{tot} which becomes small as n goes to infinity. Thus the dependence of Ω_{hb} on E is given by the Boltzmann factor, as stated.

For 'normal' systems this dependence on E is strong enough to kill the rapid growth of $\Omega(E)$ as long as the heat bath is much larger than the system. The result is that the integrand in (99) is very sharply peaked. Its maximum is obtained by differentiating the exponent i.e. solving $\frac{dS}{dE}(E) = \beta$. Physically this corresponds to the value of E for which there is an equilibrium between the system and the heat bath - the total number of states does not change to first order if one moves energy from the system to the heat bath or vica versa. The sharpness of the peak in the integrand is essential if we are to calculate the expected value of properties of the system at fixed E by simply inserting those quantities into the integrand - in other words there must be a very tight correspondence between E and T if we are to use the canonical ensemble to answer questions about the isolated system. To see how sharply peaked the integrand actually is we expand $S(E)$ to second order about the equilibrium value $\bar{E}$

$$\Omega(E) \propto e^{\frac{1}{2}\frac{d^2 S}{dE^2}(\bar{E})(E-\bar{E})^2+\ldots} \tag{100}$$

The integrand has in this approximation a Gaussian peak with width $\delta E^2 = -\frac{d^2 S}{dE^2}^{-1} = -\frac{dE}{d\beta} = T^2\frac{dE}{dT} \equiv T^2 c_V$ where c_V is the specific heat. Thus the fractional energy fluctuation is $\delta E^2/E^2 = c_V T^2/\bar{E}^2 \propto V^{-1}$ for normal systems, where the specific heat is positive and proportional to the volume. Thus as we take the volume to infinity the energy fluctuations go to zero.

I have reviewed these basic assumptions because strings will turn out to violate them. The cause of this violation is very simple. We saw in Section 3.1 that the number of states available to a string grows *exponentially* with energy (we can put all the energy into the rest mass). Thus the integrand in (99) contains the term $e^{bE-\beta E}$. For $\beta < b$ the integral diverges: this happens for

$$T > T_H \equiv b^{-1} = \sqrt{\frac{3\mu}{\pi(d-2)}} \tag{101}$$

Thus canonical thermodynamics simply does not make sense at temperatures above the Hagedorn temperature T_H. As I will emphasise this is not a fundamental inconsistency in any sense but simply means that the usual assumptions involved in the canonical approach break down.

3.4 Superstrings: Z=∞ for $T > T_H$

Surprisingly this simple fact has been ignored in some of the recent superstring literature on strings at high temperature. For example it has been claimed[110] that for heterotic strings there is a 'duality' relation $Z(\beta) =$

$Z(\pi/\beta\mu)$ which is clearly inconsistent with Z being finite below T_H and infinite above it. What has gone wrong is that some of the resummations involved in the 'proofs' of this formula are illegitimate. The far-reaching conclusions drawn from this work by Atick and Witten[111] for example seem to me to be on very shaky ground. Let me give a brief and rigorous proof that Z is indeed infinite for superstrings above T_H.

Just as in field theory we can calculate the partition function for a gas of free strings, as in Section 2.2. The only difference is that there is an extra sum over the internal mass levels for the strings;

$$\begin{aligned} Z &= e^{-\beta F} \\ -\beta F &= \sum_{m=0}^{\infty} d_m V \int d^{d-1}k(-ln(1-e^{-\beta\omega_k}) + ln(1+e^{-\beta\omega_k})) \end{aligned} \tag{102}$$

where d_m is the degeneracy of the mth mass level and the two terms come from bosons and fermions. Supersymmetry of course gives them identical degeneracies. Expanding the logarithms, we obtain

$$\begin{aligned} -\beta F &= \sum_{m=0}^{\infty} d_m \sum_{n=1}^{\infty}(1+(-1)^{n+1})V \int d^{d-1}k \frac{e^{-n\beta\omega_k}}{n} \\ &> 2\sum_{m=0}^{\infty} d_m V \int d^{d-1}k e^{-\beta\omega_k} \end{aligned} \tag{103}$$

But for the massive modes $\omega_k = \sqrt{\vec{k}^2 + m^2} < m + \vec{k}^2/(2m)$ so we obtain the further inequality

$$\begin{aligned} -\beta F &> 2\sum_{m=1}^{\infty} d_m e^{-\beta m} V \int d^{d-1}k e^{-\frac{\beta\vec{k}^2}{2m}} \\ &= 2V\sum_{m=1}^{\infty} d_m e^{-\beta m}(\frac{m}{2\pi\beta})^{\frac{d-1}{2}} \end{aligned} \tag{104}$$

Since $d_m \approx e^{bm}$ the sum is clearly divergent above T_H. Supersymmetry therefore makes little difference - the canonical partition function does not exist above T_H.

Now it might reasonably be argued that string interactions have so far been ignored and that these might well ameliorate the divergence in Z above T_H. Indeed this is precisely what happens with cosmic strings. In that case, as T approaches T_c, the critical temperature of the string - forming phase transition, the mass of the Higgs field goes to zero and thus the strings become very fat. Interactions between strings become important and the Nambu description breaks down. Both effects act to reduce the number of states available below that of free Nambu strings. The same is true of hadronic strings as one approaches the QCD phase transition.

The same phenomena could also occur in string theory but then string theory would probably not be a good name to call it! In any case we would obviously need more than free string calculations to understand it and this is really all that has been done so far. For small string coupling constant (the value of the string coupling constant depends upon the dilaton expectation value and cannot be fixed at present from low energy physics) the free string picture I will present below should indeed be reasonable. As we shall see it is not the strings but the canonical approach to thermodynamics that breaks down.

3.5 Low Temperatures and Densities: The canonical approach

At low temperatures and densities canonical thermodynamics provides a good description of a gas of strings. It is also easier than the microcanonical approach. As in the previous section we can write down the partition function for a gas of closed oriented bosonic strings

$$
\begin{aligned}
Z &= e^{-\beta F} \\
-\beta F &\approx -\int_{m_0}^{\infty} dm\rho(m) V \int d^{d-1}k \ln(1 - e^{-\beta\omega_k}) \qquad (105)
\end{aligned}
$$

where we use $\rho(m) = ce^{bm}m^{-d}$ from (94) and we include a cutoff m_0 which is a characteristic low mass in the spectrum - we shall not believe any answers precisely when they involve m_0. For fundamental closed strings the mass m of the first massive level is $\sqrt{8\pi\mu}$ and so $bm = \sqrt{\frac{8\pi^2(d-2)}{3}} >> 1$ and we should take m_0 to be at least a few times this in order for the form of $\rho(m)$ to be valid. Thus we have $bm_0 >> 1$. This will be useful.

For $\beta m_0 > bm_0 >> 1$ we can expand the logarithm and keep only the first term, and approximate w_k by its nonrelativistic value - $w_k \approx m + \vec{k}^2/(2m)$. The fact that this is a good approximation (which is easily checked) means that the configurations that dominate are those where the loops are nonrelativistic - roughly speaking the internal degrees of freedom are more important than the centre of mass degrees of freedom. This also means that we can use energy and mass more or less interchangeably. We now perform the Gaussian $\vec{k}$ integral to obtain the thermodynamic quantities

$$
-\beta F \approx \frac{cV}{(2\pi\beta)^{\frac{d-1}{2}}} \int_{m_0}^{\infty} dm \frac{e^{(b-\beta)m}}{m^{\frac{d+1}{2}}} \qquad (106)
$$

$$E = \frac{\partial}{\partial\beta}(\beta F) \approx \frac{cV}{(2\pi\beta)^{\frac{d-1}{2}}}\int_{m_0}^{\infty} dm\, \frac{e^{(b-\beta)m}}{m^{\frac{d-1}{2}}} \quad (107)$$

$$c_V \propto \overline{(E-\overline{E})^2} = \frac{\partial^2}{\partial\beta^2}(-\beta F) \approx \frac{cV}{(2\pi\beta)^{\frac{d-1}{2}}}\int_{m_0}^{\infty} dm\, \frac{e^{(b-\beta)m}}{m^{\frac{d-3}{2}}} \quad (108)$$

$$(109)$$

where I have dropped terms down by βm_0. The density (108) corresponds to a distribution of string loops with the number of loops of mass m to $m + dm$ per unit volume

$$n(m) = \frac{c}{(2\pi\beta)^{\frac{d-1}{2}}}\frac{dm}{m^{\frac{d+1}{2}}}e^{(b-\beta)m} \quad (110)$$

which is just a Boltzmann distribution at temperature β^{-1}.

The first consequence is therefore that large loops are suppressed at low temperatures. The corresponding density is approximately given from (107) as

$$\rho \approx \frac{2\pi c}{(2\pi\beta)^{\frac{d-1}{2}}}\frac{e^{(b-\beta)m_0}}{m_0^{\frac{d-1}{2}}} \quad (111)$$

we can solve this to find the temperature corresponding to a given density;

$$T \approx \frac{T_H}{1+\frac{1}{bm_0}ln(\frac{\rho_0}{\rho})}$$

$$\rho_0 \equiv m_0 c \quad (112)$$

recalling that $b^{-1} = T_H$. This shows that the thermodynamic temperature of a system of strings depends only very weakly on its density. For example for a network of cosmic strings in todays universe, $T_H \approx \mu^{\frac{1}{2}} \approx 10^{16} GeV$. Taking $bm_0 \approx 10$ and $ln(\rho_0/\rho) \approx 10^3$ we find $T \approx 10^{15}$ GeV! Of course this does not mean that the strings are literally hot but shows they are very far

from thermal equilibrium with the surrounding radiation today, and in fact become so quite soon after they form. Thus, unfortunately, we can learn very little about the distribution of cosmic string in the late universe from equilibrium calculations.

What we do learn however is that a system of strings allowed to equilibrate at low density is very different from the system of strings at formation I discussed in section 2.8. Large loops are exponentially suppressed. The phase space is dominated by configurations where essentially all the string is in the smallest possible loops.

This is consistent with what we have learned from the numerical simulations where we saw that chopping off of loops from long strings is favoured over reconnection. It also indicates that it would be inconsistent to have a string dominated universe in the following sense [99]. Imagine tha $G\mu$ was very much less than unity. If the density of long strings grew relative to the background then the scale ξ on the network would become much smaller than the Hubble radius R_H (see Section 2). If ξ is much less than R_H however the strings evolve as in flat space. The timescale for the string to approach thermal equilibrium is however given by ξ, as I discussed. In a time of order $\xi << R_H$, the expansion time, therefore, the long strings would lose a fraction of their length into loops (because this is favoured by phase space) and the density of the long strings would fall relative to the background density. Thus the scaling solution is really inevitable, barring really long time transients which would make the string system quite unique among statistical systems (a gas in a box for example equilibrates in a few mean collision times).

The result for the loop distribution in flat spacetime was calculated in [99] and shown to agree well with the flat spacetime simulations of Smith and Vilenkin [112]. In the simulations the strings are discretised and the scale b enters as the spacing between points on the string. m_0 enters as the cutoff in the smallest loops allowed. The predictions from [99] and the simulation results [112] are shown in Fig 3.1.

3.6 The Microcanonical Approach

It is clear from (108) that the canonical approach will fail at high densities for $d \geq 4$ because the canonical density tends to a finite limit

$$\rho_H \approx \frac{2}{d-3} \frac{c m_0}{(2\pi b m_0)^{\frac{d-1}{2}}} \qquad (113)$$

as the temperature approaches T_H. It is also clear that it will fail for $d = 4$ because the fluctuations are large at T_H. Thus for every d of interest it is essential to use the more fundamental microcanonical ensemble.

In the microcanonical ensemble one uses the density of states $\Omega(E)$ directly. One simply finds the configurations that dominate the total number of states, for energy fixed in the range E to $E + dE$ and for fixed volume V.

Now from (99) the partition function $Z(\beta)$ is simply the Laplace transform of $\Omega(E)$. Therefore one way to obtain $\Omega(E)$ is to invert the Laplace transform. Starting from (105), then as long as βm_0 is small we can expand the logarithm and keep only the first term. This yields

$$Z(\beta) \approx \sum_{n=1}^{\infty} \frac{1}{n!} (V \int d^{d-1}k \int_{m_0}^{\infty} dm\, \rho(m) e^{-\beta\omega_k})^n \qquad (114)$$

Now since the inverse Laplace transform of an exponential is just a delta

function we find

$$\Omega(E) = \sum_{n=1}^{\infty} \frac{1}{n!} \prod_{i=1}^{n} (V \int d^{d-1}k_i \int_{m_0}^{\infty} dm_i\, \rho(m_i)) \delta(E - \sum_{i=1}^{n} E_i) \qquad (115)$$

with $E_i = \sqrt{m_i^2 + \vec{k}_i^2}$. This is just the *classical* formula for the total number of states for an arbitrary number n strings of total energy E in a box. The $1/n!$ is just the usual 'classical statistics' overcounting factor for identical particles. It is only strictly valid when the average occupancy of each state is much less than unity, which is true here since we always have $e^{-\beta E_i} < e^{-\beta m_0} < e^{-bm_0} << 1$.

Equation (115) is the starting point for the microcanonical approach first followed by Frautschi[108] and Carlitz[109].

3.7 High Density

In reference [99] Mitchell and myself started from (115) and found the configurations that dominate the integrals directly. This is a hard calculation but there is a simpler approach that leads to the same result, based on examining the nature of the divergences in canonical expectation values at T_H which I shall explain here.

We can actually determine a lot about the function $\Omega(E)e^{-bE}$ (the integrand of the partition function at T_H) from its moments (107 - 109) and higher moments considered as T approaches T_H. This then tells us about the integrand $\Omega(E)e^{-\beta E}$ at arbitrary β. Consider the probability distribution

$$p(E)dE = \frac{\Omega(E)e^{-\beta E}dE}{Z(T_H)} \qquad (116)$$

We shall be interested in its limit as $\beta \rightarrow b$. Physically this is the probability for finding the system with energy E when it is in contact with a heat bath at T_H. The expectation value of the energy in this distribution is given by (108) at T_H;

$$\overline{E} = \frac{cV}{(2\pi b)^{\frac{d-1}{2}}} \int_{m_0}^{\infty} \frac{dm}{m^{\frac{d-1}{2}}} = \frac{2}{d-3} \frac{cVm_0}{(2\pi b m_0)^{\frac{d-1}{2}}} \tag{117}$$

which is finite in $d \geq 4$ and corresponds to a loop distribution

$$n(m) = \frac{c}{(2\pi b)^{\frac{d-1}{2}}} \frac{dm}{m^{\frac{d+1}{2}}} \tag{118}$$

For $d \geq 6$ the fluctuations around this value are small;

$$\overline{\delta E^2} = \frac{2}{d-5} \frac{cVm_0^2}{(2\pi b m_0)^{\frac{d-1}{2}}} \tag{119}$$

so the fractional fluctuation $\delta E^2/E^2 \propto V^{-1}$ which tends to zero in the infinite volume limit. However for $d < 6$ the fluctuations are large.

In $d \geq 6$ near $\overline{E}$ the probability distribution is a sharply peaked Gaussian. However (and this is how strings differ from 'normal' systems) at large E $p(E)$ is not Gaussian at all.

We can see this by calculating higher moments of E just as in (107 - 109) for β close to b. The first divergent moment is

$$\begin{aligned} \overline{\delta E^{\frac{d}{2}}} &= \frac{cV}{(2\pi\beta)^{\frac{d-1}{2}}} \int_{m_0}^{\infty} \frac{dm}{m^{\frac{1}{2}}} e^{(b-\beta)m} \\ &\approx \frac{cV}{(2\pi\beta)^{\frac{d-1}{2}}} \frac{\sqrt{\pi}}{(\beta-b)^{\frac{1}{2}}} \end{aligned} \tag{120}$$

as $\beta \to b$ which tells us that for large E and at $\beta = b$

$$p(E) \approx \frac{cV}{(2\pi b)^{\frac{d-1}{2}}} \frac{1}{E^{\frac{d+1}{2}}} \tag{121}$$

so

$$\begin{aligned} \Omega(E) &\approx Z(T_H) \frac{cV}{(2\pi b)^{\frac{d-1}{2}}} \frac{e^{bE}}{E^{\frac{d+1}{2}}} \\ Z(T_H) &= e^{-bF_H} \\ -bF_H &= \frac{cV}{(2\pi b)^{\frac{d-1}{2}}} \int_{m_0}^{\infty} \frac{dm}{m^{\frac{d+1}{2}}} \end{aligned} \tag{122}$$

The second term in Ω is simply the number of states available to a single string of energy E:

$$\Omega_1(E) = cV \int d^{d-1}p \int_{m_0}^{\infty} dm\, \frac{e^{bm}}{m^d} \delta(E - \sqrt{\vec{p}^2 + m^2}) \tag{123}$$

where one uses the nonrelativistic solution to perform the m integral: $m \approx E - \vec{p}^2/(2E)$ and thus the momentum integral to find

$$\Omega_1(E) \approx \frac{cV}{(2\pi b)^{\frac{d-1}{2}}} \frac{e^{bE}}{E^{\frac{d+1}{2}}} \tag{124}$$

Thus as one increases the density beyond ρ_H the energy just goes into a single long string. Now the first term in (122) is just the partition function at T_H

$$Z(T_H) \approx e^{-b\overline{E}} \Omega_{loops}(\overline{E}) \Delta E \tag{125}$$

in $d \geq 6$ because the integral is strongly peaked around $\overline{E}$, and we take it's width to be ΔE. Ω_{loops} is the number of states corresponding to the loop distribution (118) at T_H. Thus assuming $E >> \overline{E}$ (123) is just

$$\Omega(E) \approx \Omega_{loops}(\overline{E}) \Omega_1(E - \overline{E}) \Delta E \tag{126}$$

which is the number of states where $\overline{E}$ of the string is in loops and the remainder is in one long string. This is therefore what happens at very high density, $\rho > \rho_H$. The configurations that dominate consist of a density - independent loop distribution and a very large string which takes up a greater and greater fraction of the total energy as the density is increased.

In fact using our previous result (98) that $m \propto r^2$ we find the size distribution of loops

$$n(r)dr \propto \frac{dr}{r^d} \tag{127}$$

precisely the scale invariant distribution of loops found in simulations of cosmic string formation. The remainder of the energy is in one long string, whose energy equals

$$E_{long\ string} = E - \rho_{loops}V \tag{128}$$

where ρ_{loops} is just the energy density corresponding to the loop distribution (118).

All these results (random walks, loop distribution and predominance of long string) agree with what is found in the simulations of cosmic string formation discussed above. This is reassuring, but it must be remembered that all these results were derived for *free* strings. The cosmic string formation simulations used instead the assumption of *free Higgs field phases* and found the typical configurations. Both these calculations neglect interactions - as I have emphasised in Section 2 it is very hard to do better analytically. Happily the final scaling solution for the string network is rather insensitive to the details of the initial conditions (though it does require long strings to be present).

The predictions derived above and first made in [99] were checked in flat spacetime simulations by Sakellariadou and Vilenkin [113], extending the

work of Smith and Vilenkin cited earlier. Their results are shown in Figure 3.1. As can be seen from the Figure, the transition to long string domination is very sharp.

3.8 Microcanonical Temperatures above T_H

It is interesting to examine the transition to long string domination a little more closely. In particular let us imagine that our system is in thermal contact with a heat bath. For superstrings this would simply be the massless modes in the theory - radiation, gravity waves etc. Now the microcanonical temperature

$$\frac{1}{T_\mu} = \frac{dS}{dE} = \frac{d}{dE} ln\Omega(E) \tag{129}$$

is well defined at arbitrarily high density. It also has a simple physical meaning. It is the temperature of a heat bath which would be in equilibrium (to first order) with the system of strings (see Section 3.3). For $d \geq 6$ as we have seen the canonical approach is valid right up to T_H in the large volume limit. Thus to high accuracy we have $T_\mu(E) = T_{can}(E)$. Let me just emphasise that these two quantities are *a priori* different: to calculate T_μ one fixes E and calculates T from (129). In the canonical approach one does just the reverse, fixing T and calculating the expected value of $E(T)$. $T_{can}(E)$ is the canonical temperature required to obtain energy E.

$T_{can}(E)$ and $T_\mu(E)$ are compared in Figure 3.2. Of course T_{can} does not make sense above T_H. However at high E we can calculate T_μ from (122)

$$\frac{1}{T_\mu} = \frac{d}{dE} ln\Omega(E) = \frac{1}{T_H} - \frac{d+2}{2E} \tag{130}$$

which yields $T_\mu > T_H$! However the difference in temperature is tiny in the large volume limit - the $E-T$ curve turns vertical at T_H in the large volume limit. Differentiating (130) we would find a negative specific heat, but $c_V \propto V^2$, an infinite specific heat per unit volume.

We have come a long way in the analysis of the Hagedorn transition to long string domination. The main conclusion is simple. In the large volume limit there are more states available to a system of strings than there are to a heat bath at any temperature above T_H. Thus as one squeezes a box full of string plus radiation there comes a density above which all the energy will flood into the long string.

This may have important consequences for fundamental strings in the early universe, or in black holes, where very high densities are reached. As long as equilibrium is attained and the effects of gravity sufficiently weak to validate the above analysis, above a certain density all the energy in the radiation will flood into long strings which would fill the universe. It is important to note that the *radiation* density is always bounded above, basically to T_H^d. Thus the problem of understanding singularities in string theory involves the massive string modes crucially. This makes it so different from the usual problems with ordinary matter or radiation that it may even have a solution!. Some discussion of the quantum mechanics of the massive modes in the very early universe was given in [101] where it was argued that they may actually lead to an inflationary phase (see also [100]). Some recent work on calculating the interaction rates for massive string modes may be found in [114,115] - this is important in eventually deciding how reasonable the assumption of thermal equilibrium is in the case of fundamental strings.

Acknowledgements

I would like to thank the organisers and participants of Schladming, TASI and CCAST for making the schools so enjoyable. I thank A. Albrecht and D. Mitchell for very enjoyable continuing collaboration. This work was supported in part by the DOE and by the NASA at Fermilab.

References

[1] E. Kolb and M.S. Turner, *The Early Universe*, Addison Wesley, 1988.

[2] T.W.B. Kibble, J. Phys. A9 (1976) 1387; Physics Reports 67 (1980) 183.

[3] R. Giovanelli and M. P. Haynes, Astrophysical Journal, 87 (1982) 1355.

[4] R.F. Kirschner, A. Oemler, P.L. Shecter and S.A. Schectman, in 'Early Evolution of the Universe and its Present Structure', I.A.U. Symposium 104 (1983) 197.

[5] V. De Lapparent, M. Geller and J. Huchra, Astrophysical Journal 302 (1986) L11.

[6] S. Weinberg, *Gravitation and cosmology: Principles and Applications of the General Theory of Relativity*, John Wiley, New York, 1972.

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

[8] J. Yang, M.S. Turner, G. Steigman, D.N. Schramm and K. Olive, Ap. J. 281 (1984) 493.

[9] J. Primack, *Dark Matter in The Universe*, lectures presented at the International School of Physics, Varenna 1984, SLAC-PUB-3387.

[10] B. Bingelli, Astron. Astrophys. 107 (1982) 338.

[11] D. Koo, R. Kron, and A. Szalay, 13th Texas Symposium on Relativistic Astrophysics, Ed. P. Ulmer, World Scientific, 1986.

[12] G.O. Abell, Astrophysical Journal Supp. 3 (1958) 211; for a nice review see J. Oort, Ann. Rev. Astron. Astrophys. 21 (1983) 373.

[13] N.A. Bahcall and R.M. Soneira, Astrophysical Journal 270 (1983) 20.

[14] M. Davis and J. Huchra, Astrophysical Journal 254 (1982) 437.

[15] N.A. Bahcall and W.S. Burgett, Astrophysical Journal 300 (1986) L35.

[16] S. Schectman, Astrophysical Journal Supp. 57 (1985) 77.

[17] D. Lynden-Bell, S.M. Faber, D. Burstein, R.L. Davies, A. Dressler, R.J. Turlevich and G. Wegner, Astrophysical Journal, March 1988 in press.

[18] A. Szalay and D.N. Schramm, Nature 314 (1985) 718.

[19] J.R. Bond and G. Efstathiou, Astrophysical Journal 285 (1984) L45.

[20] D. Burstein and V.C. Rubin, Astyrophysical Journal 297 (1985) 423.

[21] M. Davis, G. Efstathiou, C. Frenk and S. White, Ap. J. 292 (1984) 371; S.D.M. White, in *Nearly Normal Galaxies, From the Planck Time to the Present*, Ed. S.M. Faber, Springer-Verlag, 1987.

[22] J.M. Bardeen, J.R. Bond, N. Kaiser, A. Szalay, Ap. J. 304 (1986) 15.

[23] C. Frenk, S. White, M. Davis and G. Efstathiou, Ap. J. 327 (1988) 507.

[24] J.R. Bond and A. Szalay, Ap. J. 274 (1984) 443.

[25] R. Brandenberger, N. Kaiser, D. Schramm and N. Turok, Phys. Rev. Lett. 59 (1987) 2371

[26] R. Brandenberger, N. Kaiser and N. Turok, Phys. Rev. D36 (1987) 335.

[27] S. Tremaine and J. Gunn, Phys. Rev. Lett. 42 (1979) 407.

[28] D.A. Kirzhnits and A.D. Linde, Phys. Lett. 42B (1972) 745.

[29] L. Dolan and R. Jackiw, Phys. Rev. D9 (1974) 3320; S. Weinberg, Phys. Rev. D((1974) 3357.

[30] A more detailed discussion is contained in K. Wilson and J. Kogut, Phys. Rep. 12 (1974) 75.

[31] E. Copeland, D. Haws and R. Rivers, Fermilab preprint 1988.

[32] T.W.B. Kibble, G. Lazarides and Q. Shafi, Phys. Rev. D26 (1982) 435; D. Olive and N. Turok, Phys. Lett. 117B (1982) 193.

[33] E. Witten, Phy. Lett. 149B (1985) 351.

[34] J. Preskill, Phy. Rev. Lett. 43 (1979) 1365.

[35] A. Vilenkin, Phys. Rep. 121 (1985) 263.

[36] E.P.S. Shellard Nuc. Phys. B283 (1987) 624.

[37] R. Matzner, University of Texas report, 1987.

[38] E.B. Bogomolny, Sov. J. Nucl. Phys. 24 (1976) 449; E.B. Bogomolny and M.S. Marinov, Sov. J. Nucl. Phys. 23 (1976) 355.

[39] L. Jacobs and C. Rebbi, Phys. Rev. B19 (1979) 4486.

[40] R. Davis, Phys. Rev. D32 (1985) 3172, A. Vilenkin and T. Vachaspati,Phys. Rev. D35 (1987) 1138.

[41] E. Witten, Nuc. Phys. B249 (1985) 557.

[42] J. Ostriker, C. Thompson and E. Witten, Phys. Lett. 180B (1986) 231.

[43] Y. Nambu, Proceedings of International Conference on Symmetries and Quark Models (Wayne State University, 1969); Lectures at the Copenhagen Summer Symposium, 1970

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

[45] K.B. Nielsen and P. Olesen, Nuclear Physics B61 (1973) 45.

[46] J. Frieman, R. Scherrer, Phys. Rev. D33 (1986) 3556.

[47] N. Turok and P. Bhattacharjee, Phys. Rev. D33 (1984) 1557.

[48] N. Turok and T.W.B. Kibble, Phys. Lett. 116B (1982) 141.

[49] A. Albrecht and N. Turok, Phy. Rev. Lett. 54 (1985) 1868; in preparation, 1988.

[50] T.W.B. Kibble, Nuc. Phys. B252 (1985) 227.

[51] D. Bennett, Phys. Rev. D33 (1986) 872.

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

[53] A.Vilenkin, Phys. Rev. Lett. 46 (1981) 1169, 1496(E).

[54] N. Turok, Phys. Lett. 123B (1983) 387.

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

[56] T. Vachaspatiand A. Vilenkin, Phys. Rev. D30 (1984) 2036; D. Garfinkle and T. Vachaspati, Phys. Rev. D36 (1987) 2229.

[57] T. Vachaspati and A. Vilenkin, Phys. Rev. D31 (1985) 3052.

[58] T. Vachaspati, Phys. Rev. D35 (1987) 1767; Gravity Research Foundation Prize Essay, 1987.

[59] C. Burden Phy. Lett. 164B (1985) 277.

[60] A. Vilenkin, Phys. Rev. D23 (1981) 852.

[61] An older reference is L. Marder, Proc. Roy. Soc. Lond. A252 (1959) 45.

[62] J.R. Gott, Ap. J. 288 (1985) 422.

[63] D. Garfinkle, Phys. Rev. D32 (1985) 1323.

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

[65] E. Bertschinger Ap. J. 316 (1987) 489.

[66] N. Turok and R. Brandenberger, Phys. Rev. D33 (1986) 2175.

[67] N. Turok, in *A Unified View of the Macro- and Micro-Cosmos*, Ed. A De Rujula, D.V. Nanopoulos and P.A. Shaver, World Scientific 1987.

[68] F. Bouchet, LBL preprint, 1988.

[69] R. Scherrer and W. Press, Smithsonian preprint, 1988.

[70] N. Turok, in *Nearly Normal Galaxies, From the Planck Time to the Present*, Ed. S.M. Faber, Springer-Verlag, 1987.

[71] A. Albrecht, A. Stebbins and N. Turok, in preparation, 1988.

[72] P.J.E.Peebles, Ap. J. 263 (1982) L1.

[73] L. Kofman and A.D. Linde, Nuc. Phys. B282 (1987) 555.

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

[75] A. Everett, T. Vachaspati and A. Vilenkin, Phys. Rev. D31 (1985) 1925.

[76] E. Copeland, D. Haws, T. Kibble, D. Mitchell and N. Turok, Nuc. Phys. B298 (1988) 445.

[77] G. Lazarides C. Panagiotakopoulos and Q. Shafi, Phys. Lett. 183 (1987) 289.

[78] R. Brandenberger, N. Kaiser, E.P.S. Shellard and N. Turok, Phys. Rev. D36 (1987) 335.

[79] J.R. Bond, in *Nearly Normal Galaxies, From the Planck Time to the Present*, Ed. S.M. Faber, Springer-Verlag, 1987.

[80] Q. Shafi and A. Vilenkin, Phys. Rev. D29 (1984) 1870.

[81] E. Vishniac, K. Olive and D. Seckel, Nuc. Phys. B289 (1987) 717.

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

[83] E. Bertschinger and P. Watts Ap. J. 328 (1988) 23.

[84] E. Harrison, Phys. Rev. D1 (1980) 2726.

[85] Ya. B. Zel'dovich, M. N. R. A. S.,160 (1982) 1P.

[86] S. Hawking, Phys. Lett. 115B (1982) 295; A. Guth and S.-Y. Pi, Phys. Rev. Lett. 49 (1982) 1110; J. Bardeen, P. Steinhardt and M. S. Turner, Phys. Rev. D28 (1983) 679.

[87] R. Brandenberger and E.P.S. Shellard, Brown preprint, 1988.

[88] N. Kaiser and A. Stebbins, Nature 310 (1984) 391; A. Stebbins, Ap. J. 327 (1988) 584.

[89] F. Bouchet, D. Bennett and A. Stebbins, Fermilab preprint, 1988.

[90] F Melchiorri, B. Melchiorri, C. Ceccarelli and L. Pietranera, Ap. J. Lett. 250 (1981) L1.

[91] L.L. Cowie and E.M. Hu, Ap. J. 318 (1987) L33.

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

[93] E. Witten, Phys. Rev. D30 (1984) 272.

[94] C. Hogan and M. Rees, Nature 311 (1984) 109.

[95] R. Brandenberger, A. Albrecht and N. Turok, Nuc. Phys. B277 (1986) 605.

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

[97] R. Davis, Phys. Lett. 161B (1285) 285; D. Bennett, Phys. Rev. D33 (1986) 872.

[98] H. Hodges and M. Turner, Fermilab preprint, 1988.

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

[100] Y.Aharonov, F. Englert and J. Orloff, Phys. Lett. B 199 (1987) 366.

[101] N. Turok, Phys. Rev. Lett. 60 (1988) 552.

[102] M. Green, J. Schwarz and E. Witten, *Superstring Theory*, Cambridge University Press, 1987.

[103] K. Huang and S. Weinberg, Phys. Rev. Lett. 25 (1970) 895.

[104] M. Bowick and L.C.R. Wijewardhana, Phys. Rev. Lett. 54 (1985) 2485; B. Sundborg, Nucl. Phys. B254 (1985) 583; E. Alvarez, Phys. Rev. D31 (1985) 418, Nucl. Phys. B269 (1986) 596; P. Salomonson and B. S. Skagerstam, Nucl. Phys. B268 (1986) 349.

[105] P. Goddard, G. Goldstone, C. Rebbi and C. Thorn, Nucl. Phys. B 56 (1973) 109.

[106] G. H. Hardy and S. Ramanujan, Proc. Lond. Math. Soc. 17 (1918) 75.

[107] R. Hagedorn, Nuovo Cimento Suppl. 3 (1965) 147.

[108] S. Frautschi, Phys. Rev. D3 (1971) 2821.

[109] S. Carlitz, Phys. Rev. D5 (1972) 3231.

[110] K. H. O'Brien and C.-I.Tan, Phys. Rev. D36 (1987) 1184.

[111] J. Atick and E. Witten, IAS preprint IAS-HEP-88/14.

[112] G. Smith and A. Vilenkin, Phys. Rev. D36 (1987) 987.

[113] M. Sakellariadou and A. Vilenkin, Phys. Rev. D37 (1988) 885.

[114] J. Polchinski, Texas preprint UTTG-07-8.

[115] D. Mitchell, N. Turok, R. Wilkinson and P. Jetzer, Fermilab preprint, 1988.

Figure Captions

Figure 2.1

The potential $V(\Phi) = \frac{\lambda}{8}(\Phi^2 - \eta^2)^2$ used in Section 2.

Figure 2.2

The effective potential for the long wavelength modes $V_{eff}(\Phi)$ calculated in Section 2.2, at $T = 0$, $T = T_c$, and $T = \sqrt{2}T_c$.

Figure 2.3

The three simplest types of defects. On the left is the spatial distribution of the fields. Arrows show the location of the fields on the minimum of the potential V_{min}, which is shown on the right.
a) a domain wall
b) a string
c) a magnetic monopole.

Figure 2.4

A pictorial demonstration that $\pi_1(G/H) = \pi_0(H)$ when G is simply connected. On the left, H has two components, the component connected to the identity, H_e, and the disconnected component H'. When G/H is constructed from G one identifies all the elements of H with the identity. This can be pictured as squeezing all the elements of H together, forming a 'tunnel' as shown on the right. This clearly has a noncontractible loop.

Figure 2.5

The profile of a cosmic string. The solid line is $| \Phi | /\eta$ where the angular dependence (in cylindrical coordinates) is simply $\Phi = | \Phi | e^{i\theta}$. The dashed line is $A_\theta gr$, with G the gauge coupling. The gauge potential A_θ asymptotically tends to $1/(gr)$ so the total 'magnetic' flux down the string is $\int \vec{B}.d\vec{S} = \int \vec{A}.d\vec{l} = 2\pi/g$. The radius is in units of $1/(g\eta)$, and the solution is shown for $\lambda = g^2$, the case where the Higgs particle mass equals the gauge particle mass in the broken phase (Section 2.7).

Figure 2.6

The approximation used in simulations of the formation of the simplest 'U(1)' cosmic strings (Section 2.8).
a) The space V_{min} (the circle) is approximated by 3 points. Each domain is assigned a value 0,1 or 2 corresponding to one of these points. The rule for

the phase of the Higgs field is to move along the shortest path on V_{min} when going from one domain to the next.
b) an edge with four surrounding domains - this edge is a string.
c) this edge is not a string.

Figure 2.7

A box of strings at formation. The strings are formed on a cubic lattice with the algorithm described in Figure 2.6. The box is tilted and the string width decreases away from the nearest side of the box to give an impression of perspective. Periodic boundary conditions are used so all of the string is in the form of closed loops- most of it is actually in a single very long loop.

Figure 2.8

Coordinates used in the derivation of the Nambu action given in Section 2.9 . It is assumed that the radius of curvature of the string, R, is much greater than it's width W.

Figure 2.9

When two strings collide they reconnect the other way (Section 2.9).

Figure 2.10

A selection of simple loop trajectories in flat spacetime which do not self-intersect at any time. The solutions plotted are given by $\vec{r}(\sigma,t) = \frac{1}{2}((1-\alpha)sin\sigma_- + \frac{\alpha}{3}sin3\sigma_- + sin\sigma_+, -(1-\alpha)cos\sigma_- - \frac{\alpha}{3}cos3\sigma_- - cos\phi cos\sigma_+, -2\sqrt{\alpha(1-\alpha)}cos\sigma_- - sin\phi cos\sigma_+)$ where $\sigma_\pm = \sigma \pm t$ and σ runs from 0 to 2π. This family of solutions satisfies the equations (52) for all α and ϕ. They are periodic, with period π.

Figure 2.11

A string network in a radiation dominated universe. The cube shown is half a Hubble radius on each side and the scale factor of the universe has doubled since the beginning of the simulation. The picture is from a forthcoming paper by Albrecht and myself.

Figure 2.12

The string density in the latest simulations by Albrecht and myself, for radiation dominated universes. The vertical axis shows the string density in units of the string tension divided by the Hubble radius squared $(\mu/(2t)^2)$. The horizontal axis shows the comoving Hubble radius in units of the initial domain size ξ when the strings are formed. This is proportional to the scale factor. The scaling density appears to be at $200\mu/(2t)^2$. The upper and lower solid lines show simulation results starting at higher and lower than scaling density. The dashed line shows the prediction of the simple model explained in Section 2.12.

Figure 2.13

The conical metric of a long straight string. If an angle $8\pi G\mu$ is 'cut out' of flat space (left) and the two edges identified, a cone is the result (right).

Figure 3.1

The equilibrium loop distribution at low density in flat spacetime found in numerical simulations by Smith and Vilenkin (squares) for different cutoffs in the size (number of points) of the loops allowed, and the prediction from the analysis of Mitchell and myself explained in Section 3.5 for the same quantity. $N \geq E$ is the number of loops with energy greater than E and δ is the energy per point on the string. μ ids the string tension. The two parameters in the predicted distribution were fitted to one curve, then used to predict the second.

Figure 3.2

The results of flat spacetime simulations by Sakellariadou and Vilenkin.
a) The equilibrium loop distribution at high density. The three sets of points show the distribution at $\rho = 0.2$ (open circles), $\rho = 0.45$ (solid circles) and $\rho = 0.8$ (triangles). The solid line is the prediction, explained in Section 3.7, that $n(E) \propto E^{-\frac{5}{2}}$.
b) the total density in loops (total density minus the density in long string) versus the total density. At low density all the string is in loops, with the distribution of Figure 3.1 . As the density is increased, a point is reached beyond which the density in loops saturates and everything goes into the long string. This is explained in Section 3.7 .

Figure 3.3

The energy versus temperature curve discussed in Section 3.8, for spacetime dimension $d \geq 4$. The dashed line shows the canonical prediction, where one fixes T and calculates E. The calculation fails above the Hagedorn temperature T_H. The solid line shows the microcanonical result, where one fixes E and calculates T. In the infinite volume limit this turns vertical beyond T_H.

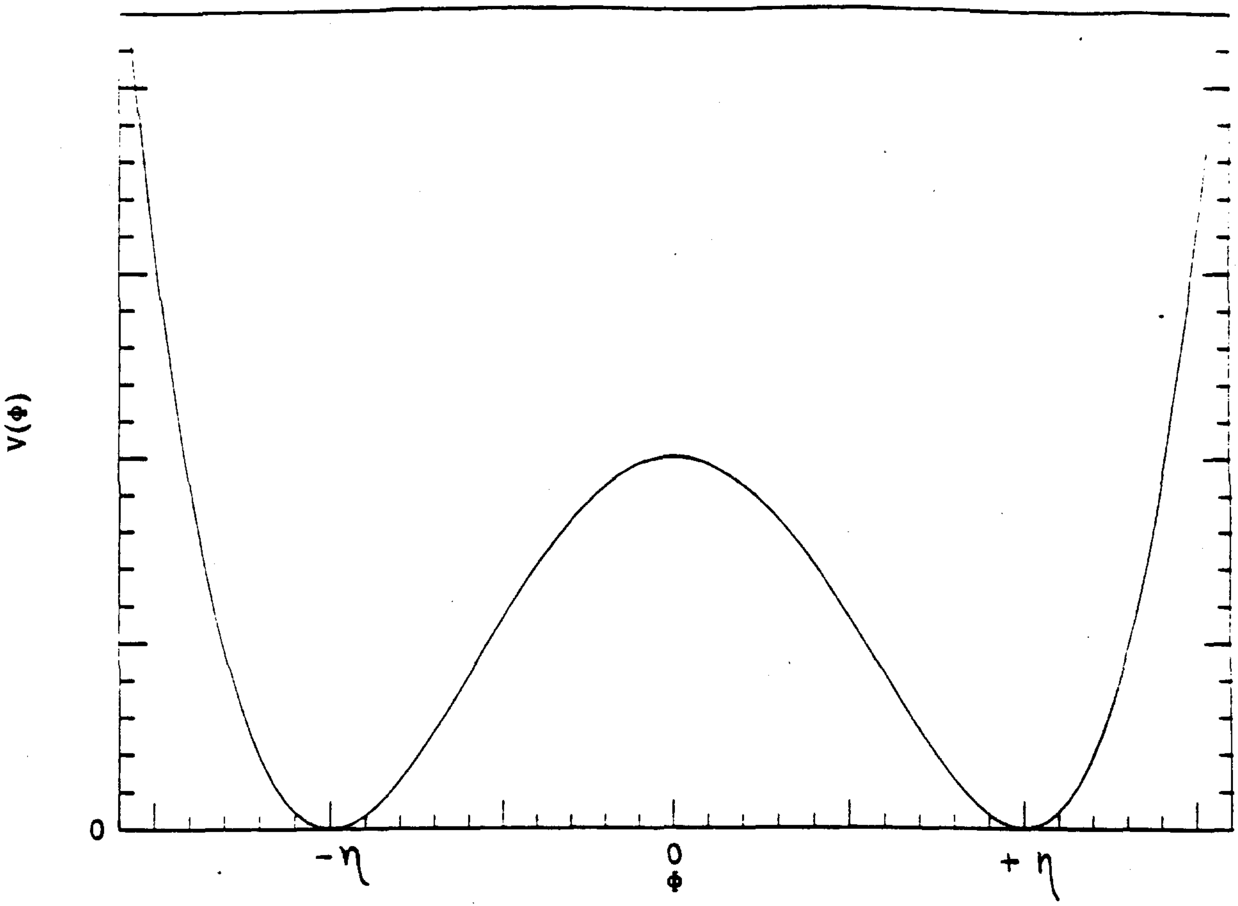


Figure 2.1

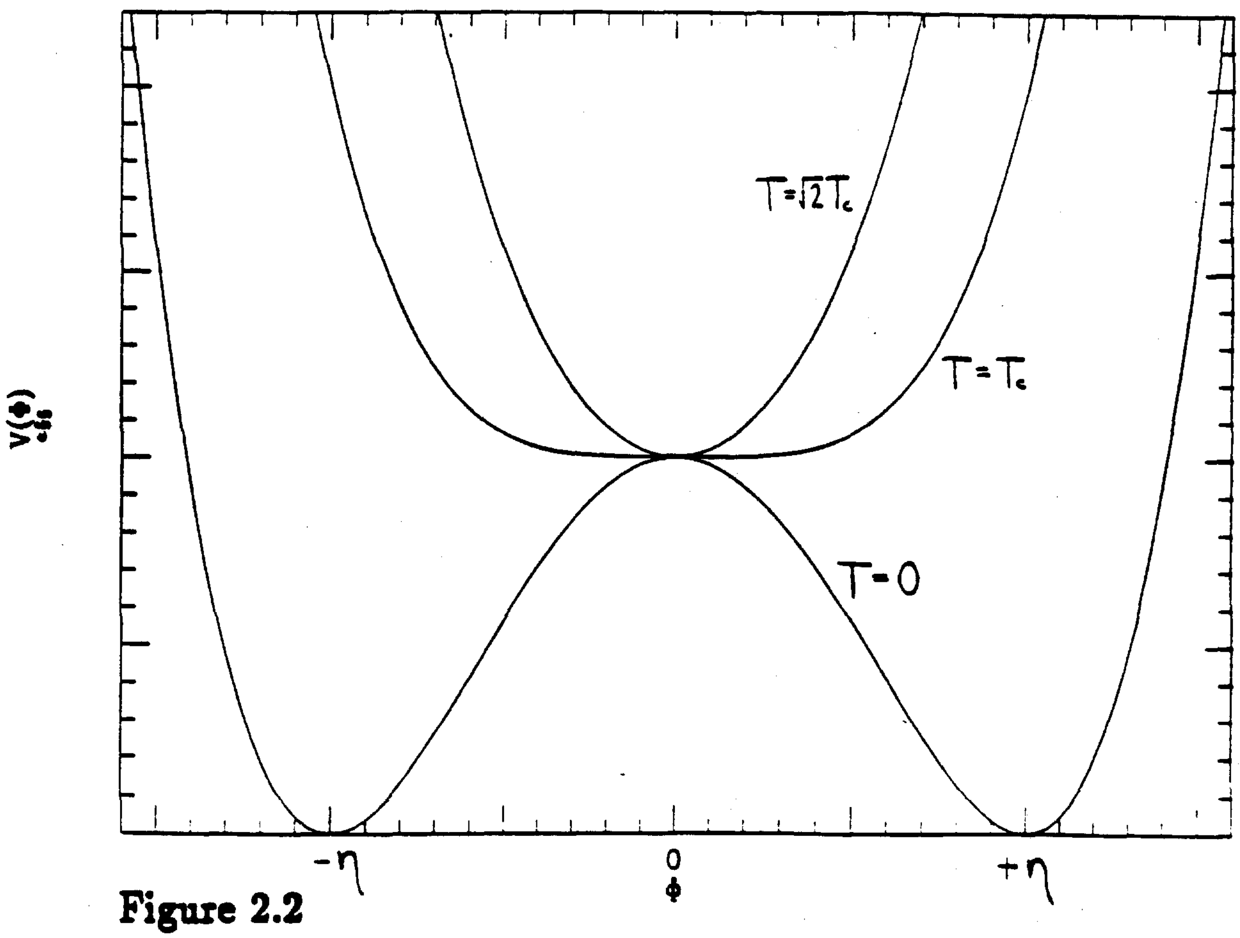


Figure 2.2

Figure 2.3

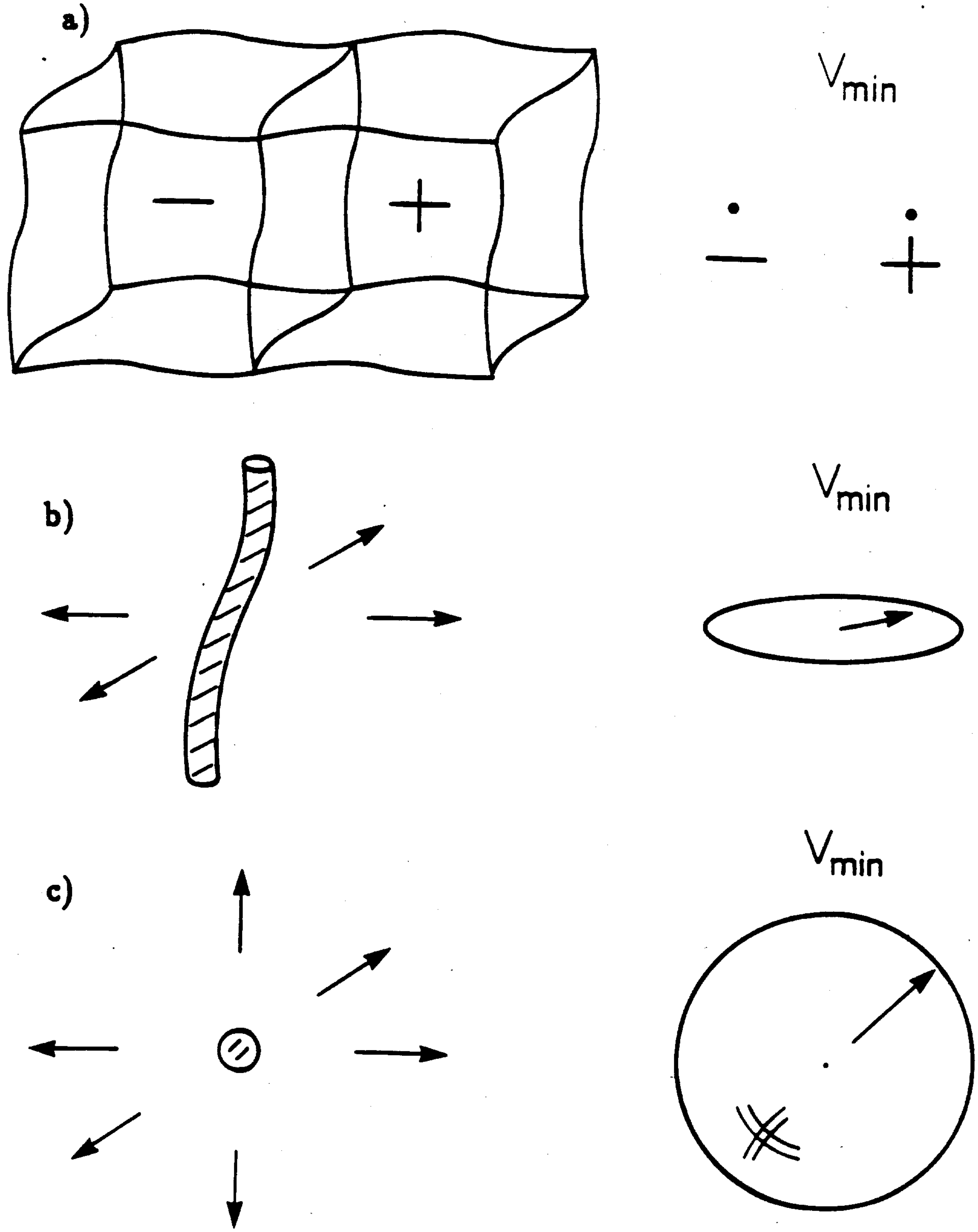
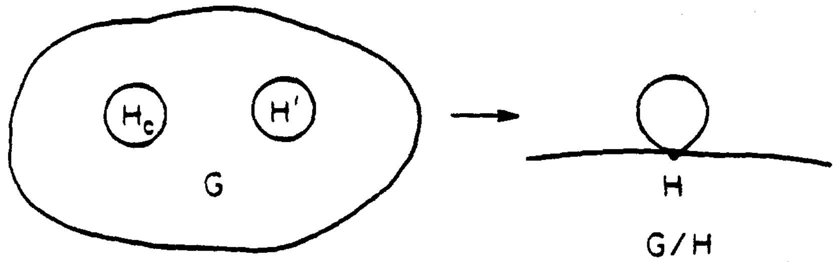


Figure 2.4



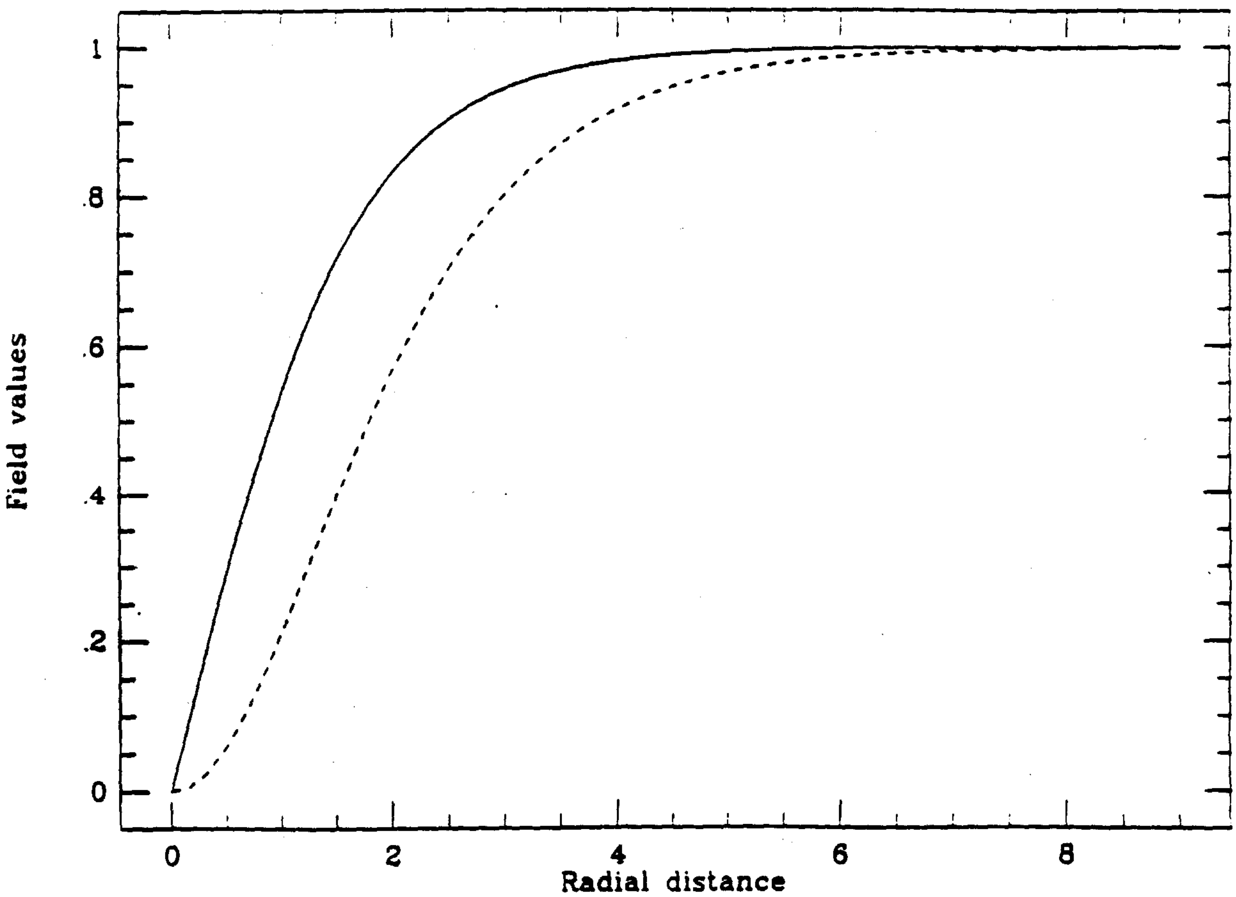


Figure 2.5

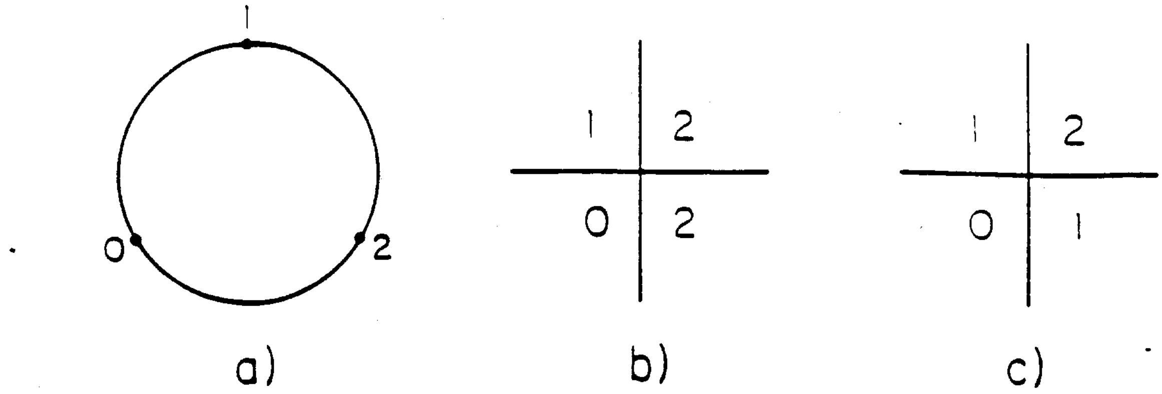


Figure 2.6

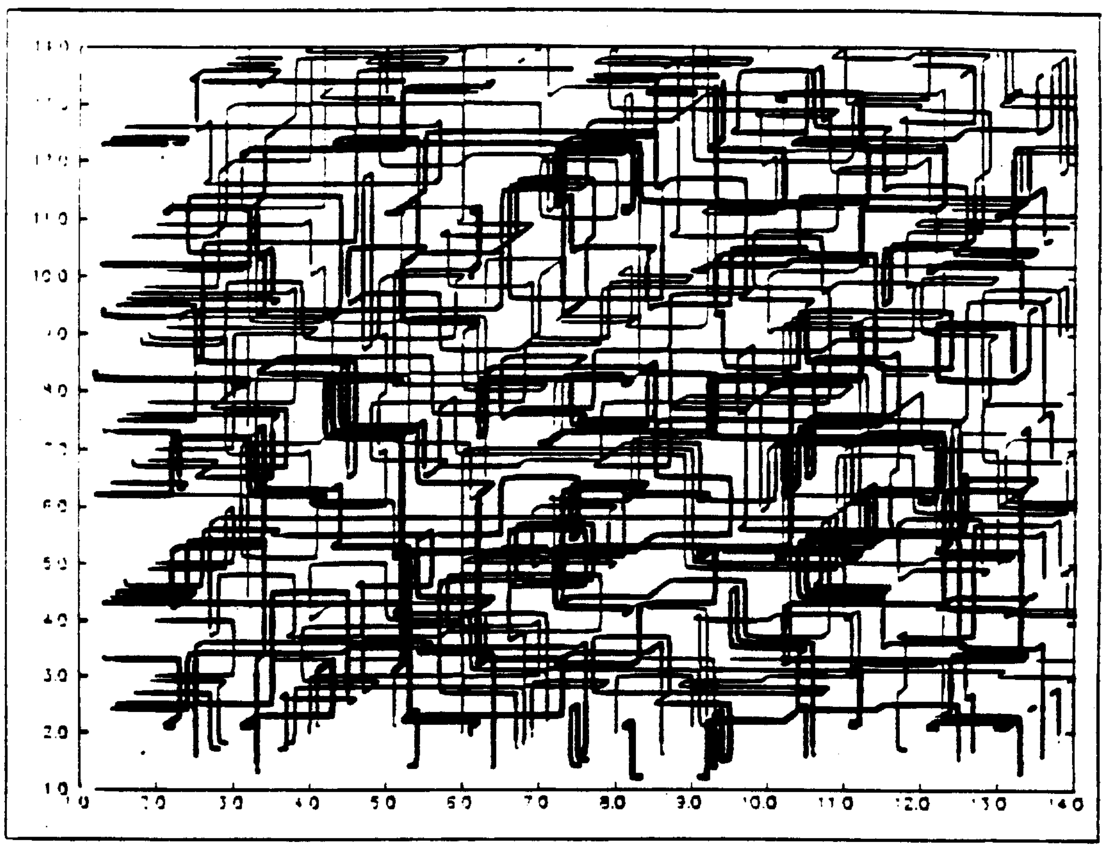


Figure 2.7

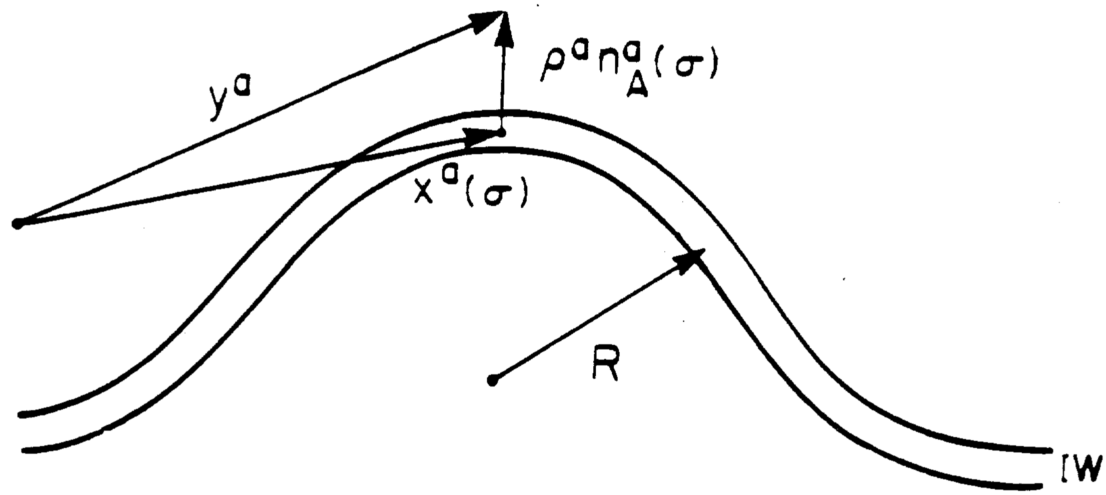


Figure 2.8

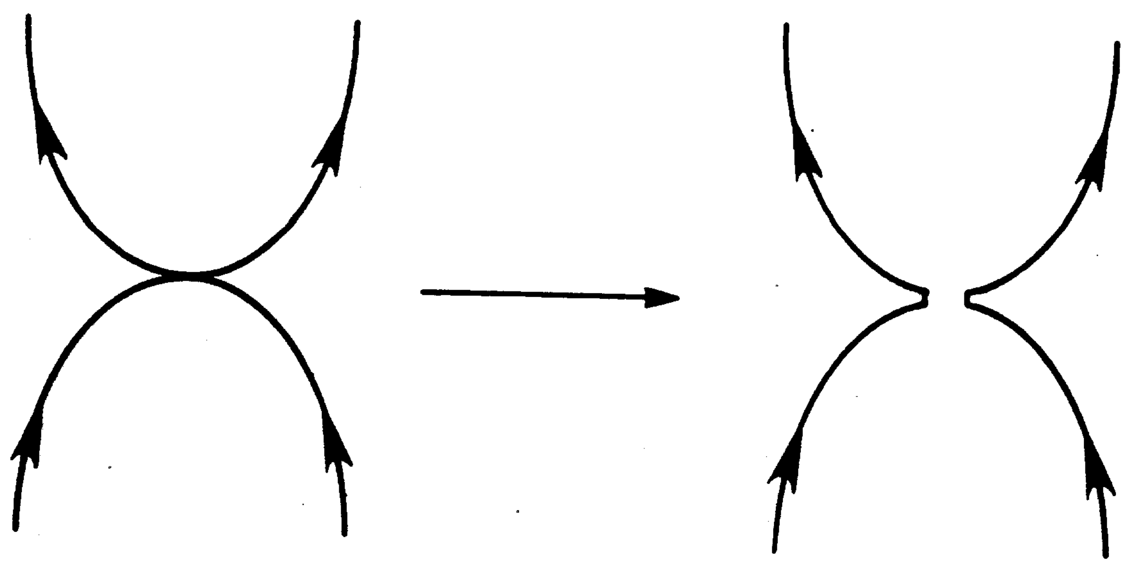


Figure 2.9

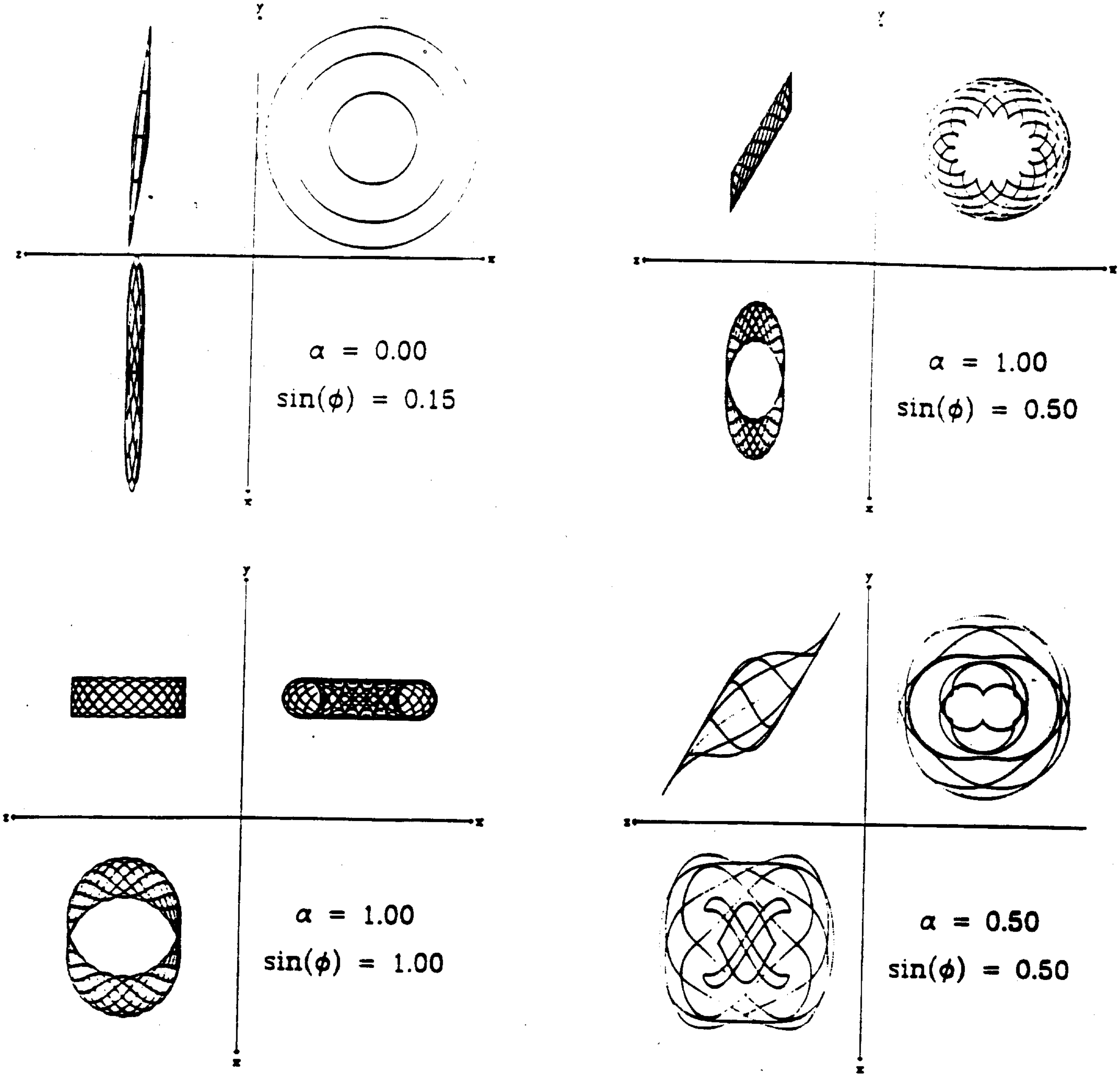


Figure 2.10

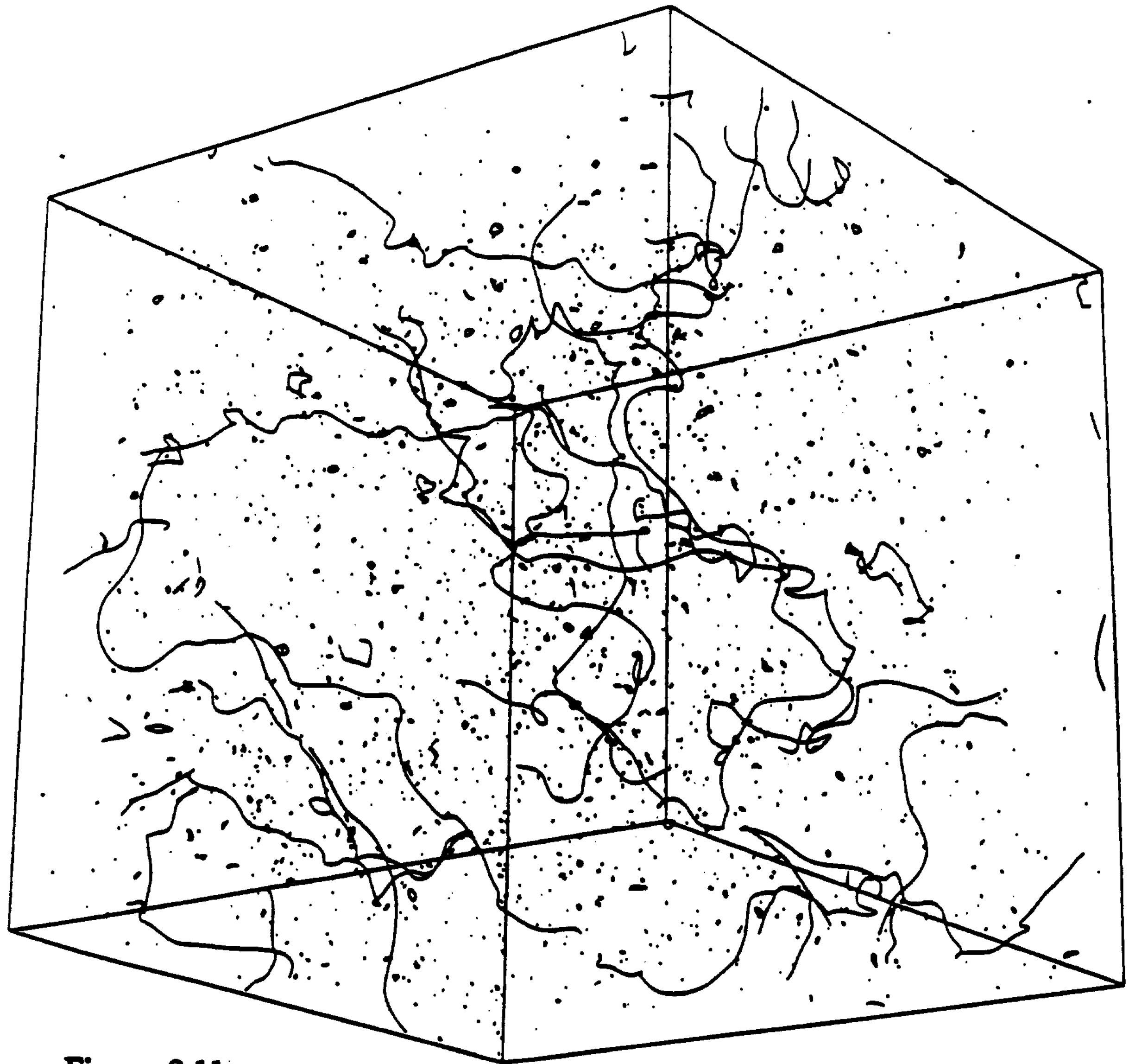


Figure 2.11