The Real-Space Two-Point Correlation Function ξ(r) and the BAO Standard Ruler

1. Introduction

The real-space two-point correlation function, denoted ξ(r), is one of the most fundamental descriptors of large-scale structure. It quantifies how galaxies (or matter) are clustered relative to a uniform random field. While the power spectrum P(k) describes clustering in Fourier space, ξ(r) provides a complementary view in physical separation r, making it ideal for highlighting the Baryon Acoustic Oscillation (BAO) feature at ~105–110 Mpc/h.

Understanding ξ(r) unifies early-universe physics, CMB anisotropies, linear perturbation theory, and galaxy survey observations (e.g., 2dF, SDSS, BOSS, and DESI). In many ways, the BAO peak in ξ(r) is the direct fossil of acoustic waves propagating in the primordial photon-baryon fluid.

2. Definition of ξ(r)

Consider a galaxy survey with mean number density ⟨n⟩. The two-point correlation function is defined:

Definition of Real-Space Two-Point Correlation Function, *ξ*

$$\xi(r) = \frac{\langle n(\mathbf{x})\, n(\mathbf{x}+\mathbf{r})\rangle}{\langle n\rangle^2} - 1$$

n(x) = number density galaxies at ***x***
⟨*n*⟩ = mean number density
r = |***r***| = Physical separation

It measures the excess probability of finding two galaxies separated by distance r.

The probability of finding a galaxy in volume dV_1 and another dV_2, separated by *r*, is

$$dP = n^2\,[1+\xi(r)]\,dV_1\,dV_2$$

ξ(r) > 0 → clustering at scale r
ξ(r) = 0 → random distribution
ξ(r) < 0 → under-dense or void-like behavior

Relationship Between P(k) and ξ(r)

ξ(r) is the 3-dimensional Fourier transform of the matter power spectrum:

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

This integral reveals:

The broadband shape of ξ(r) comes from the general slope of P(k).
The oscillatory wiggles in P(k) (acoustic peaks) transform into a single broad BAO peak in ξ(r).
Large scales (small k) dominate behavior near the BAO radius.

This Fourier relationship is central: ξ(r) provides a complementary real-space visualization of physics encoded in P(k).

Physical Origin of the BAO Peak

Acoustic oscillations before recombination
Before z ≈ 1100, baryons and photons were tightly coupled. Overdensities launched spherical sound waves, propagating outward at

$$c_s \approx \frac{c}{\sqrt{3(1+R)}}, \quad R = \frac{3\rho_b}{4\rho_\gamma}$$

At recombination, photon pressure vanished and the wavefront "froze in" at the sound horizon: $r_s \approx 105\ h^{-1}$ ***Mpc***

Result in ξ(r)

A galaxy today is slightly more likely to have a neighbor at **~105 Mpc/h**, the radius of the primordial shell.
This produces a **broad peak** in ξ(r) centered near that separation.
The BAO peak serves as a **standard ruler** for measuring the expansion history

Typical Shape of ξ(r)

1. Small scales (r < 5 Mpc/h):
   Strong clustering; ξ(r) ≫ 1.
2. Intermediate scales (5–50 Mpc/h):
   ξ(r) decays roughly as a power law, often approximated by: $\xi(r) \approx \left(\frac{r}{r_0}\right)^{-1.8}$
3. BAO scale (~100–115 Mpc/h):
   A broad peak, amplitude of a few × 10^{-2}.
4. Large scales (>150 Mpc/h):
   ξ(r) → 0 or slightly negative (approaching homogeneity).

6. A Practical Fitting Model for ξ(r)
For applications, ξ(r) is often approximated as:

ΛCDM-Like Analytic Model Approximation: $\xi_{model}(r)$

$$\xi(r) = A\,r^{-\alpha}e^{-r/r_d} + B\exp\left[-\frac{(r-r_{\rm BAO})^2}{2\sigma^2}\right]$$

where:

- $Ar^{-\alpha}e^{-r/r_d}$ → broadband clustering
- Gaussian term → BAO bump
- $r_{\rm BAO} \approx 108\,{\rm Mpc}/h$
- $\sigma \approx 10\text{–}15\,{\rm Mpc}/h$ (Silk damping + nonlinear broadening)

$$\xi_{model}(r) := 4.8\cdot 10^{4}\cdot r^{-1.73}\cdot e^{\frac{-r}{210}} + 43\cdot exp\left[\frac{-(r-108.5)^2}{2(11.5)^2}\right] + \Xi$$

This model fits SDSS-III and BOSS DR12 ξ(r) data shown below eand matches the extracted points from the BAO figure.

Measurement in Galaxy Redshift Surveys

Modern surveys measure ξ(r) by:
Counting galaxy pairs at separation r → **D–D counts**
Comparing to random catalog pairs → **R–R**
Using the Landy–Szalay estimator of *ξ(r)* :

$$\xi(r) = \frac{DD - 2DR + RR}{RR}$$

This minimizes variance and corrects for survey geometry.

Key features measured:

Position of the BAO peak → gives $D_V(z)$, H(z), $D_A(z)$
BAO peak height + width → growth of structure
Broadband slope → Ω_m, n_s, dark energy effects

What is a "galaxy pair": Landy–Szalay estimator

Take all the galaxies in a survey (e.g., SDSS, DESI).
For every pair of galaxies, compute their separation r.
Example: If galaxy A and galaxy B are 110 Mpc/h apart, that is one pair with separation 110.
A large survey may contain billions of such pairs.

D–D (Data-Data) = the number of galaxy pairs in the real data at separation r

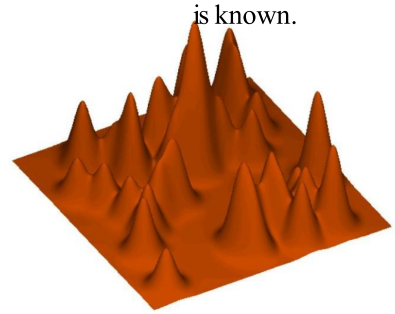
"D" stands for Data — the actual observed galaxy positions. Count how many pairs of real galaxies have separations **in each bin** (e.g., 100–101 Mpc/h).
This measures:
how often real galaxies are found at distance r.

If the Universe were completely uniform,
DD(r) would follow a predictable curve (roughly ∝ r^2).
But clustering causes **excess counts** at certain scales → that's ξ(r).

Why ξ(r) is Important

It is **the clearest observable imprint** of the **sound horizon in the late Universe**.
It provides a geometric standard ruler, independent of supernova systematics.
ξ(r) is less sensitive than P(k) to small-scale nonlinearities.
It directly connects CMB physics to galaxy clustering.

Many Superimposed Waves
Positions predicted once dark matter and baryon density is known.



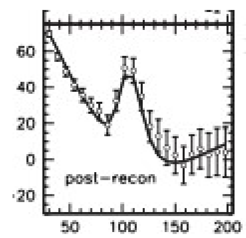
BAO measurements in ξ(r) contribute enormously to constraints on:

Dark energy equation of state w(z)
Curvature Ω_k
Hubble parameter H(z)
Matter density Ω_m

Example: BAO Data Extraction

Using the SDSS post-reconstruction plot in this paper, the extracted ξ(r) values show:

A declining broadband shape
A well-defined BAO bump near 108–110 Mpc/h
Slight negative dip near 140 Mpc/h
Rising toward 200 Mpc/h as nonlinear effects increase



These values match the analytic model above and provide a real worked example.

Summary

The real-space correlation function ξ(r):
Encodes how galaxies cluster as a function of physical separation
Is the Fourier counterpart to the power spectrum P(k)
Contains a robust BAO peak from primordial acoustic waves
Serves as an essential tool for precision cosmology
Provides a direct geometric method for testing ΛCDM and probing dark energy

Incorporating ξ(r) and the BAO standard ruler into the broader framework of cosmological modeling completes the unified picture of how early-universe physics shapes present-day structure.

DETECTION OF THE BARYON ACOUSTIC PEAK

Having established the physical origin of the sound horizon and its role in setting the characteristic acoustic scale, it is useful to examine how this primordial feature manifests itself in present-day large-scale structure. The same oscillations imprinted in the early photon–baryon fluid leave a corresponding signature in the late-time matter distribution, producing a distinct bump in the real-space two-point correlation function ***ξ(r)*** at a comoving separation of roughly 1**05 h^{-1} Mpc**. **This "BAO peak"** provides a powerful standard ruler that connects the physics of recombination to the clustering of galaxies in the low-redshift Universe. To illustrate this connection explicitly, the following dataset and plot show ξ(r) over separations of 50–250 h^{-1} Mpc, highlighting the BAO feature predicted by ΛCDM and measured in surveys such as SDSS and BOSS.

Extended Baryon Oscillation Spectroscopic Survey (eBOSS),

https://arxiv.org/pdf/astro-ph/0501171

The acoustic peaks in the cosmic microwave background (CMB) anisotropy power spectrum have emerged as one of **the strongest cosmological probes**. They measure the contents and curvature of the universe Acoustic peaks occur because the cosmological perturbations excite sound waves in the relativistic plasma of the early universe. The recombination to a neutral gas at redshift z ~ 1000 abruptly decreases the sound speed and effectively **ends the wave propagation**. In the time between the formation of the perturbations and the epoch of recombination, **modes** of different wavelength can complete different numbers of oscillation periods. **This translates the characteristic time into a characteristic length scale &** produces a **harmonic series of maxima and minima in the anisotropy power spectrum**.

The data in the file: "SDSS-III BAO Measurements.txt"
gives a table that presents a theoretical estimate of the real-space two-point correlation function, ξ(r),

for comoving separations between 50 and 250 h^{-1} Mpc in a standard ΛCDM cosmology. This function quantifies the excess probability, relative to a random distribution, of finding pairs of galaxies separated by a distance r, and it **encodes key information about the large-scale clustering of matter in the Universe.**

Research Study:

The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples,
MNRAS 441, 24–62 (2014)

BAO position in spherically averaged two-point measurements is fixed by the projection of the sound horizon at the drag epoch, r_d, and provides a measure of:

$$D_V(z) \equiv \left[cz(1+z)^2 D_A(z)^2 H^{-1}(z)\right]^{1/3}$$

Definition of Real-Space Two-Point Correlation Function, ξ

$$\xi(r) = \frac{\langle n(\mathbf{x})\, n(\mathbf{x}+\mathbf{r})\rangle}{\langle n\rangle^2} - 1$$

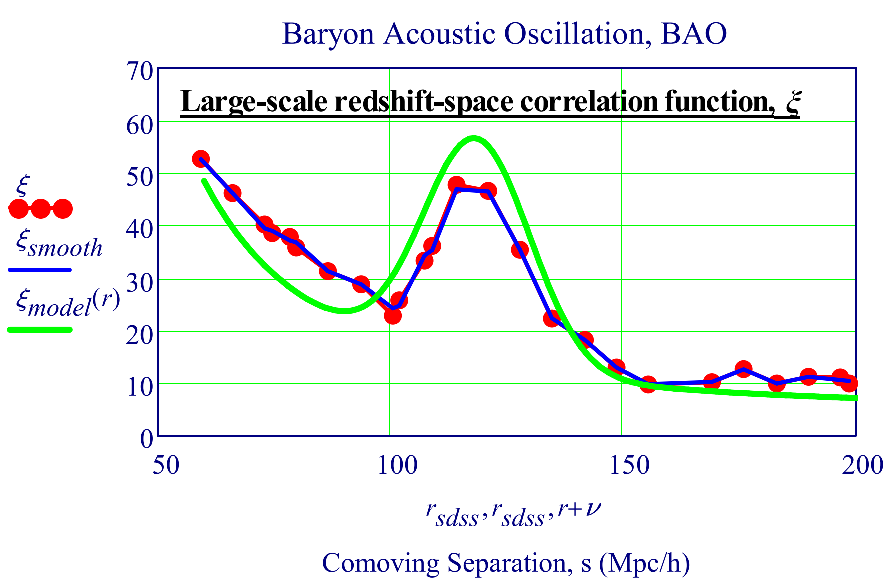
n(x) = number density galaxies at ***x***
⟨*n*⟩ = mean number density
r = |***r***| = Physical separation

Fetch the Correlation Data from the SDSS-III Paper ⟶ $Corr := READPRN(\text{"SDSS-III BAO Measurements.txt"})$

$r_{sdss} := Corr^{\langle 0\rangle}$ $\quad \xi := Corr^{\langle 1\rangle}$ $\quad \xi_{smooth} := ksmooth(r_{sdss}, \xi, 6)$

$$\xi_{model}(r) := 4.8\cdot 10^4 \cdot r^{-1.73} \cdot e^{\frac{-r}{210}} + 43\cdot exp\left[\frac{-(r-108.5)^2}{2(11.5)^2}\right]$$

Example:
At r ≈ 100Mpc/h,
ξ(r) has a small bump



Introduction: Weak Gravitational Lensing

Weak gravitational lensing uniquely probes the growth of structure rather than cosmic geometry alone. Unlike supernovae and baryon acoustic oscillations, which primarily constrain distance–redshift relations, lensing directly measures the amplitude and evolution of matter fluctuations through the coherent distortion of background galaxies. As a result, weak lensing is particularly sensitive to the parameter combination,

$$S_8 = \sigma_8 \cdot \sqrt{\frac{\Omega_m}{0.3}}$$

making it a critical test of ΛCDM consistency at late cosmic times.

Variables

- shear γ
- convergence κ
- kernels W(z)

Conceptual Overview of Weak Gravitational Lensing and the Convergence Field

This diagram shown below illustrates the physical and observational basis of weak gravitational lensing in large-scale structure. Light from distant background galaxies is deflected by intervening foreground dark-matter halos along the line of sight, producing small but coherent distortions in observed galaxy shapes.

These distortions are quantified by the **shear field** γ**,** an observable extracted statistically from ensembles of galaxy ellipticities. Through line-of-sight integration and inversion techniques, **the shear field is used to reconstruct the convergence field** $\kappa(\theta)$, which traces the projected mass density.

The resulting κ maps and their angular power spectra encode information about the underlying matter power spectrum ***P(k)***, the matter density parameter Ω_m, and the growth of structure through the growth factor ***D(z)***.

To see how it relates to ΛCDM - Refer to Section:

XXIII B Python ΛCDM Six-Parameter Base Model - GCP

Introduction for CAMB_CMB_TT and_Lensing