The Inhomogeneous Universe

The density contrast random field

First Principles

The density field of the inhomogeneous Universe is not constant everywhere, but it varies with spatial location.

(At first) the density values at different locations do not differ much from the mean density

→ they are **perturbations**.

It is usual to define the **density contrast δ (x)**:

the deviation with respect to the mean density (averaged over space)

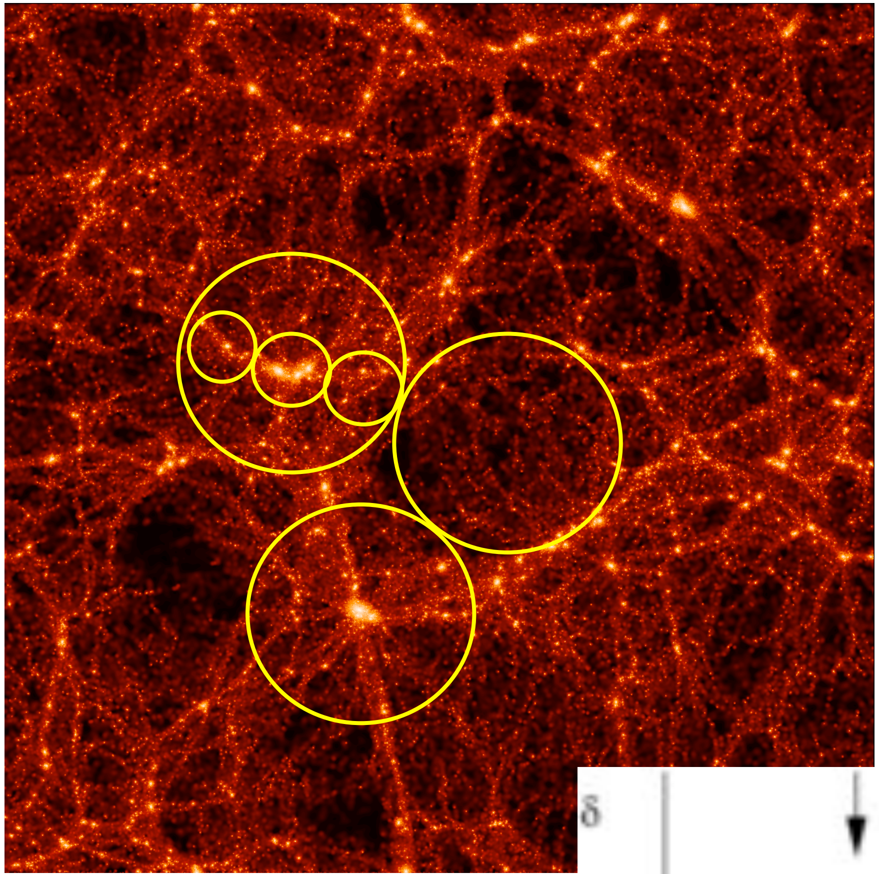
$$\delta(\vec{x}) = \frac{\rho(\vec{x})}{\bar{\rho}} - 1$$

During the **evolution of the Universe** (evolution of the mean density),
 the density contrast at each point also evolves, either increasing or decreasing, driven by gravity.

An increase of δ means clustering of matter → in practice a local region of the Universe expands slower than the global expansion.

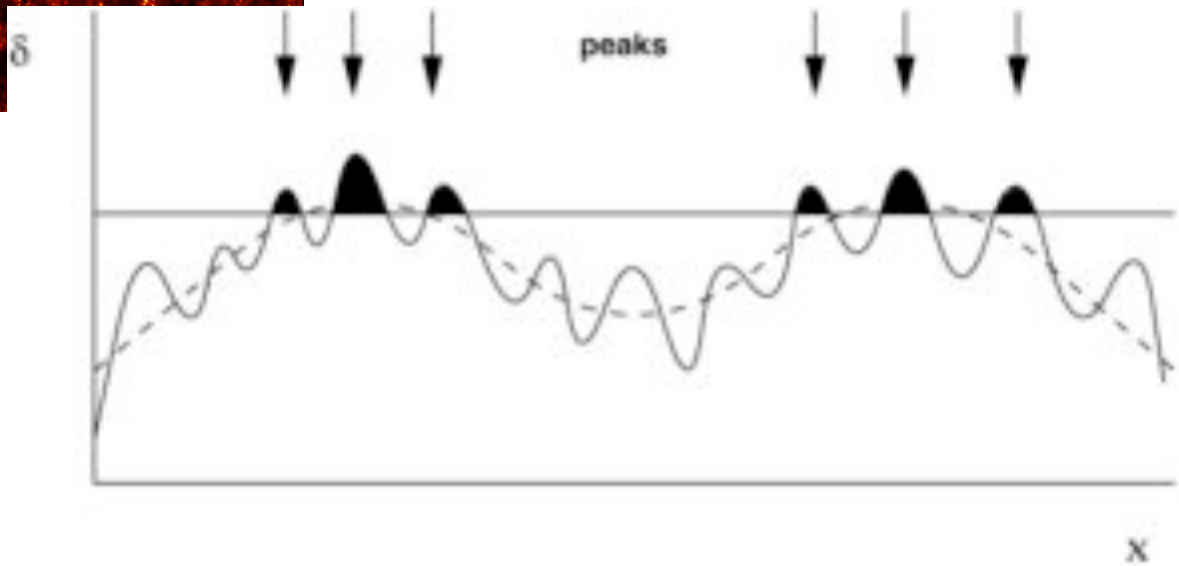
The process of evolution of the density contrast is called **structure formation**, turning density fluctuations in cosmological and astrophysical structures.

δ can become very large (not a density perturbation anymore) but the associated gravitational potential always remains a perturbation to the metric.



overdensities
and
undersdensities

on two different scales



How do initial fluctuations around the mean arise?

from quantum fluctuations of density.

In the quantum universe, there is a large number of random steps, i.e., in the very early Universe the value of density at a given location is changing all the time as the result of a stochastic (random) process. It is not possible to know the value of density at a given location at a given time, in a deterministic way.

We just know that the value is a realization of a probability distribution. Due to the large number of random processes involved, the central limit theorem tell us that the resulting probability distribution is a Gaussian

→ **the quantum density field is a Gaussian random field**.

Later, the **inflationary mechanism** makes the transition from quantum to macroscopic world

→ it produces a density field of macroscopic perturbations - called the primordial perturbations - this field is the **initial condition** for the subsequent time evolution of δ(x), but again its actual value is not known, it is a particular realization among all possible realizations.

Note that the depending on the inflationary model, the Gaussianity of the density random fields may or may not be preserved during inflation → search for possible primordial non-Gaussianity is a test of inflation.

(This is the goal of the measurements of the f_{NL} parameter in CMB observations)

Now, the value of density at a given location is then (most likely) a value taken from a Gaussian distribution.

So the actual values of δ(x) at each point are not known.

We just know that the density contrast at each point is a random variable, and its value is one among the various possible realizations of a Gaussian distribution,

$$P(\delta_1) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(\delta_1 - <\delta_1>)^2}{2\sigma_1^2}}$$

The density contrast random field is then described by the parameters of its Gaussian distribution.
As we know, a Gaussian distribution has only two parameters (its moments): mean and variance.

Note that there is one Gaussian distribution for each spatial location (hence the subscript in δ above). In principle each location may have its own mean and variance.

Consider a discretization of the density contrast field.

We need N distributions $P(\delta_i)$ (one for each position x; of course the problem is continuous N→infinity).

However, the N variables δ_1 ... δ_N are not independent

→ The value at a point depends on the values of neighboring points (due to the gravitational interactions between them).

So we cannot describe the system by considering N independent Gausian distributions, but we need an N-dimensional Gaussian:

$$f_{\mathbf{x}}(x_1,\ldots,x_k) = \frac{1}{\sqrt{(2\pi)^k|\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

(In our case the vector of k δ random fields on k locations is the random variable x with dimension k, and the k-dimension Gaussian distribution has a k-dim vector of means μ and a k x k covariance matrix Σ)

For example, if there were only 2 random variables (i.e., binning the density field such that it would have only two locations), we would need a 2-dimensional Gaussian:

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right]\right)$$

with $$\boldsymbol{\mu} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$$

where ρ is the correlation coefficient $\rho = \sigma_{xy} / (\sigma_x \sigma_y)$

Since the two random variables are not independent, the correlation coefficient is different from zero, and the covariance matrix is not diagonal.

The joint probability of having a value δ_1 at the location 1(called x in the notation above) and having at the same time a value δ_2 at the location 2 (called y in the notation above) is

$P(\delta_1, \delta_2) = P(\delta_1)\, P(\delta_2 | \delta_1)$ (where $P(\delta_2 | \delta_1)$ is the conditional probability)

It seems the complexity of the problem increases with the stochasticity!

If the problem was **deterministic**:

system described by the field $\delta(x)$ → N values

But the problem is **stochastic**:

system not described by the actual values of $\delta(x)$ but by the moments of the N-dim distribution (of which the values of δ are realizations).

The number of moments of an N-dimensional Gaussian is

→ N(N+1) (N values of mean, NxN values in the covariance matrix)

In case the correlations are symmetric, there are only N(N-1)/2 off-diagonal correlation coefficients → a total of N(N+1)/2 elements in the covariance matrix

→ a total of N(N+3)/2 moments.

So the N Gaussian random variables are described by N(N+3)/2 variables (the moments of the distribution).

Fortunately, the complexity is reduced by introducing the

Generalized cosmological principle:

"The universe is statistically homogenous and isotropic"

This means that there are perturbations to the homogeneity but they are described by a probability distribution with a homogeneous and isotropic set of moments.

→ The moments of the distribution do not depend on location or orientation.

(instead of the values of the density field themselves)

Statistical Homogeneity

implies that:

i) *The means do not depend on location* → **all N means are identical** (one for each random variable δ_i).

Can we measure the means of the distributions?

If we had a sample from the distribution, we could just measure its average in the usual way (summing the values and dividing by their number) - this is called the ensemble average. This statistic (the ensemble average) is known to give an unbiased estimate of the mean of a distribution (if the sample is large enough).

Problem: However we only have one realization - which is the Universe itself - instead of a full sample (unless there are parallel universes), i.e., we can only measure one value of δ in a given location, and we cannot repeat the experiment to get more values.

Solution: We assume that the whole Universe provides a representative set of all possibilities, i.e., the Universe includes in itself all possible realizations of the distribution.

In other words, distant parts of the field in separate parts of the Universe are independent of each other. The values of δ there are not correlated with the values of δ here. Those values are independent realizations of the same distribution that provides the values here (the distributions are the same due to statistical homogeneity).

In this way we can have access to different realizations of the same distribution, and get a sample

→ we can then make spatial averages instead of ensemble averages in order to find the moments. This is called the **ergodic hypothesis**.

Using the ergodic hypothesis, we can easily compute the mean of the distribution of δ.

From its definition, the values of δ are:

$$\delta(\vec{x}) = \frac{\rho(\vec{x})}{\bar{\rho}} - 1$$

the mean value of the distribution can then be computed by the ensemble (now equivalent to spatial) average of the values of δ across the spatial field.

The result follows immediately:

$$\langle\delta\rangle=0$$

(*Note: <> denotes ensemble or spatial averages*)

This means that the value of δ on any point of the Universe is a random value around the mean δ = 0.

This also implies that the amplitude of cosmological perturbations will not be given by the mean value of their distribution but by the variance of the distribution (a larger variance allows for the possibility of producing realizations with larger values of δ).

The N-dimensional distribution is then essentially described by the NxN covariance matrix. Its elements are:

Variance: i.e. the N terms of the diagonal (also called auto-correlation)

Covariances: i.e., the N(N-1) off-diagonal terms (also called the cross-correlations)

Statistical homogeneity further implies that:

ii) *The variances do not depend on location* → **all N terms of the diagonal are identical.**

Can we measure the variances of the distributions?

Yes, by measuring a sample of values of δ at different locations and computing the variance with the usual statistic:

$$\frac{1}{k-1}\sum_{i=1}^{k}\delta_i^2 = \langle \delta^2 \rangle$$

iii) *The correlation coefficients do not depend on location* → this does not mean that all N(N-1) terms of the off-diagonal are identical. It means that **the correlation coefficient between a pair of points separated by a given vector is the same for all pairs separated by identical vectors.**

Statistical Isotropy

implies that:

iv) *The correlation coefficients do not depend on orientation* → **the correlation coefficient between a pair of points separated by a given vector modulus (i.e. a given distance, irrespective of the orientation) is the same for all pairs separated by the same distance.**

Eg: $\sigma_{14} = \sigma_{37}$ (covariance between locations 1 and 4 and between locations 3 and 7)

Can we measure the variances of the distributions?

Yes, by measuring a sample of values of δ at different locations and computing the covariance using only pairs of points at the same separations:

$$\frac{1}{n_{\text{pairs}}}\sum_{i=1}^{n_i}\sum_{j=1}^{n_j}\delta_i\delta_j\delta_D(|i-j|-d) = \langle\delta_i\delta_j\rangle\,(d)$$

(the Dirac delta indicates the sum only includes points at a separation d from each other)

In summary, the density contrast random field (discretized in N positions of a regular grid) is described by N values:

- 1 variance (auto-correlation)
- N-1 covariances (since the condition iv reduces the original N(N-1) correlation coefficients to N-1)

Two-point functions

Correlation Function

The N-1 covariances form a function known as the **2-point correlation function** :

$$\xi_{\delta\delta}(r) = \langle \delta(x) \delta^*(x') \rangle$$

(r=|x -x' |)

(δ* accounts fot the possibility of having complex fields)

These N quantities contain the full cosmological information of a Gaussian δ(x) map.

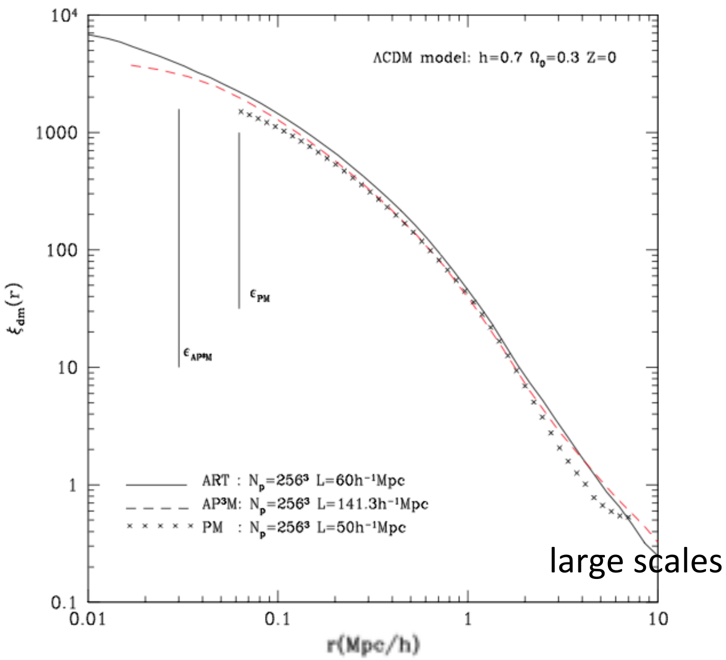
The randomness aspect and the generalized cosmological principle, make that ***the most natural spatial quantities to use in the treatment of the inhomogeneous Universe are not locations but separations between locations.***

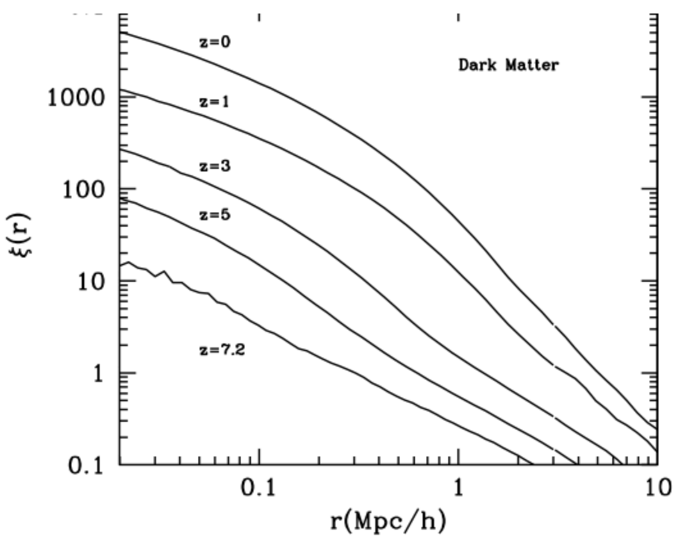
The correlation function of the density contrast field contains all the statistical information on the Gaussian density contrast field.

In particular it tells us the conditional probability of having a value δ_2 at a location "2" separated by "r" from a location "1" where there is a value δ_1 → it describes the **clustering** properties of the field

$$P(\delta_1, \delta_2) = P(\delta_1)\, P(\delta_2 | \delta_1)$$

The dark matter correlation function predicted by the ΛCDM model is **positive and decreases with separation. Its amplitude** naturally increases with structure formation (as the clustering of matter increases) → **it decreases with redshift**.





The δ field can be measured at N locations, for example, by measuring the positions x,y,z of N galaxies → **assuming the galaxies trace the locations of the overdensities**

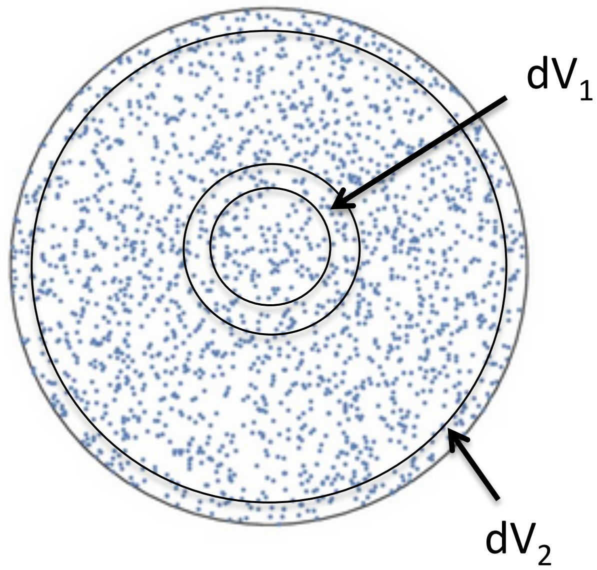
(In reality there is a **bias** between a galaxy location and a dark matter overdensity)

Let us consider N galaxies on a volume V, with a number density of n=N/V

(i) Case of an uncorrelated distribution

The probability of having a galaxy in the shell volume dV_1 is given by the number of galaxies within that volume divided by the total number of galaxies N:
$dP_1 = n\ dV_1 / N = dV_1 / V$

The probability of having a galaxy in the shell volume dV_2 is independent of dP_1 :
$dP_{2u} = n\ dV_2 / N = dV_2 / V$



Case of uncorrelated distribution

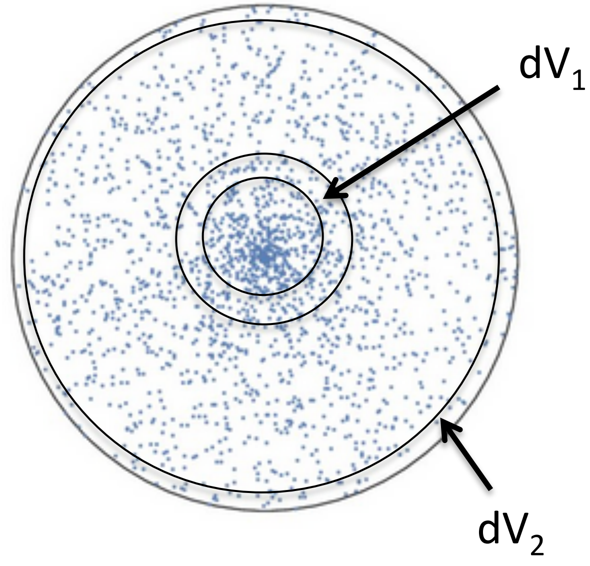
(ii) Case of a correlated distribution

The probability of having a galaxy in the shell volume dV_2 depends on dP_1 .

In other words, the value of dP_2 depends on the correlation between the locations 1 and 2,

i.e., it depends on the correlation at the separation r_{12} :

$$dP_{2c} = n\, dV_2 \,(\,1+\xi(r_{12})\,)\,/\,N = dV_2\,(1+\xi(r_{12})\,)\,/\,V$$



Case of correlated distribution

So, the number of galaxies found is no longer just a function of the size of dV_2

If there is a:

correlation, $\xi > 0 \rightarrow dP_{2c} > dP_{2u}$
(anti-)correlation, $\xi < 0 \rightarrow dP_{2c} < dP_{2u}$

The **number of galaxies as function of r**, on the full volume, is given by N times the integral of the probability dP(r).

In the uncorrelated case, the conditional probability is 1 and N(r) is just

$$N(r) = \int n \, dV = n \int dV/dr \, dr \sim r^3$$

but in the correlated case, $N(r) = n \int (1+ \xi(r)) \, dV/dr \, dr$

(Note that n dV(r) is a "distance function", the number of objects per distance bin dN (r) - the use of shell volumes dV is very practical to obtain a function of r **)**

So in general the slope will be different from r^3, depending on the correlation function slope $\xi(r)$ → the number is higher on a highly correlated area (usually on small separations).

The correlation can then be equivalently defined as the excess N(r) between the clustered and the random cases:

If we compare the probabilities dP(r) we just found for the correlated and the uncorrelated cases,

$dP_{2u} = n\, dV_2 / N$

$dP_{2c} = n\, dV_2 (1+\xi(r)) / N$

we see that **1+ξ(r) is given by the ratio of the probabilities**, i.e., by the ratio of the two “distance functions” (the number of galaxies as function of r):

$$\mathbf{1+\xi(r) = N_c (r) / N_u (r)}$$

Correlation Function in Fourier space

The correlation coefficient of 2 points separated by r tells us about **structure** - the central property of the inhomogeneous universe that we want to describe. It quantifies the clustering of the density field (**the "degree of collapse"**) - the **formation of structure**.

For example, if there is correlation on all separations up to a separation r and then the correlation drops, it shows that (on average) there are overdensity regions from x to (x+r), i.e. halos of size r

However the relation between correlation as function of separation, and size of the overdensity is not a one-to-one relation → from this example, we see that we need to know the correlation at various separations to find out if there is an overdensity of a given size.

We would like to have a function that directly shows the clustering amplitude on a given size. Is this possible?

Let us consider the Fourier transform of the density contrast field

$$\delta_k = \frac{1}{V}\int \delta(x)\, e^{-ik.x}\, d^3x \qquad\qquad \delta(x) = \frac{V}{(2\pi)^3}\int \delta_k\, e^{+ik.x}\, d^3k$$

This defines a set of Fourier modes k (3d vectors), with associated sizes 2π/k (or wave numbers)

Convention:

- we are writing the plane waves as ikx and not i2πkx → this makes a factor $(2\pi/k)^3$ to appear

- **the integrals are normalised by the volume V**, which ensures that δ_k is dimensionless if δ(x) is also dimensionless

Let us compute the 2-point correlation function in k-space :

$$\langle \delta_k \delta^*_{k'} \rangle = \frac{1}{V} \left\langle \int d^3x \; \delta(x) \, e^{i\vec{k}.\vec{x}} \frac{1}{V} \int d^3x' \; \delta^*(x') \, e^{-i\vec{k}'.\vec{x}'} \right\rangle$$

The ergodic hypothesis allows us to put the brackets inside the integrals

Inserting the definition of the correlation function, we can write:

$$= \frac{1}{V} \int d^3x \; e^{i\vec{k}.\vec{x}} \frac{1}{V} \int d^3x' \; e^{-i\vec{k}'.(\vec{x}+\vec{y})} \, \xi(|\vec{y}|) =$$

where y is the separation vector between x and x’,

for fixed x the integration over x’ is the same as an integration over y.

So we are left with an integral in x with no function with dependence on x (except the plane waves),

and an integral in y that that is a (normalised) Fourier transform of the correlation function:

$$= \frac{1}{V}\int d^3x\, e^{i\vec{x}.(\vec{k}-\vec{k}')} \frac{1}{V}\int d^3y\, \xi(|\vec{y}|)\, e^{-i\vec{k}'.\vec{y}}$$

The first integral is the (dimensionless) Dirac delta.

Recall the Dirac delta is the (standard) Fourier transform of f(x)=1:

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

The second integral is the (normalised) Fourier transform of the correlation function, which is called the **dimensionless power spectrum:**

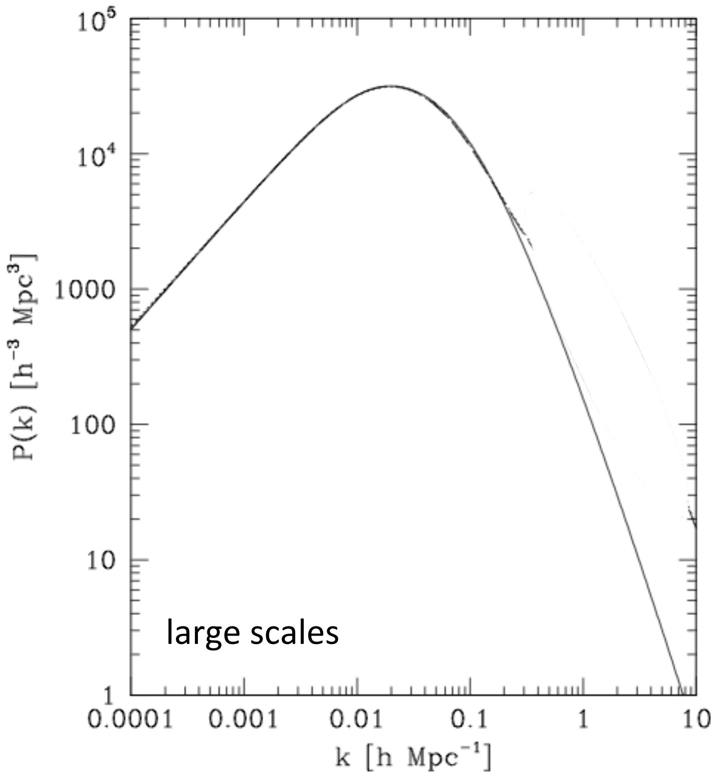
$$P_\delta(|k|) / V$$

Note that due to isotropy it only depends on the modulus of the k-mode vector.

The **power spectrum** of a random field is defined as the (standard) Fourier transform of the correlation function of the same field,

$$\xi(r) = \frac{1}{(2\pi)^3} \int P(k) e^{-ik.r} d^3k$$

(and reciprocally, the correlation function is the Fourier transform of the power spectrum)



(the ΛCDM power spectrum of the density contrast field looks like this)

So the result is $$\frac{(2\pi)^3}{V}\,\delta_D(\vec{k}-\vec{k}')P_\delta(|\vec{k}|)$$

where δ_D here is the dimensionless Dirac delta

$$\boxed{\langle \delta_k \delta^*_{k'} \rangle \;\; = \langle \delta^2_k \rangle = \frac{(2\pi)^3}{(2\pi/k)^3}\,\delta_D(\vec{k}-\vec{k}')P_\delta(|\vec{k}|) = k^3 P_\delta(k) \;= \Delta^2\,(k)}$$

where we used the fact that the length associated to a Fourier mode k is $2\pi/k$, and so the corresponding volume is $V = (2\pi/k)^3$

Notice that the power spectrum P(k) has dimensions of volume [$(\mathrm{Mpc}/h)^3$]

and $\Delta^2\,(k) = k^3\,P(k)$ is the **dimensionless power spectrum**,

also known as the power spectrum per interval of ln(k).

The important result we obtained here is that

the correlation function of the density contrast field in Fourier space is the (standard) Fourier transform of the correlation function multiplied by the Fourier volume k^3 and by a dimensionless Dirac delta function, i.e.,

it is the dimensionless power spectrum multiplied by a Dirac delta function

The presence of the Dirac delta makes the coefficients δ_k to be independent, and

the elements of the correlation function in Fourier space are independent, as are the elements of the power spectrum

It is also useful to compute the **auto-correlation function** of the density contrast field, i.e. the **variance**:

$$\sigma^2 = \langle \delta(x)\delta^*(x') \rangle = \left\langle \delta^2(x) \right\rangle$$ where x=x'

$$\sigma^2 = \frac{V}{(2\pi)^3}\frac{V}{(2\pi)^3}\int d^3k\delta_k e^{ikx}\int d^3k'\delta_{k'}e^{-ik'x}$$
$$= \frac{V}{(2\pi)^3}\frac{V}{(2\pi)^3}\int d^3k\int d^3k' \left\langle \delta_k\delta^*_{k'} \right\rangle e^{i(k-k')x}$$

Inserting the result for $\langle \delta_k\delta^*_{k'} \rangle$

$$\sigma^2 = \frac{V}{(2\pi)^3}\frac{V}{(2\pi)^3}\int d^3k\int d^3k' e^{i(k-k')x}k^3P_\delta(k)\frac{\delta_D(k-k')(2\pi)^3}{V}$$

one of the integrals is just the Fourier transform of the Dirac delta, which is 1 (and also cancels with one of the volumes);

k^3 cancels with the other volume

and we are left with:

$$\sigma^2 = \int \frac{d^3k}{(2\pi)^3} P(k)$$

So the variance of the delta field (in real space) is a 3d integral of the power spectrum. Since the power spectrum is isotropic we can integrate the angular part of

$$d^3k = k^2 \sin\theta \, dk \, d\theta \, d\phi$$

which is 4π

resulting in:

$$\sigma^2 = \int_0^\infty \frac{dk}{k} \frac{k^3 P(k)}{2\pi^2}$$

Writing k^2 as k^3/k shows explicitly that:

to integrate $k^2 P(k)$ on the linear domain dk
is equivalent to integrate the dimensionless power spectrum in the logarithmic domain dk/k

This is the reason why the dimensionless power spectrum is known as the power spectrum per interval of ln(k).

This result tells us that the variance of the density contrast field (which is its main property) has contributions from all scales of the power spectrum. Each logarithmic bin contributes with a certain value (the value of the dimensionless power spectrum of that scale)

and so **the amplitude of the dimensionless power spectrum is a direct indication of the amplitude of clustering**

$\Delta < 1$ - weak clustering, linear structure

$\Delta > 1$ - strong clustering, non-linear structure : large over-densities, or large under-densities (voids)

Power spectrum vs. Correlation function

Both descriptions - in real and Fourier space - have the same information.
Both are valid to describe the cosmological field.

The fact that the dimensionless power spectrum contains variances instead of covariances, means that it gives directly the information of a mode - or ***scale*** *- (instead of relying on separation between points).*

Note that A small value of k is called a large scale
A large value of k is called a small scale

because the inverse of the scale - $2\pi/k$ - corresponds to a physical size

So the value of the dimensionless power spectrum on a given Fourier mode, is the variance on that scale, i.e., the degree of clustering (the clustering amplitude) that exists on that scale of the Universe on average.

(Remember it is a moment of a distribution, so it does not mean that all regions of the Universe of that size will have that density contrast, it only means that their values will be realizations of that distribution with that variance).

(Recall that the dispersion of a random variable of mean zero is a direct indication of its amplitude - and not the mean! -)

Let us consider now the power spectrum as the basic quantity and compute the correlation function from it:

We need to compute the inverse Fourier transform of the power spectrum:

$$\xi(r) = \frac{1}{(2\pi)^3} \int P(k) e^{-ik.r} d^3k$$

The correlation function is real so we just need to consider:

$$\mathrm{Re}(e^{-ikr\cos\theta}) = \cos(kr\cos\theta)$$

and the power spectrum is isotropic (it depends only on the radius |k| → we can integrate over the angular part:

$$\int_0^\pi \cos(kr\cos\theta)\sin\theta\, d\theta = -\frac{\sin(kr\cos\theta)}{kr}\Big|_0^\pi = 2\frac{\sin kr}{kr}$$

(in spherical coordinates the integral element is $d^3k = k^2 \sin\theta\, dk\, d\theta\, d\phi$

The result is: $$\xi(r) = \frac{1}{2\pi^2}\int_0^\infty P(k)\,\frac{\sin kr}{kr}\,k^2\,dk$$

This means that the correlation function is a ***filtered linear combination of the power spectrum*** → one separation r is a combination of various scales k → **k are the independent and fundamental cosmological scales, the separations r are not independent.**

There is not a one-to-one correspondence between separation and scale (unless the filter in the integral, also called window function, is very narrow).

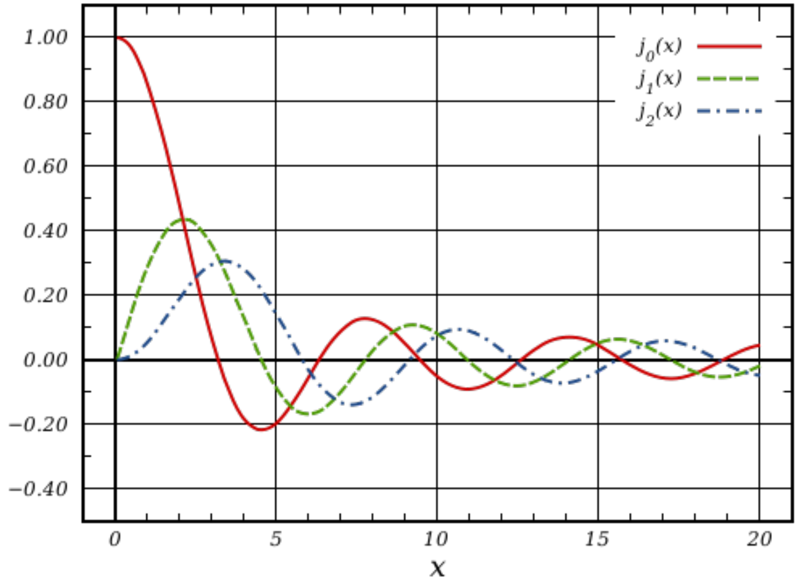
The filter (the function that multiplies k^2 P(k) in the integral) is the **spherical Bessel function** of the first kind for n=0 : j_0 (kr)

$$j_n(x) = (-x)^n \left(\frac{1}{x}\frac{d}{dx}\right)^n \frac{\sin(x)}{x}$$

$$j_0(kr) = \frac{\sin kr}{kr}$$

$$j_1(kr) = \frac{\sin kr}{(kr)^2} - \frac{\cos kr}{kr}$$

$$j_2(kr) = \frac{3\sin kr}{(kr)^3} - \frac{3\cos kr}{(kr)^2} - \frac{\sin kr}{kr}$$



The shape of j_0 (the solid line) shows that most of the contribution for the correlation at a separation r -ξ (r)- comes from larger scales: $k < 2.6/r$ (the range where the contribution is large, with filter amplitude > ~0.2)

In summary: power spectrum and correlation function have the same information, but the N components of the power spectrum are independent and give directly the amplitude of clustering as function of scale, while the N components of the correlation function do not.

So, while the original correlation function describes the density contrast field using a set of N-1 non-independent covariance (cross-correlations) variables (plus one variance) that depend on separation on the real space,

the power spectrum describes the same field using a set of N independent variance (auto-correlations) variables in the harmonic space: the set of $\langle \delta_k^2 \rangle$

Even though the 2-pt correlation function is highly correlated and does not give direct information on an individual scale, it is a useful quantity to consider because

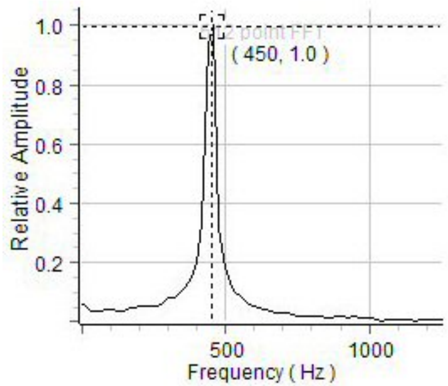
it is defined in real space → it can be **measured directly** from data measured in the sky.

(The power spectrum needs to be **estimated** from data in an indirect procedure).

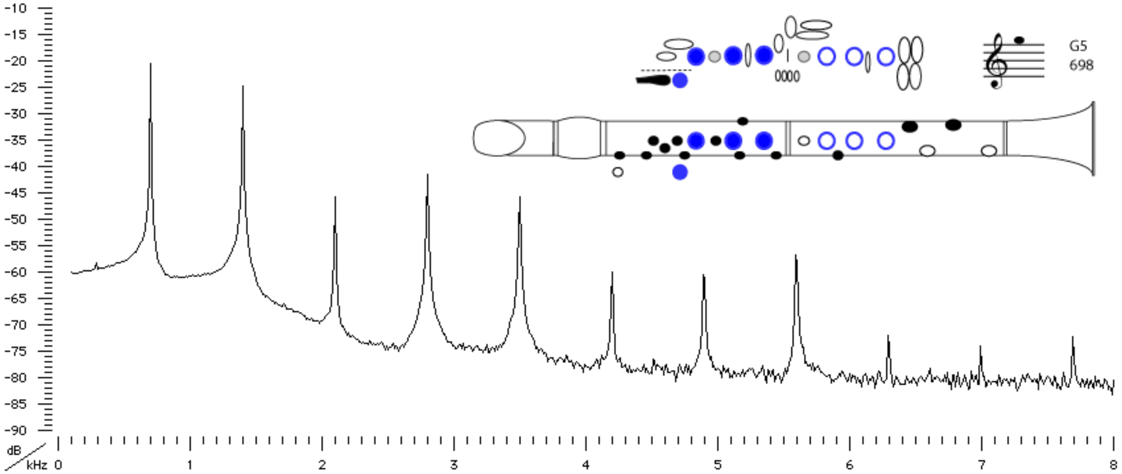
The fundamental modes in the harmonic space (i.e., the wavenumber or scale k) are thus the natural choice to define the cosmological scales.

Note: an analogy, is that the notes (A,B,C,D,...) produced by a musical instrument are not independent (they are like the separations), each one contain various fundamental notes defined by a tuning fork (which are the fundamental ones, like the cosmological scales).

Each instrument has a sound spectrum, which in fact is a Power spectrum:

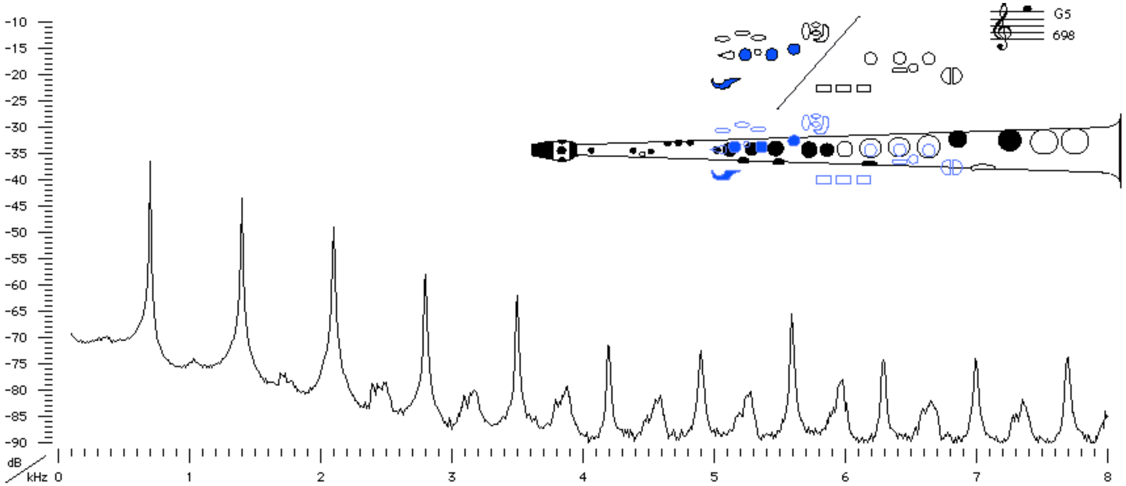


G tuning fork is independent from the other fundamental notes (the “scales”)



G clarinet is a linear combination of the fundamental notes

G saxophone is a different linear combination of the fundamental notes



(The sound power spectrum defines the timbre of the instrument)

Smooth spatial distribution: counts in cells and sigma_8

Alternatively to using discrete quantities (i.e. separations r between discrete locations x, x'), the clustering properties in the real space can be determined using a **smoother measure of density**:

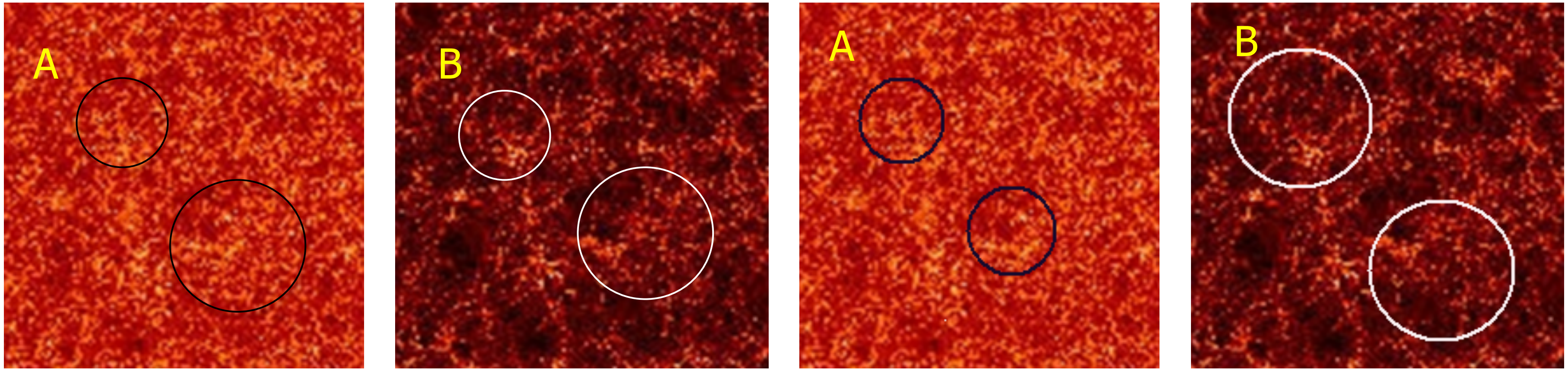
the variance of number counts in cells

Placing cells of a fixed size R on a δ map (discrete or continuous) allows to smooth the map on a scale R, defining a δ_R as a convolution of δ(x) with a window function (a filter) of size R → **δ_R is a weighted average of δ in a cell of size R.**

We can then compute the variance of this δ_R on cells R across the whole map.

Doing this for N values of R, we can define a vector of variances of δ_R.

Example: Consider two density maps A and B and two different scales R (shown by the circles).



Compute δ_R in each map for the two different values of R, obtaining 4 quantities.

Then compute the variance of each of those quantities, by moving the circles on the maps. The result is:

i) The variances in B are larger than in A (for both scales R), because B has more density contrast than A. In B the circles can fall in high-density regions or in low-density regions → large variance. While in A all regions are more similar → **B has more structure than A.**

ii) Placing the larger circle (for both A and B) it is more likely to find similar regions along the maps than with the smaller circle → the variance decreases with R → the smallest cell R to approach zero variance defines the homogeneity scale → there is no structure above that scale.

Now, since the variance of δ_R is a second-order moment, it is certainly related to the power spectrum.

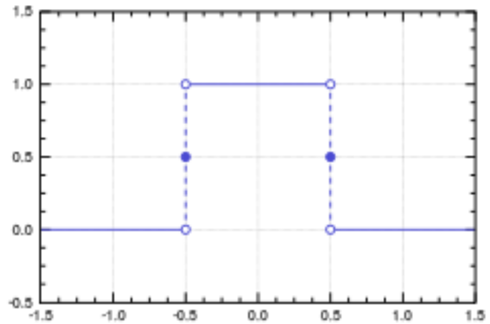
Let us derive that relation.

First, how can we write a theoretical expression for the smooth density δ_R ?

Let us consider a top-hat window function W_R, i.e., a filter of constant amplitude.

δ_R can be written as the convolution of δ with the top-hat:

$$\delta_R(x) = \int d^3y\, \delta(y)\, W_R(|x-y|)$$



The Fourier transform of the smooth field is simply the product of the Fourier transforms of δ and the top-hat:

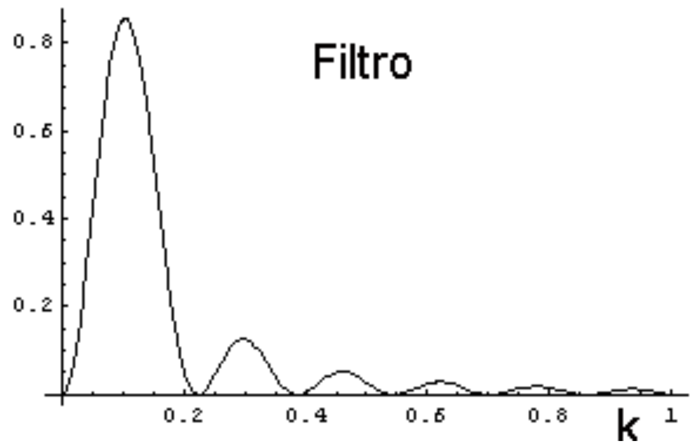
$$\delta_R(k) = \delta(k)\, W_R(k)$$

The variance of the smooth density is then,

$$\sigma_R^2 = \left\langle \delta^2(k)\, W_R^2(k) \right\rangle = \frac{1}{(2\pi)^3} \int d^3k\, W_R^2(k) P(k)$$

i.e., it is a filtered integral of the power spectrum, where the filter is the square of the Fourier transform of the top-hat $W_R(k)$:

$$W_R(k) = 3\frac{\sin kR - kR\cos kR}{(kR)^3}$$



This filter is very diferent from the j_0 Bessel function.
It is relatively narrow and peaked at k ~ 2π/R.

We conclude that a vector of σ^2_R (for various cell sizes R) is a linear combination of the power spectrum amplitudes, just like the correlation function was.

However, its components are less correlated than the correlation function ones → since the filter is very peaked, there is roughly a one-to-one correspondence between R and scale k.

For this reason, the value of σ^2_R gives a good indication of the clustering amplitude at the scale R (like the power spectrum also does).

As we will see later, to compute structure formation (i.e., the time evolution of the density contrast field), we need an initial condition for the density contrast field $\delta(x,t)$.

As we know, the field is fully represented by a 2-pt quantity. So the initial condition must be the value of a 2-pt function at a fixed time (redshift). In particular, the amplitude of an initial 2-pt function at a given scale is a **comological parameter of the inhomogeneous Universe.**

There are two alternative parameters that set the primordial amplitude of the density contrast field:

- The amplitude of the primordial power spectrum at a large scale k = 0.02 h/Mpc→ parameter A_s

- The amplitude of today's power spectrum (z=0) at a smaller scale R = 8 Mpc/h → parameter σ_8 ("sigma eight")

From early times to late times, the power spectrum evolves in amplitude and shape → the two amplitude parameters are related; the relation between the values of A_s and σ_8 depends on all cosmological parameters.

- *Why is a large scale [k=0.02 h/Mpc → R ~ 300 Mpc/h] used for early-times normalization?*

The scale factor is small → there is no resolution to access small scales

- *Why is R=8 Mpc/h used for late-time normalization?*

It is the scale where the observed dark matter power spectrum P(k,z=0) has amplitude ~1 → It is the threshold that separates linear scales (the larger ones) from non-linear scales (the smaller ones) today → so the value of σ_8 in a given model shows immediately the level of clustering in the universe today, compared with a σ_8 = 1 reference universe.