FERMILAB-Conf-86/128-A
September 1986

COSMOLOGICAL PHASE TRANSITIONS

Edward W. Kolb

Theoretical Astrophysics
Fermi National Accelerator Laboratory
Batavia, Illinois 60510 USA

1. THE EVOLUTION OF THE VACUUM

If the universe started from conditions of high temperature and density, there should have been a series of phase transitions associated with spontaneous symmetry breaking. The cosmological phase transitions could have observable consequences in the present Universe. Some of the consequences including the formation of topological defects and cosmological inflation are reviewed here.

One of the most important tools in building particle physics models is the use of spontaneous symmetry breaking (SSB). The proposal that there are underlying symmetries of nature that are not manifest in the vacuum is a crucial link in the unification of forces. Of particular interest for cosmology is the expectation that at the high temperatures of the big bang symmetries broken today will be restored, and that there are phase transitions to the broken state. The possibility that topological defects will be produced in the transition is the subject of this section. The possibility that the Universe will undergo inflation in a phase transition will be the subject of the next section.

Before discussing the creation of topological defects in the phase transition, some general aspects of high-temperature restoration of symmetry and the development of the phase transition will be reviewed.

1.1 *High Temperature Symmetry Restoration*

To study temperature effects, consider a real scalar field described by the Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - V(\phi); \quad V(\phi) = -\frac{1}{2}M^2\phi^2 + \frac{1}{4}\lambda\phi^4. \tag{1.1}$$

The minima of the potential (determined by the condition $\partial V/\partial\phi = 0$), and the value of the potential at the minima are given by

$$\langle\phi\rangle = \pm\sqrt{\frac{M^2}{\lambda}}; \quad V(\langle\phi\rangle) = -\frac{M^4}{4\lambda}. \tag{1.2}$$

Presumably, the ground state of the system is either $+\langle\phi\rangle$ or $-\langle\phi\rangle$ and the reflection symmetry $\phi \leftrightarrow -\phi$ present in the Lagrangian is not respected by the vacuum state. When a symmetry of the Lagrangian is not respected by the vacuum, the symmetry is said to be spontaneously broken.

From the stress tensor in terms of the Lagrangian, $T_{\mu\nu} = -\partial_\mu\phi\partial_\nu\phi - \mathcal{L}g_{\mu\nu}$, the energy density of the vacuum is

$$\langle T_{00}\rangle = \rho_V = -\mathcal{L} = V(\phi) = -\frac{M^4}{4\lambda}. \tag{1.3}$$

The contribution of the vacuum energy to the total energy density today must be smaller than the critical density $\rho_C = 1.88 \times 10^{-29}h^2$ g cm^{-3} $\simeq 10^{-46}$ GeV^4. Since this number is so small, it is tempting to require $\rho_V = 0$. This can be accomplished by adding to the Lagrangian a constant factor of $+M^4/4\lambda$. This constant term will not affect the equations of motion, and the sole effect is to cancel the present vacuum energy.

There are several ways to understand the phenomena of high-temperature symmetry restoration. The most physical way is to express the effective finite-temperature mass of ϕ as the zero-temperature mass, $-M^2$, and a plasma mass, $M_{plasma} \simeq a\lambda T^2$, where a is a constant of order unity. If $M_T^2 = -M^2 + M_{plasma}^2 \leq 0$, the minimum of the potential will be at $\phi \neq 0$ (SSB), while if $M_T^2 = -M^2 + M_{plasma}^2 \geq 0$, the effective mass term will be positive and the minimum of the potential will be at $\phi = 0$ (symmetry restored). There is a critical temperature, $T_c = M/(a\lambda)^{1/2}$ above which $\langle\phi\rangle = 0$ [1].

A more rigorous approach to symmetry restoration is to account for the effect of the ambient background gas in the calculation of the higher-order quantum corrections to the classical potential. The finite temperature potential will include a temperature-dependent term that represents the free energy of ϕ particles at temperature T. To one loop, the full potential is [2]

$$V_T(\phi) = V(\phi) + \frac{T^4}{2\pi^2}\int_0^\infty dx x^2 \ln\left[1 - \exp[-(x^2 + \mu^2/T^2)^{1/2}]\right], \tag{1.4}$$

where $V(\phi)$ is the zero-temperature one-loop potential, and $\mu^2 = -M^2 + 3\lambda\phi^2$. At high temperature, Eq. 1.4 has the expansion

$$V_T(\phi) = V(\phi) - \frac{\pi^2}{90}T^4 + \frac{\lambda}{8}T^2\phi^2 + \cdots \tag{1.5}$$

The term proportional to T^4 is minus the pressure of a spinless boson, which should be the leading contribution to the free energy, and the second term is the "plasma" mass term for ϕ. Eq. 1.4 has a critical temperature, $T_c = 2M/\lambda^{1/2}$, above which the symmetry is restored.

The phase transition from the symmetric to the broken phase can be either first order or higher order. If at T_c there is a barrier between $\phi = 0$ and the SSB minimum $\phi = \sigma$, the change in ϕ will be discontinuous, signalling a first order transition. If no barrier is present at T_c, the change in ϕ will be continuous, signalling a higher order transition.

In general, at some temperature $T \leq T_c$, the $\phi = 0$ phase is a metastable phase, and will be terminated by the decay of the false vacuum by quantum or thermal tunneling. Here, quantum tunneling will refer the zero-temperature part of the tunneling rate.

The quantum tunneling occurs by the nucleation of bubbles of the new phase. The probability for bubble nucleation is calculated by solving the *Euclidean* equation of motion [3]

$$\Box_E \phi - V'(\phi) = \frac{d^2\phi}{dt^2} + \nabla^2\phi - V'(\phi) = 0 \tag{1.6}$$

(where $V' \equiv dV/d\phi$) with boundary conditions $\phi = 0$ at $\vec{x}^2 + t^2 = \infty$. The probability of bubble nucleation per unit volume per unit time is $\Gamma = A\exp(-S_E)$, where S_E is the Euclidean action for the solution of Eq. 1.6

$$S_E(\phi) = \int d^4x \left[\frac{1}{2}\left(\frac{d\phi}{dt}\right)^2 + \frac{1}{2}(\nabla\phi)^2 + V(\phi)\right]. \tag{1.7}$$

The calculation of the constant A is quite complicated, but for most applications a guess of A on dimensional grounds will suffice.

Of the many possible solutions to Eq. 1.6, the one with least action is the most important. The least action solution has $O(4)$ symmetry, and the Euclidean equation of motion becomes

$$\frac{d^2\phi}{dr^2} + \frac{3}{r}\frac{d\phi}{dr} - V'(\phi) = 0, \tag{1.8}$$

with boundary conditions $\phi = 0$ at $r^2 = \vec{x}^2 + t^2 = \infty$ and $d\phi/dr = 0$ at $r = 0$. In general solutions to this equation can not be found. However in the "thin-wall" approximation, where the difference in energy between the metastable and true vacua are small compared to the height of the barrier, the "damping" term proportional to $d\phi/dr$ can be neglected. The solution for S_E is then simply

$$S_E = \int_0^\sigma d\phi\sqrt{2V(\phi)}. \tag{1.9}$$

The tunneling rate at finite temperature [4] can be found following the above procedure, remembering that field theory at finite temperature is equivalent to Euclidean field theory with the time periodic with period T^{-1}. The finite-temperature tunneling rate is found by solving the equation of motion (only considering the least-action solution, which in this case has $O(3)$ symmetry)

$$\frac{d^2\phi}{ds^2} + \frac{2}{s}\frac{d\phi}{ds} - V_T'(\phi) = 0, \tag{1.10}$$

where $s = \vec{x}^2$. The finite-temperature tunneling rate is

$$\Gamma_T = A\frac{S_3}{T}\exp(-S_3/T), \tag{1.11}$$

where S_3 is the three-dimensional action of the solution of Eq. 1.10

$$S_3 = \int d^3x \left[\frac{1}{2}(\nabla\phi)^2 + V_T(\phi)\right]. \tag{1.12}$$

1.2 *Domain Walls* [5]

The simple model of the previous section can be used to demonstrate domain walls. The Lagrangian can be written in the form

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{4}\lambda(\phi^2 - \langle\phi\rangle^2)^2; \qquad \langle\phi\rangle^2 \equiv \sigma^2 = \frac{M^2}{\lambda}. \tag{1.13}$$

The Z_2 symmetry of the Lagrangian is broken when ϕ obtains a vacuum expectation value $\phi = +\sigma$ or $\phi = -\sigma$. Imagine that space is divided into two regions. In one region of space $\phi = +\sigma$, and in the other region of space $\phi = -\sigma$. The transition region between the two vacua is called a domain wall. Domain walls should be produced, for instance, in the nucleation of bubbles. The bubbles of true vacuum will be either $\phi = +\sigma$ or $\phi = -\sigma$, with equal probability.

Imagine a wall in the $x - y$ plane at $z = 0$. At $z = -\infty$, $\phi = -\sigma$, and at $z = +\infty$, $\phi = +\sigma$. The equation of motion for ϕ is $\Box\phi + \lambda\phi(\phi^2 - \sigma^2) = 0$. The minimum energy solution to the equation of motion, subject to the boundary conditions above, is $\phi_W(z) = \sigma\tanh(z/\Delta)$ where Δ is the "thickness" of the wall, given by $\Delta = (\lambda/2)^{1/2}\sigma^{-1}$.

The finite, but non-zero, thickness of the wall is easy to understand. The terms contributing to the energy include a gradient term and a potential energy term. The gradient term is minimized by making the wall as thick as possible, and the potential term is minimized by making the wall as thin as possible, i.e., by minimizing the distance over which ϕ is away from $\pm\sigma$. The balance between these terms results in a wall of thickness Δ.

The stress tensor with $\phi = \phi_W$ is

$$T_\mu{}^\nu = \frac{\lambda}{2}\sigma^4\cosh^{-4}(z/\Delta)\mathrm{diag}(1,1,1,0). \tag{1.14}$$

From the stress tensor it is possible to define a surface tension for the wall, $\eta = \int T_0{}^0 dz = (4/3)(\lambda/2)^{1/2}\sigma^3$. It is also obvious from the stress tensor that since the (ii) component is equal to the (00) component, the gravitational interaction of the infinite wall will be non-Newtonian. This can lead to some strange interactions. For instance, two infinite walls in the $x - y$ plane will *repel* each other. This strange gravitational behavior only obtains for infinite and straight walls. The gravitational field at large distances from a spherical wall of radius R, would be that of a massive particle of mass $m \simeq R^2\sigma$.

The existence of domain walls can be ruled out today simply on the grounds of their contribution to the total mass of the Universe. A domain wall with $R \simeq R_{\rm horizon} \simeq H_0^{-1} \simeq 10^{28}$cm would contribute a mass of $M_{\rm wall} = \eta R_{\rm wall}^2 = 10^{60}$grams. This would be about a factor of 10^5 larger than the total mass within $R_{\rm horizon}$.

The simple model of this section had domain walls because of the existence of disconnected vacuum states. The general condition for the existence of domain walls in the symmetry breaking $\mathcal{G} \to \mathcal{H}$ is that $\Pi_0(\mathcal{M}) \neq I$, where $\mathcal{M}$ is the manifold of equivalent vacuum states $\mathcal{M} \equiv \mathcal{G}/\mathcal{H}$, and Π_0 is the homotopy group that counts disconnected components. In the above example, $\mathcal{G} = Z_2$, $\mathcal{H} = I$, $\mathcal{M} = Z_2$, and $\Pi_0(\mathcal{M}) = Z_2 \neq I$.

1.3 *Cosmic Strings* [6,7]

A simple model that demonstrates the existence of cosmic strings is a gauge version of the model of the previous section. The Lagrangian of the model contains a U_1 gauge field, A_μ, in addition to the complex Higgs field, ϕ,

$$\mathcal{L} = D_\mu\phi D^\mu\phi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{4}\lambda(\phi^\dagger\phi - \langle\phi\rangle^2)^2; \qquad \langle\phi\rangle^2 = \sigma\exp(i\theta) \tag{1.15}$$

Again, $\sigma^2 = M^2/\lambda$, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, and $D_\mu\phi = \partial_\mu\phi - ieA_\mu\phi$.

Since there is a local gauge symmetry, $\theta = \theta(\vec{x})$, can be position dependent. Since ϕ is single valued, the total $\Delta\theta$ around any closed path must be an integer multiple of 2π. Imagine such a closed path with $\Delta\theta = 2\pi$. As the path is shrunk to a point (and no singularities are encountered), $\Delta\theta$ cannot change from $\Delta\theta = 2\pi$ to $\Delta\theta = 0$. There must therefore be one point contained within the path where the phase θ is undefined, i.e., $\langle\phi\rangle = 0$. The region of false vacuum within the path is part of a tube of false vacuum. These tubes of false vacuum either must be closed or infinite in length, otherwise it would be possible to deform the path around the tube, and contract it to a point without encountering the tube of false vacuum. It will turn out that these tubes of false vacuum have a characteristic transverse dimension far smaller than its length, so they appear as one-dimensional objects called "strings."

The string solution to the Lagrangian in Eq. 1.15 was first found by Nielsen and Olesen [8]. At large distances from an infinite string in the z-direction, $\phi \longrightarrow \sigma\exp(in\theta)$; $A_\mu \longrightarrow -ie^{-1}\partial_\mu\left[\ln(\phi/\sigma)\right]$, where θ is the angle in the $x-y$ plane. Note this choice of A_μ and ϕ is a finite energy solution, since at large distances from the string, $D_\mu\phi \to 0$ and $F_{\mu\nu} \to 0$.

For an infinite string in the z-direction, the stress tensor takes the form $T_\mu{}^\nu = \mu\delta(x)\delta(y)\mathrm{diag}(1,0,0,1)$, where μ is the mass per unit length of the string (string tension) given by $\mu \simeq \sigma^2$.

Far from a string loop of radius R, the gravitational field of the string is that of a particle of mass $M_{\mathrm{string}} = \mu R_{\mathrm{string}}$. For a string that stretches across the present horizon, the mass would be $M_{\mathrm{string}} = 10^{18}(\sigma/\mathrm{GeV})^2$ grams. Cosmic string networks may have very interesting astrophysical consequences, including acting as seeds for the formation of large-scale structure.

String solutions will be present in the symmetry breaking $\mathcal{G} \to \mathcal{H}$, if the manifold of degenerate vacuum states $\mathcal{M} = \mathcal{G}/\mathcal{H}$ contains unshrinkable loops, i.e., if the mapping of $\mathcal{M}$ onto the circle is non-trivial. This is formally expressed by the statement that string solutions exist if $\Pi_1(\mathcal{M}) \neq I$. In the above example $\mathcal{G} = U_1$ was broken, $\mathcal{M}$ is a circle, and $\Pi_1(\mathcal{M}) = Z$, the set of integers.

Some of the cosmological and astrophysical effects of strings are discussed elsewhere in this book [9].

1.4 *Magnetic Monopoles* [10,11]

Domain walls are topological defects in two dimensions, and strings are topological defects in one dimension. Zero-dimensional defects appear in theories with SSB as magnetic monopoles. For a simple model that illustrates the existence of magnetic monopoles, consider an SO_3 gauge theory with a Higgs triplet field ϕ^a

$$\mathcal{L} = \frac{1}{2}D_\mu\phi^a D^\mu\phi^a - \frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} - \frac{1}{4}\lambda(\phi^a\phi^a - \langle\phi\rangle^2)^2; \qquad \langle\phi\rangle^2 = \sigma\hat{\sigma}, \tag{1.16}$$

where $\sigma\hat{\sigma}$ is an isovector in the SO_3 space of magnitude σ and direction $\hat{\sigma}$ ($\hat{\sigma}$ is a unit isovector). Here

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - e\varepsilon_{abc}A^b_\mu A^c_\nu; \qquad D_\mu\phi^a = \partial_\mu\phi^a - e\varepsilon_{abc}A^b_\mu\phi^c. \tag{1.17}$$

Since the theory has a local gauge symmetry, σ is a constant, but $\hat{\sigma}$ can be a function of $\vec{x}$. Imagine a configuration in which at one point $\phi^a = \sigma(0,0,1)$, at another point $\phi^a = \sigma(0,1,0)$, at another point $\phi^a = \sigma(1,0,0)$, and so forth. The lowest-energy configuration has ϕ^a = constant, and the x-dependence of ϕ^a can in general be gauged away. However there are configurations that cannot be deformed into a configuration of constant $\hat{\sigma}$ by a finite-energy transformation. An example of such a configuration is the "hedgehog" configuration, in which $\hat{\sigma} = \hat{r}$, where $\hat{r}$ is the unit vector in the radial direction. But for the obvious angular dependence, the solution is spherically symmetric at $r \to \infty$: $\phi^a(r,t) \to \sigma\hat{r}$; $A^a_\mu(r,t) \to \varepsilon_{\mu ab}\hat{r}_b/er$. The magnetic field at $r \to \infty$ corresponding to the hedgehog solution is

$$B^a_i = \frac{1}{2}\varepsilon_{ijk}F^a_{jk} = \frac{\hat{r}_i\hat{r}^a}{er^2}, \tag{1.18}$$

which is the magnetic field of a magnetic charge of $g = 1/e$. The mass of the field configuration is $M_{\text{monopole}} \simeq \sigma/e$.

There have been many experiments to look for magnetic monopoles. The limit on the average number density of magnetic monopoles in the Universe depends upon the properties of the monopoles (mass, charge, proton decay catalysis, etc.). If magnetic monopoles exist, they would have a multitude of astrophysical consequences.

Monopoles will be present in the symmetry breaking $\mathcal{G} \to \mathcal{H}$, if the manifold of degenerate vacuum states contains unshrinkable surfaces, i.e., if the mapping of $\mathcal{M}$ onto the two-sphere is non-trivial. This is formally expressed by the statement that monopole solutions exist if $\mathbf{\Pi_2}(\mathcal{M}) \neq I$. In the above example $\mathcal{G} = SO_3$, $\mathcal{H} = U_1$ and $\mathbf{\Pi_2}(\mathcal{M})$ is the set of even integers.

1.5 *The Kibble Mechanism* [6]

The existence of the above topological defects is a prediction of many gauge theories with SSB. They are inherently non-perturbative, and cannot be produced in high energy collisions. The only place they can be produced is in phase transitions in the early Universe. Although monopoles, strings, and domain walls are topologically stable, they are, of course, not the minimum energy solution. However the production of the defects in the phase transition seems unavoidable. The mechanism for the production of the defects is known as the Kibble mechanism.

The Kibble mechanism is based upon the fact that in the phase transition the correlation length is limited by the particle horizon. The particle horizon is the maximum distance over which a massless particle could propagate from the time of the bang. Imagine that a particle is emitted at coordinates ($t = 0$, $r = r_H$, $\theta = 0$, $\phi = 0$) and is detected at the origin of the coordinate system at coordinates ($t = t$, $r = 0$, $\theta = 0$, $\phi = 0$). The coordinate r_H is given by

$$\int_0^t \frac{dt'}{R(t')} = \int_0^{r_H} \frac{dr}{(1-kr^2)^{1/2}} \simeq r_H. \tag{1.19}$$

The coordinate r_H by itself is just a label. The proper distance to the horizon is given by $d_H = R(t)r_H$, so

$$d_H = R(t)\int_0^t \frac{dt'}{R(t')}. \tag{1.20}$$

If $R \propto t^n$ ($n > 1$), then $d_H = (1-n)^{-1}t$.

The correlation length in the phase transition sets the maximum distance over which the Higgs field can be correlated. In general, the calculation of the correlation length depends upon the details of the transition. However, the fact that the horizon is finite in the standard cosmology implies that at the phase transition ($t = t_c$, $T =$

T_c), the Higgs field must be uncorrelated on scales greater than the horizon, so the horizon acts as an effective upper bound to the correlation length.

Imagine that at the phase transition the Higgs field is uncorrelated on scales greater than $\xi = d_H$. The initial random nature of $\langle\phi\rangle$ is damped (remember $E_{\min}$ occurs for $\langle\phi\rangle$ = constant). However there are Higgs configurations that are topologically stable and will be frozen in as topological defects.

Consider monopoles as an example of the freezing in of topological defects [12]. The direction of the isovector Higgs field is random on scales greater than ξ. The probability that a random orientation of $\langle\phi\rangle$ will have a hedgehog structure is about 0.1, so there should be about one monopole (or antimonopole) per 10 horizon volumes, $n_M = 0.1 d_H^3 \simeq 0.1(m_{Pl}/T_c^2)^3$, using the age of a radiation-dominated Universe $t = m_{Pl}/T^2$. The entropy density at T_c is $s \simeq T_c^3$, so the monopole-entropy ratio is $n_M/s \simeq 0.1(T_c/m_{Pl})^3$. Since monopole-antimonopole annihilation is not important, if entropy is not created after monopole production, the above monopole-entropy ratio should obtain today. For $T_c = 10^{15}$GeV, $m_M = 10^{16}$ GeV as expected in grand unified theories, $n_M/s \simeq 10^{-13}$, which gives the present energy density in magnetic monopoles $\rho_{\text{monopoles}} \simeq 10^{11}\rho_C$. Obviously some mechanism must suppress monopole production, enhance monopole annihilation, or increase entropy. An increase in entropy would also dilute the abundance of strings and domain walls. It is possible that monopoles were diluted to a level accessible to observation, or that strings were produced after the dilution of monopoles. Detection of monopoles or strings would provide unique information about both particle physics and cosmology. In complicated gauge theories with several symmetry breaking steps there are often interesting hybrid creatures, such as domain walls bounded by strings, strings terminated by monopoles, monopoles with strings through them, etc. They all have unique signatures, and observation of them would provide information about the steps of symmetry breaking.

2. INFLATION

The standard FRW cosmology provides a remarkably simple and beautiful model to describe the Universe. Nevertheless, there are some aspects of the standard picture that strongly suggests that the model is not a complete one. After discussing the problems of the cosmology developed so far, a possible solution to the problems will be presented. This solution goes by the name of "inflation" [13].

2.1 *Loose Ends of the Standard Cosmology*

• *Large-Scale Smoothness*: The Robertson-Walker metric describes a space that is homogeneous and isotropic. Why is space homogeneous and isotropic? There are other possibilities, including homogeneous but anisotropic spaces, and inhomogeneous spaces. The most precise indication of the smoothness of the Universe is provided by the microwave background radiation. If the entire observable Universe was in causal contact when the radiation last scattered, it might be imagined that microphysical processes would have damped any fluctuations and a single temperature would have obtained. However in the standard cosmology the distance to the horizon increases with time. The size of the horizon is conveniently expressed in terms of the entropy within the horizon

$$S_H = s\frac{4\pi}{3}d_H^3 \simeq T^3 t^3. \tag{2.1}$$

The entropy within the horizon today is $S_H(0) \simeq 10^{88}$. In a matter-dominated Universe, $S_H = S_H(0)(1+z)^{-3/2}$, while in a radiation-dominated Universe, $S_H = S_H(0)(1+z)^{-3}$. The entropy in the horizon at recombination when the radiation last scattered was $S_H(t = t_{rec}) \simeq 10^{83}$. The Universe as presently observed consisted of about 10^5 causally disconnected regions at recombination, so causal processes could not have led to smoothness. At the time of primordial nucleosynthesis, the entropy within the horizon volume was $S_H(t_{nucleo}) \simeq 10^{53}$, or about 10^{-30} of the present Universe.

The first untidy fact about the standard cosmology is that there is no physical explanation for why the Universe is smooth.

• *Density Perturbations*: Although the Universe is smooth on large scales, there is a rich structure on small scales. It is usually assumed that the structures observed today were once small perturbations on a smooth background, and have grown as the result of the gravitational instabilities in an expanding Universe. The relic photons did not take part in the gravitational collapse, and remain as fossil evidence of the once-smooth Universe.

Density inhomogeneities are usually expressed in a Fourier expansion

$$\left(\frac{\delta\rho}{\rho}\right) = (2\pi)^{-3}\int \delta_k \exp(-i\vec{k}\cdot\vec{x})d^3k. \tag{2.2}$$

Here k is a co-moving label. The *physical* wavenumber and wavelength are related to k by $k_{ph} = k/R(t)$, $\lambda_{ph} = (2\pi/k)R(t)$. It is also convenient to express the scale of the perturbation in terms of the mass in baryons contained within the perturbation. For constant B, the baryon mass on scale λ is proportional to λ^3. The baryon mass within the horizon at time t is $M_H(t) \simeq m_p B s d_H^3 \propto S_H$. The quantity usually referred to as $(\delta\rho/\rho)$ on a given scale is the r.m.s. mass fluctuations on that scale

$$\left(\frac{\delta\rho}{\rho}\right)_k^2 = (2\pi)^{-3}k^3|\delta_k|^2. \tag{2.3}$$

The exact nature of the perturbations required for galaxy formation is unknown. A promising choice for density perturbations is that as every distance scale comes within the horizon, the r.m.s. fluctuations in the density are $10^{-4} - 10^{-5}$ *independent of the scale*. This is usually expressed as

$$\left(\frac{\delta\rho}{\rho}\right)_H \simeq 10^{-4}. \tag{2.4}$$

Here $(\delta\rho/\rho)_H$ is $(\delta\rho/\rho)$ on the scale $\lambda = d_H = t$ at time $t = d_H$.

The evolution of the perturbations within the horizon is determined by local physics, e.g., the Jeans criteria. The behavior of the perturbations outside of the horizon is complicated by the fact that there is a "gauge dependence" that reflects the freedom of the choice for a reference spacetime. Nevertheless, the growth of metric perturbations on scales larger than the horizon can be studied by using the uniform Hubble flow gauge (time slices chosen to give constant H). From the Friedmann equation with H constant, fluctuations in ρ are equivalent to fluctuations in the spatial curvature k/R^2

$$\delta\left(\frac{k}{R^2}\right) \Longleftrightarrow \delta\left(\frac{8\pi G}{3}\rho\right). \tag{2.5}$$

In a radiation-dominated (matter-dominated) Universe, $\rho \propto R^{-4}$ (R^{-3}), so

$$(\delta\rho/\rho) \propto \begin{cases} R^{-2}/R^{-4} \sim (1+z)^{-2} & \text{(RD)} \\ R^{-2}/R^{-3} \sim (1+z)^{-1} & \text{(MD)}. \end{cases} \tag{2.6}$$

Since $S_H \propto (1+z)^{-3}$ for (RD) and $S_H \propto (1+z)^{3/2}$ for (MD), $(\delta\rho/\rho) \propto S_H^{2/3} \propto M_H^{2/3}$ for both (RD) and (MD). So as any scale comes within the horizon, the growth that scale has experienced while outside the horizon depends upon the mass contained in the scale as it enters the horizon

$$\left(\frac{\delta\rho}{\rho}\right)_H \sim \left(\frac{\delta\rho}{\rho}\right)_0 (M_H(t))^{2/3}, \tag{2.7}$$

where t_0 is some arbitrary initial time. If $(\delta\rho/\rho)_0$ is proportional to $M^{-2/3}$, as each scale comes within the horizon, $(\delta\rho/\rho)$ will be a constant. Larger scales have smaller initial amplitudes, but they have a longer time to grow outside the horizon. If $(\delta\rho/\rho)_0$ is characterized by a steeper spectrum, the first scales that come within the horizon would have been non-linear. If $(\delta\rho/\rho)_0$ is characterized by a flatter spectrum, larger scales would have larger $(\delta\rho/\rho)$ at horizon crossing.

The standard model can shed no light on the origin of the density perturbations. It must simply assume that at $t = 0$ there are perturbations of the appropriate magnitude and spectrum impressed upon the metric.

• *Spatial Flatness - Age*: In the standard Friedmann cosmology, $\Omega - 1 = k/R^2H^2$. In the past, $H^2 \propto \rho$, which for a matter-dominated Universe gives $H^2 \propto R^{-3}$, and for a radiation-dominated Universe gives $H^2 \propto R^{-4}$. Since today $|\Omega - 1|$ is of order unity, at previous epochs

$$|\Omega - 1| \simeq \begin{cases} R/R_0 = (1+z)^{-1} & \text{(MD)} \\ (R/R_0)^2 = (1+z)^{-2} & \text{(RD)}. \end{cases} \tag{2.8}$$

At the time of primordial nucleosynthesis, $|\Omega - 1| \leq 10^{-16}$, and at the planck time $|\Omega - 1| \leq 10^{-60}$. Obviously Ω was *very* close to one at early times, i.e., the curvature term was small compared to H^2 and $8\pi G\rho/3$.

The smallness of the curvature term is necessary for the Universe to survive as long as it has without either re-collapsing (for $k = +1$) or becoming curvature dominated (for $k = -1$). The natural time scale in the Friedmann equation is the planck time $t_{Pl} = 2 \times 10^{-43}$ sec. The difference between the kinetic term (H^2) and the potential term $(8\pi G\rho/3)$ is the curvature term. This must be small in order for the Universe to expand for 10^{17} sec. $\sim 10^{60} t_{Pl}$.

The standard Friedmann model has no explanation for the present spatial flatness of the Universe.

• *Cosmological Constant*: The most general form of Einstein's equations includes a cosmological constant

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu} + \Lambda g_{\mu\nu}. \tag{2.9}$$

With the stress-tensor in the perfect-fluid form (U_μ is the fluid velocity vector, $U_\mu = (1,0,0,0)$ in the fluid rest frame) $T_{\mu\nu} = -pg_{\mu\nu} + (\rho + p)U_\mu U_\nu$, the effect of the cosmological constant is to add to the fluid contributions to ρ and p, terms $\rho_\Lambda = -p_\Lambda = \Lambda/8\pi G$. The generalized energy and pressure are given by $\rho^* = \rho + \rho_\Lambda$, $p^* = p + p_\Lambda$, and the Einstein equations can be written in terms of $T^*_{\mu\nu}$, which is $T_{\mu\nu}$ with $\rho \to \rho^*$, $p \to p^*$,

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T^*_{\mu\nu}. \tag{2.10}$$

If ρ^* and p^* are dominated by ρ_Λ and p_Λ, the conservation and Friedmann equations become

$$\rho^* \propto R^0 = \text{constant}; \quad H^2 = \frac{8\pi G\rho^*}{3} = \frac{\Lambda}{3}, \tag{2.11}$$

which has solution $R \propto \exp(Ht)$.

Today the contribution of a cosmological constant to the energy density of the Universe must be less than ρ_C. In useful units, $\rho_C = 8.07 \times 10^{-47}h^2$ GeV^4. Among the contributions to Λ are contributions from the condensates of Higgs particles due to SSB. During cosmological phase transitions, the vacuum energy density changes by σ^4, where σ is the zero-temperature vacuum expectation value of the Higgs field. This change in the vacuum energy is 10^9GeV^4 for the electroweak transition, and 10^{60}GeV^4 for the GUT transition. A cosmological constant of this order must be present before the transition to ensure that after all transitions are complete the energy density of the vacuum is less than about 10^{-47}GeV^4.

The standard cosmology cannot explain why the present vacuum energy density is so small.

- *Unwanted Relics*: There are a variety of particles that are expected to survive annihilation and contribute to the present energy density. Particles with very large masses typically have very small annihilation cross sections and should be abundant. This is rather unfortunate, as their contribution to the mass density typically is many orders of magnitude larger than ρ_C. The magnetic monopoles produced in the GUT phase transition are an example of such an unwanted relic.

The standard cosmology has no mechanism of ridding the Universe of unwanted particles.

The problems mentioned here do not invalidate the standard cosmology. They are accommodated by the standard cosmology, but they are not explained. The goal of cosmology is to explain the present structure of the Universe on the basis of physical law, and one hopes that physical law will one day explain the above points. Inflation is a model for such an explanation.

2.2 *Inflation - The Basic Picture* [14,15]

Consider as a model for new inflation a phase transition associated with SSB with a scalar potential given by

$$V(\phi) = \frac{1}{4}\lambda(\phi^2 - \sigma^2)^2. \tag{2.12}$$

At temperatures $T \gg T_c = 2\sigma$, $\langle\phi\rangle = 0$, and $V(\langle\phi\rangle) = \lambda\sigma^4/4 \equiv \rho_V$. At temperatures $T \ll T_c$, $\langle\phi\rangle = \sigma$, and $V(\langle\phi\rangle) = \rho_V = 0$. New inflation will occur as ϕ makes the transition from the high temperature minimum of the potential to the low temperature minimum of the potential.

At some temperature $T \leq T_c$, in some region of the Universe, the Higgs field will make the transition from $\phi = 0$ to $\phi \neq 0$. Assume that in this region of the Universe ϕ is spatially uniform. The evolution of ϕ to the low-temperature ground state is not instantaneous, but requires a time determined by the dynamics of the theory. The equation of motion for ϕ can be found from $T^{\mu\nu}{}_{;\nu} = 0$, where $T_{\mu\nu} = -\partial_\mu\phi\partial_\nu\phi - \mathcal{L}g_{\mu\nu}$. With the assumption that ϕ is spatially homogeneous $\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0$, where $V' = \partial V/\partial\phi$, and $H^2 = 8\pi G\rho/3$. The contributions to ρ include a radiation term ρ_R, a kinetic term for ϕ, and a potential term for ϕ:

$$\rho = \rho_R + \frac{1}{2}\dot{\phi}^2 + V(\phi). \tag{2.13}$$

If there is a "flat" region in $V(\phi)$, the evolution of ϕ will be "slow" and the $\ddot{\phi}$ term can be neglected in the equation of motion. In this flat region ϕ will change very slowly and $V(\phi)$ will be roughly constant. Therefore the contribution to ρ from $V(\phi)$ will be roughly constant and will rapidly come to dominate ρ_R which decreases in proportion to R^4. When ρ is dominated by potential energy the scale factor increases exponentially. If this flat region in the potential extends from ϕ_s to ϕ_e, R will increase by an amount $R(\phi_e) = R(\phi_s)\exp(H\Delta t)$, where Δt is the time it takes to make the transition from ϕ_s to ϕ_e, and $H^2 \simeq V(\phi)/m_{Pl}^2 \simeq \sigma^4/m_{Pl}^2$. For a concrete example, assume for the moment that $\Delta t = 100H^{-1}$.

Now assume that after traversing the flat region in the potential, at $\phi \geq \phi_e \simeq \sigma$ there is a "steep" region in the potential. In this steep region the oscillations in the zero momentum mode of ϕ will rapidly convert the potential energy to radiation. If this conversion is efficient, the Universe will be reheated to a temperature $T_{\rm RH}$ found by equating the potential energy density to the radiation energy density: $V(\phi) \simeq T^4$, or $T_{\rm RH} \simeq \sigma$.

This is the basis scenario for new inflation. To illustrate the scenario, take $\sigma = 10^{14}$GeV, and the initial size of the region to be the size of the horizon at T_c, $R_i = H^{-1} \simeq m_{Pl}/\sigma^2 = 10^{-23}$cm (it is reasonable to expect ϕ to be uniform on scales that are in causal contact). The initial entropy in this region is $S_i \simeq (R_iT_i)^3 \simeq 10^{14}$. The final size of the region in the example where $\Delta t = 100H^{-1}$ is $R_f = \exp(100)R_i \simeq 3\times10^{20}$cm. With efficient reheating $T_{\rm RH} = \sigma$, and the final entropy contained in the region is $S_f = (R_fT_{\rm RH})^3 \simeq 10^{144}$.

This large creation of entropy has helped with three out of four problems. *Large-Scale Smoothness*: At $T = 10^{14}$GeV, the presently observable Universe ($S = 10^{88}$) was contained in a size of 10cm, and easily fit within the smooth region after inflation. *Density Perturbations*: To see how inflation generates density perturbations it is necessary to treat the dynamics of the scalar field in greater detail than done so far. This will be done shortly. *Spatial Flatness - Age*: After inflation R has increased by $\exp(100)$ but the final temperature is close to the initial temperature. Thus, immediately after inflation the spatial curvature term k/R^2 is a factor of $\exp(-200)$ smaller, while the energy density term is unchanged. *Cosmological Constant*: Inflation does not help the cosmological constant problem. *Unwanted Relics*: The number density of particles present before inflation is decreased by a factor of $R_i^3/R_f^3 \simeq \exp(-300)$. This is true also for the original photons. It is crucial to create entropy in the termination of inflation.

In this example it was assumed that the slow-roll period lasted for 100 e-folds. The minimum number of e-folds is the number required to fit the observed entropy of 10^{88} into a single inflation region. The final entropy in the inflation region is $S_f \simeq T_{\rm RH}^3R_f^3$. The size of the final region is related to the number of e-folds by $R_f^3 = \exp(3N)R_i^3$, assuming little or no growth during the oscillation phase. The largest possible smooth initial region is the size of the horizon at the phase transition, $R_i = H^{-1}(T_c) \simeq m_{Pl}/\sigma^2$, assuming $T_c = \sigma$. The maximum reheat temperature is

$T_{RH} \simeq \sigma$, so the final entropy is $S_f \simeq \sigma^3 \exp(3N) m_{Pl}^3/\sigma^6 \simeq \exp(3N) m_{Pl}^3/\sigma^3$. The requirement $S_f \geq 10^{88}$ gives $N \geq 58 + \ln(\sigma/10^{15}\mathrm{GeV})$.

2.3 *Dynamics of Inflation*

The evolution of the spatially homogeneous scalar field (zero momentum mode of the scalar field) is crucial for inflation. If the coupling of the scalar field to other fields are included, the equation of motion for the zero-momentum mode of ϕ is (ϕ will denote the zero-momentum mode unless otherwise indicated)

$$\ddot{\phi} + 3H\dot{\phi} + \Gamma_\phi \dot{\phi} + V'(\phi) = 0, \tag{2.14}$$

where Γ_ϕ is the ϕ decay width. The decay width is typically $\Gamma_\phi = h^2 m_\phi$, where h is a coupling constant, and m_ϕ is the mass of ϕ [16]. The energy density and pressure of ϕ are given by

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi); \qquad p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi). \tag{2.15}$$

The "slow roll" regime is when the $\ddot{\phi}$ and Γ_ϕ terms in Eq. 2.14 can be neglected, and $V(\phi)$ is the dominant term in Eq. 2.13. The equation of motion during slow roll is

$$3H\dot{\phi} = -V'(\phi). \tag{2.16}$$

Neglecting $\ddot{\phi}$ is consistent if

$$|V''(\phi)| \leq 9H^2; \qquad |V'(\phi) m_{Pl}/V(\phi)| \leq (48\pi)^{1/2}. \tag{2.17}$$

These conditions will determine the duration of slow roll.

The number of e-folds of inflation while ϕ rolls from ϕ_1 to ϕ_2 during slow roll is given by

$$N(\phi_1 \rightarrow \phi_2) = \int_{\phi_1}^{\phi_2} H dt = -3 \int_{\phi_1}^{\phi_2} \frac{H^2(\phi)}{V'(\phi)} d\phi, \tag{2.18}$$

where $dt = \dot{\phi}^{-1} d\phi = -3H/V' d\phi$.

With ρ_ϕ given by Eq. 2.15, $\dot{\rho_\phi} = \dot{\phi}\ddot{\phi} + \dot{\phi}V'(\phi)$, and using Eq. 2.14, $\dot{\rho_\phi} = -3H\dot{\phi}^2 - \Gamma_\phi \dot{\phi}^2$. The two terms in the equation for $\dot{\rho_\phi}$ represent the change due to the redshift of the kinetic energy in the ϕ field (proportional to H) and the change due to decay of the ϕ field (proportional to Γ_ϕ). When ϕ starts oscillating about the minimum of the potential, the energy transfers between kinetic and potential energy until ϕ decays. Over an oscillation cycle $\langle \dot{\phi}^2 \rangle = \rho_\phi$, and $\dot{\phi}^2$ can be replaced by ρ_ϕ in the equation for $\dot{\rho_\phi}$. The energy from ϕ decay is transferred into radiation, and the equation for the evolution of ρ_R becomes $\dot{\rho_R} = -4H\rho_R + \Gamma_\phi \rho_\phi$.

The equations for $\dot{\rho_R}$ and $\dot{\rho_\phi}$ can be integrated to study reheating. If oscillation about the minimum begins at $t = t_3$ and $R = R_3$ with $\rho_\phi = \sigma^4$, the ϕ energy density will decrease as

$$\rho_\phi = \sigma^4 \left(\frac{R}{R_3}\right)^{-3} \exp[-\Gamma(t - t_3)]. \tag{2.19}$$

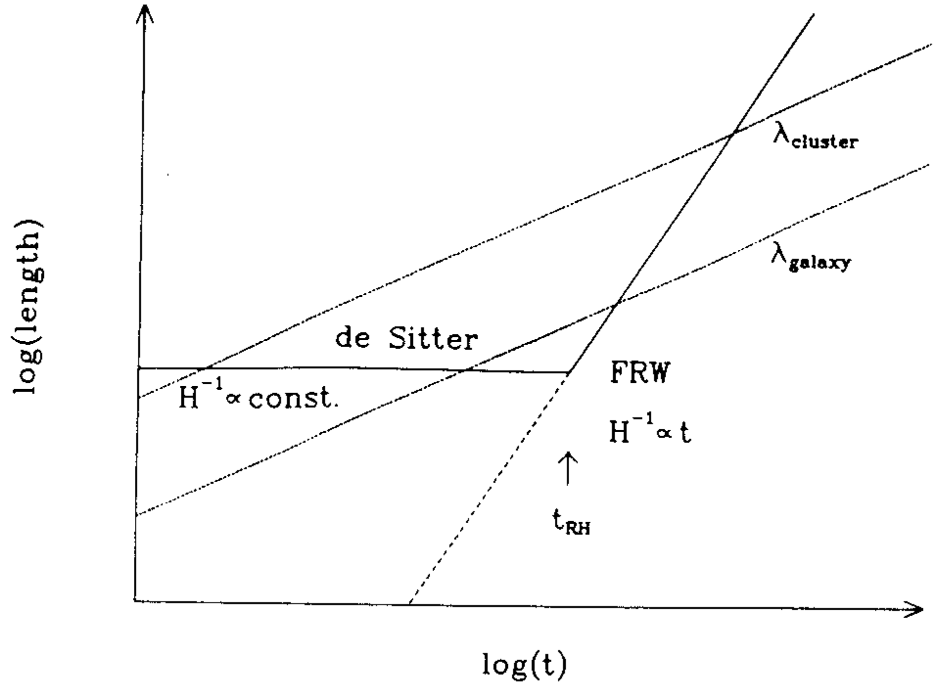

Figure 1: Physical scales cross the physics horizon twice

Until decay, the ϕ energy density decreases in expansion as the energy density for massive particles. When ϕ decays, the remaining energy is converted to radiation ($\rho_\phi \rightarrow (\pi^2/30)g_* T_{\rm RH}^4$). Obviously, the longer ϕ oscillates before decay, the less energy will be available for conversion into radiation, and the lower will be the reheat temperature. If the decay width is large compared to the expansion rate at the start of oscillation, $H_3 \simeq \sigma^2/m_{Pl}$, reheating will occur before damping of ρ_ϕ, and $T_{\rm RH} \simeq g_*^{-1/4}\sigma$. If the decay width is small compared to H_3, ϕ will oscillate until the age of the Universe is equal to the ϕ lifetime, i.e., until $\Gamma_\phi = H = \rho_\phi^{1/2}/m_{Pl}$. Then when $\rho_\phi \rightarrow g_* T_{\rm RH}^4$, the reheat temperature will be $T_{\rm RH} = g_*^{-1/4}\rho_\phi^{1/4} = g_*^{-1/4}(\Gamma_\phi m_{Pl})^{1/2}$.

Now consider the generation of fluctuations in ρ. In the FRW radiation-dominated Universe $H^{-1} \propto t$, while during the slow-roll epoch, $H^{-1} \simeq m_{Pl}/V(\phi)^2$ is constant. If H is constant, the Universe in approximately in a de Sitter phase. H^{-1} sets the scale over which microphysical processes can act. H^{-1} will be called the "physics horizon." During the slow roll phase the physics horizon is constant and physical scales increase exponentially. Eventually, physical scales once smaller than the horizon will become larger than the horizon. After termination of the slow-roll phase the Universe reheats, behaves like a FRW radiation-dominated Universe, and scales outside the horizon will eventually come (back) within the horizon. This double-cross of the physics horizon is illustrated in Fig. 1.

Notice that the *last* scales to go outside H^{-1} during the de Sitter phase are the *first* scales to come back inside H^{-1} during the FRW phase. Ignoring the σ dependence, the Hubble radius today ($\simeq$ 3000 Mpc) crossed the horizon 58 e-folds before the end of inflation. Any scale smaller than the Hubble radius today crossed the horizon $58 + \ln(\sigma/10^{15}\text{GeV}) + \ln(\lambda/3000\text{Mpc})$ e-folds before the end of inflation. Using $B = 10^{-10}$, the mass in baryons inside the horizon today is $10^{21}M_\odot$. Since $B \propto \lambda^3$, any baryon mass scale crossed the horizon $58 + \ln(\sigma/10^{15}\text{GeV}) + (1/3)\ln(M/10^{21}M_\odot)$ e-folds before the end of inflation. Scales that will eventually contain a galaxy mass ($M = 10^{11}M_\odot$) crossed the horizon 50 e-folds before the end of inflation, while scales that will eventually contain a galaxy cluster mass ($M = 10^{14}M_\odot$) crossed the horizon 53 e-folds before the end of inflation.

So far it has been assumed that the ϕ field is constant. However there are quantum fluctuations in ϕ due to the fact that during the slow-roll epoch the Universe is approximately in a de Sitter phase [18,19,20,21]. If the fluctuations $\delta\phi$ are expressed as a Fourier expansion

$$\delta\phi = (2\pi)^{-3}\int d^3k\,\delta\phi_k \exp(-i\vec{k}\cdot\vec{x}), \tag{2.20}$$

then the de Sitter fluctuations result in (note: $\Delta\phi \equiv \Delta\phi_k$)

$$(\Delta\phi)^2 \equiv (2\pi)^{-3}k^3|\delta\phi_k|^2 = \left(\frac{H}{2\pi}\right)^2. \tag{2.21}$$

These fluctuations obtain on scales less than the physics horizon during the de Sitter phase. As each scale goes outside the horizon during slow roll, it has fluctuations $(\Delta\phi)^2 = (H/2\pi)^2$. Since the energy density depends upon ϕ, the fluctuations in ϕ lead to fluctuations in ρ of $\delta\rho = (\partial V/\partial\phi)\Delta\phi$. Using $\rho \simeq V \simeq \sigma^4$ and $\partial V/\partial\phi = -3H\dot{\phi}$, fluctuations in ϕ lead to

$$\left(\frac{\delta\rho}{\rho}\right)_k \simeq \left(\frac{\dot{\phi}H^2}{\sigma^4}\right). \tag{2.22}$$

Once the scale is larger than H^{-1}, it can no longer be affected by microphysics. The behavior of the perturbation outside the horizon is gauge-dependent. However the behavior outside the horizon can be characterized by a parameter ς, given by

$$\varsigma \equiv \frac{\delta\rho}{\rho + p} \simeq \begin{cases} \delta\rho/\rho & \text{FRW} \\ \delta\rho/\dot{\phi}^2 & \text{de Sitter.} \end{cases} \tag{2.23}$$

When a scale comes back within the horizon during the FRW phase, ς is the same as when it first went outside the horizon during inflation. Therefore, $(\delta\rho/\rho)$ relevant for galaxy formation is given by [17]

$$\left(\frac{\delta\rho}{\rho}\right)_H \simeq \left(\frac{-3H\dot{\phi}\Delta\phi}{\dot{\phi}^2}\right)_H \simeq \left(\frac{H^2}{\dot{\phi}}\right)_H. \tag{2.24}$$

With the approximation that H and $\dot{\phi}$ are constant during the slow-roll phase, $(\delta\rho/\rho)$ as it re-enters the horizon will be scale free . In the slow-roll period, $\dot{\phi} = -V'(\phi)/3H$, and the equation for $(\delta\rho/\rho)$ becomes

$$\left(\frac{\delta\rho}{\rho}\right)_H \simeq \left(\frac{-3H^3}{V'(\phi)}\right). \tag{2.25}$$

2.4 *Specific Models*

The first example considered is the original attempt to implement new inflation. The model is based upon a SU_5 GUT with symmetry breaking via the Coleman-Weinberg mechanism [14,15]. The scalar field responsible for inflation (hereafter referred to as the *inflaton*) is in the 24-dimensional representation of SU_5 and is responsible for the symmetry breaking $SU_5 \rightarrow SU_3 \times SU_2 \times U_1$. Let ϕ denote the magnitude of the Higgs field in the $SU_3 \times SU_2 \times U_1$ direction. The one-loop, zero-temperature Coleman-Weinberg potential is

$$V(\phi) = B\sigma^4/2 + B\phi^4\left[\ln(\phi^2/\sigma^2) - 1/2\right], \quad (2.26)$$

where $B = 25\alpha_{GUT}^2/16 \simeq 10^{-3}$, and $\sigma = 2 \times 10^{15}$GeV. Because of the absence of a mass term, the potential is very flat near the origin (SSB arises due to one-loop radiative corrections). For $\phi \ll \sigma$, the potential may be approximated in the slow-roll regime by

$$V(\phi) \simeq B\sigma^4/2 - \lambda\phi^4/4; \qquad \lambda \simeq |4B\ln(\phi^2/\sigma^2)| \simeq 0.1. \quad (2.27)$$

For $\phi \ll \sigma$

$$V(\phi) \simeq B\sigma^4/2; \qquad H^2 = \frac{8\pi G\rho}{3} \simeq \frac{4\pi}{3}\frac{B\sigma^4}{m_{Pl}^2}. \quad (2.28)$$

The critical temperature for this potential is about 10^{14}GeV. The finite temperature potential has a small temperature-dependent barrier near the origin, and it is not until $T = 10^9$GeV or so that this barrier is low enough that the action for bubble nucleation drops to order unity. At this time the Universe will undergo "spinodal decomposition" and break up into irregularly shaped fluctuation regions within which ϕ is approximately constant.

Consider the evolution of ϕ in the slow-roll regime. Slow roll commences at ϕ_s and ends at ϕ_e. The end of slow roll is determined by the condition $|V''(\phi_e)| = 9H^2$, or $\phi_e^2 = 3H^2/\lambda$. For any ϕ in the region $\phi_s \leq \phi \leq \phi_e$, the number of e-folds from ϕ to ϕ_e (time t to time t_e) is given by

$$N(\phi \to \phi_e) = \int_t^{t_e} H dt = \int_\phi^{\phi_e} H\dot{\phi}^{-1} d\phi. \quad (2.29)$$

Using $3H\dot{\phi} = -dV/d\phi$ during slow roll,

$$N(\phi \to \phi_e) = \frac{3H^2}{2\lambda}\left(\frac{1}{\phi^2} - \frac{1}{\phi_e^2}\right). \quad (2.30)$$

The total number of e-folds in slow roll depends upon ϕ_s. To have enough inflation, $N(\phi_s \to \phi_e)$ must be greater than 58. Since λ is 10^{-1}, ϕ_s must be smaller than H in order to have sufficient e-folds. However de Sitter space fluctuations introduce uncertainties in ϕ of this order. The quantum fluctuations may prematurely terminate inflation. At the very least they suggest that the semiclassical equations of motion may be invalid.

More serious is the magnitude of the density fluctuations [18,19,20,21]. During slow roll for the Coleman-Weinberg potential $V'(\phi) \simeq \lambda\phi^3$, and Eq. 2.25 gives

$$\left(\frac{\delta\rho}{\rho}\right)_H \simeq \frac{3H^3}{\lambda\phi^3} \simeq \left(\frac{\lambda}{3}\right)^{1/2}[2N(\phi \to \phi_e)]^{3/2}, \quad (2.31)$$

where Eq. 2.30 has been used to express ϕ in terms of the number of e-folds before the end of inflation. Although $(\delta\rho/\rho)$ depends upon N to a power, N depends upon the *logarithm* of the length or mass scale, so the scale dependence of $(\delta\rho/\rho)$ is only logarithmic. The problem with the Coleman-Weinberg potential is not the spectrum, but the magnitude of the perturbations. Using $\lambda = 0.1$ and $N(\phi \to \phi_e) = 58 + (1/3)\ln(M/10^{21}M_\odot)$, $(\delta\rho/\rho)_H$ on the scale of galaxies is 182, and on the scale of clusters is 199. The spectrum is very flat, but about 10^6 too large. Notice that a smaller λ cures both problems.

Although the original model for new inflation was a failure, it pointed the way for the construction of somewhat more successful models. The potential of the original Coleman-Weinberg model was not flat enough, i.e., λ was too large. If ϕ couples to gauge fields, λ will be of order α_{GUT}^2, which is too large. If ϕ is a weakly-coupled gauge singlet, the effective λ can be small, and will remain small after radiative corrections. If $\lambda \leq 10^{-12}$, the density perturbations from Eq. 2.31 will be small enough. However a weakly-coupled inflaton will have a small decay width, and the reheat temperature will be low. If λ is also the magnitude for the coupling of the inflaton to other fields, the decay width at the minimum will be $\Gamma_\phi \simeq \lambda^2 m_\phi \simeq \lambda^2\sigma$, and the reheat temperature will be $T_{\rm RH} \simeq (\Gamma_\phi m_{Pl})^{1/2} \simeq 10^5$GeV for $\lambda = 10^{-12}$ and $\sigma = 10^{15}$GeV. A more careful calculation may give one or two orders of magnitude larger value of $T_{\rm RH}$, but it is clear that a weakly-coupled field will have a low reheat temperature. This presents a problem for baryogenesis. Any baryon asymmetry present before inflation will be diluted due to the creation of the large amount of entropy, so it is necessary to create the baryon asymmetry either during or after the reheating epoch. Many inflation models are squeezed between the requirement of a weakly coupled inflaton for a flat potential and an inflaton that has a large enough decay width to give $T_{\rm RH}$ large enough for baryogenesis.

Supersymmetric models have been proposed as a mechanism to stabilize small couplings in the inflaton potential against radiative corrections. Supersymmetric models introduce several additional potential problems. The high-temperature minimum of the potential is generally not at $\phi = 0$, and $\langle\phi\rangle$ may smoothly evolve to the zero-temperature minimum. There are two possible solutions. If the high-temperature minimum is at $\phi \leq 0$, there will always be a barrier between the high-temperature and the low-temperature minimum. The other solution is to ignore the problem. Since the inflaton must be weakly coupled, it may never be in LTE, and the initial value of ϕ may be random. Another problem with supersymmetric models is the gravitino problem. Gravitinos are weakly-interacting, long-lived particles present in supersymmetric models. They will be produced in reheating in embarrassingly large numbers unless the reheat temperature is less than about 10^9GeV. Finally, in supersymmetric models where supersymmetry breaking is done with a Polonyi field, the Polonyi field can be set into oscillations that will not decay because the Polonyi field is "hidden." Since the energy density in the oscillating field behaves like non-relativistic matter, it will eventually come to dominate the Universe.

For successful new inflation, several requirements must be fulfilled. The requirements occur during different periods of inflation [22].

- *Start Inflation*: The scalar field must be smooth in a region such that the energy density and pressure associated with spatial gradients in ϕ are smaller than the potential energy. If the average value of ϕ is ϕ_0 and the region has typical spatial dimension L, this requirement implies $(\nabla\phi)^2 = O(\phi_0/L) \ll V(\phi_0) = O(\sigma^4)$. If this requirement is not met and the $(\nabla\phi)^2$ term dominates, $R(t)$ will expand as t to a power and inflation will not occur. However once $V(\phi)$ does dominate, the gradient terms rapidly become small in the exponential expansion and can be ignored.

In supersymmetric models where LTE is obtained, the high-temperature minimum of $V(\phi)$ should be at $\phi \leq 0$ to prevent ϕ from smoothly evolving to the zero-temperature minimum without inflating.

- *Start Slow Roll*: If ϕ is not a gauge singlet it may roll in the "wrong" direction. For instance for the Coleman-Weinberg SU_5 model, the steepest direction for ϕ near the origin is toward a minimum where $SU_4 \times U_1$ is the unbroken symmetry. If ϕ is a gauge singlet there is no problem with ending up in the wrong phase.

In order to have slow roll, the potential must have a flat region in which $|V''(\phi)| \leq 9H^2$ and $|V'(\phi)m_{Pl}/V(\phi)| \leq (48\pi)^{1/2}$.

- *Roll Far Enough*: The interval of slow roll, $[\phi_s, \phi_e]$, must be large enough that quantum fluctuations do not terminate slow roll. This condition will be met if $\phi_e - \phi_s \gg H$.

The number of e-folds, $N = \int H dt$ from ϕ_s to ϕ_e, must be greater than $58 + \ln(\sigma/10^{15}\mathrm{GeV})$.

• *Small Perturbations*: The magnitude of the perturbations must be less than of order 10^{-4} on the scale of galaxies to clusters in order to avoid large fluctuations in the MBR. If the fluctuations produced in inflation are to lead to structure formation, they should be greater than of order 10^{-5}. Therefore during slow roll $H^2/\dot{\phi} \leq 10^{-4}$.

In addition to the scalar perturbations discussed so far, inflation will produce tensor perturbations. These tensor perturbations can be thought of as gravity waves. As each scale leaves the horizon during inflation the energy density of gravity waves on that scale is $\rho_{GW} \simeq H^4$. In terms of a dimensionless amplitude $h = H/m_{Pl}$ and wavelength λ, $\rho_{GW} \simeq (m_{Pl}^2 h^2/\lambda^2)_{\lambda = H^{-1}}$. These gravity waves will re-enter the horizon during the FRW phase with the same dimensionless amplitude h, and induce an anisotropy in the MBR of order h. For $\delta T/T \leq 10^{-4}$, $h = H/m_{Pl} \leq 10^{-4}$. Since $H \simeq \sigma^2/m_{Pl}$, σ must be less than about 10^{17}GeV.

• *Exit Properly*: The reheat temperature must be high enough so the Universe is radiation dominated during primordial nucleosynthesis. Using $T_{\mathrm{RH}} = (\Gamma_\phi m_{Pl})^{1/2}$, $T_{\mathrm{RH}} \geq 1$ MeV requires $\Gamma_\phi \geq 10^{-25}$GeV. If baryogenesis proceeds in the standard way, then $T_{\mathrm{RH}} \geq 10^9$GeV, which implies $\Gamma_\phi \geq 10^{-1}$GeV. In order to ameliorate the problem of low reheat temperature and baryogenesis, it has been proposed that a baryon asymmetry is created by the decay of the inflaton. The energy density in the coherent oscillations can be thought of as due to a collection of zero momentum inflatons with number density $n_\phi = \rho_\phi/m_\phi$. In reheating, $\rho_\phi \to g_* T_{\mathrm{RH}}^4$, so $n_\phi = g_* T_{\mathrm{RH}}^4/m_\phi$ at reheating. Suppose the inflaton decays into a particle, S, which, in turn, decays out of equilibrium with baryon number violation. The number density of S's that decay is the same as the number density of parent inflatons. If the CP parameter in the decay of the S is ϵ, then the asymmetry in baryons produced by the S is $n_B = \epsilon n_\phi = \epsilon g_* T_{\mathrm{RH}}^4/m_\phi$. The entropy density produced after thermalization of the inflaton decay products is $s = g_* T_{\mathrm{RH}}^3$. Therefore $B \equiv n_B/s = \epsilon T_{\mathrm{RH}}/m_\phi$. If $B \geq 10^{-10}$, then $T_{\mathrm{RH}} \geq 10^{-10} m_\phi/\epsilon$.

There is a model-dependent upper limit on T_{RH} to avoid making unwanted relics. For example, in supersymmetric models, $T_{\mathrm{RH}} \leq 10^9$GeV to avoid overproducing gravitinos.

The above problems and some possible solutions are given in Table 1. Although there are models that satisfy all the above requirements, none of them seem so compelling that they must be the final answer. In fact, in the past few years there has been increasing effort in the generalization of inflation as a phenomena that is decoupled from a cosmological phase transition.

2.5 *Present Status and Future Directions*

Although the general scenario of inflation presents a very attractive means to ameliorate at least some of the untidiness of the standard model, it is by no means clear that all (or even any) problems are solved or understood. It is now clear that there are models, both supersymmetric and non-supersymmetric, which can successfully implement the program of new inflation as outlined above. It is useful to normalize the more non-standard models of inflation by comparing them to these two "standard" models of inflation.

The non-supersymmetric model is a GUT model based upon SU_5. The model was first proposed by Shafi and Vilenkin [23], and refined by Pi [24]. In the latest version of the model the inflaton is the real part of a complex gauge-singlet field

EPOCH	PROBLEM	POSSIBLE SOLUTION
Start	ϕ Smooth	$(\nabla\phi)^2 \ll V(\phi)$
Inflation	Thermal Constraint	$\langle\phi\rangle \leq 0$
Execute	Roll in Right Direction	ϕ is gauge singlet
Slow Roll	Flat Region in $V(\phi)$	$\lvert V''(\phi)\rvert \leq 9H^2$, and
		$\lvert V'(\phi)m_{Pl}/V(\phi)\rvert \leq (48\pi)^{1/2}$
Roll Far	Quantum Fluctuations	$\phi_e - \phi_s \gg H$
Enough	Sufficient e-folds	$N = \int H dt \geq 58$
Small	Scalar Perturbations	$(H^2/\dot{\phi}) \leq 10^{-4}$
Perturbations	Tensor Perturbations	$H/m_{Pl} \leq 10^{-4}$
Exit Properly	Nucleosynthesis	$T_{\mathrm{RH}} \geq 1$ MeV
	Baryogenesis	$T_{\mathrm{RH}} \geq 10^{-10}\epsilon^{-1}m_\phi$
	Gravitinos	$T_{\mathrm{RH}} \leq 10^9$GeV

Table 1: Possible problems and solutions in new inflation

with a Coleman-Weinberg potential of the form in Eq. 2.26, with ϕ representing the magnitude of the complex field, and $B = O(10^{-14})$. It must be assumed that the couplings of the ϕ to all other fields in the theory are less than about 10^{-7} to prevent quantum corrections from spoiling the smallness of B. The real part of ϕ will be the inflaton, and the imaginary part of ϕ will be the axion. ϕ couples to the adjoint Higgs, and induces SU_5 breaking when it receives a VEV. This requires $\sigma = 10^{18}$GeV. Since B is so small (and will remain small after radiative corrections), the problems with the original Coleman-Weinberg SU_5 model vanish. The reheat temperature is barely high enough to produce a baryon asymmetry through the decay of the inflaton as discussed above. At the expense of introducing a small number, the model is simple and it works.

An example of a supersymmetric model that works was proposed by Holman, Ramond, and Ross [25]. The superpotential in their model has a "inflation sector" with superpotential $I = (\Delta^2/M)(\phi - M)^2$, where $M = m_{Pl}/(8\pi)^{1/2}$. The scalar potential in supersymmetric models is typically an expansion in ϕ/M, given in this case by

$$V(\phi) = \Delta^4(1 - 4\phi^3/M^3 + 6.5\phi^4/M^4 - 8\phi^5/M^5 + \cdots\ ). \tag{2.32}$$

For $\Delta/M \simeq 10^{-4}$, ($\Delta \simeq 2 \times 10^{14}$GeV), density fluctuations are small enough and sufficient e-folds obtain. The decay width of the ϕ (which has only gravitational coupling to other fields) is $\Gamma_\phi \simeq \Delta^6/M^5$, which for Δ small enough to satisfy the perturbations constraint, leads to $T_{\mathrm{RH}} \simeq (\Gamma_\phi m_{Pl})^{1/2} \simeq 10^6$GeV. With the baryon asymmetry produced via inflaton decay, this is large enough. At the expense of the introduction of a sector whose sole purpose is inflation, the model is simple and it works.

Both the above models have two potential problems. The first problem is that to this point the calculations of the evolution of the scalar field have been semi-classical. It may be that a true quantum calculation of the evolution of ϕ, including production of density perturbations, will give a result much different than the semi-classical result. Preliminary work on this problem suggests that the semi-classical approximations are reasonable. The second potential problem has to do with the initial value of ϕ. Both fields are extremely weakly coupled and are unlikely ever to be in LTE. There is no reason to assume $\phi \simeq 0$ for an initial condition (in fact, it may not even be the high-temperature minimum for the supersymmetric example). It is tempting to say that this is not a problem, and that it is only necessary for $\phi \simeq 0$ in some region of the Universe where the kinetic contributions to ρ are small enough to start inflation.

The above two models are existence proofs that it is possible to implement new inflation. Whether new inflation is the final answer will be discussed briefly after mentioning some other approaches for inflation that do not involve SSB.

For weakly coupled scalar fields there is no reason to believe the inflaton will be in LTE at high temperature, and the value of ϕ at high temperature might be random (hence the name "chaotic inflation"). Imagine a simple scalar potential of the form $V(\phi) = \lambda\phi^4$, with minimum at $\langle\phi\rangle = 0$. Assume as initial conditions that $\phi = \phi_0 \neq 0$, and that ϕ is sufficiently smooth in a large enough region to inflate. The number of e-folds of inflation is

$$N(\phi \to 0) = \int_\phi^0 H\,dt \simeq \pi\left(\frac{\phi}{m_{Pl}}\right)^2. \tag{2.33}$$

In order to obtain 58 e-folds of inflation, $\phi_0 \geq 4.3 m_{Pl}$. The density perturbations are

$$\left(\frac{\delta\rho}{\rho}\right)_H \simeq \left(\frac{3H^3}{V'(\phi)}\right) \simeq \lambda^{1/2}\left(\frac{\phi}{m_{Pl}}\right)^3 \simeq \lambda^{1/2} N(\phi \to 0)^{3/2}. \tag{2.34}$$

Again, using $N = 50$, λ must be smaller than about 10^{-14} for sufficiently small density perturbations. Since Linde [26] originally proposed this model several refinements have been made. First, it has been shown that it is possible to use a $m^2\phi^2$ potential rather than a $\lambda\phi^4$ potential. Some work has been done in examining and formalizing what exactly is meant by "chaotic" initial conditions, and which regions of phase space will inflate. Linde's model is an example of how general inflation is, and that it is possible, perhaps even desirable, to separate inflation from SSB phase transitions. Chaotic inflation (at least for the $\lambda\phi^4$ case) has the possible problem of using classical gravity in the regime $\phi \geq m_{Pl}$. At present it also has the undesirable feature of involving the dynamics of a scalar field introduced for the sole purpose of inflation.

A model even further from the original idea of an SSB phase transition is a pure gravity model based upon including an ϵR^2 term in the gravity Lagrangian. Such higher-derivative terms are expected to be present in theories with extra dimensions. Mijić, Morris, and Suen [27] have examined this possibility in detail, including questions of density perturbations and reheating and find that all constraints can be met for $10^{11} \leq \epsilon^{-1/2} \leq 10^{13}$GeV.

Yet further from the original idea of inflation is the possibility that the inflaton is related to the size of extra dimensions. This will not be discussed here. A possibility discussed elsewhere in this volume is the role of quantum gravity and the program of the "wave function of the Universe."

In a Universe without inflation, the space of initial conditions that give the Universe we observe is a set of measure zero. The inflationary Universe enlarges the space of initial data that will lead to the observable Universe. However, it does not imply that every imaginable set of initial data will lead to inflation. A trivial example is a closed Universe that becomes curvature dominated, and collapses before the vacuum energy dominates and causes inflation. The question "is inflation inevitable" has not yet been completely answered. Inflation may be the final answer, part of the final answer, or none of the final answer. This is discussed further in other lectures [28].

If inflation did occur there are two general predictions. The first prediction is that Ω is very close to 1. It would be hard to imagine that *exactly* 58 e-folds of inflation occurred. With all models that give small density perturbations, the number of e-folds of inflation is enormous, and the intrinsic curvature will only appear on scales far larger than our present horizon. Of course, scale-free density perturbations would appear on the horizon today, so a $(\delta\rho/\rho) \simeq 10^{-4}$ would lead to $\Omega = 1 \pm 10^{-4}$. The second prediction is that of scale-free density perturbations. At present there is no convincing data to support either prediction. Dynamical measurements of Ω seem to give $\Omega = 0.1 \rightarrow 0.3$. This has (at least) three possible explanations. Either there are systematic uncertainties in all the measurements, there is unclustered matter (like massless particles) that give the unseen part of Ω, or there is a present vacuum energy that can account for spatial flatness (the actual prediction of inflation) and $\Omega \neq 1$. None of these explanations are compelling. If the recent determination of the velocity field on large-scales are correct, it is evidence against a scale-free spectrum. Possible ways out are the measurements are wrong, cosmic strings, and double inflation.

The last point is that some explanation must be found for the present smallness of the cosmological constant.

ACKNOWLEDGEMENTS

I would like to thank Frank Accetta, Richard Holmam, and So-Young Pi for carefully looking through these notes. The section on inflation has been strongly influenced by Michael Turner's views and writings on the subject [29]. This work has been partially supported by the Department of Energy and the National Aeronautics and Space Administration.

REFERENCES

1. D. A. Kirzhnits and A. D. Linde, Sov. Phys. JETP **40**, 628 (1974).
2. L. Dolan and R. Jackiw, Phys. Rev. D. **9**, 3320 (1974); S. Weinberg, Phys. Rev. D. **9**, 3357 (1974).
3. S. Coleman, Phys. Rev. D **15**, 2929 (1977); C. Callan and S. Coleman, Phys. Rev. D **16**, 1762 (1977); S. Coleman and F. De Luccia, Phys. Rev. D **21**, 3305 (1980). The tunneling rate is associated with a classical motion in imaginary time because the decay rate is related to the imaginary part of the energy. This is because the wave function oscillates in the classically allowed region, but is exponentially damped in the classically forbidden region.
4. A. Linde Nucl. Phys. **B216**, 421 (1983).
5. Ya. B. Zel'dovich, I. Yu. Kobzarev, and L. B. Okun, Sov. Phys. JETP **40**, 1 (1975).
6. T. W. B. Kibble, J. Phys. A **9**, 1387 (1976). This is an excellent paper that is required reading in the subject.
7. A. Vilenkin, Phys. Rep **121**, 263 (1985). This is a detailed and well-written review that contains many details not included here.
8. H. B. Nielsen and P. Olesen, Nucl. Phys. **B61**, 45 (1973).
9. See D. Eardley, this volume.

10. G. 't Hooft, Nucl. Phys. **B79**, 276 (1974); A. M. Polyakov, JETP Lett. **20**, 194 (1974).

11. J. Preskill, Ann. Rev. Nucl. Part. Sci. **34**, 461 (1984).

12. Preskill, Phys. Rev. Lett. **43**, 1365 (1979).

13. A. Guth, Phys. Rev. D **23**, 347 (1981).

14. A. Albrecht and P. Steinhardt, Phys. Rev. Lett. **48**, 1220 (1982).

15. A. Linde, Phys. Lett. **108B**, 389 (1982).

16. It is crucial to remember that $m_\phi \equiv \partial^2 V(\phi)/\partial\phi^2$ is a function of ϕ, and will be small in the flat region of the potential.

17. There should be no confusion between the sub-H which indicates the quantity is to be evaluated at the time of horizon crossing, and the expansion rate H.

18. A. Starobinsky, Phys. Lett. **117B**, 175 (1982).

19. J. M. Bardeen, P. Steinhardt, and M. S. Turner, Phys. Rev. D **28**, 679 (1983).

20. A. Guth and S.-Y. Pi, Phys. Rev. Lett. **49**, 1110 (1982).

21. S. Hawking, Phys. Lett. **115B**, 295 (1982).

22. P. Steinhardt and M. S. Turner, Phys. Rev. D **29**, 2162 (1984).

23. Q. Shafi and A. Vilenkin, Phys. Rev. Lett. **52**, 691 (1984).

24. S.-Y. Pi, Phys. Rev. Lett. **52**, 1725 (1984).

25. R. Holman, P. Ramond, and G. Ross, Phys. Lett. **137B**, 343 (1984).

26. A. D. Linde, Phys. Lett. **129B**, 177 (1983).

27. M. B. Mijić, M. S. Morris, and W.-M. Suen, "The R^2 Cosmology - Inflation Without a Phase Transition," Caltech Report CATT-68-1320, (Feb. 1986).

28. See the contributions of J. Hartle and J. Barrow in this volume.

29. M. S. Turner, in *Proceedings of the 1984 Cargese School on Fundamental Physics and Cosmology*, ed. J. Audouze and J. Tran Thanh Van (Editions Frontieres, Gif-Sur-Yvette, 1985).