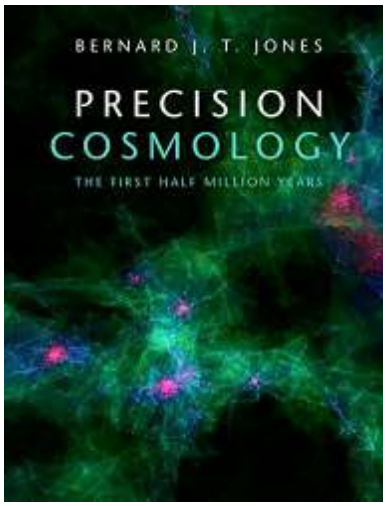
Random Fields

A Supplement to “Precision Cosmology”

Bernard Jones



Here we review some aspects of the theory if random fields and how such fields are handled in cosmology.

This is one of a set of Supplementary Notes and Chapters to “Precision Cosmology”.
Some of these Supplements might have been a chapter in the book itself, but were regarded either as being somewhat more specialised than the material elsewhere in the book, or somewhat tangential to the main subject matter.
The are mostly early drafts and have not been fully proof-read.
Please send comments on errors or ambiguities to “PrecisionCosmology(at)gmail.com”.

Contents

1 Supplement: Random Fields

1.1 Describing density inhomogeneities

We focus here on the issue of the linear growth of fluctuations on sub-horizon scales in a variety of Friedmann-Lemaître models. That has the advantage of keeping the treatment relatively straightforward (it avoids relativistic issues) and the assumption of linearity is certainly adequate at earlier times. First we walk a well trodden path describing structure evolution in models without a Λ-term, both in the case of zero pressure which is relevant to the times following the recombination of the cosmic plasma, and in the case of non-zero pressure which is relevant to the fireball phase of the expansion. Then we look at modifications of these equations to handle situations where there is a significant Dark Energy component in the Universe.

do we??

Our model for the early universe is, when viewed on the very largest scales, the homogeneous and isotropic Universe of Friedmann and Lemaître. However, as we look more closely on increasingly smaller scales the deviation from inhomogeneity increases. The question arises as to how to handle that mathematically.

The earliest inhomogeneous cosmological models dealt with spherically symmetric inhomogeneities in an otherwise homogeneous and isotropic background universe. Those models based on the Bondi-Tolman representation of the inhomogeneities can consist of an arbitrary spatial distributions of any number of such *Swiss Cheese* holes provided they do not overlap. However, such models are rather special in that the inhomogeneity does not influence the surrounding volume: the lump is shielded by the surrounding hole.

1.1.1 Relative density fluctuations

More general density distributions are best described as random fluctuations in density, having zero mean value, superposed on the otherwise homogeneous background density field. Such a density field can be described by the equation

$$\rho(\mathbf{r}) = \rho_0(1 + \delta(\mathbf{r})) \tag{1.1}$$

$$\langle\delta(\mathbf{r})\rangle = 0, \tag{1.2}$$

$$\langle\rho(\mathbf{r})\rangle = \rho_0 \tag{1.3}$$

where ρ_0 is the mean density at some time and $\delta(\mathbf{r})$ is the relative density fluctuation at position $\mathbf{r}$. [1] The angular brackets $\langle \dots \rangle$ denote spatial averages.[2] Equation (1.2) expresses the fact that the mean value of the fluctuations is zero, and equation (1.3) follows from taking the spatial average of (1.1).

An alternative way of presenting this is to write (1.1) as

$$\delta(\mathbf{r}) = \frac{\rho(\mathbf{r}) - \rho_0}{\rho_0} = \frac{\delta\rho(\mathbf{r})}{\rho_0} \tag{1.4}$$

The function $\delta\rho(\mathbf{r})$ is the fluctuation in density relative from the mean density, ρ_0, at the point $\mathbf{r}$.

While equation (1.1) looks rather innocuous, it hides a considerable amount of complexity. As it stands all it tells us is that the density is non-uniform, we do not know how non-uniform it is, nor anything about the nature of the irregular density field described by the function $\delta(\mathbf{r})$. We cannot picture what the density field looks like without more information.

1.2 Statistical descriptors of random fields

It is appropriate to describe random fields statistically.[3] A random field $\delta(\mathbf{r})$ is a function of position whose value at any point, $\mathbf{r}$, or in the neighbourhood of any point, is *a priori* not deterministic. If the function is the result of some underlying physical process, we might reasonably expect that its values at various points might be described statistically. Furthermore, if the underlying physical process is the same over the entire domain of $\{\mathbf{r}\}$ values we might expect the statistical distribution of values for a single point to be independent of position, and for the values on a set of points to depend only on the relative positions of the points and not on their absolute position.

This, in effect, describes the properties of what is known as a *homogeneous random process*, or *stationary random process*.[4] The former term is generally used in the spatial domain, while the latter is used in the time domain (*e.g.* time series).

[1] We will, insofar as is convenient, use $\mathbf{r}$ to denote position in physical 3-space, and reserve the scalars like x to denote position on the real line. Variables like $T, v, delta, \dots$ will denote variables taking on scalar values. Of course, the scalar δ may be a function of position in 3-space, $\delta(\mathbf{r})$. Likewise for vector-valued functions: $\mathbf{v}(\mathbf{r})$. In more general mathematical discussions, vectors such as $\mathbf{x}, \mathbf{y}, \dots$ will be used to denote the independent variates in arbitrary numbers of dimensions.

[2] The random process $\delta(\mathbf{r})$ is considered as being stationary and ergodic. This average can equally be thought of as being an average over multiple realisations of the random process, or simply as a spatial average. We have only one Universe and so astronomers tend to think more in terms of spatial averaging. Formally, what we mean here by the average $\langle X \rangle$ of a quantity X, is the expectation value $\mathbb{E}[X]$ given the underlying probability distribution of X. This is discussed in section 1.4.

[3] See Appendix **??** for a basic Probability and Statistics Primer. This is an enormously large field and here it is only possible to touch on some of the fundamental concepts. For more see also the books of Batchelor (1982); Monin and Yaglom (1971); Panchev (1971); Adler (1981); Papoulis and Pillai (2002), which are the texts used in putting together this section.

[4] It is a premise of probability theory that a random function $\delta(\mathbf{x})$ is determined statistically by the set of joint probability distributions of the function at any n values of the argument $\mathbf{x}$ (Batchelor, 1982, §2.2).
Suppose that $f(\mathbf{x}_1, \dots, \mathbf{x}_n)$ is the joint probability density of the values $\delta_i = \delta(\mathbf{x}_i)$ of $\delta(\mathbf{x})$ at n points $\mathbf{x}_1, \dots, \mathbf{x}_n$.

The statistics of a stationary random function can be thought of in one of two ways. Firstly, we can think of multiple *realisations* of the random process that generates the function. In this picture, if we fix attention on one point we can determine the mean, variance and other statistical moments of the function simply by taking suitable averages of the values measured in all of the realizations at that point. Alternatively, we can look at many points in one realisation of the process and take the averages of values measured at all of the points. If the underlying process is *ergodic* then the averages computed by either of these methods will yield the same values.

1.2.1 The 1-point distribution function

A random variable X taking values $x \in X$ is said to have a distribution function $F(x)$ if the probability that X takes on a value $X \leq x$ is

$$F_X(x) = \mathbb{P}[X \leq x] = \int_{-\infty}^{x} f_X(x)dx \tag{1.9}$$

$f_X(x)$ is called the *probability density* of X and clearly $f(x) = F'(x)$ wherever $F_X(x)$ is differentiable.[5]

The values of random function $\delta(\mathbf{r})$ at a single point may take on values that are drawn from some probability distribution$F_\delta(x)$ telling us the probability that a value of δ sampled at a point, $\mathbf{r}$, is greater some particular value δ_c. The standard notation for this is

$$F_\delta(\delta_c) = \mathbb{P}[\delta \leq \delta_c], \quad f_\delta(x) = F'_\delta(x) \tag{1.10}$$

We should note that, by construction, the mean value of the field $\delta(\mathbf{r})$, its expectation value,

Then we can can define the average $\langle F \rangle$ of any function $F(\delta_1, \ldots, \delta_n)$ as

$$\langle F \rangle = \int F f(\delta_1, \ldots, \delta_n)\, d\delta_1, \ldots, d\delta_n \tag{1.5}$$

where the integral is taken over the domain of possible values of $\delta(\mathbf{x})$.
The random field $\delta(\mathbf{x})$ is said to be a *homogeneous random process* if the joint probability $f(\delta_1, \ldots, \delta_n)$ is independent of position in the sense that

$$f(\delta(\mathbf{x}_1), \ldots, \delta(\mathbf{x}_n)) = f(\delta(\mathbf{x}_1 + \mathbf{y}), \ldots, \delta(\mathbf{x}_n + \mathbf{y})), \quad \forall\, \mathbf{y} \tag{1.6}$$

and in this case, the average $\langle F \rangle$ of $F(\delta_1, \ldots, \delta_n)$ is independent of position. Note that $F(\delta_1, \ldots, \delta_n)$ may nevertheless depend on the *relative* configuration of the points $\mathbf{x}_1, \ldots, \mathbf{x}_n$.
If the process $\delta(\mathbf{x})$ is ergodic, we can evaluate our average $\langle F \rangle$ via the limiting process

$$\langle F \rangle = \lim_{V \to \infty} \frac{1}{V} \int_V F(\delta_1, \ldots, \delta_n)\, dV \tag{1.7}$$

provided the limit exists and is independent of V, and where $F(\delta_1, \ldots, \delta_n)$ is given for a single realisation of the $(\delta_1, \ldots, \delta_n)$. It can be shown that this limit exists if $\delta(\mathbf{x})$ is a *mean square continuous* random process:

$$\mathbb{E}\left[|\,\delta(\mathbf{x})|^2\right] < +\infty \quad \text{and} \quad \lim_{\mathbf{x} \to \mathbf{y}} \mathbb{E}\left[|\,\delta(\mathbf{x}) - \delta(\mathbf{y})|^2\right] = 0 \tag{1.8}$$

(Monin and Yaglom, 1971, Chapter 2, especially §4.5. In that edition this is the third chapter.)

[5] See Appendix **??** for a Probability and Statistics Primer. For simplicity we can suppose that, when X denotes a random field, $\mathbb{P}[x > x_1]$ denotes the fraction of samples of the random variable X whose value exceeds a given value, x_1.

is [6]

$$\mu_\delta \equiv \mathbb{E}[\delta] = \int_{-\infty}^{\infty} x f_\delta(x) = 0 \tag{1.11}$$

When modelling simple physical processes the simplest assumption about $F(x)$ would be that it has the probability density of a zero-mean Gaussian random variable:

$$f_\delta(x) \equiv \frac{dF_\delta(x)}{dx} = \frac{1}{\sqrt{2\pi}\sigma_\delta} \exp\left(-\frac{\delta^2}{2\sigma_\delta^2}\right) \tag{1.12}$$

Here σ_δ denotes the variance of the random variable δ. This is formally defined as the expectation value of the square of the zero-mean random variable δ:

$$\sigma_\delta^2 = \mathbb{E}[\delta^2]. \tag{1.13}$$

Since the mean of the fluctuations is zero, σ_δ is the first descriptor of the random field $\delta(\mathbf{r})$: σ_δ quantifies how much the sample amplitude varies about the mean, though it tells us nothing about the character of that variation.

1.2.2 The 2-point distribution function

Consider the statistical relationship between the values , $\delta_1 = \delta(\mathbf{r}_1)$ and $\delta_2 = \delta(\mathbf{r}_2)$, of the random field $\delta(\mathbf{r})$ at two points $\mathbf{r}_1$ and $\mathbf{r}_2$. The joint distribution function of δ_1 and δ_2 is

$$F_{12} \equiv F_{\mathbf{r}_1\mathbf{r}_2}(\delta_1, \delta_2) = \mathbb{P}[\delta(\mathbf{r}_1) < \delta_1,\ \delta(\mathbf{r}_2) < \delta_2\,] \tag{1.14}$$

and their joint probability density is

$$f_{12} \equiv f_{\mathbf{r}_1\mathbf{r}_2} = \frac{\partial^2 F_{12}(\delta_1, \delta_2)}{\partial\delta_1\,\partial\delta_2} \tag{1.15}$$

provided this derivative exists.

Although we could work with the joint distribution of $\delta(\mathbf{r})$ at two points, it proves more useful to describe this joint distribution via its statistical moments. We gain some simplification of the equations since the mean value of $\delta(\mathbf{x})$ is zero, in which case the 2-point joint statistical moment of $\delta(\mathbf{r})$ is

$$Q_{12}(\mathbf{r}_1, \mathbf{r}_2) = \langle\delta(\mathbf{r}_1)\delta(\mathbf{r}_2)\rangle = \mathbb{E}[\delta(\mathbf{r}_1)\delta(\mathbf{r}_2)] \tag{1.16}$$

This is the generalisation of equations (**??**) and (**??**) of Appendix **??** to the case of a random field. $Q_{12}(\mathbf{r}_1, \mathbf{r}_2)$ is referred to as the *cross correlation* of the random fields $\delta(\mathbf{r}_1)$ and $\delta(\mathbf{r}_2)$.[7]

[6] Since the density $\rho(\mathbf{x}) \geq 0$, we have $-1 \leq \delta < +\infty$, and so, strictly speaking, the range of integration in equation (1.11) should be $[-1, \infty]$. However, this is 'small perturbation theory' where the δ-values are supposed to be arbitrarily small: $|\,\delta\,| \ll 1$. This justifies the convenience of using distributions functions $F_\delta(x)$ that have infinite range, such as the Gaussian.

[7] The two-point correlation function of a general homogeneous random field $u(\mathbf{r})$ is the second moment of the field values at two points: $\langle u(\mathbf{r}_1)u(\mathbf{r}_1)\rangle$. The 2-point *central moment* is $\langle(u(\mathbf{r}_1) - u_0)(u(\mathbf{r}_1) - u_0\rangle$, where $u_0 = \langle u(\mathbf{r})\rangle$ is the mean field value (which is the same at all points because this is a homogeneous random field). Since our density fluctuation field $\delta(\mathbf{r})$ has zero mean there is no difference between its central and non-central moments.

If Q_{12} is normalised by the variances σ_1, σ_2 at the points $\mathbf{r}_1, \mathbf{r}_2$ we have what is called the *covariance function*

$$q_{12} = \frac{Q_{12}}{\sigma_1\,\sigma_2} \tag{1.17}$$

In a homogeneous random process, $\sigma_1 = \sigma_2 = \sigma_\delta$ (*cf.* equation 1.13).

1.3 Isotropic random fields

1.3.1 Isotropic random scalar fields

There is an important situation in which the correlation function between the values of the field at two points $\mathbf{r}_1, \mathbf{r}_2$ depends only on the relative distance between the points $\boldsymbol{\tau} = \mathbf{r}_1 - \mathbf{r}_2$, and not on the direction or location of $\boldsymbol{\tau}$. In that case

$$Q_{12}(\mathbf{r}_1, \mathbf{r}_2) = Q_{12}(|\,\boldsymbol{\tau}\,|), \quad \boldsymbol{\tau} = \mathbf{r}_1 - \mathbf{r}_2. \tag{1.18}$$

Such a field $\delta(\mathbf{r})$ is said to be a *statistically isotropic random field*.[8] There are some simple results that follow:

$$Q_{12}(\mathbf{r}_1, \mathbf{r}_2) = Q_{12}(\mathbf{r}_2, \mathbf{r}_1), \quad Q_{12}(\mathbf{r}, \mathbf{r}) = \sigma(\mathbf{r})^2, \tag{1.19}$$

$$Q_{12}(\mathbf{r}_1, \mathbf{r}_2) \to 0 \quad \text{if } \langle\delta(\mathbf{r})\rangle = 0. \tag{1.20}$$

In the case of our zero mean random field $\delta(\mathbf{r})$

$$R(\tau) \equiv \mathbb{E}\,[\delta(\mathbf{r})\,\delta(\mathbf{r}+\boldsymbol{\tau})] = Q_{12}(|\,\boldsymbol{\tau}\,|) \tag{1.21}$$

With this, taking $\boldsymbol{\tau} = 0$, in which case the two points $\mathbf{r}_1$ and $\mathbf{r}_2$ coincide:

$$R(0) = \sigma_\delta^2 \tag{1.22}$$

which follows from equations (1.13) and (1.16).

There are a few elementary results that follow from (1.20) and (1.19):

$$R(\tau) = R(-\tau), \quad |\,R(\tau)\,| \le R(0) = \sigma^2, \quad R(\infty) = 0 \tag{1.23}$$

It can also be shown that $R(\tau)$ is continuous if and only if the underlying field $\delta(\mathbf{r})$ is statistically continuous.[9]

[8] This equation is the generalisation of equation (**??**) with $\delta(\mathbf{r}_1)$ replacing the random variable X and $\delta(\mathbf{r}_2)$ replacing the random variable Y.

[9] The definition of statistical continuity of a function $f(\mathbf{r})$ is that for a given ϵ, there is a δ such that for $|\,\tau\,| < \delta$

$$\mathbb{E}\,[|\,f(\mathbf{r}+\boldsymbol{\tau}) - f(\mathbf{r})\,|^2] \le \epsilon \tag{1.24}$$

1.3.2 Isotropic Gaussian random fields

The simplest, and arguably the most important, case arises when the underlying n-point distribution is a Gaussian distribution of the density fluctuation value at n points $\mathbf{r}_1, dots, \mathbf{r}_n$. The 1-point and 2-point probability densities are then

$$f_1(\delta) = \frac{1}{\sqrt{2\pi}\sigma_\delta} \exp\left(-\frac{\delta}{2\sigma_\delta^2}\right) \tag{1.25}$$

$$f_1(\delta_1, \delta_2) = \frac{1}{2\pi\sigma_\delta^2\sqrt{1-R(\tau)^2}} \exp\left[-\frac{\delta_1^2 - 2R(\tau)\delta_1\delta_2 + \delta_2^2}{2\sigma^2(1-R(\tau)^2)}\right] \tag{1.26}$$

It can be shown that all of the n-point Gaussian probability densities, $f_n(\mathbf{r}_1, \ldots, \mathbf{r}_n,)$ are expressible in terms of the correlation function $R(\tau)$. It can further be shown that the derivatives of the Gaussian random field also have a Gaussian distribution.

1.3.3 Multi-dimensional isotropic random fields

Multi-dimensional random vectors $\mathbf{X} = (X_1, \ldots, X_n)$, where the X_i are random variables, are described in Appendix **??**, (see equation **??** *et seq.*). A multi-dimensional random field is a vector of random-valued functions in which each of the X_i is a random function of position: $X_i = X_i(\mathbf{r})$. We denote this by $\mathbf{X}(\mathbf{r}) = (X_1(\mathbf{r}), \ldots, X_n(\mathbf{r}))$. Now we can, for example, cross-correlate any pair of functions $X_i(\mathbf{r})$ and $X_j(\mathbf{r})$ at the same point $\mathbf{r}$, or we can cross-correlate $X_i(\mathbf{r})$ with itself at two different points. The most general 2-point case is the following.

For an isotropic homogeneous multi-dimensional random field $\mathbf{X}(\mathbf{r}) = (X_1(\mathbf{r}), \ldots, X_n(\mathbf{r})$ we have

Statistically isotropic random field

$$\mathbb{E}\,[X_i(\mathbf{r}] = m_i, \qquad \mathbb{E}\,[(X_i(\mathbf{r}) - m_i)^2] = \sigma_i^2, \tag{1.27}$$

$$Q_{ij}\Big(X_i(\mathbf{r}), X_j(\mathbf{s})\Big) = C_{ij}(\mid \mathbf{r} - \mathbf{s} \mid). \tag{1.28}$$

where m_i, the mean of X_i, σ_i, the variance of X_i, and C_{ij} are independent of position and isotropic.

1.4 N-point descriptors of random fields

1.4.1 N-point distributions: scalar fields

There is the obvious generalisation of this to the mean of m quantities evaluated at n points (clearly $m \geq n$). This brings with it a significant amount of mathematical complexity and also a substantial amount of computational complexity for n values large than 4 or 5 unless

Fig. 1.1 Distinct configurations of $n-1$ bonds in an n-point graph in which the points are labelled. (Adapted from Fry (1984, his figure 4.))

attention is restricted to specific configurations of the points (*e.g.* equilateral triangles of 3 points).

This is not so simple - needs rephrasing! Is it only connected graphs or what?

It is important to remember that, for our homogeneous random fields where these correlations depend only on the configuration of the points and not the location of the configuration, the configuration of n points is determined by a graph of $n-1$ vectors joining pairs of points. There are then only $n-1$ independent quantities specifying the configuration of points. For $n = 2, 3$ points the linking vectors have single configuration, but when $n \geq 4$ there are multiple configurations. The first few configurations for $N = 2, 3, 4, 5$ points are depicted in Figure 1.1 (based on Fry (1984, his figure 4.)).

To keep things simple we should fix attention on zero-mean fields, like our density fluctuation field $\delta(\mathbf{r})$ defined at points $\mathbf{r} \in \mathbb{R}^3$. In this case we can imagine taking the expectation value of products of powers of $\delta(\mathbf{r}_i)$, $i = 1, \ldots, n$:

$$Q^{k_1,\ldots,k_n}_{12\ldots n} = \mathbb{E}\,[\delta(\mathbf{r}_1)^{k_1}\delta(\mathbf{r}_2)^{k_2}\ldots\delta(\mathbf{r}_n)^{k_n}], \quad n \leq k = k_1 + \cdots + k_n \tag{1.29}$$

This is referred to as the n-point moment of k^{th} order. In practice we deal with the n^{th} moments of order $k = n$:

$$Q(\mathbf{r}) = \mathbb{E}\,[\delta(\mathbf{r})^2], \quad Q(\mathbf{r}_1, \mathbf{r}_2) = \mathbb{E}\,[\delta(\mathbf{r}_1)\delta(\mathbf{r}_2)],$$
$$Q(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \mathbb{E}\,[\delta(\mathbf{r}_1)\delta(\mathbf{r}_2)\delta(\mathbf{r}_3)] \tag{1.30}$$

It should be noted that, despite the way these are written, in a homogeneous random field the values of these functions will depend only on the relative position of the reference points. Thus $Q(\mathbf{r})$ will be independent of position, and $Q(\mathbf{r}_1, \mathbf{r}_2)$ will depend only on $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$. The three-point function $Q(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ will depend on only two of the vectors $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$. If the field is also statistically isotropic, the functions will be independent of the orientation of the graph joining the points, and so $Q(\mathbf{r}_1, \mathbf{r}_2)$ will be a function of $|\,\mathbf{r}_1 - \mathbf{r}_2\,|$, and so on. This is the generalisation of the 2-point results (1.18).

Peebles (1980, Ch. III §§31-36) discusses and presents results for the correlations up to fourth order of the galaxy distribution and the moments of counts in cells (see also Fry and Peebles (1978)).

1.4.2 N-point distributions: vector fields

The field $\delta(\mathbf{r})$ is a scalar field and it is important to generalise (1.29) to vector and tensor fields. In hydrodynamic turbulence, for example, it is necessary to deal with correlations of vector functions such as velocities (Batchelor (1982, §2.3), Monin and Yaglom (1971, §4.2)).

Consider a random field of vectors $\boldsymbol{\nu}(\mathbf{r})$ with $\mathbf{r} \in \mathbb{R}^3$, so that each vector has three components. Without any loss of generality we will assume that our frame of reference is chosen so that $\langle \boldsymbol{\nu}(\mathbf{r}) \rangle = \mathbb{E}[\boldsymbol{\nu}(\mathbf{r})] = 0$. With this we can write down correlation functions between the components of the individual vectors $\boldsymbol{\nu}(\mathbf{r}_i),\ i = 1, \ldots, n$:

$$Q^{(n)}_{pq\ldots s}(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n) = \mathbb{E}[\nu_p(\mathbf{r}_1), \nu_q(\mathbf{r}_2), \ldots, \nu_s(\mathbf{r}_n)] \tag{1.31}$$

The $p, q, \ldots, s$ are labels for the n points and $n \leq k$. There are 3^k different moments of order k for vector fields in $\mathbb{R}^3$. As in the scalar case (1.29), the value of $Q^{(n)}_{pq\ldots s}$ depends only on $n-1$ of the point positions (or linear combinations thereof).

The notation is a problem here. For illustration purposes, we can write examples of two-point second, third and fourth order correlations as

$$R_{ij}(\mathbf{r}, \mathbf{r}') = \mathbb{E}[\nu_i(\mathbf{r})\, \nu_j(\mathbf{r}')] \tag{1.32}$$

$$R_{ij,k}(\mathbf{r}, \mathbf{r}') = \mathbb{E}[\nu_i(\mathbf{r})\, \nu_j(\mathbf{r})\, \nu_k(\mathbf{r}')] \tag{1.33}$$

$$R_{ij,km}(\mathbf{r}, \mathbf{r}') = \mathbb{E}[\nu_i(\mathbf{r})\, \nu_j(\mathbf{r})\, \nu_k(\mathbf{r}')\, \nu_m(\mathbf{r}')] \tag{1.34}$$

$$R_{ijk,m}(\mathbf{r}, \mathbf{r}') = \mathbb{E}[\nu_i(\mathbf{r})\, \nu_j(\mathbf{r})\, \nu_k(\mathbf{r})\, \nu_m(\mathbf{r}')] \tag{1.35}$$

where the comma in $R_{ij,k}$ *etc.* denotes which indices refer to which of the two points. An important practical example from the theory of turbulence is the two-point velocity component correlation function:

$$R_{ij}(\mathbf{r}, \mathbf{r}') = \mathbb{E}[u_i(\mathbf{r})\, u_j(\mathbf{r}')] \tag{1.36}$$

where $u_i(\mathbf{r})$ is the i^{th} component of the zero-mean velocity field $\mathbf{u}(\mathbf{r})$ at the point $\mathbf{r}$.

1.4.3 Robertson's analysis of statistically symmetric fields

Robertson (1940)[10] was able to provide a substantial simplification of equations (1.31) in circumstances where the random field exhibited some statistical symmetry, such as spherical or axial symmetry (see Batchelor (1982, §3.3) for a detailed discussion).

In the case where the random field has both spherical symmetry and reflection symmetry then the 2-point $Q^{(n)}_{pq\ldots s}$ take on the form

$$Q_i(\mathbf{r}) = A(r^2)\, r_i \tag{1.37}$$

$$Q_{ij}(\mathbf{r}) = A(r^2)\, r_i r_j + B(r^2)\, \delta_{ij} \tag{1.38}$$

$$Q_{ijk}(\mathbf{r}) = A(r^2)\, r_i r_j r_k + B(r^2)\, r_i \delta_{jk} + C(r^2)\, r_j \delta_{ki} + D(r^2)\, r_k \delta_{ij} \tag{1.39}$$

The functions A, B, C, D here are even functions of r and are to be determined from the physics underlying the random field.

There is a particularly important 3-point function that arises in turbulence theory:

$$S_{ij,k}(\mathbf{r}) = \mathbb{E}[u_i(\mathbf{x})\, u_j(\mathbf{x})\, u_k(\mathbf{x}+\mathbf{r})] \tag{1.40}$$

$$= A\, r_i r_j r_k + B\,(r_i \delta_{jk} + r_j \delta_{ik}) + D r_k \delta_{ij}, \tag{1.41}$$

[10] The same H.P. Robertson who had previously exploited symmetry principles to shed light on homogeneous and isotropic solutions of the Einstein field equations. Robertson exploited his understanding of the invariants of the rotation group in three dimensions to derive these remarkable results that proved so valuable in simplifying earlier discussions of the statistical theory of turbulence. Robertson's paper *loc. cit.* is a model of clarity.

where, as before, A, B, C, D are even functions of r. This is just equation (1.33) expressed in the case of a statistically isotropic field and comes from (1.39) by virtue of the symmetry on the indices i, j.

In the slightly less restrictive case that the random field has spherical symmetry, but without reflection symmetry, we have, for example

$$Q_{ij}(\mathbf{r}) = A(r^2)\, r_i r_j + B(r^2)\, \delta_{ij} + C(r^2)\, \epsilon_{ijk} r_k \tag{1.42}$$

where ϵ_{ijk} is the antisymmetric permutation symbol.

1.5 Variance and correlation function

At the simplest level we can ask how irregular is the density field. Empirically we would do this by measuring the density at a large number of points and plotting a histogram of the values. The mean of the histogram would be an estimate of the mean density, and the width of the histogram would tell us the variance of the density, indicating by how much on average the density deviates from its mean value.[11]

In the limit of a very large sample of points we have

$$\langle \delta(\mathbf{x}) \rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \delta(\mathbf{x_i}) = 0 \tag{1.43}$$

$$\langle \delta(\mathbf{x})^2 \rangle = \lim_{N\to\infty} \frac{1}{N-1} \sum_{i=1}^{N} \delta(\mathbf{x_i})^2 = \sigma^2 \tag{1.44}$$

The first of these follows from the requirement that $\delta(\mathbf{x})$ has zero mean, and the second is telling us something about how jiggly the random function is. If we want to know how symmetric the fluctuations $\delta(\mathbf{x})$ are about the mean we take a suitable average of $\delta(\mathbf{x})^3$. These are the statistical moments of the distribution of the zero-mean random function $\delta(\mathbf{x})$. Note that if we do not take the limit, these become mere estimators of the mean and variance and will vary from sample to sample.

1.5.1 Spatial autocorrelation function

These moments tell us about the statistical variations in density at a single point, they do not tell us anything about how the densities at two neighbouring points may be related.

[11] It should be noted that, in taking this approach, we are making the important assumption that the statistics of the density field at one place is the same as in any other place. In other words, if we were to take sufficiently large samples in large volumes that are very far apart, the histograms of the sampled density distributions would be consistent with having been drawn from the same underlying statistical distribution.

This is achieved by calculating the function

$$\xi^{(N)}(\mathbf{r}) = \frac{1}{N}\sum_{i=1}^{N}\delta(\mathbf{x}_i)\,\delta(\mathbf{x}_i+\mathbf{r}) \tag{1.45}$$

$$\xi(\mathbf{r}) = \lim_{N\to\infty}\xi^{(N)}(\mathbf{r}) \tag{1.46}$$

The sum is understood as being the sum of the products of N pairs of δ values separated by a (vector) distance $\mathbf{r}$. The superscript (N) indicates that the average is taken using N positions $\mathbf{x}_i$. If we consider a sample volume V and let the number of sample points tend to infinity, we can replace (1.46) by an integral:

Spatial auto-correlation of a zero-mean random field $\delta(\mathbf{x})$

$$\xi(\mathbf{r}) = \lim_{V\to\infty}\int\delta(\mathbf{x})\,\delta(\mathbf{x}+\mathbf{r})\,d\mathbf{x} \tag{1.47}$$

where the integral is taken to be over all space $\{\mathbf{x}\}$.[a]

[a] If $\delta(\mathbf{x})$ were a complex-valued field then this would be $\xi(\mathbf{r}) = \int\delta(\mathbf{x})^*\,\delta(\mathbf{x}+\mathbf{r})\,d\mathbf{x}$, where the star, *. denotes complex conjugation.

In cosmology, $\xi(\mathbf{r})$ is called the *correlation function* of the random field $\delta(\mathbf{x})$. It is the simplest descriptor of the spatial nature of the random field $\delta(\mathbf{x})$ and plays a very important role in modern cosmology. Note that $\xi(\mathbf{0}) = \sigma^2$ (compare equations (1.46) and (1.44)). So, just as the variance is a measure of the amplitude of the variations in $\delta(\mathbf{x})$ measured at a sample of points, $\xi(\mathbf{r})$ measures the size of the difference between the $\delta(\mathbf{x})$ measured at two points separated by some (vector) distance.

1.5.2 The Fourier description

It is useful to think of a random field as a superposition of sine and cosine waves having amplitudes and phases chosen at random from some specified distributions or, equivalently, as a superposition of functions of the form $e^{i\mathbf{k}\cdot\mathbf{x}}$. Roughly speaking , we can represent the density field $\delta(\mathbf{x})$ in some finite sample volume V as the sum of a set of waves having wave-numbers $\mathbf{k}$ and complex amplitude $\delta_{\mathbf{k}}$ [12]:

$$\delta(\mathbf{x}) = \sum_{\mathbf{k}}\delta_{\mathbf{k}}\,e^{i\mathbf{k}\cdot\mathbf{x}} \tag{1.48}$$

We are to think in terms of the field $\delta(\mathbf{x})$ occupying an infinite volume and so the sum is to be taken in the limit as the volume becomes infinite.[13]

[12] The notation using the symbols $\delta(\mathbf{x})$ and $\delta_{\mathbf{k}}$ as a Fourier transform pair is perhaps unfortunate, but this is standard in cosmological perturbation theory.

[13] The more mathematically minded reader might be justifiably concerned that we are dealing with random functions, $f(\mathbf{x})$, that are certainly not integrable in the strict Riemann sense. Such functions are not even differentiable in the normal sense. There are conditions under which the standard Fourier integral is mathematically valid: either the field $f(\mathbf{x})$ is periodic or $f(\mathbf{x})$ is such that $\int|f(\mathbf{x})|d\mathbf{x}$ is bounded. Neither of these conditions is realised for a stationary random field of density perturbations (Batchelor, 1982, §2.5).

The fact that the field $\delta(\mathbf{x})$ takes real number values imposes conditions on the complex amplitudes $\delta_{\mathbf{k}}$. We can write the complex amplitude in terms of its modulus and phase as $\delta_{\mathbf{k}} = |\delta_{\mathbf{k}}|\, e^{i\phi(\mathbf{k})}$, and then

$$\delta(\mathbf{x}) = \sum_{\mathbf{k}} \delta_{\mathbf{k}}\, e^{i\mathbf{k}\cdot\mathbf{x}} = \sum_{\mathbf{k}} \delta^*_{\mathbf{k}}\, e^{-i\mathbf{k}\cdot\mathbf{x}} \quad \text{and} \quad \phi(-\mathbf{k}) = -\phi(\mathbf{k}) \tag{1.49}$$

Here, $\delta^*_{\mathbf{k}}$ denotes the complex conjugate of $\delta_{\mathbf{k}}$.

1.5.3 Spatially infinite homogeneous random functions

We are interested in inverting equation (1.48) in order to see what is the make-up of the field $\delta(\mathbf{x})$ in terms of sinusoidal components. Since equation (1.48) is a Fourier representation of a statistically homogeneous random field we can in principle invert it by the standard Fourier techniques. However, this is technically not entirely straightforward[13] (there is a somewhat more technical discussion in Appendix **??** and in particular section **??**).

One way to deal with this simply is to consider the limit of a cubic box of volume V covered by a regular $n \times n \times n$ grid having $N = n^3$ cells. We then consider a limit in which the grid becomes finer, $N \to \infty$ and the volume $V \to \infty$. Then we can write things like

$$\delta(\mathbf{x}) = \sum_{N}^{(3)} \delta_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}} \tag{1.50}$$

where by $\sum_N{}^{(3)}$ is a short-hand for summing over the $N = n^3$ grid points in three dimensions, where the $\mathbf{x}$ and $\mathbf{k}$ are to be taken at those grid points in the triple sum. Formally, for any N and V this is the discrete Fourier representation of the field $\delta(\mathbf{x})$: there are no approximations involved and no issues of convergence. It has a well defined inverse

$$\delta_{\mathbf{k}} = \frac{1}{N} \sum_{N}^{(3)} \delta(\mathbf{x})\, e^{-i\mathbf{k}\cdot\mathbf{x}} \tag{1.51}$$

An alternative way of presenting this point is to regard (1.48) as a Fourier series for which the individual terms are

$$\delta_{\mathbf{k}} = \lim_{V\to\infty} \frac{1}{V} \int_V \delta(\mathbf{x})\, e^{-i\mathbf{k}\cdot\mathbf{x}}\, d\mathbf{x} \tag{1.52}$$

In practice, the integral has to be evaluated on a grid in a computer, and then it comes down to dealing with equation (1.51).

It should be noted that equations (1.51) and (1.50) are exact on every rectangular $\mathbf{x}$-space and $\mathbf{k}$-space $N = n \times n \times n$ sampling grid and so we should perhaps write these fields as $\delta^{(N)}_{\mathbf{x}}$ and $\delta^{(N)}_{\mathbf{k}}$. If these sums converged in the limit $N \to \infty$ these $\delta^{(N)}$'s should become the same δ's as in (1.52).

Using the fact that the mean of the integral is the integral of the mean, equation (1.51) shows us that

$$\langle \delta_{\mathbf{k}} \rangle = 0 \tag{1.53}$$

In a practical sense we can think of a finite box in which $\delta(\mathbf{x})$ is sampled on a regular grid and on which the field $\delta(\mathbf{x})$ has periodic boundary conditions. Formally our Fourier representation is then just the discrete Fourier transform (see Appendix **??**). The limit is achieved by considering an ever-larger box with an ever finer grid resolution.

since, by construction, $\langle\delta(\mathbf{x})\rangle = 0$. The Fourier components have zero mean.

1.5.4 Correlation function in Fourier space

The $1/N$'s are probably wrong!!!

We can now calculate the correlation function $\xi(\mathbf{r})$ of the zero-mean real random field $\delta(\mathbf{x})$ using (1.47) and (1.48):

$$\xi(\mathbf{r}) = \frac{1}{V}\int_V \delta(\mathbf{x})\,\delta(\mathbf{x}+\mathbf{r})d\mathbf{x} \overset{V}{=} \frac{1}{N}\sum_{\mathbf{k}} |\delta_{\mathbf{k}}|^2\, e^{i\mathbf{k}\cdot\mathbf{r}} \tag{1.54}$$

The averaging has crept in with no comment!

The second '=' sign is marked with a 'V' to remind us that this result pertains to a finite sample volume V. If we let the volume $V \to \infty$ we can write

$$\xi(\mathbf{r}) = \lim_{N\to\infty} \frac{1}{N}\sum_{\mathbf{k}} \left\langle |\delta_{\mathbf{k}}|^2 \right\rangle\, e^{i\mathbf{k}\cdot\mathbf{r}} \tag{1.55}$$

We note that setting $\mathbf{r} = 0$ in this equation and using (1.44) and (1.46) gives us

$$\xi(0) = \frac{1}{N}\sum_{\mathbf{k}} \left\langle |\delta_{\mathbf{k}}|^2 \right\rangle = \sigma^2 \tag{1.56}$$

where σ^2 is the variance of the fluctuations in the random field as given by equation (1.44).

Equation (1.56) provides an important insight to the nature of the Fourier representation. If the amplitudes $\delta_{\mathbf{k}}$ are chosen from some statistical distribution having zero mean and variance $\left\langle |\delta_{\mathbf{k}}|^2 \right\rangle$, then the total variance of the random field, σ, is just the sum of the variances of the individual Fourier components.

1.5.5 Power spectral density

We now introduce the important quantity

Power spectral density

$$\langle\delta(\mathbf{k}_1)\delta^*(\mathbf{k}_2)\rangle = \mathcal{P}(\mathbf{k}_1)\delta^D(\mathbf{k}_1 - \mathbf{k}_2) \tag{1.57}$$

The function $\delta^D(\mathbf{x})$ is the standard Dirac delta function. The superscript has been added to avoid confusion with the symbol $\delta(\mathbf{x})$ that is conventionally used for the relative density fluctuation field.

Note that the $\mathcal{P}(\mathbf{k}) \geq 0$: the power spectrum is never negative.

This is called the *Power Spectrum* or *Spectral density* of the random process $\delta(\mathbf{x})$.[14] When $\delta(\mathbf{k})$ is the Fourier transform of a real-valued field, $\delta^*(\mathbf{k}) = \delta(-\mathbf{k})$, (1.57) can be written

$$\langle\delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\rangle = \mathcal{P}(\mathbf{k}_1)\delta^D(\mathbf{k}_1 + \mathbf{k}_2) \tag{1.59}$$

[14] This is often written, with lack of mathematical rigour, as

$$\mathcal{P}(\mathbf{k}) = \left\langle |\,\delta_{\mathbf{k}}\,|^2 \right\rangle \tag{1.58}$$

and is, in fact, incorrect. See Bertschinger (1992, §5.8.1) for a good discussion of this.

With this equation (1.55) can be written as

$$\xi(\mathbf{r}) = \sum_{\mathbf{k}} \mathcal{P}(\mathbf{k})\, e^{i\mathbf{k}\cdot\mathbf{r}} \tag{1.60}$$

which by Fourier transforming can be shown to be equivalent to

$$\mathcal{P}(\mathbf{k}) = \lim_{V\to\infty} \frac{1}{V} \int_V \xi(\mathbf{r})\, e^{-i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r} \tag{1.61}$$

This important equation relating the correlation function $\xi(\mathbf{r})$ to the moduli of the Fourier amplitudes $\delta_{\mathbf{k}}$ is called the *Weiner-Khintchine theorem* which says that for any random field, the power spectrum is the Fourier transform of the correlation function.

We note, as an aside, that this theorem tells us that not every function $\xi(\mathbf{r})$ can be the correlation function of some random process. $\xi(\mathbf{r})$ must be such that the integral (1.61) is nowhere negative.

We should also note that it is generally supposed that the random process $\delta(\mathbf{r})$ is isotropic, by which we mean that the correlation function is independent of direction:

$$\xi(\mathbf{r}) = \xi(|\mathbf{r}|) = \xi(r) \tag{1.62}$$

The fluctuations $\delta(\mathbf{x})$ and $\delta(\mathbf{x}+\mathbf{r})$ at widely separated points, ie. large $|\mathbf{r}|$, should be independent of one another and so:

$$\xi(r) \to 0, \quad r \to \infty \tag{1.63}$$

This is in fact related to an underlying assumption that the random process $\delta(\mathbf{r})$ is *ergodic*.

1.6 The 3-point function and the Bispectrum

The Bispectrum is the Fourier Transform of the 3-point correlation function. This is analogous to the the power spectrum being the Fourier transform of the 2-point correlation function.

Discussion of the Bispectrum goes back to the 1980's and before with the work on the clustering of galaxies in catalogues (see Peebles (1980, §43)). Much of that work was devoted to the analysis of counts of galaxies in cells on the sky, or the analysis of discrete point sets and the goals were generally the analysis of the sky-projected data that was available at the time. So we find, for example, extra terms in the equations that are due to the discrete nature of the sample.

For a zero-mean field $\delta(\mathbf{x})$, the 2-point and 3-point spatial correlation functions of the field are respectively

$$\xi_{12} = \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\rangle \tag{1.64}$$

$$\zeta_{123} = \langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\delta(\mathbf{x}_2)\rangle \tag{1.65}$$

The 2-point function was found to be well approximated by a simple isotropic power law, $\xi(r) \propto r^{-\gamma}$ over a wide rangle of separations $r = |\,\mathbf{r}_1 - \mathbf{r}_2\,|$. Moreover, the galaxy 3-point

function ζ_{123} had been found to be expressible in terms of the 2-point function, ξ_{12} by the relationship

$$\zeta_{123} = Q\,(\xi_{12}\xi_{13} + \xi_{12}\xi_{23} + \xi_{13}\xi_{23}) \tag{1.66}$$

for some data-determined constant Q. The special feature of equation (1.66) is the absence of a triple-term $\xi_{12}\xi_{13}\xi_{23}$ (which would be easy to identify since it would dominate at small separations).

1.6.1 Fourier transform of the 3-point function

We define the *Bispectrum* as the expectation value of products of three Fourier components, $\delta(\mathbf{k})$ of the density field, $\delta(\mathbf{x})$:

$$\langle\delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)\rangle = (2\pi)^3 B(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)\delta^D(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3) \tag{1.67}$$

Likewise we can define higher order spectra, such as the *Trispectrum* $\langle\delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)\delta(\mathbf{k}_3)\rangle$ and so on. The *Normalised Bispectrum* is simply

$$Q(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) = \frac{B(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3)}{\mathcal{P}(k_1)\mathcal{P}(k_2)+\mathcal{P}(k_2)\mathcal{P}(k_3)+\mathcal{P}(k_3)\mathcal{P}(k_1)} \tag{1.68}$$

This normalisation simply removes dependence on overall size and concentrates on triangle shape. For simple models of random functions this is analytically calculated for triangles having special configurations of their sides $\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3$, such as equilateral triangles. For a thorough discussion of the Power Spectrum and Bi-Spectrum as applied to galaxy redshift surveys see Jeong (2010, Ch. 7).

Fry and Seldner (1982) showed that the Bispectrum corresponding specifically to the 3-point function (1.66) is simply related to the Power spectrum $\mathcal{P}(k)$ by

$$B(k_1,k_2,k_3) = Q\,(\mathcal{P}(k_1)\mathcal{P}(k_2)+\mathcal{P}(k_2)\mathcal{P}(k_3)+\mathcal{P}(k_3)\mathcal{P}(k_1)) \tag{1.69}$$

for the same value of Q as in (1.66). This point was then taken up, within the framework of linear perturbation theory, in the seminal paper of Fry (1984) who showed the remarkable result that[15]

$$B(k_1,k_2,k_3) = \left[\frac{10}{7} + \frac{\mathbf{k_1}\cdot\mathbf{k_2}}{k_1k_2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right) + \frac{4}{7}\left(\frac{\mathbf{k_1}\cdot\mathbf{k_2}}{k_1k_2}\right)^2\right]\mathcal{P}(k_1)\mathcal{P}(k_2) + \text{sym.} \tag{1.70}$$

This was derived for an Einstein-de Sitter universe with zero cosmological constant. For $\Omega \neq 1$ the factors $10/7$ and $4/7$ become $1+\kappa$ and $1-\kappa$ respectively, where $\kappa \simeq 3\Omega^{-2/63}/7$, and, for a flat non-zero Λ model, $\kappa \simeq 3\Omega_m^{-1/143}/7$ (Matarrese et al., 1997; Scoccimarro et al., 1998; Scoccimarro, 2000).

For equilateral triangle configurations, equation (1.70) gives $Q = \frac{4}{7}$. Somewhat later, Baumgart and Fry (1991) were able to verify this for a variety of the then-available redshift catalogues. The quantity in square brackets is referred to as the *gravitational instability*

[15] See also the subsequent, more general discussions, of Goroff et al. (1986), Bernardeau (1992), Catelan and Moscardini (1994a,b) using higher order perturbation theories to describe the evolution of non-Gaussianity in the density and velocity fields. Matarrese et al. (1997) brought this forward by taking account of galaxy bias.

kernel and is characteristic of the theory of gravity. A different gravitational theory would yield a different kernel.

1.6.2 Bispectrum as a measure of non-Gaussianity

The Bispectrum is the lowest order statistic that that can discriminate between a Gaussian and a non-Gaussian field having the same variance. A non-zero Bispectrum indicates a component of skewness in the underlying distribution.

Verde et al. (2000) considered a model in which the power spectrum of the density fluctuation field $\delta(\mathbf{x})$ grows nonlinearly with redshift z, following an equation of the form

$$\mathcal{P}_0(k,z) = \mathcal{M}_k(z)^2\,\mathcal{P}_\delta(k) \tag{1.71}$$

where $\mathcal{P}_\Phi(k)$ is the primordial scale invariant spectrum. The function $\mathcal{M}_k(z)$ relates to the matter transfer function $T(k)$ via

$$\mathcal{M}_k(z) = \frac{2}{3}\frac{D(z)}{H_0^2\Omega_{m,0}}k^2T(k) \tag{1.72}$$

where H_0 and $\Omega_{m,0}$ are respectively the current values of the Hubble constant and the matter density parameter.

The density field $\delta(\mathbf{x})$ is divided into two components, one Gaussian and the other non-Gaussian. The simplest representation of such a field is

$$\delta(\mathbf{x}) = \delta_0(\mathbf{x}) + f_{NL}[\delta_0(\mathbf{x})^2 - \langle\delta_0(\mathbf{x})^2\rangle] \tag{1.73}$$

where $\delta_0(\mathbf{x})$ is Gaussian distributed. We see that as $\delta_0(\mathbf{x})$ increases towards the nonlinear regime, the amplitude $\delta(\mathbf{x})$ develops an increasingly important non-Gaussian component.

The outcome from this simple model is that (Verde et al., 2000):

$$B(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) \simeq \left[f_{NL}\frac{\mathcal{M}(k_3)}{\mathcal{M}(k_1)\mathcal{M}(k_2)} + J(\mathbf{k}_1,\mathbf{k}_2)\right]\mathcal{P}(k_1)\mathcal{P}(k_2) + \text{sym.} \tag{1.74}$$

Here the function $J(\mathbf{k}_1,\mathbf{k}_2)$ is a gravitational instability kernel (*cf.* equation 1.70 and comments thereafter) for the relevant cosmological model.

Other factors such a bias can be included in this model. This is important because in the analysis of galaxy redshift catalogues we observe only the luminous galaxy component $\delta_g(\mathbf{x})$ of the matter distribution, and not the underlying dark matter component $\delta_{DM}(\mathbf{x})$. This simple model can be generalised by writing

$$\delta_g(\mathbf{x}) = b_1(z)\delta_{DM}(\mathbf{x}) + b_2(z)[\delta_{DM}(\mathbf{x})^2 - \langle\delta_{DM}(\mathbf{x})^2\rangle] \tag{1.75}$$

with which (1.74) becomes (Matarrese et al., 1997):

$$B_g(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) \simeq b_1(z)^3\left[\frac{b_2}{b_1^3}\frac{\mathcal{M}(k_3)}{\mathcal{M}(k_1)\mathcal{M}(k_2)} + J(\mathbf{k}_1,\mathbf{k}_2) + \frac{b_2(z)}{2b_1(z)}\right]\mathcal{P}(k_1)\mathcal{P}(k_2) + \text{sym.} \tag{1.76}$$

The additional complication, not included in this model, is the redshift-space distortion (Sefusatti and Komatsu, 2007).

1.6.3 Bispectrum as a clue to inflation theories

Inflationary theories of the very early universe (Salopek and Bond, 1990) are expected to generate random curvature fluctuations, $\Phi(\mathbf{x})$, having a non-Gaussian curvature component. This provided the motivation for writing the field $\Phi(\mathbf{x})$ as the sum of a Gaussian distributed field, $\Phi_0(\mathbf{x})$, and a non-Gaussian component of the form:

$$\Phi(\mathbf{x}) = \Phi_0(\mathbf{x}) + f_{NL}[\Phi_0^2(\mathbf{x} - \langle\Phi_0^2(\mathbf{x})\rangle], \tag{1.77}$$

where the parameter f_{NL} quantifies the nonlinearity. The prospect of measuring f_{NL} may provide a test for theories of inflation.

The Fourier transform of equation (1.77) can be written as

$$\Phi(\mathbf{k}) = \Phi_L(\mathbf{k}) + \Phi_{NL}(\mathbf{k}) \tag{1.78}$$

It can be shown that (Komatsu and Spergel, 2001) that the cross-bispectrum between the linear and nonlinear components of $\Phi(\mathbf{x})$ is

$$\langle\Phi_L(\mathbf{k}_1)\Phi_L(\mathbf{k}_2)\Phi_{NL}(\mathbf{k}_3)\rangle = 2(2\pi)^3 f_{NL}\mathcal{P}_\Phi(k_1)\mathcal{P}_\Phi(k_1)\delta^D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \tag{1.79}$$

Here, $\mathcal{P}_\Phi(k_1)$ is the power spectrum of the linear component $\Phi_L(\mathbf{x})$.

1.7 Homogeneous and isotropic random fields

This is somewhat out of place now

In the limit of a statistically homogeneous and isotropic random field, the correlation function $\xi(\mathbf{r})$, defined as a statistical expectation $\xi(\mathbf{r}) = \mathbb{E}[\delta(\mathbf{x}\delta(\mathbf{x}+\mathbf{r})]$, has a Fourier transform $\mathcal{P}(\mathbf{k})$ such that

$$\xi(\mathbf{r}) = \int \mathcal{P}(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}}d\mathbf{k}, \quad \mathcal{P}(\mathbf{k}) = \frac{1}{(2\pi)^3}\int \xi(\mathbf{r})e^{-i\mathbf{k}\cdot\mathbf{r}}d\mathbf{r} \tag{1.80}$$

Provided that $\xi(\mathbf{r}) \to 0$ sufficiently fast as $|\,\mathbf{r}\,| \to 0$, both of these integrals exist whether or not the underlying field $\delta(\mathbf{r})$ has a Fourier transform.

The field $\delta(\mathbf{r})$ is a zero-mean real valued field and so $\xi(\mathbf{r}) = \xi(-\mathbf{r})$ and $\mathcal{P}(\mathbf{k}) = \mathcal{P}(-\mathbf{k})$. Consequently, equations (1.80) can be written

$$\xi(\mathbf{r}) = \int \mathcal{P}(\mathbf{k})\cos(\mathbf{k}\cdot\mathbf{r})\,d\mathbf{k}, \quad \mathcal{P}(\mathbf{k}) = \frac{1}{(2\pi)^3}\int \xi(\mathbf{r})\cos(\mathbf{k}\cdot\mathbf{r})\,d\mathbf{r} \tag{1.81}$$

A further simplification occurs if $\delta(\mathbf{r})$ is statistically isotropic, in which case $\xi(\mathbf{r}) = \xi(r)$. In this case it can be shown that $\mathcal{P}(\mathbf{k}) = \mathcal{P}(k)$: the power spectral density is also isotropic. Equations (1.81) then become

$$\xi(r) = 4\pi\int_0^\infty \frac{\sin kr}{kr}\mathcal{P}(k)\,k^2dk, \quad \mathcal{P}(k) = \frac{1}{2\pi^2}\int_0^\infty \frac{\sin kr}{kr}\xi(r)\,r^2dr \tag{1.82}$$

To obtain these results it is necessary to convert to spherical polar coordinates and integrate out the angular components.[16]

1.7.1 A mathematical diversion

This is somewhat out of place now

The more mathematically minded reader might justifiably be concerned that we are dealing with random functions, $u(\mathbf{x})$, that are certainly not integrable in the strict Riemann sense. Such functions are not even differentiable in the normal sense. There are conditions under which the standard Fourier integral is mathematically valid: either the field $u(\mathbf{x})$ is periodic or $u(\mathbf{x})$ is such that $\int |u(\mathbf{x})| d\mathbf{x}$ is bounded. Neither of these conditions is realised for a stationary random field of density perturbations (Batchelor, 1982, §2.5).

The "way out" of this is to assume that the field exists in some finite box having periodic boundary conditions, and to let the box size go to infinity. This is the concept that has implicitly been used in the preceding discussion.

The mathematically correct way to handle this was first presented by Wiener (1930). Formally one should write the Fourier representation of a random function $u(\mathbf{x})$ as a *Fourier-Stieltjes* integral:

$$u(\mathbf{x}) = \int e^{i\mathbf{k}\cdot\mathbf{r}} dZ(\mathbf{k}) \tag{1.84}$$

One should note, however, that even this approach does not work for $1/f$ noise.

Short biography: **Thomas Stieltjes (1856-1894** Thomas Stieltjes was born in Holland and died in France at the age of 38. He had worked as an assistant to the director of the Leiden observatory, until 1883 when he gave up that position to become a mathematician at the University of Delft. The following year he applied for a position in Groningen, but was turned down by the education department because he lacked the necessary diplomas. Stieltjes' friend and fellow mathematician Charles Hermite solved this problem by getting the University of Leiden to grant Stieltjes an honorary doctorate (Lorentz gave his support for the award). In 1885 he and his family went to Paris and shortly thereafter he completed a PhD thesis at the École Normale Superieur under the stewardship of Hermite and Darboux. In 1889 he took up a professorship in Toulouse, France.

[16] In two dimensions on a plane there are analogous results:

$$\xi(r) = 2\pi \int_0^\infty J_0(kr)\mathcal{P}(k)\, k dk, \quad \mathcal{P}(k) = \frac{1}{2\pi} \int_0^\infty J_0(kr)\xi(r)\, r dr \tag{1.83}$$

where $J_0(x)$ is the standard Bessel function and r and k are the lengths of the 2-vectors $\mathbf{r} = (x_1, x_2)$ and $\mathbf{k} = (k_1, k_2)$: $r = (x_1^2 + x_2^2)^{1/2}$ and $k = (k_1^2 + k_2^2)^{1/2}$.

References

Adler, R.J. 1981. *The geometry of Random Fields.* John Wiley & Son.

Batchelor, G.K. 1982. *The Theory of Homogeneous Turbulence*. Cambridge Science Classics. Cambridge University Press. First published 1953.

Baumgart, D. J., and Fry, J. N. 1991. Fourier spectra of three-dimensional data. ApJ, **375**(July), 25–34.

Bernardeau, F. 1992. The gravity-induced quasi-Gaussian correlation hierarchy. ApJ, **392**(June), 1–14.

Bertschinger, E. 1992. Large-Scale Structures and Motions: Linear Theory and Statistics. Page 65 of: Martinez, V. J., Portilla, M., and Saez, D. (eds), *New Insights into the Universe*. Lecture Notes in Physics, Berlin Springer Verlag, vol. 408.

Catelan, P., and Moscardini, L. 1994a. Kurtosis and large-scale structure. ApJ, **426**(May), 14–18.

Catelan, P., and Moscardini, L. 1994b. Kurtosis as a non-Gaussian signature of the large-scale velocity field. ApJ, **436**(Nov.), 5–10.

Fry, J. N. 1984. The Galaxy correlation hierarchy in perturbation theory. ApJ, **279**(Apr.), 499–510.

Fry, J. N., and Peebles, P. J. E. 1978. Statistical analysis of catalogs of extragalactic objects. IX - The four-point galaxy correlation function. ApJ, **221**(Apr.), 19–33.

Fry, J. N., and Seldner, M. 1982. Transform analysis of the high-resolution Shane-Wirtanen Catalog - The power spectrum and the bispectrum. ApJ, **259**(Aug.), 474–481.

Goroff, M. H., Grinstein, B., Rey, S.-J., and Wise, M. B. 1986. Coupling of modes of cosmological mass density fluctuations. ApJ, **311**(Dec.), 6–14.

Jeong, D. 2010 (Aug.). *Cosmology with high (z¿1) redshift galaxy surveys*. Ph.D. thesis, University of Texas at Austin.

Komatsu, E., and Spergel, D. N. 2001. Acoustic signatures in the primary microwave background bispectrum. Phys. Rev. D, **63**(6), 063002.

Matarrese, S., Verde, L., and Heavens, A. F. 1997. Large-scale bias in the Universe: bispectrum method. MNRAS, **290**(Oct.), 651–662.

Monin, A.S., and Yaglom, A.M. 1971. *Statistical Fluid Mechanics - vol 1: Mechanics of Turbulence*. MIT Press.

Panchev, S. 1971. *Random Functions and Turbulence*. Pergamon Press.

Papoulis, A., and Pillai, S.U. 2002. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill Higher Education. 4th. edition.

Peebles, P.J.E. 1980. *Large-Scale Structure of the Universe*. Princeton Series in Physics. Princeton University Press.

Robertson, H. P. 1940. The invariant theory of isotropic turbulence. *Proceedings of the Cambridge Philosophical Society*, **36**, 209.

Salopek, D. S., and Bond, J. R. 1990. Nonlinear evolution of long-wavelength metric fluctuations in inflationary models. Phys. Rev. D, **42**(Dec.), 3936–3962.

Scoccimarro, R. 2000. The Bispectrum: From Theory to Observations. ApJ, **544**(Dec.), 597–615.

Scoccimarro, R., Colombi, S., Fry, J. N., Frieman, J. A., Hivon, E., and Melott, A. 1998. Nonlinear Evolution of the Bispectrum of Cosmological Perturbations. ApJ, **496**(Mar.), 586.

Sefusatti, E., and Komatsu, E. 2007. Bispectrum of galaxies from high-redshift galaxy surveys: Primordial non-Gaussianity and nonlinear galaxy bias. Phys. Rev. D, **76**(8), 083004.

Verde, L., Wang, L., Heavens, A. F., and Kamionkowski, M. 2000. Large-scale structure, the cosmic microwave background and primordial non-Gaussianity. MNRAS, **313**(Mar.), 141–147.

Wiener, N. 1930. Generalized harmonic analysis. *Acta Mathematica*, **55**, 117–258.