# Lesson 9 : Inflation

*Notes from Prof. Susskind video lectures publicly available on YouTube*

## Introduction

Before we use it to explain what probably happened during the very first moments of the universe, let's review what we already saw in the last lesson concerning the inflaton field.

In the universe there are several fields : electric fields, magnetic fields, gravitational fields, etc. We have postulated the existence of another field, yet undiscovered directly, and which is a scalar field.

Fields have energy. Electric fields have energy, magnetic fields have energy, scalar fields have energy. And the energy depends on the value of the field obviously.

The inflaton is an undiscovered field. We don't know how the energy depends on the field. But let's hypothesize it, and plot it. Figure 1 shows first of all the coordinate axes.


Figure 1 : Coordinate axes to plot the energy density of the inflaton field.

Vertically will be plotted the energy of the field $V(\phi)$, per unit volume, and horizontally the field $\phi$ itself. Just think about the energy density as what would be stored in space just by virtue of the fact that the field $\phi$ has a certain value.

If the field is one value, then the energy is one thing. If the field has a different value, the energy is a different thing, and so forth. In that respect it is very much like the electric energy or magnetic energy that is stored in an electromagnetic field. But there is a difference : it is a scalar field. And that makes a big difference for some things.

So let's plot $V(\phi)$ as a function of $\phi$. Since nobody has ever experimentally detected it, we are just going to make up a graph, figure 2.

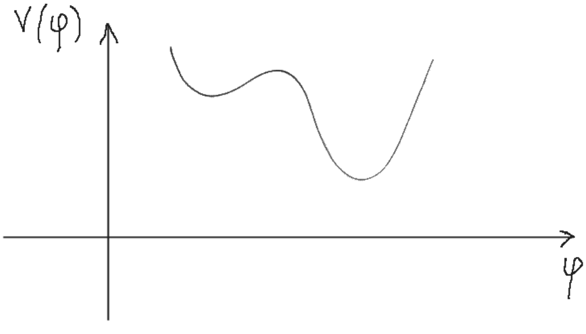


Figure 2 : Energy density $V(\phi)$ as a function of $\phi$.

Now you might think that this is funny. If the field is zero, shouldn't the energy be zero ? But where a field takes the value zero is somewhat arbitrary.

Remember that when we use the analogy with a gravitational potential energy $V(\phi)$, the field $\phi$ itself is simply the height above the Earth, and the height zero is wherever we like. Alternatively – and here it is even more appropriate – we could use the analogy of a compressed spring [1] where the potential energy is maximum when its length is minimum, and that length would naturally be labeled zero.

We call the energy density $V(\phi)$ the *potential* energy density because it is the energy that is not associated with the time dependence of the field. It is not kinetic energy that is associated with a time derivative. And remember that, since the field $\phi$ is uniform in space, see figure 12 of chapter 8, there are no relevant space derivatives. It is the field itself that can be viewed as a spatial coordinate.

The time dependence of the field defines a certain kind of "velocity". It is not the velocity of something in space. It is the velocity of the field itself, in other words the time derivative of the field, which we denote $\dot{\phi}$.

If at time $t$ the field has a certain value $\phi(t)$, at time $t + dt$ it has a slightly different value $\phi(t + dt)$, and by definition

$$\dot{\phi} = \lim_{dt \to 0} \frac{\phi(t + dt) - \phi(t)}{dt} \tag{1}$$

to use the fully-fledged notation of mathematicians.

1. It creates a repulsive force which is a better analogy with the inflaton than the attractive force of gravitation.

That is the analog of the time derivative of the position of something. We must think of the value of the field as a coordinate. It is not the coordinate of a particle, it is the value of the field itself, take one more look at figure 12 of chapter 8. And it is changing with time.

For instance, think of the compressed spring. It is a kind of field taking its values in a one-dimensional spatial world. And it could be vibrating. If we call $\phi$ the length of the spring, there is a $\dot{\phi}$.

Then, by virtue of the fact that the spring is moving, there is kinetic energy. And it is proportional to $\dot{\phi}^2$. In fact we usually define it as

$$\frac{\dot{\phi}^2}{2} \tag{2}$$

to follow the customary definition of the ordinary kinetic energy of a mass in motion [2]. Equation (2) is the *analog* of the kinetic energy of the field.

Now if the field were the position of a particle we would say that we recognize this : it is like $1/2\ mv^2$. Where is the $m$ ? The answer is : even if we put a factor, it would not be a genuine mass. Let's just call it a parameter, say $k$. But then we could redefine our field $\phi$ by working with $\sqrt{k}\ \phi$. And we would be back with no parameter. So let's just not bother with anything in front of $\dot{\phi}/2$.

2. Recall that the factor 2 in the denominator of $1/2\ mv^2$ is totally arbitrary. It simply relates the units of mass to the units of energy, distance and time.

We have said there is also a potential energy $V(\phi)$. So the total field energy at any time is

$$E = \frac{\dot{\phi}^2}{2} + V(\phi) \tag{3}$$

As usual, even if we don't always specify it, this is the energy *density*. In other words it is the energy contained in a small unit box in space. And we remember that the field is uniform in space.

So let's follow a small box the field energy contained in which is given by equation (3). The mathematics of it is isomorphic to – meaning, it is the same as – the mathematics of a particle with a kinetic energy and a potential energy respectively $1/2\ \dot{\phi}^2$ and $V(\phi)$. Or it is very close to it as we will see. And if $\phi$ was the coordinate of a particle, we would know what equation to write down.

Let's go through the steps. The steps are to say $E$ is the energy, kinetic plus potential. Corresponding to it there is a lagrangian $\mathcal{L}$. It is kinetic *minus* potential energy.

$$\mathcal{L} = \frac{\dot{\phi}^2}{2} - V(\phi) \tag{4}$$

And then there is Lagrange's equations. The field $\phi$, or the coordinate of the particle, as a function of time, satisfies the functional equation

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \frac{\partial \mathcal{L}}{\partial \phi} \tag{5}$$

where the unknown is the function $\phi(t)$.

Now, to find what is $\phi$, we work out Lagrange's equation – here there is only one. On the left-hand side, we first compute the derivative of the lagrangian with respect to $\dot{\phi}$. That is just $\dot{\phi}$. Then taking the time derivative yields $\ddot{\phi}$. On the left-hand side we simply have $-\partial V/\partial\phi$. And since $V$ depends on only one independent variable, namely $\phi$, we can write it as a plain derivative. Thus Lagrange's equation becomes

$$\ddot{\phi} = -\frac{dV}{d\phi} \tag{6}$$

If $\phi$ was the position of a particle, that would be the usual $F = ma$, where "$F$" is $-dV/d\phi$, "$m$" is one, and "$a$" is $\ddot{\phi}$.

So we see that $\ddot{\phi}$ is a kind of acceleration : it is the second derivative of $\phi$ with respect to time.

We shall also keep the notation $F$, although it is not a force, for minus the gradient of $V$ with respect to $\phi$.

$$F = -\frac{dV}{d\phi} \tag{7}$$

It could be constant. For instance a falling object in the gravitational field near the surface of the Earth experiences a constant force pulling it down, because the potential energy is decreasing with height with a constant gradient [3]. But the $F$, defined by equation (7), doesn't always have to be

3. Once more, make sure you see that when we use the analogy of a particle falling on Earth, the field $\phi$ we are talking about is the height of the particle. And the usual gravitational field is what we denote $V(\phi)$, also called the potential gravitational energy.

constant.

So, the equation of evolution of the field $\phi$ is

$$\ddot{\phi} = -\frac{dV}{d\phi} = F(\phi) \tag{8}$$

We think of the field as being in a box[4], and we assume that it is not varying very much in space. The reason we can get away with that in cosmology is because space has expanded a lot, or we are assuming space has expanded a lot.

The variations of the field have gotten stretched out, and eventually they got so stretched out that over a small region of space, over the box, the field is very smooth and very flat. That is the assumption.

But equation (8) is the field equation for a scalar field with a field potential energy $V(\phi)$. And it looks like Newton's equations.

The quantity $E$ defined by equation (3) is the energy density. If we actually want the energy in the box, we have to multiply it by the volume of the box. We consider an expanding box following the expansion of the universe. So its volume is proportional to $a^3$.

The scale factor would not matter if it didn't depend on time. It would just then be a numerical constant. And

4. In other words, we do add spatial independent variables which the field depends on. Hence we could write $\phi(t,\ x,\ y,\ z)$. But we immediately add that the field doesn't vary over space. So the variables $x$, $y$ and $z$ don't play any significant role.

the numerical constant, when working with the lagrangian, would cancel out on both sides. So it would not matter.

But $a$ is time dependent. It is $a(t)$. Therefore the energy in the expanding box, for which we keep the same notation $E$, now is

$$E = a(t)^3 \left[ \frac{\dot{\phi}^2}{2} + V(\phi) \right] \tag{9}$$

And the same applies to the lagrangian. Its new form is

$$\mathcal{L} = a(t)^3 \left[ \frac{\dot{\phi}^2}{2} - V(\phi) \right] \tag{10}$$

Then let's go back and redo Lagrange's equation, taking into account $a(t)^3$. Lagrange's equation is now

$$\frac{d}{dt}\, a(t)^3\, \dot{\phi} = -a(t)^3\, \frac{\partial V}{\partial \phi} \tag{11}$$

It is clear that if $a$ was constant, we could pull it out of the time derivative on the left, factor it out, and get rid of it. But now it participates to the quantity we want to differentiate with respect to time. That is the new thing. That is what the expansion of the universe does to the equation of motion for a scalar field.

Since we want to use the new lagrangian given by equation (10), $\dot{a}$ will appear in the equation of motion for $\phi$. Lagrange's equation (11) yields

$$a^3 \ddot{\phi} + 3a^2 \dot{a}\, \dot{\phi} = -a^3\, \frac{\partial V}{\partial \phi} \tag{12}$$

The next step, as we know – we are still in the review of last lesson –, is to divide by $a^3$. We get

$$\ddot{\phi} \; + \; 3 \; \frac{\dot{a}}{a} \; \dot{\phi} \; = \; -\frac{\partial V}{\partial \phi} \tag{13}$$

And $\dot{a}$ over $a$ is what ? The Hubble "constant" $H$, which is constant over space, but can vary over time. So equation (13) rewrites

$$\ddot{\phi} \; + \; 3 \; H \; \dot{\phi} \; = \; F(\phi) \tag{14}$$

For convenience we revert to the notation $F$ for minus the gradient on the right-hand side. But keep in mind what it is. It is the derivative of the potential energy density with respect to the field itself.

So if we go back to the representation of $V(\phi)$ and add a little ball representing the field $\phi$, figure 3, it would be as if there was a force pushing or pulling the field down the hill.

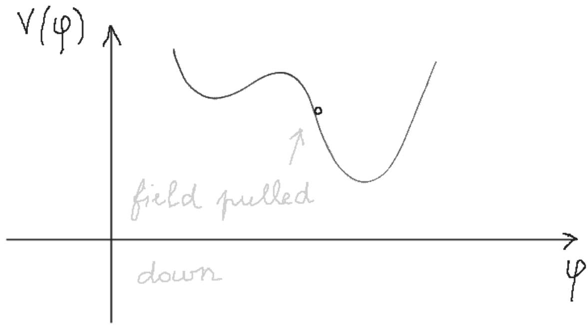


Figure 3 : Potential energy density, and its gradient pulling the field down the hill.

Question : Is there a physical interpretation of the force $F$ and a way to visualize it ?

Answer : It is just the derivative of the potential. Since what we are doing is a general form of mechanics – the mechanics of a field now –, mathematically it has all the properties of a force. But it is not a physical force in space pushing or pulling some object. It is an *effect*, if you like, which by analogy we call a force, pulling the field toward lower potential energy. It is a tendency for the field to accelerate toward lower potential energy.

So we have our equation of motion. And the interesting thing of course is the new term $3H\dot{\phi}$. And this new term has exactly the properties of a viscosity term.

Equation (14) is of course the equation of motion of a field, not a particle. Yet think of the field as the position of a rock falling through a viscous fluid, just as an analogy, figure 4

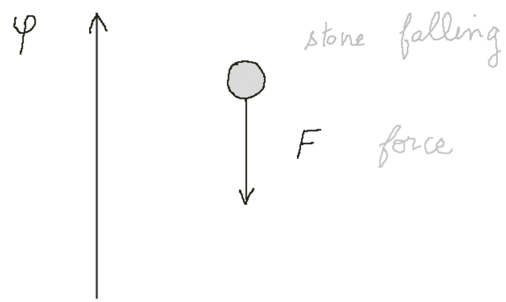


Figure 4 : Rock falling through a viscous fluid.

The height of the rock is called $\phi$. It is the height above the surface of the Earth or whatever. And the rock is falling, let's say under the influence of a force $F$, which in this case is pushing downward exactly like the force due to gravity.

That means the potential energy is increasing upward. And therefore the force, which is minus its gradient, points downward.

If that is all there was, we would expect the equation to be acceleration = force. Indeed we set the mass equal to 1 by rescaling of $\phi$ if necessary. So the object in figure 4 would accelerate. In a few seconds it would be moving a hundred meters a second. Well, it takes some time to get to 100 $m/s$ – the reader can calculate it – but not much time, and it would just whizz away.

But what if it is falling through water, or, even better, something more viscous than water : honey ? Moving through honey, it would gradually sink, but it won't continue to accelerate. Now we know what $H$ stands for in equation (14). It stands for honey.

The rock would experience an additional force opposite to the direction of motion and proportional to the velocity. Viscous forces depend on the velocity. There is no viscous force on a thing at rest. It is only when it is moving that it experiences a viscous force. It experiences the viscous force opposite to the direction of motion.

To see this in equation (14), just shift $3H\dot{\phi}$ to the right-hand side.

$$\ddot{\phi} \;=\; -\,3\,H\,\dot{\phi} \;+\; F(\phi) \tag{15}$$

We now see that the acceleration is the result of two forces : $F$ and $-3H\dot{\phi}$. So if $H$ is positive the drag is opposite to $\dot{\phi}$. Imagine space filled with gooey substance, and the viscous coefficient proportional to $3H$ in this case. Then equation (15) is the equation of motion of the rock.

Now, in this particular case, if $H$ does depend on time, it would be as though the viscosity depended on time. But we are going to be interested in situations where the viscosity or where $H$ does not depend very much on time. Does the viscosity of honey depend on time ? Sure, since it depends on temperature, if it is in the summer it is less viscous, if it is in the winter it is more viscous [5]. So, yes, we can think of viscosity as depending on time. It makes perfect sense.

But for the moment let's suppose the viscosity did not depend on time. And let's watch what happens. In the beginning, meaning if we release the stone from rest, there is no viscous force because $\dot{\phi}$ is zero. So initially the equation is only $\ddot{\phi} \;=\; F(\phi)$. It accelerates.

The stone picks up some velocity according to Newton's law. But as the velocity increases, the term $3H\dot{\phi}$ becomes more important. Eventually we will get to a point where the viscous factor $3H\dot{\phi}$ matches $F(\phi)$, and the right-hand side of equation (15) becomes 0.

5. Notice too that it depends on the amount of energy absorbed by the honey, which is transformed into heat, and therefore warms up the substance. In other words, there is room to build more elaborate models of viscosity than just the equation $\ddot{\phi} = -\gamma\dot{\phi} - \frac{dV}{d\phi}$.

When that happens the stone stops accelerating. That doesn't mean the velocity stops. It doesn't mean the object comes to rest. It means it just falls at a constant velocity called the *terminal velocity*. It is the final velocity that the stone reaches, where the force pulling downward and the viscous force opposite to the motion balance.

We can calculate the terminal velocity. Let's not throw another symbol in, but just write

$$\dot{\phi} = \frac{F}{3H} \tag{16}$$

That is the terminal velocity reached by the stone when viscosity balances force.

The bigger the viscosity, the slower the terminal velocity. If the viscosity is very strong, the terminal velocity would be very slow. The stone would just move very slowly through the viscous fluid.

In that case, if the stone is moving really slowly, it is likely that $\dot{\phi}$ would be very small compared to $V(\phi)$. Indeed if a rock or a stone is falling through very viscous stuff, then the kinetic energy is generally negligible compared with the potential energy. Let's keep that in mind. It is going to come back.

Equations (15) and (16), however, are not the equations for the motion of a stone, but the equations for the evolution of a scalar field $\phi$ in a cosmological context.

Question : By using the lagrangian, are we invoking the principle of least action ?

Answer : Yes we are. Go back to the discussion of field theory in volume 3 of the collection *The Theoretical Minimum* on special relativity and classical field theory. We described the use of the lagrangian and Lagrange's equations to find the evolution of a field $\phi(t, x)$, where $x$ can be multidimensional, figure 5.

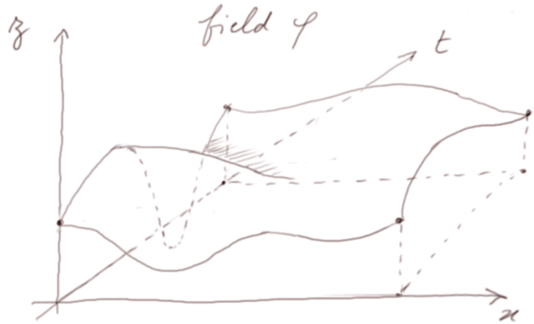


Figure 5 : General scalar field depending on time and space.

But if you don't want to think about lagrangians, just take it as a fact that the time-dependence in the energy

$$E = a(t)^3 \left[ \frac{\dot{\phi}^2}{2} + V(\phi) \right] \tag{17}$$

introduces a term akin to viscosity in the equation of motion of the field :

$$\ddot{\phi} + 3 H \dot{\phi} = F(\phi) \tag{18}$$

Now we might wonder what happened to energy conservation. Consider an object moving through a viscous fluid. Let's suppose for a moment that there was no force $F$. We just give it a push through honey. What happens to the object ? It slows down and comes to rest. And what happens to energy conservation ?

In the case of real genuine honey, we know what happens : the stone heats the honey a little bit.

But there isn't any gooey stuff involved in $\ddot{\phi}+3H\dot{\phi} = F(\phi)$. There is just the equation of motion of a scalar field. There are no molecules that get heated up or anything. And so we should ask : isn't the term $3H\dot{\phi}$ going to suck energy out of the system ? The answer is : yes it is.

So the question becomes : why isn't energy conserved for a mathematical system with a lagrangian that looks like this ?

$$\mathcal{L} = a(t)^3 \left[ \frac{\dot{\phi}^2}{2} - V(\phi) \right] \tag{19}$$

It is because

*Energy conservation is a consequence of time translation invariance of the equations. Furthermore when time translation invariance is not satisfied, energy in not conserved.*

The lagrangian of equation (19) being time dependent, it is no longer time translation invariant. And there is no energy conservation for a lagrangian like this. There just isn't any.

Hence, not surprisingly, that shows up as some sort of term in equation (18) which mimics a mechanism which allows energy to be lost by the system.

This is only one half of the story of the equations for an expanding universe in the presence of a scalar inflaton field. The expansion affects the equation of motion of $\phi$ through the viscous term. But now we can ask how $\phi$ affects the expansion. In fact what does it do to the expansion ?

For that purpose, let's go back to the cosmological equation, not the field equation for $\phi(t)$, but the equation for $a(t)$. That is the *Friedmann equation.*

We are going to simplify it and just take the flat space case. You can put back the curvature if you like. It would make no difference to what we are doing, because it is small by comparison with the things we are going to keep.

Friedmann equation is

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho \tag{20}$$

The left-hand side is also what we call $H^2$. This equation can be viewed as an equation for $H$ if we like. It is an equation for several things. On the right-hand side, we have the constant factor which as usual we write $8\pi G/3$, where $G$ is Newton's constant, times the energy density which we called $\rho$ in the past [6].

6. Go to chapter 1 to review how we derived Friedmann equation, which is equation (30) of that chapter, and why the factor in front of $\rho$ has this form.

But we now have an expression for $\rho$. It is the quantity in the square brackets of equation (17), disregarding the factor $a(t)^3$ in front.

$$\rho = \frac{\dot{\phi}^2}{2} + V(\phi) \tag{21}$$

So we can rewrite equation (20) as

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3} \left[ \frac{\dot{\phi}^2}{2} + V(\phi) \right] \tag{22}$$

We are now going to make an assumption, which we will come back to in a little while to motivate it. We assume that the potential energy $V(\phi)$ is fairly large in some units we will see, but it has a shallow slope, figure 6.

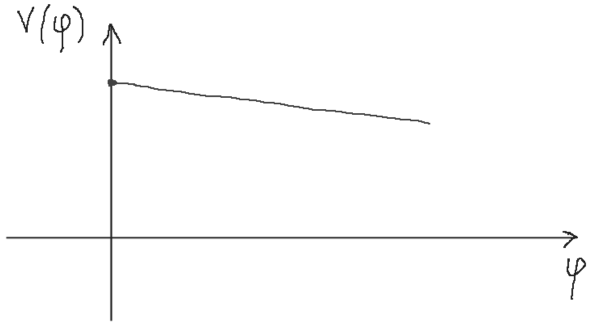


Figure 6 : Potential energy density $V(\phi)$.

It is tilted down to the right. If was increasing to the right, we would just redefine $\phi$ by changing its sign, so we would be back to this picture. Thus it is tilted down to the right, and it does so with a very moderate slope.

Remember, this hypothetical scalar field is called the *inflaton*. Has it been discovered yet ? Only indirectly [7]. Does it exist ? Probably.

**Questions / answers session**

Q. : Is the inflaton a scalar field which somehow only has to do with energy density ?

A. : All scalar fields have to do with energy density. In fact, every field has to do with energy density, period.

---

7. Notice that this raises the question of what exists and what doesn't ? When do we see something directly and do we see it only indirectly ? The difference is not as sharp as it may seem. In fact it can be argued that there is no difference at all. Our perception of the world, and of things existing in it, is only via models in our minds buit from our senses and reflexion.

Only when lots of traces have accumulated do we say that something exists. For instance lots of wolf prints in the snow, plus lots of sheep attacked, plus some vague moving shadows in the distance, etc. are proof of the wolf, even if we have never seen one in front of us. Yet we can then surmise that someday we will be able to shoot one and see it from close. Same with the inflaton : there are many indirect traces of it.

It is a discussion which deserves more than a footnote, but which doesn't have its place in this book. Just notice that man is a creature fond of symbols. Perhaps what we call reality is only symbols turned into practical and efficient representations.

Every field has and contributes to energy density. But the inflaton has been particularly cooked up to do a certtain job.

Q. : So it is a contribution to boson fields ?

A. : It *is* a boson field. Scalar fields are always boson fields.

Every boson could be like this. The inflaton is tailored to do a certain job – a job that other scalar fields would not do because they just don't have the right properties.

For intance, one could think of the Higgs field [8]. It is a boson field. But it doesn't have the right properties. In particular its potential is not shallow as in figure 6. And the energy density associated with exciting the Higgs field is not very high. We want much more energy density there.

But let's leave these considerations. So far, this is just a hypothetical description of the universe expanding with a field $\phi$ that has a certain property.

So we suppose that the inflaton field $\phi$ has a very gradual

8. named after Peter Higgs (born in 1929), British theoretical physicist. In 1964, Higgs, together with Belgian theoretical physicits François Englert (born in 1932) and Robert Brout (1928 - 2011), as well as other scientists, hypothesized the existence of an as yet unknown boson field and corresponding boson particle which would explained the mechanism of mass. The Higgs boson, sometimes called the BEH particle, was confirmed experimentally at the CERN in summer 2012.

tilt to it as shown in figure 6. Think of $\phi$ as a stone falling. It is falling to the right. Things always fall in the direction of decreasing potential energy.

When, in figure 4, we drew $\phi$ on the vertical axis, we wanted to think of the stone falling through space vertically. As said, it always goes to lower potential energy. When there is no other energy expanding it, potential energy always tends to decrease if it has the freedom to do so.

In the case when $\phi$ is drawn on the horizontal axis, and $V(\phi)$ on the vertical one, the point moving on the curve to the right still represents the stone falling vertically, figure 7.

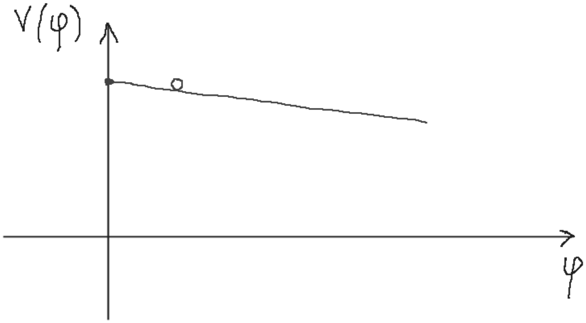


Figure 7 : Stone or $\phi$ falling vertically.

So toward the right are decreasing heights or decreasing values of $\phi$.

Now, if the tilt is gradual, it means that in equation (18), reproduced below, the force $F(\phi)$ is small.

$$\ddot{\phi} \ + \ 3 \ H \ \dot{\phi} \ = \ F(\phi) \tag{23}$$

In a moment we will come back to the viscosity-like parameter $H$ (or $3H$) when in fact we get an equation for it.

So far, from Friedmann equation, we know that

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3} \left[ \frac{\dot{\phi}^2}{2} + V(\phi) \right] \tag{24}$$

Whatever $H$ is, we already know that it is not zero, simply because the right-hand side is not zero. In particular there is $V(\phi)$. We will come back to it.

So $H$ is not necessarily small. And, in the context that we are going to be thinking of, it is rather large : big viscosity, big viscous coefficient, very gradual tilt. So it is like thick honey with a weak force pulling us down.

What happens then ? We very quickly get to the terminal velocity. According to equation (16), which says that $\dot{\phi}_{term} = F/3H$, the force being small, and the viscosity large, the terminal velocity is small.

The ball represented in figure 7, which is really the value of the field, very slowly, gradually creeps down the hill. Consequently the term $\dot{\phi}/2$ on the right-hand side of equation (24) can be neglected.

It can be checked with equation (16) that, whatever we are choosing for the parameters, $\dot{\phi}/2$ is negligible compared to the potential energy density and can be omitted. Let's suppose we have done that. Then to high approximation, as long as it is moving very slowly, we can write

$$H^2 \;=\; \frac{8\pi G}{3}\, V(\phi) \tag{25}$$

Moreover, $V(\phi)$ is not varying very quickly. The field $\phi$ is varying very slowly. And $V(\phi)$, the height of the potential, is varying with time very slowly too.

In other words, to a pretty good approximation, the field is standing still, and the value of the potential energy is just whatever it is on the ramp in figure 7. It is approximately just a fixed number.

So let's keep that equation in mind

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\, V(\phi) \tag{26}$$

How does the universe expand under these circumstances ? Let's assume that, because $V(\phi)$ is so slowly varying, as a first shot at it we can just say that $H$ is constant over a range of time until the rock in figure 7 slides to an appreciably different height. So as long as the rock on the ramp is maintaining its height, or its potential energy, we can just say $V$ is a constant. And in that case of course $H$ is a constant because it is linked to $V$ by equation (26).

What is the solution of the equations of cosmology with a constant $H$ ? Exponential expansion. Why ? Equation (26) can be rewritten

$$\frac{da}{dt} = Ha \tag{27}$$

The solution of that equation is an exponential expansion. If the time derivative is proportional to the same object with a constant coefficient, or approximately constant coefficient, then the object $a$ expands like some constant times $e$ to the power $Ht$.

$$a(t) = ke^{Ht} \tag{28}$$

And, from the second equality in equation (26), we also know what $H$ is.

$$H = \sqrt{\frac{8\pi G}{3}\ V(\phi)} \tag{29}$$

Of course we might want to fit this to experiment. If, at the time at which this was taking place, we could monitor how fast the universe was expanding, we would simply fit $V$ to observation. But the things we are interested in here took place in the very early universe.

Let's be clear about it : this is a theory of the very early universe, a theory of the inflationary early starting point of the universe. At that time $V$ was believed to be rather large in any kind of sensible units.

If this is a correct theory, the universe was exponentially expanding with a large coefficient. How large ? The guess that would be natural in cosmology would be the doubling of the size of the universe every $10^{-32}$ seconds.

That is really going ! The universe was really accelerating, really expanding at that time.

Pay attention to the fact that this is not the theory of the *current accelerating universe*. It is the theory of the early universe.

In a while we will describe what happened later. But this is the exponential expansion during *inflation*, which is the name given to the phenomenon that took place in the very early moments of the universe.

Question : Large viscosity and low gradient on the one hand, and rapidly expanding universe on the other hand, seem to be opposite.

Answer : It is $\phi$ that is moving slowly ; not $a$.

The field $\phi$ moving slowly is related to the potential energy $V(\phi)$ staying more or less constant.

But it is the value of the potential energy which is driving the expansion of the universe through the constant $H$.

In other words, it is the *derivative* of the potential which is making $\phi$ move, equation (23). That derivative is small.

But it is the *value* of the potential energy itself which is driving the acceleration of $a$, equation (26). That potential energy is large.

These are quite different things.

It is not the scale factor $a$ which is experiencing viscosity ;

it is the inflaton field $\phi$. So there is no viscous drag on the expansion of the universe. It is not what is going on here. The viscous drag is on the evolution of the field, and it is keeping it pretty constant.

Why should we be interested in a theory in which the universe was expanding very rapidly for possibly some significant length of time? Time now is measured in units of $10^{-32}$ seconds?

The original guess about this, which went back to Alan Guth[9] in 1980, did not seem terribly compelling. Yet cosmologists and theoretical physicists caught on to it very quickly. By now it has been confirmed. But it is one of those stories of guessing in theoretical physics. And at first it was an outlandish guess.

There were two puzzles, related to each other, that this theory was addressing. The first puzzle had do with particles called *magnetic monopoles.*

Before we go into that, let's stress that this inflation theory, i.e. this early very fast exponential expansion, is by now well confirmed. We now know with some degree of confidence that it really happened that way. And it lasted long enough to cause the universe to expand by a factor of at least $e^{60}$. We will come to how we know it.

So why should one be interested in the universe stretching

9. Alan Guth (born in 1947), American theoretical physicist and cosmologist.

so much ? The first reason, as said, is monopoles. What are monopoles ? They are particles that once again have not been discovered.

The reader may wonder why should we introduce one crazy theory with a totally undiscovered inflaton field to worry about particles which are also undiscovered ? The reason is that, although physicists haven't yet been able to discover them – i.e. to "see" them in manner that would be satisfactory –, these hypothetical magnetic monopoles play a very important role in physics :

*Magnetic monopole plays the same role with respect to magnetic field as electric charge plays with respect to electric field.*

A magnetic monopole is simply a source, a point source or an approximately point source of magnetic field. The magnetic field should be radiating out from it with a Coulomb pattern that would be the exact analog of an electric charge, figure 8.

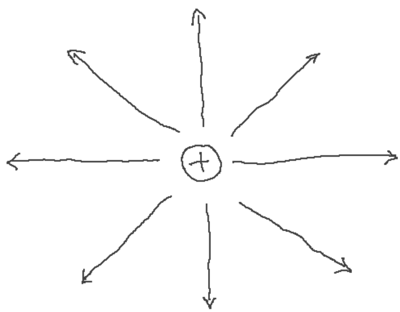

Figure 8 : Coulomb field created by a single electric charge.

An electric charge $Q$, in the center of the picture, creates a field $E$ in space. At a distance $r$ from the charge, the electric field has the value

$$E(r) = \frac{Q}{r^2} \tag{30}$$

And the corresponding potential field follows a law in $1/r$.

We would expect to meet in nature, or in the laboratory, magnetic monopoles with the analog behavior for magnetism. However so far we have only met magnetic dipoles whose field has the following shape [10].

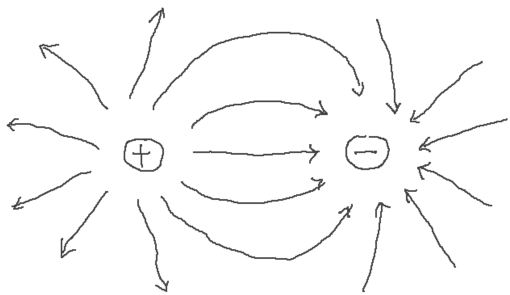

Figure 9 : Magnetic field created by a magnetic dipole.

The sought for magnetic monopoles, according to current theories, have never been discovered in the laboratory because they are very heavy. How heavy ? Somewhere up near

10. When we split a small magnetic bar, which has a plus pole at one end and a minus pole at the other end, into two shorter magnetic bars, we obtain two magnetic bars each with a plus pole at one end and a minus pole at the other end.

the Planck mass.

That is to put it all in a nutshell : magnetic monopoles are believed to exist ; they exist in most unified theories that we would think of ; and in most theories they are very heavy by comparison with ordinary particles.

The fact that they are very heavy – perhaps 19 orders of magnitude heavier than the proton – would easily account for the fact that they have never been seen in the laboratory.

But now let's go back to the very early universe. Remember that the early universe was hot. And when it was hot enough, there was enough kinetic energy in the form of photons and other particles that you didn't need an accelerator to make elementary particles.

At one point the temperature was high enough that photons would scatter off each other and produce electron-positron pairs.

Even earlier it was hot enough that photons would scatter off each other and produce protons and antiprotons.

If we go way way back to what physicists think of as the very beginning, the temperature, according to theory, was hot enough that it could have created magnetic monopoles – pairs of them.

Just like you create pairs of electrons, that is electrons and positrons, when the temperature was 26 orders of magnitude higher than the temperature for electron-positron pairs, it should have made pairs of monopoles.

Now the reader can think that this is wild speculation. Nevertheless every theory that we know about seems to contain these monopoles.

So the question is : where were they ?

First of all, monopoles cannot easily disappear. They have a magnetic charge, so they can't easily disappear. What could they disappear into ? They cannot become electrons and photons and other things, because those don't have a magnetic charge. Magnetic charge would be conserved, just like electric charge.

So, unless these particles anihilated, and one can estimate how likely it is that they all anihilated – not likely –, in the course of the universe as it evolved we would have expected to find some population of magnetic monopoles floating around in space.

We might not expect a large density of them. The universe expanded a lot since the time of decoupling [11]. It expanded by a factor of 1000 since then, figure 10.

Of course, since the time that it was hot enough to create magnetic monopoles according to the standard theory, it would have expanded a lot more than that. So they would be pretty dilute.

11. See figure 6 of chapter 7 to review what decoupling is.

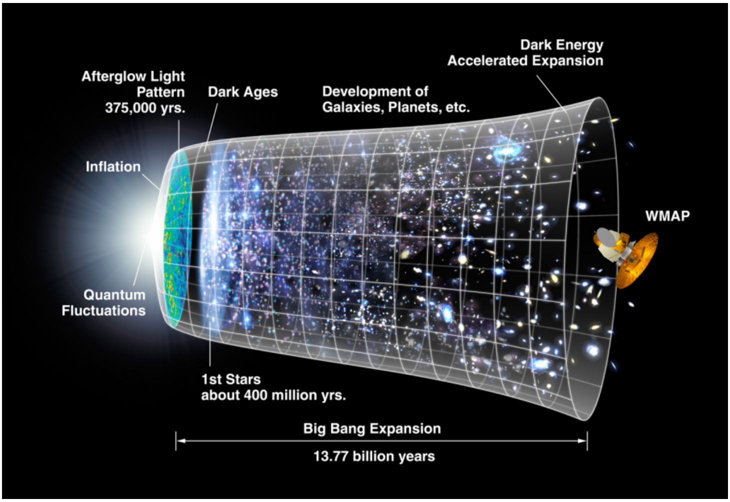


Figure 10 : Chronology of the universe.
Source : NASA/WMAP Team.

But they don't seem to be there. Now why don't they seem to be there ? Because, they would have an interesting effect on the magnetic fields of galaxies.

What happens if there is an electric field around and you have a population of electrically charged particles and antiparticles ? They discharge the field.

Suppose we have a capacitor with an electric field between the positive and the negative plates, figure 11.

And, as shown in the figure, there are also positive charges and negative charge in between the plates. Let's say electrons and positrons.

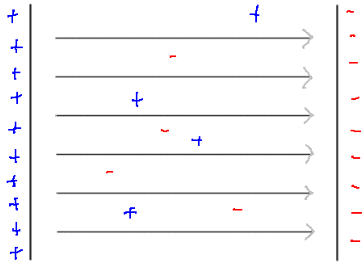

Figure 11 : Capacitor and the electric field between the plates. On the left is the positive plate, on the right the negative one. And in between are also charges floating freely around.

What happens ? The plus side of the capacitor pulls over the negative charges ; the minus side of the capacitor pulls over the positive charges ; and this discharges the field.

In the same way, if the population of monopoles was as large as theorists have estimated, and if they lived forever and sat there in the magnetic fields of galaxies, then, over the age of the universe, they would have discharged the magnetic fields of galaxies.

Yet the magnetic fields of galaxies have not been discharged, they are there, they exist.

So it became pretty clear that there is no significant population of monopoles out there, even though the standard cosmological theories around 1980 predicted that the universe should have been hot enough at the beginning to create monopoles, and then the standard expansion should not have completely diluted them away.

That was one set of facts. The other set of facts concerned the extraordinary homogeneity of the universe, not today's homogeneity but *the homogeneity at the time of decoupling.* How do we know the universe was so homogeneous then ? By looking at the cosmic microwave background in different directions of space.

Since the nineteen sixties, the temperature of the microwave background became better and better measured. And it was done in every directions. The flux of radiation coming in, as well as the temperature of it, and its distribution were measured to very high precision all over the sky.

By the nineteen eighties, it has been found that it was absolutely, or almost absolutely, uniform. To one part in $10^4$, or even better, it was uniform over the entire sky.

That raised another question : why was everything, at the time of decoupling, so darn smooth ? No standard theory of cosmology could explain it.

If you start with a small universe very very hot, it is likely to have all kinds of fluctuations, and bumps and lumps on it and so forth. And standard expansion would not have flattened it out and smoothed it to the degree it was known to be smooth and homogeneous.

The guess that Guth made was that very early, after an initial period, however, during which the monopoles had been produced and perhaps the universe had a highly complicated texture, an enormous amount of expansion happened which diluted the monopoles.

Enough expansion would dilute anything, so the presence of monopoles was just diluted to the point where they were much fewer of them per volume than had been estimated. This inflation also diluted the textural lumps and bumps that one might have expected to be there.

Nobody can prove that they were there prior to the inflation. You could just say : well, the universe just started very smooth. And it stayed very smooth. You could have also said : the universe started with no monopoles, and that is why we don't see any today.

But there was no good reason for the almost perfect homogeneity and the absence of detectable monopoles at the time of decoupling – with the exception of one shot at an explanation. The idea of a very early extremely big expansion can solve both problems. It smoothes everything out, just like when we inflate a balloon its surface becomes very smooth, see figure 8 of chapter 8, and it also dilutes the density of the monopoles.

Estimations were made which vary somewhat. But they all point to the fact that, if you want to make the universe as smooth and flat as it is [12], from an unknown starting point, except that this starting point was not terribly smooth, you must have approximately 60 $e$-foldings. An $e$-folding is a multiplication by $e$, which is approximately 2.7.

So, over the course of this extremely big inflation, we have

12. Or as it was at the time of decoupling, whose minute remaining inhomogeneities lead to today's universe, with its galaxies and stars and so forth.

$$\frac{a(t_2)}{a(t_1)} = e^{H(t_2-t_1)} \approx e^{60} \tag{31}$$

It implies that, in $a = e^{Ht}$, the exponent $Ht$ must be at least 60. That is what is needed to make consistent a small inhomogeneous initial universe and its observed flatness, i.e. absence of curvature, and smoothness, i.e. homogeneity, at the time of decoupling.

That, in itself, tells us that the ramp shown in figure 12 must have been rather shallow. The universe had to stay on the ramp for a sufficiently long time for it to have grown by a factor of $e^{60}$.

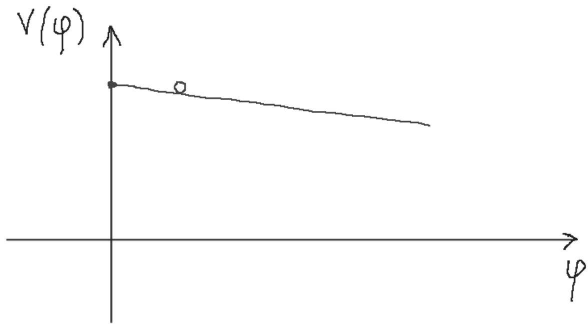


Figure 12 : Shallow ramp on which the inflaton field evolved during inflation.

It doesn't tell us any details. There are all sorts of ways that we can change one thing and change the other ones. But the working number is that there had to be 60 $e$-foldings of inflation in order to simultaneously account for the lack of monopoles and for the lack of structure in the universe at the time the CMB began.

Q. : Is there any good reason why the universe should have been rough and bumpy at the outset, before inflation took place ?

A. : It is better to ask if there is any good reason why it could not have been rough and bumpy.

Anyway, sixty $e$-foldings would be enough for something starting rather chaotic and rough to be stretched out and smoothed into what we observe in the remnant CMB.

The factor 60 is not hard and fast. It seems to be a minimum. If you make your assumptions worse, you need more, say 70 $e$-foldings. In fact we don't have an upper bound. It could be $e^{60000}$ or whatever.

On the other hand, by changing the theory in this way or that way, you can probably bring the minimum number of $e$-foldings necessary down to 40 or 50.

But that is not the point. The point is that we can say with reasonable confidence that a huge inflation took place at the beginning of the universe, of most probably 60 $e$-foldings or more. And that inflation should not be confused with the Hubble expansion we observe today.

Nothing in the observations bars a much bigger factor, like $e^{60000}$ for instance, during this early inflation. It would be very remarkable, though, because to do that we would have to make the ramp in figure 12 *extremely* shallow. We would have to fine tune the shallowness of it and so forth. But still there is no known upper bound.

And as far as a lower bound is concerned, as said, an order of magnitude of 60 $e$-foldings for a minimum increase in the scale factor seems plausible. It is necessary to understand the flatness and smoothness at decoupling.

So that is where things stood around 1980. Guth postulated this inflation. In order to drive the inflation, he postulated a potential which looked something like figure 12. The first round of theory was not quite right. It got into the form described above after a year or two, following the contribution of Linde [13] and others.

We don't want to go back into too much of the history, because some of the physics was wrong. But ultimately it got straightened out to where the theory looked pretty much like in figure 12, with a potential following a shallow ramp which supported the inflationary tendency for at least 60 $e$-foldings.

Now one interesting thing is : if the number of $e$-foldings was just barely 60, that would mean we were on the edge of being able to see exactly what it is that the inflation was invented to get rid of. Inflation was invented to get rid of something. The factor in equation (31) is how much we

13. Andrei Linde (born in 1948), Russian-American theoretical physicist.

needed to flatten the initial rough an bumpy universe to its current flatness.

What if it were right on that edge? If it were right on the edge, it means we should be on the threshold of being able, in the next round of experiments, to see some additional curvature in the universe.

So one of the things that will be hunted for by the Planck satellite, and other rounds of observational cosmology, will be to see if the universe really is flat. It might have *spatial curvature*, a positive one, or a negative one. And the fact is, inflation doesn't tell us whether there was more $e$-foldings than 60. It just tells us there must have been at least 60 in order to account for the present data from observation.

Always, in such circumstances, if you are on the edge, it means that you have a chance in the next round of discovering more. We will come to why this is so important.

Q. : How is this rapid expansion during inflation related to the temperature of the universe?

A. : The first impression, which is close to the truth, is that this rapid expansion cooled the universe.

If it was just classical physics, and you had heat in a box, and you exponentially expanded the box, very quickly, it would cool to negligible temperatures. So the first answer to this is, at least in a classical universe, that the rapid expansion will have cooled it down.

So our first statement, let's say, is that it was cold, very cold, during this inflation, because the heat got diluted [14]. Everything got diluted.

The space was expanding rapidly. So it was empty of any kind of structures – empty in the sense that everything had been diluted. Ordinarily particles were diluted. Monopoles were diluted. Everything else was diluted. It was empty as hell, and cold.

Now this could not go on forever. Why not ? No theoretical reason. If the potential in figure 12 had some other shape, it could go on arbitrarily long. But we know it did not go on arbitrarily long : we are here ! The universe is not doubling every $10^{-32}$ seconds. So something happened.

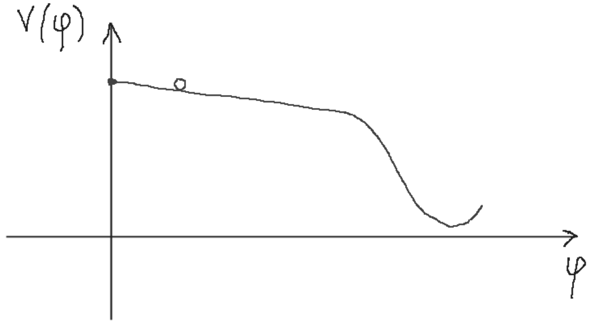


Figure 13 : Potential energy density of the inflaton field.

14. Remember that the wavelength of a photon in an expanding box increases, see the section Radiation-dominated universe of chapter 2. Therefore its energy decreases. And temperature is energy.

The natural guess is that the potential followed some curve like the one shown in figure 13. It had a plateau. And it came to the end of the plateau.

Why it had a plateau, and why the end of the plateau looks like figure 13, is a mystery today. We don't know.

What we know is that it is a pretty good representation of cosmological data.

And what about today ? What about the bottom on the right of the figure ? The bottom of the curve represents today. The potential energy of the field is very small. We will come back to why in a little while.

If this is the model that works, how does it work ? The first question is : how did the universe get to where the ball is shown in figure 13 in the first place ? Nobody knows. That is part of the initial conditions. And, of course, initial conditions in this context mean : that is how far we can trace back with anything like observational data.

So we know from observational data, CMB data and so forth, that there was a period of inflation. We must have started up at some location on the left part of the ramp. How we got there ? We do not know. It is a rather completely arbitrary thing to start us up there. And it requires a theory of some cosmological setting.

Did we really start there, or did we come down from some other place ? We don't know. But tracing back as far as we can go, it looks like a good bet that we started somehow like the ball in figure 13.

Then we rolled very gradually down the ramp [15]. It took long enough that 60 $e$-foldings happened.

Now what happens when we get to the end of the plateau? If it is steep enough, viscosity becomes less important.

Viscosity – which is the way we view the Hubble parameter $H$ in this setting – becomes less important for two reasons.

a) First of all, since $H$ is proportional to the square root of the potential energy, equation (29), as we go down the hill, $V(\phi)$ gets smaller and $H$ gets smaller.

a) Secondly, the force $F$ gets bigger as we go down the hill, equation (23). So again the viscosity becomes less important.

The expected result is that, once we got to the edge of the plateau, we would sort of slide down relatively quickly. Inflation would stop.

Moreover in the process of sliding down the hill in figure 13, energy was released. The potential energy of the field disappeared.

Even though in an expanding universe energy is not strictly conserved [16], there is still some memory. If it is not expanding too fast, you still do get to use, at least approximately, energy conservation.

15. "Very gradually" not in the sense that it took a long time, but in the sense that during the inflation the potential didn't decrease very much. Its gradient remained almost flat.

16. Because time invariance doesn't hold. Therefore we cannot apply Noether's theorem to energy.

Consequently, if the potential energy suddenly disappears, it will be made up for by something. In other words, it is thought that, as we went over the hill, the potential energy $V(\phi)$ got converted to something else.

One thing that it would get converted to is the velocity of the field $\phi$, obviously. But then eventually we get down near the bottom of the curve.

Let's assume there is a bottom, as in figure 13. And we get stuck at the bottom. Where is the energy ?

The current thinking is that the energy somehow got converted to particles and to heat. And this end of the inflation, and creation of particles and heat from $V(\phi)$, is usually taken as the *starting point of the Big Bang.*

When this vast amount of potential energy, which is believed to have been there, got released fairly suddenly and got converted in part into heat, that heat was the starting temperature of the universe when it began as a hot Big Bang.

So the Big Bang has to do with what happened as we slid down the steep slope on the right of the almost flat ramp, figure 14.

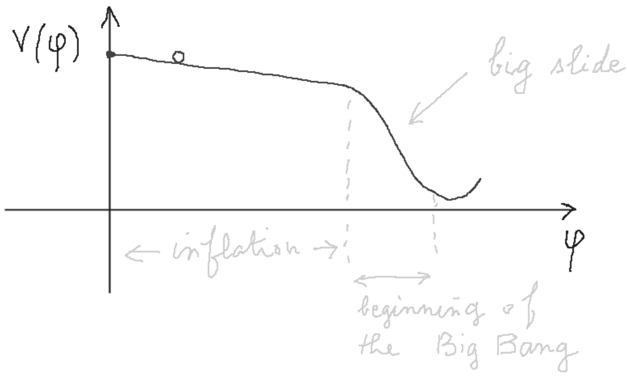


Figure 14 : Inflation and beginning of the Big Bang.

The heat that was generated during what we could call the big slide replaced the right-hand side of equation (26), reproduced below, by perhaps radiation-domination.

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\ V(\phi) \tag{32}$$

And that is when standard cosmology starts, see chapter 2.

**Questions / answers session (2)**

Q. : In the Higgs field that was mentioned earlier, we have a constant field and the expectation of that field is the Higgs boson. So, when we get to the bottom value of $V(\phi)$ on the right of figure 14, are inflatons created ?

A. : Yes, exactly. Well, I mean, when I say exactly, that is one nice theory. That is one nice way things could have happened.

The inflaton field is itself a field which has quanta. Those quanta are called inflaton particles, or simply inflatons.

So as you slid down from the edge of the plateau to the bottom of $V(\phi)$, you could either think of the inflaton field $\phi$ as increasing, or you could think of the phenomenon as the production of lots of inflaton particles.

If these inflaton particles themselves are unstable, they could decay into electrons, positrons, photons and other things. And that would be one explanation of how all that energy got converted to heat and particles.

So, yes, that is one story, which in some form is probably the right story.

Q. : Is the bottom of the trough, in figure 14, where space actually began ?

A. : That is not where space began. But somewhere along the big slide is where the Big Bang, as we normally think of it, the hot Big Bang, began.

– So, to the left of the slide, space did not exist ?

– Why do you say that space did not exist ?

– Because the Big Bang did not expand into a space that existed. It created space, didn't it ?

– No, it didn't. *A huge amount of space was created by the inflation.* That is what the inflation did [17]. It blew up a sort of fictitious balloon.

I am not sure exactly what creating space means. But let's suppose what it means is it makes space bigger.

We don't understand the starting point. But let's take as a starting point some tiny little 3-sphere of space [18] in want of anything else.

What would we mean by saying creating space ? We mean we made a bigger [19]. So that we would be able to move around in it, and have stars and galaxies and so forth.

---

17. In figure 14, the horizontal axis is not time but $\phi$. In fact, as we said, the inflation happened in about 60 times $10^{-32}$ seconds. That is an order of magnitude where we assimilate one $e$-folding and one doubling of size.

18. Remember that the *surface* of the Earth is a big 2-sphere. And the surface of a marble, or of the ball of a ball-pen, is a tiny 2-sphere. As we saw in volume 4, as well as in the present volume 5, chapter 3, this geometry can be extended to three-dimensional spaces. We then speak of 3-spheres. The reader must make sure that he or she doesn't mix up a 3-sphere with the suface plus the interior of an ordinary ball in a 3D Euclidean space.

19. The reader should not imagine the universe as embedded into some larger infinite Euclidean space. If we were in two dimensions, we would think of a 2-sphere, at first small, then larger, but in each case without any other dimension. A 2-sphere is finite, yet has no boundary. If the 2-sphere is very big, like the surface of the Earth, it is locally like a Euclidean plane. The same is true with 3-spheres in three dimensions and the ordinary 3D Euclidean space..

2-spheres and 3-spheres have interesting spatial relations called triangles and things of that nature. When the spheres are very big, they are locally like Euclidean spaces of the same number of dimensions. In that case triangles which are not too big have the sum of their angles equal to 180 degrees. But when the triangles are big this is no longer true.

So creating space simply involves expanding the initial tiny universe by a factor of $e^{60}$, that is approximately $10^{26}$. After 26 orders of magnitude of expansion, much space was created. And it was essentially empty.

But it was full of potential energy. So to say that is was empty is not quite right, because of this potential energy. Now this potential energy was actually, at that time, in the form roughly speaking of *dark energy* – except that it was a much much larger value than today's dark energy.

It was the potential energy of the inflaton field $\phi$ when we are still on the ramp in figure 14, but near the right edge.

That potential energy $V(\phi)$ got eliminated when we rolled down the hill. Something had to replace it. What replaced it, according to theory, was particles, heat, light and all the stuff that the universe began with – "began with" in the sense of standard cosmology.

It rolled down to the bottom. What about the bottom ? Was there any potential energy left over at the bottom ? Yes there was some. And what do we call it today ? Dark energy.

The universe is still believed today to be expanding exponentially but with a tiny Hubble constant. Now it doubles (or is multiplied by $e$) not every $10^{-32}$ seconds, but once every 10 billion years. That requires a very very small potential energy in the trough of figure 14. And that is what we today call the cosmological constant.

Q. : So where are the magnetic monopoles ?

A. : The magetic monopoles where created before the particles that make up our universe. The latter where formed *after the inflation* was over and $V(\phi)$ was transformed into matter and other kinds of energy. The former were created *before the inflation.*

The monopoles were diluted out of existence during the inflation. They are still around, according to current theory, but their density is so weak that it is as if they didn't exist.

Q. : Are we on the edge of seeing monopoles ?

A. : No, I don't think we are on the edge [20] of seeing some monopoles. But we are on the edge of seing some curvature.

I don't think there is any chance of seeing monopoles. We know that the galactic magnetic fields were not unwound. But the precision with which we can say anything about

20. "On the edge", here, is not in the sense that it is finally going to happen, but in the sense that it is almost possible.

monopoles from the fact that the magnetic fields were not unwound is not very good.

What we can do, however, is measure the curvature of space. It has been measured to some precision. So far it is zero with a precision of the order of 1% in some units. In the next round of experiment the tolerance in the precision will go down possibly two orders of magnitude. And we may discover that space is curved.

The tightest constraint imposing a large number of $e$-foldings during inflation is the fact that space looks flat [21]. In other words there was enough inflation that it flattened things out to a degree such that at present we cannot measure any curvature of the space.

Q. : So the potential energy at the bottom of the trough in figure 14 is a potential energy still present today in the universe ?

A. : Yes. It is the *vacuum state energy* of today, also called the *dark energy* of today.

---

21. It is always a bit weird to think that the universe might not be a flat 3D Euclidean space, but might have, in three dimensions, a positive, or a negative, curvature. Yet, observe two things :

1) If it is not flat, it is certainly locally flat – locally meaning within a few billion light years of where we stand.
2) In Antiquity the Earth was thought to be a flat plane, until the Greeks understood that it was more like a sphere, or part of a sphere. And that it was a complete sphere was confirmed by Magellan's circumnavigation (1519 - 1522).

So in some sense figure 14 is a nice unified picture of inflation and dark energy, inflation having taken place while the universe was on the slightly tilted ramp, and dark energy being what remains today, after the big slide, at the bottom of the trough, of the initial $V(\phi)$ which was mostly converted into light, matter and heat.

On the other hand it is a complete mystery why the value of $V(\phi)$ today, down at the bottom of the trough, is 120 orders of magnitude smaller than it was up on the ramp.

That is one of the numerous mysteries of cosmology. In figure 14, we represented the lowest value of $V$ still somewhat above 0, but it should be remembered that it is 120 orders of magnitude smaller than on the ramp. So it should all but touch the horizontal axis.

That, of course, is the mystery of the small cosmological constant.

Q. : You say that during the expansion there were no radiation or anything, just space that expanded, but you also say that the expansion smoothed things out. What was there to smooth out ?

A. : The curvature of space. The 3D geometric cuvature of space was flattened out, just like, in 2D, the curvature of the Earth is less than that of a basket ball.

– But how does this stand with the monopoles that got created later on ?

– No, no. no. No monopoles were created later on. If you think that monopoles could have been created during or after inflation, you are in trouble.

The question is : when the point [ $\phi$, $V(\phi)$ ] fell off in the big slide, after the ramp in figure 14, how hot did it heat the universe ?

Did it heat it to the Planck temperature ? In that case, we are back exactly to where we started : plenty of energy to create monopoles.

Or did it heat it to something more moderate and modest ? That depends on the details.

So yes, you are right, there is a question here.

The big slide in figure 14 is called *reheating*, incidentally. It could have just been called heating. But, whatever its name, it created a certain temperature. It is the temperature at the beginning of the Big Bang.

It is essential that that temperature is low enough to be below the threshold for creating monopoles. Otherwise we would be back where we started. If monopoles were created on the big slide leg of the journey, we are back in the soup.

Q. : Are monopoles condidered particles ?

A. : Yes, they are – very heavy ones.

– And they existed before expansion happened ?

A. : They may have existed before expansion happened. Or they may have failed to have been produced, for the simple reason that it was not hot enough.

Now, you have to juggle numbers to get this to work out. But the numbers aren't crazy.

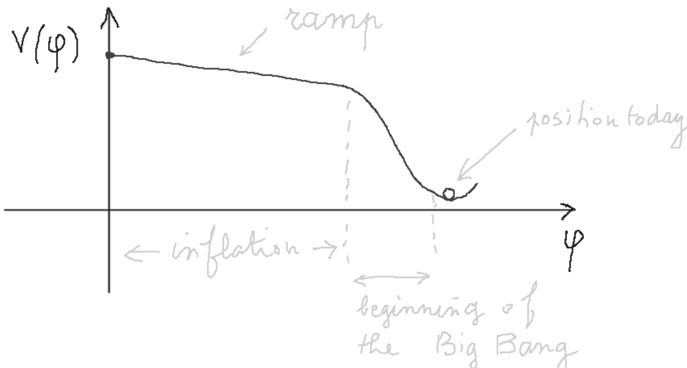


Figure 15 : Position of the universe today.

Q. : Where are we now on the curve ?

A. : We are at the bottom of the trough, figure 15.

– Is there some overshooting expected that would send us upward on the other side ?

– No no. Well, maybe a little bit. The reason that it doesn't overshoot is because of the friction.

Remember that the equations of cosmology have a friction term in them. The expansion of the universe today is big enough to provide a friction that would make it very unlikely for the point $[\,\phi,\ V(\phi)\,]$ to spontaneously go up the hill.

Another way of saying it is that it was the friction term that brought it to rest. It may have overshot. It may have even rolled back and forth a couple of times, like a ball in a bowl with friction.

But the friction term due to the expansion of the universe would have slowed it down. It would have come to rest. And then like any kind of object at rest in an equilibrium position, a rock sitting at the bottom of the sea, it is not going to suddenly spontaneously move up.

Q. : What caused that curve to suddenly drop ?

A. : That is the question. We just don't know !

The curve in figure 15 is a parameterization that fits observation.

We don't know why it was so flat along the ramp. And we don't know why it didn't continue to be flat.

There is no theory of it.

Q. : Even if there is no clear explanation for it, do we believe that there was *a cause* which stopped inflation ? And what could it be ?

A. : Let's put it this way. I think the only known answer to any of these questions, and when I say "known answer" I don't mean to say that it is known to be right, but the only answer around is the *anthropic principle.*

It is the principle according to which the universe may be extremely big, and may have many many different environments of all different kinds. And by happenstance we are in one environment which enables life and creatures like us.

In other words, you can ask : here we are, what kind of environment was necessary for us to exist ? The thing is, a lot of finely tuned parameters are necessary for us to exist.

And of course, to start with, we certainly couldn't exist if the universe was as big as a peanut [22]. So it had to have evolved and expanded.

I don't want to get into a discussion of fine-tuning and the anthropic principle in this lesson [23]. Notice however that the general structure of the curve in figure 15 is dictated not only by observation, but most of it, or large parts of it,

22. We stress once again that this is an image embedded in 3D Euclidean space. *What is meant is a small 3-sphere on its own.* A peanut, whether we consider only its surface or its surface and its interior, is not a 3-sphere. It is just something small. And it is in that sense that we use it as an image.

23. For a discussion of fine-tuning and the anthropic principle, see the final section of chapter 10 of volume 6 in the collection *The Theoretical Minimum* on statistical mechanics.

are necessary for the universe to have evolved to something that looks like what we live in.

We will come back to it. As of 2013, other than the anthropic principle, we have no explanation of why the features of the curve are like this.

Q. : Doesn't everything that happens in the universe have a cause ? That would be the first example of something with no cause ?

A. : We don't know. Some things are unknown.

Concerning things with no cause, in quantum mechanics when we are interested in one observable, if the state of the system, before the observation, is in a linear superposition of eigenvectors of the observable, after the observation we will measure one eigenvalue and the system will be in the one corresponding eigenvector.

As the reader remembers from volume 2, those eigenvalue and corresponding eigenvector are random. The same exact setting and experiment can produce another eigenvalue and its corresponding eigenvector. It is the so-called wave function collapse. We studied it and the probabilities of the various eigenvalues, and won't go over it here. But in that sense the measurement obtained has no cause.

Q. : The horizontal axis in figure 15 is the value of the inflaton field $\phi(t)$. Is $\phi$ increasing with time, or could it become smaller ?

A. : Time isn't on the graph. But we can mark it off, figure 16. The value of the field is indeed increasing with time. This is the model. Since $\phi$, so far, is a made-up field, if the model lead to a decreasing field, we would just look at its opposite, and it would be increasing with time.

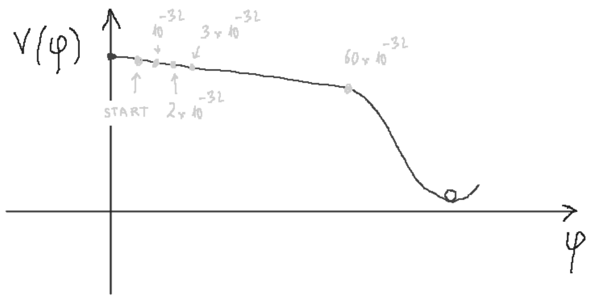


Figure 16 : Marks of time.

As time passed, the field $\phi(t)$ progressed on the ramp to the right. Every $10^{-32}$ seconds the universe doubled in size (or more precisely was multiplied by $e$). At the edge of the ramp, before the big slide, the time is at least 60 x $10^{-32}$ seconds, corresponding to at least 60 $e$-foldings. It could be more.

Then, relatively quickly it went down the hill. How long did it take to go from the edge of the ramp to the bottom of the trough – which is where we are today ? Probably more than $10^{-32}$ seconds, even a lot more, but, on the scale of ordinary time, still pretty quickly.

Q. : Do we have an idea how the temperature varied while the point [ $\phi$, $V(\phi)$ ] ran its course on the curve ?

A. : Yes. The temperature was very cold during the inflation period, when the point was on the ramp.

Then, after it toppled over the edge, during the big slide, temperature went up, because the potential energy $V(\phi)$ of the inflaton field got transformed into light, matter and heat. We don't know with any detail to what temperature the universe was heated up by the time it got down to the bottom.

But what we do know is it had to have been hot enough to create all the particles that we know and love, quarks and everything else. So it had to be in the range of huge particle-physics-temperatures.

Q. : How many degrees ?

A. : I always have to work this out :

Room temperature is a 40th of an electron-volt.

We are talking about temperatures which were high enough to create proton-antiproton pairs, or very likely much hotter than that.

1/40 eV, for room temperature, that is 300 degrees Kelvin. So 1 eV is about 12 000 degrees Kelvin.

A billion electron-volts in necessary to create proton-antiproton pairs. So that is a billion times 12 000 K. We are in the range of $10^{13}$ degrees, or hotter. $10^{13}$ degrees is probably on the very low end.

Remember, you must have heated the universe up enough to create the baryon excess, to create proton-antiproton pairs and that sort of things. But it was probably much hotter than that.

Q. : Are the boson axion fields created during this process [24] ?

A. : Yes. During the slide down axions are created. Then they decay. And in the process of decaying they produce the usual particles : photons, electrons, quarks, anything that creates ordinary energy density.

And they also decay into dark matter for that matter – probably more dark matter than anything else.

Q. : You said that the field $\phi$ always moves in the direction of decreasing potential energy. And now we are at stationary point of potential energy at the bottom of the trough.

24. See chapter 6 for a brief presentation of axions.

Does this imply that the potential energy that is in the universe today is here to stay and will never change ?

A. : That is right. It looks like, at the present time, we are stuck at the minimum. And unless something else happens – that is not built into the equations at the moment –, we will stay there forever.

Now it is interesting to ask : what happens if we stay there forever ? The universe exponentially expands. True it doesn't exponentially expand very rapidly. But still, every 10 billion years or so it doubles.

In another few dozen billion years it will have expanded so much that basically the galaxies will have all receded from each other beyond their respective horizons. Then our galaxy will be alone.

The case of the Andromeda galaxy [25] is special. By then it will probably have merged with ours.

Everything else is moving apart. And we will truly be alone out there. There will be nothing but the Milky Way.

The cosmic microwave background will also have cooled to a temperature which is way below detectable. Everything else in between galaxies will have stretched out and diluted.

This is not due to the sudden, huge and very brief inflation corresponding to the ramp in figure 15, but to the

25. The Andromeda galaxy is 2.5 million light-years from the Earth. It is our nearest neighboring galaxy.

current Lemaître[26] -Hubble expansion which is very slow – the point [ $\phi$, $V(\phi)$ ] not moving at all anymore.

In a few dozen billion years, I think it is a good deal less than that, but let's say in a couple dozen billion years, our galaxy will be absolutely isolated. And physicists who live in it at that time will have no way of knowing that they are part of a larger population of galaxies.

They will not be able to measure anything about astronomy outside their own galaxy. They will not even be able to tell that their universe is undergoing an accelerated expansion.

Remember : how did we discover the current dark energy ? We discovered it by astronomy, by looking at the cosmic microwave background radiation. But, as said, the CMB will have cooled to a temperature which makes it undetectable.

Everything in astronomy will have departed and gone through the horizon. And so we won't be able – not we, but they –, they will not be able to tell that they are in an inflating accelerating universe.

If they come on the scene without history books and so forth, at least for some period of time, let's imagine, they will look around at a universe which is 50 billion light years big. They won't see anything. They won't be able to tell it

26. Georges Lemaître (1894 - 1966), Belgian jesuit, astronomer and cosmologist. Lemaître published a paper describing the expansion of the universe, and its rate, in 1927, two years before Edwin Hubble. The Russian physicist and mathematician Alexander Friedmann (1888 - 1925) had also began to figure out even earlier in the 20's that the universe might be expanding, see chapter 1.

is 50 billion light years in size.

They will send out probes. And the probes will go out a few thousand light years and come back. They will bring out nothing because there will be nothing out there for them to see.

So they will discover themselves to be a unique phenomenon sitting at the center – they will of course call themselves the center – of the universe. And they will be puzzled.

Then some smart young person will come along and say : you know, maybe there is a lot more things out there. But the universe just expanded and those other things maybe just went away.

Q. : Is it clear that all this started from a point source ? Couldn't there have been other point sources that extended back to us ?

A. : What does a point source mean ? Point sources in what ?

– A point without any space, when there was only time, and no space, before the creation of space itself.

– Ok, but if there is no space, except only this tiny little nugget, what is a point source in ?

You are imagining that this thing – this tiny 3-sphere, for instance – is in some other space. But that is the wrong idea.

Q. : Couldn't there be more than one $\phi$ space?

A. : There could have been more than one. But you will have to ask somebody else.

Q. : In a few dozen billion years from now, when we will not see the next galaxy, won't it be analogous to today's situation where we can't see beyond our horizon?

A. : Yes it will be analogous, except that *everything*, apart from our own galaxy, will be out beyond that point.

Q. : Do we have any idea what these inflaton particles might have looked like?

A. : We don't have much idea. The mass of the inflaton particle has to do with the curvature of the curve down at the bottom in figure 15. And we know nothing about it.

The only thing we know is that it is not so light that we would have produced it in an accelerator. That is about all.

Q. : Is this $\phi$ also function of spatial coordinates?

A. : Oh, absolutely. Again, however, $\phi$ might have started with lots of ripples. But in the process of sudden huge inflation those ripples got stretched out.

So after a couple of $e$-foldings, the ripples would have been pretty much eliminated locally, and you can forget them, although we will come back to that in the next lesson.

Q. : Will you talk about those ripples related to quantum fluctuations ?

A. : Not in the present lesson. We will come back to quantum fluctuations.

For the moment, let's just imagine a classical wave form which gets stretched out. That is the primary thing that happened.

Superimposed on top of that will be quantum fluctuations, and they will play an enormously big role when we come back to it.

Q. : Can we today see galaxies slipping over the horizon ?

A. : No, no, no. The horizon is so redshifted, that all we see out there is cosmic microwave background.

When we look back that far, first of all we are seeing the universe at an early time before galaxies even formed.

So we look way out there, and the farthest that we can see is to the surface of decoupling. It is called the surface of last scattering, but that means the place where the universe was opaque. That is pretty much out near the horizon.

We can't see farther than decoupling [27]. But even if we could, we would be looking back so early that there would have been no galaxies.

So the answer is no, we can't see galaxies going through the horizon.

Q. : So it is only at a later time that these objects are going to slip over the horizon ?

A. : But we will never see them slip over [28]. When we look out we will see pretty much the same thing.

Q. : Question relating to information theory : considering that no information is ever lost, what information could be contained in the universe in a couple dozen billion years that could contribute to the knowledge of its origin ?

---

27. Before decoupling, photons were constantly scattered, absorbed by atoms and reemitted. So they don't carry information on what went on before.

28. Go to volume 4 of the collection *The Theoretical Minimum*, on general relativity, to review what happens to objects going through the horizon of an observer, as seen by that observer.

A. : If observers still exist in a couple dozen billion years – if stars exist and so forth –, what they will have is their own little galaxy in an othewise completely empty space.

It is very hard to see how they could do any observations which would directly tell them that they were in a space which was doubling once every 10 billion years or once every 20 billion years.

I can't offhand think of anything that they could do, short of sending out probes which would take of the order of billions of years to go out there and come back, and make measurements on the whole global structure of the universe. Short of that, I can't see how they could tell what the expansion properties of the universe were or anything else.

They could do that. They could take tens of billions of years, and explore. But they can't go too far, because if they go too far, they go outside their horizon and they cannot get back.

They can only go out to distances of order 20 or 30 billion light years, something like that, send signals back and forth which would take billions and billions of years, triangulate, discover the fact that things are receding away from each other. But they would have to come back before they go too far.

So in principle they can do geometric measurements, discover this curvature of spacetime, discover the fact that there is a dark energy. But it would take billions of years, at least with the level of precision of their instruments. I think even with better levels of precision, they will have trouble.

In summary, if their astronomers are anything like our astronomers [29], it will take them billions of years the reconstruct the fact that they are in an inflating universe.

Q. : I guess it is a similar question. Do you think that dark energy can carry information ?

A. : No.

Q. : Concerning gravitational attraction of galaxies, aren't there larger structures such that galaxies will be closer to each other sometimes in the future ?

A. : Yeah, I actually once asked some observational cosmologists : is it really true that all that we will see will be us and the Andromeda ? And I think it is true. I think everything else is participating in the outward flow enough that they will recede.

Q. : Then, if the universe expansion is accelarating, what will be happening to us ?

A. : We are down there and just sitting there. Nothing is

29. Notice that, by then, it will have been a long time since they moved from the Solar system to some more hospitable location in the Milky Way.

happening in this universe.

Q. : Can you express this field $\phi$ as a function of time ?

A. : Why write an equation ? A highly speculative equation, as far as the actual coefficients are concerned, involving time won't add anything.

So I'm saying that [ $\phi$, $V(\phi)$ ] started at point $A$, slowly[30] drifted to $B$, suddenly picked up and plunged down to $C$, figure 17.

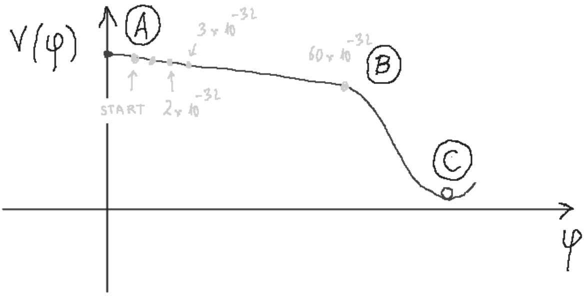


Figure 17 : Main events in the evolution of [ $\phi$, $V(\phi)$ ].

Between $A$ and $B$, it is pretty linear in $\phi$ as well as in $t$. Then, at $B$, it becomes non linear. Between $B$ and $C$, it is non linear in time. Finally today at $C$, it is constant in time. Indeed, after some possible small past oscillations

30. Slowly in the sense of a small gradient of $V(\phi)$ with respect to $\phi$, not with respect to time. Indeed, this theoretical inflation was sudden, huge and of short duration.

quickly damped by friction, the point [ $\phi$, $V(\phi)$ ] doesn't move anymore.

Q. : You mentioned that in a distant future astronomers could actually send probes out. But dark energy[31] is such that at some point in space from us the universe is expanding faster than the speed of light.

A. : Yes, the probes must not go past the horizon if the astronomers want to get them back or even to be able to receive their signals.

But they can certainly go a few billion light years out and then come back. And that should be enough for them to be able to measure carefully properties of space (geometry and expansion). The astronomers would not just send one probe, but a few probes emitting light signals back and forth between each other and base camp. These signals themselves would take billions of years to travel.

They could probably do some triangulation. I haven't thought about it a lot. It is not an urgent thing on my agenda :-)

Q. : Does the horizon recede ?

A. : No. The horizon stays where it is, roughly 50 or 60 billion light years away from us. I don't remember exactly

31. Remember that dark energy, that is the density of energy of vacuum in space, is the same thing as the cosmological constant $\Lambda$.

the value.

Q. : There is less space out there, isn't it ?

A. : Less space ? Oh no ! It is the same space everywhere. The farther you go the more space there is [32].

But there is less stuff in it. There is a lot of space out there, but not much stuff in it.

Q. : How do you define that boundary horizon ? Is that where the recession velocity of galaxies approaches the speed of light ?

A. : Yes.

If you are sitting at the point $C$ in the trough, the potential energy $V(\phi)$ is constant over time. Then, according to equation (29), reproduced below, the Hubble constant $H$ is constant over time.

$$H = \sqrt{\frac{8\pi G}{3}\ V(\phi)} \tag{33}$$

Since $H = \dot{a}/a$, it implies that the universe is expanding

32. If we live in a 3-sphere, it is meaningful to say, though, that at the other extremity of the universe from us space is shrinking. But that is way beyond our horizon. So for practical purposes, space has more and more room, so to speak, as we go farther away. Think, in 2D, of the surface of the Earth to imagine this comfortably.

like this

$$a(t) = e^{Ht} \tag{34}$$

Therefore the receding velocity of any point from us is equal to its distance from us times the Hubble constant [33]. We can write this as follows

$$V = DH \tag{35}$$

Now we can solve for $D$ the equation : at what distance away from us are things receding with the speed of light ?

$$C = DH \tag{36}$$

We find that the speed of light $C$ (no relation with the point $C$ in figure 17) divided by Hubble is that distance, i.e. the distance of the horizon.

The calculation gives something of order 15 billion light years – from the known numbers. That is where things are receding away from us with the speed of light.

Notice that past that horizon, they recede from us at speeds higher than the speed of light.

Q. : By the time any object outside our own galaxy has

33. Remember that the term "Hubble constant" refers to the fact that it is always constant in space. When it happens to also be constant in time, as in the above example, we usually specify it – whence funny sentences like "when the Hubble constant is constant".

receded beyond our horizon, won't our galaxy have disappeared as such too ? Won't all the stars be dead ? And same for the other galaxies, even if they are far away ?

A. : You will have to ask a galactic specialist. But I don't think so. Perhaps the galaxy will have been diluted somewhat, but not a lot.

How many stars alive will there be ? What density of stars ? I don't know. But I think that will survive for a trillion years.

I think there will still be galaxies. There will still be stars. And with any luck there will still be planets, some of them hosting intelligent creatures.

Q. : You spoke about the inflaton as a particle. Is it correct to think of space as having been then filled with many inflatons, like it is filled today with many ordinary particles and energy ?

A. : Yes. That is exactly right. To some extent you can imagine that all the energy was stored in a kind of condensation of the inflatons. And it was released as we came down from point $B$ to point $C$, in figure 17, and the inflatons decayed.

Again, much of what I'm telling you would have to be true in order to explain why the universe is the way it is today, and was the way it was during the observable period. We do not understand with any great insight why is this the picture.

In other words, the inflation theory is a hypothesis trying to explain what we observe. Significant evidence now is gathering which confirms it. And no observation so far appears to starkly contradict it. It raises of course as many questions as it solves. It is live science, with debates, critics, stakes, and emotions.

Let's close this long and interesting questions / answers session.

## Structures in the universe

The next thing on our agenda is to understand why, after all this inflation, isn't the universe completely smooth.

The amount of inflation, that took place in order to make it as flat as it is, would have made it so smooth that there would be no explanation for the structures that are out there – the inhomogeneities that are out there today.

So the next topic we want to come to is the fact that the universe is not homogeneous. And it has a structure. It shows all kinds of structures under various scales.

Let's begin with a very naive, but still more or less right, explanation of why the structure of the universe appears the way that it does.

Now we are not talking about what it traces back to in the early universe. For the moment we are talking about today.

And when we say structures, we don't mean galaxies now. Galaxies are small potatoes on a small scale. We mean on a much bigger scale, not even clusters of galaxies, but whole superclusters of superclusters of galaxies. And they form a structure that fills the universe.

It is not completely homogeneous on those scales. That structure is quite fascinating, and, we think, largely understood. But it is mostly understood by simulations.

The simulations start by assuming the universe was not precisely smooth. So they start with a certain pattern of non-smoothness called the *fluctuation spectrum* – which we are going to study a bit. Then they allow gravity and just the motion of objects and so forth to evolve and they show that it creates a structure.

Now, on the biggest scales of super duper clusters of galaxies, gravity is not all that important. On the scale of a galaxy, and certainly on the scale of a star or of a planet, gravity is what pulls things together. But on bigger scales the structure that is seen is largely not, or to some extent not coming from gravity. It would be there even if there was no gravity.

But what the structure requires to be there is to start with some inhomogeneity, some fluctuation in space. So we are going to talk a little bit about the theory of how that structure got formed. And that theory is extremely simple.

Let's look at three pictures. The first one is an actual photograph of the deep sky, showing things very far away, figure 18.

Figure 18 : Real photograph of the deep sky.
Source : NASA, Hubble telescope.

This picture is very deep astronomy, billions of light years out, high redshifts, etc.

Beyond some stars of the Milky Way, and some galaxies not too far from us, appearing as points, we can see a kind of filamentary structure. The filaments are not single galaxies, there are thousands and thousands, maybe millions of galaxies. The filamentary structures intersect and form a complicated web, at least photographically on the sky.

This filamentary structure is sketched in figure 19. It is an exaggeration to focus on what we are interested in. At the intersection of the apparent filaments, the nodes are particularly bright. And there are big voids in between the filaments.

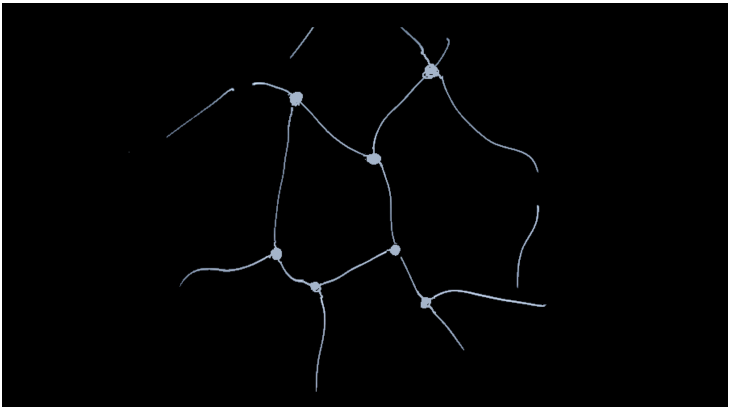

Figure 19 : Filamentary structure appearing in the deep sky.

This structure is on a scale much much bigger than a galaxy, the picture width is half a billion light years or so.

The second picture is the product of a computer simulation made by people, figure 20.

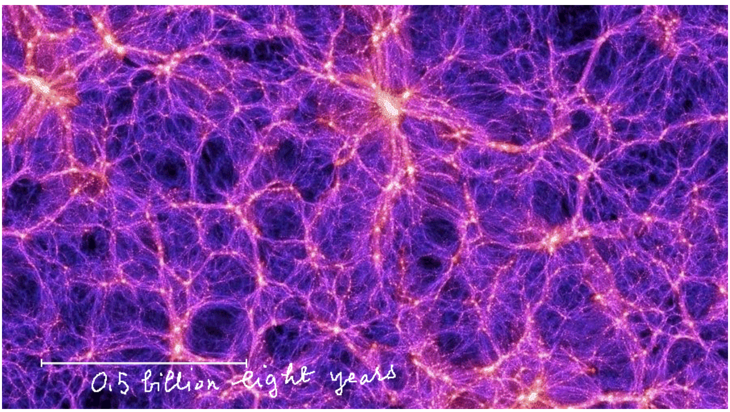


Figure 20 : Computer simulation creating a similar filamentary structure. Source : Volker Springel, Virgo Consortium.

What do they do ? They start with a bunch of particles. Well these are galaxies, but let's speak of them as particles. Initially the particles have some kind of spectrum of fluctuations. Fluctuations means some randomness in their distribution, in particular in their velocity distribution. In one region a space they might be moving a little more this way, in another region moving a little more that way, traced back to some primordial variation which we want to understand. That is what we start with, and we let it evolve for a while.

So a good starting point for this would be a uniform spatial distribution of stuff, but with small variations in velocity. It is the velocity variations which are the most important things. And what do we see after a time in the simulations ? We see a picture which looks just like the real filamentary structure in the sky, except sharper, clearer. Why ? Because it is simulation. We are not looking through a telescope at billions of light years away or more. So we have great detail.

Q. : What interaction are you allowing in those ?

A. : At the moment, none. I am going to expound the theory that we would get if there were no forces.

And the third and last picture is a beautiful blue colour photograph of the bottom of a swimming pool on a sunny day. Light forms patterns at the bottom of the swimming pool, figure 21.

Figure 21 : Light patterns at the bottom of a swimming pool.

If we set picture 2 and picture 3 in the same colour, you would have a hard time telling the bottom of the swimming pool from the simulation of galactic structure. They look very similar. Of course one is in three dimensions, and that makes a big difference. But of course what we see is only two-dimensional.

To summarize, the swimming pool, the real universe, and simulations have a remarkable similarity.

What is the explanation of this ? It stems from something called *caustics*. Caustics were first encountered as an optical phenomenon, for instance the bottom of the swimming pool phenomenon [34] . And we can make a simple one-dimensional version of it.

34. With proper lighting, you can even see some – different from those at the bottom of the swimming pool – with your mug of coffee or tea.

If the world were one-dimensional, caustics would just be spots. If the bottom of the swimming pool was replaced by a one-dimensional bottom, instead of seeing lines like in figure 21, we would just see spots of high intensity, figure 22.

Figure 22 : One-dimensional caustics phenomenon
(for convenience light corresponds to black.)

The phenomenon of caustics is slightly different depending on the number of dimensions :

a) In one dimension, the caustics are the spots of high intensity.
b) In two dimensions, the bottom of the swimming pool shows a web not of spots but of lines of high intensity.
c) And in three dimensions we shall find out what they are replaced by.

Let's explain the phenomenon of caustics. For simplicity we consider the one-dimensional case. Let's imagine a one-dimensional swimming pool, with some length, some depth but almost no width. So it is a thin vertical rectangular tank of water. One horizontal side is the bottom of the swimming pool, the other horizontal side is the top. The other two sides are lateral walls which don't play any role in the explanation. It is represented in figure 23.

Furthermore, the bottom of the swimming pool is represented at the top, and the vertical axis, which is the depth (upside down), is also the time.

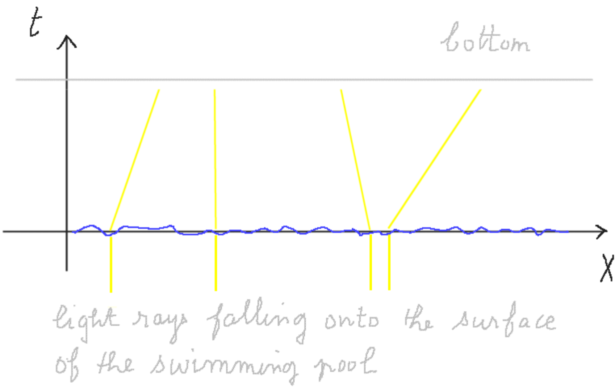


Figure 23 : One-dimensional swimming pool, upside down. The surface of the swimming pool (at the bottom) has small waves which refract light rays in different directions.

There is a little bit of waviness on the surface of the swimming pool (shown at the bottom in the figure). It is sort of a rough surface. So, when light rays come in, they are refracted in directions which vary from place to place.

The different directions of the light rays in the swimming pool, on their way to hit the bottom, are due to the random fluctuations in the slope of the water surface, which act like plenty of lenses. As a result, at the bottom of the swimming pool, we see spots of high intensity separated by intervals of lower intensity.

Let's dig further into the explanation and extend it to the sky. Let's call $X$ the horizontal axis corresponding to the surface of the swimming pool.

Notice we can either think of figure 23 as representing the light waves from the top of the pool to the bottom of the pool (upside down), or we could think of it in a totally different way. We could think of particles starting with different velocity at $t = 0$.

The horizontal axis, instead of being the top of the pool, would be the time $t = 0$. And instead of the slightly different angles that the light refracts through, the particles would have slightly different velocities moving along the horizontal axis. Figure 23 then becomes the usual two-dimensional representation of motions on a one-dimensional horizontal axis, the vertical axis, instead of being the depth of the pool, now being simply time.

So let's start down at $t = 0$, taking the point of view of particles moving. And let's assume for simplicity that initially the particles are uniformly distributed along the $X$-axis, with no tendency to bunch up like the points in figure 22. To sum up : uniformly distributed but slightly different velocities as we move from place to place.

Let's define

$$\frac{dn}{dX} \tag{37}$$

as the number of particles per unit $X$ as we move along the $X$-axis. It is what we would call the *density of particles* on the $X$-axis. It depends on $X$ and on $t$. At $t = 0$ let's take this density to be equal to 1 for all $X$. We can take it anything we want, but let's just say it is uniform, featureless. The magnitude of it is not important. What is important is that it is featureless. So we initially have $dn/dX = 1$.

Now what happens to each particle? Consider a particle which started at position $X$, and let's see where it is at time $t$, figure 24.

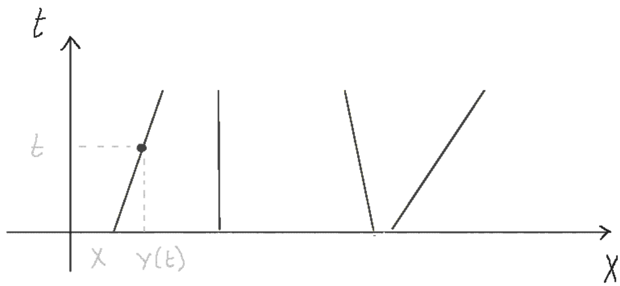


Figure 24 : Motions of particles.

We call $Y(t)$ its position as a function of time. It the particles were not moving, we would just say for each particle

$$Y(t) = X \tag{38}$$

The position of a particle after a certain time would be the same as its position when it started. And $X$ is just a label that labels where the particle was when it started. That is what we would say if the particles weren't moving.

But if the particles are each moving with a velocity $V$, and $V$ depends on $X$, then equation (38) becomes

$$Y(t) = X + V(X)\ t \tag{39}$$

It starts at position $X$, and after a time $t$ its new position is $X$ plus the distance it moved over, namely $V(X)$ times $t$.

Now let's try to compute, at a later time, how many particles there are per unit $Y$. The variable $Y$ is the new coordinate of the particle at time $t$. We are interested in the density at this later time $t$. What we want to compute is

$$\frac{dn}{dY} \tag{40}$$

What do we do ? What is the trick ? The trick is differentiate equation (39) not with respect to time but with respect to $X$. We get

$$\frac{dY}{dX} = 1 + V'(X)\, t \tag{41}$$

We simply differentiated both sides of equation (39) with respect to $X$.

What is the next step ? We have $dn/dX$ ; we have $dY/dX$ ; what we want is $dn/dY$. We can write

$$\frac{dn}{dY} = \frac{dn}{dX}\,\frac{dX}{dY} \tag{42}$$

But $dn/dX$ is 1, and $dX/dY$ is just the inverse of what we calculated in equation (41). So let's put that in

$$\frac{dn}{dY} = \frac{1}{1 + V'(X)t} \tag{43}$$

Where do you think these bright spots are after time $t$ ? They are where the denominator is equal to zero. The ratio

$dn/dY$ is the density after time $t$. Of course in the geometric optics approximation, when the denominator is 0, we are going to get something infinite. In the wave theory it is not really infinite. Anyway we are going to get especially bright spots, at the places where $1+V'(X)t$ is equal to zero. Let's see if we can figure out what that looks like.

Let's plot $-V'(X)$, figure 25.

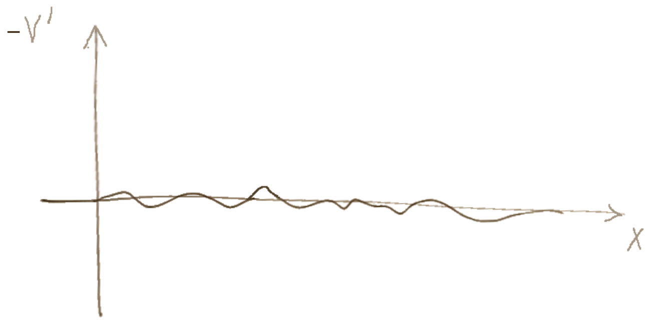


Figure 25 : Plot of $-V'(X)$.

$V'(X)$ varies around its mean. We can assume that its mean is zero. And $V'(X)$ fluctuates around 0. In some places it is positive ; in some places it is negative.

And when does $1 + V'(X)t$ become infinite ? It is when

$$- V'(X)\ t\ =\ 1 \tag{44}$$

The height, that is the time $t$ in this equation, acts like a scaling factor. When $t$ is small, $-V'(X)t$ is small, and never reaches 1. When $t$ increases, after say 10 seconds, the scaling factor increases. Now the question is when $-V'(X) = 1/10$.

Eventually, after a certain amount of time, a first peak in $-V'(X)t$ will reach 1, figure 26.

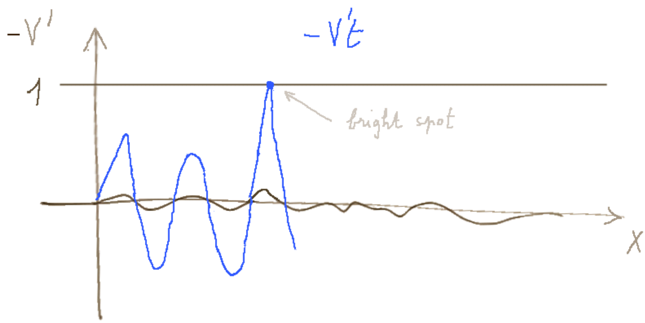


Figure 26 : Plot of $-V'(X)$ and of $-V'(X)t$ for a given $t$.

What happens when $-V'(X)t = 1$ ? Then the denominator on the right-hand side of equation (43) becomes 0. That is where we get a lot of intensity.

We use the term intensity in a generic sense. If it was light rays, it would actually be intensity ; we would get a bright spot. If it is particles, we get a huge enhancement of the particle density of that point.

If we go further in time, at the first bright spot in figure 26 the curve $-V'(X)t$ will now go through the ceiling and create two spots. Elsewhere there will be more bright spots. These are caustics. Caustic means sharp. The caustics are the places where we see sharp features. At the bottom of a one dimensional swimming pool the caustics are these points where we get sharp light spots.

We can do the same exercise in two dimensions. We would start with a two-dimensional space. The coordinate $X$ would actually represent two spatial dimensions, corresponding either to a two-dimensional world or to the top of an ordinary swimming pool. And we would evolve either in time or in height along the third vertical dimension.

What we would see is basically figure 19. There would be caustic lines of high intensity. And at the intersection of the lines there would be points of extra high intensity. That is indeed what we see at the bottom of the swimming pool, figure 21. So we get a network of lines and spots, the highest intensity being at the spots, high intensity along the lines, and big voids of light in the regions in between.

In three dimensions we can play exactly the same game. Now we are starting with a three dimensional world and a uniform distribution of particles – in the sense of evenly spread – but a velocity dispersion of the same kind as above. What we will get is surfaces. The surfaces will be intersecting other surfaces in a complicated form, like an aggregate of soap bubbles [35].

The caustics will be surfaces. The surfaces from time to time and place to place will intersect along lines. And the lines will intersect along points. That is what we see when

35. But the surface tension of soap bubbles will usually make them distinguishable from a computer simulation of galactic caustic structures.

we look at a computer simulation as in figure 20. We cannot clearly see the surfaces because the image is in 2D. But we do see the lines, and the hotspots where they intersect. And it is easy to imagine in 3D the surfaces. They are caustics.

So this is what we would expect in a world without gravity. No gravity has been invoked in the model. All that was invoked is a somewhat random distribution of very small velocities at the starting point.

Now it would not hurt to also have some variation in density along the $X$-space at time $t = 0$ or, equivalently, at the initial height. It wouldn't make much difference though. It would basically be the same thing.

That is, if you like, a very simple but not wholly inaccurate theory of the very large scale structures, caustics, that we can observe when we look at the universe very very far – when we look at what is called the deep sky.

On smaller scales of course gravity plays a big role. It causes things to tend to condense into galaxies and so forth.

But the reason why I emphasize this story of caustics is because we have to have a theory of this variation in the initial velocity distribution, and where did it come from. We said that we stretch the universe out to the point where it is incredibly flat and uniform. So where does the starting "random"[36] distribution of velocities come from ?

36. We use the word "random" for convenience. But remember that randomness is essentially attached to experiments which can be reproduced, and which would produce different outcomes.

That in itself is an inhomogeneity. So the next thing we are going to study is the quantum origin of these inhomogeneities in the universe, and how it lead to the starting point for the filamentary structures and everything else in the sky.

These notes come from the site of the notetaker at `https://www.lapasserelle.com/cosmology` The next notes are `https://www.lapasserelle.com/statistical_mechanics`