Week 9: Statistics of CMB Fluctuations

April 5, 2012

1 Introduction

Last week we discussed relativistic cosmological perturbation theory. CMB fluctuations provide one of the principal motivations for developing relativistic cosmological perturbation theory. We will shortly see how this works. First, though, we need to discuss how CMB fluctuations (temperature and polarization) are quantified; i.e., what you measure and what the theory is supposed to predict.

2 CMB Temperature Fluctuations: Flat Sky

Suppose we are given a temperature map on a small region of the sky. The entire sky is a spherical surface (a two-sphere), but a sufficiently small region can be approximated as flat. In this case, measurements provide a map $T(\vec{\theta})$ as a function of $\vec{\theta}$ on the sky. The description of this temperature field is analogous to that for the projected galaxy distribution that we dealt with earlier. First, we expand $T(\vec{\theta})$ in terms of Fourier modes,

$$\tilde{T}(\vec{l}) = \int d^2\vec{\theta}\, T(\vec{\theta}) e^{i\vec{\theta}\cdot\vec{l}}. \tag{1}$$

The theory does not predict what $T(\vec{\theta})$ nor what $\tilde{T}(\vec{l})$ are—rather, theory makes predictions for the *statistics* of this temperature field. The simplest such statistic is (in Fourier space) the *power spectrum* $C_l^{\rm TT}$ defined by

$$\left\langle \tilde{T}(\vec{l})\tilde{T}(\vec{l}')\right\rangle = (2\pi)^2 \delta_D(\vec{l}+\vec{l}')C_l^{\rm TT}, \tag{2}$$

where the average is over all realizations of this field. Since $T(\vec{\theta})$ is real, $\tilde{T}^*(\vec{l}) = T(-\vec{l})$. Thus, Eq. (3) can equivalently be written,

$$\left\langle \tilde{T}(\vec{l})\tilde{T}^*(\vec{l}')\right\rangle = (2\pi)^2 \delta_D(\vec{l}-\vec{l}')C_l^{\rm TT}, \tag{3}$$

In principal, a precise measurement of $C_l^{\rm TT}$ would require that we average over many realizations of the temperature field. Nature, though, provides us with only one realization; i.e., *our* Universe.

There is thus always some finite precision, some "cosmic variance," with which this power spectrum can be measured. We'll make that statement more precise later.

The presence of the Dirac delta function is a consequence of, or statement of, statistical homogeneity; i.e., that there is no preferred location in the Universe. This can be understood as follows: Each of these plane waves can be thought of as a state of momentum $\vec{\ell}$. In QFT, invariance of a theory under spacetime invariance leads to momentum conservation, and that conservation of momentum introduces a momentum-conserving delta function in correlation functions. The same argument applies here. Likewise, the fact that the power spectrum $C_l^{\rm TT}$ depends only on the magnitude $l \equiv |\vec{l}|$, and not on the direction of $\vec{l}$, is a consequence of statistical isotropy; i.e., there is no preferred direction in the Universe.

A given power spectrum implies a two-point correlation function **(check signs and factors of two in last step)**,

$$\begin{aligned} C_{\rm TT}(\alpha) &\equiv \left\langle T(\vec{\theta})T(\vec{\theta}+\vec{\alpha})\right\rangle \\ &= \int \frac{d^2\vec{l}}{(2\pi)^2} C_l^{\rm TT} \\ &= \int_0^\infty \frac{l\,dl}{2\pi} J_0(l\theta) C_l^{\rm TT}. \end{aligned} \tag{4}$$

Note that the two-point correlation function is independent of position $\vec{\theta}$, a consequence of statistical homogeneity, and independent of the direction of α, a consequence of statistical isotropy.

If the temperature field is *Gaussian*, then each Fourier amplitude $\tilde{T}(\vec{l})$ is selected from a Gaussian distribution of variance $C_l^{\rm TT}$. More precisely, Gaussianity implies that the two-point probability distribution function $P(T_1, T_2)$ for the temperature to take on a value T_1 at position $\vec{\theta}_1$ and T_2 at $\vec{\theta}_2$ is a multivariate Gaussian with variances $C_{\rm TT}(0)$ and covariance $C_{\rm TT}(|\vec{\theta}_1 - \vec{\theta}_2|)$. This then implies that all odd-number higher correlation functions vanish:

$$\left\langle T(\vec{\theta}_1)T(\vec{\theta}_2)\cdots T(\vec{\theta}_n)\right\rangle = 0 \qquad \text{for } n = \text{odd}. \tag{5}$$

It also implies that all even-number higher correlations are disconnected; e.g., the four-point correlation function is

$$\begin{aligned} \left\langle T(\vec{\theta}_1)T(\vec{\theta}_2)T(\vec{\theta}_3)T(\vec{\theta}_4)\right\rangle &= C_{\rm TT}(|\vec{\theta}_1-\vec{\theta}_2|)C_{\rm TT}(|\vec{\theta}_3-\vec{\theta}_4|) \\ &\quad + C_{\rm TT}(|\vec{\theta}_1-\vec{\theta}_3|)C_{\rm TT}(|\vec{\theta}_2-\vec{\theta}_4|) + C_{\rm TT}(|\vec{\theta}_1-\vec{\theta}_4|)C_{\rm TT}(|\vec{\theta}_2-\vec{\theta}_3|). \end{aligned} \tag{6}$$

A six-point correlation function would consist of all 15 permutations of products of two-point correlation functions, etc.

In Fourier space, Gaussianity implies that all odd-numbered higher spectra are zero: $\left\langle \tilde{T}(\vec{l}_1)\tilde{T}(\vec{l}_2)\cdots\tilde{T}(\vec{l}_n)\right\rangle = 0$ for all odd n. E.g., the "bispectrum," the three-point correlation function in Fourier space, is zero. The even-numbered spectra functions are simply related to the power spectrum. So, for example, the "trispectrum" $\mathcal{T}(\vec{l}_1, \vec{l}_2, \vec{l}_3, \vec{l}_4)$, the four-point correlation function in Fourier space, defined by

$$\left\langle \tilde{T}(\vec{l}_1)\tilde{T}(\vec{l}_1)\tilde{T}(\vec{l}_1)\tilde{T}(\vec{l}_1)\right\rangle = (2\pi)^2 \delta_D(\vec{l}_1+\vec{l}_2+\vec{l}_3+\vec{l}_4)\mathcal{T}(\vec{l}_1, \vec{l}_2, \vec{l}_3, \vec{l}_4), \tag{7}$$

is

$$\mathcal{T}(\vec{l}_1,\vec{l}_2,\vec{l}_3,\vec{l}_4) = C_{l_1}C_{l_3}\delta_D(\vec{l}_1+\vec{l}_2)\delta_D(\vec{l}_3+\vec{l}_4) \quad + \quad 2\text{permutations}. \tag{8}$$

The simplest single-field slow-roll (SFSR) models of inflation predict that primordial perturbations should be extremely close to Gaussian, and current measurements constrain departures from Gaussianity to be extremely small, no more than $O(0.1\%)$—we will make this statement more precise below. Still, models for inflation beyond the simplest SFSR models often predict departures from Gaussianity. If so, then, e.g., the bispectrum may be predicted to be nonzero, although small. There are also a variety of late-time effects that may make the *observed* temperature map non-Gaussian, even if the primordial map is Gaussian. For example, weak gravitational lensing induces small departures from Gaussianity, as we will see below, and there may be other more exotic late-time effects that induce non-Gaussianity.

For example, some inflationary models (e.g., the curvaton, in which it is quantum fluctuations in a spectator field, the "curvaton," not the inflaton, that give rise to primordial perturbations) predict non-Gaussianity to be approximated by "local-model" non-Gaussianity. In this model, the temperature $T(\vec{\theta})$ is written in terms of a Gaussian field $t(\vec{\theta})$ as

$$T(\vec{\theta}) = t(\vec{\theta}) + 3f_{\rm nl}\left\{[t(\vec{\theta})]^2 - \left\langle [t(\vec{\theta})]^2\right\rangle\right\}, \tag{9}$$

where $f_{\rm nl}$, the non-Gaussianity parameter, quantifies the departure from Gaussianity. In the limit that $f_{\rm nl}\to 0$, $T(\vec{\theta})$ becomes Gaussian. Noting that $\left\langle [T(\vec{\theta})]^2\right\rangle \sim (10^{-5})^2$, the non-Gaussian term, $f_{\rm nl}t^2$, is smaller than the Gaussian part by a factor of $\sim f_{\rm nl}\langle T^2\rangle$. Current constraints are $|f_{\rm nl}|\lesssim 100$, implying that non-Gaussianity is empirically at the $\lesssim 0.1\%$ level.

If we have non-Gaussianity of the form, it induces a bispectrum $B(l_1,l_2,l_3)$, defined by

$$\left\langle \tilde{T}(\vec{l}_1)\tilde{T}(\vec{l}_2)\tilde{T}(\vec{l}_3)\right\rangle = (2\pi)^2\delta_D(\vec{l}_1+\vec{l}_2+\vec{l}_3)B(l_1,l_2,l_3), \tag{10}$$

(note the momentum-conserving delta function, a consequence of statistical homogeneity), with

$$B(l_1,l_2,l_3) = 6f_{\rm nl}(C_{l_1}C_{l_2}+C_{l_2}C_{l_3}+C_{l_1}C_{l_3}), \tag{11}$$

which follows if $t(\theta)$ is Gaussian with power spectrum C_l. In other models for non-Gaussianity, the bispectrum may have a different functional dependence on l_1,l_2,l_3. Note, however, that statistical isotropy/homogeneity dictates that the bispectrum be a function only of the *magnitudes* l_1, l_2, and l_3, not their directions. In other words, the bispectrum depends only on the size and shape of the triangle, not its orientation in space. There will then be a nonzero three-point correlation function (in configuration) space associated with this bispectrum.

The local model also induces a trispectrum a new connected part (i.e., the part in addition to the disconnected part given above) of the trispectrum given by

$$\begin{aligned}\mathcal{T}_c(\vec{l}_1,\vec{l}_2,\vec{l}_3,\vec{l}_4) &= f_{\rm nl}^2\Big[P^{l_1l_2}_{l_3l_4}(|\vec{l}_1+\vec{l}_2|) \\ &+ P^{l_1l_3}_{l_2l_4}(|\vec{l}_1+\vec{l}_3|)+P^{l_1l_4}_{l_2l_3}(|\vec{l}_1+\vec{l}_4|)\Big],\end{aligned} \tag{12}$$

where

$$P^{l_1 l_2}_{l_3 l_4}(|\vec{l}_1+\vec{l}_2|) \equiv 4C_{|\vec{l}_1+\vec{l}_2|}\left[C_{l_1}C_{l_3}+C_{l_1}C_{l_4} +C_{l_2}C_{l_3}+C_{l_2}C_{l_4}\right]. \quad (13)$$

The trispectrum is nonvanishing only for $\vec{l}_1+\vec{l}_2+\vec{l}_3+\vec{l}_4=0$, that is, only for quadrilaterals in Fourier space.

We have not yet discussed measurement of the power spectrum, bispectrum, correlation functions, etc. That is more easily understood in the full-sky formalism to which we now turn.

3 CMB Temperature: Full-sky formalism

Suppose now that we have data from a satellite experiment, like COBE, WMAP, or Planck, that provides the temperature $T(\hat{n})$ as a function of the direction $\hat{n}=(\theta,\phi)$ on the full sky. A spherical-harmonic transform,

$$T(\hat{n}) = \sum_{lm} a^{\rm T}_{lm} Y_{lm}(\hat{n}), \quad (14)$$

$$a^{\rm T}_{lm} = \int d\hat{n}\, T(\hat{n}) Y^*_{lm}(\hat{n}), \quad (15)$$

provides the full-sky analog of a Fourier transform. Statistical isotropy and homogeneity now imply

$$\left\langle a^{\rm T}_{lm}(a^{\rm T}_{l'm'})^*\right\rangle = C_l \delta_{ll'}\delta_{mm'}. \quad (16)$$

Here the "momentum-conserving" Dirac delta function, a consequence of statistical homogeneity, is replaced by an "angular-momentum–conserving" Kronecker delta. The lack of any dependence on m is a consequence of statistical isotropy; i.e., no preferred m implies no preferred direction (or *vice versa*). Reality of $T(\hat{n})$ now implies $a^*_{lm}=(-1)^m a_{l,-m}$.

Suppose now that we had a perfect (i.e., no instrumental noise, full sky coverage, and perfect angular resolution) map $T(\hat{n})$. We would then construct the $a^{\rm T}_{lm}$s and our estimator for the power spectrum would then be

$$\widehat{C^{\rm TT}_l} = \frac{1}{2l+1}\sum_{m=-l}^{l} |a^{\rm T}_{lm}|^2. \quad (17)$$

Let's think about this measurement a bit more carefully, though. Theory predicts the expectation values $\left\langle |a^{\rm T}_{lm}|^2\right\rangle = C^{\rm TT}_l$ when averaged over a ensemble of universes, or assuming ergodicity, a spatial average over all observer positions in the Universe. However, we only observe a single realization of the ensemble from a single location. Therefore, even if we had an ideal (full-sky coverage, no foreground contamination, infinite angular resolution, and no instrumental noise) experiment, the accuracy with which the estimator above could recover the power spectrum would be limited by a sample variance known as "cosmic variance." Moreover, there will be finite angular resolution and instrumental noise to be taken into account, even if we have a full-sky map. The naive estimator above must then be altered to take these effects into account, and the noise with which we can measure the power spectrum may then be increased beyond the irreducible cosmic variance.

Consider a temperature map $T^{\rm map}(\hat{\mathbf{n}})$ of the full sky which is pixelized with $N_{\rm pix}$ pixels. If we assume that each pixel subtends the same area on the sky then we can construct multipole coefficients of the temperature map using

$$d^{\rm T}_{lm} = \int d\hat{\mathbf{n}} \left(\frac{T^{\rm map}(\hat{\mathbf{n}})}{T_0}\right) Y_{lm}(\hat{\mathbf{n}}) \simeq \frac{1}{T_0}\sum_{j=1}^{N_{\rm pix}} \frac{4\pi}{N_{\rm pix}} T^{\rm map}_j \, Y_{lm}(\hat{\mathbf{n}}_j), \tag{18}$$

where $T^{\rm map}_j$ is the measured temperature perturbation in pixel j, and $\hat{\mathbf{n}}_j$ is its direction. The difference between d_{lm} and a_{lm} is that the former includes the effects of finite beam size and detector noise. The observed temperature $T^{\rm map}_j = T_j + T^{\rm noise}_j$ is due to a cosmological signal T_j and a pixel noise $T^{\rm noise}_j$. If we assume that each pixel has the same rms noise, and that the noise in each pixel is uncorrelated with that in any other pixel, and is uncorrelated with the cosmological signal, i.e. $\left\langle T^{\rm noise}_i T^{\rm noise}_j \right\rangle = T_0^2 \left(\sigma^{\rm T}_{\rm pix}\right)^2 \delta_{ij}$ and $\left\langle T_i \, T^{\rm noise}_j \right\rangle = 0$, then

$$\begin{aligned} \left\langle d^{\rm T}_{lm} d^{{\rm T}\,*}_{l'm'} \right\rangle &= \left\langle a^{\rm T}_{lm} a^{{\rm T}\,*}_{l'm'} \right\rangle + \left\langle a^{\rm T,noise}_{lm} \left(a^{\rm T,noise}_{l'm'}\right)^* \right\rangle \\ &= |W^{\rm b}_l|^2 C^{\rm T}_l \delta_{ll'}\delta_{mm'} + \left\langle a^{\rm T,noise}_{lm} \left(a^{\rm T,noise}_{l'm'}\right)^* \right\rangle, \end{aligned} \tag{19}$$

where we have written the expectation value of the cosmological signal in terms of that, , $C^{\rm TT}_l$, predicted by theory multiplied by a (Fourier-space) window function $|W^{\rm b}_l|^2$. This window function arises because the thing we measure is not the temperature in direction $\vec{\theta}$, but rather the smoothed temperature field, $\int d^2\vec{\theta}' T(\vec{\theta}') W(|\vec{\theta} - \vec{\theta}'|)$, where $W(\alpha)$ is a window function which we model as a Gaussian of full-width half-maximum of θ_{fwhm}. Thus, if the observed temperature is a convolution of the actual temperature with a window function $W(\alpha)$, then in Fourier space, the observed Fourier moments are multiplied by the Fourier transform $W^{\rm b}_l$ which, if the beam is Gaussian, is given by $W^{\rm b}_l \approx \exp(-l^2\sigma_{\rm b}^2/2)$, with $\sigma_{\rm b} = \theta_{\rm fwhm}/\sqrt{8\ln 2} = 0.00742\,(\theta_{\rm fwhm}/1^\circ)$.

The second term in Eq. (19) is

$$\begin{aligned} \left\langle a^{\rm T,noise}_{lm} \left(a^{\rm T,noise}_{l'm'}\right)^* \right\rangle &= \frac{1}{T_0^2} \sum_{i=1}^{N_{\rm pix}} \sum_{j=1}^{N_{\rm pix}} \left(\frac{4\pi}{N_{\rm pix}}\right)^2 \left\langle T^{\rm noise}_i T^{\rm noise}_j \right\rangle Y_{lm}(\hat{\mathbf{n}}_i)\, Y^*_{l'm'}(\hat{\mathbf{n}}_j) \\ &= \left(\sigma^{\rm T}_{\rm pix}\right)^2 \left(\sum_{i=1}^{N_{\rm pix}} \frac{4\pi}{N_{\rm pix}} Y_{lm}(\hat{\mathbf{n}}_i)\, Y_{l'm'}(\hat{\mathbf{n}}_j)\right) \frac{4\pi}{N_{\rm pix}} \\ &= \frac{4\pi \left(\sigma^{\rm T}_{\rm pix}\right)^2}{N_{\rm pix}} \delta_{ll'}\, \delta_{mm'} \equiv C^{\rm TT,noise}_l \delta_{ll'}\delta_{mm'}, \end{aligned} \tag{20}$$

where the last line defines the noise power spectrum $C^{\rm TT,noise}_l$. Therefore, the moments measured by the map are distributed with a variance

$$\left\langle d^{\rm T}_{lm} d^{{\rm T}\,*}_{l'm'} \right\rangle = D^{\rm TT}_l \delta_{ll'}\delta_{mm'}, \tag{21}$$

where

$$D^{\rm TT}_l \equiv C^{\rm TT}_l |W^{\rm b}_l|^2 + C^{\rm TT,noise}_l, \tag{22}$$

is the power spectrum for the map.

So now let's suppose we're given a map—i.e., we're provided with measured temperatures T_j in each pixel. We then do the (discretized) spherical-harmonic transform in Eq. (18) to get the $d_{lm}^{\rm T}$, and from those we get an estimator

$$\widehat{D_l^{\rm TT}} = \frac{1}{2l+1} \sum_{m=-l}^{l} |d_{lm}^{\rm T}|^2, \tag{23}$$

for the map power spectrum. Once we have this, we then obtain an estimator for the CMB power spectrum $C_l^{\rm TT}$ from

$$\widehat{C_l^{\rm TT}} = \left(\widehat{D_l^{\rm TT}} - C_l^{\rm TT,noise}\right) |W_l^{\rm b}|^{-2}. \tag{24}$$

We now evaluate the measurement error associated with this estimator for $C_l^{\rm TT}$. Recall that if x_i are Gaussian random variables with variances $\langle x_i x_j \rangle = \sigma_{ij}^2$, then $\left\langle x_i^2 x_j^2 \right\rangle = \sigma_{ii}^2 \sigma_{jj}^2 + 2\sigma_{ij}^2$, and $\langle x_i^3 x_j \rangle = 3\sigma_{ii}^2 \sigma_{ij}^2$.

$$\begin{aligned}
\left\langle \left(\widehat{D_l^{\rm TT}}\right) \right\rangle &= \sum_{mm'} \frac{\left\langle |d_{lm}^{\rm T}|^2 |d_{lm'}^{\rm T}|^2 \right\rangle}{(2l+1)^2} \\
&= \sum_{mm'} \frac{\left(D_l^{\rm TT}\right)^2}{(2l+1)^2} \left[(1-\delta_{mm'}) + 3\delta_{mm'}\right] \\
&= \left(D_l^{\rm TT}\right)^2 \left[1 + \frac{2}{(2l+1)^2}\right].
\end{aligned} \tag{25}$$

Thus, the variance with which $C_l^{\rm TT}$ can be measured is

$$\begin{aligned}
\left(\Delta C_l^{\rm TT}\right)^2 &\equiv \left\langle \left(\widehat{C_l^{\rm TT}}\right)^2 \right\rangle - \left\langle \left(\widehat{C_l^{\rm TT}}\right) \right\rangle^2 \\
&= \left\{ \left\langle \left(\widehat{D_l^{\rm TT}}\right)^2 \right\rangle - \left\langle \left(\widehat{D_l^{\rm TT}}\right) \right\rangle^2 \right\} |W_l^{\rm b}|^{-4} \\
&= \frac{2}{2l+1} \left(C_l^{\rm TT} + C_l^{\rm TT,noise} |W_l^{\rm b}|^{-2} \right)^2.
\end{aligned} \tag{26}$$

Note that even if we are given a perfect map—i.e., no noise, $C_l^{\rm TT,noise} = 0$—there is a cosmic-variance error $\Delta C_l^{\rm TT} = \sqrt{2/(2l+1)} C_l^{\rm TT}$ with which we can measure the power spectrum. Instrumental noise increases the error. Note also that the presence of the window-function factor $|W_l^{\rm b}|^{-2} \propto e^{l^2 \sigma_b^2}$ implies that the noise contribution grows expontially for $l \gtrsim \sigma_b^{-1}$ which makes sense: you can't measure structure effectively on angular scales smaller than the beam.

Let's now turn to the bispectrum on the full sky. The CMB bispectrum is given by

$$B_{l_1 l_2 l_3}^{m_1 m_2 m_3} \equiv \langle a_{l_1 m_1} a_{l_2 m_2} a_{l_3 m_3} \rangle . \tag{27}$$

The reduced or angle-averaged bispectrum is defined as

$$b_{l_1 l_2 l_3} \equiv \sum_{m_1 m_2 m_3} \left(\begin{array}{ccc} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{array} \right) B_{l_1 l_2 l_3}^{m_1 m_2 m_3} \left[\sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \left(\begin{array}{ccc} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{array} \right) \right]^{-1} \tag{28}$$

where the matrices are Wigner-3j symbols. The bispectrum $B^{m_1 m_2 m_3}_{l_1 l_2 l_3}$ must satisfy the triangle conditions and selection rules: $m_1+m_2+m_3=0$, $l_1+l_2+l_3=\text{even}$, and $|l_i-l_j|\le l_k\le l_i+l_j$ for all permutations of indices. Thus, $B^{m_1 m_2 m_3}_{l_1 l_2 l_3}$ consists of the Gaunt integral, $\mathcal{G}^{m_1 m_2 m_3}_{l_1 l_2 l_3}$, defined by

$$\begin{aligned}\mathcal{G}^{m_1 m_2 m_3}_{l_1 l_2 l_3} &\equiv \int d^2\hat{\mathbf{n}} Y_{l_1 m_1}(\hat{\mathbf{n}}) Y_{l_2 m_2}(\hat{\mathbf{n}}) Y_{l_3 m_3}(\hat{\mathbf{n}}) \\ &= \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}. \end{aligned} \tag{29}$$

This Gaunt integral $\mathcal{G}^{m_1 m_2 m_3}_{l_1 l_2 l_3}$ is real, and satisfies all the conditions mentioned above.

The bispectrum can be written in terms of the reduced bispectrum through

$$B^{m_1 m_2 m_3}_{l_1 l_2 l_3} = \mathcal{G}^{m_1 m_2 m_3}_{l_1 l_2 l_3} b_{l_1 l_2 l_3}, \tag{30}$$

using used the identity,

$$\sum_{m_1 m_2 m_3} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \mathcal{G}^{m_1 m_2 m_3}_{l_1 l_2 l_3} = \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}. \tag{31}$$

With a little effort, you can convince yourself that the reduced bispectrum $b_{l_1 l_2 l_3}$ is equivalent in the flat-sky limit ($l_i \gg 1$) to the flat-sky bispectrum $B(l_1, l_2, l_3)$ discussed above.

Eq. (28) provides the recipe to measure the bispectrum. Ignoring instrumental noise, the estimator $\widehat{b_{l_1 l_2 l_3}}$ is constructed from that equation by replacing $B^{m_1 m_2 m_3}_{l_1 l_2 l_3}$ by $a^{\mathrm{T}}_{l_1 m_1} a^{\mathrm{T}}_{l_2 m_2} a^{\mathrm{T}}_{l_3 m_3}$. The effects of noise can be dealt with following a sequence of steps analogous to those for the power spectrum.

You can show that the functional form for the full-sky bispectrum for the local model is identical to that (derived above) for the flat sky.

The skewness,

$$S_3 \equiv \left\langle \left(\frac{\Delta T(\hat{\mathbf{n}})}{T} \right)^3 \right\rangle \tag{32}$$

is the simplest statistic characterizing the non-Gaussianity. It is expanded in terms of $b_{l_1 l_2 l_3}$ from

$$S_3 = \frac{1}{2\pi^2} \sum_{2\le l_1 l_2 l_3} \left(l_1+\frac{1}{2}\right)\left(l_2+\frac{1}{2}\right)\left(l_3+\frac{1}{2}\right) \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}^2 b_{l_1 l_2 l_3}. \tag{33}$$

This is the honest-to-goodness skewness of the CMB. However, if we are measuring the CMB with an experiment with window function W^{b}_l then the skewness of the map is obtained by including a factor $W^{\mathrm{b}}_{l_1} W^{\mathrm{b}}_{l_2} W^{\mathrm{b}}_{l_3}$ in the right-hand side. The $l=0$ and $l=1$ modes are excluded from the sum as they cannot be measured.