Advanced Cosmology : Primordial non-Gaussianities

Eugene A. Lim

This is a set of lectures given to the Cambridge DAMTP Math Tripos 2012 Advanced Cosmology Class (Lent Term). What it really needs are figures right now.

I. GAUSSIAN RANDOM FIELDS AND CHARACTERIZATION OF NON-GAUSSIAN FIELDS

Consider a field $f(\mathbf{x})$ living on some space S, where $\mathbf{x}$ is a coordinate on S. We can Fourier transform this

$$f(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} f(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} \ , \ f(\mathbf{k}) = \int d^3x f(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} \tag{1}$$

where, without loss of generality, we can parameterize the Fourier coefficients as

$$f(\mathbf{k}) = a_{\mathbf{k}} + ib_{\mathbf{k}}, \tag{2}$$

with amplitude $|f(\mathbf{k})| = \sqrt{a_{\mathbf{k}}^2 + b_{\mathbf{k}}^2}$. Reality of $f(\mathbf{x})$ imposes $a_{\mathbf{k}} = a_{-\mathbf{k}}$, $b_{\mathbf{k}} = -b_{-\mathbf{k}}$.

In words, for a given configuration in real space $f(\mathbf{x})$, a set of real numbers $(a_{\mathbf{k}}, b_{\mathbf{k}})$, parameterized by the vector $\mathbf{k}$, completely (and uniquely) describe it. Different configurations of $f(\mathbf{x})$ are described by different set of numbers. Suppose now we want to *randomly* generate a field configuration $f(\mathbf{x})$, then one way to do it is to prescribe a probability distribution function for the set $(a_{\mathbf{k}}, b_{\mathbf{k}})$. If, furthermore, we want $f(\mathbf{x})$ to be a *Random Gaussian Field* configuration, then this put a tight constraint on what is the PDF for $(a_{\mathbf{k}}, b_{\mathbf{k}})$. Let's begin by studying the PDF for one particular mode $\mathbf{k}$.

Definition : A Gaussian distribution for a mode is one such that $a_{\mathbf{k}}$ and $b_{\mathbf{k}}$ is drawn from the following Gaussian distribution *in k space* with zero mean, and k-dependent variance σ_k

$$P(a_{\mathbf{k}}, b_{\mathbf{k}}) = \frac{1}{\pi\sigma_k^2} \exp\left[-\frac{a_{\mathbf{k}}^2 + b_{\mathbf{k}}^2}{\sigma_k^2}\right]. \tag{3}$$

where we have normalized this distribution such that

$$\int_{-\infty}^{\infty} da_{\mathbf{k}} \int_{-\infty}^{\infty} db_{\mathbf{k}} \ \frac{1}{\pi\sigma_k^2} \exp[-|f(\mathbf{k})|^2/\sigma_k^2] = 1. \tag{4}$$

Note that the integral is over *all possible values* of $a_{\mathbf{k}}$ and $b_{\mathbf{k}}$ for the mode $\mathbf{k}$. The double Gaussian integral results in the normalization $1/\pi\sigma_k^2$ instead of $\sqrt{1/\pi\sigma_k^2}$.

Generalizing to all the modes, then the formal definition of the PDF is via the *functional*

$$P[f(\mathbf{k})] = \frac{1}{\pi\sigma_k^2} \exp[-|f(\mathbf{k})|^2/\sigma_k^2]. \tag{5}$$

We have made quite a large leap between Eqn. (3) and Eqn. (5), so let's decode a bit. A *functional*, as you may have learned, is a map that eats a function, and spits out a scalar field. Eqn. (5) is the just high brow way of saying that all the fields $f(\mathbf{k})$ are drawn from a distribution that obeys Eqn. (3).

Notice that, in general we can have a directional dependence in the variance, i.e. $\sigma_{\mathbf{k}}^2$ instead of σ_k^2. But if we assume *statistical isotropy*, then the variance becomes just a function of the amplitude k – this also of course means that we can have non-isotropic but still Random Gaussian fields.

Now we encounter the first instance of the much abused angled brackets : $\langle\ \rangle$. What is $\langle Q\rangle$? Here what we are doing is "taking the *expectation value*" of some observable Q over an *ensemble* of possible realizations of this observable. Consider the following simple example. A plane drops many parachutists from the sky, who will land on a line x on the ground. The parachutists are scattered around some normalized distribution $\rho(x)$.

The expectation value of x is then defined to be

$$\langle x\rangle = \int_{-\infty}^{\infty} x\rho(x)dx. \tag{6}$$

In words, the *expectation value* is the average of a large (formally infinite) number of drawings from this distribution, i.e. our "best" guess value. If you like, for example, $\rho(x)$ could be some normalized Gaussian distribution with mean x_0 and variance σ

$$\rho(x) = \sqrt{\frac{1}{\pi\sigma^2}}\exp[-(x-x_0)^2/\sigma^2]. \tag{7}$$

We can also calculate the expectation value of any function of x, say $g(x)$ i.e.

$$\langle g(x)\rangle = \int_{-\infty}^{\infty} g(x)\rho(x)dx. \tag{8}$$

Given a functional distribution $P[\phi(\mathbf{k})]$ for some scalar field $\phi(\mathbf{k})$, then we can generalize the idea : the expectation value of finding the *functional* of some field configuration $\phi(\mathbf{k})$, $Q[\phi(\mathbf{k})]$, is given by (remember, a functional spits out a scalar)

$$\langle Q[\phi(\mathbf{k})]\rangle = \int \mathcal{D}\phi\ Q[\phi(\mathbf{k})]P[\phi(\mathbf{k})] \tag{9}$$

where the integral $\mathcal{D}\phi$ is over *all possible field configurations* in $\mathbf{k}$-space. (Of course, here we are working in fourier space – in general it can be any space that ϕ lives in, including real/configuration space $\mathbf{x}$).

With all these definitions, we can now define the notion of a *Gaussian Random Field*: Let $P[f(\mathbf{k})]$ be a PDF of the form Eqn. (5). Such a distribution is called *Gaussian Random*, and such a field $f(\mathbf{k})$ is called a Gaussian Random Field.

The expectation value for some functional $Q[f(\mathbf{k})]$ is given by

$$\langle Q[f(\mathbf{k})]\rangle = \Pi_{\mathbf{k}}\int da_{\mathbf{k}}\int db_{\mathbf{k}} Q[f(\mathbf{k})]\frac{1}{\pi\sigma_k^2}\exp[-(a_{\mathbf{k}}^2+b_{\mathbf{k}}^2)/\sigma_k^2]. \tag{10}$$

Note that the integral over all possible configurations has become $\mathcal{D}f \to \Pi_{\mathbf{k}}\int da_{\mathbf{k}}\int db_{\mathbf{k}}$, since $a_{\mathbf{k}}$ and $b_{\mathbf{k}}$ parameterize the configurations. Sometimes you hear the words "both the amplitudes and the phases of $f(\mathbf{k})$ are drawn from a Gaussian PDF" with variance σ_k^2. Now you can see this is clearly wrong : $|f(\mathbf{k})|^2 = a_{\mathbf{k}}^2 + b_{\mathbf{k}}^2$, so for each mode in a GRF, the *amplitude* is drawn from a Gaussian PDF while the *phase* is essentially drawn from a flat distribution – this is kinda obvious since you can't really define a Gaussian (with infinite support) over a compact phase, but these words are uttered with such religious fervor that we tend to repeat it without thinking.

Despite the scary looking equation Eqn. (10), with all the integrals and product sums and all, it is actually very easy to use it. Consider the simple example $Q[f(\mathbf{k})] = a_{\mathbf{k}}a_{\mathbf{k}'}$, then

$$\begin{aligned}\langle a_{\mathbf{q}}a_{\mathbf{q}'}\rangle &= \Pi_{\mathbf{k}}\int da_{\mathbf{k}}\int db_{\mathbf{k}}\ a_{\mathbf{q}}a_{\mathbf{q}'}\frac{1}{\pi\sigma_k^2}\exp[-(a_{\mathbf{k}}^2+b_{\mathbf{k}}^2)/\sigma_k^2] \\ &= \frac{\sigma_k^2}{2}\delta(\mathbf{q}+\mathbf{q}') + \mathbf{q}' \leftrightarrow -\mathbf{q}.\end{aligned} \tag{11}$$

The point is that since the distribution is even under both $a_{\mathbf{k}}$ and $b_{\mathbf{k}}$, all odd products of a's and b's vanish while, except for $\mathbf{k} = \mathbf{q}$ and $\mathbf{k}' = \mathbf{q}'$, the rest are simply trivial Gaussian integrals. Even-ness means that $\mathbf{q} = \mathbf{q}'$ gives us the delta function (recall that $a_{\mathbf{k}} = a_{-\mathbf{k}}$ by reality).

Let's do a simple but important example : what is the *expectation* of a two-point correlation function[1] $Q[f(\mathbf{k})] = f(\mathbf{k})f(\mathbf{k}')$:

$$\langle f(\mathbf{k})f(\mathbf{k}')\rangle = \langle a_{\mathbf{k}}a_{\mathbf{k}'}\rangle - \langle b_{\mathbf{k}}b_{\mathbf{k}'}\rangle = \sigma_k^2\delta(\mathbf{k}+\mathbf{k}') \tag{12}$$

where we have used the fact that the cross terms vanishes and reality imposes $b_{\mathbf{k}} = -b_{-\mathbf{k}}$ and

$$\langle a_{\mathbf{k}}a_{\mathbf{k}'}\rangle = \langle b_{\mathbf{k}}b_{-\mathbf{k}'}\rangle = \int d^3k\ a_{\mathbf{k}}a_{\mathbf{k}}P(a_{\mathbf{k}}, b_{\mathbf{k}})\delta(\mathbf{k}+\mathbf{k}') = \frac{\sigma_k^2}{2}\delta(\mathbf{k}+\mathbf{k}') \tag{13}$$

---

[1] Recall that the expectation value of any variable x or any product of variables $f(x)$, given its PDF $P(x)$, is $\langle f(x)\rangle = \int dx f(x)P(x)$.

This definition carries through to configuration space : to compute something like the power spectrum $\langle f(\mathbf{x})f(\mathbf{x}')\rangle$, Fourier transform $f(\mathbf{x})$, and plug into Eqn. (10). Let's do a famous example

$$\begin{aligned}\langle f(\mathbf{x}_1)f(\mathbf{x}_2)\rangle &\equiv \int \frac{d^3k_1}{(2\pi)^3}\frac{d^3k_1}{(2\pi)^3}e^{i\mathbf{k}_1\cdot\mathbf{x}_1+i\mathbf{k}_2\cdot\mathbf{x}_2}\left(\langle a_{\mathbf{k}_1}a_{\mathbf{k}_2}\rangle-\langle b_{\mathbf{k}_1}b_{\mathbf{k}_2}\rangle\right) \quad (14)\\ &= \int \frac{d^3k_1}{(2\pi)^3}\frac{d^3k_1}{(2\pi)^3}e^{i\mathbf{k}_1\cdot\mathbf{x}_1+i\mathbf{k}_2\cdot\mathbf{x}_2}\delta(\mathbf{k}_1+\mathbf{k}_2)\sigma^2_{k_1} \quad (15)\\ &= \int \frac{d^3k}{(2\pi)^3}e^{i\mathbf{k}\cdot(\mathbf{x}_1-\mathbf{x}_2)}\frac{\sigma_k^2}{(2\pi)^3} \quad (16)\end{aligned}$$

which is of course the two-point correlation function in real space. You probably have seen Eqn. (16) – it is the definition of the *power spectrum* $P(k)\equiv\sigma_k^2/(2\pi)^3$ or two-point correlation function. σ_k is a k dependent power spectrum, and tells us about the *amplitude* of the correlations at k.

Let us state a few famous true-facts using our high-brow tool:

- **Scale Invariance** : A *rescaling* is defined to be $\mathbf{x}\to\lambda\mathbf{x}$ where $\lambda>0$ is some constant. Then a *scale invariant power spectrum* obeys the following relation

$$\langle f(\mathbf{x})f(\mathbf{x}')\rangle = \langle f(\lambda\mathbf{x})f(\lambda\mathbf{x}')\rangle. \quad (17)$$

  Using $\delta(\lambda(\mathbf{k}+\mathbf{k}'))=\lambda^{-3}\delta(\mathbf{k}+\mathbf{k}')$, then it follows that for a power spectrum to be scale invariant, it must obey

$$P(k)=\frac{\sigma_k^2}{(2\pi)^3}\propto\frac{1}{k^3}. \quad (18)$$

  This should be familiar to you.

- **White Noise** : On the other hand, if $\sigma_k^2=\text{const}$ for all k, the spectrum is known as white noise. Note it is not scale-invariant!

- **Correlations vs Gaussianity** : Notice that Gaussianity is not a statement about two-point correlations! For example, if I take a random Gaussian sky (say the CMB), and then rearrange the cold and hot spots in such a way that it looks like Mickey Mouse[2], the sky will then be highly correlated in a very Mickey Mouse way, but *it will still be completely Gaussian random* – what I have changed in my magical rearrangment is the power spectrum $\sigma^2_{\mathbf{k}}$, not the underlying distribution $P[f(\mathbf{k})]$. Mickey Mouse is of course not very isotropic, so hence the directional dependent power spectrum.

It is clear that since the Gaussian PDF Eqn. (3) is even under parity around the mean 0, any odd expectation value vanishes

$$\langle f(\mathbf{k}_1)f(\mathbf{k}_2)f(\mathbf{k}_3)\rangle=0 \quad (19)$$

while any even expectation value is simply a product of the respective power spectra which we can compute, for example for a 4-pt function

$$\langle f(\mathbf{k}_1)f(\mathbf{k}_2)f(\mathbf{k}_3)f(\mathbf{k}_4)\rangle=\sigma^2_{k_1}\sigma^2_{k_3}\delta(\mathbf{k}_1+\mathbf{k}_2)\delta(\mathbf{k}_3+\mathbf{k}_4)+1\leftrightarrow3+1\leftrightarrow4. \quad (20)$$

You can do the algebra here to convince yourself (Example Sheet), but this kind of "contraction" algebra occurs very often so it's good to practice doing the combinatorics. In QFT it's called **Wick's Theorem**, although according to Wikipedia, it's also called **Isserlis theorem** : if $f(\mathbf{k})$ is a Gaussian Random Field, then

$$\langle f(\mathbf{k}_1)f(\mathbf{k}_2)f(\mathbf{k}_3)f(\mathbf{k}_4)...\rangle=\sum\text{All Possible Two Point Contractions.} \quad (21)$$

You can show similarly the expectation values in real space follow the same relation, i.e. odd correlations vanish, e.g.

$$\langle f(\mathbf{x}_1)f(\mathbf{x}_2)f(\mathbf{x}_3)\rangle=0 \quad (22)$$

[2] Thanks to I-S.Yang for this very descriptive example.

On the other hand, if a field $f(\mathbf{x})$ is *non-Gaussian*, then *odd* correlations do not vanish. What is non-Gaussianity? One can think of breaking non-Gaussianities by changing the k-dependent PDF Eqn. (3), say by adding a k-dependent *skewness parameter* $S_k < 1$ to the distribution[3] $a_{\mathbf{k}}$

$$P(a_{\mathbf{k}}, b_{\mathbf{k}}) = \frac{1}{\pi\sigma_k^2}\exp\left[-\frac{a_{\mathbf{k}}^2 + b_{\mathbf{k}}^2}{\sigma_k^2}\right]\left(1 + \sigma_k S_k \frac{a_{\mathbf{k}}}{\sigma_k}\right). \tag{23}$$

It is easy to show that the two-point correlation function remains unchanged $\langle f(\mathbf{k})f(\mathbf{k}'))\rangle = \sigma_k^2\delta(\mathbf{k}+\mathbf{k}')$. However the 1-point correlation function

$$\langle f(\mathbf{k})\rangle = S_k\frac{\sqrt{\pi}\sigma_k}{2}, \tag{24}$$

no longer vanishes by virtue of the odd term. It is not hard to show that all the odd correlations no longer vanish too.

One can think of the PDF Eqn. (23), if $S_k \ll 1$, to be a slightly skewed Gaussian distribution, hence S_k is known as the *skewness* parameter. Eqn. (23) is usually known as the *Edgeworth* expansion, which is an expansion in the variance σ_0 from a Gaussian PDF.

Hers is something more interesting

$$P(a_{\mathbf{k}}, b_{\mathbf{k}}) = \frac{1}{\pi\sigma_k^2}\exp\left[-\frac{a_{\mathbf{k}}^2 + b_{\mathbf{k}}^2}{\sigma_k^2}\right]\left(1 + \sigma_k \frac{S_k}{6} H_3\left(\frac{a_{\mathbf{k}}}{\sigma_k}\right)\right). \tag{25}$$

where H_3 is the 3rd Hermite polynomial, given by

$$H_3(x) = -12x + 8x^3 \tag{26}$$

which you might note that is an *odd* function of x. This is a very special PDF since its 1-point correlation function vanishes (example sheet), and the two-point correlation function reproduces the usual Gaussian result Eqn. (12). On the other hand, the 3-point auto-correlation is

$$\langle f(\mathbf{k})f(\mathbf{k})f(\mathbf{k})\rangle = S_k\sigma_k^4. \tag{27}$$

Notice that Eqn. (27) is proportional to the power spectrum squared $(\sigma_k^2)^2$. This will be a repeating theme : the next to leading order non-Gaussianity, i.e. the 3 point correlations function, is usually cast as some parameter (here we used S_k) times the power spectrum squared. However, one can think more generally and consider an underlying PDF which depends on more than a single k. In other words, if the underlying PDF is a function of three different $\mathbf{k}$'s but constrained by "momentum conservation" $P(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = F(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)\delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)$ instead of Eqn. (3), then it is clear that a 3-pt correlation function do not vanish. "Momentum conservation" comes out the fact that the 3-pt correlation in real space $\langle f(\mathbf{x}_1)f(\mathbf{x}_2)f(\mathbf{x}_3)\rangle$ must form a triangle so in Fourier space this translates to a triangle in momentum space.

Finally, as in the power spectrum case, **scale invariance** of the 3-pt correlation function is defined such that

$$\langle f(\mathbf{x})f(\mathbf{x}')f(\mathbf{x}'')\rangle = \langle f(\lambda\mathbf{x})f(\lambda\mathbf{x}')f(\lambda\mathbf{x}'')\rangle. \tag{28}$$

As we have studied in inflationary physics, the *initial spectrum of perturbations* are Gaussian by the Central Limit theorem. Our goal in this first part of these lectures is to study how such a Gaussian initial perturbations *stay almost, but not quite, Gaussian* during inflation. We will see that Gaussianity is a very strong prediction of inflation, yet *slight* deviations from Gaussianity provide to us a powerful probe of the theory of inflation. There are several very strong model-independent predictions of the kind of non-Gaussianity, in addition to some classes of non-Gaussianity which is semi model-independent.

In these lectures, we will study the formalism in which we can calculate, from first principles, the non-Gaussianity of a generic single-field inflation theory. This formalism, first presented by Maldacena, has since become the standard prescription of primordial non-Gaussian calculations. The goal is to present Maldacena's original derivation, with commentary and additional material, such that at the end of the lectures, you will be able to do your own calculations of different models.

[3] One can do the same for $b_{\mathbf{k}}$ but it is a bit more cumbersome due to the need to keep the distribution real.

A. An example : Komatsu-Spergel Local form

Before we begin, let us look at a common parameterization of the 3-pt correlation function, introduced by Komatsu and Spergel

$$\Phi(\mathbf{x}) = \Phi_G(\mathbf{x}) + f_{NL}^{local}(\Phi_G(\mathbf{x})^2 - \langle\Phi_G(\mathbf{x})^2\rangle) \tag{29}$$

where $\Phi_G(\mathbf{x})$ is a Gaussian random field in the sense that its Fourier transform coefficients $\Phi(\mathbf{k})$

$$\Phi_G(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3}\Phi_G(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}} \tag{30}$$

are drawn from a Gaussian PDF as in the previous section. I.e. the angle brackets $\langle Q\rangle$ means expectation value. The $\Phi(\mathbf{x})$ are to be considered as temperature anisotopies $\Delta T/T$ at a point in the sky $\mathbf{x}$. In other words, we split the field $\Phi(\mathbf{x})$ into its linear Gaussian part $\Phi_G(\mathbf{x})$ and its non-linear part which is the square of its local value $\Phi(\mathbf{x})^2$ minus the *variance* of its Gaussian part $\langle\Phi_G(x)^2\rangle$. The f_{NL}^{local} is a constant "coupling" parameter which does not depend on scale k. We will soon encounter many others. It is easy to show that

$$\langle\Phi_G(\mathbf{x})\rangle = 0. \tag{31}$$

Following Eqn. (16), the two-point correlation function, or the *linear* power spectrum is then

$$\langle\Phi_G(\mathbf{x})\Phi_G(\mathbf{x}')\rangle \equiv \int \frac{d^3k}{(2\pi)^3}P(k)e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} \tag{32}$$

Now, we can "sort of derive" the above by assuming that the angle brackets $\langle\ \rangle$ on the LHS denotes *sum over all points in space* $\mathbf{x}$ and defining $P(k) = |\Phi_G(k)|^2$, then $\langle\Phi_G(\mathbf{x})^2\rangle$ is the spatial variance of the field $\Phi_G(\mathbf{x})$ viz

$$\int \frac{d^3k_1}{(2\pi)^3}\ P(k_1) = \int \frac{d^3k_1}{(2\pi)^3}\frac{d^3k_2}{(2\pi)^3}\ \Phi_G(\mathbf{k}_1)\Phi_G(\mathbf{k}_2)(2\pi)^3\delta(\mathbf{k}_1-\mathbf{k}_2) \tag{33}$$

$$= \int d^3x \int \frac{d^3k_1}{(2\pi)^3}\frac{d^3k_2}{(2\pi)^3}\ \Phi_G(\mathbf{k}_1)\Phi_G(\mathbf{k}_2)(2\pi)^3 e^{i(\mathbf{k}_1-\mathbf{k}_2)\cdot\mathbf{x}} \tag{34}$$

$$= \int d^3x\Phi_G(\mathbf{x})\Phi_G(\mathbf{x}) = \langle\Phi_G(\mathbf{x})\Phi_G(\mathbf{x})\rangle \tag{35}$$

Also,

$$\int d^3k_2 P(k_1)\delta(\mathbf{k}_1-\mathbf{k}_2) = \int \frac{d^3k_2}{(2\pi)^3}\ \langle\Phi_G(\mathbf{k}_1)\Phi_G^*(\mathbf{k}_2)\rangle \tag{36}$$

$$\langle\Phi_G(\mathbf{k}_1)\Phi_G^*(\mathbf{k}_2)\rangle = (2\pi)^3P(k_1)\delta(\mathbf{k}_1-\mathbf{k}_2). \tag{37}$$

and using $\Phi(\mathbf{k}) = \Phi^*(-\mathbf{k})$,

$$\langle\Phi_G(\mathbf{k}_1)\Phi_G(\mathbf{k}_2)\rangle = (2\pi)^3P(k_1)\delta(\mathbf{k}_1+\mathbf{k}_2). \tag{38}$$

At a pinch, thinking $\langle\ \rangle$ as an integration over space will do. But there is a sneaky assumption made here, i.e. that the all the points in space is drawn from the same standardized distribution (e.g. Eqn. (5)). This is fine when we do CMB or LSS physics since we are assuming that each point in the sky provide to us an honest drawing from the distribution, but bear in mind the formal definition Eqn. (10) – i.e. the 3-pt correlation function we are going to compute is an expectation value of an *ensemble*. This is known as the "Ergodic Theorem"[4].

---

[4] The Ergodic Theorem states that, under reasonable conditions, the average over all space and *ensemble* averages are essentially the same. The conditions are that the regions of the fields are (1) uncorrelated at large distances and (2) the correlations stays the same under time and space translations – basically a "statistical" Copernican Principle. In equations the first condition is, for any given constant z

$$\langle\Phi(\mathbf{x}_1+z)\Phi(\mathbf{x}_2+z)...\Phi(\mathbf{x}_n+z)\rangle = \langle\Phi(\mathbf{x}_1)\Phi(\mathbf{x}_2)...\Phi(\mathbf{x}_n)\rangle \tag{39}$$

where if z is space it is called "Homogenous" and if z is time, it is called "stationary". The second condition is

$$\langle\Phi(\mathbf{x}_1+\mathbf{u})\Phi(\mathbf{x}_2+\mathbf{u})...\Phi(\mathbf{y}_1-\mathbf{u})\Phi(\mathbf{y}_2-\mathbf{u})...\rangle \overset{|\mathbf{u}|\to\infty}{=} \langle\Phi(\mathbf{x}_1+\mathbf{u})\Phi(\mathbf{x}_2+\mathbf{u})...\rangle\langle\Phi(\mathbf{y}_1-\mathbf{u})\Phi(\mathbf{y}_2-\mathbf{u}...)\rangle \tag{40}$$

which states that the correlations are uncorrelated as the distance $\mathbf{u}$ becomes large.

With this in mind, the Fourier transform of Eqn. (29) is

$$\Phi(\mathbf{k}) = \Phi_G(\mathbf{k}) + f_{NL}^{local}\left(\int \frac{d^3p}{(2\pi)^3}\ \Phi_G(\mathbf{k}+\mathbf{p})\Phi_G^*(\mathbf{p}) - (2\pi)^3\delta(\mathbf{k})\langle\Phi_G(\mathbf{x})^2\rangle\right) \equiv \Phi_G(\mathbf{k}) + \Phi_{NL}(\mathbf{k}) \tag{41}$$

where we have inserted a $1 = \int d^3k\delta(\mathbf{k})$ in the last term. The three-point correlation in Fourier space is then

$$\langle\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi(\mathbf{k}_3)\rangle = \langle\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi_{NL}(\mathbf{k}_3)\rangle + \mathrm{sym} + ... \tag{42}$$

where the ... means higher order terms and sym means an interchange of the indices (3 of them). Each term has the form

$$\langle\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi_{NL}(\mathbf{k}_3)\rangle = f_{NL}^{local}\langle\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\left(\int \frac{d^3p}{(2\pi)^3}\ \Phi_G(\mathbf{k}_3+\mathbf{p})\Phi_G^*(\mathbf{p}) - (2\pi)^3\delta(\mathbf{k}_3)\langle\Phi_G(\mathbf{x})^2\rangle\right)\rangle \tag{43}$$

$$= f_{NL}^{local}\langle\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\rangle\left(\int \frac{d^3p}{(2\pi)^3}\ \langle\Phi_G(\mathbf{k}_3+\mathbf{p})\Phi_G^*(\mathbf{p})\rangle - (2\pi)^3\delta(\mathbf{k}_3)\langle\Phi_G(\mathbf{x})^2\rangle\right) \tag{44}$$

$$+f_{NL}^{local}\int \frac{d^3p}{(2\pi)^3}\langle\Phi(\mathbf{k}_1)\Phi^*(\mathbf{p})\rangle\langle\Phi(\mathbf{k}_2+\mathbf{p})\Phi(\mathbf{k}_3)\rangle + 2 \to 3 \tag{45}$$

where we have used the fact that the constituent fields Φ_G are Gaussian random, so we can contract them using Wick's Theorem, i.e. the 4-pt is a sum of all possible combinations of two points (Eqn. (20)

$$\langle\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi(\mathbf{k}_3)\Phi(\mathbf{k}_4)\rangle = \langle\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\rangle\langle\Phi(\mathbf{k}_3)\Phi(\mathbf{k}_4)\rangle + \mathrm{sym} \tag{46}$$

Using the definition for the power spectrum, the first term in Eqn. (45) vanishes while the 2nd term becomes $(2\pi)^3P(k_1)P(k_2)\delta(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)$, so the total is

$$\langle\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi_{NL}(\mathbf{k}_3)\rangle = f_{NL}^{local}2(2\pi)^3P(k_1)P(k_2)\delta(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3) \tag{47}$$

so the final 3-pt correlation function for the Komatsu-Spergel local form is of the *scale-invariant* form (you will be asked to show this in an Example)

$$\langle\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi(\mathbf{k}_3)\rangle = f_{NL}^{local}\left[2(2\pi)^3P(k_1)P(k_2)\delta(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3) + \mathrm{sym}\right]. \tag{48}$$

In other words, the 3-pt correlation function can be expressed as a sum over a product of power spectrum times some non-linear coupling f_{NL}^{local}. This form is very special, because the non-linear field $\Phi(\mathbf{x})$ is constructed out of *local* (i.e. located at the same point $\mathbf{x}$) Gaussian random fields. However, it is a very important quantity − as its form is very close to what a single field slow-roll featureless inflationary model predicts. Indeed, such a theory predicts

$$f_{NL}^{local} \sim \epsilon \tag{49}$$

where $\epsilon \ll 1$ is the slow roll parameter, i.e, a very small number. But how small is small? Recall that, $\Phi(\mathbf{x}) \sim 10^{-5}$ for the CMB, so from Eqn.(48), we can see that $f_{NL}^{local} \sim 1$ implies a non-Gaussianity of 1 part in 10^5. So this is very tiny amount of non-Gaussianity indeed, and is the source of much folklore regarding how inflation will predict an unobservable amount of non-Gaussianity. However, as we will see in these lectures, there exist many perfectly viable models of inflation which predict a much larger amount of non-Gaussianity, and probing for non-Gaussianity allow us to sift through the myriad of inflationary models in the market.

Current best bounds on local form non-Gaussianity is given by the WMAP satellite

$$74 > f_{NL}^{local} > -10 \tag{50}$$

to 95% confidence. The PLANCK satellite can probe up to $\Delta f_{NL}^{local} \sim 5$, so the window is closing in on non-Gaussianity.

II. IN-IN FORMALISM AND HIGHER ORDER PERTURBATION THEORY

A. Two Approaches to Hell : Canonical Perturbation Theory vs Background Field Method

Since the dawn of time, cosmologists have pondered the burning question : given a inflationary model, what is the primordial non-Gaussianity generated? The basic idea is as follows. Consider the action of a scalar field minimally coupled to gravity

$$S = \int d^4x\sqrt{g}\left[\frac{M_P^2 R}{2} - \frac{1}{2}(\nabla\phi)^2 - V(\phi)\right] \tag{51}$$

and its perturbations

$$\phi(\mathbf{x},t) \rightarrow \bar{\phi}(t) + \delta\phi(\mathbf{x},t)\ ,\ g_{\mu\nu} \rightarrow \bar{g}_{\mu\nu} + \delta g_{\mu\nu}. \tag{52}$$

Restricting to scalar perturbations, and choosing a gauge such that $\delta\phi = 0$, we can then expand the metric in terms of the (Bardeen) curvature perturbation ζ

$$\delta g_{ij} = a^2(\gamma_{ij} + 2\zeta\delta_{ij}). \tag{53}$$

We know from linear theory that, given an inflationary model with slow roll parameters (with dots denoting derivatives w.r.t. cosmic time)

$$\epsilon \equiv -\frac{\dot{H}}{H^2}\ ,\ \eta \equiv \frac{\dot{\epsilon}}{H\epsilon} \tag{54}$$

the linear power spectrum at the end of inflation is

$$\langle\zeta(\mathbf{k}_1,t_0)\zeta^*(\mathbf{k}_2,t_0)\rangle = (2\pi)^3 P_\zeta\delta(\mathbf{k}_1 - \mathbf{k}_2)\ ,\ P_\zeta = \frac{H^2}{8\pi\epsilon}. \tag{55}$$

Our goal is to calculate the higher correlation function – in particular, we want to calculate the leading correction to the 2-pt correlation, i.e. the 3-pt correlation function

$$\langle\zeta(\mathbf{k}_1,t_0)\zeta(\mathbf{k}_2,t_0)\zeta(\mathbf{k}_3,t_0)\rangle \tag{56}$$

where we add in t_0 as time at the end of inflation to emphasise that the 3-pt correlation is an *evolving* quantity. Now, inflationary initial conditions are Gaussian random, as every high school (Russian) student knows. However we know that gravity is non-linear, and the expansion (and acceleration)of the universe is driven by gravity, so the question is : how did the universe stay almost (but not quiet) Gaussian during inflation?

$$\text{GRF initial conditions} \stackrel{\text{non-linear dynamics}}{\longrightarrow} \text{Non-Gaussian end of inflation} \tag{57}$$

The non-vanishing of Eqn. (56) tells us that the modes $\mathbf{k}_1$,$\mathbf{k}_2$ and $\mathbf{k}_3$ are *correlated* – but from linear theory we know that the modes evolved independently. To calculate this quantity then, we need to move beyond linear perturbation theory, and consider non-linear perturbation theory.

The so-called "canonical" approach would be simply to perturb variables to second order

$$\phi \rightarrow \bar{\phi} + \delta\phi^{(1)} + \delta\phi^{(2)} + \ldots\ ,\ g_{\mu\nu} \rightarrow \bar{g}_{\mu\nu} + \delta g^{(1)}_{\mu\nu} + \delta g^{(2)}_{\mu\nu} + .. \tag{58}$$

and insert this into the equation of motion and solve order by order iteratively

$$M_p^2\bar{G}_{\mu\nu} = \bar{T}_{\mu\nu}\ ,\ M_p^2\delta G^{(1)}_{\mu\nu} = \delta T^{(1)}_{\mu\nu}\ ,\ M_p^2\delta G^{(2)}_{\mu\nu} = \delta T^{(2)}_{\mu\nu}. \tag{59}$$

The first order terms are then sourced by products of second order terms, heuristically the EOM is then

$$\ddot{\phi}^{(1)}(\mathbf{x}) - \nabla^2\phi^{(1)}(\mathbf{x}) \sim g\phi^{(2)}(\mathbf{x})\phi^{(2)}(\mathbf{x}) \Longrightarrow \ddot{\phi}^{(2)}(\mathbf{k}) + k^2\phi^{(2)}(\mathbf{k}) \sim g\int\frac{d^3k'}{(2\pi)^3}\phi^{(2)}(\mathbf{k})\phi^{(2)}(\mathbf{k}-\mathbf{k}') \tag{60}$$

where g is some non-linear coupling. This is historically the method which was first applied (by Acquaviva et al, 2002), but as one can imagine, it is a huge mess (e.g. Acquaviva didn't quite get it right despite a heroic calculation). Not only is it algebraic cumbersome, there are much confusion in the literature on "second order gauge transformation". Furthermore, if you want to go beyond 2nd order, you have to perturb to 3rd order, and the iteration become tedious. Instead, we are going to use a second, more modern and elegant approach, the so-called "in-in formalism" to compute correlation functions.

The "in-in" formalism is really a big word for something you already know from your QFT days : calculating correlation functions using a perturbative expansion around some background field configuration, hence it is a form of background field method. Now you may ask, isn't what we want to do is simply *perturbative classical field theory*, what has QFT to do with all this? The answer is that a lot of "quantum" calculations in QFT are really "classical" calculations – tree level diagrams for example[5], insofar as that the once the initial conditions are specified, we are simply evolving them using classical field equations.

In this section, we will lay out the general idea of the this formalism, and the nitty gritty derivation details for later. But first, more pontificating.

B. What is $\langle\rangle$? (Quantum Edition)

To understand the idea of using "in-in" QFT technique to calculate correlations, we need to get to grip with our old foe "What is $\langle\rangle$"? Previously, we argue that it is the expectation value of some observable Q over an ensemble of realizations. Now, let's set that idea aside for the moment, and decode a statement we have made way back in Lecture 2 : "Now, inflationary initial conditions are Gaussian random, as every high school (Russian) student knows". What that means is that, for a given mode *amplitude* k, there are an *infinite* number of realizations (remember for each k there are an infinite number of directions), and each is drawn from the distribution Eqn. (3) with some power σ_k^2 given by some inflationary theory.

But we have learned that these initial conditions are "quantum fluctuations" of the inflaton field ϕ in its ground state $|0\rangle$. Let's see how this lead to the Gaussian initial conditions that we have asserted. We'll ignore backreaction to gravity and consider a massive scalar field ϕ in some curved space background, say FRW. We can then quantize it as usual – give it hats $\phi \rightarrow \hat{\phi}$, and calculate the 2-point correlation function of its vacuum state:

$$\langle\hat{\phi}(\mathbf{k}_1,t_0)\hat{\phi}(\mathbf{k}_2,t_0)\rangle = \langle 0|\hat{\phi}(\mathbf{k}_1,t_0)\hat{\phi}(\mathbf{k}_2,t_0)|0\rangle = |u_k(t_0)|^2\delta(\mathbf{k}_1-\mathbf{k}_2), \tag{61}$$

where $u_k(t)$ is the mode function of the operator $\hat{\phi}$, which we allow it to evolve in whatever way (depending on the potential for ϕ)

$$\hat{\phi}(\mathbf{x},t) = \int \frac{d^3k}{(2\pi)^3}\left[a(\mathbf{k})u_k(t)e^{i\mathbf{k}\cdot\mathbf{x}} + a^\dagger(\mathbf{k})u_k^*(t)e^{-i\mathbf{k}\cdot\mathbf{x}}\right], \tag{62}$$

with commutator

$$[a(\mathbf{k}), a^\dagger(\mathbf{k}')] = (2\pi)^3\delta(\mathbf{k}-\mathbf{k}') \tag{63}$$

and the vacuum state is one annihilated by the lower operator

$$a(\mathbf{k})|0\rangle = 0. \tag{64}$$

Now, if we look at equation Eqn. (61), it is tempting to identify

$$|u_k(t)|^2 \Longleftrightarrow \sigma_k^2. \tag{65}$$

In fact this is exactly what we do (to make contact with usual inflation calculation you recall $|u_k|^2 = P(k)^2$). The question is : why are we allowed to do that? The answer is both trivial and deep.

The trivial part of the answer goes something as follows : the field ϕ is really a quantum field[6] and each of its mode is a quantum harmonic oscillator, and by construction each mode started out in its ground state and hence by

---

[5] In fact, in "in-in" calculations, loops are not necessary quantum corrections – in the language of correlations, loops are simply integrals over momenta $\int d^3k$ (or space if you are working in configuration space) and that could mean an average over states. This goes way beyond the scope of our class though – go bug somebody smart like Daniel Baumann or Xingang Chen if you'd like to learn more.

[6] Somebody famous once said that we should say "classical approximation" instead of "classical". I don't remember who this famous person is, so s/he must not be that famous.

linearity stays in its ground state (remember that in the quantum SHO the ground state is a stationary state). Then the probability of finding, say the state of $\phi(\mathbf{k}) \equiv a_{\mathbf{k}} + ib_{\mathbf{k}}$ is, as we learn from doing quantum harmonic oscillator, is a Gaussian "wavepacket". In the $\phi(\mathbf{k})$ representation[7]

$$\langle\phi(\mathbf{k})|0\rangle \propto \exp[-|\phi(\mathbf{k})|^2/|u_k|^2] \propto \exp[-(a_{\mathbf{k}}^2 + b_{\mathbf{k}}^2)/|u_k|^2] \tag{66}$$

and the normalized distribution for $\phi(\mathbf{k})$ is some Gaussian with the variance given by the mode function $|u_k|^2$. So, now as we make observations of the sky, mode by mode, we are drawing from this same distribution of the same mode *amplitude* – hence it makes sense to see the quantum expectation value as the statistical expectation value. You probably have heard some version of this story somewhere.

The deep part of the answer, and less often pondered over, is the notion that when we make an observation, say measure the amplitude of some previously unobserved mode in the CMB, *we are really making an observation over a hitherto "uncollapsed" wavefunction of the universe.* In other word, everytime we point our telescope into the sky, or stare out deep into the cosmos, we are collapsing some part of the wave function of the universe – and once we do that, the value we get is an eigenvalue of our observable (say the temperature) and the wavefunction collapses to an eigenfunction. This is so mind-blowing, that it sometimes keep me up at night (ok, it doesn't really, but it gives me goosebumps).

C. An Overview of "In-in"

In the previous section, we have argued that the expectation value of two fields over an ensemble is equivalent to computing the two-point correlation function of the quantized field over the some state. Notice that we have specified previously that the state we are taking expectation over is the vacuum state, and that led to a Gaussian distribution. But the point is that *we don't have to take the expectation value over the vacuum state!* We also said that, quickly in passing, that a state started out in vacuum will stay in vacuum in linear theory since there is no interaction that will excite it – all the k modes do not know about the existence of each other.

However, if we introduce an interaction between the different modes, then it is not surprising that the states can now mix, and hence there is no reason that the vacuum states will stay in vacuum. This is the main idea of the "in-in" formalism : we want to calculate how the states, which started out as vacuum, evolves in the presence of non-linear interactions. In the past two paragraphs, we have sneaked in the important code word : "interaction". Recall that when you quantize the Harmonic oscillator, the first thing you calculate is its Hamiltonian, and that if we want to do interesting to the oscillator we add "interaction" Hamiltonians into it. This is roughly what we are going to do here. So, with all this fun and games over, let's now go to the hard and formal definition of "in-in".

Historically, the "in-in" formalism was developed first Schwinger (1961) and then by Keldysh (1964) for condensed matter systems (where they care about correlation functions instead of S-matrix scattering amplitudes). The formalism is a lot more powerful than what we are going to do – it is used to treat non-equilibrium systems and critical phenomenon. It was then first applied to cosmology by R. Jordan (1986) and Calzetta+Hu (1987), the latter under the name "Closed-Time-Path" formalism (you will learn later why it is called that). Maldacena used the formalism, wrongly at first[8] in 2002 to calculate higher order cosmological calculations in his now famous paper. S. Weinberg then put the whole formalism, applied to cosmology by formalizing classical-background-quantum-fluctuation split and time-dependent couplings in 2005[9].

1. From "in-out" to "in-in" (*)

Let's motivate our discussion from what we know in standard "in-out" QFT – you can skip this section if you want to get to doing calculations. Recall from your QFT class, that given an interacting field theory, we can split the Hamiltonian H into its "free part" H_0 and an interaction term H_{int}

$$H = H_0 + H_{int}. \tag{67}$$

---

[7] Recall that in a simple harmonic oscillator with Hamiltonian $H = (1/2)p^2 + (1/2)x^2$, in the x representation, $\langle x|0\rangle \propto e^{-x^2}$.

[8] I was told this story by the guy who corrected him, Steven Weinberg, so it must be true.

[9] My own small part in this business is that in 2008, with P. Adshead and R. Easther, we formalized the canonical quantization form (as opposed to path integrals) of this calculation and clarified a subtle issue in Weinberg's paper dealing with contours. Our paper has an error in it, dealing with dimensional regularization, which was in turn corrected by L. Senatore and M. Zaldarriaga in 2009.

For the moment, we assume that the interaction Hamiltonian H_{int} is time-independent. The "in" (+, at $t \to -\infty$) and "out" (-, at $t \to +\infty$) states $|\Psi_\alpha^\pm\rangle$ are Lorentz invariant and hence describe the states throughout all spacetime. The subscript α denotes the entire collection of all particles and we use the Weinbergian shorthand

$$\int d\alpha ... \equiv \sum_{all\ particles} \int d^3\vec{P} \tag{68}$$

to sum over all particles and momenta. Let $|\Psi_\alpha^\pm\rangle$ be eigenstates of the full Hamiltonian H

$$H|\Psi_\alpha^\pm\rangle = E_\alpha |\Psi_\alpha^\pm\rangle \tag{69}$$

which obey completeness

$$\langle \Psi_\alpha^\pm | \Psi_{\alpha'}^\pm \rangle = \delta(\alpha - \alpha'). \tag{70}$$

Eigenstates are not localized, so we consider wave packets $g(\alpha)$

$$|\Psi_g^\pm(t)\rangle \equiv \int d\alpha\ e^{-iE_\alpha t} g(\alpha) |\Psi_\alpha^\pm\rangle \tag{71}$$

Sadly, the eigenstates of H are not easy to find since it is an interacting theory, so this is the end of the story...but wait! If we can consider the interaction term H_{int} to be a *perturbation*, then we can do perturbation theory by working in the Hilbert space of the free Hamiltonian H_0. Notice that I have just condensed entire chapters of QFT into two sentences. Anyhow, let's define "free states" $|\Phi_\alpha^\pm\rangle$ which are eigenstates of H_0

$$H_0|\Phi_\alpha^\pm\rangle = E_\alpha |\Phi_\alpha^\pm\rangle \tag{72}$$

and its wavepacket

$$|\Phi_g^\pm(t)\rangle \equiv \int d\alpha\ e^{-iE_\alpha t} g(\alpha) |\Phi_\alpha^\pm\rangle. \tag{73}$$

We now want $|\Psi^\pm\rangle$ to contain the same particles at $t \to \pm\infty$ as the free eigenstates $|\Phi^\pm\rangle$, so the condition is

$$\int d\alpha\ e^{-iE_\alpha t} g(\alpha) |\Psi_\alpha^\pm\rangle \to \int d\alpha\ e^{-iE_\alpha t} g(\alpha) |\Phi_\alpha^\pm\rangle \tag{74}$$

or

$$e^{-iHt} \int d\alpha\ g(\alpha) |\Psi_\alpha^\pm\rangle \to e^{-iH_0 t} \int d\alpha\ g(\alpha) |\Phi_\alpha^\pm\rangle. \tag{75}$$

We can write the formal solution of Ψ as a function of Φ and H_{int} (the Lippmann-Schwinger Equation)

$$|\Psi_\alpha^\pm\rangle = |\Phi_\alpha^\pm\rangle + \frac{1}{E_\alpha - H_0 + i\epsilon} H_{int} |\Psi_\alpha^\pm\rangle. \tag{76}$$

The $i\epsilon$ is added in make the inverse meaningful. But we don't have to fret about why we put it in as long as we can show that we recover the condition Eqn. (75) then this is one formal prescription. Expanding $|\Psi\rangle$ with a complete set of free particle states and then integrating Eqn. (76) with $\int d\alpha\ g(\alpha)$, we get

$$|\Psi_g^\pm(t)\rangle = |\Phi_g^\pm(t)\rangle + \int d\alpha \int d\beta \frac{\langle \Phi_\beta | H_{int} | \Psi_\alpha^\pm \rangle e^{-iE_\alpha t} g(\alpha)}{E_\alpha - E_\beta \pm i\epsilon} |\Phi_\beta\rangle. \tag{77}$$

At $t \to \pm\infty$, the double integral on the right must vanish. Consider the in (+) vacuum $t \to -\infty$ and do the α integral: we can close the contour using the upper half plane. All the contributions from that contour, as well as any singularities that appear in $g(\alpha)$ and $\langle \Phi_\beta | H_{int} | \Psi_\alpha^\pm \rangle$ are killed by $e^{-iE_\alpha t}$. And since we closed the contour above the real axis, we missed the residue at $E_\alpha = E_\beta - i\epsilon$ which lies below the real axis. So the entire integral vanish at $t \to -\infty$. Similarly for the out (-) vacuum, we close the integral below the real axis. This is of course the usual prescription in "picking the vacuum" in QFT.

Now let us consider just the "in" vacuum $|\Psi_\alpha^+\rangle$. The condition Eqn. (75) still apply (for +). In "in-in", we are computing correlations, so let us consider the scalar product of Ψ at $t \to -\infty$. According to Eqn. (75) this must equal to the free state correlation i.e. unity. Since $\langle\Psi|$ is simply the conjugate of $|\Psi^+\rangle$, we can again expanding Eqn. (76) in a complete set, and then integrating over $\int d\alpha g(\alpha)$ (the cross terms vanish as above by the same contour chosen for + vacuum, and dropping the + superscript)

$$\langle\Psi_g(t)|\Psi_g(t)\rangle = \langle\Phi_g(t)|\Phi_g(t)\rangle + \int d\alpha d\alpha' d\beta d\beta' g(\alpha)g(\alpha') \frac{\langle\Psi_{\alpha'}|H_{int}|\Phi_{\beta'}\rangle\langle\Phi_\beta|H_{int}|\Psi_\alpha\rangle}{(E_\alpha - E_\beta + i\epsilon)(E_{\alpha'} - E_{\beta'} - i\epsilon)}\langle\Phi_\beta|\Phi_{\beta'}\rangle e^{-iE_\alpha t + iE_{\alpha'} t} \quad (78)$$

using completeness, we get rid of one of the β' integral

$$\langle\Psi_g(t)|\Psi_g(t)\rangle = \langle\Phi_g(t)|\Phi_g(t)\rangle + \int d\alpha d\alpha' d\beta g(\alpha)g(\alpha') \frac{\langle\Psi_{\alpha'}|H_{int}|\Phi_\beta\rangle\langle\Phi_\beta|H_{int}|\Psi_\alpha\rangle}{(E_\alpha - E_\beta + i\epsilon)(E_{\alpha'} - E_\beta - i\epsilon)} e^{-iE_\alpha t + iE_{\alpha'} t}. \quad (79)$$

For the LHS to be the same as the RHS, the integral must vanish at $t \to -\infty$. Now for the $\int d\alpha$ integral, we close the E_α contour in the upper half plane as above. On the other hand, the $\int d\alpha'$ integral, to kill off the contributions from the contour, we need $\Im(E_{\alpha'}) < 0$ since the exponent has a + sign, so we close it in the lower half plane instead. This also misses the singularity at $E_\beta + i\epsilon$, and hence the integral vanish hence we are done.

All this is a bit heavy going and rather opaque – the point is that "in-in" is nothing but roughly "$|\text{QFT}|^2$", with two following subtleties (which we will discuss in words but not too explicit about it – check out any QFT book if you like to learn about picking vacuum by clever contouring)

- We only need to pick the "in" vacuum state, so the "out" contour never enter the game.
- You can think of doing the $\int d\alpha$ and $\int d\alpha'$ as integrals over some contour in the complex time plane. Then when integrating over the correlator $\langle\Phi_\beta|H_{int}|\Psi_\alpha\rangle$, we take the contour from $-\infty$ to t in the positive imaginary plane, and then when we do the integral over the $\langle\Psi_{\alpha'}|H_{int}|\Phi_\beta\rangle$ correlator, we loop over at t and then back to $-\infty$ in the negative imaginary plane. The loop never formally "close" at $-\infty$, although this didn't stop Calzetta + Hu from calling it the "Closed-Time-Path" formalism. We will see this loop explicitly later. You can, if you like, cast the whole thing in path integral language, but we won't do it in this class.

2. *General Formula for "in-in" Correlators*

Now let us generalize this (to cosmology, in particular). Don't worry if all this is still opaque – we will state the formula without proof and then prove it later. In general, we want to consider *quantized fields on some classical background*. Then instead of splitting the Hamiltonian into free and interacting parts, we split the Hamiltonian into 3 part

$$H = H_b + H_0 + H_{int} \quad (80)$$

where the background H_b constructed out of *classical* fields in the background, the "free field" perturbation H_0 constructed out of perturbations to quadratic order and the interaction Hamiltonian H_{int} constructed out of perturbations to 3rd and higher order. Without proof, let's state the 3-pt correlation function

$$\langle\zeta(\mathbf{k}_1,t)\zeta(\mathbf{k}_2,t)\zeta(\mathbf{k}_3,t)\rangle = \left\langle \left[\bar{T}\exp\left(i\int_{-\infty(1-i\epsilon)}^t dt'\ H_{int}(t')\right)\right]\zeta_I(\mathbf{k}_1,t)\zeta_I(\mathbf{k}_2,t)\zeta_I(\mathbf{k}_3,t)\left[T\exp\left(-i\int_{-\infty(1+i\epsilon)}^t dt'\ H_{int}(t')\right)\right]\right\rangle \quad (81)$$

where the subscript I denotes *interaction picture* fields, i.e. the values for the fields are evolved using the free Hamiltonian H_0. The interaction Hamiltonian H_{int} is also constructed out of the interaction picture fields ζ_I (and the background evolution of course), and T and $\bar{T}$ are the time and anti-time ordering operator respectively. More generally, one can imagine calculating the correlation functions of a product of field operators $W(t)$

$$\langle W(t)\rangle = \left\langle\left[\bar{T}\exp\left(i\int_{-\infty(1-i\epsilon)}^t dt'\ H_{int}(t')\right)\right]W^I(t)\left[T\exp\left(-i\int_{-\infty(1+i\epsilon)}^t dt'\ H_{int}(t')\right)\right]\right\rangle \quad (82)$$

To be able to calculate this quantity, we need to know how to do the following :

- Derive the Eqn. (82). I.e. we will show the formal derivation of the in-in formalism with a time-dependent background.
- Calculate H_{int} to 3rd and higher order. To do this we will need to extract the true degrees of freedom from the coupled scalar-gravity theory.
- Solve Eqn. (81) given the H_{int}.

D. In-In Formalism of Generic Fields

In this section, we will discuss the "in-in" formalism in calculating cosmological perturbations. Those who has some experience with calculating S-matrices in QFT will find this familiar, but there are differences in the calculation.

- We are calculating correlation functions at some time $t > t_0$, given boundary conditions at time t_0, instead of "in-out" matrix element. To see the connection, insert complete sets of "out" states α and β into Eqn. (81)

$$\langle W(t)\rangle = \int d\alpha \int d\beta \langle 0| \left(Te^{-i\int_{t_0}^{t} H_{\rm int}(t')dt'}\right)^{\dagger} |\alpha\rangle\langle\alpha|W(t)|\beta\rangle\langle\beta| \left(Te^{-i\int_{t_0}^{t} H_{\rm int}(t'')dt''}\right)|0\rangle \,. \tag{83}$$

  It is clear from this picture that the correlation function is the product of vacuum transitions "in-out" amplitudes, and a matrix element $\langle\alpha|W(t)|\beta\rangle$, summed over all possible "out" states. This also means that "disconnected" diagrams, i.e. "vacuum fluctuation" terms in in-in automatically cancel.

- There is no time flow in the calculation, unlike "in-out" – each insertion of the vertex is associated with not just momentum conservation, but also a time integral.

- More technically, which we will discuss further later: since we are not calculating propagation of fields, the relevant Greens function is the Wightman function, instead of Feynman.

With this preamble, let us consider a Lagrangian with generic fields ϕ_a,

$$S = \int d^4x \mathcal{L}(\phi_a(\mathbf{x},t), \dot{\phi}_a(\mathbf{x},t)) = \int dt L \tag{84}$$

In a generic system, a range over all the fields in the theory, be it the metric, matter scalar fields etc. We will keep the discussion general in this section, and specialize to a FRW cosmology in the next section. The canonical momenta for this system is defined as usual

$$\pi_a = \frac{\partial \mathcal{L}}{\partial \dot{\phi}_a} \tag{85}$$

and hence the Hamiltonian is then

$$H[\phi_a(t), \pi_a(t)] = \int d^3x \ \dot{\phi}_a \pi_a - L. \tag{86}$$

In the quantum theory, the variables obey the equal time commutator relations

$$[\phi_a(\mathbf{x},t), \pi_b(\mathbf{y},t)] = i\delta_{ab}\delta(\mathbf{x}-\mathbf{y}) \tag{87}$$

and other commutators vanish as usual. The Heisenberg equations of motion are then

$$\dot{\phi}_a(\mathbf{x},t) = i[H[\phi(t),\pi(t)], \phi_a(\mathbf{x},t)] \ , \ \dot{\pi}_a(\mathbf{x},t) = i[H[\phi(t),\pi(t)], \pi_a(\mathbf{x},t)]. \tag{88}$$

So far, just the usual story. Now we want to consider the following system. We want to split the fields into a "classical" part $\bar{\phi}$ and its "quantum" part $\delta\phi$

$$\phi_a(\mathbf{x},t) = \bar{\phi}_a(\mathbf{x},t) + \delta\phi_a(\mathbf{x},t) \ , \ \delta\pi_a(\mathbf{x},t) = \bar{\pi}_a(\mathbf{x},t) + \delta\pi_a(\mathbf{x},t) \tag{89}$$

where the classical part obey the classical Hamiltonian equations of motion

$$\dot{\bar{\phi}}_a(\mathbf{x},t) = \frac{\delta H[\bar{\phi}(t),\bar{\pi}(t)]}{\delta\bar{\pi}_a(\mathbf{x},t)} \ , \ \dot{\bar{\pi}}_a(\mathbf{x},t) = -\frac{\delta H[\bar{\phi}(t),\bar{\pi}(t)]}{\delta\bar{\phi}_a(\mathbf{x},t)}. \tag{90}$$

The notation can be tricky – what we are doing is varying the Hamiltonian *density*

$$\dot{\bar{\phi}}(\mathbf{x},t) = \int d^3y \ \frac{\delta\mathcal{H}[\bar{\phi}(\mathbf{y},t),\bar{\pi}(\mathbf{y},t)]}{\delta\bar{\pi}(\mathbf{y},t)}\delta(\mathbf{x}-\mathbf{y}). \tag{91}$$

Since classical part is just a c-number, its commutator with everything vanishes, e.g.

$$[\bar{\phi}(\mathbf{x},t), H] = [\bar{\pi}(\mathbf{x},t), H] = 0. \tag{92}$$

In the application to cosmological perturbation theory, the classical quantities are just what we normally call "background quantities". In FRW, these would be the the scale factor $a(t)$ and its derivatives (e.g. the Hubble parameter $H(t)$), and the background field quantities for the inflaton ϕ. Plugging these into Eqn. (87), we see that the perturbations obey the commutator

$$[\delta\phi_a(\mathbf{x},t),\delta\pi_b(\mathbf{y},t)] = i\delta_{ab}\delta(\mathbf{x}-\mathbf{y}) \tag{93}$$

The way to think about this system is that it is a *theory of quantized perturbations living on the classical time-dependent background*. This should be familiar to you in the context of linear cosmological perturbation theory, but in this section we will formalize it, and consider beyond linear order perturbation theory. We should emphasise here that the classical-quantum split Eqn. (89) is *completely arbitrary* – we simply made the choice. What we are doing here is that we assume that the background fields are sitting in their (possibly time-dependent) "vacuum expectation value', i.e. $\langle\hat{\phi}_a\rangle = \phi_a$. In terms of cosmology, these background fields are the metric and the homogenous background scalar field(s).

Proceeding, we can expand these fields in their raising and lowering operators, and construct a Hilbert space if you like. But as we will see later, the more useful Hilbert space to work in will be the interaction picture space, so we'll postpone this step and plow on.

For simplicity, let's drop the subscript a and consider a single field system – you can generalize this to multifields if you like by adding indices. Now, let us expand the Hamiltonian around the classical background

$$H[\phi(t),\pi(t)] = H[\bar{\phi}(t),\bar{\pi}(t)] + \frac{\delta H[\bar{\phi}(t),\bar{\pi}(t)]}{\delta\bar{\phi}(\mathbf{x},t)}\delta\phi(\mathbf{x},t) + \frac{\delta H[\bar{\phi}(t),\bar{\pi}(t)]}{\delta\bar{\pi}(\mathbf{x},t)}\delta\pi(\mathbf{x},t) + \tilde{H}[\delta\phi(t),\delta\pi(t);t] \tag{94}$$

where

$$\tilde{H}[\delta\phi(t),\delta\pi(t);t] = \tag{95}$$

$$\frac{1}{2}\left[\frac{\delta^2 H[\bar{\phi}(t),\bar{\pi}(t)]}{\delta\bar{\phi}(\mathbf{x},t)\delta\bar{\phi}(\mathbf{y},t)}\delta\phi(\mathbf{x},t)\delta\phi(\mathbf{y},t) + \frac{\delta^2 H[\bar{\phi}(t),\bar{\pi}(t)]}{\delta\bar{\pi}(\mathbf{x},t)\delta\bar{\pi}(\mathbf{y},t)}\delta\pi(\mathbf{x},t)\delta\pi(\mathbf{y},t) + 2\frac{\delta^2 H[\bar{\phi}(t),\bar{\pi}(t)]}{\delta\bar{\phi}(\mathbf{x},t)\delta\bar{\pi}(\mathbf{y},t)}\delta\phi(\mathbf{x},t)\delta\pi(\mathbf{y},t) + ...\right] \tag{96}$$

where ... denotes 3rd and higher order terms. In other words $\tilde{H}$ contains terms of quadratic and higher order in the quantized perturbations. Several comments :

- The variation of the Hamiltonian $H[\bar{\phi}(t),\bar{\pi}(t)]$ is with respect to $\bar{\phi}(\mathbf{x},t)$ – note that $\delta\bar{\phi}(\mathbf{x},t)$ in the denominator is *not* the perturbation $\delta\phi(\mathbf{x},t)$ – see Eqn. (91).
- The perturbation Hamiltonian $\tilde{H}[\delta\phi(t),\delta\pi(t);t]$ term is a functional of the *perturbations*, with an explicit time-dependence on the background values $(\bar{\phi}(t),\bar{\pi}(t))$ which we denote by appending a "$;t$" at the end. To make contact with what you have learned in linear perturbation theory, let's consider the Mukhanov-Sasaki action for the perturbation variable $v(\mathbf{x},t)$

$$S_{MS} = \frac{1}{2}\int d^3x d\tau \,\left[v'^2 - (\partial v)^2 - \frac{z''}{z}v^2\right], \tag{97}$$

where $z(\tau) = a(\tau)\sqrt{-\frac{\dot{H}}{H^2}}$ denotes the background "classical" fields. The canonical momentum is

$$\delta\pi(\mathbf{x},\tau) = v'(\mathbf{x},\tau) \tag{98}$$

so the $\tilde{H}$ is

$$\tilde{H}[v(\tau),\delta\pi(\tau);\tau] = \frac{1}{2}\int d^3x\,\left[\delta\pi^2 + (\partial v)^2 + \frac{z''}{z}v^2\right] \tag{99}$$

where the "$;\tau$" at the end denotes explicit time-dependence due to $z(\tau)$.

Plugging the expansion Eqn. (94) into the Heisenberg EOM Eqn. (88), we get

$$\begin{aligned}\dot{\bar{\phi}}(\mathbf{x},t) + \delta\dot{\phi}(\mathbf{x},t) &= i\left[H[\bar{\phi}(t),\bar{\pi}(t)] + \int d^3y\left\{\frac{\delta\mathcal{H}\bar{\phi}(\mathbf{y},t),\bar{\pi}(\mathbf{y},t)]}{\delta\bar{\phi}(\mathbf{y},t)}\delta\phi_a(\mathbf{y},t) + \frac{\delta\mathcal{H}[\bar{\phi}(\mathbf{y},t),\bar{\pi}(\mathbf{y},t)]}{\delta\bar{\pi}(\mathbf{y},t)}\delta\pi_a(\mathbf{y},t)\right\}\right.\\ &\quad\left.+\tilde{H},\bar{\phi}_a(\mathbf{x},t)+\delta\phi(\mathbf{x},t)\right]\\ \dot{\bar{\phi}}(\mathbf{x},t) + \delta\dot{\phi}(\mathbf{x},t) &= i\left[\int d^3y\left\{\frac{\delta\mathcal{H}[\bar{\phi}(\mathbf{y},t),\bar{\pi}(\mathbf{y},t)]}{\delta\bar{\phi}(\mathbf{y},t)}\delta\phi_a(\mathbf{y},t) + \frac{\delta\mathcal{H}[\bar{\phi}(\mathbf{y},t),\bar{\pi}(\mathbf{y},t)]}{\delta\bar{\pi}(\mathbf{y},t)}\delta\pi_a(\mathbf{y},t)\right\} + \tilde{H},\delta\phi(\mathbf{x},t)\right]\end{aligned} \tag{100}$$

where we have used the fact that the background quantities commute with everything. Focusing on the first order commutators, we see that $[\delta\phi(\mathbf{x},t),\delta\phi(\mathbf{y},t)] = 0$ commute, and the 2nd term is

$$\left[\int d^3y \frac{\delta\mathcal{H}}{\delta\bar{\pi}(\mathbf{y},t)}\delta\pi(\mathbf{y},t),\delta\phi(\mathbf{x},t))\right] = \dot{\bar{\phi}}(\mathbf{x},t) \tag{101}$$

which we have used the background classical EOM Eqn. (91) and the commutator Eqn. (93). This term cancels the same term on the LHS of Eqn. (100). Similar computations can be done for $\delta\dot{\pi}$, giving us the final Heisenberg EOM for the quantized perturbations

$$\delta\dot{\phi}(\mathbf{x},t) = i[\tilde{H}[\delta\phi(t),\delta\pi(t);t],\delta\phi(\mathbf{x},t)]\ ,\ \delta\dot{\pi}(\mathbf{x},t) = i[\tilde{H}[\delta\phi(t),\delta\pi(t);t],\delta\pi(\mathbf{x},t)]. \tag{102}$$

In other words, we have shown that the perturbations are evolved using the *perturbative* Hamiltonian $\tilde{H}$, which is time-dependent in two ways : the time-dependence of the perturbation fields and also explicit time-dependence due to the time-changing background fields.

So the prescription is : find the Hamiltonian of the original action Eqn. (84), expand it powers of $\delta\phi$ and $\delta\pi$, and the evolution Hamiltonian for the perturbations consists of terms that are 2nd and higher order in these perturbations. Equivalently, one can expand the action Eqn. (84) to 2nd and higher order in perturbations, and find the Hamiltonian of the perturbations for the 2nd and higher order terms. We will follow the latter in the next section.

We are halfway through our derivation of Eqn. (82). The solution to the Heisenberg EOM Eqn. (102) is well known from QFT: it is given by

$$\delta\dot{\phi}(\mathbf{x},t) = U^{-1}(t,t_0)\delta\phi(\mathbf{x},t_0)U(t,t_0)\ ,\ \delta\dot{\pi}(\mathbf{x},t) = U^{-1}(t,t_0)\delta\pi(\mathbf{x},t_0)U(t,t_0) \tag{103}$$

where the unitary operator $U(t,t_0)$ is solves the equation[10]

$$\frac{d}{dt}U(t,t_0) = -i\tilde{H}[\delta\phi(t),\delta\pi(t);t]U(t,t_0) \tag{105}$$

We can set the initial condition to be $U(t_0,t_0) = 1$ (it is also clear that it is unitary $U^{-1}(t,t_0)U(t,t_0) = 1$). Again, as in standard QFT, we want to split $\tilde{H}$ into a free-field Hamiltonian H_0 and an interaction Hamiltonian H_{int}

$$\tilde{H} = H_0 + H_{int}. \tag{106}$$

To this end, we *define* the interaction picture fields $\delta\phi_I(\mathbf{x},t)$ and $\delta\pi_I(\mathbf{x},t)$ with initial conditions such that they equal the "full theory" fields at some initial time t_0 (in inflationary applications, this would be the Bunch-Davies initial conditions)

$$\delta\phi_I(\mathbf{x},t_0) = \delta\phi(\mathbf{x},t_0)\ ,\ \delta\pi_I(\mathbf{x},t_0) = \delta\pi(\mathbf{x},t_0) \tag{107}$$

and do it in such a way that we can construct H_{int} out of the interaction picture fields. We want to find the unitary operators $U_0(t,t_0)$ and $F(t,t_0)$ associated with the Hamiltonians H_0 and H_{int}, constructed out of these interaction picture fields. By construction, the interaction picture fields obey the following Heisenberg EOM

$$\delta\dot{\phi}_I(\mathbf{x},t) = i[H_0[\delta\phi_I(t),\delta\pi_I(t);t],\delta\phi_I(\mathbf{x},t)]\ ,\ \delta\dot{\pi}_I(\mathbf{x},t) = i[H_0[\delta\phi_I(t),\delta\pi_I(t);t],\delta\pi_I(\mathbf{x},t)]. \tag{108}$$

Now we use the fact that in Eqn. (108), we can choose the time at where we evaluate the Hamiltonian $H_0[\delta\phi_I,\delta\pi_I;t]$ – this is a consequence of the fact that H_0 commutes with itself and hence we can evolve it to anytime we want inside the commutator, in particular let's evolve it to the t_0, so then we can replace it with

$$H_0[\delta\phi_I(t),\delta\pi_I(t);t] \to H_0[\delta\phi(t_0),\delta\pi(t_0);t] \tag{109}$$

---

[10] We learned from QFT, this equation has the solution

$$U(t,t_0) = T\exp\left[-i\int_{t_0}^{t} dt'\ \tilde{H}(t')\right] \tag{104}$$

where T is the time-ordering operator. (The time ordering is required for time-dependent Hamiltonians.) But this form is not very useful, since it is hard to calculate $\delta\phi(\mathbf{x},t)$.

hence the U_0 obey the equation

$$\frac{d}{dt}U_0(t,t_0) = -iH_0[\delta\phi(t_0),\delta\pi(t_0);t]U_0(t,t_0). \tag{110}$$

Notice that this sleight of hand does not affect the its dependence on the background $;t$, so it keeps this explicit dependence on t. This equation has the solution

$$U_0(t,t_0) = T\exp\left[-i\int_{t_0}^{t} dt\ H_0(t)\right] \tag{111}$$

and initial condition $U_0(t_0,t_0) = 1$. Hence we can use Eqn. (111) to evolve the interaction picture fields – in practice this is nothing more than simply saying $\delta\phi_I$ and $\delta\pi_I$ are calculated using the free-field Hamiltonian. In application to inflationary perturbations, this means that the linear solutions given by solving the Mukhanov-Sasaki equation of motion are the interaction picture fields.

Now let's turn to the task of finding the $F(t,t_0)$, which is the unitary operator associated with the interaction Hamiltonian, i.e. we want it to solve

$$\frac{d}{dt}F(t,t_0) \overset{?}{=} -iH_{int}(t)F(t,t_0). \tag{112}$$

The question is what is H_{int} constructed out of?

Since $H_{int} = \tilde{H} - H_0$, we can guess that $F(t,t_0) = U_0^{-1}(t,t_0)U(t,t_0)$ and so plugging this into Eqn. (112) and using Eqns. (105) and (110) we get

$$\begin{aligned}
\frac{d}{dt}F(t,t_0) = \frac{d}{dt}(U_0^{-1}(t,t_0)U(t,t_0)) &= -U_0^{-2}\frac{dU_0}{dt}U + U_0^{-1}\frac{dU}{dt} && (113)\\
&= iU_0^{-1}(H_0[\delta\phi(t_0),\delta\pi(t_0);t] - \tilde{H}[\delta\phi(t_0),\delta\pi(t_0);t])U && (114)\\
&= -iU_0^{-1}H_{int}[\delta\phi(t_0),\delta\pi(t_0);t]U_0U_0^{-1}U && (115)\\
&= -iH_{int}^{I}[\delta\phi_I(t_0),\delta\pi_I(t_0);t]F(t,t_0) && (116)
\end{aligned}$$

where we have inserted a $U_0U_0^{-1}$ in the 3rd line, and rewritten

$$H_{int}^{I}[\delta\phi_I(t_0),\delta\pi_I(t_0);t] = U_0^{-1}(t,t_0)H_{int}[\delta\phi(t_0),\delta\pi(t_0);t]U_0(t,t_0). \tag{117}$$

In other words, $F(t,t_0)$ solves the equation

$$\frac{d}{dt}F(t,t_0) = -iH_{int}^{I}(t)F(t,t_0). \tag{118}$$

i.e. with the interaction Hamiltonian constructed out of the interaction picture fields. This equation has the usual time-ordered solution

$$F(t,t_0) = T\exp\left[-i\int_{t_0}^{t} dt\ H_{int}^{I}(t)\right]. \tag{119}$$

Putting all these together, for the expectation value of any product of operators $W(t)$,

$$\begin{aligned}
\langle W(t)\rangle &= \langle U^{-1}(t,t_0)W(t_0)U(t,t_0)\rangle && (120)\\
&= \langle U^{-1}(t,t_0)U_0(t,t_0)W^{I}(t)U_0^{-1}(t,t_0)U(t,t_0)\rangle && (121)\\
&= \langle F^{-1}(t,t_0)W^{I}(t)F(t,t_0)\rangle && (122)\\
&= \left\langle\left[\bar{T}\exp\left(i\int_{-\infty(1-i\epsilon)}^{t} dt'\ H_{int}^{I}(t')\right)\right]W^{I}(t)\left[T\exp\left(-i\int_{-\infty(1+i\epsilon)}^{t} dt'\ H_{int}^{I}(t')\right)\right]\right\rangle && (123)
\end{aligned}$$

which is Eqn. (82). Notice that although we have shown the above derivation inside the bra-kets, the state at which the correlation is evaluated remains arbitrary.

One thing we didn't show, but hope that you remember from your QFT days, is that we have appended an $\pm i\epsilon$ in the limit of the integration. This effectively "turns off" the interaction H_{int} at the infinite far past – presumably when inflation begins. There are two subtleties here, when compared to the usual QFT "in-out" set up. Firstly, since both states are "in" states, unitary implies that

$$T\exp\left(-i\int_{-\infty(1+i\epsilon)}^{t}dt'\ H_{int}(t')\right)|\text{in}\rangle \leftrightarrow \langle\text{in}|\left(T\exp\left(-i\int_{-\infty(1+i\epsilon)}^{t}dt'\ H_{int}(t')\right)\right)^{\dagger} \tag{124}$$

i.e. the bra is just the conjugation of the ket, and the conjugation operator also acts on the integration limits and hence flip the sign. In terms of a path integral, we then take the contour from $-\infty(1-i\epsilon)$ to t (where you evaluate the correlation) and then back to $-\infty(1+i\epsilon)$ – in other words the contour doesn't close. This also tells us that, in field theory language, when we contract operators, we don't contract into Feynman Green's Functions which are manifestly complex – instead we contract into the manifestly real *Wightman* Green's Functions. We will prove this statement later, but let's motivate it by connecting this to an old question : what is $\langle\ \rangle$?

Remember, now we are in the quantum theory, the question has become "what is the state we want to calculate the expectation value at"? This, of course depends on the physical situation. In the case of inflationary theory, we implicitly assume that the universe stays in its ground state Bunch-Davies state, i.e. where $|0\rangle$ is annihilated by the positive frequency part of the *interaction picture* fields $\delta\phi_I$ and $\delta\pi_I$. In other words, we expand the interaction picture fields in their (interaction picture) raising and lowering operators a_I and $a_I^{\dagger}$

$$\delta\phi_I(\mathbf{x},t) = \int\frac{d^3k}{(2\pi)^3}\left[a_I(\mathbf{k})u_k(t)e^{i\mathbf{k}\cdot\mathbf{x}} + a_I^{\dagger}(\mathbf{k})u_k^*(t)e^{-i\mathbf{k}\cdot\mathbf{x}}\right] = \delta\phi_I^{+}(\mathbf{x},t)+\delta\phi_I^{-}(\mathbf{x},t) \tag{125}$$

and then the states are those that obey

$$\delta\phi_I^{+}|0\rangle = \delta\pi_I^{+}|0\rangle = 0. \tag{126}$$

This makes sense, as the *power spectrum* at some time t is, with $W(t)=\zeta(\mathbf{k}_1,t)\zeta(\mathbf{k}_2,t)$ is simply a trivial contraction

$$(2\pi)^3P(k_1,t)\delta(\mathbf{k}_1+\mathbf{k}_2) = \langle 0|\zeta^I(\mathbf{k}_1,t)\zeta^I(\mathbf{k}_2,t)|0\rangle \tag{127}$$

a result which you should be familiar with. Hence, it follows that if we want to calculate any higher order correlation functions, we should sandwich Eqn. (123) around the Bunch-Davies vacuum state.

One can imagine other kinds of states – for example one that is not annihilated by the $\delta\phi_I^{+}$. These are so-called "excited states", and we will discuss the non-Gaussian signatures of such states in future lectures.

III. THE ARNOWITT-DESER-MISNER FORMALISM AND 3RD ORDER INTERACTION HAMILTONIAN

The next step in the calculation of non-Gaussianities is the computation of the interaction Hamiltonian H_{int}. In the previous section, we have kept the formalism general – the only separation we have assumed is the existence of "classical" background fields and their quantized perturbations.

We will use ADM to decompose the metric such that (i) the dynamical degrees of freedom are manifest and (ii) there exist an obvious foliation so that we have a working notion of "time". In the literature, you will see that people simply take the ADM metric

$$ds^2 = -N^2dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt) \tag{128}$$

and plug into the action, and then begin to compute. There will be (unfortunately) a lot of algebra, and this algebra is messy and hide the beautiful idea behind the decomposition – the calculation can be very mysterious too. I hope to delve a little bit deeper into this formalism and explain a little bit more, before we plunge right into the mess.

A. The Revenge of Hamilton : A Toy Model (*)

In this section, we will lay out the philosophy behind the ADM formalism. You may have seen some of the equations here in Prof. Shellard's class when he talked about the 3+1 covariant formalism, and heard the words "constraints" before. We will discuss a little bit more about this, so what follows will be a bit less mysterious.

We have talked a lot about Hamiltonians in the previous section, and in particular in the section on deriving the correlators, we have assumed that the Hamiltonian given $H(\pi_a, \phi_a)$ is constructed out of the *true degrees of freedom in the theory.* What we mean by that is that π_a and ϕ_a are all physical d.o.f. – they are measurable quantities like particles or cats. Let's forget field theory for the moment, and think about the theory of a system of M particles labeled by i. The dynamics of this system of particles is then encoded in its Hamiltonian $H(p, q)$. One can use the Legendre transformation to write down the action in Lagrangian form

$$S = \int dt\ L = \int dt\ \left(\sum_{i=1}^{i=M} p_i\dot{q}_i - H(p,q)\right). \tag{129}$$

One can then think of p_i and q_i as *independent* variables, and to get their equations of motion we can vary the action w.r.t. to p_i and q_i to get the EOM and of course we know these as the "Hamiltonian" E.O.M.

$$\dot{q}_i = \frac{\delta H}{\delta p_i}\ ,\ \dot{p}_i = -\frac{\delta H}{\delta q_i}. \tag{130}$$

Notice that the EOM for p_i and q_i are *first order* in time derivatives. An action written in the form Eqn. (129) is called the "canonical form". With such a set of equations, then we can do the happy things like quantize the theory either by promoting Eqn. (129) into a path integral or by identifying the Poisson Brackets of the canonical variables with the commutators of the quantized variables *ala.* Dirac. In particular, in the canonical form, the EOM are *unconstrained* – roughly this means that every equation we get from varying the action w.r.t. the variables are *first order* in time derivatives, hence evolve true d.o.f.

What about GR? As we know in GR, instead of p_i and q_i, information on the physical d.o.f. are encoded in the metric, with the Lagrangian being simply the Ricci scalar $R(g_{\mu\nu})$. The metric $g_{\mu\nu}$ in 4D is a symmetric 4×4 object, hence has 10 components. The action has gauge invariances due to reparameterization invariance[11] of the 3 space and 1 time coordinates.

[11] Some people like to use the even more high-brow word "diffeomorphism". A way to think about this is that "physics must be invariant under a change of coordinate system", i.e.

$$x^\mu \to x^\mu + \xi^\mu(x^\mu) \tag{131}$$

must not change physics. The metric transforms under this coordinate becomes

$$g^{\mu\nu} \to g^{\mu\nu} + 2\partial^{(\mu}\xi^{\nu)}. \tag{132}$$

A fun thing to do is to compare it to the vector potential A^μ and *its* gauge invariance $A^\mu \to A^\mu + \partial^\mu\alpha$. How do we break the gauge invariance there? What is the symmetry associated with this, when compared to the reparameterization symmetry of a metric theory?

These gauge invariance pose a problem when we want to write down the theory in canonical form – some of the variables are not true d.o.f. but instead encoded our freedom in "choosing coordinates". Operationally, when we vary the action w.r.t. with the metric, *some* of the equations are redundant – these are are the constraint equations i.e. equations which must be satisfied by the solutions but do not in themselves evolve the dynamical d.o.f. Such a system is known as "constrained Hamiltonian". These unphysical d.o.f. are often called "gauge modes".

Our aim hence is now to find a way to tease out the true dynamical d.o.f. To see how this is done, we will now show a simple system to demonstrate that reparameterization invariance is associated with a constraint – we will show this for time reparameterization and leave the space reparameterization invariance for the reader to find out him/herself. Consider again the action of the system of particles Eqn. (129). If we regard the time coordinate t as an *independent* variable, let's call it q_{M+1}, then Eqn. (129) can be written in the so-called *parameterized* form, with τ as an (affine) parameter

$$S = \int d\tau\ \tilde{L} = \int d\tau\ \left(\sum_{i=1}^{i=M+1} p_i q_i'\right). \tag{133}$$

where $q' \equiv dq/d\tau$. However, in addition to the EOM obtained from varying this equation w.r.t. p_i and q_i, to reproduce the dynamics of the original Hamiltonian $H(p,q)$, we need to add in by hand a *constraint equation*[12]

$$p_{M+1} = -H(p,q) \tag{134}$$

where the p and q in $H(p,q)$ runs only to M. Notice that since Eqn. (134) has no time derivative on p_{M+1} it doesn't really evolve the variable – this is one of the main hallmark of a constraint equation. We can explicitly introduce the constraint Eqn. (134) into the action, we can do this via the use of a Lagrange Multiplier $N(\tau)$

$$S = \int d\tau\ \tilde{L} = \int d\tau\ \left(\sum_{i=1}^{i=M+1} p_i q_i' - N(p_{M+1} + H(p,q))\right). \tag{135}$$

The form Eqn. (135) is now explicitly invariant under any time parameterization $\tilde{\tau} = \tilde{\tau}(\tau)$ – remember that $N(\tau)$ transform like $N(\tau)d\tau \rightarrow N(\tilde{\tau})d\tilde{\tau}$. We can "undo" the parameterization Eqn. (133) by inserting the solution $p_{m+1} = -H(p,q)$ back into the action to get

$$\begin{aligned} S &= \int d\tau\ \left(\sum_{i=1}^{i=M} p_i q_i' - H(p,q) q_{M+1}'\right), && (136)\\ &= \int dq_{M+1}\ \left(\sum_{i=1}^{i=M} p_i (dq_i/dq_{M+1}) - H(p,q)\right), && (137)\end{aligned}$$

and we get back the original action by identifying $q_{M+1} = t$. In other words, *if a system of dynamics is described by a Lagrangian which has some coordinate reparameterization built into it, this freedom of reparameterization is encoded as* **constraint equation(s)**. Hence, since GR has 4 reparam invariances, we expect that the EOM of GR has 4 constraint equations, and these are the so-called Hamiltonian (1) and Momentum constraints[13] Another way of saying this is that the Hamiltonian of GR is manifestly of the *constrained Hamiltonian* form Eqn. (133).

So it follows that if we want to separate out the true d.o.f. of the action from the constraints, we want to cleverly choose to write the metric in a way that it is manifestly of the form Eqn. (135) – i.e. such that the constraint equations appear explicitly as $N \times$ (constraint equation) where N is a Lagrange Multiplier.

The ADM decomposition is exactly such a decomposition of the metric such that, when plugged into the action, decomposes the action into the form Eqn. (135). This is what we will study next – applying the ADM decomposition into an inflationary action.

---

[12] For the hardcore, this is a primary first class constraint.
[13] The Hamiltonian constraint encode the time reparam and the Momentum constraints encoded the space reparam.

B. ADM in Inflationary Cosmology

We now want to specialize to the case we care about in inflationary cosmology – that of a scalar field interacting minimally with gravity, i.e. the action

$$S = \int dx^4 \sqrt{g} \left[M_p^2 \frac{R}{2} - \frac{1}{2}(\nabla\phi)^2 - V(\phi) \right] \tag{138}$$

In other words, we want to find the H_{int} for the effective degrees of freedom, i.e. the $\delta\phi_a$ of the previous section. To do this, we need to do two things. Firstly, we want to figure out what are the effective degrees of freedom, i.e. the $\delta\phi_a$ of the previous section. Secondly, we want to be able to write them in a form such that they have an explicit "time" and "space" separation so that we can construct its Hamiltonian. The steps are

- Use ADM to decompose the metric, perturb the variables and choose a gauge.
- Calculate the action (in the Lagrangian picture) to 3rd order in perturbation variables.
- Calculate the free Hamiltonian H_0 and the interaction Hamiltonian H_{int} given the 3rd order Lagrangian.

1. *Cosmological Perturbation Theory*

As you have learned in the first half of the class taught by Prof. Shellard, for small fluctuations, the gauge transformations can be linearized. The coordinate reparameterization associated with these gauge transformations are (with $i = 1, 2, 3$ being the space index)

$$t = t + \xi^0(t, x^i) \ \text{ for time reparameterization} \tag{139}$$

and

$$x^i = x^i + \xi^i(t, x^i) \ \text{ for space reparameterization.} \tag{140}$$

In linear theory then, it can contain information about 6 degrees of freedom, 2 each of which transform like spin-0 "scalar", spin-1 "vector" and spin-2 "tensor" modes. In GR as we know the only physical *propagating* degrees of freedom are the 2 spin-2 modes. Adding in matter to the action will "turn on" the other degrees of freedom, hence we expect the gravity + scalar field action in Eqn. (138) to contain two spin-2 and a single spin-0 physical propagating degrees of freedom.

For *scalar* spin-0 modes, it is acted upon the two of the gauge invariances ξ^0 and α where $\xi^i = \partial^i\alpha + \beta^i$ such that $\partial_i\beta^i = 0$, i.e. the tranverse component of ξ^i. The two remaining gauges act on the spin-1 vector modes, while the spin-2 tensor modes are gauge-invariant[14].

All this high-brow language aside, how do we compute stuff? A brute force way is to simply fix a gauge by imposing 4 constraint equations on the metric. But as we have learned from E+M, not all gauges are born equal and some are cleverer than others. There is also an additional complication in gravity, which is that there is a freedom to choose spatial foliations, hence a bad choice of foliation combined with a bad gauge can lead to unnecessary complications.

As we discussed previously, these gauge modes associated with the coordinates reparam are "built-in" into the Einstein Hilbert action as constraints. What we want is choose a decomposition of the metric, such that once inserted into the action exposes as manifest the constraints, and also provide us a suitable foliation. A famous decomposition is provided by the Arnowitt-Deser-Misner decomposition

$$ds^2 = -N^2 dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt) \tag{141}$$

where N and N^i are functions of spacetime and are called the *lapse* and the *shift* respectively[15]. Comparing it to Eqn. (132), one can see that in this form the spatial coordinate reparameterization is an explicit symmetry, while the time reparam is not so obvious since the N^i mixes up the space and time coordinates. Now plug Eqn. (141) into the action Eqn. (138) to obtain, as you have learned in Prof. Shellard's lectures

$$S = \frac{1}{2}\int d^4x\sqrt{h} \ \left[M_p^2 N R^{(3)} - 2NV + N^{-1}(E_{ij}E^{ij} - E^2) + N^{-1}(\dot{\phi} - N^i\partial_i\phi)^2 - N h^{ij}\partial_i\phi\partial_j\phi \right] \tag{142}$$

---

[14] Funfact : This is due the fact that diffeomorphisms are generated by vector fields, which themselves cannot contain spin-2 information.
[15] Sorry for the sign change compared to Prof. Shellard's lectures on N_i.

where (with K_{ij} being the extrinsic curvature)

$$E_{ij} = \frac{M_p}{2}(\dot{h}_{ij} - 2\nabla_{(i}N_{j)}) = NK_{ij} \tag{143}$$
$$E = E^i_i = h^{ij}E_{ij} \tag{144}$$

i.e. we use the 3-metric h_{ij} to raise and lower indices, with the condition that the inverse metric $h^{ij}h_{jk} = \delta^i_k$. One can think of $E_{ij}E_{ij} - E^2$ as the "kinetic term" that governs the dynamics of the evolution of the extrinsic curvatur if you like. Furthermore, notice that in the action, N and N^i do not have any (time) derivatives acting on them, and hence are Lagrange multipliers. Compare this to our prototypical constrained action in the previous section Eqn. (135) – as promised the ADM decomposition make the constraints of GR manifest.

To proceed we want to find the canonical form of the Hamiltonian by "undoing" the pre-built-in constraints. From our toy model in Section III A, it is clear how we proceed : vary the action w.r.t. with N and N^i to find the constraints

$$M_p^2R^{(3)} - 2V - h^{ij}\partial_i\phi\partial_j\phi - N^{-2}\left[E_{ij}E^{ij} - E^2 + (\dot{\phi} - N^i\partial_i\phi)^2\right] = 0 \tag{145}$$

$$\nabla_i\left[N^{-1}(E^i_j - E\delta^i_j)\right] = 0 \tag{146}$$

and then plug the algebraic solutions for N and N^i back into the action, leaving h_{ij} and ϕ as the dynamical variables. This is the cleverness of the ADM formalism : not only have we automatically get the constraints, we have also obtained an explicit space and time split, which allow us to employ much of our physical intuition and tools. In the language of Hamiltonian analysis, Eqns. (145) and (146) are Hamiltonian and Momentum constraints[16].

Before we do that, let's deal with the issue of gauge – note that the ADM decomposition is *not* a specification of gauge! If you like, a gauge specification in our Toy Model in Section III A is equivalent to choosing $q_{M+1} = t$ and a gauge transformation is equivalent to reparameterizing $\tilde{\tau}(\tau)$. Thus our freedom of choosing a gauge in GR is our freedom of in choosing the function $\tilde{\tau}$.

Since we are doing perturbation in cosmology, so let's specialize on a flat (hence $R^{(3)} = 0$) Friedman-Robertson-Walker background and consider perturbations. A choice of gauge (called *comoving gauge*) is

$$\delta\phi = 0\ ,\ h_{ij} = a^2\exp[2\zeta]\exp[\gamma]_{ij} = a^2[(1+2\zeta)\delta_{ij} + \gamma_{ij}] + \mathcal{O}(2) + ...\ ,\ \partial_i\gamma_{ij} = 0\ ,\ \gamma_{ii} = 0 \tag{147}$$

to linear order. A good exercise is to check that, in this gauge, ζ (sometimes called the Bardeen variable) is constant to all orders outside the horizon. Similarly, we also expand the shift and lapse

$$N = N^{(0)} + N^{(1)} + N^{(2)} + ...\ ,\ N_i = N_i^{(0)} + N_i^{(1)} + N_i^{(2)} + ... \tag{148}$$

The shift can further be decomposed into its spin-0 scalar and spin-1 vector components

$$N_i^{(n)} = \partial_i\psi^{(n)} + \beta_i^{(n)}. \tag{149}$$

The strategy is now to solve the constraint equations Eqns. (145) and (146) order by order in terms of the actual dynamical variables (ϕ, h_{ij}), and then plug in the solutions back into the action.

At $\mathcal{O}(0)$ order, a solution is given by $N^{(0)} = 1$ and $N_i^{(0)} = 0$, and we find that Eqn. (146) vanishes identically, and Eqn. (145) give us the Friedman equation

$$H^2 = \frac{1}{3M_p^2}\left(V + \frac{1}{2}\dot{\phi}^2\right). \tag{150}$$

Hence we have recovered the fact that the Friedman equation is really *not* an equation of motion for the scale factor. Instead, it is a Hamiltonian constraint[17].

---

[16] For the hardcore, these are first class constraints since they secretly originated from gauge freedom.

[17] This means that if one attempts to solve the equations of motion for the $a(t)$, if the initial conditions do not satisfy the friedman equations the solution never converges to the right answer – there exist papers in the literature which this mistake is made so be careful! The momentum constraint vanishes trivially because of the underlying isotropy of the zeroth order background.

Let's do this for $\mathcal{O}(1)$ order for the *scalar* modes. The traceless and tranverse conditions on γ_{ij} ensures that it cannot carry a scalar degree of freedom, hence we will drop it for now. Then the metric for the scalar d.o.f. becomes

$$h_{ij} = a^2 e^{2\zeta}\delta_{ij}\ ,\ h^{ij} = a^{-2}e^{-2\zeta}\delta^{ij} \tag{151}$$

$$\dot{h}_{ij} = 2a^2(H+\dot{\zeta})e^{2\zeta}\delta_{ij}\ ,\ \dot{h}^{ij} = -2a^{-2}(H+\dot{\zeta})\delta^{ij} \tag{152}$$

and the connection is

$$\Gamma^k_{ij} = \frac{1}{2}h^{kl}(\partial_j h_{ik} + \partial_i h_{kj} - \partial_k h_{ij}) = \delta^{kl}(\partial_j\zeta\delta_{ik} + \partial_i\zeta\delta_{jk} - \partial_k\zeta\delta_{ij}) \tag{153}$$

The extrinsic curvature is then (Exercises)

$$E_{ij} = M_p\left[a^2e^{2\zeta}\delta_{ij} - \partial_{(i}N_{j)} + (2N_{(i}\partial_{j)}\zeta - N_k\partial_k\zeta\delta_{ij})\right] \tag{154}$$

$$E^{ij} = h^{ik}h^{jl}E_{kl}a^{-4} = a^{-4}e^{-4\zeta}\delta^{ik}\delta^{jl}E_{kl} \tag{155}$$

$$E = h^{ij}E_{ij} = 3(H+\dot{\zeta}) - a^{-2}e^{-2\zeta}(\partial_k N_k + N_k\partial_k\zeta) \tag{156}$$

and thus the "kinetic terms" are

$$E^{ij}E_{ij} - E^2 = -6(H+\dot{\zeta})^2 + \frac{4e^{-2\zeta}}{a^2}(H+\dot{\zeta})(\partial_i N_i + N_i\partial_i\zeta) \tag{157}$$

$$-\frac{e^{-4\zeta}}{a^4}\left[(\partial_i N_i)^2 + 2(\partial_i N_i\zeta)^2 - \left(\partial_{(i}N_{j)} - (\partial_i N_i + N_i\partial_i\zeta)\right)^2\right] \tag{158}$$

$$R^{(3)} = -2a^{-2}e^{-2\zeta}[2\partial^2\zeta + (\partial\zeta)^2] \tag{159}$$

When exercising your index juggling kungfu, here are some helpful things to keep in mind:

- Resist the urge to expand $e^{2\zeta} = (1 + 2\zeta + 2\zeta^2 + ...)$ even though you are doing perturbation theory! Often $e^{2\zeta}$ is easier to keep track off than the series expansion.
- Also notice that there is no "2nd order" perturbation variables for the dynamical variables, i.e. no such thing as $\zeta^{(2)}$ or $\delta\phi^{(2)}$ (if we haven't chosen the $\delta\phi = 0$ gauge, we would have to keep track of the matter perturbation), ζ and $\delta\phi$ are all there is. This might be confusing for those who are used to cranking 2nd order perturbation theory equations *on the level of the equation of motion*. However, these are not present here : the constraints have "taken care of the higher order perturbations" for you if you like. This is one of the beauty of the ADM/background field theory formalism.
- $\delta^{ij}\partial_j V_i = \partial_i V_i$, i.e. resist the urge to use the delta functions to raise and lower indices. The i's are still summed over but *not contracted*.
- Corollary : $\partial^2 \equiv \delta^{ij}\partial_i\partial_j$, *not* $\partial^i\partial_i$, i.e. they are different by a factor of the space metric.
- The inverse of the time derivative of the spatial metric is h_{ij} has its sign flipped $h^{im}h_{jm} = \delta^i_j$, $\dot{h}^{im}h_{jm} = -h^{im}\dot{h}_{jm}$.

Also note that to first order $\nabla_i N_j \to \nabla_i(N_j^{(0)} + N_j^{(1)} + ...)$ and since $N_i^{(0)} = 0$ by the zeroth order constraints, and $\nabla_i N_j \to \partial_i N_j^{(1)} - ...$, i.e. the christoffel symbols are of $\mathcal{O}(\zeta)$. At higher order this is not true, so one has to be careful. Plugging these into Eqn. (145), we obtain (you can use the trick that since we are dealing with scalar d.o.f., the spin-1 component β_i will trivially vanish), we get

$$\nabla_i(HN^{(1)} - \dot{\zeta}) = 0 \tag{160}$$

and

$$a^{-2}\partial^2\psi^{(1)} = -a^{-2}\frac{\partial^2\zeta}{H} + \frac{\dot{\phi}^2}{2H^2}\dot{\zeta} \tag{161}$$

with solutions (Exercise : prove this is true by showing that they solve Eqns. (145) and (146))

$$N^{(1)} = \frac{\dot{\zeta}}{H}\ ,\ N_i^{(1)} = \partial_i\psi \text{ where } \psi = -\frac{\zeta}{H} + a^2\frac{\dot{\phi}^2}{2H^2}\partial^{-2}\dot{\zeta}. \tag{162}$$

(For those who are following Maldacena's paper, note the difference in the indexing of his definition for N_i, which leads to extra factors of a^2.)

2. The Quadratic Action

Before we compute the 3rd order action we need to calculate the non-Gaussianities, let's warm up by doing the quadratic action. Plugging in N and N_i into the action, and expanding it to second order we get

$$S = \frac{1}{2}\int dt\ ae^{\zeta}\left(1+\frac{\dot{\zeta}}{H}\right)(-4\partial^2\zeta - 2(\partial\zeta)^2 - 2Va^2e^{2\zeta}) \tag{163}$$

$$+e^{3\zeta}a^3\left(1-\frac{\dot{\zeta}}{H}+\frac{\dot{\zeta}^2}{H^2}\right)[-6(H+\dot{\zeta})^2+\dot{\phi}^2+4a^{-2}e^{-2\zeta}(H+\dot{\zeta})(\partial_i\psi\partial_i\zeta+\partial^2\psi)] \tag{164}$$

where we have used integration by parts and dropping 3rd orders terms to get rid of the $(\partial_{(i}\partial_{j)})^2$ terms in $E^{ij}E_{ij}-E^2$.

Now you will do some integration by parts and toss away the boundary terms (more later), and use the background equations of motion $3H^2 = V + (1/2)\dot{\phi}^2$ and $\ddot{a}/a + 2H^2 = V$ to reduce the equation for the quadratic terms to (Exercise)

$$S_2 = \int dtd^3x\ 4\frac{\dot{\phi}^2}{H^2}\left[\frac{a^3}{2}\dot{\zeta}^2 - \frac{a}{2}(\partial\zeta)^2\right] \tag{165}$$

Also, you can use the relation Eqn. (162) to convert a $\dot{\zeta}$ to a ∂^2(stuff) term which will make some integration by parts easier.

If you plug in $v = \zeta a\sqrt{2\epsilon} \equiv z\zeta$ to the 2nd order term, you will find that this reduces to the Mukhanov-Sasaki action which you might have seen before (with τ being conformal time now)

$$S_{ms} = \frac{1}{2}\int d\tau d^3x\left(v'^2 - (\partial v)^2 - \frac{z''}{z}v^2\right). \tag{166}$$

Again some hints :

- "do a lot of integration by part" (J. Maldacena). The whole calculation is an exercise in doing a lot of integration parts, which can be a cyclic exercise if we do not have a target "form". The rule of thumb is to eliminiate higher *time* derivatives on variables in the action by "spreading" them around, i.e. terms like vv'' can be recast as v'^2 etc. Note that the number of derivatives are always "conserved" so use this as an aid to keep track of things.
- You can do this calculation in various ways. The most straightforward is to expand e^{ζ} to 2nd order in ζ, and then take the 2nd order terms and then integrate by parts some of the terms to get into the form of Eqn. (165). If you decide to keep $e^{2\zeta}$ general, you will find that linear terms of $\dot{\zeta}$ will end up multiplying the zero order constraint as expected and the zeroth order terms cancel via the background EOM.

All this algebra to write it in the form of Eqn. (165) makes it clear that the action itself is proportionate to $\dot{\phi}^2/H^2 \equiv -2\epsilon$, i.e.the slow roll parameter. As $\epsilon \to 0$, i.e. as we approach dS space, the action vanishes – consistent with the fact that in pure dS space, the scalar mode is simply a gauge mode and hence vanish by the background constraint.

Now, a completely valid question is : why do we have to integrate by parts and toss away the boundary terms? The simple answer is : we don't have to. In fact, it is a tricky business to toss away boundary terms if we are really interested in the value of the *action* (such as when calculating the Euclidean bounces and tunneling rates) and not simply the equation of motion[18]. In the latter case, the boundary terms do not contribute when you vary the action with respect to the field ζ. However, especially in a physical model like inflation, the boundary is at the time t_e when we evaluate ζ and the boundary term do not always vanish as $\dot{\zeta}(t_e)$ is not always zero. Having said all that, to make contact with literature, we will follow the standard derivation – the usual argument being that the Mukhanov-Sasaki action Eqn. (166) looks "nice" like a scalar field with a time dependent mass, so we can quantize it[19].

---

[18] Think of GR : when we vary the Einstein-Hilbert action to get the Einstein Equations, we toss away boundary terms *which are not formally zero* since there is no reason that the derivatives of $\delta(\partial_\sigma g_{\mu\nu})$ to be zero. These are additional constraints. To "fix" this deficiency, we can add the Hawking-Gibbons-York term into the action which formally cancel the boundary term which consist of derivatives. But for most of the rest of the world which cares only about the equations of motion, we don't have to worry about this formal difference.

[19] Even though we can equally well quantize the theory without all the integration by parts.

3. 3rd order action

Now, we want to compute things to the next order in the action. As we mentioned above, we only need up to 1st order in the constraints N and N_i so we plug Eqn. (162) into the action, and expand to 3rd order

$$S_3 = \int ae^{\zeta}\left(1+\frac{\dot{\zeta}}{H}\right)(-4\partial^2\zeta - 2(\partial\zeta)^2 - 2Va^2e^{2\zeta}) \tag{167}$$

$$+e^{3\zeta}a^3\left(1-\frac{\dot{\zeta}}{H}+\frac{\dot{\zeta}^2}{H^2}\right)[-6(H+\dot{\zeta})^2+\dot{\phi}^2+4a^{-2}e^{-2\zeta}(H+\dot{\zeta})(\partial_i\psi\partial_i\zeta+\partial^2\psi) \tag{168}$$

$$-a^{-4}e^{-4\zeta}(\partial^2\psi)^2+a^{-4}e^{-4\zeta}\left((\partial_i\partial_j\psi)^2-2(\partial_i\psi\partial_j\zeta+\partial_j\psi\partial_i\zeta)(\partial_i\partial_j\psi)\right)]. \tag{169}$$

As before, we also want to write it in the form where all the terms are explicitly proportional to the slow roll parameters as what we did to the second order action Eqn. (165). This is a lot of algebra (sorry) involving integration by parts, which we find at the end

$$S_3 = \int dtd^3x\ \left(a^3\epsilon^2\zeta\dot{\zeta}^2 - 2a\epsilon^2\dot{\zeta}(\partial\zeta)(\partial\psi)+a\epsilon^2\zeta(\partial\zeta)^2\right. \tag{170}$$

$$\left.+\frac{a}{2}\epsilon\dot{\eta}\zeta^2\dot{\zeta}+\frac{1}{2}\frac{\epsilon}{a}\partial\zeta\partial\psi\partial^2\psi+\frac{\epsilon}{4a}\partial^2\zeta(\partial\psi)^2+2f(\zeta)\frac{\delta L_2}{\delta\zeta}\right) \tag{171}$$

where L_2 is the integrand of the quadratic action Eqn. (165) $S_2 = \int L_2$

$$\frac{\delta L_2}{\delta\zeta} = (a^3\epsilon\dot{\zeta})^{\cdot} - \epsilon a\partial^2\zeta \tag{172}$$

is the first order equation of motion, while

$$f(\zeta) = \frac{\eta}{4}\zeta^2+\frac{1}{H}\zeta\dot{\zeta}+\frac{1}{4a^2H^2}\left[-(\partial\zeta)^2+\partial^{-2}(\partial_i\partial_j(\partial_i\zeta\partial_j\zeta))\right]+\frac{1}{2a^2H}\left[(\partial\zeta)(\partial\psi)-\partial^{-2}(\partial_i\partial_j(\partial_i\zeta\partial_j\psi))\right]. \tag{173}$$

This is Maldacena's original derivation, copied countless times since his ground breaking 2002 paper. He proceeded to remove the terms proportional to $\delta L_2/\delta\zeta$ by doing a field redefinition

$$\zeta_n = \zeta - f(\zeta). \tag{174}$$

You can show that by plugging in Eqn. (174) into Eqn. (171) above that this field redef kills off the term proportional to the equation of motion. We will follow his derivation here.

The important point here is that while ζ does not evolve outside the horizon, ζ_n does. This means that $\dot{\zeta}_n$ does not vanish at the point when we evaluate the 3-pt – i.e. at the *boundary* $t \to 0$. We have to take this into account when we evaluate the 3-pt correlation function, i.e. given *any field redef* $\zeta = \zeta_n + \alpha\zeta_n^2 + ...$, the first order correction to the three point is

$$\langle\zeta(\mathbf{x}_1)\zeta(\mathbf{x}_2)\zeta(\mathbf{x}_3)\rangle = \langle\zeta_n(\mathbf{x}_1)\zeta_n(\mathbf{x}_2)\zeta_n(\mathbf{x}_3)\rangle + 2\alpha\left[\langle\zeta_n(\mathbf{x}_1)\zeta_n(\mathbf{x}_2)\rangle\langle\zeta_n(\mathbf{x}_1)\zeta_n(\mathbf{x}_3)\rangle + \text{sym}\right] + ... \tag{175}$$

For now, we will drop the subscript n to keep notation simplicity, bearing in mind that in the final calculation we have to add in the extra terms in Eqn. (175).

C. Calculating the Interaction Hamiltonian

Given the action, and the degree of freedom ζ, we know want to calculate the interaction Hamiltonian H_{int}. In the literature, you will often see the substitution $\mathcal{H}_{int} = -\mathcal{L}_{int}$, where $\mathcal{L}_{int}$ is the part of the Lagrange density that is 3rd order and higher. This is not incorrect, but it is not true in general – the reason is that the definition of the canonical momentum $\pi = \partial\mathcal{L}/\partial\dot{\zeta}$ is changed when the 3rd order terms contain $\dot{\zeta}$ terms. In our case though, the substitution $\mathcal{H}_{int} = -\mathcal{L}_{int}$ turns out to induce higher order (in the perturbations) corrections as follows.

Recall that $\tilde{H}$ is the sum of all second order and higher terms, let's combine the second action Eqn. (165) and third order action (171), so we get the Lagrange density $\mathcal{L}$

$$S_2 + S_3 = \int d^4x\ \tilde{\mathcal{L}}. \tag{176}$$

The canonical momentum is as usual

$$\pi = \frac{\partial \tilde{\mathcal{L}}}{\partial \dot{\zeta}} \tag{177}$$

then the Hamiltonian density $\tilde{H}$ is

$$\tilde{\mathcal{H}} = \pi\dot{\zeta} - \tilde{\mathcal{L}} \equiv \mathcal{H}_0 + \mathcal{H}_{int} \tag{178}$$

where we have split the full Hamiltonian $\tilde{\mathcal{H}}$ into a free part $\mathcal{H}_0$ which consists of 2nd order terms and an interaction part. Now, if $\tilde{\mathcal{L}}$ has terms higher order in $\dot{\zeta}$, it is usually not possible to find the invert $\pi(\dot{\zeta})$. To proceed, we assume that the perturbation canonical variables $|\zeta| \sim |\pi|$ are of the same order, and then *iteratively* calculate the Hamiltonian[20], and we will find that to 3rd order, the substitution $\mathcal{H}_{int} = -\mathcal{L}_{int}$. The best way to see this through an example.

Consider the Lagrange density

$$\mathcal{L} = \frac{1}{2}\dot{\zeta}^2 - V(\zeta, \partial\zeta) + \alpha\dot{\zeta}^3 \tag{179}$$

where α is some dimensionful constant (which may or may not be small), and we have collected all the terms that have no time derivatives into V. The canonical momentum is

$$\pi = \dot{\zeta} + 3\alpha\dot{\zeta}^2 \tag{180}$$

so

$$\dot{\zeta} = \pi - 3\alpha\dot{\zeta}^2 \tag{181}$$

and substituting this into

$$\begin{aligned}
\tilde{\mathcal{H}} &= \pi\dot{\zeta} - \tilde{\mathcal{L}} && (182)\\
&= \pi^2 - 3\alpha\dot{\zeta}^3 - \frac{1}{2}(\pi^2 + 9\alpha^2\dot{\zeta}^4 - 6\alpha\pi\dot{\zeta}^2) - \alpha\pi^3 + V(\zeta, \partial\zeta) + \mathcal{O}(4) && (183)\\
&= \frac{1}{2}\pi^2 + V(\zeta, \partial\zeta) - \alpha\pi^3 + \mathcal{O}(4) && (184)\\
&\equiv \mathcal{H}_0 + \mathcal{H}_{int} && (185)
\end{aligned}$$

where the free hamiltonian density is (V_0 is the quadratic part of V)

$$\mathcal{H}_0 = \frac{1}{2}\pi^2 + V_0(\zeta, \partial\zeta) \tag{186}$$

and the interaction hamiltonian is (V_{int} is the 3rd order part of V)

$$\mathcal{H}_{int} = -\alpha\pi^3 - V_{int}(\zeta, \partial\zeta) + \mathcal{O}(4) = -\mathcal{L}_{int} + \mathcal{O}(4) \tag{187}$$

One can show that the substitution $\mathcal{H}_{int} = -\mathcal{L}_{int}$ is fine for general FRW background as long as the perturbative expansion in ζ is valid, up to 3rd order. This gets trickier when we consider 4th order terms – but it turns out in inflation the corrections will be proportional to the slow roll parameters so the subtitution $\mathcal{H}_{int} = -\mathcal{L}_{int}$ is still OK. It will fail in a non-inflationary background though, so one has to consider extra terms in general when going to 4th order and beyond.

Harking back to our original 2nd and 3rd order action Eqns. (165) and (171), we then have

$$\mathcal{H}_0 = 2\epsilon\left[\frac{1}{2}\pi^2 - \frac{a}{2}(\partial\zeta)^2\right] \tag{188}$$

---

[20] This statement – that we are working in perturbation theory and that the higher order derivative terms can be solved iteratively, is non-trivial. In genereal, a general higher-order derivative theory suffers from Ostragradski Instability, which is that there exist negative energy *classical* states which will render the theory unstable. Assuming that the higher derivative terms can be iteratively calculated is tantaumount to assuming that the theory is only valid in perturbation theory.

and

$$\mathcal{H}_{int} = -\Big[a^3\epsilon^2\zeta\pi^2 + 2a\epsilon^2\zeta(\partial\zeta)^2 \tag{189}$$

$$-2a\epsilon\pi(\partial\zeta)(\partial\chi) + \frac{a^3}{2}\epsilon\dot{\eta}\zeta^2\pi + \frac{1}{2}\frac{\epsilon}{a}\partial\zeta\partial\chi\partial^2\chi + \frac{\epsilon}{4a}\partial^2\zeta(\partial\chi)^2\Big]. \tag{190}$$

Finally, one last sleight of hand. Recall that we want to construct the Hamiltonian out of the Interaction picture fields, so we replace $\zeta \to \zeta_I$ and $\pi \to \pi_I$ where

$$\pi_I = \frac{\partial\mathcal{H}_0}{\partial\pi}\bigg|_{\pi=\pi_I} = \dot{\zeta}_I. \tag{191}$$

D. Evaluating the Correlation, Wick's theorem

We now have all the pieces required to calculate the correlation Eqn. (82)

$$\langle W(t)\rangle = \langle\left[\bar{T}\exp\left(i\int_{-\infty(1-i\epsilon)}^{t} dt'\ H^I_{int}(t')\right)\right] W^I(t)\left[T\exp\left(-i\int_{-\infty(1+i\epsilon)}^{t} dt'\ H^I_{int}(t')\right)\right]\rangle. \tag{192}$$

Firstly, all the quantities are constructed out of the interaction picture field ζ_I and the background quantities. Let's quantize ζ_I

$$\zeta_I(\mathbf{x},t) = \int\frac{d^3k}{(2\pi)^3}\left[a_I(\mathbf{k})u_k(t)e^{i\mathbf{k}\cdot\mathbf{x}} + a_I^\dagger(\mathbf{k})u_k^*(t)e^{-i\mathbf{k}\cdot\mathbf{x}}\right] = \zeta_I^+(\mathbf{x},t) + \zeta_I^-(\mathbf{x},t) \tag{193}$$

where the raising and lowering operators obey the commutator

$$[a_I(\mathbf{k}), a_I(\mathbf{k}')^\dagger] = (2\pi)^3\delta(\mathbf{k}-\mathbf{k}'). \tag{194}$$

The usual Bunch-Davies vacuum state is one that is annihilated by the positive frequency modes a_I

$$\zeta_I^+|0\rangle = 0. \tag{195}$$

Meanwhile, the mode functions $u_k(\tau)$ solve the *linear* order Mukhanov-Sasaki equation of motion

$$u_k'' + \left(k^2 - \frac{z''}{z}\right)u_k = 0 \tag{196}$$

where $z = a\sqrt{2\epsilon}$, with corresponding "Bunch-Davies" initial conditions, with $v_k = zu_k$

$$v_k(\tau_0) = \sqrt{\frac{1}{2k}}\ ,\ v_k'(\tau_0) = -i\sqrt{\frac{k}{2}}. \tag{197}$$

During inflation, the mode functions u_k can be approximated as the so-called "de Sitter modes"

$$u_k(\tau) = \frac{H}{\sqrt{4\epsilon k^3}}(1-ik\tau)e^{ik\tau} \tag{198}$$

Then, it is easy to show that the two point correlation function of ζ_I, at the end of inflation when $\tau \to 0$

$$\langle 0|\zeta_I(\mathbf{k}_1)\zeta_I(\mathbf{k}_2)|0\rangle = \langle 0|(\zeta_I^+(\mathbf{k}_1) + \zeta_I^-(\mathbf{k}_1))(\zeta_I^+(\mathbf{k}_2) + \zeta_I^-(\mathbf{k}_2))|0\rangle = \frac{H^2}{2(2\pi)^3\epsilon k_1^3}\delta(\mathbf{k}_1 - \mathbf{k}_2). \tag{199}$$

All very familiar.

As usual in QFT, we will expand in the interaction Hamiltonian, so we use the Dyson series

$$T\exp\left(-i\int_{-\infty(1+i\epsilon)}^{t} dt'\ H^I_{int}(t')\right) = 1 - i\int_{-\infty(1+i\epsilon)}^{t} dt'\ H^I_{int}(t') + \frac{i^2}{2!}\int_{-\infty(1+i\epsilon)}^{t} dt'\int_{-\infty(1+i\epsilon)}^{t'} dt''\ H^I_{int}(t')H^I_{int}(t'') + ... \tag{200}$$

and the anti-time ordering is simpy the conjugate. Pay careful attention to the limits of the integration.

Each order in H_{int} is a vertex, and carries a time integral t and the space integral (that H^I_{int} comes equipped with) which in Fourier space enforces momentum conservation. At a single vertex we have two terms,

$$\langle W(t)\rangle = \left\langle \left(1 + i\int_{-\infty(1-i\epsilon)}^{t} dt'\ H^I_{int}(t')\right) W^I(t) \left(1 - i\int_{-\infty(1+i\epsilon)}^{t} dt'\ H^I_{int}(t')\right)\right\rangle \tag{201}$$

expanding, and using the fact that the 3-pt correlation of the interaction picture fields vanish (can you see why?), the next to leading order term is

$$\begin{aligned}\langle W(t)\rangle &= \left\langle \left(i\int_{-\infty(1-i\epsilon)}^{t} dt'\ H^I_{int}(t')\right) W^I(t) - W^I(t)\left(i\int_{-\infty(1+i\epsilon)}^{t} dt'\ H^I_{int}(t')\right)\right\rangle \\ &= \mathrm{Re}\langle\left[-2iW^I(t)\int_{-\infty(1+i\epsilon)}^{t} dt'\ H^I_{int}(t')\right]\rangle \end{aligned}\tag{202}$$

where the 2nd line we used Hermiticity.

For those steep in QFT lore, it is clear that the next step is then to plug in the interaction Hamiltonian Eqn. (190), expand $W(t)$ and H_{int} in terms of interaction picture fields, and then *contract* them into a product of Green's Functions. In particular, in "in-out" calculation, we use Wick's Theorem, given $\phi_1 \equiv \phi(x_1)$

$$\langle 0|T\{\phi_1\phi_2...\phi_n\}|0\rangle =: \phi_1\phi_2...\phi_n : + \text{all possible contractions} \tag{203}$$

where a contraction is defined to be

$$\overline{\phi_1\phi_2} \equiv D_F(x_1 - x_2) \tag{204}$$

where D_F is the Feynman's Greens Function. On the other hand, what we want is a similar calculational technique for correlation functions. Say , in $k-space$, where $(\zeta_i = \zeta(\mathbf{k}_i, t_i))$

$$\langle 0|\zeta_1\zeta_2...\zeta_n|0\rangle =: \zeta_1\zeta_2...\zeta_n : + \text{all possible contractions} \tag{205}$$

as rule for the contraction. In both in-out and in-in cases, what we want is to commute all the lowering (raising) operators to the right (left) so as to annhilate the zero ket (bra). So, let's define *normal ordering* as usual $: \zeta_1\zeta_2...\zeta_n := \zeta_1^-\zeta_2^-....\zeta_n^+$. Then the contraction can be defined as follows[21]

$$\overline{\zeta_1\zeta_2} \equiv [\zeta_1^+, \zeta_2^-], \tag{206}$$

but we know that

$$\langle\zeta_1(t_1)\zeta_2(t_2)\rangle = \langle[\zeta_1^+(t_1), \zeta_2^-(t_2)]\rangle = (2\pi)^3 u(\mathbf{k}_1, t_1)u^*(\mathbf{k}_2, t_2)\delta(\mathbf{k}_1 - \mathbf{k}_2). \tag{207}$$

In other words, we contract into the absolute value square of the *linear mode function.* Now, one can prove that Wick's Theorem Eqn. (205) is obeyed – we will do this below – with the two major differences compared to in-out

- There is no time ordering operator inside. Remember that each vertex in the in-in calculation has a time integral associated with it, so even though they are equal time commutators you usually have to carry around t in the mode functions and let the integrals sort out the time dependence. Note that the time dependence occur in the *mode function*, not the raising/lowering operator.
- The contractions are into Wightman Function, or more prosaicly, the absolute square of the linear mode function.

On the other hand, similar to QFT in-out, we drop all the contractions which do not "fully" contract.

This ends our discussion on the calculation of the free and interaction Hamiltonians H_0 and H_{int} using the ADM formalism for cosmological perturbations. In the next section, we will use all our accumulated knowledge to calculate some non-Gaussianities at last.

---

[21] Compare this to the time ordering dependent contraction in the usual in-out picture.

E. Wick's Theorem

Let's complete our discussion here by proving Wick's Theorem. This is almost a symbol-by-symbol proof copied from Peskin + Schroeder.

First, let's use Q instead of ζ, because it's easier for me to tex. Hence the contraction is defined to be

$$\overline{Q_1 Q_2} \equiv [Q_1^+, Q_2^-], \tag{208}$$

where

$$Q(\mathbf{x},t) = \int \frac{d^3k}{(2\pi)^3} \left[a_I(\mathbf{k}) u_k(t) e^{i\mathbf{k}\cdot\mathbf{x}} + a_I^\dagger(\mathbf{k}) u_k^*(t) e^{-i\mathbf{k}\cdot\mathbf{x}} \right] = \int \frac{d^3k}{(2\pi)^3} \left[Q^+(\mathbf{k},t) + Q^-(\mathbf{k},t) \right] e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{209}$$

and the subscripts $1, 2...i$ denote some dependence on $\mathbf{k}$ (or $\mathbf{x}$ if you like – the Theorem does not depend on the labeling though the result of the contraction does).

It's clear that for $n = 2$, Wick Theorem's is obeyed. Now we want to prove it by induction. Assume that it is obeyed for $n = i - 1$, then

$$\begin{aligned}
Q_1 Q_2 ... Q_i &= (Q_i^+ + Q_1^-)(: Q_1 Q_2 .... Q_i + \text{all contractions, no } Q_1 :) \\
&= Q_i^+(: Q_1 Q_2 .... Q_i + \text{all contractions, no } Q_1 :) + Q_1^-(: Q_1 Q_2 .... Q_i + \text{all contractions, no } Q_1 :) \\
&= [Q_1^+, (: Q_1 Q_2 .... Q_i + \text{all contractions, no } Q_1)] + (: Q_1 Q_2 .... Q_i + \text{all contractions, no } Q_1 :) Q_1^+ \\
&\quad + Q_1^-(: Q_1 Q_2 .... Q_i + \text{all contractions, no } Q_1 :).
\end{aligned}$$

Let's look at the following term, which is representative of all the other terms. We want to commute Q_1^+ through to the other side, viz

$$\begin{aligned}
[Q_1^+, : Q_2 ... Q_i :] &= [Q_1^+, Q_2^- ... Q_i^- Q_2^+ ... Q_i^+] + \text{all permutations} \\
&= [Q_1^+, Q_2^-] Q_3^- ... Q_i^- Q_2^+ ... Q_i^+ + Q_2^- Q_1^+ Q_3^- ... Q_i Q_2^+ ... Q_i^+ + \text{all permutations} \\
&= \sum_{k=2}^{k=i} Q_2^- ... Q_{k-1} [Q_1^+, Q_k^-] Q_{k+1}^- ... Q_i^- Q_2^+ ... Q_i^+ + \text{all permutations} \\
&= \text{terms with one contraction inv. one } Q_1
\end{aligned} \tag{210}$$

Now we can do similar calculations for the other contractions which has no Q_1, and so on. Adding them up we then get

$$Q_1 Q_2 ... Q_i =: Q_1 Q_2 ... Q_i : + : \text{terms with one contraction inv. one } Q_1 + \text{terms with one contraction inv. no } Q_1 : \tag{211}$$

and hence we are done.

IV. NON-GAUSSIANITIES, SHAPES AND FORMS

For this section, all the fields are Interaction picture fields, so we will drop the subscript I from now on.

A. The Slow-Roll Single Scalar Field Model : Your First Primordial Non-Gaussianity

We have spent the last three chapters talking about formalism, and now finally we will put everything together to calculate some non-Gaussianity by evaluating Eqn. (202). We will consider a single scalar field inflationary model, with slow-roll parameters

$$\epsilon = -\frac{\dot{H}}{H^2} \ , \ \eta = \frac{\dot{\epsilon}}{\epsilon H} \tag{212}$$

Furthermore, we will assume that $\epsilon \ll 1 \approx$ const, hence $H \approx$ const. Note that since $\epsilon(\phi)$ and $\eta(\phi)$ are implicitly functions of the scalar field, this ansatz defines the classical background solution for the scalar inflaton field. This means that the background solution for the metric is in near de Sitter space – we can safely assume that it is in dS space so the solution for the metric is given by

$$a \approx -\frac{1}{H\tau} \ , \ \ -\infty < \tau < 0. \tag{213}$$

Also, as the mode functions are oscillatory in conformal time τ space, we also convert the dots into primes $d/dt = a^{-1}d/d\tau$.

Now, each term in the interaction Hamiltonian Eqn. (190) contributes some *mode-dependent* amount of non-Gaussianity. This mode-dependence is called *shape*, and ultimately will lead to form of $F(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$. Notice that all the terms in the interaction Hamiltonian is of the form

$$\mathcal{H}_{int} \sim \text{slow roll parameters} \times \zeta^3 \tag{214}$$

i.e. the couplings are functions of the slow roll parameters. In a slow-roll model, these are small, e.g. $\epsilon \sim 0.01$. Since non-Gaussianities are sourced by these interaction terms, we expect them to be small as we have asserted way back in Section I. But enough words! Let's calculate something now.

The integration of Eqn. (202) is straightforward, just insert any of the interaction terms in Eqn. (190) into Eqn. (202), Wick contract them into their mode functions using Eqn. (205,207), integrate over the momenta and then integrate over the time τ. Since we are assuming that $H \approx$ const, we can use the de Sitter mode functions Eqn. (198).

Using these, let's calculate the non-Gaussianity generated by one of the terms $H^A_{int} = \int d^3x\ a\epsilon^2\zeta(\mathbf{x},\tau)(\zeta'(\mathbf{x},\tau))^2$ to show you how the mechanic work. Specifically, we want to evaluate the 3-pt at time $\tau_e \to 0$, such that $k\tau_e \ll 0$ i.e. they are far outside the Horizon hence has "frozen out"

$$\langle\zeta(\mathbf{k}_1,0)\zeta(\mathbf{k}_2,0)\zeta(\mathbf{k}_3,0)\rangle = \mathrm{Re}\left[-2i\zeta(\mathbf{k}_1,0)\zeta(\mathbf{k}_2,0)\zeta(\mathbf{k}_3,0)\int_{-\infty(1+i\epsilon)}^{\tau} d\tau' d^3x\ a^2\epsilon^2\zeta(\mathbf{x},\tau')\zeta'(\mathbf{x},\tau')\zeta'(\mathbf{x},\tau')\right] \tag{215}$$

We expand the fields inside the interaction Hamiltonian in their respective modes $\zeta(\mathbf{q}_1), \zeta(\mathbf{q}_2), \zeta(\mathbf{q}_3)$. We then contract the modes outside with the modes inside of the time integral – it's easy to show that the other contractions do not contribute[22], leading to

$$\begin{aligned}\langle\zeta(\mathbf{k}_1,0)\zeta(\mathbf{k}_2,0)\zeta(\mathbf{k}_3,0)\rangle &= \mathrm{Re}\left[-2i\zeta(\mathbf{k}_1,0)\zeta(\mathbf{k}_2,0)\zeta(\mathbf{k}_3,0)\int \frac{d^3q_1}{(2\pi)^3}\frac{d^3q_1}{(2\pi)^3}\frac{d^3q_1}{(2\pi)^3}\right.\\ &\quad \left.\times\int_{-\infty(1+i\epsilon)}^{\tau} d\tau' d^3x\ a^2\epsilon^2\zeta(\mathbf{q}_1,\tau)\zeta'(\mathbf{q}_2,\tau)\zeta'(\mathbf{q}_3,\tau)\right] e^{-i(\mathbf{q}_1+\mathbf{q}_2+\mathbf{q}_3)\cdot\mathbf{x}}\\ &= 2\times\mathrm{Re}\left[-2iu_{k_1}(0)u_{k_2}(0)u_{k_3}(0)\int\frac{d^3q_1}{(2\pi)^3}\frac{d^3q_1}{(2\pi)^3}\frac{d^3q_1}{(2\pi)^3}\int_{-\infty(1+i\epsilon)}^{\tau} d\tau' d^3x\ a^2\epsilon^2\right.\\ &\quad \left.\times u^*_{q_1}(\tau')\frac{d}{d\tau}u^*_{q_2}(\tau')\frac{d}{d\tau}u^*_{q_3}(\tau')(2\pi)^9\delta(\mathbf{k}_1-\mathbf{q}_1)\delta(\mathbf{k}_2-\mathbf{q}_2)\delta(\mathbf{k}_3-\mathbf{q}_3)e^{-i(\mathbf{q}_1+\mathbf{q}_2+\mathbf{q}_3)\cdot\mathbf{x}}\right]\\ &\quad +1\to 2+1\to 3\end{aligned} \tag{216}$$

[22] In QFT language, they are "vacuum fluctuations", which in in-out contribute a phase, but here in in-in, vanish.

where the extra 2 in front is due to the two identical terms we get from a choice of contracting to either of the two ζ' terms in the interaction Hamiltonian. The last two terms come from the symmetry with labeling the non-primed ζ term. Now using

$$\frac{d}{d\tau}u_k(\tau) = \frac{H}{\sqrt{4\epsilon k^3}}k^2\tau e^{ik\tau}, \tag{217}$$

and then integrating over the $\int d^3x\ e^{-i(\mathbf{q}_1+\mathbf{q}_2+\mathbf{q}_3)\tau} = (2\pi)^3\delta(\mathbf{q}_1+\mathbf{q}_2+\mathbf{q}_3)$, followed by integrating over the $k's$ to enforce momentum conservation, we get (using $a = -1/(H\tau)$ for near dS space expansion)

$$\begin{aligned}\langle\zeta(\mathbf{k}_1,0)\zeta(\mathbf{k}_2,0)\zeta(\mathbf{k}_3,0)\rangle &= 2\times\mathrm{Re}\left[-2i\frac{H^6}{(4\epsilon)^3}\frac{1}{(k_1k_2k_3)^3}\int_{-\infty(1+i\epsilon)}^{\tau}d\tau'\ \frac{1}{(H\tau')^2}\epsilon^2\right.\\ &\left.\times(k_2k_3)^2\tau'^2(1+ik_1\tau')e^{-i(k_1+k_2+k_3)\tau'}(2\pi)^3\delta(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\right] + \mathrm{sym.}\end{aligned} \tag{218}$$

The τ^2 from a^2 neatly cancels the τ^2 from the u's which it has to from dimensional analysis. Doing the time integral, we get, using $K = k_1+k_2+k_3$

$$\int_{-\infty(1+i\epsilon)}^{\tau}d\tau'(k_2k_3)^2(1+ik_1\tau')e^{-iK\tau'} = (k_2k_3)^2\left(\frac{ik_1}{K^2}+\frac{i}{K}-\frac{k_1\tau'}{K}\right)e^{-iK\tau'}\Big|_{-\infty(1+i\epsilon)}^{\tau} \tag{219}$$

and the $i\epsilon$ prescription cancels out the contributions at far infinity as promised. This leaves us with a wholy imaginary part, and combining this with the i in front of the total integral, leaves us with a manifestly real quantity as expected.

$$\langle\zeta(\mathbf{k}_1,0)\zeta(\mathbf{k}_2,0)\zeta(\mathbf{k}_3,0)\rangle = \frac{H^4}{16\epsilon}\frac{1}{(k_1k_2k_3)^3}(2\pi)^3\delta(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)(k_2k_3)^2\left(\frac{1}{K}+\frac{k_1}{K^2}+1\to 2+1\to 3\right) \tag{220}$$

which is of course of the promised form with the shape

$$F(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) = \frac{H^4}{16\epsilon}\frac{1}{(k_1k_2k_3)^3}(2\pi)^3(k_2k_3)^2\left(\frac{1}{K}+\frac{k_1}{K^2}+1\to 2+1\to 3\right). \tag{221}$$

You will be ask to calculate the contributions from the other terms in the Example sheet. The only subtle point to note is that space derivatives brings down a factor of $\partial\to\mathbf{k}$, so $(\partial\zeta)^2\to\mathbf{k}_1\cdot\mathbf{k}_2\zeta^2$.

The final answer for the entire Hamiltonian, after all the algebra, is

$$\langle\zeta(\mathbf{k}_1,0)\zeta(\mathbf{k}_2,0)\zeta(\mathbf{k}_3,0)\rangle = \tag{222}$$

$$(2\pi)^3\delta(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\frac{H^4}{M_p^4}\frac{1}{(k_1k_2k_3)^3}\frac{1}{4\epsilon^2}\left[\frac{\eta}{8}\sum k_i^3+\frac{\epsilon}{8}\left(-\sum k_i^3+\sum_{i\neq j}k_ik_j^2+\frac{8}{K}\sum_{i>j}k_i^2k_j^2\right)\right] \tag{223}$$

where we have dropped the $\mathcal{H}_{int} = \dot{\eta}\epsilon\zeta^2\dot{\zeta}$ interaction term as $\dot{\eta}$ can be neglected in the slow roll approximation. Having said that, this term turns out to be extremely crucial in models of inflation which generates large non-Gaussianities, as we will see later.

B. The Squeezed Limit and the Local Form, and a Consistency Condition

Now we are ready to prove the statement that $f_{NL}^{local}\approx\epsilon$ that we asserted way back in the first lecture. Recall Eqn. (48) again

$$\langle\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi(\mathbf{k}_3)\rangle = f_{NL}^{local}\left[2(2\pi)^3P(k_1)P(k_2)\delta(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)+\mathrm{sym}\right] \tag{224}$$

and for scale-invariant power spectrum $P(k_1)\approx\frac{H^2}{4\epsilon M_p^2}k_1^{-3}$, we have the following scaling (dropping some factors of $\mathcal{O}(1)$ and a minus sign which comes from converting $\zeta = -(5/3)\Phi$ etc.)

$$\langle\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi(\mathbf{k}_3)\rangle \approx f_{NL}^{local}\frac{H^4}{16M_p^4}\frac{1}{\epsilon^2}\delta(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\frac{1}{(k_1k_2k_3)^3}(\sum k_i^3). \tag{225}$$

We see that *if* one of the modes is much longer (i.e. lower momentum) than the others e.g. $k_1 \ll k_2, k_3$ and hence $k_2 \approx k_3$, then the k dependence in (223) sums to $\sum(k_2^3 + k_2^3)$, i.e.

$$\frac{\eta}{8}\sum k_i^3 + \frac{\epsilon}{8}\left(-\sum k_i^3 + \sum_{i\neq j} k_i k_j^2 + \frac{8}{K}\sum_{i>j} k_i^2 k_j^2\right) \rightarrow \frac{\eta+2\epsilon}{8}\sum k_i^3, \tag{226}$$

hence

$$\langle\zeta(\mathbf{k}_1,0)\zeta(\mathbf{k}_2,0)\zeta(\mathbf{k}_3,0)\rangle \stackrel{\text{squeezed}}{=} (2\pi)^3\delta(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\frac{H^4}{M_p^4}\frac{1}{(k_1k_2k_3)^3}\frac{1}{4\epsilon^2}\left[\frac{\eta+2\epsilon}{8}\sum k_i^3\right]. \tag{227}$$

This is known as the **squeezed limit**, since the $k_1 \ll k_2, k_3$ triangle looks like a very squeezed triangle. Comparing this to Eqn. (225), we see that in *the squeezed limit, the slow-roll shape coincide with the local shape* with

$$f_{NL}^{local} = \frac{\eta}{2} + \epsilon. \tag{228}$$

If $\epsilon \gg \eta$ then we have proven our assertion in Section I that $f_{NL}^{local} \approx \epsilon$ for single field slow roll model. The slow-roll shape is dominated by the squeezed limit term – this is a consequence of the fact that the 3-pt correlation is dominated by the factor $(k_1k_2k_3)^3$, hence will blow up as one of the momenta goes to zero.

However, there is another more important point regarding the squeezed limit. In this limit, one of the mode k_1 is much longer than the other two, hence *acts like a perturbation of the background for the two short modes k_2 and k_3*. The long wavelength mode k_1 will cross the horizon much earlier than the short wavelengthe modes k_2 and k_3 so would be constant when the latter two modes cross the horizon. This k_1 perturbation changes the time t_* when this happen by $\delta t_* = -\zeta_1/H$. One can then see that the 3-pt correlation $\langle\zeta_1\zeta_2\zeta_3\rangle$ in the squeezed limit is a correlation between the long wavelength mode with the *power spectrum of the long wavelength mode* k_1, i.e. roughly $\langle\zeta(k_1)P(k_2)\rangle$, and by assumption of weak non-Gaussianity, leads to a correlation with the rate of change of the power spectrum of the short wavelength modes, i.e.

$$\begin{aligned}
\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle &\stackrel{k_1\ll k_2,k_3}{\approx} \langle\zeta(\mathbf{k}_1)\langle\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle_{k_4}\rangle\delta^{-3}(\mathbf{k}_2+\mathbf{k}_3+\mathbf{k}_4)\\
&= \left\langle\zeta(\mathbf{k}_1)(\langle\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle_{k_4} + \frac{d\langle\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle_{k_4}}{d\zeta(\mathbf{k}_4)}\zeta(\mathbf{k}_4)\right\rangle\delta^{-3}(\mathbf{k}_2+\mathbf{k}_3+\mathbf{k}_4)\\
&= \langle\zeta(\mathbf{k}_1)\zeta(-\mathbf{k}_1)\rangle|^{\dagger}\frac{1}{H}\frac{d}{dt}\langle\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle^{*}\delta^{-3}(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)
\end{aligned} \tag{229}$$

where † indicates the time evaluated when the mode k_1 cross the horizon (i.e. earlier) while * indicates the same for modes k_2 and k_3 later, i.e. $t_* > t_\dagger$. In the first line, we have assumed Ergodicity of the long and short wavelength modes hence we are correlating between the long wavelength modes with the *power spectrum* of the short wavelength modes averaged over some length scale $1/k_4$. In the second line, we expand the power spectrum inside as a function of k_4. In the third line, we drop the 3-point correlation (by assumption of Gaussianity of the individual fields), and use the fact that $d\zeta(\mathbf{k}_4) \sim -Hdt \sim -Had\tau \sim d\tau/\tau \sim d\log k|_*$.

But recall that the scalar index for the power spectrum $n_s - 1$ is

$$n_s - 1 = \frac{d\log P_k}{d\log k} = -2\epsilon - \eta = \frac{1}{H}\frac{d}{dt}\log(\langle\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle)|_* \tag{230}$$

where $*$ indicates evaluation at horizon crossing for the mode $k_1 \approx k_2$. Then

$$\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle \sim (-2\epsilon^* - \eta^*)P_{k_1}P_{k_2}\delta(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3). \tag{231}$$

Note that the two power spectra are of the short and long modes respectively and the slow roll parameters are evaluated at the time the short wavelength modes cross the horizon. In other words, there is a **consistency relationship** between the index of the scalar power spectrum and the 3-pt correlation function in the squeezed limit.

This is a fairly powerful relationship : the argument laid out above relies on the fact that the long wavelength mode ζ_3 remains frozen out outside the horizon and does not evolve – this is a feature of *single scalar field* models regardless of the exact details of the potential. Hence if this consistency relationship is not obeyed via observations, then *we can rule out single scalar field models.*

In practice this is a particularly difficult observation to do : the squeezed limit requires $k_1 \ll k_2, k_3$, hence we need a very long wavelength mode. But the CMB only has about 3 octaves, and we only have very few long wavelength mode in the sky due to cosmic variance, hence our ability to observe the squeezed limit of a 3-pt is basically constrained by cosmic variance[23].

Statistical Isotropy. One thing that you might have noticed is that examples of the function $F(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ that we have derived so far contain only *amplitudes* of k's, and not vectors. The reason that this is so is that we have sneakily assumed that the *perturbations are statistically isotropic*, i.e. u_k is simply a function of the amplitude k and not of the vector $\mathbf{k}$. This is an eminently reasonable assumption. The equivalent statement for the power spectrum is that P_k is just a function of k and not $\mathbf{k}$. Physically, it means that all triangle configurations in the CMB are assumed to be drawn from the same distribution, regardless of their orientation[24].

C. What Shape? Theory-based and observational-based approaches

It is clear that, even assuming statistical isotropy, there is still an infinite freedom in the functional form of $F(k_1, k_2, k_3)$. We have argued that the local form is physically relevant to standard inflationary theories. The question is : what other shapes are important?

This turns out to be a very tricky question, and there are several approaches to dealing with this issue. The "theory-based" approach is to ask the question : *What do we expect from "reasonable" models of inflation?* This is the historic approach, and we will take this track in these lectures. One can ask a separate "observations-based" question : *Given a data set, what are the best constrained shapes?* The latter is very recently being pioneered by people at Cambridge.

The advantage of the former method is that it provides a direct link between observations and theory. The disadvantage is that it is *too direct* – since we don't have a true model of inflation, the entire exercise has to be repeated each time a theory of inflation is constructed. One can try to instead *classify* models of inflation – Xingang Chen in Cambridge is one of the world's expert in this endeavour so ask him! The main advantage of the latter method is that it is model independent. However, this very model independence makes it hard for us to try to connect to realistic models of inflation.

An obvious middle ground exists : perhaps one can try to find some "linear combinations" of shape functions, which can encompass a large class of possible shapes. Such linear combinations exist, and are not unique – again work pioneered by people in Cambridge. Historically, the earliest "linear" combination of shapes have two "modes" – the local shape described earlier, and the "equilateral shape". These two shapes are roughly orthogonal – but definitely very incomplete. However, the equilateral shape has a very important physical intepretation, hence we will talk about it next.

D. Equilateral Shape

The equilateral shape is defined by the following function

$$F(k_1, k_2, k_3) = f_{NL}^{equil} 6P_k^2 \left(-\sum_{i>j} \frac{1}{(k_i k_j)^3} - \frac{2}{(k_1 k_2 k_3)^2} + \frac{1}{k_1 k_2^2 k_3^3} + (5 \text{ permutations}) \right). \tag{232}$$

The shape is called "equilateral" because it peaks when $k_1 = k_2 = k_3$, and tapers off when this condition is not fulfilled. The shape is normalized such that at some fixed pivot k_*, its numerical value is equal to the local shape. Current bounds on the equilateral shape from the WMAP7 data is

$$f_{NL}^{equil} = 26 \pm 140. \tag{233}$$

Physically, a strong 3-pt correlation that peaks when the mode amplitudes are the same implies that the non-linearities must be generated without special choice of scale and direction. In other words, it implies *spherical*

---

[23] We can try to extend our "pivot" arm by considering other observations beyond the CMB – one particular promising case is to observe the clustering statistics of "halos", or galactic cluster sized clumps of dark matter.

[24] Note that in the *trispectrum*, i.e. the 4-point correlation function, statistical isotropy means that the 4-point shape function $F_4(k_1, k_2, k_3, k_4, \theta, \phi)$ is a function of 6 numbers – the amplitudes of the edges, and two angles. This set of 6 number parameterize the all the possible tetrahedroids – i.e. any 3D solid with 4 plane faces. Statistical isotropy means that the trispectrum does not depend on how the tetrahedroid is orientated in the sky.

collapse. Given any self-gravitating system of average density ρ, one can define a so-called **free fall time**

$$T_{ff} \equiv \frac{1}{\sqrt{4\pi G\rho}}. \tag{234}$$

The free fall time is, roughly speaking, the time needed for a spherical blob of matter to "fall" into a point under its own self-gravity. Roughly speaking, given some perturbation $\delta(t)$, then any instability suggests that it will grow exponentially over this timescale i.e. $\delta(t) \propto e^{t/T_{ff}}$, hence obey an equation of motion

$$\frac{d^2\delta}{dt^2} \approx \frac{\delta}{T_{ff}^2}. \tag{235}$$

On the other hand, if there exist a counter pressure *gradient* that keeps the spherical blob from self-collapse, then one can characterize this pressure gradient by its rate of change with respect to the change of the density,

$$c_s^2 \approx \frac{\delta p}{\delta \rho} \tag{236}$$

which is known as the **sound speed**. Without going too far afield, the sound speed characterize the propagation velocity of any perturbation over this self-gravitating (and self-supporting) system. By dimensional analysis, the EOM for the perturbation with some pressure support would look like

$$\frac{d^2\delta_k}{dt^2} + \left(c_s^2 k^2 - \frac{1}{T_{ff}^2}\right)\delta_k = 0 \tag{237}$$

where k is the wave number for this perturbation. This means that if $1/k \gg c_s T$, then perturbations over the length scale $1/k \equiv L_J$ is *unstable* and will collapse. $L_J = c_s/\sqrt{4\pi G\rho}$ is called the **Jeans Length**.

Now in inflation, spacetime is stretched by exponential expansion, so perturbations which collapse under gravity must compete against this effect. One can think of the exponential expansion as a "pressure" term in a spherical collapse scenario where the "free-fall" timescale is similarly given by $T_{ff} = 1/\sqrt{4\pi G\rho} \approx 1/H$, and hence the important thing one should look for is the Jeans length c_s/H of the perturbations, where c_s is the sound speed of the perturbations.

As it turns out, in standard single scalar field models, the sound speed for the perturbations $c_s^2 = 1$, i.e. the Jeans length of the perturbations is the Hubble horizon – only wavelengths of Hubble scale are unstable. In the language that you are more familiar with – instability of perturbations means that the perturbation stops behaving like a wave and "freeze-out". On the other hand, one can immediately ask "what happens if the Jeans length is smaller?" If such a model exist, the perturbations become unstable *earlier*, and hence freezes-out before it crosses the Horizon $1/H$ – instead it freezes out when it crosses the *sound horizon* c_s/H.

Now how do we arrange for such a model? It turns out that there exists a generic model of single scalar field inflation that allows for a varying Jeans length, called **k-fluid** where the action looks like

$$S_\phi = \int d^4x\sqrt{g}\,[P(X,\phi)] \tag{238}$$

where $X = -(1/2)(\partial\phi)^2$ is the kinetic term. The "k" stands for kinetic, i.e. one generates a dynamic by having non-standard kinetic term in the theory. The extra function $P(X,\phi)$ here introduces a new dynamic into the system besides the slow roll parameters ϵ and η, namely the perturbation velocity c_s^2. In general c_s is *background dependent*, but in a FRW backround ϕ is homogenous and hence is X is timelike. In general, the functional form of c_s^2 for $P(X,\phi)$ when X is timelike is given by

$$c_s^2(X,\phi) = \frac{P_X}{P_X + 2XP_{XX}} \tag{239}$$

where the subscript means partial derivatives (really functional derivatives) w.r.t. X. If you are brave, you can *derive* this by following the ADM formalism laid out in previous lectures, and find that the 2nd order "Mukhanov-Sasaki" action is

$$S_{MS} = \frac{1}{2}\int d^3x d\tau\,\left[v'^2 - c_s^2(\partial v)^2 - \frac{z''}{z}v^2\right], \tag{240}$$

where $z = (a/c_s)\sqrt{2\epsilon}$ and $v = -z\zeta$ as usual. The extra c_s in the action means that the mode function now evolves (in near de Sitter space) as

$$u_k(\tau) = \frac{H}{\sqrt{4\epsilon c_s k^3}}(1 - ikc_s\tau)e^{ic_s k\tau}. \tag{241}$$

Since the perturbation in such a model now freezes out at the sound horizon c_s/H, so the power spectrum of this model can be shown to be

$$P(k) = \frac{H^2}{8\pi\epsilon}\frac{1}{c_s}. \tag{242}$$

The extra $1/c_s$ term ultimately comes from the $\sqrt{c_s}$ in the denominator of the mode function Eqn. (241). It is also intuitively clear that non-linearities of a more unstable system will also be bigger, to see this we have to then calculate the higher order corrections to the equation of motion of the perturbations – information which is encoded in the bispectrum of the perturbations.

Hence you have to be very brave, as then you need to calculate the 3rd order action. It can be calculated very similarly to the standard case, with just even more extra algebra to keep track of the non-trivial c_s^2 term. Fortunately, this has been done, and it looks like

$$\mathcal{H}_{int} = -\left[-a^3(\Sigma(1-\frac{1}{c_s^2}) + 2\lambda)\pi^3 + a^3\frac{\epsilon}{c_s^4}(\epsilon - 3 + 3c_s^2)\zeta\pi^2 + 2a\frac{\epsilon}{c_s^2}(\epsilon - 2s + 1 - c_s^2)\zeta(\partial\zeta)^2\right. \tag{243}$$

$$\left. -2a\frac{\epsilon}{c_s^2}\pi(\partial\zeta)(\partial\chi) + \frac{a^3}{2}\frac{\epsilon}{c_s^2}(\eta/c_s^2)^{\cdot}\zeta^2\pi + \frac{1}{2}\frac{\epsilon}{a}\partial\zeta\partial\chi\partial^2\chi + \frac{\epsilon}{4a}\partial^2\zeta(\partial\chi)^2\right]. \tag{244}$$

We have defined the higher derivatives of P (in terms of X) as

$$\Sigma \equiv \frac{H^2\epsilon}{c_s^2} = XP,X + 2X^2P_{,XX} \tag{245}$$

$$\lambda = X^2P_{,XX} + \frac{2}{3}X^3P_{,XXX} \tag{246}$$

$$s = \frac{\dot{c}_s}{Hc_s}. \tag{247}$$

You can check by setting $c_s^2 = 1$ for $P(X,\phi) = X - V$, that we recover the 3rd order action for canonical fields Eqn. (171). By eye, you can see that the bunch of c_s^2 in the denominators (not to mention in the mode functions, although one do have to remember that each dot and ∂ brings down a factor of c_s too), means that there are possibilities of generating large non-Gaussianities if c_s is small enough. This should not be a surprise : smaller soundspeed means that the Jeans length is shorter and hence perturbations are much more unstable at long wavelengths and hence form non-linearities at a greater rate. It is worth emphasizing that this is a *generic* observation, regardless of background cosmology : *gravitational instability of perturbations lead to spherical collapse and hence equilateral type non-Gaussianities*. This effect is seen in both inflationary perturbations, post re-entry evolution and late time large scale formation. This is especially pronounced in the latter during the matter domination phase of the universe as has been verified in large-scale N-body numerical simulations.

By the above scaling argument, one can easily check that the first three terms of Eqn. (244) will dominate over the others. One can then use the technique laid out in the previous section, plug in H_{int} into Eqn. (81) to calculate the non-Gaussianities – the only difference is that we substitute in Eqn. (241) instead of the usual de Sitter mode functions Eqn. (198)) when contracting. These terms contribute the following shape

$$F(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = (2\pi)^3\delta^3(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)\frac{H^4}{16\epsilon^2 c_s^2}\frac{1}{(k_1k_2k_3)^3}(S_\lambda + S_c) \tag{248}$$

where

$$S_\lambda = \left(\frac{1}{c_s^2} - 1 - \frac{\lambda}{\Sigma}\right)\frac{3k_1k_2k_3}{2K^3} \tag{249}$$

$$S_c = \left(\frac{1}{c_s^2} - 1\right)\frac{1}{k_1k_k2k_3}\left(-\frac{1}{K}\sum_{i>j}k_i^2k_j^2 + \frac{1}{2K^2}\sum_{i\neq j}k_i^2k_j^3 + \frac{1}{8}\sum_i k_i^3\right). \tag{250}$$

The shape Eqn. (248) is *roughly* of the equilateral form Eqn. (232) – in the sense that both of them *peak* at $k_1 = k_2 = k_3$. Note that they are *not* the same! The equilateral shape Eqn. (232) is a *template*. On the other hand, one can argue that a detection of the equilateral shape is a strong indication of $P(X,\phi)$ models of inflation. A particularly popular form of this model is the so-called Dirac-Born-Infeld inflation model inspired by string models

$$S_{DBI} = \int d^4x\sqrt{g}\left[h^{-1}(\phi)\sqrt{1+h(\phi)X} + V(\phi)\right] \tag{251}$$

which clearly has the form $P(X,\phi)$. Inflation is driven by the potential $V(\phi)$ term which has a minima.

By contrast, the original *k-inflation* no explicit potential term. This is a large class of theories – an example is a "power-law" k-inflation $a \propto t^{2/3\gamma}$

$$P(X,\phi) = \frac{4}{9}\frac{4-3\gamma}{\gamma^2}\frac{X^2 - M_p^4 X}{(\phi-\phi_*)^2} \tag{252}$$

where the ϕ_* is some attractive fixed point and $0 < \gamma < 2/3$.

E. Folded Shape

The third important shape is the so-called folded shape, where the non-Gaussianity peak at $k_1 = 2k_2$ and $k_2 = k_3$. The template for this shape is given by

$$F_{folded}(k_1,k_2,k_3) = 6\left(\frac{k_1^2}{k_2k_3} + \text{sym}\right) - 6\left(\frac{k_1}{k_2} + \text{sym}\right) + 18. \tag{253}$$

The primary generator for this kind of shape is that if the in-state is not the Bunch-Davies state, but some form of *excited initial state*. There are all sorts of theories why the universe should be in its excited state instead of its vacuum state – one popular theory suggests that if one take some co-moving mode with some co-moving length $1/k$ and evolve it backwards, at some point t_*, this mode will be Planck scale, i.e. $a(t_*)/k < 1/M_{pl}$. At this point, one can argue that we don't really have control over QFT – hence if we decide to set initial conditions at this point there is no reason to assume that it is in the vacuum state[25].

In the previous case, we have consider the in-state to be one annihilated by the positive frequency operator $|\text{in}\rangle = |\text{BD}\rangle$ and

$$\zeta_I^+|\text{BD}\rangle = 0. \tag{254}$$

However, one can equally consider the in-state to be any other *excited* state which we can construct simply by hitting it with a bunch of negative frequency operators

$$|\text{in}\rangle = \zeta_I^-(\mathbf{k}_1)\zeta_I^-(\mathbf{k}_2)...|\text{BD}\rangle. \tag{255}$$

Operationally, this means that the mode function Eqn. (198) now gets a negative frequency part, i.e.

$$\tilde{u}_k(\tau) = \alpha_k u_k(\tau) + \beta_k u_k^*(\tau) \tag{256}$$

where α_k and β_k are complex **Bogoliubov coefficients** which can be a function of k and

$$u_k(\tau) = \frac{H}{\sqrt{4\epsilon k^3}}(1-ik\tau)e^{ik\tau}. \tag{257}$$

α_k and β_k are normalized such that $|\alpha_k|^2 - |\beta_k|^2 = 1$ and they encode how excited the in-state is, i.e. they depend on the exact set of negative frequency operators in Eqn. (255). In general, they are any functions of k, but one must be careful as initial state which is overly excited may contain too much energy and hence overclose the universe. This is called the **Hadamard condition**[26], and usually means that $|\beta_k|/|\alpha_k| \ll 1$.

---

[25] These kind of arguments are called "Transplanckian" theories, which was all the rage about 10 years ago and now is undergoing some form of minor renaissance which occurs when people run out of things to do and start recycling old ideas.

[26] Technically this means that β_k has to fall faster than $1/k^2$, although I never quite figure out why it is called "Hadamard" in the first place.

Aplpying Wick's Theorem, we now contract to the Wightman functions constructed out of $\tilde{u}$ instead of u

$$\begin{aligned}\langle\zeta_1\zeta_2\rangle &= \langle[\zeta_1^+,\zeta_2^-]\rangle = (2\pi)^3\tilde{u}(\mathbf{k}_1,t)\tilde{u}^*(\mathbf{k}_2,t)\delta(\mathbf{k}_1-\mathbf{k}_2) &(258)\\ &= (|\alpha_k|^2+|\beta_k|^2)|u_k|^2+(\alpha_k\beta_k^*u_k^2+\text{c.c}). &(259)\end{aligned}$$

It is then straightforward to follow the calculation in Section IV to calculate non-Gaussianities – one simply replace u with $\tilde{u}$, and substituting in Eqn. (259). Let's see how this will change the result by considering the interaction term we calculated in that section, $H_{int}^A = \int d^3x\ a\epsilon^2\zeta(\mathbf{x},\tau)(\zeta'(\mathbf{x},\tau))^2$. To simplify matters, we consider the case where $|\beta_k|/|\alpha_k| \ll 1$, hence we will only keep terms to first order in β_k. We pick up the derivation from Eqn. (216), and substituting in $\tilde{u}$ instead of u, we get

$$\begin{aligned}\langle\zeta(\mathbf{k}_1,0)\zeta(\mathbf{k}_2,0)\zeta(\mathbf{k}_3,0)\rangle &= \text{Re}\left[-2i\zeta(\mathbf{k}_1,0)\zeta(\mathbf{k}_2,0)\zeta(\mathbf{k}_3,0)\int\frac{d^3q_1}{(2\pi)^3}\frac{d^3q_1}{(2\pi)^3}\frac{d^3q_1}{(2\pi)^3}\right.\\ &\quad\left.\int_{-\infty(1+i\epsilon)}^{\tau}d\tau' d^3x a^2\epsilon^2\zeta(\mathbf{q}_1,\tau)\zeta'(\mathbf{q}_2,\tau)\zeta'(\mathbf{q}_3,\tau)\right]e^{-i(\mathbf{q}_1+\mathbf{q}_2+\mathbf{q}_3)\cdot\mathbf{x}}\\ &= 2\times\text{Re}\left[-2i\tilde{u}_{k_1}(0)\tilde{u}_{k_2}(0)\tilde{u}_{k_3}(0)\int\frac{d^3q_1}{(2\pi)^3}\frac{d^3q_1}{(2\pi)^3}\frac{d^3q_1}{(2\pi)^3}\int_{-\infty(1+i\epsilon)}^{\tau}d\tau' d^3x\ a^2\epsilon^2\right.\\ &\quad\left.\times\tilde{u}_{q_1}^*(\tau')\frac{d}{d\tau}\tilde{u}_{q_2}^*(\tau')\frac{d}{d\tau}\tilde{u}_{q_3}^*(\tau')(2\pi)^9\delta(\mathbf{k}_1-\mathbf{q}_1)\delta(\mathbf{k}_2-\mathbf{q}_2)\delta(\mathbf{k}_3-\mathbf{q}_3)e^{-i(\mathbf{q}_1+\mathbf{q}_2+\mathbf{q}_3)\cdot\mathbf{x}}\right]\\ &\quad +1\rightarrow 2+1\rightarrow 3 \qquad (260)\end{aligned}$$

as usual. Now, using Eqn. (259), we will get something really long but straightforward to compute. To simplify our life for this calculation, let's focus only on a single leading order term in β_{k_1} – we'll see how the folded shape comes out clear enough. Eqn. (260) then becomes

$$\begin{aligned}\langle\zeta(\mathbf{k}_1,0)\zeta(\mathbf{k}_2,0)\zeta(\mathbf{k}_3,0)\rangle &= \mathcal{O}(\alpha_{k_1})\text{ term}+2\beta_{k_1}\times\text{Re}\left[-2i\frac{H^6}{(4\epsilon)^3}\frac{1}{(k_1k_2k_3)^3}\int_{-\infty(1+i\epsilon)}^{\tau}d\tau'\ \frac{1}{(H\tau)^2}\epsilon^2\right.\\ &\quad\left.\times(k_2k_3)^2\tau^2(1-ik_1\tau)e^{-i(-k_1+k_2+k_3)\tau}(2\pi)^3\delta(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\right]+.... \qquad (261)\end{aligned}$$

Comparing the above equation to Bunch-Davies case in Eqn. (218), we see that the addition of the negative frequency mode β_{k_1} means that we have swapped the sign $+ik_1\rightarrow -ik_1$ inside the integrands (in two instances). Note that we have *not* swapped the sign of the wavevector $i\mathbf{k}_1$ (can you see why?)! The time integral now leads to

$$\begin{aligned}\int_{-\infty(1+i\epsilon)}^{\tau}d\tau'(k_2k_3)^2(1+ik_1\tau)e^{-i(-k_1+k_2+k_3)\tau} &= (k_2k_3)^2\left[\frac{-ik_1}{(-k_1+k_2+k_3)^2}+\frac{i}{(-k_1+k_2+k_3)}-\text{Real Term}\right]\\ &\quad\times e^{-i(-k_1+k_2+k_3)\tau}|_{-\infty(1+i\epsilon)}^{\tau}. \qquad (262)\end{aligned}$$

Note that the denominator $k_1+k_2+k_3\rightarrow -k_1+k_2+k_3$ compared to the Bunch Davies case – this comes from the negative frequency contribution in the exponential term. Hence although $|\beta_k|\ll|\alpha_k|$, the expression blows up when $k_1=k_2+k_3$ and by the triangle condition $k_2=k_3$, i.e. the when the triangle is of the "folded shape". Of course a singularity is unphysical, we usually cut off the integral by substituting

$$k_1 = k_2+k_3+\delta \qquad (263)$$

where δ is some co-moving positive mass-scale (so the triangle is never "fully folded"). This mass-scale usually depends on the model in question. The final answer is then, for the order β_k term

$$F_{folded}^{\beta_k}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)=\beta_{k_1}\frac{H^4}{16\epsilon}\frac{1}{(k_1k_2k_3)^3}(2\pi)^3(k_2k_3)^2\left[\frac{1}{(-k_1+k_2+k_3+\delta)}-\frac{k_1}{(-k_1+k_2+k_3+\delta)^2}+1\rightarrow 2+1\rightarrow 3\right]. \qquad (264)$$

It's not hard to see that Eqn. (264) is peaked when the denominators $\rightarrow 0$, i.e. when the shape is folded.

F. General Non-Gaussianities from Features in the Inflationary Potential (*)

Finally, we will briefly touch on the fact that breaking slow-roll can generate large non-Gaussianities. Breaking slow-roll is tricky business – to ensure inflation we require that $\epsilon < 1$ for at least 60-efolds, this means that η cannot deviate from 1 for too long. However, notice that one of the interaction term in Eqn. (171) is

$$\mathcal{H}_{int} = \dot{\eta}\epsilon\zeta^2\dot{\zeta} \tag{265}$$

i.e. this term is $\propto \dot{\eta}$. So one can always engineer potentials which maintain effective slow-roll, while generating a non-trivial amount of non-Gaussianity from this term since we can have extremely large $\dot{\eta}$ while keeping ϵ small. There are generally two large classes of model which do this : **Step Models** and the **Resonant Models**. There are many variations of such models, including the *ramp* model where the step occurs at the derivative level and the *speedbump* model where we have a blister on the potential. The only constraint to such model building is the number of graduate students to calculate them.

1. *Step Models*

In this model, we introduce a small but sharp "step" in the potential

$$V(\phi) = V_0(\phi)\left(1 + c\tanh\frac{\phi-\phi_s}{d}\right)\ ,\ c \ll 1\ ,\ d/M_p \ll 1. \tag{266}$$

Roughly speaking c parameterize the size of the step, d parameterize the steepness of the step while ϕ_s parameterize the position of the step.

By tuning the parameters c,d and ω, one can make η' big while keeping $\Delta\epsilon$ small. For the step potential Eqn. (266), one can estimate that

$$\dot{\eta} \sim 10H\frac{c^2}{d^2\epsilon}. \tag{267}$$

For a model with $c = 0.0018$,$d = 0.022$ and $V_0 = (1/2)m^2\phi^2$ with $m = 3\times 10^{-6}M_p$, one can get roughly $\mathcal{O}(1000)$ boost in the bispectrum when compared to the other terms with $\epsilon \sim 0.01$. The shape of the bispectrum is non-trivial, but has the following rough sinodoidal form

$$F(k_1,k_2,k_3) = A\sin(K/k_s)\times(K/k_s)^n e^{-\alpha K/k_s} \tag{268}$$

where k_s is the scale of the step, and α and n are numbers of order 1 which encodes the "enveloping" of the bispectrum ringing. Since the *location* of the step defines a scale, it is not surprising that the bispectrum is not scale-invariant.

This is quite a large violation of the slow-roll condition and it turns out that the *power spectrum* also undergoes this ringing and hence violate scale-invariance – and the power spectrum is very well measured. Surprisingly, although one might expect that a small violation of scale-invariant power spectrum would rule out the model, due to the projection effects of the power spectrum on the spherical CMB sky and the fact that small "ringing" is actually present in the CMB power spectrum, there is some wiggle room for inflation model building using such step models. This has prompted an invasion of theorists to calculate and recalculate such models in recent months.

2. *Resonant Models*

On the other hand, one can introduce a small oscillatory component into the potential

$$V(\phi) = V_0(\phi)(1 + c\sin(\phi/\lambda))\ ,\ c \ll 1\ ,\ \lambda \ll \sqrt{\epsilon}M_p. \tag{269}$$

For the resonant-type models, the oscillatory nature of the potential means that there exist an oscillatory component to $\dot{\eta}$

$$\dot{\eta} \supset A\sin(\omega t)\ ,\ \omega \approx \frac{\dot{\phi}}{2\pi\lambda}, \tag{270}$$

where A is some function weakly dependent on time. Now we also know that the product of the wave functions in the iteraction term Eqn. (265) is oscillatory themseles $\propto \sin((k_1+k_2+k_3)\tau)$ – hence when $\omega \sim (k_1+k_2+k_3)/a$ we expect *resonance* to occur. This resonance can constructively generate large amount of non-Gaussianities. Destructive resonance can also occur, and hence one expect the non-Gaussianities to exhibit the form

$$F(k_1,k_2,k_3) \propto f_{NL}^{res}\sin(C\ln K/k_s + \text{phase}) \tag{271}$$

where C is some constant depending on the model. Notice that since all modes have to pass through the resonance scale as it is deep inside the horizon $\omega \gg H$, the spectra of non-Gaussianities *is present at all scales*. Since the non-Gaussianities are built up over several oscillation time scales, the size of the oscillations c can be much smaller than the step model, and hence violation of slow roll is even tinier. This means that the power spectrum for this model is scale invariant (albeit modulated by a small oscillatory component which one can tune to be unobservable).

This innocuous looking model was cooked up during a boring cold evening at Yale University by your lecturer – who was convinced at that time that it is so contrived that nobody will take it seriously. However, a little more careful thought indicated that it is rather devious and interesting – when cast in high brow language. In usual inflation, by the arguments of effective theory, any mass scale $M \gg H$ is expected to play little or no role into the dynamics of inflation hence the fact that we have an introduce a high mass scale $\omega \gg H$ into the inflation theory with interesting dynamics and an *observable* signature is a very non-trivial statement. One way to think about such an introduction of scale is think in terms of a breaking of symmetry. The statement that the inflaton is "slowly rolling" is the same statement as this shift symmetry

$$\phi \rightarrow \phi + c, \tag{272}$$

is *weakly* broken – leading to the *almost* scale invariant power spectrum we observed. Hence the introduction of this high mass scale as an *oscillatory* scale ω breaks the *continuous* shift symmetry into a *discrete* shift symmetry.

Naturally, after the cold boring evening at Yale, the resonant mechanism has become rather fasionable among model builders.

V. EPILOGUE

I asked the moon to join me for a drink, and my shadow made us a threesome...now I am drunk, we part ways, but perhaps we may meet again somewhere along the Milky Way.

Li Bai (AD 701-761)

While this set of lectures have been supposedly designed to teach you how to calculate primordial non-Gaussianities given some theory of inflation, my secret plan is to use this goal as a way to introduce some advanced techniques in modern cosmology – In-in formalism and ADM. While it is almost impossible to read a paper on primordial non-Gaussianities without encountering these techniques, their use is a lot more far-reaching than to calculate mere non-Gaussianities. The "in-in" Schwinger-Keldysh formalism is a standard tool in condensed matter calculations for 60 years now, and it is somewhat surprising that cosmologists just started to pick it up less than a decade ago.

Anyhow, hopefully having gone this far, the lectures have not completely bored the brains out of your head. This set of lectures is constantly evolving – next up are figures. There are still plenty of typos to be had. Okay, the last sentence does not inspire confidence, but I hope you have enjoyed reading these notes as much as I had making them.

VI. EXAMPLES

1.*Contraction Algebra*
For a Gaussian GRF, calculate

$$\langle f(\mathbf{k}_1)f(\mathbf{k}_2)f(\mathbf{k}_3)f(\mathbf{k}_4)\rangle \tag{273}$$

$$\langle f(\mathbf{k}_1)f(\mathbf{k}_2)f(\mathbf{k}_3)f(\mathbf{k}_4)f(\mathbf{k}_5)f(\mathbf{k}_6)\rangle. \tag{274}$$

Given the PDF

$$P(a_{\mathbf{k}},b_{\mathbf{k}}) = \frac{1}{\pi\sigma_k^2}\exp\left[-\frac{a_{\mathbf{k}}^2+b_{\mathbf{k}}^2}{\sigma_k^2}\right]\left(1+\sigma_k\frac{S_k}{6}H_3(a_{\mathbf{k}}/\sigma_k)\right), \tag{275}$$

show that the one-point correlation function

$$\langle f(\mathbf{k})\rangle = 0 \tag{276}$$

vanishes and hence calculate

$$\langle f(\mathbf{k})f(\mathbf{k})f(\mathbf{k})\rangle. \tag{277}$$

For fun, calculate

$$\langle f(\mathbf{k})f(\mathbf{k})f(\mathbf{k})f(\mathbf{k})\rangle. \tag{278}$$

2.*Scale Invariance of bispectrum*
Show that the Komatsu-Spergel Local form

$$\langle\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi(\mathbf{k}_3)\rangle = f_{NL}^{local}\left[2(2\pi)^3P(k_1)P(k_2)\delta(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)+\mathrm{sym}\right], \tag{279}$$

is scale invariant provided that the power spectrum $P(k)$ is also scale-invariant.
3.*Interaction Hamiltonian*
In class, we showed that for a 3rd order Lagrange density of the form

$$\mathcal{L} = \frac{1}{2}\dot{\zeta}^2 - V(\zeta,\partial\zeta) + \alpha\dot{\zeta}^3 \tag{280}$$

the Hamiltonian density can be divided into

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{int} \tag{281}$$

where $\mathcal{H}_0 = (1/2)\pi^2 + V_0(\zeta,\partial\zeta)$ and

$$\mathcal{H}_{int} = -\mathcal{L}_{int} + \mathcal{O}(4) \tag{282}$$

where $\mathcal{L}_{int}$ consists of the 3rd order terms of $\mathcal{L}$ while V_0 contains terms quadratic in ζ. Show that for the general Lagrange Density of the form

$$\mathcal{L} = \frac{1}{2}\dot{\zeta}^2 - V(\zeta,\partial\zeta) + \alpha\dot{\zeta}^3 + \beta\zeta\dot{\zeta}^2 + \gamma\zeta^2\dot{\zeta} \tag{283}$$

the substitution $\mathcal{H}_{int} = -\mathcal{L}_{int} + \mathcal{O}(4)$ holds as long as $|\zeta| \sim |\pi|$.

4. *Calculation of non-Gaussian Terms using "in-in" Formalism*

Calculate the non-Gaussianities for the following terms at $\tau \to 0$

- $\mathcal{H}_{int} = a\epsilon^2\zeta(\partial\zeta)^2$
- $\mathcal{H}_{int} = a^3\epsilon\zeta^2\dot{\zeta}$
- $\mathcal{H}_{int} = a^3\epsilon^2 H^{-1}\dot{\zeta}^3$
- $\mathcal{H}_{int} = a\epsilon^2(\partial\zeta)\cdot(\partial(\partial^{-2}\dot{\zeta}))\dot{\zeta}$

5. *To what order in $N^{(n)}$ do we need to expand to calculate the action to order n? (*)*

In this guided problem, we will show that to calculate the action to $\mathcal{O}(\zeta^n)$, we need to expand the lapse and shift to order $n-2$ for $n \geq 3$ and to first order for $n=2$ as claimed in the notes and lectures.

In order to simplify notation, we will simply denote the lapse N and shift N^i by just N – there will be sufficient indices to write down as it is. We begin with the action

$$S = \int d^4x \mathcal{L}(N, \partial_i N). \tag{284}$$

The variation of the action w.r.t to N is

$$\delta S = \int d^4x \left[\frac{\partial\mathcal{L}}{\partial(\partial_i N)}\partial_i \delta N + \frac{\partial\mathcal{L}}{\partial(N)}\delta N\right]. \tag{285}$$

As usual, we expand $N = N^{(0)} + \Delta N = N^{(0)} + N^{(1)} + N^{(2)} + ...$ By ignoring cross terms such as

$$\frac{\partial^2\mathcal{L}}{\partial(\partial_i N)\partial(N)} \tag{286}$$

show that

$$\frac{\partial\mathcal{L}}{\partial(\partial_i N)} = \left.\frac{\partial\mathcal{L}}{\partial(\partial_i N)}\right|_0 + \left.\frac{\partial^2\mathcal{L}}{\partial(\partial_i N)\partial(\partial_i N)}\right|_0 \partial_j \Delta N + ... \tag{287}$$

$$\frac{\partial\mathcal{L}}{\partial(N)} = \left.\frac{\partial\mathcal{L}}{\partial(N)}\right|_0 + \left.\frac{\partial^2\mathcal{L}}{\partial N^2}\right|_0 \Delta N + ... \tag{288}$$

where $|_0$ means that the term is evaluated with $\Delta N = 0$ as per the usual rule for Taylor expansions. Now here is the subtlety – even though we have Taylor expanded around $N^{(0)}$, the terms still contain *perturbation variables ζ to all orders*, and hence one can further expand them in this perturbation order

$$\left.\frac{\partial\mathcal{L}}{\partial(N)}\right|_0 = \left.\frac{\partial\mathcal{L}}{\partial(N)}\right|_{0,\zeta^0} + \left.\frac{\partial\mathcal{L}}{\partial(N)}\right|_{0,\zeta} + \left.\frac{\partial\mathcal{L}}{\partial(N)}\right|_{0,\zeta^2} + ... \tag{289}$$

and similarly for the 2nd derivative term. Using these facts, expand the action Eqn. (285) to *zeroth* order in perturbation in both N and ζ and show that it gives the background equation of motion. Then show that to *first* order in perturbation, Eqn. (285) is given by

$$\int d^4x \left\{\left[\left.\frac{\partial\mathcal{L}}{\partial(\partial_i N)}\right|_{0,\zeta} + \left.\frac{\partial^2\mathcal{L}}{\partial(\partial_i N)\partial(\partial_i N)}\right|_{0,\zeta^0}\partial_j N^{(1)}\right]\partial_i\delta N \right.$$
$$\left. + \left[\left.\frac{\partial\mathcal{L}}{\partial(N)}\right|_{0,\zeta} + \left.\frac{\partial^2\mathcal{L}}{\partial N^2}\right|_{0,\zeta^0} N^{(1)}\right]\delta N\right\} = 0. \tag{290}$$

Show that, after integrating Eqn. (290) by parts, one obtain the *constraint equation* for ΔN to order ζ, i.e. $N^{(1)}$. Comment on how one would obtain the constraint equations for $N^{(n)}$ for $n > 1$. To obtain the action to $\mathcal{O}(\zeta^3)$, then one naively simply solve the constraints to $N^{(3)}$, and plug them back into the action. However, now we show that in fact, we just need information up to $N^{(1)}$ in the following manner.

First, expand the action around the background solution, $S = S_0 + \Delta S$ – note that ΔS is *not* the variation of the action δS ! Since we know that the background solution S_0 is simply the action with the constraint equations $N^{(0)}$ plugged back in, Taylor expand S around this (and hence expand in terms of ΔN) to obtain

$$\Delta S = \int d^4x \left[\left.\frac{\partial \mathcal{L}}{\partial(\partial_i N)}\right|_0 \partial_i \Delta N + \frac{1}{2!} \left.\frac{\partial^2 \mathcal{L}}{\partial(\partial_i N)\partial(\partial_j N)}\right|_0 (\partial_i \Delta N)(\partial_j \Delta N) + \left.\frac{\partial \mathcal{L}}{\partial(N)}\right|_0 \Delta N + \frac{1}{2!} \left.\frac{\partial^2 \mathcal{L}}{\partial N^2}\right|_0 (\Delta N)^2 \right]. \tag{291}$$

Calculate all the terms that contain $N^{(n)}$. Show that, after integration by parts, the terms is simply proportional to the *zeroth* order constraint equation, and hence vanish.

Finally, show that the terms that contain $N^{(n-1)}$ at order n is given by

$$\Delta S = \int d^4x \left[\left.\frac{\partial \mathcal{L}}{\partial(\partial_i N)}\right|_{0,\zeta} \partial_i N^{(n-1)} + \frac{1}{2!} \left.\frac{\partial^2 \mathcal{L}}{\partial(\partial_i N)\partial(\partial_j N)}\right|_{0,\zeta^0} (\partial_i N^{(n-1)})(\partial_j N^{(1)}) + \left.\frac{\partial \mathcal{L}}{\partial(N)}\right|_{0,\zeta} N^{(n-1)} + \frac{1}{2!} \left.\frac{\partial^2 \mathcal{L}}{\partial N^2}\right|_{0,\zeta^0} N^{(n-1)} N^{(1)} \right]. \tag{292}$$

By integrating by parts, and using Eqn. (290) show that this term also vanishes. Hence we have proven that to calculate the action to order $n = 2$ and $n = 3$, we need only to solve the constraints N to order $n = 1$ in ζ. Comment on what order in ζ that we need to solve N to if the action required is $n > 3$.