8 Inflation: background

8.1 Motivation

Inflation is a scenario in which there was a period of accelerated expansion in the very early universe. While it has not been established beyond reasonable doubt that inflation took place, inflation, like dark matter, is a very successful hypothesis. Inflationary models have made several detailed predictions that have been observationally verified, and no competing scenario has had the same success.

The word "inflation" also refers to the period of accelerated expansion, and also to the accelerated expansion itself. For example, we say that the "the universe inflates" to mean that the expansion accelerates. The motivation for inflation came from some unresolved issues of the "Big Bang model", i.e. the homogeneous and isotropic FRW model with matter consisting of a gas of particles, which we have considered thus far.

8.1.1 Homogeneity and isotropy problem, or the horizon problem

One question concerns the origin of the symmetry of the FRW model. There is no unique way to define a measure on the "set of spacetimes" (though attempts have been made), but the homogeneous and isotropic universes seem very special. Homogeneity and isotropy of the universe have two distinct aspects. First, departures from homogeneity and isotropy are small, i.e. the differences in any physical quantity calculated at two different spatial points are small. This is the case in the early universe, as we know from the fact that the amplitude of the CMB perturbations is only $\sim 10^{-5}$ (apart from the dipole component which is 10^{-3} and which is presumably overwhelmingly due to our motion)[1]. Today, perturbations are locally large in the sense that differences in the local value of the energy density, expansion rate and so on are large. The density of a galaxy can typically be 10^6 times larger than the mean density, and a void can expand faster than the mean by tens of percent. However –and this is the second aspect of homogeneity and isotropy– the universe is still statistically homogeneous and isotropic today, because the initial distribution of small perturbations had this symmetry (we will discuss this in more detail later). So the homogeneity and isotropy question is related to the origin of the seeds of structure. Where did the small deviations come from, and why is their distribution the observed one? A particular aspect of this the *horizon problem*. The CMB anisotropies are correlated on all observed scales. However, at the time of last scattering any regions which are today separated by more than about 1° had not had time to interact in the Big Bang model (i.e. assuming the FRW metric and ordinary matter).

[1] Strictly speaking, this is the amplitude of deviations in the distribution of photons. As baryonic matter was in equilibrium with the photons, it was also close to homogeneous and isotropic. However, departures from homogeneity and isotropy were larger in the dark matter, which is decoupled from photons. This is crucial for structure formation, as we will discuss in chapter 11.

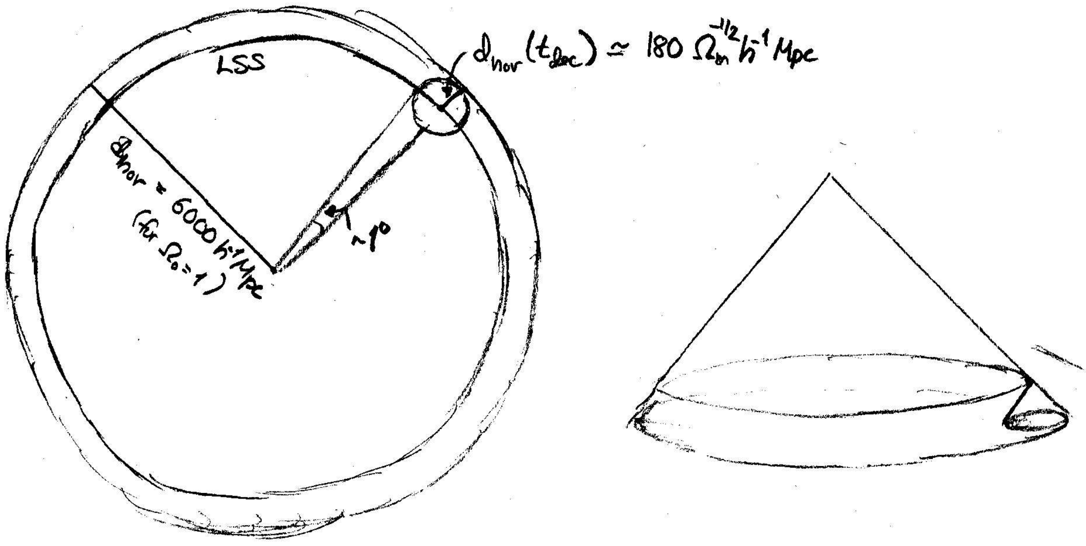

Figure 1: The horizon problem.

8.1.2 The flatness problem

Another issue is the spatial flatness of the universe. The density parameter is

$$\Omega - 1 = \frac{K}{(aH)^2} \ . \tag{8.1}$$

If Ω is unity at some time, it is always unity. If $\Omega \neq 1$ at any time, it evolves in time. Assuming that the spatial curvature is initially small, we have

$$\text{mat. dom.} \qquad a \propto t^{2/3},\ H \propto t^{-1} \Rightarrow \frac{1}{aH} \propto t^{1/3} \qquad \Rightarrow |1-\Omega| \propto t^{2/3} \propto a \tag{8.2}$$

$$\text{rad. dom.} \qquad a \propto t^{1/2},\ H \propto t^{-1} \Rightarrow \frac{1}{aH} \propto t^{1/2} \qquad \Rightarrow |1-\Omega| \propto t \propto a^2 \ . \tag{8.3}$$

If the spatial curvature is positive, it will quickly dominate over matter or radiation, and the expansion will stop and turn around, and the universe will collapse. If the spatial curvature is negative, the universe will quickly become empty and cold. The flatness problem is thus also called the oldness problem. In the ΛCDM model, the curvature today is very small, $|\Omega(t_0) - 1| < 10^{-2}$. Therefore, at BBN we have $|\Omega(t_{\rm BBN}) - 1| \lesssim 10^{-19}$, which seems like a strong tuning. (Of course, if the spatial curvature is initially zero, it will remain zero. However, this is a special value, for which we would like to have an explanation.)

8.1.3 The relic problem

At early times in the Big Bang model the temperature is very high. In grand unified theories of particle physics there are phase transitions at high temperatures (in the Standard Model, there are no such phase transitions, only the QCD and electroweak crossovers). These phase transitions can generate topological defects such as magnetic monopoles, cosmic strings and domain walls, which correspond to (approximately) zero-, one- and two-dimensional topological defects. Just like (given a specific model) we can calculate the relic density of dark matter particles, we can calculate the density of these relics. In some models the energy density of monopoles today would be much higher than the observed energy density. This is related to the fact that monopoles are typically very massive, with masses of the order of the grand unified scale. The presence of cosmic strings and domain walls would also be problematic, as they are also typically very heavy, and their energy density relative to ordinary matter increases with time (i.e. it goes down more slowly). In supersymmetric models one particular problem is the overproduction of gravitinos, the supersymmetric partners of the graviton. If gravitinos are not stable, their lifetime is very long, because they interact only gravitationally, so they typically decay after BBN, and ruin its observational success. It is however also possible that the gravitinos form the dark matter. The constraint on the temperature of the universe from gravitinos is of the order $T \lesssim 10^7$ GeV. However, it may be that the grand unified theories or supersymmetric models do not describe reality, in which case this is not a problem. Observationally, from BBN, we know only that the universe has been at least as hot as 1 MeV.

8.1.4 What is needed

All of the above problems are solved if we have a mechanism which produces an "initial condition" for the universe at $T > 1$ MeV, where the universe is homogeneous and isotropic up to small perturbations that are correlated on all observable scales, where the spatial curvature is very small and matter is in thermal equilibrium (at least the part which consists of baryons, photons and neutrinos). In specific theories of particle physics, there may be an upper limit on the temperature.

8.2 Inflation introduced

Inflation is not a replacement for the Hot Big Bang model, but an add-on, occurring at very early times (somewhere between the energy scales of MeV and 10^{16} GeV, in most models closer to the upper end) without disturbing its successes. Inflation is defined as accelerating expansion:

$$\text{inflation} \Leftrightarrow \ddot{a} > 0 \ . \tag{8.4}$$

Often the term inflation is used to refer only to a period of acceleration expansion in the early universe, and not to the recent phase of accelerated expansion.

Consider how the flatness and horizon problems can be solved with inflation. The origin of the flatness problem is that $|\Omega - 1| = |K|/(aH)^2 = |K|\dot{a}^{-2}$ grows with time because $\dot{a}$ falls, i.e. the universe decelerates. If the expansion instead accelerates, Ω is driven towards unity starting from any value. (This is the case for an expanding universe. If the universe contracts, the behaviour is reversed.)

Consider now the horizon problem. The problem is that in the standard Big Bang model the horizon at the time of photon decoupling is small compared to the part of the universe we can see today. In standard Big Bang picture the universe is first radiation-dominated and then becomes matter-dominated somewhat before photon decoupling. (Recall that $t_{\text{eq}} \approx 50\ 000$ years and $t_{\text{dec}} \approx 380\ 000$ years.) In the radiation-dominated era, the horizon is $d_{\rm hor}(t) = 2t = H^{-1}$. In the matter-dominated era, we have $d_{\rm hor}(t) = 3t = 2H^{-1}$. The horizon at decoupling is between these values, $H^{-1} < d_{\rm hor}(t_{\rm dec}) < 2H^{-1}$. The size of the observable universe today is of the order of the present Hubble length $d_{\rm hor}(t_0) \sim H_0^{-1}$ (the precise prefactor depends on the vacuum energy density, but the order of magnitude is enough for us). The presently observable universe was a factor of $a_{\rm dec}/a_0$ smaller at decoupling.

In order to compare size of the horizons at different times we can use comoving lengths, where this change is taken into account. The comoving horizon at decoupling is

$$d^c_{\rm hor}(t_{\rm dec}) = (1+z_{\rm dec})d_{\rm hor}(t_{\rm dec}) \sim (1+z_{\rm dec})H^{-1}_{\rm dec} = (a_{\rm dec}H_{\rm dec})^{-1} \ , \tag{8.5}$$

and $d_{\rm hor}(t_0) = d^c_{\rm hor}(t_0) \sim H_0^{-1}$. The horizon problem arises because the first horizon is much smaller than the second,

$$\frac{d^c_{\rm hor}(t_{\rm dec})}{d^c_{\rm hor}(t_0)} \sim \frac{a_0 H_0}{a_{\rm dec}H_{\rm dec}} \sim (1+z_{\rm dec})\frac{t_{\rm dec}}{t_0} \approx 0.03 \ll 1 \ , \tag{8.6}$$

where we have for clarity inserted a_0, even though it is equal to unity, and used $t_{\rm dec} = 380\,000$ yr, $t_0 = 14 \times 10^9$ yr and $z_{\rm dec} = 1090$. In other words, the presently observed universe was about 30 times larger than the particle horizon at decoupling, so it contained about 10^5 regions that had never been in causal contact with each other. Thus the problem is again that aH decreases with time,

$$\frac{d}{dt}(aH) = \ddot{a} < 0 \ , \tag{8.7}$$

so a period with $\ddot{a} > 0$ might solve the problem.

Recall that the particle horizon refers to the maximum distance that light could in principle have travelled from the beginning of the universe until time t. If we add a new period in the early universe to the matter-dominated era and the radiation-dominated era, such as like accelerating expansion, the calculation of $d^c_{\rm hor}$ will depend on it. We always have $d^c_{\rm hor}(t_0) > d^c_{\rm hor}(t_{\rm dec})$ since $t_0 > t_{\rm dec}$, and the interval $(0, t_{\rm dec})$ is included in the interval $(0, t_0)$. However, in the horizon problem, the relevant present-day quantity is not actually the distance from which we could in principle have received signals, but the distance from which we actually measure signals. Because the universe is opaque before decoupling, the size of the present observable universe is given by the distance photons have travelled in the interval $(t_{\rm dec}, t_0)$, and this is not affected by what happens before $t_{\rm dec}$. Thus the relevant present-day scale is always $\sim H_0^{-1}$.

Note that the comoving Hubble parameter is equal to the conformal Hubble parameter,

$$aH = \frac{1}{a}\frac{da}{d\eta} = \dot{a} \ , \tag{8.8}$$

where η is conformal time. The Hubble length is

$$l_H \equiv H^{-1} \ , \qquad \text{where } H \equiv \frac{\dot{a}}{a} \ , \tag{8.9}$$

and the *comoving Hubble length* is

$$l^c_H \equiv \frac{1}{a} l_H = \frac{1}{aH} = \frac{1}{\dot{a}} \ . \tag{8.10}$$

If aH decreases, then $(aH)^{-1}$ increases, and vice versa. So we can say that inflation is any epoch when the comoving Hubble length shrinks:

$$\text{inflation} \Leftrightarrow \frac{d}{dt}\left(\frac{1}{aH}\right) < 0 \ . \tag{8.11}$$

It has unfortunately become customary in cosmology to use the word "horizon" also for the Hubble distance, particularly with regard to inflation. We adopt this lamentable practice when referring to subhorizon or superhorizon modes (to be defined a bit later), but will otherwise try to be careful not to confuse the two concepts.

Let us consider an example of accelerated expansion that we are already familiar with from the discussion on dark energy, namely exponential expansion, corresponding to the vacuum energy equation of state $w = -1$, $a(t) \propto e^{Ht}$, with constant H. We will shortly see that this a first approximation for the expansion law during inflation. The horizon distance is

$$d_{\text{hor}}(t) = a(t) \int_0^t \frac{\mathrm{d}t'}{a(t')} = H^{-1} e^{Ht} (1 - e^{-Ht}) \simeq H^{-1} e^{Ht} \ , \tag{8.12}$$

where the last limit is for $t \gg H^{-1}$. So in contrast to the radiation- and matter-dominated eras, the physical particle horizon grows exponentially, and the comoving particle horizon stays almost constant, $d^c_{\text{hor}}(t) \simeq H^{-1}$. The present observable universe has evolved from a small patch of a much larger causally connected region. See figure 2.

However, even though the particle horizon grows exponentially, the distance over which it is possible to send signals does not grow. If we consider a light ray, we have $0 = \mathrm{d}s^2 = -\mathrm{d}t^2 + a(t)^2 \mathrm{d}r^2$, so the comoving coordinate separation (which is the comoving distance, since the universe is spatially flat) between emission at t_1 and reception at t_2 is (taking $a = e^{Ht}$)

$$\Delta r = \int_{t_1}^{t_2} \frac{\mathrm{d}t'}{a(t')} = H^{-1} (e^{-Ht_1} - e^{-Ht_2}) < H^{-1} \ . \tag{8.13}$$

If the coordinate separation between two points is more than the Hubble length, it is not possible to send signals between them. In this sense, the Hubble length gives the comoving size of the region during inflation over which it is possible to retain causal connection. If the universe before inflation is matter-dominated, for example, observers with separation $2(aH)^{-1}$ have been able to send signals to each other, so causal connection is lost. Also, regardless of what happened before inflation, during inflation a signal sent at t_1 cannot travel a longer coordinate distance than $H^{-1}e^{-Ht_1}$, and this distance gets smaller as t_1 grows, so causal connection is lost during inflation. Note that the particle horizon, which expresses the maximum distance at which parts of the universe can have been in causal contact always grows as a function of time, it never shrinks. What changes during inflation is just that regions that once were in causal contact cannot send signals to each other any more.

The Friedmann equations are

$$3\frac{\dot{a}^2}{a^2} = 8\pi G_{\text{N}} \rho - 3\frac{K}{a^2} \tag{8.14}$$

$$3\frac{\ddot{a}}{a} = -4\pi G_{\text{N}} (\rho + 3p) \ . \tag{8.15}$$

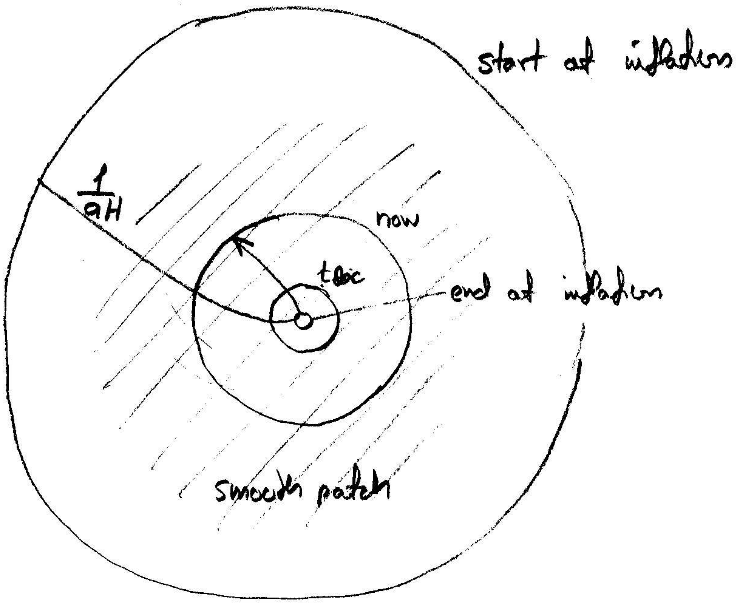

Figure 2: Evolution of the comoving Hubble radius (length, distance) during and after inflation (schematic).

Thus, in general relativity and assuming the FRW metric, accelerating expansion requires negative pressure:

$$\text{inflation} \;\Leftrightarrow\; \rho + 3p < 0 \; . \tag{8.16}$$

Note that the energy density of matter for which $p/\rho < -\frac{1}{3}$ falls down in an expanding universe slower than a^{-2}, i.e. it grows relative to the spatial curvature. (If $p/\rho < -1$, the energy density actually rises as the universe expands.) The flatness problem of the Big Bang model is simply the feature that for matter composed of a gas of particles we have $p \geq 0$, so the energy density falls at least as fast as a^{-3}, and the curvature term will at some point overtake the energy density. In inflationary models, the energy density typically remains nearly constant during a period in which the scale factor grows by a huge factor, typically by a factor e^{60} or more. Thus inflation predicts that $\Omega_0 = 1$ to extremely high accuracy[2]. See figure 3.

As for the relic problem, if unwanted relics are produced before inflation, they are diluted to practically zero density by the expansion, just like spatial curvature. However, we have to be careful that they are not produced after inflation, i.e. the reheating temperature (see below) has to be small enough. This is one constraint on models of inflation. At the end of inflation, matter is produced in *reheating*[3],

[2] If it were discovered by observations that $\Omega_0 \neq 1$, this would be a blow to the credibility of inflation. However, there is a version of inflation, called *open inflation*, for which it is natural that $\Omega_0 < 1$. The existence of such models of inflation have led critics of inflation to complain that inflation is "unfalsifiable" in the sense that no matter what the observation, a model of inflation can be found that agrees with it. Nevertheless, most models of inflation give similar "generic" predictions, including $\Omega_0 = 1$ to great accuracy, and thus far the observations have been in good agreement with them.

[3] "Reheating" may turn out to be as much a misnomer as "recombination", as it is not clear whether matter was ever in a thermal state before inflation.

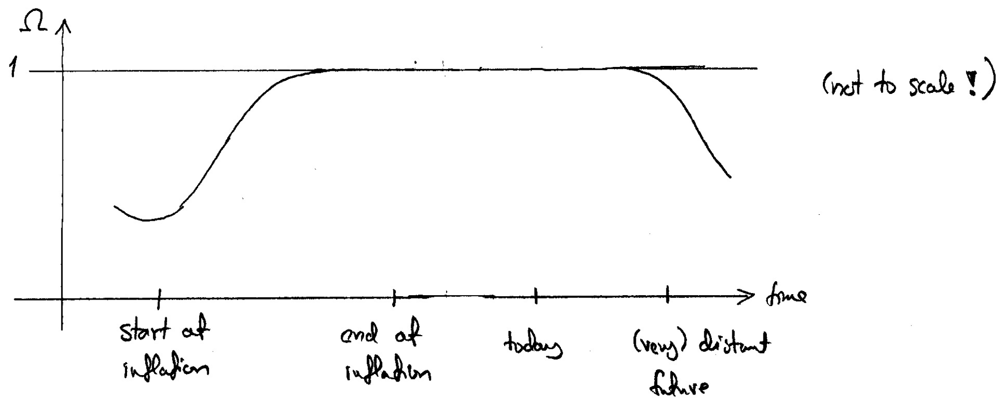

Figure 3: Solving the flatness problem. This figure is for a universe with no dark energy, where the expansion keeps decelerating after inflation ended in the early universe. Present observational evidence indicates that actually the expansion began accelerating again a few billion years ago. Thus the universe is, technically speaking, inflating again, and Ω is again being driven towards 1. However, this current epoch of inflation is not enough to solve the flatness problem, or the other problems, since the universe has only expanded by about a factor of 2 during it.

which produces the gas of particles that is the initial condition for the hot Big Bang model.

Inflation is better called a scenario rather than a theory. It is an idea of a certain kind of behaviour of the universe, which is realised in hundreds of different models. Some of the models are related to extensions of the Standard Model of particle physics or extensions of general relativity, and some of them are just "toy models", which have the right features and are presumably at most an approximate description of some more complicated physics. One noteworthy inflationary model is based on the Standard Model Higgs boson coupled to gravity in a non-standard way.

The important point is that inflation makes many "generic" predictions, i.e. predictions that are independent of the particular model of inflation, for most models. (Though exceptional models can be found that would violate one or more of these general features.) There are also numerical predictions of cosmological observables that differ from one model of inflation to another, allowing observations to rule out models (and some have been already ruled out). Present observational data agrees with the generic predictions (which were made before the advent of the observations of the CMB anisotropies, which are the most direct way of testing the models), while alternatives to inflation have not managed to explain the observations in an equally simple and successful way. Most cosmologists thus consider it likely that inflation took place in the primordial universe. To quote the cosmologist Douglas Scott, *"something like inflation is something like proven"*.

Exercise: Assume that at the beginning of inflation we have $|\Omega_K| = 0.1$. Calculate, as a function of the reheating temperature T_{reh}, how many e-folds of inflation are required to reduce present-day spatial curvature to $|\Omega_{K0}| < 10^{-2}$. (Assume $h = 0.7$ and that neutrinos are massless.) Approximate that the expansion rate at the beginning of inflation is completely dominated by the inflaton, that the inflaton field value does not change during inflation and that reheating happens instantaneously. In which directions do the above approximations change the result? What is the number of e-folds for $T_{\rm reh} = 10^7$ GeV?

8.2.1 Starting inflation

In the discussion, we have already assumed that we can use the FRW metric, i.e. that the universe is homogeneous and isotropic. In order to explain how inflation produces homogeneity and isotropy and solves the horizon problem, we should consider how inflation gets started from some generic initial conditions. Inflation certainly makes the homogeneity and isotropy problem "exponentially smaller" in the sense that it produces a large homogeneous and isotropic causally connected patch from a small one. However, the issue of how to get inflation started remains an open question. There are some ideas and studies of this, but as we have no solid theoretical understanding (and no observations at all) of the pre-inflationary era, the issue remains rather speculative. We will comment on this a bit more after discussing the simplest inflationary models.

We will assume that sufficient inflation has already taken place to make the universe (within a horizon volume) spatially flat, homogeneous and isotropic, and follow inflation in detail after that. Thus we will work in the flat FRW universe. In any case, from the modern point of view, the most important (and testable) aspect of inflation is the generation of the seeds of structure, which we will discuss in chapter 10, as it makes deviation from homogeneity and isotropy quantitative into a quantitative issue.

8.3 The inflaton field

As we saw in section 8.2, inflation requires negative pressure. In chapter 5 we considered systems of particles where interaction energies can be neglected (the ideal gas approximation). For such systems the pressure is always non-negative[4]. However, the particle picture is not fundamental. In the early universe, at high energy densities, we have to consider the more fundamental entities, *fields*. Particles are just excitations of fields. The mean value of a field can have negative pressure, even if a gas consisting of the particles corresponding to the field does not. The simplest form of matter which has a negative pressure is a scalar field, so the simplest inflationary models involve just a single scalar field. The field responsible for inflation (and the corresponding spin 0 particle) is called the *inflaton*.

The starting point is the *Lagrangean density* $\mathcal{L}(\varphi, \partial^\mu\varphi)$, where φ is the inflaton field. In the simplest case where the kinetic term of the field has the *canonical* form and the field is minimally coupled (see below), we have

$$\mathcal{L} = -\frac{1}{2}g^{\mu\nu}\partial_\mu\varphi\,\partial_\nu\varphi - V(\varphi)\ . \tag{8.17}$$

where $V(\varphi)$ is the *potential* of the field. The *action* is correspondingly

$$S = \int d^4x\sqrt{-g}\mathcal{L}\ , \tag{8.18}$$

where g is the determinant of the metric. The effect of spacetime curvature is manifested via the metric in the kinetic term and the determinant of the metric in

[4]A gas of interacting particles could have negative pressure.

the integration measure. This case is called the *minimal coupling* to gravity. It is also possible to include in the Lagrangean density terms which couple the scalar field to quantities built from the derivatives from the metric. (Such a non-minimal coupling is important if the Higgs field is the inflaton.) We will not discuss non-minimal coupling.

If the field is free, we have

$$V(\varphi) = \frac{1}{2}m^2\varphi^2 \ , \tag{8.19}$$

and the mass of the particle corresponding to the field φ is m. If the potential has higher order terms, these describe self-interactions of the field. Even when the potential is more complicated than in eq. (8.19), we define the quantity $m^2(\varphi) \equiv V''(\varphi)$. For $m^2 > 0$, this gives the mass of the particles when the field has the value φ. In the case $m^2 < 0$, the field configuration is unstable, and small perturbations no longer describe particles with mass m. We also use the notation

$$V'(\varphi) \equiv \frac{dV}{d\varphi} \qquad \text{and} \qquad V''(\varphi) \equiv \frac{d^2V}{d\varphi^2} \ . \tag{8.20}$$

Minimisation of the action leads to the *Euler–Lagrange equation*

$$\frac{\partial(\sqrt{-g}\mathcal{L})}{\partial\varphi} - \partial_\mu \frac{\partial(\sqrt{-g}\mathcal{L})}{\partial[\partial_\mu\varphi]} = 0 \ . \tag{8.21}$$

For the Lagrangean density (8.17) we get the field equation

$$-\frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu\varphi) + V' = 0 \ . \tag{8.22}$$

In flat spacetime (Minkowski space), we have $g^{\mu\nu} = \mathrm{diag}(-1,1,1,1)$, so we get

$$\ddot{\varphi} - \nabla^2\varphi + V' = 0 \ . \tag{8.23}$$

For a free field, we have $V'(\varphi) = m^2\varphi$, and the equation of motion reduces to the *Klein-Gordon equation*. For the spatially flat FRW metric we have (in Cartesian coordinates) $g^{\mu\nu} = \mathrm{diag}(-1, a^{-2}, a^{-2}, a^{-2})$, so we get

$$\ddot{\varphi} + 3H\dot{\varphi} - a^{-2}\nabla^2\varphi + V' = 0 \ , \tag{8.24}$$

where $\nabla^2\varphi \equiv \delta^{ij}\partial_i\partial_j$. During inflation, the field (like the space) is almost homogeneous, so we can take $\partial_i\varphi = 0$ for the background evolution (we will consider perturbations in chapters 9 and 10). In fact, inflation makes the inflaton field more homogeneous, as the coefficient a^{-2} falls. A sufficient level of initial homogeneity of the field is required to get inflation started. We start our discussion when a sufficient level of inflation has already taken place to make the gradients negligible, so that the field can be considered homogeneous.

The Lagrangean density also gives us the energy-momentum tensor

$$T_{\mu\nu} = -\frac{\partial\mathcal{L}}{\partial(\partial^\mu\varphi)}\partial_\nu\varphi + g_{\mu\nu}\mathcal{L} \ , \tag{8.25}$$

which for the Lagrangean density (8.17) is

$$T_{\mu\nu} = \partial_\mu\varphi\partial_\nu\varphi - g_{\mu\nu}\left(\frac{1}{2}g^{\alpha\beta}\partial_\alpha\varphi\partial_\beta\varphi + V\right) . \tag{8.26}$$

For the FRW metric, the energy density and pressure measured by an observer comoving with the FRW metric are[5]

$$\rho = -T^0{}_0 = \frac{1}{2}\dot{\varphi}^2 + V \tag{8.27}$$

$$p = T^i{}_i = \frac{1}{2}\dot{\varphi}^2 - V , \tag{8.28}$$

The field has negative pressure when the potential dominates over the kinetic term, i.e. when the field is moving slowly. The equation of state parameter $w \equiv p/\rho$ is

$$w = \frac{\dot{\varphi}^2 - 2V(\varphi)}{\dot{\varphi}^2 + 2V(\dot{\varphi})} = \frac{1 - 2V/\dot{\varphi}^2}{1 + 2V/\dot{\varphi}^2} , \tag{8.29}$$

so

$$-1 \leq w \leq 1 . \tag{8.30}$$

If the kinetic term $\frac{1}{2}\dot{\varphi}^2$ dominates, $w \approx 1$; if the potential term $V(\varphi)$ dominates, $w \approx -1$. Different inflaton models have different potentials $V(\varphi)$. From (8.27), we can form the useful combinations

$$\begin{aligned} \rho + p &= \dot{\varphi}^2 \\ \rho + 3p &= 2\left(\dot{\varphi}^2 - V\right) . \end{aligned} \tag{8.31}$$

We have the equation of motion of the field from (8.24). Alternatively, we could just insert the energy density and pressure from (8.27) into the continuity equation

$$\dot{\rho} = -3H(\rho + p) . \tag{8.32}$$

This gives the same result,

$$\ddot{\varphi} + 3H\dot{\varphi} = -V' . \tag{8.33}$$

This is the field equation for a homogeneous field in a spatially flat FRW universe. The effect of expansion is to add the term $3H\dot{\varphi}$, which acts like friction and slows down the evolution of φ.

The condition for inflation, $\rho + 3p < 0$, is satisfied if

$$\dot{\varphi}^2 < V . \tag{8.34}$$

Let us assume that φ is initially far from the minimum of $V(\varphi)$. The potential then pulls φ towards the minimum (see figure 4). If the potential has a suitable (sufficiently flat) shape, the friction term soon makes $\dot{\varphi}$ small enough to satisfy (8.34), even if it was not satisfied initially.

[5] Those used to the Einstein summation convention should note that there is no summation over i in (8.28).

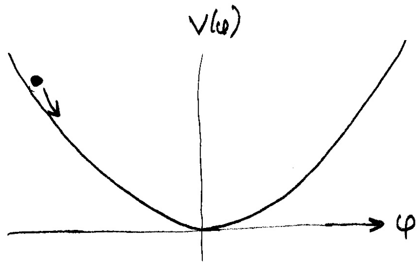

Figure 4: An example of inflaton potential.

We also need the Friedmann equation,

$$H^2 = \frac{8\pi G}{3}\rho = \frac{1}{3M_{\rm Pl}^2}\rho \ . \tag{8.35}$$

Inserting the energy density from (8.27), we have

$$\boxed{H^2 = \frac{1}{3M_{\rm Pl}^2}\left[\frac{1}{2}\dot\varphi^2 + V\right] \ .} \tag{8.36}$$

We have ignored other contributions to the energy density and pressure besides the inflaton. During inflation, the inflaton moves slowly, so the inflaton energy density, which is dominated by $V(\varphi)$, also changes slowly. If there are matter and radiation components in the energy density, they decrease fast, $\rho \propto a^{-3}$ or $\propto a^{-4}$, and soon become negligible, like the spatial curvature. The presence of extra matter can put some constraints on the initial conditions for inflation to get started and the inflaton to become dominant. But once inflation begins, we can soon forget components other than the inflaton.

8.4 Slow-roll inflation

The friction (expansion) term tends to slow down the evolution of φ, so the system easily reaches a situation where the following conditions hold:

$$\dot\varphi^2 \ll V \tag{8.37}$$

$$|\ddot\varphi| \ll 3H|\dot\varphi| \tag{8.38}$$

These are the *slow-roll conditions*. If the slow-roll conditions are valid, we may approximate (the *slow-roll approximation*) (8.33) and (8.36) by the *slow-roll equations*:

$$H^2 = \frac{V}{3M_{\rm Pl}^2} \tag{8.39}$$

$$3H\dot\varphi = -V' \ . \tag{8.40}$$

The shape of the potential $V(\varphi)$ determines the *slow-roll parameters*, defined as

$$\varepsilon(\varphi) \equiv \frac{1}{2}M_{\rm Pl}^2\left(\frac{V'}{V}\right)^2 \tag{8.41}$$

$$\eta(\varphi) \equiv M_{\rm Pl}^2\frac{V''}{V} \ . \tag{8.42}$$

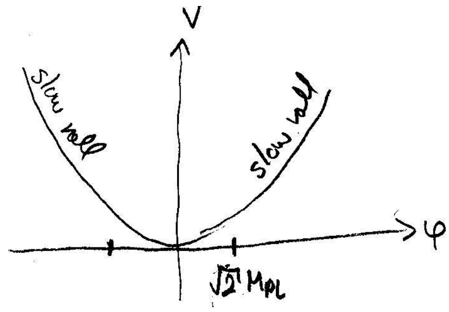

Figure 5: The potential $V(\varphi) = \frac{1}{2}m^2\varphi^2$ and its two slow-roll sections.

Exercise: Show that

$$\varepsilon \ll 1 \quad \text{and} \quad |\eta| \ll 1 \qquad \Leftarrow \qquad (8.37) \text{ and } (8.38) \tag{8.43}$$

Note that the implication goes only in this direction. The conditions $\varepsilon \ll 1$ and $|\eta| \ll 1$ are necessary, but not sufficient for the slow-roll approximation (i.e. the slow-roll conditions) to be valid. The conditions are not sufficient, because they only constrain the form of the potential, and identify from the potential a *slow-roll section*, where the slow-roll approximation *may* be valid. Since the field equation (8.33) is second order, it accepts arbitrary φ and $\dot{\varphi}$ as initial conditions. Thus (8.37) and (8.38) may not hold initially, even if φ is in the slow-roll section. However, it turns out that the *slow-roll solution*, the solution of the slow-roll equations (8.39) and (8.40), is an *attractor* of the full equations, (8.33) and (8.36). This means that the solution of the full equations rapidly approaches it, if the initial conditions that are in the *basin of attraction*. To be in the basin of attraction means that φ must be in the slow-roll section; if $\dot{\varphi}$ is large, φ needs to be deep in the slow-roll section.

Once the system has reached the attractor, where (8.39) and (8.40) hold, $\dot{\varphi}$ is determined by φ. In fact everything is determined by φ (assuming a fixed potential $V(\varphi)$). The value of φ is the single parameter describing the state of the universe, and φ evolves down the potential $V(\varphi)$ as specified by the slow-roll equations.

The ideas of "attractor" and "basin of attraction" can be taken further. **If** the universe (or a region of it) finds itself initially (or enters) the basin of attraction of slow-roll inflation, meaning that: there is a sufficiently large region, where the curvature is sufficiently small, the inflaton makes a sufficient contribution to the total energy density, the inflaton is sufficiently homogeneous, and lies sufficiently deep in the slow-roll section, **then** this region begins inflating, it becomes rapidly very homogeneous and flat, all other contributions to the energy density besides the inflaton become negligible, and the inflaton begins to follow the slow-roll solution.

Thus inflation *erases all memory of initial conditions*, and we can predict the later history of the universe just from the shape of $V(\varphi)$ and the assumption that φ started out far enough in the slow-roll part of it. Note the similarity to thermal equilibrium. In the stages of the universe we discussed earlier, things were calculable because in thermal equilibrium, it is sufficient to know the temperature, masses of particles and conserved quantum numbers in order to have full information about the system. In the case of inflation, knowing the inflaton field value (and the shape of the potential) is enough, because of a rather different kind of attractor behaviour.

Example:

$$V(\varphi) = \frac{1}{2}m^2\varphi^2 \qquad \Rightarrow \qquad V'(\varphi) = m^2\varphi\,, \quad V''(\varphi) = m^2 \tag{8.44}$$

$$\left.\begin{aligned} \varepsilon(\varphi) &= \frac{1}{2}M_{\rm Pl}^2\left(\frac{2}{\varphi}\right)^2 \\ \eta(\varphi) &= M_{\rm Pl}^2\frac{2}{\varphi^2} \end{aligned}\right\} \quad \Rightarrow \quad \varepsilon = \eta = 2\left(\frac{M_{\rm Pl}}{\varphi}\right)^2 \tag{8.45}$$

and

$$\varepsilon, \eta \ll 1 \quad \Rightarrow \quad \varphi^2 \gg 2M_{\rm Pl}^2 \tag{8.46}$$

See figure 5.

8.4.1 Relation between inflation and slow-roll

$$H = \frac{\dot a}{a} \qquad \Rightarrow \qquad \dot H = \frac{\ddot a}{a} - \frac{\dot a^2}{a^2} \qquad \Rightarrow \qquad \frac{\ddot a}{a} = \dot H + H^2 \tag{8.47}$$

Thus the condition for inflation is $\dot H + H^2 > 0$. This would be satisfied if $\dot H > 0$, but this is not possible here, since it would require $p < -\rho$, i.e., $w \equiv p/\rho < -1$, which is not possible for a minimally coupled scalar field, see (8.29).[6] Thus we have $\dot H \leq 0$ and:

$$\text{inflation} \;\Leftrightarrow\; -\frac{\dot H}{H^2} < 1\;. \tag{8.48}$$

If the slow-roll approximation is valid,

$$H^2 = \frac{V}{3M_{\rm Pl}^2} \quad \Rightarrow \quad 2H\dot H = \frac{V'\dot\varphi}{3M_{\rm Pl}^2} \quad \Rightarrow \quad H^2\dot H \;=\; \frac{V'H\dot\varphi}{6M_{\rm Pl}^2} \overset{3H\dot\varphi=-V'}{=} -\frac{V'^2}{18M_{\rm Pl}^2}$$
$$\Rightarrow \quad -\frac{\dot H}{H^2} = \frac{V'^2}{18M_{\rm Pl}^2}\frac{9M_{\rm Pl}^4}{V^2} = \frac{1}{2}M_{\rm Pl}^2\left(\frac{V'}{V}\right)^2 = \varepsilon \ll 1\;.$$

So if the slow-roll approximation is valid, inflation is guaranteed. This result also shows that during slow-roll inflation, the Hubble parameter changes slowly (while the scale factor changes almost exponentially). As we have noted, slow-roll conditions are not necessary for inflation, it is possible to have inflation even when the slow-roll parameters are not small (called fast-roll inflation). However, when we consider perturbations in chapter 10, we will see that slow-roll inflation automatically produces a spectrum of perturbations that is in close agreement with observations, unlike fast-roll inflation.

[6] From the Friedmann equations,

$$\left.\begin{aligned} \left(\frac{\dot a}{a}\right)^2 &= \frac{8\pi G}{3}\rho - \frac{K}{a^2} \\ \frac{\ddot a}{a} &= -\frac{4\pi G}{3}(\rho + 3p) \end{aligned}\right\} \quad \Rightarrow \quad \dot H = \frac{\ddot a}{a} - \frac{\dot a^2}{a^2} = -4\pi G\left(\rho + p - \frac{K}{3a^2}\right)$$

Thus $\dot H > 0$ requires $\rho + p - \frac{K}{3a^2} < 0$. In the above, we assume that spatial curvature can already be neglected, i.e. we can take $K = 0$.

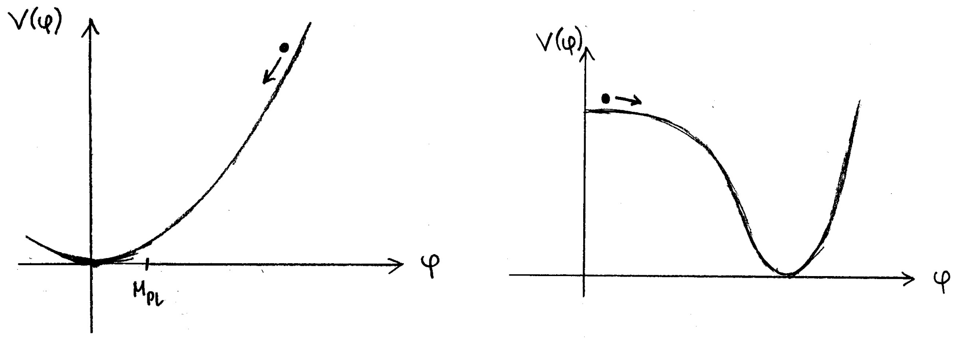

Figure 6: Potential for (a) large field and (b) small field inflation. For a typical small-field model, the entire range of φ shown is $\ll M_{\rm Pl}$.

8.5 Models of inflation

A scalar field model of inflation consists of the potential for the inflation and its couplings to other fields. In most models, couplings to other fields don't matter during inflation, and only the inflaton is dynamically important. However, these couplings usually come into play when inflation ends. Inflation can end because the slow-roll approximation is no longer valid, as the field has rolled down the potential. In this case inflation ends when either $\varepsilon(\varphi)$ or $|\eta(\varphi)|$ becomes of order unity. Another possibility is that inflation ends while the inflaton undergoes slow-roll, because other fields coupled to the inflaton become dynamically important and terminate inflation. An example of this is *hybrid inflation*, where there is an extra scalar field in addition to the inflaton. Inflation models can be divided into two classes:

1. Small field inflation: $\Delta\varphi < M_{\rm Pl}$ in the slow-roll section.

2. Large field inflation: $\Delta\varphi > M_{\rm Pl}$ in the slow-roll section.

Here $\Delta\varphi$ is the change of φ during (the observationally relevant part of) inflation.

Example: Consider a simple potential of the form $V(\varphi) = A\varphi^n$. This is a large field model, since $V'/V = n/\varphi \quad \Rightarrow \quad \varepsilon \ll 1$ requires $\varphi^2 \gg \frac{1}{2}n^2 M_{\rm Pl}^2$.

See figure 6 for typical shapes of potentials of large field and small field models.

8.5.1 An exact solution

Usually the slow-roll approximation is sufficient. In single-field models it fails near the end of inflation, but this is usually not a large correction. It is also much easier to solve the slow-roll equations, (8.39) and (8.40), than the full equations, (8.33) and (8.36). However, it is illustrative to consider and exact solution to the full equations. For some special cases, exact analytical solutions exist. One example is *power-law inflation*, where the potential is

$$V(\varphi) = V_0 \exp\left(-\sqrt{\frac{2}{p}}\frac{\varphi}{M_{\rm Pl}}\right), \qquad p > 1, \tag{8.49}$$

where V_0 and p are constants.

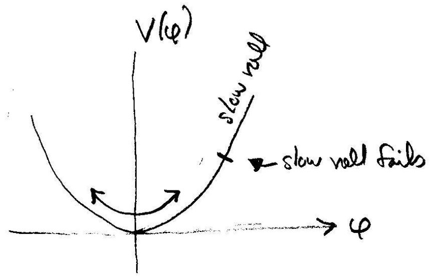

Figure 7: After inflation, the inflaton field is left oscillating at the bottom.

An exact solution for the full equations, (8.33) and (8.36), is

$$a(t) = a_0 t^p \tag{8.50}$$

$$\varphi(t) = \sqrt{2p} M_{\rm Pl} \ln \left(\sqrt{\frac{V_0}{p(3p-1)}} \frac{t}{M_{\rm Pl}} \right) . \tag{8.51}$$

The slow-roll parameters for this model are

$$\varepsilon = \frac{1}{2}\eta = \frac{1}{p} , \tag{8.52}$$

independent of φ. In this model inflation never ends unless other physics intervenes.

8.6 Reheating

During slow-roll inflation, practically all energy density in the universe is in the inflaton potential $V(\varphi)$. As inflation ends, this energy is transferred in the reheating process to a thermal bath of particles produced in the reheating. Thus reheating creates, from $V(\varphi)$, all the stuff there is in the later universe. The conversion of the inflaton energy density into a thermal gas of particles does not affect the spectrum of density perturbations in single field models of inflation (at least on superhorizon scales; see section 8.7 below). (It does change the relationship between the relation of φ_k and k/H_0 given in (8.62), i.e. the amount that the perturbations are stretched between the end of inflation and today.) The main constraint on reheating is that the reheating temperature must be above 1 MeV, but sufficiently low so as not to produce unwanted relics – where "sufficiently" depends on the theory under consideration. For typical supersymmetric theories the constraint on the reheating temperature is $T_{\rm reh} \lesssim 10^7$ GeV.

8.6.1 Scalar field oscillations

After inflation, the inflaton field φ begins to oscillate at the bottom of the potential $V(\varphi)$, see figure 7. The inflaton field is still homogeneous, $\varphi(t,\vec{x}) = \varphi(t)$, so it oscillates in the same phase everywhere (the oscillation is *coherent*). The oscillation period soon becomes much shorter than the expansion time scale H^{-1}.

Assume the potential can be approximated as $V(\varphi) = \frac{1}{2}m^2\varphi^2$ near the minimum of $V(\varphi)$, where the amplitude of φ is small. The equation of motion is then

$$\ddot{\varphi} + 3H\dot{\varphi} \;\; = -m^2\varphi \; . \tag{8.53}$$

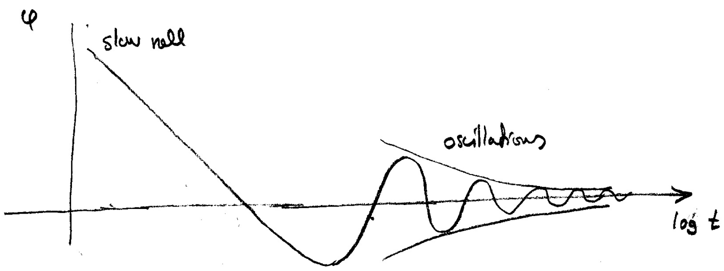

Figure 8: The time evolution of φ as inflation ends.

In the limit $m \gg H$, we can neglect the friction term, and the field undergoes oscillations with frequency m. We can write the energy continuity equation as

$$\dot{\rho} + 3H\rho \;\; = \;\; -3Hp = -\frac{3}{2}H\left(m^2\varphi^2 - \dot{\varphi}^2\right) \; . \tag{8.54}$$

The oscillating factor on the right hand side averages to zero over one oscillation period (in the limit where the period is $\ll H^{-1}$), so on average the energy density goes like $\rho \propto a^{-3}$, just like in a matter-dominated universe. The fall in the energy density shows as a decrease of the oscillation amplitude, see figure 8.

8.6.2 Inflaton decay

When the inflaton field is oscillating around the minimum of the potential, the energy stored in the inflaton field is transferred into particles, both by decay into quanta of the inflaton field, which subsequently decay, and direct decay into other fields via coupling between them and the inflaton. There can be tension between achieving efficient reheating and having a long period of inflation. To have a long duration of inflation, the inflaton field must be weakly coupled, but couplings to other degrees of freedom are required for reheating[7].

If the decay is slow, inflaton energy density satisfies the equation

$$\dot{\rho}_\varphi + 3H\rho_\varphi = -\Gamma_\varphi \rho_\varphi \, , \tag{8.55}$$

where $\Gamma_\varphi = 1/\tau_\varphi$, the *decay width*, is the inverse of the inflaton decay time τ_φ, and the term $-\Gamma_\varphi\rho_\varphi$ represents energy transfer to other particles.

If the inflaton can decay into bosons, the decay may be very rapid, involving a mechanism called *parametric resonance*. The produced particles are far from thermal equilibrium (only certain bands in momentum space become populated, and their occupation numbers are huge). In realistic models of inflation, the inflaton can decay via a mixture of different decay methods. The process by which the inflaton transfers its energy into particles is called *preheating* and the thermalisation of the gas of particles is called *reheating*. However, terminology varies, and often the term reheating is used to refer just to the energy transfer, even if the final state is not in thermal equilibrium.

[7] In fact, if the scale of inflation is sufficiently high, it is possible to reheat without any couplings between the inflaton and the Standard Model degrees of freedom by producing particles gravitationally out of the vacuum. This is called *gravitational reheating*, and it is one of the many delicacies of inflation we will not have time to sample!

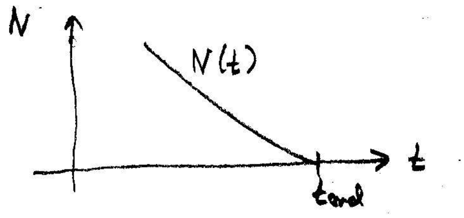

Figure 9: Remaining number of e-foldings $N(t)$ as a function of time.

8.6.3 Thermalisation

The particles produced from the inflaton will interact, create other particles through particle reactions, and the resulting soup will eventually reach thermal equilibrium with some temperature $T_{\rm reh}$. This *reheating temperature* is determined by the energy density $\rho_{\rm reh}$ at the end of the reheating epoch:

$$\rho_{\rm reh} = \frac{\pi^2}{30} g_*(T_{\rm reh}) T_{\rm reh}^4 \,. \tag{8.56}$$

Necessarily $\rho_{\rm reh} < \rho_{\rm end}$ (end = end of inflation). If reheating takes a long time, we may have $\rho_{\rm reh} \ll \rho_{\rm end}$. The evolution of the gas of particles into a thermal state can be quite involved, and it has been studied in various models. Usually it is just assumed that it happens eventually, since the particles are able to interact. However, it is possible that some particles (such as gravitinos) never reach thermal equilibrium, since their interactions are too weak. In any case, as long as the momenta of the particles are much higher than their masses, the energy density of the universe behaves like radiation, regardless of the momentum space distribution. So the background expansion rate is the same. After thermalisation of at least the baryons, photons and neutrinos is complete, the standard Hot Big Bang era begins.

8.7 Scales of inflation

8.7.1 Amount of inflation

During inflation, the scale factor $a(t)$ grows by a huge factor. We define the *number of e-foldings* from time t to end of inflation $t_{\rm end}$

$$N(t) \equiv \ln \frac{a(t_{\rm end})}{a(t)} \,. \tag{8.57}$$

See figure 9. We can calculate $N(t) \equiv N(\varphi(t)) \equiv N(\varphi)$ from the shape of the potential $V(\varphi)$ and the value of φ at time t:

$$N(\varphi) = \quad \ln \frac{a(t_{\rm end})}{a(t)} = \int_t^{t_{\rm end}} H(t) dt = \int_\varphi^{\varphi_{\rm end}} \frac{H}{\dot\varphi} d\varphi \overset{\text{slow roll}}{\approx} \boxed{\frac{1}{M_{\rm Pl}^2} \int_{\varphi_{\rm end}}^{\varphi} \frac{V}{V'} d\varphi} \,, \tag{8.58}$$

where we have used $\frac{da}{a} = d\ln a = H dt = H \frac{d\varphi}{\dot\varphi}$.

8.7.2 Evolution of scales

When discussing the evolution of density perturbations and formation of structures in the universe (to which we will get later), we will be interested in the history of each comoving distance scale, or each *comoving wave number* k (from Fourier expansion in comoving coordinates).

$$k = \frac{2\pi}{\lambda} , \qquad k^{-1} = \frac{\lambda}{2\pi}$$

An important question is whether a distance scale is larger or smaller than the Hubble length at a given time. A scale is said to be

- superhorizon, when $k < aH \quad \left(k^{-1} > (aH)^{-1}\right)$
- at horizon (exiting or entering horizon), when $k = aH$
- subhorizon, when $k > aH \quad \left(k^{-1} < (aH)^{-1}\right)$.

Note that *large* length scales (large k^{-1}) correspond to *small* k, and *vice versa*, although we often talk about "scale k". This can easily cause confusion, so be careful with wording! Notice also that we are here using the word "horizon" to refer to the Hubble length: more correct terminology would be "sub-Hubble" and "super-Hubble"[8]. Recall that $(aH)^{-1}$ shrinks during inflation, and grows during all other eras. See figures 10 and 11.

We shall later find that the amplitude of primordial density perturbations on a given comoving scale freezes as this scale exits the horizon during inflation. The largest observable scales are of the size of the horizon today. (Since the universe has recently began accelerating again, these scales have just barely entered, and are now exiting again.)

To identify the distance scales *during inflation* with the corresponding distance scales in the *present universe*, we need a complete history from inflation to the present. We divide it into the following periods:

1. **From** the time the scale k of interest exits the horizon during inflation **to** the end of inflation (t_k to t_{end}).
2. **From** the end of inflation **to** the time when thermal equilibrium at high temperature (Hot Big Bang conditions) is achieved, i.e. reheating. We assume that the universe behaves as if matter-dominated, $\rho \propto a^{-3}$, during this period, as discussed in section 8.6.1 (t_{end} to t_{reh}).
3. **From** reheating **to** matter-radiation equality (the radiation era, $\rho \propto a^{-4}$) (t_{reh} to t_{eq}).
4. The matter era, $\rho \propto a^{-3}$ **from** t_{eq} **to** t_0.

[8] As discussed in the first part of the course, there are (at least) three different usages for the word "horizon":

1. particle horizon
2. event horizon
3. Hubble length

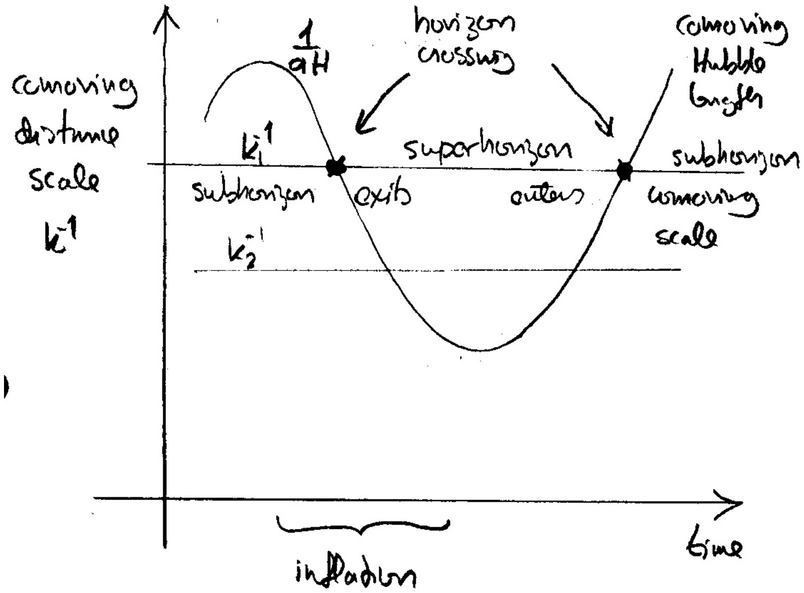

Figure 10: The evolution of the Hubble length, and two scales, k_1^{-1} and k_2^{-1}, seen in comoving coordinates.

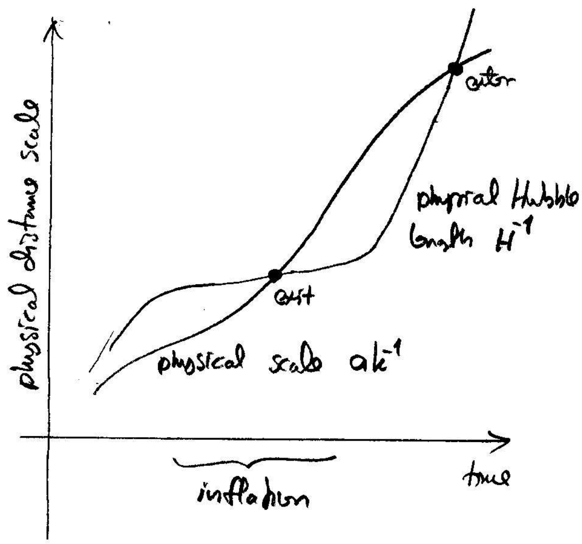

Figure 11: The evolution of the Hubble length, and the scale k^{-1} seen in terms of physical distance.

Consider a scale k that exits at $t = t_k$, when $a = a_k$ and $H = H_k$

$$\Rightarrow \quad k = (aH)_k = a_k H_k \,.$$

To find out how large this scale is today, we relate it to the present "horizon", i.e., the Hubble scale (for clarity, we insert a_0 here, even though we have chosen it to be equal to unity):

$$\begin{aligned} \frac{k}{a_0 H_0} &= \frac{a_k H_k}{a_0 H_0} \\ &= \frac{a_k}{a_{\text{end}}} \frac{a_{\text{end}}}{a_{\text{reh}}} \frac{a_{\text{reh}}}{a_0} \frac{H_k}{H_0} \\ &= e^{-N(k)} \left(\frac{\rho_{\text{end}}}{\rho_{\text{reh}}}\right)^{-\frac{1}{3}} \left(\frac{\rho_{\text{reh}}}{\rho_{r0}}\right)^{-\frac{1}{4}} \left(\frac{V_k}{\rho_{c0}}\right)^{\frac{1}{2}} , \end{aligned} \tag{8.59}$$

where we have used the relation $\frac{a_k}{a_{\text{end}}} = e^{-N(k)}$, where $N(k)$ is the number of e-foldings until the end of inflation after the scale k exits the horizon. We have also taken into account that from the end of inflation until reheating we have approximately $\rho \propto a^{-3}$ and that from reheating until today the radiation component evolves like $\rho \propto a^{-4}$. This is slightly inaccurate, since $\rho_r \propto a^{-4}$ does not take into account the change in g_*. However, the approximation is good enough for us here, as the error will only enter logarithmically in the number of e-foldings[9]. (We assume that almost all energy density goes into particles with masses smaller than the reheating temperature.) Finally, we have used $H \propto \sqrt{\rho}$, which follows from the Friedmann equation, and noted that during slow-roll inflation $\rho_k \approx V_k$, where the subscript k again refers to the time when the mode with wavenumber k exits the horizon.

We can rewrite (8.59) as

$$\frac{k}{a_0 H_0} = e^{-N(k)} \frac{10^{16}\ \text{GeV} \rho_{r0}^{\frac{1}{4}}}{\rho_{c0}^{\frac{1}{2}}} \frac{V_k^{\frac{1}{4}}}{10^{16}\ \text{GeV}} \left(\frac{V_k}{V_{\text{end}}}\right)^{\frac{1}{4}} \left(\frac{\rho_{\text{reh}}}{V_{\text{end}}}\right)^{\frac{1}{12}} , \tag{8.61}$$

where we have inserted the comparison scale 10^{16} GeV, taken into account that $\rho_{\text{end}} \approx V_{\text{end}}$ (if inflation ends due to the slow-roll approximation being violated, this

[9]Accurately this would go as:

$$g_{*s} a^3 T^3 = \text{const.} \quad \Rightarrow \quad \frac{a_{\text{reh}}}{a_0} = \left[\frac{g_{*s}(T_0)}{g_{*s}(T_{\text{reh}})}\right]^{\frac{1}{3}} \frac{T_0}{T_{\text{reh}}} \,. \tag{8.60}$$

We approximated this with

$$\left(\frac{\rho_{r0}}{\rho_{\text{reh}}}\right)^{\frac{1}{4}} = \left[\frac{g_*(T_0)}{g_*(T_{\text{reh}})}\right]^{\frac{1}{4}} \frac{T_0}{T_{\text{reh}}}$$

Taking $g_{*s}(T_{\text{reh}}) = g_*(T_{\text{reh}}) \sim 100$, the ratio of these two becomes

$$\frac{g_{*s}(T_0)^{\frac{1}{3}}}{g_*(T_0)^{\frac{1}{4}} g_*(T_{\text{reh}})^{\frac{1}{12}}} \approx \frac{3.909^{\frac{1}{3}}}{3.363^{\frac{1}{4}} 100^{\frac{1}{12}}} = 0.79 \sim 1 \,.$$

Note that $a \propto \rho_r^{-1/4}$ is a better approximation than $a \propto T^{-1}$, since these two differ by

$$\left[\frac{g_*(T_{\text{reh}})}{g_*(T_0)}\right]^{\frac{1}{4}} \sim \left(\frac{100}{3.363}\right)^{\frac{1}{4}} \sim 2.33 \,.$$

will only be true up to factors of order unity, which we neglect) and rearranged some of the terms. We don't know the energy scale of inflation, but there is an upper limit of approximately 10^{16} GeV from the lack of observation of primordial gravity waves, whose amplitude provides a measure of the inflationary energy scale. Inserting the values $\rho_{r0} = 4.18 \times 10^{-5} h^{-2} \rho_{c0}$ (assuming massless neutrinos) and $\rho_{c0}^{1/4} = (\sqrt{3} H_0 M_{\rm Pl})^{1/2} \approx 3.0 \times 10^{-12} h^{1/2}$ GeV, and taking $h = 0.7$, we obtain for the number of e-folds

$$N(\varphi_k) = -\ln \frac{k}{a_0 H_0} + 61 - \frac{1}{3} \ln \frac{V_{\rm end}^{1/4}}{\rho_{\rm reh}^{1/4}} + \ln \frac{V_k^{1/4}}{V_{\rm end}^{1/4}} - \ln \frac{10^{16}\ \mathrm{GeV}}{V_k^{1/4}} \ , \tag{8.62}$$

where $\varphi_k \equiv \varphi(t_k)$. The terms have been arranged such that the quantities in the logarithms are bigger than unity. The second term depends on the efficiency of reheating: if all of the inflaton potential energy is converted into radiation degrees of freedom instantaneously, it is zero. The third term is expected to be small, since the potential varies slowly during slow-roll: the dependence on k in the first term is expected to dominate. The last factor can however be large if the inflation scale is much lower than 10^{16} GeV. For example, inflation at the TeV scale would give -30.

For any given present scale, given as a fraction of the present Hubble distance[10], (8.62) identifies the value φ_k the inflaton had, when this scale exited the horizon during inflation. The last three terms give the dependence on the energy scales connected with inflation and reheating. In typical inflation models, they are relatively small. Usually, the precise value of N is not that important; we are more interested in the derivative dN/dk, or rather $d\varphi_k/dk$.

Anyway, we see that typically (for high scale inflation) about 60 e-foldings of inflation occur *after* the largest observable scales exit the horizon. There is no similar constraint on the number of e-folds before these scales exited the horizon, and the number varies from a few to 10^8 (or more) between different models.

Exercise: Assume that at the beginning of inflation we have $|\Omega_K| = 0.1$. a) Calculate, as a function of the reheating temperature $T_{\rm reh}$, how many e-folds of inflation are required to reduce present-day spatial curvature to $|\Omega_{K0}| < 10^{-2}$. (Assume $h = 0.7$ and that neutrinos are massless.) Approximate that the expansion rate at the beginning of inflation is completely dominated by the inflaton, that the inflaton field value does not change during inflation and that reheating happens instantaneously. b) In which directions do the above approximations change the result? c) What is the number of e-folds for $T_{\rm reh} = 10^7$ GeV?

8.8 Before inflation

As we discussed earlier, inflation erases all memory of the initial conditions before inflation, and on the theoretical side we do not have a good theoretical understanding of what happened in that era. However, there are some ideas. During inflation, the universe is expanding and (in most models) the energy density is decreasing. We thus expect that the energy density is higher before inflation than during it or after it. Often it is assumed that inflation begins right at the Planck scale, $\rho \sim M_{\rm Pl}^4$, which is the limit to how high energy densities we can extend our discussion, which is based

[10]For example, $k/H_0 = 10$ means that we are talking about a a scale corresponding to a wavelength λ such that $\lambda/2\pi$ is one tenth of the Hubble distance.

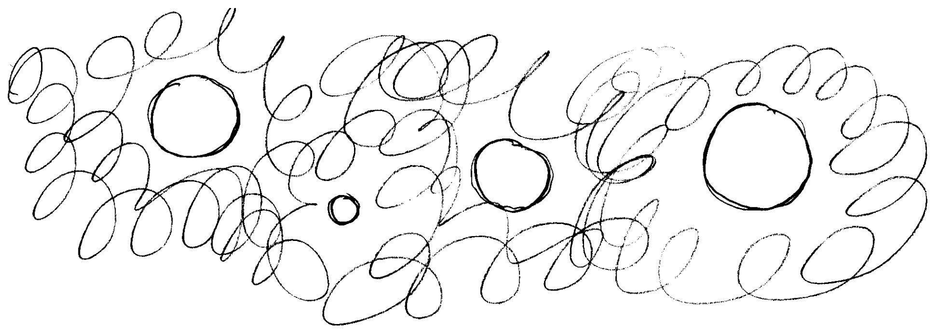

Figure 12: Classical regions emerging from spacetime foam.

on classical general relativity, and quantum gravitational effects are expected to be important. One idea is that the universe at that time, the *Planck era*, is some kind of a "spacetime foam", where the fabric of spacetime itself is subject to large quantum fluctuations. When the energy density of some region, larger than H^{-1}, falls below M_{Pl}^4, spacetime in that region begins to behave in a classical manner. See figure 12.

The initial conditions, i.e. conditions at the time when our Universe emerges from the spacetime foam, are usually assumed *chaotic* (this word does not refer to chaos theory!), i.e. φ takes random values at different regions. Since $\rho \geq \rho_\varphi$, and

$$\rho_\varphi = \frac{1}{2}\dot{\varphi}^2 + \frac{1}{2}\frac{1}{a^2}\partial_i\varphi\partial_i\varphi + V\,, \tag{8.63}$$

we must have

$$\dot{\varphi}^2 \lesssim M_{\mathrm{Pl}}^4\,, \quad \frac{1}{a^2}\partial_i\varphi\partial_i\varphi \lesssim M_{\mathrm{Pl}}^4\,, \quad V \lesssim M_{\mathrm{Pl}}^4 \tag{8.64}$$

in a region for it to emerge from the spacetime foam.

Inflation may begin at many different parts of the spacetime foam. Our observable universe is just one small part of one such region which has inflated to a huge size.

It is also possible that during inflation, for some part of the potential, quantum fluctuations of the inflaton (not the spacetime!) dominate over the classical evolution and push φ higher in some regions. These regions will then expand faster, and dominate the volume. This gives rise to *eternal inflation*, where, at any given time, most of the volume of the universe is inflating. (Whether or not this can happen depends on the shape of the potential and the field value during inflation.) But our observable Universe is part of a region where φ rolled down and came to a region of the potential, where the quantum fluctuations of φ were small and the classical behaviour began to dominate and inflation ended.

Thus the ultra-large scale structure of the universe may be very complicated. However, this is not observable to us, and all the features of the universe we see can be explained in terms of what happened in our patch during and after inflation. These ideas of the spacetime foam and eternal inflation are rather speculative, and there are also different suggestions for the initial stages of the universe.