Cosmological large-scale structure

Lecture 2

L. Amendola

WS2024

The structure of the large scale structure

large scales

small scales

super-horizon scales	linear scales	mildly non-linear scales	strongly non-linear scales
relativistic corrections	*linear pert. theory*	*non-linear pert. theory*	*N-body simulations*

$k \approx aH \approx$0.002 h/Mpc

$\lambda \approx 3000\ Mpc/h$

$k \approx 0.1\ h/Mpc$

$\lambda \approx 60\ Mpc/h$

$k \approx 0.3\ h/Mpc$

$\lambda \approx 20\ Mpc/h$

Recap

Collisionless Boltzmann equation

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_{n=1}^{6} \frac{\partial w_n}{\partial t}\frac{\partial f}{\partial w_n} = \frac{\partial f}{\partial t} + \frac{\partial q_i}{\partial t}\frac{\partial f}{\partial q_i} + \frac{\partial p_i}{\partial t}\frac{\partial f}{\partial p_i} = 0$$

Vlasov-Poisson equation

$$\frac{\partial f}{\partial t} + \frac{\mathbf{p}}{m}\cdot\nabla f - m\nabla\Phi\cdot\frac{\partial f}{\partial \mathbf{p}} = 0$$

final set of eqs
in comoving coords
and conf time

$$\dot{\theta} = -\nabla^2 c_s^2 \delta - \nabla^2 \phi - \mathcal{H}\theta$$
$$\dot{\delta} = -\theta$$

Fourier space

$$\begin{aligned}\dot{\delta} &= -\theta \\ \dot{\theta} &= -\mathcal{H}\theta + c_s^2 k^2 \delta + k^2 \phi \\ k^2\phi &= -\frac{3}{2}\mathcal{H}^2\Omega_m \delta\end{aligned}$$

Recap

Master equation

$$\ddot{\delta} + \mathcal{H}\dot{\delta} + \left(k^2 c_s^2 - \frac{3}{2}\mathcal{H}^2\Omega_m\right)\delta = 0$$

For pressureless matter

$$\delta'' + \frac{1}{2}\delta' - \frac{3}{2}\delta = 0$$

$$\delta_+ = Aa^1, \quad \delta_- = Ba^{-3/2}$$

For ΛCDM

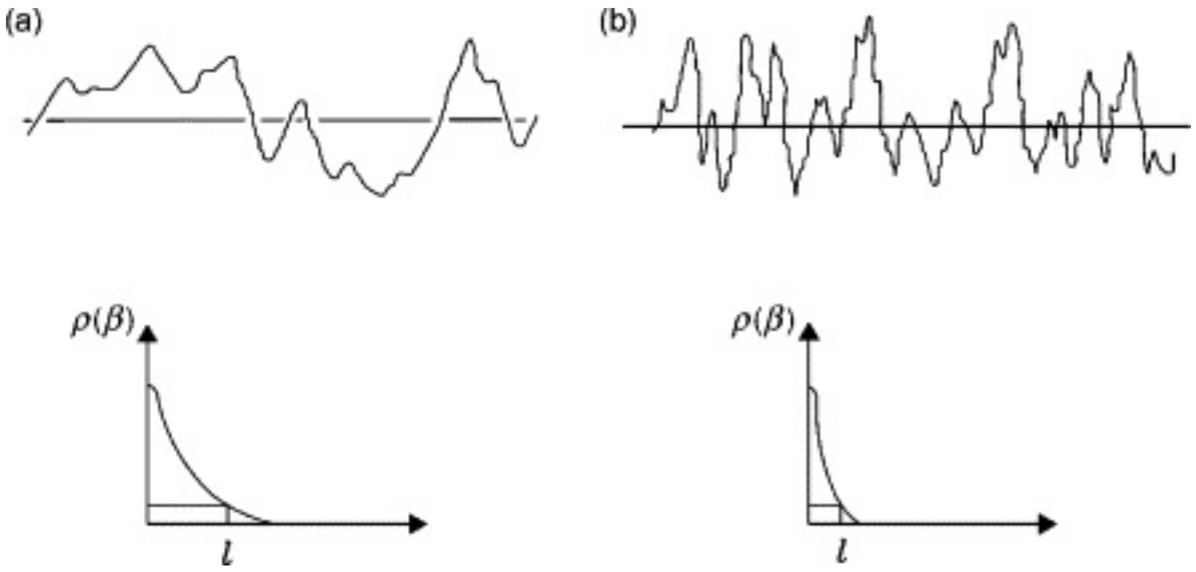
$$f \equiv \frac{d\log\delta_m}{d\log a} \approx \Omega_m^\gamma(z)$$

$$\gamma \approx 0.55$$

Roadmap for today

- Measures of clustering of a distribution of points.
- The correlation function describes the clustering of a distribution of points in space.
- The power spectrum is the Fourier conjugate of the correlation function
- Correlation function and power spectrum are two-point descriptors.
- Generalization to n-point descriptors.

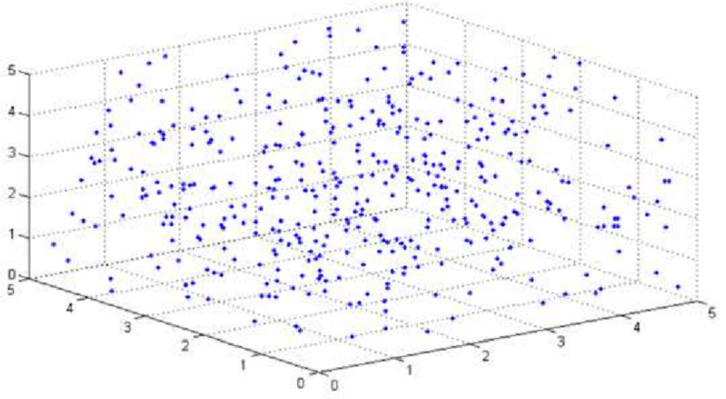
Correlation of continuous random fields in 1D



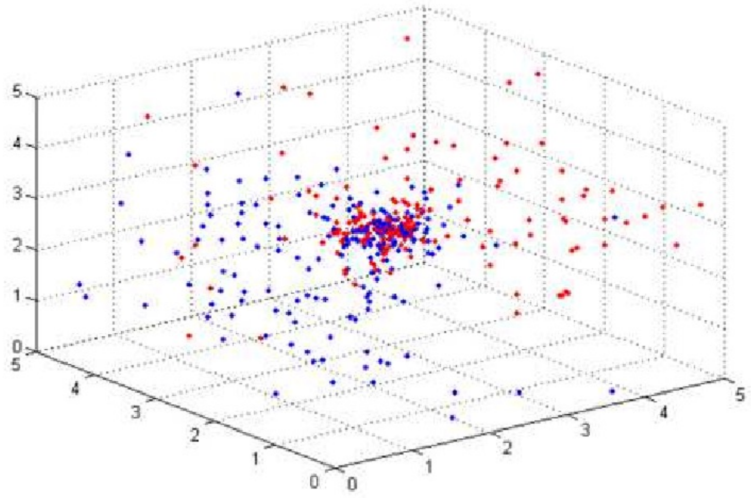
Correlation up to large scales

Correlation up to small scales

Correlation of discrete random fields in 3D



random distribution



clustered distribution

Definition of moments

$$\int f(x)dx = 1$$

$$\langle x \rangle = \int x f(x) dx$$

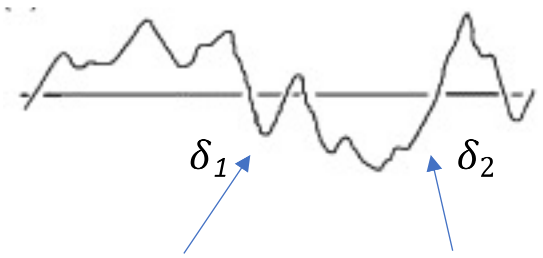
$$\langle x^2 \rangle = \int x^2 f(x) dx$$

moment $$\bar{M}_n = \int x^n f(x) dx$$

central moment $$M_n = \int (x - \langle x \rangle)^n f(x) dx$$

Definition of moments for multivariate distributions

random field $\delta(x)$



every point is a random variable!

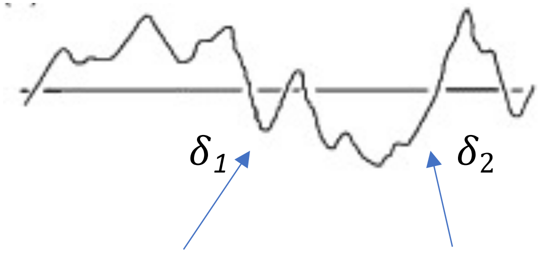
multivariate distribution $f(\delta_1, \delta_2 \ldots)$

prob. that δ_1 is in d δ_1
δ_2 in $d\,\delta_2$, etc $f(\delta_1, \delta_2 \ldots) d\,\delta_1\, d\,\delta_2 \ldots$

δ_i is independent
(uncorrelated)
of δ_j only if $f(\delta_1, \delta_2 \ldots) = f_1(\delta_1) f_2(\delta_2) \ldots .$

Definition of moments for multivariate distributions

random field $\delta(x)$



every point is a random variable!

first moment (mean vector)

$$\langle \delta_i \rangle = \delta_{0i} = \int \delta_i f(\delta_1, \delta_2 \ldots) d^n \delta$$

second central moment (covariance matrix)

$$c_{ij} = \int (\delta_i - \delta_{0i})(\delta_j - \delta_{0j}) f(\delta_1, \delta_2 \ldots) d^n \delta$$

if the variables are independent, the covariance becomes diagonal

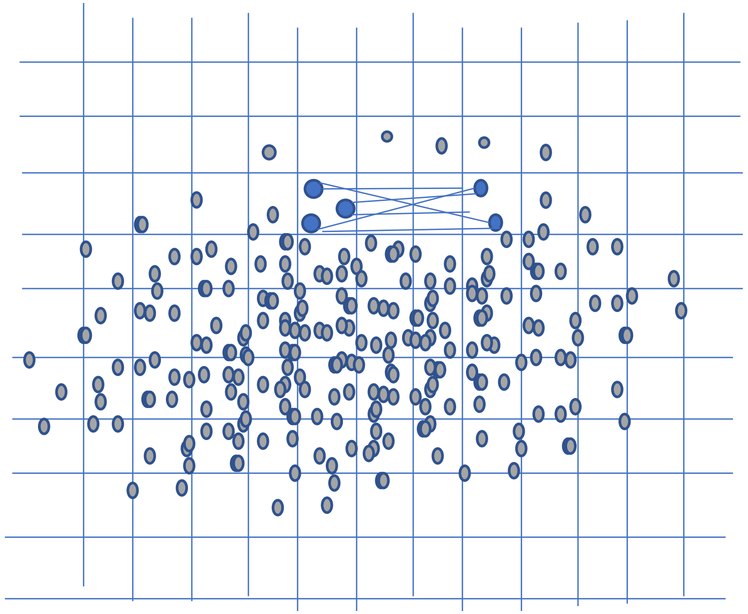
$$c_{ij} = \int (\delta_i - \delta_{0i}) f_i(\delta_i) d\delta_i \int (\delta_j - \delta_{0j}) f_j(\delta_j) d\delta_j \prod_m^{n-2} \int f_m(\delta_m) d\delta_m = 0 \quad i \neq j$$

$$c_{ij} = \int (\delta_i - \delta_{0i})^2 f_i(\delta_i) d\delta_i \prod_m^{n-1} \int f_m(\delta_m) d\delta_m = \int (\delta_i - \delta_{0i})^2 f_i(\delta_i) d\delta_i = \sigma_i^2 \quad i = j$$

Correlation of random fields

3 particles in cell *a*
2 particles in cell *b*

3x2 = 6 pairs



Average number in cells of volume dV $\langle n \rangle = \rho_0 dV$

average pairs
if uncorrelated:

$$\langle n_i n_j \rangle = \int n_i f_i(n_i) d n_i \int n_j f_j(n_j) d n_j \prod_m^{n-2} \int f_m(n_m) d n_m = \langle n_i \rangle \langle n_j \rangle$$

Average number of pairs
in a random (uncorrelated)distribution

$$\langle n_a n_b \rangle = \langle n_a \rangle \langle n_b \rangle = \rho_0^2 d V_a d V_b$$

Definition of correlation function

Average number in cells dV $\langle n \rangle = \rho_0 dV$

Average pairs in cells a,b $dN_{ab} = \langle n_a n_b \rangle$

Average number of pairs:
definition of corr function $dN_{ab} = \langle n_a n_b \rangle = \rho_0^2 dV_a dV_b (1 + \xi(r_{ab}))$

Number density contrast $\delta(r_a) = n_a/(\rho_0 dV) - 1$

What is its average?

The correlation function
is a second order moment,
i.e. the variance of $\delta(x)$

$$\xi(r_{ab}) = \frac{dN_{ab}}{\rho_0^2 dV_a dV_b} - 1 = \frac{\langle n_a n_b \rangle}{\rho_0^2 dV_a dV_b} - 1 = \langle (\delta_a + 1)(\delta_b + 1) \rangle - 1 = \langle \delta(r_a)\delta(r_b) \rangle$$

Definition of correlation function

The corr function is the covariance of the density contrast

$$\xi(r_{ab}) = \frac{dN_{ab}}{\rho_0^2 dV_a dV_b} - 1 = \frac{\langle n_a n_b \rangle}{\rho_0^2 dV_a dV_b} - 1 = \langle (\delta_a + 1)(\delta_b + 1) \rangle - 1 = \langle \delta(r_a)\delta(r_b) \rangle$$

In practice, easier to use a conditional probability

$$dN_b = dN_{ab}/dN_a = \rho_0^2 dV_a dV_b (1 + \xi(r_{ab}))/dN_a = \rho_0 dV_b (1 + \xi(r_b))$$

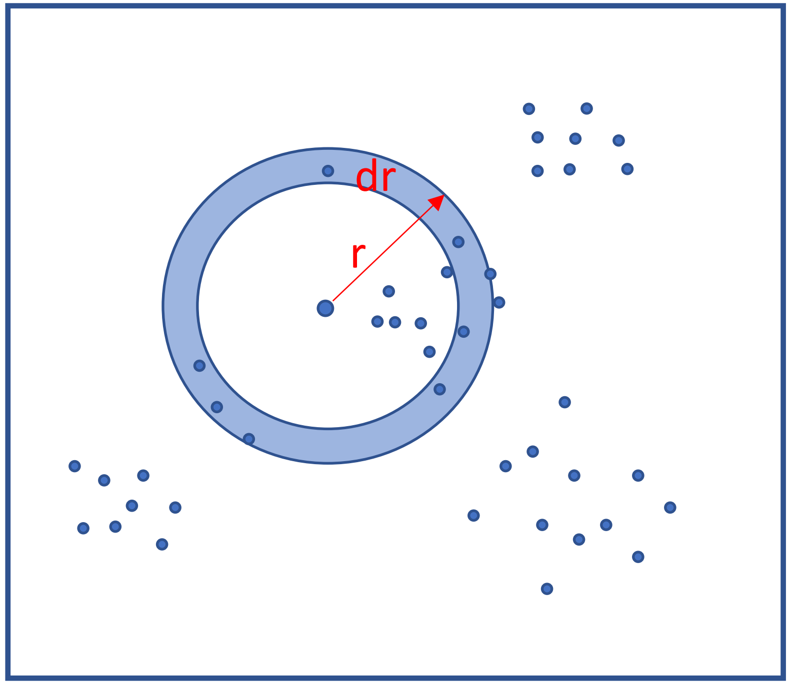
The corr function is also the average number of neighbors!

$$\xi(r) = \frac{dN_c(r)}{\rho_0 dV} - 1 = \frac{\langle \rho_c \rangle}{\rho_0} - 1$$

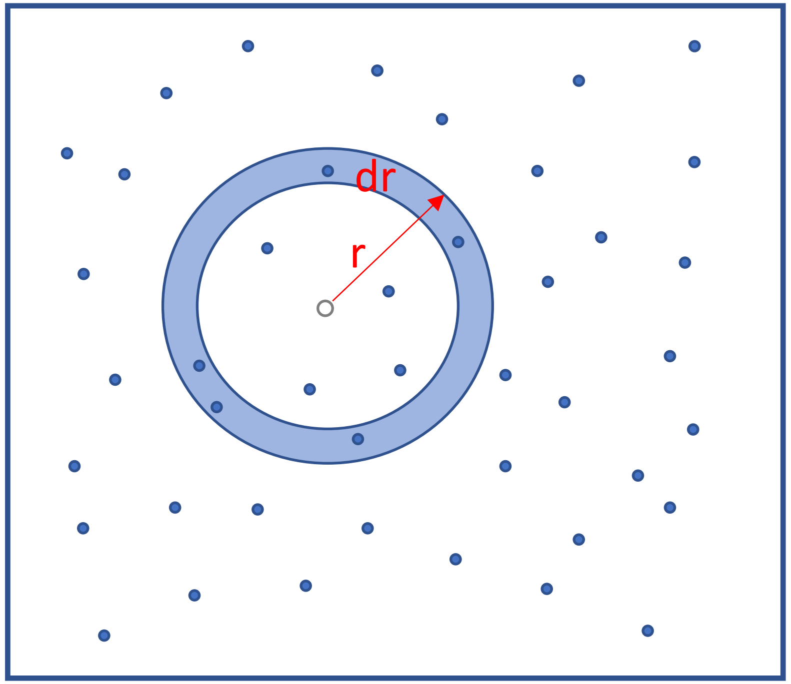
Correlation function in a real data catalog

$$\xi = \frac{DD}{DR} - 1$$

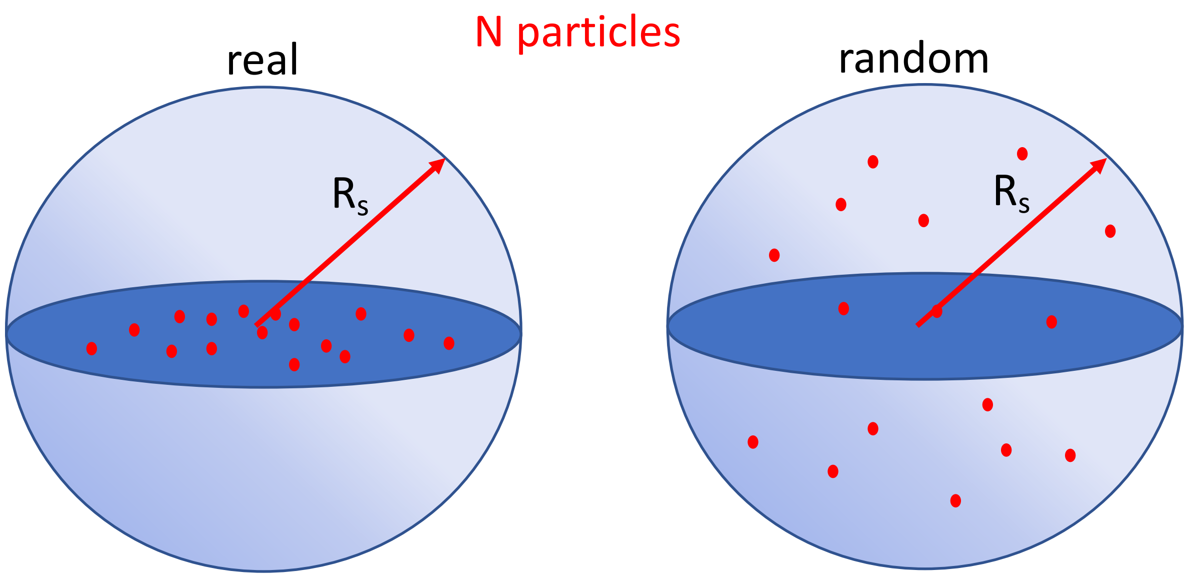
real data: DD

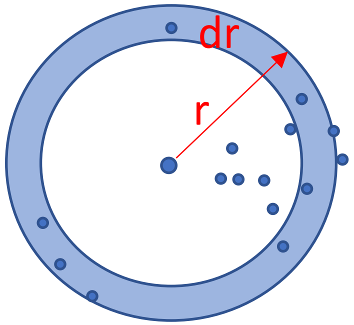


random simulation: DR



Correlation function for a pancake





$$DD = \text{superf. density} \times 2\pi r dr = \frac{N}{\pi R_s^2} 2\pi r dr$$

In the uniform world we have

$$DR = \text{density} \times 4\pi r^2 dr = \frac{3N}{4\pi R_s^3} 4\pi r^2 dr$$

Then we get

$$\xi = \frac{2R_s}{3} r^{-1} - 1$$

Depends only on r,
not on location:
Statistical homogeneity!

Galaxy correlation function

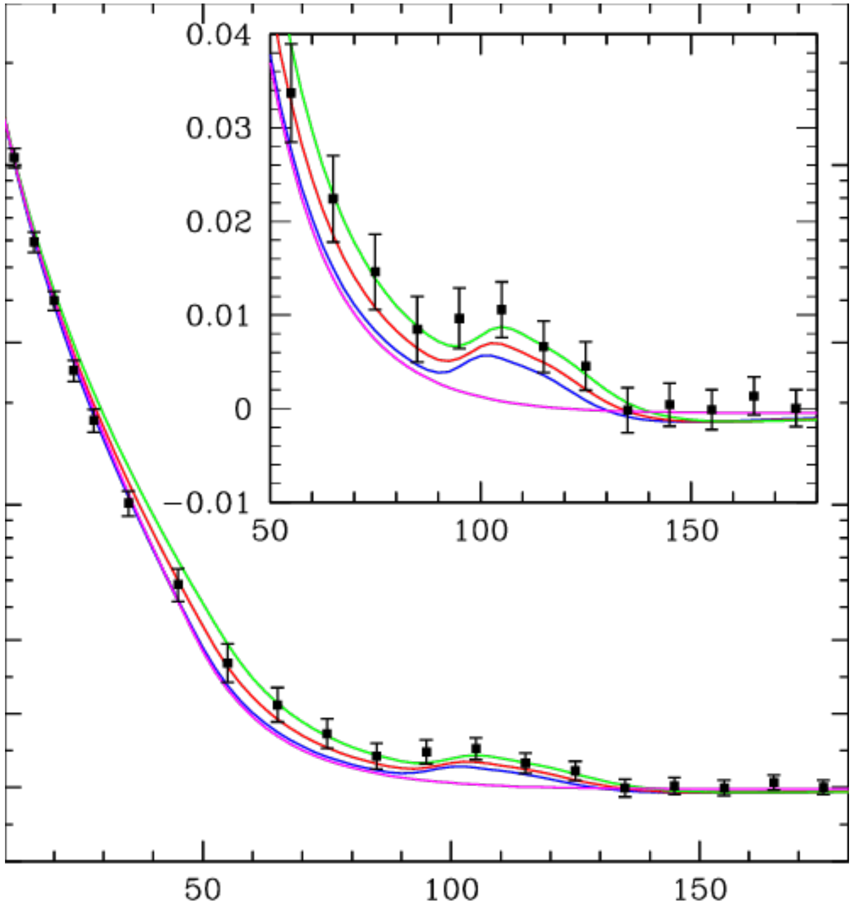
$\xi > 0$ positive correlation: more clustering than random

$\xi = 0$ no correlation: same clustering as random

$\xi < 0$ anticorrelation: less clustering than random

Integral constraint

$$\int_0^R \xi_s(r)dV = N/\rho_0 - V = 0$$



Yun Wang, Observational Probes of Dark Energy
AIP Conference Proceedings

N-point correlation function

Two-point

$$\xi(r_{ab}) = \frac{dN_{ab}}{\rho_0^2 dV_a dV_b} - 1 = \frac{\langle n_a n_b \rangle}{\rho_0^2 dV_a dV_b} - 1 = \langle (\delta_a + 1)(\delta_b + 1) \rangle - 1 = \langle \delta(r_a)\delta(r_b) \rangle$$

Three-point

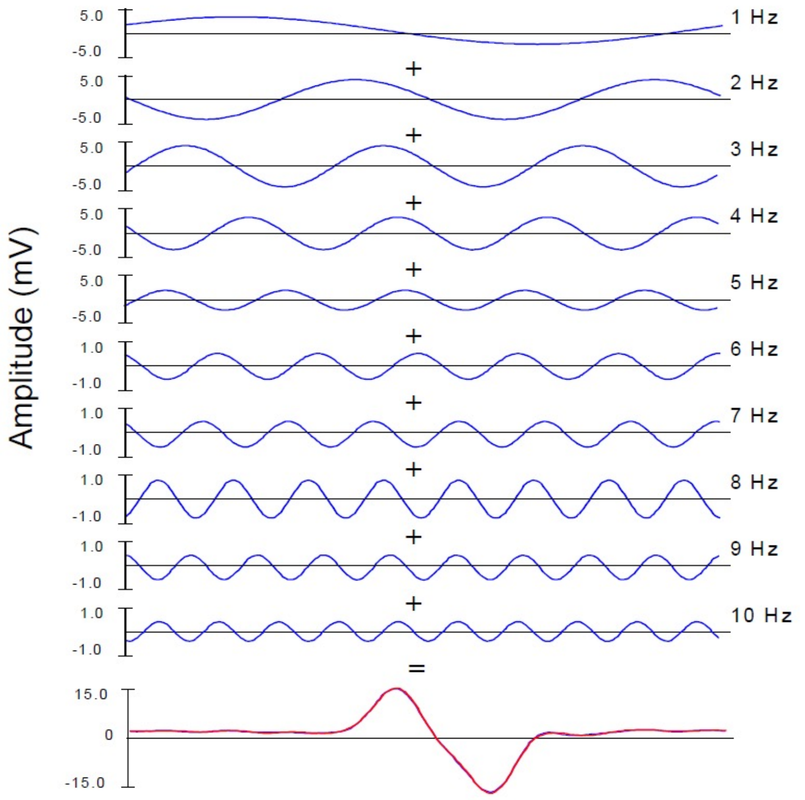
$$\varsigma(r_a, r_b, r_c) = \langle \delta(r_a)\delta(r_b)\delta(r_c) \rangle$$

$$\begin{aligned} \varsigma(r_a, r_b, r_c) &= \langle \left(\frac{n_a}{\rho_0 dV_a} - 1 \right) \left(\frac{n_b}{\rho_0 dV_b} - 1 \right) \left(\frac{n_c}{\rho_0 dV_c} - 1 \right) \rangle \\ &= \frac{\langle n_a n_b n_c \rangle}{\rho_0^3 dV_a dV_b dV_c} - \xi_{ab} - \xi_{bc} - \xi_{ac} - 1 \end{aligned}$$

Fourier space

f(k)exp(-ikx)

f(x)



De Luca, C. J. . (2003). Fundamental concepts in EMG signal acqusition. In D. Inc. (Ed.).

A function decomposed into Fourier modes

Correlation in Fourier space

Fourier convention

$$\begin{aligned} f(x) &= \frac{V}{(2\pi)^3}\int f_k e^{ikx} d^3k \\ f_k &= \frac{1}{V}\int f(x) e^{-ikx} d^3x \end{aligned}$$

Dirac delta

$$\delta_D(\mathbf{k}) = \frac{1}{(2\pi)^3}\int e^{-i\mathbf{kx}} d^3x$$

It diverges for k=0
and goes to zero for k$\neq$ 0

The Power Spectrum

definition
$$P(\mathbf{k}) = \int \xi(r) e^{-i\mathbf{kr}} dV$$

conversely
$$\xi(\mathbf{r}) = (2\pi)^{-3} \int P(k) e^{i\mathbf{kr}} d^3k$$

$$\left.\begin{array}{c} \\ \\ \end{array}\right\} \text{Fourier conjugates}$$

ensemble average
$$V \langle \delta_k \delta^*_{k'} \rangle = \frac{1}{V} \int \langle \delta(y+r) \delta(y) \rangle e^{i(k-k')y + ikr} dV_r dV_y$$

The power spectrum is the amplitude of the second-order moment in Fourier space
i.e. the variance of $\delta(k)$

$$V \langle \delta_k \delta^*_{k'} \rangle = \frac{1}{V} \int \xi(r) e^{i(k-k')y + ikr} dV_r dV_y = \frac{(2\pi)^3}{V} P(k) \delta_D(k-k')$$

Since ξ depends only on **r**,
if $k \neq k'$ the modes are uncorrelated !
(statistical homogeneity)

Correlation in Fourier space:
The power spectrum from the data

Given a distribution in real space, we calculate the Fourier coefficients

$$\delta_k = \frac{1}{V}\int \delta(x) e^{-ikx} dV$$

Definition of power spectrum for a single realization

$$P(k) = V\,\delta_k \delta_k^*$$

$$P(k) = \frac{1}{V}\int \delta(x)\delta(y) e^{-ik(x-y)} dV_x dV_y$$

Correlation in Fourier space: The Power spectrum

$$P(k) = \frac{1}{V}\int \delta(x)\delta(y)e^{-ik(x-y)}dV_x dV_y$$

$$r = x - y;$$

Average over a volume

$$\xi(r) = \langle\delta(y+r)\delta(y)\rangle_V = \frac{1}{V}\int \delta(y+r)\delta(y)dV_y$$

$$\Rightarrow \quad P(\mathbf{k}) = \int \xi(r)e^{-i\mathbf{kr}}dV$$

Conversely

$$\xi(\mathbf{r}) = (2\pi)^{-3}\int P(k)e^{i\mathbf{kr}}d^3k$$

Fourier conjugates

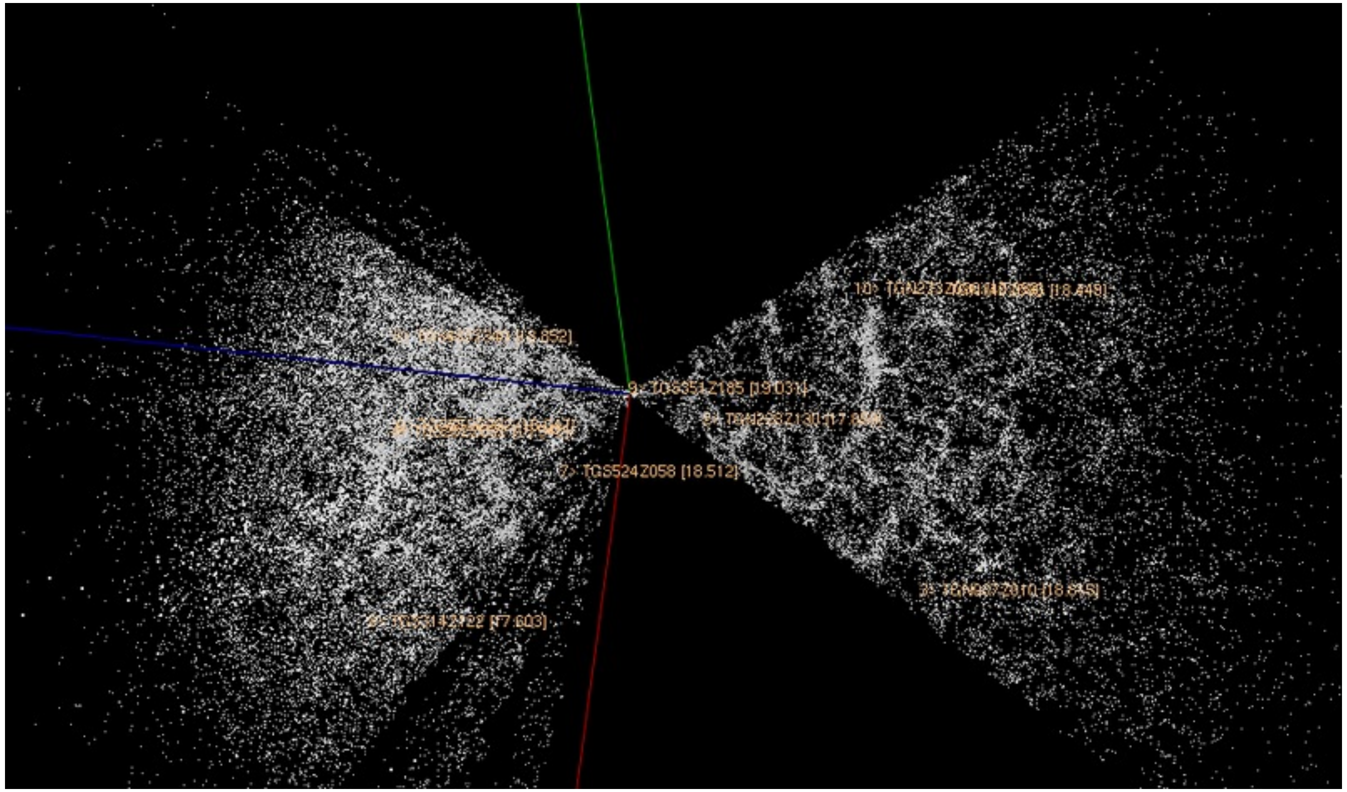
Assuming statistical homogeneity and isotropy

$$P(k) = 4\pi\int \xi(r)\frac{\sin kr}{kr}r^2 dr$$

Quiz time

1. Why two point correlations are so important?
2. What would be the correlation function for filaments?
3. Why is Fourier space so useful?
4. Why we assume Gaussian fields?
5. What could be the complications in real life?

Two complications: Finite size and discreteness



Two complications: Finite size and discreteness

Finite size

Window function

$$W(x) = \text{constant inside the survey}$$
$$W(x) = 0 \quad \text{outside}$$

Conventional normalization

$$\int W(x)dV = 1$$

Inside the survey

$$W(x) = 1/V$$

Redefined fluctuation field

$$\delta_s = \delta(x)VW(x)$$

Two complications: Finite size and discreteness

$$\boxed{\delta_s = \delta(x) V W(x)}$$

Discreteness

normalization

$$\int \rho(\mathbf{x}) d^3x = M$$

$$\sum_i m_i \int \delta_D(\mathbf{x} - \mathbf{x}_i) d^3x = \sum_i m_i = M$$

density field of discrete points

$$\rho(\mathbf{x}) = \sum_i m \delta_D(\mathbf{x} - \mathbf{x}_i)$$

density contrast
of discrete points

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x})}{\rho_0} - 1 = \frac{\sum_i m \delta_D(\mathbf{x} - \mathbf{x}_i)}{\rho_0} - 1 = \frac{V}{N} \sum_i m \delta_D(\mathbf{x} - \mathbf{x}_i) - 1$$

Two complications: Finite size and discreteness

$$\boxed{\delta_s = \delta(x) V W(x)}$$

$$\boxed{\delta(\mathbf{x}) = \frac{\rho(\mathbf{x})}{\rho_0} - 1 = \frac{\sum_i m\delta_D(\mathbf{x} - \mathbf{x}_i)}{\rho_0} - 1 = \frac{V}{N}\sum_i m\delta_D(\mathbf{x} - \mathbf{x}_i) - 1}$$

Density contrast
for a finite-size
discrete distribution

$$\delta_s(x) = \left(\frac{\rho(x)}{\rho_0} - 1\right) V W(x) = \frac{V}{N}\sum_i m_i w_i \delta_D(x - x_i) - V W(x)$$

$$w_i = V W(x_i).$$ (1 inside, 0 outside survey)

Two complications: Finite size and discreteness

$$\boxed{\delta_s(x) = \left(\frac{\rho(x)}{\rho_0} - 1\right) VW(x) = \frac{V}{N}\sum_i m_i w_i \delta_D(x - x_i) - VW(x)}$$

$$\boxed{w_i = VW(x_i).}$$

Fourier transform

Fourier modes
for a finite-size
discrete distribution

$$\delta_k = \frac{1}{V}\int \left(\frac{V}{N}\sum_i m_i w_i \delta_D - VW(x)\right) e^{ikx} dV = \frac{1}{N}\sum_i m_i w_i e^{ikx_i} - W_k$$

window function
In k-space

$$W_k = \int W(x) e^{ikx} dV$$

Two complications: Finite size and discreteness

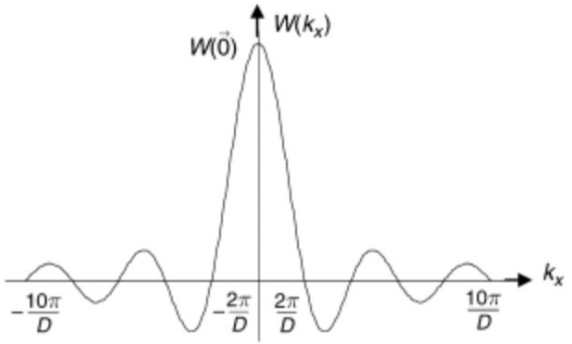
window function
In k-space

$$W_k = \int W(x) e^{ikx} dV$$

Spherical top-hat
window function

$$\begin{aligned} W(x) &= 1/V \text{ inside a spherical volume } V \text{ of radius } R \\ W(x) &= 0 \text{ outside} \end{aligned}$$

$$\begin{aligned} W_k &= \int W(x) e^{i\mathbf{kx}} dV = V^{-1} \int e^{i\mathbf{kx}} dV \\ &= \frac{3}{4\pi} R^{-3} \int^R r^2 dr \int_{-\pi}^{\pi} e^{ikr\cos\theta} d\cos\theta d\phi \\ &= \frac{3}{2} R^{-3} \int^R \left(e^{ikr} - e^{-ikr}\right) \frac{r^2}{ikr} dr \\ &= 3R^{-3} \int^R \frac{r \sin kr}{k} dr = 3 \frac{\sin kR - kR\cos kR}{(kR)^3} \end{aligned}$$



Victor Giurgiutiu, in Structural Health Monitoring, 2008

The window function is only important for scales as large as the survey ($k \to 0$)

Two complications: Finite size and discreteness

Fourier modes for a finite-size discrete distribution

$$\delta_k = \frac{1}{V}\int\left(\frac{V}{N}\sum_i m_i w_i \delta_D - VW(x)\right)e^{ikx}dV = \frac{1}{N}\sum_i m_i w_i e^{ikx_i} - W_k$$

Quiz: where is the random variable?

We don't know the galaxy masses; so we use number counts, i.e. m_i=1

$$P(k) = V\langle\delta_k\delta_k^*\rangle$$

Volume average

$$\langle\frac{1}{N}\sum_i^N w_i e^{ikx_i}\rangle_V = \langle\frac{V}{N}\sum_i^N W(x)e^{ikx_i}\rangle_V = \frac{V}{N}\sum_i^N \frac{1}{V}\int W(x)e^{ikx_i}dV$$

$$= \frac{1}{N}\sum_i^N W_k = W_k$$

Two complications: Finite size and discreteness

$$\delta_k = \frac{1}{V}\int\left(\frac{V}{N}\sum_i m_i w_i \delta_D - VW(x)\right)e^{ikx}dV = \frac{1}{N}\sum_i m_i w_i e^{ikx_i} - W_k$$

Power spectrum
for a finite-size set
of particles

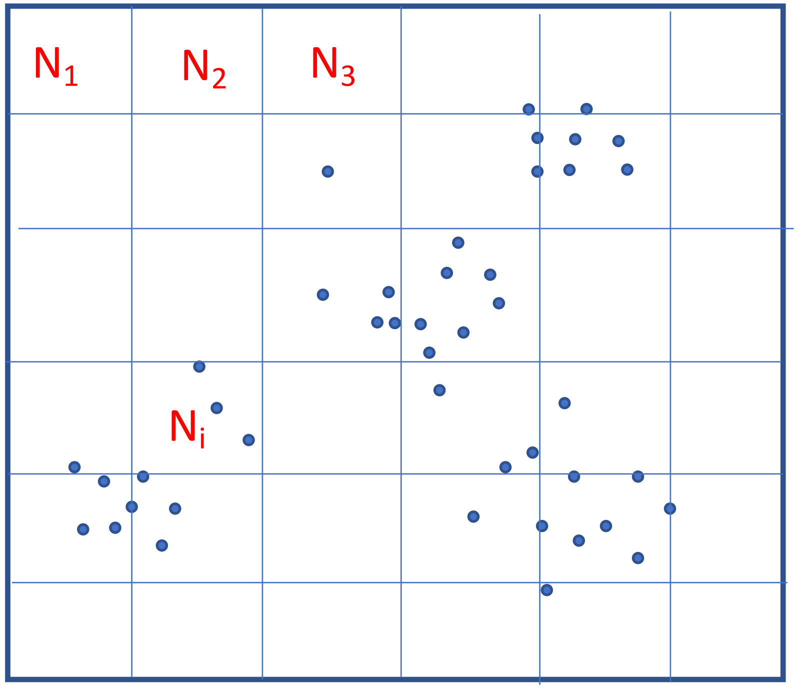
$$P(k) = \frac{V}{N^2}\sum_{ij} w_i w_j \langle e^{ik(x_i - x_j)}\rangle - VW_k^2$$

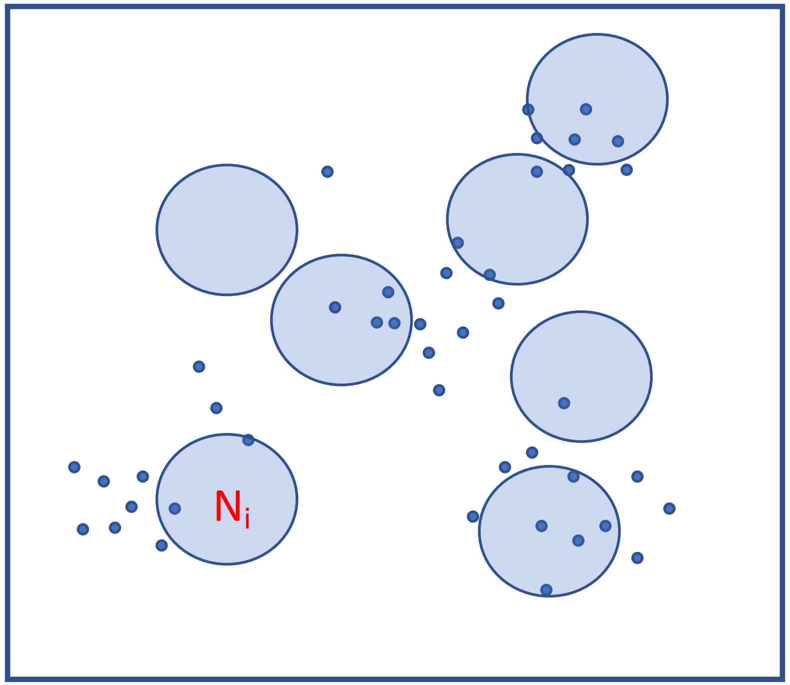
If positions are uncorrelated $\langle e^{ik(x_i - x_j)}\rangle$ vanishes except for $x_i = x_j$

Shot-noise spectrum

$$P_{sn}(k) = \frac{V}{N^2}\sum_i^N w_i^2 = \frac{V}{N^2}\sum_i^N 1 = \frac{V}{N}$$

A related statistic: Counts in Cells





$$m_1 = \langle N_i \rangle = \frac{\sum N_i}{n},$$

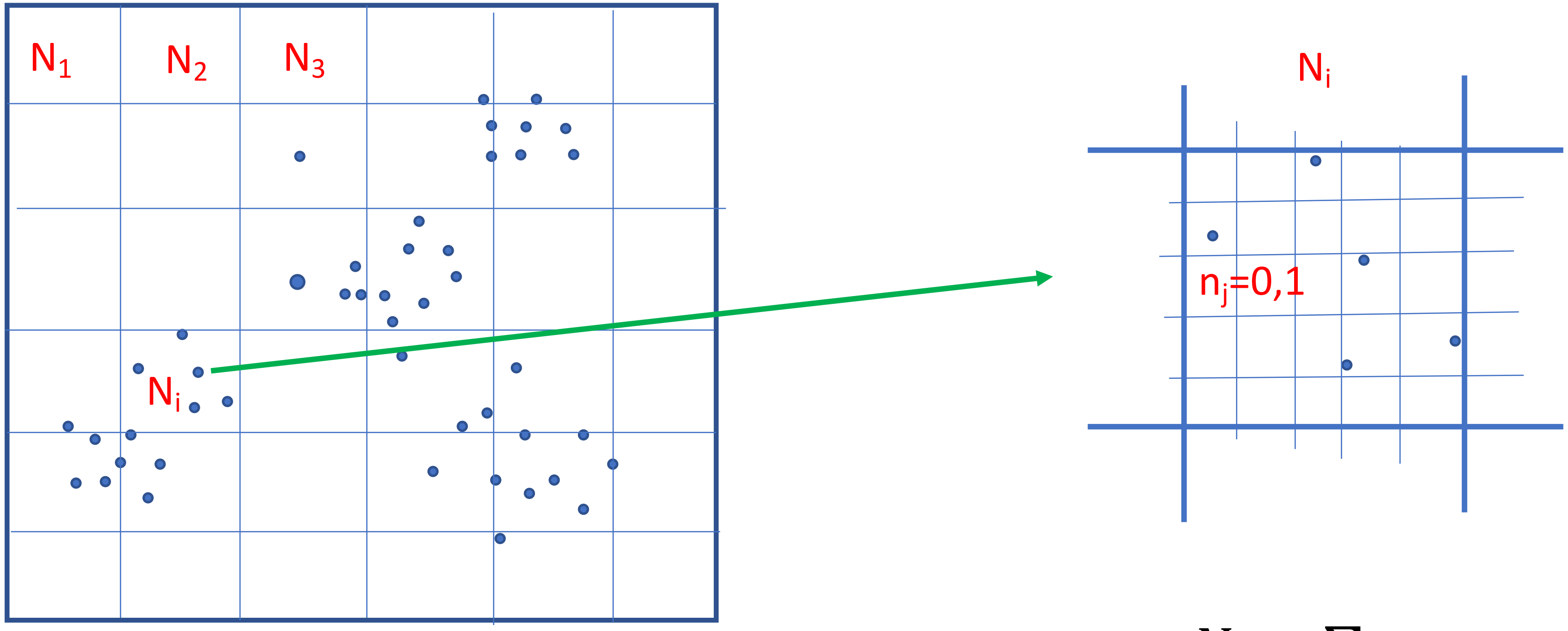
moments

$$m_2 = \langle N_i^2 \rangle = \frac{\sum N_i^2}{n}$$

$$m_{2(c)} = \langle (N_i - m_1)^2 \rangle$$

…

Infinitesimal cells



$$N_i = \sum_j n_j$$

Integral moments

$$\langle N^2 \rangle = \langle \sum n_i \sum n_j \rangle = \sum \langle n_i^2 \rangle + \sum \langle n_i n_j \rangle =$$
$$N_0 + N_0^2 \int dV_i dV_j W_i W_j [1 + \xi_{ij}]$$

$$W_i dV_i = dV_i^*$$

define
$$\sigma^2 = \int dV_1^* dV_2^* \xi_{12}$$

then
$$\langle N^2 \rangle = N_0 + N_0^2 + N_0^2 \sigma^2$$

and the variance is
$$M_2 = N_0^{-2}(\langle N^2 \rangle - N_0^2) = N_0^{-1} + \sigma^2$$

Shot noise

Integral moments

volume average
of correlation function

$$\boxed{\sigma^2 = \int dV_1^* dV_2^* \xi_{12}}$$

$$\sigma^2 = (2\pi)^{-3} \int P(k) e^{i\mathbf{k}(\mathbf{r}_1 - \mathbf{r}_2)} W_1 W_2 d^3k d^3r_1 d^3r_2$$

And finally

$$\sigma_R^2 = (2\pi^2)^{-1} \int P(k) W_R^2(k) k^2 dk$$

Relation between moments of the counts in cells and the power spectrum.

The higher is σ the stronger the clustering

$$\sigma^2 \approx \delta(x)^2$$

$\sigma \approx 1$ implies strong non-linearity

Today we measure $\sigma_{8Mpc/h} \approx 0.8$

The structure of the large scale structure

large scales — **small scales**

super-horizon scales	linear scales	mildly non-linear scales	strongly non-linear scales
relativistic corrections	*linear pert. theory*	*non-linear pert. theory*	*N-body simulations*

$k \approx aH \approx$ 0.002 h/Mpc

$\lambda \approx 3000\ Mpc/h$

$k \approx 0.1\ h/Mpc$

$\lambda \approx 60\ Mpc/h$

$k \approx 0.3\ h/Mpc$

$\lambda \approx 20\ Mpc/h$