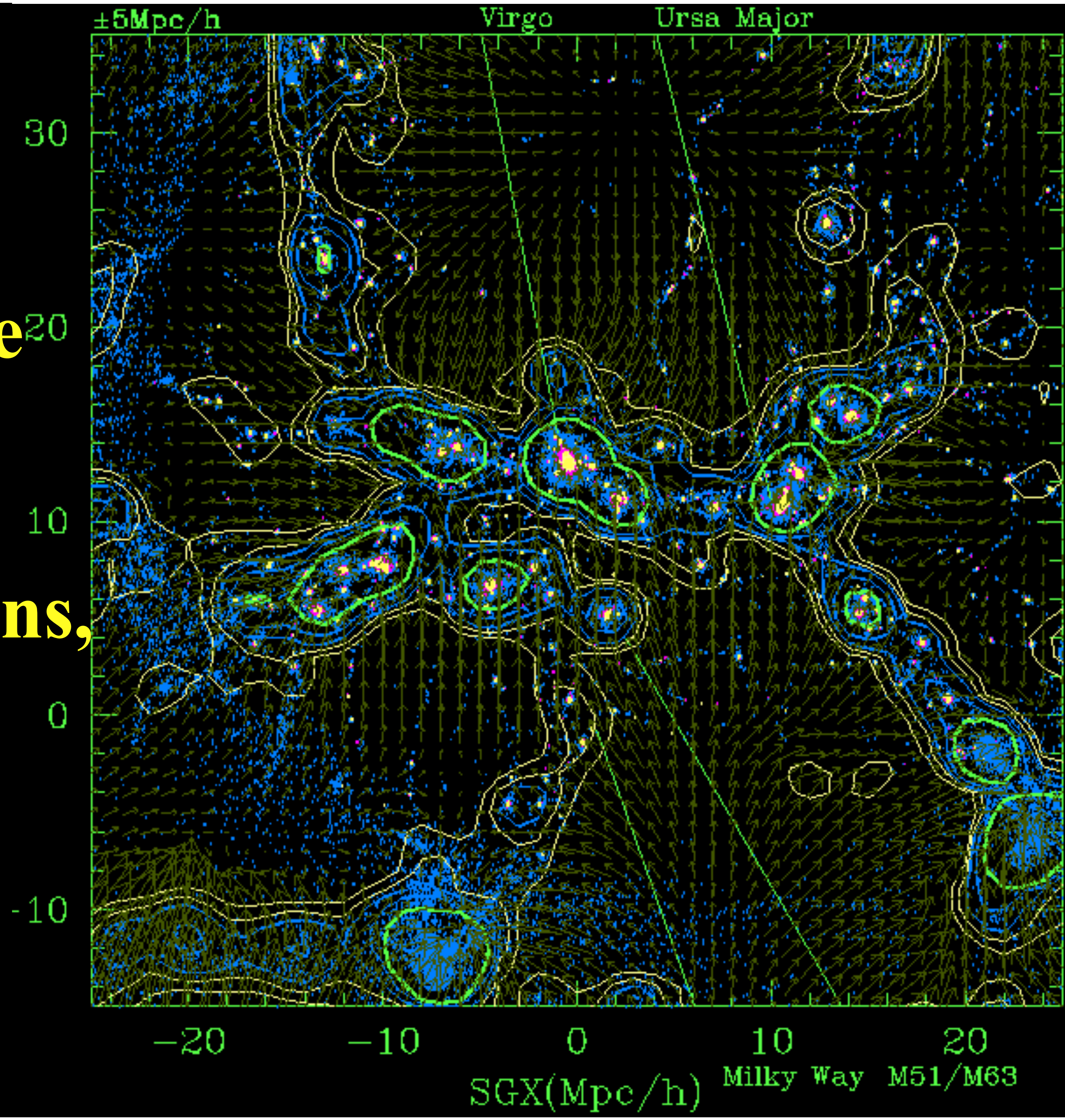
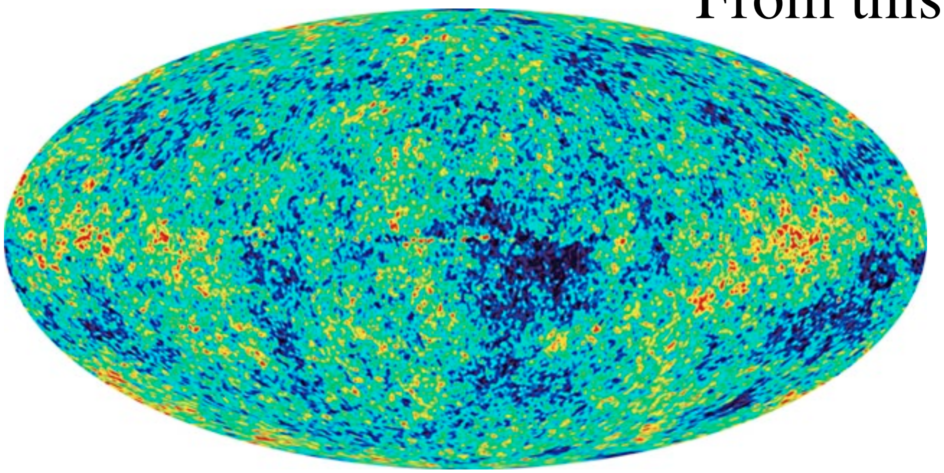
Ay 127

Large Scale Structure: Basic Observations, Redshift Surveys, Biasing

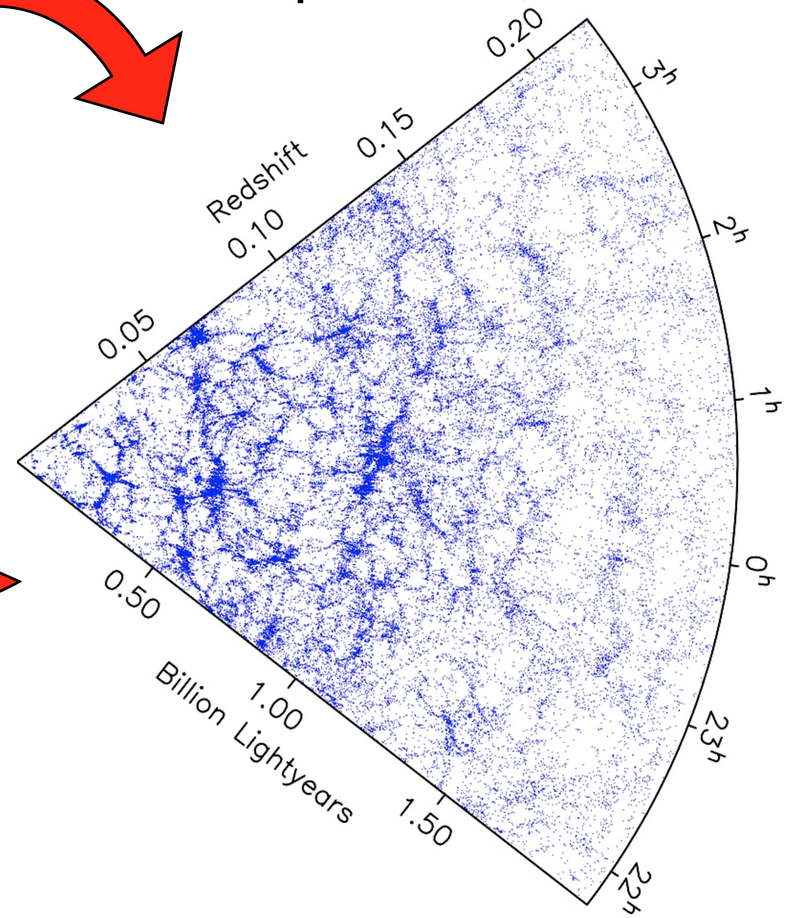


Structure Formation and Evolution

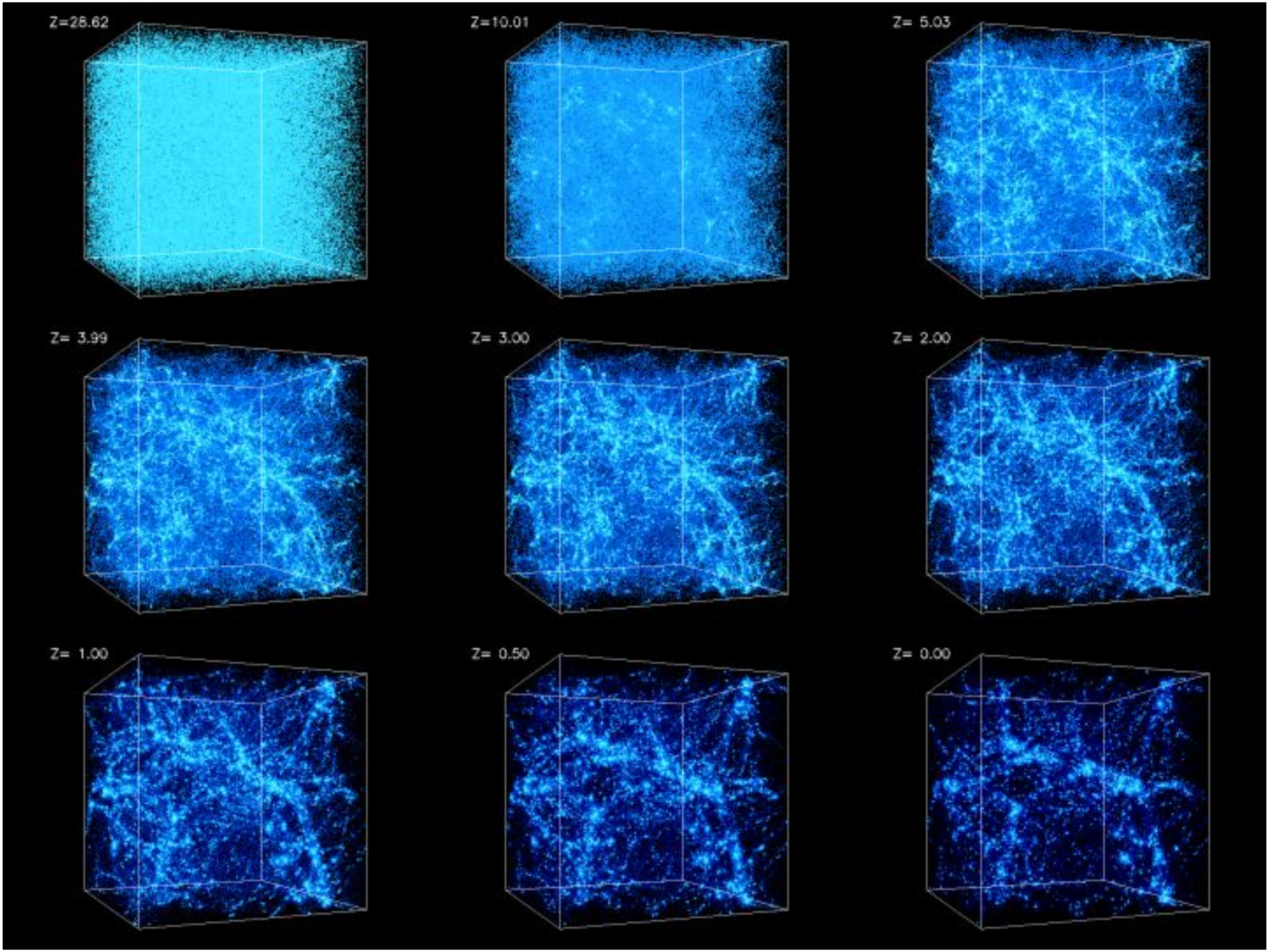
From this ($\Delta\rho/\rho \sim 10^{-6}$)



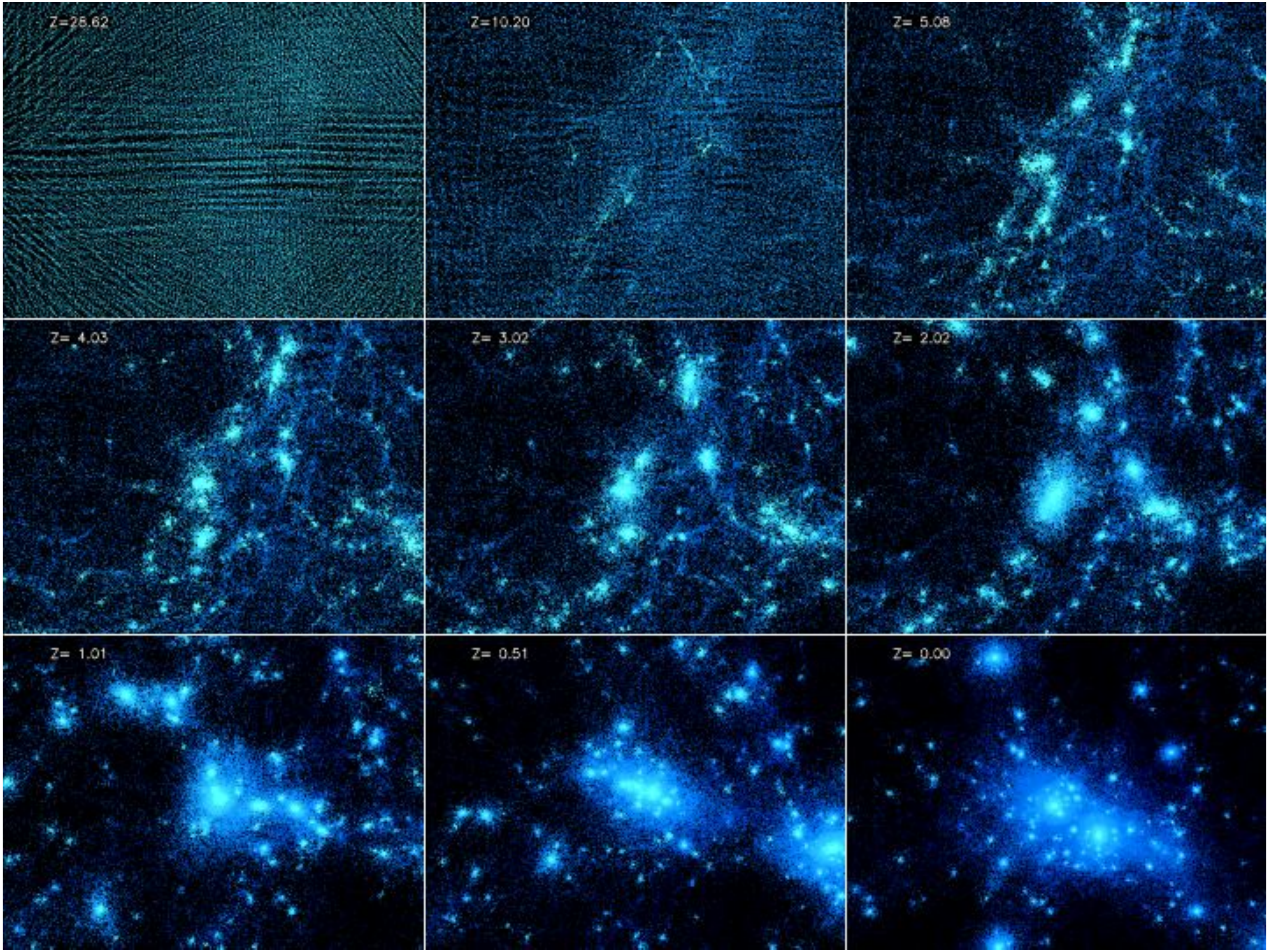
to this ($\Delta\rho/\rho \sim 10^{+2}$)



to this ($\Delta\rho/\rho \sim 10^{+6}$)

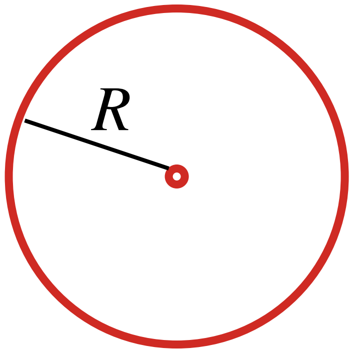


Z=28.62
Z=10.01
Z= 5.03
Z= 3.99
Z= 3.00
Z= 2.00
Z= 1.00
Z= 0.50
Z= 0.00


Z=28.62
Z=10.20
Z= 5.08
Z= 4.03
Z= 3.02
Z= 2.02
Z= 1.01
Z= 0.51
Z= 0.00

How Long Does It Take?

The (dissipationless) gravitational collapse timescale is on the order of the free-fall time, t_{ff}:



The outermost shell has acceleration $g = GM/R^2$
It falls to the center in:

$$t_{ff} = (2R/g)^{1/2} = (2R^3/GM)^{1/2} \approx (2/G\rho)^{1/2}$$

Thus, low density lumps collapse more slowly than high density ones. More massive structures are generally less dense, take longer to collapse. For example:

For a galaxy: $t_{ff} \sim 600$ Myr $(R/50\text{kpc})^{3/2}$ $(M/10^{12}M_\odot)^{-1/2}$

For a cluster: $t_{ff} \sim 9$ Gyr $(R/3\text{Mpc})^{3/2}$ $(M/10^{15}M_\odot)^{-1/2}$

So, we expect that galaxies collapsed early (at high redshifts), and that clusters are still forming now. This is as observed!

Large-Scale Structure

- Density fluctuations evolve into structures we observe: galaxies, clusters, etc.
- 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 ~ a million galaxies
- 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

Discovery of the Large Scale Structure

1930' s: H. Shapley, F. Zwicky, and collab.
1950' s: Donald Shane, Carl Wirtanen, others
1950' s - 1970' s: Gerard de Vaucouleurs, first redshift surveys
1970' s - 1980' s: CfA, Arecibo, and other redshift surveys

Zwicky et al., galaxy overdensities map

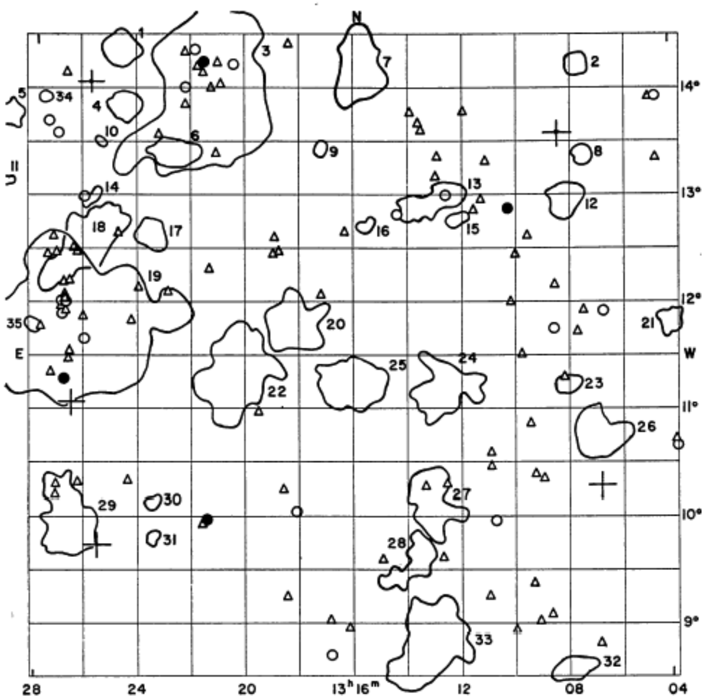
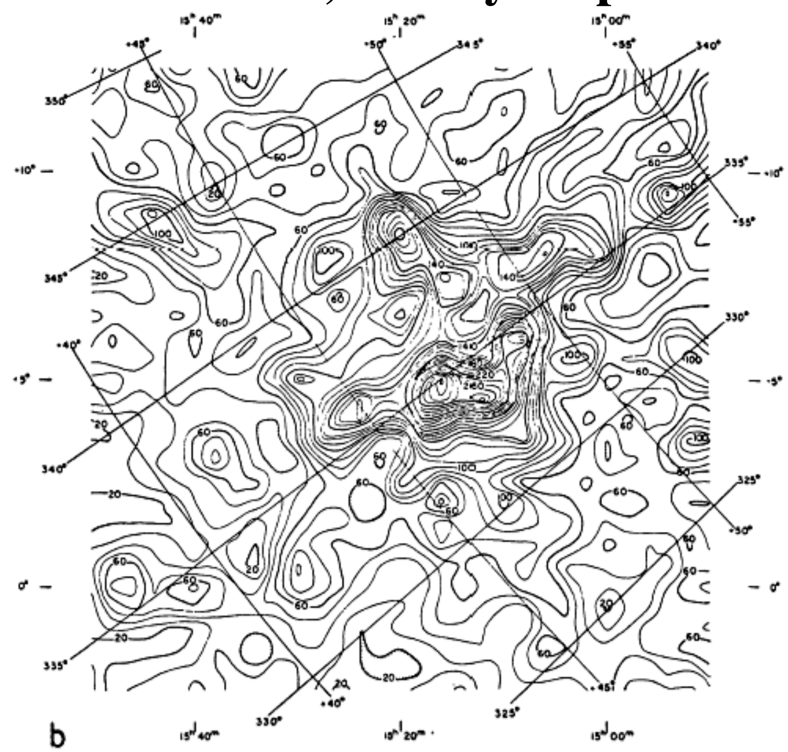


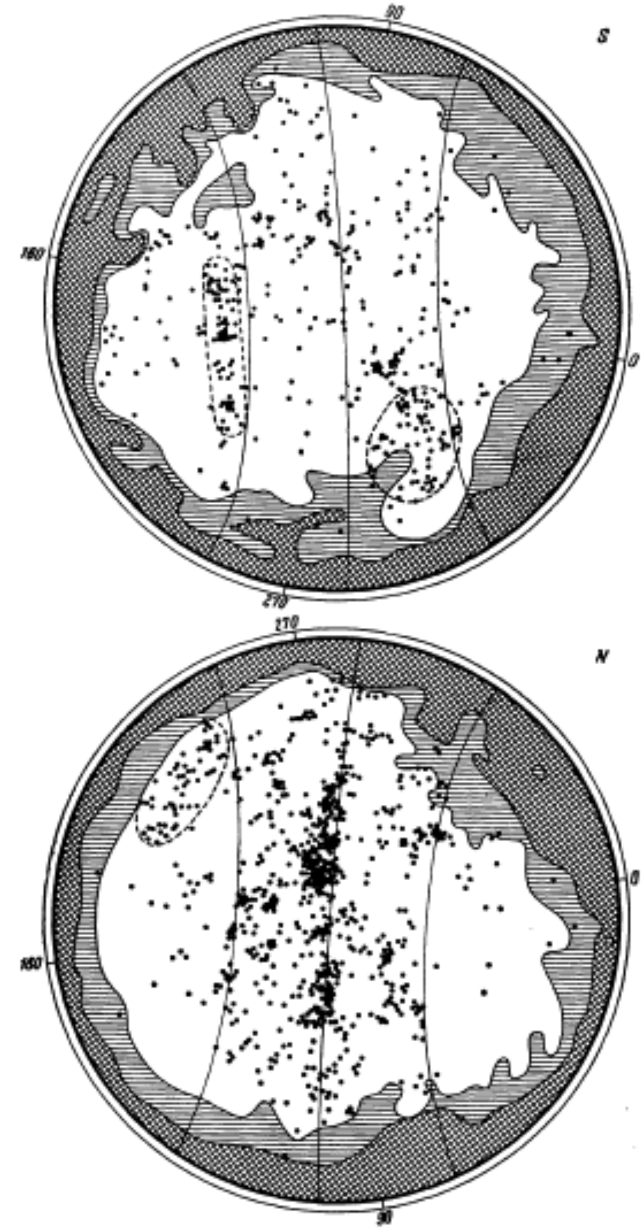
FIGURE 4

Shane-Wirtanen galaxy counts, density map

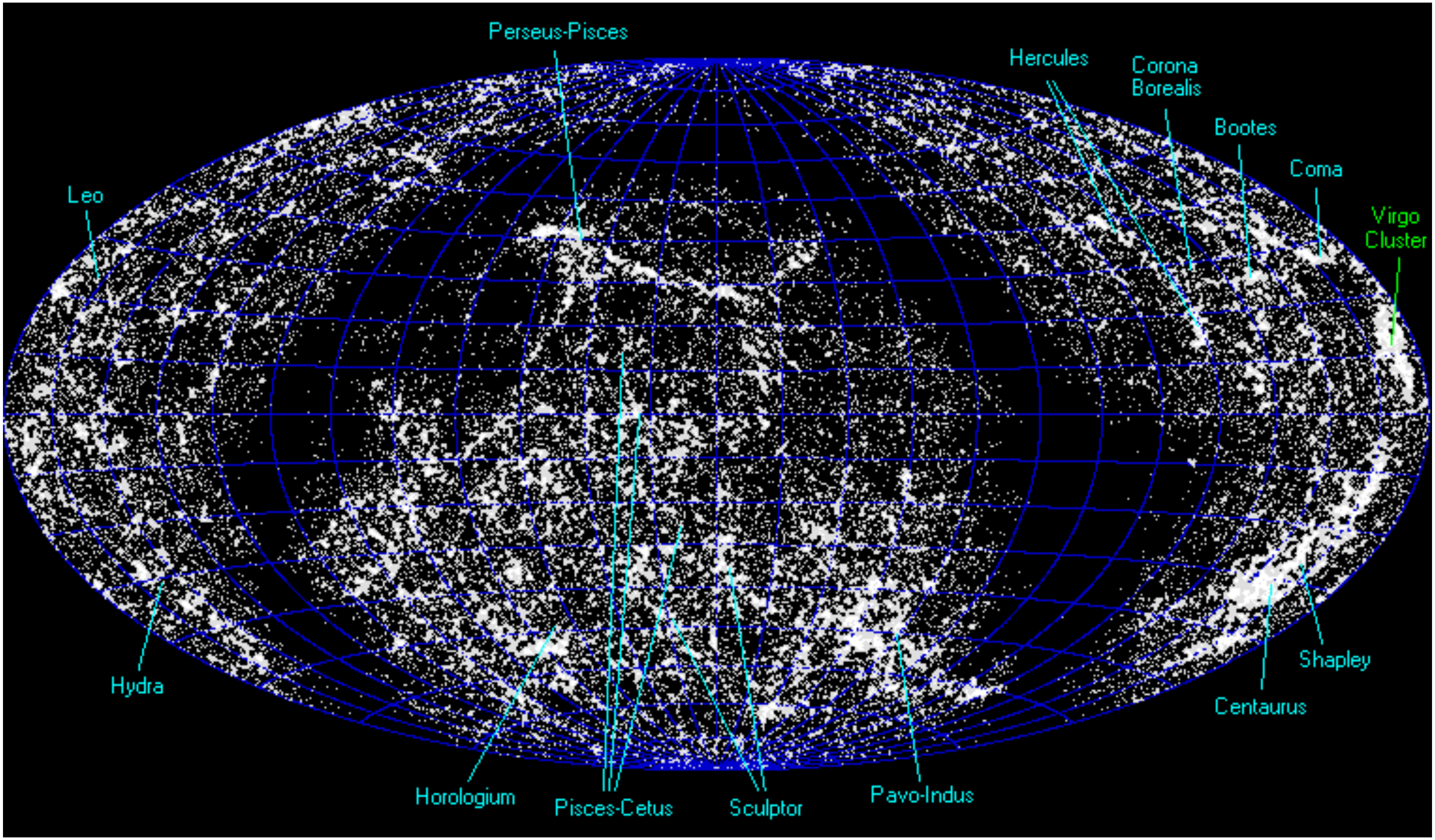


The Local Supercluster

- 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

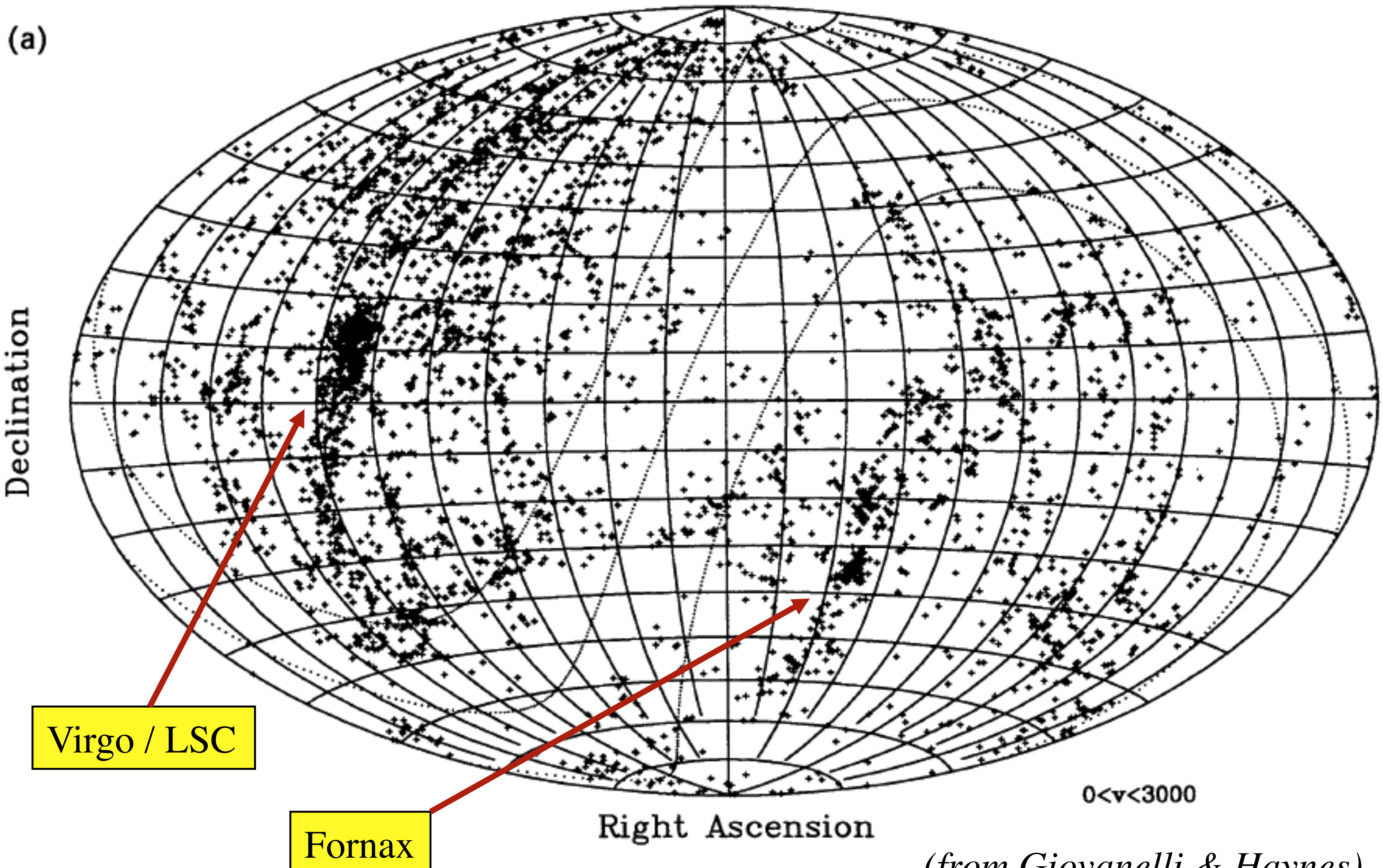


6000 Brightest Galaxies on the Sky



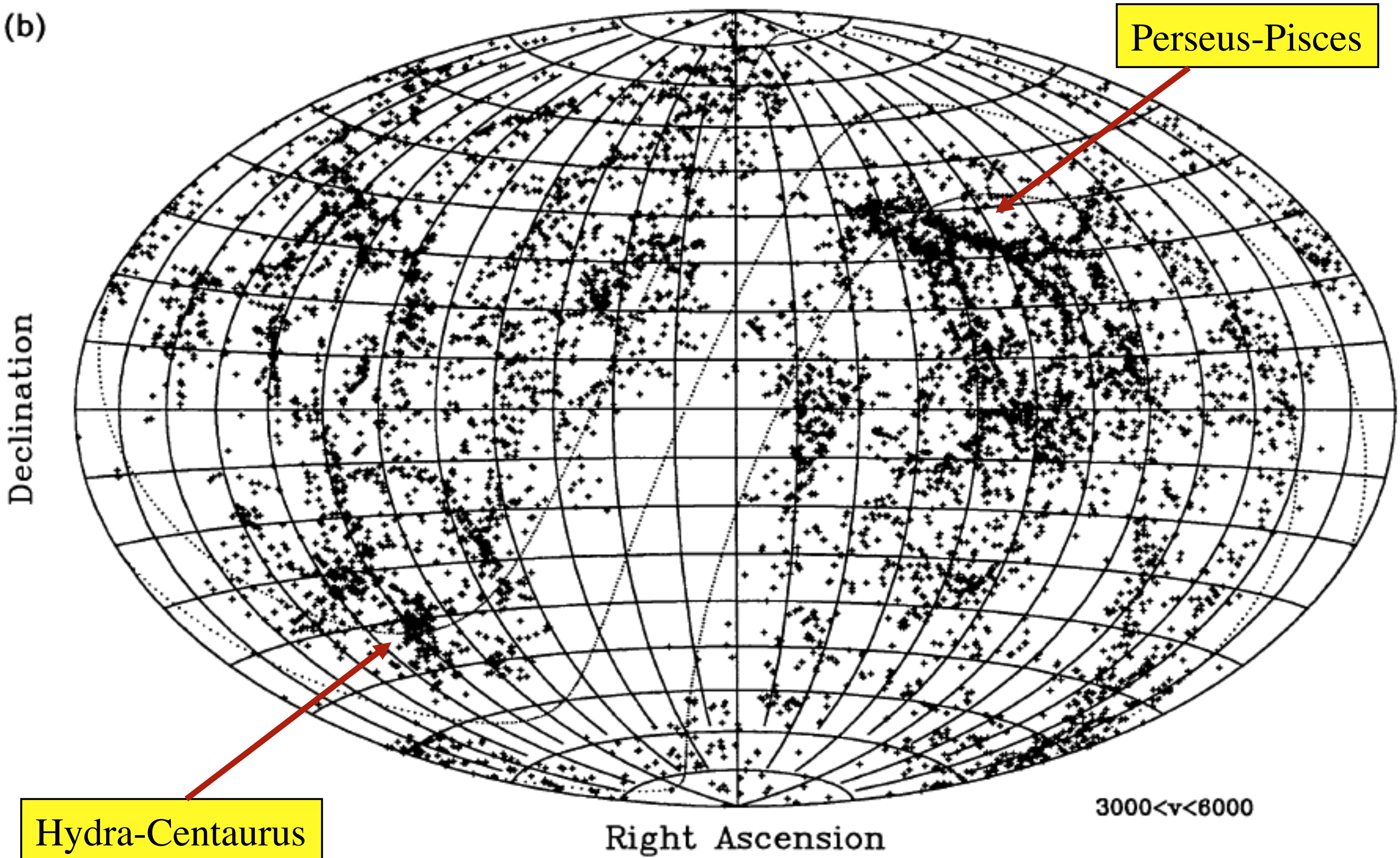
How would the picture of the 6000 brightest stars on the sky look?

Stepping Out in Redshift Slices



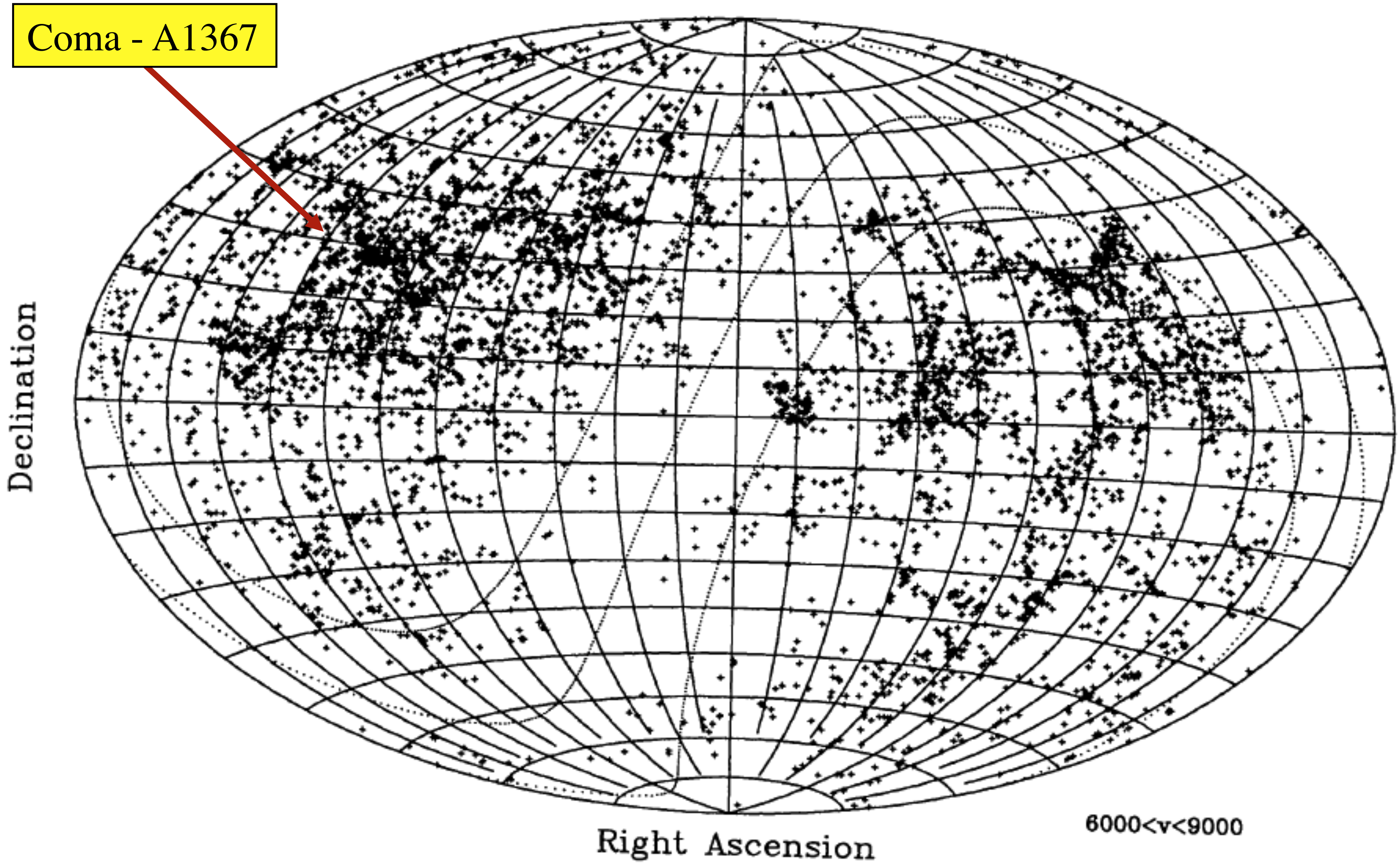
(from Giovanelli & Haynes)

Stepping Out in Redshift Slices



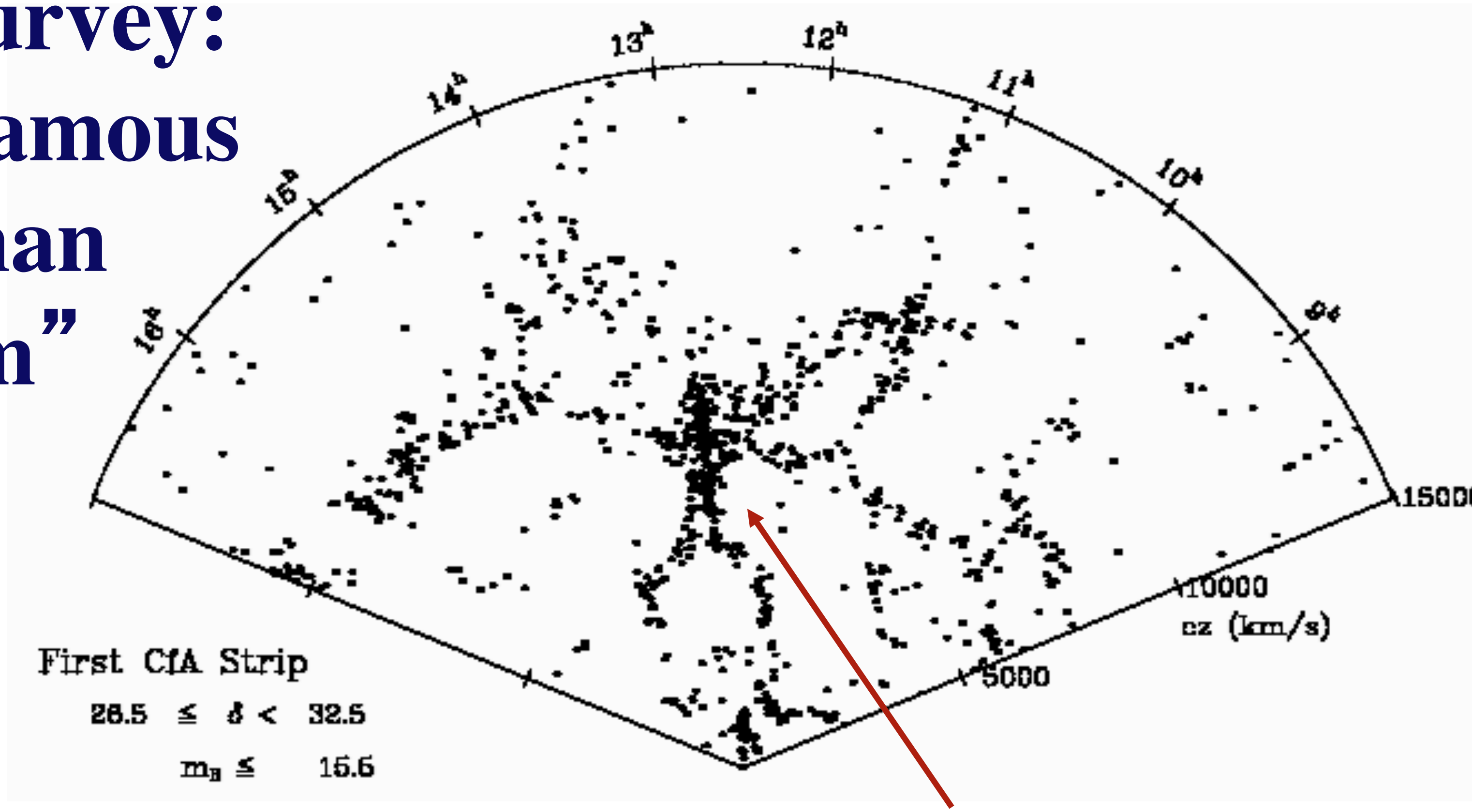
(from Giovanelli & Haynes)

Stepping Out in Redshift Slices



(from Giovanelli & Haynes)

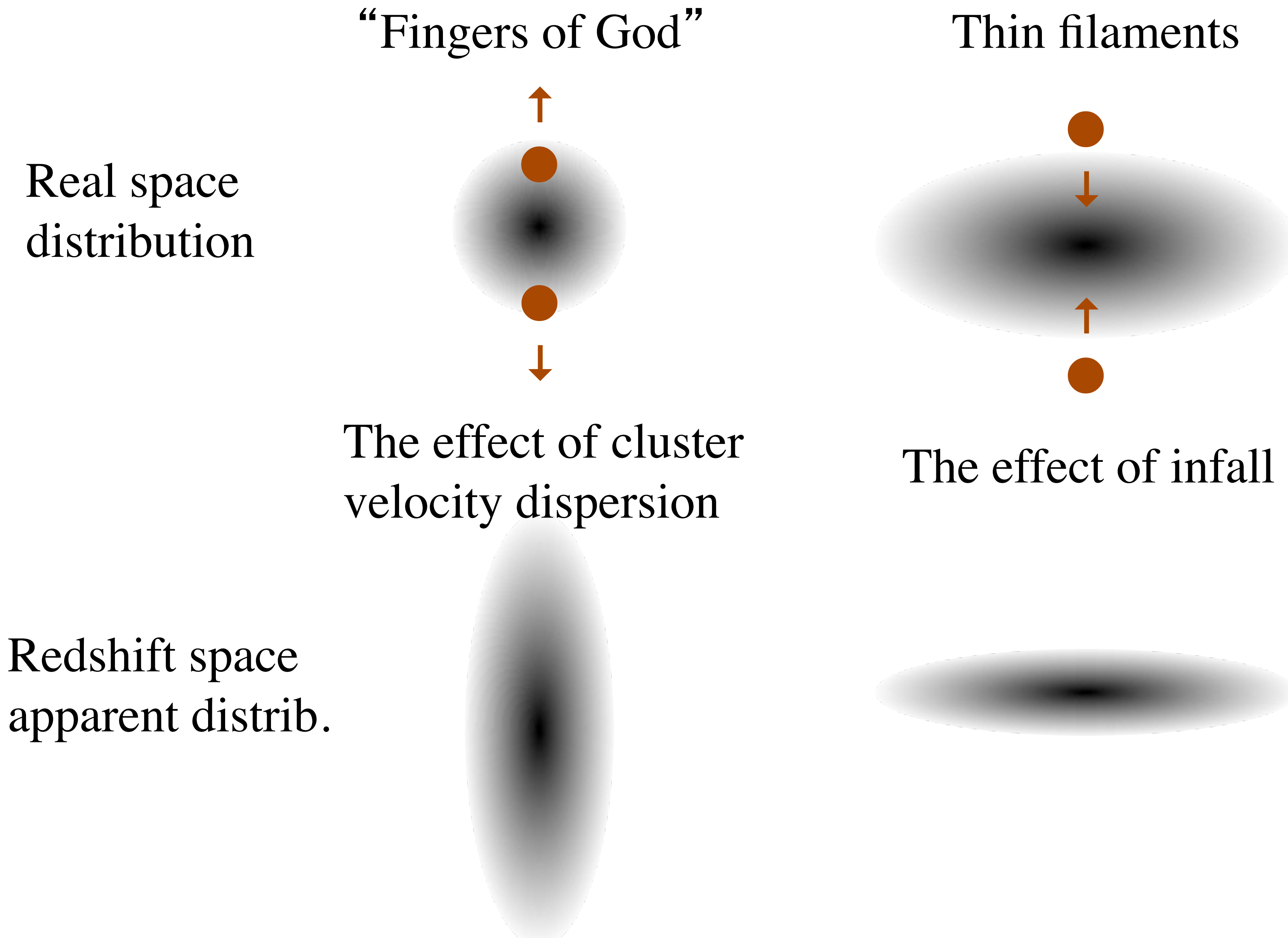
CfA2 Survey: The Infamous "Stickman Diagram"



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

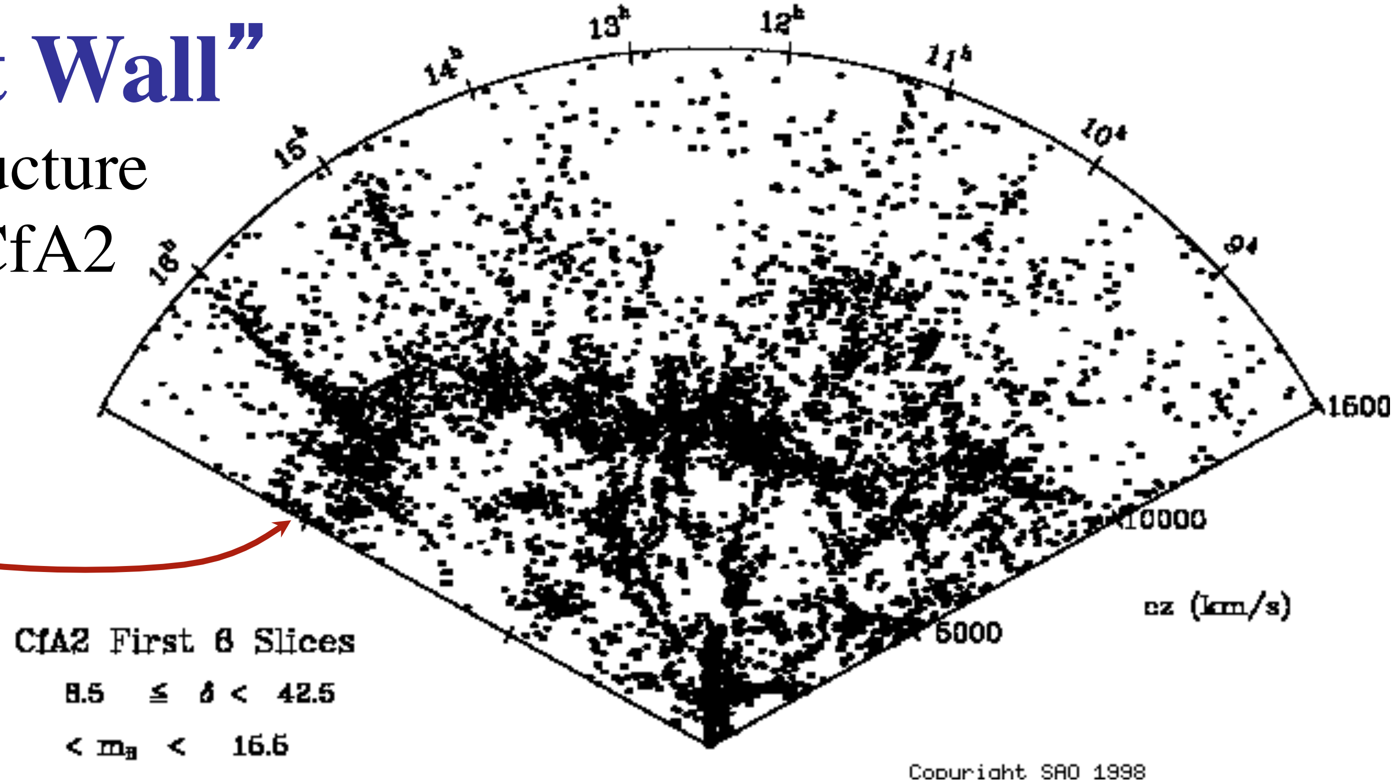
The 2nd generation redshift surveys were often done in slices which were thin in Dec and long in RA, thus sampling a large dynamical range of scales. This also helped reveal the large-scale topology (voids, walls, filaments).

Redshift Space vs. Real Space



The “Great Wall”

a ~ 100 Mpc structure revealed in the CfA2 redshift survey

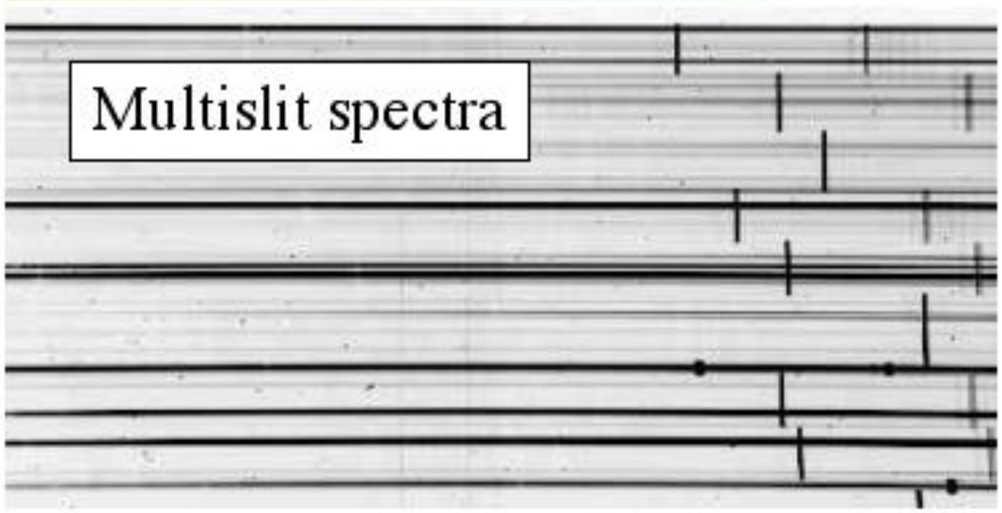


Up until then, redshift surveys revealed structures as large as can be fitted within the survey boundaries - but 100 Mpc turned out to be about as large as they come.

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

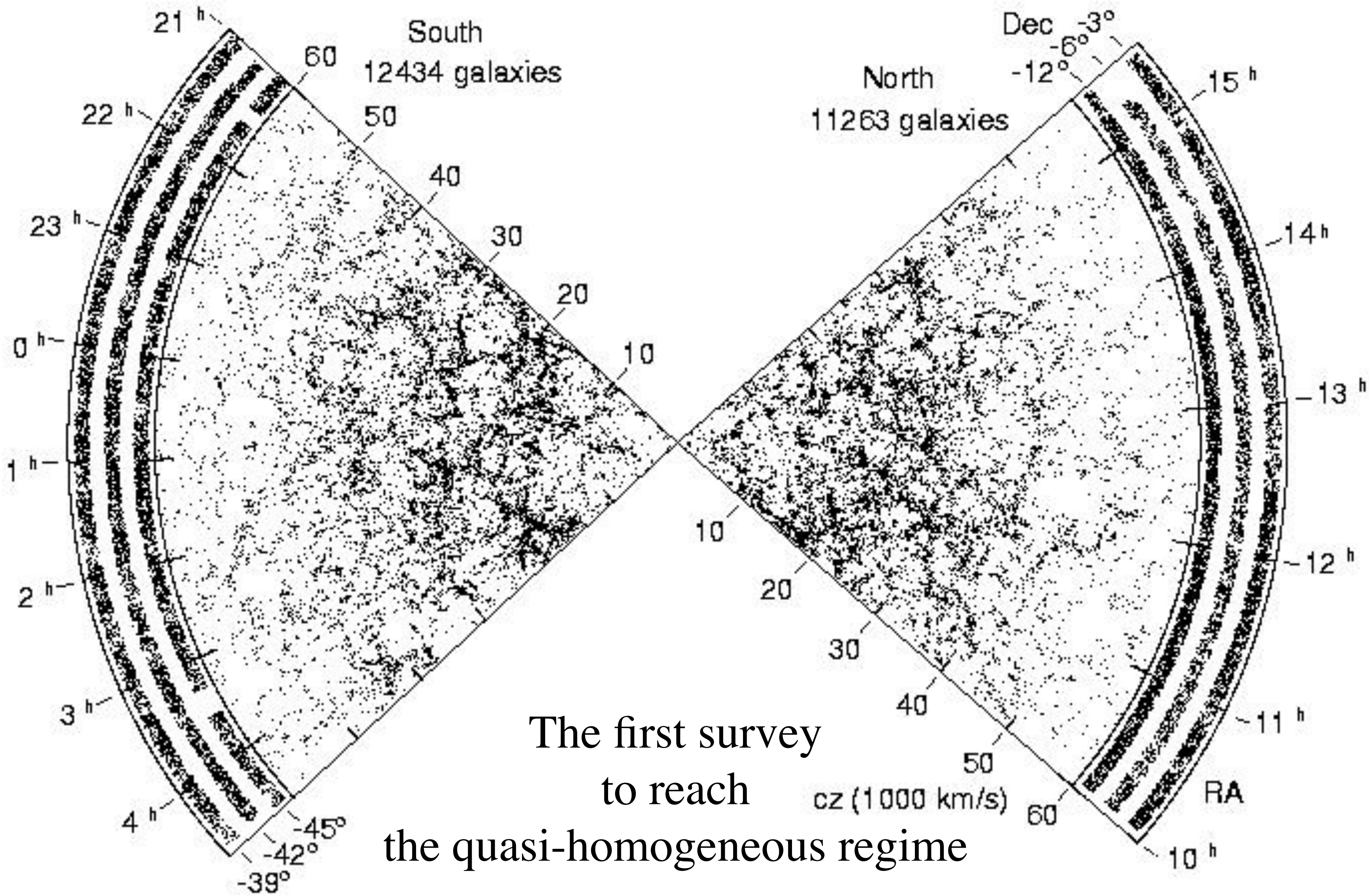
Keck DEIMOS slitmask (detail)



Multislit spectra

2dF multifiber spectrograph

Las Campanas Redshift Survey



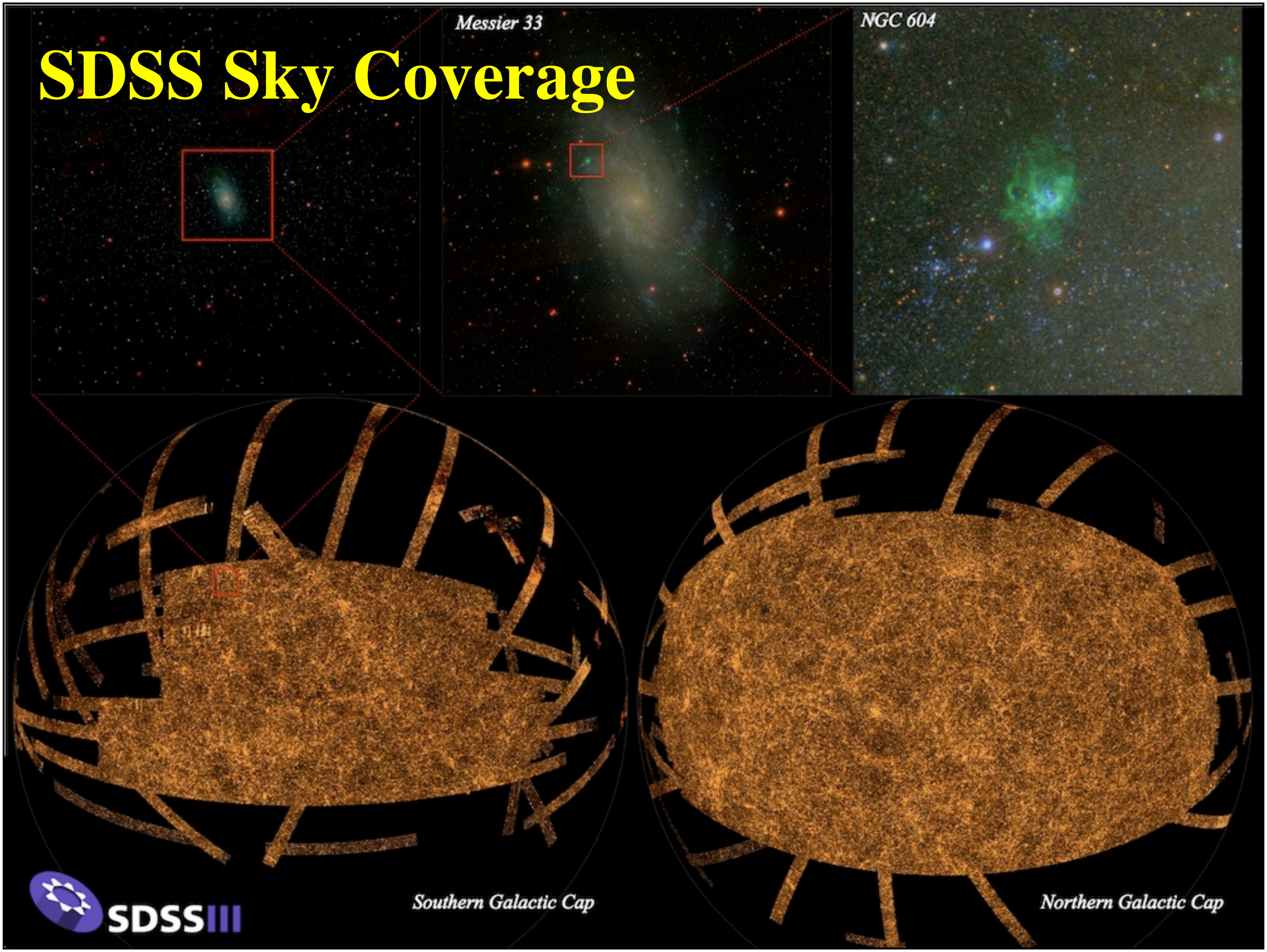
The first survey
to reach
the quasi-homogeneous regime

Huge Redshift Surveys

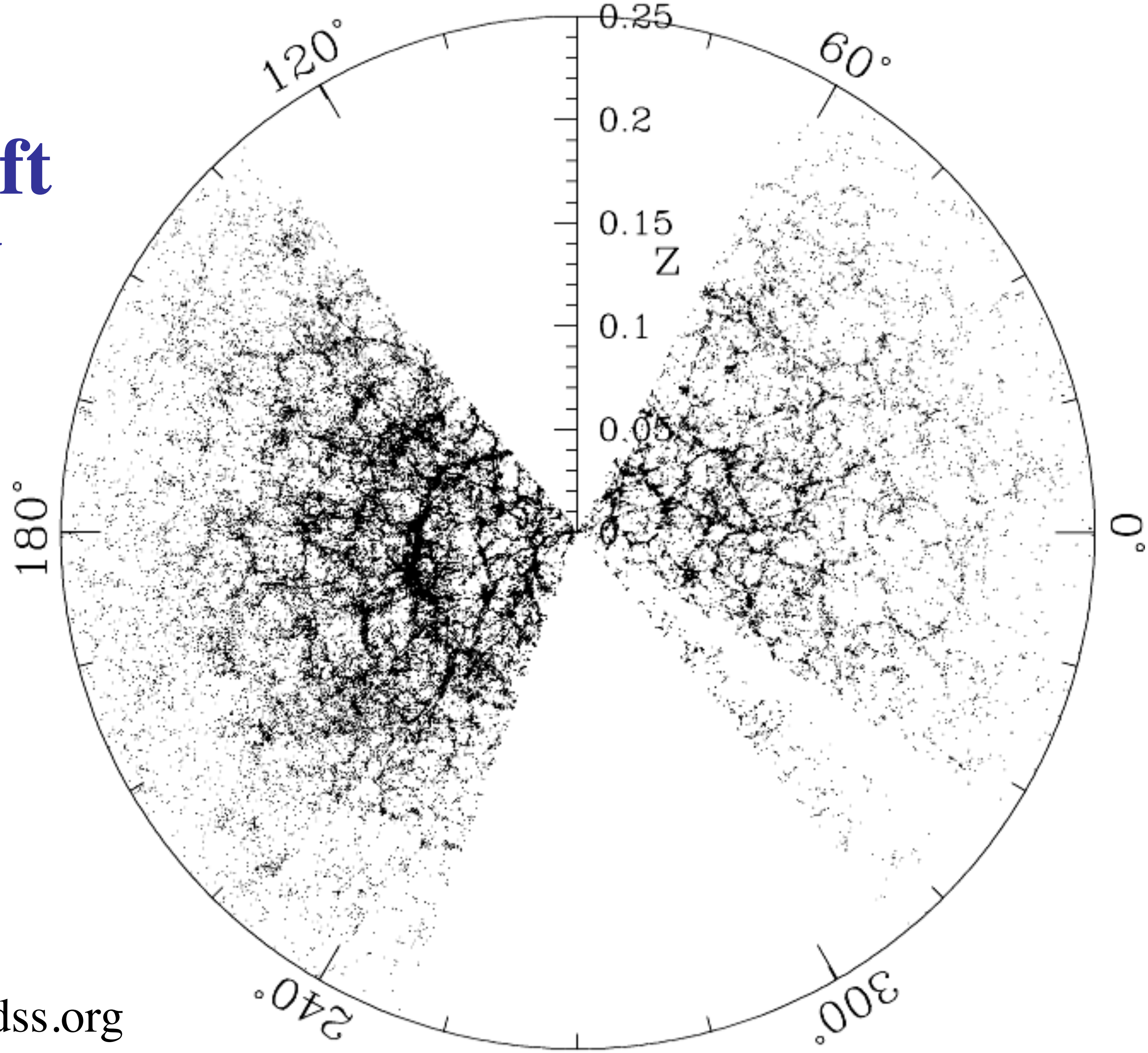
Recently, two very large redshift surveys have been undertaken:

- 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 ~ 0.4, ~ 600 spectra at a time
  - Over a million spectra, including ~ 930,000 galaxies
  - Also spectra of ~ 120,000 quasars out to $z \sim 6.4$

SDSS Sky Coverage

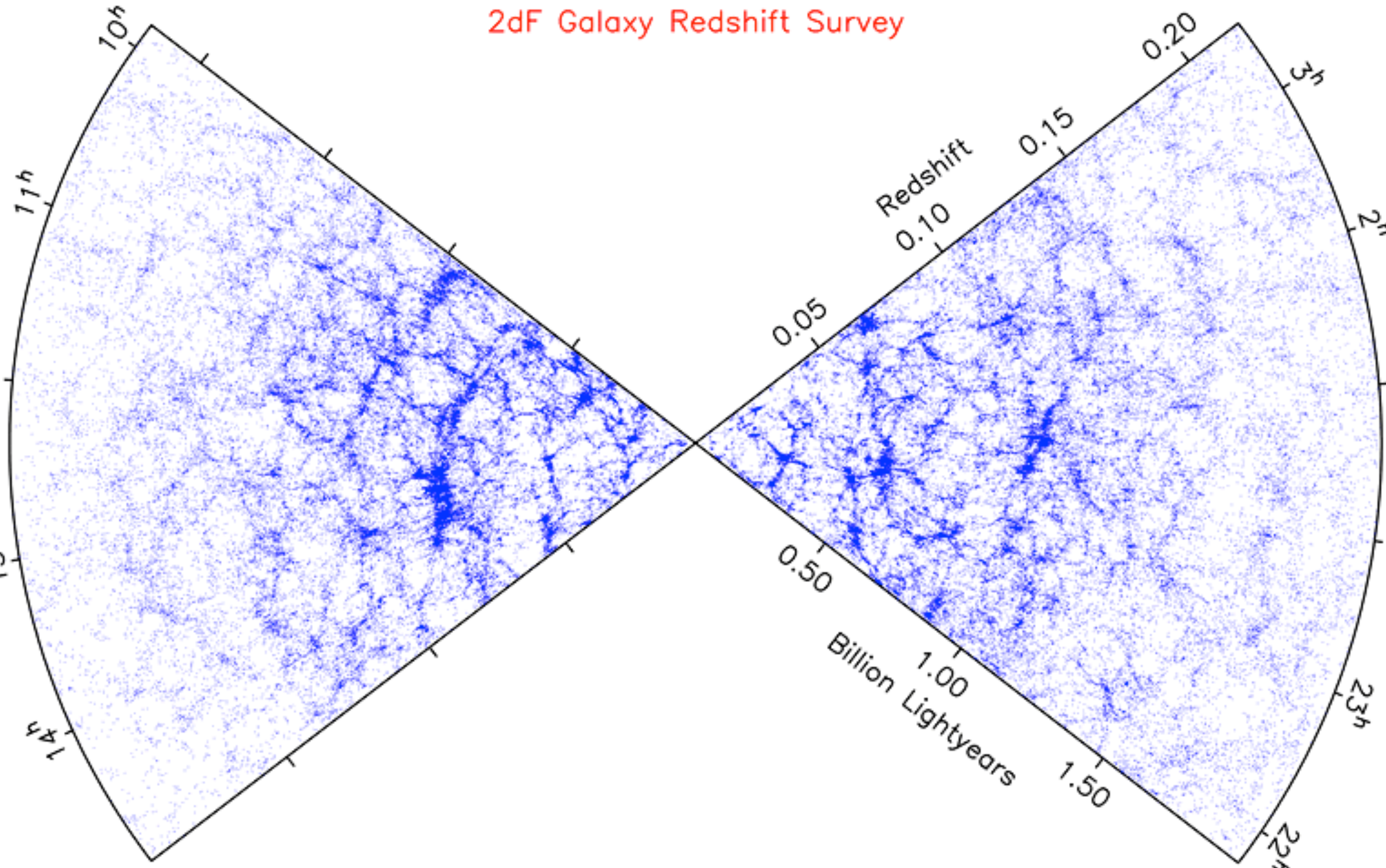


SDSS Redshift Survey



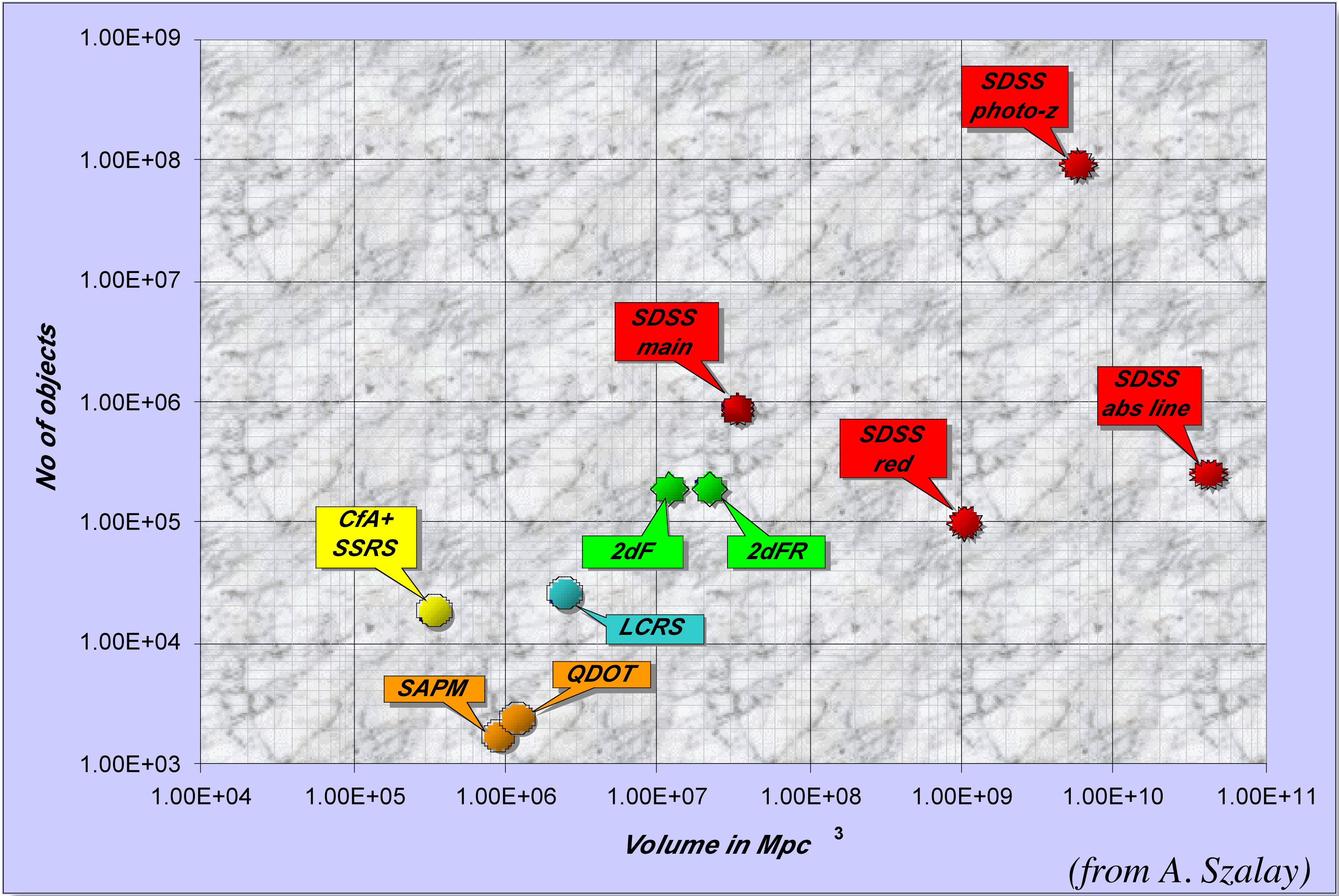
http://www.sdss.org

2dF Galaxy Redshift Survey



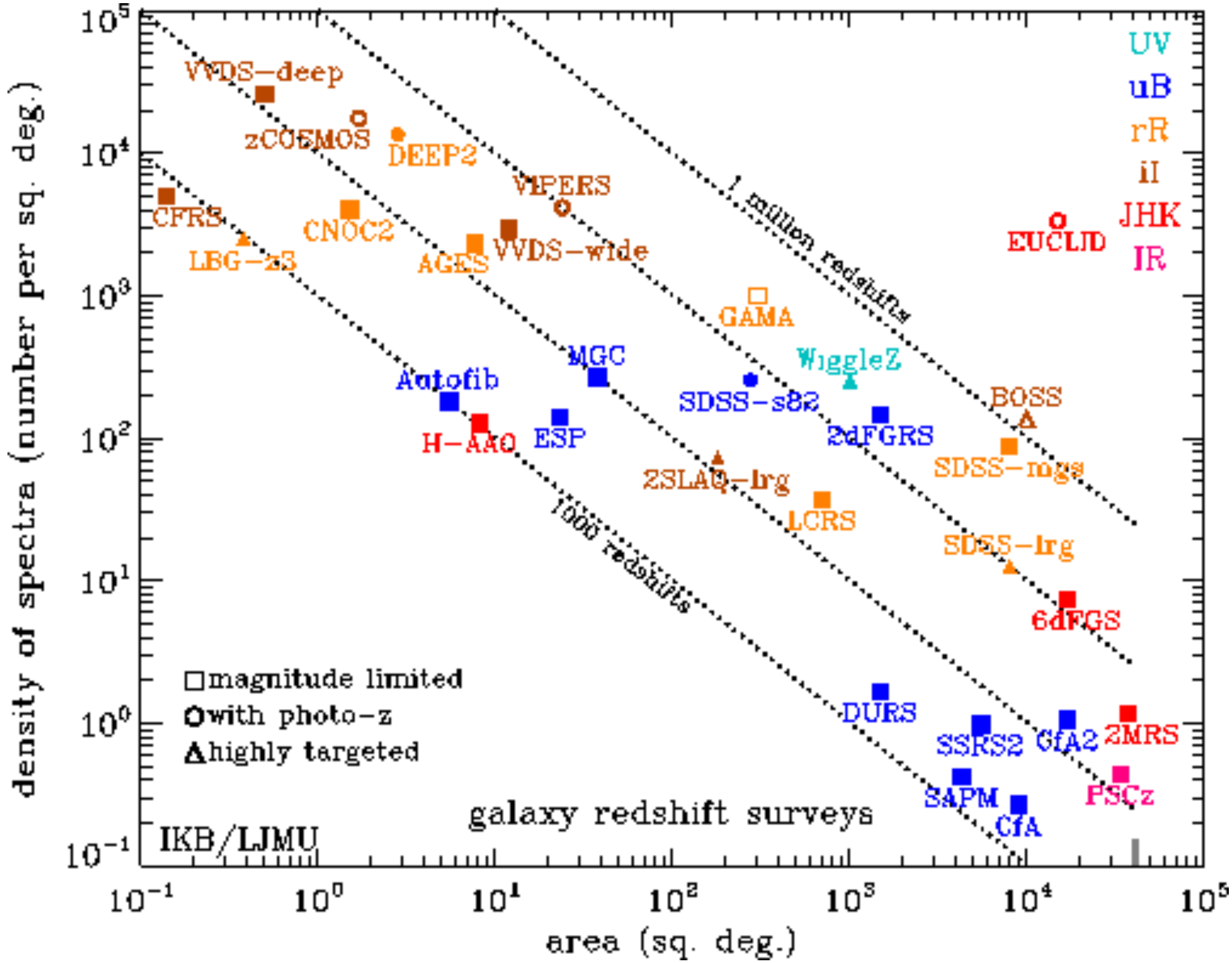
http://www.mso.anu.edu.au/2dFGRS/

Area and Size of Redshift Surveys



(from A. Szalay)

Comparing Redshift Surveys



Galaxy Distribution and Correlations

- If galaxies are clustered, they are "correlated"
- This is usually quantified using the ***2-point correlation function***, $\xi(r)$, 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:

$$dN(r) = \rho_0 \left(1 + \xi(r)\right) dV_1 dV_2$$

- Usually represented as a power-law:

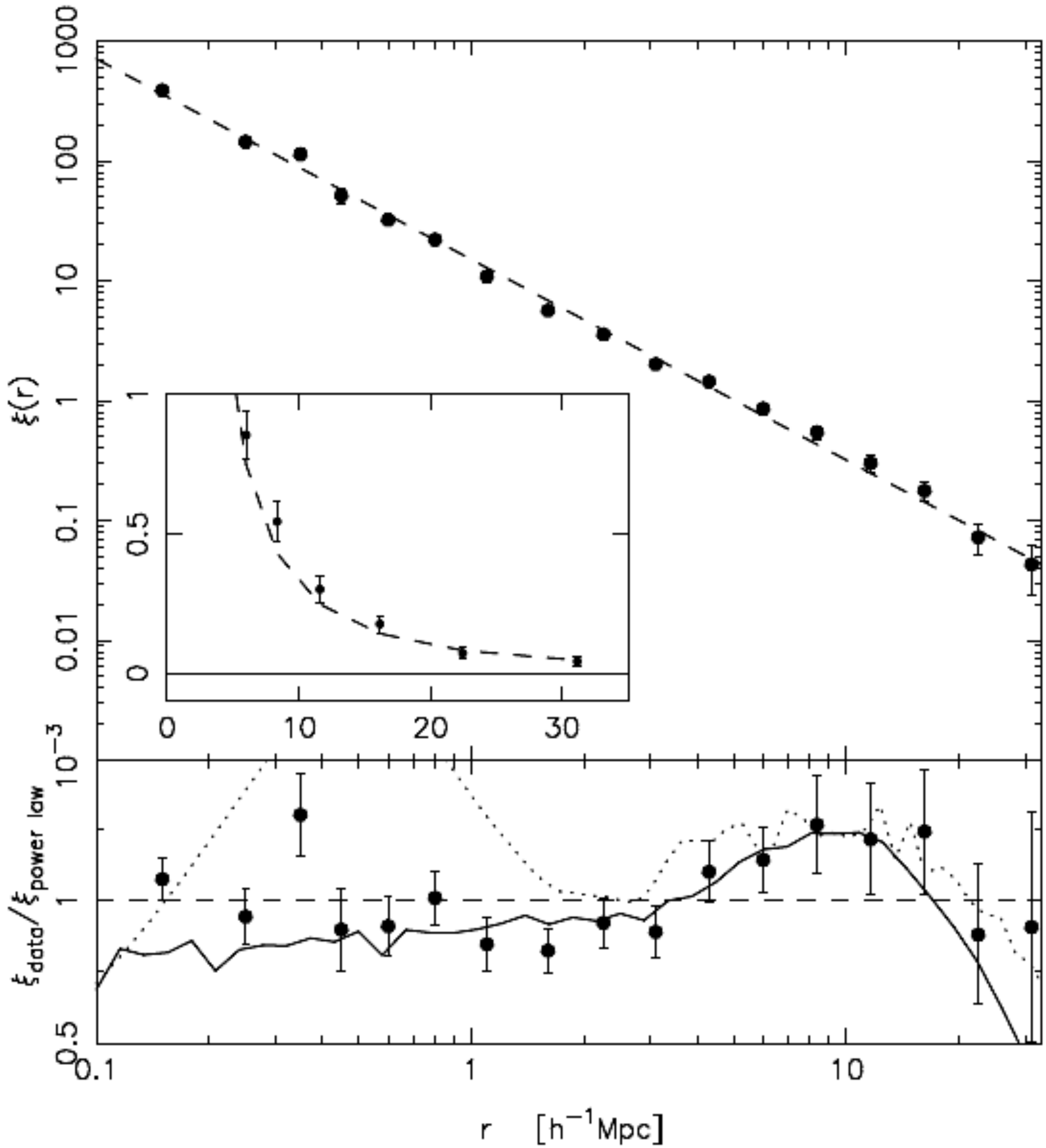
$$\xi(r) = \left(r / r_0\right)^{-\gamma}$$

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

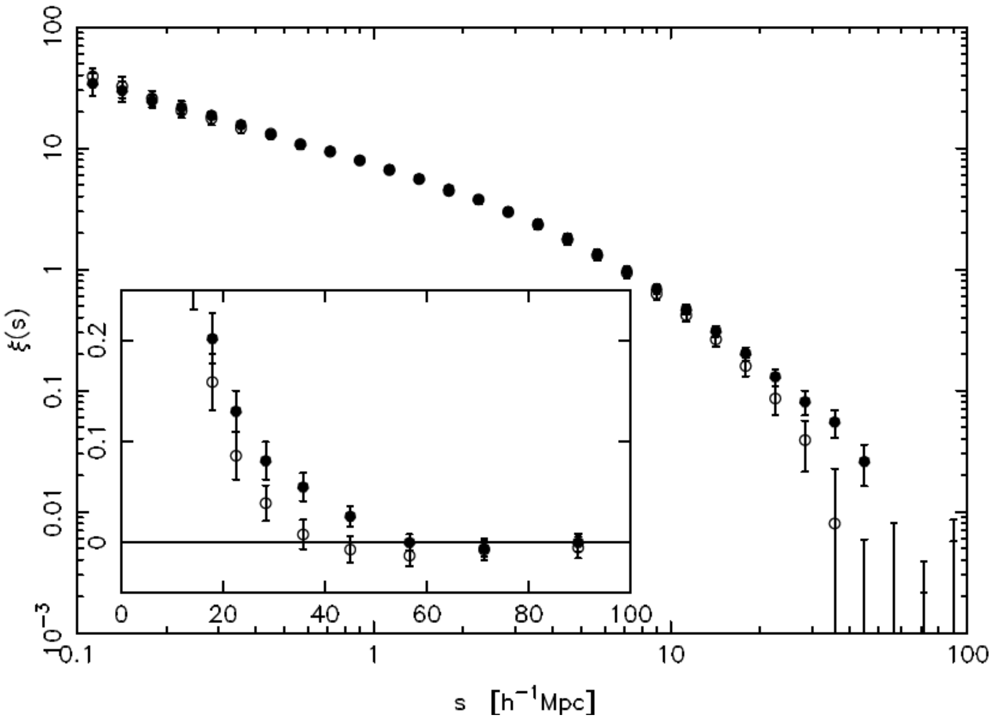
Galaxy Correlation Function

As measured by the 2dF redshift survey

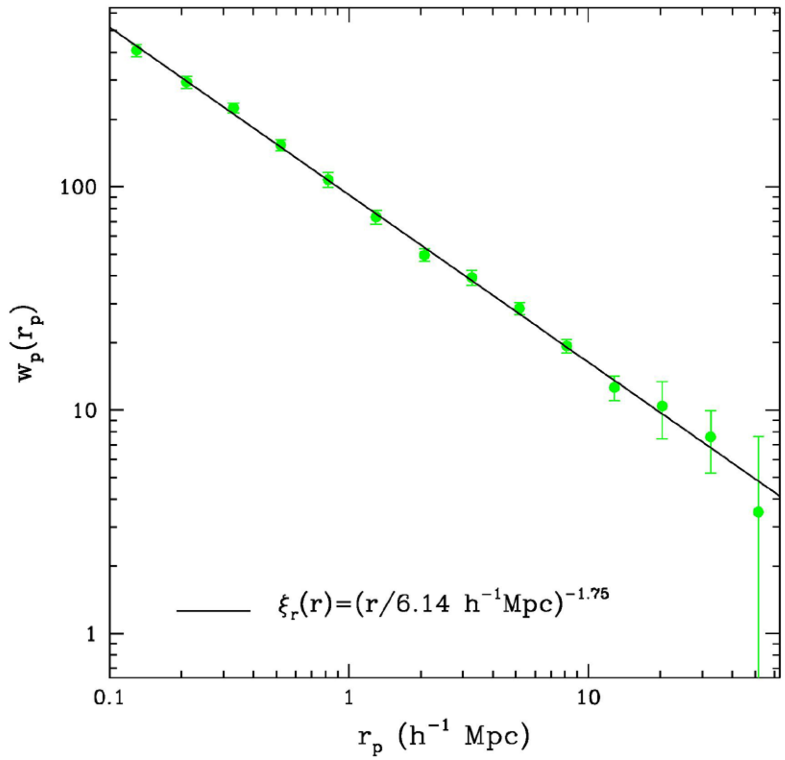
Deviations from the power law:



Galaxy Correlation Function



At sufficiently large scales, e.g., voids, $\boldsymbol{\xi}\,(\boldsymbol{r})$ must turn negative



- If only 2-D positions on the sky are known, then use angular separation $\boldsymbol{\theta}$ instead of distance $\boldsymbol{r}$:

$$w(\theta) = (\theta/\theta_0)^{-\beta}, \ \ \beta = \gamma - 1$$

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, $\langle RR \rangle$:

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

- A better (Landy-Szalay) estimator is:

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

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

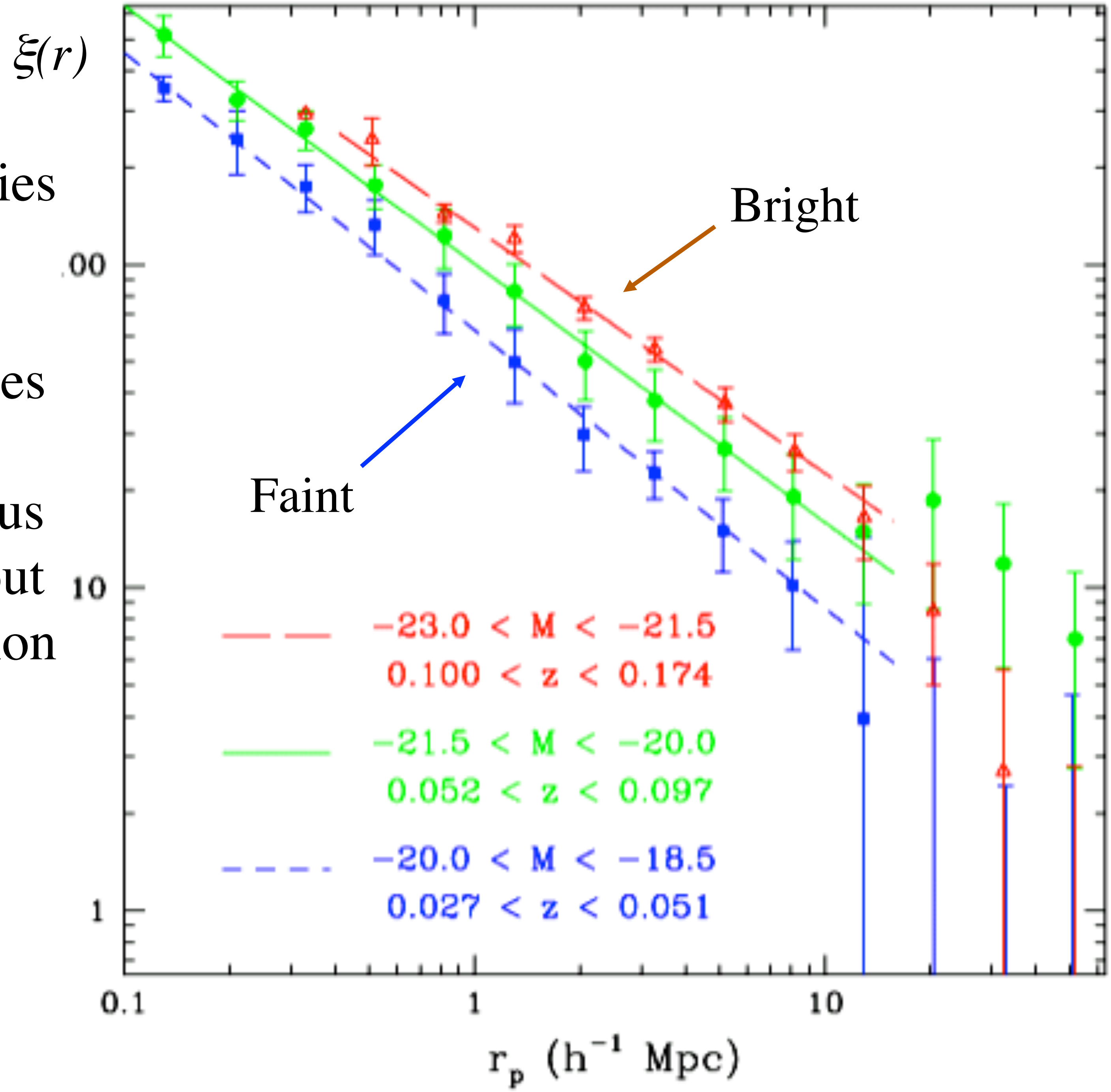
- This takes care of the edge effects, where one has to account for the missing data outside the region sampled, which can have fairly irregular boundaries

Another Definition of $\xi(r)$

- We can also measure it through the overdensity: where $\langle n \rangle$ is the mean density
$$\delta(\mathbf{r}) = \frac{n - \langle n \rangle}{\langle n \rangle}$$
- In case of discrete galaxy catalogs, define counts in cells, N_i
$$\delta_i(\mathbf{r}) = \frac{N_i - \langle N_i \rangle}{\langle N_i \rangle}$$
- Then $\boldsymbol{\xi}\,(\boldsymbol{r})$ is the expectation value:
$$\xi(\mathbf{x}_1, \mathbf{x}_2) = \left\langle \delta(\mathbf{x}_1)\delta(\mathbf{x}_2) \right\rangle$$
- Note that we have considered a correlation of a single density field with itself, so strictly speaking $\boldsymbol{\xi}\,(\boldsymbol{r})$ 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, $\zeta = \left\langle \delta_1\, \delta_2 \delta_3 \right\rangle$,
$\eta = \left\langle \delta_1\, \delta_2 \delta_3\, \delta_4 \right\rangle$ … etc.

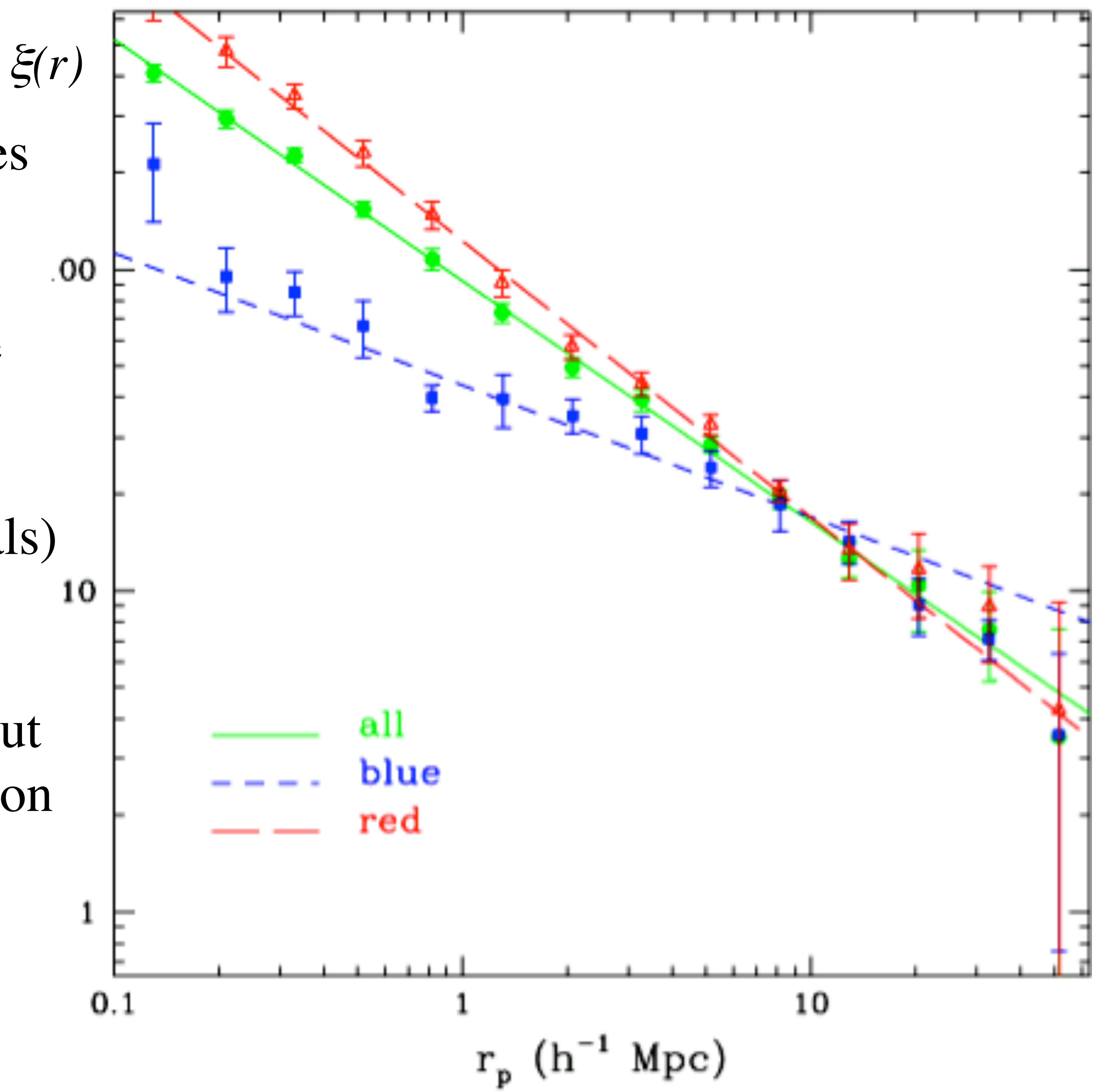
Brighter galaxies are clustered more strongly than fainter ones

This is telling us something about galaxy formation



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

That, too, says something about galaxy formation



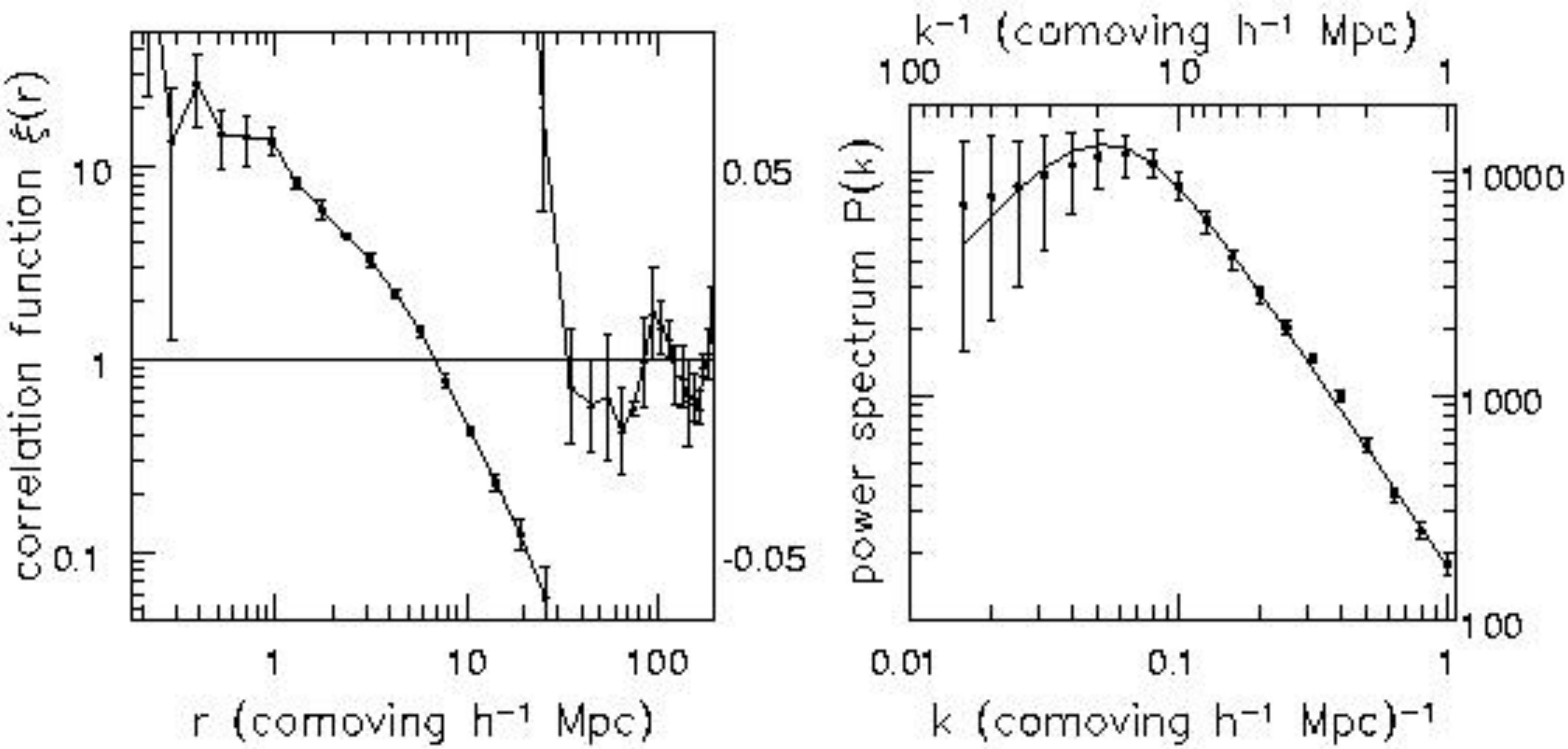
Correlation Function and Power Spectrum

- 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 $P(\mathbf{k}) = |\delta(\mathbf{k})|^2$
- Then $$\xi(r) = \frac{1}{4\pi^2} \int d\ln k\; j_0(kr) \left[k^3 P(k)\right]$$

Correlation function and power spectrum are a Fourier pair

An Example from Las Campanas Redshift Survey



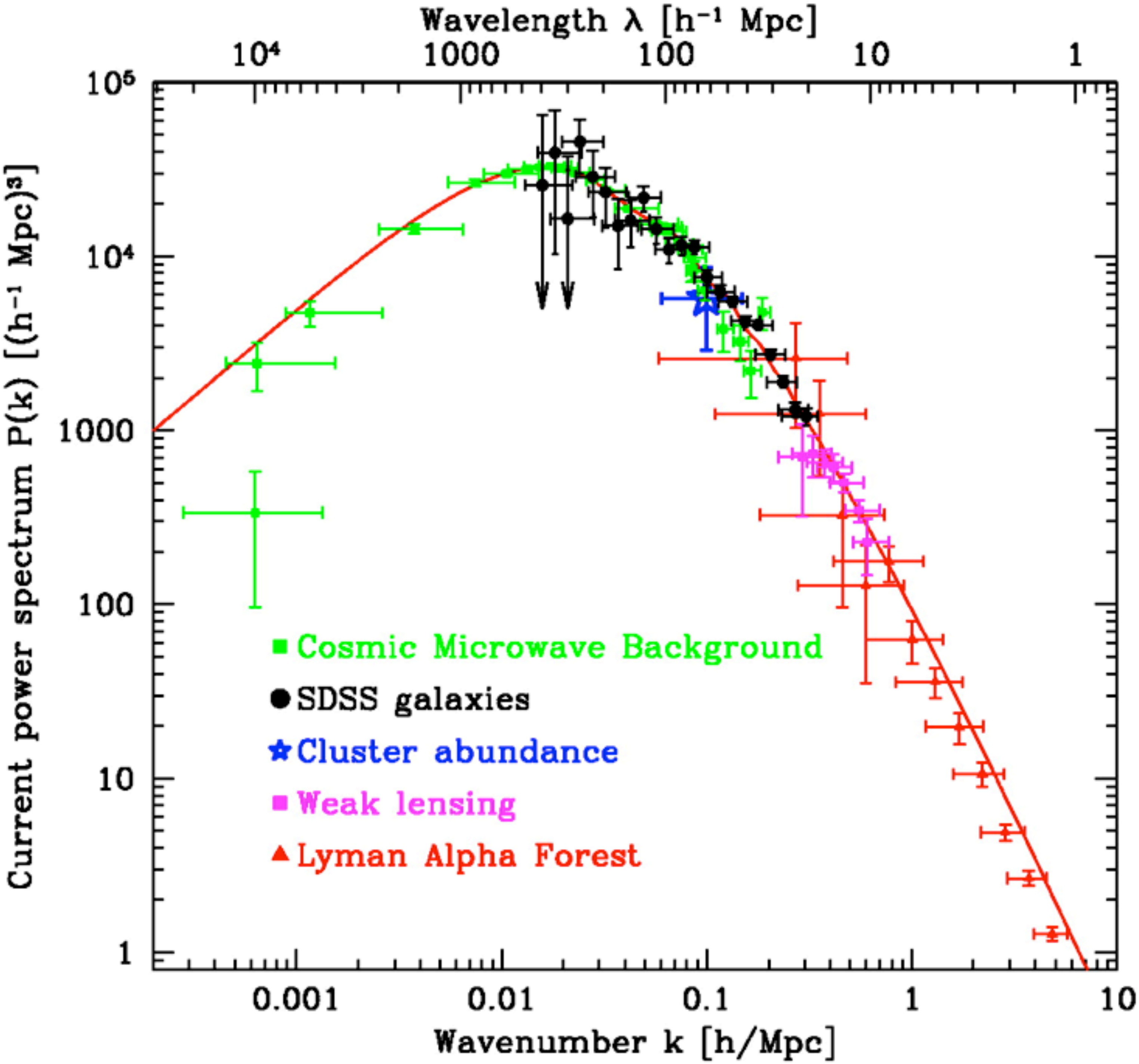
Correlation function is easier to evaluate, but power spectra is what we need to compare with the theory

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 $\sigma \approx 1$
- 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, $$\sigma_R{}^2 = \frac{1}{4\pi^2}\int d\ln k\,\left[k^3 P(k)\left|K_R(k)\right|^2\right]$$ where K_R is a convolving kernel, a spherical top-hat with a radius R:

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

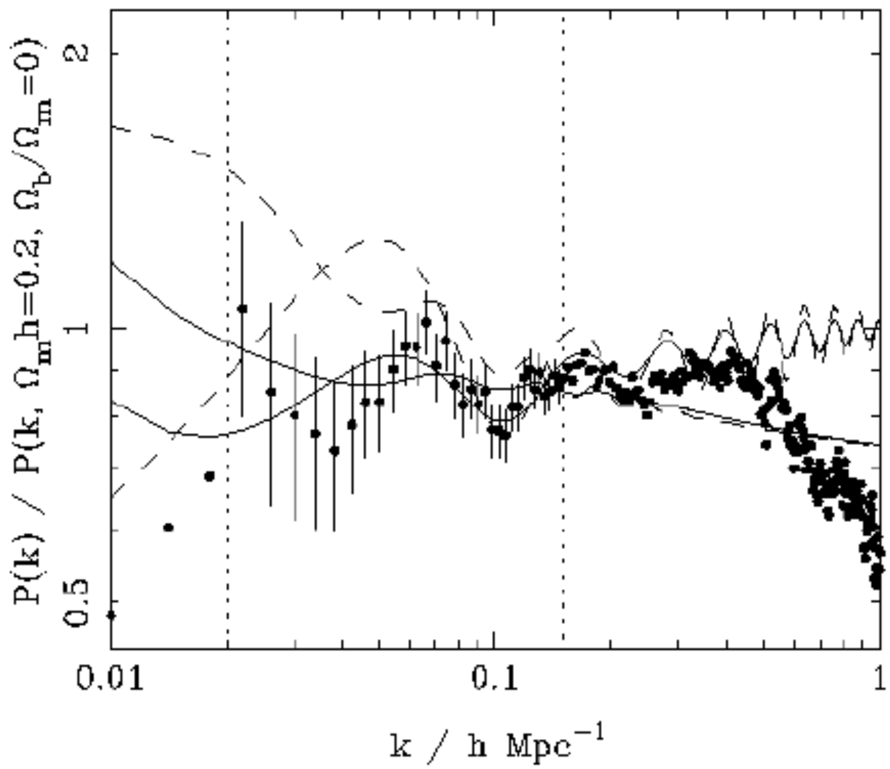
The Observed Power Spectrum



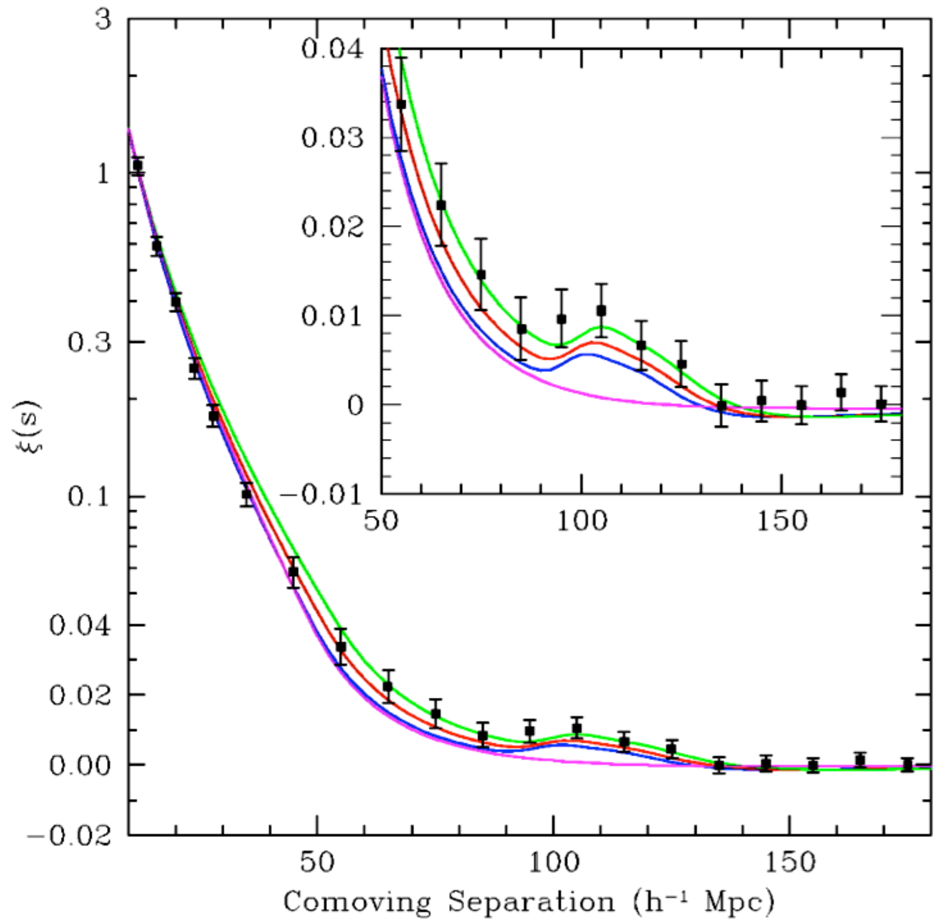
(Tegmark et al.)

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



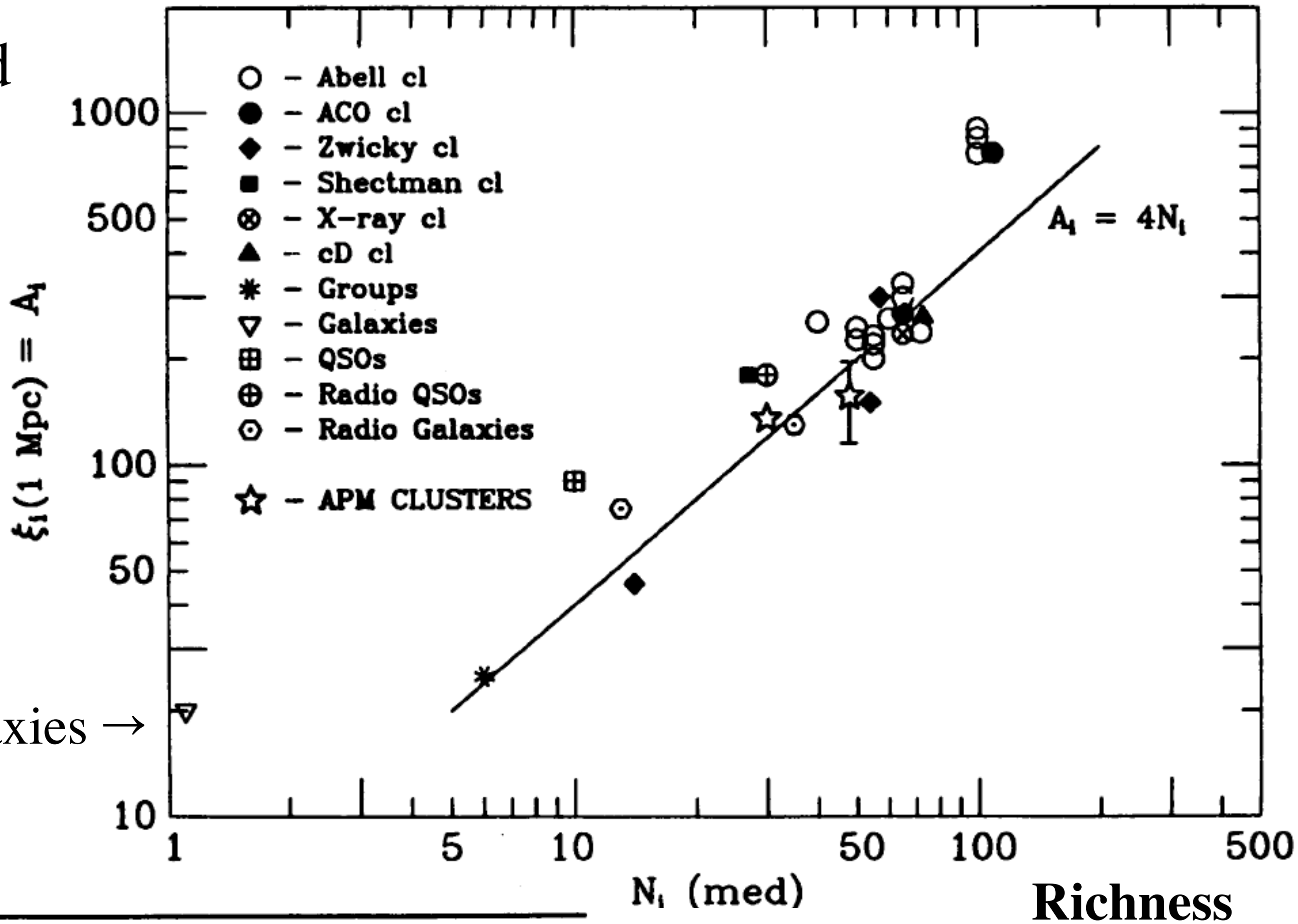
2dF (Percival et al.)



SDSS (Eisenstein et al.)

Cluster-Cluster Clustering

Clusters are clustered more strongly than individual galaxies, and rich ones more than the poor ones



Field galaxies →

Richness

Table 1. Cluster correlations

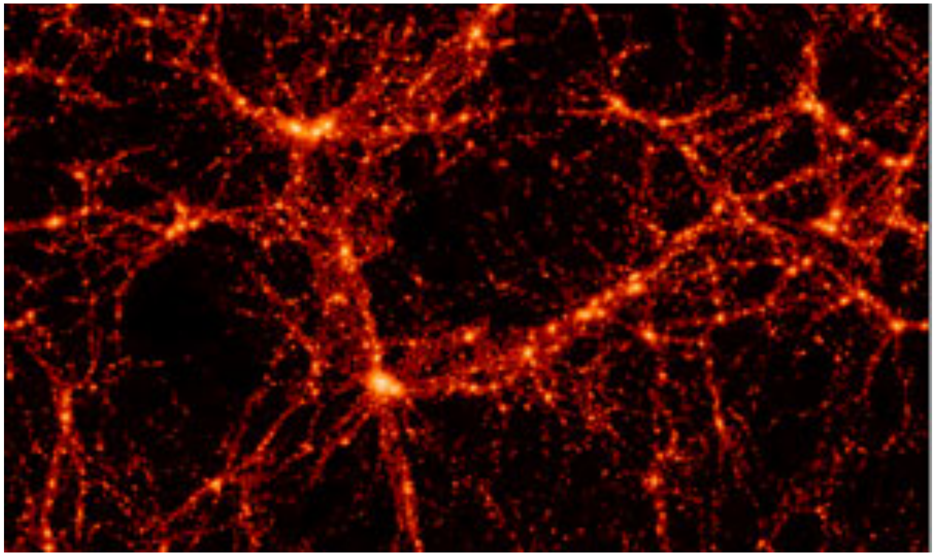
Catalog	n_c, h^3 Mpc^{-3}	d, h^{-1} Mpc	r_o(obs)	$r_o = 0.4d$
Abell, $R \geq 2$	1.2×10^{-6}	94	42 ± 10	37.6
Abell, $R \geq 1$	6×10^{-6}	55	22 ± 3	22.0
EDCC	15×10^{-6}	40.5	16 ± 4	16.2
APM	24×10^{-6}	34.7	13 ± 2	13.9

(from N. Bahcall)

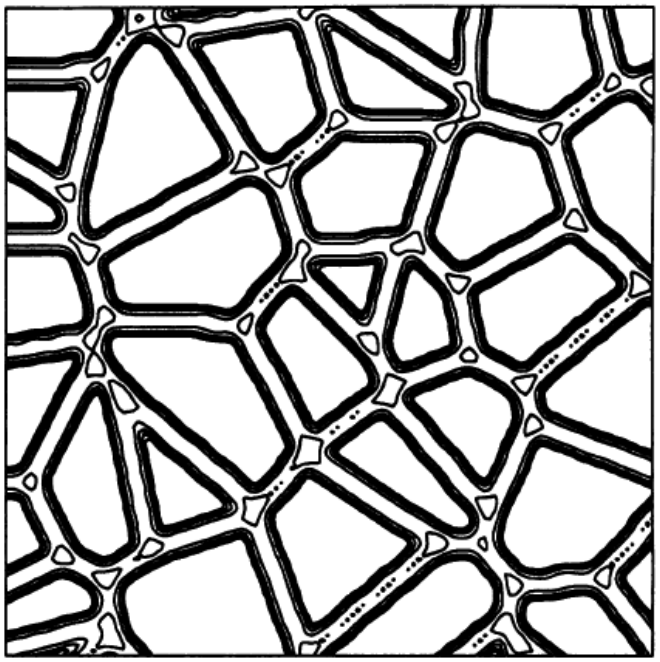
Is the Power Spectrum Enough?

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

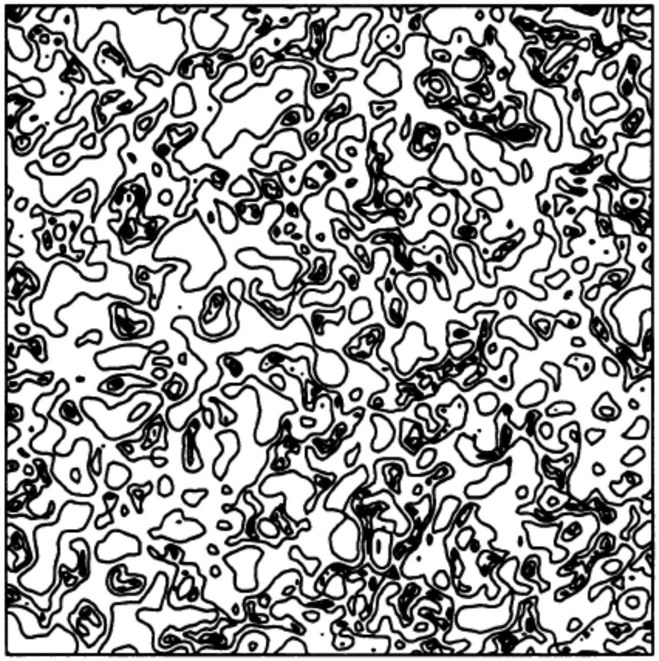
The power spectrum alone does not capture the phase information: the coherence of cosmic structures (voids, walls, filaments …)



Voronoi foam, R=1.6, smoothed original



Voronoi foam, R=1.6, random phases

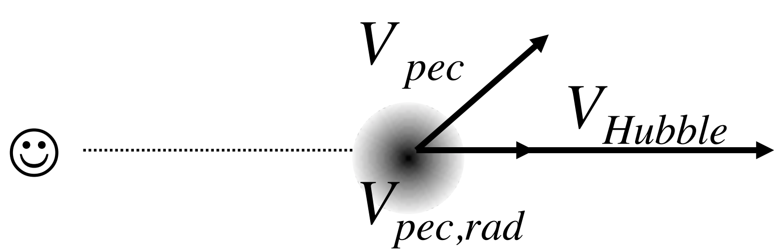


LSS Observations Summary

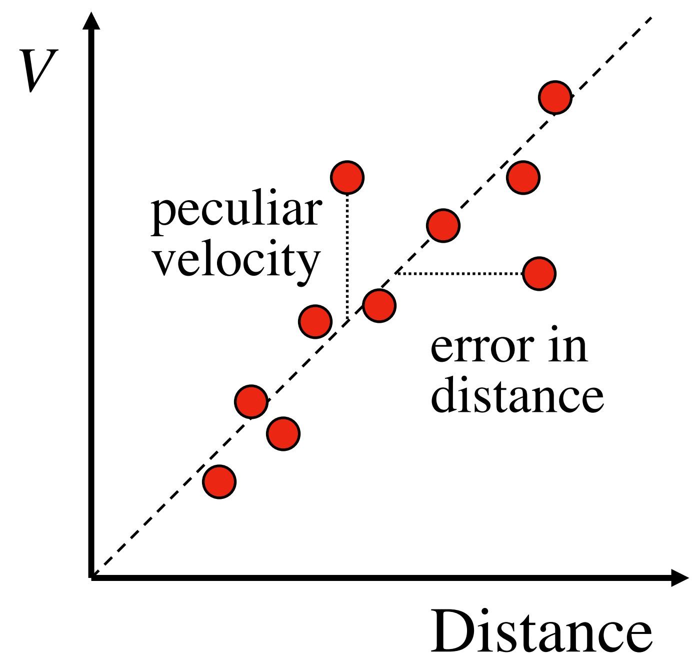
- A range of structures: galaxies (~10 kpc), groups (~ 0.3 - 1 Mpc), clusters (~ few Mpc), superclusters (~ 10 - 100 Mpc)
- Redshift surveys are used to map LSS; ~ 10^6 galaxies now
- LSS topology is prominent: voids, sheets, filaments…
- LSS quantified through 2-point (and higher) correlation function(s), well fit by a power-law:
$$\xi(r) = \left(r / r_0\right)^{-\gamma}$$
typical $\gamma \sim 1.8$, $r_0 \sim$ few Mpc
- Equivalent description: power spectrum *P(k)* - useful for comparisons with the theory
- CDM model fits the data over a very broad range of scales
- Objects of different types have different clustering strengths
- Generally more massive structure cluster more strongly

Peculiar Velocities

- It means velocities of galaxies in addition to their Hubble flow velocities, i.e., relative to their comoving coordinates restframe
- They act as a noise on the $V = cz$ axis in the Hubble diagram, and could bias the measurements of the H_0 (which is why we want the "far field" measurements)
- They result from the large-scale density field, and thus can be used to probe it
- A good review: Strauss & Willick 1995, *Phys. Rep.* **261**, 271

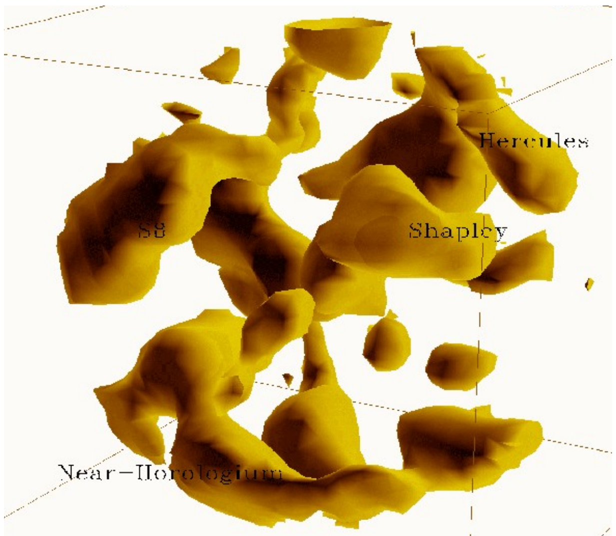


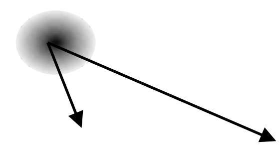
$$V_{total} = V_{Hubble} + V_{pec,rad}$$



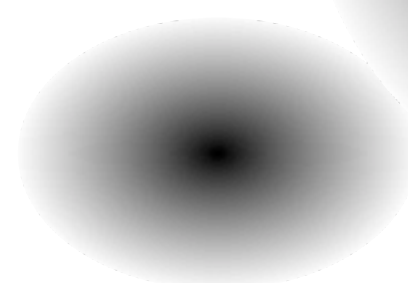
Large-Scale Density Field Inevitably Generates a Peculiar Velocity Field

The PSCz survey
local 3-D density field





A galaxy is accelerated towards the nearby large mass concentrations

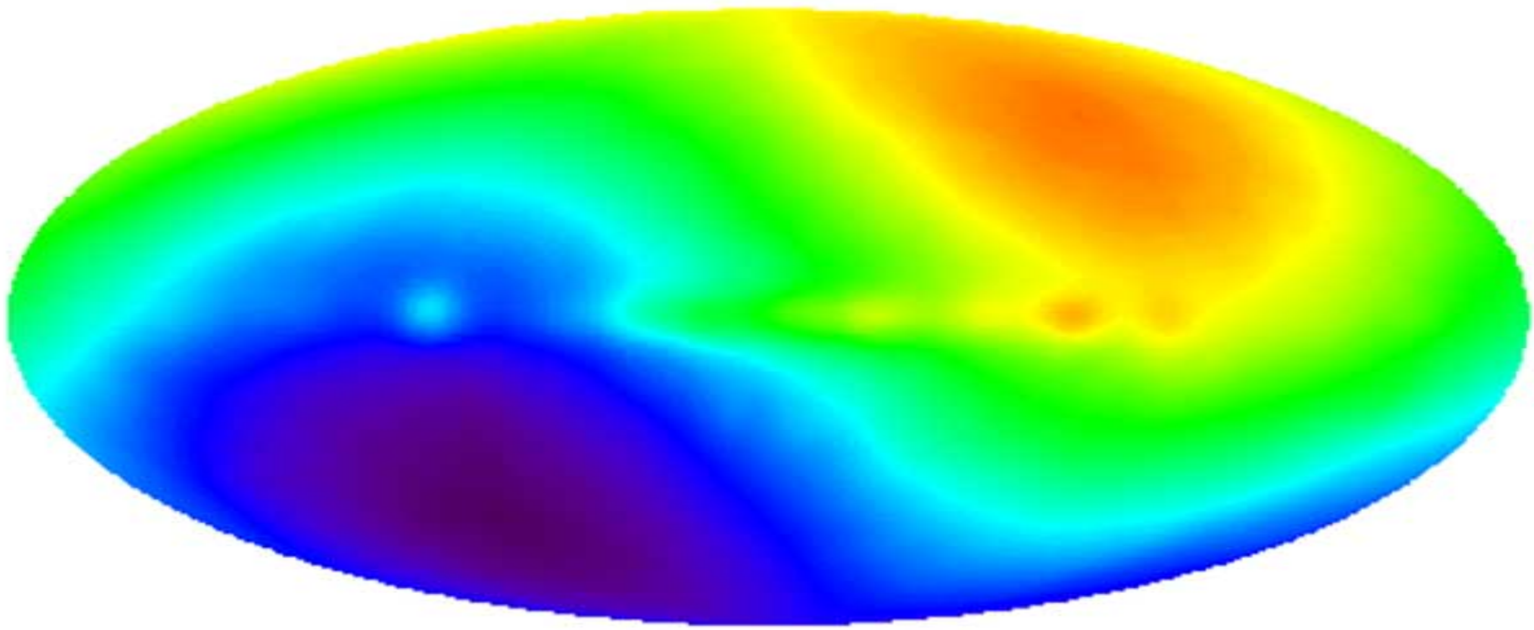


Integrated over the Hubble time, this results in a peculiar velocity

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

CMBR Dipole:

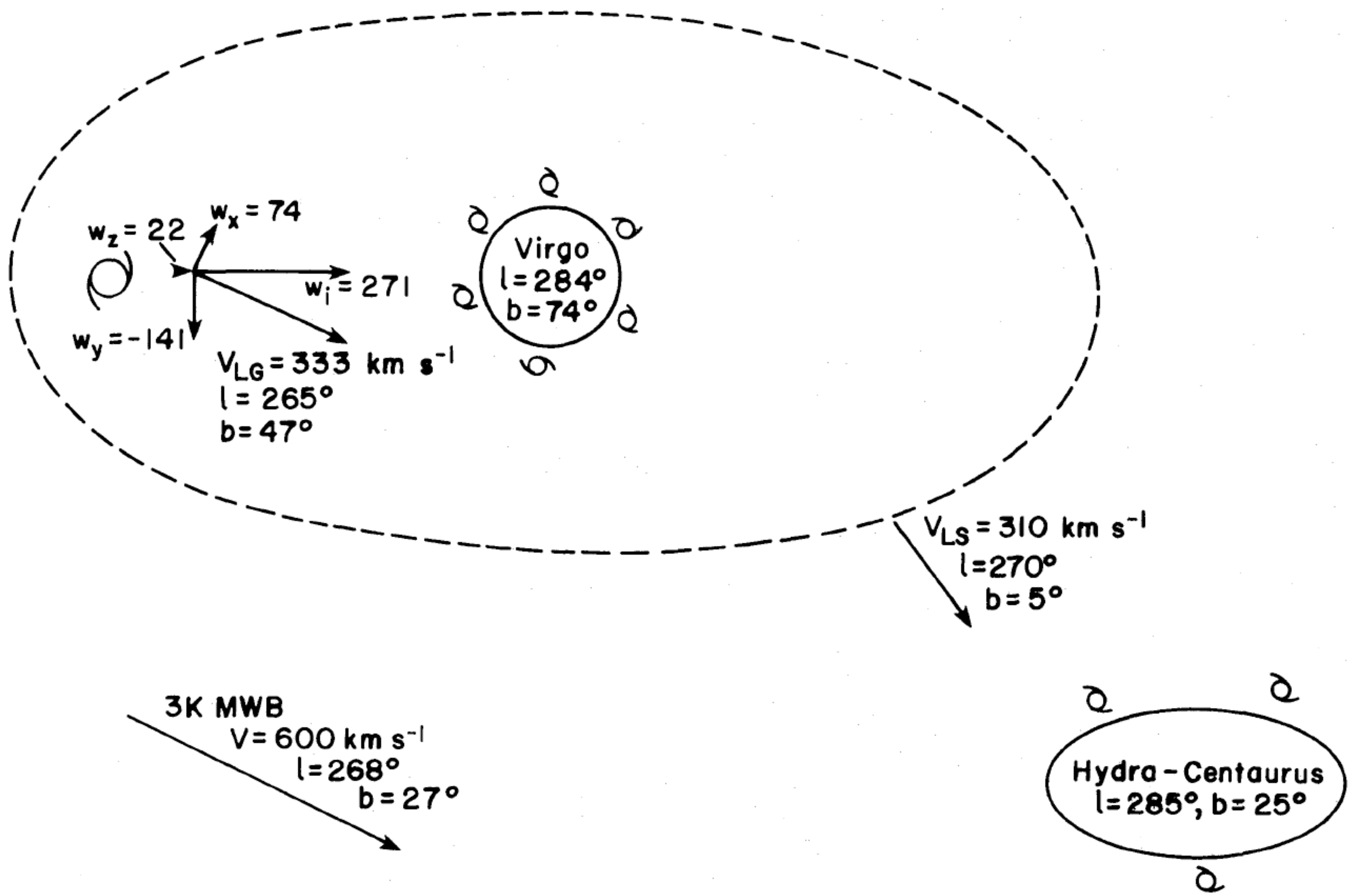
The One Peculiar Velocity We Know Very Well



We are moving wrt. to the CMB at ~ 620 km/s towards $b=27°$, $l=268°$

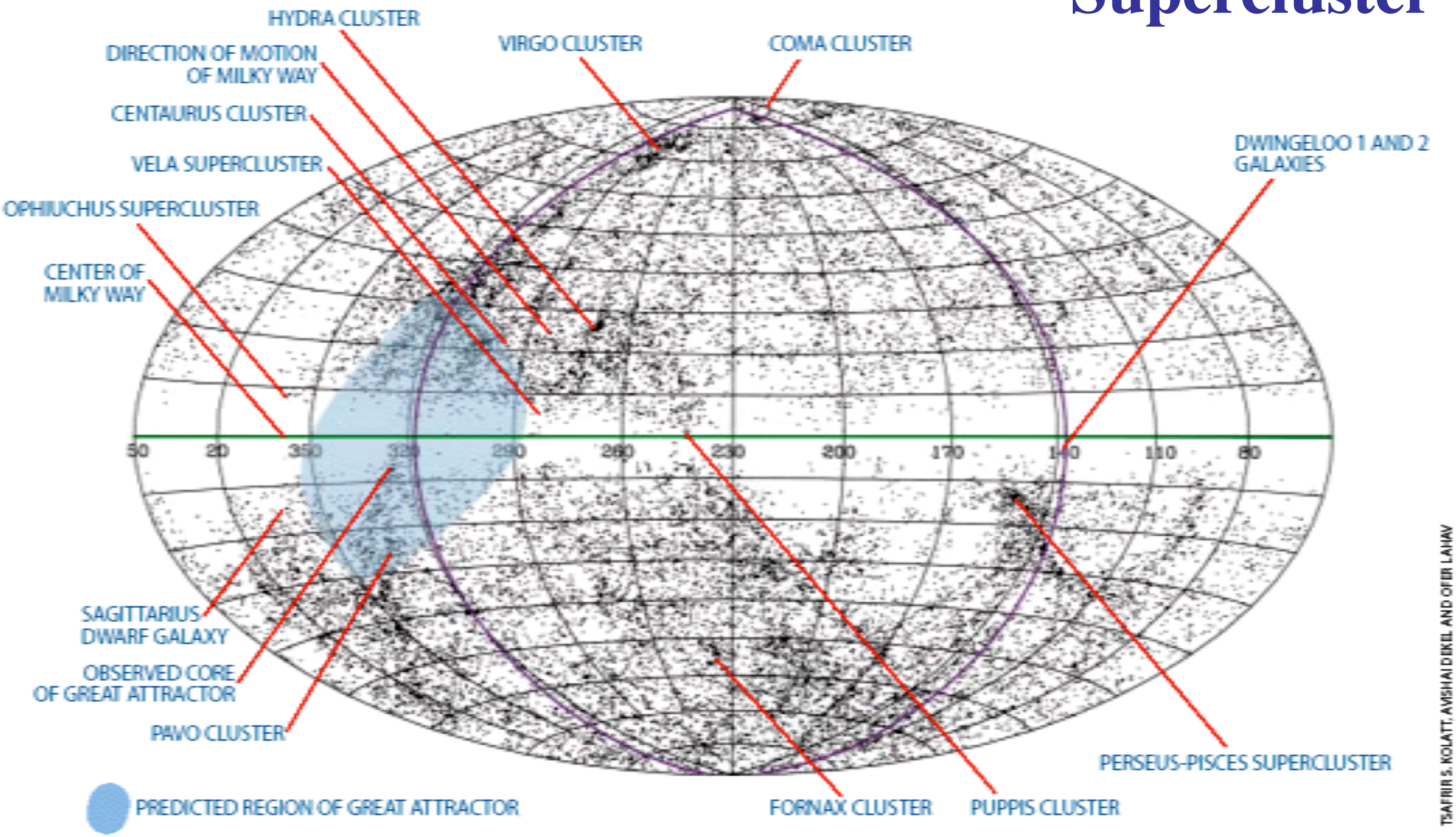
This gives us an idea of the probable magnitude of peculiar velocities in the local universe. Note that at the distance to Virgo (LSC), this corresponds to a ~ 50% error in Hubble velocity, and a ~ 10% error at the distance to Coma cluster.

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



… and to the Shapley Concentration beyond? The "dark Flow"?

The “Great Attractor” aka the Hydra-Centaurus Supercluster



30,000 GALAXIES, culled from three standard astronomical catalogues, are shown as dots on this map. The galaxies appear all over the sky except in the so-called zone of avoidance, which corresponds to the plane of our Milky Way galaxy (*green horizontal center line*). Outside the zone, the galaxies tend to clump near a line that traces out the Supergalactic Plane (*purple line*).

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

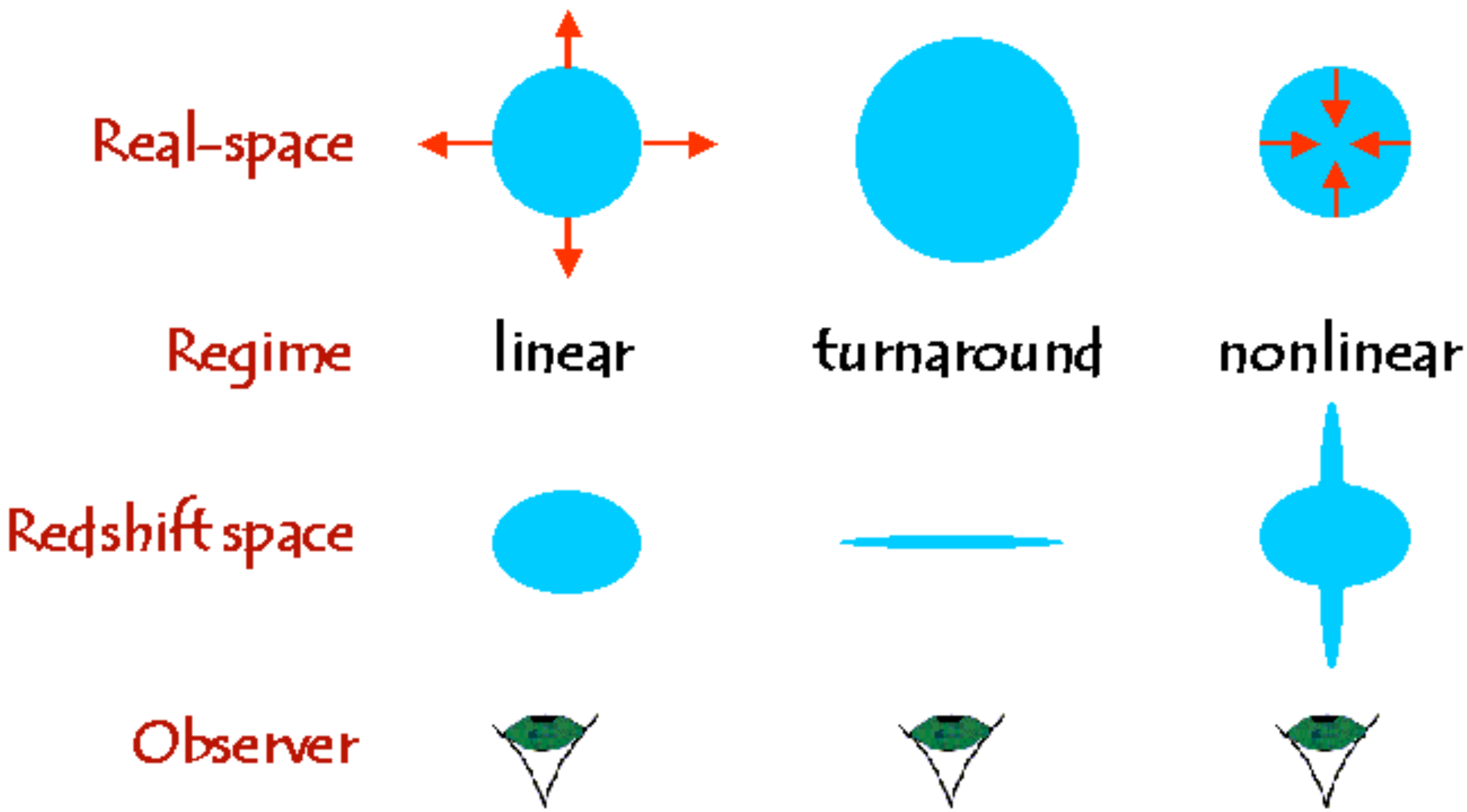
2. Statistically from a redshift survey

- Model-dependent

Redshift-Space Distortions

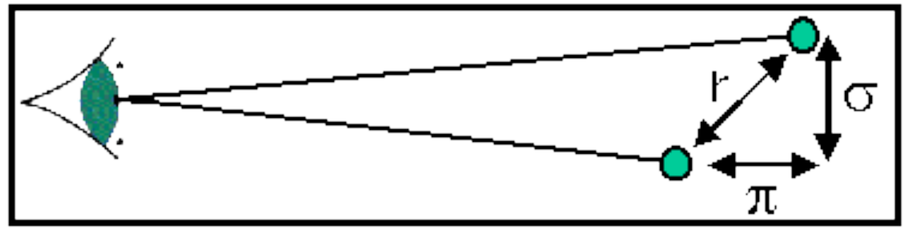
$z_{obs} = z_{true} + v_{pec}/c$ where $v_{pec} \propto \Omega^{0.6}\,\delta\rho/\rho = (\Omega^{0.6}/b)\,\delta n/n$

bias

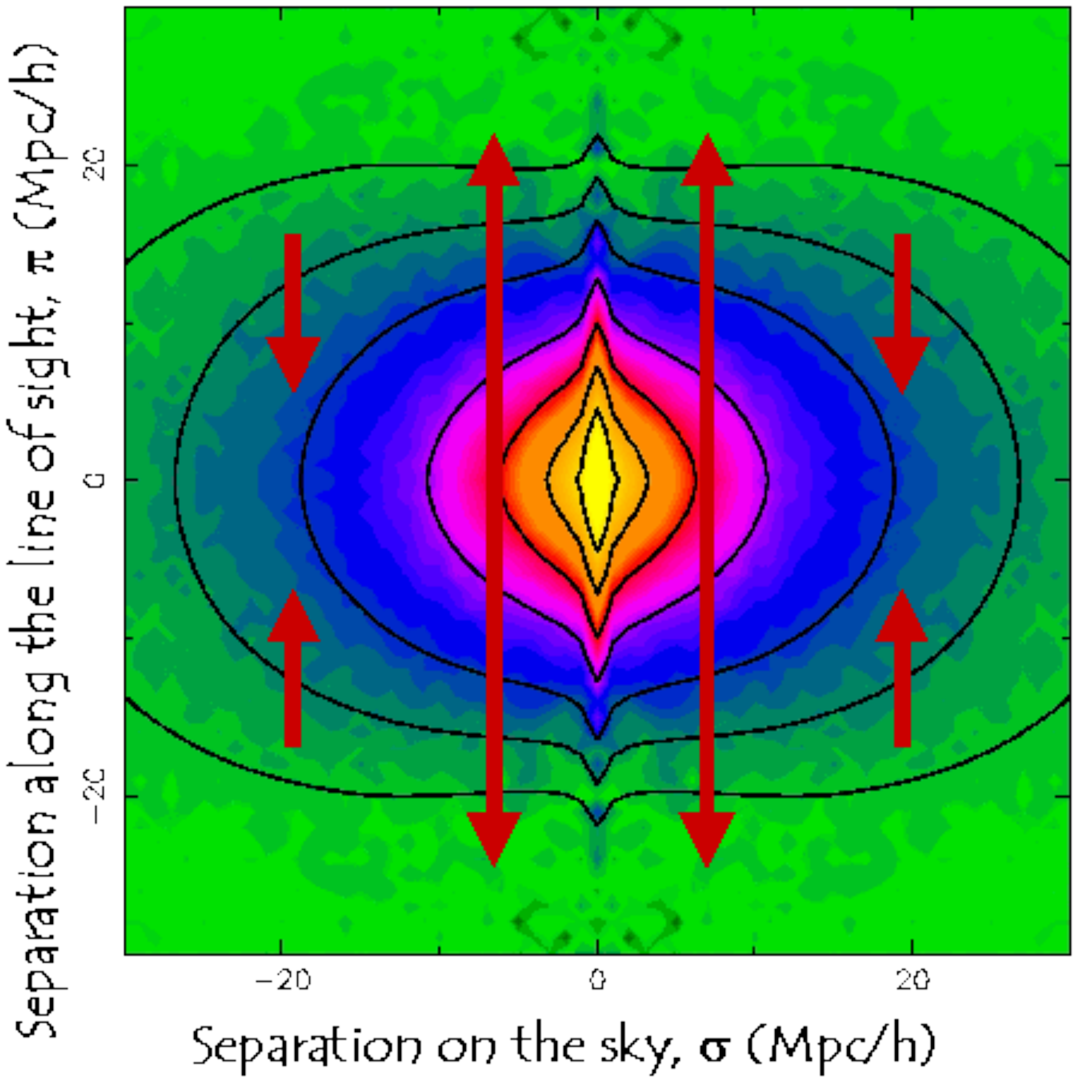


(from 2dF team)

Redshift-Space Correlation Function



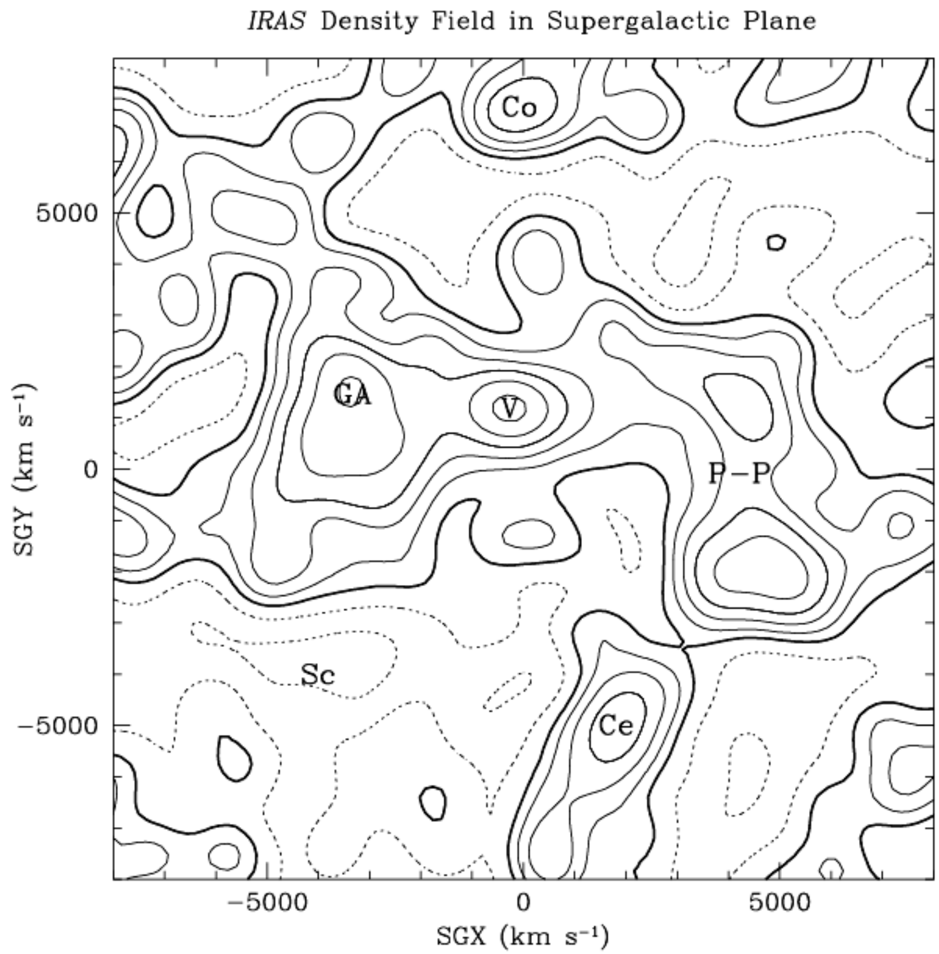
Hawkins et al. (2002), astro-ph/0212375
2dFGRS: $\beta = 0.47 \pm 0.09$



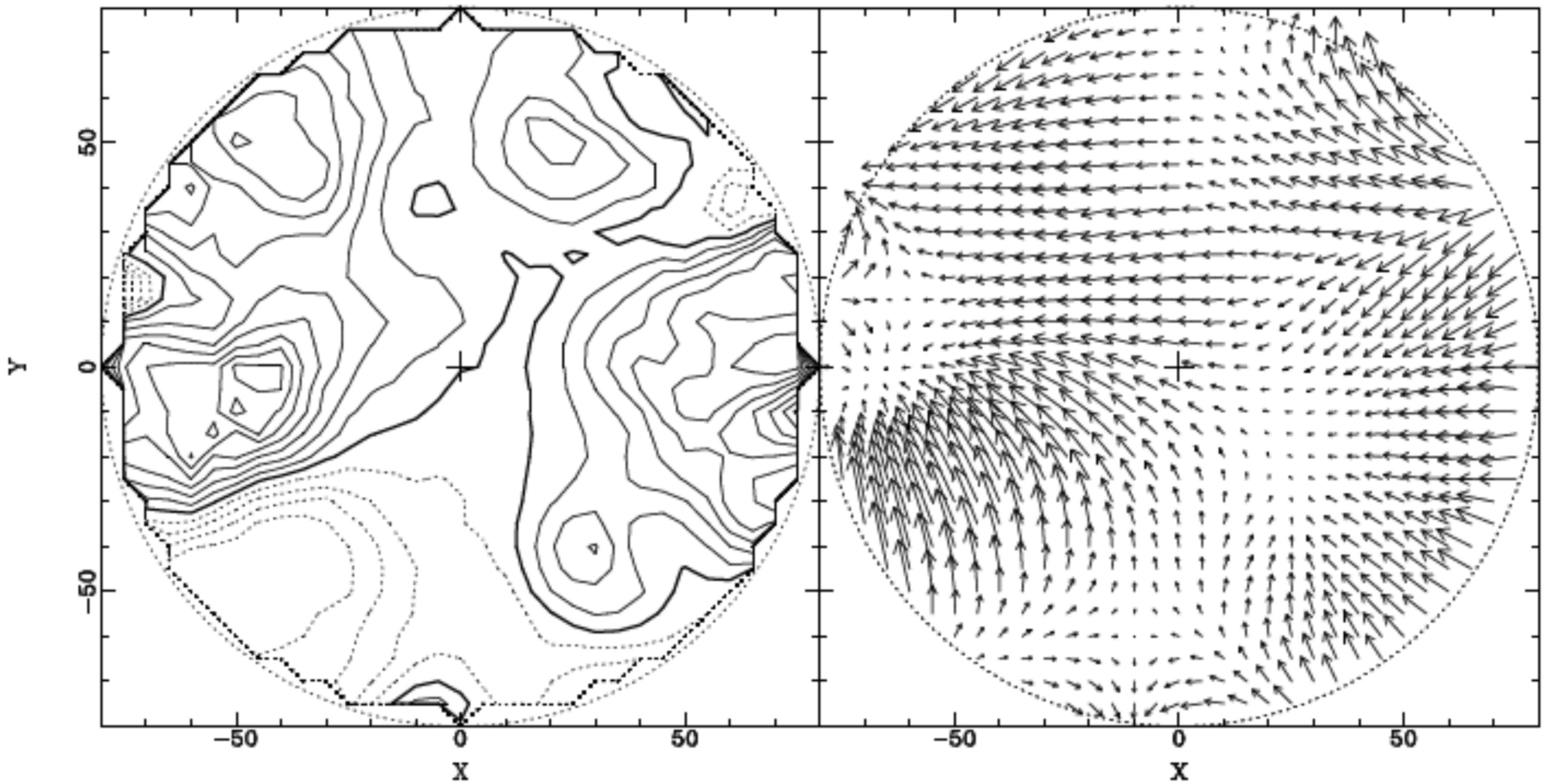
- Small $\sigma \Rightarrow$ non-linear 'Finger-of-God' effect
- Large $\sigma \Rightarrow$ flattening due to coherent infall
- Fit to r = 8–30 h^{-1} Mpc; (nb: $\langle z \rangle = 0.15$, $\langle L \rangle = 1.4L^*$)
- Distortion parameter $\beta = \Omega^{0.6}/b = 0.47 \pm 0.09$
- Pair-wise vel. dispersion $\sigma_p = 495 \pm 52$ km/s
- If $b \approx 1 \Rightarrow \Omega \approx 0.21$
- If $\Omega \approx 0.3 \Rightarrow b \approx 1.2$

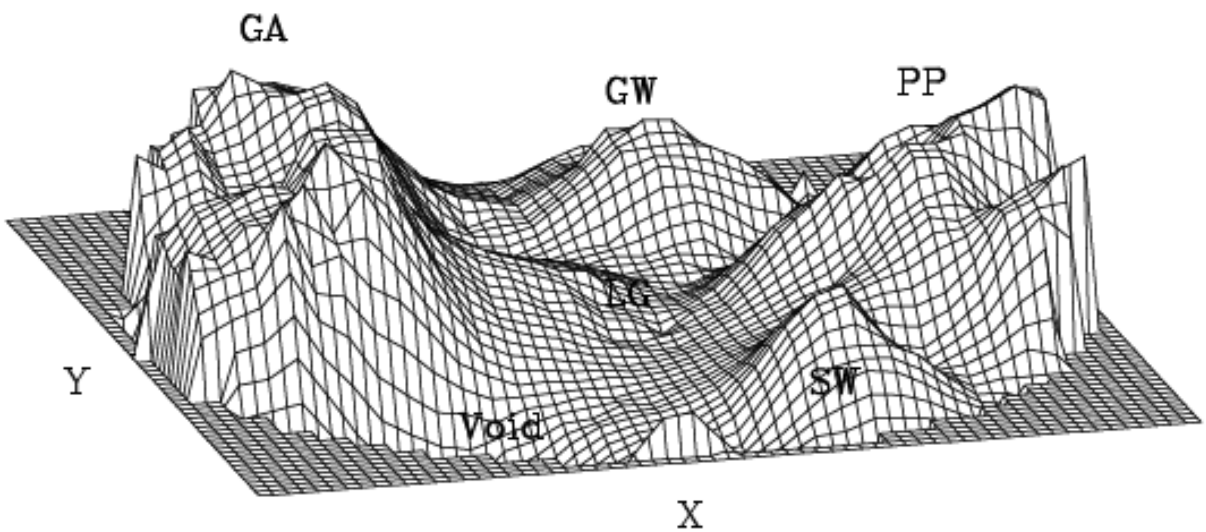
Measuring Peculiar Velocity Field Using a Redshift Survey

- 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 esimated peculiar velocities
- Update the positions according to new Hubble velocities
- Iterate until the convergence
- You get a consistent density and velocity field

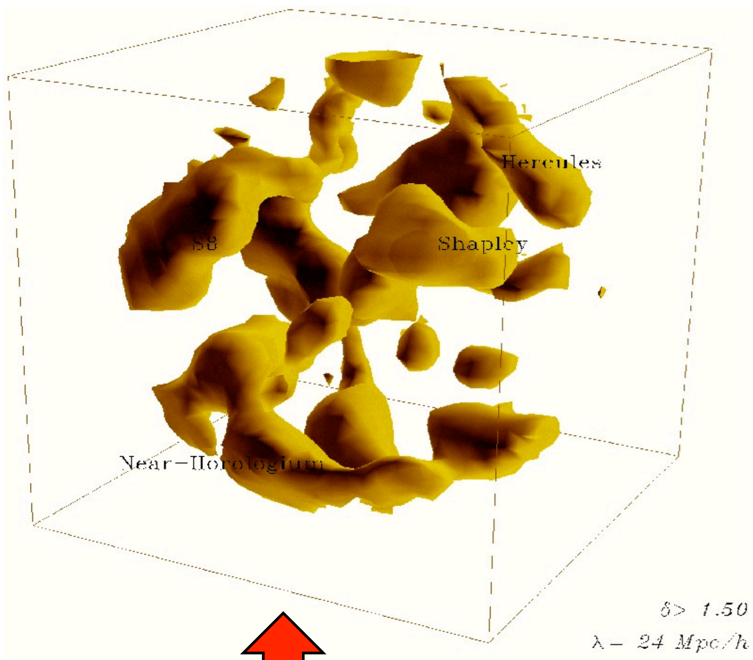


Local Density and Velocity Fields From Peculiar Velocities of Galaxies



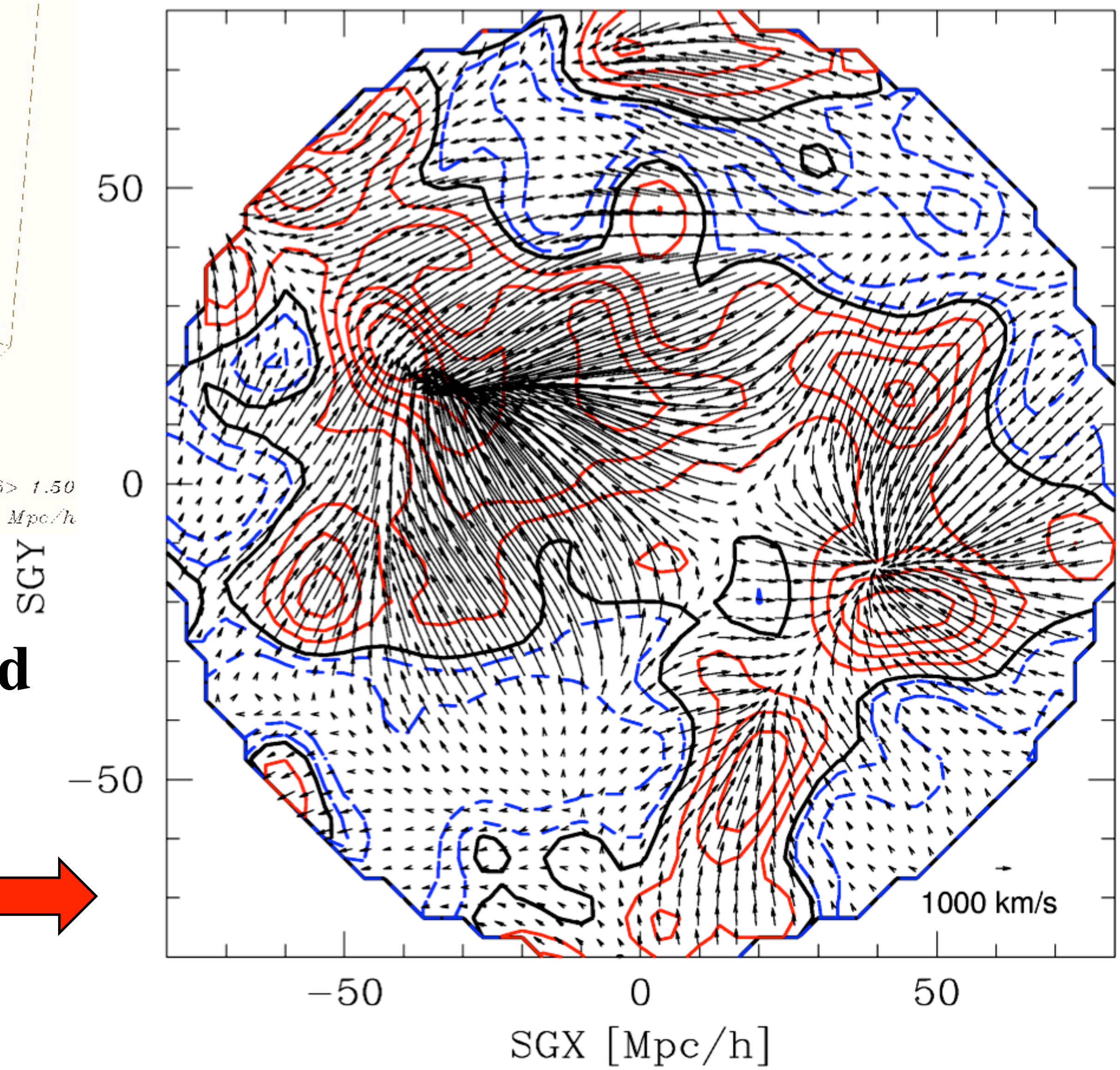


An Example: The PSCz Redshift Survey



Local 3-D density field

The corresponding velocity field



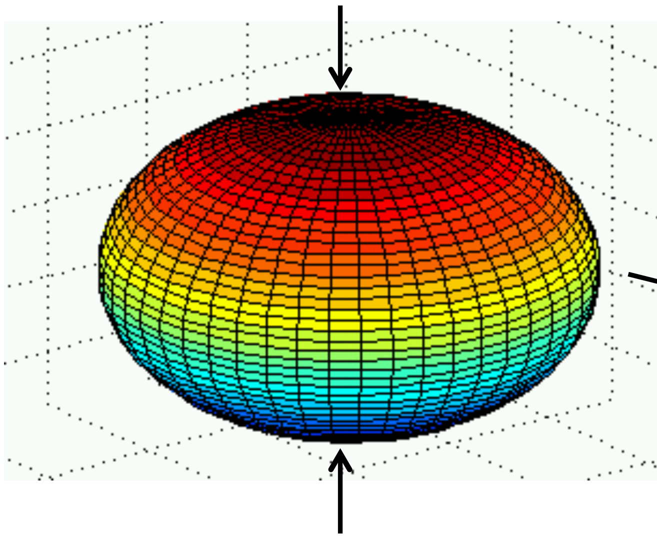
Peculiar Velocities: Summary

- Measurements of peculiar velocities are very, very tricky
  - Use (relative) distances to galaxies + Hubble flow, to infer the peculiar velocities of individual galaxies. Systematic errors?
  - Use a redshift survey + numerical modeling to infer the mass density distribution and the consistent peculiar velocity field
- Several general results:
  - We are falling towards Virgo with ~ 300 km/s, and will get there in about 10 - 15 Gyr
  - Our peculiar velocity dipole relative to CMB originates from within ~ 50 Mpc
  - The LSC is falling towards the Hydra-Centaurus Supercluster, with a speed of up to 500 km/s
  - The whole local ~ 100 Mpc volume may be falling towards a larger, more distant Shappley Concentration (of clusters)
- The mass and the light seem to be distributed in the same way on large scales (here and now)

Supplementary Slides

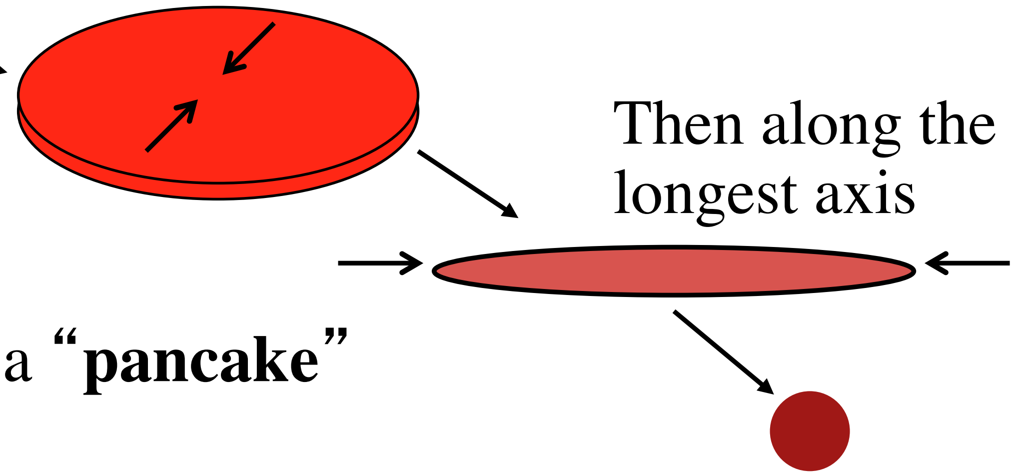
Non-Spherical Collapse

Real perturbations will not be spherical. Consider a collapse of an ellipsoidal overdensity:



The expansion turns into collapse along the shortest axis first …

Then along the intermediate axis



Then along the longest axis

- Perturbation first forms a "**pancake**"
- Then forms **filaments**
- Then forms **clusters**

This kinds of structures are seen both in numerical simulations of structure formation and in galaxy redshift surveys

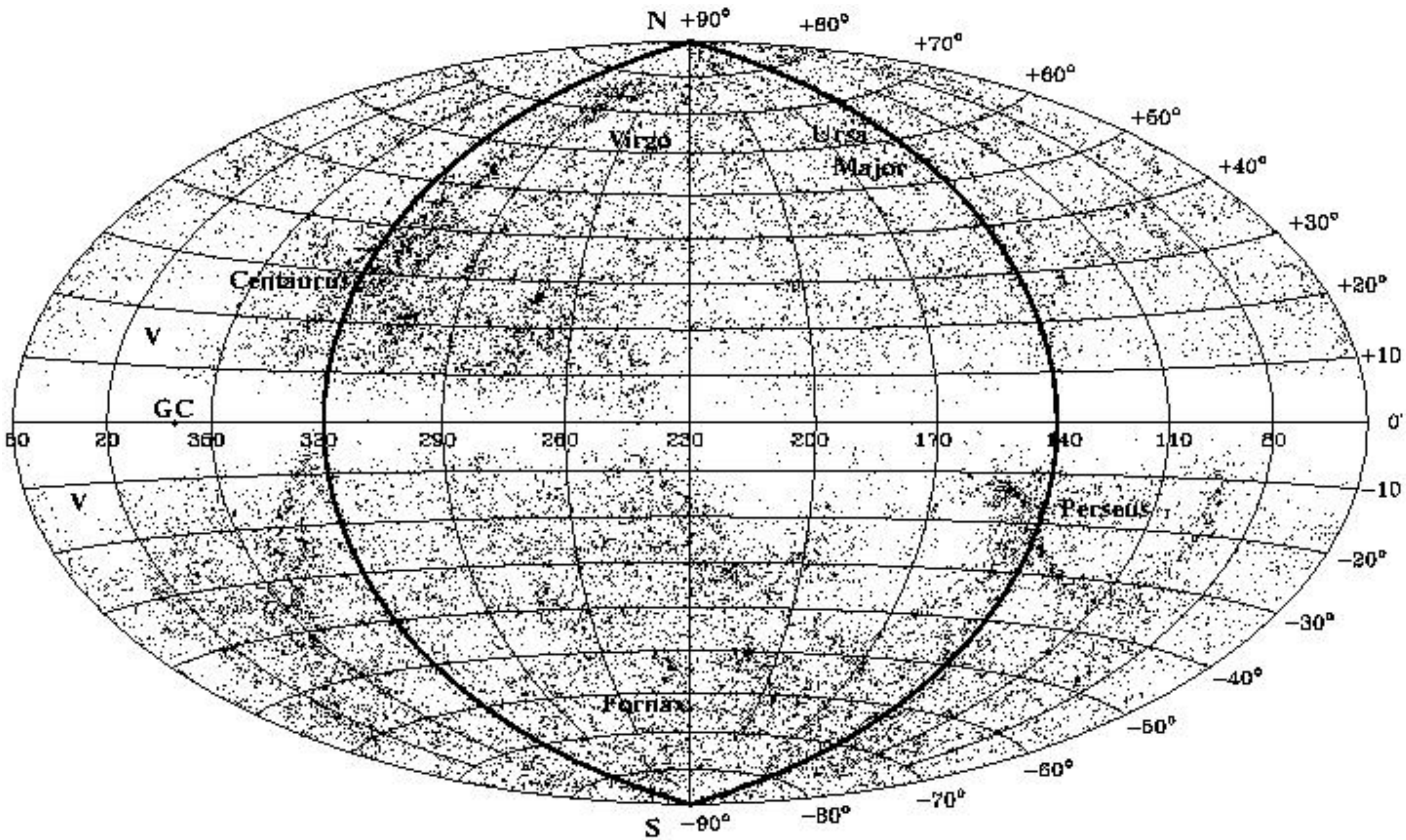
Structure Formation Theory: A Summary of the Key Ideas (1)

- Structure grows from initial density perturbations in the early universe, via gravitational infall and hierarchical merging
- Initial conditions described by the primordial density (Fourier power) spectrum $P(k)$, often assumed to be a power-law, e.g., $P(k) \sim k^n$, $n = 1$ is called a Harrison-Zeldovich spectrum
- Dark matter (DM) plays a key role: fluctuations can grow prior to the recombination; after the recombination, baryons fall in the potential wells of DM fluc's (proto-halos)
- Damping mechanisms erase small-scale fluctuations; how much, depends on the nature of the DM: HDM erases too much of the high-freq. power, CDM fits all the data
- Collapse occurs as blobs ➙ sheets ➙ filaments ➙ clusters

Structure Formation Theory: A Summary of the Key Ideas (2)

- Pure gravitational infall leads to overdensities of ~ 200 when the virialization is complete
- Free-fall time scales imply galaxy formation early on (t_{ff} ~ a few $\times 10^8$ yrs), clusters are still forming (t_{ff} ~ a few $\times 10^9$ yrs)
- Characteristic mass for gravitational instability is the Jeans mass; it grows before the recombination, then drops precipitously, from ~ $10^{16}\ M_\odot$, to ~ $10^5\ M_\odot$
- Cooling is a key concept:
  - Galaxies cool faster than the free-fall time: formation dominated by the dissipative processes, achieve high densities
  - Groups and clusters cool too slowly: formation dominated by self gravity, lower densities achieved
  - The cooling curve separates them

~15000 brightest galaxies on the sky, in Galactic coordinates. Solid line defines the Supergalactic plane

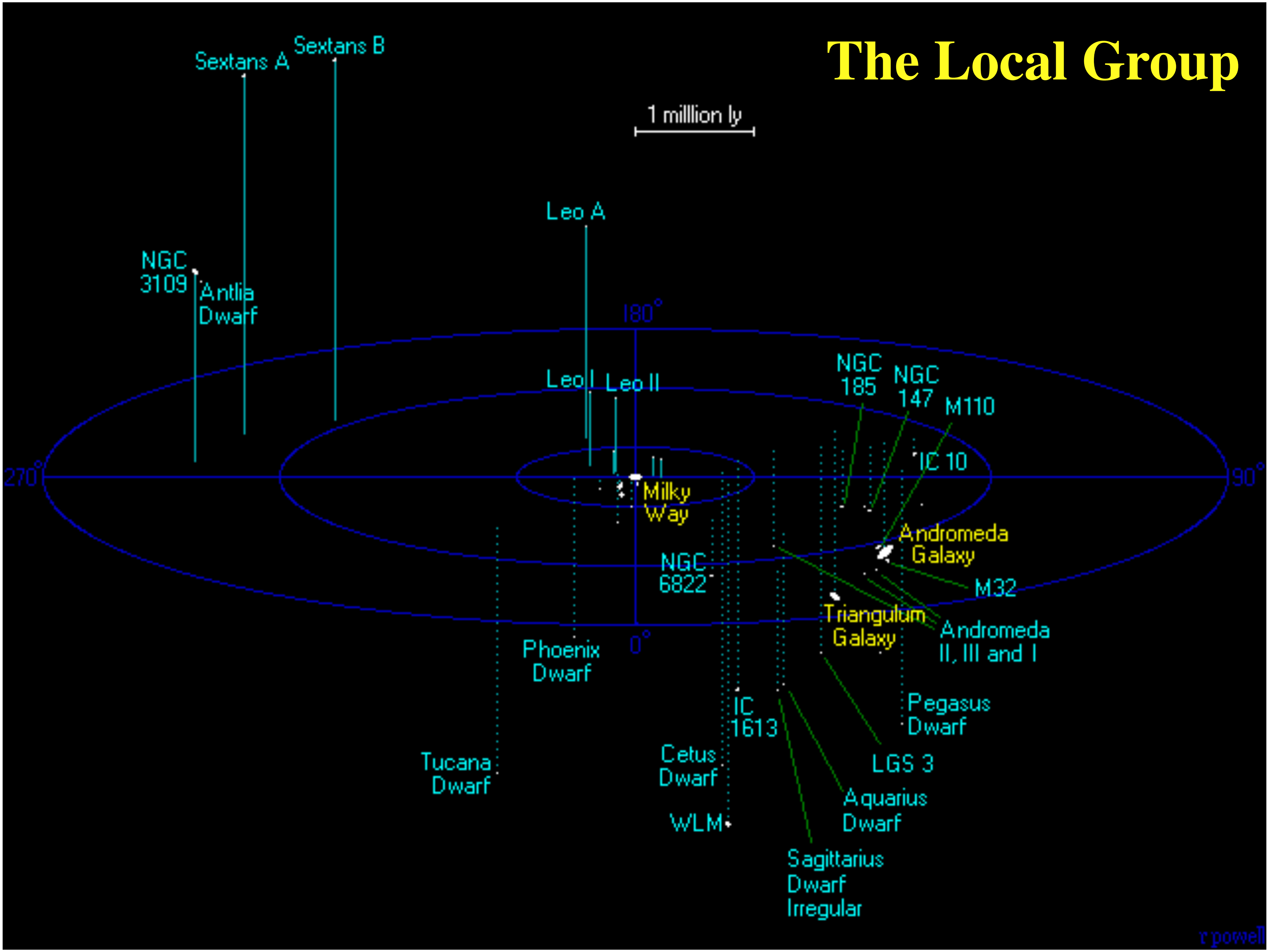


Representative Galaxy Catalogs/Surveys

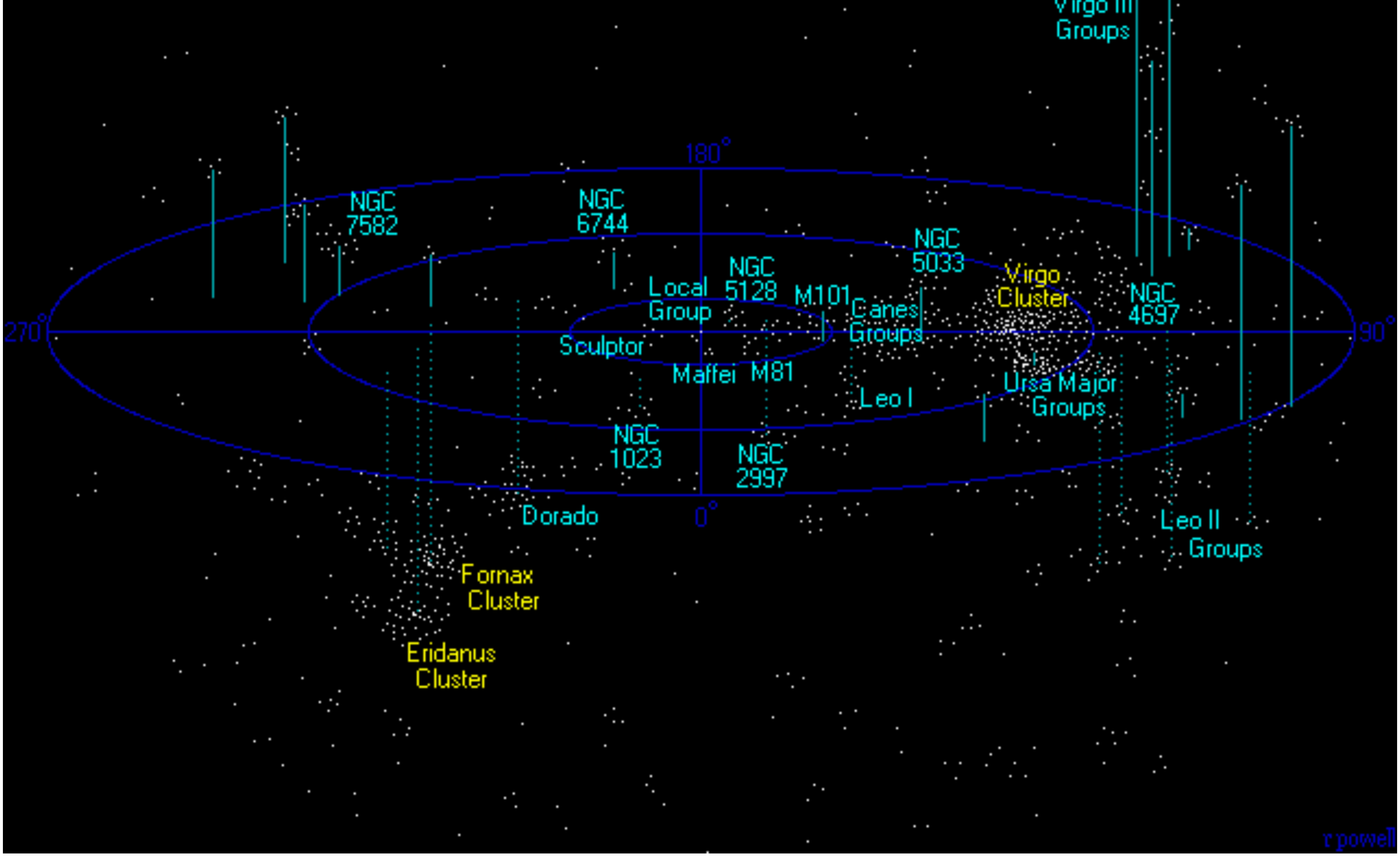
1970	Lick (Shane-Wirtanen)	1 M galaxies
1990	APM	2 M
1995	DPOSS	50 M
2005	SDSS	200 M
2015?	LSST	2000 M

Representative Local Redshift Surveys

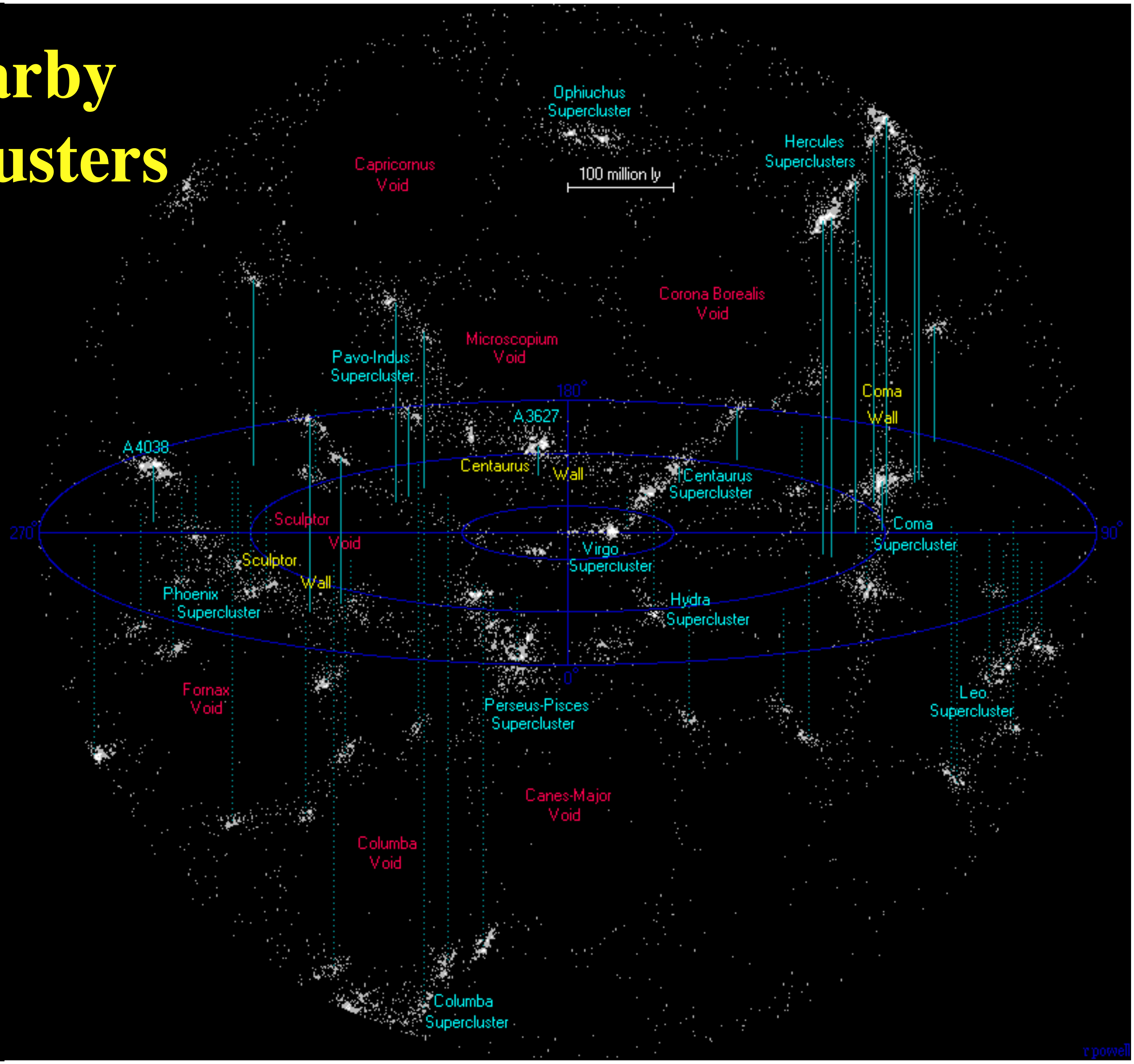
1985	CfA	2,500 galaxies
1992/5	IRAS	9,000
1995	CfA2	20,000
1996	LCRS	23,000
2003	2dF	250,000
2005/8	SDSS	900,000


The Local Group
1 milllion ly
Sextans A
Sextans B
NGC 3109
Antlia Dwarf
Leo A
180°
270°
90°
0°
Leo I
Leo II
NGC 185
NGC 147
M110
IC 10
Milky Way
Andromeda Galaxy
NGC 6822
M32
Triangulum Galaxy
Andromeda II, III and I
Phoenix Dwarf
IC 1613
Pegasus Dwarf
Tucana Dwarf
Cetus Dwarf
LGS 3
Aquarius Dwarf
WLM
Sagittarius Dwarf Irregular
r.powell

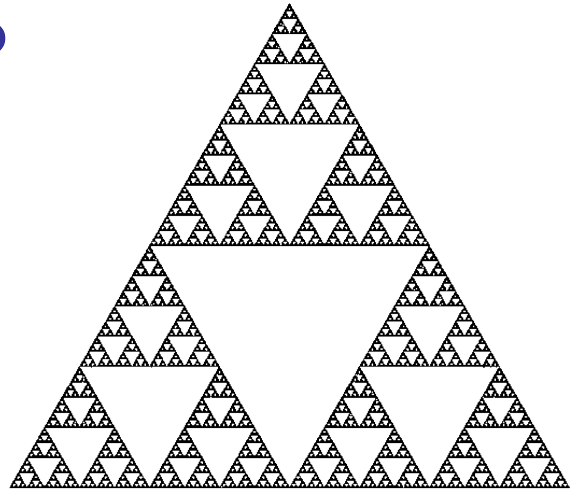
The Local Supercluster



The Nearby Superclusters



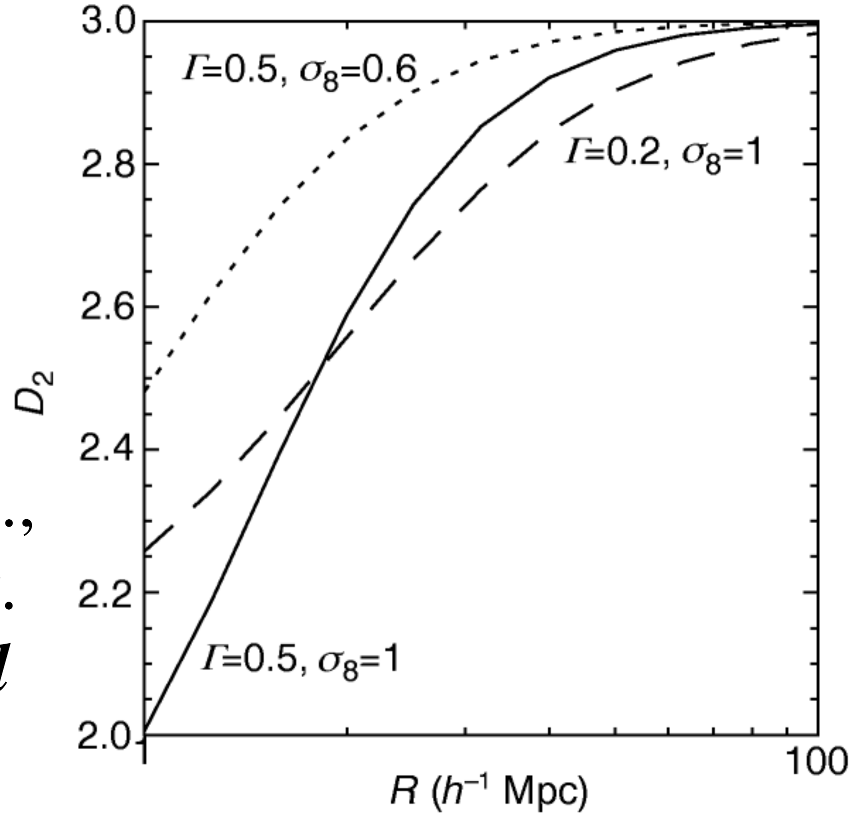
So, is the Universe Fractal?



For each galaxy, count the number of galaxies within distance R from it, $N(<R)$. If

$$N(<R) \propto R^{D_2}$$

…then the distribution can be described as a fractal with correlation dimension $D_2 = const.$ If $D_2 = 3$, then the distribution is consistent with being simple, homogenous in 3-D.



But in the real universe $D_2 \neq const.$, since $\boldsymbol{\xi}(\boldsymbol{r})$ is *not* a pure power-law. ***Thus, the universe is not a fractal*** (although it is pretty close)