Probes of Structure Formation

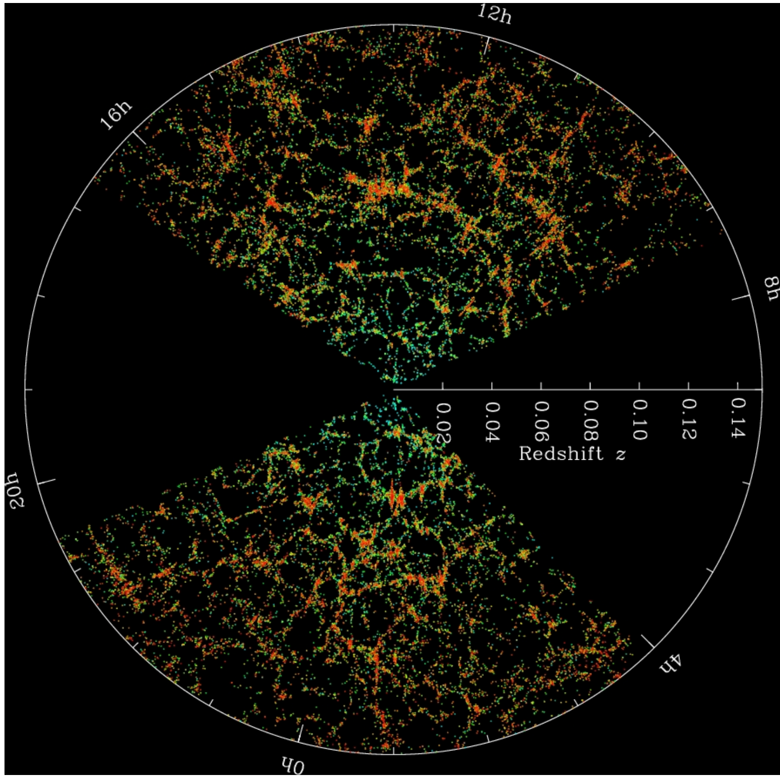
Galaxy Clustering

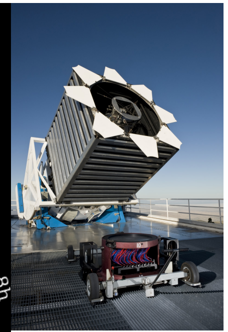
Galaxy Clustering

Observationally →

the **angular correlation function in real space**, the **correlation function in redshift space**, and the **BAO peak** are three cosmological functions measured from the observed spatial distribution of galaxies (observable)

Theoretically → need to compute the matter power spectrum and its BAO peak, and the redshift-space distortions





SDSS

$\delta_{dm} \rightarrow \delta_g$

The observed δ_g is a biased tracer of δ_{dm},
i.e., light only follows matter in an approximate way.

Galaxy clustering: theoretical predictions

The number density of galaxies represent the density contrast of galaxies δ_g, which traces the dark matter density contrast δ, with two caveats:

- **Galaxies are a biased tracer of dark matter**

There is a bias, defined as the ratio between the galaxy density contrast and the dark matter density contrast: $\delta_g = b\,\delta$

The bias depends on galaxy formation. The linear biasing model assumes that galaxies form at the peaks of matter density and only form when those peaks exceed a certain threshold → b > 1 (the amplitude of the galaxies power spectrum is larger than the dark matter power spectrum) and is determined by the growth rate:

$$b(t) - 1 = \frac{D(t_i)}{D(t)}(b_i - 1) \qquad D(t)/D(t_i) = \delta_M(\boldsymbol{r}, t)/\delta_M(\boldsymbol{r}, t_i)$$

In other models the bias is scale- and redshift-dependent: b = b (k,z)

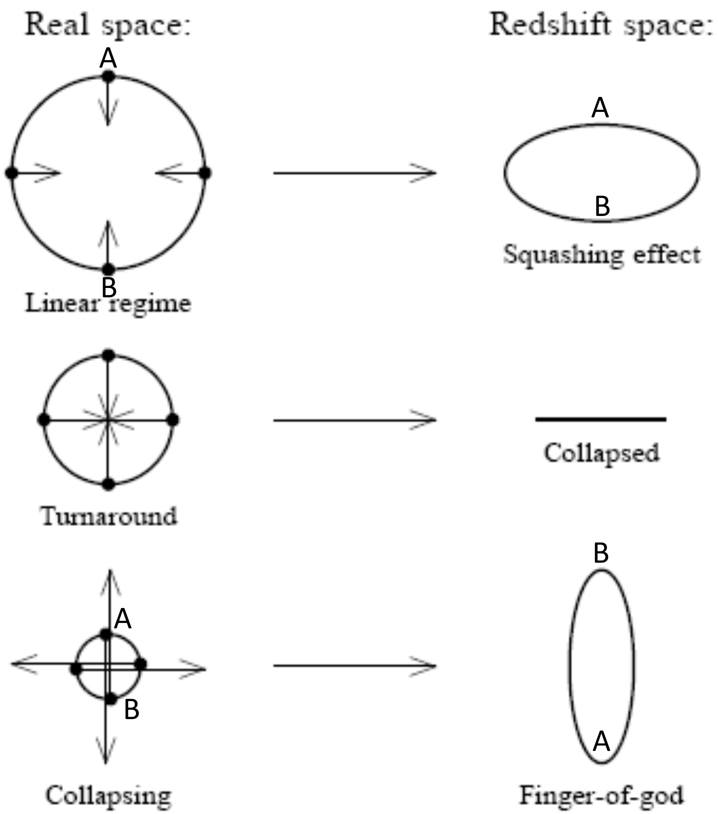
- **3D positions of galaxies are not known in real space, but only in redshift space**

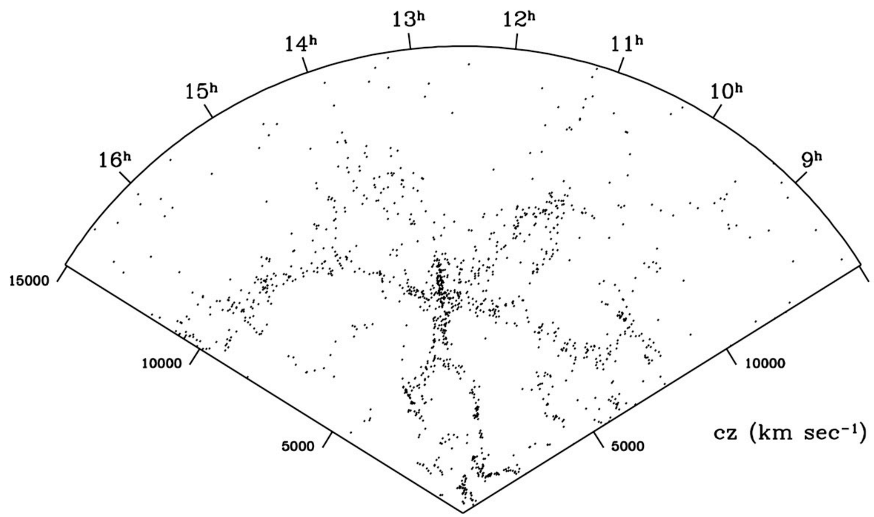
The clustering process produces peculiar velocities which lead to **redshift-space distortions** **(RSD)**

Redshift space distortions (correlation function in redshift space)

Because of **peculiar velocities**, the redshift is not a direct indicator of distance - especially at low z and for non-virialized objects where the closest point has a larger redshift than the farthest point!

$$z = \frac{D\,H_0 + v}{c}$$





The famous finger-of-God in the Coma cluster

To go from **redshift z-axis position** (s_3) to **coordinate space z-axis position** (r_3) we need to correct for the peculiar velocity (v_3).

$$r_3 = s_3 - \frac{v_3}{H_0}$$ → **the galaxy density contrast in the redshift space is a biased estimator of the galaxy density contrast in real space**

$$\delta_g^s = \delta_g - \frac{1}{H_0}\frac{dv_3}{dr_3}$$

Now, note that **the peculiar velocity term can be written in terms of the density contrast** (no need for extra observations of the velocities).

This is because a peculiar velocity is due to the gravitational field created by the overdensities → it is a gradient of the potential, and so it is related with δ

$$v = \nabla\psi$$

On the other hand, in the matter-dominated epoch, we know that **δ grows $\sim a^f$**

where
$$f = \frac{d\ln\delta}{d\ln a} = \frac{\dot{\delta}}{\delta H}$$

We can insert this solution in the first-order continuity equation that involves δ and v, and **obtain a solution for div v** :

$$\frac{\partial\delta}{\partial t} + \frac{1}{a}\nabla.v = 0 \quad \rightarrow \quad \text{div v} = - a H_0 f \delta$$

Assuming dv_3/dr_3 is given by div v, we get a solution for the relation between the galaxies overdensity in redshift-space and the galaxies overdensity in coordinate-space:

$$\delta_g^s = \delta_g + f\delta$$

By introducing a linear bias b, δ can be written in terms of δ_g, and the relation can be written in terms of δ_g only:

$$\delta_g^s = \delta_g + f/b\, \delta_g = (1+\beta)\,\delta_g \qquad \text{with} \quad \beta = \frac{f}{b}$$

In reality, the derivative dv_3 / dr_3 is not exactly the same as the div.v, we need to consider the velocity direction.

The correct expression is

$$\delta_g^s = \left(1 + \beta\mu^2\right)\ \delta_g$$

where

$$\mu^2 = \frac{\vec{k}.\vec{n}}{k}$$

Introducing again the bias, we can also relate the galaxy power spectrum of RSD with the matter power spectrum:

$$P_g^s(k) = b^2\left(1 + \beta\mu^2\right)^2\ P_\delta(k)$$

This is the theoretical prediction of RSD → the measured redshift-space galaxy power spectrum is a function of the dark matter power spectrum (dependent on the bias and on the growth function)

Since the amplitude of P_δ is given by σ_8^2 → **the amplitude of the RSD power spectrum is sensitive to the combination f σ_8**

As we saw already, there are other two quantities also measured in galaxy surveys:

Angular correlation function in real space

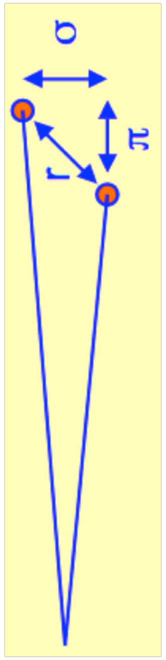
BAO peak

So galaxy clustering has 3 independent quantities, sensitive to different information:
full correlation function, BAO feature (sensitive to background), RSD feature (sensitive to growth)

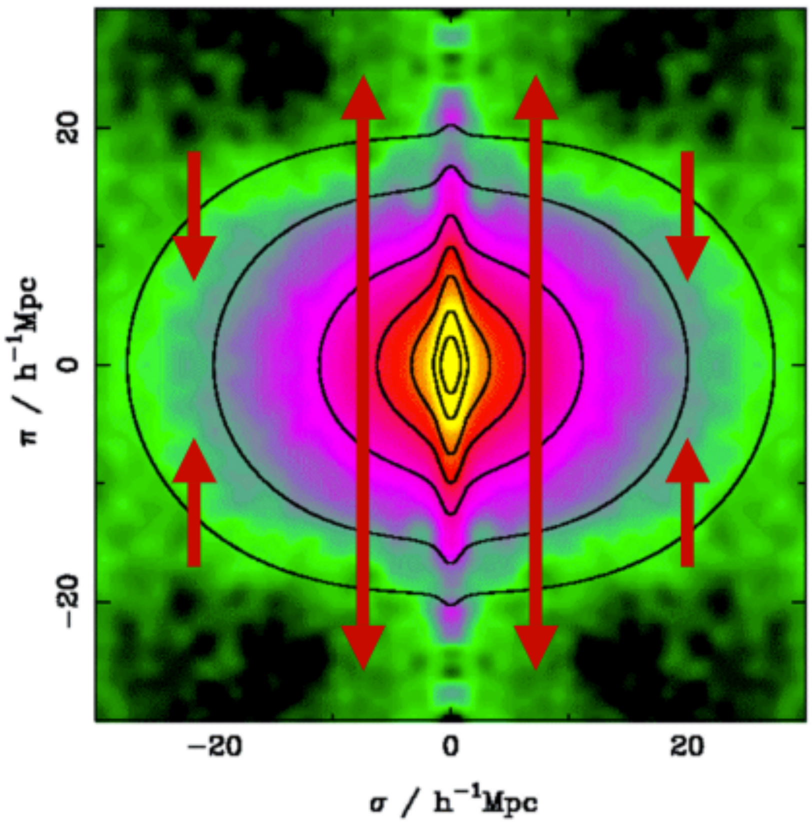
Galaxy Clustering: estimators

Redshift Space Distortions

The correlation function in redshift space is not isotropic $\xi(\sigma,\pi)$



Observer



(zoom on small scales)

Near the center of the overdensity (small angular separations): structures are elongated (Finger-of-God) → virialized motions (decoupled from the Hubble flow)

Farther from the center of the overdensity (larger angular separations: structures are flattened → coherent flow

In RSD the estimator is usually defined for the quantity $f\sigma_8$, and not for the correlation function or power spectrum.

In order to do this, **the theoretical coordinate-space correlation function is expanded on moments of the RSD correlation function** (in Legendre polynomials), which are the observables.

$$\xi(r_p,\pi) = \xi_0(s)P_0(\mu) + \xi_2(s)P_2(\mu) + \xi_4(s)P_4(\mu)$$

$$\xi_0(s) = \left(1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2\right)\xi(r),$$

$$\xi_2(s) = \left(\frac{4}{3}\beta + \frac{4}{7}\beta^2\right)\left[\xi(r) - \bar{\xi}(r)\right],$$

$$\xi_4(s) = \frac{8}{35}\beta^2\left[\xi(r) + \frac{5}{2}\bar{\xi}(r) - \frac{7}{2}\bar{\bar{\xi}}(r)\right];$$

$$\bar{\xi}(r) \equiv 3r^{-3}\int_0^r \xi(r')r'^2 dr',$$

$$\bar{\bar{\xi}}(r) \equiv 5r^{-5}\int_0^r \xi(r')r'^4 dr'.$$

Each moment can be related to the theoretical correlation function.

Taking ratios between the various moments of the measured RSD correlation function, the dependence on the coordinate-space correlation cancels out, and we are left with polynomials in the β parameter.

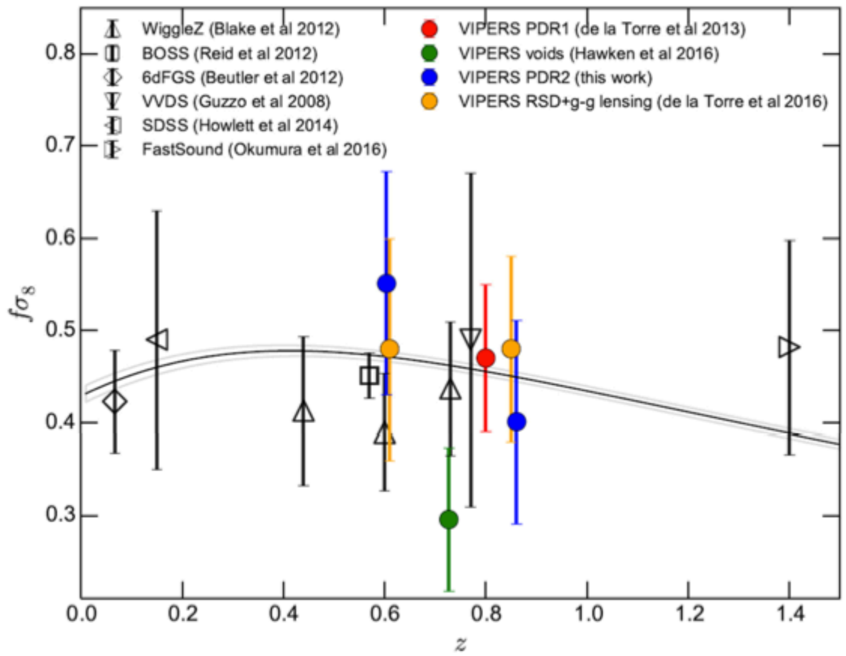
$$\frac{\xi_2(s)}{\xi_0(s)} = \frac{\gamma - 3}{\gamma}\,\frac{\frac{4}{3}\beta + \frac{4}{7}\beta^2}{1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2}$$

where

$$\gamma = \frac{\ln f}{\ln \Omega_m}$$

From the ratios we get β and can get f - **if the bias (b) is known** -
In the simple linear bias model, b is modeled as b = σ_gal /σ_8

The correlation function is also measured in various redshift bins → **the measurement (data points) can thus be shown as $f\sigma_8$ (z)**, instead of a correlation function ξ(r).



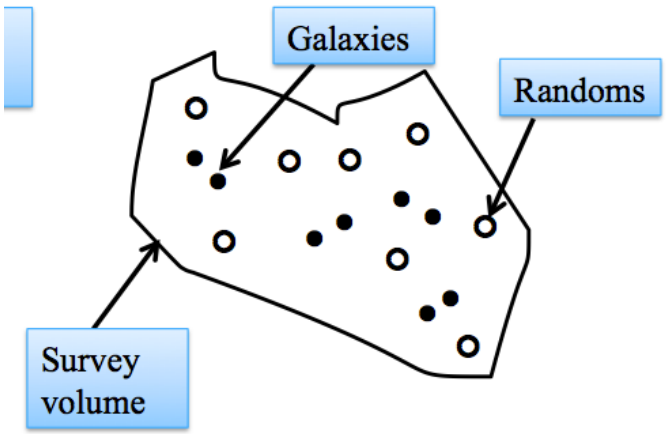
So, **basically RSD is a feature (like BAO)**. The goal of RSD measurements is not to compute RSD correlations on a wide range of separations, but to focus on the small scales where peculiar velocities are large, to get information on the growth factor.

Correlation function in real space

We already saw that the correlation function may be measured by counting galaxy pairs as function of separation and comparing this count with the number count in the case of no correlation (random point distribution).

This procedure requires that we build a sample of mock galaxies (the “randoms”), in the same survey volume and geometry, with the same spatial sampling as the data sample, but with no correlation.

DD (r) - number of galaxy-galaxy pairs as function of separation
RR (r) - number of mock-mock pairs as function of separation
DR (r) - number of galaxy-mock pairs as function of separation



Several estimators of the correlation function were proposed, based on different ways of making the data-random comparison:

$$1 + \xi_1 = \frac{\langle DD \rangle}{\langle RR \rangle}$$

$$1 + \xi_2 = \frac{\langle DD \rangle}{\langle DR \rangle},$$

$$1 + \xi_3 = \frac{\langle DD \rangle \langle RR \rangle}{\langle DR \rangle^2},$$

$$1 + \xi_4 = 1 + \frac{\langle (D - R)^2 \rangle}{\langle RR \rangle^2}$$

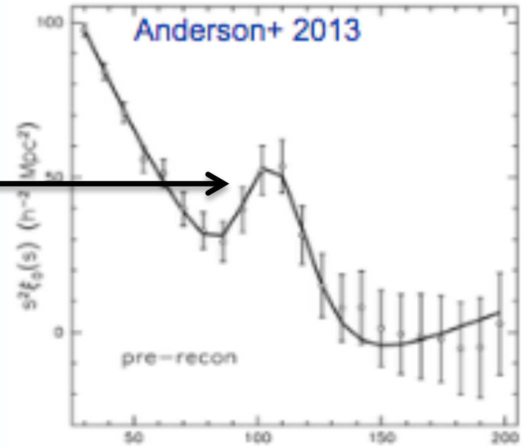
The 4 estimators have different noise properties. The last one has the best signal-to-noise ratio.

The typical result obtained for the correlation function (of galaxies positions) is a power-law, with slope γ= 1.7 (where r_0 is a critical separation that depends on the population of galaxies, a typical value is r_0 ~ 5 Mpc/h)

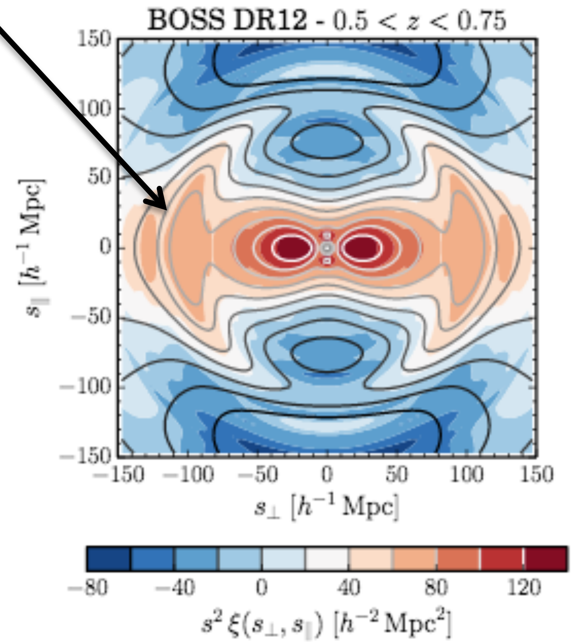
$$\xi(r) = \left(\frac{r}{r_0} \right)^{-\gamma}$$

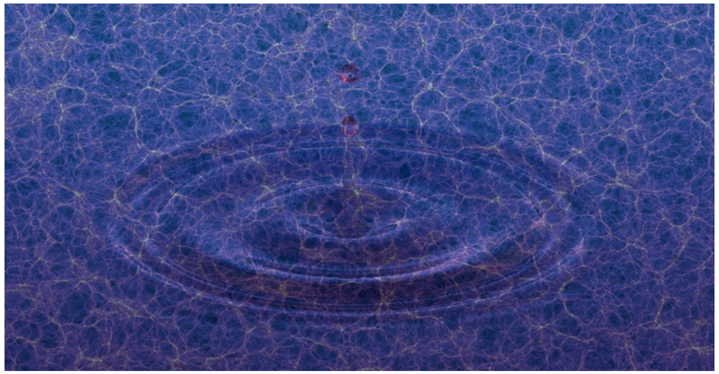
BAO

The BAO peak can be detected as a peak in the **isotropic angular real-space correlation function,** but it is also detectable as an anisotropic ring in the **RSD correlation function**.



The systematics are different in the 2 cases.
In the RSD case they are not yet fully studied.
They include non-linear smearing effects from non-linear corrections and random velocity of galaxies.





In both cases, the BAO detection is used to constrain $D_A(z)$, $H(z)$ (geometric probe)

The ring appears at ~100 Mpc/h, on quasi-linear scales where the RSD distortion is weaker.