Week 5: Inflation basics

February 24, 2017

1 Motivation

The standard hot big-bang model with an (flat) FRW spacetime accounts correctly for the observed expansion, the CMB, BBN, etc. However, it leaves a number of questions unanswered.

The flatness problem. First of all, there is the flatness problem. I.e., why is our Universe so close to flat? One possibility is that it simply began as flat; i.e., with zero curvature. In some sense, this is strange because even if we started with a precisely flat Universe, then if there were even tiny density perturbation in the initial state, then an observable region of the Universe would have locally a density different than the critical density. Thus, imagine there was some initial nonzero curvature, $k \neq 0$ and ignore the cosmological constant since it remains dynamically negligible until a redshift $z \sim 1$. Then, at redshifts $z \gtrsim 1$, Ω_m is extremely close to 1. The Friedmann equation can be rearranged to give

$$\Omega_m - 1 = \frac{k}{a^2 H^2}.$$

During matter (radiation) domination, $a \propto t^{2/3}$ ($a \propto t^{1/2}$), and $H \propto 1/t$ at all times, so if Ω_m is not precisely equal to 1, then it diverges from 1 with the expansion of the Universe. For $\Omega_m \simeq 0.3$ today, the matter density at BBN must have been $|\Omega_m(t_{\rm bbn}) - 1| \lesssim 10^{-16}$, and at the time of the quantum-gravity event that presumably gave rise to the FRW Universe, $|\Omega_m(t_{\rm Pl}) - 1| \lesssim 10^{-60}$. In other words, the Universe would have had to be *extremely* close to flat in the initial state, or put another way, could have tolerated no more than the very tiniest density fluctuations, no more than 1 part in 10^{60}. Put another way, if the Universe were born at the Planck time with equal energy density in the curvature and matter degrees of freedom, it would have survived no longer than a Planck time; the problem is sometimes then phrased as "why is the Universe so old?" This is also sometimes referred to as the "Dicke coincidence," although it was noted presumably much earlier by Einstein, who therefore concluded that the Universe must be precisely flat.

Horizon problem. We know that CMB photons last scattered at a redshift $z_{lss} \simeq 1100$ when the Universe was $t_{\rm lss} \simeq 380,000$ years old, and that today it is $t_0 \simeq 13.8$ billion years old and very close to flat. We can thus infer that a causally connected region at the surface of last scatter subtends an angle $\theta \simeq (1 + z_{lss})(t_{\rm lss}/t_0) \sim 1^\circ$. However, there are 4π steradians$\simeq 40,000$ square degrees on the sky. We are therefore looking at roughly 40,000 causally disconnected patches of the early Universe when we look at the CMB. Yet each has a temperature that is the same to one part in 10^5. How did these causally disconnected regions of the early Universe know to have the same temperature? This is the horizon, causality, or smoothness problem. A related problem is that the Universe must have also been very smooth on very small scales. The horizon at the time of BBN enclosed roughly a solar mass of material, more than 20 orders of magnitude less mass than the horizon encloses today. The predicted light-element abundances are nonlinear functions of the baryon density. If there were density fluctuations of order unity on solar-mass scales at the time of BBN, then the observed light-element abundances would be different from those that are observed and that are observed to be in good agreement with the predictions. We therefore know that the Universe must have been smooth on small scales as well as large.

Monopole problem. Grand unified theories predict the existence of magnetic monopoles, topological defects with masses $\sim M_{\rm GUT} \sim 10^{15}$ GeV. According to the Kibble mechanism, roughly one such monopole is produced in every Hubble volume at the GUT phase transition near $T \sim 10^{15}$ GeV. You will show in a homework problem that this would result in a monopole density many orders of magnitude greater than the critical density.

Acausal primordial perturbations. We see in the CMB primordial density fluctuations $(\delta\rho/\bar{\rho}) \sim 10^{-5}$ with a nearly scale-invariant spectrum. We also see that they are seemingly acausal; i.e., there are Fourier modes of the perturbations that have wavelengths larger than the horizon size at the surface of last scatter. Where did these come from?

2 Homogeneous evolution

2.1 Kinematics

An expanding isotropic and homogeneous Universe is described by a Friedmann-Robertson-Walker (FRW) spacetime, with line element $ds^2 = -dt^2 + a^2(t)d\boldsymbol{x}^2$, in terms of a scale factor $a(t)$ that parametrizes the physical distance that corresponds to a given comoving distance. As the Universe expands [i.e., the scale factor $a(t)$ increases with time t], the Hubble length H^{-1}, where $H \equiv \dot{a}/a$ is the Hubble or growth rate, increases. During radiation and matter domination, $(d/dt)(aH)^{-1} > 0$, and so the Hubble distance H^{-1} increases more rapidly than the scale factor. As a result, with time, an observer sees larger comoving volumes of the Universe, and objects and information enter the horizon. This observation leads to the horizon problem: if the Universe began with a period of radiation domination, then how did the $\sim 40,000$ causally disconnected patches of CMB sky know to have the same temperature (to roughly one part in 100,000)?

If, however, $(d/dt)(aH)^{-1} < 0$, then an observer sees with time a smaller comoving patch (even though the physical or proper size of the observable patch may still be increasing), and objects/information/perturbations exit the horizon. In this way, the Universe becomes increasingly smooth, thus explaining the remarkable large-scale homogeneity of the Universe.

The requirement $(d/dt)(aH)^{-1} = \left[(\dot{H}/H^2) + 1\right]/a < 0$ implies that we must have $\epsilon \equiv -\dot{H}/H^2 < 1$ for inflation. Most generally, $\dot{H} \neq 0$ (so that inflation can end, if for no other reason). As we will see, however, theory and measurement suggest $\epsilon \ll 1$, implying that the scale factor grows almost exponentially, $a(t) \propto e^{Ht}$, during inflation.

If we assume the validity of general relativity, as we do here then the time evolution of the scale factor satisfies the Friedmann equations, $H^2 = \rho/(3{M_{\rm Pl}}^2)$ and $\dot{H} + H^2 = -(\rho + 3p)/(6{M_{\rm Pl}}^2)$, where p and ρ are the pressure and energy density of the cosmic fluid, respectively. We work in particle-physics units, with $\hbar = c = 1$ and have written Newton's constant $G = (8\pi {M_{\rm Pl}}^2)^{-1}$ in terms of the reduced Planck mass, $M_{\rm Pl} = 2.435 \times 10^{18}$ GeV. These two Friedmann equations imply that

$$\epsilon = (3/2)\,(1 + p/\rho)\,, \tag{1}$$

from which we infer that the equation-of-state parameter $w \equiv p/\rho$ must be $w < -1/3$ in order for inflation to occur.

2.2 Scalar-Field Dynamics

In the simplest paradigm for inflation, and that on which we focus, this exotic equation of state is provided by the displacement of a scalar field ϕ, the "inflaton," from the minimum of its potential $V(\phi)$. The homogeneous time evolution of the scalar field then satisfies, in an FRW spacetime, the equation of motion, $\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0$, where the dot denotes a derivative with respect to time and prime a derivative with respect to ϕ. We thus see that the expansion acts as a friction term. The scalar field has energy density $\rho = (1/2)\dot{\phi}^2 + V(\phi)$ (a kinetic-energy density and a potential-energy density) and pressure $p = (1/2)\dot{\phi}^2 - V(\phi)$. If $V(\phi)$ is nonzero and sufficiently flat and the friction term in the ϕ equation sufficiently large, then the kinetic-energy density will be $(1/2)\dot{\phi}^2 < 2V(\phi)$, in which case $p < -\rho/3$ and inflation ensues (see **Figure 1**).

This condition is made more precise by solving the scalar-field equation of motion along with the Friedmann equation, $H^2 = (\dot{a}/a)^2 = \left[V(\phi) + (1/2)\dot{\phi}^2\right]/(3m_{\rm Pl}^2)$. During inflation ϕ varies monotonically with time t and can thus be used as the independent variable (rather than t). Let us suppose that the field and potential are defined so that $\dot{\phi} > 0$ during inflation. We then differentiate the Friedmann equation with respect to time, obtaining $2H\dot{H} = \dot{\phi}\left[V'(\phi) + \ddot{\phi}\right]/(2m_{\rm Pl}^2)$. Then rearranging the scalar-field equation of motion, $-3H\dot{\phi} = \ddot{\phi} + V'(\phi)$, we get $\dot{H} = \dot{\phi}^2/(2m_{\rm Pl}^2)$. We thus infer that

$$\epsilon = 3\frac{\dot{\phi}^2/2}{V + \dot{\phi}^2/2} \simeq \frac{{M_{\rm Pl}}^2}{2}\left(\frac{V'}{V}\right)^2, \tag{2}$$

where the last expression is the result of the slow-roll approximation, $\epsilon \ll 1$, in which $\dot{\phi}^2/2 \ll V$. Note that in much of the literature, ϵ is *defined* in terms of V and V' through this relation, rather than through $\epsilon = -\dot{H}/H^2$, as is done here, a distinction whose subtlety will be unimportant in this article, although it can be important for quantitative conclusions given the precision of current measurements. We also define a second slow-roll parameter,

$$\eta = -2\frac{\dot{H}}{H^2} - \frac{\dot{\epsilon}}{2H\epsilon} \simeq {M_{\rm Pl}}^2\frac{V''}{V}, \tag{3}$$

which will become important below; the approximation in Equation 3 is valid during slow-roll, when $\eta \ll 1$.

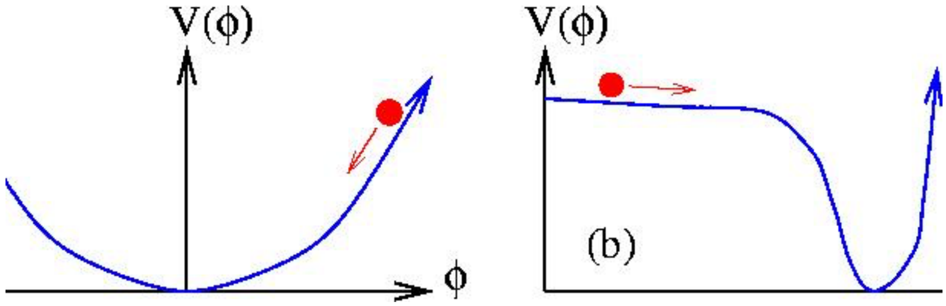

Figure 1: Inflation postulates that at some point in the early history of the Universe, the cosmic energy density was dominated by the vacuum energy associated with the displacement of some scalar field ϕ (the inflaton) from the minimum of its potential. Shown here for illustration are two toy models for the inflaton potential: on the left, a quadratic potential and on the right, a hilltop potential

2.3 Duration of inflation and evolution of scales

The number of e-folds of inflation between the end of inflation and a time t during inflation is

$$N(t) \equiv \ln \frac{a(t_{\rm end})}{a(t)} = \int_t^{t_{\rm end}} H\, dt = -\frac{1}{2m_{\rm Pl}^2} \int_{\phi_t}^{\phi_{\rm end}} \frac{H}{H'}\, d\phi = \int_{\phi_{\rm end}}^{\phi_t} \frac{d\phi}{M_{\rm Pl}} \frac{1}{\sqrt{2\epsilon(\phi)}}. \tag{4}$$

The largest comoving scales exit the horizon first during inflation, and they are the last to re-enter the horizon later during matter or radiation domination. To evaluate the number of e-foldings required to solve the horizon problem, consider a physical wavenumber $k_{\rm phys}$. Its ratio to the Hubble scale today is

$$\frac{k_{\rm phys}}{a_0 H_0} = \frac{a_k H_k}{a_0 H_0} = \frac{a_k}{a_{\rm end}} \frac{a_{\rm end}}{a_{\rm reh}} \frac{a_{\rm reh}}{a_{\rm eq}} \frac{a_{\rm eq}}{a_0} \frac{H_k}{H_0}, \tag{5}$$

where a_k and H_k are the scale factor and Hubble parameter when this particular wavenumber exits the horizon; $a_{\rm end}$ is the scale factor at the end of inflation; $a_{\rm eq}$ is the scale factor at matter-radiation equality; and $a_{\rm eh}$ is the scale factor at the time of reheating. Plugging in numbers, we find that the number of e-foldings between the end of inflation and the time at which the wavenumber k exits the horizon is

$$N(k) = 62 - \ln \frac{k_{\rm phys}}{a_0 H_0} - \ln \frac{10^{16}\,{\rm GeV}}{V_k^{1/4}} + \ln \frac{V_k^{1/4}}{V_{\rm end}} - \frac{1}{3} \ln \frac{V_{\rm end}^{1/4}}{\rho_{\rm reh}^{1/4}}, \tag{6}$$

where $\rho_{\rm reh}^{1/4}$ is the energy density at reheating. If the energy scale of inflation is near the current upper limit $V^{1/4} \lesssim 10^{16}$ GeV (see below), but higher than the energy scale of electroweak symmetry breaking ($V_k \gtrsim 10^3$ GeV), then the number N of e-folds between the time that the largest observable scales today exited the horizon and the end of inflation falls in the range $30 \lesssim N \lesssim 60$. Recent treatments that consider different families of inflationary potentials, include current constraints to the scalar spectral index n_s (see below), as well as plausible reheating scenarios, find a range $40 \lesssim N \lesssim 60$. More conservatively, the near scale-invariance of primordial density perturbations over the ~ 3 orders of magnitude over which they have been measured tells us that $N \gtrsim 10$ at the very least.

3 Exact Solutions

There are a few models for which the scalar-field and scale-factor EOMs can be solved exactly. The first example is power-law inflation, which features a potential,

$$V(\phi) = V_0 \exp\left(-\sqrt{\frac{2}{p}}\frac{\phi}{m_{\rm Pl}}\right).$$

This model has a scale factor $a(t) = a_0 t^p$ and the scalar field moves with time t according to

$$\frac{\phi}{m_{\rm Pl}} = \sqrt{2p}\ln\left(\sqrt{\frac{V_0}{p(3p-1)}}\frac{t}{m_{\rm Pl}}\right).$$

There is inflation for $p > 1$, and the slow-roll parameters are $\epsilon = \eta/2 = 1/p$, and $w = 2/3p$. In this model, there is no end to inflation.

Intermediate inflation has a scale factor $a(t) \propto \exp[At^f]$ with $0 < f < 1$ and $A > 0$. This is obtained from rolling down a potential,

$$V(\phi) \propto \left(\frac{\phi}{m_{\rm Pl}}\right)^{-\beta}\left(1 - \frac{\beta^2}{6}\frac{m_{\rm Pl}^2}{\phi^2}\right),$$

where $\beta = 4(f^{-1} - 1)$.

4 Hamilton-Jacobi formulation

The idea here is to use the fact that ϕ changes monotonically during inflation to replace the time variable as the independent variable in the equations of motion by ϕ. This then allows us to combine the scalar-field and scale-factor equations of motion. Take $\dot{\phi} > 0$. Then differentiate the Friedmann equation with respect to time:

$$\frac{d}{dt}\left[H^2 = \frac{8\pi}{3m_{\rm Pl}^2}\left(V(\phi) + \frac{1}{2}\dot{\phi}^2\right)\right].$$

This yields

$$2H\dot{H} = \frac{8\pi}{3m_{\rm Pl}^2}\left(V'(\phi)\dot{\phi} + \dot{\phi}\ddot{\phi}\right).$$

Then rearranging the scalar-field EOM, $-3H\dot{\phi} = \ddot{\phi} + V'(\phi)$, we get $\dot{H} = 4\pi\dot{\phi}^2/m_{\rm Pl}^2$, but since $d/dt = \dot{\phi}(d/d\phi)$, we also have $\dot{\phi} = H'(\phi)$. We then plug this back into the Friedmann equation to get the Hamilton-Jacobi equation,

$$\left[H'(\phi)\right]^2 - \frac{12\pi}{m_{\rm Pl}^2}H^2(\phi) = -\frac{4\pi}{m_{\rm Pl}^4}V(\phi).$$

This is an *exact* equation that is equivalent to the combined scalar-field and scale-factor EOMs. The advantage of this formalism is that it allows one to generate exact solutions. So, for example, if I take $H(\phi) \propto \phi^{-\beta/2}$, I get the potential for intermediate inflation, and if I take $H(\phi) \propto \exp\left[-\sqrt{1/2p}(\phi/m_{\rm Pl})\right]$, then I recover power-law inflation.

The other advantage is that it also makes the slow-roll conditions for inflation more precise. Define

$$\epsilon_H = 2m_{\rm Pl}^2 \left(\frac{H'}{H}\right)^2,$$

and

$$\eta_H = \frac{2m_{\rm Pl}^2}{8\pi}\frac{H''}{H}.$$

Then

$$\epsilon_H = 3\frac{\dot{\phi}^2/2}{V + \dot{\phi}^2/2} = -\frac{d\ln H}{d\ln a},$$

and

$$\eta_H = -3\frac{\ddot{\phi}}{3H\dot{\phi}} = -\frac{d\ln\dot{\phi}}{d\ln a} = -\frac{d\ln H'}{d\ln a}.$$

We now see that $\ddot{a} > 0$ corresponds to $\epsilon_H < 1$ *exactly.* Note also that with this formalism, the number of e-foldings of inflation is

$$N \equiv \ln\frac{a(t_{\rm end})}{a(t)} = \int_t^{t_{\rm end}} H\,dt = -\frac{1}{2m_{\rm Pl}^2}\int_\phi^{\phi_{\rm end}} \frac{H}{H'}\,d\phi.$$

Inflation as attractor. The HJ formalism also allows us to justify earlier statements that inflation produces the requirements for inflation. Suppoose $H_0(\phi)$ is a solution to the HJ equation of motion,

$$\left[H'(\phi)\right]^2 - \frac{12\pi}{m_{\rm Pl}^2}H^2(\phi) = -\frac{4\pi}{m_{\rm Pl}^4}V(\phi).$$

Now consider another solution, $H(\phi) = H_0(\phi) + \delta H(\phi)$. Then

$$H_0\delta H' \simeq \frac{2\pi}{m_{\rm Pl}^2}H_0\delta H,$$

or

$$\delta H(\phi) = \delta H(\phi_i)\exp\left(\frac{12\pi}{m_{\rm Pl}^2}\int_{\phi_i}^{\phi}\frac{H_0(\phi)}{H_0'(\phi)}\,d\phi\right).$$

The quantity in the exponent is negative, and this therefore implies that all perturbations die away during inflation. In particular, during inflation, $\epsilon_H < 1$, so

$$\delta H(\phi) < \delta H(\phi_i)\exp\left[-\frac{3}{\sqrt{2}}\frac{\phi-\phi_i}{m_{\rm Pl}}\right] = \delta H(\phi_0)\exp\left[-3(N_i - N)\right].$$