Galaxy formation and evolution
PAP 318, 5 op, autumn 2022

BK106, Exactum

Lecture 6: Correlation functions and the spectrum of the initial fluctuations, 12/10/2022

On this lecture we will discuss

1. The two-point correlation function of galaxies.
2. The perturbation spectrum. The relationship between the two-point correlation function and the power spectrum of perturbations.
3. The form of the initial perturbation spectrum and the Harrison-Zeldovich spectrum.
4. Transfer functions for adiabatic cold and hot dark matter models.
5. Biasing of the galaxy population and baryonic acoustic oscillations in the galaxy power spectrum.
6. The lecture notes correspond to: L: p. 385-419 (§14.1-14.7)
   MBW: p. 196-208, 262-270 (§4.3-4.4,§6.1)

6.1 The two-point correlation function for galaxies I

- To make a quantitative comparison between theories of galaxy formation and the observed distribution of galaxies, we need to quantify the spectrum of density perturbations in the Universe.
- The simplest quantitative description of the statistical distribution of galaxies on large scale is provided by the two-point correlation function, which describes the excess probability of finding a galaxy at distance *r* from a galaxy selected at random over that expected in a uniform, random distribution:

$$dN(r) = N_0[1 + \xi(r)]dV$$

- N_0 is the background number density and the correlation function $\xi(r)$ can also be written in terms of finding pairs of galaxies as:

$$dN_{\rm pair} = N_0^2[1 + \xi(r)]dV_1 dV_2$$

The two-point correlation function for galaxies II

- The two-point correlation function can be directly related to the density contrast $\Delta=\delta\rho/\rho$ by writing $\rho=\rho_0[1+\Delta(\mathbf{x})]$:

$$dN_{\rm pair}(\vec{r}) = \rho(\vec{x})dV_1\rho(\vec{x}+\vec{r})dV_2$$

$$dN_{\rm pair}(\vec{r}) = \rho_0^2[1+\Delta(\vec{x})][1+\Delta(\vec{x}+\vec{r})]dV_1dV_2$$

- Taking averages over a large number of elements, the mean value of $\Delta(x)$ and $\Delta(x+r)$ are zero by definition and we get:

$$dN_{\rm pair}(\vec{r}) = \rho_0^2[1+\langle\Delta(\vec{x})\Delta(\vec{x}+\vec{r})\rangle]dV_1dV_2$$

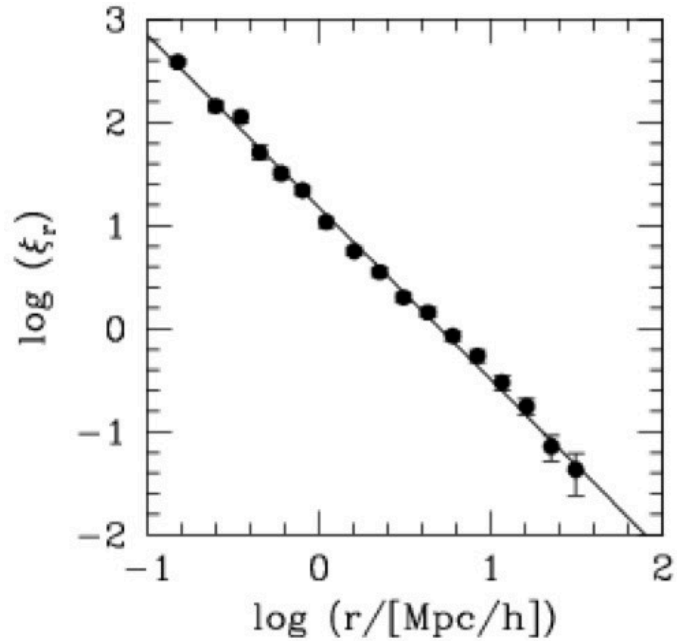
- From which we can identify: $\xi(r) = \langle\Delta(\vec{x})\Delta(\vec{x}+\vec{r})\rangle$
- Observationally we usually work with the angular two-point correlation function $w(\theta)$, which can be measured on the sky and then be converted to $\xi(r)$: $N(\theta)d\Omega = n_g[1+w(\theta)]d\Omega$

The two-point correlation function: Observations

- Observationally the two-point correlation function can be well described in the range [100h^{-1} kpc,10h^{-1} Mpc] by a power law:

$$\xi(r) = \left(\frac{r}{r_0}\right)^{-\gamma}, \; r_0 = 5h^{-1}\mathrm{Mpc}, \; \gamma = 1.8$$

1. The correlation function is quite smooth, there are no obvious preferred scales.
2. There is a characteristic scale r_0=5h^{-1} Mpc, which defines the scale at which the density of galaxies is twice the background density (i.e. the ~non-linear scale).
3. Also on larger scales, the richest galaxy clusters and brightest quasars are correlated with r_0≈(15-25) h^{-1} Mpc, although these perturbations are still in the linear stage of development.



The observed two-point correlation function from the 2dF survey.

The perturbation spectrum I

- In deriving the relationship between $\xi(r)$ and the power spectrum of fluctuations it is natural to work with wavevectors $\mathbf{k}=(2\pi/\lambda)\mathbf{i}_k$ and Fourier transforms.
- Using Parseval's theorem we can relate the integrals of the squares of $\Delta(\mathbf{r})$ and its Fourier transform Δ_k:

$$\frac{1}{V}\int \Delta^2(\vec{r})d^3x = \frac{V}{(2\pi)^3}\int |\Delta_k|^2 d^3k$$

- The left-hand side of this equation is the mean square of the amplitude of the fluctuation within volume V and $|\Delta_k|^2$ is the power spectrum of the fluctuations, which is often written simply as P(k):

$$\langle\Delta^2\rangle = \frac{V}{(2\pi)^3}\int |\Delta_k|^2 d^3k = \frac{V}{(2\pi)^3}\int P(k)d^3k$$

The perturbation spectrum II

- Since the correlation function is spherically symmetric, the volume element of **k**-space can be written as $d^3k=4\pi k^2 dk$.
- The final step is to relate $<\Delta^2>$ to the two-point correlation function:

$$\Delta(\vec{x}) = \sum_k \Delta_k e^{-i\vec{k}\cdot\vec{x}} \Rightarrow \xi(r) = \left\langle \sum_k \sum_{k'} \Delta_k \Delta^*_{k'} e^{-i(\vec{k}-\vec{k'})\cdot\vec{x}} e^{i\vec{k'}\cdot\vec{r}} \right\rangle$$

- Multiplying out the cross terms in this summation, they all vanish except for those which **k=k'**:

$$\xi(r) = \sum |\Delta_k|^2 e^{i\vec{k}\cdot\vec{r}} = \frac{V}{(2\pi)^3} \int |\Delta_k|^2 e^{i\vec{k}\cdot\vec{r}} d^3k$$

- Finally we notice that $\xi(r)$ is a real function and we are only interested in the real part of $e^{i\mathbf{k}\cdot\mathbf{r}}$, that is $\cos(\mathbf{k}\cdot\mathbf{r})=\cos(kr\cos\theta)$

The perturbation spectrum III

- Performing the integral over an isotropic probability distribution of angles θ on a sphere, that is we integrate cos(kr cos θ) over (½)sin θdθ gives the final result:

$$\xi(r) = \frac{V}{2\pi^2}\int |\Delta_k|^2 \frac{\sin kr}{kr} k^2 dk = \frac{V}{2\pi^2}\int P(k)\frac{\sin kr}{kr} k^2 dk$$

- The function (*sin* kr)/kr allows only wavenumbers $k \leq r^{-1}$ to contribute to the amplitude of fluctuations on the scale r. Fluctuations with larger wavenumbers corresponding to smaller scales, average out to zero on the scale r.
- The inverse relation for the power spectrum can also be derived:

$$P(k) = \frac{1}{V}\int_0^\infty \xi(r)\frac{\sin kr}{kr} 4\pi r^2 dr$$

6.2 The initial power spectrum I

- The smoothness of the two-point correlation suggest that the spectrum of initial fluctuations must have been very broad with no preferred scales and it is natural to begin with a power law form:

$$P(k) = |\Delta_k|^2 \propto k^n \Rightarrow \xi(r) \propto \int \frac{\sin kr}{kr} k^{(n+2)} dk$$

- Because of the (*sin*kr)/kr has value unity for kr<<1 and decreases rapidly to zero when kr>>1. We can integrate *k* from 0 to k_{max}≈1/r to estimate the dependence of the amplitude of the correlation function on the scale r: $\xi(r) \propto r^{-(n+3)}$
- Since the mass of the fluctuation is proportional to r^3, this result can also be written in terms of the mass within the fluctuations on the scale r, M~ρr^3: $\xi(M) \propto M^{-(n+3)/3}$

The initial power spectrum II

- To relate ξ to the root-mean-square density fluctuation on the mass scale M, Δ(M), we take the square root of ξ:

$$\Delta(M) = \frac{\delta\rho}{\rho}(M) = \langle\Delta^2\rangle^{1/2} \propto M^{-(n+3)/6}$$

- Before the perturbations came through their particle horizons and before the epoch of equality of matter and radiation energy densities, the density perturbations grew as $\Delta(M) \propto a^2$, as the perturbations of the gravitational potential were frozen-in (see end of Lecture 4).
- Therefore, the development of the spectrum of density perturbations can be written as:

$$\Delta(M) \propto a^2 M^{-(n+3)/6}$$

The initial power spectrum III

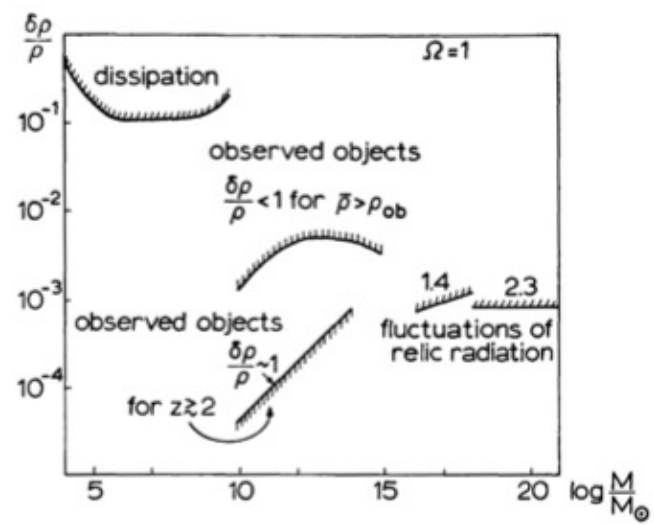
- A perturbation of scale r came through the horizon when r≈ct, and so the dark matter mass within it was $M_D \approx \rho_D (ct)^3$. During the radiation-dominated phases $a \propto t^{1/2}$ and the number density of dark matter particles, which will eventually form bound structures, varied as $N_D \propto a^{-3}$.
- Therefore, the horizon dark matter mass increased as $M_H \propto a^3$ or $a \propto M_H^{1/3}$. The mass spectrum $\Delta(M_H)$ when the fluctuations came through the horizon at different cosmic epochs was

$$\Delta(M) \propto M^{2/3} M^{-(n+3)/6} = M^{-(n-1)/6}$$

- Thus, if n=1, the density perturbations $\Delta M = \delta\rho/\rho(M)$ all had the same amplitude when they came through their particle horizons during the radiation-dominated era.

The Harrison-Zeldovich power spectrum

- The rather special value of n=1 is known as the Harrison-Zeldovich spectrum. In particular Sunyaev and Zeldovich found that in order to produce the observed structure today, the initial fluctuations had to have a scale-invariant spectrum (i.e. n=1) $\delta\rho/\rho=10^{-4}$ on mass scales 10^5 -10^{20} $M_\odot$.
- Harrison studied the form the primordial spectrum must have in order to prevent the overproduction of excessively large amplitude perturbations on small and large scales. A power spectrum of the form P(k) $\propto$k does not diverge on large physical scales and so is consistent with the observed large-scale isotropy of the Universe.
- The inflation model predicts n slightly below 1, as inflation ends at slightly different times in different regions. The Planck satellite measured n=0.9649$\pm$0.042, over 8σ away from n=1.



6.3 Transfer functions: Processing of the initial power spectrum

- Today we do not observe the initial power-spectrum except on the largest physical scales. The transfer function T(k) describes how the shape of the initial power-spectrum $\Delta_k(z)$ in the dark matter is modified by different physical processes through the relation:

$$\Delta_k(z=0) = T(k) f(z) \Delta_k(z)$$

- $\Delta_k(z=0)$ is the power spectrum at the present epoch and $f(z) \propto a \propto t^{2/3}$ is the linear growth factor between the scale factor at redshift z and the present epoch in the matter-dominated era.
- The form of the transfer function is largely determined by the fact that there is a delay in the growth of the perturbations between the time when the perturbations came through the horizon and began to grow again. For example in the CDM picture, the oscillations in the photon-baryon plasma were dynamically more important than those in dark matter before the epoch of matter-radiation equality.

Adiabatic cold dark matter I

- We adopt a standard power-law power spectrum: $P(k)=|\Delta_k|^2 \propto k^n$.
- Before the perturbations entered the horizon during the radiation-dominated era, their density contrast grew as $\Delta_k \propto a^2$ on all scales. If the perturbations came through the horizon during the radiation-dominated phase, the dark matter perturbations were gravitationally coupled to the radiation-dominated plasma and their amplitudes were stabilized (Mézaros effect, see Problem sheet 3). After matter-radiation equality, all perturbations grew as $\Delta_k \propto a$. The net result is a flattening of the perturbations on small scales:

$$\Delta_k = \left(\frac{\delta\rho}{\rho}\right)_k \propto k^{n/2}\left(\frac{a_H}{a_{\rm eq}}\right)^2 \begin{cases} T_k = 1, & M \geq M_{\rm eq}, k \leq k_{\rm eq} \\ T_k \propto k^{-2}, & M \leq M_{\rm eq}, k \geq k_{\rm eq} \end{cases}$$

Adiabatic cold dark matter II and Adiabatic hot dark matter

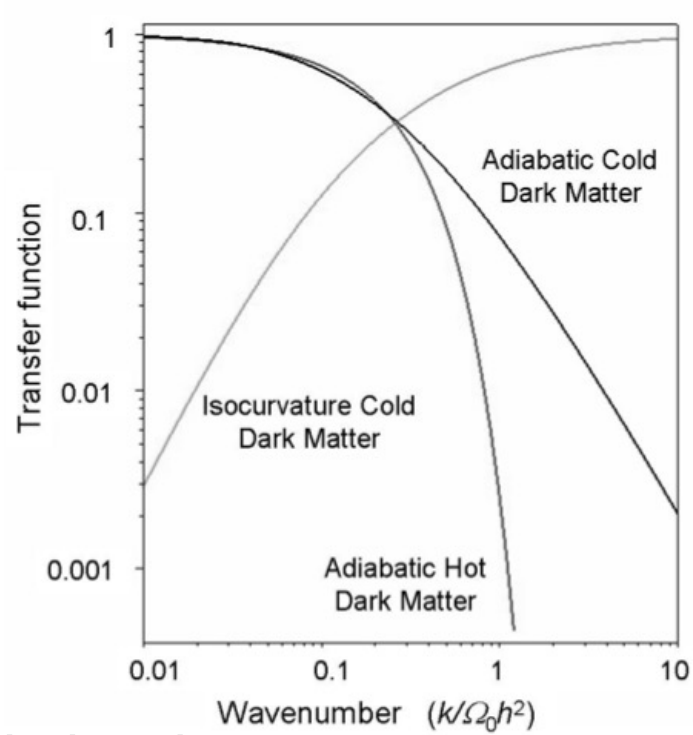
- Detailed calculations of the transfer functions have to be carried out using the full apparatus of the coupled Boltzmann and Einstein field equations.
- It is traditional to provide convenient analytic formulae, which accurately describe the general form of the transfer function. For the adiabatic CDM model the form by Bardeen et al. can be used:

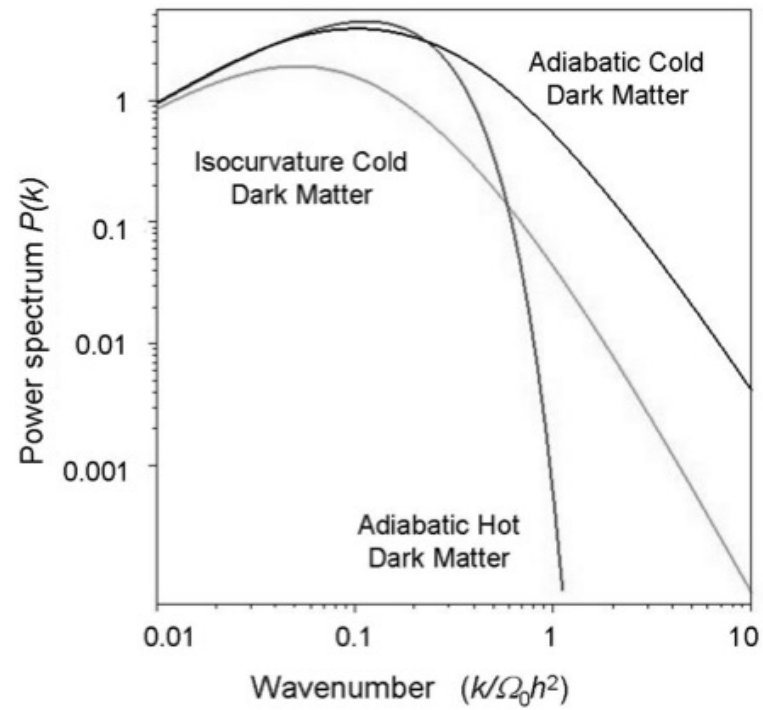
$$T_k = \frac{\ln(1+2.34q)}{2.34q}\left[1+3.89q+(16.1q)^2+(5.46q)^3+(6.71q)^4\right]^{-1/4}, \; q = k/(\Omega_0 h^2 \mathrm{Mpc}^{-1})$$

- In the case of adiabatic hot dark matter model with massive neutrinos, small-scale perturbations are damped by free-streaming of neutrinos. The spectrum cuts of exponentially below $M_{FS} \sim 10^{15}\, M_\odot$

$$T_k = e^{-3.9q-2.1q^2}, \; q = k/(\Omega_0 h^2 \mathrm{Mpc}^{-1})$$

Transfer functions and the resulting power spectra





- Notice that on very large scales (small wavenumbers) the spectrum is unprocessed. On the scales of galaxies and clusters, the spectrum has been strongly modified.

Subsequent evolution

- Linear growth of the spectrum of density fluctuations continue until the perturbations become non-linear, Δ_D~1. Perturbations with the largest amplitudes attained Δ_D~1 first and then began to collapse to form bound systems. The spectrum of density fluctuations was given by:
$$\Delta_k \propto [k^3 P(k)]^{1/2}$$
- In the case of adiabatic cold dark matter models perturbations on small scales became non-linear first and collapsed to form the low-mass bound systems.
- In the hot dark matter picture, elongated and flattened structures were formed very effectively, in fact too effectively, as everything collapsed into large clusters.

Biasing I

- So far we have implicitly assumed that the visible parts of galaxies trace the distribution of dark matter. However this might not always be the case and the generic term for this is called biasing:

$$\xi_{\rm gal}(r) = b^2 \xi_D(r)$$

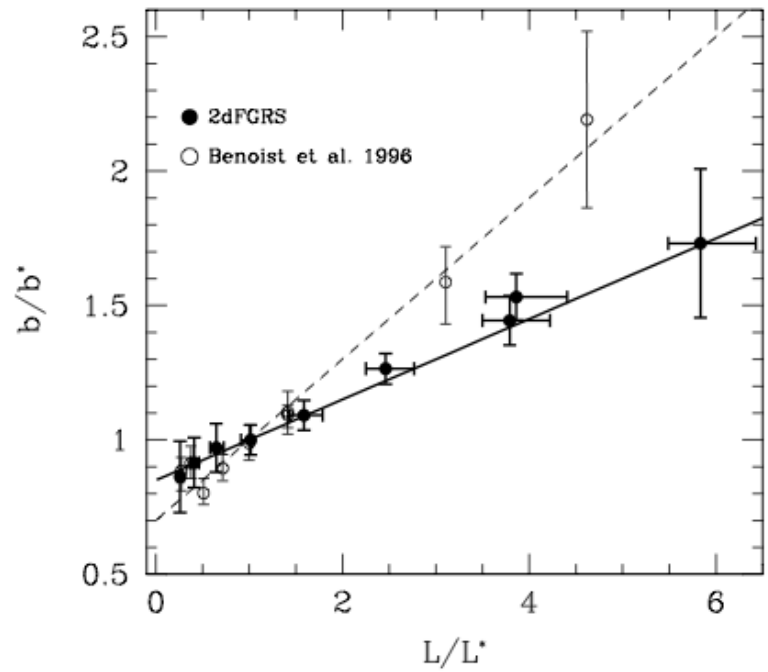
$$P_{\rm gal}(k) = b^2 P_D(k) \qquad \left(\frac{\delta\rho}{\rho}\right)_{\rm gal} = b\left(\frac{\delta\rho}{\rho}\right)_{\rm D} \qquad \Delta_{\rm gal} = b\Delta_{\rm D}$$

- Here *b* is the bias-parameter, if the b>1 galaxies are a biased representation of the underlying dark matter.
- This parameter was especially important in early cosmological models, where the observations of galaxies had to be reconciled with Ω_0=1 (predicted by inflation theory). Now that we know that Ω_0=0.3 recent analysis have shown that b=1.04±0.11 (consistent with b=1).

Biasing II

- Although the distribution of galaxies overall is unbiased on large scales, this does not exclude the possibility that there is bias on small scales or for different classes of galaxies, which must be the case for the morphology-clustering relation to be valid.
- The bias parameter varies with galaxy luminosity in the sense that the most luminous galaxies are more strongly correlated than galaxies in general:

Norberg et al. (2001)



$$b/b^* = 0.85 + 0.15 L/L^*$$

6.4 Reconstructing the processed initial power spectrum

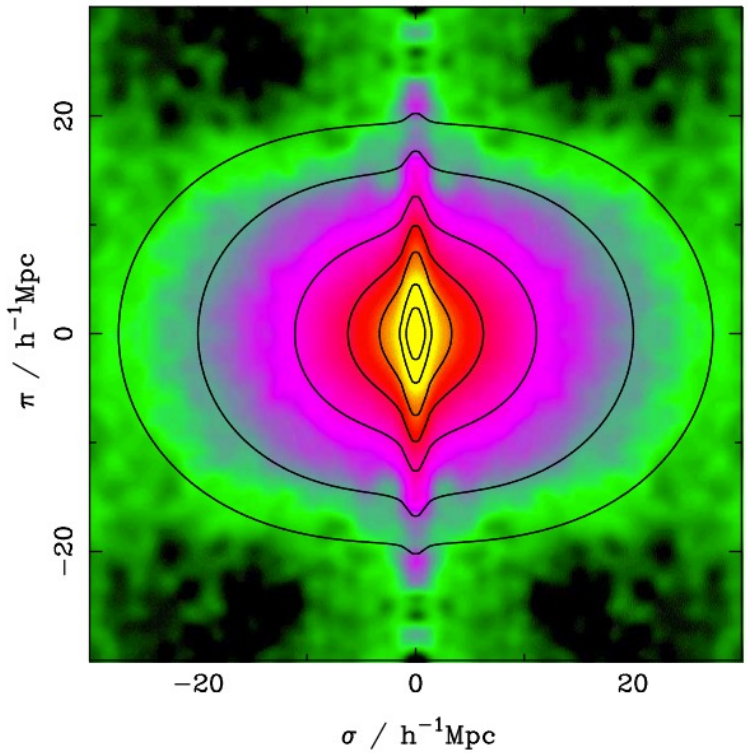
- We are now in a position to attempt to invert the observational data and to determine the form of the processed initial power spectrum and compare it with predictions of models such as those illustrated on page 16 of this lecture.
- The observational results must be derived using large galaxy surveys and we will here utilize the results of the 2dF Galaxy redshift survey that measured the redshifts of 221 414 galaxies and the Sloan Digital Sky Survey from which we use their data on 46 748 luminous red galaxies in the redshift range z=0.16-0.47.
- A key factor in the analysis of the data is to first quantitatively account for the many selection effects, which are inevitably present.

Redshift biases

- Distances are typically measured from the redshift of galaxies using the Hubble law, but these distance estimates must be corrected.

1. ‘Fingers of God’ effect alter the cosmological redshift of galaxies in clusters due to their peculiar velocities. In this way the galaxies are displaced from their true positions.
2. Large-scale density perturbations induce potential motions and as a result galaxies are observed falling into large-scale density perturbations.

Peacock et al. (2001)

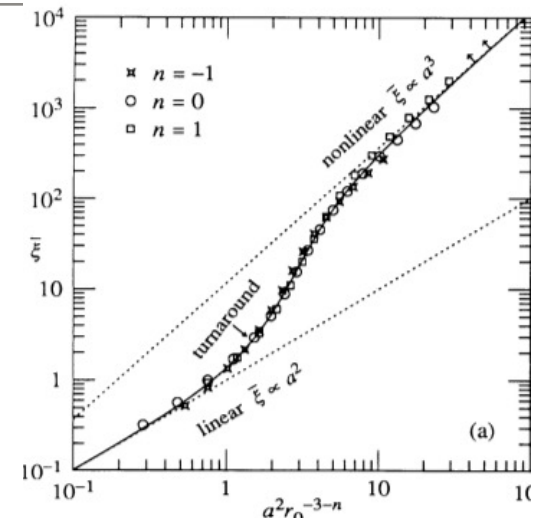


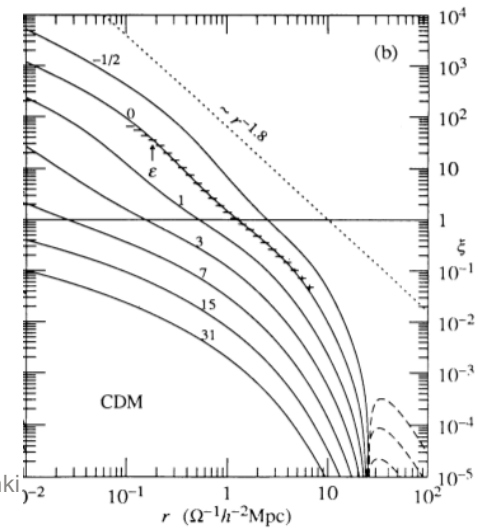
The 2D correlation function for galaxies.

Non-linear development of density perturbations

Hamilton et al. (1991)

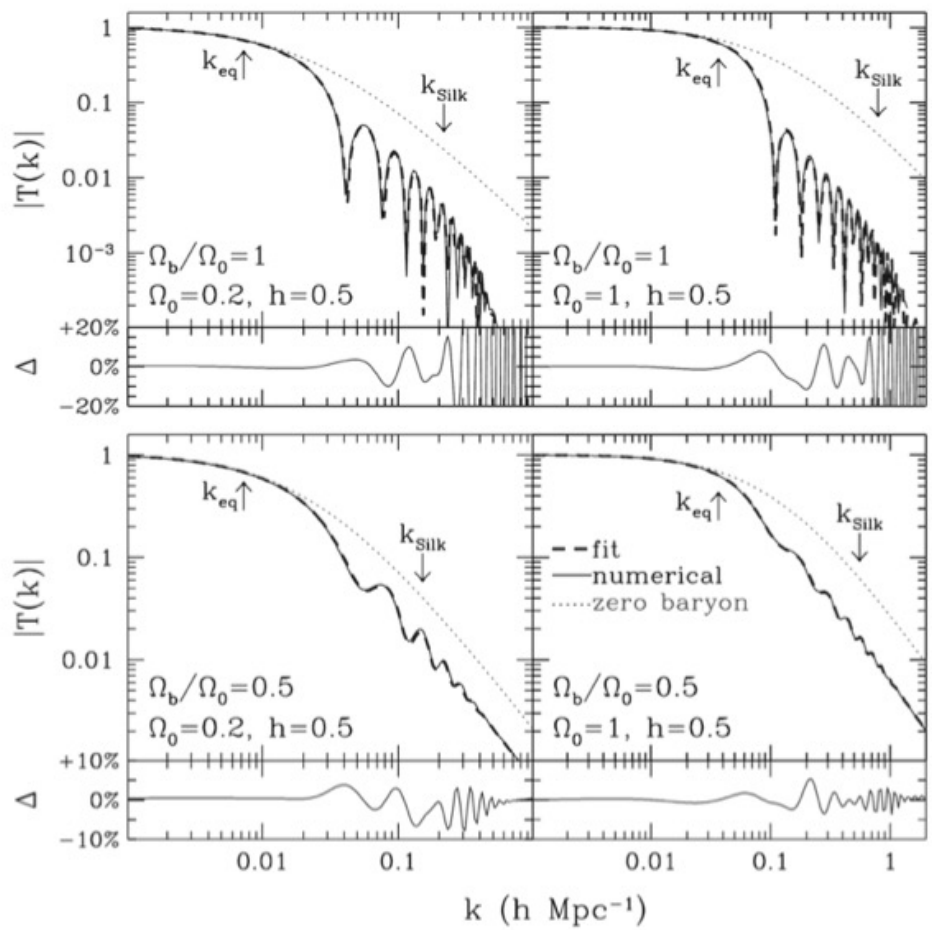
- It is evident from the power-law form of the two-point correlation function that on scales much larger than the characteristic length scale r_0, the perturbations are still in the linear stage of development.
- For smaller scales than r_0 the perturbations have become non-linear. In the linear matter dominated regime $\delta\rho/\rho \propto a$, so $\xi \propto a^2$. In the highly non-linear regime clustering becomes stationary, so that at fixed radius ξ grows as $\xi \propto a^3$.
- The plots show the evolution of ξ as a function of the square of the scale factor (top) and radius for different redshifts, as indicated in the figure (bottom).





The role of baryon perturbations

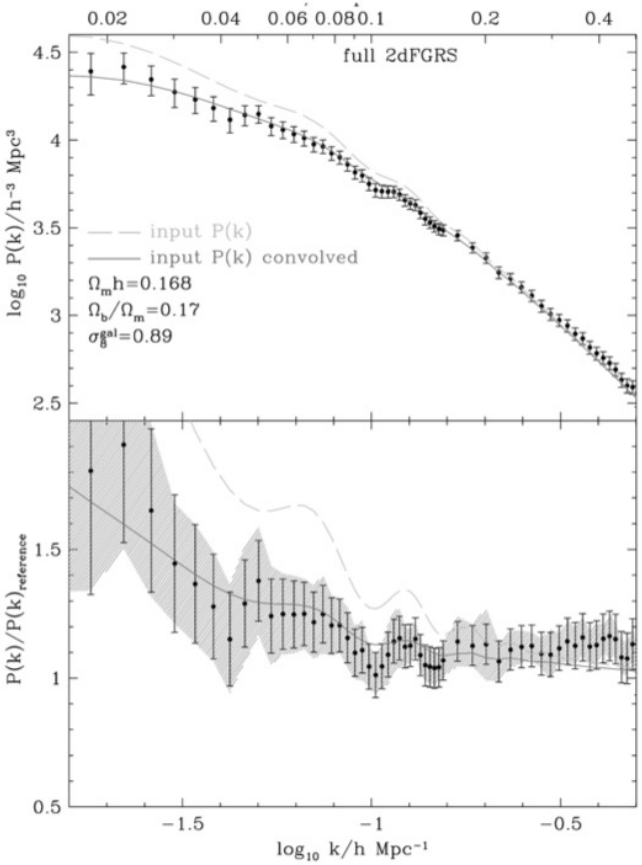
- Four examples of the transfer function for models with baryons only (top pair of diagrams) and with mixed cold and baryonic models (bottom pair of diagrams)
- The numerical results are shown as solid lines and their fitting functions by dashed lines.
- The baryons, which make up ~20% of the total mass leave imprints upon the galaxy power spectrum.



Acoustic peaks in the power spectrum of galaxies: 2dF

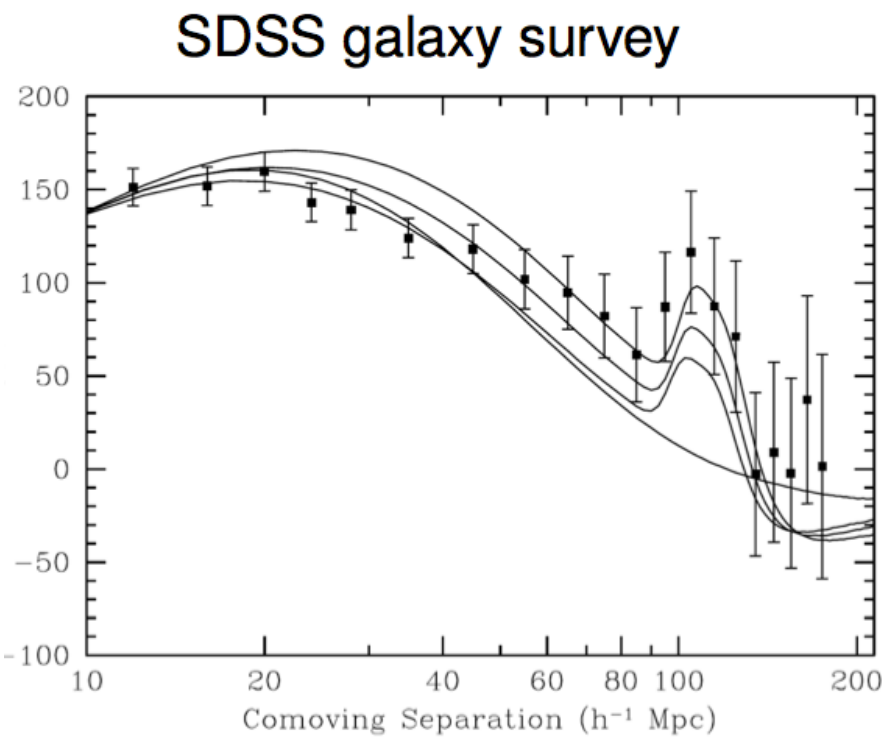
- The baryonic oscillations from the Cosmic Microwave background can be observed in the present galaxy power spectrum as Baryonic Acoustic Oscillations (BAOs) in the sense that there is excessive power on the scales corresponding to BAO scale.
- In the 2dF data the first and second peaks have been detected at wavenumbers of k=0.06 and k=0.12 corresponding to the first peak at $100h^{-1}$ and the second peak at $60h^{-1}$ Mpc.

AAT 2dF galaxy survey



Acoustic peaks in the power spectrum of galaxies: SDSS

- The same result is confirmed by the observations using the SDSS.
- The key use of BAOs in cosmological parameter estimations is that the fluctuation scale can be considered to be a standard rod of a known size, by measuring the BAO scale as a function of redshift we can determine the cosmological parameters.
- The galaxy surveys show that Ω_0=0.24-0.27 irrespective of the CMB results.



The bottom smooth line shows a DM only model with no BAOs.

What have we learned?

1. The two-point correlation function describes the excess probability of finding a galaxy at distance *r* from a galaxy selected at random over that expected in a uniform, random distribution.
2. The initial power spectrum of perturbations was a smooth power law with an index close to the scale-invariant Harrison-Zeldovich model, n=1.
3. The transfer function describes how the initial perturbation spectrum is modified until the present-day due to the combined baryonic and dark matter processes.
4. The initial perturbation spectrum can be reconstructed from careful observations. The perturbation spectrum still bears the imprints of the baryonic acoustic oscillations from the formation of the CMB.