INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

SMR.762 - 15

SUMMER SCHOOL IN HIGH ENERGY PHYSICS AND COSMOLOGY

13 June - 29 July 1994

STRUCTURE FORMATION IN THE UNIVERSE

S. MATARRESE
Department of Physics
University of Padova
Padova, Italy

Please note: These are preliminary notes intended for internal distribution only.

MAIN BUILDING STRADA COSTIERA, 11 TEL. 22401 TELEFAX 224163 TELEX 460392 ADRIATICO GUEST HOUSE VIA GRIGNANO, 9 TEL.224241 TELEFAX 224531 TELEX 460449
MICROPROCESSOR LAB. VIA BEIRUT, 31 TEL. 224471 TELEFAX 224600 TELEX 460392 GALILEO GUEST HOUSE VIA BEIRUT, 7 TEL. 22401 TELEFAX 2240310 TELEX 460392

Structure Formation in the Universe

Trieste, summer '94

S. Matarrese,
Physics Dept.
Padova Univ., ITALY

Contents

- The standard model

units
$c = \hbar = k_B = 1$

spatial homogeneity & isotropy =>

$$ds^2 = dt^2 - a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2)\right]$$

Robertson-Walker metric

$$k = \begin{cases} +1 & \text{closed} \\ 0 & \text{flat ("Einstein-de Sitter")} \\ -1 & \text{open} \end{cases}$$

$a(t) \equiv$ "scale factor" $= \mathcal{O}$(length)

$H(t) \equiv \frac{d}{dt} \ln a(t) = \frac{\dot{a}}{a} =$ "Hubble constant"
$\stackrel{!}{=} \mathcal{O}$(inverse time)

$q(t) \equiv -\left(\frac{\ddot{a}}{a}\right) / H^2 =$ "deceleration parameter"
$= \mathcal{O}(\#)$

... kinematics $\longrightarrow$ dynamics ...

– The Friedmann–Robertson–Walker model

Einstein's eqs. $\Longrightarrow \begin{cases} H^2 = \dfrac{8\pi G}{3}\rho - \dfrac{k}{a^2} \\ \dfrac{\ddot{a}}{a} = -\dfrac{4\pi G}{3}(\rho + 3p) \end{cases}$

$\boxed{\text{FRIEDMANN EQS.}}$

energy mom. continuity $\Longrightarrow \{ \dot{\rho} = -3H(\rho + p)$

SOURCE:

perfect fluid with pressure p and energy density ρ. Eq. of state, e.g.:

$$\boxed{p = (\gamma - 1)\rho \equiv w\rho}$$

$\gamma \equiv$ "adiabatic index"

ex.:

$p = \rho$	$\gamma = 2$	stiff matter
$p = \frac{1}{3}\rho$	$\gamma = 4/3$	radiation ← ultra-relativistic
$p \approx 0$	$\gamma = 1$	matter ≈ non-relativistic
$p = -\frac{1}{3}\rho$	$\gamma = 2/3$	cosmic strings
$p = -\frac{2}{3}\rho$	$\gamma = 1/3$	domain walls
$p = -\rho$	$\gamma = 0$	vacuum (cosm. constant)

speed of sound: $c_s^2 \equiv \left.\dfrac{\partial p}{\partial \rho}\right|_{s=\text{const}} = \gamma - 1 \quad 0 \leq c_s^2 \leq 1$

En. Density vs. time

$\longrightarrow$ $\rho \propto a^{-3\gamma}$ $\Leftarrow$ from continuity equation

ex.: $\begin{cases} \rho \propto a^{-4} & \text{radiation} \\ \rho \propto a^{-3} & \text{matter (dust)} \\ \rho \simeq \text{const.} & \text{vacuum (cosmological constant)} \end{cases}$

$\longrightarrow$ if $k=0$ or the spatial curvature is negligible: $\frac{8\pi G}{3}\rho \gg \frac{k}{a^2}$.

$\longrightarrow$ $a(t) \propto t^{2/(3\gamma)}$ $\Leftarrow$ from $H^2 \approx \frac{8\pi G}{3}\rho$

ex.: $\begin{cases} a \propto t^{1/2} & \text{radiation} \\ a \propto t^{2/3} & \text{matter} \\ a \propto t & \text{curvature: } \rho=0,\ k=-1 \end{cases}$

i.e. free expansion

in all cases (except $\gamma = 1$)

$\rho \propto t^{-2}$

Critical density

$\rho = \rho_c$ if $k = 0 \Longrightarrow$

$$\rho_c = 3H^2/8\pi G$$

more in general we define a

Density parameter

$$\boxed{q = \frac{1}{2}\Omega(3\gamma - 2)}$$

deceleration parameter

$$\Omega(t) \equiv \frac{\rho(t)}{\rho_c(t)} = \frac{8\pi G \rho(t)}{3H^2(t)}$$

(for vanishing cosmological constant, $\Lambda = 0$)

Ω is related to spatial curvature, k:

$\Omega > 1$	$\Longleftrightarrow$	$k = 1$
$\Omega = 1$	$\Longleftrightarrow$	$k = 0$
$\Omega < 1$	$\Longleftrightarrow$	$k = -1$

$\Longrightarrow$ If $\Omega \gtreqless 1$ at t_0 ("now") $\Longrightarrow$
$\Omega \gtreqless 1$ always: k = constant of motion !!

Hubble law within the standard cosmological model

$$ds^2 = dt^2 - a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2)\right]$$

For a photon ($ds=0$) moving along the radial direction:

$$\frac{dt}{a(t)} = -\frac{dr}{\sqrt{1-kr^2}}$$

Using $\delta t_e \propto \lambda_e$, $\delta t_0 \propto \lambda_0$ one gets

$$\frac{\lambda_e}{\lambda_0} = \frac{a(t_e)}{a_0}$$ e = emission in the past

0 = observation now

$$\Rightarrow \quad 1+z = \frac{\lambda_0}{\lambda} = \frac{a(t_0)}{a(t)} \Rightarrow \frac{a(t_0)}{a(t_e)} > 1$$

$$\rightarrow \quad cz = H_0 d \qquad\qquad d_L = r a_0 (1+z)$$

Also:

$$d = ar \rightarrow \dot{d} = \dot{a}r = \frac{\dot{a}}{a} d \Rightarrow$$

$$v_r = Hd$$

The Hubble constant, H_0

$H_0 = H(t_0)$: its present value

$\boxed{\text{Hubble law}}$: $\quad cz = H_0 d_L$

$z \equiv$ "redshift" $\equiv \dfrac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}}$

$$d_L \equiv \left(\frac{\overbrace{\mathcal{L}}^{\text{object luminosity}}}{4\pi \underbrace{\mathcal{F}}_{\text{observed flux}}}\right)^{1/2} = \text{"luminosity distance"}$$

$H_0 \equiv$ Hubble constant $= 100\, h$ km sec^{-1} Mpc^{-1} *

$0.4 \lesssim h \lesssim 1$ $\leftarrow$ systematic uncertainty

Sandage & co. $\qquad$ de Vaucouleurs & co.

$v_r = cz \;\; \longrightarrow$ 1 Mpc distance corresponds to $v_r = 100\, h$ km/sec

also: $H_0^{-1} = 9.78\, h^{-1} \times 10^9$ yr

$cH_0^{-1} = 3{,}000\, h^{-1}$ Mpc

---

* $1\,\text{pc} \simeq 3\,\text{l.yr} \simeq 3 \times 10^{13}\,\text{km} \Rightarrow 1\,\text{Mpc} \simeq 3 \times 10^{19}\,\text{km}$

The cosmological distance ladder

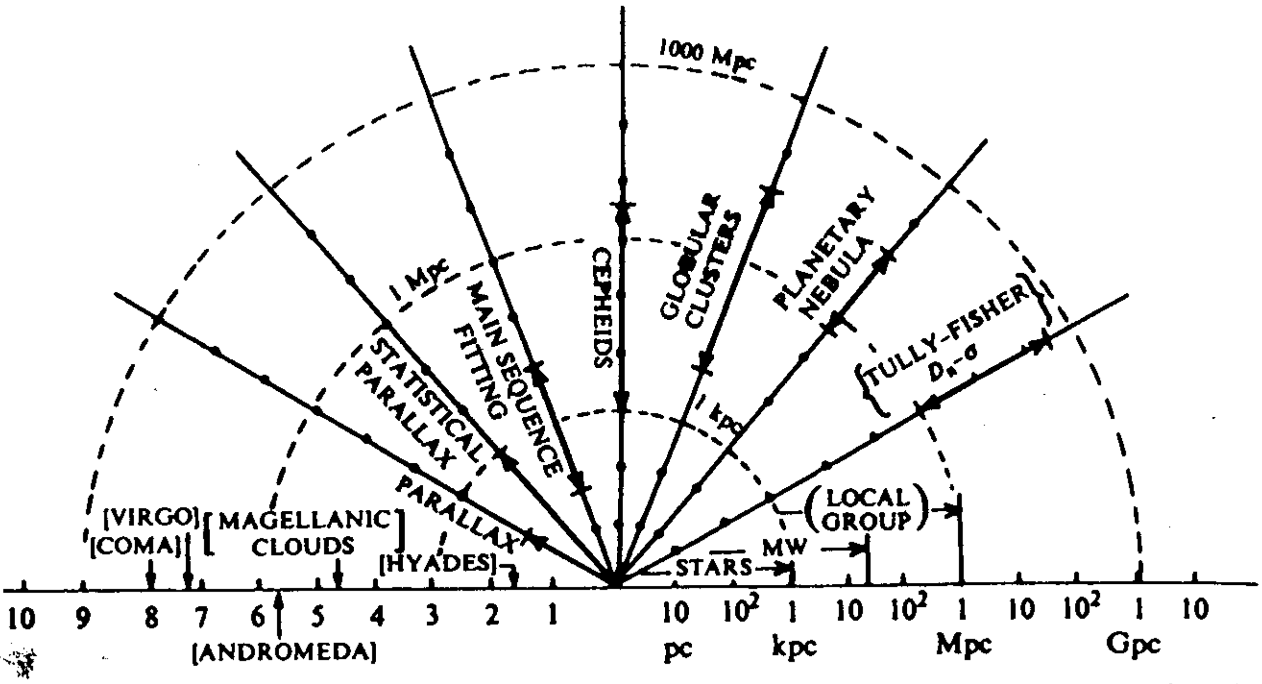


Fig. 1.4. The various distance measurement techniques are summarized. The radial lines are marked by thick, double-arrowed, segments indicating the range over which each of the method is used. The horizontal line on the right indicates the actual distance scale and the typical sizes of various structures discussed in the text. The locations of several astronomical objects are shown in the horizontal axis on the left. [from Padmanabhan, 1993]

H_0 determination: measuring extragalactic distances

THE COSMOLOGICAL DISTANCE LADDER
(a method of calibration)

1. $d \lesssim 3pc$ — For nearby stars in our galaxy one uses trigonometric methods (using earth's orbital motion): parallax

2. $50pc \lesssim d \lesssim 10\,kpc$ — Determine by (statistical) parallax the distance of Hyades star cluster [(46±2)pc] and use it to calibrate distance of ZAMS stars (by abs. luminosity). This is a stellar evolution based method

3. $d \lesssim 1 \div 3 Mpc$ — For larger distances use variable stars like Cepheids, for which there is a fixed relation between period and lumin. To calibrate the relation one uses C. in Magellanic clouds ($\simeq 20\%$ errors)

$d \leq 20 Mpc$ Similarly for Planetary nebulae.

4. $d \leq 10^2 Mpc$ — Stellar evolution based are the brightest stars of a galaxy used as distance indicators

5. $d \gtrsim 10^2 Mpc$ — On the largest distances one can use brightest galaxies in g. clusters

Galaxy distance determination

For spiral galaxies:

TULLY FISHER (in the infrared!)

$$v_c = 220 \left(\frac{L}{L_*}\right)^{0.22} \frac{km}{sec}$$

v_c: circular velocity

L_*: characteristic gal. luminosity

For elliptical galaxies:

FABER JACKSON

$$\sigma_r = \frac{220}{\sqrt{2}} \left(\frac{L}{L_*}\right)^{0.25} \frac{km}{sec}$$

σ_r: 1-D velocity dispersion

L_*: characteristic gal. luminosity

Energy density of the Universe

$$\rho_{oc} = \frac{3H_0^2}{8\pi G} \simeq 1.9 \times 10^{-29} h^2 \, g/cm^3$$

$$\simeq 2.78 \times 10^{11} M_\odot \, Mpc^{-3}$$

$$\boxed{\text{CONTRIBUTIONS TO } \Omega_0 = \rho_0/\rho_{oc}}$$

- The Cosmic Microwave Background (CMB).
  The most recent measurement of the temperature $T_{0\gamma}$ (Mather et al. '94)

  $$T_{0\gamma} = 2.726 \pm 0.010 \quad {}^\circ K$$ by COBE FIRAS (95% CL)

  $$\rho_{0\gamma} = \frac{\pi^2}{15} T_{0\gamma}^4 \simeq 4.8 \times 10^{-34} g/cm^3$$

  $$\rightarrow \boxed{\Omega_{0\gamma} \simeq 2.6 \times 10^{-5} h^{-2}} \lll 1$$

- Massless neutrinos give a similar contribution but $T_{0\nu} = T_{0\gamma}\left(\frac{4}{11}\right)^{1/3} = 1.9\,{}^\circ K$

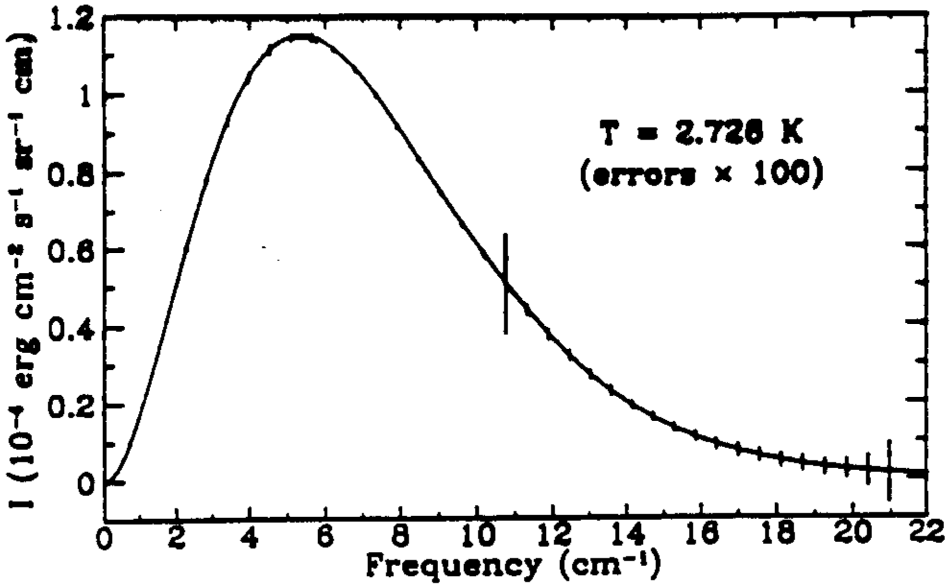
  $$\rho_{0\nu} \simeq \underbrace{N_\nu}_{\text{\# of massless } \nu \text{ flavours}} 10^{-34} g/cm^3$$

  $$\rightarrow \Omega_{0\nu} \lll 1$$

$1 \, M_\odot \simeq 1.99 \times 10^{33} g$

$1 \, L_\odot \simeq 3.9 \times 10^{33} erg \, sec^{-1}$

The CMB spectrum



CMBR brightness vs. frequency (inverse wavelength) measured by COBE
(Mather et al. '94) FIRAS

$$T_0 = 2.726 \pm 0.010 \quad (95\% \, CL)$$

- Baryonic matter : the theory of cosmological nucleosynthesis of light elements (D, ^3He, ^4He, ^7Li, ..) implies a bound on the amount of baryonic material

$$0.0057 h^{-2} \leq \Omega_B \leq 0.011 h^{-2}$$

Recent measurements of D by Keck implies

$$\Omega_B \simeq 0.0068 h^{-2}$$ [Krauss & Kernan '94]

$$\longrightarrow \quad \Omega_B \ll 1$$

- Mass in galaxies :

i) LUMINOUS MATTER The luminous matter in galaxies can be estimated as

$$\Omega_g = \frac{\rho_g}{\rho_{oc}} = \frac{\langle \mathcal{L}_g \rangle}{\rho_{oc}} \left(\frac{M}{L}\right) \simeq 0.86\, h^{-1} \times 10^{-3} \left(\frac{M}{L}\right) \Big/ \left(\frac{M_\odot}{L_\odot}\right)$$

mean luminosity

$$\langle \mathcal{L}_g \rangle = \int_0^\infty d\mathcal{L}\, \mathcal{L}\, \Phi(\mathcal{L}) \simeq 2.4\, h \times 10^8 L_\odot \, Mpc^{-3}$$

↳ galaxy luminosity function

$\Phi(\mathcal{L})$ = mean # of gal. / vol. in the $\mathcal{L}$ range $(L, L+dL)$

$$\stackrel{!}{=} \phi_* (\mathcal{L}/\mathcal{L}_*)^\alpha \exp(-\mathcal{L}/\mathcal{L}_*) \quad \leftarrow \text{Schechter}$$

$\alpha = -1.07 \pm 0.05$, $\mathcal{L}_* = 1.0 \times 10^{10} h^{-2} L_\odot$, $\phi_* = 0.010 h^3 Mpc^{-3}$

The M/L issue

one needs $M/L = 1{,}200\, h\, M_\odot / L_\odot$ to get $\Omega_g = 1$

but observations give $M/L \simeq 5$ $\Rightarrow$

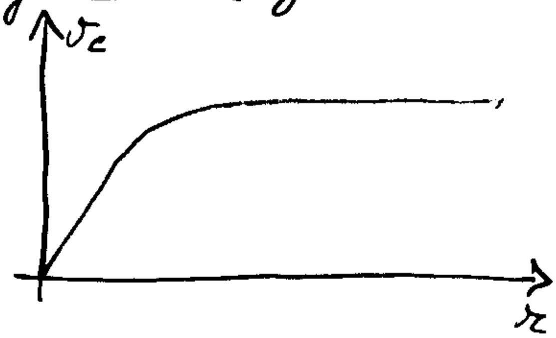
$$\Omega_g \simeq 0{,}004\, h^{-1} \quad \text{(luminous)}$$

- GALAXY & CLUSTER HALOS

evidence for **DARK MATTER**

rotation curves of spiral galaxies give

$$M(r) \simeq \frac{r\, v_c^2(r)}{G}$$



this mass ~ 10 times the luminous mass

similar results obtained from v_σ of ellipticals

galaxy groups: **$M/L \sim 200\, h$**

galaxy clusters **$M/L \sim 300\, h$**

$\longrightarrow \Omega \sim 0{,}1 \div 0{,}2$

Ω_0 from peculiar motions

Linear theory provides a local relation between the density fluctuation δ and the peculiar (i.e. after Hubble flow subtraction) velocity:

at t_0 and measuring v in km/s $\rightsquigarrow \vec{\nabla} \cdot \vec{v} = -f(\Omega_0)\,\delta$

$\hookrightarrow \simeq \Omega_0^{0.6}$

Once $\delta M/M$ and $\vec{v}$ are known one gets Ω_0

However ...

$$\frac{\delta N}{N} \neq \frac{\delta M}{M}$$

observed galaxy # fluctuations

theoretical mass fluctuation

$\Longrightarrow$ **BIAS** (Kaiser '84, ..)

$$\frac{\delta N}{N} = b\,\frac{\delta M}{M}$$

b = "bias factor" ≥ 1

$$\longrightarrow \; -\frac{\vec{\nabla}\cdot\vec{v}}{\delta_g} = \frac{f(\Omega_0)}{b} \longleftarrow$$

"measured" quantity

$f(\Omega)'' = 1.3^{+0.4}_{-0.3}$

$\hookrightarrow \Omega_{dyn}$ compatible with 1

INFLATION (?)

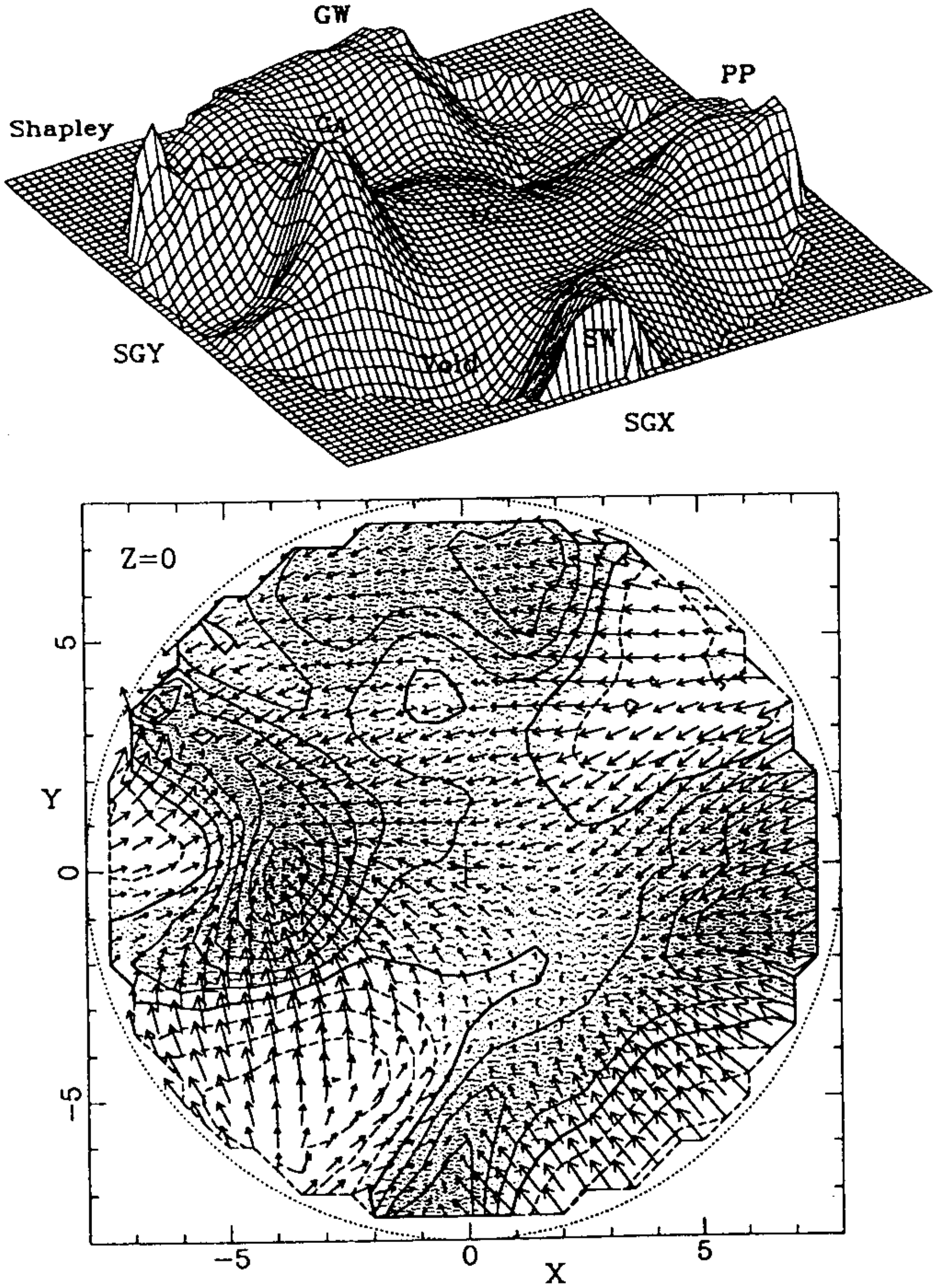


Figure 3: The fluctuation fields of velocity and *mass*-density in the Supergalactic plane as recovered by POTENT from the Mark III velocities of $\sim$ 3000 galaxies with 12 h^{-1}Mpc smoothing. The vectors shown are projections of the 3D velocity field in the CMB frame. Distances and velocities are in 1000 km s^{-1}. Contour spacing is 0.2 in δ, with the heavy contour marking $\delta = 0$ and dashed contours $\delta < 0$. The LG is at the center. The GA is on the left, PP on the right, and Coma is at the top. The grey-scale in the contour map and the height of the surface in the landscape map are proportional to δ (Dekel *et al.* 1994).

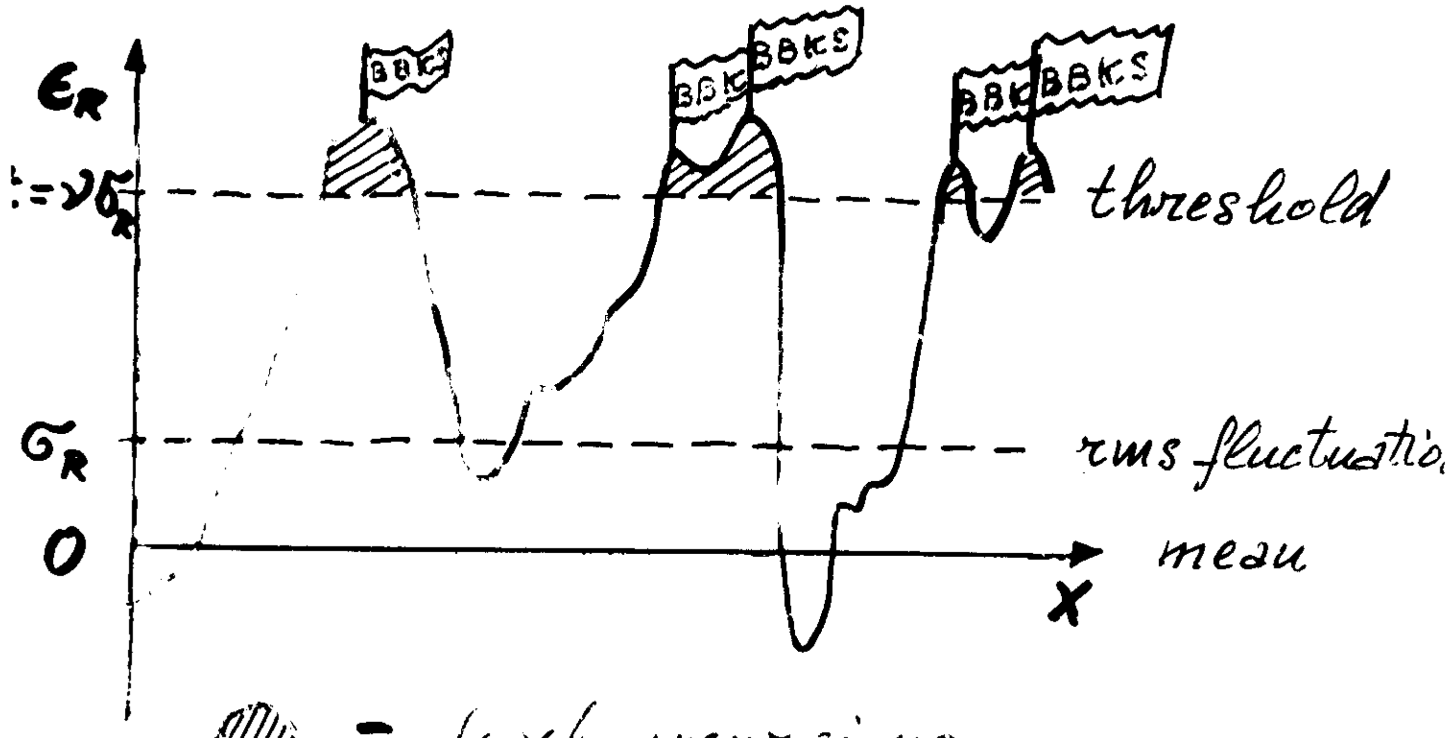
RARE EVENTS (= BIAS)

REFS.

Gaussian random fields: Doroshkevich, 1970; Kaiser, 1984; Politzer and Wise, 1984; Peacock and Heavens, 1985; Bardeen et al. (BBKS), 1986; Jensen and Szalay, 1986; Couchman, 1987; Cline et al., 1987; ...

non-Gaussian random fields: Grinstein and Wise, 1986; S.M., Lucchin and Bonometto, 1986; Fry, 1986; Lucchin and S.M., 1987; [illegible]

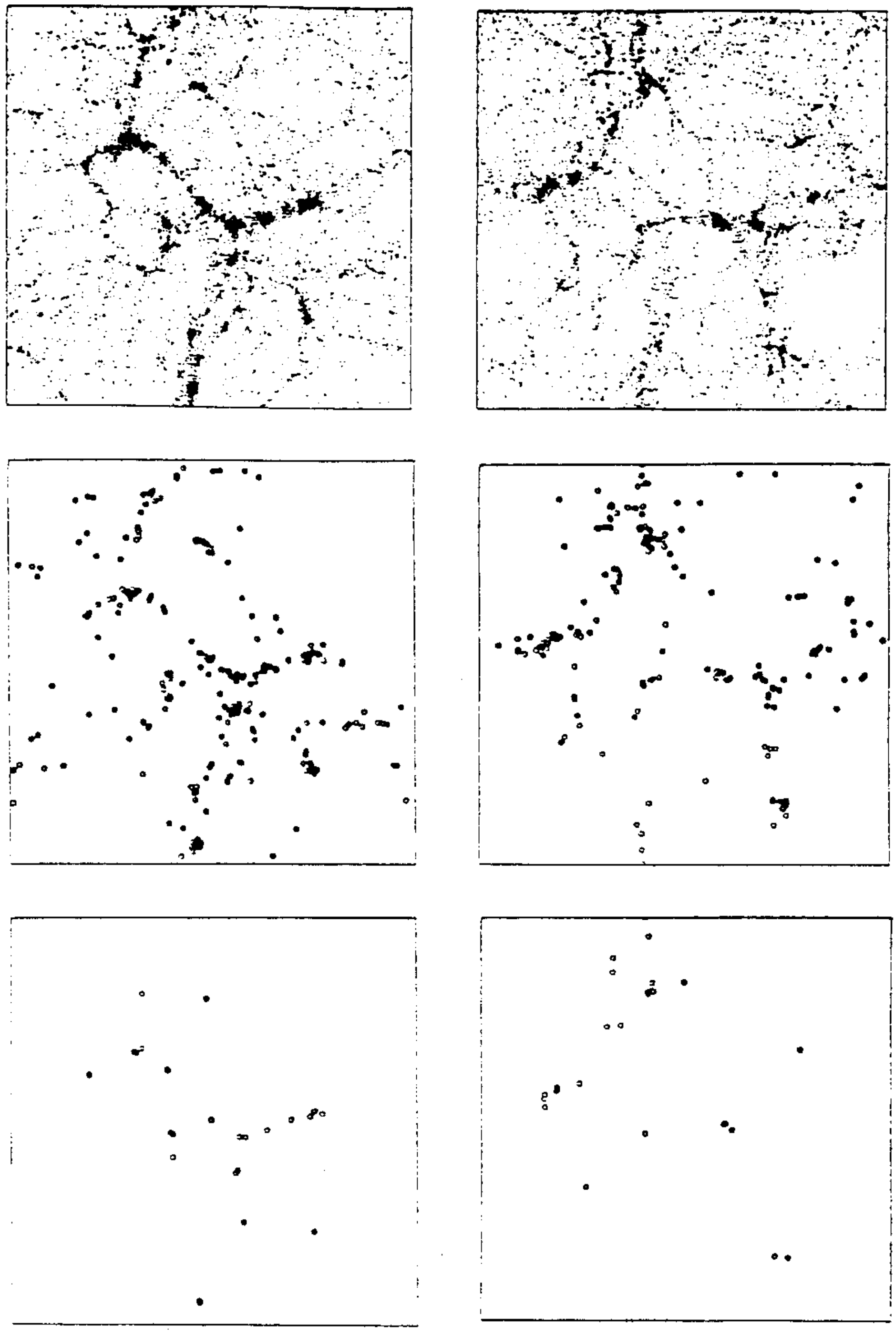
+ papers in 2-dim (CMB statistics)



≡ level excursions

≡ peaks (≡ local maxima)

effect of bias



from : White, Davis, Efstathiou and Frenk, 1987.

Age of the Universe

remind $H_0^{-1} = 9.78\, h^{-1} \times 10^9$ yr Hubble time

FRW models: $t_0 = f(\Omega_0) H_0^{-1}$

$f(\Omega_0)$ larger with smaller Ω_0

$\Omega_0 = 1 \quad \to \quad f = 2/3$

$\boxed{\text{OBSERVATIONAL LIMITS}}$:

Nuclear cosmochronology (radioactive dating of r-process elements)

Re/Os & Ur/Th $\to t_0 = 10-20$ Gyr

White dwarf cooling $\quad t_0 = 10 \pm 2$ Gyr

Age of oldest globular clusters in the Galaxy

$t_0 = (13-15) \pm 3$ Gyr

from: Nicolaci da Costa et al., 1994, ApJ., 424, L1

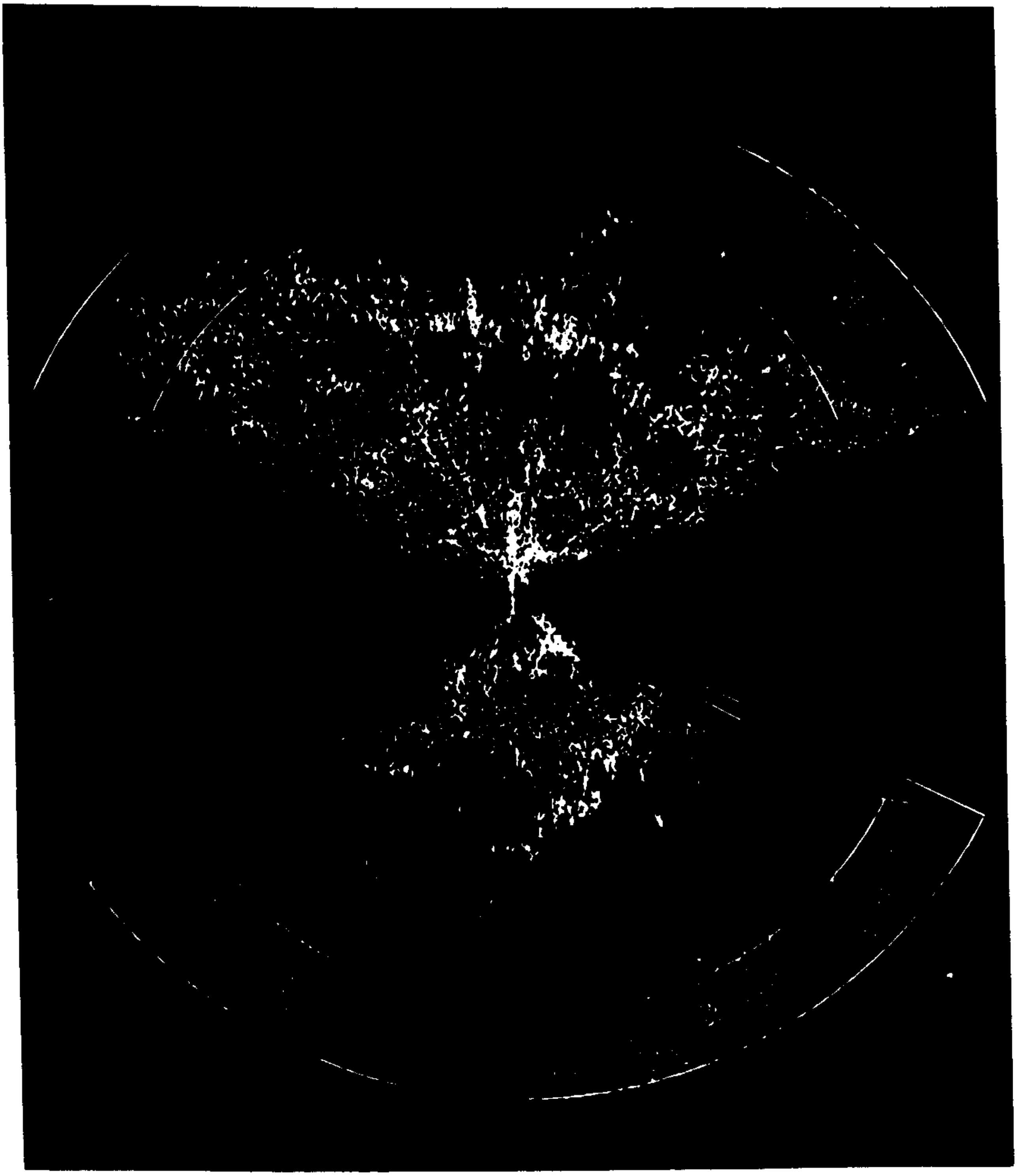


FIG. 2.—Display of CfA2 (north, *upper portion*) and SSRS2 (south, *lower portion*) for $cz < 12{,}000$ km s^{-1}. For CfA2 the box shows the declination limits $8.5° \le \delta \le 44.5°$ and the right ascension limits $8^h \le \alpha \le 17^h$. The inner arc is $\delta = 44.5°$. In the south the box limits are $-40° \le \delta \le -2.5°$ and the right ascension limits are $20^h8 \le \alpha \le 4^h$. The inner arc is $\delta = -2.5°$. There are 9325 galaxies (*points*) in the image. (Graphics by E. E. Falco and M. J. Geller.)

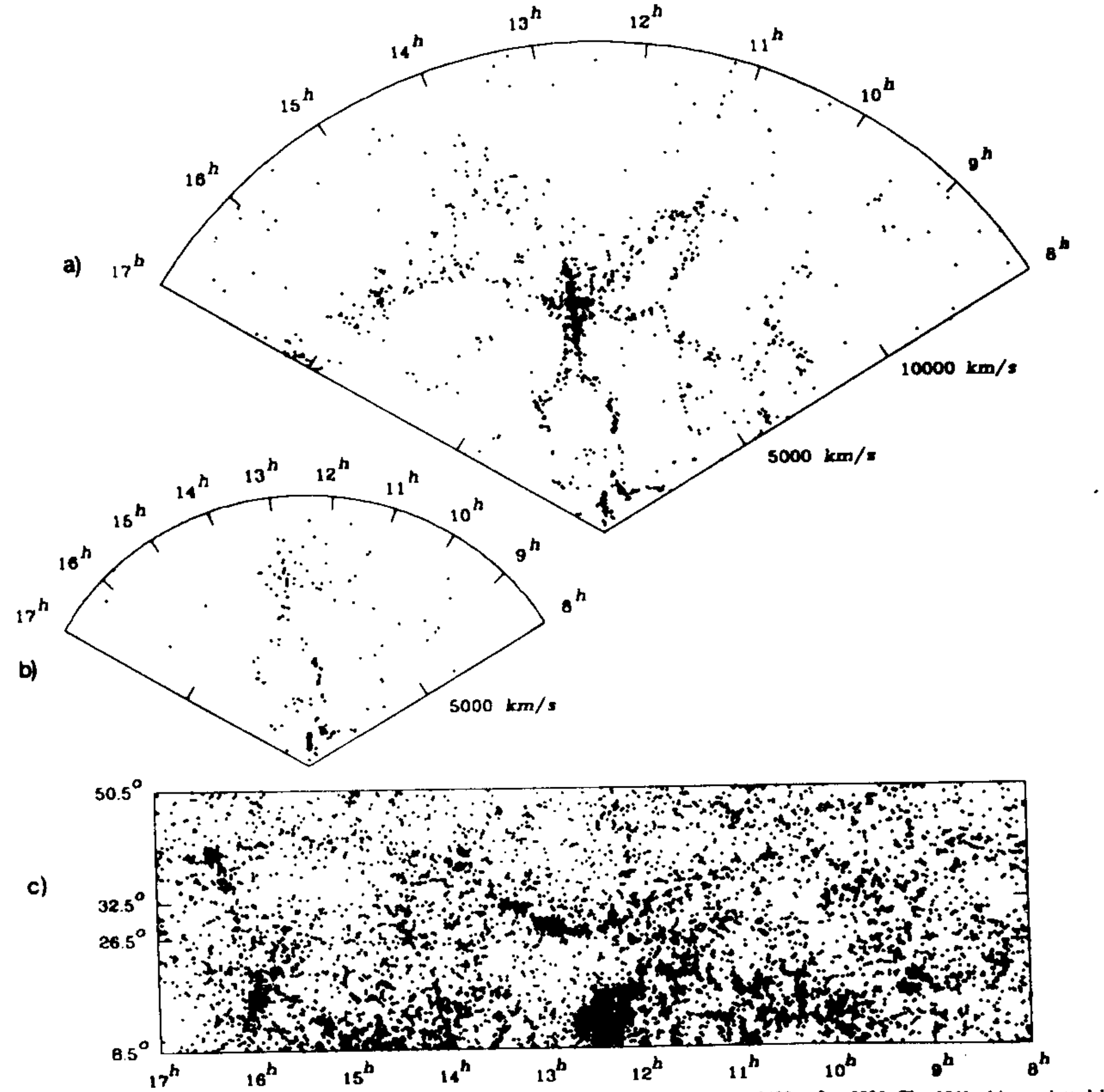


FIG. 1.—(*a*) Map of the observed velocity plotted vs. right ascension in the declination wedge $26°5 \le \delta \le 32°5$. The 1061 objects plotted have $m_B \le 15.5$ and $V \le 15{,}000\ \mathrm{km\ s^{-1}}$. (*b*) Same as Fig. 1*a* for $m_B \le 14.5$ and $V \le 10{,}000\ \mathrm{km\ s^{-1}}$. The plot contains 182 galaxies. (*c*) Projected map of the 7031 objects with $m_B \le 15.5$, listed by Zwicky *et al.* in the region bounded by $8^h \le \alpha \le 17^h$ and $8°5 \le \delta \le 50°5$.

galaxies appear to be on the surfaces of bubble-like structures. The bubbles have typical diameter of $\sim 25\ h^{-1}$ Mpc.

"Great Wall" in CfA survey (Geller & Huchra 1989)

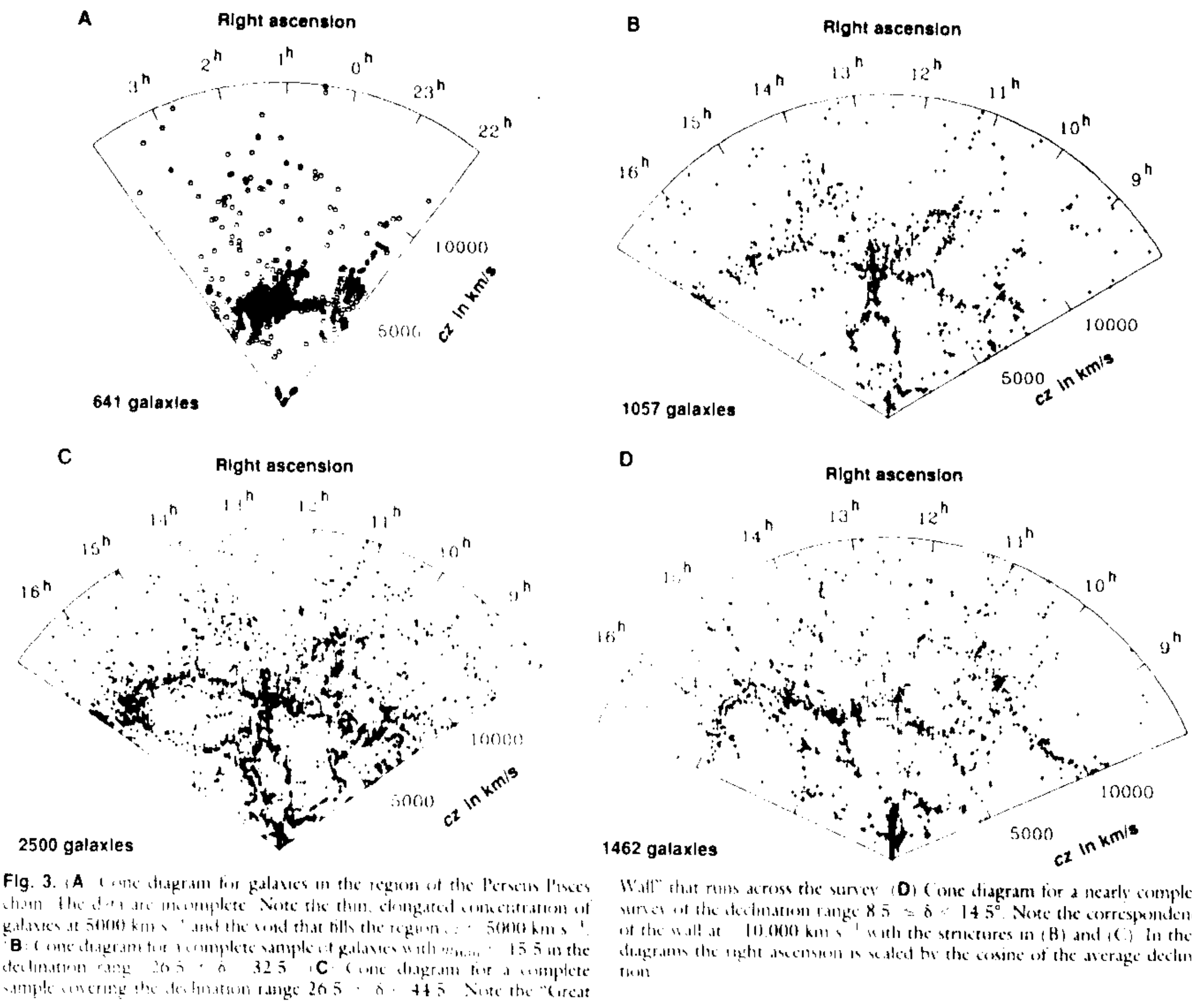


Fig. 3. (**A**) Cone diagram for galaxies in the region of the Perseus-Pisces chain. The data are incomplete. Note the thin, elongated concentration of galaxies at 5000 km s^{-1} and the void that fills the region $cz < 5000$ km s^{-1}. (**B**) Cone diagram for a complete sample of galaxies with $m_{B(0)} \leq 15.5$ in the declination range $26.5° \leq \delta < 32.5°$. (**C**) Cone diagram for a complete sample covering the declination range $26.5° \leq \delta < 44.5°$. Note the "Great Wall" that runs across the survey. (**D**) Cone diagram for a nearly complete survey of the declination range $8.5° \leq \delta < 14.5°$. Note the correspondence of the wall at ~10,000 km s^{-1} with the structures in (B) and (C). In these diagrams the right ascension is scaled by the cosine of the average declination.

"Pencil-beam" survey (Broadhurst et al. '90)

faint

bright

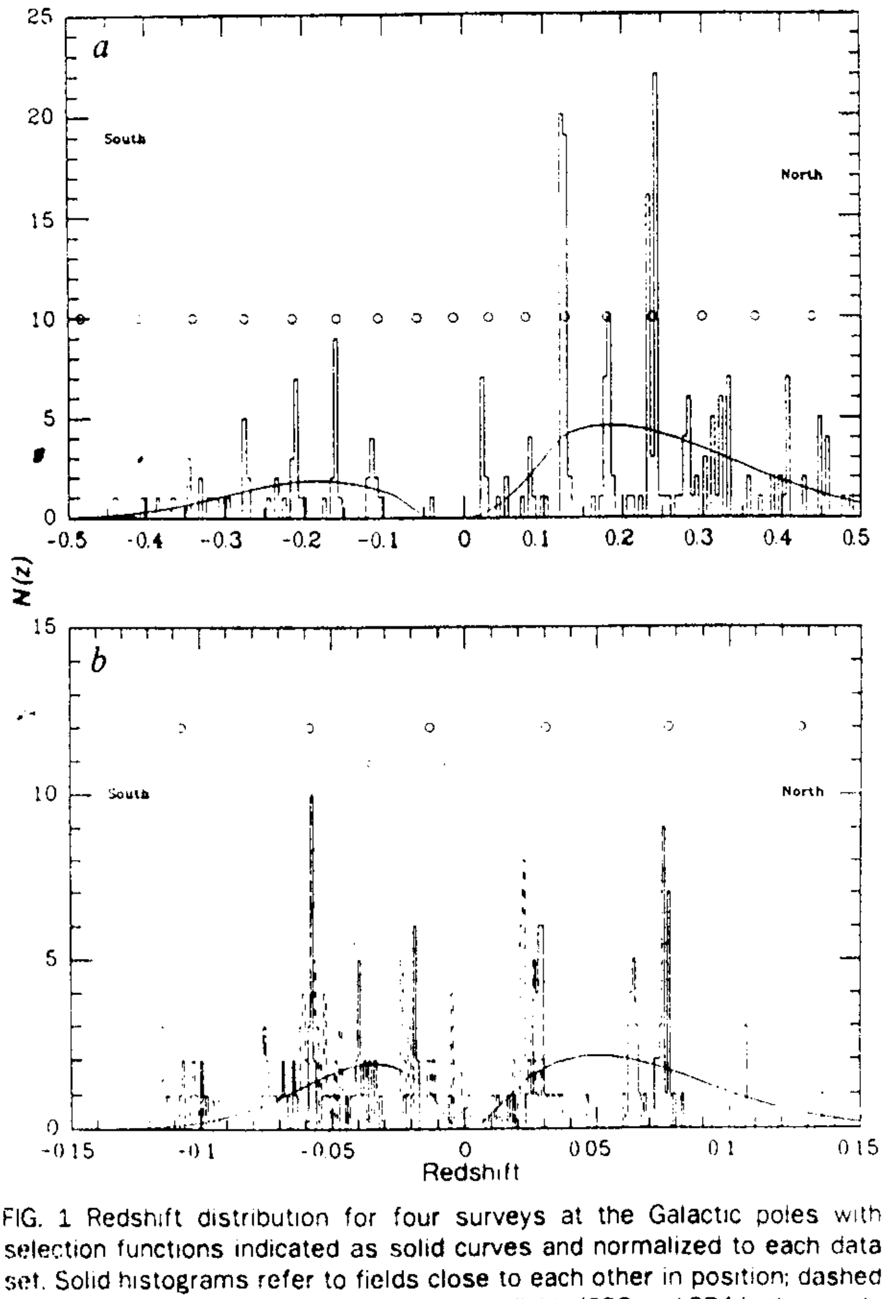


FIG. 1 Redshift distribution for four surveys at the Galactic poles with selection functions indicated as solid curves and normalized to each data set. Solid histograms refer to fields close to each other in position; dashed histograms refer to more widely distributed fields (SP3 and SP4 in the south. NP4 in the north[3,4]). *a* refers to faint samples[5,6] and *b* to bright surveys[2-4]. Circles indicate a best-fit constant comoving separation of 128 h^{-1} Mpc for $q_0 = 0.5$.

23

⇒ The large-scale structure of the Universe

As we already noticed, on small scales the ~~[illegible]~~ matter distribution tends to deviate from a merely homogeneous and isotropic one. One sees structures such as planets, stars, galaxies, groups, clusters of galaxies, superclusters, voids, filaments and sheets of galaxies, ...
Studying the spatial distribution of galaxies and galaxy systems is one of the aims of cosmology.
To this aim one has to make use of either two-dimensional (i.e. sky-projected) or three-dimensional (i.e. containing the

redshift as third spatial dimension) catalogs of galaxies or galaxy systems.

Two-dimensional catalogs include:

- The Zwicky sample, containing the angular position of 3753 galaxies
- The Lick sample. Here galaxies are summed in 10'×10' cells.
- The Jagellonian field, including more than 10,000 galaxies
- The APM (Automatic Plate Machine) Galaxy Survey, consists of some 5 million galaxies

Samples that include redshifts are:

- The CfA (Center for Astrophysics) survey[8] (now containing ~ 10,000 galaxies) in its various releases.
- the Southern Sky Redshift Survey (SSRS) includes 1657 galaxies.

Other catalogs have been obtained from the IRAS (Infra Red Astronomical Satellite) source catalog, using suitable selection criteria to disentangle galaxies from other possible sources. This types of catalogs contain ~2,000 or so galaxies with redshift

- The (spatial) clustering of extragalactic objects like galaxies (or galaxy systems) can be studied (and it has indeed been studied) by a number of statistical tools such as
  - N-point correlation functions
  - mass functions & number counts
  - percolation statistics and filament detecting techniques
  - multifractal analyses
  - topological statistics (e.g. genus of iso-density contours)
  - void probability functions (or nearest neighbour statistics, ...)

- Most of these statistical tools, when applied to cosmology, rely upon a basic hypothesis, known as the FAIR SAMPLE (or ERGODIC) HYPOTHESIS, which can be stated as follows (Peebles 1980):
« Matter certainly is distributed in a strongly clumped fashion [...], so the homogeneity and isotropy assumption must be supposed to apply in a statistical sense, in the average over large enough regions: the U. is assumed to be a homogeneous and isotropic random process [...]. One imagines that the matter distribution has been determined by

some physical process involving a lengthy and complex sequence of events, so that what is found at any particular place is the result of many slight variations of parameters within the sequence. Samples from well separated spots are uncorrelated, and the collection of such samples is a statistical ensemble generated by many independent applications of the process. Statistics such as the N-point correlation functions can be considered to be averages across the ensemble. The *fair sample hypothesis* states first ~~that~~ it makes sense to think of well separated parts of the U. as being independent realizations of the same physical process, second that within the visible part of the U. there are many independent samples that can be lumped to approximate a statistical ensemble, and third that averages across the ensemble are unaffected by a rotation.»

In these statistical analyses space curvature and expansion are ignored (not for QSOs, for instance) and usual Euclidean geometry is assumed; this comes to the fact that most catalogs probe distances $\lll cH_0^{-1}$. Thus, the fact that galaxies are observed along the past light cone and not at fixed instant of constant time can be ignored.

<u>N-point correlation functions</u>. It is often useful to think of the matter distribution as a distribution of point-like objects, where we are only interested in the positions $\vec{x}_j$ of our $j=1,..N$ objects. In this case n-point correlation functions (spatial) are a useful tool (see Peebles 1980, Ch. III)

- The probability that an object is found in the infinitesimal volume δV is

$$\delta P = n \delta V \qquad (1)$$

where the mean number density n is position-independent. This can be understood as an average over the ensemble: if M realizations are examined, an object is found in δV in $N = M n \delta V$ cases. The probability is proportional to δV because we assume that the chance to find more than one object in a small volume δV is negligible. The mean number of objects found within the finite volume V is the integral of Eq. (1), i.e. $\langle N \rangle = n$

- The <u>two-point correlation function</u> (or "<u>covariance</u>" or "<u>autocorrelation function</u>") is defined by the joint probability of finding an object in both the volumes δV_1 and δV_2 at separation $\vec{r}_{12}$:

$$\delta P = n^2 \delta V_1 \delta V_2 [1 + \xi(\vec{r}_{12})]$$

Homogeneity and isotropy (i.e. "<u>stationarity</u>") of the random process underlying the galaxy distribution implies that $\xi(\vec{r}) \equiv \xi(r)$.

- (translational and rotational invariance); $\xi(r)$ can also be defined through the conditional probability of finding one object in δV given that ~~the~~ object has been randomly selected from the ensemble

$$\delta P = n\delta V[1+\xi(r)]$$

[Remember Bayes theorem: $P(A/B) = P(A,B)/P(B)$]

- On the other hand, ~~for~~ a Poisson process (i.e. a purely random one) one has

$$\delta P = n^2 \delta V_1 \delta V_2 \quad \text{(Poisson)}$$

implying $\xi \equiv 0$; if the objects are correlated $\xi > 0$, if their positions are anticorrelated $-1 \leq \xi < 0$. Then ξ measures the excess (positive or negative) of probability of finding two objects at relative distance r_{12} over a pure Poisson process.
One can also calculate the mean number of neighbours of a randomly chosen object, within some distance R, this results

$$\langle N \rangle_n = nV + 4\pi n J_3(R),$$

where $J_3(R) \equiv \int_0^R \xi(r) r^2 dr$.

- One can continue with this kind of definitions in introducing the three-point correlation function by the joint probability of finding objects in each of the three elements $\delta V_1, \delta V_2, \delta V_3$

$$\delta P = n^3 \delta V_1 \delta V_2 \delta V_3 \left[1 + \xi(r_{12}) + \xi(r_{13}) + \xi(r_{23}) + \right.$$
$$\left. + \zeta(r_{12}, r_{23}, r_{31}) \right]$$

Homogeneity and isotropy imply that the "reduced" (or "connected") 3-point function ζ is a symmetric function of the three lengths.

- Similarly one can define a four-point function through

$$\delta P = n^4 \delta V_1 \delta V_2 \delta V_3 \delta V_4 \left[1 + \right.$$
$$+ \xi(r_{12}) + \cdots \qquad \text{6 terms}$$
$$+ \zeta(r_{12}, r_{23}, r_{31}) + \cdots \qquad \text{4 terms}$$
$$+ \xi(r_{12})\xi(r_{34}) + \cdots \qquad \text{3 terms}$$
$$\left. + \eta(r_{12}, r_{23}, r_{34}, r_{41}, r_{31}, r_{24}) \right]$$

The reduced 4-point function is a function of the 6 variables $r_{12}, r_{23}, r_{34}, r_{41}, r_{31}, r_{24}$ needed to fix the relative positions of the four points (4×3 coordinates minus 3 translations, minus 3 rotations).

- In principle one could go on up to the number N of "points" in the catalogue. Check:
- Let me also mention that $\xi(0)$ is called the "variance", $\zeta(0,0,0)$ the "skewness", $\eta(0,0,0,0,0,0)$ the "kurtosis" of the random process

As we already mentioned most data are presently stored in 2-dim. catalogs. Because of this it is also useful to define similar quantities in projection on the sky. In particular one defines a mean density $\mathcal{N}$ of objects in the sky, so that the probability of finding an object in the element of solid angle $\delta\Omega$ is

$$\delta P = \mathcal{N}\,\delta\Omega .$$

The mean number of objects in the finite cell Ω is $\langle N\rangle = \mathcal{N}\Omega$. The <u>angular two-point function</u> $w(\vartheta)$ is therefore defined by the joint probability of finding objects in both of the elements of solid angle $\delta\Omega_1$ and $\delta\Omega_2$, placed at angular separation ϑ_{12}:

$$\delta P = \mathcal{N}^2 \delta\Omega_1 \delta\Omega_2 [1 + w(\vartheta_{12})].$$

Of course, all the previous discussion on spatial correlations could be repeated here.

⇒ <u>Scaling relation</u>. A very important property of the angular two-point f. is known as scaling relation. Let D be the typical depth of a 2-dimensional catalog, then w varies with D according to the law

$$w_D(\vartheta) = D^{-1} W(D\vartheta) \qquad (2)$$

This relation, which could be easily demonstrated, allows to relate $w(\vartheta)$ of different surveys, usually characterized by different depths D, and to check for the universality of the correlation properties. Since the previous scaling law is based on the assumption of hom. and isotropy, a (large) deviation from this law would indicate a failure of this hypothesis or a truly fractal behaviour in the gal. distribution. So far, people have found very good agreement of data with this law![10]

- The Limber equation. Based on the assumption of spatial homogeneity and isotropy, as well as on the following two approximations
  - Small separation approximation ($r_{12} \ll r_1 \sim D$), which is a very good one
  - Decoupling of luminosity ("magnitude") and position, i.e. the assumption that the luminosity of a galaxy is statistically independent of its position relative to other galaxies: this is not necessarily true,

one is able to relate $w(\theta)$ to $\xi(r)$ and viceversa. The Limber equation reads:

$$W(x) = \frac{\int_0^\infty y^4 \phi^2(y)\,dy \int_{-\infty}^{\infty} du\, \xi\left[(u^2+(xy)^2)^{1/2}\right]}{\left[\int_0^\infty y^2 \phi(y)\,dy\right]^2}, \quad (3)$$

where $W(x)$ is the function defined in Eq.(2) and $\phi(y)$ (related to the "luminosity function") gives the probability that a galaxy at distance and similarly for the higher-order (spatial) correlation functions. This continuous representation will turn out to be very useful when we shall try to relate the observed clustering properties of the U. to processes occurring in the early U. history (like e.g. "inflation").

⇒ Observational data. There have been so far many independent determinations of the galaxy-galaxy correlation function ξ_{gg}, both from angular samples and from redshift complete ones. All of them agree on the fact that, at least on some range of separations r, one has the following power-law:

$$\xi_{gg}(r) = \left(\frac{r}{r_{0g}}\right)^{-\gamma}$$

with $\gamma \simeq 1.8$ and $r_{0g} \simeq 5\,h^{-1}\,Mpc$, for (according to the Davis & Peebles (1983) determination) $r \lesssim 10\,h^{-1}\,Mpc$.
- Analysis of the cluster two-point function have generally shown that clusters too are correlated but more strongly than galaxies:

$$\xi_{cc}(r) = \left(\frac{r}{r_{0c}}\right)^{-\gamma},$$

again with $\gamma \simeq 1.8$, but with $r_{0c} = 25\,h^{-1}\,Mpc$ (according to Bahcall & Soneira 1983 and Klypin & Kopilov 1983); other authors claim a smaller value of r_{0c} ($\simeq 16\,h^{-1}\,Mpc$). This time the power-law extends to larger scales: $r \lesssim 80\,h^{-1}\,Mpc$.

There have been also determinations of ξ for other kinds of objects, like galaxy groups and superclusters, or peculiar objects such as QSOs. The least we can say is that all of them appear to be correlated over some range of scales. The better determination of the galaxy 2-point function at present is that of the APM group (Maddox et al. 1990) who got $w(\vartheta)$ from their catalog, and found a relevant break in its shape at $\vartheta \sim 2.5°$. This determination has been of great interest since it poses a big challenge to current theories of galaxy formation (namely the Cold Dark Matter (CDM) scenario).

Finally let us mention that a number of authors have also measured higher order correlations both for galaxies and for clusters. The interesting result for galaxies is that a so-called "hierarchical form" holds:

$$\zeta(r_{12}, r_{23}, r_{31}) \approx Q\left[\xi(r_{12})\xi(r_{23}) + \xi(r_{13})\xi(r_{23}) + \xi(r_{12})\xi(r_{13})\right],$$

with $Q \simeq 1$. Similar forms have also been found for the 4-point and 5-point functions.

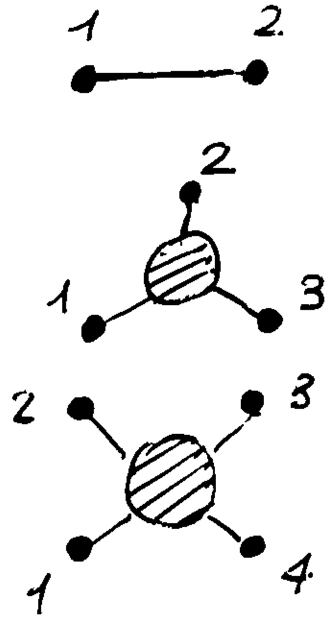
N-point correlation functions:

$\xi(\vec{x}_1, \vec{x}_2)$

$\xi_c^{(3)}(\vec{x}_1, \vec{x}_2, \vec{x}_3)$

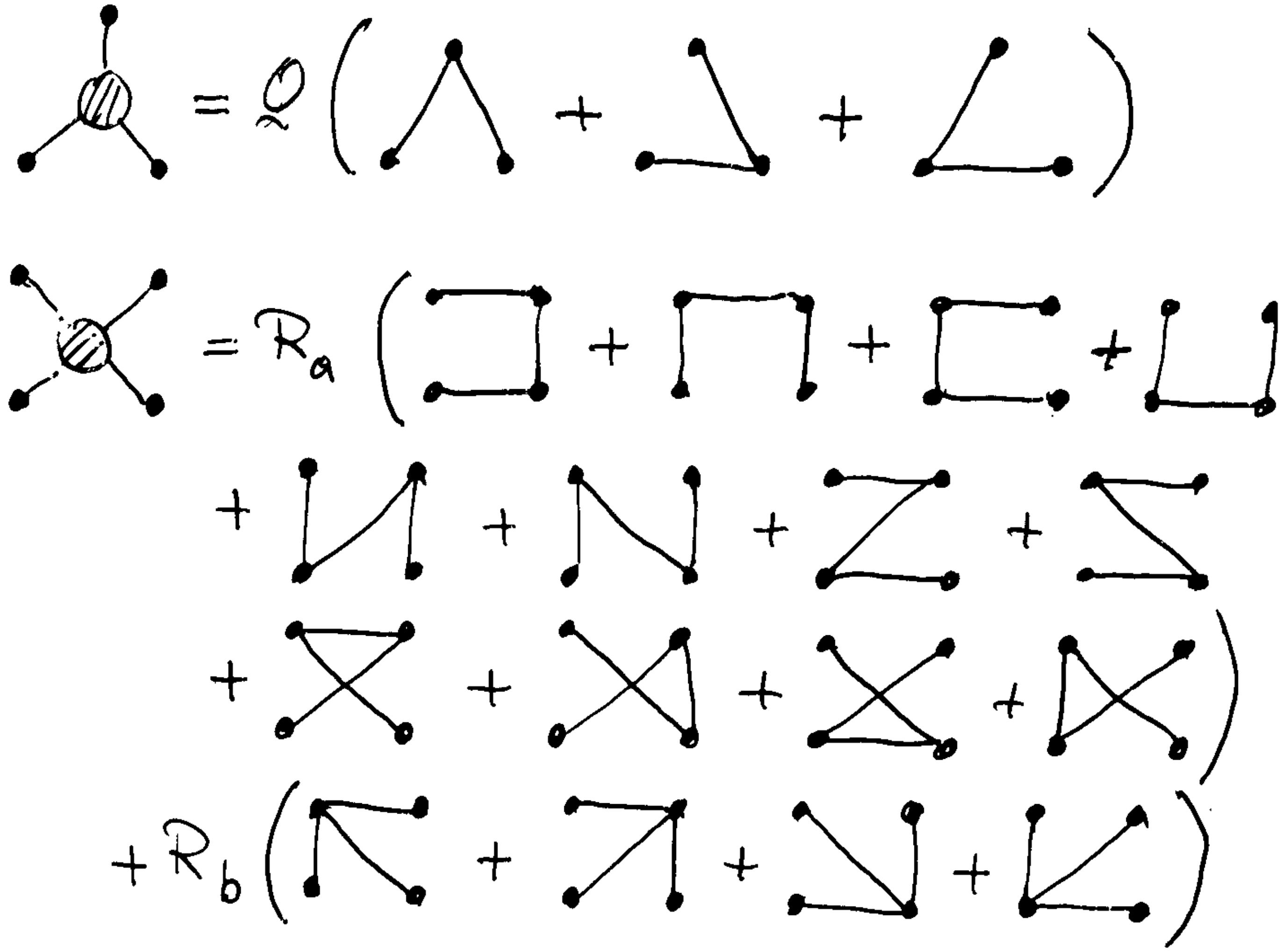
$\xi_c^{(4)}(\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4)$

. . . .



a connected (reduced) correlation function of order N is that contribution to ∂P_N that vanishes when <u>any</u> of the N points is at infinity.

Ex 1: Hierarchical statistics



⇒ Structures on very large scales. Very large structures seem to exist in the U, these are for instance the superclusters, loosely-bound, non-virialized objects with densities about twice the mean density of the U., containing many rich clusters. Nearby superclusters include our own Local Supercluster (centred on the Virgo cluster), Hydra-Centaurus, and Pisces-Cetus.
- Several galaxy surveys have identified large voids in the distribution of bright galaxies. For example the Kirshner et al. (1981) survey showed the existence of a void in Boötes of diameter about $50 h^{-1}$ Mpc, and the CfA slices of the Universe by de Lapparent, Geller & Huchra (1986) seem to indicate that voids of size ~ $20 h^{-1}$ Mpc are quite common, according to these authors, the galaxy distribution appears "bubbly", with galaxies concentrated on sheet-like structures surrounding nearly empty voids. the largest of such sheet-like structures, the "Great Wall" observed by Geller & Huchra (1989) has sides ~ $100 h^{-1}$ Mpc or so.
Quite long "filaments" of galaxies seem to exist. A very pronounced one is located in the Perseus-Pisces region
- Altogether these observations show that the galaxy (matter ?) distribution is clustered well above the distance implied by the correlation length for galaxies ($\lesssim 5 h^{-1}$ Mpc).