where we enter the discussion on cosmological perturbations we need to review a few key concepts on statistics. The reason is that cosmology we do not predict the evolution of single systems but rather the statistical properties of a large set of them.

In our case the "set" we are after are basically density perturbations. In cosmology there is no way of predicting in which particular point the universe should be over dense or under dense from first principles, it rather we can predict what is the probability that at any point \( x_1 \) is over dense, that point \( x_2 \) is under dense, etc. This is particularly so as the source of density perturbations is thought to be quantum fluctuations from inflation, all we can make are statistical predictions. Therefore, we will think of the observable universe as being a particular realization out of a statistical ensemble of possibilities.

The way we model this is by using the concept of a random field - this is the natural generalization of a random variable, say \( r \), which has some probability distribution function (PDF) \( P(r) \). The statistical properties of \( r \) are fully determined by \( P(r) \). For example,

\[
\langle r \rangle = \int r P(r) \, dr : \text{average} \quad \left( \int P(r) \, dr = 1 \right)
\]

\[
\langle r^2 \rangle = \int r^2 P(r) \, dr : \text{second moment}
\]

\[
\langle r^2 \rangle = \langle r^2 \rangle - \langle r \rangle^2 : \text{variance}
\]

\[
\langle r^3 \rangle = \int r^3 P(r) \, dr : \text{third moment}
\]

\[
\sigma_r = \langle r^2 \rangle - \langle r \rangle^2
\]

Under sufficientlynice conditions, knowing all the moments allows one to reconstruct the PDF - we get to measure moments, as we’ll discuss...
Run from this we can say something about \( P(t) \), for example. We said, a random field is a generalization of random variable, in particular a random field \( \Phi(x) \) means that we have a random variable at each point in space \( x \), so we have an infinite number of random variables. We write any quantity that fluctuates in terms of its mean value plus fluctuations away from this mean:

\[
\Phi(x,t) = \langle \Phi(x,t) \rangle + \delta \Phi(x,t) = \overline{\Phi} + \delta \Phi(x,t)
\]

We have made the assumption that the mean depends only on time, it is so because under conditions we discuss below. First note that physically one defines a dimensionless fluctuation amplitude:

\[
\delta \Phi = \frac{\Phi - \overline{\Phi}}{\overline{\Phi}} = \frac{\delta \Phi}{\overline{\Phi}}
\]

And sometimes it is easier to discuss the statistical properties of \( \delta \Phi \), which has zero mean:

\[
\langle \delta \Phi \rangle = \langle \frac{\Phi - \overline{\Phi}}{\overline{\Phi}} \rangle = \frac{1}{\overline{\Phi}} \langle \Phi - \overline{\Phi} \rangle = \frac{1}{\overline{\Phi}} \langle \Delta \Phi \rangle = 0
\]

The picture is:

A few key points. First, in cosmology we are only capable of observing a single universe (unfortunately!), therefore, unlike usual experiments...
is is key assumption, as otherwise it would be hard to do much in
mology. Note that by definition we get that the average of \( \Phi \) is
a function of time \( \bar{\Phi}(t) \).

What constraints does ergodicity impose? When is it valid? Mathematically,ergodicity holds for Gaussian (we'll define this shortly) random fields with
continuous power spectra which are statistically homogeneous
we we have an infinite number of random variables in a random field, we
take averages of "different random variables", i.e. of products
the field at different points. Thus defines the correlation function

\[
\langle \delta q(r_1) \delta q(r_2) \rangle : \quad \text{two-point correlation function}
\]

\[
\langle \delta q(r_1) \delta q(r_2) \delta q(r_3) \rangle : \quad 3\text{-point correlation function}
\]

\[
\langle \delta q(r_1) \cdots \delta q(r_N) \rangle : \quad N\text{-point correlation function}
\]

in points coincide, they reduce to moments, i.e. \( \langle \delta q(r)^2 \rangle \), \( \langle \delta q(r)^3 \rangle \).

In our case we are thinking about physics, these correlation
others will go to zero as you take some point to so compared to the
thers, these correlation functions will vanish as some inverse
or of separation (more accurate statement on this later, when we
consider connected correlation functions).

If the fluctuations were independent from point to point, these
correlators would vanish (this is the case for Poisson fluctuations).

Again, we assume ergodicity so we will calculate this using spatial
averages, instead of the ensemble average which would be:

\[
\langle \delta_\phi(x_1) \cdots \delta_\phi(x_N) \rangle = \int \mathcal{P} \left[ \delta_\phi(x_1), \ldots, \delta_\phi(x_N) \right] \delta_\phi(x_1) \cdots \delta_\phi(x_N) \, dx_1 \cdots dx_N
\]

where PDF for \( N \) random variables,

\( \delta_\phi(x) \) \ldots \( \delta_\phi(x_N) \)

we turn ergodicity into a tool for calculating corr. function will be discussed

First, we need to introduce another pair of key properties.

Random Fields in cosmology are typically statistically homogenous
and isotropic. A natural extension of the ideas of homogeneity
and isotropy to correlation functions. It means that correlation
notions are invariant under translations (stat. homogeneity)
and rotations (stat. isotropy).

<br>\text{stat. homog.}\quad<br>\text{stat. homog.}\quad\textbf{stat. homog.}\quad<br>\text{stat. homog.}

that is, \( f \) is only a function of the relative

\[ \langle \delta_\phi(x_1) \cdots \delta_\phi(x_N) \rangle = \int f(\mathbf{X}_2, \mathbf{X}_3, \ldots, \mathbf{X}_N) \]

\( X_{ij} \in X_2 \cap X_4 \)

\( X_{ij} \in X_3 \cap X_4 \)

\[ \varepsilon = \mathbf{X}_N - \mathbf{X}_i \]

\[ \text{Separation between points} \]

Then, if we do a translation \( \mathbf{X}_i \rightarrow \mathbf{X}_i + \varepsilon \mathbf{X}_0 \quad \varepsilon = \mathbf{X}_N \)

although the fields themselves are not homogeneous: \( \delta_\phi(x) \neq \delta_\phi(x + \varepsilon \mathbf{X}_0) \),

that's because precisely they are fluctuating!
the correlation function is invariant.

**Isotropy:** Correlation functions are invariant under rotations!

This means that the physics that created and evolved these fluctuations is not distinguishable from another (statistical homogeneity) or particular direction (statistical isotropy).

Note this simplifies correlation functions in a dramatic way:

- **Two-point function in 3D:**
  \[ \langle \delta_\Phi(x_1) \delta_\Phi(x_2) \rangle = \frac{1}{V} \int d^3x \delta_\Phi(x_1 - x_2) \]
  a function of 6 variables
  in principle

- **Function:**
  \[ \langle \delta_\Phi(x_1) \delta_\Phi(x_2) \delta_\Phi(x_3) \rangle = \frac{1}{V} \int d^3x \delta_\Phi(x_1 - x_2) \delta_\Phi(x_1 - x_3) \]
  function of 9 variables.

We are ready to say how one would take advantage of ergodicity, statistical homogeneity, and isotropy in observations to measure correlation functions - for the two-point function:

\[ \delta_\Phi(r) = \int \frac{d^2\Omega}{4\pi} \int \frac{d^3p}{V} \delta_\Phi(p') \delta_\Phi(p' + p) \]

Add these from right to left: take a point \( p' \), evaluate \( \delta_\Phi \) there, then take another point at \( p' + p \) (where \( p \) has magnitude \( r \), the distance at which one wants to calculate the two-point function) evaluate \( \delta_\Phi \) multiply together to correlate, then move to another \( p' \) and keep going.
until you use all your $V$.  Note that so far we just used translation invariance.  We can further average over all different pieces of $\mathcal{F}$, that's the outer integral over Fourier space.

It is natural to deal with fluctuations in Fourier space rather than "real space" as we have been doing so far -- the reason for this is that small-amplitude fluctuations can be linearized and that each Fourier mode evolves independently. Statistically, translation invariance singles out Fourier modes as they are two-point-wise correlated, as we shall see.

The Fourier transform is defined as (be aware of different conventions!)

$$
A(k) = \int e^{i k \cdot x} A(x) \frac{d^3x}{(2\pi)^3}
$$

and the inverse: $A(x) = \int e^{i k \cdot x} A(k) \, d^3k$.

On the first of these we see that if $A(x)$ is real $\Rightarrow A(-k) = A^*(k)$, this means that $A(-k)$ has the same information as $A(k)$ (it is just the complex conjugate).

It is the same "Fourier mode".

Also note when dealing with inverses, it is useful to remember the delta rad function $\delta_D(x)$, which can be written as

$$
(2\pi)^3 \delta_D(x) = \int e^{i k \cdot x} \, d^3k
$$

Here $\int \delta_D(x) \, d^3x = 1$, $\delta(x) = 0$ unless $x = 0$, $\int \delta_D(x) F(x) \, d^3x = F(0)$

or more generally $\int F(x) \delta_D(x - x_0) \, d^3x = F(x_0)$.
when we decompose any random field \( \delta(x) \) in Fourier space:

\[
\widehat{\delta(x)} = \int e^{i k \cdot x} \, d^3 k \, \delta(k)
\]

and, similarly:

\[
\langle \delta_q(k) \delta_q(k') \rangle = \int \frac{d^3 k}{(2\pi)^3} \frac{d^3 k'}{(2\pi)^3} \langle \delta_q(k) \delta_q(k') \rangle = \int \frac{d^3 k}{(2\pi)^3} \frac{d^3 k'}{(2\pi)^3} \langle \delta_q(k) \delta_q(k') \rangle
\]

so outside average over PDF because they are not random variables,

\[
\langle \delta_q(k) \delta_q(k') \rangle = \delta_q(\mathbf{k} + \mathbf{k}') P_q(k)
\]

One thing to note is that this correlator, unlike in real space, is only nonzero \( k = -k' \), which means that Fourier modes are uncorrelated, (Recall \( \delta(k) \)) and can be understood just from translation invariance.

if we make a translation:

\[
\mathbf{r} \rightarrow \mathbf{r} + \mathbf{k} \quad \delta_q(k) \rightarrow \int e^{i k \cdot \mathbf{r}} \frac{d^3 k}{(2\pi)^3} \delta_q(\mathbf{k}) = e^{i k \cdot \mathbf{r}} \delta_q(k)
\]

\[
\Rightarrow \langle \delta_q(k) \delta_q(k') \rangle \rightarrow e^{i k \cdot \mathbf{r}} \langle \delta_q(\mathbf{k} + \mathbf{r}) \delta_q(\mathbf{k'}) \rangle
\]
order to have translation invariance, then need $h(x) = 0$, that is $\delta = 0$. So $\langle \delta \rangle$ appears. Note also $\mathcal{P}(h)$ depends only on $|k|^2$, due to isotropy.

Similarly, if you calculate 3-pt correlator in Fourier space,

$$\langle \delta_1 \delta_2 \delta_3 \rangle = \frac{S_0(k_1 + k_2 + k_3)}{B(k_1, k_2, k_3)}$$

again, translation invariance. Bispectrum: FT of 3-pt function by isotropy depends only on 3 vars.

So let's go back to basics for a second, and ask the question whether 4-pt correlations are independent. Let's start with a field $\phi$ with zero mean:

$$\langle \phi \rangle = 0$$

in principle this could be non-zero just because $\langle \phi \rangle$ and $\langle \phi \rangle$ are equal to the mean, i.e. there is no "true" correlation. For this reason we define the unconnected correlation $\langle \phi \rangle_c$ as:

$$\langle \phi_1 \phi_2 \rangle = \langle \phi_1 \rangle \langle \phi_2 \rangle + \langle \phi_1 \phi_2 \rangle_c$$

We see the splitting that makes sense physically. As point 1 gets far away

$$\langle \phi_1 \phi_2 \rangle_c \rightarrow 0$$

but $\langle \phi_1 \rangle \rightarrow 0$, so a

in-zero $\langle \phi_1 \phi_2 \rangle$ in this limit is not independent of $\langle \phi \rangle$ but $\langle \phi_1 \phi_2 \rangle \rightarrow 0$. It makes more sense to work with connected correlations.

we way to avoid this at the 2-pt level is to work with $\delta_\phi$ instead of $\phi$:

$$\langle \delta_\phi \rangle = 0$$

$$\langle \delta_\phi(1) \delta_\phi(2) \rangle = \langle \delta_\phi(1) \delta_\phi(2) \rangle_c$$

and some:

$$\langle \delta_\phi(1) \delta_\phi(2) \delta_\phi(3) \rangle = \langle \delta_\phi(1) \delta_\phi(2) \delta_\phi(3) \rangle_c$$

but:

$$\langle \delta_\phi(1) \delta_\phi(2) \delta_\phi(3) \delta_\phi(4) \rangle = \langle \delta_\phi(1) \delta_\phi(2) \rangle \langle \delta_\phi(3) \delta_\phi(4) \rangle + \langle \delta_\phi(1) \delta_\phi(2) \rangle_c \langle \delta_\phi(3) \delta_\phi(4) \rangle_c + \langle \delta_\phi(1) \delta_\phi(2) \rangle_c \langle \delta_\phi(3) \rangle_c \langle \delta_\phi(4) \rangle_c$$
which again defines the part of the 4-point function that is desired. Indeed, if we take, say, point 1 \( \rightarrow \infty \), \( \phi(1) \cdots \phi(n) \rangle \) will typically vanish, whereas \( \langle \phi(1) \cdots \phi(4) \rangle \approx \langle \phi(1) \phi(4) \rangle \phi(2) \phi(3) \%

Clearly connected correlations contain all the information.

**Gaussian Random Fields**

Gaussian Random Fields are the simplest example of random fields, and are defined by saying that all connected correlations of order larger than 2 are zero.

Therefore a Gaussian field is completely characterized by its power spectrum, 2-point function (and also the mean, but typically we take this to be zero). In a field at a single point a Gaussian PDF is

\[
P_{\phi}(\phi) = \frac{1}{\sqrt{2\pi \sigma^2}} \exp \left[ -\frac{1}{2} \frac{(\phi - \mu)^2}{\sigma^2} \right]
\]

where \( \sigma^2 = \langle \phi^2 \rangle \) is the only correlation function (or moment) in this case that enters, as it should for a Gaussian.

This generalizes trivially to \( N \) points, i.e. a multi-variate Gaussian is

\[
P_{\phi}(\phi_1, \cdots, \phi_N) = \frac{1}{2\pi \sqrt{\det \Sigma}} \exp \left[ -\frac{1}{2} \left( \phi - \mu \right)^T \Sigma^{-1} \left( \phi - \mu \right) \right]
\]

where \( \Sigma_{ij} = \langle \phi_i \phi_j \rangle \%

Now, in Fourier space things look even simpler, because the two-point correlator is diagonal, then

\[
P_{\phi} \left[ \Phi(k) \right] \approx \frac{1}{(2\pi)^4} \frac{1}{\sqrt{2\pi P(k)}} \cdot \frac{1}{\text{variance in Fourier space}} \cdot \text{variance in \( \Phi \)}}^2}{2 \langle \Phi(1) \rangle}
\]

product of independent Gaussians

\( \langle \Phi(1) \rangle \)
Typically power spectra are \( P(k) \sim k^n \)

is then known as the spectral index. If \( P(k) \) is more generic function

we can define a local, or effective, spectral index as

\[
\text{Neff}(k) = \frac{\text{d} \ln P(k)}{\text{d} \ln k}
\]

which reduces to \( n = \text{const.} \) if \( P(k) \sim k^n \).

\( P(k) \) measures the amplitude of Fourier coefficients @ \( k \), in 3D

are a lot of modes at a given \( k \), so a more physical measure

the **amplitude of fluctuations** is a dimensionless quantity:

\[
\Delta(k) = \sqrt{\frac{4\pi^2 k^3 P(k)}{\text{# of modes}}} \tag{v}
\]

in 3D

\( \Delta(k) \) has units of volume

\[
k^3 \sim \frac{1}{V}
\]

we can write the variance in real space in terms of an integral over the

over spectrum:

\[
\begin{align*}
\text{amplitude squared} : \quad & \sigma^2 = \langle \delta^2 \rangle = \int e^{-i k_1 \cdot \mathbf{x}} e^{i k_2 \cdot \mathbf{x}} \, d^3 k_1 \, d^3 k_2 \\
\text{fluctuations in} & \quad \text{Real space} \\
\Rightarrow \quad & \sigma^2 = \int d^3 k \, P(k) = \int \text{d} \ln k \, \frac{4\pi^2 k^3 P(k)}{\Delta(k)}
\end{align*}
\]

\( \Delta(k) \): amplitude squared

of all modes, in 3D