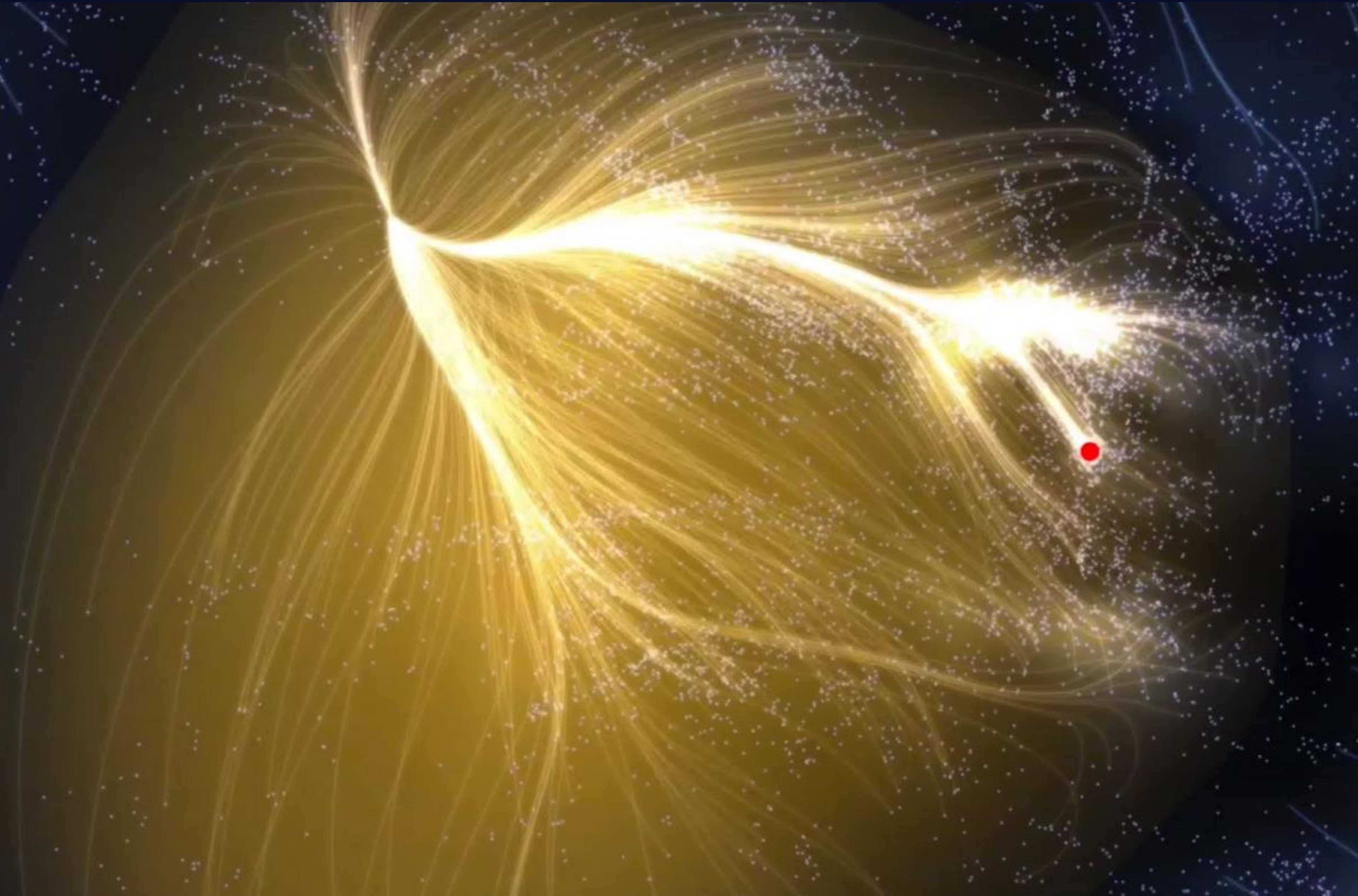
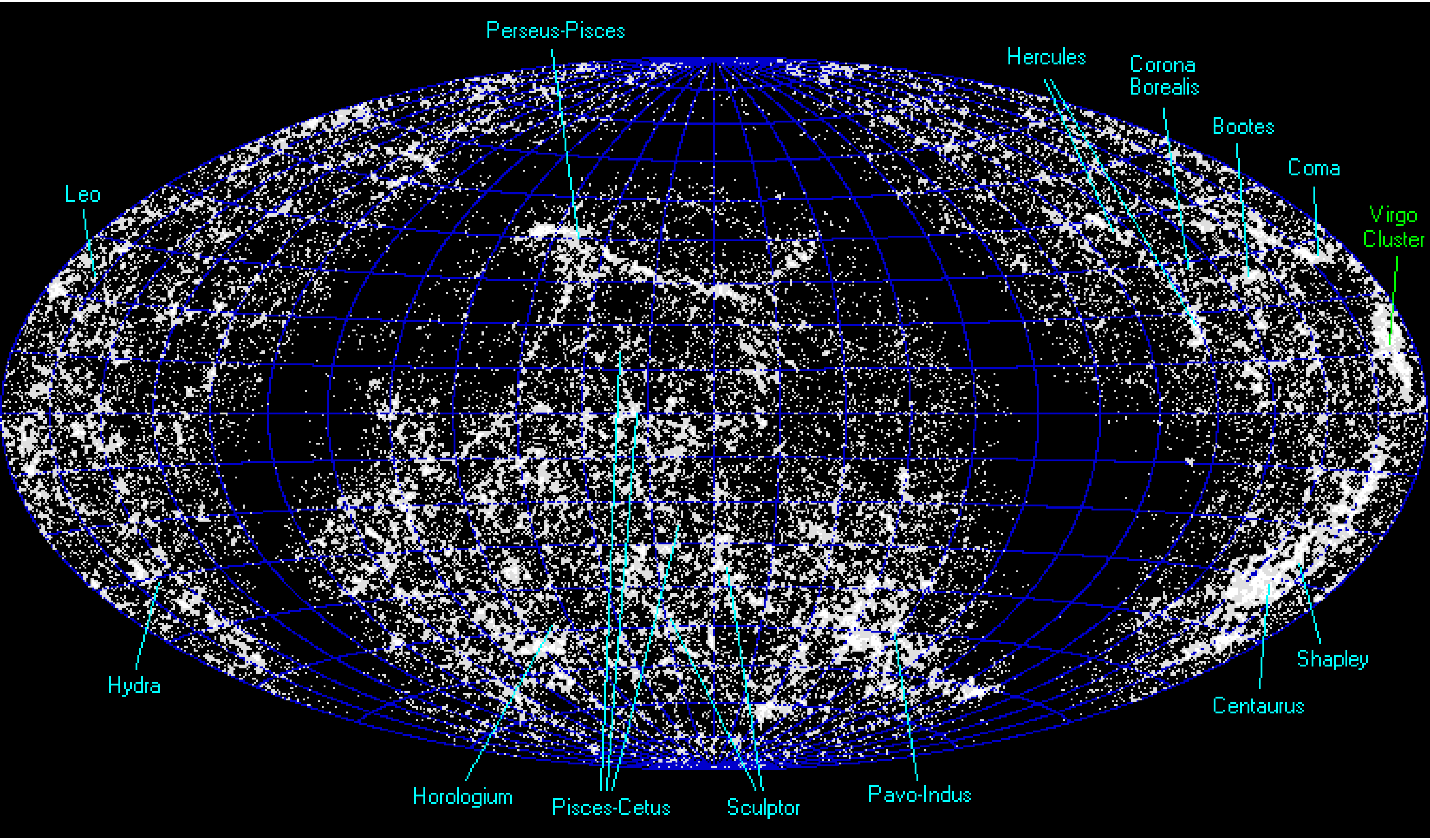
Ay 21: Large-Scale Structure: Observations



Large-Scale Structure

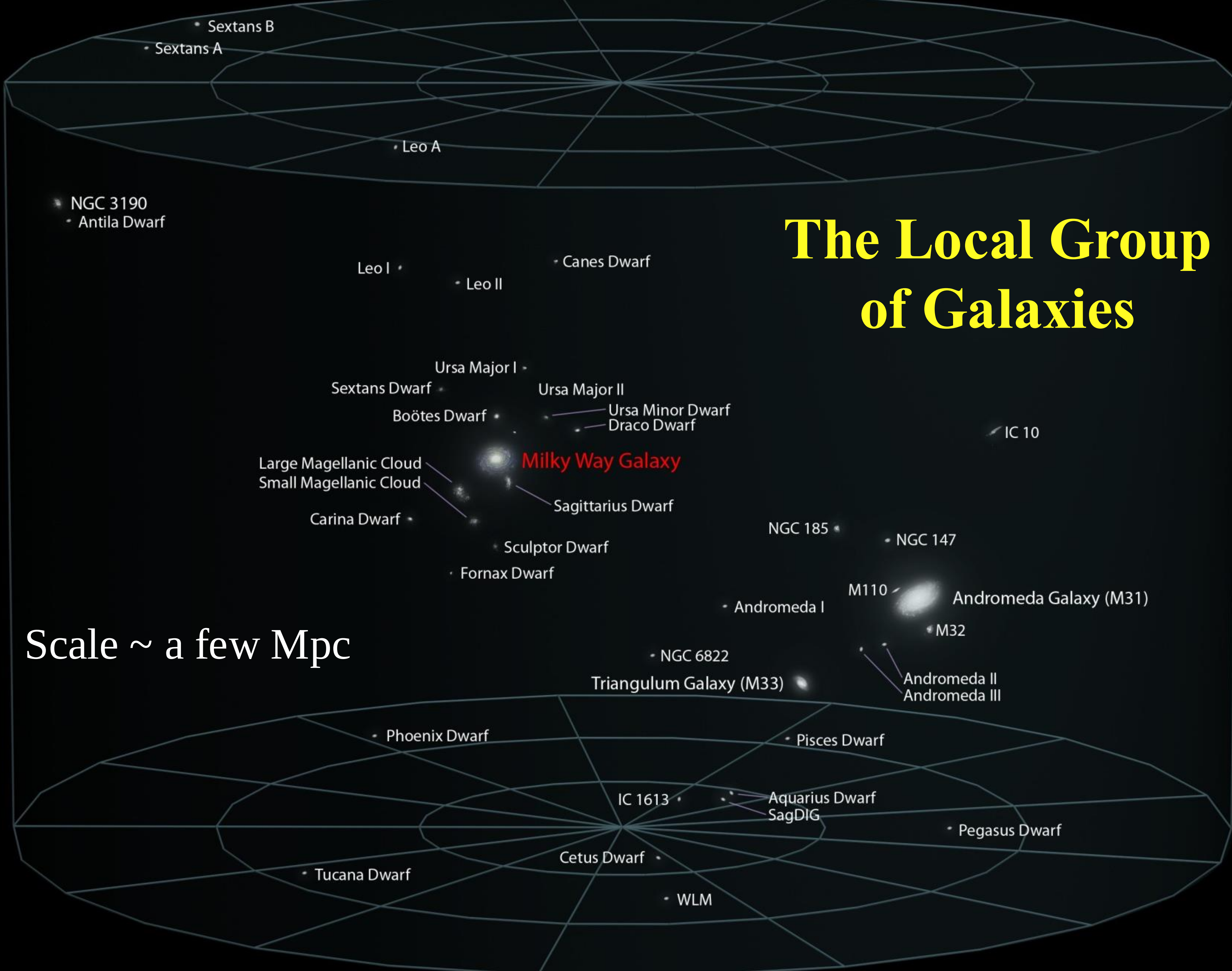
- Density fluctuations evolve into structures we observe: galaxies, clusters, etc.
- While the existence of clusters was recognized early on, it took a while to recognize that galaxies are not distributed in space uniformly randomly, but in coherent structures
- On scales > galaxies, we talk about the **Large Scale Structure (LSS);** groups, clusters, filaments, walls, voids, superclusters are the elements of it
- To map and quantify the LSS (and compare with the theoretical predictions), we need **redshift surveys**: mapping the 3-D distribution of galaxies in the space
  - Today we have redshifts measured for $> 5 \times 10^8$ galaxies
- The largest structures are on the scale of the first peak of the baryonic acoustic oscillations (BAO), ~ 100 h^{-1} Mpc

6000 Brightest Galaxies on the Sky



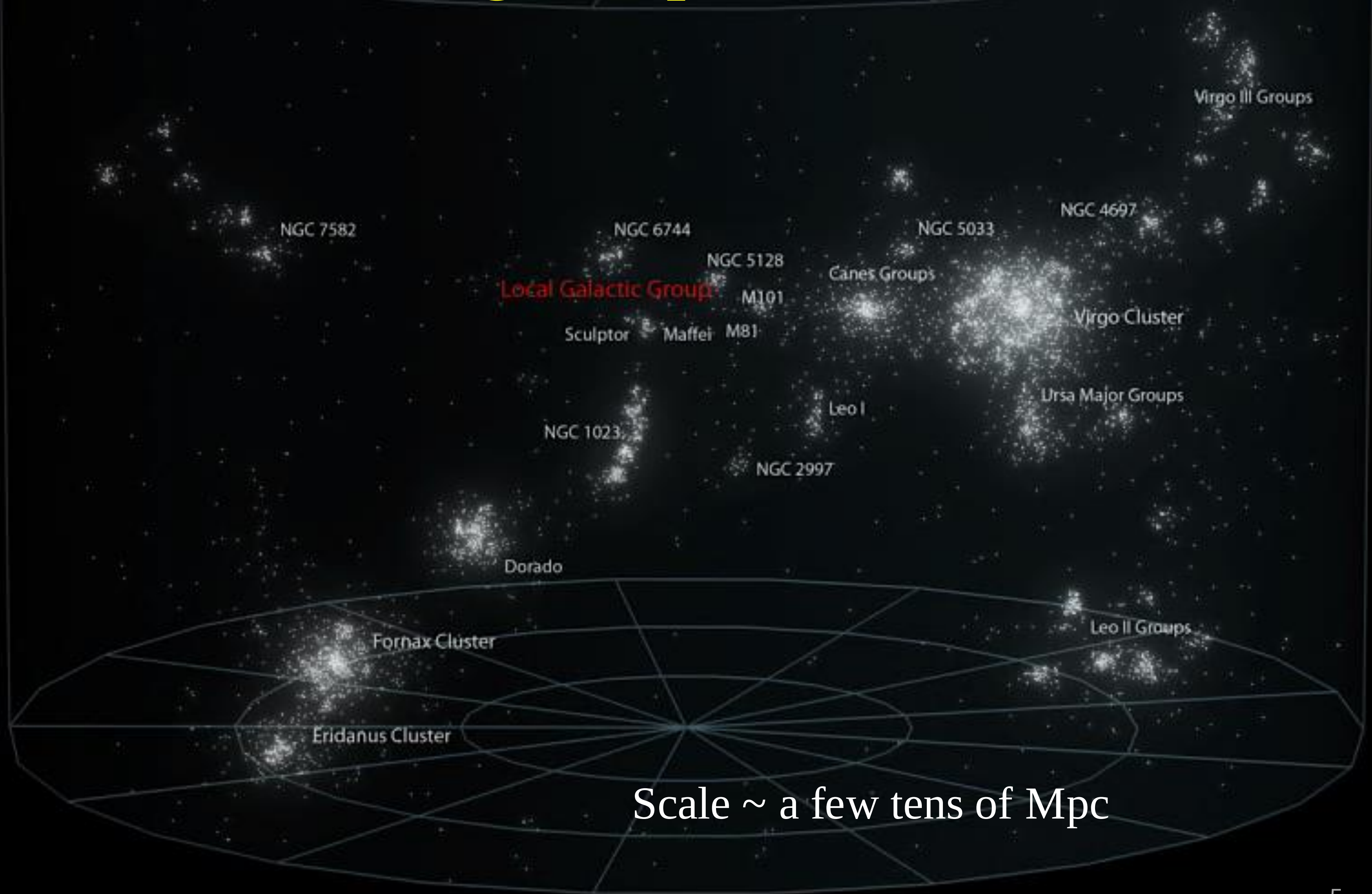
How would the picture of the 6000 brightest stars on the sky look?

The Local Group of Galaxies

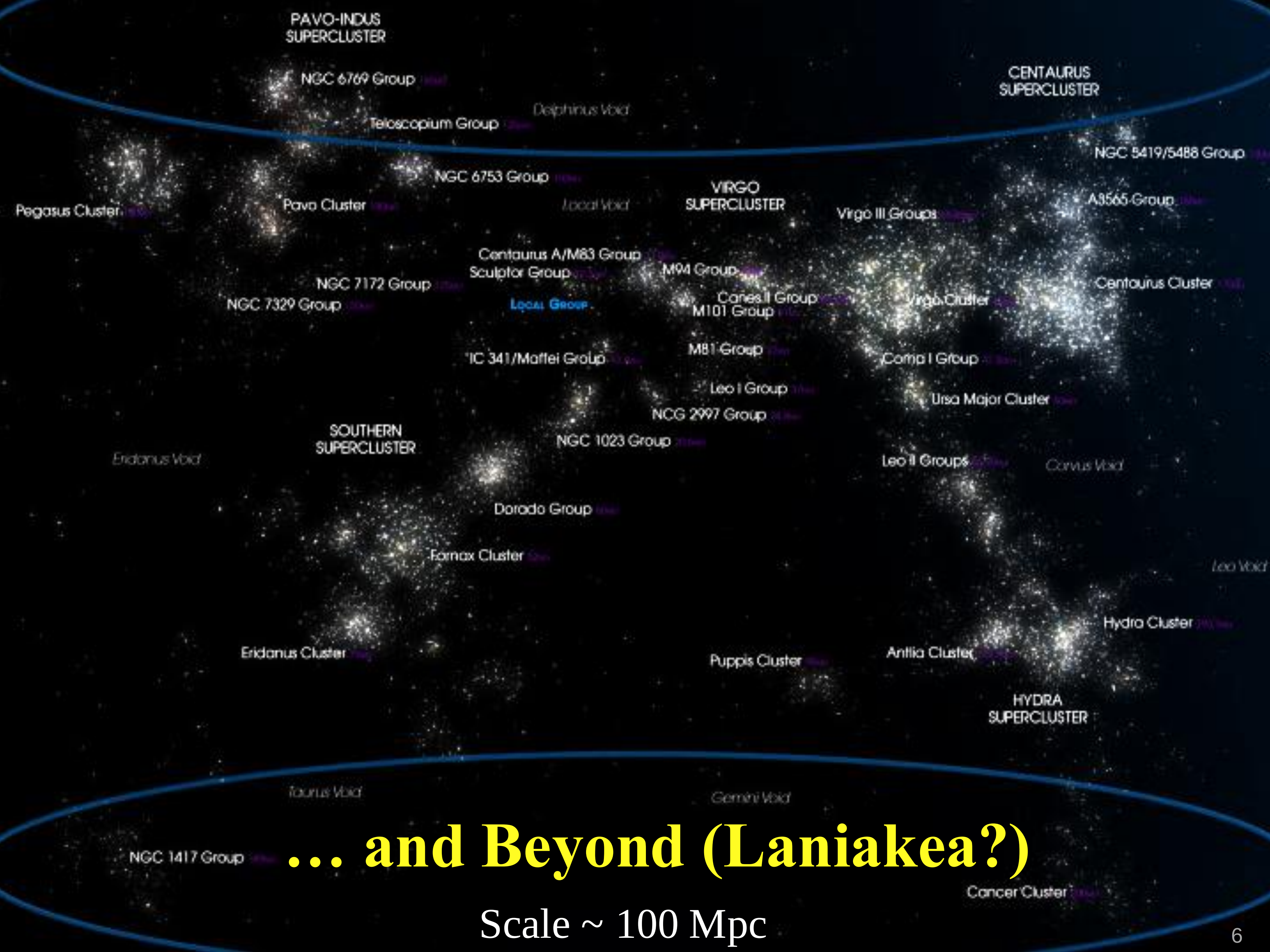


Scale ~ a few Mpc

The Local (Virgo) Supercluster



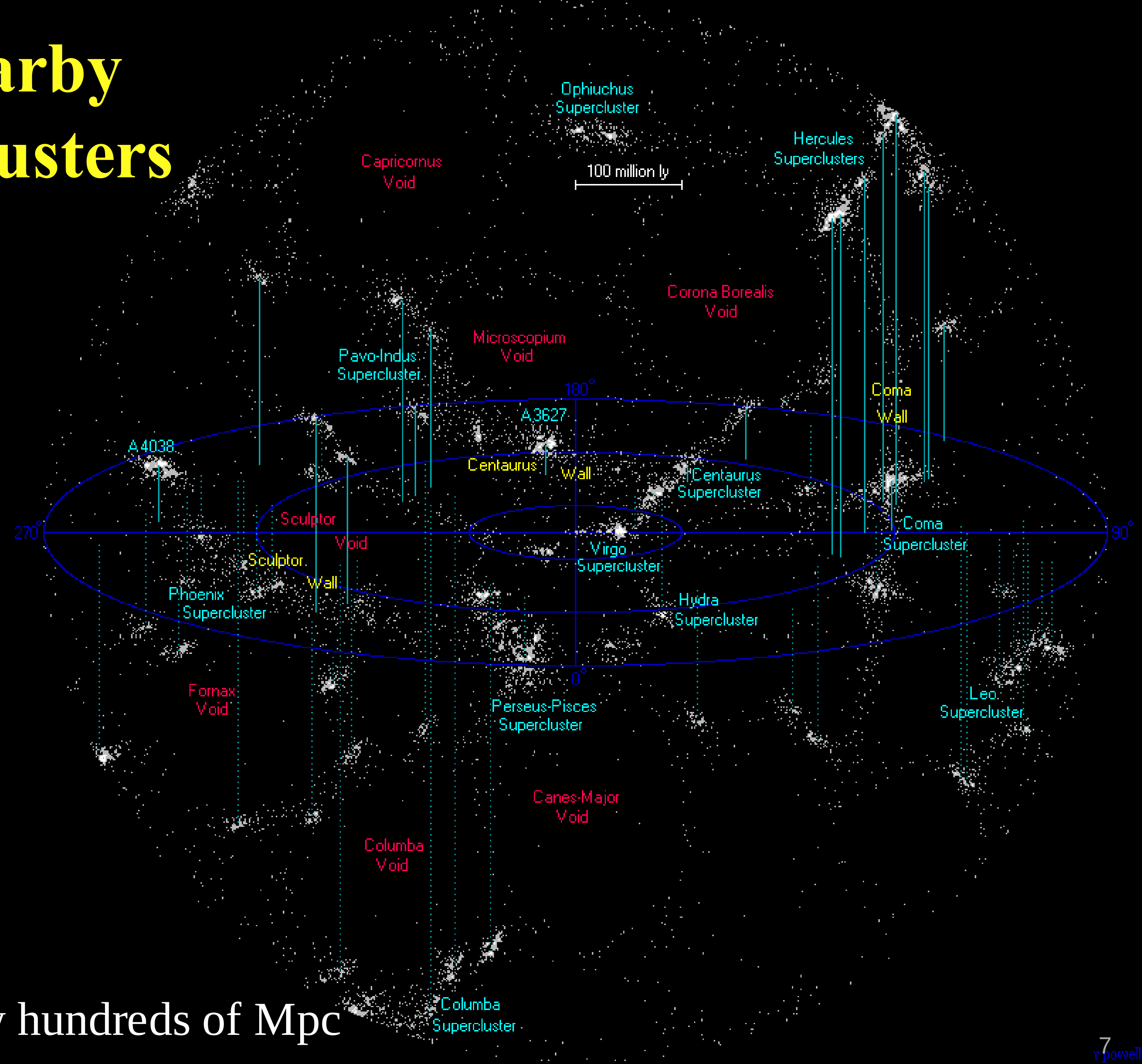
Scale ~ a few tens of Mpc



… and Beyond (Laniakea?)

Scale ~ 100 Mpc

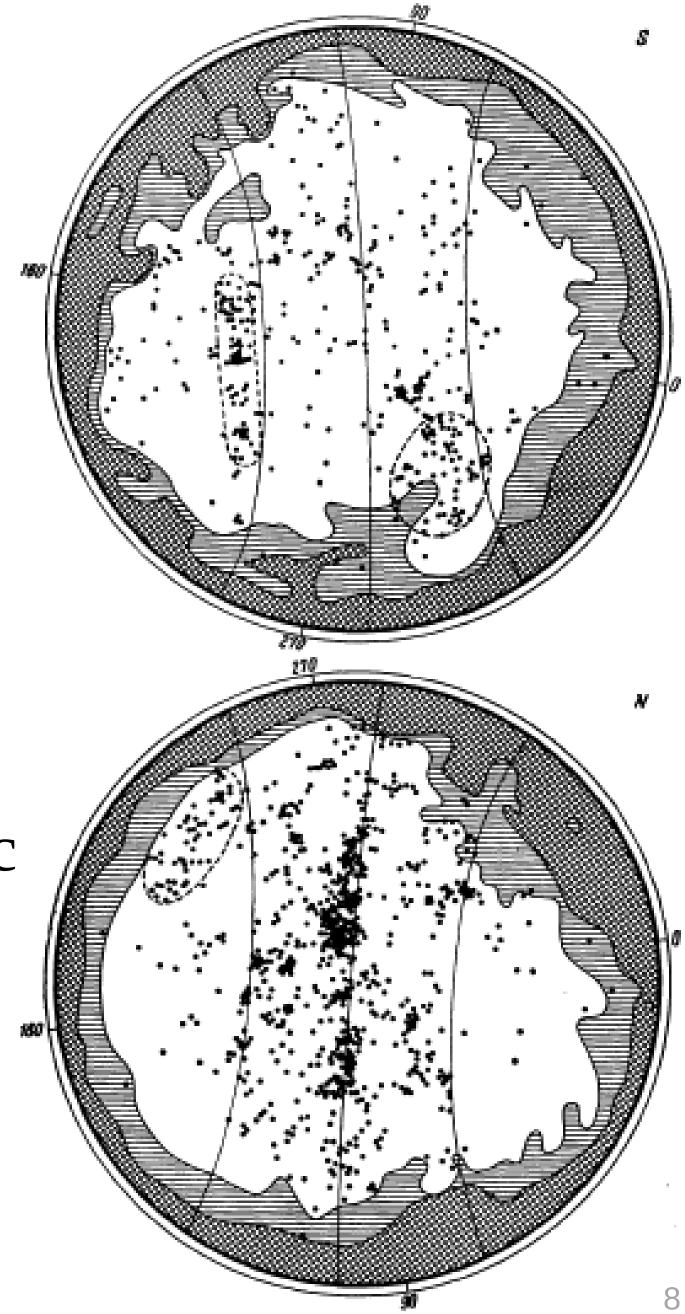
The Nearby Superclusters



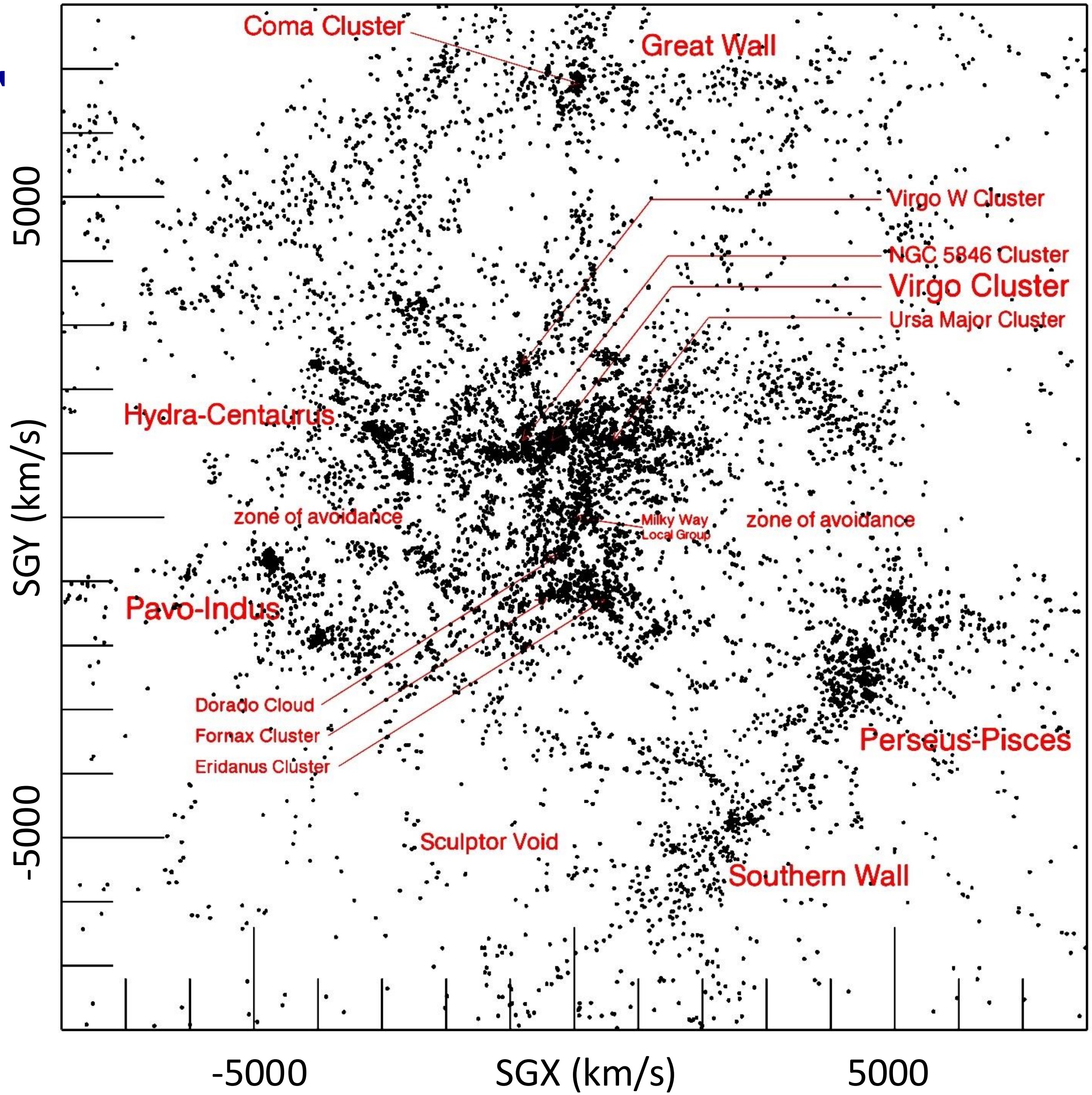
Scale ~ a few hundreds of Mpc

The Local Supercluster

- Hinted at by H. Shapley (and even earlier), but really promoted by G. de Vaucouleurs
- Became obvious with the first modern redshift surveys in the 1980's
- A ~ 60 Mpc structure, flattened, with the Virgo cluster at the center; the Local Group is at the outskirts
- Its principal axes define the supergalactic coordinate system (XYZ)
- Many other superclusters known; and these are the largest (~ 100 Mpc) structures known to exist



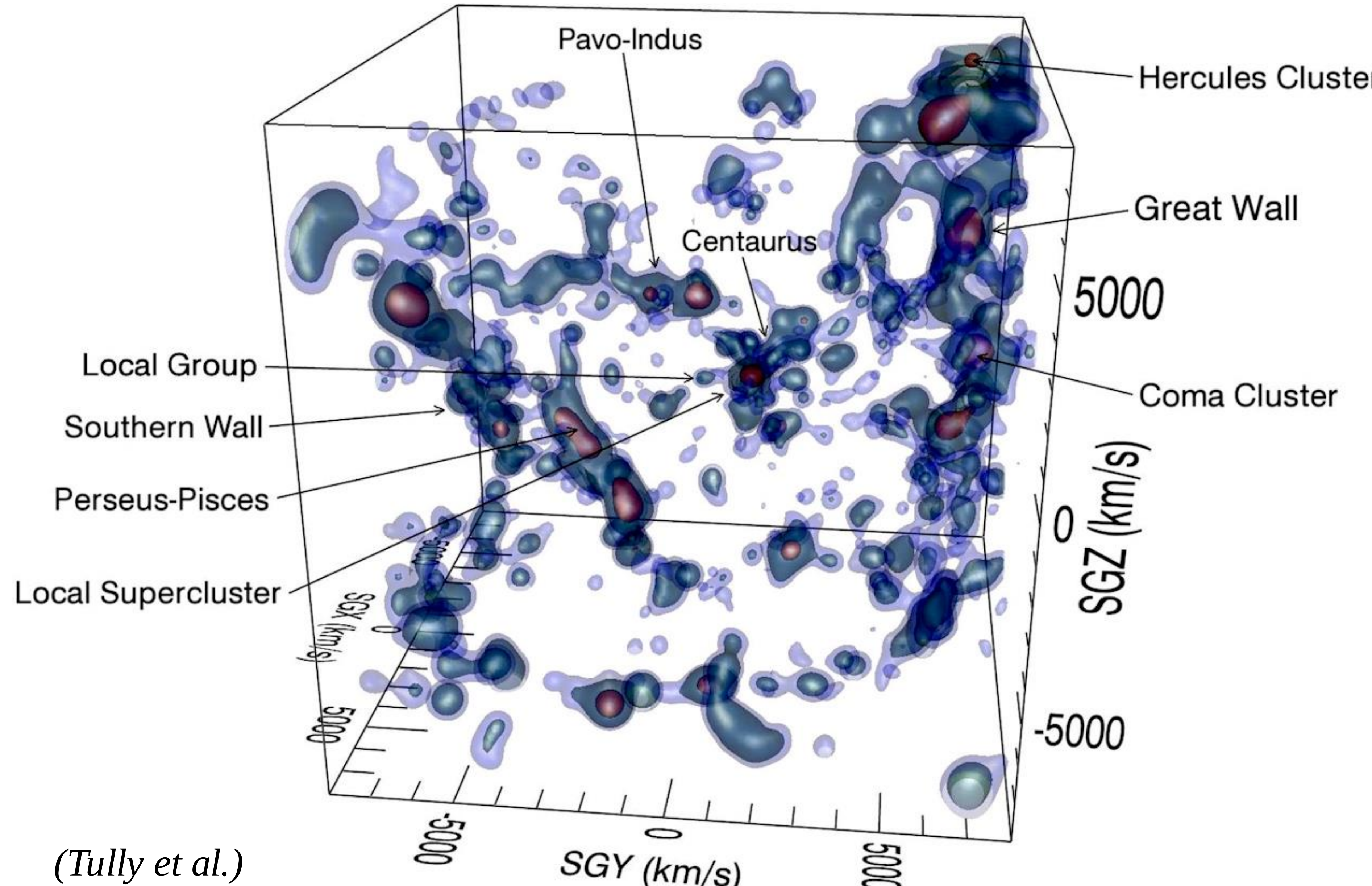
The Local Supercluster and Vicinity



In the redshift space, projected on the supergalactic plane

(Tully et al.)

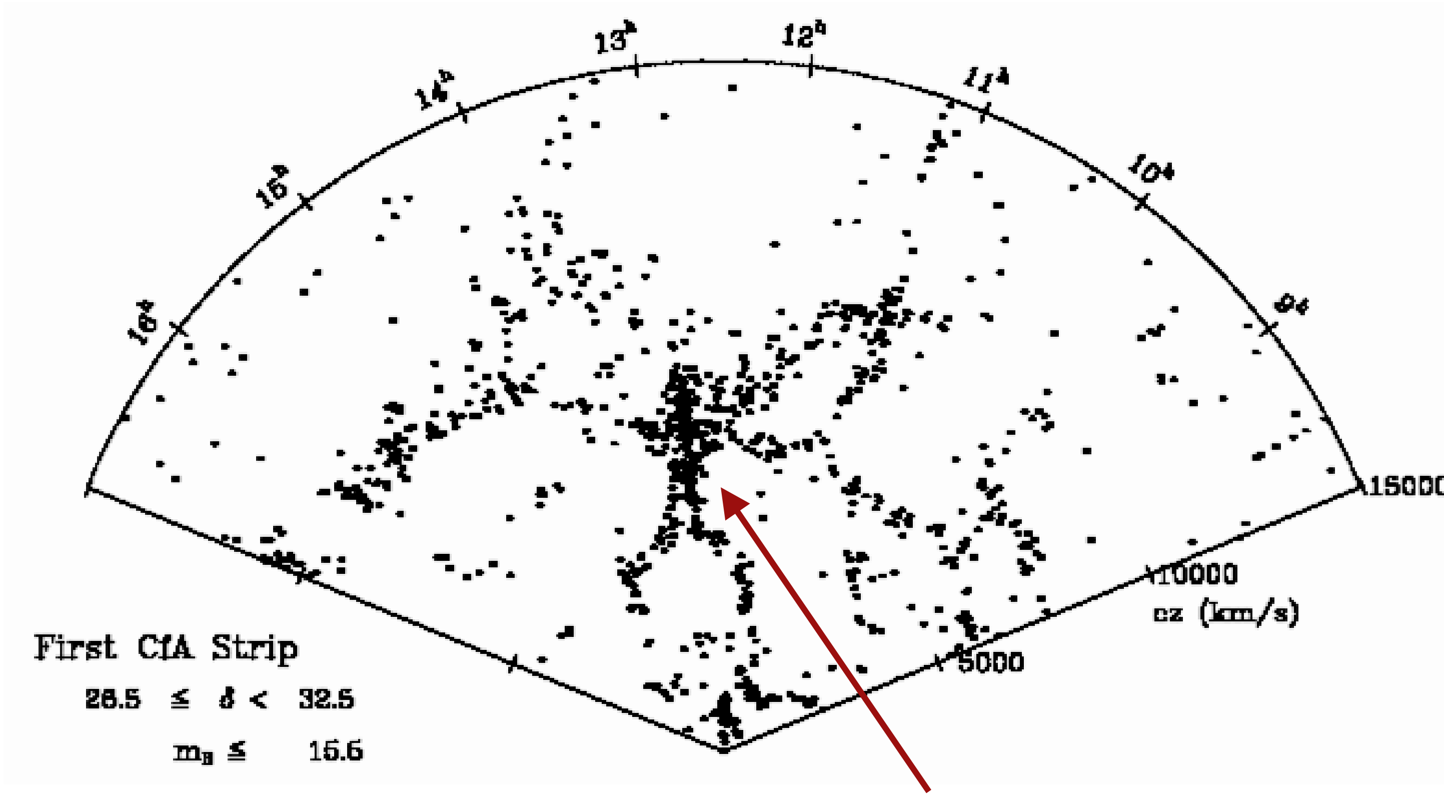
3D View in the Supergalactic Coordinates



(Tully et al.)

The Early Redshift Surveys

CfA2 Survey: The Infamous "Stickman Diagram"



Coma cluster:
Note the "finger of God" effect, due to the velocity disp. in the cluster

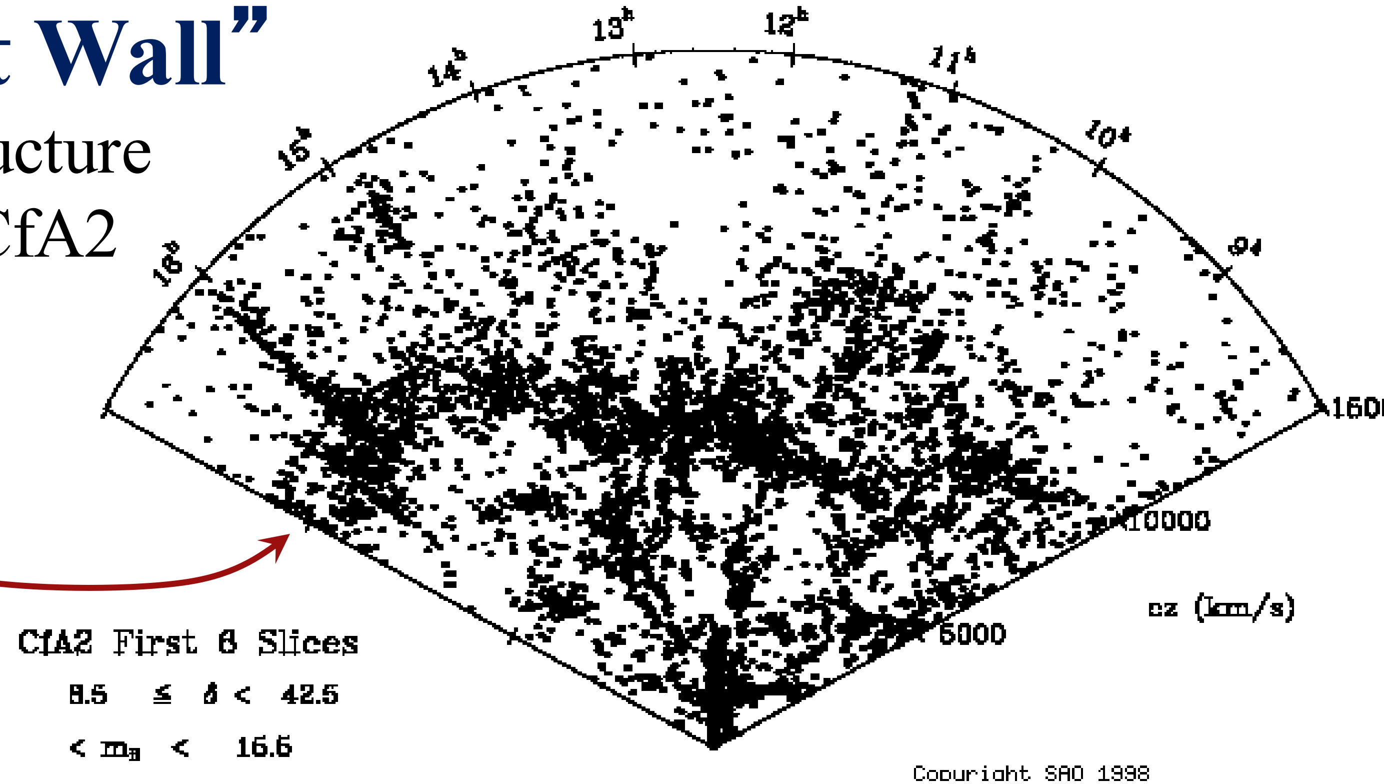
At first galaxies were observed one by one.

The 2nd generation redshift surveys were often done in slices which were thin in Dec and long in RA, thus sampling a large dynamical range of scales.

This also helped reveal the large-scale ***topology*** (voids, walls, filaments).

The “Great Wall”

a ~ 100 Mpc structure revealed in the CfA2 redshift survey

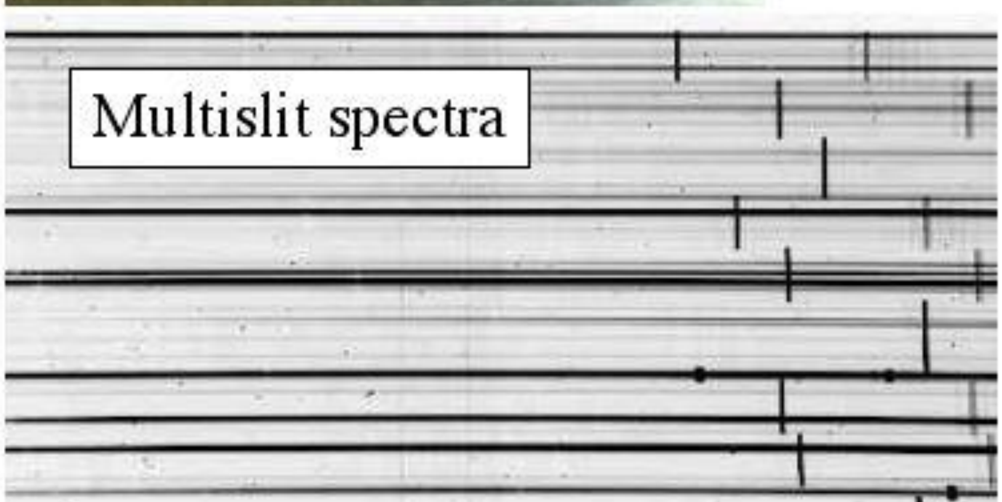


Up until then, redshift surveys revealed structures as large as can be fitted within the survey boundaries - but 100 Mpc turned out to be about as large as they come.

The next generation of surveys sampled *3-D volumes* (rather than thin slices), sometimes with a *sparse sampling* (measure redshift of every *n*-th galaxy), and often used *multi-object spectrographs*.

Tools of the Trade: Multiobject Spectrographs

Keck DEIMOS slitmask (detail)



Multislit spectra

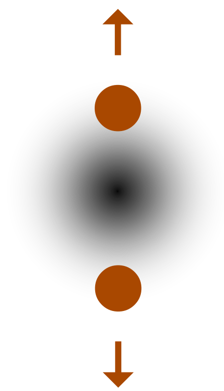
2dF multifiber spectrograph

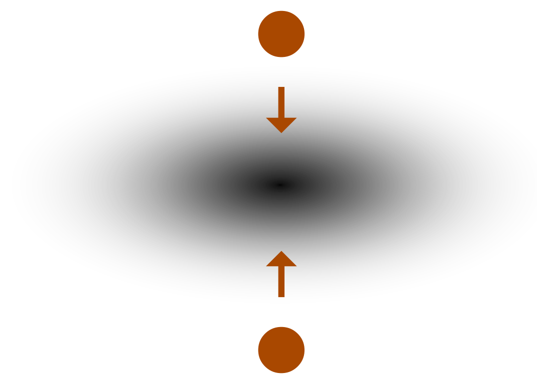
Redshift Space vs. Real Space

Because the observed velocity = cosmic expansion + the peculiar velocity

"Fingers of God" | Thin filaments

Real space distribution

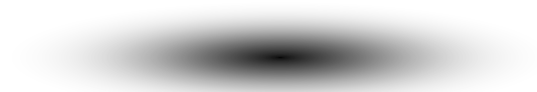




The effect of cluster velocity dispersion

The effect of infall

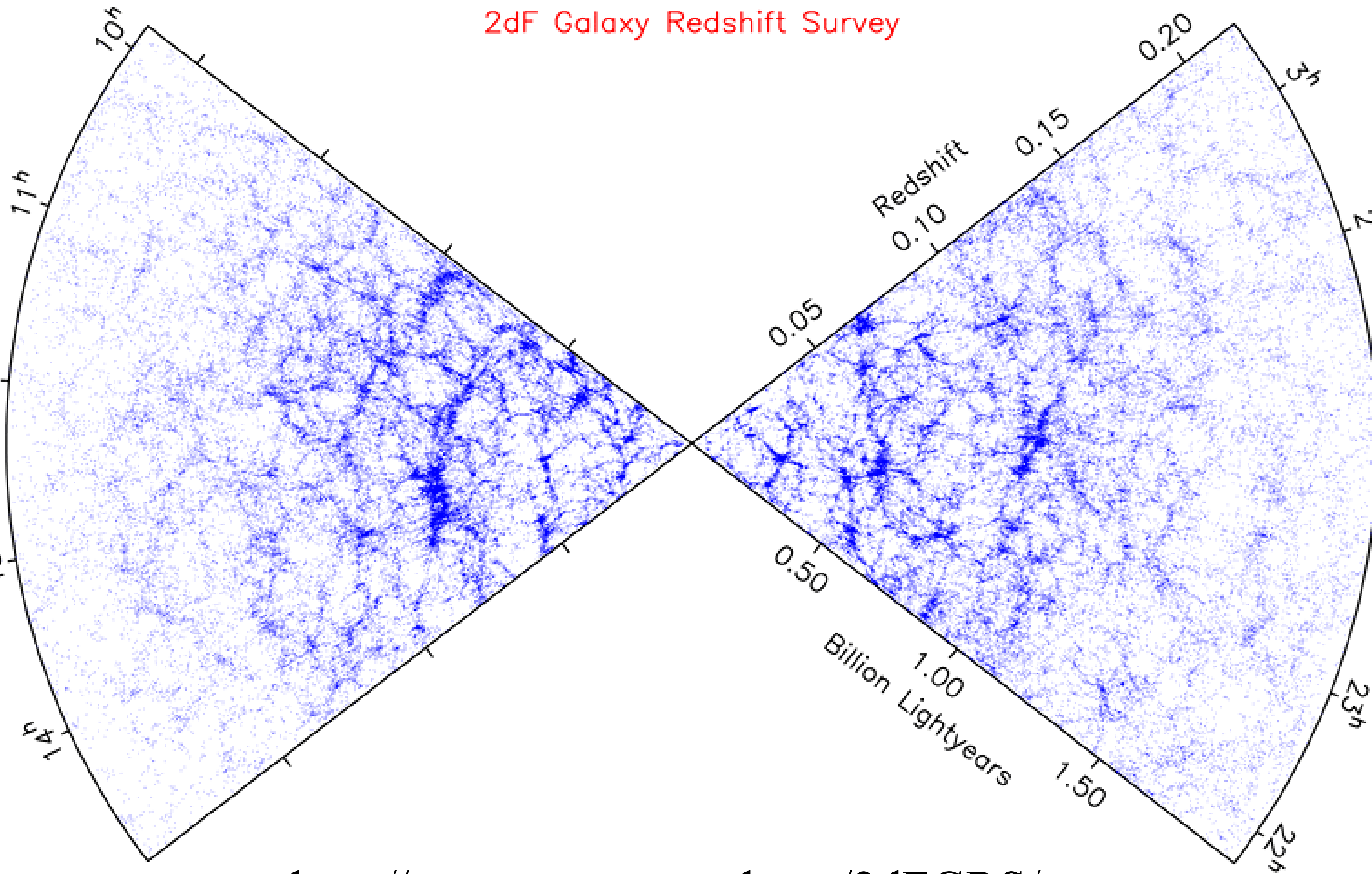
Redshift space apparent distrib.



The First Large Redshift Surveys

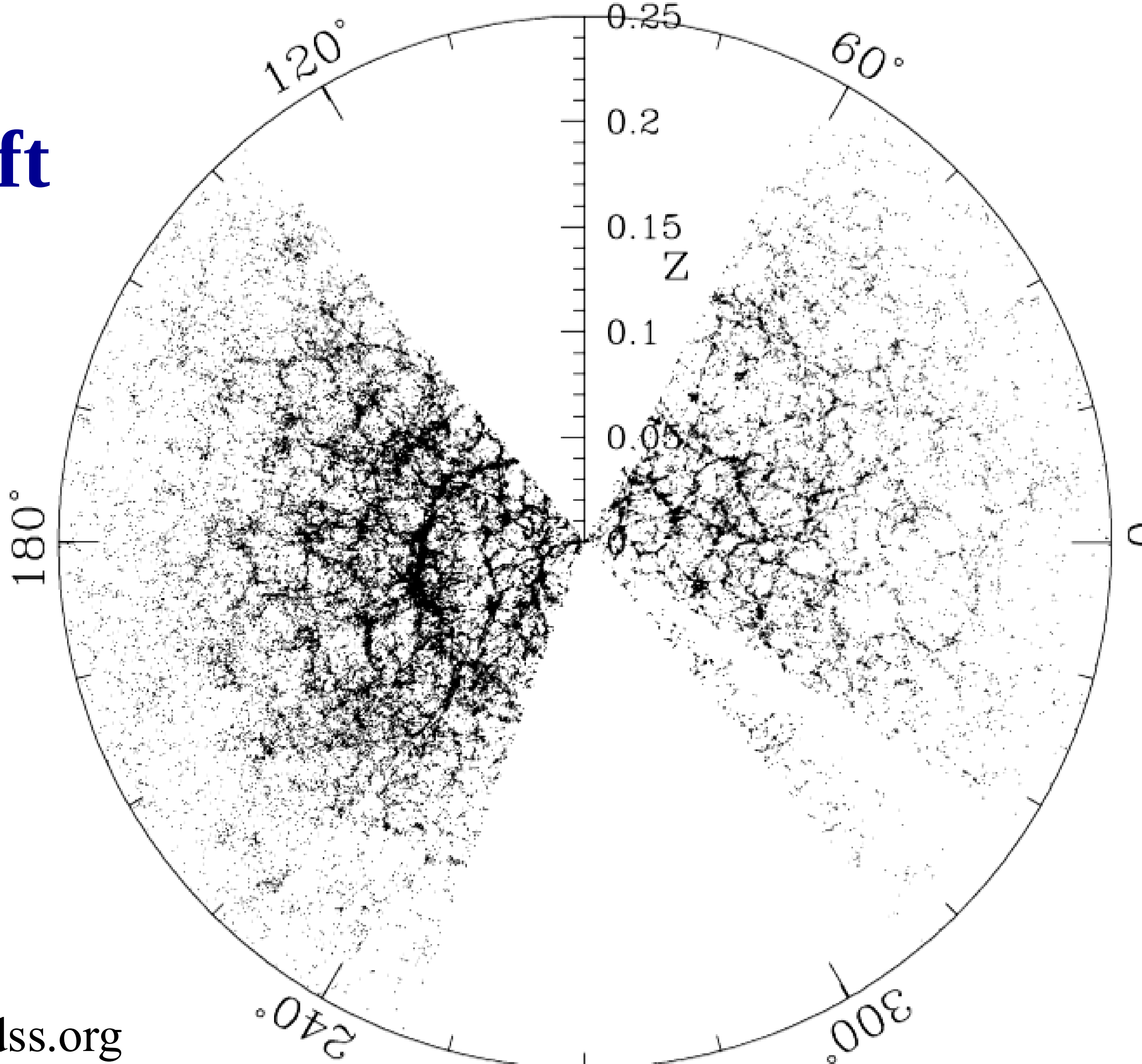
- **The 2dF (2 degree Field) redshift survey** done with the 3.9-m Anglo-Australian telescope by a UK/Aus consortium
  - Redshifts of ~ 250,000 galaxies with $B < 19.5$ mag, covering 5% of the sky reaching to $z \sim 0.3$
  - Spectrograph can measure 400 redshifts at a time
  - Also spectra of ~ 25,000 QSOs out to z ~ 2.3
- **The Sloan Digital Sky Survey (SDSS)** done with a dedicated 2.5-m telescope at Apache Point Observatory in New Mexico
  - Multicolor imaging to $r \sim 23$ mag, and spectra of galaxies down to $r < 17.5$ mag, reaching to $z \sim 0.4$, obtaining ~ 600 (now ~ 1,000) spectra at a time, covering ~ 14,000 deg^2
  - To date (SDSS+extensions): ~ 1 *billion* detected sources, 2+ *million* galaxy spectra, > 700,000 stellar spectra, > 300,000 quasar spectra (reaching out to $z \sim 6.4$)

2dF Galaxy Redshift Survey



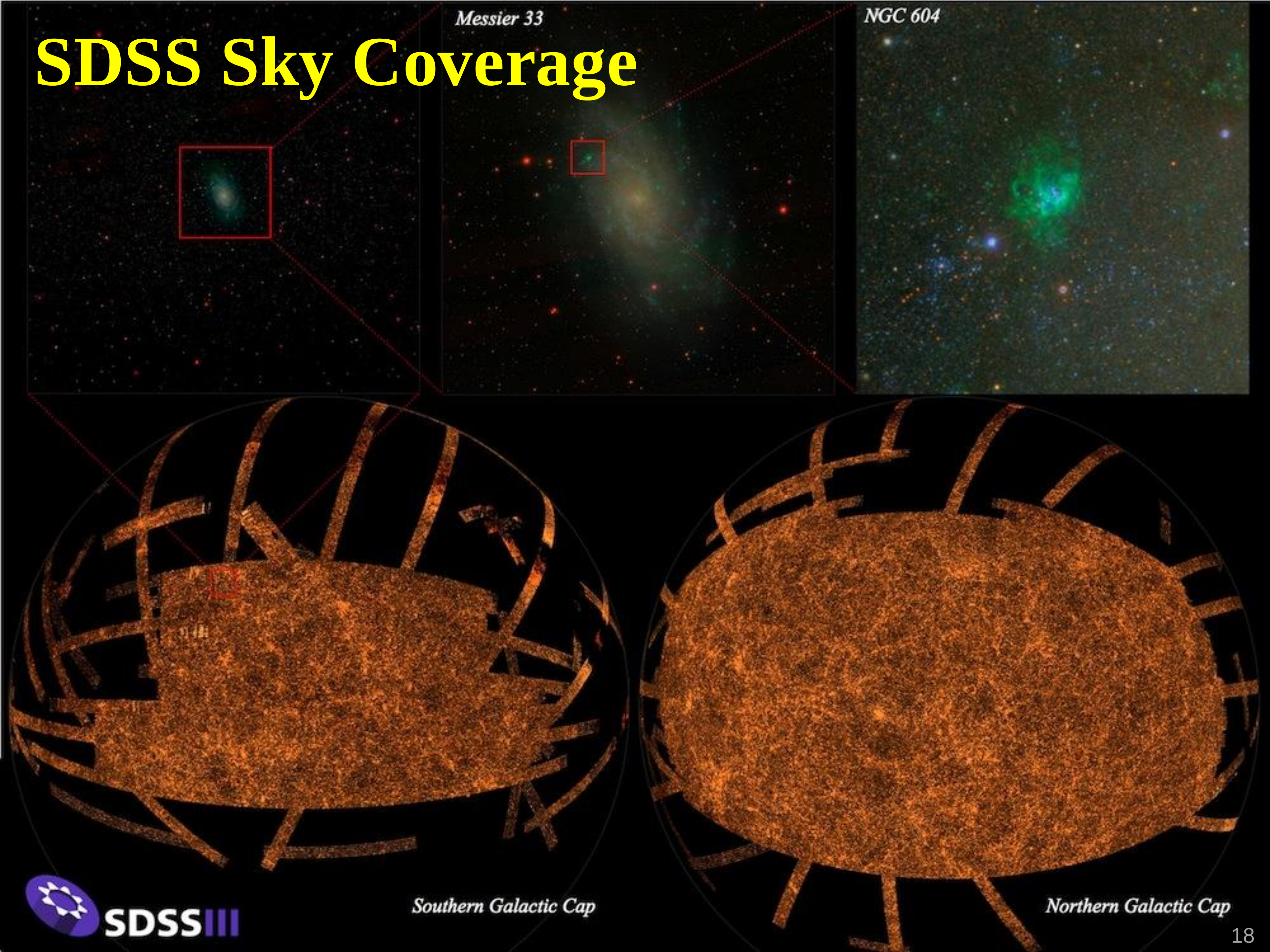
http://www.mso.anu.edu.au/2dFGRS/

SDSS Redshift Survey



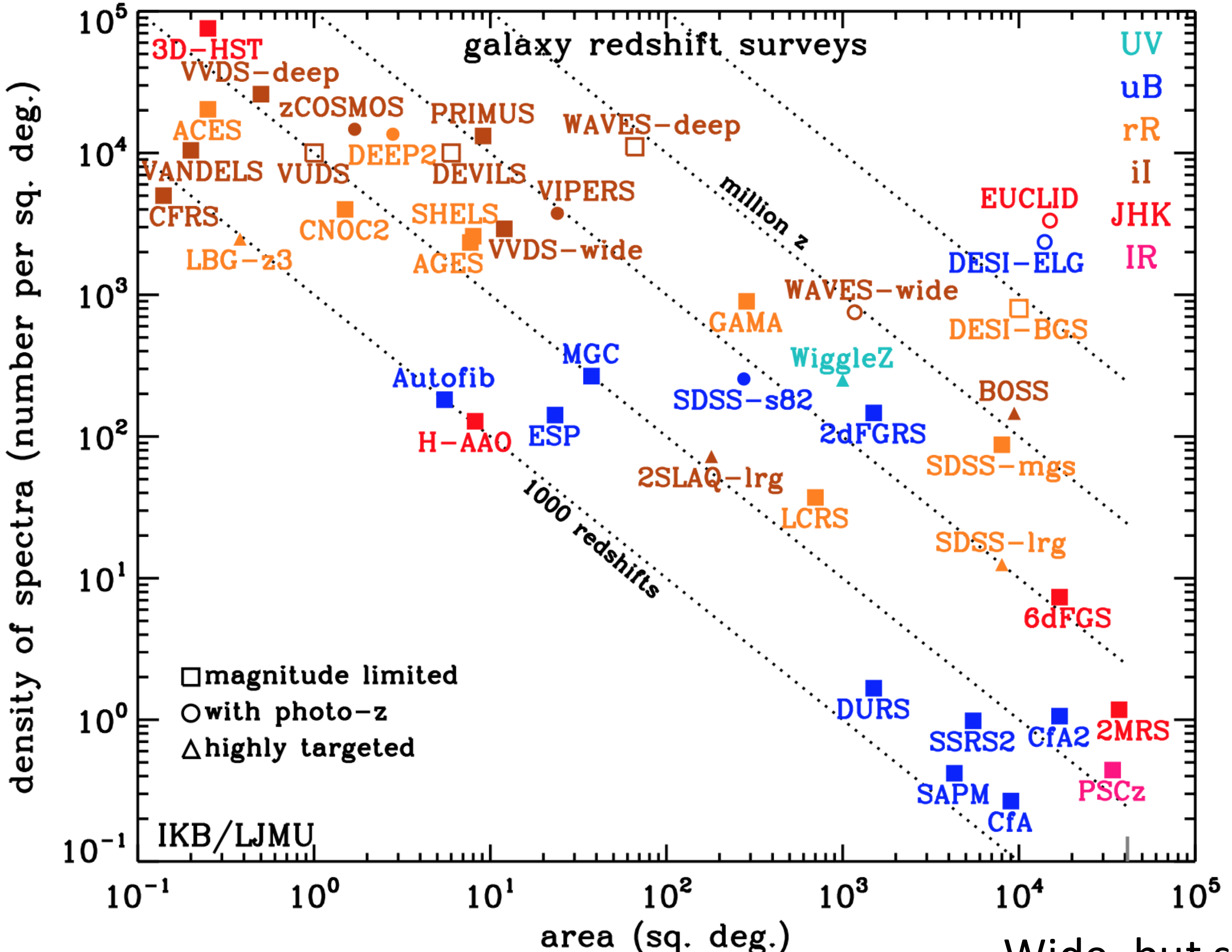
http://www.sdss.org

SDSS Sky Coverage



Comparing Redshift Surveys

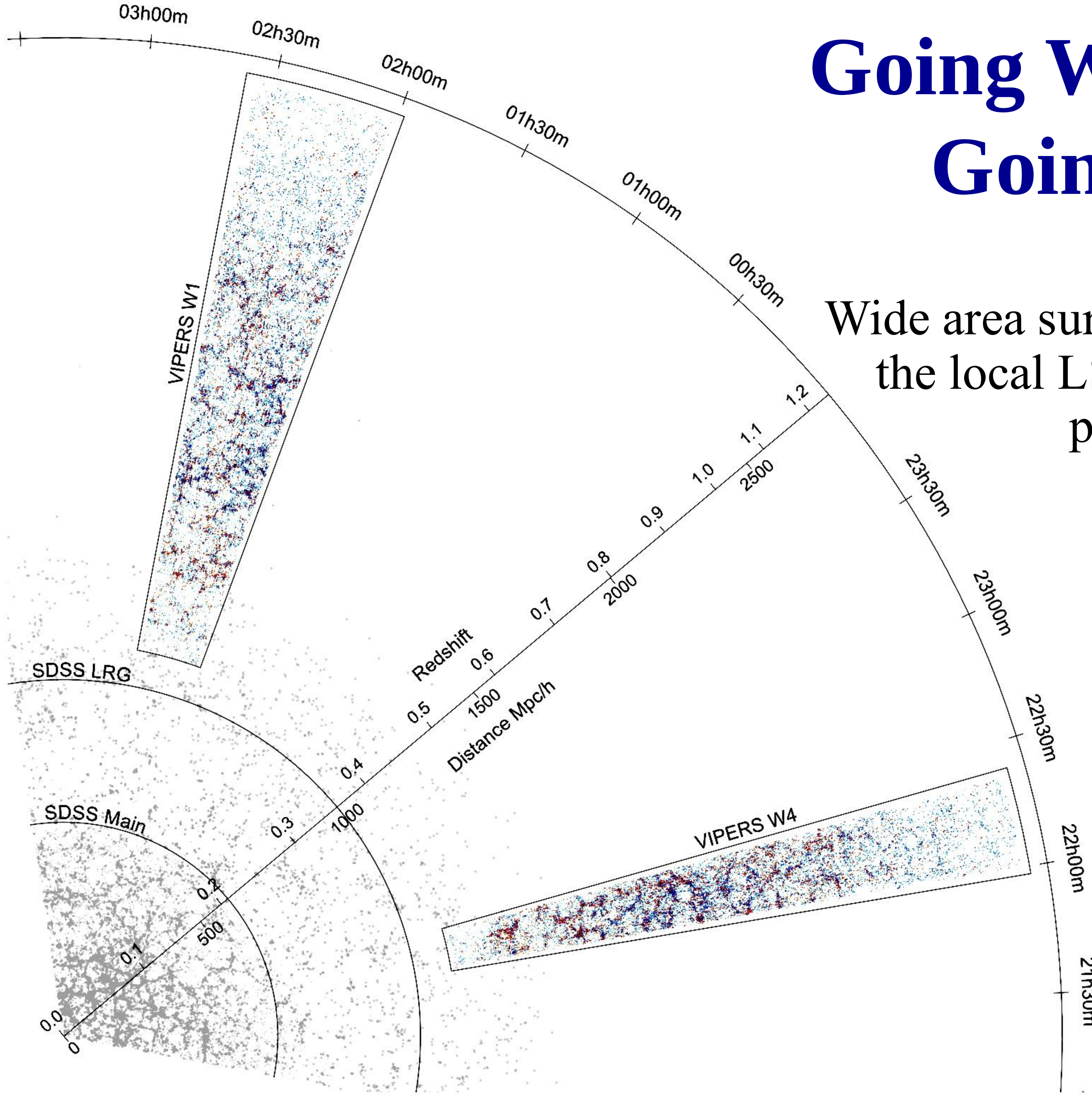
Small area, but deep



Wide, but shallow

Going Wide vs. Going Deep

Wide area surveys characterize the local LSS, but deep ones probe its evolution

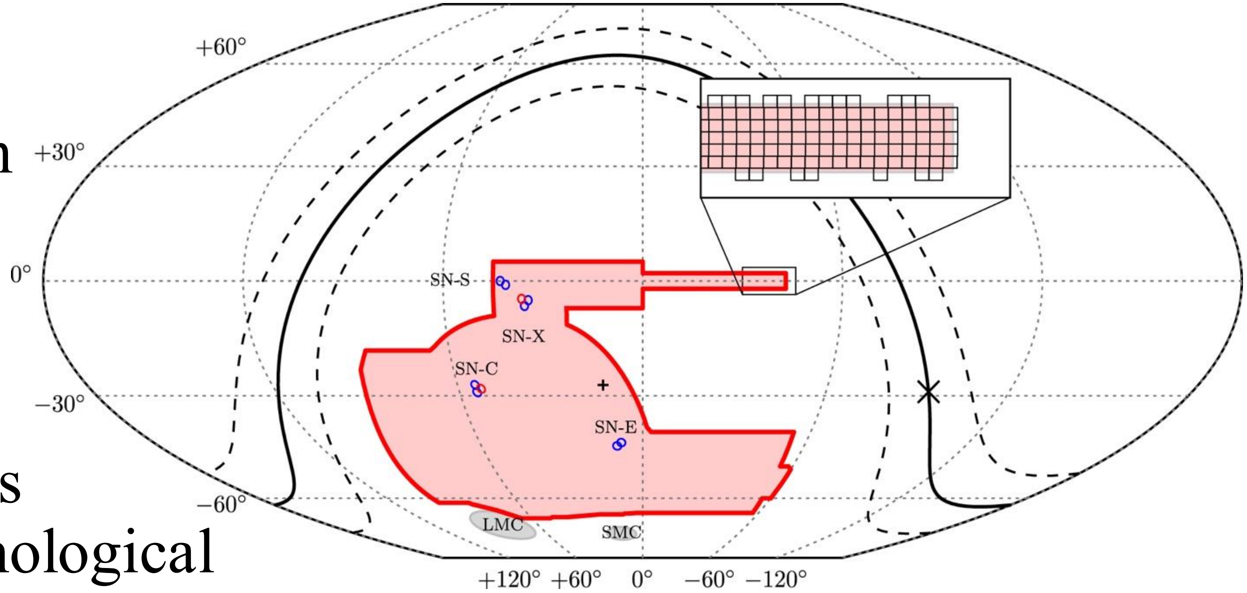


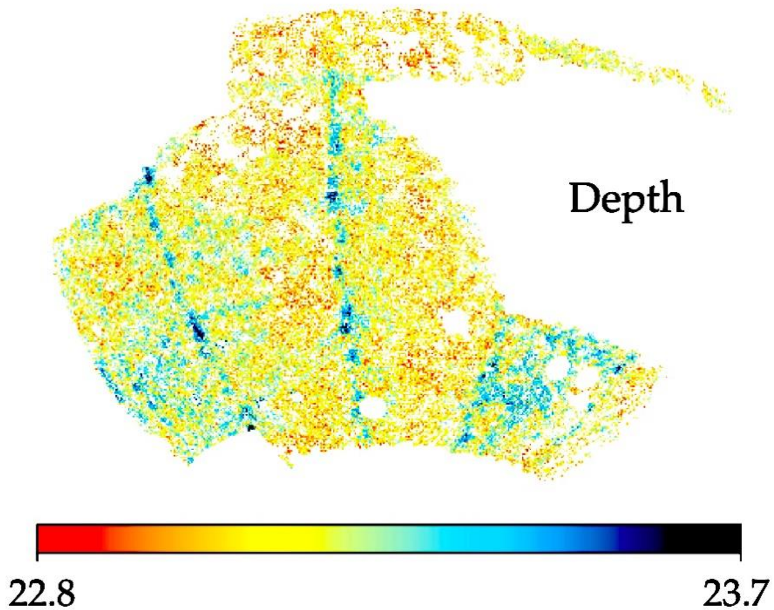
The Dark Energy Survey (DES)

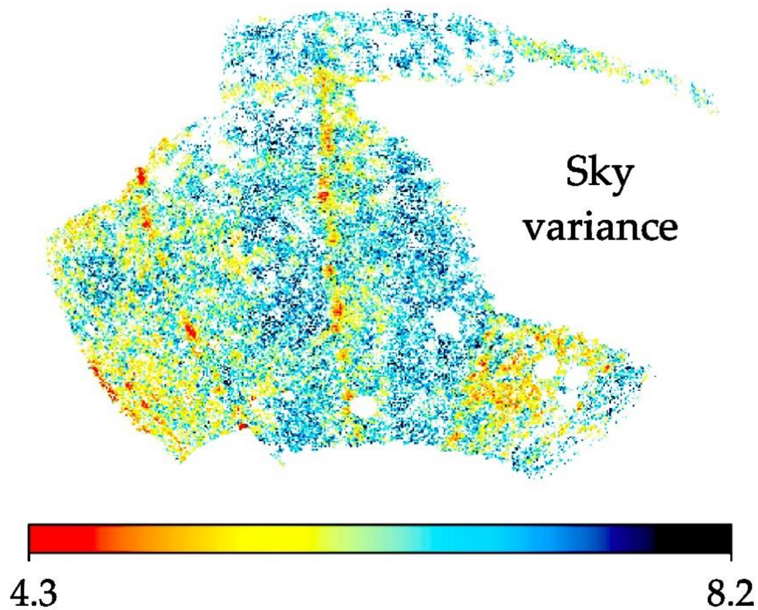
Covers ~ 5000 deg^2

Spectra of ~ 500 million galaxies, some reaching to $z > 1$, plus many quasars

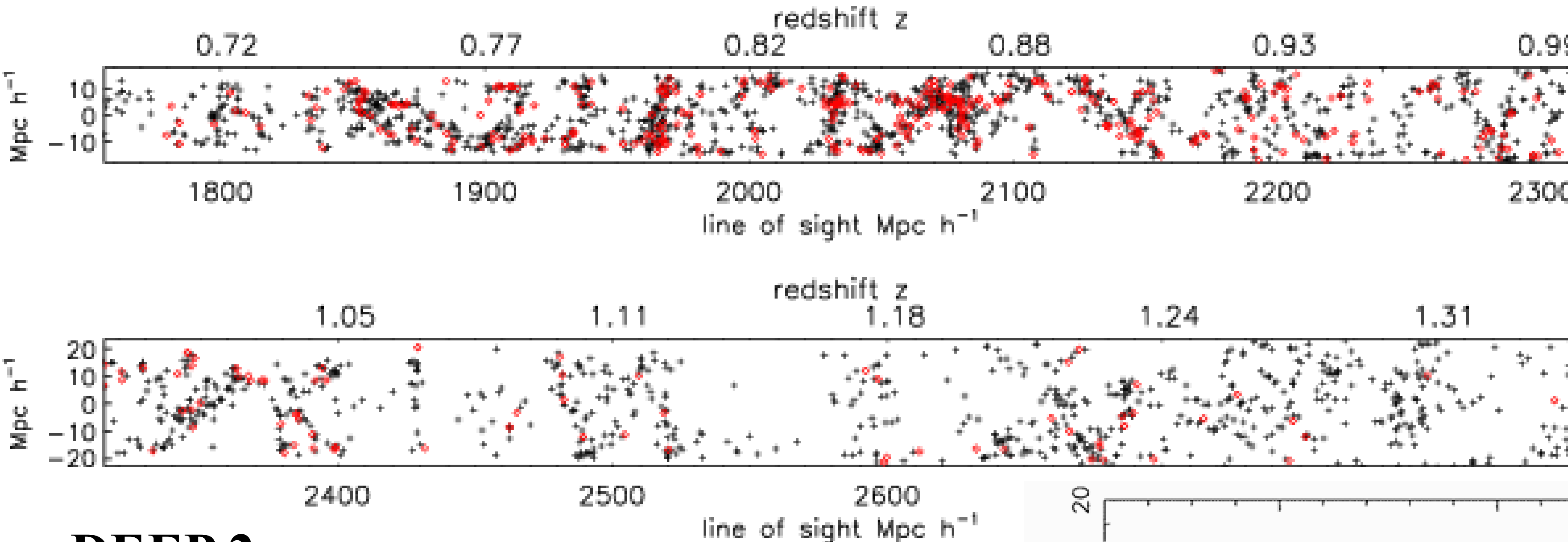
Combines measurements of the LSS and the cosmological parameters







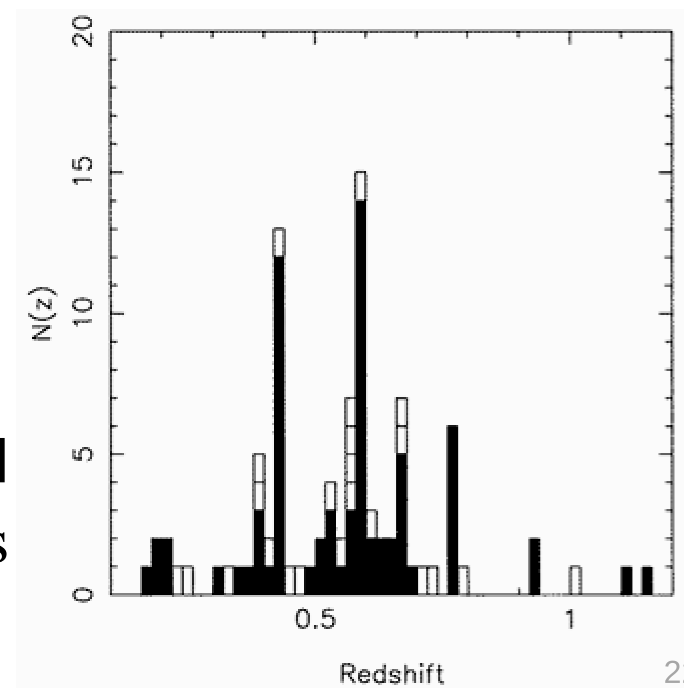
Structures in Deep Redshift Surveys



DEEP 2
02 hr field

red=emission-line
black=absorption

Spikes in the redshift histogram □
as line of sight intersects walls or filaments



Galaxy Distribution and Correlations

- If galaxies are clustered, they are "correlated"
- This is usually quantified using the ***2-point correlation function, ξ (r),*** defined as an "excess probability" of finding another galaxy at a distance ***r*** from some galaxy, relative to a uniform random distribution; averaged over the entire set:

$$dN(r) = r_0\left(1 + x(r)\right)dV_1 dV_2$$

- Usually represented as a power-law:

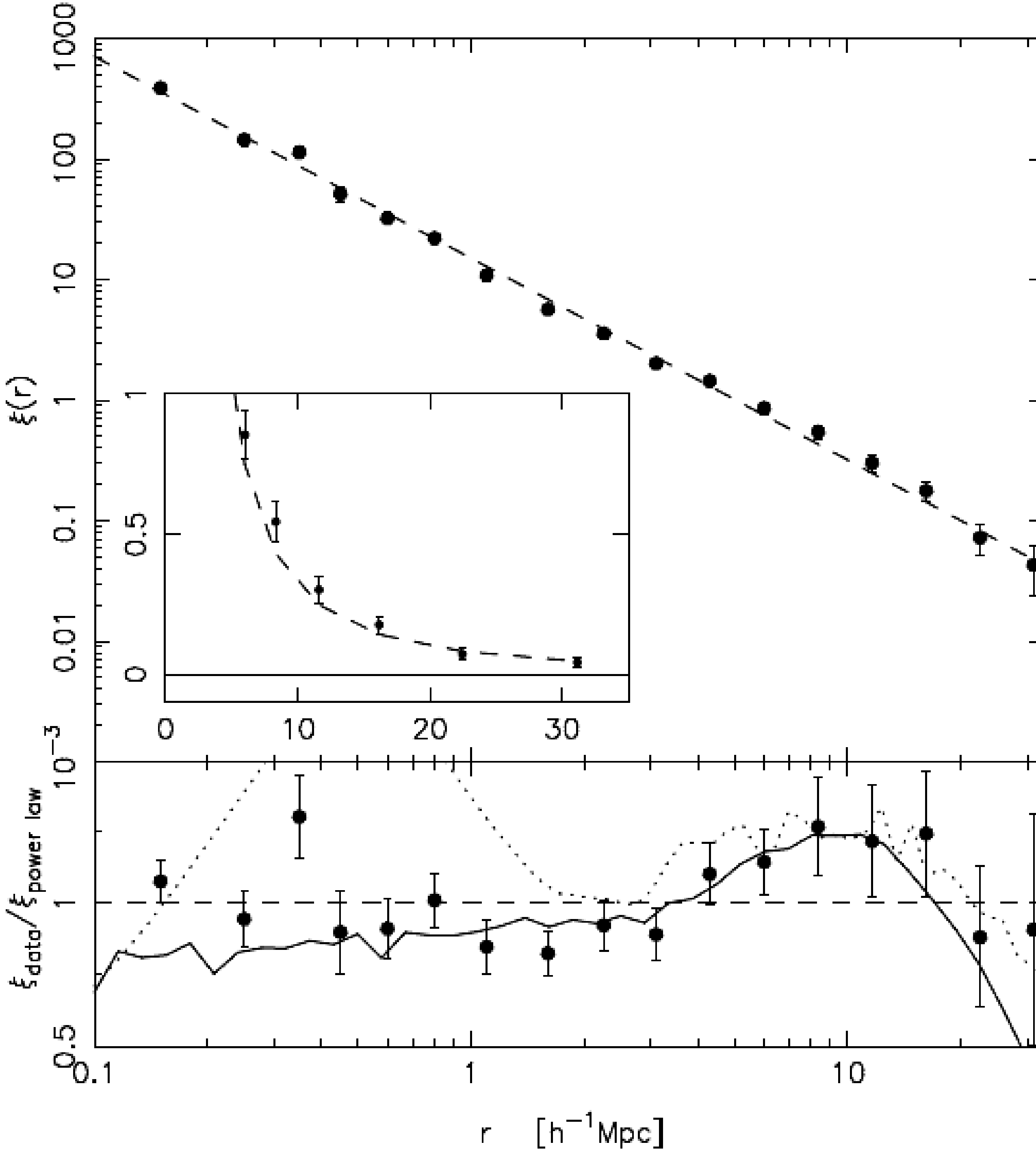
$$x(r) = \left(r / r_0\right)^{-g}$$

- For galaxies, typical ***correlation or clustering length*** is $\boldsymbol{r_0} \sim 5$ h^{-1} Mpc, and typical slope is $\gamma \approx 1.8$, but these are functions of various galaxy properties; clustering of clusters is stronger

Galaxy Correlation Function

As measured by the 2dF redshift survey

Deviations from the power law:



How to Measure ξ(r)

- Simplest estimator: count the number of data-data pairs, $\langle DD \rangle$, and the equivalent number in a randomly generated (Poissonian) catalog, $\langle RR \rangle$:

$$x(r)_{est} = \frac{\langle DD \rangle}{\langle RR \rangle} - 1$$

- A better (Landy-Szalay) estimator is:

$$x(r)_{est} = \frac{\langle DD \rangle - 2\langle RD \rangle + \langle RR \rangle}{\langle RR \rangle}$$

where $\langle RD \rangle$ is the number of data-random pairs

- This takes care of the edge effects, where one has to account for the missing data outside the region sampled, which can have fairly irregular boundaries

Another Definition of $\xi(r)$

- We can also measure it through the overdensity: $$d(\mathbf{r}) = \frac{n - \langle n \rangle}{\langle n \rangle}$$ where $\langle n \rangle$ is the mean density
- In case of discrete galaxy catalogs, define counts in cells, N_i $$d_i(\mathbf{r}) = \frac{N_i - \langle N_i \rangle}{\langle N_i \rangle}$$
- Then $\xi(\boldsymbol{r})$ is the expectation value:

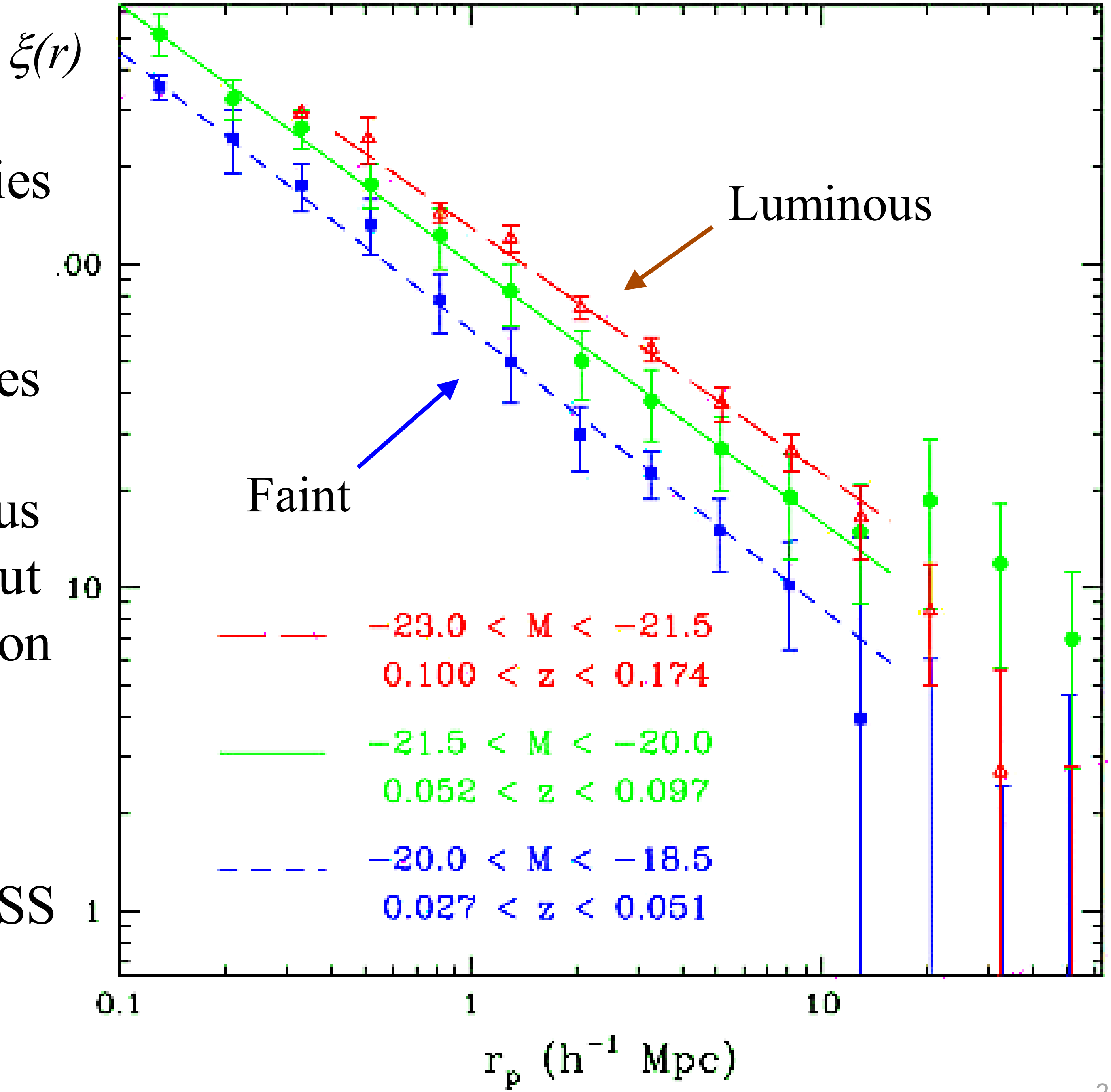
$$x(\mathbf{x}_1, \mathbf{x}_2) = \langle d(\mathbf{x}_1) d(\mathbf{x}_2) \rangle$$

- Note that we have considered a correlation of a single density field with itself, so strictly speaking $\xi(\boldsymbol{r})$ is the *autocorrelation* function, but in general we can correlate two different data sets, e.g., galaxies and quasars
- One can also define n-point correlation functions, $z = \langle d_1 d_2 d_3 \rangle$, $h = \langle d_1 d_2 d_3 d_4 \rangle$ … etc.

Brighter galaxies are clustered more strongly than fainter ones

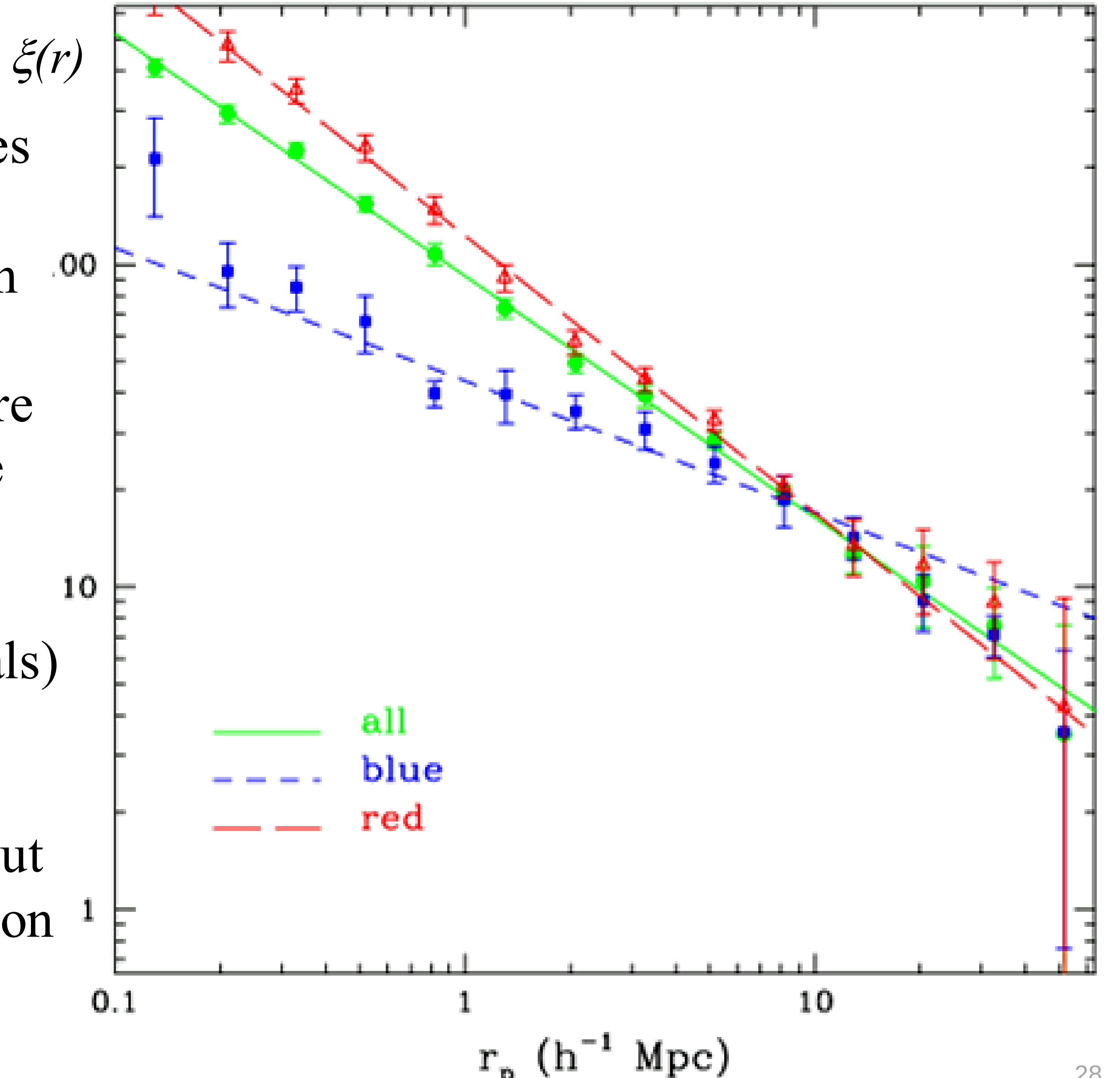
This is telling us something about galaxy formation

Formation and evolution of galaxies and LSS are coupled



Redder galaxies (or early-type, ellipticals, with older stellar populations) are clustered more strongly than bluer ones (or late-type, spirals)

That, too, says something about galaxy formation



Correlation Function and Power Spectrum

- Given the overdensity field $\delta(\mathbf{x}) = \dfrac{n(\mathbf{x})}{\langle n \rangle} - 1$
- Its Fourier transform is $\delta(\mathbf{x}) = \dfrac{1}{(2\pi)^3} \int d^3\mathbf{k}\, e^{i\mathbf{kx}} \delta(\mathbf{k})$
- Its inverse transform is $\delta(\mathbf{k}) = \int d^3\mathbf{x}\, e^{-i\mathbf{kx}} \delta(\mathbf{x})$

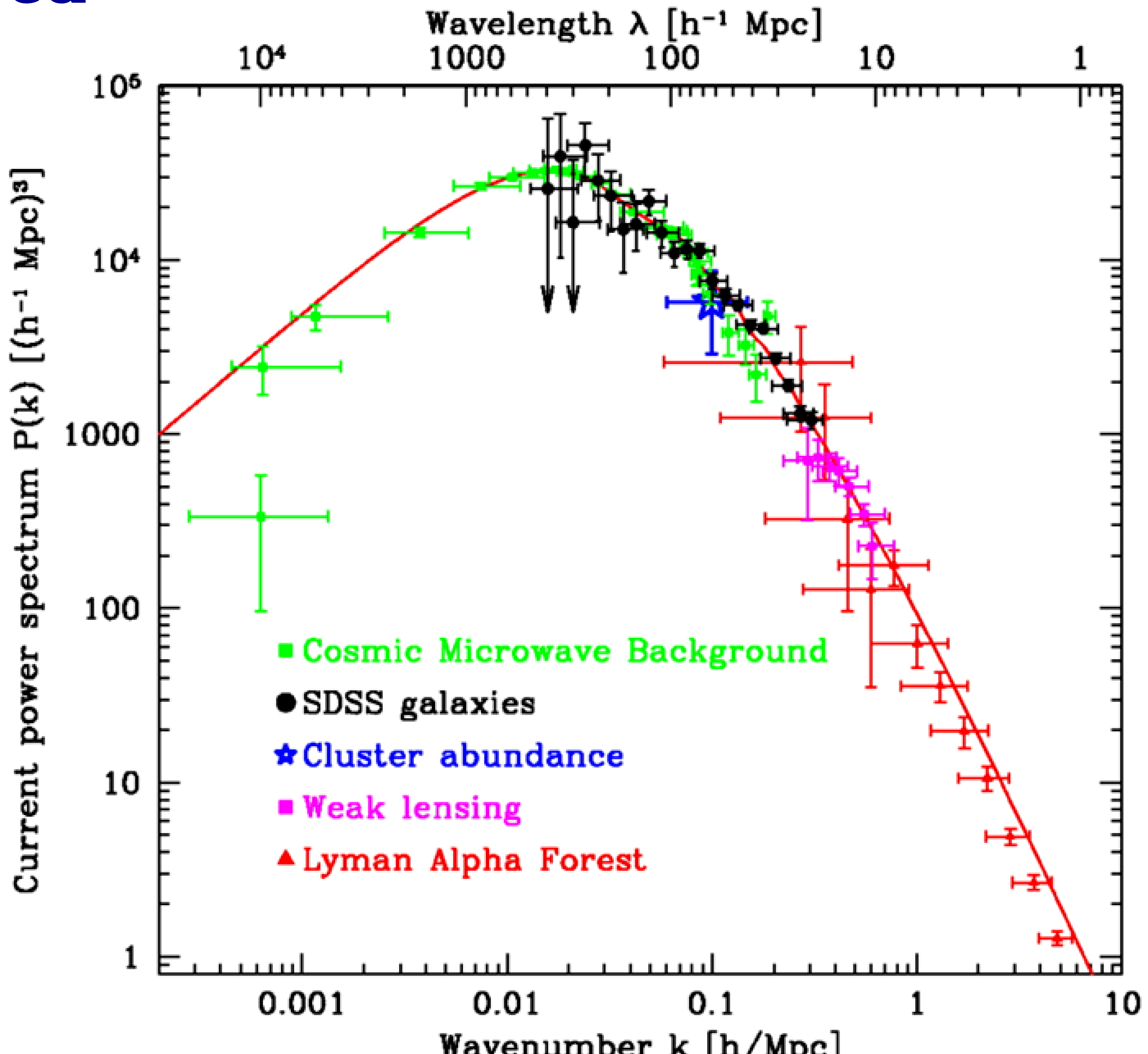
  where $k = \dfrac{2\pi}{\lambda}$ is the wave number
- The power spectrum is $P(\mathbf{k}) = \left|\delta(\mathbf{k})\right|^2$
- Then $\xi(r) = \dfrac{1}{4\pi^2} \int d\ln k\; j_0(kr) \left[k^3 P(k)\right]$

Correlation function and power spectrum are a Fourier pair

The Observed Power Spectrum

Power spectrum is more directly comparable to theoretical models of LSS formation

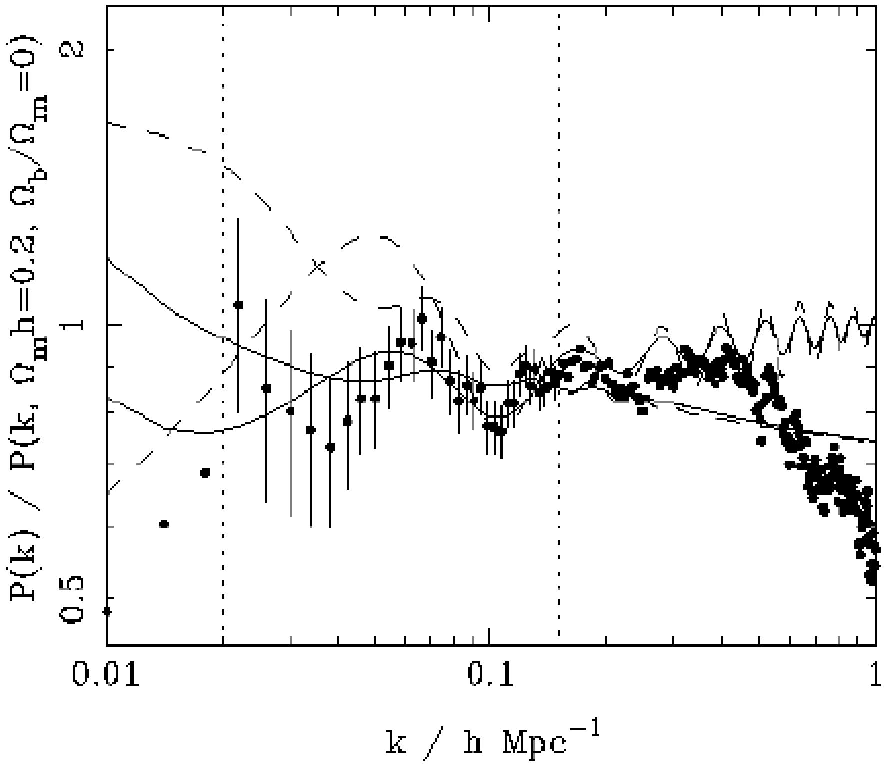
Use different probes at different physical scales



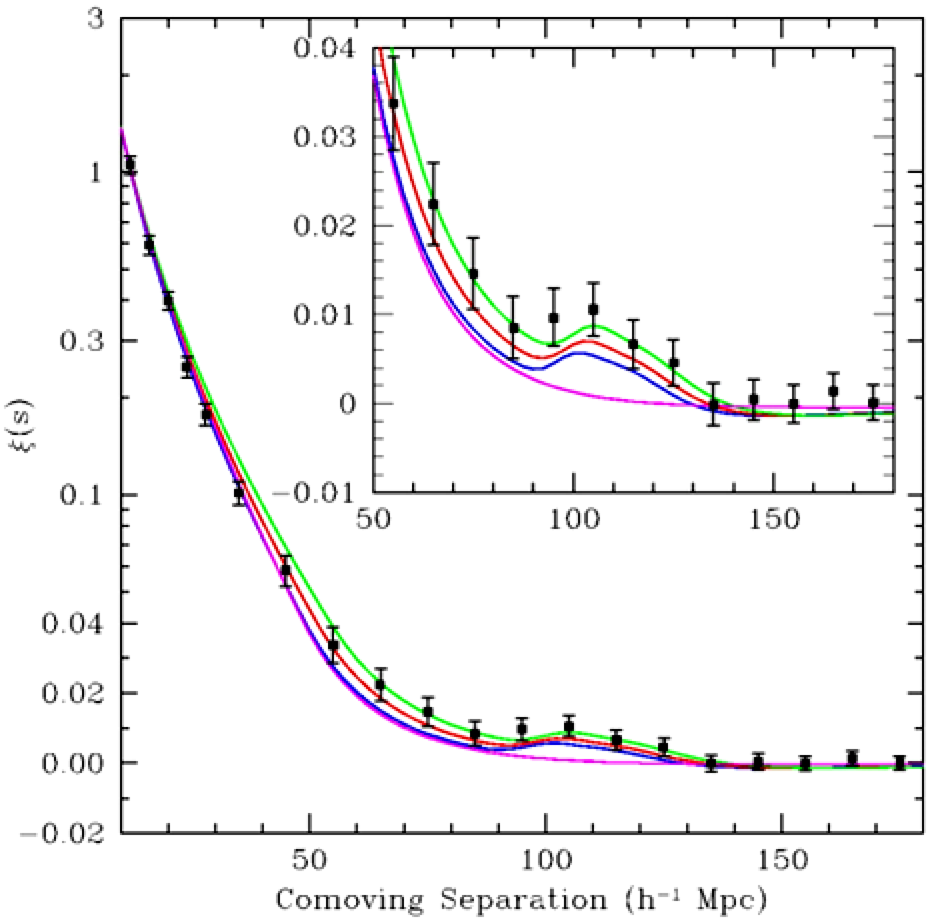
(Tegmark et al.)

Baryonic oscillations seen in the CMBR are detected in the LSS at lower redshifts

Thus, we can use the first peak as a stardard ruler at more than one redshift



2dF (Percival et al.)



SDSS (Eisenstein et al.)

Normalizing the Power Spectrum

- Define σ_R as the *r.m.s. of mass fluctuations* on the scale R
- Typically a sphere with a radius $R = 8\ h^{-1}\ Mpc$ is used, as it gives $\sigma_8 \approx 1$ (actually, closer to 0.8…)
- So, the amplitude of $P(k)$ is ~ 1 at k = 2π / (8 h^{-1} Mpc)
- This is often used to normalize the spectrum of the PDF
- Mathematically, where K_R is a convolving kernel, a spherical top-hat with a radius R:

$$\sigma_R^2 = \frac{1}{4\pi^2} \int d\ln k \left[k^3 P(k) \left| K_R(k) \right|^2 \right]$$

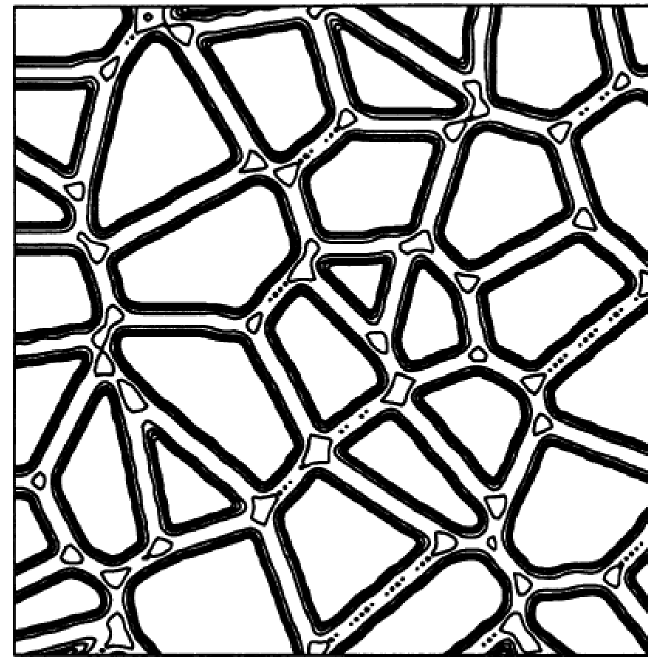
$$K_R(r) = \begin{cases} 1, & \text{if } r < R \\ 0, & \text{if } r \geq R \end{cases} \qquad K_R(k) = \left[\frac{j_1(kr)}{kr} \right]$$

- Alternatively, we can use the CMB fluc's to normalize it

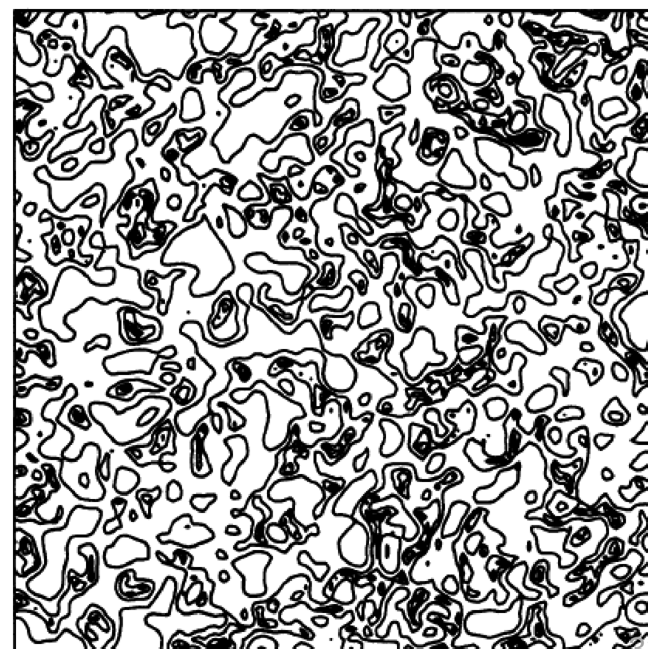
Is the Power Spectrum Enough?

These two images have ***identical power spectra*** (by construction)

The power spectrum alone does not capture the *phase information*: the coherence of cosmic structures (voids, walls, filaments …)



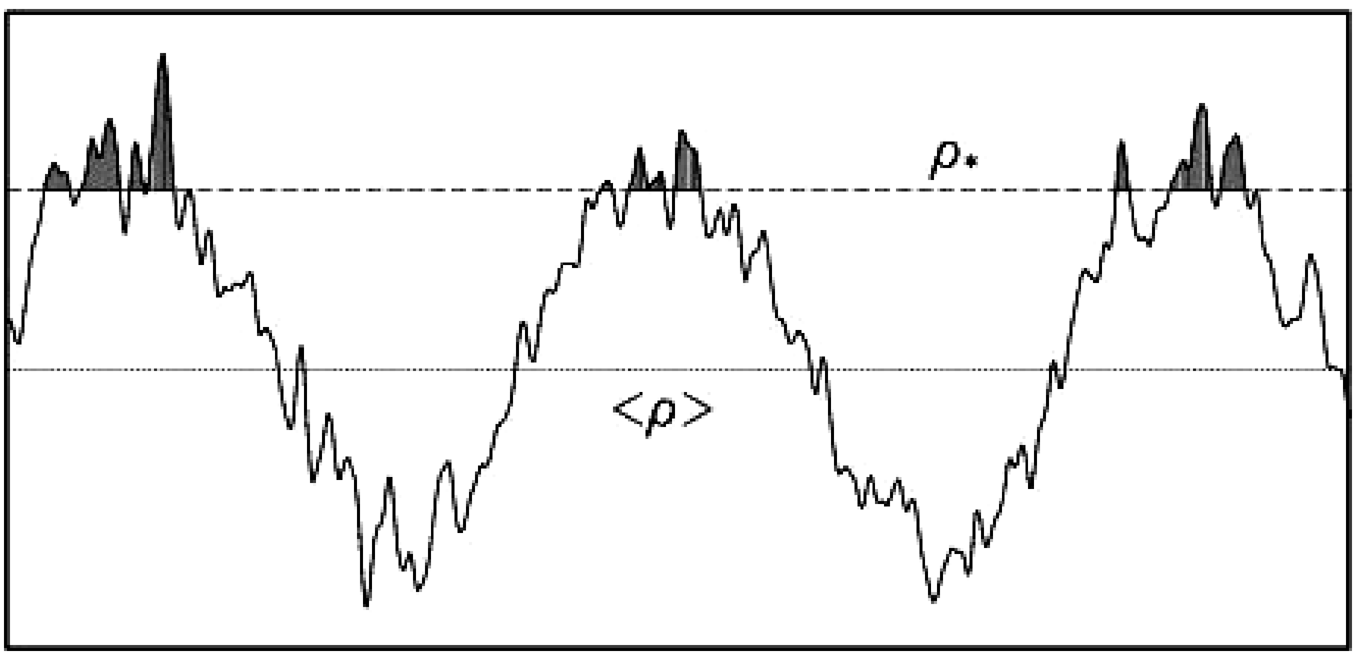
Voronoi foam, R=1.6, smoothed original



Voronoi foam, R=1.6, random phases

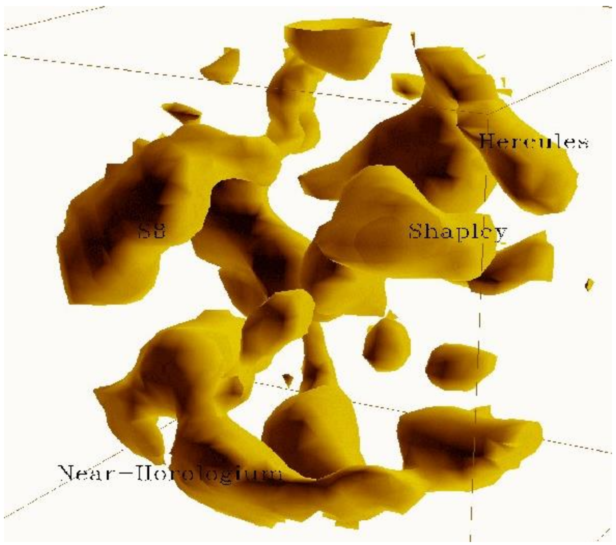
Clustering of Different Structures

- The more massive systems (e.g., elliptical vs. spiral galaxies; groups and clusters of increasing richness) cluster more strongly
- They correspond to increasingly higher peaks of the density field, and thus increasingly rare
- This can be naturally explained with the concept of ***biasing***

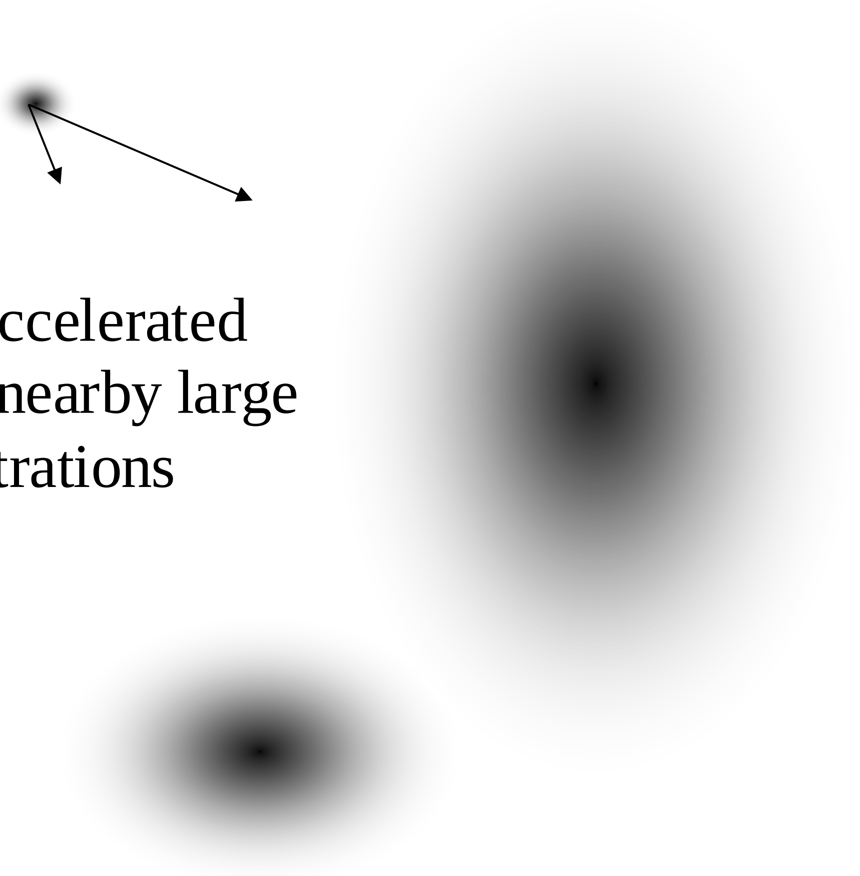


Large-Scale Density Field Inevitably Generates a Peculiar Velocity Field

The PSCz survey
local 3-D density field



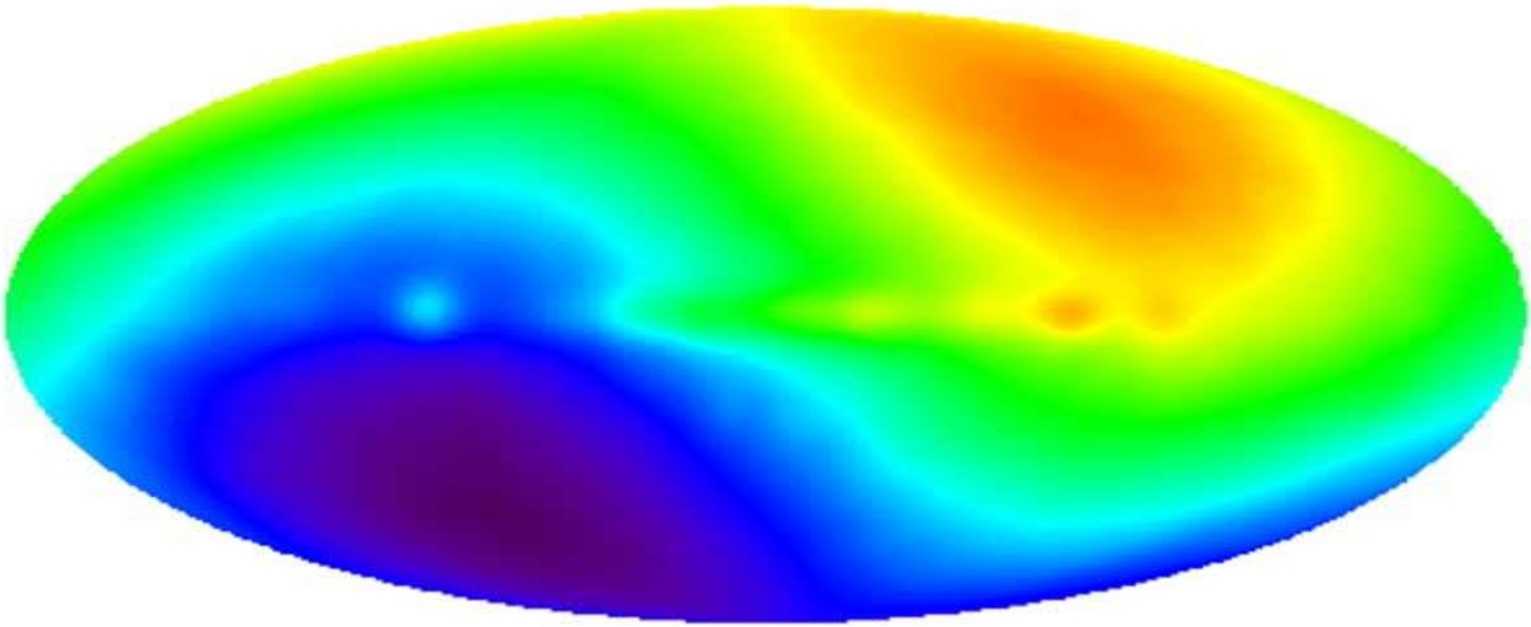
A galaxy is accelerated towards the nearby large mass concentrations



Integrated over the Hubble time, this results in a peculiar velocity

The pattern of peculiar velocities should thus reflect the underlying mass density field

CMBR Dipole:
The One Peculiar Velocity We Know Very Well



We are moving wrt. to the CMB at ~ 620 km/s towards *b=27°*, *l=268°*. This gives us an idea of the probable magnitude of peculiar velocities in the local universe. Note that at the distance to Virgo (LSC), this corresponds to a ~ 50% error in Hubble velocity, and a ~ 10% error at the distance to Coma cluster.

How to Measure Peculiar Velocities?

1. Using distances and residuals from the Hubble flow:

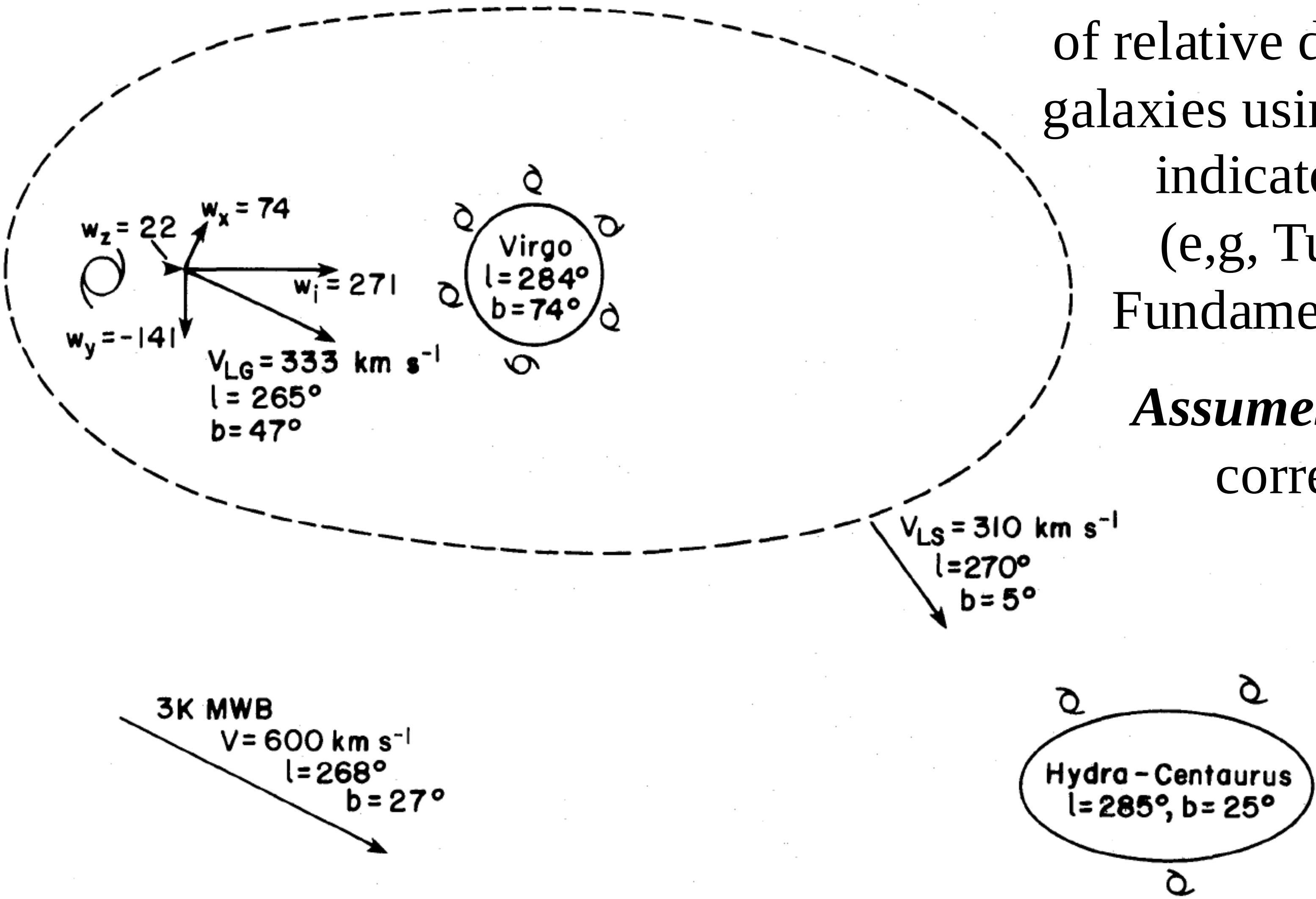
$$V_{total} = V_{Hubble} + V_{pec} = H_0 D + V_{pec}$$

- So, if you know relative distances, e.g., from Tully-Fisher, or D_n-σ relation, SBF, SNe, …you could derive peculiar velocities
- A problem: distances are seldom known to better than ~10% (or even 20%), multiply that by V_{Hubble} to get the error of V_{pec}
- Often done for clusters, to average out the errors
- But there could be systematic errors - distance indicators may vary in different environments

2. Statistically from a redshift survey

- Model-dependent

Virgo Infall, and the Motion Towards the Hydra-Centaurus Supercluster

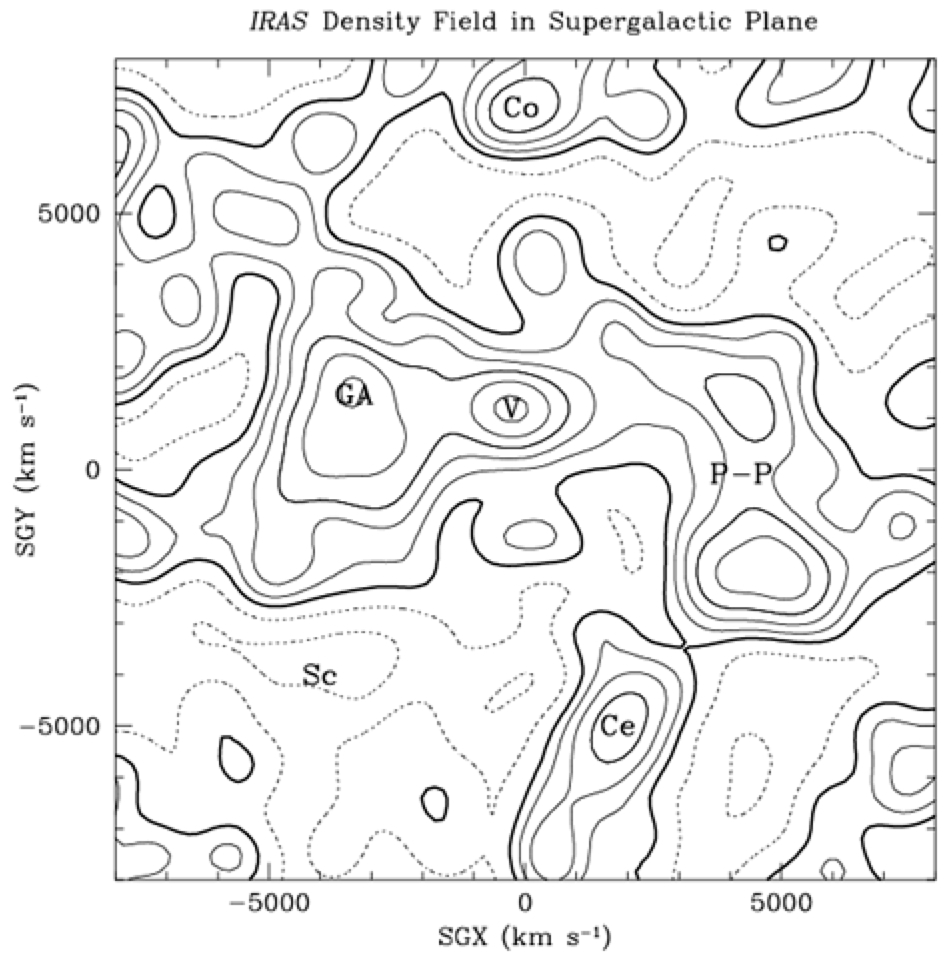


Based on the measurements of relative distances to galaxies using distance indicator relations (e,g, Tully-Fisher, Fundamental Plane)

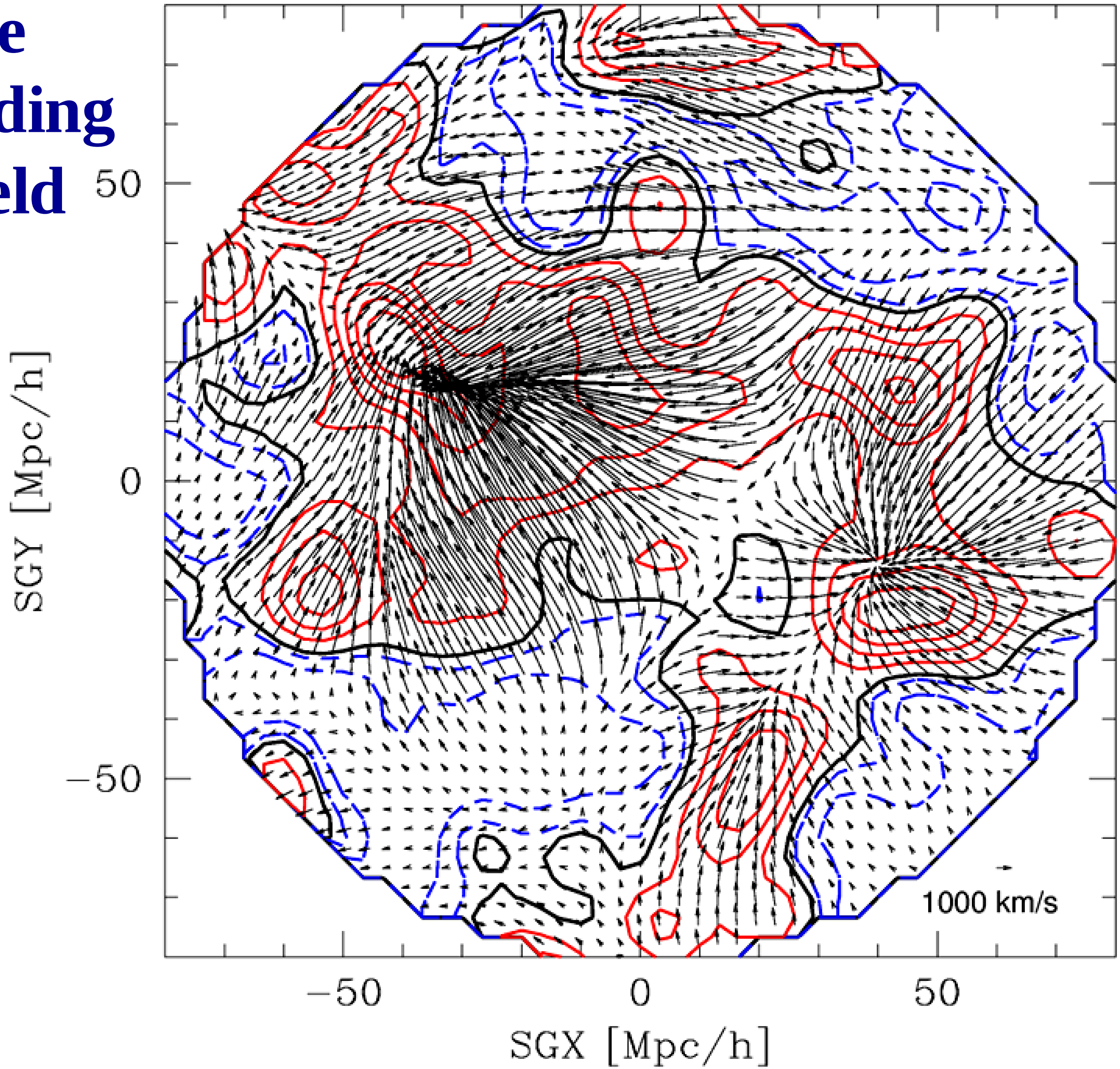
Assumes that these correlations are universal

Measuring Peculiar Velocity Field Using a Redshift Survey

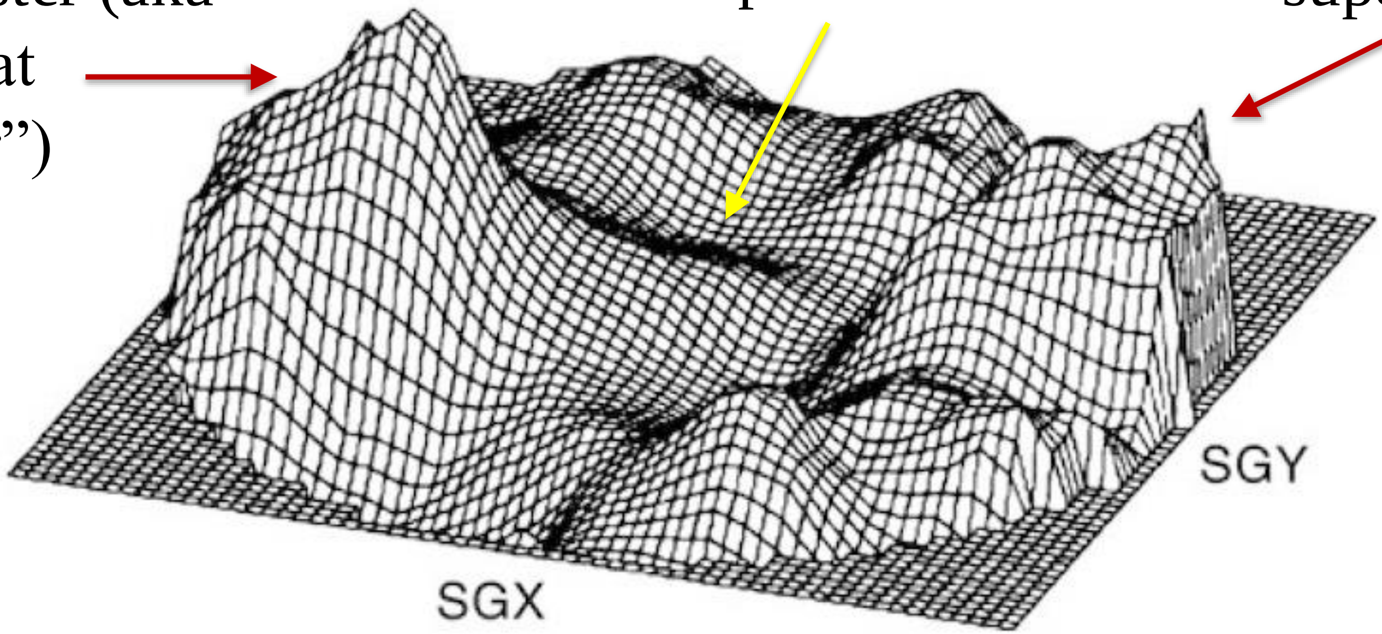
- Assume that galaxies are where their redshifts imply; this gives you a density field
- You *need a model on how the light traces the mass*
- Evaluate the accelerations for all galaxies, and their estimated peculiar velocities
- Update the positions according to new Hubble velocities
- Iterate until the convergence
- You get *consistent density and velocity fields*



PSCz: The corresponding velocity field



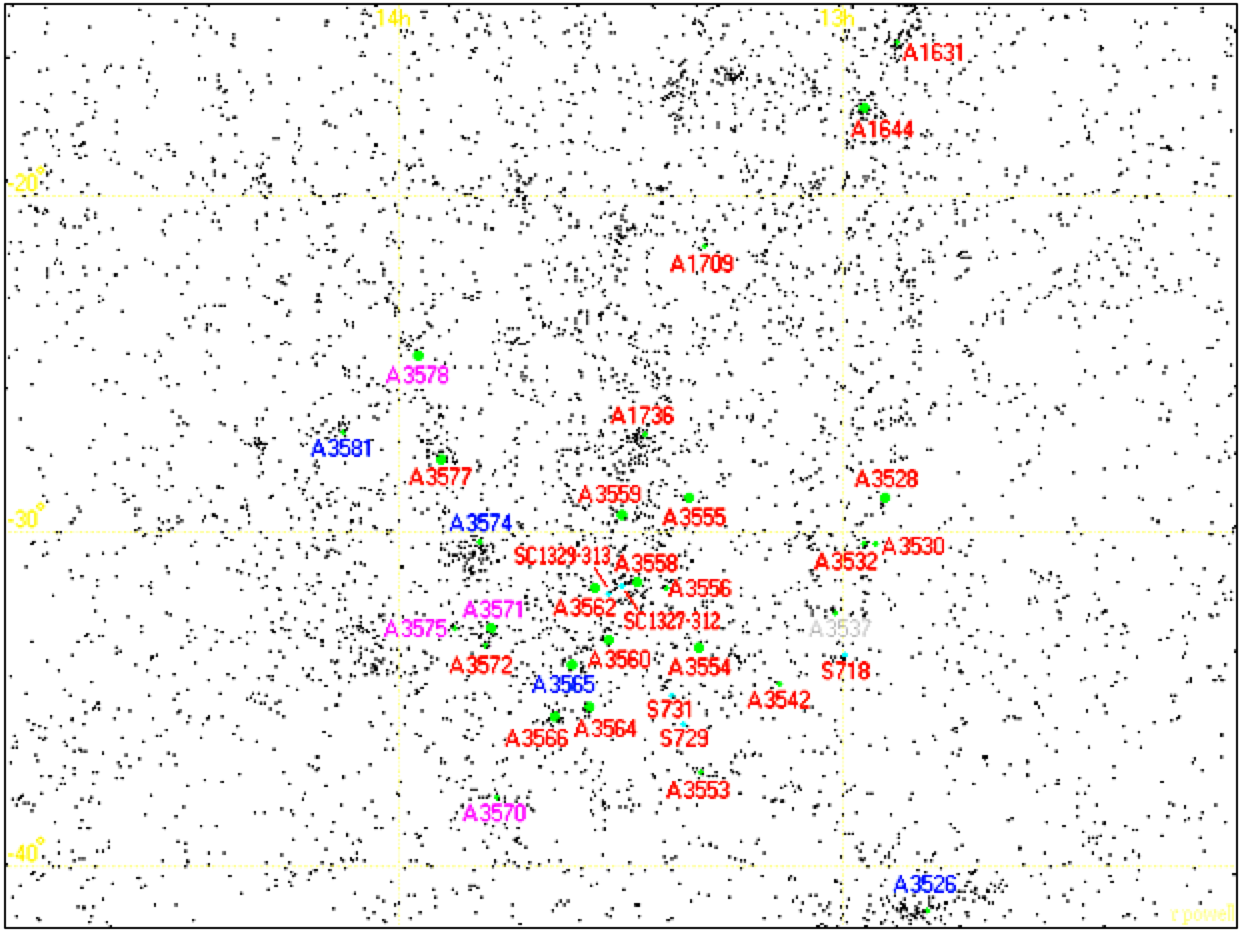
Local Density and Velocity Fields From Peculiar Velocities of Galaxies



This assumes that the distance indicator relations do not vary in different environments, and that the distribution of light and mass is the same everywhere

The Flow Continues?

The Shapley Concentration of clusters at ~ 200 Mpc, beyond the Hydra-Centaurus may be responsible for at least some of the large-scale bulk flow



LSS Observations Summary

- **A range of structures**: galaxies (~10 kpc), groups (~ 0.3 - 1 Mpc), clusters (~ few Mpc), superclusters (~ 10 - 100 Mpc)
- **Redshift surveys** are used to map LSS; ~ 5×10^8 galaxies now
- LSS **topology** is prominent: voids, sheets, filaments…
- LSS quantified through 2-point (and higher) **correlation function**(s), well fit by a **power-law:**

  typical $\gamma \sim 1.8$, $r_0 \sim$ few Mpc $$\xi(r) = (r / r_0)^{-\gamma}$$
- Equivalent description: **power spectrum** *P(k)* - useful for comparisons with the theory
- Objects of different types have different clustering strengths: **more massive structures cluster more strongly**
- Mass density field generates the peculiar velocity field, but its measurements require some assumptions