The early Universe: Inflation and symmetry breaking

(Dated: July 15, 2014)

I. THE EXPANDING UNIVERSE

Until the beginning of the 20-th century the sky was assumed to be essentially immutable[1]. In 1916, right after formulating the his theory of General Relativity, Einstein applied it to the entire Universe, in what could be considered as the birth of modern theoretical cosmology. To do that, he generalized the length of the 4-interval (that is, the space-time metric) of special relativity, $ds^2 = -c^2\,dt^2 + dx^2 + dy^2 + dz^2$, to the most general form that agrees with the *cosmological principle*: the Universe should look the same in all places (spatial homogeneity) and in all directions (spatial isotropy).

A. The scale factor of the Universe

Here we will consider the following simplified version (that however agrees with observations) of such most general metric metric:

$$ds^2 = -c^2\,dt^2 + a(t)^2\,\left(dx^2 + dy^2 + dz^2\right)\,, \tag{1}$$

where $a(t)$ is a function of time (called the *scale factor* of the Universe) that will be determined by dynamics.

First, let us note that the cosmological principle, and the choice of the metric (1) selects a preferred class of frames connected by a spatial translation, those where the Universe looks homogeneous and isotropic. Technically we say that we have broken the full symmetry $SO(3,\,1)$ of the Lorentz group down to the $SO(3)$ group of spatial rotations. If we are not in one of such frames we will observe stars moving towards us along a specific direction. Then we can go to the preferred frame by performing a boost with a velocity equal to that of the stars.

The choice of the metric (1) has then the following meaning. If I have two test bodies "at rest" (that is, with respect to whom the Universe looks isotropic) that at a time t_1 are at a distance L, then at time t_2 the distance of the two bodies will be $L \times a(t_2)/a(t_1)$. Note that this definition allows to measure only $a(t_2)/a(t_1)$, while the absolute value of a at some given time is conventional (and usually set to be equal to unity for $t = t_0$, where t_0 is the current value of t). To complete the above definition in a rigorous way, one has to define the way one measures - at least in thought experiment - the distance between the two bodies. The standard way of doing this is to send a light ray from the first to the second body, have it reflected on the second body, and see how much time Δt it takes for the light ray to come back to the first body. Then the distance between the two bodies is $c\,\Delta t/2$. Note that this definition makes sense only as long as L is such that $\Delta t \ll t_1 - t_0$, otherwise one has to perform more complicated calculations. However this is sufficient for a rigorous operative definition of what we mean by the metric (1).

The first observational evidence that $a(t)$ is not constant and is actually increasing with time came in 1929 from the measurement of the redshift of the light from far galaxies performed by Edwin Hubble. Hubble effectively measured L as the wavelength of the light emitted by a distant galaxy. He found that the lightwaves emitted by far galaxies received more stretching than those emitted by nearby galaxies, that was evidence for an expanding Universe.

B. Friedmann's equation

The behavior of $a(t)$ is determined by solving Einstein's equations, that determine in general the behavior of geometry in the presence of matter. For the geometry (1) and assuming a homogeneous and isotropic distribution of matter with energy density ρ (energy per unit volume), the behavior of $a(t)$ is determined by the *Friedmann equation*

$$\frac{\dot{a}^2}{a^2} = \frac{8\,\pi\,G}{3}\,\frac{\rho}{c^2}\,, \tag{2}$$

---

[1] Even if already in the 19-th century the so called *Olbers' paradox* was showing that such an assumption might be wrong: if there were stars everywhere in an eternal and infinite Universe, then there would be a star along each line of sight from the earth, and therefore the night sky should be infinitely bright.

with G Newton's constant and an overdot denoting derivative with respect to the time t.

Let us consider a gas of N particles, all with the same mass m and the momentum p in a volume V. Then the total energy density of the system, computed at a reference time t_1 is $\frac{N}{V}\sqrt{m^2\,c^4+p^2\,c^2}$. At a later time t_2 the volume V will have increased by a factor $\left(a(t_2)/a(t_1)\right)^3$ whereas the momentum will have decreased by a factor $a(t_1)/a(t_2)$ (since $p \propto \lambda^{-1}$ with $\lambda \propto a$ the wavelength of the particle).

Therefore we have

$$\rho(t_2) = \begin{cases} \rho(t_1)\left(\frac{a(t_1)}{a(t_2)}\right)^3 & \text{nonrelativistic matter,} \qquad p \ll m\,c \\ \rho(t_1)\left(\frac{a(t_1)}{a(t_2)}\right)^4 & \text{relativistic matter,} \qquad p \gg m\,c \end{cases} . \tag{3}$$

It is customary to define an *equation of state parameter* $w \equiv p/\rho$, such that $w=0$ for nonrelativistic matter and $w=1/3$ for ultrarelativistic matter. Then the relations (3) can be generalized to

$$\rho(t) = \rho(t_0)\,\left(\frac{a(t_0)}{a(t)}\right)^{3(1+w)} , \tag{4}$$

for w =constant, where we have set t_0 as the current value of t. Note that eq. (4) implies that for $w < -1$ the energy density of matter would increase as the Universe expands, violating even the weakest requirements of energy conservation one can think of. We will therefore assume, as customary, that $w \geq -1$.

C. Solving Friedmann's equation – the Big Bang

Insterting eq. (4) into eq. (2) and solving for $a(t)$ we obtain

$$a(t) = \left[a(t_0)^{\frac{3}{2}(1+w)} + \frac{3}{2}\,(1+w)\;H_0\,(t-t_0)\right]^{\frac{2}{3(1+w)}} , \tag{5}$$

where we have defined

$$H_0 = \sqrt{\frac{8\,\pi\,G}{3}\,\frac{\rho(t_0)}{c^2}\,a(t_0)^4}\,. \tag{6}$$

Eq. (5) shows that there is a value of t for which a crosses zero. This is the *Big Bang singularity*. It is a mathematical that is present only if we assume that the Friedmann equation (2) is unchanged as $a \to 0$, that is for diverging energy densities. Since we expect that Einstein gravity receives corrections in the presence of very high energy densities, this assumptions is not guaranteed, and the Big Bang singularity should be seen as a merely mathematical singularity, the physical singularity being taken care by unknown physics taking over at high energy scales. It is convenient to consider

By defining t_0 in such a way that the singularity in eq.(5) accurs at $t=0$, we obtain

$$a(t) = \left(\frac{3}{2}\,(1+w)\;H_0\,t\right)^{\frac{2}{3(1+w)}} , \tag{7}$$

In particular,

$$a(t) = \begin{cases} \left(\frac{3}{2}\,H_0\,t\right)^{\frac{2}{3}} & \text{nonrelativistic matter,} \qquad w=0 \\ \left(2\,H_0\,t\right)^{\frac{1}{2}} & \text{relativistic matter,} \qquad w=1/3 \end{cases} . \tag{8}$$

II. THE HORIZON PROBLEM

Let us now see to what amounts the horizon problem, that is one of the main motivations for inflation. We will first see a heuristic argument and then a rigorous one.

A. A heuristic derivation

Figure 1 shows a section of the sky observed in a large galaxy survey. We are at the center of the plot and the distance from center of each dot corresponds to the redshift z of each galaxy. The redshift z is related to the time $t_{\rm em}$ at which the light we see was emitted by the galaxy by $1 + z = (a(t_0)/a(t_{\rm em}))$ with t_0 the current age of the Universe, about 14 billion years. Using the expression (8) during matter domination (not quite an accurate approximation for the current Universe, but still good for our purposes) we obtain $t_{\rm em} = t_0/(1+z)^{3/2}$.

Therefore a galaxy at $z \simeq 1$ was the way we see it when the Universe was a factor $2^{-3/2} \simeq .35$ younger than now, that is when the Universe was only about 5 billion years old. Let us forget, for the time being, the expansion of the Universe and let us simply assume that the Universe was created infinitely extended and not expanding at $t = 0$. Then the galaxy G_1 in the left pane of figure 1 emitted its light at $t = 5$ billion years. This light reached us after a trip of $14 - 5 = 9$ billion years. At the time of the emission of light, only a a circle of radius 5 billion light years from G_1 knows about its existence, since no information can travel faster than light and no light could be emitted by G_1 before $t = 0$, that is 5 billion years before. However we see galaxies on the opposite side of the sky, in particular galaxy G_2, that are at a distance $9 \times 2 = 18$ billion light years from G_1. Galaxies around G_2 look pretty much like galaxies around G_1 even if, at the time at which we are seeing them, they could not know about each other.

Now if the Universe was created by some random phenomenon there is no reason why causally disconnected region should look the same way. This is the *horizon problem*: why parts of the Universe that in the past were beyond each other's horizon seem to know about each other, to the point that they look very similar?

Two comments about the above argument: *(i)*, maybe the fact that we are ignoring the expansion of the Universe is an exceedingly rough approximation and *(ii)* after all this is a problem of initial conditions, and there is nothing bad by assuming that the Universe was set-up to look very homogeneous from the beginning. Concerning *(i)*, as we will see below, one can make a rigorous statement and show that the problem is also present in a matter- or radiation-dominated expanding Universe. However *(ii)* is a valid objection. Yet, "coincidences" sometimes mean that Nature is trying to tell us something. As we will see, inflation, invented to solve the horizon problem – or, if one prefers – to explain the coincidence of a homogeneous yet apparently causally disconnected Universe, will not only solve this problem, but make other predictions that are tested experimentally.

B. A rigorous derivation

A rigorous derivation of the horizon problem is the following. Let us assume that the Universe is radiation dominated (this is a good approximation for the Early Universe, since the energy densities were much larger then, leading to larger momenta and relativistic particles). Then according to the derivation of the previous section $a(t) = (2\, H_0\, t)^{1/2}$. It is now convenient to define a new time coordinate τ, called *conformal time*, defined as

$$\tau(t) - \tau_1 \equiv \int_{t(\tau_1)}^{t} \frac{dt'}{a(t')}\,. \tag{9}$$

The advantage of this definition is that in terms of conformal time the metric reads

$$ds^2 = a(\tau)^2\,\left(-c^2\,d\tau^2 + dx^2 + dy^2 + dz^2\right)\,, \tag{10}$$

that is the same form as the Minkowski metric times an overall function of time. Since lightrays move along lines for which $ds^2 = 0$, one can apply to the expanding Universe the techniques of Minkowski diagrams, that are usually used to study the causality in flat, infinite space.

What is the function $a(\tau)$? Let us compute it explicitly in the case of a radiation dominated Universe. We have then, from eq. (9)

$$\tau = \sqrt{\frac{2\,t}{H_0}}\,, \qquad \text{radiation domination} \tag{11}$$

where we have set the integration constant to have $\tau = 0$ at the Big Bang $t = 0$. By substituting $t(\tau)$ into $a(t)$ we find

$$a(\tau) = H_0\,\tau\,. \tag{12}$$

The latter equation has an important implication: the time τ does not extend to $\tau < 0$. This implies that, when drawing Minkowski diagrams for the spacetime metric (10) we can only use the top portion of the diagram. This has

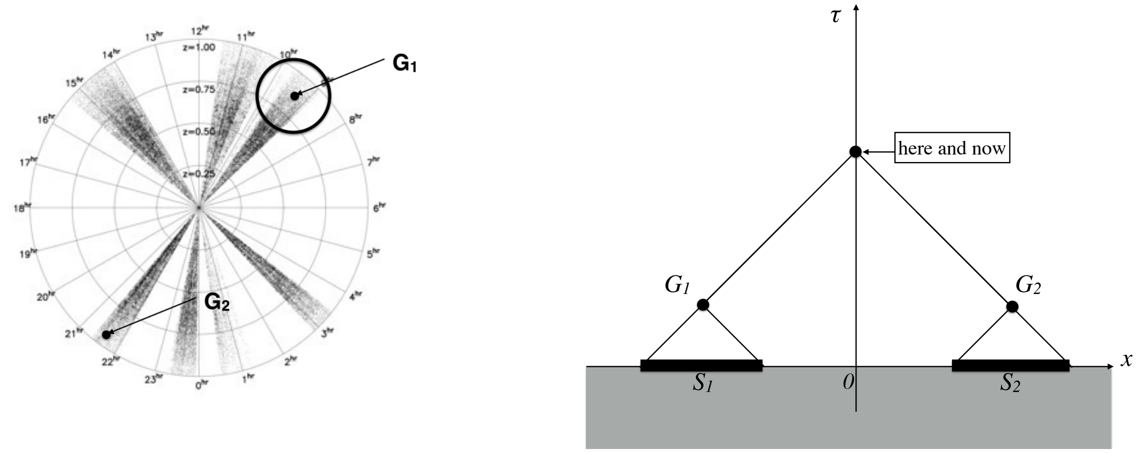

FIG. 1: Left: a map of the sky from the BOSS galaxy survey (each dot is a galaxy). The galaxies G_1 and G_2 are shown surrounded by the area that was in causal contact with them at the time they emitted the light we are seeing today. The two causally connected regions do not overlap, and therefore we should not expect to have similar properties (same density of galaxies etc...). Right: a Minkowski diagram showing the situation described in the left panel.

important implications for the causal structure of the spacetime. Suppose we live at $x = 0$ in the Minkowski diagram of figure 2 and the event "observation of light from galaxies G_1 and G_2" is associated to the point marked "here and now" in the diagram. The lines connecting G_1 and G_2 to the point "here and now" are at 45° in the diagram, since they are associated to lightrays.

Now the points G_1 and G_2 can be affected only from the regions S_1 and S_2 respectively since they cannot be affected by anything traveling faster than light. Therefore there is no reason *a priori* why the region around G_1 should look similar to that around G_2, since they have never been in causal contact.

III. INFLATION AS A SOLUTION OF THE HORIZON PROBLEM

By looking at figure 2 we see how to solve the horizon problem. If we were allowed to extend the diagram below the line $\tau = 0$ then the lines below starting from G_1 and G_2 and extending downwards would cross at some points, determining a region of spacetime that in the past of both G_1 and G_2 and that could have created the conditions form which G_1 and G_2 can be similar.

To see whether it is possible to perform such an extension of the Minkowski diagram, let us calculate $a(\tau)$ for arbitrary values of w. We will not keep all of the w-dependence in what follows, but only the main features that are necessary to outline the argument. Form eqs. (9) and $a(t) \sim t^{\frac{2}{3\,(1+w)}}$ we obtain

$$\tau \simeq \frac{3\,(1+w)}{1+3\,w}\, t^{\frac{1+3\,w}{3\,(1+w)}} + \text{constant}\,. \tag{13}$$

Now, as long as $1+3\,w > 0$, we can set $\tau = 0$ for $t = 0$ and proceed as in the case of a radiation dominated Universe discussed in the previous section, obtaining that τ cannot be negative.

However, when $1+3\,w < 0$, one has that $\tau \to -\infty$ as $t \to 0$: the Big Bang singularity in this case is mapped to $\tau \to -\infty$ and τ does in this case extend to negative values! Therefore the solution to the horizon problem is simply to glue, in figure 2, a period of expansion with $w < -1/3$ below the radiation dominated epoch.

Note that, from $a \propto t^{\frac{2}{3\,(1+w)}}$, it follows that the condition $w < -1/3$ corresponds to the condition $\ddot{a} > 0$. Therefore we will have the following general definition:

Inflation is a period of accelerated expansion

.

Besides leading to the solution of the horizon problem, inflation erases any primordial inhomogeneity. In fact, while in Minkowski space a region containing a matter overdensity tends to attract more matter and therefore to become even more overdense, the expansion of the Universe counteracts this effect by pulling particles away from each other. Inflation is maximally efficient at this, leaving a smooth Universe after a sufficiently long period of accelerated expansion.

IV. SLOW ROLL INFLATION

From here on we will use "natural" units $c = \hbar = 1$. We will also trade Newton's constant for the *reduced Planck mass* M_P defined as

$$M_P \equiv \frac{1}{\sqrt{8\,\pi\,G}} \simeq 2.4 \times 10^{18}\ \mathrm{GeV}\,. \tag{14}$$

A. The inflaton

The simplest example of matter leading to an inflating Universe is a *cosmological constant*, defined as a form of matter with $w = -1$. The solution of Friedmann equation is in the case given by $a(t) = e^{H_0\,t}$ with $H_0 = \sqrt{\rho/3\,M_P^2}$ =constant, where we have set the scale factor equal to unity at $t = 0$. In terms of conformal time τ we have $a(\tau) = -\frac{1}{H_0\,\tau}$, where $a = 1$ at $\tau = -1/H_0$.

A cosmological constant represents an excellent source of inflation. Unfortuntely it is an overkill, since inflation is so perfect in this case that it never ends!

What we really need is slow roll inflation: a system that behaves almost like a cosmological constant, but that in a sufficiently long time leads to the end of inflation. The most studied example of such a system is provided by a scalar field ϕ with a potential $V(\phi)$. In order for ϕ to mimic a cosmological constant, $V(\phi)$ must be extremely flat, at least for some range of values of ϕ. A typical potential for the inflaton is plotted in figure 3.

Analogously to a one-dimensional problem in classical mechanics, the field ϕ has a energy density

$$\rho_\phi = \frac{\dot\phi^2}{2} + V(\phi)\,, \tag{15}$$

and satisfies the Klein-Gordon equation in an expanding Universe

$$\ddot\phi + 3\,\frac{\dot a}{a}\,\dot\phi + \frac{dV(\phi)}{d\phi} = 0\,. \tag{16}$$

Now in order for the field to mimic a cosmological constant, the kinetic energy must be negligible with respect to the potential energy: $\dot\phi^2 \ll 2\,V(\phi)$. Moreover this condition must be maintained for long enough, that is, the first term in eq. (16) must be negligible with respect to the remaining two.

This is the so-called *slow-roll approximation* where the equations of motion reduce to

$$\begin{aligned} 3\,H\,\dot\phi + \frac{dV(\phi)}{d\phi} &\simeq 0\,,\\ H^2 &\simeq \frac{V(\phi)}{3\,M_P^2}\,, \end{aligned} \tag{17}$$

that can be solved by quadratures for any given $V(\phi)$.

It is usual to define two *slow roll parameters*

$$\epsilon \equiv \frac{M_P^2\,(dV/d\phi)^2}{2\,V(\phi)^2}\,, \qquad \eta \equiv \frac{M_P^2\,(d^2V/d\phi^2)}{V(\phi)}\,, \tag{18}$$

such that, using the slow-roll equations (17) we obtain

$$\frac{\dot\phi^2}{2\,V} \simeq \frac{\epsilon}{3}\,, \qquad \frac{\ddot\phi}{dV/d\phi} = \frac{\eta - \epsilon}{3} \tag{19}$$

so that the slow roll approximation boils down to the slow roll conditions

$$\epsilon \ll 1, \qquad |\eta| \ll 1\,. \tag{20}$$

It is possible to show that one has accelerated expansion (that is, inflation) if and only if the slow-roll conditions are satisfied. The advantage of the formulation (20) of the slow-roll conditions is that they depend only on the potential and not on the dynamics. For instance for a monomial potential $V(\phi) \propto \phi^a$ we have $\epsilon = n^2\,M_P^2/(2\,\phi^2)$, $\eta = n\,(n-1)\,M_P^2/\phi^2$ so that inflation occurs as long $|\phi| \gg M_P$, and the larger $|\phi|$ the better the slow-roll conditions are satisfied.

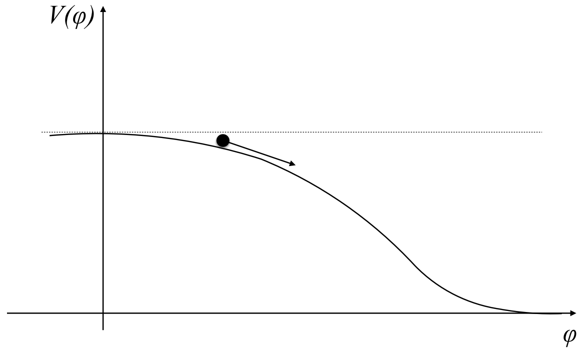

FIG. 2: A typical shape of an inflationary potential, with a flat region able to support inflation and a minimum where $V \simeq 0$. For comparison, the thinner line represents a cosmological constant: good to do inflation (it approximates the potential $V(\phi)$ in the region where inflation occurs), but not good to describe the vacuum today.

B. Solving the slow-roll equations

During inflation, it is useful to use, instead of time, a a variable known as the *number of efoldings* N_e *from the end of inflation*. It is defined as

$$N_e \equiv \log\left(\frac{a_{\rm end}}{a(t)}\right) , \tag{21}$$

where $a_{\rm end}$ is the scale factor of the Universe at the end of inflation. It is easy to show that

$$N_e(t) = \int_t^{t_{\rm end}} \frac{\dot{a}}{a}\, dt \tag{22}$$

By trading t for ϕ as an integration variable via $dt = d\phi/\dot{\phi}$ and by using the slow roll equations (17) one can write

$$N_e(\phi) = \frac{1}{M_P^2} \int_{\phi_{\rm end}}^{\phi} \frac{V(\varphi)}{dV(\varphi)/d\varphi}\, d\varphi \tag{23}$$

that allows to determine, for any given potential $V(\phi)$, the value of the inflaton ϕ corresponding to a given number of efoldings. In particular, the solution of the horizon problem requires about 60 efoldings of inflation. Nevertheless most models yield many, many more efoldings of inflation.

C. The end of inflation

How does inflation end? At some point the slow-roll conditions are not satisfied any more, and the field ϕ reaches the bottom of its potential, around which performs several oscillations while decaying into lighter particles. At some point the energy density in the Universe is dominated by these decay products of the inflaton rather than by the inflaton itself. Once such decay products thermalize, we can say that the process of *reheating after inflation* is concluded, and we enter the radiation dominated era of the Hot Big Bang.

V. INFLATION, SYMMETRY BREAKING, AND THE STANDARD MODEL

The fact that the inflaton ϕ can take large excursions implies that it can lead to phenomena similar to the symmetry breaking present in the Standard Model. For instance, if the inflaton is electrically charged, then during inflation it will have a non vanishing expectation value, and electromagnetism will be broken ($\Rightarrow$ no massless photons). Of course, since today photons are massless, the minimum of the potential of such a charged inflaton should lie at $\phi = 0$.

Usually, however, the inflaton is considered to be a gauge singlet (even if there is no real reason for this to be the case). As a consequence it can couple to the Standard model either through the Higgs mass term or through dimension-5 and higher operators. Given the fact that the inflaton can get very expectation values, one in general expects that the Lagrangian of the Standard Model gets large corrections during inflation.

For instance, adding to the standard Higgs potential

$$V_{\text{Higgs}} = \frac{\lambda}{4}\left(|h|^2 - v^2\right)^2 , \tag{24}$$

a small coupling to the inflaton ϕ of the form

$$\delta V_{\text{Higgs}} = \frac{\lambda'}{2}\,|h|^2\,\phi^2 , \tag{25}$$

gives a contribution to the Higgs potential $\sim \lambda'\,|h|^2\,M_P^2$ during inflation, when $\phi \gtrsim M_P$, that, as long as $\lambda' \gtrsim v^2/M_P^2 \simeq 10^{-32}$ dominates the standard negative quadratic term of the Higgs potential and leads to restoration of the electroweak symmetry during inflation if $\lambda' > 0$. Of course, if $\lambda' < 0$, the the electroweak symmetry is broken even more during inflation.

Because of this, the behavior of the Standard Model during inflation is extremely model dependent: a problem for predictivity, but an advantage for model builders!

VI. COSMOLOGICAL PERTURBATIONS DURING INFLATION

For the time being we have focused on the dynamics of the homogeneous inflating Universe. Now we will discuss the evolution of the quantum mechanical fluctuations of the inflaton. It is convenient to start our discussion by considering a test massless scalar field in an expanding Universe.

The action for such a field is obtained by the standard action for a massless scalar on Minkowski space $\mathcal{S} = \frac{1}{2}\int d^3x\,dt\left(\dot{\psi}^2 - (\nabla\psi)^2\right)$ by replacing dx by $a\,dx$. It is also convenient to use conformal time, that is to replace dt by $a\,d\tau$ everywhere. Let us also Fourier transform in the three spatial dimensions so that we use

$$\tilde{\psi}(\mathbf{k},\,\tau) = \int \frac{d^3\mathbf{x}}{(2\,\pi)^{3/2}} e^{-i\,\mathbf{k}\,\mathbf{x}}\,\psi(\mathbf{x},\,\tau) \tag{26}$$

Then we obtain the action

$$\mathcal{S} = \frac{1}{2}\int d^3\mathbf{k}\,d\tau\,a^4(\tau)\left(\frac{\tilde{\psi}'^2}{a(\tau)^2} + \frac{k^2}{a(\tau)^2}\,\tilde{\psi}^2\right) , \tag{27}$$

where we denote by a prime the derivative with respect to conformal time: $' \equiv \frac{d}{d\tau}$.

To proceed we define a new field $\tilde{\psi}_c(\mathbf{k},\,\tau) \equiv a(\tau)\,\tilde{\psi}(\mathbf{k},\,\tau)$ that is canonically normalized (that is, the coefficient of the kinetic term is $\frac{1}{2}$). The field $\tilde{\phi}_c$ satisfies the following equation

$$\tilde{\psi}_c'' + k^2\,\tilde{\psi}_c - \frac{a''}{a}\,\tilde{\psi}_c = 0 , \tag{28}$$

that we can rewrite in a more suggestive form as

$$-\tilde{\psi}_c'' + \frac{a''}{a}\,\tilde{\psi}_c = k^2\,\tilde{\psi}_c , \tag{29}$$

that is precisely the same form of the time-independent Schrödinger equation of a "particle" on mass $m = 1/2$, once we identify k^2 with the energy and a''/a with the potential. The main difference with respect to the standard Schrödinger equation is that in our case the "coordinate" is (conformal) time, rather than space.

We assume for simplicity that the equation of state parameter during inflation is exactly $w = -1$, and that inflation is immediately followed by a period of radiation domination, so that

$$a(\tau) = \begin{cases} (-H_i\,\tau)^{-1} , & \text{for} -\infty < \tau < \tau_{\text{end}} < 0 \\ a(\tau_{\text{end}}) + H_r\,(\tau - \tau_{\text{end}}) & \text{for} \tau > \tau_{\text{end}} \end{cases} , \tag{30}$$

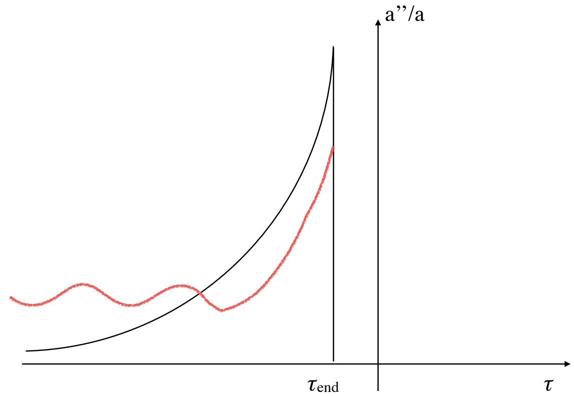

FIG. 3: The "effective potential" a''/a for a period of inflation followed by an epoch of radiation domination. Also, a schematic representation of a wave function $\tilde{\psi}_c(\tau)$.

with H_i and H_r constants. Therefore the "potential" a''/a equals $2/\tau^2$ for $\tau < \tau_{\rm end}$ and is vanishing for $\tau > \tau_{\rm end}$, as shown in figure 3.

By looking at figure 3, we see that the Schrödinger problem is analogous to that of scattering on/tunneling through a potential barrier, with a crucial difference: in the usual problem of tunnelling through a barrier, one assumes an incoming and a reflected wave on the left of the wall as well as a trasmitted wave to the right, and flux conservation implies that the transmitted wave cannot be larger to the incident one. In our case, we cannot have a reflected wave, since this would imply that the modes go "backwards in time". Therefore we cannot impose flux conservation and the outgoing wave for $\tau > \tau_{\rm end}$ can be much larger than the incoming one. This amplification of the incoming wave is interpreted as the generation of large, classical fluctuations of the ϕ field starting from its quantum, vacuum fluctuations.

Lte us now solve eq. (29) in the presence of the potential of figure 4. An exact solution is possible, however it is more instructive to find an approximate solution. We see immediately that we can solve the equation in three regions:

1. in the first region τ large and negative and the potential is negligible with respect to k^2. This first region ends at $\tau \simeq -1/k$;

2. in the intermediate region $-1/k \lesssim \tau < \tau_{\rm end}$ the "potential" dominates over the k^2 term;

3. for $\tau > \tau_{\rm end}$ there is no potential and the evolution is controlled by the term in k^2 again. This part of evolution is not that interesting and we will not care about it here.

Now in the region 1. the solution of the Schr̈odinger equation will be given by plane waves $\tilde{\psi}_c = \tilde{\psi}_c^1\, e^{-ik\tau}$, with $\tilde{\psi}_c^1$ the overall normalization of the wave function that we will determine below. In region 2., by neglecting the contribution from k^2 on the right hand side of eq. (29) we obtain that the growing solution is simply $\tilde{\psi}_c = \tilde{\psi}_c^2\, a(\tau)$, where the constant $\tilde{\psi}_c^2$ is determined by imposing that $\tilde{\psi}_c$ is continuous when we go from region 1. to region 2. Therefore, up to a phase, $\tilde{\psi}_c^2 = \tilde{\psi}_c^1/a(\tau = -1/k)$. After the end of inflation the evolution of $\tilde{\psi}_c$ is trivially given by oscillations, and we will not care about it.

How do we determine the normalization $\tilde{\psi}_c^1$? Going ahead with our quantum mechanical analogy, the functions $\tilde{\psi}_c(\mathbf{k},\, \tau)$ can be seen as wave functions beloning to different eigenvalues k^2 of the "energy". As such they are orthogonal, so that we can set $\langle \tilde{\psi}_c(\mathbf{k}_1) | \tilde{\psi}_c(\mathbf{k}_2) \rangle \propto \delta^{(3)}(\mathbf{k}_1 - \mathbf{k}_2)$. The proportionality constant can be determined by dimensional analysis: since ψ has mass dimension 1, $\tilde{\psi} \sim (\text{mass})^{-3}\, \psi$ must have mass dimension -2. Since $\delta^{(3)}(\mathbf{k}_1 - \mathbf{k}_2)$ has mass dimension -3, then the proportionality constant must have mass dimension -1. We fix the constant at early times $\tau \to -\infty$ when the system is supposed to be in vacuum. In this regime the only dimensionful scale is k_1 (that equals k_2 courtesy of the Dirac δ function). Therefore we must have

$$\langle \tilde{\psi}_c(\mathbf{k}_1,\, \tau \to -\infty) | \tilde{\psi}_c(\mathbf{k}_2,\, \tau \to -\infty) \rangle = C\, \frac{\delta^{(3)}(\mathbf{k}_1 - \mathbf{k}_2)}{k_1}\,. \tag{31}$$

with C a dimensionless constant of order unity (that turns out to be equal to $1/2$ after a more rigorous analysis - therefore we will set $C = 1/2$ from here on).

By the analysis above, at the end of inflation we have that $\tilde{\psi}_c(\mathbf{k},\,\tau_{\rm end}) \simeq \tilde{\psi}_c(\mathbf{k},\,\tau \to -\infty)\, a_{\rm end}/a(\tau = -1/k)$, so that, remembering that $\tilde{\psi}_c \equiv a(\tau)\,\tilde{\psi}$,

$$\langle \tilde{\psi}(\mathbf{k}_1,\,\tau_{\rm end}) | \tilde{\psi}(\mathbf{k}_2,\,\tau_{\rm end}) \rangle = \frac{1}{a_{\rm end}^2}\,\frac{\delta^{(3)}(\mathbf{k}_1 - \mathbf{k}_2)}{2\,k_1}\,\frac{a_{\rm end}^2}{a(\tau = -1/k)^2} = \frac{H_i^2}{2\,k_1^3}\,\delta^{(3)}(\mathbf{k}_1 - \mathbf{k}_2)\,, \tag{32}$$

where we have used $a(-1/k) = k/H_i$.

The main observable we care about is the *two point function*, in coordinate space, of the field ϕ. We can define this as[2]

$$\langle \psi(\mathbf{x},\,\tau_{\rm end})) | \psi(\mathbf{y},\,\tau_{\rm end})) \rangle = \int \frac{d^3\mathbf{k}_1\,d^3\mathbf{k}_1}{(2\pi)^3} e^{-i\mathbf{k}_1\mathbf{x} + i\mathbf{k}_2\mathbf{y}} \langle \tilde{\psi}(\mathbf{k}_1,\,\tau_{\rm end}) | \tilde{\psi}(\mathbf{k}_2,\,\tau_{\rm end}) \rangle = \int \frac{d^3\mathbf{k}}{4\pi\,k^3}\left(\frac{H_i}{2\,\pi}\right)^2 e^{-i\mathbf{k}(\mathbf{x}-\mathbf{y})} \tag{33}$$

that shows that the two point function of ϕ has the important property of being *scale invariant*: a rescaling of $|\mathbf{x} - \mathbf{y}|$ by a constant factor can be absorbed by a rescaling in the integration variable k and does not affect the two point function.

In particular, by taking $\mathbf{x} = \mathbf{y}$ we obtain the variance

$$\langle \psi(\mathbf{x},\,\tau_{\rm end})) | \psi(\mathbf{x},\,\tau_{\rm end})) \rangle = \int \frac{dk}{k}\left(\frac{H_i}{2\,\pi}\right)^2\,, \tag{34}$$

where the integral is on all the momenta the underwent amplification, that is $H_i\,a_{\rm in} < k < H_i\,a_{\rm end}$ with $a_{\rm in}$ and $a_{\rm end}$ the initial and the final scale factor of inflation respectively.

The above result (34) is extremely important. It tell that a massless field in an inflating space with Hubble parameter H experiences a random amplification of its quantum fluctuations whose sign is undefined, whereas its typical amplitude is $\frac{H}{2\pi}$ per logarithmic interval in k.

Now how does this convert into observables? Up to now we have assumed that ψ was a massless test field. However our whole analysis is valid also if ψ denotes the fluctuation of the inflaton ϕ about its background value: $\psi = \delta\phi$.

Let us now consider two distant regions of the Universe right at the end of inflation. Inflation will end when the inflaton ϕ reaches some given value $\phi_{\rm end}$ when the slow roll conditions are violated. If $\phi = \phi_{\rm end}$ in one region of reference, then, as a consequence of the mechanism discussed above, we will have $\phi = \phi_{\rm end} - \delta\phi$ in a different region, with $\delta\phi \sim H_i/2\pi$. Now in the region where $\phi = \phi_{\rm end} - \delta\phi$ inflation will end a bit later, by an amount of time $\delta t \simeq \delta\phi/\dot{\phi}_0$. The scale factor a during inflation is given approximately by $a \propto \exp\{H_i\,t\}$ so that the region where inflation end later will have expanded for an extra time δt and will be slightly larger, by a factor $\simeq (1 + H_i\,\delta t) \simeq (1 + H_i\,\delta\phi/\dot{\phi}_0)$, than our reference region. As a consequence energy density of radiation will be smaller, by a factor $\simeq (1 + H_i\,\delta\phi/\dot{\phi}_0)^{-4}$.

We conclude therefore that *quantum fluctuations during inflation generate fluctuations of the energy density in matter at the end of inflation*. This implies that the Universe at the end of inflation is now perfectly homogeneous, but on the contrary has small inhomogeneities. These inhomogeneities, whose initial amplitude is of

$$\frac{\delta\rho}{\rho} \simeq \frac{H_i}{\dot{\phi}}\,\delta\phi \simeq \frac{H_i^2}{2\,\pi\,\dot{\phi}} \tag{35}$$

eventually evolve and grow during the subsequent expansion of the Universe, and give rise to the galaxies, clusters of galaxies, stars and planets that now we inhabit.

An important property of these fluctuations, as we stated above, is that their spectrum is scale invariant. One can check that structures we inhabit originate from a (quasi) scale invariant spectrum of perturbations by taking a picture of the early Universe, that is by mapping the Cosmic Microwave Background radiation. This also allows to measure the amplitude of the fluctuations $\delta\rho/\rho$. Using eq. (35) above we obtain the relation

$$\frac{H^2}{2\,\pi\,\dot{\phi}} \simeq 5 \times 10^{-5} \tag{36}$$

where H and $\dot{\phi}$ have to be evaluated about 60 efoldings before the end of inflation. This relation allows to determine a relation between the parameters in the inflationary potential.

---

[2] Notice that since we are studying the amplitude of fields, and not probabilities, the observable we care about is the amplitude $\langle ...|... \rangle$ rather than the probability $|\langle ...|... \rangle|^2$.

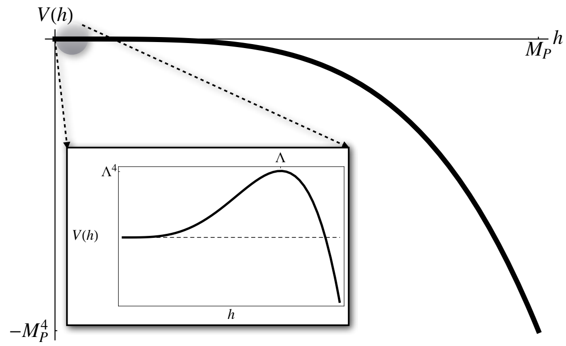

FIG. 4: A schematic representation of the Higgs potential after taking into account loops from the top quark.

VII. INFLATION AND THE HIGGS

Let us conclude with a natural question: could the Higgs field be the inflaton? At large values of the Higgs field h its potential is given by $V(h) \simeq \frac{\lambda}{4} h^4$. This is a monomial potential that, as we have seen above, can satisfy the slow-roll conditions if $|h| \gg M_P$. Now how about the fluctuations? Let us just perform an order-of-magnitude estimate. Observations give $\delta\rho/\rho \simeq 10^{-5}$. We have shown that $\delta\rho/\rho \simeq H^2/\dot{\phi}$. During slow-roll $\dot{\phi} \simeq (dV/d\phi)/H$, so that $\delta\rho/\rho \simeq H^3/(dV/d\phi)$. Using again the slow-roll condition $H \simeq \sqrt{\lambda}\,\phi^2/M_P$ we eventually find $\delta\rho/\rho \simeq \lambda^{1/2}\,\phi^3/M_P^3$. Since $\phi \gg M_P$, the obtain that $\lambda \ll (\delta\rho/\rho)^2 \simeq 10^{-10}$. Finally, since the mass of the Higgs is related to λ via $m_H \simeq \sqrt{\lambda}\,v$ with $v \simeq 100$ GeV, we obtain that for the Higgs to drive inflation we would need $m_H \ll 1$ MeV, that clearly is ruled out. Therefore the Standard Model Higgs cannot be the inflaton.

For the currently measured values of the mass of the Higgs and of the top quark it appears that the Higgs potential $V(h)$ turns negative at $h = \Lambda \simeq 10^{12}$ GeV and is unbounded from below (see figure 4). Therefore if for some reason $|h|$ gets larger than Λ during or after inflation, then the Higgs will end in a vacuum different from ours, leading to a Universe completely different from ours - possibly leading to a collapsing Universe that falls down an infinitely deep vacuum with negative energy. An unpleasant outcome of inflation.

While classically one could just invoke the possibility that we started with $|h| < \Lambda$, quantum mechanically we must take into account the fluctuations of the Higgs field. Based on the discussion above, these fluctuations are generated with a typical amplitude $\langle \delta h^2 \rangle = (H_i/2\pi)^2$ per efolding of inflation. Therefore even if we start inflation with $h = 0$, eventually the value of H will drift away. The exercise below will allow to make a quantitative statement on the mass of the Higgs based on the requirement that inflation does not lead to the end of the world by bringing the Higgs to its unstable region.

EXERCISE

Chaotic inflation and the mass of the Higgs. The simplest model of inflation is known as *chaotic inflation* and its potential has the form $V(\phi) = \frac{m^2}{2}\,\phi^2$.

We will assume that this is the correct model of inflation, and we will infer a constraint on the mass of the Higgs by assuming that quantum fluctuations during inflation do not bring the Higgs field to the region where its potential is negative.

To do this:

1. Compute the mass m of the inflaton: assuming that inflation ended at $\phi = M_P$, use the slow-roll approximation to compute the value of ϕ 60 efoldings before the end of inflation. Use this result to compute the mass of the inflaton after imposing

$$\frac{\delta\rho}{\rho} = \frac{H^2}{2\,\pi\,\dot{\phi}} = 5\times 10^{-5}\,. \tag{37}$$

2. As seen above, the variance $\sqrt{\langle h^2\rangle}$ of a test field h and the end of inflation is given by

$$\langle h(\mathbf{x},\,\tau_{\rm end}))|h(\mathbf{x},\,\tau_{\rm end}))\rangle = \int_{a_{\rm in}}^{a_{\rm end}} \frac{dk}{k}\,\left(\frac{H_i}{2\,\pi}\right)^2\,, \tag{38}$$

   From this you can compute, as a function of Λ, the probability that the Higgs field (assumed to be a massless field, that is most probably a good approximation in this context) takes a expectation value between $-\Lambda$ and Λ. Compute the value of Λ_{95}, defined as the value of Λ for which there is 95% probability that the Higgs is in the region $-\Lambda_{95} < h < \Lambda_{95}$, assuming $\langle h\rangle = 0$.

3. The one-loop corrected Higgs potential reads approximately

$$V(h) \simeq \frac{\lambda}{4}\,\left(h^2 - v^2\right)^2 - \frac{3}{8\,\pi^2}\,y_t^4\,h^4\,\log\frac{h}{v}\,, \tag{39}$$

   where $v = 246$ GeV and y_t is the top quark Yukawa coupling. It is clear that for large values of h the potential turns negative. Estimate, as a function of the mass of the Higgs m_h in the electroweak vacuum the critical value h_c of h where this happens.

   [Neglect, for simplicity, the correction from the top quark to the Higgs potential in the calculation of m_h].

4. From the above expression of Λ_{95} derive a bound of the mass of the Higgs by requiring that $h_c < \Lambda_{95}$. That is, that inflation had only 5% probability to bring the Higgs into the unstable region.

5. Compare your result with the observed mass of the Higgs. Is there a contradiction? If yes, how would try to change things to obtain an agreement between the measured mass of the Higgs and the theoretical bound from vacuum stability during inflation?

SOLUTION

1. Slow roll approximation gives

$$N_e = \frac{1}{M_P^2} \int_{M_P}^{\phi} \frac{V(\phi')}{V'(\phi')} d\phi' = \frac{1}{4\,M_P^2} \left(\phi^2 - M_P^2\right) \,, \tag{40}$$

so that

$$\phi = 2\,M_P \sqrt{\frac{1}{4} + N_e} \simeq 15.5\,M_P \tag{41}$$

where in the last equality we have used $N_e = 60$.

Now in slow roll approximation $|\dot{\phi}| = |V'(\phi)/3\,H$, so that

$$5 \times 10^{-5} = \frac{\delta\rho}{\rho} = \frac{H^2}{2\,\pi\,\dot{\phi}} \simeq \frac{3\,H^3}{2\,\pi\,V'} = \frac{1}{4\,\pi\,\sqrt{6}} \frac{m\,\phi^2}{M_P^3} \tag{42}$$

where we have used $H \simeq m\,\phi/\sqrt{6}\,M_P$ from the slow relations. From this equation we obtain, using the result (41) above, $m = \frac{\pi\,\sqrt{6}}{1/4+N_e} \times 5 \times 10^{-5}\,M_P \simeq 1.5 \times 10^{13}$ GeV.

2. In first approximation, one can just assume $H_i \simeq$constant during inflation, as set it equal to its value at 60 efoldings before the end of inflation. In this case the integral is simply $\int_{H\,a_{\rm in}}^{H_i\,a_{\rm end}} \frac{dk}{k} = \log(\frac{a_{\rm end}}{a_{\rm in}}) = N_e = 60$. Taking H computed for the value of phi given in eq. (41) we obtain

$$\langle h^2 \rangle \simeq \frac{N_e}{4\,\pi^2} \frac{m^2}{6} \left(1 + 4\,N_e\right) \simeq \left(1.2 \times 10^{14}\ \text{GeV}\right)^2 \,. \tag{43}$$

A more accurate estimate can be obtained by using $dk/k = dN_e$ and $H^2(N_e) = \frac{m^2}{6}\,(1 + 4\,N_e)$ so that

$$\langle h^2 \rangle \simeq \frac{m^2}{24\,\pi^2} \int_0^{N_e} dN_e' \left(1 + 4\,N_e'\right) = \frac{m^2}{24\,\pi^2} \left(N_e + 2\,N_e^2\right) \simeq \left(8.5 \times 10^{13}\ \text{GeV}\right)^2 \,. \tag{44}$$

Then Λ_{95} corresponds to a $2\,\sigma$ value, that is $\Lambda_{95} = 1.7 \times 10^{14}$ GeV.

3. The mass of the Higgs in vacuum is given by $m_H^2 = 2\,\lambda\,v^2 \simeq (126\ \text{GeV})^2$ with $v = 246$ GeV. We also have that the top Yukawa is $y_t = m_t/v$ with $m_t \simeq 173$ GeV.

For $h \gg v$, when the log becomes important, we can approximate the potential by keeping only the terms in h^4 and $h^4 \log h$. Then the potential turns negative for $\frac{\lambda}{4} - \frac{3}{8\,\pi^2}\,y_t^4$
, $\log\frac{h}{v} = 0$ that is

$$h_c = v\ \exp\left\{\frac{2\,\pi^2}{3} \frac{\lambda}{y_t^4}\right\} = v\ \exp\left\{\frac{\pi^2}{3} \frac{m_H^2\,v^2}{m_t^4}\right\} \tag{45}$$

4. By requiring stability, that is $h_c > \Lambda_{95}$ we obtain

$$m_H > \frac{\sqrt{3}}{\pi} \frac{m_t^2}{v} \sqrt{\log\frac{\Lambda_{95}}{v}} \simeq 350\,\text{GeV} \,. \tag{46}$$

5. The Higgs being much lighter that ~ 250 GeV implies that: either the model of inflation is not given by the $\frac{m^2}{2}\phi^2$ potential, or that the potential of the Higgs is not simply given by eq. (39). It is easy to modify the latter hypothesis as discussed in the main text, by assuming for instance a sizable direct coupling $|h|^2\,\phi^2$ that gives a large mass to the Higgs during inflation. In this case the analysis of fluctuations of a field during inflation, that hold for a massless field, do not apply any more and the LHC can agree with chaotic inflation.