

Lecture notes on Cosmology (ns-tp430m)

by Tomislav Prokopec

Part III: Cosmic Inflation

In this and in the subsequent chapters we use natural units, that is we set to unity the Planck constant, $\hbar = 1$, the speed of light, $c = 1$, and the Boltzmann constant, $k_B = 1$. We do however maintain a dimensionful gravitational constant, and define the reduced Planck mass and the Planck mass as,

$$M_{\text{P}} = (8\pi G_N)^{-1/2} \simeq 2.4 \times 10^{18} \text{ GeV}, \quad m_{\text{P}} = (G_N)^{-1/2} \simeq 1.23 \times 10^{19} \text{ GeV}. \quad (1)$$

Cosmic inflation is a period of an accelerated expansion that is believed to have occurred in an early epoch of the Universe, roughly speaking at an energy scale, $E_I \sim 10^{16} \text{ GeV}$, for which the Hubble parameter, $H_I \sim E_I^2/M_{\text{P}} \sim 10^{13} \text{ GeV}$.

The dynamics of a homogeneous expanding Universe with the line element,

$$ds^2 = dt^2 - a^2 d\vec{x}^2 \quad (2)$$

is governed by the FLRW equation,

$$\frac{\ddot{a}}{a} = -\frac{4\pi G_N}{3}(\rho + 3\mathcal{P}) + \frac{\Lambda}{3}, \quad (3)$$

from which it follows that an accelerated expansion, $\ddot{a} > 0$, is realised either when the active gravitational energy density is negative,

$$\rho_{\text{active}} = \rho + 3\mathcal{P} < 0, \quad (4)$$

or when there is a positive cosmological term, $\Lambda > 0$. While we have no idea how to realise Λ in laboratory, a negative gravitational mass could be created by a scalar matter with a nonvanishing potential energy. Indeed, from the energy density and pressure in a scalar field in an isotropic background,

$$\begin{aligned} \rho_{\varphi} &= \frac{1}{2}\dot{\varphi}^2 + \frac{1}{2}(\nabla\varphi)^2 + V(\varphi) \\ \mathcal{P}_{\varphi} &= \frac{1}{2}\dot{\varphi}^2 + \frac{1}{6}(\nabla\varphi)^2 - V(\varphi) \end{aligned} \quad (5)$$

we see that

$$\rho_{\text{active}} = 2\dot{\varphi}^2 + (\nabla\varphi)^2 - 2V(\varphi), \quad (6)$$

can be negative, provided the potential energy is greater than twice the kinetic plus gradient energy, $V > \dot{\varphi}^2 + (\nabla\varphi)^2$, $\nabla = \vec{\partial}/a$. Up to date no experimental proposals have been put forward, based on which one would realise in laboratory setting a matter with a measureable negative gravitational mass.

The energy of a static gravitational field is negative, and is made more negative by bringing a point mass m into the system. For example, bringing a test particle of mass m into gravitational field of a point mass M reduces the gravitational energy by

$$m\phi_N = -G_N \frac{mM}{r} < 0, \quad (7)$$

such that, in principle, energy can be supplied by gravitational fields for building matter fields. (The general relativistic expression of this is of course the covariant stress-energy conservation, $\nabla_\mu T^{\mu\nu} = 0$.) This lends support to the possibility that the total energy of the Universe, that is the energy of matter fields plus gravitational fields, may be zero, such that the Universe might have been created from nothing by a tunneling event. We know of no way of acquiring information about whether such a creation event took place and in what state the Universe was after the creation. Yet we are free to speculate. If we take the possibility of cosmic inflation seriously, than one should consider more seriously primordial states which could have lead to cosmic inflation. The Liouville hypothesis tells us that, whatever the initial state was, it is quite likely that large sections of the available phase space were exploited (the principle of maximising entropy). Hence, *primordial chaos* may be a fair name for such a state. The question is then, under what conditions could have such a primordial chaos lead to cosmic inflation. Andrei Linde has observed that as long as the energy density of such a state is much below the Planck energy density, such that we can trust field theoretical descriptions we are familiar with, the Universe would begin accelerating, provided the active gravitational mass is negative in a sufficiently big patch. This can be seen by considering the local FLRW equations,

$$\begin{aligned} H^2 &\equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{\rho}{3M_P^2} + \frac{\Lambda}{3} - \frac{\mathcal{K}}{a^2} && \text{(Hawking - Ellis)} \\ \frac{\ddot{a}}{a} &= -\frac{1}{6M_P^2}(\rho + 3\mathcal{P}) + \frac{\Lambda}{3} - \frac{1}{3}\nabla_i\left(\frac{\nabla_i\mathcal{P}}{\rho + \mathcal{P}}\right) && \text{(Raychaudhuri),} \end{aligned} \quad (8)$$

where $H = H(\vec{x}, t)$, $a = a(\vec{x}, t)$, $\rho = \rho(\vec{x}, t)$, $\mathcal{P} = \mathcal{P}(\vec{x}, t)$ and $\mathcal{K} = \mathcal{K}(\vec{x}, t)$ are all functions of space and time, and $\nabla_i = \partial_i/a$. Cosmic inflation starts in a patch, $V \gg V_H = (4\pi/3)H^{-3}$, around a space-time point $\{\vec{x}, t\}$, provided active gravitational energy is negative in that patch, and provided local spatial curvature \mathcal{K} is not large and positive, such that $\ddot{a} > 0$ in the patch. This may be realised for example by dominance of the potential energy of a scalar field over the gradient and kinetic energy of all other fields in that patch. How likely is that to happen, is subject of a debate, and we shall not attempt to estimate it here. We remark in passing that, smaller the energy density in comparison to the Planck energy density, smaller is the likelihood. Since kinetic and gradient energy redshift more rapidly than potential energy, once inflation starts, it will continue, at least for some time.

A. Problems inflation can solve

Since gravitation is an attractive force, it has an inherent stability problem: the Hamiltonian is not bounded from below. This makes it very hard to find within the standard theory of gravitation a natural explanation for the observed homogeneity of the Universe on large scales (Einstein's Cosmological Principle), unless one postulates that the Universe began in a very special state with a large degree of homogeneity. Cosmic inflation naturally provides a natural explanation for this, and for several other puzzling facts about the Universe, which we now list:

1. *Why is the Universe so big and old*
2. *Large scale homogeneity and isotropy*
3. *The horizon problem*
4. *The flatness problem*
5. *Cosmological relics*
6. *Cosmological perturbations*

We shall now discuss in some detail each of these issues.

1. *Why is the Universe so big and old*

The visible Universe contains on average about 1 baryon per 4 m^3 and 412 photons per cm^3 , which implies that there are about 4×10^{88} photons and 2×10^{78} baryons in the Hubble volume, $V_H = (4\pi/3)R_H^3 \simeq 9 \times 10^{78} \text{ m}^3$, where

$$R_H = \frac{c}{H_0} = 4200 \pm 200 \text{ Mpc} \simeq 1.3 \times 10^{26} \text{ m} \quad (9)$$

denotes the Hubble radius today. The age of the Universe is (WMAP 2003),

$$t_0 = 13.7 \pm 0.2 \text{ Gy}. \quad (10)$$

On the other hand, we expect that the Universe began with 'natural' initial conditions, characterised by a state with a large entropy, and hence inhomogeneous. Due to the attractive nature of gravitation, initial inhomogeneities tend to grow, such that large density regions are enhanced, while low density regions are rarefied. This scenario suggests that these 'natural' initial conditions would lead to a Universe which quickly recollapses, at least in overdense regions. If a typical energy density is of the order ρ_{in} ,

then we expect that the recollapse takes place on a time scale given by, $t \sim t_H$, where $H \sim \rho_{\text{in}}^{1/2}/M_P$. Given that we expect ρ_{in} to be quite large, t_H very short in comparison with the actual age of the Universe (10), defining the age problem. The question why is the Universe so big is closely related to the age problem, since the Hubble radius (9), $R_H \sim ct_0$.

The solution provided by inflation is very simple. During inflation the Universe expands by a large factor, which results in a big universe, even if the initial universe was very small. Moreover, inflation smoothens out any initial inhomogeneities, and makes the Universe spatially flat (see sections A 2 and A 4 below), such that the evolution of the Universe on large scales can be well approximated by the homogeneous FLRW equations, leading to a big and old Universe.

2. Large scale homogeneity and isotropy

Figure II.4, which shows density contrast (relative matter density fluctuations), $\delta = \Delta\rho/\rho$, as a function of scale, illustrates nicely how the Universe becomes homogeneous on large scales. In particular, direct observations of large scale structures (galaxies and clusters) by the SDSS and 2dF survey indicate that on the scale of about 200 Mpc, the density contrast,

$$\frac{\delta\rho}{\rho} \sim 10^{-1}, \quad (r \sim 200 \text{ Mpc}), \quad (11)$$

representing a direct evidence for homogeneity of the visible baryonic matter. Since numerical simulations of clustering (driven by the Newtonian dynamics in an expanding Universe) suggest that baryonic matter is a good tracer of dark matter, Eq. (11) can be interpreted as an indirect evidence for homogeneity of the ‘invisible’ baryonic and dark matter on large scales, lending support to the Einstein Cosmological Principle.

On the other hand, temperature anisotropies of cosmic microwave background radiation (CMBR) imply that, at the Hubble scale, inhomogeneities are at the level,

$$\frac{\delta\rho}{\rho} \sim 10^{-4}, \quad (r \sim R_H \simeq 4200 \text{ Mpc}). \quad (12)$$

This can be inferred as follows. The temperature of CMBR is $T_0 = 2.725 \pm 0.001$ K, while the anisotropies are at the level (see figure II.11.-II.13.), $\delta T \sim 30 \mu\text{K}$, such that $\delta T/T_0 \sim 10^{-5} \sim (1/4)\delta\rho/\rho$.

On average our Universe is alike in all directions, and hence it is isotropic. If it is isotropic for all inertial observers, then it is homogeneous. Since we do not have information from other distant observers, strictly speaking we cannot make this inference. Yet, an isotropic but inhomogeneous Universe would imply that we take a special position in the Universe, which is theoretically unappealing. Thus, on

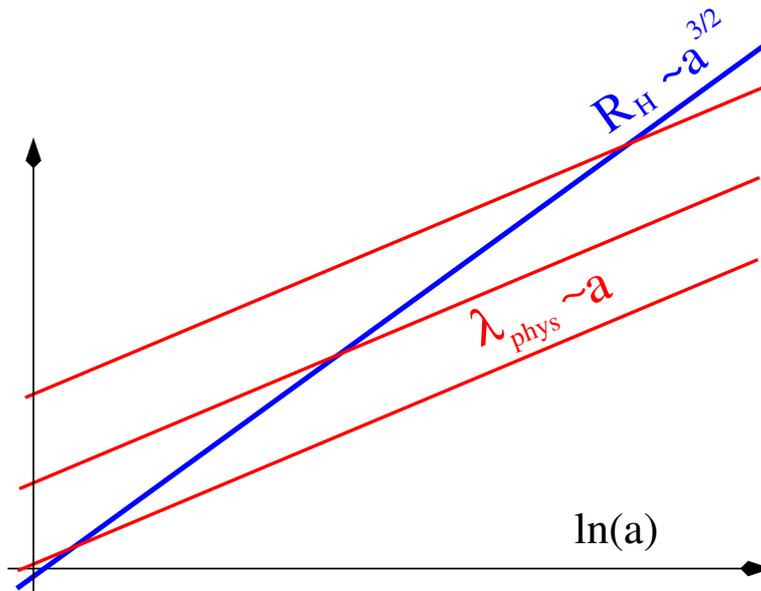


FIG. 1: Evolution of the Hubble radius, $R_H \propto a^{3/2}$, and the physical length, $\lambda_{\text{phys}} \propto a$, in matter era. Since at later times larger and larger scales become visible (since they fall within the Hubble radius), from the point of view of a late time observer, the Hubble scale at an earlier time appears as a fraction of the Hubble scale at a later time, implying the horizon problem.

theoretical grounds, it is quite safe to make the inference that the apparent isotropy of the Universe has as a ‘consequence’ its homogeneity.

The striking homogeneity of the Universe on large scales is hard to reconcile with the standard theory of gravitation, which is attractive and hence leads to a growth of inhomogeneities. A notable exception are small scale inhomogeneities in regions of a significant pressure, for which there is a critical scale, known as the Jeans length, above which inhomogeneities grow, but below which inhomogeneities are damped.

Cosmic inflation solves the homogeneity and isotropy problem as follows. The growth of inhomogeneities is in general slower in an expanding background, due to the Hubble damping term, which tends to limit the growth. Inflation is an extreme case of an accelerated rapid growth, such that for sufficiently distant regions the effect of expansion *wins* over the attraction, and gravitation becomes effectively repulsive. One is even tempted to speak of an effective ‘antigravitation’, which is in this case induced by a rapidly expanding space-time background.

3. Horizon problem

The horizon problem may be stated as follows. When an astronomer observes the early Universe (for example, by observing CMB photons), if the information comes to him/her uninterrupted, he or she

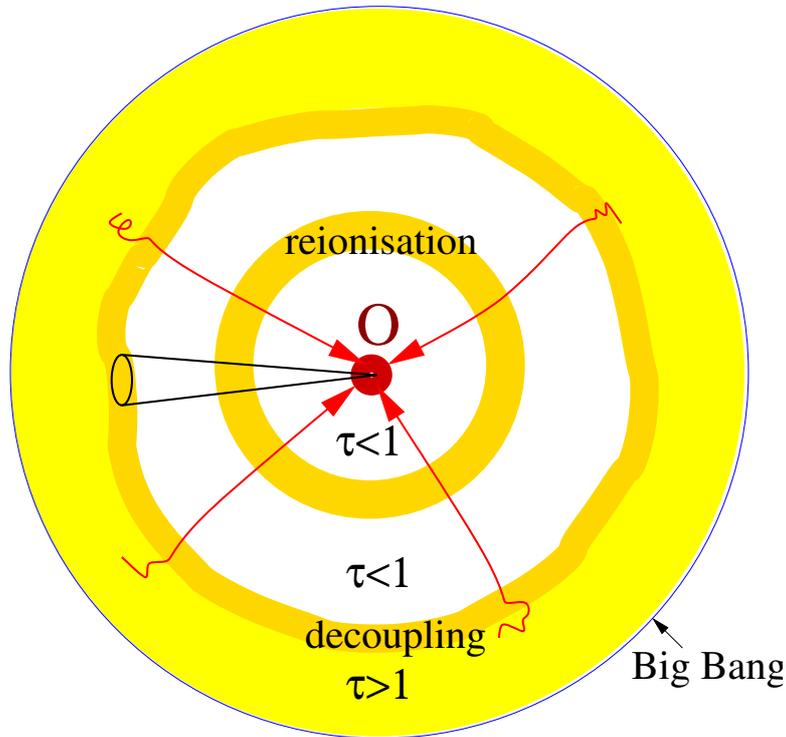


FIG. 2: A conformal diagram of an expanding Universe (see also figure II.64) . The observer O is placed at the origin, $z = 0$. The Universe begins with a hot Big Bang, represented by the large outer circle at a redshift, $z \rightarrow \infty$. Before decoupling, $z \gg z_{\text{dec}} \simeq 1089$, photons rapidly Compton-scatter, and the optical depth, $\tau \gg 1$. From the decoupling epoch photons stream mostly freely, $\tau \ll 1$. At $z_{\text{reion}} \sim 17$ the Universe reionises, and about 17% photons re-scatter. The horizontal cone represents the horizon at decoupling as seen by the observer today.

may be able to see regions that were naïvely out of causal contact at the time of emission. Why this is so can be clearly seen in figure 1, where we show the Hubble radius R_H and the physical length λ_{phys} as a function of the scale factor a during matter era. Since in matter era, $R_H \propto a^{3/2}$, while $\lambda_{\text{phys}} \propto a$, as time passes, larger and larger scales come within the Hubble volume. Hence, from the point of view of a later time observer, when measured in the units of physical scale, the early Universe Hubble radius appears smaller, and he or she will be able to acquire information about the scales which were, from the point of view of an earlier observer, beyond the Hubble radius, and hence naïvely beyond causal contact.

This situation is illustrated by the conformal diagram of the Universe in figure 2, in which an observer O regards CMB photons carrying information about the state of the Universe at decoupling, $z_{\text{dec}} = 1089$. From the point of view of a late time observer, the Hubble radius at the time of decoupling, denoted by the ellipse on the last scattering surface, comprises only a small fraction of the entire sky.

In order to be more quantitative, let us consider a homogeneous FLRW conformal metric, describing

a Universe in which both radiation and (nonrelativistic) matter play a role, such that the scale factor is given by,

$$a(\eta) = a_1\eta + a_2\eta^2, \quad (13)$$

where η denotes conformal time. This represents a good description of the Universe during radiation and matter era, up to the recent epoch of acceleration. Since decoupling of photons from electrons occurs in matter era, to a reasonably good approximation we have, $a_{\text{dec}} \simeq a_2\eta_{\text{dec}}^2$.

When calculated at decoupling, the comoving particle horizon is then,

$$\ell_{\text{dec}} \simeq \int_0^{t_{\text{dec}}} \frac{dt'}{a(t')} = \eta_{\text{dec}} \simeq \ell_0 \sqrt{\frac{a_{\text{dec}}}{a_0}}, \quad (14)$$

where ℓ_0 and a_0 denote the particle horizon and scale factor today, respectively. On the other hand, the comoving Hubble radius today is,

$$R_{Hc} = \frac{1}{\mathcal{H}_0} \simeq \frac{\eta_0}{2} = \frac{\ell_0}{2}. \quad (15)$$

This implies that the number of causally disconnected regions in the entire sky ($\Omega = 4\pi$) seen by an observer today, who has access to all photons coming towards him/her from within the Hubble volume, is

$$\frac{N_{\text{causal}}}{V_H} \sim 4\pi \left(\frac{R_{Hc}}{\ell_{\text{dec}}} \right)^2 \simeq \pi \left(\frac{\ell_0}{\ell_{\text{dec}}} \right)^2 \simeq \pi z_{\text{eq}} \sim 3 \times 10^3. \quad (16)$$

This then implies that, in a universe filled with radiation and matter, an observer today sees in the entire microwave background sky about 3×10^3 regions that were causally disconnected at the time of decoupling. A typical region that corresponds to a causally connected patch at the last scattering surface is illustrated by the ellipse on the last scattering surface shown in figure 2.

The flatness problem can then be stated as follows: How is it possible that regions in the microwave background sky, which are seemingly out of causal contact, have the same temperature up to a very high accuracy of, $\delta T/T \sim 10^{-5}$, which is the amplitude of the observed large scale temperature anisotropies?

Inflation provides an elegant solution to the horizon problem, simply by dramatically increasing the particle horizon with respect to the Hubble radius. In order to see how this works in detail, let us consider de Sitter inflation, with the scale factor, $a = -1/(H_I\eta)$, which lasts from $\eta \rightarrow \eta_{\text{in}} < 0$ up to $\eta = \eta_e$.

The comoving particle horizon is then simply,

$$\ell_c(\eta) = \eta - \eta_{\text{in}} = \frac{1}{H_I} \left(\frac{1}{a_{\text{in}}} - \frac{1}{a} \right), \quad (17)$$

where η_{in} and $a_{\text{in}} = -1/(H_I\eta_{\text{in}})$ denote conformal time and scale factor at the beginning of inflation, respectively. The conformal particle horizon approaches a constant during inflation, while the physical

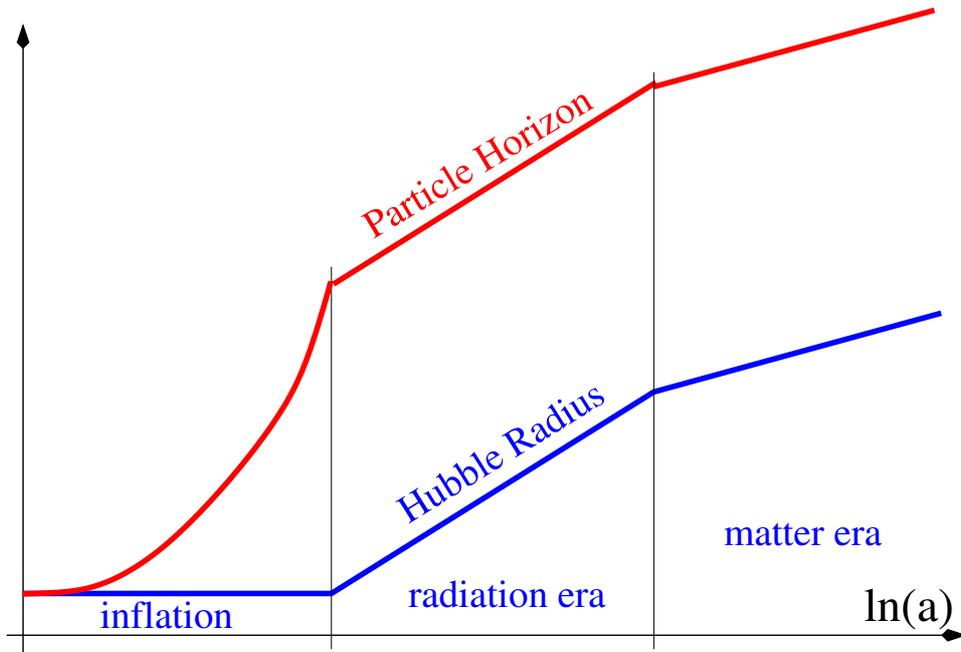


FIG. 3: Evolution of the Hubble radius and the particle horizon during de Sitter inflation, radiation and matter era. During inflation the Hubble radius is constant, it grows as $R_H \propto a^2$ during radiation era, and $R_H \propto a^{3/2}$ during matter era. On the other hand, the particle horizon grows as, $\ell_{\text{phys}} \propto a$ (inflation), $\ell_{\text{phys}} \propto a^2$ (radiation era) and $\ell_{\text{phys}} \propto a^{3/2}$ (matter era).

particle horizon,

$$\ell_{\text{phys}} = \frac{1}{H_I} \left(\frac{a}{a_{\text{in}}} - 1 \right) \quad (18)$$

grows linearly with the scale factor, $a = a_{\text{in}} e^{H_I(t-t_{\text{in}})}$. On the other hand, the Hubble radius remains constant, $R_H = 1/H_I$, such that their ratio grows exponentially with time (*cf.* figure II.35 and Eqs. II.117–II.121),

$$\frac{\ell_{\text{phys}}}{R_H} = \frac{a}{a_{\text{in}}} - 1, \quad (\text{de Sitter inflation}). \quad (19)$$

This is to be contrasted with the corresponding ratio during a pure radiation and matter era, which is of the order of unity,

$$\frac{\ell_{\text{phys}}}{R_H} \simeq 2, \quad (\text{matter era}) \quad (20)$$

$$\frac{\ell_{\text{phys}}}{R_H} \simeq 1, \quad (\text{radiation era}). \quad (21)$$

In figure 3 we show evolution of the physical particle horizon and the Hubble radius as a function of the scale factor during inflation, radiation and matter era. Note that, due to its exponential growth with time, the particle horizon is typically much bigger today and at decoupling than the horizon radius today. Since the particle horizon defines the region that can be in causal contact, when radiation era is preceded by a period of inflation, the complete last scattering surface as seen by an observer today can be in causal contact, resolving thus the horizon problem.

The number of e-foldings (or bouts) of inflation required to resolve the horizon problem is simple to estimate. (The number of e-foldings is defined as, $N(t) = \ln(a_e/a(t))$, where a_e denotes the scale factor at the end of inflation, and $a = a(t)$ is the scale factor at a time t .) Since $\ell_{\text{dec}}/\ell_0 \simeq 1/33$, we need to increase the particle horizon (with respect to the Hubble radius) during inflation by a factor, $a/a_{\text{in}} = e^N \sim 33$. This means that the number of e-foldings required to resolve the horizon problem is,

$$N \geq \ln\left(\frac{\ell_0}{\ell_{\text{dec}}}\right) \simeq 3.5. \quad (22)$$

4. Flatness problem

Let us consider the Friedmann equation,

$$H^2 = \frac{\rho_{\text{tot}}}{3M_P^2} - \frac{k}{a^2}, \quad \Omega_{\text{tot}} = \frac{\rho_{\text{tot}}}{\rho_{\text{crit}}}, \quad \rho_{\text{crit}} = 3M_P^2 H^2, \quad \rho_{\text{tot}} = \rho_{\text{m}} + \rho_{\gamma} + M_P^2 \Lambda + \rho_Q + \dots \quad (23)$$

where ρ_{m} , ρ_{γ} , ρ_{Λ} , and ρ_Q denote the energy density in matter, radiation, cosmological term and quintessence, respectively. Equation (23) can be recast as,

$$\Omega_k \equiv \Omega_{\text{tot}} - 1 = \frac{k}{a^2 H^2}, \quad (24)$$

where the total energy density is constrained by the CMBR,

$$\Omega_{\text{tot}} = 1.02 \pm 0.02. \quad (25)$$

This then implies for the radius of curvature,

$$R_{\text{curv}} \equiv \frac{1}{\sqrt{|k|}} = \frac{1}{\sqrt{|\Omega_{\text{tot}} - 1|}} \frac{1}{aH}. \quad (26)$$

Making use of $H_0 = 71 \pm 3$ km/s/Mpc, $a(t_0) = a_0 = 1$ and (25), we obtain for the radius of curvature of the Universe (see sections II.B.6 and II.B.9) the following lower bound,

$$\infty \geq R_{\text{curv}} \geq 21 \pm 1 \text{ Gpc}. \quad (27)$$

This implies that the Universe we live in is very flat, in the sense that the sum of angles in a triangle with sides smaller or of the order 20 Gpc does not deviate significantly from 180° (see figure II.36). While the early Universe was much smaller ($R_H \propto t$), the radius of curvature remained the same as today. This discord between the Hubble scale and the curvature scale during the early Universe's epochs comprises the flatness problem. In other words, a causal physics could not have possibly arranged balance between the energy density and the rate of expansion to such a high precision, that the curvature contribution was so tiny.

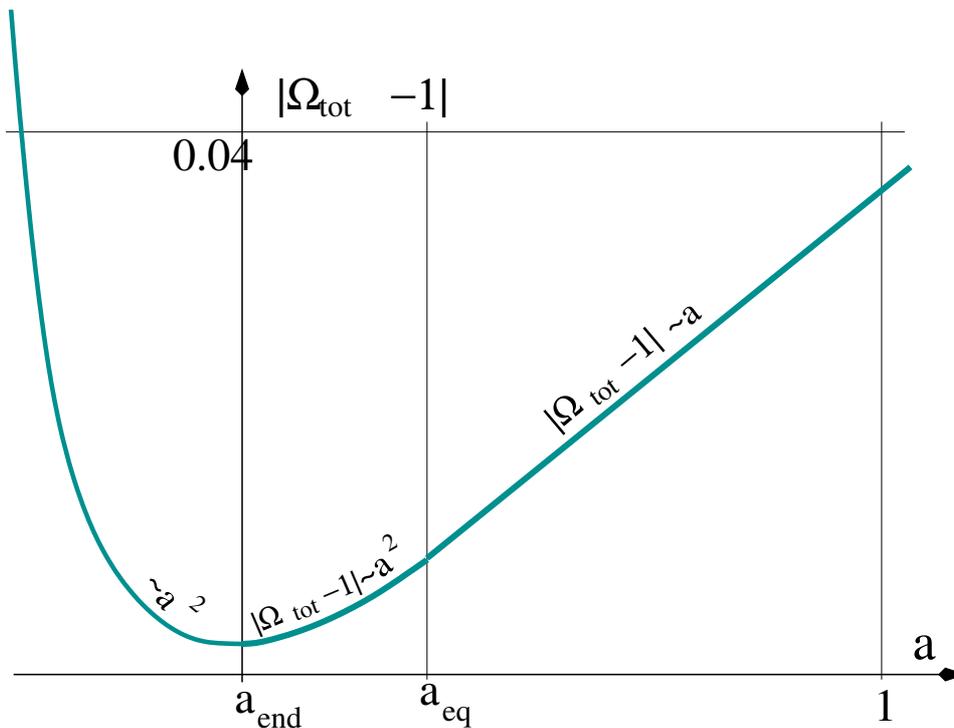


FIG. 4: Evolution of the curvature density, $|\Omega_k| = |\Omega_{\text{tot}} - 1|$, with the scale factor a . In matter era, $|\Omega_k| \propto a$, while in radiation era, $|\Omega_k| \propto a^2$, approaching a tiny value at the beginning of radiation era. When preceded by an inflationary epoch, during which $|\Omega_k| \propto a^{-2}$, $|\Omega_k| = |\Omega_{\text{tot}} - 1|$ can become large at early epochs during inflation, resolving thus the flatness problem. The current experimental bound, $|\Omega_{\text{tot}} - 1| < 0.04$ at $a = a_0 = 1$ is also shown.

To be more quantitative, consider now matter era, during which the scale factor, $a = (t/t_0)^{2/3}$ ($a_0 = 1$), such that, $H = 2/(3t)$ and $Ha = H_0/\sqrt{a}$, where $H_0 = H(t_0)$ denotes the Hubble parameter today. When written in matter era, Eq. (24) then becomes,

$$\Omega_{\text{tot}}(t) - 1 = \frac{k}{H_0^2} a, \quad (\text{matter era}). \quad (28)$$

Similarly, we find that during radiation era, in which $a \propto t^{1/2}$,

$$\Omega_{\text{tot}}(t) - 1 = \frac{k}{(H_{\text{eq}} a_{\text{eq}})^2} \left(\frac{a}{a_{\text{eq}}}\right)^2, \quad (\text{radiation era}), \quad (29)$$

where $a_{\text{eq}} = a(t_{\text{eq}})$ and $H_{\text{eq}} = H(t_{\text{eq}})$ denote the scale factor and Hubble parameter at the matter-radiation equality, respectively. Finally, during de Sitter inflation, $a \propto e^{H_I t}$, and we have,

$$\Omega_{\text{tot}}(t) - 1 = \frac{k}{H_I^2} \frac{1}{a^2}, \quad (\text{de Sitter inflation}). \quad (30)$$

The scaling laws (28–30) and the current bound (25) on $\Omega_{\text{tot}} - 1$ are shown in figure 4 as a function of the scale factor.

The reader has for sure noticed that during early radiation era $|\Omega_{\text{tot}} - 1|$ becomes indeed tiny. Inflation again comes to rescue by reversing the trend, and reducing $\Omega_{\text{tot}} - 1$ during inflation (in the positive time

direction) from some natural value of the order unity at the beginning of inflation to the value given at the end of radiation era. Causality is not violated, since the particle horizon becomes large during inflation.

We now estimate the number of e-foldings (bouts) of inflation N required to solve the flatness problem. N is defined as

$$N(t) = \ln \left(\frac{a_e}{a(t)} \right), \quad (31)$$

where a_e denotes the scale factor at the end of inflation. The flatness problem is considered to be solved, if $\Omega_{\text{tot}} - 1$ becomes of the order unity. We begin by noting that Eqs. (28–29) imply that at the end of inflation,

$$(\Omega_{\text{tot}} - 1)_e = \frac{z_{\text{eq}} + 1}{(z_e + 1)^2} (\Omega_{\text{tot}} - 1)_0 \quad (32)$$

where $z_{\text{eq}} \simeq 3230$ and z_e denotes the redshift at the end of inflation, which can be estimated as follows. The Hubble parameter today is, $H_0 = 2.1332h \times 10^{-33} \text{ eV} \simeq 1.5 \times 10^{-33} \text{ eV}$ ($h = 0.71 \pm 0.03$), while at the end of inflation, $H_I \sim 10^{13} \text{ GeV}$. On the other hand, from the scaling of H is during matter and radiation era, $H \propto a^{3/2}$ and $H \propto a^{3/2}$, respectively, one finds for the redshift at the end of inflation,

$$(z_e + 1)^2 \simeq \frac{H_I}{H_0} (z_{\text{eq}} + 1)^{\frac{1}{2}}. \quad (33)$$

Upon inserting this into (32) we get,

$$(\Omega_{\text{tot}} - 1)_e = \frac{H_0}{H_I} (z_{\text{eq}} + 1)^{\frac{1}{2}} (\Omega_{\text{tot}} - 1)_0 \quad (34)$$

and taking into account (30) and (31), we find for the number of e-foldings required to resolve the flatness problem,

$$N = 63.4 + \frac{1}{2} \ln \left(\frac{10^{-2}}{(\Omega_{\text{tot}} - 1)_0} \right) + \frac{1}{2} \ln \left(\frac{H_I}{10^{13} \text{ GeV}} \right). \quad (35)$$

A natural further question to pose is whether we expect to see any deviation in Ω_{tot} from unity in future measurements, and if yes at what level? There is nothing special about inflation lasting about 60 e-foldings, as it is required to solve the flatness problem. In fact we expect that inflation lasts much longer, which then implies that $|\Omega_{\text{tot}} - 1|$ today should be very close to zero, only if we were able to measure it across the whole universe. Our measurements are however limited to our Hubble volume. We know that within the Hubble volume matter density fluctuates, and that at the Hubble scale matter fluctuations are of the order $\delta\rho_{\text{m}}/\rho_{\text{m}} \simeq 4 \times (\delta T/T)_{\gamma} \sim 5 \times 10^{-5}$. Since matter makes up about 26% of the total energy content of the Universe, and it is reasonable to assume that dark energy does not fluctuate, we find $\delta\rho_{\text{tot}}/\rho_{\text{tot}} \sim 10^{-5}$, which also represents the expected level of fluctuations in Ω_{tot} ,

$$\delta|\Omega_{\text{tot}} - 1| \sim 10^{-5}. \quad (36)$$

Even if we were able to observe deviations in Ω_{tot} from unity at that level, it would not tell us a great deal about inflation, except that this deviation means that locally (in our Hubble volume) the Universe is not perfectly flat. If on the other hand one would find that Ω_{tot} deviates much less or much more from unity than indicated in Eq. (36), that would be surprising, and difficult to explain within the current inflationary paradigm.

5. Cosmic relics

There are different types of relics that may have been produced in the early Universe, for example during thermally induced phase transitions (*cf.* section II.C.3). We shall briefly discuss taxonomy of cosmic relics, how they may have been created, and the role of inflation, if any.

Relics can be conveniently divided into topological defects and particles. Topological defects comprise,

- * domain walls
- * cosmic strings (global and local)
- * cosmic textures
- * magnetic monopoles

Particle relics include,

- * neutrinos and photons
- * gravitons and (their supersymmetric partners) gravitinos
- * axions and (supersymmetric partners) axinos
- * lightest supersymmetric particles (LSP),
- * moduli (Polony) fields

Some of the particle relics are useful in the sense that they represent viable dark matter candidates, *e.g.* LSP, gravitino, axino, *etc.* Primordially created photons, neutrinos and gravitons carry a great deal of information about the early Universe.

In the 1990s, topological defects were considered as an alternative explanation for the origin of large scale structures. Modern CMB measurements have ruled them out as a primary candidate for structure

formation. At this moment no observations support the existence of topological defects in the Universe. Nevertheless, since topological defect represent an interesting theoretical possibility, and they may be observed in the future, they are a valid subject of study in cosmology.

Kibble mechanism

Topological defects can be produced after thermally induced phase transitions by the Kibble mechanism. Quite generally, let us consider a theory given by a Lagrangian which respects a symmetry group G , which is spontaneously broken by the vacuum to G' , such that the symmetry of the vacuum manifold \mathcal{M} is the coset group,

$$\mathcal{M} = G/G'. \quad (37)$$

According to the Kibble mechanism, topological defects form after a phase transition which breaks the symmetry, if one of homotopy groups, π_n ($0 \leq n \leq 3$) of the vacuum manifold is nontrivial. Thus, when $\pi_0(\mathcal{M}) \neq \mathbf{1}$, then domain walls form, when $\pi_1(\mathcal{M}) \neq \mathbf{1}$, cosmic strings form, when $\pi_2(\mathcal{M}) \neq \mathbf{1}$, monopoles form, and finally when $\pi_3(\mathcal{M}) \neq \mathbf{1}$, cosmic textures form. Elements of π_0 are the disconnected parts of the vacuum manifold \mathcal{M} . Elements of the fundamental group π_1 are inequivalent loops that can be constructed around a point P of \mathcal{M} , *i.e.* loops which cannot be deformed to each other by a set of continuous deformations. If all loops on \mathcal{M} at P can be deformed to a point, the fundamental group is trivial, $\pi_1(\mathcal{M}) = \mathbf{1}$. Next, $\pi_2(\mathcal{M})$ is nontrivial if there are homotopically inequivalent 2-spheres at a point P of \mathcal{M} . More generally, $\pi_n(\mathcal{M}) \neq \mathbf{1}$ is nontrivial if there are homotopically inequivalent n -dimensional spheres, S^n , at a point $P \in \mathcal{M}$.

The following theorem may be helpful when calculating homotopy groups. If $\pi_n(G) = \pi_{n-1}(G) = \mathbf{1}$, then

$$\pi_n(G/G') = \pi_{n-1}(G'). \quad (38)$$

In particular, this theorem applies when a simple group G breaks to a group G' . For example, when $O(3) \rightarrow O(2)$, such that $\mathcal{M} = O(3)/O(2) = S^2$. Since $\pi_2(O(3)) = \mathbf{1}$ and $\pi_1(O(3)) = \mathbf{1}$, the theorem (38) applies and we conclude that $\pi_2(\mathcal{M}) = \pi_2(S^2) = \pi_1(O(2)) = \pi_1(S^1) = \mathbb{Z}$.

In order to illustrate the Kibble mechanism, consider an N component real scalar field with the following $O(N)$ symmetric potential,

$$V = \frac{\lambda}{4} \left(\sum_{a=1}^N (\phi^a)^2 - \mu^2 \right)^2. \quad (39)$$

At high temperatures the symmetry is restored, and $\phi^a = 0$ ($\forall a$), such that the symmetry group G of

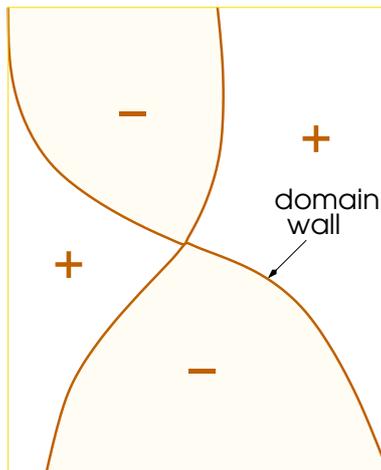


FIG. 5: In the potential with one real scalar field, after the spontaneous symmetry breaking, $\mathbb{Z}_2 \rightarrow \mathbf{1}$, domains of the vacua $\phi = +\mu$ and $\phi = -\mu$ form. Domain walls form at the interface between two different vacua.

the Lagrangian,

$$\sqrt{-g}\mathcal{L} = \sqrt{-g} \sum_{a=1}^N \frac{1}{2} g^{\mu\nu} (\partial_\mu \phi^a) (\partial_\nu \phi^a) - V \quad (40)$$

is

$$G = O(N). \quad (41)$$

Below the critical temperature, the field acquires an expectation value, breaking G spontaneously to G' . At zero temperature, the (classical) vacuum manifold of the theory is given by,

$$\sum_{a=1}^N (\phi^a)^2 = \mu^2, \quad (42)$$

in which the potential energy (39) vanishes, $V|_{\text{vac}} = 0$. Since one of the fields, say ϕ^b , has a definite value in the vacuum, the remaining unbroken symmetry of the vacuum is

$$G' = O(N-1), \quad (43)$$

such that the symmetry of the vacuum manifold is,

$$\mathcal{M} = O(N)/O(N-1) = S^{N-1}. \quad (44)$$

This vacuum manifold has a nontrivial $(N-1)$ -st homotopy group,

$$\pi_{N-1}(\mathcal{M}) = \pi_{N-1}(S^{N-1}) \neq \mathbf{1}. \quad (45)$$

The simplest case is a real scalar field, for which $N = 1$, and the vacuum manifold consists of two points, $\phi = +\mu$ and $\phi = -\mu$, such that

$$\mathcal{M} = S^0 \equiv \mathbb{Z}_2, \quad (46)$$

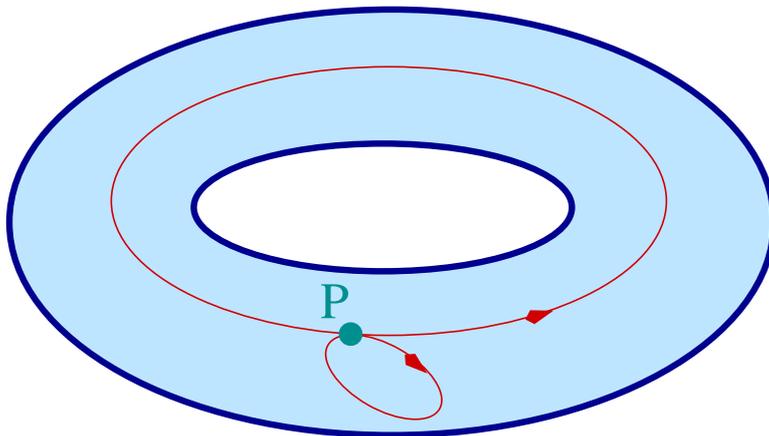


FIG. 6: Two homotopically inequivalent loops on a two dimensional vacuum manifold with a nontrivial fundamental group, π_1 . The smaller loop can be deformed to a point, and hence belongs to the class of homotopically trivial loops. The larger loop is a representative of the class of loops which wind once anticlockwise around the hole of the manifold.

and the homotopy group

$$\pi_0(\mathcal{M}) = \pi_0(S^0) = \{-1, 1\} \quad (47)$$

has two (inequivalent) elements. As a consequence, domains of both vacua, $\phi = +\mu$ and $\phi = -\mu$, form after the symmetry breaking, as illustrated in figure 5. Domain walls form at the interface between the two vacua. Since the symmetry is restored at the center of domain walls, they carry energy, which is characterised by the surface tension (energy per unit area),

$$\sigma = \int dx_{\perp} \rho_{\text{dw}}, \quad (48)$$

where x_{\perp} is the coordinate orthogonal to the wall.

The second example is $N = 2$, which corresponds to a two component real scalar field or, equivalently, to a single complex scalar field. In this case the vacuum manifold is a circle,

$$\mathcal{M} = O(2)/O(1) = S^1, \quad (49)$$

such that the first homotopy group (fundamental group) of the vacuum manifold is

$$\pi_1(\mathcal{M}) = \pi_1(S^1) = \mathbb{Z}. \quad (50)$$

Inequivalent elements of $\pi_1(S^1)$ comprise loops which wind around the circle n times in one or the other direction, such that $\pi_1(S^1)$ is isomorphic to the set of integers, \mathbb{Z} . An example of a two dimensional vacuum manifold with a nontrivial fundamental group is shown in figure 6. Two homotopically inequivalent loops at a point P on \mathcal{M} are shown. The fundamental group of this manifold is isomorphic to

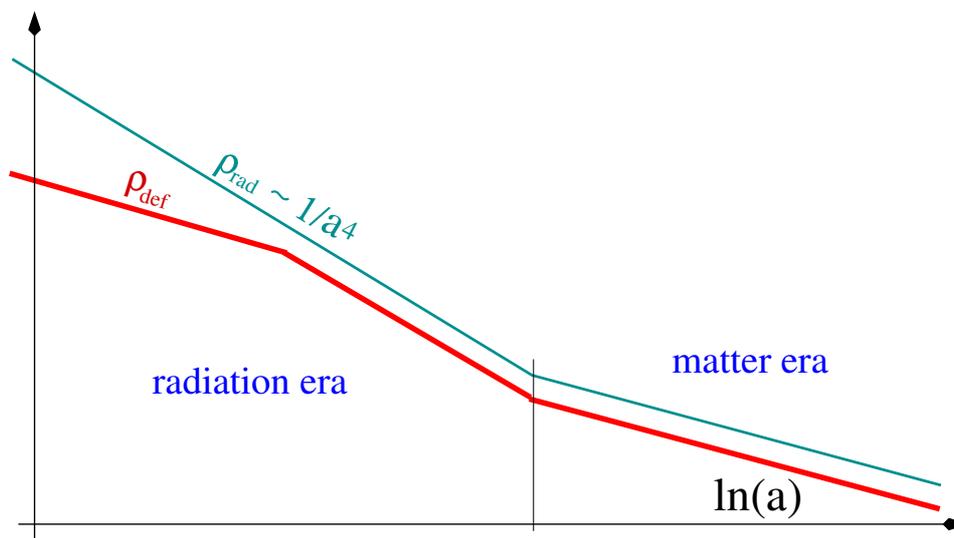


FIG. 7: Energy density in cosmic strings, ρ_{def} , as a function of the scale factor a . At early times, the scaling of defects (strings) differs from the scaling of radiation, $\rho_{\text{rad}} \propto 1/a^4$. After they reach a scaling regime, the scaling of defects becomes identical to that of the dominant energy component.

\mathbb{Z} . The dynamics of cosmic strings (cs) is characterised by the corresponding energy density per unit length, defined as,

$$\mu_{\text{cs}} = \int d^2x_{\perp} \rho_{\text{cs}}. \quad (51)$$

Inflation and topological defects

Inflation dilutes any preexisting relics, including topological defects, to ‘nothingness’, implying that any relics formed before inflation are unobservable. More concretely, during inflation the following scaling applies,

$$\rho_{\text{def}} \propto a^{-n}, \quad (52)$$

with $n = 1$ (domain walls), $n = 2$ (cosmic strings), $n = 3$ (magnetic monopoles and nonrelativistic particle relics) and $n = 4$ (cosmic texture and relativistic particle relics). On the other hand, the energy density driving inflation, ρ_I , scales very slowly during inflation, implying that $\rho_{\text{def}}/\rho_I \sim 1/a^n$ decreases during inflation as a positive power of the scale factor.

Topological defects after inflation

If produced at a phase transition after inflation, cosmic strings and textures reach a scaling solution. The scaling of energy density in cosmic strings is sketched in figure 7, in which strings are initially created in an abundance which does not correspond to a scaling solution. After some time they reach

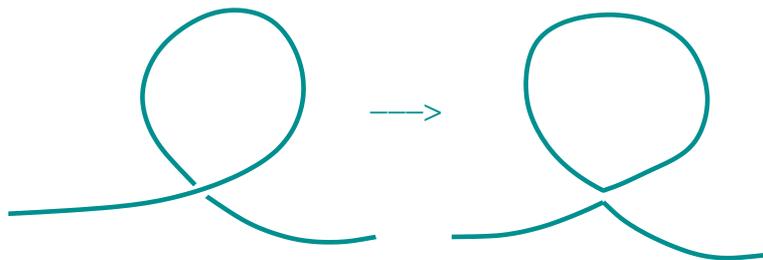


FIG. 8: Intercommuting at a long cosmic string, which is the dominant mechanism for energy loss of cosmic strings. After intercommuting takes place, a small string loop is created. The loop oscillates and loses energy through gravitational radiation.

a scaling solution, during which the energy density in strings, which form string networks, remains a constant fraction of the total energy density,

$$\frac{\rho_{\text{string}}}{\rho_{\text{tot}}} = \text{const.} \quad (\text{scaling solution}). \quad (53)$$

Since the natural scaling of cosmic strings is, $\rho_{\text{string}} \propto 1/a^2$, and the dominant energy component scales faster with a , there must be a mechanism by which strings lose energy during scaling regime. This mechanism indeed exists, and it is known as intercommuting, which we illustrate in figure 8. Strings typically move with relativistic speeds, and their different parts move independently, such that often different parts of the same long string cross each other. When they cross, it is often the case that they intercommute, such that a string loop and a shorter string are formed. The place at which intercommuting takes place usually develops a cusp, and the newly formed loop oscillates, both sourcing gravitational radiation, by which the loop eventually decays. Intercommuting represents the main energy loss mechanism of cosmic string networks, and it is responsible for scaling of energy density of cosmic string networks with the dominant energy component.

Due to the scaling of cosmic string networks and cosmic textures, they will quite generically generate scale invariant fluctuation of matter and radiation. In the 1990s this was considered a very attractive feature of defect networks, and proclaimed as an alternative mechanism for structure formation to cosmic inflation. Since the mechanism by which defects imprint inhomogeneities into cosmic microwave background radiation is causal, photons of CMB do not exhibit acoustic oscillations, which have been observed. Hence topological defects are considered ruled out as the main source density fluctuations in matter and CMB radiation. Since defects naturally form at phase transitions, some cosmologists suggest that they still may contribute to a small but not insignificant fraction of CMB anisotropies.

By failing to explain the observed anisotropies in the CMBR, topological defects have lost their main appeal in cosmology. Nevertheless, they still remain an interesting theoretical possibility.

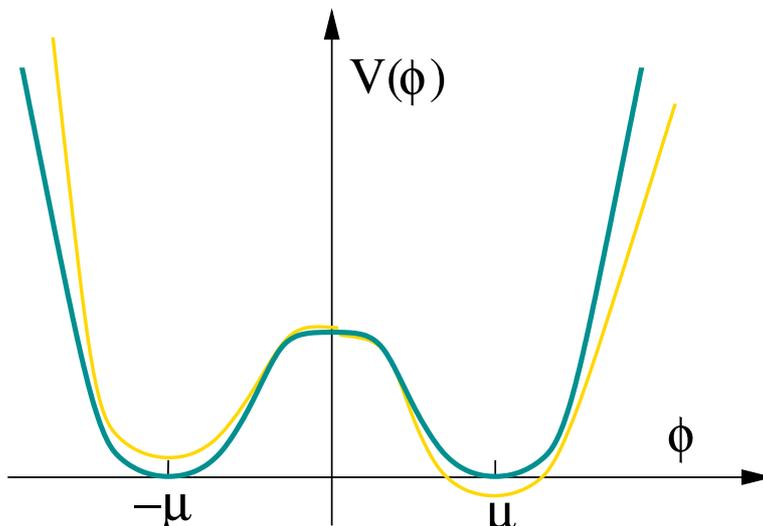


FIG. 9: When the domain wall potential is tilted a little, the degeneracy between the two vacua is broken. This represents a possible solution to the domain wall problem.

Domain walls

Since there is no efficient energy loss mechanism for domain walls, they do not reach scaling solution. Since they scale with the scale factor as,

$$\rho_{\text{dw}} \propto \frac{1}{a} \quad (54)$$

even producing a tiny amount of domain walls at a phase transition in an early Universe would be disastrous, since they would eventually end up dominating the energy density of the Universe, which we know it is not the case. Many solutions have been proposed to the domain wall problem. Probably the simplest one is to break the symmetry between the two vacua by a small amount, as shown in figure 9. The right vacuum tunnels to the left vacuum when the tunneling rate becomes larger than the expansion rate of the Universe, resulting in disappearance of domain walls.

Magnetic monopoles

The original motivation that lead Allan Guth in 1981 to the discovery of inflationary paradigm was to get rid of the monopole problem. Magnetic monopoles form after a thermally induced phase transition where a local gauge group G breaks to a subgroup G' , when

$$\pi_2(G/G') = \pi_2(\mathcal{M}) \neq \mathbf{1}, \quad (55)$$

which is realised for example by a three-component real scalar field, in which case $\mathcal{M} = O(3)/O(2) = S^2$, and $\pi_2(S^2) = \mathbb{Z}$. More realistically, if a grand unified gauge symmetry group, $G = SU(5)$, breaks to the

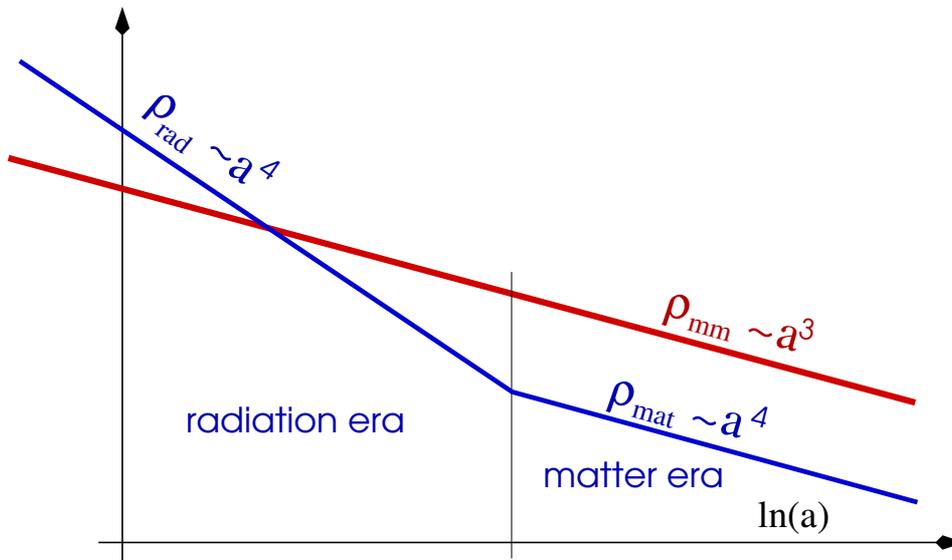


FIG. 10: Scaling of magnetic monopoles during radiation and matter era. If they form at the grand unified phase transition, monopoles eventually dominate the energy density of the Universe, even when one monopole per Hubble volume forms, as required by the causality bound. This defines the monopole problem.

Standard Model group,

$$SU(5) \rightarrow G_{\text{SM}} = SU(3)_c \times SU(2)_W \times U(1)_Y, \quad (56)$$

the second homotopy group of the resulting vacuum manifold, $\mathcal{M} = SU(5)/G_{\text{SM}}$, is nontrivial. This can be easily shown by making use of theorem (38), which applies to this case since $\pi_2(SU(5)) = \mathbf{1}$ and $\pi_1(SU(5)) = \mathbf{1}$. According to (38) we have,

$$\pi_2(\mathcal{M}) = \pi_1(G_{\text{SM}}) = \pi_1(U(1)) = \pi(S^1) = \mathbb{Z}, \quad (57)$$

such that the appearance of an $U(1)$ factor in the symmetry of the Standard Model has as a consequence formation of magnetic monopoles at the grand unified phase transition. Inflation can ‘solve’ the monopole problem, by diluting any preexisting monopoles. The mechanism works, provided the grand unified symmetry is not restored after inflation, which can be achieved by setting the scale of inflation below the scale of grand unification.

The mass of a magnetic monopole (confined within a volume V_m) can be estimated as follows,

$$m_{mm} = \int_{V_m} d^3x V(\phi^a) \sim \frac{4\pi\mu}{g} \sim 10^{17} \text{ GeV}, \quad (58)$$

where $\mu \sim 10^{16}$ GeV is the scale of grand unification, and g is the (grand unified) gauge coupling constant, which is given by $\alpha_g = g^2/(4\pi) \sim 1/30$. When monopoles form, the Universe is typically radiation dominated such that, $\rho \simeq \rho_{\text{rad}} \propto 1/a^4$. On the other hand, since the temperature after the grand unified phase transition is lower than the monopole mass (58), the monopoles are nonrelativistic,

and the corresponding energy density scales as,

$$\rho_{\text{mm}} \simeq \frac{m_{\text{mm}}}{V_{\text{mm}}} \propto \frac{1}{a^3}, \quad (59)$$

where $V_{\text{mm}} \sim \xi_{\text{mm}}^3$ represents the volume in which one monopole forms, and ξ_{mm} is the correlation length, representing the relevant domain size. Note that by a simple causality argument, ξ_{mm} is bounded from above by the Hubble radius, $\xi_{\text{mm}} \leq R_H$, and $V_{\text{mm}} < V_H$. This means that during radiation era, the fraction of energy density in monopoles grows with the scale factor,

$$\frac{\rho_{\text{mm}}}{\rho_{\text{rad}}} \propto a, \quad (60)$$

dominating eventually the energy density of the Universe. We illustrate the monopole scaling in figure 10. Note that during matter era, $\rho_{\text{mm}}/\rho_{\text{mat}} \simeq \text{const.}$ The monopoles that form at the grand unified scale will eventually dominate the energy density before the onset of matter domination, even when the causality bound is saturated, $V_{\text{mm}} \sim V_H$, defining the monopole problem.

6. Cosmological perturbations

Inflation amplifies vacuum fluctuations of the following fields,

- tensor metric fluctuations
- light scalar matter fluctuations and consequently scalar metric fluctuations
- matter and gauge field fluctuations, provided conformal symmetry is broken during inflation
- exotica, which include scalar gravitational fields, dark matter candidates, *etc.*

This amplification is important since scalar metric fluctuations are considered responsible for large scale structure formation of the Universe, as well as for imprinting anisotropies in the CMB radiation. Dark matter and baryons (after decoupling) fall into the potential wells formed by the scalar gravitational perturbations. Since a typical amplitude of potential fluctuations is independent on the size of the region considered, these primordial gravitational potential fluctuations are said to be (nearly) scale invariant. In figure 11 we show how dark matter particles fall into the primordial potential wells formed during inflation, which makes the wells deeper. Baryons are tightly coupled through the Coulomb scattering to electrons, such they start falling into the wells only after decouplings, when most of free electrons become captured in the neutral hydrogen. At that time the epoch of structure formation begins, which ends up in formation of galaxies, stars, *etc.*



FIG. 11: Dark matter particles fall into primordial potential wells formed during inflation, deepening the wells. Before decoupling baryons are tightly coupled (*via* Coulomb scattering) to electrons. After decoupling baryons start falling into the deepened potential wells, triggering structure formation.

Amplitude of cosmological perturbations: a simple estimate

We now present a simple heuristic estimate of the amplitude of cosmological perturbations imprinted on quantum fields during inflation. While the following discussion is very useful for developing of an intuitive picture on how the amplification mechanism works, we warn that a genuine quantitative estimate is possible only by performing quantisation of the relevant fields during inflation. In particular, in order to study the generation of the primordial gravitational potentials, a quantisation of the metric tensor is required.

Let us start by considering a (scalar) field, which can be split into a homogeneous and a fluctuating part as follows,

$$\varphi = \varphi_0(t) + \delta\varphi(\vec{x}, t) \quad (61)$$

When averaged over the Hubble volume, the gradient of the fluctuating part, when evaluated in a FLRW background, is simple to estimate,

$$\nabla\delta\varphi = \frac{\vec{\partial}}{a}\delta\varphi \sim H\delta\varphi, \quad (62)$$

where H is the Hubble parameter during inflation. This just means that, as can be seen in figure 12, the modes of a wavelength $\lambda \sim 2\pi R_H$, principally determine the gradient in Eq. (62).

The gradient energy density is then simply,

$$\rho_{\text{grad}} = \frac{1}{2}(\nabla\delta\varphi)^2 \sim H^2\delta\varphi^2 \quad (63)$$

which implies a gradient energy,

$$E_{\text{grad}} \sim \frac{4\pi}{3}R_H^3\rho_{\text{grad}} \sim \frac{(\delta\varphi)^2}{H}. \quad (64)$$

An analogous analysis for the kinetic energy gives, $E_{\text{kin}} \sim (\delta\varphi)^2/H$, such that, $E \simeq E_{\text{kin}} + E_{\text{grad}} \sim (\delta\varphi)^2/H$.

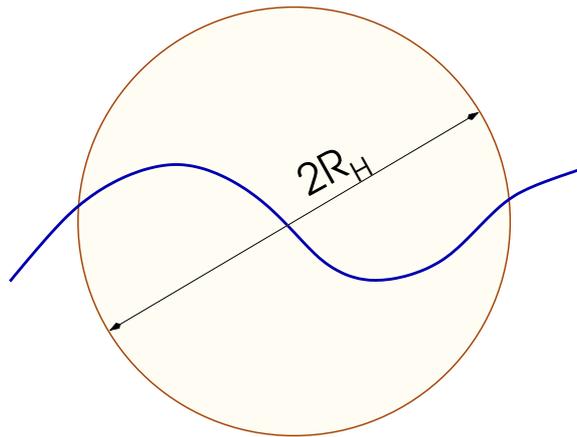


FIG. 12: An example of a mode whose wavelength λ is comparable to the Hubble radius, $\lambda \sim 2\pi R_H$. These are the largest modes that contribute to averaging of fluctuations, $\delta\phi$, over the Hubble volume.

We now make use of the Heisenberg uncertainty principle, according to which energy can fluctuate by an amount $\delta E \sim E \sim (\delta\phi)^2/H$, provided the following relation is not violated,

$$\delta E \delta t \leq 1 \quad (\text{Heisenberg uncertainty relation}), \quad (65)$$

from which it follows,

$$\delta\phi \sim H. \quad (66)$$

Note that, by evoking the uncertainty principle, we made use of quantum physics. It is not possible to make an estimate of the amplitude of field fluctuations (during inflation) without making use of quantum physics. A more accurate estimate, based on the field quantisation, gives for the amplitude of field fluctuations at the horizon scale,

$$\delta\phi \simeq \frac{H}{2\pi}. \quad (67)$$

This type of reasoning, when applied to Minkowski space-time, implies for the amplitude of quantum fluctuations in a field on a scale R to be at the level,

$$\delta\phi|_R \sim \frac{1}{R}. \quad (68)$$

The crucial difference between quantum fields in a Minkowski space-time background and fields in an inflationary space-time is that quantum fluctuations in Minkowski space-time are *virtual*, while fluctuations in inflation can become *real* particles. As we shall see in a moment, the important consequence of this is that on large scales, $R \gg 1/H$, the amplitude of vacuum fluctuations (of nonconformally coupled fields) freezes out in inflation at the level,

$$\delta\phi|_R \simeq \frac{H}{2\pi}, \quad (R \gg R_H = 1/H). \quad (69)$$

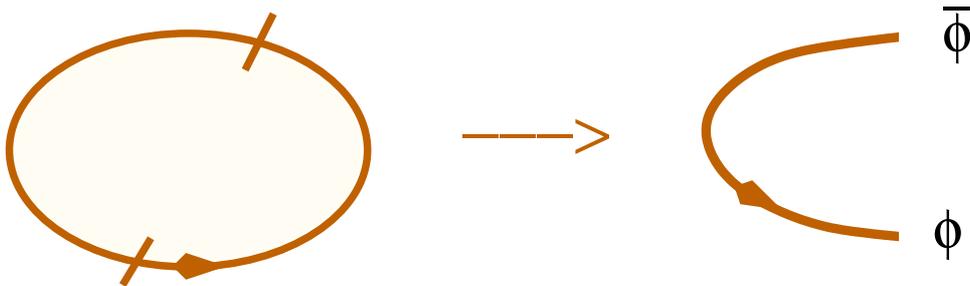


FIG. 13: The one-loop diagram contributing at leading order in perturbation theory to the pair creation amplitude for a field ϕ . Cutting the loop correspond to constraining the loop propagator on the energy shell. While the projected amplitude vanishes in Minkowski space, the diagram can give a contribution to the amplitude for pair creation during inflation, implying that gravitation can assist pair creation and make particle pair excitations out of the vacuum a real process.

In order to decide whether a fluctuation corresponds to a virtual or real particle, let us consider the process of pair creation out of the vacuum, which is diagrammatically described by the one-loop diagram shown in figure 13. ‘Cutting’ the loop propagator corresponds to constraining the propagator onto the energy shell. The amplitude does not vanish if there is an on-shell contribution to the cut diagram, or equivalently the on-shell propagator. A simple way of deciding whether such a contribution exists is to consider the following generalisation of the uncertainty principle,

$$\int_{t_0}^t dt' 2E(k, t') \leq 1, \quad E = \sqrt{(p/a)^2 + m^2}, \quad (70)$$

written for the particle pair creation, each particle carrying an energy E and a physical momentum, $p_{\text{phys}} = p/a$, where p denotes the comoving momentum of a particle, and a the scale factor. We are interested in estimating the time t for which the uncertainty relation (70) saturates. Let us for simplicity consider a massless particle, $m = 0$ (which is the most optimistic case), for which Eq. (70) becomes,

$$2p \int_{t_0}^t \frac{dt'}{a(t')} = 2p(\eta - \eta_0) \leq 1, \quad (71)$$

where η and η_0 denote conformal times. In de Sitter inflation, $a = -1/(H_I \eta)$, Eq. (71) becomes,

$$p \left(\frac{1}{a_0} - \frac{1}{a} \right) = p_{\text{phys}}(t_0) - p_{\text{phys}}(t) \leq \frac{H}{2}. \quad (72)$$

Since during de Sitter inflation the physical momentum gets reshifted exponentially with time, $p_{\text{phys}} \propto 1/a \propto e^{-H_I t}$, after some time Eq. (72) can be reinterpreted as follows. As long as at an initial time, $t = t_0$ (corresponding to the time of pair creation) the physical momentum of a particle is smaller than about,

$$p_{\text{phys}}(t_0) \leq \frac{H}{2}, \quad (73)$$

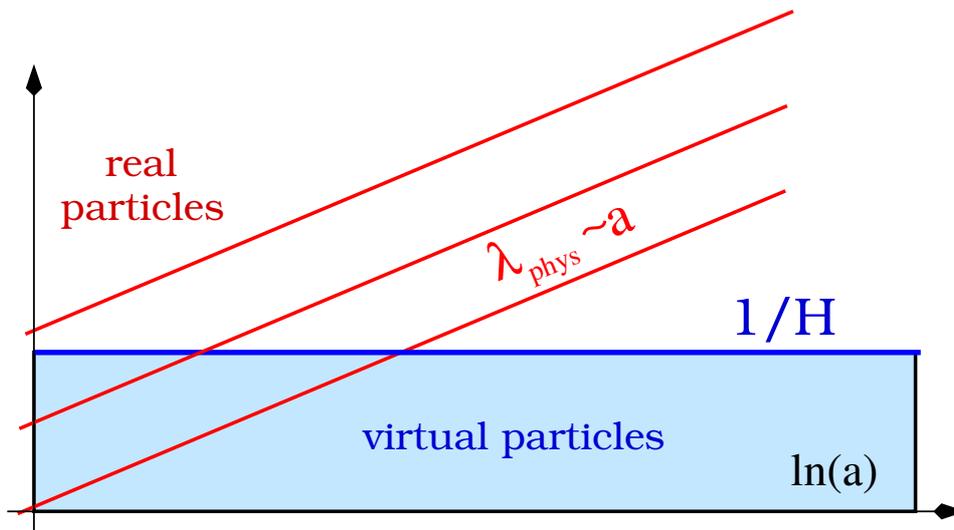


FIG. 14: Evolution of the physical scale in inflation. Perturbations of a massless field on scales smaller than about the Hubble radius, $R_H = 1/H$, are virtual, while perturbations on scales larger than R_H correspond to real particles and live indefinitely.

particle pairs created out of the vacuum by the process in figure 13 live forever during de Sitter inflation, and thus correspond to real particles. This in contrast to particle pairs in Minkowski vacuum, in which particle pairs of momentum p live typically for a time interval, $\delta t \leq 1/(2p)$, before they get annihilated. While Minkowski vacuum does not allow spontaneous creation of real particle pairs out of the vacuum, this process is allowed during inflation because of the energy and momentum redshift.

In figure 14 we show the evolution of physical scales during inflation. Fluctuations whose wave length is smaller than about $R_H = 1/H$ can be considered virtual, while fluctuations whose wavelength is larger than about R_H correspond to real particles.

While this explains what is special about an inflationary space-time when compared with Minkowski space-time, it still does not explain why and when fluctuations freeze out on superhorizon scales, $\delta\varphi|_{R \gg R_H} \simeq H/2\pi$. A heuristic argument can be constructed by considering the Feynman path integral of a massless minimally coupled scalar field, φ , in an expanding space time,

$$\sum_{\varphi} e^{iS[\varphi, a]}, \quad (74)$$

where $S[\varphi, a]$ is the classical scalar action,

$$S[\varphi, a] = \int d^4x \left[a^3 \frac{1}{2} \dot{\varphi}^2 - \frac{a}{2} (\vec{\partial}\varphi)^2 \right], \quad (75)$$

where \sum_{φ} denotes a sum over all field configurations. It is well know that the sum in (74) dominated by the solution to the classical trajectory, given by vanishing of the variation of the classical action,

$$\delta_\varphi S[\varphi_{\text{cl}}, a] = 0, \quad (76)$$

which determines φ_{cl} as the solution of the classical equation of motion,

$$\ddot{\varphi}_{\text{cl}} + 3\frac{\dot{a}}{a}\dot{\varphi}_{\text{cl}} - \frac{\vec{\partial}^2}{a^2}\varphi_{\text{cl}} = 0. \quad (77)$$

At late times, $a \rightarrow \infty$, we can neglect the gradient term, $\vec{\partial}^2\varphi_{\text{cl}}/a^2 \rightarrow 0$, and Eq. (77) is solved by,

$$\dot{\varphi}_{\text{cl}} = \frac{\dot{\varphi}_0}{a^3}, \quad (78)$$

implying that $\dot{\varphi}_{\text{cl}}$ is suppressed as the Universe expands. In other words, the classical solution, ϕ_{cl} , corresponds to a field which freezes out on super-Hubble scales, at which the gradient contribution to the equation of motion (77) can be neglected. Indeed, in de Sitter inflation, $a = a_0 e^{H_I t}$, and on super-Hubble scales, $k/a \ll H_I$, we have $\varphi_{\text{cl}} = \varphi_0 + \dot{\varphi}_0/(3H_I a^3) \approx \varphi_0$, so the dominant contribution to the Feynman path integral is a solution that is approximately constant, explaining thus the fact that the field freezes out.

An exception to this rule are conformally coupled fields. For example, for a conformally coupled scalar field, the action reads,

$$S_c[\varphi, a] = \int d^4x \left[\dot{\varphi}_c^2 - (\vec{\partial}\varphi_c)^2 \right], \quad \varphi_c = a\varphi, \quad (79)$$

such that, just like in Minkowski space, vacuum fluctuations of a conformal scalar field, and in fact any other conformally coupled field, which may be fermions, gauge fields, *etc.*, exhibit the conformal scaling during inflation,

$$\delta\varphi_c|_R \sim \frac{1}{R}, \quad (80)$$

which applies to any scale R . This is identical to the scaling in Minkowski space (68). Note that the above argument, which is based on finding a dominant contribution to the Feynman path integral, fails for a conformally coupled field, because in this case one cannot neglect the gradient term contribution, even on super-Hubble scales. In other words, for conformally coupled fields nothing special happens at the Hubble crossing, because they do not sense it in any way.

To summarise, we have argued that vacuum fluctuations of nonconformally coupled fields get amplified during inflation. The mechanism can be summarised as follows. The amplitude of vacuum fluctuations for nonconformally coupled fields tends to get frozen at the level $\delta\varphi|_R \sim H/2\pi$ ($R \gg 1/H$), while for conformally coupled fields the field fluctuations are not amplified, and exhibit a conformal scaling, $\delta\varphi|_R \sim 1/R$ on all scales, just like in Minkowski vacuum.

Examples of nonconformally coupled fields are,

- (a) light scalar fields, with $\xi \neq 1/6$, where ξ is the coupling to the curvature scalar, $\delta\mathcal{L} \sim -\xi R\varphi^2/2$
- (b) gravitons
- (c) massive fermions
- (d) massive gauge fields, *etc.*

In the case of matter (fermionic) fields and gauge fields, conformal invariance can be broken either by a mass term, or by coupling to a nonconformally coupled scalar field. For a discussion of conformal coupling of gauge and fermionic fields see section J of Part I.

When mass $m \gg H$, the amplification process is exponentially suppressed. Indeed, the uncertainty relation (70) gives in this case, $\int_{t_0}^t dt' 2E \sim 2m(t - t_0) \leq 1$, which implies that no real particle pairs are created by vacuum fluctuations. Indeed, just like in Minkowski space, all fluctuations are virtual, and can live only,

$$\Delta N \simeq H\Delta t \leq \frac{H}{2m} \ll 1 \quad (81)$$

e-foldings during inflation. A more careful WKB analysis reveals that an exponentially small number of particle pairs can nevertheless be created. The probability for particle pair creation is then suppressed by, $e^{-2\pi m/H}$.

To summarise, according to inflationary paradigm, the following basic picture of the origin of large scale structure formation emerges. During inflation, when the Hubble scales changes slowly, quantum fluctuations are generated and, as the Universe expand, the physical length of fluctuations grows as, λ_{phys} , crossing eventually the Hubble scale at $a = a_{1x}$, defined by $k/a = H$, as shown in figure 15. Modes of different wavelength approximately decouple (the effective coupling is weak, and it is characterised by, H_I^2/M_P^2). Later during radiation or matter era, the modes cross the horizon for the second time at $a = a_{2x}$, at the moment defined by, $k/a = H$. The amplitude of fluctuations freezes, that is it does not change between the first and second horizon crossing,

$$\delta\varphi|_{1x} \simeq \delta\varphi|_{2x} \sim \frac{H}{2\pi}. \quad (82)$$

For example, for gravitational radiation at super-Hubble scales during inflation, the amplitude of vacuum fluctuations freezes at,

$$\mathcal{P}_g \simeq \frac{H^2}{\pi^2 M_P^2}, \quad (83)$$

where H is to be evaluated at the first horizon crossing, $a = a_{1x}$, when $k/a = H$.

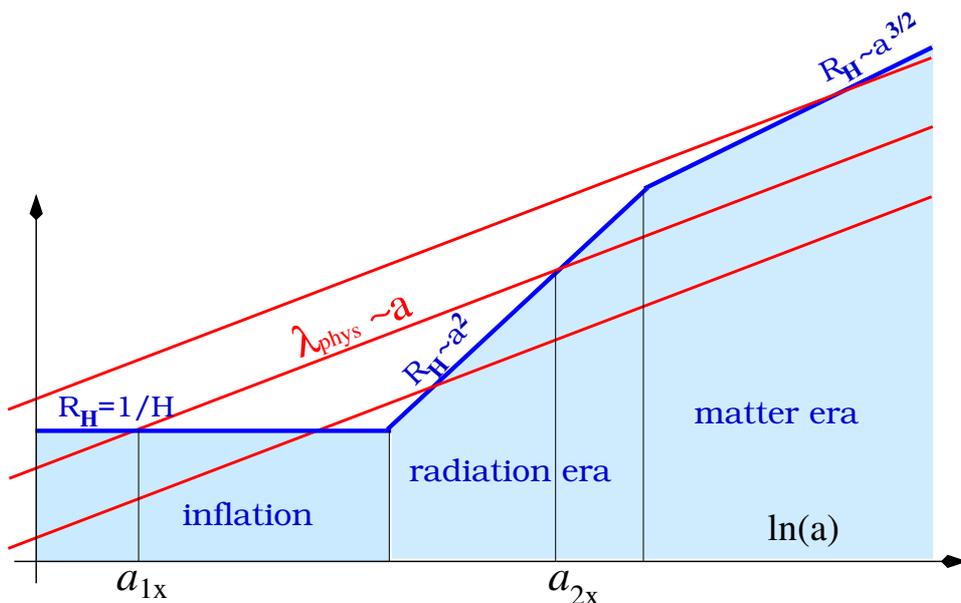


FIG. 15: Evolution of physical scale during inflation, radiation and matter era. Quantum fluctuations are generated at a scale M , probably smaller than the Planck scale M_P , by an unknown mechanism. During inflation the physical scale of the fluctuations grows with the Universe expansion, $\lambda_{\text{phys}} \propto a$, while the Hubble scale remains approximately constant, $R_H = 1/H \simeq \text{const.}$, such that the scale of quantum fluctuations becomes larger than the Hubble radius after the first Hubble crossing, $a = a_{1x}$. While the amplitude of fluctuations on sub-Hubble scales, $R < 1/H$, scales as, $\delta\varphi|_R \simeq 1/R$, it freezes out at super-Hubble scales, $\delta\varphi|_R \simeq H/2\pi$ ($R \gg 1/H$), and remains approximately constant until the second horizon crossing, $a = a_{2x}$, in radiation or matter era, $\lambda_{\text{phys}} \simeq 1/H$, at which the modes enter the Hubble radius and begin again oscillating.

The mechanism for generation of scalar cosmological perturbations during inflation relies on the assumption that a scalar field condensate (inflaton) exists. The energy density of such a field is given by,

$$\rho_\varphi = T_{\varphi_0}^0(\varphi) = \frac{1}{2}\dot{\varphi}^2 + \frac{1}{2}\left(\frac{\partial_i\varphi}{a}\right)^2 + V(\varphi). \quad (84)$$

Consider now small scalar perturbations, $\delta\varphi$, around the homogeneous field, φ_0 , defined by,

$$\varphi(\vec{x}, t) = \varphi_0(t) + \delta\varphi(\vec{x}, t). \quad (85)$$

Expanding (84) to first order in $\delta\varphi$ results in

$$\rho_\varphi \simeq \rho_{\varphi_0} + \delta_1\rho_\varphi \quad (86)$$

where

$$\rho_{\varphi_0} = \frac{1}{2}\dot{\varphi}_0^2 + V(\varphi_0) \quad (87)$$

and

$$\delta_1\rho_\varphi = \dot{\varphi}_0\delta\dot{\varphi} + V'(\varphi_0)\delta\varphi, \quad (88)$$

where $V' = dV(\varphi_0)/d\varphi_0$. This then implies for the density contrast,

$$\delta_\varphi \equiv \frac{\delta\rho_\varphi}{\rho_\varphi} \simeq \frac{\delta_1\rho_\varphi}{(\rho_\varphi)_0} \simeq \frac{\sqrt{2\epsilon}}{M_P}\delta\varphi \quad (89)$$

where $\epsilon = (M_P^2/2)(V'/V)^2$ is a slow roll parameter. With $\delta\varphi \simeq H/(2\pi)$, we arrive at the spectrum of the scalar density contrast during a slow roll regime on super-Hubble scales,

$$\mathcal{P}_{\delta_\varphi} \sim \frac{\epsilon}{M_P^2}\mathcal{P}_\varphi \sim \epsilon\left(\frac{H}{2\pi M_P}\right)^2, \quad (\text{first horizon crossing}) \quad (90)$$

This is calculated by assuming, $\dot{\varphi}\delta\varphi \ll V'\delta\varphi$ and $\dot{\varphi}_0^2 \ll V$, which is justified during a slow roll inflation.

On the other hand, after inflation we have $V \simeq 0$ and $V' \simeq 0$, and Eq. (90) reduces to,

$$\delta_\varphi \sim \frac{\delta\dot{\varphi}}{\dot{\varphi}_0}. \quad (91)$$

On super-Hubble scales, $\delta\dot{\varphi} \simeq H\delta\varphi$, such that

$$\mathcal{P}_{\delta_\varphi} \sim \frac{H^2}{\dot{\varphi}_0^2}\mathcal{P}_{\delta\dot{\varphi}} \sim \frac{H^2}{\dot{\varphi}_0^2}\frac{H^2}{4\pi^2} \quad (92)$$

It is not clear however to make a reasonable estimate of $\dot{\varphi}_0^2$. If we make the following slow roll estimate, $\dot{\varphi}_0^2 \sim V'^2/(9H^2) = 2\epsilon M_P^2 H^2$, we then get,

$$\mathcal{P}_{\delta_\varphi} \sim \frac{1}{\epsilon}\left(\frac{H}{2\pi M_P}\right)^2 \quad (\text{radiation/matter era}), \quad (93)$$

which gives the correct order of magnitude estimate for scalar perturbations. It is however unclear why should be the slow roll approximation justified, which was used in getting the estimate (93). A comparison of (93) and (90) illustrates the difficulties involved in making an heuristic estimate of the amplitude of scalar cosmological perturbations, since we have neglected the backreaction effects from gravitational potentials. In Part IV we make a proper quantitative calculation of the amplitude and spectrum of scalar fluctuations.

B. de Sitter space

Inflation is usually realised through models which exhibit an approximately power law expansion,

$$a \propto t^\alpha, \quad \alpha \gg 1, \quad (94)$$

which is known as quasi-de Sitter space, since in the limit, $\alpha \rightarrow \infty$, the space-time (94) becomes de Sitter space,

$$a \propto e^{H_I t}. \quad (95)$$

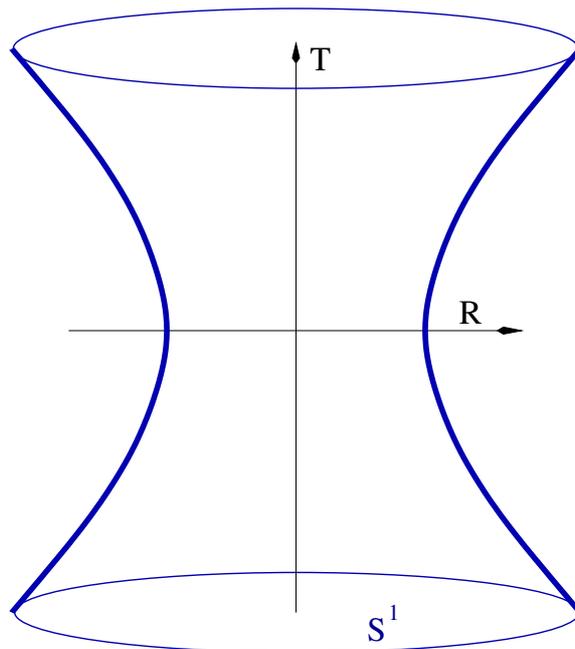


FIG. 16: The embedding of de Sitter space into a five dimensional flat space-time. The vertical line corresponds to the time coordinate, $X_0 = T$, and the radial coordinate, $R = \sqrt{X_1^2 + X_2^2 + X_3^2 + X_4^2}$. At each point (T, R) there is a unit 3-sphere, S^3 . Since we cannot easily draw 3-spheres, we mimic the 3-spheres by the circles, corresponding to a S^1 at each point (T, R) . Hubble radius, $R_H = 1/H_I$, which characterises the radius of curvature of the manifold, is the distance R of the hyperboloid from the origin at $T = 0$.

This motivates a study of de Sitter space as a prototype model for inflationary space-times, with the caveat that so far no realistic particle physics models of inflation have been constructed which exhibit an exact de Sitter space. Namely, in proposed working models of inflation, inflationary era is followed by a radiation era, and it is not known how to convert a de Sitter space into a radiation dominated universe. This is also known as the graceful exit problem of de Sitter inflationary models.

De Sitter space exhibits de Sitter symmetry, which corresponds to a local

$$SO(4, 1) \tag{96}$$

group Lorentz group. The symmetry can be most easily appreciated by making use of a flat embedding of de Sitter space into a flat five dimensional space-time,

$$ds_5^2 = dX_0^2 - dX_1^2 - dX_2^2 - dX_3^2 - dX_4^2, \tag{97}$$

in which de Sitter space carves out a five dimensional hyperboloid, defined by,

$$X_0^2 - X_1^2 - X_2^2 - X_3^2 - X_4^2 = -\frac{1}{H_I^2}, \tag{98}$$

where H_I denotes the Hubble parameter. The space-time diagram is shown in figure 16. This is an extrinsic embedding of de Sitter space, in which the curvature of space-time is seen as motion on a

curved hypersurface. This is in contrast with intrinsic embeddings to be discussed below, in which the de Sitter hyperboloid is parametrised by coordinates, whose number corresponds to the dimensionality of the hyperboloid, and the information about curvature is in the nontrivial metric.

One can define a de Sitter invariant distance function,

$$Z(X, X') \equiv H_I^2 \sum_{a,b=0}^4 \eta_{ab} X_a X'_b = -1 - \frac{1}{2} H_I^2 \sum_{a,b=0}^4 \eta_{ab} (X_a - X'_a)(X_b - X'_b), \quad (99)$$

such separations between two points X and X' are lightlike when,

$$Z(X, X') = -1, \quad (100)$$

separations are spacelike when

$$Z(X, X') > -1, \quad (101)$$

and separations are timelike when

$$Z(X, X') < -1. \quad (102)$$

Similarly to its lower dimensional cousin, the Lorentz group $SO(3, 1)$, which has 6 generators, 3 of which are boosts and 3 are rotations, $SO(4, 1)$ has 10 generators, 4 of which are boosts, and 6 are rotations. Different 3+1 dimensional coordinatisations of de Sitter manifold are possible, with the symmetry of spatial sections,

$$E(3), \quad SO(3, 1), \quad O(4) \subset SO(4, 1), \quad (103)$$

each of which is a subgroup of $SO(4, 1)$ with 6 generators.

1. Flat coordinates

The following coordinate transformation,

$$\begin{aligned} X_0 &= \frac{1}{H_I} \sinh(H_I t) + \frac{H_I}{2} e^{H_I t} \|\vec{x}\|^2, & -\infty < t < \infty \\ X_4 &= \frac{1}{H_I} \cosh(H_I t) - \frac{H_I}{2} e^{H_I t} \|\vec{x}\|^2 \\ X_i &= e^{H_I t} x_i, & -\infty < x_i < \infty, & \quad (i = 1, 2, 3) \end{aligned} \quad (104)$$

These coordinates cover 1/2 of de Sitter manifold shown in figure 17, as can be seen from

$$X_0 + X_4 = \frac{1}{H_I} e^{H_I t} \geq 0. \quad (105)$$

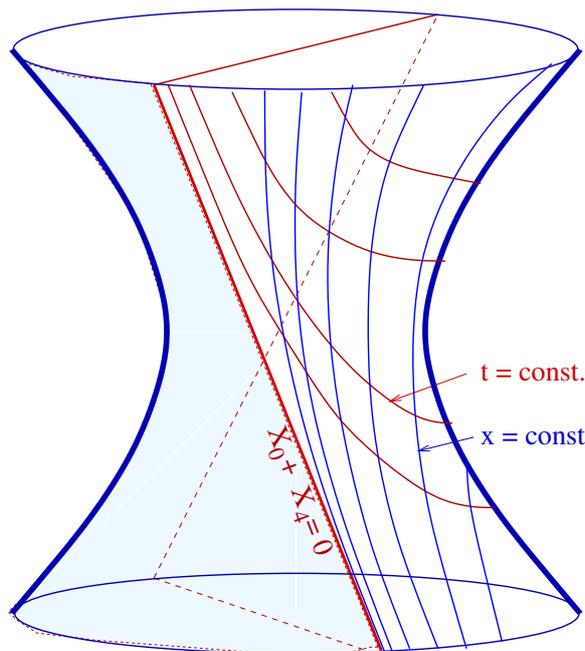


FIG. 17: The embedding of de Sitter space into a five dimensional flat space-time in flat coordinates, which cover half of de Sitter manifold, $X_0 + X_4 > 0$. The surfaces (lines) $t = \text{const.}$ and $r = \sqrt{\sum_i x_i^2} = \text{const.}$ are also indicated.

In order to find out what is the line element in flat coordinates, we take the differential of (104), to obtain

$$\begin{aligned}
 dX_0 &= \cosh(H_I t) dt + \frac{H_I^2}{2} e^{H_I t} \|\vec{x}\|^2 dt + H_I e^{H_I t} \vec{x} \cdot \vec{d}\vec{x} \\
 dX_4 &= \sinh(H_I t) dt - \frac{H_I^2}{2} e^{H_I t} \|\vec{x}\|^2 dt - H_I e^{H_I t} \vec{x} \cdot \vec{d}\vec{x} \\
 dX_i &= H_I e^{H_I t} x_i dt + e^{H_I t} dx_i,
 \end{aligned} \tag{106}$$

square them and add or subtract, to obtain the line element,

$$ds^2 = dt^2 - a^2 d\vec{x}^2, \quad a(t) = e^{H_I t}, \tag{107}$$

which clearly corresponds to a metric with flat spatial sections of FLRW form. Eq. (107) can be easily transform to conformal coordinates,

$$dt = a d\eta \rightarrow \eta = -\frac{1}{H_I} e^{-H_I t}, \quad a = -\frac{1}{H_I \eta}, \tag{108}$$

in which the line element (107) becomes

$$ds^2 = a(\eta)^2 (d\eta^2 - d\vec{x}^2). \tag{109}$$

Conformal coordinates, $-\infty < x_i < \infty$, $-\infty < \eta < 0$, cover a half of the de Sitter manifold. Since the scale factor,

$$a = -\frac{1}{H_I \eta} \quad (\eta < 0), \tag{110}$$

increases with increasing η , the spatial sections are expanding as η increases. Also since

$$\frac{a''}{a} = \frac{2}{\eta^2}, \quad (111)$$

space-time undergoes accelerated expansion, which is a necessary condition for an inflationary Universe.

On the other hand, the conformal coordinates, $-\infty < x_i < \infty$, $0 < \eta < \infty$ cover the second half of the de Sitter manifold, with the scale factor,

$$a = \frac{1}{H_I \eta} \quad (\eta > 0), \quad (112)$$

which decreases as η increases, implying that the spatial sections are contracting as η increases.

Flat coordinates are of relevance for cosmology, since the curvature radius (26–27) of the Universe is large,

$$R_{\text{curv}} = \frac{1}{\sqrt{|\mathcal{K}|}} \geq 21 \text{ Gpc} \gg \frac{1}{H_I}. \quad (113)$$

This is indeed much larger than the Hubble radius during inflation, $R_{H_I} = 1/H_I \sim 10^{-28}$ m, as well as the Hubble radius today, $R_H \simeq 4$ Gpc, such that, to a good precision, the spatial sections of the Universe can be considered as flat. This means that, when one studies evolution of quantum fields during inflation and the subsequent radiation and matter era, matching of flat inflationary sections onto flat radiation era sections suffices for most purposes. If, on the other hand, one is interested in studying subtle differences arising due to the spatial curvature of the Universe, and which may arise on Hubble scales today, then open and closed coordinates may be useful, which is what we discuss next.

2. Closed coordinates

Closed coordinates are defined by the following coordinate transformation,

$$\begin{aligned} X_0 &= \frac{1}{H_I} \sinh(H_I t), & [-\infty < t < \infty] \\ X_1 &= \frac{1}{H_I} \cosh(H_I t) \cos(\chi), & [0 \leq \chi \leq \pi] \\ X_2 &= \frac{1}{H_I} \cosh(H_I t) \sin(\chi) \cos(\theta), & [0 \leq \theta \leq \pi] \\ X_3 &= \frac{1}{H_I} \cosh(H_I t) \sin(\chi) \sin(\theta) \cos(\phi), & [0 \leq \phi \leq 2\pi] \\ X_4 &= \frac{1}{H_I} \cosh(H_I t) \sin(\chi) \sin(\theta) \sin(\phi). & \end{aligned} \quad (114)$$

These coordinates cover fully the de Sitter manifold shown in figure 18, where we indicate lines (hypersurfaces) of constant χ and t . For negative times, $t < 0$, spatial sections of space-time contract with

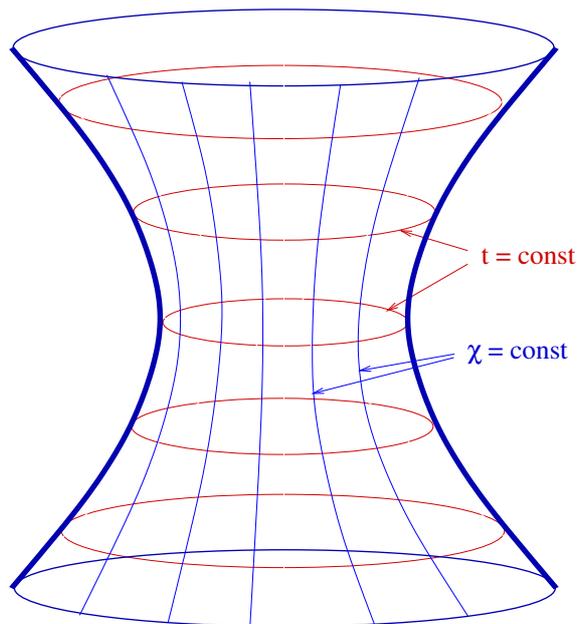


FIG. 18: de Sitter space covered by closed coordinates, which represent a full cover of the de Sitter manifold. The lines (surfaces) $t = \text{constant}$ and $\chi = \text{constant}$ are indicated.

increasing t , while for positive times, $t \geq 0$, spatial sections expand as t increases. The underlying symmetry group of spatial sections of this coordinate covering is $O(4)$.

After differentiating (114) and inserting the result into Eq. (97), we arrive at the line element

$$ds^2 = dt^2 - \frac{a^2(t)}{H_I^2} [d\chi^2 + \sin^2(\chi) d\Omega^2], \quad a = \cosh(H_I t), \quad d\Omega^2 = d\theta^2 + \sin^2(\theta) d\phi^2, \quad (115)$$

which corresponds to a FLRW metric with positively curved spatial sections. The line element (115) can be transformed into conformal coordinates,

$$dt = ad\eta/H_I \rightarrow \eta = 2 \arctan \left(e^{H_I t} \right), \quad a(\eta) = \frac{1}{\sin(\eta)} \quad (0 \leq \eta \leq \pi), \quad (116)$$

in which the line element acquires the following conformal form,

$$ds^2 = a^2(\eta) \left(d\eta^2 - [d\chi^2 + \sin^2(\chi) d\Omega^2] \right), \quad d\Omega^2 = d\theta^2 + \sin^2(\theta) d\phi^2. \quad (117)$$

These coordinates are referred to as the conformal cover of de Sitter space, and the corresponding Carter-Penrose diagram is shown in figure 19. Since they represent a full cover of de Sitter space, and they are all of finite range, closed coordinates are often used in studies of de Sitter space. They are also of use in cosmology, for example when one is interested in studying spectra of fields produced during de Sitter inflation on scales comparable to the Hubble radius today.

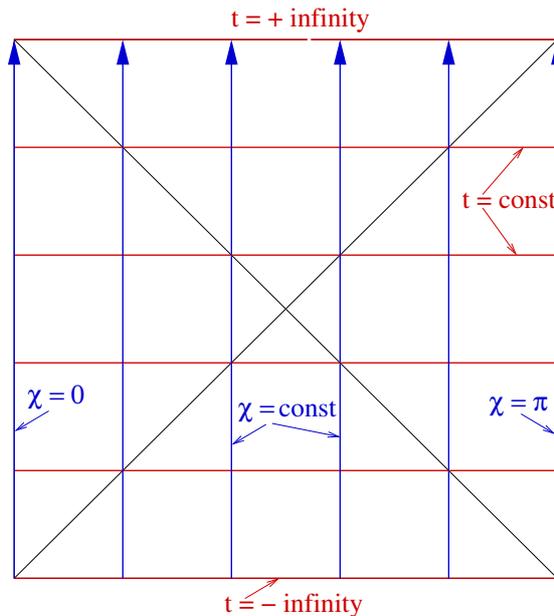


FIG. 19: The Carter-Penrose diagram of de Sitter space as covered by closed coordinates. Vertical lines correspond to $\chi = \text{constant}$, which are also geodesics, horizontal lines to $t = \text{constant}$. The lines $t = +\infty$ and $t = -\infty$ are also indicated. This diagram correspond to the diagram shown in figure 18 cut along $\chi = 0 = \pi$.

3. Open coordinates

Open coordinates are defined by the following coordinate transformation,

$$\begin{aligned}
 X_0 &= \frac{1}{H_I} \sinh(H_I t) \cosh(\lambda), & [0 \leq \lambda < \infty] \\
 X_1 &= \frac{1}{H_I} \cosh(H_I t), & [-\infty < t < \infty] \\
 X_2 &= \frac{1}{H_I} \cosh(H_I t) \sinh(\lambda) \cos(\theta), & [0 \leq \theta \leq \pi] \\
 X_3 &= \frac{1}{H_I} \cosh(H_I t) \sinh(\lambda) \sin(\theta) \cos(\phi), & [0 \leq \phi \leq 2\pi] \\
 X_4 &= \frac{1}{H_I} \cosh(H_I t) \sinh(\lambda) \sin(\theta) \sin(\phi). &
 \end{aligned} \tag{118}$$

After differentiating (118) and inserting the result into Eq. (97), we obtain,

$$ds^2 = dt^2 - \frac{a^2(t)}{H_I^2} [d\lambda^2 + \sinh^2(\lambda) d\Omega^2], \quad a = \sinh(H_I t), \tag{119}$$

which corresponds to a FLRW metric with negatively curved spatial sections (spatial sections are open hyperboloids), whose symmetry group is the Lorentz group, $SO(3, 1)$. One can show that these coordinates cover 1/4 of the de Sitter manifold.

When transformed into conformal coordinates, defined by

$$dt = a d\eta \quad \rightarrow \quad a = -\frac{1}{\sinh(H_I \eta)} = \sinh(H_I t) \quad (-\infty < \eta < 0), \tag{120}$$

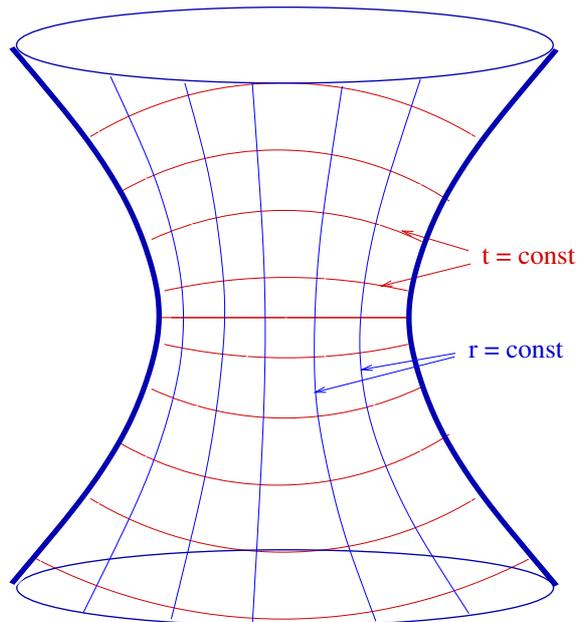


FIG. 20: The de Sitter manifold covered by static coordinates. The lines (surfaces) $t = \text{constant}$ and $r = \text{constant}$ are indicated.

the line element (119) becomes

$$ds^2 = a^2(\eta) \left(d\eta^2 - \frac{1}{H_I^2} [d\lambda^2 + \sin^2(\lambda) d\Omega^2] \right). \quad (121)$$

These coordinates cover 1/2 of de Sitter manifold, in which the spatial sections expand with growing η . The other half of the manifold is covered by the scale factor,

$$a = \frac{1}{\sinh(H_I \eta)}, \quad (0 < \eta < \infty), \quad (122)$$

such that spatial sections shrink with increasing η . The de Sitter manifold covered by static coordinates is shown in figure 20. In the limit when $|\eta| \ll 1/H_I$, open coordinate covering of de Sitter space is well approximated by flat coordinates (108–109).

Open coordinates can be of use in cosmology for studies of quantum field spectra produced during inflation on Hubble scales.

4. Static coordinates

The following coordinate transformation defines static coordinates,

$$\begin{aligned} X_0 &= \sqrt{\frac{1}{H_I^2} - r^2} \sinh(H_I t), & [0 \leq r < 1/H_I] \\ X_1 &= \sqrt{\frac{1}{H_I^2} - r^2} \cosh(H_I t), & [-\infty < t < \infty] \end{aligned}$$

$$\begin{aligned}
X_2 &= r \sin(\theta) \cos(\phi), & [0 \leq \theta \leq \pi] \\
X_3 &= r \sin(\theta) \sin(\phi), & [0 \leq \phi \leq 2\pi] \\
X_4 &= r \cos(\theta). &
\end{aligned} \tag{123}$$

Note that these coordinates break spatial homogeneity of space-time. These coordinates are analogous to Rindler coordinates, which break the spatial homogeneity of Minkowski space-time, and which correspond to Minkowski space-time as seen by a uniformly accelerated observer.

Equations (123) imply,

$$X_0 + X_1 > 0, \tag{124}$$

and hence these coordinates cover at most one half of the de Sitter manifold. In fact, one can show that when r is allowed to range from 0 to ∞ , then 1/2 of the de Sitter manifold is covered by static coordinates. When $r > 1/H_I$, Eqs. (123) should be replaced by,

$$\begin{aligned}
X_0 &= \sqrt{r^2 - \frac{1}{H_I^2}} \cosh(H_I t), & [1/H_I < r < \infty] \\
X_1 &= \sqrt{r^2 - \frac{1}{H_I^2}} \sinh(H_I t), &
\end{aligned} \tag{125}$$

with X_2 , X_3 and X_4 unchanged.

When expressed in terms of static coordinates (123), the line element (97) reads,

$$ds^2 = (1 - H_I^2 r^2) dt^2 - \frac{dr^2}{1 - H_I^2 r^2} - r^2 d\Omega^2. \tag{126}$$

Since the metric is static, this line element possesses a time like Killing vector, ∂_t . The line element (126) is valid for $0 \leq r < \infty$. Similarly as in the case of Schwarzschild space, there is an event horizon, where, from the point of view of an observer at the origin, $r = 0$, the metric (126) becomes singular. Since this singularity is not present in other coordinates, we conclude that this is a coordinate singularity. The existence of an event horizon has led to a spur of activity on de Sitter thermal radiation with a Hawking temperature, $T_H = H_I/2\pi$, which is the temperature of radiation seen by an observer in static coordinates at $r = 0$. This is at best indirectly related to the spectra of fields produced by vacuum fluctuations during de Sitter epoch of relevance for structure formation, since these spectra show the important scale invariant behaviour on super-Hubble scales, $r > 1/H_I$, which are inaccessible to an observer in static coordinates. In particular, a static observer sees a thermal spectrum of conformally coupled fields on sub-Hubble scales, and hence would have no way of telling that conformally coupled fields are irrelevant for structure formation.

C. Evidence for inflation

Here we briefly summarise evidence supporting that the Universe underwent a period of cosmic inflation.

- (1) Inflation predicts a nearly scale invariant spectrum of cosmological perturbations, which is supported by the formation of large scale structures of the Universe and by the observed anisotropies in the cosmic microwave background radiation;
- (2) The near spatial flatness of the Universe, expressed by the lower limit on the radius of curvature of the Universe,

$$R_{\text{curv}} \geq 20 \text{ Gpc}. \quad (127)$$

Sometimes one also mentions the adiabatic gaussian nature of cosmological perturbations as a generic prediction of inflation, even though one could quite easily design inflationary models, where the dominant component corresponds to isocurvature perturbations, and which can be, for example, created by energy density perturbations in one of the matter components, which are nor adiabatic nor Gaussian.

Furthermore, acoustic peaks in cosmic microwave background radiation are a consequence of Sakharov oscillations in gravitational potentials, and hence can be considered as a part of the general framework of nearly scale invariant adiabatic cosmological perturbations generated during inflation.

Even though the evidence for inflation is indirect, and not fully conclusive, it is nevertheless quite suggestive. One should not forget that the alternative explanations for structure formation, which comprise topological defects and isocurvature perturbations, have been both ruled out by the recent CMB anisotropy observations.

D. Scalar inflationary models

Let us consider the scalar action in a curved space-time background,

$$S_\phi = \int d^4x \sqrt{-g} \mathcal{L}_\phi, \quad (128)$$

with the Lagrangian

$$\sqrt{-g} \mathcal{L}_\phi = \sqrt{-g} \frac{1}{2} g^{\mu\nu} (\partial_\mu \phi)(\partial_\nu \phi) - \sqrt{-g} V(\phi), \quad (129)$$

where $V = V(\phi)$ denotes a scalar potential, $g^{\mu\nu}$ is the inverse of the metric tensor $g_{\mu\nu}$ and $g = \det(g_{\mu\nu})$. This Lagrangian assumes a canonical kinetic term, which in Minkowski space reduces to the standard form, $(1/2)\eta^{\mu\nu}(\partial_\mu \phi)(\partial_\nu \phi)$, $\eta^{\mu\nu} = \text{diag}(1, -1, -1, -1)$. Other forms of kinetic term have been considered,

which in the one field case can be quite generally (up to possible singular points of the transformation) mapped onto the form (129). In the case of several scalar fields however, such a map is not always possible, leading to qualitatively different behaviour of multi-inflaton models.

The corresponding equation of motion is easily obtained by varying the action (128) with respect to ϕ ,

$$\frac{1}{\sqrt{-g}}\partial_\mu\left(g^{\mu\nu}\sqrt{-g}\partial_\nu\phi\right) + V' = 0, \quad (130)$$

where $V' = dV/d\phi$.

In a homogeneous FLRW space time, with the metric tensor,

$$g_{\mu\nu} = \text{diag}(1, -a^2, -a^2, -a^2), \quad (131)$$

where $a = a(t)$ denotes the scalar factor, the equation of motion (130) for the homogeneous component of $\phi = \phi(t)$, becomes

$$\ddot{\phi} + 3H\dot{\phi} + V' = 0, \quad (132)$$

where $H = \dot{a}/a$, and $\dot{\phi} = d\phi/dt$. Eq. (132) together with the Friedmann equation,

$$H^2 = \frac{1}{3M_P^2}\rho_\phi, \quad \rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi) \quad (133)$$

fully determines the dynamics of $\phi = \phi(t)$ and $a = a(t)$ during inflation. In particular, both the energy conservation equation,

$$\dot{\rho}_\phi + 3H(\rho_\phi + \mathcal{P}_\phi) = 0, \quad \mathcal{P}_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi) \quad (134)$$

and the second FLRW equation,

$$\frac{\ddot{a}}{a} = -\frac{1}{6M_P^2}(\rho_\phi + 3\mathcal{P}_\phi) \quad (135)$$

can be derived from (132) and (133).

Slow roll approximation

Recall that during inflation the active gravitational mass,

$$\rho_\phi + 3\mathcal{P}_\phi = 2(\dot{\phi}^2 - V) < 0 \quad (136)$$

must be negative. It is often the case that during inflation, to a good approximation potential energy dominates kinetic energy by a large factor. This is one of the conditions defining slow roll regime of inflation, which can be defined by the following conditions,

$$\ddot{\phi} \ll 3H\dot{\phi}, V', \quad \dot{\phi}^2 \ll V. \quad (137)$$

More generally, during slow roll inflation the following conditions are satisfied,

$$\epsilon \ll 1, \quad \eta \ll 1, \quad \xi \ll 1, \quad \dots, \quad (138)$$

where slow roll parameters are defined as,

$$\epsilon = \frac{1}{2} M_P^2 \left(\frac{V'}{V} \right)^2 \quad (139)$$

$$\eta = M_P^2 \frac{V''}{V} \quad (140)$$

$$\xi^2 = M_P^4 \frac{V''' V'}{V^2}. \quad (141)$$

For example, we can easily integrate Eq. (134) in slow roll regime of inflation, to obtain

$$\int_{\phi_0}^{\phi} d\phi' \frac{\sqrt{V}}{V'} = -\frac{M_P}{\sqrt{3}}(t - t_0) \quad (142)$$

where $V = V(\phi')$, $V' = dV/d\phi'$.

Next, it is convenient to define the number of e-foldings of inflation,

$$N(t) = \int_t^{t_e} H(t') dt' \quad (143)$$

where t_e denotes the end of inflation. In slow roll regime, Eq. (143) becomes,

$$N = \frac{1}{\sqrt{3} M_P} \int_t^{t_e} \sqrt{V} dt' = \frac{1}{M_P^2} \int_{\phi_e}^{\phi} \frac{V}{V'} d\phi', \quad (144)$$

where $\phi_e = \phi(t_e)$.

For example, in a chaotic model of inflation realised by the quadratic potential

$$V = \frac{1}{2} m_\phi^2 \phi^2 \quad (145)$$

and by a large scalar expectation value, $\phi \gg M_P$, one finds by integrating (142),

$$\phi = \phi_0 - \sqrt{\frac{2}{3}} M_P m_\phi (t - t_0) \quad (146)$$

where $\phi_0 = \phi(t_0) \gg M_P$ is the initial value of the scalar field (at an initial time $t = t_0$). With this we can write the Friedmann equation (133) as,

$$\frac{d}{dt} \left[\ln \left(\frac{a}{a_0} \right) \right] = \frac{\sqrt{V}}{\sqrt{3} M_P} = \frac{1}{\sqrt{6}} \frac{m_\phi}{M_P} \left(\phi_0 - \sqrt{\frac{2}{3}} M_P m_\phi (t - t_0) \right) \quad (147)$$

which can be integrated to give,

$$a = a_0 \exp \left[\frac{\phi_0 m_\phi (t - t_0)}{\sqrt{6} M_P} \left(1 - \sqrt{\frac{1}{6}} \frac{M_P m_\phi (t - t_0)}{\phi_0} \right) \right] = a_0 \exp \left[\frac{\phi_0^2 - \phi^2}{4 M_P^2} \right]. \quad (148)$$

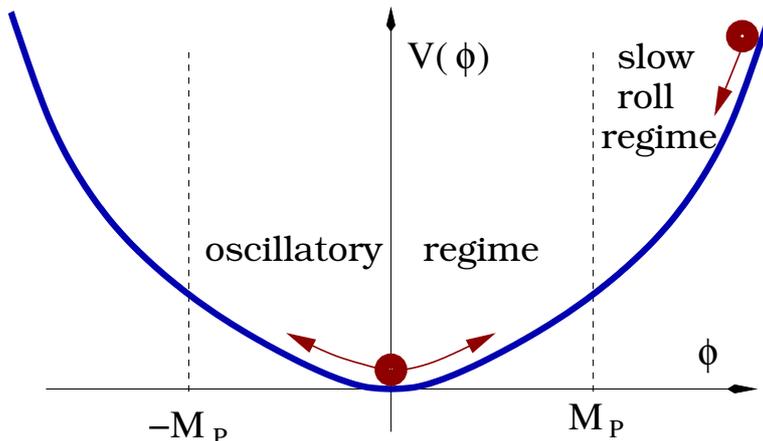


FIG. 21: Slow roll ($\phi \gg M_P$) and oscillatory regime ($\phi \leq M_P$) of a single scalar field inflationary model.

As long as the piece in the exponent, which evolves quadratically with time, can be neglected with respect to the linear term, the solution (148) represents an exponentially expanding universe, motivating the term ‘quasi-de Sitter inflation’.

The number of e-foldings (144) in slow roll approximation is,

$$N = \frac{1}{2M_P^2} \int_{\phi}^{\phi_e} \phi' d\phi' = \frac{\phi^2 - \phi_e^2}{4M_P^2}, \quad (149)$$

where $\phi = \phi(t)$ is defined in Eq. (146), and $\phi_e = \phi(t_e)$, where t_e denotes the end of inflation, which is to a good approximation given by, $\ddot{a}/a \simeq 0$, which in many cases corresponds to the time when one or more of the slow roll conditions (138) get violated.

Oscillatory regime and inflaton decay

Roughly speaking, slow roll inflation ends when $\phi \sim M_P$, below which ϕ begins oscillating. The two regimes are illustrated in figure 21. During the oscillatory regime, it is useful to rescale the field,

$$\phi = \frac{\varphi}{a^{3/2}}, \quad (150)$$

such that Eq. (132) becomes,

$$\ddot{\varphi} + \left[-\frac{3}{4} \left(\frac{\dot{a}}{a} \right)^2 - \frac{3\ddot{a}}{2a} \right] \varphi + a^{3/2} V'(\varphi/a^{3/2}) = 0, \quad (151)$$

In the simple case, when $V'(\phi) = m_\phi \phi$, immediately after inflation the Universe enters to a good approximation an era of matter domination,

$$a \propto t^{2/3} \quad (152)$$

such that

$$-\frac{3}{4} \left(\frac{\dot{a}}{a} \right)^2 - \frac{3\ddot{a}}{2a} = 0, \quad (153)$$

and Eq. (151) simplifies to

$$\ddot{\varphi} + m_\phi^2 \varphi = 0. \quad (154)$$

This is solved by the harmonic function,

$$\phi = \frac{\varphi}{a^{3/2}} = \frac{A \cos(m_\phi t + \xi_0)}{a^{3/2}}, \quad (155)$$

where A and ξ_0 are two integration constants. With this the scalar energy density (133) scales as,

$$\rho_\phi \propto \frac{1}{a^3}, \quad (156)$$

implying a self-consistency of the solution (152). This means that, before inflaton decays, during the oscillatory regime of the inflationary model governed by the quadratic potential (145), the Universe is in an intermediate matter era. A phenomenological description of the inflaton decay can be obtained by adding to Eq. (132) a phenomenological damping term, $\Gamma_\phi \dot{\phi}$, resulting in

$$\ddot{\phi} + 3H\dot{\phi} + \Gamma_\phi \dot{\phi} + V' = 0, \quad (157)$$

which mimics the inflaton decay into other matter fields, upon which the Universe thermalises. Roughly speaking, the solution (155) gets modified to

$$\phi \sim e^{-\Gamma_\phi t} \frac{A \cos(m_\phi t + \xi_0)}{a^{3/2}}, \quad (158)$$

such that, apart from being redshifted by $a^{-3/2}$, the scalar field decays exponentially. From the equation of motion (132), it is clear that the the damping becomes important when,

$$\Gamma_\phi \sim H \simeq \frac{g_*^{1/2}}{3} \frac{T^2}{M_P}, \quad (159)$$

based on which one can estimate the thermalisation (reheat) temperature,

$$T_{\text{reh}} \simeq g_*^{-1/4} \sqrt{3M_P \Gamma_\phi}. \quad (160)$$

1. Guth's model

Inspired by a phase transition in the scalar sector of a grand unified theory, Allan Guth proposed in 1981 an inflationary model, whose potential we sketch in figure 22. Already in the original article Guth realised that a temporary inflationary phase can resolve the magnetic monopole, homogeneity, isotropy and flatness problems of cosmology. But he also understood that his inflationary model is flawed: the Universe gets stuck in the false vacuum forever. Namely, the Universe in the false vacuum state expands

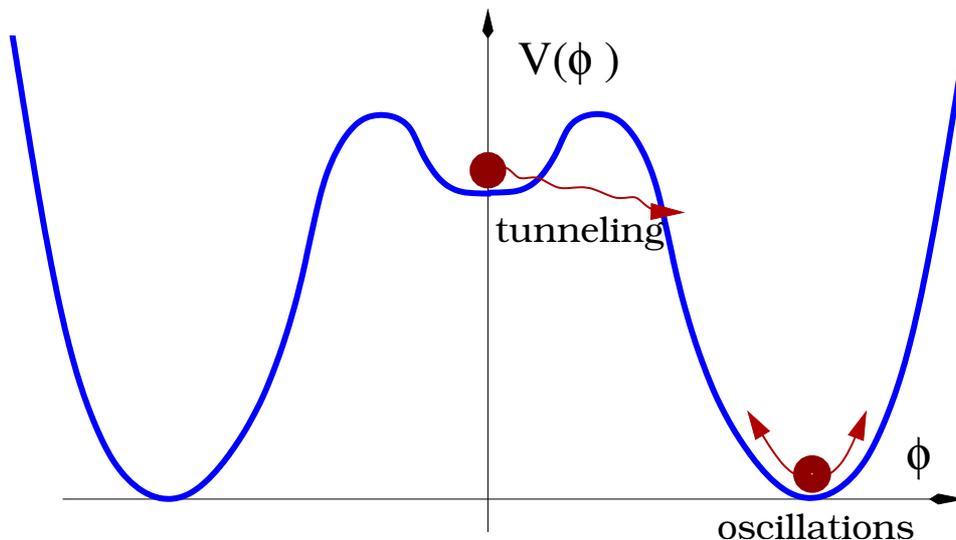


FIG. 22: A sketch of the scalar field potential for Guth’s inflationary model. Inflation begins by a scalar field trapped in a false vacuum, whose vacuum energy, $V = V_0$, drives a de Sitter-like inflation, $H_I = \sqrt{V_0/3}/M_P$. After some time the field tunnels to the true vacuum, oscillates and decays. The potential energy in the true vacuum is, $V \simeq 0$.

so rapidly, that it overcomes nucleation of bubbles, implying that, in spite of a continuous nucleation, the amount of space in the false vacuum grows. This problem is known as the “graceful exit problem.” It is not known in general how the Universe can exit out of a de Sitter phase.

To illustrate the graceful exit problem, we sketch in figure 23 a sequence of rapidly expanding space-times in conformal coordinates. Initially at conformal time, $\eta = \eta_0$, the whole spatial section is in the false vacuum state. At $\eta = \eta_1$ the space-time has tunneled with a probability, $P = 1/3$, into the true vacuum state, indicated by *vac*. the space-time has expanded by a factor, $a_1/a_0 = 3$, which means that physical distances between objects have grown by the same factor. At a time $\eta = \eta_2$ a further $P = 1/3$ -rd of the space-time has tunneled into the true vacuum state *vac*, and finally at a time $\eta = \eta_3$ yet a further $P = 1/3$ -rd of space-time has tunneled, *etc.* The volume of the false vacuum at $\eta = \eta_0$ is V_0 . At $\eta = \eta_1$, the volume of false vacuum has increases to,

$$V_1 = V_0 \times 3 - \frac{1}{3} \times 3 \times V_0 = 2V_0 \quad (161)$$

Similarly, at $\eta = \eta_2$ and $\eta = \eta_3$, we get,

$$V_2 = 2^2V_0, \quad V_3 = 2^3V_0, \quad \text{etc.}, \quad (162)$$

implying that, in spite of the tunneling, the amount of space in the false vacuum expands exponentially, posing a graceful exit problem.

The graceful exit problem is most elegantly solved by power law inflationary models, in which space-time expands as a power law of the scale factor. Since bubble nucleation is an exponentially fast process,

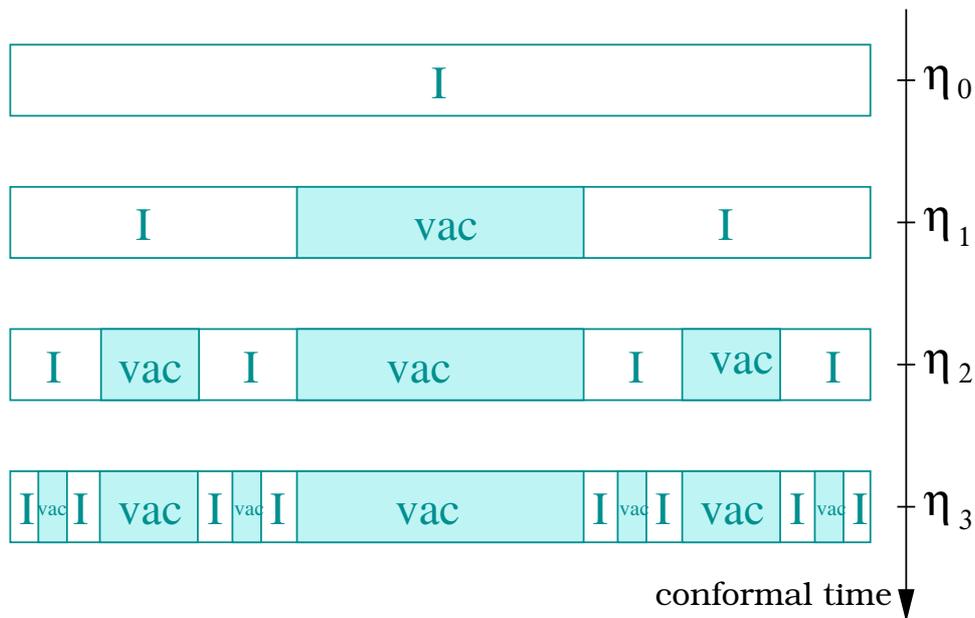


FIG. 23: A conformal diagram of an exponentially expanding inflationary space-time. Snapshots of a comoving volume taken at, $\eta = \eta_0 < \eta = \eta_1 < \eta = \eta_2 < \eta = \eta_3$, illustrate the graceful exit problem of de Sitter inflation. In spite of a continuous nucleation of a true vacuum (vac), when measured in physical units, the volume of space in false vacuum (I) continues to grow exponentially.

the nucleation eventually wins over the expansion. A first model of inflation which solves the graceful exit problem was constructed by Andrei Linde and by Albrecht and Steinhardt in 1982, and it was termed (somewhat optimistically) new inflation, which we study next.

2. New inflation

Consider a potential shown in figure 24, which contains a very flat section around the origin, $V' \simeq 0$, for $\varphi \approx 0$. Inflation is in this model realised by the scalar field being placed initially at, $\varphi \approx 0$, but $\dot{\varphi} \neq 0$, with $\dot{\varphi} = 0$, or by $\varphi = 0$ and $\dot{\varphi} = \dot{\varphi}_0 \neq 0$. While this problem solves naturally the graceful exit problem, it possesses the following naturalness problem. A very flat potential is required in order to realise more than, $N \geq 60$, e-foldings of inflation, needed to solve the flatness and homogeneity problems of cosmology. A solution to this naturalness problem is provided by the so-called flat potential directions, which occur in supersymmetric models. We shall return to this question below, where we study the F-term hybrid inflationary model. Another unpleasant feature of the new inflationary scenario is that, in order to get a sufficiently long inflation, it is necessary to fine tune initial conditions, such that to a high accuracy $\varphi \approx 0$, and $\dot{\varphi}_0 \approx 0$. Given that a natural state of the preinflationary Universe is a primordial chaos, a small value of both kinetic term and the field expectation value over a relatively large region space

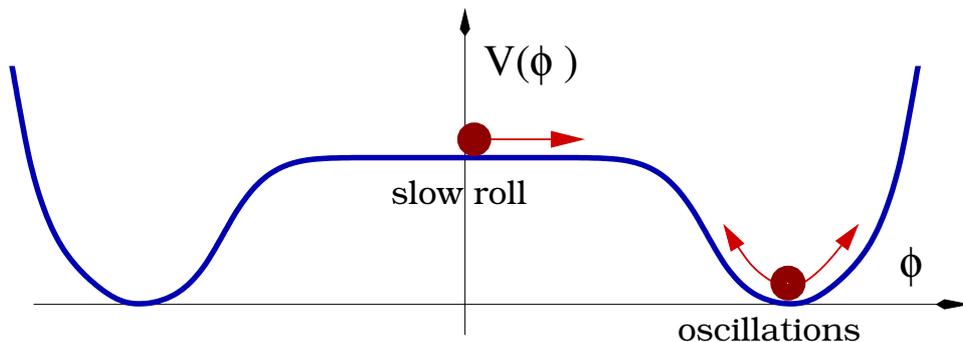


FIG. 24: Potential of a new inflationary model. Inflation begins at $\varphi = 0$, $\dot{\varphi}\dot{\phi}_0 \neq 0$ (or at $\phi_0 \neq 0$, $\dot{\varphi}_0 = 0$). Since the potential is very flat around the origin, the field takes a long time to roll down to the true minimum, leading to inflation.

is very unlikely. Andrei Linde has proposed chaotic inflationary models, which address both of these problems.

3. Chaotic inflation

Chaotic inflation is usually realised with one of the following two potentials,

$$V = \frac{1}{2}m_\varphi^2\varphi^2, \quad V = \frac{\lambda_\varphi}{4!}\varphi^4. \quad (163)$$

In order for inflation to start, the field has to be displaced away from the origin by, $\varphi \gg M_P$, such that

$$V \gg \frac{1}{2}\dot{\varphi}^2 \quad (164)$$

and hence

$$\rho_\varphi + 3\mathcal{P}_\varphi = 2(\dot{\varphi}^2 - V) < 0 \quad (165)$$

in the region of space which is much larger than the Hubble volume,

$$V \gg V_H = \frac{4\pi}{3} \frac{1}{H^3}, \quad H^2 = \frac{\rho_\varphi}{3M_P^2} - \frac{\mathcal{K}}{a^2} \quad (166)$$

where $\mathcal{K} = \mathcal{K}(x^\mu)$ denotes a (local) spatial curvature, $R_{\text{curv}} = |\mathcal{K}|^{-1/2}$. Assuming that before inflation the Universe is in a primeval chaotic state of unknown statistical properties, it is really difficult to make any statement on how probable such a state would be. A naive estimate gives though,

$$p \propto \exp\left(-VM_P^3\right). \quad (167)$$

For a grand unified inflation, $V \gg V_H \sim 3^{3/2}M_P^3/E_{\text{GUT}}^6$, we get the estimate,

$$p \sim \exp\left[-\mathcal{O}(10)\left(\frac{M_P}{E_{\text{GUT}}}\right)^6\right] \sim e^{-10^{15}}. \quad (168)$$

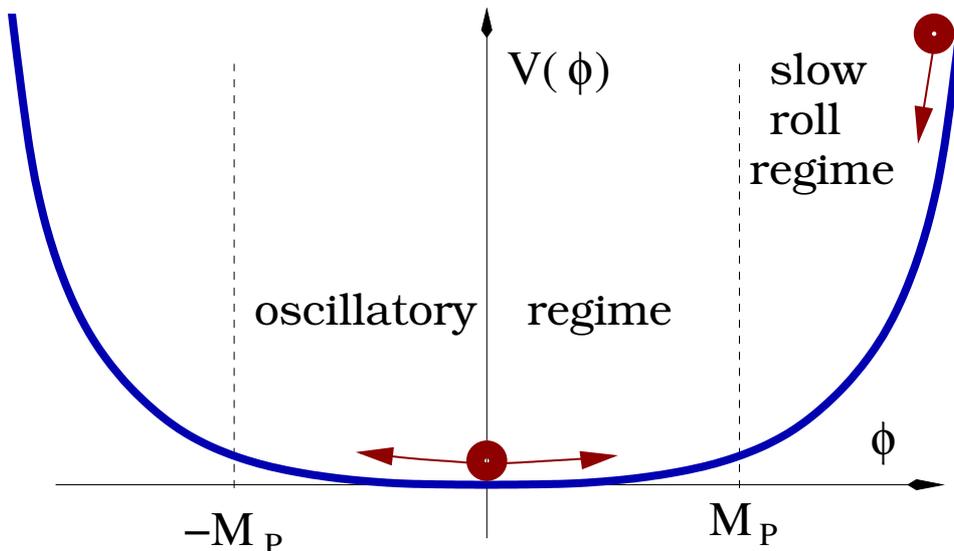


FIG. 25: Potential of a chaotic inflationary model. Inflation begins at $\varphi \gg M_P$, and ends at $\varphi \sim M_P$. When $\varphi \leq M_P$, the inflaton oscillates and decays.

which is an extremely small number. Linde argues that this problem can be resolved by increasing the initial scale of inflation to a scale which is close to the Planck scale, and still sufficiently below the Planck scale, such that one can trust classical gravitation. The probability (168) would increase accordingly, resolving thus the problem of initial conditions for inflation.

Chaotic inflation still suffers from the problem of naturalness of the potential, in the sense that the potential required to give curvature perturbations which accord with large scale CMB anisotropies is, $m_\varphi \sim 10^{13}$ GeV and $\lambda_\varphi \sim 10^{-13}$. Both of these parameters seem fine tuned in the following sense. Since the inflaton potential is built such to couple primarily to gravitation, there is a natural dimensionless parameter, $(m_\varphi/M_P)^2 \sim 10^{-11}$, which is very small (close to zero). Likewise, $\lambda_\varphi \sim 10^{-13}$ is very close to zero, posing a naturalness question of such small couplings. Yet some argue that, given the fact that we do not understand the origin of inflaton potential, there is not naturalness question.

Finally, we note that, based on a relatively large deviation from scale invariance, which is in the $\lambda_\varphi \varphi^4$ model (see Problem 3.2 and Part IV),

$$n_s - 1 = -6\epsilon + 2\eta = -\frac{7}{1+N}, \quad \epsilon = \frac{2}{3}\eta = \frac{1}{1+N}, \quad N \sim 50 - 60 \text{ e-foldings}, \quad (169)$$

where, $n_s - 1 = d \ln[\mathcal{P}_\zeta]/d \ln(k)$ defines by the scalar spectral index, the model is disfavoured by the one year WMAP measurements (excluded at two standard deviation), representing thus a first popular inflationary model strongly disfavoured by observations. On the other hand, the $m_\varphi^2 \varphi^2$ model yields,

$$n_s - 1 = -6\epsilon + 2\eta = -\frac{4}{1+2N}, \quad \epsilon = \eta = \frac{1}{1+2N}, \quad (170)$$

and hence it is still a viable model of inflation.