Fermi National Accelerator Laboratory

FERMILAB-Conf-87/206-A
November 1987

Cosmology for High Energy Physicists *

Andreas Albrecht
NASA/Fermilab Astrophysics Center
MS209 Fermi National Accelerator Laboratory
P.O.B. 500
Batavia, IL 60510-0500

November 12, 1987

1 Introduction

The idea that our universe endured a hot dense past before expanding and cooling to its present state has caused a special relationship to develop between the fields of high energy physics and cosmology. In order to describe our high temperature past, an understanding of high energy physics is required. On the other hand, our understanding of high energy physics is incomplete, and the early universe has proven to be a useful "laboratory" in which to test new ideas. A new high energy theory can be used to model earlier epochs. If this new high energy physics affects the evolution of the universe in a way that can be detected today, a test of that theory results. In addition, applying high energy theories to studying the early universe has introduced interesting new ideas into the field of cosmology.

This field is too vast to cover in three lectures. (For more detailed discussions see, for example, references [1], and [2].) My goal is, by presenting a selection of topics, to convey my enthusiasm for the field and to point out areas which could prove fruitful for further research.

*Lectures presented at the 1987 Theoretical Advanced Studies Institute, Santa Fe, New Mexico

When one applies theories of the high energy physics to the early universe one can take different approaches. The most conservative approach is to simply make sure that your beautiful theory of particles does not have any cosmological disasters hidden in it. As we shall see, there are ways in which a model can have a cosmology which is inconsistent with our observed universe, even if the particle physics looks great. On the other hand, there are unsolved issues in cosmology which seem to look naturally toward high energy physics for resolutions. There is a lot of work to be done just in exploring what high energy physics can and can not do to address these issues. Part of the fun, of course, is that you may start out with one of these approaches but there is no telling in which area you may end up contributing. It pays to keep your eyes open to all possibilities. As these lectures progress I will be stopping to emphasize points that could cause cosmological stumbling blocks for models of high energy physics, as well as to present some important cosmological issues waiting to be solved.

In a certain sense cosmology lacks a definiteness found in other areas of physics. One can not go back in time and determine what did and did not happen the way one can measure a cross section or a critical temperature in the laboratory. Still, our curiosity drives us to understand the early universe as best we can, and there are constraints which serve as tests of cosmological models. First of all one must have internal consistency. For example, a proposed change at the grand unification scale could have unacceptable consequences at a later epoch when nuclear physics plays a dominant role. Then there is the obvious need for consistency with current observations. As we shall see, events in the early universe can leave traces (sometimes dramatic, sometimes subtle) which can appear in, and possibly contradict, current observations. The fact that many new and more detailed observations are expected in the coming years makes this a particularly exciting time to be studying cosmology.

I will start these lectures with a presentation of the standard big bang model of cosmology. [1] Although not perfect, its many successes make it a good starting point for most discussions of cosmology. I will point to places where well understood laboratory physics is incorporated into the big bang,

[1] For a more complete discussion of the standard big bang model see, for example, reference [3]

leading to successful predictions. Then I will move on to much less established aspects of high energy physics and discuss some of the new ideas they have introduced into the field of cosmology.

2 The Standard Big Bang: Evolution of Spacetime

Cosmologists find it attractive to assume that, when viewed on large enough scales, our region of the universe looks just the same as any other. That is, on large scales the universe is homogeneous, no point is singled out. It is almost as if we are still bitter over the death of geocentrism, and declare that if *we* can not be the center of the universe then no one else can either! Cosmologists also like to assume that the universe is isotropic, so that there is no preferred direction in the universe just as there is no preferred point. These two assumptions together are known as "the cosmological principle". This may sound like something deep, but of course the reason this principle is popular is that it leads to a successful model.

2.1 Friedmann-Robertson-Walker Cosmologies

Any homogeneous and isotropic spacetime can be described by the Robertson-Walker metric in which the invariant length ds is given by [2]

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

The coordinates r, θ and ϕ provide a parameterization of the three space dimensions. The coordinate t is a measure of time, and $a(t)$ is called the "scale factor" and gives an overall scale as a function of time. If Santa Fe were expanding uniformly the number of blocks from this building to the plaza would remain constant, while the length of a block, and the total distance, would be increasing. The number of blocks would be the analogue of r, θ and ϕ while the block size would be the analogue of $a(t)$.

[2]Throughout I use units where $c = \hbar = k_{Boltzmann} = 1$.

For Robertson-Walker spacetimes the curvature scalar of three dimensional space is given by

$$R(t) = ka(t)^{-2} \tag{2}$$

and is independent of position since space is homogeneous. The curvature scalar distinguishes between three types of space:

Type	Curvature	Spatial Volume
closed	$R > 0$	$2\pi^2 R^{-3/2}$
flat	$R = 0$	infinite
open	$R < 0$	infinite

If $R \neq 0$, I will measure $a(t)$ in units of $|R|^{-1/2}$ so $k = \pm 1$. A flat universe appears in the $a \to \infty$ limit in these units. In the case of the closed universe there exist homogeneous and isotropic models with different global topologies, but locally they look the same.

In almost all modern theories of cosmology it is assumed that general relativity correctly describes the dynamics of spacetime, at least over length scales greater than the Planck length ($\equiv G_{Newton}^{1/2}$). Thus, to calculate the dynamics one needs to know the stress- energy tensor ($\mathbf{T}$) of the matter in the universe. Homogeneity and isotropy restrict $\mathbf{T}$ to the form:

$$T^{\mu}_{\nu} = \begin{pmatrix} \rho(t) & 0 & 0 & 0 \\ 0 & p(t) & 0 & 0 \\ 0 & 0 & p(t) & 0 \\ 0 & 0 & 0 & p(t) \end{pmatrix} \tag{3}$$

again, independent of position. The parameter ρ is the energy density and p is called the pressure. When the matter takes on a familiar form such as that of an ideal gas, p is the pressure we are used to defining.

Einstein's equations in a Robertson-Walker spacetime give

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \tag{4}$$

which is known as the Friedmann equation. Here G is Newton's constant. Note that (for $\rho > 0$) in an open or flat ($k = -1$ or 0) universe $\dot{a}$ can not go through zero. In this case an expanding universe will expand forever. For a closed universe ($k=1$) $\dot{a}$ can go through zero. Typically an expanding closed universe will eventually re-contract.

To solve Eqn (4) for $a(t)$ one needs to know $\rho(a)$. The conservation of stress-energy ($\nabla \mathbf{T} = 0$) gives

$$\frac{d}{da}\left(\rho a^3\right) = -3pa^2. \tag{5}$$

If $p(\rho)$ is given then Eqn (5) can be solved for $\rho(a)$ and the evolution of $a(t)$ can be calculated.

2.2 Input From Microscopic Physics

The one place where the physics of microscopic processes comes in to calculating $a(t)$ is in determining the relationship between the pressure and the energy density. The equation which gives $p(\rho)$ is called the equation of state. In the standard big bang there are two equations of state which appear.

2.2.1 Non-relativistic matter

Today most of the matter we observe in the universe appears to be non-relativistic:

$$\text{Kinetic Energy} \ll \text{Rest Energy}. \tag{6}$$

Because the kinetic energy is relatively small

$$p = 0 \tag{7}$$

is a very good approximation to the equation of state. This is known as a non-relativistic, dust dominated, or matter dominated equation of state. Plugging into Eqn (5) gives

$$\frac{d(\rho a^3)}{da} = 0 \tag{8}$$

which is solved by

$$\rho = \frac{B}{a^3}\ . \tag{9}$$

Here B is a constant of integration. Equation (9) can be understood by noting that the amount of energy in a co-moving volume remains the same while the volume scales as a^3.

2.2.2 Relativistic matter

In most models the particles in the hot early universe are highly relativistic:

$$\text{Kinetic Energy} \gg \text{Rest Energy}. \tag{10}$$

As with a relativistic ideal gas, the equation of state is

$$p = \frac{1}{3}\rho \,. \tag{11}$$

This is called a relativistic or"radiation dominated" equation of state. In this case Eqn (5) gives

$$\frac{d(\rho a^3)}{da} = -\rho a^2. \tag{12}$$

Equation (12) is solved by

$$\rho = \frac{D}{a^4} \tag{13}$$

where D is a constant of integration. Equation (13) can be understood by realizing that for relativistic matter, as with photons for example, the energy in a co-moving volume scales as a^{-1} or "redshifts", while the volume still scales as a^{-3}.

2.3 General Properties of the Big bang

I will now introduce a crude version of the standard big bang. At very early times the matter was very hot, so it is taken to be relativistic. As the universe expands and cools the kinetic energy of the matter redshifts, until eventually it becomes non-relativistic. I will assume that this transition occurs instantaneously at a time t_*. Corrections to this approximation will not significantly alter the qualitative discussion which follows.

This model has three adjustable parameters, D (the integration constant in the relativistic era), k (the curvature), and t_*. The integration constant (B) in the non-relativistic era can be calculated by writing

$$\rho(t_*) = \frac{B}{a(t_*)^3} = \frac{D}{a(t_*)^4} \tag{14}$$

and solving for B. The value of t_* comes from microphysics. It is the masses of the particles and the spectrum of bound states which tell us at what stage the matter becomes non-relativistic. That leaves two "cosmological" parameters, k and D, which we must fit to our universe.

Before fitting the big bang model to our universe we can make some interesting general remarks. The Friedmann equation (4) can be re-written as

$$\dot{a} = \sqrt{\tfrac{8\pi}{3}G\rho a^2 - k} \qquad (15)$$

$$\uparrow$$

$$\sim a^{-2} \text{ or } a^{-1}$$

were we take the positive root because the universe is observed to be expanding today. Using expression (9) or (13) for ρ one finds that the first term in the square root goes as a^{-2} or a^{-1} respectively. Since $\dot{a}$ is a decreasing function of a, an upper bound on a in the past can be found by taking $\dot{a}$ constant at today's value:

$$a(t) < a_{ub}(t) \equiv a_{now} \cdot \left(\frac{t - t_0}{t_{now} - t_0}\right) \quad \text{for } t < t_{now}. \qquad (16)$$

The fact that a_{ub} goes to zero at t_0 means that a itself was zero at some finite time in the past (no earlier than t_0). A Robertson-Walker spacetime with $a = 0$ is singular and is not physically understood. Such singularities are known to appear generically in most spacetimes, not just in Robertson-Walker[4]. The real frontier of our understanding, however, lies at small but finite values of a which correspond to very large values of ρ and require a knowledge of ultra-high energy physics to understand. It is possible that new ideas from this front could allow such singularities to be avoided or understood.

Still, the Standard Big Bang does have an $a = 0$ singularity. We usually choose $t = 0$ at this point and call it the beginning of the universe. Using a_{ub} one can get an upper bound on the age of the universe:

$$t_{now} = \int_0^{a_{now}} \dot{a}^{-1} da < \int_0^{a_{now}} \dot{a}_{ub}^{-1} da = \left.\frac{a}{\dot{a}}\right|_{now}. \qquad (17)$$

Any evidence that the universe is older than this bound would contradict the standard big bang model.

2.4 Fitting D and k to Observations

The value of $\dot{a}/a$ today is called H, the Hubble constant. The recession speed s_r of an object a distance d away due to the expanding universe is

$$s_r = Hd. \tag{18}$$

The observed value of H is written[5]

$$H = h \times 100 km\, sec^{-1}\, Mpc^{-1} \simeq \frac{h}{10^{10} yr} \tag{19}$$

where $1 Mpc$(megaparsec) $= 3 \times 10^{24} cm$. The parameter h is bounded by

$$1 \geq h \geq 0.4 \tag{20}$$

and represents our uncertainty in H.

In order to fit the parameter k in the big bang one must measure the spatial curvature of the universe. It is convenient to define ρ_{crit} (using Eqn (4))

$$\left(\frac{\dot{a}}{a}\right)^2 = H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \equiv \frac{8\pi G}{3}\rho_{crit}. \tag{21}$$

If $\rho = \rho_{crit}$ then the expansion is driven entirely by ρ, and $k = 0$. Today $\rho_{crit} \simeq 2h^2 \times 10^{-29} gm/cm^3$. It is often convenient to work with $\Omega \equiv \rho/\rho_{crit}$. Then

$$\frac{k}{a^2} = \frac{8\pi G}{3}\rho_{crit}\,(1 - \Omega) \tag{22}$$

gives the spatial curvature.

If one just considers the luminous matter (stars) in the universe one finds that[6]

$$\Omega_{lum} \equiv \frac{\rho_{lum}}{\rho_{crit}} \simeq .01 \tag{23}$$

giving $\Omega \geq .01$. On the other hand one can also estimate the masses of dynamical systems by inferring the gravitational forces necessary to explain the motions of observed objects. When this technique is applied to galaxies, one gets[7]

$$\Omega_{dyn} \equiv \frac{\rho_{dyn}}{\rho_{crit}} \simeq 0.1 - 0.3 \tag{24}$$

which gives a higher lower bound on the Ω.

The discrepancy between Eqns (23) and (24) is the starting point for a very interesting and active area of current research: The search for the "dark matter" or "missing mass". Theoretically this matter could be some exotic new particles from grand unification, or a more familiar form of dark matter such as planets. This issue provides a classic opportunity for high energy physics to interact with cosmology and it is the subject of Blumenthal's lecture in this volume.

We are not yet done with putting bounds on Ω. The age of the universe is related to Ω, as one can see qualitatively by examining Eqn (15). The curvature contribution to the right side does not decrease with a while the ρ contribution does. The more the curvature dominates, the more slowly $\dot{a}$ decreases and the faster the universe evolves. One can make arguments based on geology, as well as stellar evolution and the evolution of nuclear abundances, which put lower bounds on the age of the universe[8]. There is general agreement that the age of the universe is no less than 10 billion years. This conservative bound gives[9]

$$\Omega \leq 1.1 \frac{1}{h^2} \leq 6.9. \tag{25}$$

Putting Eqns (24) and (25) together one gets

$$0.1 \leq \Omega \leq 6.9. \tag{26}$$

The fact that we do not know if Ω is greater than or less than one means that we do not know what sign to give the spatial curvature in the big bang model. Thus, we can not predict whether the universe will go on expanding forever, or eventually re-contract. Equation (26) does tell us, however, that ρ is not "too far" from ρ_{crit} or that the spatial curvature is not "too large". In fact

$$-\frac{8\pi G}{3} 5.9 \rho_{crit} \leq \frac{k}{a^2} \leq \frac{8\pi G}{3} 0.9 \rho_{crit} \ (today) \tag{27}$$

are the limits on the curvature.

We have already passed over an important success of the big bang model which is often under-emphasized. The big bang gives a finite upper bound on the age of the universe. In determining which values of Ω are consistent with the lower bounds on the age of the universe, it is perfectly possible that a contradiction would arise rather than Eqn (25). If the big bang were to allow the universe to be a maximum of a few thousand years old, it might be adopted by creationists, but it would have no place in modern science. It is still possible that as these lower and upper bounds on the age of the universe are refined, the big bang model might become inconsistent. There are some who believe we are on the brink of such an inconsistency today.

Now that we have determined the spatial curvature as best we can, let us try to fit the one remaining parameter, the integration constant D. A simple lower bound on D comes from looking at the microwave background photons. Currently we observe the universe to be filled with a thermal distribution of photons at about 3K. These microwave photons contribute a relativistic component ρ_{mw} to ρ today, as they did in the relativistic era. Thus

$$\rho_{mw} = \frac{D_{mw}}{a^4} \text{ (all the time)}. \tag{28}$$

In the relativistic era the photons were not the only component to the energy density so

$$D > D_{mw}. \tag{29}$$

Using the standard formula for ρ of a relativistic thermal gas and the bound on a implied by Eqn (27) gives

$$\begin{aligned} D &> D_{mw} \\ &= \rho_{mw} a^4 \Big|_{\text{today}} \\ &= (3K)^4 \frac{2\pi^2}{30} a^4 \Big|_{\text{today}} \\ &\geq 6.3 \times 10^{114}. \end{aligned} \tag{30}$$

What does such a large value of D mean? It is, after all, just a free parameter in the big bang model and it can be what ever it likes. Many people choose, however, not to simply leave it at that. I will now discuss several puzzles or problems that people associate with such large values of D. In the process I hope the physical meaning of the dimensionless number D will become clearer.

2.4.1 Large numbers puzzle

Some people balk at any theory in which dimensionless parameters must be set very far from one. They take it as a sign that something important is being swept under the rug, and argue that if a model could be found without this problem it would have considerably greater appeal.

2.4.2 Large entropy puzzle

The large numbers puzzle is often re-cast as the large entropy puzzle. For a relativistic gas of particles in equilibrium the energy density (ρ) and the entropy density (s) are given by

$$\rho = \tilde{g}\frac{\pi^2}{30}T^4 \tag{31}$$

$$s = \tilde{g}\frac{2\pi^2}{45}T^3 \tag{32}$$

where $\tilde{g}$ is the number of Bose plus 7/8 the number of Fermi degrees of freedom. On can then see that

$$\begin{aligned} D^{3/4} &= \left(\rho a^4\right)^{3/4} \\ &= \left(\tilde{g}\frac{\pi^2}{30}T^4 \cdot a^4\right)^{3/4} \\ &= 1.8(\tilde{g}\pi^2)^{-1/4} \cdot \left[s a^3\right] \end{aligned} \tag{33}$$

so $D^{3/4}$ is roughly the entropy contained in a cube one radius of curvature on a side. Thus large D can be thought of as large entropy.

2.4.3 Flatness puzzle

Certainly the large value of D has something to do with the degree of flatness of the universe. Since a flat universe appears as the $a \to \infty$ limit in our units and $\rho = D/a^4$, a nearly flat universe with finite ρ must necessarily have a large D. The relationship between large D and flatness can be explored by rewriting Eqn (21) as

$$\frac{\rho_{crit}}{\rho} = 1 - \frac{k}{\rho a^2 \frac{8\pi}{3} G}. \tag{34}$$

Using the definition of the Planck length ($l_P^2 \equiv G$) and Eqn (13) for ρ in the relativistic era this becomes

$$\left|\frac{\rho_{crit}}{\rho} - 1\right| = \frac{\frac{3}{8\pi}}{D}\left(\frac{a}{l_P}\right)^2 < 3 \times 10^6 \left(\frac{a}{a_{now}}\right)^2. \tag{35}$$

Bounds on D and a_{now} from Eqns (30) and (27) were used to get the inequality. If D were $O(1)$ Eqn (35) would give a large deviation from flatness for a universe where $a/l_P \gg 1$. The large value of D means our universe is "flat for its size". Equation (35) also says that for our universe to be as flat as it is today, it must have been much more flat in the past. For example, at the era when grand unification physics would be relevant, $a/a_{now} \simeq 10^{-29}$. Equation (35) then gives

$$\left|\frac{\rho_{crit}}{\rho} - 1\right| < 3 \times 10^{-52} \tag{36}$$

for the grand unification era.

How much one chooses to be puzzled by these puzzles is to some extent a matter of personal taste. As I said earlier, one can regard D as a free parameter which must be set, and leave it at that. Some find it appealing to declare, as a new principle, that $k = 0$. In this case ρ is identically ρ_{crit} at all times and $D = \infty$. Of course, it is still possible that such a principle would be contradicted by future observations.

Probably the strongest basis for concern comes from demanding that the adjustable parameters of today's models become the derived parameters of tomorrow's. From this point of view *all* free parameters in present theories are puzzles awaiting resolution. To some it seems more likely that parameters can be derived in some new super theory when they are all $O(1)$, and numbers like 10^{114} seem less easy to come by. One way to demonstrate this point is to suppose that some fundamental development caused us to modify the big bang model at very early times and left us with the task of matching up to the standard big bang at, say, the end of the grand unification era. According to Eqn (36) we would be required to have $\rho = \rho_{crit}$ with an accuracy of one part in 10^{52} ! If the requirement were simply $\rho \simeq \rho_{crit}$ the task would appear much easier. Critics of this line of thinking argue, however, that it is ridiculous to talk about what is easy or hard to do with a new theory until one knows what it is[10].

2.5 Causality Structure of the Big Bang

The region in causal contact with an event is determined by how far photons which start at that event can travel. For a photon $ds = 0$ so Eqn (1) gives

$dt^2 = a^2 dl^2$ where l is the coordinate distance that the photon has traveled. The physical distance d a photon travels between times t_1 and t_2 is then given by

$$d = a(t_2)l = a(t_2)\int_{t_1}^{t_2} a^{-1}dt \tag{37}$$

so given $a(t)$ one can solve for d. We have seen (from Eqn (35)) that for most of the history of the universe the curvature of the universe can be neglected. This makes it particularly easy to solve the the Friedmann equation (Eqn (4)) for $a(t)$. In the relativistic era (using Eqn(13))

$$a = \left(2t\sqrt{\frac{8\pi}{3}GD}\right)^{1/2} \tag{38}$$

which gives

$$d = 2(t_2 - \sqrt{t_1 t_2}) \stackrel{\text{large}\Delta t}{\rightarrow} 2t_2. \tag{39}$$

In the non-relativistic era (using Eqn (9))

$$a = \left(\frac{3}{2}t\sqrt{\frac{8\pi}{3}GB}\right)^{2/3} \tag{40}$$

which gives

$$d = \frac{3}{1}(t_2 - t_1^{1/3}t_2^{2/3}) \stackrel{\text{large}\Delta t}{\rightarrow} \frac{3}{1}t_2. \tag{41}$$

This means that throughout the evolution of the universe the largest regions in causal contact have a size of roughly t.

One result of the above discussion is that at $t = 0$ there is no causal contact in the universe. At first sight this may seem strange since $a(0) = 0$ and the distance between any two points in the universe is zero. Of course, one must be careful with such a singularity. As one works back toward $t = 0$ the causal regions decrease in size as t, while the separation between points decreases more slowly (see Eqns (38) and (40)). Thus the initial singularity can be thought of as many arbitrarily small causally separate regions which are arbitrarily close together.

3 The Standard Big Bang: Evolution of Matter

3.1 Thermal Equilibrium

It is usually assumed that the matter in the early universe was in thermal equilibrium. Today the matter is clearly not in thermal equilibrium. We see hot stars and cold comets, clusters of galaxies as well as voids. These structures are believed to be the result of gravitational collapse. Gravitational collapse starts to play a serious role in the big bang model after the end of the relativistic era. Before then the high pressure due to the relativistic particles counteracts gravitational collapse, a typical particle having escape velocity from an initial gravitational perturbation. All the effects of gravitational collapse in cosmology are not completely worked out. It seems clear, however, that in the relativistic era the predecessors of the currently observed galaxies and other structures were probably tiny perturbations on the matter distribution. I will neglect these perturbations in most of the following discussion.

A thermal distribution of very relativistic particles ($T \gg m$, where m is the mass of a particle) has an energy density given by Eqn (31). In the very non-relativistic limit ($T \ll m$)

$$\rho = mn \tag{42}$$

where the number density n is given by

$$n = \tilde{g}\left(\frac{mT}{2\pi}\right)^{3/2} e^{-m/T}. \tag{43}$$

Of course thermal equilibrium means equilibrium over a "reasonable" part of phase space, and understanding what degrees of freedom really might be in equilibrium is an important part of the problem. Changes in what is in and out of equilibrium play an important role in many interesting aspects of the big bang model.

During the relativistic era (using Eqns (31) and (13)

$$\rho \sim T^4 \sim \frac{1}{a^4} \Rightarrow T \sim \frac{1}{a} \tag{44}$$

For equilibrium to occur the interactions which maintain equilibrium must proceed quickly compared to the expansion rate ($1/t_{exp}$) of the universe. Using Eqns (31) and (4) one finds

$$\frac{1}{t_{exp}} \equiv \frac{\dot{a}}{a} = O(1) \times \frac{T}{m_P} T \tag{45}$$

where $m_P \equiv \sqrt{1/G} \simeq 10^{19} GeV$. At temperatures much below $10^{19} GeV$ the expansion rate is much slower (by a factor T/m_P) compared with the inverse time scale given by T, which should play an important role in the interactions of relativistic matter. These rough dimensional arguments hint that equilibrium might reasonably occur. Calculating what degrees of freedom really are in equilibrium at a given era demands more work. Here I will report some of the more interesting results of such calculations.

3.2 Photon Decoupling

One interesting event which the big bang predicts is "photon decoupling". At temperatures somewhat above $1/3GeV$ one expects to find photons, as well as e^-, H^+, D^+, $^3He^{++}$, and $^4He^{++}$ particles kept in equilibrium by electromagnetic interactions. As the temperature drops below $1/3GeV$, energies are low enough to allow the electrons and the nuclei to bind into neutral atoms. The absence of free charged particles means that the interactions of the photons with other matter drop off dramatically. The photons at this stage are effectively decoupled from the other matter. The only thing which affects the evolution of the photons is redshifting of their momentum, $\vec{k}$ due to the expansion of the universe:

$$\vec{k}(t) = \vec{k}_d \frac{a_d}{a(t)} \tag{46}$$

(the subscript d refers to the time of decoupling). The thermal nature of the distribution is preserved because the Boltzmann factor for a photon with some initial k_d is the same as one for the redshifted $k(t)$ at some rescaled temperature $\tilde{T}$:

$$exp(\frac{k_d}{T_d}) = exp(\frac{k(t)}{\tilde{T}}) \tag{47}$$

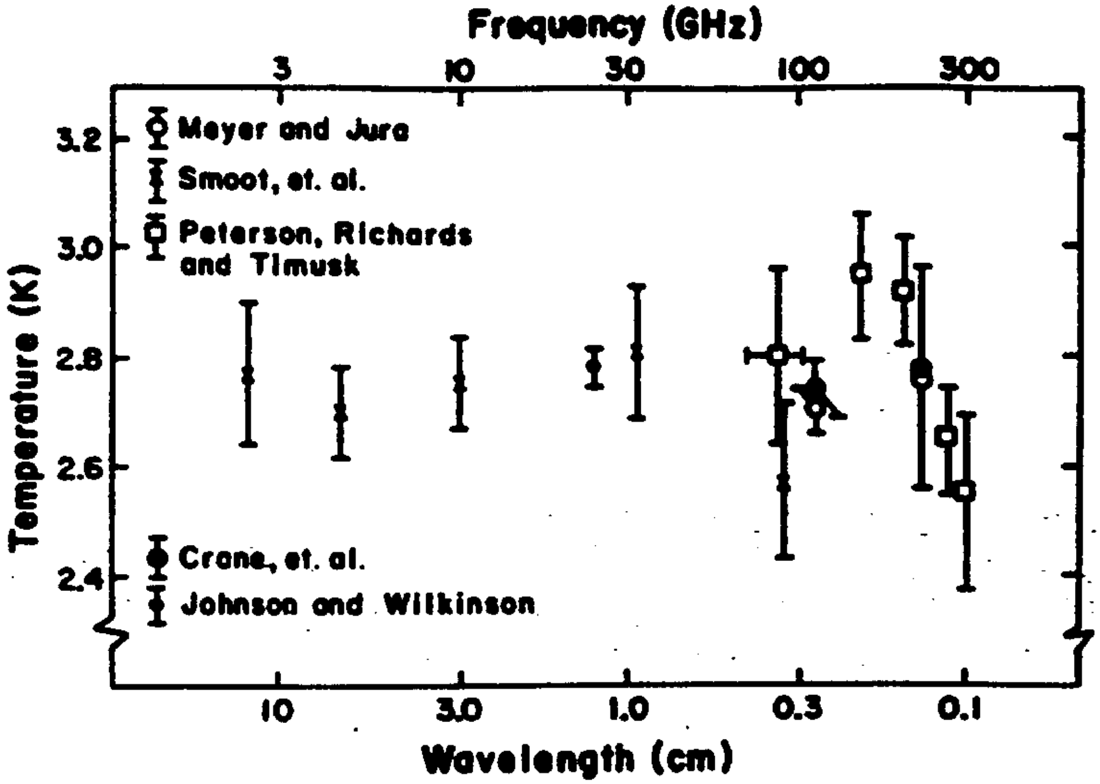

Figure 1: Recent measurements of the microwave radiation. The plot shows equivalent black body temperature as a function of wavelength

where

$$\tilde{T} \equiv T_d \frac{a_d}{a(t)}. \tag{48}$$

Thus the big bang model predicts that there should be a thermal distribution of photons present today, left over from the time when they decoupled from other matter. Probably the most significant experimental result in cosmology is that such a spectrum of photons is indeed observed.

Figure [1] shows the results of several measurements of this so called "microwave background" radiation made at different wavelengths. The vertical axis gives the corresponding temperature assuming that light of that wavelength is in a black body distribution. The results are consistent with a true black body spectrum at a temperature of $2.8K$.

Observations of the microwave background can provide important information to cosmologists, and the results can prove to be serious constraints on cosmological models. For example, the microwave radiation can probe the isotropy of the universe. Radiation reaching us from different directions

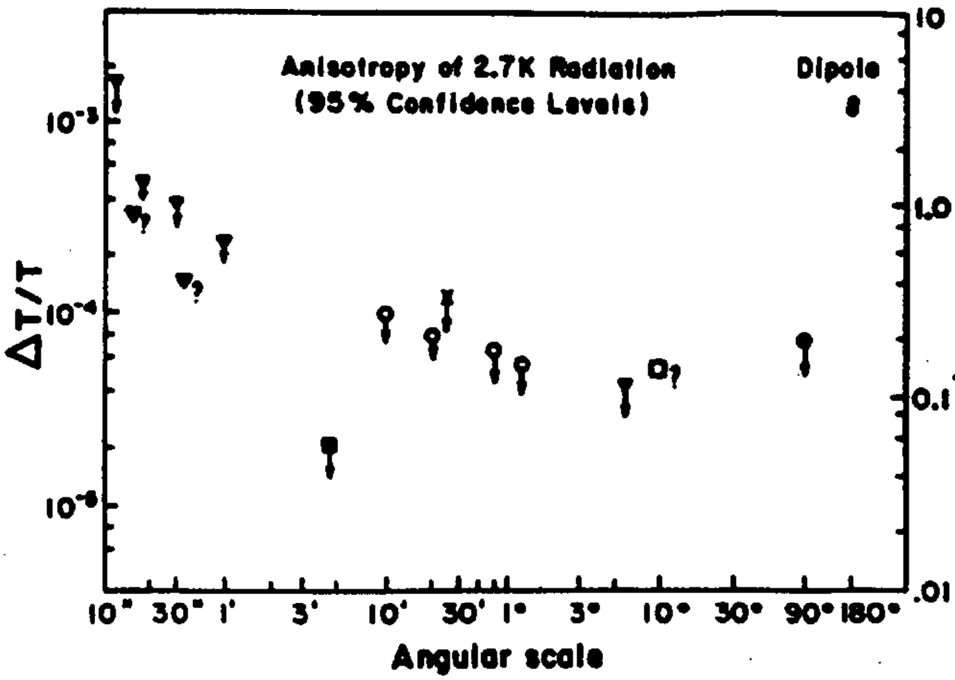

Figure 2: Measurements of the microwave anisotropy.

decoupled at very different locations in the universe. Variations in the temperature observed in different directions can be interpreted as evidence for anisotropy on the corresponding distance scale. Figure [2] shows measurements of variations in the temperature of the microwave background as a function of the angular separation between observing directions. All the certain results (without"?'s") are upper bounds except the dipole measurement. The dipole measurement might represent a large scale anisotropy,[11] but it is simpler (and standard practice) to assume that it is the result of our peculiar motion with respect to the microwave background. With the dipole accounted for in this manner Fig [2] shows a striking confirmation of the isotropy assumption made in the big bang model.

On closer inspection, however, Fig [2] shows that the universe is *too* isotropic for the big bang model. In the previous section we concluded that the the size of a region of the universe which is in causal contact is always roughly t. This fact means that microwave photons originated in causally separate regions if their directions differ by more than a few degrees[3]. Although some degree of isotropy is part of the initial assumptions of the big bang, the data show that a truly striking degree of isotropy is actually present.

What caused such isotropy on larger angular scales? In the big bang model it could not have been causal processes. The striking isotropy of the universe on large scales (at the time of decoupling) can not be explained in the big bang model and is just another aspect which must be set "by hand" to fit the observations.

It is important to keep the microwave background in mind when doing any work in cosmology. There are many examples of ways in which physics, even at much earlier times, could have enough of an effect on the background radiation to contradict current observations. For example, explicit anisotropies in the matter at decoupling might be produced. Also, the microwave background could appear anisotropic after traversing a universe full of gravitational radiation produced at an earlier era.

I should remark that it is possible that the microwave background is not the radiation predicted by the big bang model. Other explanations do exist[12], but they seem rather contrived to me. Perhaps it is a good exercise is to try and imagine alternative sources of the microwave radiation, even if one only ends up convinced that there are no serious contenders.

3.3 Primordial Nucleosynthesis

At sufficiently early times the temperatures were so high that nuclei were easily dissociated. As the universe cooled it eventually became favorable for the protons and neutrons to combine into nuclei. This process has been extensively analysed using large computer codes. I will outline the results here. For a detailed review of this subject see reference [13].

When the temperatures were around $10MeV$ the energy density was still dominated by photons, electrons and positrons, and the light neutrinos and antineutrinos. Equilibrium was sustained by these electromagnetic interactions: $e^+ + e^- \leftrightarrow \gamma + \gamma$, $\gamma + p \leftrightarrow \gamma + p(etc)$, as well as the electroweak interactions $n \leftrightarrow p + e + \nu$, $\gamma + n \leftrightarrow p + e$, and $e + n \leftrightarrow p + \gamma$ which maintain the neutrons and protons in roughly equal numbers. One might think that with binding energies of several MeV there would be many light nuclei around at this stage. However, the equations for nuclear statistical equilibrium give the number densities of deuterons and He^{++} compared with that of baryons to be

$$n_D/n_b \simeq \eta \left(\frac{T}{m_N}\right)^{3/2} e^{B_D/T} \simeq 10^{-13} \tag{49}$$

and

$$n_{^4He}/n_b \simeq \eta \left(\frac{T}{m_N}\right)^{3/2} e^{B_{^4He}/T} \simeq 4 \times 10^{-36} \tag{50}$$

To a good approximation the nucleon mass (m_N) may be taken to be the same for both protons and neutrons in these formulas. The respective nuclear binding energies, B_D and $B_{^4He}$, are $2.2MeV$ and $28MeV$ so the Boltzmann factors do not account for the small values of these number densities. The ratio η, ($\equiv n_b/n_\gamma$) can be determined from observations today since baryon number should be conserved between the $T \sim 10MeV$ era and the present. Its value is found to be about 10^{-10} and it is clearly this small number which is responsible for the low relative abundances of light nucleons at these temperatures.

A while later, when the temperature drops to around $1MeV$, the weak interaction rates become too slow to compete with the expanding universe and maintain equilibrium. The ratio of neutrons to protons "freezes out" at this time since the nucleons are no longer in equilibrium, and the neutrinos decouple, just as the photons did in the earlier discussion. If the corresponding "cosmic background neutrinos" could be detected we would have a probe, similar to the microwave photons, with which to look even deeper into our cosmic past.

Turning back to the nucleons, the relative numbers of protons and neutrons at neutrino decoupling is given by their mass difference (Δm):

$$\frac{n_N}{n_p} = e^{\Delta m/T_{freeze}} \simeq \frac{1}{6}. \tag{51}$$

Of course this ratio changes after neutrino decoupling because the neutrons can decay. The value of this ratio at later times is given by

$$\frac{n_n}{n_p} \simeq \frac{1}{6} e^{-t/\tau_n} \tag{52}$$

where τ_n is the neutron lifetime.

At a temperature of about $0.3 MeV$ equilibrium would prefer ^4He nuclei to free protons and neutrons, but the production rates are too slow for equilibrium to occur yet. The low densities of ^3He, D, and ^3H (traceable to the small value of η) which are intermediate states for ^4He production cause the production rates to be low.

Finally, as the temperature reaches $0.1 MeV$ the production rate becomes sufficient to bind essentially all the nucleons it can into ^4He. By this time free neutron decay has reduced the neutron to proton ratio to

$$\frac{n_n}{n_p} \sim \frac{1}{7} \tag{53}$$

so the mass fraction of ^4He winds up being about 1/4. Small amounts of ^7Li and ^7Be are also produced but the production rates are extremely low for heavier elements.

The nucleosynthesis calculations are sensitive to several factors. The larger η is, the greater the abundance of intermediate states and the earlier ^4He production begins in earnest. If there were additional contributions to ρ, the expansion rate would be faster at a given temperature (see Eqns (31) and (4)) so freeze out would occur at a higher temperature. Both these effects would increase the neutron to proton ratio at the time of Helium production and change the resulting abundances.

Still, the standard big bang model can give definite predictions of primordial nuclear abundances. One obstacle to testing these predictions is the difficulty of determining "observed" primordial abundances based on observations today. Various production and depletion mechanisms are expected to have operated between primordial production and now. There have been attempts to calculate all these effects, and there are established values for "observed" primordial abundances[14]. The value of the parameter η is also not known precisely enough, so the standard procedure is to use some of the data to fit η precisely and than compare the rest of the predictions. Amazingly enough, the predictions are quite successful.

There are many ways the standard nucleosynthesis calculations could be upset. Anisotropies in spacetime have been shown to have a noticeable effect on the results. Also, inhomogeneities in the nucleon distribution which might be introduced at the quark-hadron phase transition could dramatically affect the outcome[15,16]. Again, there could be important processes affecting nuclei at temperatures lower than $0.1MeV$ which are not included in standard calculations. Because of all these issues, primordial nuclear abundance predictions are not as brilliant a success of the standard big bang as the microwave background is. Still, they are a more modest success, and it interesting to test new proposals in particle physics and cosmology to see if the standard calculations are changed. A failure to preserve the standard results might bring the burden on the new model to explain nuclear abundances. Instead, one could point to the many weak areas in the standard calculations and hope that when those issues are understood better everything would work out.

4 Incorporating More Speculative Physics Into the Standard Big Bang

In the the previous section we discussed efforts to incorporate well established laboratory physics (e.g. atomic and nuclear physics) into the big bang model. The results produced interesting consequences which, at least to some extent, could be tested today. Now we turn to applications of less well established physics to the big bang.

4.1 Baryon Number Violation

Most models of grand unification have baryon number changing interactions. In these models the baryon number of the universe is no longer a free parameter which must be adjusted to fit current observations. Rather, it has dynamics and its evolution must be understood. Simple thermal equilibrium arguments would tell us that at temperatures where baryon number violation occurs easily, interactions would maintain the baryons and anti-baryons in equal numbers, leaving the net baryon number equal to zero. Fortunately, it has been shown that as temperatures fall below the scale where baryon number violations occur readily, *non*-equilibrium processes can occur which can leave a residual net baryon number. This is an example of free parameter of an old model becoming a derived parameter in a new one. The details of these calculations will be discussed in Hall's lecture. I wish to stress here, however, that once one introduces baryon non-conservation into high energy

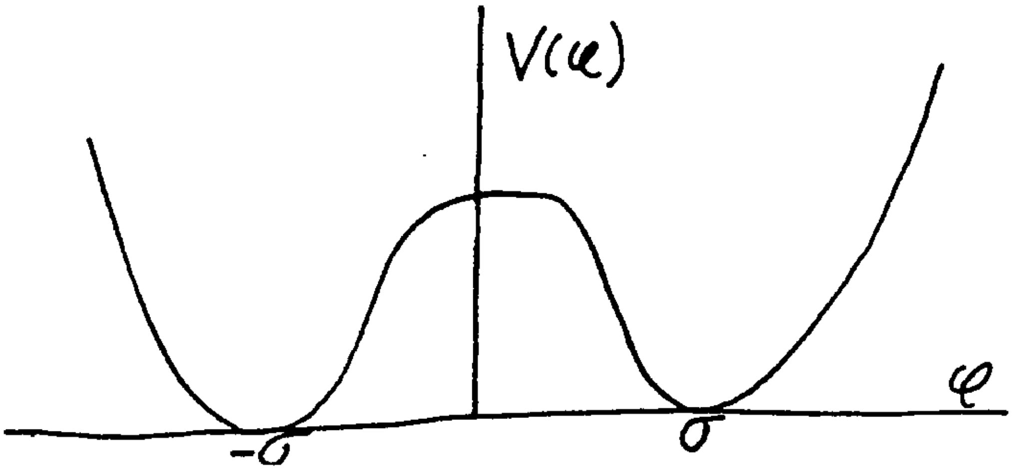

Figure 3: The double well potential

physics, one shoulders the responsibility of correctly accounting for the net baryon number observed today. If the number comes out wrong, one would have grounds for rejection of that particular model.

4.2 Strange Objects

Field theory can do an excellent job of describing particles interacting in a laboratory. It is then quite conventional to assume that it can equally well describe particles interacting in the early universe. However, field theories contain states which do not look at all like particles. Many such states have manifested themselves when field theories are applied to condensed matter physics, but in the realm of high energy physics none are clearly observed to play a role in laboratory experiments. Still, if one takes field theory seriously in its entirety, these states could be there and one would expect them to play a very significant role in the early universe.

4.2.1 Domain Walls

Consider a single real field ϕ with a potential like the one in Fig [3]. The potential has two degenerate minima at $\phi = \pm\sigma$. Particles of the ϕ field correspond to small oscillations in the field around one of the minima. However, this model also contains states which do not look like particles. Let us consider boundary conditions which are plane symmetric in the $x - y$ plane and let us require $\phi(z = -\infty) = -\sigma$ and $\phi(z = \infty) = \sigma$. Somewhere in between $z = \pm\infty$, ϕ must cross over from one minimum to the other. In doing so the configuration must acquire some non-zero energy. In fact, the minimum energy configuration consistent with these boundary conditions contains a sheet-like region in the $x - y$ plane with an energy per unit area of roughly σ^3. and a width of about $1/\sigma$. Such a field configuration is called a domain wall. A single infinite domain wall is stable, in the sense that it takes an infinite amount of energy to stimulate its decay. In particular, one must move the field in half of space up over the barrier between the two minima in order to get rid of a domain wall (this takes an infinite amount of energy for every *finite* section of domain wall which is removed). Of course if a domain wall encloses a finite size region rather than extending to infinity, that wall can disappear, because the inside region can shrink to zero size, leaving the field near only one minimum in all of space. Still, if the interior region is large enough a finite domain wall can play an interesting dynamical role before decaying.

It takes an infinite amount of energy to get rid of an infinite domain wall, and likewise, one must move the field from one minimum to the other in half of space to create one. That procedure would also use an infinite amount of energy. However, when it comes to cosmology there is always an infinite amount of energy available because there is a finite energy density in a space of infinite extent. (Of course in a closed universe the volume of space is finite, but then so is the energy needed to create a domain wall.) Since the energy density diverges as one works back toward $t = 0$ in the big bang, there must be a time were the energy density is large compared with the height of the barrier separating the two minima. Under these circumstances domain walls could be freely created and destroyed. In fact, at high enough energy densities it would be hard to distinguish domain walls as such because there would be many field excitations with shorter wavelength than the thickness of a domain wall.

The cosmic production of domain walls is usually discussed in terms of the "Kibble mechanism"[17]: Initially the energy density in ϕ is much greater than the barrier separating the two minima, but as the universe expands and cools, the fields evolve closer and closer to a minimum energy configuration. In the case where there are more than one minimum energy states available (such as the current example), different regions (or "domains") of the universe would approach different choices of minimum energy state. The exact size of these domains is hard to calculate but they are not expected to be larger than the causal horizon size at the time of domain formation. (One would expect some causal contact would be necessary to communicate which minimum a given region is approaching.) On boundaries between domains, domain walls necessarily form. Thus the causal structure of the big bang guarantees that if a field with a potential such as the one in Fig [3] exists, domain walls will form as the universe evolves. Domain walls can be thought of as "two dimensional defects" in the field configuration. In some sense the field would "rather" approach one minimum everywhere but the causal dynamics do not allow this to happen.

4.2.2 Nielsen-Olesen Strings

Nielsen-Olesen strings[18] are the one dimensional analogue of the (two dimensional) domain walls. Again, the degeneracy of minima of the potential allows defects to form as the universe cools, but the nature of the defects is line-like rather than plane-like. A simple example is given by a complex scalar field with the "Mexican hat" potential shown in Fig [4] In this case there are not just two, but a continuum of degenerate minima, corresponding to the circle at the bottom of the Mexican hat ($|\phi| = \sigma$). An infinite straight Nielsen-Olesen string can be constructed by imposing the right cylindrically symmetric boundary conditions: The field at large distances from the axis of symmetry must be in a minimum of the potential, and the choice of which minium varies smoothly around the bottom of the Mexican hat as one circles around the axis of symmetry. Any continuous field configuration with these boundary conditions must have some region where the field is at the top of the potential. The minimum energy configuration occurs when this region runs straight down the axis of symmetry, like a string.

The model just described is a model with a broken global $U(1)$ symmetry. It is well known that such a model has a massless particle in the spectrum called a Goldstone boson (corresponding to excitations of the field in direc-

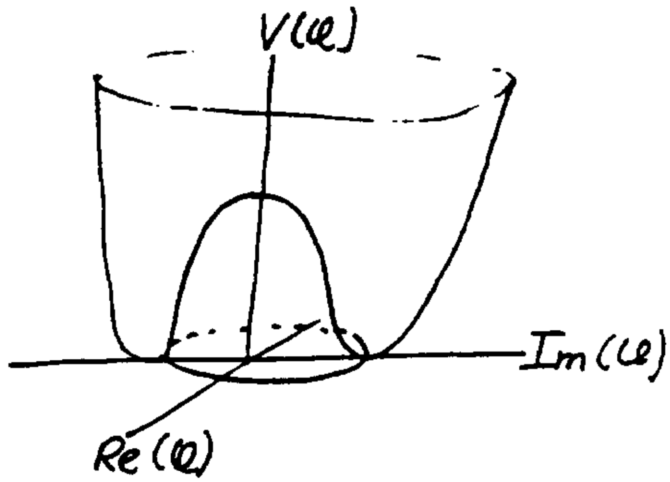

Figure 4: The Mexican hat potential. Note that all points on the circle at the bottom of the hat are degenerate absolute minima.

tions where the potential does not increase, *around* the Mexican hat). This massless particle produces a long range force field around the string, and the mass per unit length and radius of the string are actually infinite. If the $U(1)$ symmetry is gauged, there are no massless particles and the string has mass per unit length $\sim \sigma^2$ and width $\sim 1/\sigma$. As with domain walls, an infinite amount of energy is required to get rid of an infinite string, and strings are stable in that sense. Finite closed loops of string are unstable because they could collapse to a point and annihilate, but in practice a large loop of string might have a long lifetime.

The key to the presence of Nielsen-Olesen strings in a field theory is a suitable degeneracy of the manifold of ground states ($\equiv \mathcal{M}$). There must be a non-trivial mapping from $\mathcal{M}$ to the physical space around the string (represented by the circle). The equivalent statement in group theory is $\Pi_1(\mathcal{M}) \neq 0$. Any times these conditions exist one can expect strings to arise via the Kibble mechanism, the same way domain walls can. One expects at least of order one piece of string per causal horizon size to appear when the energy density drops below that of the barrier in the potential.

4.2.3 Monopoles

Monopoles are the "point-like" analogue of the strings and walls. The potential for a simple model is hard to draw, but again the key is non-trivial mapping from the manifold of ground states to the physical space around the monopole, represented by the surface of a sphere (E.g. $\Pi_2(\mathcal{M}) \neq 0$). One expects monopoles to be produced by the Kibble mechanism in the early universe if this condition is met. The size and mass of the monopole are characterized by the mass scale in the potential in the usual way.

One thing that makes monopoles particularly interesting is that essentially all theories of grand unification have them. That is because we know that the $U(1)$ symmetry of electromagnetism is a good symmetry of the world today, and there turns out to be no way of breaking a simple unification symmetry, leaving an unbroken $U(1)$, without producing a non-trivial $\Pi_2(\mathcal{M})$. These monopoles have a net magnetic charge (thus their name). Domain walls and Nielsen-Olesen strings appear in many models of grand unification but certainly not in all. No monopoles, strings, or walls appear in the standard model of electroweak interactions.

Because at a distance magnetic monopoles look like point particles, their evolution is easy to calculate. What follows is a very rough calculation of the contribution of magnetic monopoles to the energy density. All relevant mass scales are assumed of order M_G, the grand unification mass scale, and coupling constants are assumed $O(1)$. Assuming one monopole per causal horizon gives the energy density in monopoles (ρ_m) at the time of formation t_m to be

$$\rho_m \sim (\frac{M_G}{t_m^3}). \tag{54}$$

Using Eqns (4) and (38) and neglecting curvature gives

$$\frac{1}{t_m} \sim H = \sqrt{\frac{8\pi}{3} G\rho_{tot}} \sim \sqrt{1000 G T_m^4} \simeq \sqrt{1000 G M_G^4} \tag{55}$$

where Eqn (31) was used to introduce T_m, the temperature at which the monopoles form. Thus

$$\left.\frac{\rho_m}{\rho_{tot}}\right|_{t_m} \sim 100 \left(\frac{m_g}{m_P}\right) \sim 10^{-10}. \tag{56}$$

After their production the monopoles are essentially non-relativistic ($\rho_m \sim a^{-3}$) while most of the other matter is relativistic ($\rho_{tot} \sim a^{-4}$) . This means

$$\frac{\rho_m}{\rho_{tot}} \approx 10^{-10} \left(\frac{a}{a_m}.\right) \quad (57)$$

Because the energy density in non-relativistic matter decreases more slowly than that of relativistic matter with the expansion of the universe, the monopoles can eventually come to dominate. Equation (57) indicates that monopole domination would occur at $T \approx 10^4 GeV$ (using Eqn (44)). Not only would such monopole domination disrupt the big bang model at the eras of nucleosynthesis and photon decoupling, it would completely contradict present observations since at most one magnetic monopole has been observed! More sophisticated calculations which include monopole-anti-monopole annihilation fail to alter this result in a qualitative way[19]. This is a very serious problem with grand unification since all models of grand unification have magnetic monopoles.

I should remark that problems such as the one just discussed can occur if any stable heavy particles are produced in sufficient numbers in the early universe. A particle need not be as exotic as a magnetic monopole to cause problems, and every model of high energy physics must be checked for such problems. Of course, one must wonder if there can be domain wall and string problems similar to the monopole problem. I shall briefly return to the issue of strings in section 6.

In a recent paper[20] a dynamical process was presented which would cause monopoles to annihilate with anti-monopoles fast enough to avoid the monopole problem. The authors are convinced, however, that their proposed dynamics, although causal, is not a realistic dynamics for the field theory in question. Thus the monopole problem and the Kibble mechanism appear to remain intact.

4.3 Potential Dominated States

We have discussed states in field theory where the fields deviate far from the fluctuations around potential minima which correspond to traditional particles. In fact, the field is actually at a local maximum of the potential at a point (for monopoles), on a line (for Nielsen-Olesen strings), or on a plane (for domain walls). It is interesting to study what would happen if a field were at a maximum of the potential over some large three dimensional volume in space. Furthermore, let us assume that the potential energy density (V) dominates over energy density due to gradients and time variation. There is nothing that makes such a state really stable the way monopoles etc. are, but it may have a non-trivial lifetime. Later on I will discuss how such a state might come about. For now I will focus on the cosmological consequences of a potential dominated state, should it come about. We shall see that these consequences are very interesting and provide motivation for finding out if such states really can arise.

It is easy to check the effect of a potential dominated state on the evolution of the universe by writing down the stress-energy of a field neglecting non-potential terms:

$$T^{\mu}_{\nu} = \begin{pmatrix} V & 0 & 0 & 0 \\ 0 & -V & 0 & 0 \\ 0 & 0 & -V & 0 \\ 0 & 0 & 0 & -V \end{pmatrix}. \tag{58}$$

Comparing Eqn (58) with the standard Robertson-Walker form for the stress-energy (Eqn (3)) gives

$$p = -\rho = -V. \tag{59}$$

Equation (59) represents a completely different equation of state compared with the two used in the standard big bang model. The resulting evolution of spacetime is very different as well. Applying the conservation of stress-energy (Eqn (5)) to a potential dominated state gives

$$\frac{d}{da}\left(\rho a^3\right) = 3\rho a^2 = \rho \frac{d}{da}\left(a^3\right) \tag{60}$$

which implies

$$\frac{d}{da}\rho = 0. \tag{61}$$

For a potential dominated state the energy density does not drop as the universe expands. The Friedmann equation is

$$\frac{\dot{a}}{a} = \sqrt{\frac{8\pi}{3}GV} \equiv H_V \tag{62}$$

so the expansion rate is constant and

$$a \sim e^{H_V t}. \tag{63}$$

Here we have the universe expanding exponentially, as opposed to the power law evolution encountered in the standard big bang.

Let us modify the standard big bang model in a simple way in order to introduce a period where the equation of state of matter is the potential dominated equation of state. Suppose that early in the relativistic era the matter became potential dominated at a time t_1 and then returned to the relativistic equation of state at a time t_2. The evolution of the scale factor and ρ would look like this:

$t > t_1$	$a \sim t^{1/2}$	$\rho = D_1/a^4$
$t_1 > t > t_2$	$a \sim e^{H_V t}$	$\rho = const$
$t_2 > t > t_*$	$a \sim t^{1/2}$	$\rho = D_2/a^4$
$t_* > t$	$a \sim t^{2/3}$	$\rho = B/a^3$

. (64)

The integration constants D_1 and D_2 (for the two different relativistic eras) can be related by using Eqn (61) to equate $\rho(t_1)$ with $\rho(t_2)$:

$$D_2 = D_1 e^{4H_V(t_2 - t_1)}. \tag{65}$$

What is striking about this expression is that D_2 can be very large even if D_1 is not. In this model it is D_2 which is known to be large based on the discussion in section 2.4. If $H_V(t_2 - t_1) \gtrsim 66$ then one can have $D_2 \geq 10^{114}$ with D_1 only $O(1)$. Thus this modification of the big bang appears to offer an explanation for the large value of D discussed earlier.

Models in which a potential dominated period occurs are called "inflationary" cosmologies, and when $a \sim e^{H_V t}$ the universe is said to be "inflating". The first mention of this sort of cosmology in the literature is reference [21]. Inflationary cosmologies are interesting for several reasons and the next section is devoted to discussing them.

5 Inflation

In section 4.3 we saw that introducing an inflationary period into the standard big bang model allowed the integration constant D to be large after inflation even if it is chosen $O(1)$ before. The spatial curvature is exponentially reduced as the universe expands during inflation, while the energy density remains constant. Thus, the relative importance of energy density and curvature in the Friedmann equation is greatly shifted away from the curvature, and the universe appears to be much more flat than before inflation. This fact offers hope that the large value of D (and the associated puzzles) might be explained in inflationary cosmologies. I will return to this issue after mentioning some other attractive aspects of inflationary cosmologies.

The field whose stress-energy becomes potential dominated during inflation is called the inflaton. During the inflationary period the energy density of the inflaton holds constant, while relativistic, and non-relativistic matter have their energy densities reduced by $e^{-4H_V(t_2-t_1)}$ and $e^{-3H_V(t_2-t_1)}$ respectively. At the end of inflation the inflaton potential energy is released into radiation composed of inflatons and whatever else the inflatons couple to. If the amount of expansion is as large (of order e^{66}) as would be needed to give a large enough D, then essentially all the matter in the universe is inflatons and their decay products.

Thus, inflation offers a way out of the monopole problem. If monopoles are produced before the inflationary period, and temperatures after inflation are not high enough to produce more of them, then monopoles would be present only in negligible numbers. Grand unified models which allow for an appropriate period of inflation can avoid the monopole disaster. This aspect of inflation is a useful tool, and it can be used to get rid of other troublesome objects should they arise in high energy models. The other side of this issue is that things that you *do* want must be produced by inflaton decay. If one's model has baryon number violation, one's baryon number generating scheme must operate after inflation has occurred.

Inflation also drastically alters the causality structure of the big bang. We saw in section 3.2 that causal processes could not account for the observed isotropy of the cosmic microwave radiation in the standard big bang. With inflation the story is very different. Regions that are in causal contact before inflation are increased in size by exponential factors. In inflationary models the entire universe fits inside one causally connected region. I should stress that this fact alone does not guarantee sufficient isotropy. What the causal forces actually do to isotropy is a separate issue. (Of course, the causal regions around events after the inflationary period are the same as in the standard big bang.)

There is yet another important aspect of inflationary cosmologies. As the inflationary period proceeds all field excitations get red-shifted away. Eventually the most significant field fluctuations are the inherent quantum fluctuations of fields in an exponentially expanding universe. It is then only these fluctuations which provide deviations from a perfectly isotropic and homogeneous distribution of matter. The inflationary period essentially wipes the slate clear of primordial fluctuations and introduces its own, calculable fluctuations in their place. This process has been investigated in some detail, and there is hope that the spectrum of fluctuations that emerge from inflation might be just what is needed to explain the formation of galaxies and other structure in the universe.

Let us return to the issue of large D. Clearly, inflation does not really elevate D to the status of a derived parameter. The value after inflation is simply much larger than the initial value. In an inflationary cosmology, the need to explain the large value of D, or why the universe is "flat for its size", is reduced to the need to explain why D_1 is $O(1)$ initially. In fact, what happens after inflation depends so little on what happened before it, that all one really cares about is that an inflationary period is entered. The inflationary era itself can then produce a universe very much like that described by the standard big bang.

Many non-traditional cosmic entrances to inflation have been proposed, but it is impossible to arrange things so that *any* early cosmology will enter an inflationary period. As we shall see, the initial cosmology does have a lot to do with whether an inflationary era is entered. An inflationary period acts as a funnel which evolves many diverse cosmologies into the (large D) standard big bang bottle, but there are still many others which spill over the edge. Even so, this property, along with the others I have just discussed, makes inflation a powerful tool for cosmologists searching for a deeper understanding of our universe.

5.1 Getting Inflation to Happen

The above discussion was based on the assumption that it is possible for the matter to enter a potential dominated state. Furthermore, it was assumed that this state would last long enough that $H_V(t_2 - t_1) \geq O(100)$. I now turn to the issue of how this may come about. I will only be able to offer a summary here, I direct you to the references for details.

At first glance, it may not seem like much to ask the potential dominated state to last $O(100)$ expansion times, but really it is. This fact can be seen by means of a simple dimensional argument. I assume that the microphysics which causes an inflationary state to be entered is characterized by a mass scale m_I, and that all couplings are $O(1)$. One would expect that V, the energy density during inflation would be $O(m_I^4)$, and $(t_2 - t_1)$, the lifetime of the inflationary state would be $O(m_I^{-1})$. One then finds (using Eqn(62)) that

$$H_V(t_2 - t_1) \approx \frac{m_I}{m_P}. \tag{66}$$

As long as $m_I < m_P$, our estimate says that inflation occurs for only a fraction of an expansion time. The number we want to be $O(100)$ comes out to be the *small* number m_I/m_P. The details of the inflationary period must therefore be sufficiently complex to invalidate this simple dimensional argument.

5.1.1 Old inflation

When Alan Guth introduced his "old" inflationary model[22] he had no trouble making inflation occur for many expansion times. In his model the inflaton had a potential such as the one depicted in Fig [5]. The key feature is a "false" minimum, separated by a barrier from the lower "true" minima. Guth argued that as the universe cooled from high temperatures the inflaton would become trapped in the false minimum. The expansion of the universe would then redshift away all contributions to the stress energy except that of the potential, leaving a potential dominated state. The decay of the inflaton into the true minimum is a tunneling process with the usual exponentially suppressed rate. In this scheme $t_2 - t_1$ is $m_I^{-1}e^{O(100)}$ so the inflationary period is plenty long enough.

The main problem with old inflation has to do with the nature of the decay process[23]. A uniform field in a false minimum will decay via bubble formation: A finite region will decay through the barrier and approach the true minimum. Unfortunately the size of such a region is typically around m_I^{-1}, and it has been shown that the decay process never really is complete

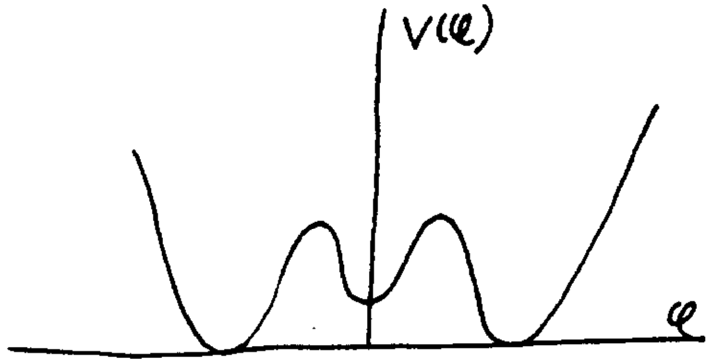

Figure 5: The form of the potential for old inflation. There is a local minimum which is distinct from the global minimum.

over all space. Furthermore, the regions that do not remain inflating are dominated by large bubbles with all their energy concentrated in the bubble walls. Essentially nowhere is there a region which could evolve into the universe we observe today. Old inflation has no trouble getting inflation for a long enough period, the problem lies in getting it to stop.

I should remark that old inflation actually represents a possible pitfall for any model of high energy physics. All models should be checked for potentials of the form shown in Fig [5]. The presence of such a potential could cause the universe to inadvertently enter an era of old inflation, and cause the model to have an unacceptable cosmology.

5.1.2 New inflation

Another form of inflation, called "new inflation"[24,25], offers a more graceful exit from an inflationary period. In new inflation the inflaton has a potential similar to the one in Fig [3]. The inflationary state is one in which the inflaton is everywhere balanced at the top of the potential hump. The decay of the inflationary state occurs as the field "rolls off" the local maximum.

This decay process is not limited to finite sized regions the way the bubble formation was in old inflation. Although there are still "bubble walls" between regions which roll in different directions, inflation continues even as the rolling starts, so such walls are pushed very far apart.

If all dimensionless couplings are $O(1)$ the "rollover" time is $O(m_I^{-1})$ and the dimensional argument presented above applies. The inflationary period is only made long enough by adjusting parameters in the potential to increase the rollover time.

Considerable effort has gone into understanding the spectrum of deviations from pure homogeneity and isotropy which emerge from quantum fluctuations in a new inflationary cosmology[26,27,28,29,30] . The resulting spectrum of perturbations has been called "almost scale invariant". If all couplings are $O(1)$ the magnitude of the perturbations comes out too large to be consistent with bounds on the isotropy of the microwave background. (Remember, causal contact is not enough to guarantee isotropy.) The couplings of some models can be adjusted to accommodate this problem, and typically a dimensionless coupling must be set $O(10^{-11})$ to reduce the perturbations to acceptable levels. It this adjustment is the most extreme needed, and once it is made the duration of inflation easily comes out long enough.

In order to maintain the necessary small coupling the effective interactions between the inflaton and the other fields must all be very small. This fact makes it a challenge to couple energy out of the inflaton and into all the familiar forms of matter, once the inflationary period is over. None the less, models have been constructed in which the familiar forms of matter do eventually get sufficiently excited after inflation is over.

I have yet to address the issue of how the inflaton can get balanced on top of the potential hump at the beginning of inflation. Initially it was hoped that thermal effects could play a significant role, but the small couplings of the inflaton preclude thermal equilibrium before inflation as a reasonable possibility. Some current thinking on this subject centers on the flat nature of inflaton potentials. The adjustments one makes to produce slow rollover and small quantum perturbations cause the inflaton potential to be very flat near the local maximum. If fields are near the local maximum the potential produces only a very slight force toward a potential minimum, while the expanding universe is constantly redshifting away the gradient and time derivative contributions to the stress-energy. The redshifting process can proceed faster than the slow slippage toward a minimum and cause a potential dominated state to be entered. Numerical work[31] indicates that this could happen even with a rather large amount of initial space and time variations present in the inflaton field. With a fairly wide range of initial conditions the adjustments to the inflaton potential required to avoid excessive quantum fluctuations were sufficient to allow inflation to start. On the other hand, plenty of perfectly reasonable initial conditions do not inflate in these models.

I feel that despite the existence of models that "work", our understanding of the onset of inflation leaves a lot to be desired. No convincing examples exist of models which very generally enter an inflationary period. The interested student should study the literature with a critical eye and an open mind. The field is in need of some fresh new thinking on this difficult problem.

5.1.3 Other ideas

Some interesting alternatives to standard new inflation have already been proposed. Generally, as indicated by the dimensional argument presented earlier, things grow easier as the m_I approaches m_P. Of course the closer one gets to the Planck scale the more nervous one gets about one's understanding of gravity.

In his "chaotic inflation" scenario[32] Linde abandons the local maximum of the potential in favor of a more general region of a potential. Linde shows that large ϕ fluctuations at the Planck era for an inflaton with a simple ϕ^4 potential can end up inflating for a sizable period. Due to the proximity of the Planck scale no small couplings are needed for sufficient inflation, but the usual adjustments must be made to prevent excessive anisotropies. In addition, the initial fluctuation must be sufficiently uniform in order for inflation to start. The exact uniformity requirements are thought by some to arise "naturally", while others are not convinced.

Another interesting idea involves depicting the birth of the universe as a quantum gravity fluctuation, and showing that some initial states appear to naturally fluctuate into an inflating spacetime in simple models[33]. Other proposals suggest that quantum corrections to Einstein's equations could cause an inflationary period near the Planck era[34]. All these proposals would look better if gravity were really understood at the quantum level.

5.2 Current Status of Inflation

From a cosmologist's point of view inflation appears to be a very powerful tool. It can take a variety of initial cosmologies and evolve them to a cosmology consistent with our observed universe. It can produce large values of the integration constant D, get rid of monopoles and other junk, improve the causality structure, and inflation can introduce its own spectrum of energy density perturbations in place of whatever there was before. From the standpoint of microscopic physics inflation is not easy to come by. For some reviews see references [35], [36] and [37]. All models in which inflation occurs seem very contrived. At this point it is fair to question what one achieves by gaining the benefits of inflation at the expense of contrived microphysics. Fortunately for inflation, our understanding of physics at the appropriate energy scales is very poor. One can hope that, should a better understanding emerge, inflation will appear to fit more naturally into the picture.

I should remark that a typical model in which inflation occurs does not have "just enough" inflation, but considerably more. The scale factor increases by much more than $e^{O(100)}$ so $D \gg 10^{114}$ and the universe is extremely flat. If it is ever determined that the curvature is non-negligible today, all inflationary models except those with just enough inflation would be ruled out.

It is important to emphasize that a discussion of inflation touches on some of the deepest unsolved problems in physics. At present we have only a classical theory of gravity, whereas we describe matter quantum mechanically. How this quantum theory produces a c-number stress-energy to go into Einstein's equations is not absolutely clear. In particular, the "cosmological constant" term ($= \Lambda g_{\mu\nu}$) in the stress energy is set very small today to match observations, but no one understands the physics behind it. The cosmological constant is intimately connected with inflation since it can cancel or enhance the effects of the potential ($= V g_{\mu\nu}$) in the stress-energy. It is possible that a deeper understanding of the cosmological constant could radically change our understanding of inflation. [38,39]

The process by which quantum fluctuations can emerge as classical objects is also not understood, yet this process plays a central role in the calculation of energy density fluctuations produced by inflation. As Murray Gell-Mann likes to say, the universe is full of Schrödinger's cats. The nature of the perturbations which emerge from inflation will not be fully comprehended until the quantum to classical transition is understood.

Inflation is closely tied up with the question "what are the initial conditions of the universe?". At this point we do not even know if the question makes sense, let alone has a clear answer. We will never know what role inflation may have played in the early universe unless progress is made on this question, or the onset of inflation is made sufficiently independent of what went before.

Inflation has a very clear appeal from a cosmologist's perspective. There are many outstanding issues regarding its implementation, but these all lie in areas of physics which are incomplete in their own right. Inflation's promise for cosmology makes it well worth pursuing these unresolved issues vigorously, and makes any advances on these fronts all the more rewarding.

6 Cosmic Strings

Unfortunately, I do not have time in these lectures to discuss the topic I am most actively investigating. I will just give a plug for it here, and refer you to the literature for more details. The evolution of Nielsen-Olesen strings after their formation in the early universe is now being actively investigated[40,41]. There is hope that regardless of the details of the initial network, with time it will approach a so called "scaling solution"[42,43,44]. Such a string network would not cause problems the way the monopoles did, yet it might have interesting effects on the evolution of the universe. Loops which break off a network of infinite string could become the seeds of gravitational collapse. The work is still in progress, but there is some indication that this process could account for much of the structure (e.g galaxies and clusters) that we now observe in the universe[45,46,47,48]. This scheme is particularly appealing because the nature of perturbations made by the string network should depend very little on what happened in the very early universe. It has also been suggested that radiation from currents on superconducting cosmic strings could cause explosions which might be resposible for the formation of voids and filaments[49].

7 Conclusions

In these lectures I have shown some of the interesting issues which arise at the frontier between high energy physics and cosmology. I have argued that all high energy physicists must take an interest in cosmology in order to fully test their theories. I also pointed to opportunities which exist for high energy physicists to make interesting contributions to cosmology. We have seen how gross features of our universe such as its size, the relative absence of monopoles, and the microwave background, result in important tests and constraints for cosmology. We live in a time when more detailed observations are coming in at a rapid rate, and new challenges for cosmology will no doubt arise. I feel that the interface between cosmology and high energy physics has a very exciting future, and I hope you all will participate in one way or another.

8 Acknowledgments

I thank the students at TASI for their stimulating questions which made giving these lectures worthwhile. I would also like to thank N. Turok and D. Haws for their comments on the manuscript. This work was supported in part by the DOE and by the NASA at Fermilab.

References

[1] M. Turner. Cosmology and particle physics. In *Les Houches, Session XLIV, 1985 – Architecture of fundamental interactions at short distances*, Elsevier Science Publishers, 1987.

[2] E. W. Kolb. Particle physics and cosmology. In *Proceedings of the 1986 Theoretical Advanced Studies Institute*, World Scientific, 1986.

[3] S. Weinberg. *Gravitation and Cosmology.* Wiley, New York, 1972.

[4] S. W. Hawking and G.F.R. Ellis. *The large-scale structure of space-time.* Cambridge University Press, 1973.

[5] G. A. Tammannn. The value of the hubble constant. In M. P. Ulmer, editor, *13th Texas Synposium on Relativistic Astrophysics*, World Scientific, 1987.

[6] S. Faber and J. Gallagher. *Ann. Rev. Astron. Astrophys*, 17:135, 1979.

[7] See reference [1], section 5.2.

[8] W. A. Fowler. *Quarterly Journal of the Roy. Astron. Soc.*, 28:87, 1987.

[9] See reference [1], section 1.

[10] N. Turok. Private communication.

[11] R. Matzner. *Ap. J.*, 241:851, 1980.

[12] D. Layser and R. Hively. *Ap. J.*, 179:361, 1973.

[13] A. M. Boesgaard and G. Steigman. *Ann. Rev. Astron. Astrophys.*, 23:319, 1979.

[14] Part viii, big-bang nucleosynthesis abundances. In M. P. Ulmer, editor, *13th Texas Synposium on Relativistic Astrophysics*, World Scientific, 1987.

[15] G. J. Mathews G. M. Fuller and C. R. Alcock. In G. J. Mathews, editor, *The Origin and Distribution of the Elements*, 1987.

[16] T. Rothman R. Matzner, J. M. Centrella. In G. J. Mathews, editor, *The Origin and Distribution of the Elements*, 1987.

[17] T. W. B. Kibble. *J. Phys*, A9:1387, 1986.

[18] H.B. Nielsen and P. Olesen. *Nucl. Phys.*, B61:45, 1973.

[19] J. P. Preskill. *Phys. Rev. Lett.*, 43:1365, 1979.

[20] T. W. B. Kibble D. Mitchell N. Turok E. Copeland, D. Haws. Monopoles connected by string and the monopole problem. Nucl. Phys B, in press.

[21] F. Englert R. Brout and E. Gunzig. *Proc. Roy. Soc.*, A360:117, 1978.

[22] A. H. Guth. *Phys. Rev. D*, 23:347, 1981.

[23] A. Guth and E. Weinberg. *Nucl. Phys.*, B212:321, 1983.

[24] A. D. Linde. *Phys. Lett.*, 108B:389, 1982.

[25] A. Albrecht and P. Steinhardt. *Phys. Rev. Lett.*, 48:1220, 1982.

[26] A. Guth and S.-Y. Pi. *Phys. Rev. Lett.*, 49:1110, 1982.

[27] S. Hawking. *Phys. Lett.*, 115B:295, 1982.

[28] A. Starobinski. *Phys. Lett.*, 117B:175, 1982.

[29] P. Steinhardt J. Bardeen and M. Turner. *Phys. Rev. D*, 28:679, 1983.

[30] W. Ratara W. Fischler and L. Susskind. *Nucl. Phys.*, B259:730, 1985.

[31] R. Brandenberger A. Albrecht and R. Matzner. *Phys. Rev. D*, 35:429, 1987.

[32] A. Linde. *Phys. Lett.*, 129B:177, 183.

[33] J. B. Hartle. Initial conditions and quantum cosmology. In M. P. Ulmer, editor, *13th Texas Synposium on Relativistic Astrophysics*, page 80, World Scientific, 1987. Thi is a review with further references.

[34] N.-M. Suen M. B. Mijic, M. S. Morris. *Phys. Rev. D*, 34:2934, 1986.

[35] A. Linde. *Physics Today*, September:61, 1987.

[36] A. Linde. *Rep. Prog. Phys.*, 47:925, 1984.

[37] R. Brandenberger. *Rev. Mod. Phys.*, 57:1, 1985.

[38] L. H. Ford. Stability of de sitter space and the cosmological constant problem. In M. P. Ulmer, editor, *13th Texas Synposium on Relativistic Astrophysics*, page 97, World Scientific, 1987. A discussion of some quantum aspects of the cosmological constant and further references may be found here.

[39] T. Banks. *Nucl. Phys.*, B249:332, 1985. This paper contains an interesting idea regarding the cosmological constant.

[40] A. Albrecht and N. Turok. Evolution of cosmic strings. *Phys. Rev. Lett.*, 54:1868, 1985.

[41] D. Bennett and F. Bouchet. Simulations of cosmic strings. 1987. To be submitted to Phys. Rev. Lett.

[42] T. W. B. Kibble. *Phys. Rep.*, 67:183, 1980.

[43] D. Bennett. *Phys. Rev. D*, 33:872, 1986.

[44] D. Bennett. *Phys. Rev. D*, 34:3592, 1986.

[45] Ya. B. Zel'dovich. *Mon. Not. Roy. Astron. Soc.*, 192:663, 1980.

[46] Neil Turok. *Phys. Rev. Lett.*, 55:1801, 1985.

[47] R. Brandenberger and N. Turok. *Phys. Rev. D*, 33:2182, 1986.

[48] N. Turok. Two lectures on the cosmic string theory of galaxy formation. Imperial College preprint Imperial/TP/86-87/23.

[49] J. Ostriker, C. Thompson, and E. Witten. *Phys. Lett.*, B180:231, 1986.