The Large Scale Structure of the Universe 2: Theory and Measurement

Eusebio Sánchez Álvaro
CIEMAT

UniCamp Winter School on Observational Cosmology
Sao Paulo
July 2018
The LSS: Theory and Measurement

Gravitational collapse
Matter Perturbations in the Universe and their growth
The CMB: Brief description and results
Non-linear growth: Spherical collapse
Simulations

How to measure LSS,
Correlation function, power spectrum
Galaxy bias
Observational effects. Random samples and other effects
We have treated the Universe as smooth. But it contains structures! Galaxies are not randomly distributed, but clustered.

We are able to explain how this structure formed and evolved.
The properties of the initial fluctuations determine the properties of the LSS

Important point: Inflaton is a quantum field → We cannot predict the specific value of the fluctuations, but only their statistical properties → Our predictions for the LSS are statistical
Summary of the formation and evolution of structure in the Universe

- **Quantum Fluctuations during inflation**
  - $10^{-35}$ s
  - $V(\phi)$

- **Perturbation Growth: Pressure vs. Gravity**
  - $\Omega_M$, $\Omega_r$, $\Omega_b$, $f_\nu$
  - $\approx 10^5$ years

- **Photons freestream: Inhomogeneities turn into anisotropies**
  - $z_{\text{reion}}$, $\Omega_\Lambda$, $w$

- **Matter perturbations grow into non-linear structures observed today**
Fluctuations are small. We can use perturbation theory.

\[ g_{\mu\nu}(\eta, \mathbf{x}) = \bar{g}_{\mu\nu}(\eta) + \delta g_{\mu\nu}(\eta, \mathbf{x}) \]
\[ T_{\mu\nu}(\eta, \mathbf{x}) = \bar{T}_{\mu\nu}(\eta) + \delta T_{\mu\nu}(\eta, \mathbf{x}) \]

\[ \Rightarrow \Phi, \delta \rho_m, \delta \rho_r \]

2 types of perturbations: metric perturbations, density perturbations

Remember: Spacetime tells matter how to move, matter tells spacetime how to curve.
**Matter perturbations**

Use newtonian gravity → Good approximation of general relativity in cosmology on scales well inside the Hubble radius and when describing non-relativistic matter (for which the pressure $P$ is much less than the energy density $\rho$).

\[
\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0
\]

\[
\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = - \frac{1}{\rho} \nabla p - \nabla \phi
\]

\[
\nabla^2 \phi = 4\pi G \rho.
\]

These equations are used in all cosmological N-body simulations of the non-linear growth of structure.

3 regimes:

- $\delta \ll 1$: **linear theory**
- $\delta \sim 1$: need specific assumptions (i.e. spherical symmetry)
- $\delta \gg 1$: **non-linear regime**. Solve numerically, simulations (also higher order perturbations)

In general: Universe is lumpy on small scales and smoother on large scales – consider inhomogeneities as a perturbation to the homogeneous solution.
Using the Fourier transform, we can write eqs. For the Fourier modes:

\[ \delta (\mathbf{x}, t) = \sum_k \delta_k (t) e^{i \mathbf{k} \cdot \mathbf{x}} \]

\[ \delta_k (t) = \frac{1}{V} \int \delta (\mathbf{x}, t) e^{-i \mathbf{k} \cdot \mathbf{x}} \, d^3 x \]

\[ \ddot{\delta}_k + 2H \dot{\delta}_k = \left( 4\pi G \rho_0 (t) - \frac{k^2 c_s^2}{a^2} \right) \delta_k. \]

\[ \ddot{\delta}_k + 2H \dot{\delta}_k - 4\pi G \rho_m (t) \delta_k = 0 \]

For baryonic matter

For dark matter
Matter perturbations

We can linearize this equation because $\delta$ is very small. The linear regime is very important:

- On all scales, primordial fluctuations were extremely small, $\delta \ll 1$. The seeds of structure formation were linear.
- The linear stage of structure formation is a relatively long lasting one.
- One may always find large scales where the density and velocity perturbations are still linear. Today, scales larger than $\sim 10 \, h^{-1} \, \text{Mpc}$ behave linearly.
- CMB measurements have established the linear density fluctuations at the recombination era. By studying the linear structure growth, we are able to translate these into the amplitude of fluctuations at the current epoch, and compare these predictions against the measured LSS in the Galaxy distribution.

$$\delta(x, t) = \frac{\rho(x, t) - \rho_0(t)}{\rho_0(t)}$$
Matter perturbations

in a matter dominated Universe

\[ H = \frac{2}{3t} \]

\[ \ddot{\delta} + \frac{4}{3t} \dot{\delta} - \frac{2}{3t^2} \delta = 0 \]

\[ \delta = A(x) t^{2/3} + B(x) t^{-1} \]

linear growth

\[ \delta = \delta_0(x) a \]
Matter perturbations

in a lambda dominated Universe

\[ H^2 = \frac{\Lambda c^2 \pi G}{3} \]

\[ \ddot{\delta} + 2H \dot{\delta} = 0 \]

\[ \delta = A(x) + B(x) e^{-2Ht} \]

frozen fluctuations
In a radiation dominated universe

\[
\ddot{\delta}_k + 2H \dot{\delta}_k - 4\pi G \bar{\rho}_m \delta_k = 0
\]

\[
a \propto t^{1/2} \quad \Rightarrow \quad H = \frac{1}{2t} = \sqrt{\frac{8\pi G}{3\bar{\rho}}} \]

\[
H^{-2} \ddot{\delta}_k + 2H^{-1} \dot{\delta}_k = \frac{3}{2} \frac{\bar{\rho}_m}{\bar{\rho}} \delta_k \quad \bar{\rho}_m \ll \bar{\rho}
\]

At most logarithmic growth during radiation domination.

1) The increased expansion rate due to the presence of a smooth component slows down the growth of perturbations.
2) There is no significant growth during the radiation dominated period.

\[
\delta_k(t) = A + B \ln t
\]
Matter perturbations

- Matter dominated case
  \[ \delta = (A(x) t^{2/3}) + B(x) t^{-1} \] Linear growth

- Radiation dominated case
  \[ \delta_k(t) = A + B \ln t \] No significant growth

- Lambda dominated case
  \[ \delta = A(x) + B(x) e^{-2Ht} \] Frozen fluctuations

- General case
  \[ \delta = \delta_0(x) a g(a, \Omega_{m0}) \]

  g is constant at early times and scales as 1/a at late times

  for our cosmology, the action ended around z=0.5
Matter perturbations

Baryon photon fluid
Jeans length and scales for collapse

\[
\ddot{\delta}_k + 2H \dot{\delta}_k = \left( 4\pi G \rho_0 (t) - \frac{k^2 c_s^2}{a^2} \right) \delta_k.
\]

GRavity \quad PRESSure

Jeans Length: Both effects are equal

\[
\frac{c_s^2 k^2}{a^2} > 4\pi \rho_0 \quad \rightarrow \quad \text{Oscillating solution}
\]

\[
\frac{c_s^2 k^2}{a^2} < 4\pi \rho_0 \quad \rightarrow \quad \text{Perturbations grow}
\]

\[
k_J = \frac{a}{a_0} \frac{\sqrt{4\pi G \rho}}{c_s}
\]

Jeans wavenumber

\[
\lambda_J = \frac{2\pi}{k_J}
\]

Jeans wavelength

the perturbations grow exponentially (if no expansion) with time or oscillate as sound waves depending on whether their wave number is greater than or less than the Jeans wave number.

For \( k > k_J \) we have sound waves, for \( k < k_J \) we have collapse. The expansion adds a sort of friction term on the left-hand side: The expansion of the universe slows the growth of perturbations down.
CMB shows that at $z \sim 1100$, perturbations are of the order $10^{-5}$. If they grow as $\delta \sim t^{2/3}$, then for $z=0$ they grow a factor of 1000, becoming of the order 1%.

Dark matter provides a solution. Perturbations in dark matter do not couple to radiation and can be much larger than gas (baryons) perturbations without perturbing the CMB. By starting with much larger perturbations, we can reach the $\delta \sim 1$ regime much earlier, allowing to form the observed structures $\rightarrow \textbf{LSS formation NEEDS dark matter!}$

It is useful to express the perturbation as $\delta(z) = \delta_0 D_+(z)$

linear growth factor, $D_+(z)$

$$D_+(z) = \frac{1}{1 + z} \frac{5\Omega_m}{2} \int_0^1 \frac{da}{a^3 H(a)^3}$$
Matter perturbations: Comparing the theory to the observations

The current standard model of cosmology includes the inflation as primordial perturbations generator. In any case, the initial perturbations are Gaussian.

The density contrast $\delta$ is a homogeneous, isotropic Gaussian random field (Fourier modes are uncorrelated)

Its statistical properties are completely determined by 2 numbers: mean and variance.

*The variance is described in terms of a function called the **POWER SPECTRUM***

$$\left\langle \hat{\delta}(\vec{k}) \hat{\delta}^*(\vec{k}') \right\rangle \equiv (2\pi)^3 P(k) \delta_D \left( \vec{k} - \vec{k}' \right)$$

The initial power spectrum has the Harrison-Zel’dovich form:

$$P(k) \propto k^{n_S}, \quad n_S \sim 1$$

Spectral index
Matter perturbations

These assumptions have been very precisely verified in the CMB. Spectral index can be related to some parameters of the inflationary field.

The power spectrum is splitted into a linear and a non-linear part

\[ P(k) = P_L(k) + P_{NL}(k) \]

The linear power spectrum corresponds to the linear overdensity field and is given by

\[ P_L(t, k) = A_0 \, k^{n_S} T^2(k) D_+(t) \]

Where \( D_+(t) \) is the growth factor and \( T(k) \) is the transfer function and takes into account the transformation from the density fluctuations from the primordial spectrum.

Through the radiation domination epoch
Through the recombination epoch
To the post recombination power spectrum

And contains the messy physics of the evolution of density perturbations.

It is computed by solving the Boltzmann equation for the primordial multicomponent cosmic plasma numerically (for example, using CAMB (Lewis & Challinor 2011)).
Transfer function and power spectrum for several models:

CDM: Cold dark matter
HDM: Hot dark matter
MDM: mixed dark matter (cold+hot)
Matter perturbations

The full power spectrum shape

\[ P_\delta(k) \]

- \( P_\delta(k) \sim k \) for large scales
- \( P_\delta(k) \sim k^{-3} \) for small scales

~ \( k_{eq} \)
Matter perturbations

Measured power spectrum for different cosmological tracers

Linear approximation

ΛCDM is a good description
The shape of the matter power spectrum is sensitive to the matter density on the Universe through the position of the turnover scale.

The turnover scale is the one that enters the horizon at the epoch of matter-radiation equality.

$$P(k) \propto k^n T^2(k) \propto \begin{cases} k^n & \text{Large scales} \\ k^{n-3} \ln^2(k) & \text{Small scales} \end{cases}$$

Log since structure grows slightly during radiation era when potential decays.

$$k_{EQ} = 0.073 \Omega_m h^2 \text{Mpc}^{-1}$$

The shape of the power spectrum is sensitive to the matter density on the Universe through the position of the turnover scale.
Neutrinos affect the large scale structure

They do not participate in collapse for scales smaller than
The freestreaming scale

\[ k_{fs}^{-1} \approx \frac{vt}{a} \approx \left( \frac{T}{m} \right) H^{-1} \]

This scale is \( \sim 0.02 \text{ Mpc}^{-1} \) for a 1 eV neutrino. Power on smaller scales is suppressed

Even for a small neutrino mass, a large impact on structure. The power spectrum is an excellent probe of neutrino masses
The Cosmic Microwave Background radiation (CMB)

One of the first and most important application of the perturbation theory and fluctuations description is the study of the CMB anisotropies.

One of the main sources of information about cosmology

Measurements of the CMB provide the most precise results of the comological parameters up to date (Galaxy surveys start to reach a similar precisión level now)
Fluctuations in the baryon-photon plasma

Equation of motion for the baryon-photon fluid

\[
\delta'' + \frac{\mathcal{H} R}{1 + R} \delta' + c_s^2 k^2 \delta = -\frac{4}{3} k^2 \Phi + 4 \Phi'' + \frac{4 R'}{1 + R} \Phi' \quad \mathcal{H} = a H
\]

↑ pressure  ↑ gravity

Until “decoupling” at \( z \approx 1100 \), there is a strong coupling between electrons and photons, which behave like a single fluid.

Baryon to photon ratio

\[
R \equiv \frac{3 \bar{\rho}_b}{4 \bar{\rho}_\gamma} = 0.6 \left( \frac{\Omega_b h^2}{0.02} \right) \left( \frac{a}{10^{-3}} \right)
\]

Sound speed

\[
c_s^2 \equiv \frac{1}{3(1 + R)}
\]

For sub-horizon modes

before MRE \( a \ll a_{eq} \)

• gravitational potential small

\[
\delta'' - \frac{1}{3} \nabla^2 \delta \approx 0
\]

after MRE \( a \gg a_{eq} \)

• grav. potential as external force

\[
\delta'' - \frac{1}{3} \nabla^2 \delta = \frac{4}{3} \nabla^2 \Phi = \text{const.}
\]
Calculation more complicated.
Need to take into account all the physics of the photon-electron plasma
Well-known physics → allows a precise prediction of the CMB power spectrum
The shape of the power spectrum has a lot of cosmological information
Before recombination
• Early Universe
• High temperature
  – *Electrons are free*
  – *Light interacts with them*

Recombination
• Late Universe
• Lower temperature
  – *e− y p+ form hydrogen*
  – *Light travels freely*
Cosmic Microwave Background (CMB)

Thermal radiation from the atom formation period ~380000 years after the BB or .... 13800 Myears ago!!

Discovered in 1965
In 1992 Discovery of its non-uniformity. Its small anisotropies are the imprint of the seeds of all the structure of the Universe.

The most precise measurement of the cosmological parameters come from the CMB.

The Cosmic Microwave Background Radiation's "surf of last scatter" is analogous to the light coming through the clouds to our eye on a cloudy day.
The frequency spectrum is a perfect black body at 2.725K.

The Universe was in thermodynamic equilibrium before the recombination: *The collision rate was much smaller than the expansion rate.*
On the CN non-discovery


Herzberg (1950) in Spectra of Diatomic Molecules, p 496:

“From the intensity ratio of the lines with $K=0$ and $K=1$ a rotational temperature of $2.3^\circ$ K follows, which has of course only a very restricted meaning.”

There went Herzberg’s [second] Nobel Prize.
CMB Temperature vs. z

$\beta = -0.007 \pm 0.027$

- COBE
- CO Molecule lines
- SZ Effect
- C atom lines

A Big Media Splash in 1992:
THE TIMES
25 April 1992

Prof. Stephen Hawking of Cambridge University, not usually noted for overstatement, said: “It is the discovery of the century, if not of all time.”
Dipole anisotropy from the Earth movement

Solar System: $v = 368 \pm 2 \text{ km/s}$
Towards the constellation of Leo
Planck: The most recent satellite

From May 2009 to October 2013
Much more precise than previous
Able to measure polarization
Arrived at L2 in July 2009.
Final results in July 2018.
Final results released on 17 July 2018

Highest precisión confirmation of ΛCDM
Statistical Properties

Expansion in spherical harmonics (Fourier transform on the sphere)

Quantifies clustering at different scales

$T_0 = 2.726\text{K}$

$\Delta T(\theta,\phi) = T(\theta,\phi) - T_0$

$$\frac{\delta T}{T_0}(\theta,\phi) = \sum a_{\ell m} Y_{\ell m}(\theta, \phi)$$

$$a_{\ell m} = \int Y_{\ell m}^*(\theta, \phi) \frac{\delta T}{T_0}(\theta, \phi) d\Omega$$

$$C_\ell \equiv \langle |a_{\ell m}|^2 \rangle = \frac{1}{2\ell + 1} \sum_m \langle |a_{\ell m}|^2 \rangle$$
Spherical Harmonics:

$l=1$

by Matthias Bartelmann
$l=2$
$l=3$

by Matthias Bartelmann
\[ l = 4 \]
by Matthias Bartelmann

l=6
$l=7$
Higher $l$ means smaller scales; $l \sim \pi / \theta$

$l = 8$
Example of a map reconstruction

l=1
$l=1 + l=2$
l = 1-4

by Matthias Bartelmann
$l = 1 - 6$
$l = 1 - 7$
$l = 1 - 8$
Original map

by Matthias Bartelmann
3 zones in the power spectrum

There is a characteristic scale, $\theta \sim 1^\circ$
Planck vs. $\Lambda$CDM
PLANCK 2018

- Dark Energy: 68.5%
- Ordinary Matter: 4.9%
- Dark Matter: 26.6%
Spherical collapse, non-linear evolution

New concept: **HALO**

Halos are the self-gravitating systems in the Universe

Peaks in the density field above $\delta_c$

Sites for Galaxy formation (gastrophysics, virialization)

Halos are non-linear peaks in the dark matter density field whose self-gravity has overcome the Hubble expansion
Spherical collapse, non-linear evolution

Spherical model: Overdense sphere ↔ closed sub-universo

Friedmann equation in a closed universe

\[
\frac{1}{a} \frac{da}{dt} = H_0 \left( \Omega_m a^{-3} + (1 - \Omega_m) a^{-2} \right)^{1/2}
\]

\[
A(1 - \cos \theta) \quad A = r_0 \Omega_m / 2(\Omega_m - 1)
\]

\[
B(\theta - \sin \theta) \quad B = H_0^{-1} \Omega_m / 2(\Omega_m - 1)^{3/2}
\]

\[
\delta \equiv \frac{\rho_m - \bar{\rho}_m}{\bar{\rho}_m} \approx \frac{3}{20} \left( \frac{6t}{B} \right)^{2/3}
\]

density perturbation within the sphere

3 epochs:

1) Turnaround: Sphere breaks away from the general expansión and reaches a maximum radius (at $\theta = \pi$, $t = \pi B$)
   Density enhancement $\rho/<\rho> \sim 5.55$ and $\delta \sim (3/20)(6\pi)^{2/3} \sim 1.06$

2) Collapse: Sphere will collapse to a singularity at $\theta = 2\pi$ (in reality it virializes due to non-gravitational physics)

3) Virialization: Interactions $\rightarrow$ Convert kinetic energy of collapse into random motions, $V=-2K$
   Density enhancement at collapse: $\rho/<\rho> \sim 178$; $\delta_C \sim 1.686$

$\theta = H_0 \eta (\Omega_m - 1)^{1/2}$
Solve with the development angle (Scaled conformal time $\eta$) as the parameter
Gravitational collapse

![Diagram of gravitational collapse](image)

- **Initial seed**
- **overdense region**
- **Expansion (Hubble flow)**
- **Colapsed region (a galaxy)**

**Time**
The over-density for linear theory is \( \delta_c \approx 1.69 \). The density at linear turnaround is \( d_{\text{lin}} = 1.69 \). The maximum density \( \rho_{\text{max}} / \bar{\rho} \approx 5.5 \). The virial equilibrium radius is \( \rho_{\text{vir}} / \bar{\rho} \approx 178 \). The scale factor is \( a^{-3} \). The universe expands by \( 2^{2/3} \) less dense by a factor of 4.
To quantify this distributions, define the mass function: Number of halos with a mass above some threshold

\[
f(\sigma, z; X) \equiv \frac{M}{\bar{\rho}} \frac{dn_X(M, z)}{d\ln \sigma^{-1}}
\]

\[
s^2(M, z) = \frac{b^2(z)}{2\pi} \int k^2 P(k) W^2(k; M) dk
\]

Mass function is parameterized in terms of fluctuations in the mass field

Press-Schechter:
The fraction of mass in halos \(M\) → the fraction of volume with density above threshold \(\delta_c\)

\[
\delta_c = 1.686
\]

Many formulae:

Press & Schechter 1974
Sheth & Tormen 1999, 2001
Jenkins et al 2001
Reed et al 2005
Warren at al 2005

\[
f(\sigma, PS) = \sqrt{\frac{2}{\pi}} \frac{\delta_c}{\sigma} \exp(-\frac{\delta_c^2}{2\sigma^2})
\]
Non-linear growth and N-body simulations

Numerical N-body simulations are the best tool to understand the nature of non-linear dynamics, and to test methods and compare with observations.

Simulations use dark matter halos and evolve them using only gravity, evolving into a nonlinear gravitational clustering.

Galaxies are included in dark matter halos using semi-analitical and phenomenological methods, and matching them to observations (reproduce clustering, bias...)

<table>
<thead>
<tr>
<th>z=18.3</th>
<th>z=5.7</th>
<th>z=1.4</th>
<th>z=0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.21 Gyr</td>
<td>1.0 Gyr</td>
<td>4.7 Gyr</td>
<td>13.6 Gyr</td>
</tr>
</tbody>
</table>
How to measure the Large Scale Structure of the Universe
How to measure LSS: Characterizing structure

If one Galaxy has comoving coordinate \( x \), then the probability of finding another Galaxy in the vicinity of \( x \) is not random. They are correlated.

Consider two comoving points \( x \) and \( y \). If \(<n>\) is the average number density of galaxies, probability of finding a Galaxy in the volume element \( dV \) around \( x \) is

\[
P_1 = <n> \, dV
\]

In practice, assume \( dV \) is small so that \( P_1 << 1 \) and the probability of finding >1 galaxies in \( dV \) is negligible.

The probability of finding a Galaxy in \( dV \) around \( x \) and finding a Galaxy in \( dV \) around \( y \) is

\[
P_2 = (\langle n \rangle \, dV)^2 \left[ 1 + \xi_G(x,y) \right]
\]

If the probabilities were uncorrelated, \( P_2 = P_1^2 \). Because they are correlated include an extra term \( \xi_G(x,y) \), which is the correlation function.
How to measure LSS: Characterizing structure

Many methods:

- The Spatial Correlation Function
- The Angular Correlation Function
- Power spectrum
- Counts in Cells
- Void Probability Functions
- Higher order statistics

Generally we want to measure how a distribution deviates from the Poisson case.

How can we distinguish between a random and a clustered distribution?

2pt-correlation function or power spectrum are the main observables to study the structure of the universe.
How to measure LSS: Characterizing structure

2 possible measurements:

- **Spatial correlation function** $\xi(r,z)$ (clustering in 3D)
- **Angular correlation function** $w(\theta,z)$ (projected sky)

Excess of probability with respect to a uniform distribution to find two galaxies separated by $r$ or $\theta$

$$dP = n \left( 1 + \xi(r,z) \right) dV ; \xi(r,z) > -1 ; \xi(r,z) \to 0 \text{ when } r \to \infty$$

**Definition of the correlation function**

$$\xi(r_1, r_2) = < \delta(r_1) \delta(r_2)> =$$

$$= \xi(|r_1-r_2|)$$

In practice: the correlation function is calculated by counting the number of pairs around galaxies in a sample volume and comparing with a Poisson distribution.
How to measure LSS: Characterizing structure

Compare the data with a homogeneous randomly distributed (no clustering) distribution of points, that has the same spatial sampling as galaxies.

Estimators of the Correlation Function

\[ w(\theta) = (DD/RR) - 1 \quad \text{Natural} \]

\[ w(\theta) = (2DD/DR) - 1 \quad \text{Standard} \]

\[ w(\theta) = (DD-2DR+RR)/RR \quad \text{Landy-Szalay} \]

\[ w(\theta) = 4(DD\times DR)/(DR^2-1) \quad \text{Hamilton} \]

DD(r) number of pairs data-data
RR(r) number of pairs random-random
DR(r) number of pairs data-random

Using the random sample one can take into account practical difficulties like the partial covering of the sky with observations or the different depth of the observations for different points in the sky.
Comparing measurements to theory: The correlation function is the Fourier transform of the power spectrum. The power spectrum and correlation function contain the same information; accurate measurement of each give the same constraints on cosmological models.

\[
\delta(x) = \sum_k \delta_k e^{i k \cdot x} \quad \delta_k = \frac{1}{V} \int_V \delta(x) e^{-i k \cdot x} \, d^d x
\]

\[
\langle \delta^*_k \delta_k' \rangle = \frac{1}{V^2} \int d^d x e^{i k \cdot x} \int d^d x' e^{-i k' \cdot x'} \langle \delta(x) \delta(x') \rangle
\]

\[
= \frac{1}{V^2} \int d^d x e^{i k \cdot x} \int d^d r e^{-i k' \cdot (x + r)} \langle \delta(x) \delta(x + r) \rangle
\]

\[
= \frac{1}{V^2} \int d^d r e^{-i k' \cdot r} \xi(r) \int d^d x e^{i(k-k') \cdot x}
\]

\[
= \frac{1}{V} \delta_{kk'} \int d^d r e^{-i k \cdot r} \xi(r) \equiv \frac{1}{V} \delta_{kk'} P(k),
\]
How to measure LSS: Characterizing structure

Since $\xi(r)$ is independent of the $r$ direction, the angular integrals can be calculated:

\[
P_g(k) = \int_0^{2\pi} d\phi \int_0^{\pi} \sin \theta d\theta \int_0^{\infty} dr r^2 \xi_g(r) e^{-ikr \cos \theta}
\]

\[
= 2\pi \int_0^{\infty} dr r^2 \xi_g(r) \int_0^{\pi} d\theta \sin \theta e^{-ikr \cos \theta}
\]

\[
= 2\pi \int_0^{\infty} dr r^2 \xi_g(r) \frac{1}{i kr} \int_{-ikr}^{ikr} dx e^{-x}
\]

\[
= 4\pi \int_0^{\infty} dr r^2 \xi_g(r) \frac{\sin(kr)}{kr},
\]
How to measure LSS: Characterizing structure

The correlation function (or the power spectrum) contains the full statistical information only for Gaussian distributions.

This is the 2-point correlation function.

Higher order statistics to obtain more information: 3, 4 ... points correlations functions (instead of pairs, consider triangles, quadrangles...) \(\rightarrow\) non-Gaussianity
How to measure LSS: Characterizing structure

To measure the correlation function, we need a catalog of objects (usually galaxies)

2 main kinds of catalogs:

**Spectroscopic:** Obtain the spectrum for a selected group of galaxies. This gives an accurate determination of the redshift, and allows to measure the full spatial distribution.

**Photometric:** Obtain images in different colors for all the objects. This gives a not so precise determination of the redshift (photometric redshift or photoz). Measure the angular (projected) distribution of galaxies for several redshift intervals.

We can define angular quantities that behave like the full spatial ones:

Angular correlation function \( w(\theta, z) \), angular power spectrum \( C_l \)

\[
\omega(\theta) = \int_0^\infty dz_1 \phi(z_1) \int_0^\infty dz_2 \phi(z_2) \xi(r; \bar{z})
\]

\[
\bar{z} = \frac{z_1 + z_2}{2}
\]

\[
r = \sqrt{\chi(z_1)^2 + \chi(z_2)^2 - 2\chi(z_1)\chi(z_2) \cos \theta}
\]

\[
\chi(z) = \frac{c}{H_0} \int_0^z dz \sqrt{\Omega_m (1+z)^3 + \Omega_\Lambda (1+z)^3(1+w)}
\]
How to measure LSS: Characterizing structure

Non-trivial to compare observation to theory

The observables, $\delta_{\text{gal}}$, are complicated \textit{functionals} of the easy-to-predict linear matter density field, $\delta_L$.

$\delta_L \rightarrow \delta_{\text{NL}} \rightarrow \delta_{\text{gal}}$

- N-Body interactions in Newtonian gravity
- Galaxy formation including hydro, feedback from SN, star formation, ...
How to measure LSS: Characterizing structure

Difficulties: Biasing. We observe galaxies, not dark matter.

How well do galaxies trace the underlying perturbations in the matter?

Correlation function depends on galaxy properties: Brighter, more massive galaxies have a larger bias than fainter, lower mass galaxies

The different clustering properties of these galaxies tell us something about how they form

Use these dependencies in the data analysis to obtain information about bias and control systematic errors
The galaxies we observe do not perfectly trace the underlying mass distribution in the universe (i.e., light does not trace mass)

Expect galaxies to be found preferentially in the most prominent high-mass peaks
How to measure LSS: Galaxy Bias

Express fluctuations in the number of observed galaxies in terms of fluctuations in the mass density times biasing factor:

$$\delta_{\text{Galaxies}} = b \delta_{\text{Matter}}$$

*linear bias*  
(in general, more complicated)

In general, $b \geq 1$

Bias depends on the properties of the selected Galaxy sample
How do we compile these galaxy samples?

I. Obtaining multi-colour images of a large area of the sky

II. Create a catalogue and then select the sources over some range of brightness (and perhaps using some other criteria)

III. Measure redshifts for sources (to add third dimension)
How to measure LSS: Characterizing structure

Example: Galaxy density map for DES Science Verification data (Benchmark Sample)
How to measure LSS: Characterizing structure

Survey conditions maps that can affect the clustering of galaxies
How to measure LSS: Characterizing structure

Measured correlation functions

M. Crocce et al.,
How to measure LSS: Characterizing structure

From the previous analysis one can obtain the Galaxy bias and its evolution with the redshift.

Linear bias
DES–SV bench–mark sample ($i<22.5$)

- BPZ (template)
- TPZ (machine learning)

$\sigma_8$ ($0.83$)

$z$

CFHTLS ($i<22.5$)
Of course, there are other techniques as well for quantifying clustering:

Counts In Cells -- Divide the Space into Discrete Grid Points “Cells” and Calculate the Variation in the # of Sources per Grid Point

Void Probability Function -- Probability of Finding Zero Galaxies in a Volume of Radius R

Higher order statistics – 3pt correlation functions

*We will not describe them in this course*
For next sessions: How to do cosmology with the correlation function

Baryon acoustic oscillations

Redshift space distortions

Other probes and combinations
Non-linear regime

Redshift space distortions

Baryon Acoustic Oscillations

$z=0$

$\xi(r)$

$r$ [Mpc/h]

Linear theory

Non-linear theory

Simulations