Lectures 4,5: Quantum Generation of Gravitational Waves

1 The geometrical essence of inflation (and the bouncing alternatives)

The quantity $1/|aH|$ is the "**comoving Hubble radius**," *i.e.* the radius of the "**comoving Hubble sphere**" or "**comoving Hubble volume**". This quantity grows during decelerated expansion ($\dot{a} > 0$, $\ddot{a} < 0$, as in the ordinary radiation or matter dominated epochs of the hot Big Bang model), but shrinks during a period of accelerated expansion ($\dot{a} > 0$, $\ddot{a} > 0$, as in inflation) or a period of decelerated contraction ($\dot{a} < 0$, $\ddot{a} < 0$, as in bouncing models that include a contracting phase prior to the Big Bang). In both inflationary and bouncing models, the idea is that, if the factor by which $1/|aH|$ initially shrinks during the inflationary or contracting phase is greater than the factor by which it subsequently grows during the ordinary decelerated (matter or radiation dominated hot Big Bang) phase, then a number of cosmological puzzles and observations may be explained in a unified way.

2 Single scalar field inflation

What dominated the energy density during the hypothetical inflationary era? We don't know, but for concreteness we will here consider the simplest and most successful possibility: a single scalar field φ with potential $V(\varphi)$, so that the action is

$$S = \int d^4x\sqrt{-g}\left[\frac{R}{16\pi G_N} - \frac{1}{2}g^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi - V(\varphi)\right]. \qquad (1)$$

It will be convenient to define the "**slow-roll parameters**"

$$\epsilon = -\frac{\dot{H}}{H^2} \approx \frac{1}{2}M_{pl}^2\left[\frac{V'(\varphi)}{V(\varphi)}\right]^2 \qquad (2a)$$

$$\eta = M_{pl}^2\frac{V''(\varphi)}{V(\varphi)} \qquad (2b)$$

where M_{pl} is the so-called "reduced Planck mass" ($8\pi G_N = M_{pl}^{-2}$). We will assume that the potential $V(\varphi)$ is such that the field can "roll" monotonically down its potential from a "slow-roll region" (with ϵ and η both $\ll 1$), to a local minimum at $\varphi = \varphi_{min}$ with vanishing potential $V(\varphi_{min}) = 0$. This is the framework for "**single-field slow-roll inflation**." Note that the $V(\varphi) = (1/2)m^2\varphi^2$ is an example of a potential satisfying these requirements: when $|\varphi| > m_{pl}$, we have $\epsilon \ll 1$ and $\eta \ll 1$.

2.1 Homogeneous evolution; slow-roll approximation

A homogeneous scalar field φ has energy density ρ and pressure p given by

$$\rho = \frac{1}{2}\dot{\varphi}^2 + V(\varphi) \tag{3a}$$

$$p = \frac{1}{2}\dot{\varphi}^2 - V(\varphi) \tag{3b}$$

Substituting these results into the continuity equation $\dot{\rho} = -3H(\rho + p)$ leads to the equation of motion

$$\ddot{\varphi} + 3H\dot{\varphi} + V'(\varphi) = 0. \tag{4}$$

This is like the equation of motion for a particle experiencing two different forces: a potential force (the V' term) and a friction force (the $3H\dot{\varphi}$ term). The field starts in the slow-roll regime: it is strongly overdamped and quickly relaxes to its "terminal velocity," where its acceleration $\ddot{\varphi}$ is negligible in Eq. (4) and the drag approximately balances the potential forcing

$$\dot{\varphi} \approx -V'(\varphi)/3H. \tag{5}$$

This means that the condition $\epsilon \ll 1$ becomes $\dot{\varphi}^2 \ll V(\varphi)$, which implies $w \approx -1$ and hence $\ddot{a} > 0$. Eventually, when the field gets close enough to the minimum so that $V(\varphi) \approx (1/2)m^2\varphi^2 + \ldots$ and $H < m$, the field begins underdamped oscillations $\varphi(t) \propto a^{-3/2}\cos(mt)$, and with energy density decaying as $\rho \propto a^{-3}$, which implies $w = 0$ and hence $\ddot{a} < 0$. So the field starts in its slow-roll regime, gradually oozing down its potential as the universe accelerates; finally, close enough to the minimum, the slow-roll conditions cease to hold and the φ begins underdamped oscillations about its minimum, as the universe stops accelerating and begins to decelerate.

2.2 Perturbations

In electromagnetism, the gauge field A_μ naively has 4 components, but closer inspection reveals that it only describes two physical degrees of freedom (the 2 polarizations of an electromagnetic wave). Similarly, in pure gravity (with no additional matter fields) if we consider small perturbation about Minkowski space, $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, we have seen that, although the symmetric 4×4 matrix $h_{\mu\nu}$ has 10 components, it only contains $10 - 4 - 4 = 2$ independent physical degrees of freedom. Now, when we add a single scalar field to the picture, we have one additional degree of freedom. In the ADM language, we can write

$$ds^2 = -N^2dt^2 + \gamma_{ij}(dx^i + N^i dt)(dx^j + N^j dt) \tag{6}$$

and choose a gauge in which

$$\delta\varphi = 0, \qquad \gamma_{ij} = a^2(t)e^{2\zeta(t,\mathbf{x})}e^{2h_{ij}(t,\mathbf{x})} \tag{7}$$

where h_{ij} is in TT gauge $\delta^{ij}h_{ij} = \delta^{ij}\partial_i h_{jk} = 0$), and then constraint equations determine the rest of the metric, $N(t,\vec{x})$ and $N^i(t,\mathbf{x})$. The 3 physical modes are ζ (which, in this $\delta\varphi = 0$ gauge, physically corresponds to the perturbation of the Ricci 3-curvature of the spatial slices) and the two independent components of a traceless transverse 3×3 symmetric matrix h_{ij} (which physically corresponds to the two polarizations of a gravitational wave). You will hear ζ referred to as a "primordial scalar perturbation" or "primordial curvature perturbation;" it is observed more-or-less directly by observing the fluctuation in the temperature of the CMB between different points on the sky; it is the primordial perturbation that is ultimately responsible for all of the density perturbations and structure (*e.g.* galaxies, clusters, voids, stars, you, me,...) that we see in the universe today. And you will hear the 2 independent components of h_{ij} referred to as "primordial tensor perturbations;" they are primordial gravitational waves predicted by inflation, but at a level difficult to observe; it is hoped that they will be observed in the coming years through observations of the polarization pattern of the CMB – if and when that happens, it could be the observational result that clinches the case for inflation, and converts it from a possible explanation to an overwhelmingly likely one.

2.3 Defining the spectra

I will compute the "scalar" (ζ) and "tensor" (h_{ij}) power spectra in parallel, since the calculations are very similar. The starting point for all of our subsequent calculation is the quadratic action obtained by expanding our original action (1) up to quadratic order in the small perturbations ζ and h_{ij}. Even though the calculation of this quadratic part should be straight forward, it gets messy if you don't do it right; Maldacena explains a good way to do this calculation (in Section 2 of arXiv:astro-ph/0210603) and I simply quote the result in Eq. (9) below. First we split ζ and h_{ij} into Fourier components as follows

$$\zeta(\eta,\vec{x}) = \int \frac{d^3\vec{k}}{(2\pi)^{3/2}}\zeta_{\vec{k}}(\eta)e^{i\vec{k}\vec{x}} \tag{8a}$$

$$h_{ij}(\eta,\vec{x}) = \int \frac{d^3\vec{k}}{(2\pi)^{3/2}}\epsilon^s_{ij}(\vec{k})h^s_{\vec{k}}(\eta)e^{i\vec{k}\vec{x}}, \tag{8b}$$

where we sum over polarizations $s = \pm$ in (8b), and the 3×3 polarization tensors ϵ^+_{ij} and ϵ^-_{ij} are real ($\epsilon^*_{ij} = \epsilon_{ij}$), symmetric ($\epsilon_{ij} = \epsilon_{ji}$), traceless ($\epsilon_{ii} = 0$), transverse ($\epsilon_{ij}k^j = 0$), even-parity [$\epsilon_{ij}(\vec{k}) = \epsilon_{ij}(-\vec{k})$] and "orthonormal" [$\epsilon^s_{ij}(\vec{k})\epsilon^{s'}_{ij}(\vec{k}) = 4\delta^{ss'}$]. Of course, our final expression for the tensor spectrum will not depend on the normalization of the polarization tensors ϵ^s_{ij}, but this particular choice is convenient, because it canonically normalizes the scalar fields $h_\pm$ (see below).

At quadratic order, ζ and h_{ij} are governed by the actions

$$S_\zeta = -\frac{1}{2}\int \mathrm{d}\eta \mathrm{d}^3\vec{x}\left[(M_{pl}z)^2\eta^{\mu\nu}\zeta_{,\mu}\zeta_{,\nu}\right] \tag{9a}$$

$$S_h = -\frac{1}{8}\int \mathrm{d}\eta \mathrm{d}^3\vec{x}\left[(M_{pl}a)^2\eta^{\mu\nu}h_{ij,\mu}h_{ij,\nu}\right] \tag{9b}$$

$$= -\frac{1}{2}\int \mathrm{d}\eta \mathrm{d}^3\vec{x}\left[(M_{pl}a)^2\eta^{\mu\nu}h^s_{,\mu}h^s_{,\nu}\right], \tag{9c}$$

where $\vec{x}$ is a comoving coordinate, η is conformal time, $\eta^{\mu\nu}$ is the Minkowski metric (diagonal$\{-1,+1,+1,+1\}$), a prime ($'$) denotes $\mathrm{d}/\mathrm{d}\eta$, and we must sum over $s=\pm$ in (9c). I have also defined the background function

$$z \equiv \sqrt{2\epsilon}\, a, \tag{10}$$

along with the two dimensionless "scalar" fields

$$h_s(\eta,\vec{x}) = \int \frac{\mathrm{d}^3\vec{k}}{(2\pi)^{3/2}} h^s_{\vec{k}}(\eta)\mathrm{e}^{i\vec{k}\vec{x}}. \tag{11}$$

The action (9a) describes a massless scalar field $M_{pl}\zeta$, minimally coupled to an unperturbed FRW background — but with z playing the role of the scale factor. And the action (9b) for h_{ij} is equivalent to the action (9c) for 2 massless, canonically-normalized scalar fields, $M_{pl}h_+$ and $M_{pl}h_-$, minimally coupled to the usual unperturbed FRW background with scale factor $a(\eta)$.

From the actions (9) we obtain the equations of motion

$$\zeta'' + 2(z'/z)\zeta' - \nabla^2\zeta = 0 \tag{12a}$$

$$h_s'' + 2(a'/a)h_s' - \nabla^2 h_s = 0 \tag{12b}$$

and the conjugate momenta

$$\pi_\zeta = \partial\mathcal{L}_\zeta/\partial\zeta' = (M_{pl}z)^2\zeta' \tag{13a}$$

$$\pi_s = \partial\mathcal{L}_h/\partial h_s' = (M_{pl}a)^2 h_s' \tag{13b}$$

Now promote the classical fields (ζ,π_ζ) to quantum fields $(\hat{\zeta},\hat{\pi}_\zeta)$ satisfying the canonical commutation relations

$$\left[\hat{\zeta}(\eta,\vec{x}),\hat{\pi}_\zeta(\eta,\vec{x}\,')\right] = i\delta^{(3)}(\vec{x}-\vec{x}\,') \tag{14a}$$

$$\left[\hat{\zeta}(\eta,\vec{x}),\ \hat{\zeta}\ (\eta,\vec{x}\,')\right] = \left[\hat{\pi}_\zeta(\eta,\vec{x}),\hat{\pi}_\zeta(\eta,\vec{x}\,')\right] = 0. \tag{14b}$$

And similarly, let the fields $(h_s,\pi_s)\to(\hat{h}_s,\hat{\pi}_s)$ satisfy the commutation relations

$$\left[\hat{h}_s(\eta,\vec{x}),\hat{\pi}_{s'}(\eta,\vec{x}\,')\right] = i\delta^3(\vec{x}-\vec{x}\,')\delta_{ss'} \tag{15a}$$

$$\left[\hat{h}_s(\eta,\vec{x}),\hat{h}_{s'}(\eta,\vec{x}\,')\right] = \left[\hat{\pi}_s(\eta,\vec{x}),\hat{\pi}_{s'}(\eta,\vec{x}\,')\right] = 0. \tag{15b}$$

Since $\hat{\zeta}$ is a real (hermitean) field, its Fourier components satisfy $\hat{\zeta}_{\vec{k}} = \hat{\zeta}^{\dagger}_{-\vec{k}}$. Similarly, since $\hat{h}_s$ is hermitean [this follows from the hermiticity of $\hat{h}_{ij}$ and our definition of ϵ^s_{ij}], its fourier components satisfy $\hat{h}^s_{\vec{k}} = \hat{h}^{s\dagger}_{-\vec{k}}$. Thus, we may write

$$\hat{\zeta}_{\vec{k}}(\eta) = \zeta_k(\eta) a_{\vec{k}} + \zeta^*_k(\eta) a^{\dagger}_{-\vec{k}} \tag{16a}$$

$$\hat{h}^s_{\vec{k}}(\eta) = h_k(\eta) a^s_{\vec{k}} + h^*_k(\eta) a^{s\dagger}_{-\vec{k}}. \tag{16b}$$

Here the ζ creation and annihilation operators ($a^{\dagger}_{\vec{k}}$ and $a_{\vec{k}}$), and the h_s creation and annihilation operators ($a^{s\dagger}_{\vec{k}}$ and $a^s_{\vec{k}}$), satisfy the usual commutation relations

$$\left[a_{\vec{k}}, a^{\dagger}_{\vec{k}'}\right] = \delta^{(3)}(\vec{k} - \vec{k}') \tag{17a}$$

$$\left[a_{\vec{k}}, a_{\vec{k}'}\right] = \left[a^{\dagger}_{\vec{k}}, a^{\dagger}_{\vec{k}'}\right] = 0 \tag{17b}$$

$$\left[a^s_{\vec{k}}, a^{s'\dagger}_{\vec{k}'}\right] = \delta^{(3)}(\vec{k} - \vec{k}')\delta^{ss'} \tag{18a}$$

$$\left[a^s_{\vec{k}}, a^{s'}_{\vec{k}'}\right] = \left[a^{s\dagger}_{\vec{k}}, a^{s'\dagger}_{\vec{k}'}\right] = 0, \tag{18b}$$

while the classical mode functions, $\{\zeta_k(\tau), \zeta^*_k(\tau)\}$ and $\{h_k(\tau), h^*_k(\tau)\}$, are linearly-independent solutions of the (fourier-transformed) equations of motion

$$\zeta''_k + 2(z'/z)\zeta'_k + k^2\zeta_k = 0 \tag{19a}$$

$$h''_k + 2(a'/a)h'_k + k^2 h_k = 0. \tag{19b}$$

Note that homogeneity allows us to choose mode functions that do not depend on the direction of $\vec{k}$ (or on the polarization s), but only on the magnitude $k = |\vec{k}|$ and on η. Consistency of the 2 sets of commutation relations, (14) and (17) in the scalar case, or (15) and (18) in the tensor case, fixes the normalization of the mode-function Wronskians:

$$W(\zeta_k, \zeta^*_k) \equiv \zeta_k(\eta)\zeta^{*\prime}_k(\eta) - \zeta^*_k(\eta)\zeta'_k(\eta) = \frac{i}{(M_{pl}z)^2} \tag{20a}$$

$$W(h_k, h^*_k) \equiv h_k(\eta)h^{*\prime}_k(\eta) - h^*_k(\eta)h'_k(\eta) = \frac{i}{(M_{pl}a)^2}. \tag{20b}$$

To understand the behavior of Eqs. (19a, 19b), it is often helpful to switch to new variables v_k and μ_k:

$$\zeta_k = v_k/z \tag{21a}$$

$$h_k = \mu_k/a \tag{21b}$$

which obey equations of motion

$$v''_k + [k^2 - (z''/z)]v_k = 0 \tag{22a}$$

$$\mu''_k + [k^2 - (a''/a)]\mu_k = 0. \tag{22b}$$

From here we see that, in the far past, when modes are inside the Hubble horizon ($k^2 \gg |z''/z|$ or $k^2 \gg |a''/a|$, respectively), the solution of Eqs. (19a,19b) is

$$\zeta_{_k}(\eta) \approx \frac{1}{(M_{pl} z)\sqrt{2k}} \left[A_1(k)\mathrm{e}^{-ik\eta} + A_2(k)\mathrm{e}^{+ik\eta}\right] \tag{23a}$$

$$h_{_k}(\eta) \approx \frac{1}{(M_{pl} a)\sqrt{2k}} \left[B_1(k)\mathrm{e}^{-ik\eta} + B_2(k)\mathrm{e}^{+ik\eta}\right] \tag{23b}$$

where the Wronskian conditions (20a, 20b) require that the constants (A_1, A_2) and (B_1, B_2) must satisfy

$$|A_1|^2 - |A_2|^2 = 1, \tag{24a}$$

$$|B_1|^2 - |B_2|^2 = 1. \tag{24b}$$

Different choices for (A_1, A_2) or for (B_1, B_2) generally define physically distinct vacua. The standard choice is $(A_1 = 1,\ A_2 = 0)$ and $(B_1 = 1, B_2 = 0)$, so that the mode functions $\zeta_{_k}$ and $h_{_k}$ are "purely positive frequency." Physically, this corresponds to placing the scalar and tensor fluctuations in the natural vacuum of a comoving observer in the far past:

$$\zeta_{_k} \to \frac{1}{(M_{pl} z)\sqrt{2k}}\mathrm{e}^{-ik\eta}, \tag{25a}$$

$$h_{_k} \to \frac{1}{(M_{pl} a)\sqrt{2k}}\mathrm{e}^{-ik\eta}. \tag{25b}$$

So the ζ and h_{ij} power spectra are given by

$$P_\zeta(k,\eta) \equiv \frac{\mathrm{d}\langle 0|\,\hat{\zeta}^2\,(\eta,\vec{x})|0\rangle}{\mathrm{d}\ln k} = \frac{k^3}{2\pi^2}|\zeta_{_k}|^2 \tag{26a}$$

$$P_h(k,\eta) \equiv \frac{\mathrm{d}\langle 0|\hat{h}_{ij}^2(\eta,\vec{x})|0\rangle}{\mathrm{d}\ln k} = 8\frac{k^3}{2\pi^2}|h_{_k}|^2, \tag{26b}$$

where $\zeta_{_k}(\eta)$ and $h_{_k}(\eta)$ are obtained by solving the equations of motion (19a, 19b), subject to the initial conditions (25a, 25b).

Now, in the slow-roll approximation, we can calculate extremely useful approximate formulae by the following argument. Focus on one particular Fourier mode, of comoving wavelength $\lambda = 2\pi/k =$ constant. Initially, $\lambda/2\pi$ is smaller than the comoving Hubble radius $1/aH$; but $1/aH$ is rapidly shrinking during inflation, and at some moment $\eta = \eta_{\rm exit}$ during inflation it becomes smaller than $\lambda/2\pi$; and then, after inflation, $1/aH$ re-expands until, at some moment $\eta = \eta_{\rm re-entry}$, it becomes larger than $\lambda/2\pi$. Cosmologists commonly describe this situation (very confusingly) by saying that the corresponding mode was initially "inside the Hubble horizon," then "exited the Hubble horizon" at the time $\eta_{\rm exit}$ during inflation; and finally "re-entered the Hubble horizon" at the time $\eta_{\rm re-entry}$ during the radiation or matter era. Before horizon exit, the mode functions $\zeta_k(\tau)$ and $h_{_k}(\eta)$ may be approximated by Eqs. (25a, 25b); and then after horizon exit the friction terms in Eqs. (19a, 19b) take over and pin $\zeta_k(\eta)$ and $h_k(\eta)$ to constant values – approximately their values at the moment η_* of horizon exit when the friction kicked in. So the power spectra (26a, 26b) also become time-independent constants

$$P_\zeta(k) = \left[\frac{k}{2\pi(M_{pl}z)}\right]^2 \approx \frac{1}{2\epsilon}\left[\frac{H}{2\pi M_{pl}}\right]^2 \tag{27a}$$

$$P_h(k) = 8\left[\frac{k}{2\pi(M_{pl}a)}\right]^2 = 8\left[\frac{H}{2\pi M_{pl}}\right]^2 \tag{27b}$$

where all of the quantities on the right-hand-side of these equations are understood to be evaluated at the moment of horizon exit $\eta = eta_{\rm exit}$ when $(\lambda/2\pi = 1/aH \;\Rightarrow\; k = aH)$. It is this η-independent (but k-dependent) primordial power spectrum that we can measure directly by making a map of the temperature and polarization of the CMB over the whole sky. According to inflation, these measurements are giving us direct information about ϵ and H as a function of time (and hence $V(\varphi)$) during this incredibly early and high energy epoch, prior to the start of radiation domination. If inflation is correct, this is probably the best glimpse of ultra-high-energy physics that we will get.

3 Relating the primordial and present-day gravitational wave spectra

From the equation of motion (19b) for h_k, we see that "outside the horizon" (where k^2 is negligible), h_k quickly damps to a constant value; and then, from the equation of motion written in the form (22b), we see that after the mode re-enters the horizon (where $k^2 \gg |a''/a|$) it behaves like $h_k \propto {\rm e}^{-ik\eta}/[a(\eta)]$. Therefore, the value of h_k today is simply its primordial constant value (when it was outside the horizon), divided by the redshift $(1+z_k)$ since it re-entered the horizon. And since the tensor spectrum $P_h(k)$ is proportional to $|h_k|^2$ squared, we roughly have

$$P_h(k,\text{today}) = P_h(k,\text{primordial})/(1+z_k)^2. \tag{28}$$

Next note that, from the continuity equation $\dot\rho = -3H(\rho+p)$, we see that when the ratio $w = p/\rho$ is a constant, ρ scales as $\rho \propto a^{-3(1+w)}$. Then, from the Friedmann equation in a flat FRW universe we have $H \propto \rho^{1/2} \propto a^{-3(1+w)/2}$, and hence $a \propto (1/aH)^{2/(1+3w)} \Rightarrow (1+z_k) \propto k^{2/(1+3w_k)}$. When talking about a stochastic background today, it is convenient to convert the strain spectrum $P_h(k)$ to an energy spectrum

$$\Omega_{gw}(k,\eta) \equiv \frac{1}{\rho_{crit}(\eta)}\frac{d\langle 0|\rho_{gw}(\eta)|0\rangle}{d\ln k} \tag{29}$$

where we can obtain $\rho_{gw}(\tau)$ by first using the formula

$$T_{\alpha\beta} = -2\frac{\delta L}{\delta \bar g^{\alpha\beta}} + \bar g_{\alpha\beta}L \tag{30}$$

to compute the gravitational wave stress-energy, and then using

$$\rho_{gw} = -T^0_0 = \frac{1}{64\pi G_N}\frac{(h'_{ij})^2 + (\vec{\nabla} h_{ij})^2}{a^2} \tag{31}$$

and averaging over multiple wavelengths to obtain

$$\Omega_{gw}(k,\tau) = \frac{1}{12}\frac{k^2 P_h(k,\tau)}{a^2(\tau)H^2(\tau)}. \tag{32}$$

Since the gravitational wave stress energy tensor $T_{\mu\nu}$ is proportional to $h_{ij,\mu}h_{ij,\nu}$, rather than $h_{ij}h_{ij}$, $\Omega_{gw}(k)$ is bluer than $P_h(k)$ by two powers of k. Putting it all together, we find that

$$\Omega_{gw}(k, \text{today}) \propto P_h(k, \text{primordial}) k^{2(3w_k - 1)/(3w_k+1)}. \tag{33}$$

As I will sketch on the board, this equation may be used to see why the current CMB upper bound on the primordial tensor power spectrum implies an upper bound $\Omega_{gw}(k) \lesssim 10^{-15}$ for modes that re-entered the horizon during the radiation-dominated era after standard single-field slow-roll inflation. This is orders of magnitude too small to be detected by Pulsar Timing Arrays, LISA, or LIGO, although a future satellite experiment like BBO or DECIGO (basically advanced versions of LISA, operating at slightly higher frequencies) migth someday be able to achieve sensitivities of $\Omega_{gw} \sim 10^{-17}$ and thereby directly detect the inflationary background.