Lecture on "Non-Gaussianity"

given for "School on Cosmology & Gravitational Waves,"
at Inter University Centre for Astronomy & Astrophysics
(IUCAA), Pune, India.

By Eiichiro Komatsu (University of Texas at Austin)

Q. What is "non-Gaussianity"?
A. Deviation from a Gaussian distribution.
Q. Why is it important?
A. Because detection of any forms of **Primordial** non-Gaussianity is a breakthrough in cosmology.
Q. Why?
A. I will tell you ...

Outline of the lecture

1. Gaussian and non-Gaussian statistics (Dec. 1)
2. Non-Gaussianity as a test of inflation (Dec. 2-3)
3. Measuring non-Gaussianity from the cosmic microwave background (Dec.
4. Measuring non-Gaussianity from the large-scale structure (Dec. 4) (if we have time)

References

- E. Komatsu, Classical and Quantum Gravity, 27, 124010 (2010) (arXiv: 1003.6097)
- E. Komatsu, astro-ph/0206039
- E. Komatsu et al., arXiv: 0902.4759

1. Gaussian and non-Gaussian Statistics

1.1. Gaussian statistics

Consider the simplest case in which we have a random variable, x, drawn from a probability density function (PDF), $P(x)$. This function is normalized such that

$$\int_{-\infty}^{\infty} dx\, P(x) = 1$$

with zero mean

Now, x obeys Gaussian statistics, if $P(x)$ takes on the following form:

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2}$$

The coefficient, $(2\pi\sigma^2)^{-1/2}$, is necessary for $\int_{-\infty}^{\infty} dx\, P(x) = 1$ to be satisfied.

This PDF specifies all the statistical properties of x. Various statistical quantities can now be computed:

$$\langle x \rangle \equiv \int_{-\infty}^{\infty} dx\, x P(x) = 0 \qquad \text{[zero mean]}$$

$$\langle x^2 \rangle \equiv \int_{-\infty}^{\infty} dx\, x^2 P(x) = \sigma^2 \qquad \text{[variance]}$$

$$\langle x^3 \rangle \equiv \int_{-\infty}^{\infty} dx\, x^3 P(x) = 0 \qquad \text{[zero "skewness"]}$$

$$\langle x^4 \rangle \equiv \int_{-\infty}^{\infty} dx\, x^4 P(x) = 3\sigma^4$$

$$\langle x^5 \rangle \equiv \int_{-\infty}^{\infty} dx\, x^5 P(x) = 0$$

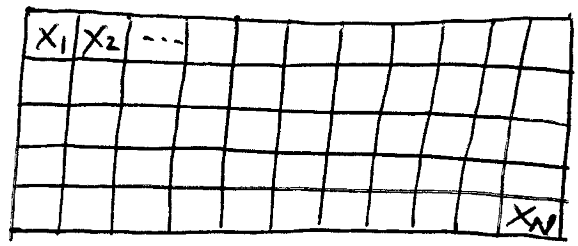
$$\vdots$$

[zero "kurtosis," defined as $\kappa_4 \equiv \langle x^4 \rangle - 3\langle x^2 \rangle^2 = 0$]

Thus, the odd moments, $\langle x^{2n-1} \rangle$, vanish, and the even moments, $\langle x^{2n} \rangle$, are given in terms of σ^{2n}.

σ determines everything!

Next, consider many random variables, $\{x_1, x_2, \cdots, x_N\}$, spread over in space :



These variables may obey a "multi-variate Gaussian distribution"

$$P(x_1, x_2, \cdots, x_N) = \frac{1}{(2\pi)^{N/2} |\xi|^{1/2}} e^{-\frac{1}{2}\sum_{ij} x_i (\xi^{-1})_{ij} x_j}$$

where $\xi_{ij} \equiv \langle x_i x_j \rangle$ describes the **co**variance of x_i & x_j. We call this quantity either a "covariance matrix," or "2-point correlation function."

This PDF is again normalized such that

$$\int_{-\infty}^{\infty} dx_1 dx_2 \cdots dx_N \, P(x_1, x_2, \cdots, x_N) = 1$$

The moments are

$$\langle x_i \rangle = \int_{-\infty}^{\infty} dx_1 \cdots dx_N \quad x_i P(\{x_k\}) = 0$$

$$\langle x_i x_j \rangle = \int_{-\infty}^{\infty} dx_1 \cdots dx_N \quad x_i x_j P(\{x_k\}) = \xi_{ij}$$

$$\langle x_i x_j x_l \rangle = \int_{-\infty}^{\infty} dx_1 \cdots dx_N \quad x_i x_j x_l P(\{x_k\}) = 0$$

$$\langle x_i x_j x_l x_m \rangle = \int_{-\infty}^{\infty} dx_1 \cdots dx_N \; x_i x_j x_l x_m P(\{x_k\}) = \xi_{ij}\xi_{lm} + \xi_{il}\xi_{jm} + \xi_{im}\xi_{jl}$$

$\vdots$

→ This, and higher-order moments results, are know as "Wick's Theorem," which allows us to write all the even moments in terms of ξ_{ij}. (e.g.,)

$$\langle x_1 x_2 x_3 x_4 \rangle = \langle x_1 x_2 \rangle \langle x_3 x_4 \rangle + \langle x_1 x_3 \rangle \langle x_2 x_4 \rangle + \langle x_1 x_4 \rangle \langle x_2 x_3 \rangle$$
$$= \xi_{12}\xi_{34} + \xi_{13}\xi_{24} + \xi_{14}\xi_{23}.$$

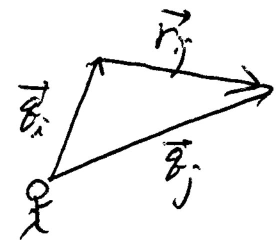
For a single-variate Gaussian, we recover $\langle x^4 \rangle = \langle x^2 \rangle^2 + \langle x^2 \rangle^2 + \langle x^2 \rangle^2 = 3\sigma^2$ on the previous page

1.2 Homogeneity and isotropy

Let $\vec{q}$ be spatial coordinates. Then, random variables at a given spatial position will be given by $X(\vec{q})$.
The 2-point correlation function between 2 points $\vec{q}_i$ & $\vec{q}_j$ is

$$\xi_{ij} = \langle X(\vec{q}_i) X(\vec{q}_j) \rangle$$

Let $\vec{r}_{ij}$ be a vector connecting 2 points.



$$\xi_{ij} = \langle X(\vec{q}_i) X(\vec{q}_i + \vec{r}_{ij}) \rangle$$

Thus, ξ_{ij} would depend on $\vec{q}_i$ as well as $\vec{r}_{ij}$.

☆ Homogeneity (translational invariance)
- Statistical homogeneity demands ξ_{ij} **not** depend on $\vec{q}_i$. $\therefore$ $\xi_{ij} \neq \xi_{ij}(\vec{q}_i, \vec{r}_{ij})$, but $\xi_{ij} = \xi_{ij}(\vec{r}_{ij})$

☆ Isotropy (rotational invariance)
- When seen from a given point $\vec{q}_i$, ξ_{ij} does not depend on the direction of $\vec{r}_{ij}$. $\therefore$ $\xi_{ij} = \xi_{ij}(\vec{q}_i, |\vec{r}_{ij}|)$

☆ Homogeneity & Isotropy :
- If ξ_{ij} is isotropic from **any** points in space, then it must be also homogeneous. We believe that we live in a universe which is isotropic & homogeneous. (But we must keep testing this hypothesis - it's fundamental enough for it to be tested repeatedly!)

$$\boxed{\xi_{ij} \equiv \xi_{ij}(|\vec{r}_{ij}|)}$$

translational & rotational invariance

Let us consider a Fourier transform of $x(\vec{q})$,

$$\tilde{x}(\vec{k}) \equiv \int d^3q\, e^{-i\vec{k}\cdot\vec{q}}\, x(\vec{q})$$

and consider the covariance matrix given by

$$C_{ij} \equiv \langle \tilde{x}(\vec{k}_i)\, \tilde{x}^*(\vec{k}_j) \rangle$$

The translational invariance demands

$$C_{ij} = (2\pi)^3\, \delta_D^{(3)}(\vec{k}_i - \vec{k}_j)\, \underbrace{P(\vec{k}_j)}_{\text{"power spectrum"}}$$

(not to be confused with PDF!)

The rotational invariance further demands

$$P(\vec{k}_j) \rightarrow P(|\vec{k}_j|)$$

Proof

$$C_{ij} = \int d^3q \int d^3q'\, e^{-i\vec{k}_i\cdot\vec{q}}\, e^{i\vec{k}_j\cdot\vec{q}'} \langle x(\vec{q})\, x(\vec{q}') \rangle$$

$\vec{q}' = \vec{q} + \vec{r}$

$$= \int d^3q \int d^3r\, e^{-i(\vec{k}_i - \vec{k}_j)\cdot\vec{q}}\, e^{i\vec{k}_j\cdot\vec{r}} \langle x(\vec{q})\, x(\vec{q}+\vec{r}) \rangle$$

Translational Invariance : $\langle x(\vec{q})\, x(\vec{q}+\vec{r}) \rangle = \xi(\vec{r})$

$$= \underbrace{\int d^3r\, e^{i\vec{k}_j\cdot\vec{r}}\, \xi(\vec{r})}_{\equiv P(\vec{k}_j)}\ \underbrace{\int d^3q\, e^{-i(\vec{k}_i - \vec{k}_j)\cdot\vec{q}}}_{= (2\pi)^3 \delta_D^{(3)}(\vec{k}_i - \vec{k}_j)}$$

Rotational Invariance : $\xi(\vec{r}) = \xi(|\vec{r}|) \Rightarrow P(\vec{k}_j) = P(|\vec{k}_j|)$

Therefore, $C_{ij} = (2\pi)^3 \delta_D^{(3)}(\vec{k}_i - \vec{k}_j)\, P(|\vec{k}_j|)$

Q.E.D.

We can write down the PDF of $\tilde{x}(\vec{k})$ as well :

$$P(\tilde{x}(\vec{k}_1), \cdots, \tilde{x}(\vec{k}_N)) = \frac{1}{(2\pi)^{N/2}|C|^{1/2}} e^{-\frac{1}{2}\sum_{ij} \tilde{x}(\vec{k}_i)(C^{-1})_{ij}\tilde{x}^*(\vec{k}_j)}$$

For a translationally and rotationally invariant x, we have

$$P(\tilde{x}(\vec{k}_1), \cdots, \tilde{x}(\vec{k}_N)) = \frac{1}{(2\pi)^{N/2}\left(\prod_i P(|\vec{k}_i|)\right)^{1/2}} e^{-\frac{1}{2}\sum_i \frac{|\tilde{x}(\vec{k}_i)|^2}{P(|\vec{k}_i|)}}$$

which is much simplified.

For this reason, we often deal with the Fourier quantities such as the power spectrum, $P(|\vec{k}_i|)$, when we analyze the cosmological data sets, as we belive that we live in a statistically homogeneous and isotropic universe.

= Bispectrum =

We can expand the above argument, and define the "bispectrum," given by

$$\langle \tilde{x}(\vec{k}_i)\tilde{x}(\vec{k}_j)\tilde{x}(\vec{k}_\ell)\rangle \quad (= 0 \text{ for a Gaussian } \tilde{x})$$

Translational and rotational invariance then demand that this quantity be

$$\langle \tilde{x}(\vec{k}_i)\tilde{x}(\vec{k}_j)\tilde{x}(\vec{k}_\ell)\rangle = (2\pi)^3 \delta_D^{(3)}(\vec{k}_i + \vec{k}_j + \vec{k}_\ell) B(|\vec{k}_i|, |\vec{k}_j|, |\vec{k}_\ell|)$$

More later.

1.3. Non-Gaussian Statistics

Gaussian statistics is nice because we know the form of PDF, either in real space or Fourier space (or both). What do we do when we have to deal with "non-Gaussian" statistics?

In general, there is no optimal approach when we do not know anything about PDF.

However, in cosmology, we know that the statistical properties of the primordial fluctuations (e.g., as seen in CMB) are well-described by a Gaussian distribution, and thus a departure from a Gaussian distribution, if any, must be small.

In this case, we can obtain a reasonable guess for the PDF of non-Gaussian fluctuations by <u>perturbing a Gaussian PDF</u>.

Recap: Taylor expansion of a function $f(x)$ around $x=0$ is given by

$$f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \frac{d^n f}{dx^n}$$

We shall apply this to a Gaussian PDF.

Gram-Charlier Expansion

[See Blinnikov & Hoessner, Astronomy & Astrophysics Supplement, 130, 193 (1998) for a review.]

Consider a Gaussian PDF with a unit variance, $\langle x^2 \rangle = 1$ (and zero mean). Let's call it $G(x)$:

$$G(x) \equiv \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Then, suppose that we have a PDF of a weakly non-Gaussian random variable, and want to describe this new PDF, $P(x)$, as a perturbation to $G(x)$. We obtain

$$P(x) = \sum_{n=0}^{\infty} C_n \frac{d^n G}{dx^n}$$

$$= G(x) \left[C_0 + (-1) C_1 x + C_2 (x^2 - 1) + C_3 (-1)(x^3 - 3x) + \cdots \right.$$

This is the so-called "Gram-Charlier expansion."

Or, defining the Chebyshev-Hermite polynomials as

$$He_n(x) \equiv (-1)^n \frac{1}{G(x)} \frac{d^n G}{dx^2}$$

$$= (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}$$

we have

$$He_0(x) = 1$$
$$He_1(x) = x$$
$$He_2(x) = x^2 - 1$$
$$He_3(x) = x^3 - 3x$$
$$He_4(x) = x^4 - 6x^2 + 3.$$

go to next page!! $\Longrightarrow$

The PDF can be written as

$$P(x) = G(x)\left[\sum_{n=0}^{\infty} c_n (-1)^n He_n(x)\right] \quad -(*)$$

In other words, we are expanding the ratio, $P(x)/G(x)$, in terms of the Chebyshev-Hermite polynomials times $(-1)^n$.

$$\frac{P(x)}{G(x)} = \sum_{n=0}^{\infty} c_n \underline{(-1)^n He_n(x)}$$

↳ basis function for P/G.

A nice property of $He_n(x)$. It satisfies the following relation:

$$\int_{-\infty}^{\infty} dx\, G(x) He_n(x) He_m(x) = m!\,\delta_{nm}.$$

which allows us to systematically derive the coefficients, c_n.

First, multiply both sides of $(*)$ by $He_n(x)$, and integrate over x.

$$\begin{aligned}\int_{-\infty}^{\infty} dx\, P(x) He_n(x) &= \sum_{m=0}^{\infty} c_m (-1)^m \int_{-\infty}^{\infty} dx\, G(x) He_m(x) He_n(x) \\ &= \sum_{m=0}^{\infty} c_m (-1)^m\, m!\,\delta_{nm} \\ &= (-1)^n n!\, c_n\end{aligned}$$

$$\boxed{\therefore\ c_n = \frac{(-1)^n}{n!} \int_{-\infty}^{\infty} dx\, P(x) He_n(x)}$$

next page ⟹

Recalling $He_0(x) = 1$; $He_1(x) = x$; $He_2(x) = x^2 - 1$; $He_3(x) = x^3 - 3x$
$He_4(x) = x^4 - 6x^2 + 3$; etc., we get

$$C_0 = \int_{-\infty}^{\infty} P(x)\, dx = 1 \quad \text{[normalization of PDF]}$$

$$C_1 = -\int_{-\infty}^{\infty} dx\, x P(x) = 0 \quad \text{[zero mean]}$$

$$C_2 = \frac{1}{2!}\int_{-\infty}^{\infty} dx\, (x^2-1) P(x) = \frac{1}{2}(\langle x^2 \rangle - 1)$$

$$= 0 \quad \text{[for unit variance]}$$

$$C_3 = \frac{-1}{3!}\int_{-\infty}^{\infty} dx\, (x^3 - 3x) P(x) = -\frac{1}{6}\langle x^3 \rangle \equiv -\frac{1}{6}\kappa_3$$

[where κ_3 is the "skewness"]
$\kappa_3 \equiv \langle x^3 \rangle$

$$C_4 = \frac{1}{4!}\int_{-\infty}^{\infty} dx\, (x^4 - 6x^2 + 3) P(x)$$

$$= \frac{1}{24}(\langle x^4 \rangle - 3) \equiv \frac{1}{24}\kappa_4$$

[where κ_4 is the "kurtosis"]
$\kappa_4 \equiv \langle x^4 \rangle - 3\sigma^4$

$\vdots$

Therefore, the expansion coefficients are the moments of $P(x)$.
Collecting all terms the Gram-Charlier expansion of a zero-mean, unit-variance PDF is given by

$$\boxed{P(x) = G(x)\left[1 + \frac{1}{6}\kappa_3 He_3(x) + \frac{1}{24}\kappa_4 He_4(x) + \cdots\right]}$$

Edgeworth Expansion

[See, e.g., Bernardeau & Kofman, Astrophysical Journal, 443, 479 (1995)]

Now, let us extend this result to the case with $\sigma \neq 1$.
It turns out that the following quantity:

$$S_n \equiv \frac{\kappa_n}{\sigma^{2n-2}}$$

is a good expansion parameter. (Note that S_n is not necessarily dimensionless, if κ_n is not dimensionless.)

Let y be a new random variable.
The expansion then becomes:

$$P(y) = \frac{1}{\sigma} G(y/\sigma) \left[1 + \sigma \frac{S_3}{6} He_3(y/\sigma) + \sigma^2 \frac{S_4}{24} He_4(y/\sigma) + \cdots \right]$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma} e^{-y^2/2\sigma^2} \left[1 + \sigma \frac{S_3}{6}\left(\frac{y^3}{\sigma^3} - 3\frac{y}{\sigma}\right) + \sigma^2 \frac{S_4}{24}\left(\frac{y^4}{\sigma^4} - 6\frac{y^2}{\sigma^2} + 3\right) + \cdots \right]$$

This is known as the "Edgeworth Expansion."

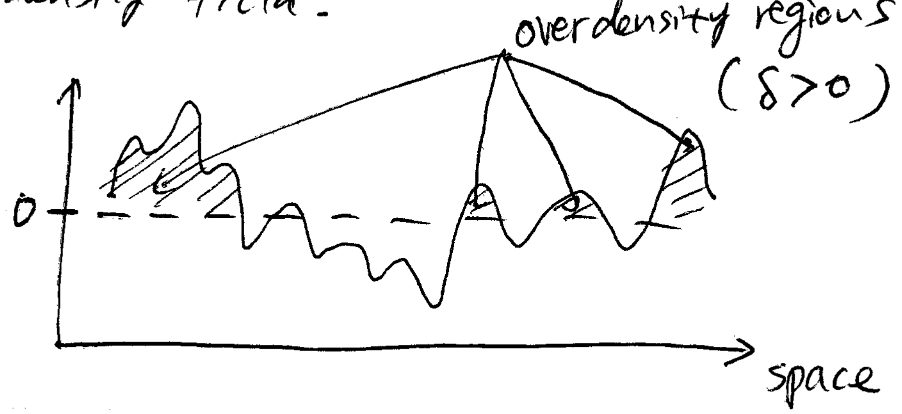
1.4. Application : Mass Function

[See Press & Schechter, Astrophysical Journal, 187, 425 (1974), LoVerde et al, Journal of Cosmology and Astroparticle Physics, 04, 014 (2008)]

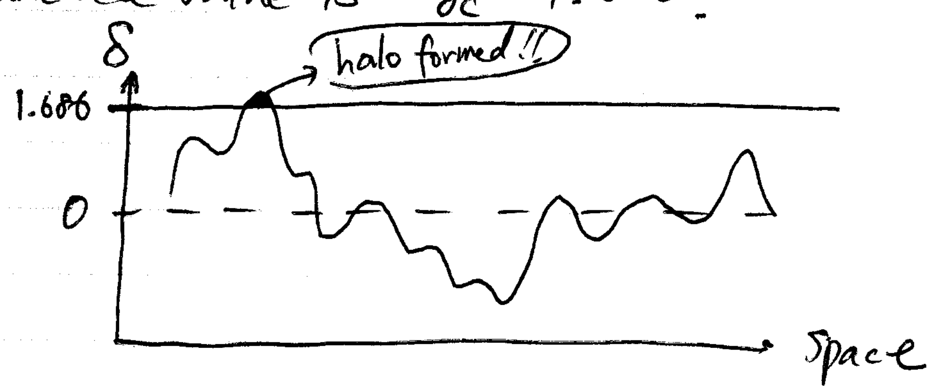
★ In cosmology, we believe that collapsed "objects" such as dark-matter haloes are formed at the locations of high density peaks.

Thus, within this picture, the (statistical) distribution of objects such as the number density of halos can be derived from the PDFs of the underlying density field.

$$\delta \equiv \frac{\rho - \langle\rho\rangle}{\langle\rho\rangle}$$



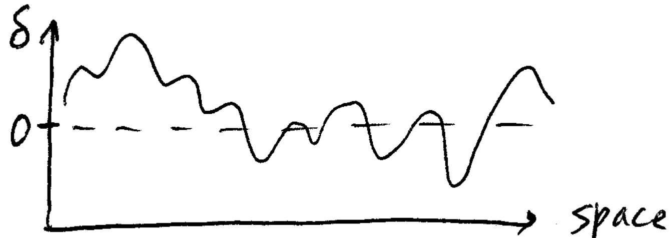
In the simplest picture, halos are formed at the locations where δ exceeds a certain "threshold value," δ_c, and the numerical value is $\delta_c \cong 1.686$.



But, what about a mass? The standard approach is to go to sufficiently early time, such that δ is very small, $\delta \ll 1$. Then, the mass enclosed within a region with a radius R is given by

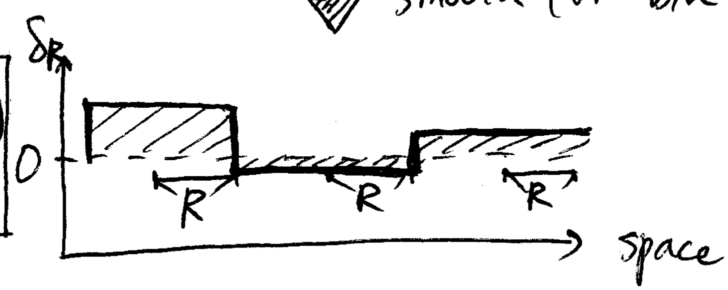
$$M = \frac{4\pi}{3}\langle\rho\rangle R^3 .$$

Then, we smooth the density field with a top-hat function with a radius R. (In other words, we bin the density field with a radius R.)

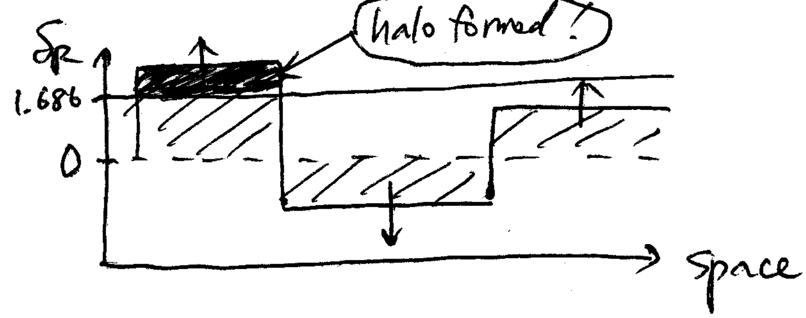


smooth (or bin)

$$\boxed{\delta_R(\vec q) \equiv \frac{1}{\frac{4\pi R^3}{3}} \int_{|\vec r| \le R} d^3r\, \delta(\vec q + \vec r)}$$



Then, watch δ_R growing, until δ_R exceeds $\delta_c = 1.686$. An object with a mass $M = \frac{4\pi}{3}\langle\rho\rangle R^3$ is formed.



According to Press & Schechter, the mass function (the number density of collapsed objects per unit mass bin) is given by

$$\frac{dn}{dM} = -2\frac{\langle\rho\rangle}{M}\frac{d}{dM}\int_{\delta_c}^{\infty} d\delta_R\, P(\delta_R)$$

For example, for a Gaussian distribution $P(\delta_R)=\frac{1}{\sqrt{2\pi}\sigma_R}e^{-\delta_R^2/2\sigma_R}$,

$$\boxed{\frac{dn}{dM} = 2\frac{\langle\rho\rangle}{M}\delta_c\frac{d\sigma_R^{-1}}{dM}\frac{1}{\sqrt{2\pi}}e^{-\delta_c^2/2\sigma_R^2}}$$

This is the so-called the Press-Schechter mass function.
This mass function is normalized such that

$$\int_0^\infty dM\, M\frac{dn}{dM} = \langle\rho\rangle.$$

That is to say, all the mass in the universe is accounted for by halos, once we integrate down to zero mass. This assumption can be debated.

Now, let us see what we would get if we use an Edgeworth-expanded PDF:

$$P(\delta_R) = \frac{1}{\sqrt{2\pi}\sigma_R}e^{-\delta_c^2/2\sigma_R^2}\left[1+\sigma_R\frac{S_3}{6}\left(\frac{\delta_R^3}{\sigma_R^3}-3\frac{\delta_R}{\sigma_R}\right)+\cdots\right]$$

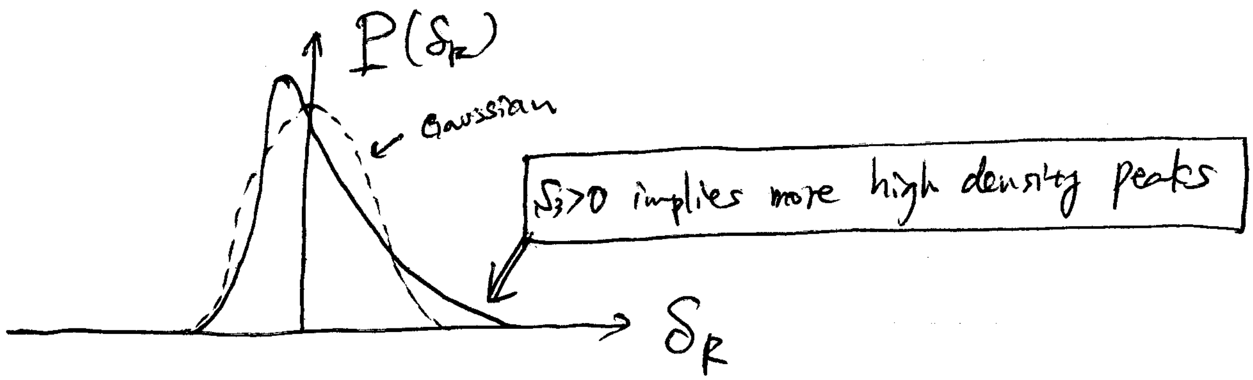
$\Longrightarrow$ next page

For a non-Gaussian PDF, we obtain

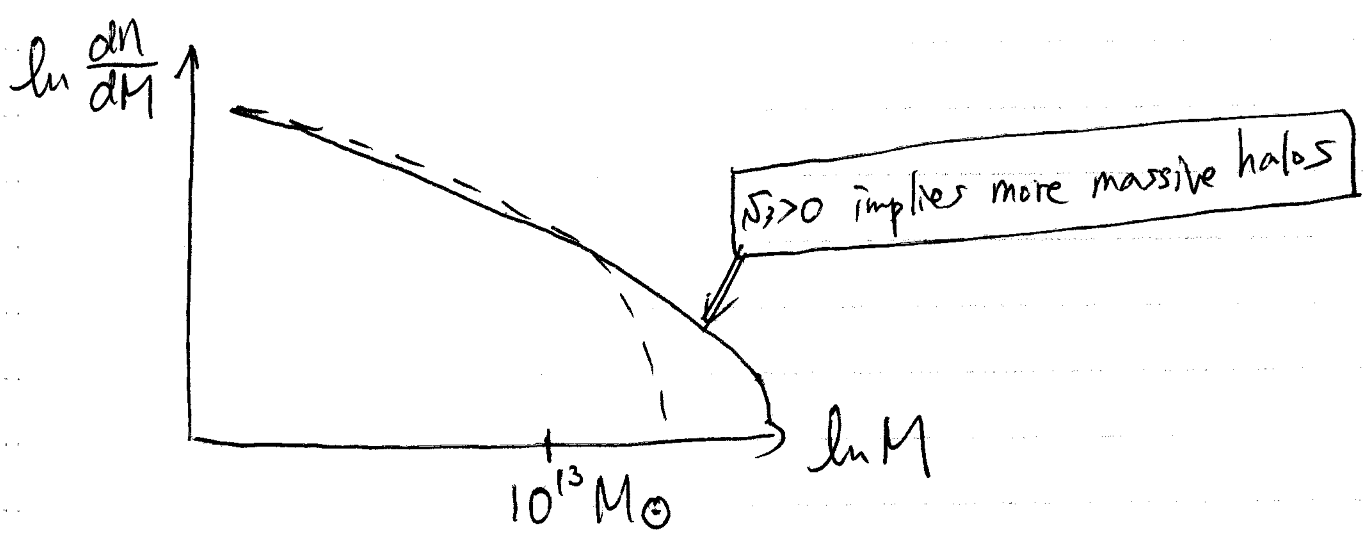
$$\boxed{\frac{dn}{dM} = \sqrt{\frac{2}{\pi}}\frac{\langle\rho\rangle}{M} e^{-\delta_c^2/2\sigma_R^2}\left\{\delta_c \frac{d\sigma_R^{-1}}{dM}\left[1+\sigma_R\frac{S_3}{6}\left(\frac{\delta_c^3}{\sigma_R^3}-2\frac{\delta_c}{\sigma_R}-\frac{\sigma_R}{\delta_c}\right)\right] - \sigma_R\frac{1}{6}\frac{dS_3}{dM}\left(\frac{\delta_c^2}{\sigma_R^2}-1\right)\right.}$$

[LoVerde et al., JCAP, 04, 014 (2008)]

★ Positive Skewness, $S_3 > 0$



The mass function looks like (at $z=0$):



2. Non-Gaussianity as a test of inflation

2.1. Power Spectrum

Inflation produces perturbations in the spatial part of the metric, δg_{ij}. The scalar part of δg_{ij} will seed the structures we see today; while the tensor part of δg_{ij} will propagate as gravitational waves.

Let us write: $$g_{ij} = e^{2\zeta}\gamma_{ij}\, a^2(t)$$

Here,
$$\begin{cases} e^{2\zeta} = 1 + 2\underbrace{\zeta}_{=\text{"Curvature Perturbation"}} + \cdots \\ \gamma_{ij} = \delta_{ij} + \underbrace{h_{ij}}_{=\text{"gravitational wave"}} \end{cases}$$

> **NOTE on subtlety!** We normally write the metric as $g_{ij} = e^{2\Phi}\gamma_{ij}a^2(t)$ and define ζ as $\zeta = \Phi + \int \frac{d\rho}{3(\rho+p)}$. But, we use the uniform density gauge.

In this lecture, we will study the statistical properties of ζ. It is convenient to recall the following relation between ζ and the observables.

(a) CMB Temperature Anisotropy on very large angular scales where the "Sachs-Wolfe" approximation is valid:
$$\frac{\delta T}{T}(\hat{n}) = -\frac{1}{5}\zeta(\hat{n}r_*) \quad \left(\begin{array}{l}\text{where } r_* \text{ is the conformal} \\ \text{distance to } z_* = 1090\end{array}\right)$$

(b) Newtonian gravitational potential during the matter era:
$$\tilde{\Psi}_{\vec{k}} = -\frac{3}{5}\tilde{\zeta}_{\vec{k}}\, T(k) \quad \left(\begin{array}{l}\text{where } T(k) \text{ is the} \\ \text{linear transfer function}\end{array}\right)$$

(tildes denote variables in Fourier space)

Inflation usually predicts a nearly power-law power spectrum. Defining

$$\langle \tilde{\zeta}_{\vec{k}} \tilde{\zeta}^{*}_{\vec{k}'} \rangle = (2\pi)^3 \delta_D^{(3)}(\vec{k}-\vec{k}') P_\zeta(\vec{k})$$

We write $P_\zeta(k) \equiv P_\zeta(|\vec{k}|) \propto k^{n_s-4}$.

Inflation predicts $n_s \simeq 1$, and it is called a "scale-invariant spectrum." Why?

To see why $n_s = 1$ is called "scale-invariant," let us calculate the real-space 2-point correlation function:

$$\xi_\zeta(r) \equiv \langle \zeta(\vec{x})\zeta(\vec{x}+\vec{r}) \rangle = \int \frac{d^3k}{(2\pi)^3} P_\zeta(k) e^{i\vec{k}\cdot\vec{r}}$$

$$= \int_0^\infty \frac{dk}{k} \cdot \frac{k^3 P_\zeta(k)}{2\pi^2} \cdot \frac{\sin(kr)}{kr}$$

Therefore, $\xi_\zeta(r)$ is given by

$$\boxed{\xi_\zeta(r) \simeq \left.\frac{k^3 P_\zeta(k)}{2\pi^2}\right|_{k \simeq \frac{1}{r}} \propto r^{1-n_s}.}$$

Thus, for $n_s = 1$, $\xi_\zeta(r)$ does not depend on r.
(To be more precise $\xi_\zeta(r)$ depends on r logarithmically.)
Hence the term "scale invariant."

The latest observations gave $\underline{\underline{n_s = 0.96 \pm 0.01 \ (68\% \ \mathrm{CL})}}$

(e.g., Keisler et al. arXiv:1105.3182)

2.2. Bispectrum

Simple models of inflation satisfying all of the following conditions

1. Single field.
2. Canonical kinetic term,
3. Slow-roll, and
4. Vacuum initial state

produce a tiny amount of non-Gaussianity – too tiny to be detectable by any experiments. Specifically, writing the bispectrum as

$$\langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_3} \rangle = (2\pi)^3 \delta_D^{(3)}(\vec{k}_1 + \vec{k}_2 + \vec{k}_3) B_\zeta(k_1, k_2, k_3),$$

the above simple models tend to produce

$$\frac{B_\zeta(k_1, k_2, k_3)}{P_\zeta(k_1)P_\zeta(k_2) + (\text{permutations})} = \mathcal{O}(10^{-2})$$

which happens to be unobservably small.

However, if any of the above four conditions were violated, inflation could produce much larger non-Gaussianity.

IMPORTANT

This is an exciting prospect: detection of non-Gaussianity above $\mathcal{O}(10^{-2})$ (in the above sense) would rule out the simplest models, and call for richer physics of inflation!!

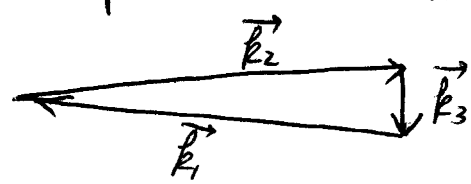
Since $B_\zeta(k_1, k_2, k_3)$ has 3 variables to play with, we have many <u>shapes</u> to play with.

Imposing a scale invariance allows us to reduce the number of degrees of freedom to 3:

$$\frac{k_2}{k_1}, \quad \frac{k_3}{k_1}, \quad \text{and angle between } k_2 \,\&\, k_3.$$

Among various shapes considered in the literature, let us pick 3 which are associated with the conditions mentioned in the previous page:

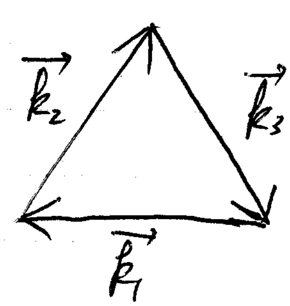
~~Single field~~ 1. "Squeezed Shape" $|\vec{k}_3| \ll |\vec{k}_2| \approx |\vec{k}_1|$



Detection of this shape would rule out <u>all</u> single-field models

~~Canonical Kinetic Term~~ or other field interactions

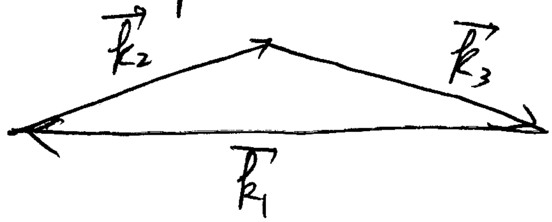
2. "Equilateral Shape" $|\vec{k}_3| \approx |\vec{k}_2| \approx |\vec{k}_1|$



Detection of this shape indicates enhanced field interactions at the horizon exit. This can be achieved by, e.g. non-canonical kinetic terms.

~~Initial Vacuum State~~

4. "Folded Shape" $|\vec{k}_3| \approx |\vec{k}_2| \approx |\vec{k}_1|/2$



Detection of this shape would suggest that the initial state of quantum fluctuations in scalar fields were **not** in the vacuum state.

~~Slow roll~~

3. More complex behavior

When slow roll is violated, it often induces a violation of scale-invariance, and thus more complex shapes are possible, such as an oscillating bispectrum.

In this lecture, we shall focus on 1, the squeezed triangle.

2.3. Multi-field Inflation & Local-form Bispectrum

Perhaps, the most important goal of measuring the squeezed bispectrum is to rule out single-field models of inflation.

How do we then calculate the bispectrum of the curvature perturbation, ζ, from multi-field inflation?
For this purpose, there is a very convenient and powerful method, called the "δN formalism."

Here, "N" is the number of e-folds of expansion.

$$N \equiv \ln \frac{a(t_{obs})}{a(t_{ini})}$$

By this, we really mean the time after which ζ is conserved. More later

where t_{obs} is the time of "observation," and t_{ini} is some initial time.

Now, looking at the metric (spatial components, ignoring gravitational waves),

$$g_{ij} = a^2(t) e^{2\zeta(\vec{x})} \delta_{ij} \qquad [\vec{x} \text{ is the comoving coordinates}]$$

One can define an inhomogeneous scale factor,

$$\hat{a}(t, \vec{x}) \equiv a(t) e^{\zeta(\vec{x})}$$

IMPORTANT!!

☆ Then, if we choose the initial time to be on the flat hypersurface, in which $g_{ij} = a^2(t)\delta_{ij}$, then

$$\hat{N} = \ln \frac{\hat{a}(t_{obs}, \vec{x})}{a(t_{ini})} = \ln \frac{a(t_{obs})}{a(t_{ini})} + \zeta(\vec{x})$$

Thus,

$$\boxed{\zeta(\vec{q}) = \delta N(\vec{q}) \equiv \hat{N}(t_{obs}, t_{ini}, \vec{q}) - N(t_{obs}, t_{ini})}$$

this is the δN formalism. To be more precise,

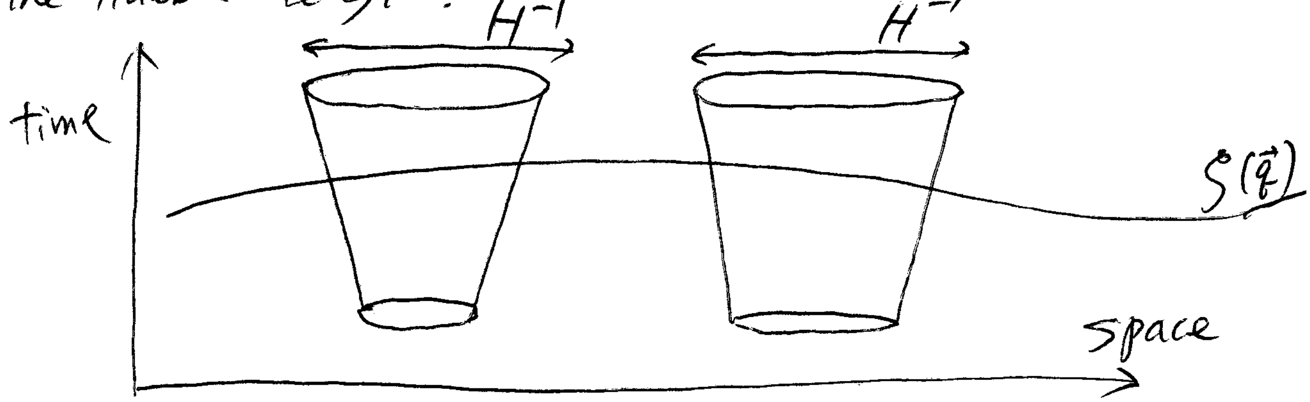
① t_{ini} is on the flat hypersurface, in which $g_{ij} = a^2(t)\delta_{ij}$.

② t_{obs} is on the uniform density hypersurface, in which $\delta\rho = 0$.

[Ref: Salopek & Bond, Physical Review D, 42, 3936 (1990)
Sasaki & Stewart, Progress in Theoretical Physics, 95, 71 (1996)
Lyth, Malik & Sasaki, JCAP, 05, 004 (2005)]

The next question is, "how do we compute $\delta N(\vec{q})$?"

To calculate $\delta N(\vec{q})$, we can use the so-called "gradient expansion method," for which we systematically ignore the terms that involve spatial derivatives. This is a valid approximation, as long as the wavelength of perturbations we deal with is much greater than the Hubble length.



To the zeroth-order in the gradient expansion, in which all the spatial derivatives are ignored, something remarkable happens.

The Friedmann equation for the background (mean)

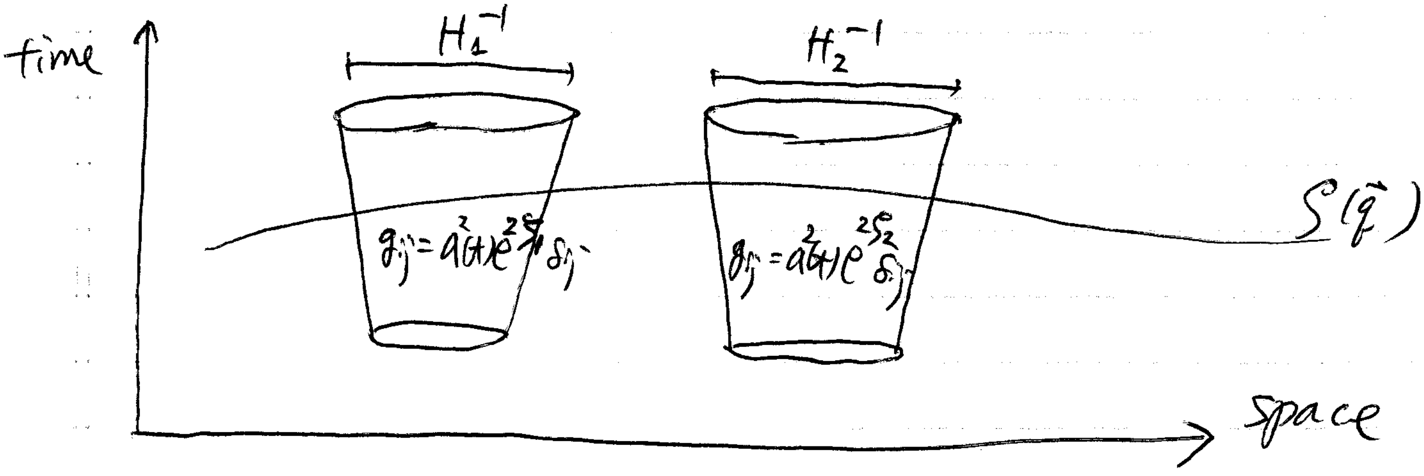
$$3M_{pl}^2 \langle H \rangle^2 = \langle \rho \rangle$$

is the same as the Friedmann equation for the perturbed quantities !!

$$3M_{pl}^2 \hat{H}^2 = \hat{\rho}$$

This is completely non-trivial result.
[Ref: Salopek & Bond, PRD, 42, 3936 (1990);
Lyth, Malik & Sasaki, JCAP, 05, 004 (2005)]

Naively, one can say "when the wavelength of ζ is much greater than the Hubble horizon, each horizon patch evolves as if they were a separate Universe."



PUNCH LINE

☆ Under this approximation, all we need to do is to solve the background evolution :

$$3M_{pl}^2 \langle H \rangle^2 = \langle \rho \rangle$$

for different initial conditions, and calculate perturbations as the differences between initial conditions.

Namely :

I^{th} field

$$\hat{N}(\vec{q}) = \hat{N}\left(\varphi^{I}_{ini}(\vec{q}),\ \dot{\varphi}^{I}_{ini}(\vec{q})\right)$$

$$= \int dt\ \hat{H}\left(\varphi^{I}_{ini}(\vec{q}),\ \dot{\varphi}^{I}_{ini}(\vec{q}), t\right)$$

In order for φ to contribute to ρ (hence H), φ should be slow-rolling. Then, $\dot{\varphi}$ is a function of φ, and thus we obtain

$$\boxed{\zeta(\vec{q}) = \delta N(\vec{q}) = \sum_I \frac{\partial N}{\partial \varphi^I_{ini}} \delta\varphi^I_{ini}(\vec{q}) + \frac{1}{2}\sum_{IJ} \frac{\partial^2 N}{\partial \varphi^I_{ini} \partial \varphi^J_{ini}} \delta\varphi^I_{ini}(\vec{q})\, \delta\varphi^J_{ini}(\quad + \cdots}$$

<u>This completes the δN formalism.</u>

Single-field check

Let us apply the δN formalism to the single-field case, to make sure that we recover the well-known result:

$$\boxed{\zeta = -\left.\frac{H}{\dot\phi}\right|_{ini} \delta\phi_{ini}}$$

(Here, $\phi \equiv \varphi^1$ because we have a single-field model.)

$$\hat{N} = \int_{t_{ini}}^{t_{obs}} dt\, \hat{H} = \int_{\phi_{ini}}^{\phi_{obs}} d\phi \frac{\hat{H}}{\dot\phi}$$

Now, t_{obs} is on the uniform density hypersurface, so

$$\phi_{obs} = \langle\phi\rangle_{obs}.$$

On the other hand, t_{ini} is on the flat hypersurface, and thus

$$\phi_{ini} = \langle\phi\rangle_{ini} + \delta\phi_{ini}.$$

Therefore

$$\hat{N} = \int_{\langle\phi\rangle_{ini}+\delta\phi_{ini}}^{\langle\phi\rangle_{obs}} d\phi \frac{\hat{H}}{\dot\phi}$$

$$\simeq \left(\int_{\langle\phi\rangle_{ini}}^{\langle\phi\rangle_{obs}} d\phi \frac{\langle H\rangle}{\dot\phi}\right) - \left.\frac{\langle H\rangle}{\dot\phi}\right|_{ini} \delta\phi_{ini} + \mathcal{O}(\delta\phi^2)$$

$$\therefore\ \zeta = \delta N = \hat{N} - \langle N\rangle = -\left.\frac{H}{\dot\phi}\right|_{ini} \delta\phi_{ini}$$

Q.E.D.

Local-form Non-Gaussianity

Look at the δN form:

$$\zeta(\vec{q}) = \frac{\partial N}{\partial \varphi^I}\delta\varphi^I(\vec{q}) + \frac{1}{2}\frac{\partial^2 N}{\partial\varphi^I\partial\varphi^J}\delta\varphi^I(\vec{q})\,\delta\varphi^J(\vec{q}) + \cdots$$

(Here, we omitted the subscript "ini" for clarity, and assumed that the repeated indices mean the summation.)

This form is "local" because the left hand side and the right hand side are evaluated at the same comoving position, $\vec{q}$.

IMPORTANT!

Also, this form implies that ζ must be non-Gaussian **EVEN IF** $\delta\varphi^I(\vec{q})$'s are Gaussian!

To see this, compute the real-space 3-point function:

$$\langle \zeta(\vec{q}_1)\zeta(\vec{q}_2)\zeta(\vec{q}_3)\rangle = \frac{\partial N}{\partial\varphi^I}\frac{\partial N}{\partial\varphi^J}\frac{\partial N}{\partial\varphi^K}\langle \delta\varphi^I(\vec{q}_1)\,\delta\varphi^J(\vec{q}_2)\,\delta\varphi^K(\vec{q}_3)\rangle$$
$$+\frac{1}{2}\frac{\partial N}{\partial\varphi^I}\frac{\partial N}{\partial\varphi^J}\frac{\partial^2 N}{\partial\varphi^K\partial\varphi^L}\langle\delta\varphi^I(\vec{q}_1)\delta\varphi^J(\vec{q}_2)\delta\varphi^K(\vec{q}_3)\delta\varphi^L(\vec{q}_3)\rangle$$
$$+\cdots$$

The 1st term vanishes if $\delta\varphi^I$'s are Gaussian. However, the 2nd term does not (recall the Wick's theorem).

Now, let's compute the bispectrum. $\Longrightarrow$ next page

In Fourier space,

$$\tilde{\zeta}_{\vec{k}} = N_I \widetilde{\delta\varphi}^I_{\vec{k}} + \frac{1}{2} N_{IJ} \int \frac{d^3p}{(2\pi)^3} \widetilde{\delta\varphi}^I_{\vec{k}-\vec{p}} \widetilde{\delta\varphi}^J_{\vec{p}} + \cdots$$

(Here, we introduce a short-hand notation, $N_I \equiv \partial N/\partial\varphi^I$, $N_{IJ} \equiv \partial^2 N/\partial\varphi^I\partial\varphi^J$, etc.)

$$\begin{aligned}\langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_3} \rangle = {} & N_I N_J N_K \langle \widetilde{\delta\varphi}^I_{\vec{k}_1} \widetilde{\delta\varphi}^J_{\vec{k}_2} \widetilde{\delta\varphi}^K_{\vec{k}_3} \rangle \\ & + \frac{1}{2} N_I N_J N_{KL} \int \frac{d^3p}{(2\pi)^3} \langle \widetilde{\delta\varphi}^I_{\vec{k}_1} \widetilde{\delta\varphi}^J_{\vec{k}_2} \widetilde{\delta\varphi}^K_{\vec{k}_3-\vec{p}} \widetilde{\delta\varphi}^L_{\vec{p}} \rangle \\ & \quad + (2 \text{ permutations}) \\ & + \cdots\end{aligned}$$

Using Wick's theorem, the second term becomes

$$\begin{aligned}(2\text{nd}) = \frac{1}{2} N_I N_J N_{KL} \int \frac{d^3p}{(2\pi)^3} \Big\{ & \langle \widetilde{\delta\varphi}^I_{\vec{k}_1} \widetilde{\delta\varphi}^J_{\vec{k}_2} \rangle \langle \widetilde{\delta\varphi}^K_{\vec{k}_3-\vec{p}} \widetilde{\delta\varphi}^L_{\vec{p}} \rangle \\ & + \langle \widetilde{\delta\varphi}^I_{\vec{k}_1} \widetilde{\delta\varphi}^K_{\vec{k}_3-\vec{p}} \rangle \langle \widetilde{\delta\varphi}^J_{\vec{k}_2} \widetilde{\delta\varphi}^L_{\vec{p}} \rangle \\ & + \langle \widetilde{\delta\varphi}^I_{\vec{k}_1} \widetilde{\delta\varphi}^L_{\vec{p}} \rangle \langle \widetilde{\delta\varphi}^J_{\vec{k}_2} \widetilde{\delta\varphi}^K_{\vec{k}_3-\vec{p}} \rangle \Big\}\end{aligned}$$

Without loss of generality, we can write

$$\langle \widetilde{\delta\varphi}^I_{\vec{k}_1} \widetilde{\delta\varphi}^J_{\vec{k}_2} \rangle = (2\pi)^3 \delta_D^{(3)}(\vec{k}_1 + \vec{k}_2) \, \delta^{IJ} P_\varphi(k_1)$$

Then,

$$\begin{aligned}(2\text{nd}) = {} & \frac{1}{2} N_I N^I N_K{}^K (2\pi)^3 \delta_D^{(3)}(\vec{k}_1 + \vec{k}_2) \, \delta_D^{(3)}(\vec{k}_3) P_\varphi(k_1) P_\varphi(k_3) \int d^3p \; 1 \\ & + N_I N_J N^{IJ} (2\pi)^3 \delta_D^{(3)}(\vec{k}_1 + \vec{k}_2 + \vec{k}_3) P_\varphi(k_1) P_\varphi(k_2)\end{aligned}$$

However, the first line (involving $\delta_D^{(3)}(\vec{k}_3)$) vanishes because of $P_\varphi(0) = 0$.

$\Rightarrow$ next page

Therefore, if $\delta\varphi^I$'s are Gaussian, the leading-order result is

$$\langle \zeta_{\vec{k}_1}\zeta_{\vec{k}_2}\zeta_{\vec{k}_3}\rangle = (2\pi)^3\delta_D^{(3)}(\vec{k}_1+\vec{k}_2+\vec{k}_3)\, N_I N_J N^{IJ}\left(P_\varphi(k_1)P_\varphi(k_2) + \text{perm.}\right)$$

On the other hand, the power spectrum of ζ is related to that of $\delta\varphi$ as

$$\langle \zeta_{\vec{k}_1}\zeta^*_{\vec{k}_2}\rangle = (2\pi)^3\delta_D^{(3)}(\vec{k}_1-\vec{k}_2)P_\zeta(k_1) = N_I N_J\langle \delta\tilde{\varphi}^I_{\vec{k}_1}\delta\tilde{\varphi}^{J*}_{\vec{k}_2}\rangle$$
$$= (2\pi)^3\delta_D^{(3)}(\vec{k}_1-\vec{k}_2)\,N_I N^I P_\varphi(k_1)$$

$$\boxed{\therefore\ P_\zeta(k) = N_I N^I P_\varphi(k)}$$

Writing the bispectrum as $\langle \zeta_{\vec{k}_1}\zeta_{\vec{k}_2}\zeta_{\vec{k}_3}\rangle = (2\pi)^3\delta_D^{(3)}(\vec{k}_1+\vec{k}_2+\vec{k}_3)B_\zeta(k_1,k_2,k_3)$, we find

$$\boxed{\frac{B_\zeta(k_1,k_2,k_3)}{P_\zeta(k_1)P_\zeta(k_2) + \text{perm.}} = \frac{N_I N_J N^{IJ}}{(N_K N^K)^2}}$$

[Ref: Lyth & Rodriguez, PRL, 95, 121302 (2005)]

Let us compute this in the single-field limit. $(\phi \equiv \varphi^1)$

$$\frac{B_\zeta(k_1,k_2,k_3)}{P_\zeta(k_1)P_\zeta(k_2) + (\text{perm})} = \frac{\partial^2 N/\partial\phi^2_{ini}}{(\partial N/\partial\phi)^2_{ini}}$$

Here, $$\partial N/\partial\phi_{ini} = -\left.\frac{H}{\dot\phi}\right|_{ini} = -\frac{1}{M_{pl}}\frac{1}{\sqrt{2\varepsilon}}$$ [where $\varepsilon \equiv \frac{1}{2M_{pl}^2}\frac{\dot\phi^2}{H^2} = \mathcal{O}(10^{-2})$ is a slow-roll parameter,]

and thus

$$\frac{\partial^2 N/\partial\phi^2_{ini}}{\partial N/\partial\phi_{ini}} = \frac{1}{2}\frac{\partial(\ln\varepsilon)}{\partial\phi_{ini}}.$$

$\Rightarrow$ Next page

Therefore, for a single field model in which $\delta\varphi$ is Gaussian, we obtain

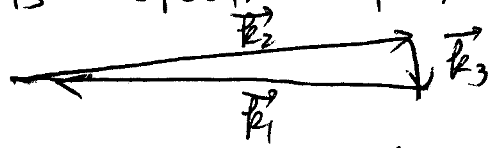
$$\boxed{\frac{B_\zeta(k_1,k_2,k_3)}{P_\zeta(k_1)P_\zeta(k_2)+(\text{perm.})} = \frac{1}{2}\,\frac{\partial(\ln\varepsilon)/\partial\phi_{ini}}{\partial N/\partial\phi_{ini}} = 2\varepsilon-\eta}$$

(where $\eta \equiv \varepsilon - \frac{\ddot{\phi}}{H\dot{\phi}}$ is another slow-roll parameter and it is also $\eta = \mathcal{O}(10^{-2})$.)

Indeed, we find $\frac{B_\zeta}{P_\zeta^2} = \mathcal{O}(10^{-2})$, which is too small to be observed, and thus detection of this quantity would rule out single-field slow-roll inflation models.

Squeezed Bispectrum

Since $P_\zeta(k)$ is nearly scale invariant, it scales as $P_\zeta(k) \propto 1/k^3$. This means that $B_\zeta(k_1,k_2,k_3)$ for the above local-form non-Gaussianity peaks when one of k_i's is very small. If we order $k_1 \geq k_2 \geq k_3$, then this bispectrum peaks when $k_3 \ll k_2 \approx k_1$.



For this limit, we find

$$B_\zeta(k_1,k_2,k_3\to 0) \to \frac{2N_I N_J N^{IJ}}{(N_K N^K)^2}\,P_\zeta(k_1)\,P_\zeta(k_3)$$

$$\left(= (4\varepsilon-2\eta)\,P_\zeta(k_1)\,P_\zeta(k_3) \text{ for single-fiel}\right.$$

(*)If this section is too confusing, ask me later.

2.4(*) Single-field Theorem (a.k.a consistency condition)

[Ref: Maldacena, JHEP, 05, 013 (2003); Creminelli & Zaldarriaga, JCAP, 10, 006 (2004); Ganc & Komatsu, JCAP, 12, 009 (2010)]

Actually, the squeezed-limit bispectrum is much more powerful than just ruling out single-field SLOW-ROLL models. In fact, it can be used to rule out ALL single-field models, regardless of the details of models.

This "single-field Theorem" states:

$$\boxed{B_\zeta(k_1, k_2, k_3 \to 0) \to (1 - n_s)\, P_\zeta(k_1) P_\zeta(k_3)}$$

for ALL single-field inflation models!

There are various ways to show this relation, but here we use a slightly non-standard approach, in order to continue on what we have been leaning so far.

⟹ next page!!

A good starting point is this formula we have seen before (see page 27)

$$\langle \zeta_{\vec{k}_1} \zeta_{\vec{k}_2} \zeta_{\vec{k}_3} \rangle = N_I N_J N_K \langle \delta\tilde{\varphi}^I_{\vec{k}_1} \delta\tilde{\varphi}^J_{\vec{k}_2} \delta\tilde{\varphi}^K_{\vec{k}_3} \rangle$$

$$+ N_I N_J N^{IJ} (2\pi)^3 \delta_D^{(3)}(\vec{k}_1 + \vec{k}_2 + \vec{k}_3) \left[P_\varphi(k_1) P_\varphi(k_2) + (\text{perm.}) \right]$$

While we have ignored the first term so far by assuming that $\delta\varphi^I$'s are Gaussian. However, for single-field models, the second term is small too (of order ε). At the level of $\mathcal{O}(\varepsilon)$, we cannot ignore the first term as there exist interactions of $\delta\varphi$ at the level of $\mathcal{O}(\varepsilon)$ during inflation, making $\delta\varphi$ slightly non-Gaussian.

While the precise computation of the first term requires Quantum field theory (and was done self-consistently by Maldacena (2003)), we make a short-cut by considering only the squeezed limit, $k_3 \ll k_2 \simeq k_1$.

To understand the situation better, let us call:

$$\left[\begin{array}{l} k_3 \Rightarrow k_L \text{ (for Long wavelength)} \\ k_1, k_2 \Rightarrow k_S \text{ (for Short wavelength)} \end{array} \right.$$

where "L" is much larger than H^{-1}, whereas "S" is much smaller.

$\Rightarrow$ next page.

Then, we can write the $\langle\rangle$ in the first term as

$$\langle \delta\tilde{\varphi}_{\vec{k}_1} \delta\tilde{\varphi}_{\vec{k}_2} \delta\tilde{\varphi}_{\vec{k}_3} \rangle \Rightarrow \left\langle \delta\tilde{\varphi}_{\vec{k}_S} \delta\tilde{\varphi}_{\vec{k}_S} \delta\tilde{\varphi}_{\vec{k}_L} \right\rangle$$

$$= \left\langle \langle \delta\tilde{\varphi}_{\vec{k}_S} \delta\tilde{\varphi}_{\vec{k}_S} \rangle_L \; \delta\tilde{\varphi}_{\vec{k}_L} \right\rangle$$

Here, $\langle \delta\tilde{\varphi}_{\vec{k}_S} \delta\tilde{\varphi}_{\vec{k}_S} \rangle_L$ is the 2-point function of the short mode, GIVEN THAT THERE EXISTS THE LONG MODE.

In single-field models, there is only one dynamical degree of freedom, ζ. Therefore, all we need to compute is $\langle \delta\tilde{\varphi}_{\vec{k}_S} \delta\tilde{\varphi}_{\vec{k}_S} \rangle$ in the presense of ζ_L.

Recalling $g_{ij} = a^2(t) e^{2\zeta_L}$, e^{ζ_L} amounts rescaling of comoving coordinates. In other words, ζ_L influences $\langle \delta\tilde{\varphi}_{\vec{k}_S} \delta\tilde{\varphi}_{\vec{k}_S} \rangle$ via the rescaling of the comoving coordinates

The result is

$$\boxed{\langle \delta\tilde{\varphi}_{\vec{k}_S} \delta\tilde{\varphi}_{\vec{k}_S} \rangle_L = \langle \delta\tilde{\varphi}_{\vec{k}_S} \delta\tilde{\varphi}_{\vec{k}_S} \rangle - \frac{d\ln k_S^3 P_\varphi(k_S)}{d\ln k_S} P_\varphi(k_S)\, \zeta_L}$$

For the single-field result $P_\varphi(k) \propto H^2/k^3$, and $dN = d\ln k$, we find

$$\frac{d\ln k^3 P_\varphi(k)}{d\ln k} = \frac{d\ln H^2}{d\phi}\frac{d\phi}{dN} = -2\varepsilon$$

$\Rightarrow$ next page

With this result, we find

$$\left(\frac{\partial N}{\partial \phi_{ini}}\right)^3 \langle \delta\varphi_{\vec{k}_1}\delta\varphi_{\vec{k}_2}\delta\varphi_{\vec{k}_3}\rangle \underset{(k_3\to 0)}{\Rightarrow} \left(\frac{\partial N}{\partial \phi_{ini}}\right)^3 (+2\varepsilon)\, P_\varphi(k_1) \langle \zeta_{\vec{k}_1+\vec{k}_2}\, \delta\varphi_{\vec{k}_3}\rangle$$

$$= +2\varepsilon\, P_\zeta(k_1) P_\zeta(k_3) \times (2\pi)^3 \delta_D^{(3)}(\vec{k}_1+\vec{k}_2+\vec{k}_3) \quad (\text{for } k_3\to 0)$$

Combining this with the second term, we finally arrive at

$$\langle \zeta_{\vec{k}_1}\zeta_{\vec{k}_2}\zeta_{\vec{k}_3}\rangle \xrightarrow{k_3\to 0} (2\pi)^3 \delta_D^{(3)}(\vec{k}_1+\vec{k}_2+\vec{k}_3) \times \left\{ +2\varepsilon\, P_\zeta(k_1)P_\zeta(k_3) + (4\varepsilon - 2\eta) P_\zeta(k_1)P_\zeta(k_3) \right\}$$

$$= (2\pi)^3 \delta_D^{(3)}(\vec{k}_1+\vec{k}_2+\vec{k}_3) \times (6\varepsilon - 2\eta) P_\zeta(k_1)P_\zeta(k_3)$$

$$= (2\pi)^3 \delta_D^{(3)}(\vec{k}_1+\vec{k}_2+\vec{k}_3) \times (1-n_s) P_\zeta(k_1)P_\zeta(k_3)$$

Q.E.D.

Alternatively, one can do this directly on $\langle \zeta^3\rangle$:

$$\langle \zeta_{\vec{k}_1}\zeta_{\vec{k}_2}\zeta_{\vec{k}_3}\rangle \underset{(k_3\to 0)}{\Rightarrow} \langle \langle \zeta_{\vec{k}_1}\zeta_{\vec{k}_2}\rangle_L\, \zeta_{\vec{k}_3}\rangle$$

$$= -\frac{d\ln k_1^3 P_\zeta(k_1)}{d\ln k_1} P_\zeta(k_1) \langle \zeta_{\vec{k}_1+\vec{k}_2}\zeta_{\vec{k}_3}\rangle$$

$$= +(2\pi)^3 \delta_D^{(3)}(\vec{k}_1+\vec{k}_2+\vec{k}_3)(1-n_s) P_\zeta(k_1)P_\zeta(k_3)$$

★ Therefore, if the squeezed-limit bispectrum is found to exceed $1-n_s = \mathcal{O}(10^{-2})$, all single-field models are ruled out.

f_{NL}

It is customary to use the so-called "non-linear parameter," f_{NL}, to characterize B_ζ/P_ζ^2, as

$$\boxed{\frac{6}{5} f_{NL} \equiv \frac{B_\zeta(k_1, k_2, k_3)}{P_\zeta(k_1)P_\zeta(k_2)+P_\zeta(k_2)P_\zeta(k_3)+P_\zeta(k_3)P_\zeta(k_1)}}$$

The current best bound on f_{NL} has been obtained from the WMAP 7-year data:

$$f_{NL} = 32 \pm 21 \quad (68\% \text{ C.L.})$$

or

$$-10 < f_{NL} < 74 \quad (95\% \text{ C.L.})$$

[Ref: Komatsu et al., ApJ Supplement, 192, 18 (2011)]

Note: the reason for a factor of "$\frac{6}{5}$" in the above definition is that f_{NL} was defined originally for the curvature perturbation during the matter-era in Newtonian gauge, $\Phi = \frac{3}{5}\zeta$. I.e.,

$$2 f_{NL} \equiv \frac{B_\Phi(k_1, k_2, k_3)}{P_\Phi(k_1)P_\Phi(k_2) + (\text{permutations})}$$

[Ref: Komatsu & Spergel, PRD, 63, 063002 (2001)]

2.5. Trispectrum & Multi-field Consistency Relation

[Ref: e.g., Suyama & Yamaguchi, PRD, 77, 023505 (2008); Sugiyama, Komatsu & Futamase, PRL, 106, 251301 (2011)]

If $f_{NL} \gg 1$ is found by, e.g, Planck, what should we do?

Single-field models are gone, Should we then study multi-field models? Is there any way we can rule out multi-field models also?

Perhaps the 4-point function can be used to test multi-field models. Let us start from the local-form (δN) non-Gaussianity

$$\zeta = N_I \delta\phi^I + \frac{1}{2} N_{IJ}\delta\phi^I\delta\phi^J + \frac{1}{6} N_{IJK}\delta\phi^I\delta\phi^J\delta\phi^K + \cdots$$

The 4-point function can be obtained from:

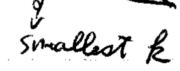
τ_{NL} ① (1st) × (1st) × (2nd) × (2nd)

g_{NL} ② (1st) × (1st) × (1st) × (3rd) → this is too small (because it involves the cubic term)

The local-form non-Gaussianity then yields:

$$\langle \zeta_{\vec{k}_1}\zeta_{\vec{k}_2}\zeta_{\vec{k}_3}\zeta_{\vec{k}_4}\rangle = (2\pi)^3\delta_D^{(3)}(\vec{k}_1+\vec{k}_2+\vec{k}_3+\vec{k}_4) \times \Big\{ \tau_{NL}\left[P_\zeta(k_1)P_\zeta(k_2)\left(P_\zeta(|\vec{k}_1+\vec{k}_3|)+P_\zeta(|\vec{k}_1+\vec{k}_4|)\right)+(\text{perm})\right] + g_{NL}\frac{54}{25}\left[P_\zeta(k_1)P_\zeta(k_2)P_\zeta(k_3)+(\text{perm.})\right]\Big\}$$

(The τ_{NL} term peaks at "small diagonal limits," e.g., 

The g_{NL} term peaks at "squeezed limits," e.g., 

)

(*) If these terms vanish for symmetry reasons, one needs to go to the next order terms.

Both τ_{NL} & g_{NL} can be given in terms of N_I, N_{IJ}, etc. The leading order contributions(*) are :

$$\boxed{\begin{aligned} \tau_{NL} &= \frac{(N_{IJ}N^J)(N^{IK}N_K)}{(N_L N^L)^3} \\ \frac{54}{25} g_{NL} &= \frac{N_{IJK}N^I N^J N^K}{(N_L N^L)^3} \end{aligned}}$$

historical factor, similar to "$\frac{6}{5}$" for f_{NL}.

Of these, τ_{NL} is of great interest. At this order, τ_{NL} is given by N_I and N_{IJ}. On the other hand, f_{NL} is also given by a combination of N_I and N_{IJ}, and thus there should be a relation between the two. (Recall $\frac{6}{5}f_{NL} = \frac{N_{IJ}N^I N^J}{(N_L N^L)^2}$.)

To see this, define the following new variables :

$$\begin{cases} a_I \equiv \dfrac{N_{IJ}N^J}{(N_L N^L)^{3/2}} \\ b_I \equiv \dfrac{N_I}{(N_L N^L)^{1/2}} \end{cases}$$

Then, we get

$$\frac{6}{5}f_{NL} = a_I b^I$$
$$\tau_{NL} = (a_I a^I)(b_J b^J)$$

The Cauchy-Schwarz inequality, $(a_I a^I)(b_J b^J) \geq (a_I b^I)^2$, gives

$$\boxed{\tau_{NL} \geq \left(\frac{6}{5}f_{NL}\right)^2}$$

[Ref: Suyama & Yamaguchi, PRD, 77, 023505 (2008)].

This is the (tree-level) Suyama-Yamaguchi inequality.

Is $\tau_{NL} \geq (\frac{6}{5} f_{NL})^2$ general?

The derivation of this relation relied on the so-called "tree-level approximation." Namely, for

$$\zeta = N_I \delta\phi^I + \frac{1}{2} N_{IJ} \delta\phi^I \delta\phi^J + \frac{1}{6} N_{IJK} \delta\phi^I \delta\phi^J \delta\phi^K + \cdots,$$

The bispectrum (f_{NL}) comes from (1st) $\times$ (1st) $\times$ (2nd), and τ_{NL} comes from (1st) $\times$ (1st) $\times$ (2nd) $\times$ (2nd).

However, it is possible to have models, for which these terms vanish (for symmetry reasons). In such cases, the leading order contributions may arise from:

f_{NL} : (2nd) $\times$ (2nd) $\times$ (2nd)

τ_{NL} : (2nd) $\times$ (2nd) $\times$ (2nd) $\times$ (2nd)

Then it is not obvious if any inequalities exist between these terms, in general.

PUNCH LINE

☆ The latest studies show that $\tau_{NL} \gtrsim (\frac{6}{5} f_{NL})^2$ may still hold for general theories of multi-field inflation, and thus this inequality can be used as the fundamental test of the inflationary mechanism for generating fluctuations.

[Ref: Sugiyama, Komatsu & Futamase, PRL, 106, 251301 (2011); Smith, LoVerde & Zaldarriaga, PRL, 107, 191301 (2011)]

3. Measuring Non-Gaussianity from CMB

3.1. Warm Up [Ref: Babich, PRD, 72, 043003 (2005)]

How have we determined $f_{NL} = 32 \pm 21$ from the WMAP data? Before we talk about CMB, it would be useful to work out a simpler example: estimating a skewness, $\kappa_3 \equiv \langle x^3 \rangle$.

Recalling the "Taylor-expanded" PDF (up to κ_3), we have

$$P(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-x^2/2\sigma^2} \left[1 + \frac{\kappa_3}{6} \left(\frac{x^3}{\sigma^6} - 3\frac{x}{\sigma^4} \right) + \cdots \right]$$

(see page 11)

The best-fit κ_3 can be found by maximizing this PDF with respect to κ_3. [NOTE: by doing this we are implicitly using Bayes' theorem, interpreting $P(x|\kappa_3)$ as $P(\kappa_3|x)$, with a flat prior on κ_3.]

$$\boxed{\left\langle \frac{d \ln P(x)}{d\kappa_3} \right\rangle = 0}$$

Let us first evaluate $\ln P(x)$. Expanding it up to κ_3^2,

$$\ln P(x) = \text{const} + \frac{1}{6}\left(\frac{x^3}{\sigma^6} - 3\frac{x}{\sigma^4} \right)\kappa_3 - \left[\frac{1}{6}\left(\frac{x^3}{\sigma^6} - 3\frac{x}{\sigma^4} \right) \right]^2 \frac{\kappa_3^2}{2} + \cdots$$

Then $d\ln P/d\kappa_3 = 0$ yields

$$\left\langle \frac{1}{36}\left(\frac{x^6}{\sigma^{12}} - 6\frac{x^4}{\sigma^{10}} + 9\frac{x^2}{\sigma^8} \right) \right\rangle \kappa_3 = \left\langle \frac{1}{6}\left(\frac{x^3}{\sigma^6} - 3\frac{x}{\sigma^4} \right) \right\rangle$$

⇒ next page

The left hand side becomes (note that $\langle x^6\rangle = 15\sigma^6$ & $\langle x^4\rangle = 3\sigma^4$)

$$(\text{l.h.s.}) = \frac{1}{6}\frac{1}{\sigma^6}.$$

Thus, the estimator is

$$\boxed{K_3 = \langle x^3\rangle - 3\sigma^2\langle x\rangle}$$

As one might expect. For a zero-mean variable, which is what we assumed, the second term vanishes and we obtain $K_3 = \langle x^3\rangle$. Of course, we donot have access to ensemble average, and thus the actual estimator would take the form of

$$K_3 = \frac{\frac{1}{6}\sum_i \left(\frac{x_i^3}{\sigma_i^6} - 3\frac{x_i}{\sigma_i^4}\right)}{\frac{1}{6}\sum_i \frac{1}{\sigma_i^6}}$$

[where "i" refers to an "ith measurement"].

This is quite a suggestive form, which has a direct contact with the actual measurement performed on the WMAP data.

3.2. PDF for CMB

In order to write down the PDF of CMB temperature (or polarization) anisotropy, we need to know how to generalize the Gram-Chalier (or Edgeworth) expansion to a multivariate case.

First, let us start from a Gaussian PDF. Temperature anisotropy (i.e., temperature in a given direction $\hat{n}$ minus the mean temperature), $\delta T(\hat{n})$, obeys

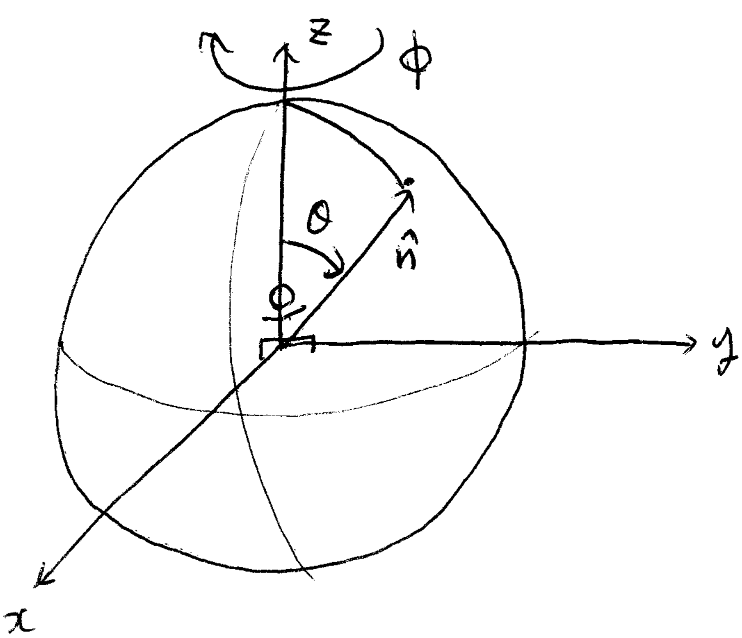
$$P(\{\delta T(\hat{n}_i)\}) = \frac{1}{(2\pi)^{N_{pix}/2}|\xi|^{1/2}} e^{-\frac{1}{2}\sum_{ij} \delta T(\hat{n}_i)(\xi^{-1})_{ij}\,\delta T(\hat{n}_j)}$$

where

$$\xi_{ij} \equiv \langle \delta T(\hat{n}_i)\,\delta T(\hat{n}_j)\rangle$$

and

$$\hat{n} = (\sin\theta\cos\phi,\ \sin\theta\sin\phi,\ \cos\theta)$$



It is more convenient to work in "Fourier" space; however, since we deal with a field on 2-sphere, we cannot use the usual Fourier transform on the full sky (unless we deal with a small section on the sky which may be approximated as a flat surface). We thus use the spherical harmonic transform:

$$\delta T(\hat{n}) = \sum_{lm} a_{lm} Y_{lm}(\hat{n})$$

or

$$a_{lm} = \int d\Omega\, \delta T(\hat{n}) Y^*_{lm}(\hat{n}).$$

Then a Gaussian PDF becomes

$$P(\{a_{lm}\}) = \frac{1}{(2\pi)^{N_{harm}/2} |C|^{1/2}} e^{-\frac{1}{2}\sum_{lm, l'm'} a^*_{lm} (C^{-1})_{lm, l'm'} a_{l'm'}}$$

where

$$C_{lm, l'm'} \equiv \langle a^*_{lm} a_{l'm'} \rangle$$

For translationally and rotationally invariant temperature anisotropy, we have

$$C_{lm, l'm'} = C_l \delta_{ll'} \delta_{mm'} \quad \text{(for trans. \& rot. invariant case)}$$

However, even if the CMB itself is trans. & rot. invariant, noise and foreground emission may not. Thus, for generality we use $C_{lm, l'm'}$ without approximating it to be diagonal.

Multi-variate expansion of PDF

We can now "Taylor-expand" a Gaussian PDF of $a_{\ell m}$.
[Ref: Amendola, MNRAS, 283, 983 (1996);
Taylor & Watts, MNRAS, 328, 1027 (2001)]

$$P(\{a_{\ell m}\}) = \frac{1}{(2\pi)^{N_{harm}/2}|C|^{1/2}} e^{-\frac{1}{2}\sum_{\ell m, \ell' m'} a^*_{\ell m}(C^{-1})_{\ell m, \ell' m'} a_{\ell' m'}}$$
$$\times \left\{ 1 + \frac{1}{6}\sum_{\text{all } \ell m} \langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3}\rangle \left[(C^{-1}a)_{\ell_1 m_1}(C^{-1}a)_{\ell_2 m_2}(C^{-1}a)_{\ell_3 m_3}\right.\right.$$
$$\left.\left. - \frac{1}{6}\sum_{\text{all } \ell m} \langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3}\rangle \left[3(C^{-1})_{\ell_1 m_1, \ell_2 m_2}(C^{-1}a)_{\ell_3 m_3}\right] + \cdots \right\}$$

If we compare this with a uni-variate case (see page 38):

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2}\left[1 + \frac{\kappa_3}{6}\left(\frac{x^3}{\sigma^6} - 3\frac{x}{\sigma^4}\right) + \cdots\right]$$

The correspondence is clear.

BASIC IDEA ☆ The above PDF can be used to derive an estimator for the full bispectrum, $\langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3}\rangle$, or the "m-averaged bispectrum, $B_{\ell_1 \ell_2 \ell_3} \equiv \sum_{\text{all } m} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3}\rangle$.
However, given a low signal-to-noise ratio, it is more useful to fit the data to a model bispectrum with a fixed shape (i.e., dependence of (ℓ_1, ℓ_2, ℓ_3)), and find the amplitude, such as f_{NL}. This is what we are going to do now.

Let's say, we have a model shape $\mathcal{B}_{l_1 l_2 l_3}$, such that
$$B_{l_1 l_2 l_3} \equiv f_{NL}\, \mathcal{B}_{l_1 l_2 l_3}.$$
Then the estimator for f_{NL} is given by

$$(*)\quad \boxed{f_{NL} = \frac{\frac{1}{6}\sum_{\text{all } lm} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \mathcal{B}_{l_1 l_2 l_3} \left[(C^{-1}a)_{l_1 m_1} (C^{-1}a)_{l_2 m_2} (C^{-1}a)_{l_3 m_3} - 3 (C^{-1})_{l_1 m_1, l_2 m_2} (C^{-1}a)_{l_3 m_3} \right]}{\frac{1}{6}\sum_{\text{all } lm} \sum_{\text{all } l'm'} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \mathcal{B}_{l_1 l_2 l_3} (C^{-1})_{l_1 m_1, l_1' m_1'} (C^{-1})_{l_2 m_2, l_2' m_2'} (C^{-1})_{l_3 m_3, l_3' m_3'} \mathcal{B}_{l_1' l_2' l_3'} \begin{pmatrix} l_1' & l_2' & l_3' \\ m_1' & m_2' & m_3' \end{pmatrix}}}$$

For a diagonal C,

$$f_{NL} = \frac{\frac{1}{6}\sum_{\text{all } lm} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \mathcal{B}_{l_1 l_2 l_3} \left[\frac{a_{l_1 m_1}}{C_{l_1}} \frac{a_{l_2 m_2}}{C_{l_2}} \frac{a_{l_3 m_3}}{C_{l_3}} - 3 \delta_{l_1 l_2} \delta_{m_1 m_2} \frac{a_{l_3 m_3}}{C_{l_1} C_{l_3}} \right]}{\frac{1}{6}\sum_{\text{all } lm} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}^2 \frac{\mathcal{B}^2_{l_1 l_2 l_3}}{C_{l_1} C_{l_2} C_{l_3}}}$$

$$= \frac{1}{6}\sum_{\text{all } l} \frac{\mathcal{B}^2_{l_1 l_2 l_3}}{C_{l_1} C_{l_2} C_{l_3}}$$

[Ref: Komatsu, Spergel & Wandelt, ApJ, 634, 14 (2005);
Creminelli et al., JCAP, 05, 004 (2006)]

Comparing this to the estimator for the skewness K_3 (see page 3

$$K_3 = \frac{\frac{1}{6}\sum_i \left(\frac{x_i^3}{\sigma_i^6} - 3\frac{x_i}{\sigma_i^4} \right)}{\frac{1}{6}\sum_i \frac{1}{\sigma_i^6}}$$

The correspondence is again clear. (*) This is the estimator used by the WMAP team to obtain $f_{NL} = 32 \pm 21$.

3.3. Relation between $a_{\ell m}$ and $\tilde{\zeta}_{\vec{k}}$

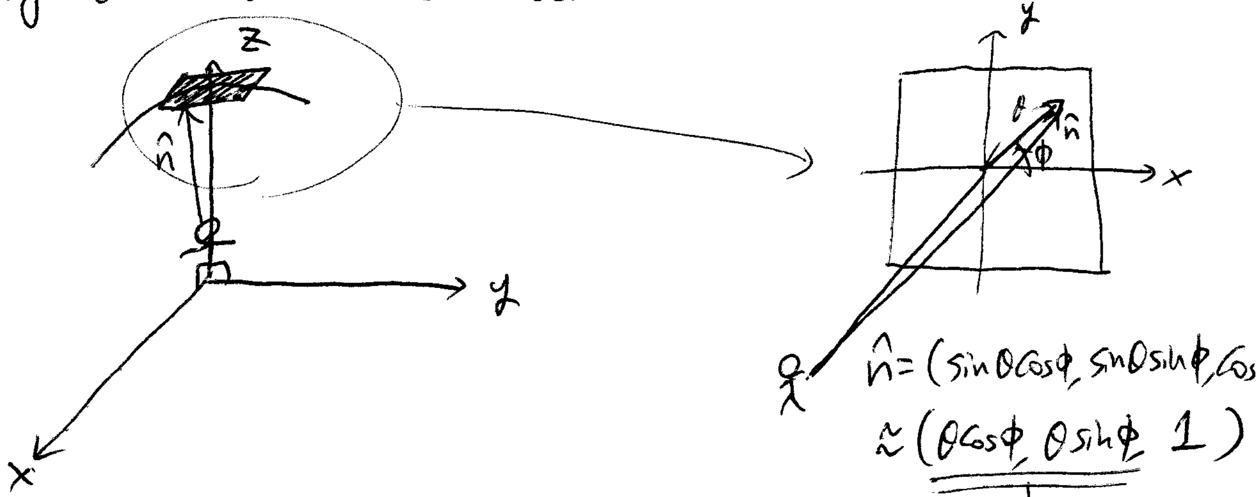
We observe $\delta T(\hat{n})$ (or $a_{\ell m}$). However, we wish to obtain information about $\tilde{\zeta}_{\vec{k}}$ from the observed $a_{\ell m}$. How are they related?

Let us begin with the simplest case. On very large angular scales (larger than the sound horizon at $z_*=1090$), the Sachs-Wolfe approximation may be used. It gives

$$\frac{\delta T(\hat{n})}{T} = -\frac{1}{5}\zeta(\hat{n}r_*, z_*) \quad \left[\begin{array}{l}\text{Ref: Sachs \& Wolfe,}\\ \text{ApJ, } \underline{147}\text{, 73 (1967)}\end{array}\right]$$

where r_* is the comoving angular diameter distance to z_*.

From this, it is clear that we are seeing only a slice of a 3d field, $\zeta(\vec{x})$. For example, if we take a tiny section of the sky centered on a $\hat{z}$ direction:



$$\hat{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$$
$$\approx \underline{(\theta\cos\phi, \theta\sin\phi,} \ 1)$$

We can perform the 2d Fourier transform:

$$\frac{\widetilde{\delta T}}{T}(\vec{\ell}) = \int d^2\theta \, \frac{\delta T}{T}(\vec{\theta}) e^{-i\vec{\ell}\cdot\vec{\theta}}$$

where $\vec{\theta} \equiv (\theta\cos\phi, \theta\sin\phi)$.

Then we find the relation between $\widetilde{\delta T}(\vec{\ell})$ and $\widetilde{\zeta}(\vec{k})$ as

$$\frac{\widetilde{\delta T}}{T}(\vec{\ell}) = \int d^2\theta\, e^{-i\vec{\ell}\cdot\vec{\theta}} \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\hat{n} r_*} \left(-\frac{1}{5}\widetilde{\zeta}_{\vec{k}}\right)$$

Using $\hat{n} \approx (\vec{\theta}, 1)$,

$$= \int d^2\theta\, e^{-i\vec{\ell}\cdot\vec{\theta}} \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}_\perp\cdot\vec{\theta} r_*} e^{ik_\parallel r_*} \left(-\frac{1}{5}\widetilde{\zeta}_{\vec{k}}\right)$$

where $\vec{k} = (\vec{k}_\perp, k_\parallel)$. The integral over $\vec{\theta}$ yields a 2d Dirac delta function:

$$= \int \frac{d^3k}{(2\pi)^3} (2\pi)^2 \delta_D^{(2)}(\vec{k}_\perp r_* - \vec{\ell}) e^{ik_\parallel r_*} \left(-\frac{1}{5}\widetilde{\zeta}_{\vec{k}}\right)$$

$$\boxed{\therefore \quad \frac{\widetilde{\delta T}(\vec{\ell})}{T} = \frac{1}{r_*^2}\int \frac{dk_\parallel}{2\pi} e^{ik_\parallel r_*} \left[-\frac{1}{5}\widetilde{\zeta}\left(\vec{k}_\perp = \frac{\vec{\ell}}{r_*}, k_\parallel\right)\right]}$$

From this result, it is clear that we are mostly sensitive to the $\vec{k}$ that are perpendicular to our line of sight, and information on $k_\parallel$ is smeared out by the integral.

This is a sad news: this limits the statistical power of CMB, which is really a 2-d object.

(It is possible to retrieve the full 3-d information by using the large-scale structure density field. More later.)

Full Sky

For the full-sky analysis, we must work with spherical harmonics. Again in the Sach-Wolfe limit,

$$a_{\ell m} = \int d\hat{n}\, Y^*_{\ell m}(\hat{n}) \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\hat{n}r_*} \left(-\frac{1}{5}\tilde{\zeta}_{\vec{k}}\right)$$

Now, using partial wave decomposition of a plane wave (called "Rayleigh's Formula"),

$$e^{i\vec{k}\cdot\hat{n}r_*} = \sum_{\ell m} 4\pi i^\ell j_\ell(kr_*) Y_{\ell m}(\hat{n}) Y^*_{\ell m}(\hat{k})$$

We obtain

$$a_{\ell m} = \int d\hat{n}\, Y^*_{\ell m}(\hat{n}) \int \frac{d^3k}{(2\pi)^3} \left(-\frac{1}{5}\right) \tilde{\zeta}_{\vec{k}} \sum_{\ell' m'} 4\pi i^{\ell'} j_{\ell'}(kr_*) Y_{\ell' m'}(\hat{n}) Y^*_{\ell' m'}(\hat{k})$$

Using $\int d\hat{n}\, Y^*_{\ell m}(\hat{n}) Y_{\ell' m'}(\hat{n}) = \delta_{\ell\ell'}\delta_{mm'}$ and summing over ℓ' & m',

$$\boxed{\therefore\ a_{\ell m} = 4\pi i^\ell \int \frac{d^3k}{(2\pi)^3} j_\ell(kr_*) Y^*_{\ell m}(\hat{k}) \left(-\frac{1}{5}\tilde{\zeta}_{\vec{k}}\right)}$$

The fact that we pick up $\vec{k}_\perp$ is obscured in this expression.

Now, let us go beyond the Sach-Wolfe limit, and include all the relevant effect. We replace

$$-\frac{1}{5} j_\ell(kr_*) \longrightarrow g^{(S)}_{T\ell}(k)$$

and write

$$\boxed{a_{\ell m} = 4\pi i^\ell \int \frac{d^3k}{(2\pi)^3} g^{(S)}_{T\ell}(k) Y^*_{\ell m}(\hat{k}) \tilde{\zeta}_{\vec{k}}},$$

where "$g_{T\ell}(k)$" is called the "radiation transfer function", which can be computed using the linear Boltzmann code such as CMBfast, CAMB, etc. The power spectrum is given by

$$C_\ell = \langle |a_{\ell m}|^2 \rangle = \frac{2}{\pi} \int k^2 dk\, P_\zeta(k)\, g^{(S)2}_{T\ell}(k).$$

3.4. Local-form CMB Bispectrum

In order to make a better contact with the literature, let us now relate $a_{\ell m}$ to Φ (curvature perturbation in matter era, Newtonian gauge):

$$a_{\ell m} = 4\pi i^\ell \int \frac{d^3k}{(2\pi)^3} \tilde{\Phi}_{\vec{k}}\, g_{T\ell}(k)\, Y_{\ell m}^*(\hat{k}).$$

Now, for the local-form bispectrum:

$$\langle \tilde{\Phi}_{\vec{k}_1} \tilde{\Phi}_{\vec{k}_2} \tilde{\Phi}_{\vec{k}_3} \rangle = (2\pi)^3 \delta_D^{(3)}(\vec{k}_1 + \vec{k}_2 + \vec{k}_3) \times 2 f_{NL} [P_\Phi(k_1) P_\Phi(k_2) + (\text{perm.})]$$

The CMB bispectrum is given by

$$\boxed{\langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} \rangle = \mathcal{G}^{\ell_1 \ell_2 \ell_3}_{m_1 m_2 m_3} \times 2 f_{NL} \times \int r^2 dr [\beta_{\ell_1}(r) \beta_{\ell_2}(r) \alpha_{\ell_3}(r) + (\text{perm.})]}$$

where

$$\mathcal{G}^{\ell_1 \ell_2 \ell_3}_{m_1 m_2 m_3} \equiv \int d\hat{n}\, Y_{\ell_1 m_1}(\hat{n}) Y_{\ell_2 m_2}(\hat{n}) Y_{\ell_3 m_3}(\hat{n})$$

$$\alpha_\ell(r) \equiv \frac{2}{\pi} \int k^2 dk\, g_{T\ell}(k)\, j_\ell(kr)$$

$$\beta_\ell(r) \equiv \frac{2}{\pi} \int k^2 dk\, P_\Phi(k)\, g_{T\ell}(k)\, j_\ell(kr)$$

[Ref: Komatsu & Spergel, PRD, 63, 063002 (2001)]

Using this in the estimator given on page 43, we can estimate f_{NL} from the data!!