RevMexAA (Serie de Conferencias), **14**, 63–67 (2002)

X Reunión Regional Latinoamericana de Astronomía Unión Astronómica Internacional (Córdoba, Argentina, septiembre 17 -21 2001)
Editores: Juan José Clariá, Diego García Lambas, y Hugo Levato

HIERARCHICAL CLUSTERING

L. Infante

Departamento de Astronomía y Astrofísica, Pontificia Universidad Católica de Chile,
Vic. Mackenna 4860, Casilla 306, Santiago 22, Chile

RESUMEN

En esta contribución presento resultados recientes sobre las propiedades de acumulación de galaxias, grupos, cúmulos y supercúmulos de bajo redshift (z$\leq$1). Presento, a su vez, lo esperado y lo medido con respecto al grado de evolución de la acumulación de galaxias. Hemos usado el catálogo fotométrico de galaxias extraído de las primeras imágenes del "Sloan Digital Sky Survey", para estudiar las propiedades de acumulación de pequeñas estructuras de galaxias, pares, tríos, cuartetos, quintetos, etc. Un análisis de la función de correlación de dos puntos, en un área de 250 grados cuadrados del cielo, muestra que estos objetos, al parecer, están mucho más acumulados que galaxias individuales.

ABSTRACT

In this contribution I present current results on how galaxies, groups, clusters and superclusters cluster at low (z$\leq$1) redshifts. I also discuss the measured and expected clustering evolution. In a program to study the clustering properties of small galaxy structures we have identified close pairs, triplets, quadruplets, quintuplets, etc. of galaxies in the Sloan Digital Sky Survey commissioning imaging data. An analysis of the 2-point angular correlation function on an area of more than 250 square deg^2 show that these objects appear to be appreciably more strongly clustered than single galaxies.

Key Words: **GALAXIES – CLUSTERS**

1. INTRODUCTION

For more than two decades the clustering properties of galaxies and of clusters of galaxies have been the subject of numerous studies. It seems that clusters are more strongly clustered than galaxies; the correlation length increases with richness (Bahcall & Soneira 1983; Bahcall 1988; Postman et al. 1992; Croft et al. 1997; Abadi et al. 1998). A central issue has been the existence of a common physical mechanism to explain clustering properties from galaxies groups to rich clusters. The richness dependence of the correlation function is generally explained in terms of high-density peak biasing of the galaxy systems (Kaiser 1984), and is seen in cosmological simulations (e.g. Bahcall & Cen 1992, Colberg et al. 2000 and references therein).

In this paper I describe the main statistical tools used in characterizing clustering and define co-moving correlation length, effective volumes and mean system separations. Then, a description of current results of clustering of galaxies at low redshift (Sloan Digital Sky Survey) and its evolution (CNOC2 redshift survey, Shepherd et al. 2001) are discussed. Finally, I present the clustering properties of small groups and richer clusters and discuss our recent results on the correlation properties of pairs and triplets of galaxies from the SDSS commissioning survey (Infante et al. 2002).

2. CHARACTERIZING CLUSTERING

Great efforts have been made to develop statistical tools to describe and analyze the manner in which galaxies and groups of galaxies cluster. The most common statistics are the N point spatial correlation functions $\xi(r)$. (For details refer to Limber 1953, 1954; Groth & Peebles 1977; &; Peebles 1980). Unfortunately, these functions are hard to obtain - both in terms of telescope time and data reduction - if one wants to reach the faint magnitudes needed to study large volumes and the evolution of clustering. An alternative, which requires a much smaller investment of telescope time, is to use angular correlation functions, (*e.g.* Peebles 1980, pp. 180 - 195). Given

X Reunión Regional Latinoamericana de Astronomía Unión Astronómica Internacional (Córdoba, Argentina, septiembre 17 -21 2001)
Editores: Juan José Clariá, Diego García Lambas, y Hugo Levato

a knowledge of the redshift distribution of the sample and making a few simple assumptions one can invert the angular function to estimate the spatial correlation function. In spite of the fact that many details are lost due to the integral nature of the angular correlation function, it is clearly the preferred tool for studying the evolution of the covariance function to large look-back times, and for measuring spatial correlation functions over large volumes of space.

2.1. *Definitions and Estimators*

Let's consider a random distribution of points. If n is the number density of points, the probability of finding one point in volume dV is,

$$dP = ndV.$$

The joint probability of finding one point in dV_1 and a second one in dV_2 separated by a distance r is,

$$dP_r = n^2 dV_1 dV_2.$$

And, in general, the joint probability of finding one point in dV_1, a second one in dV_2, a third one in dV_3, etc. is,

$$dP_r = n^N \; dV_1 dV_2 ... dV_N.$$

Now, if the distribution of points in space is clustered, the joint probability of finding one point in dV_1 and a second one in dV_2 separated by a distance r is,

$$dP_r = n^2(1+\xi(r))dV_1 dV_2,$$

where, $\xi(r)$ is the spatial covariance function.

On the other hand, if the distribution is a continuous function $f(x)$, then

$$< f(x_1)f(x_2) > = < f >^2 \; (1+\xi(x_{12})),$$

where the fourier transform of the covariance function is,

$$\xi(r) = \int \vec{k} \; e^{i\vec{k}\cdot\vec{r}} P(\vec{k}) d^3\vec{k},$$

where P is the power spectrum. Since P depends only on k,

$$\xi(r) = 4\pi \int_0^\infty k^2 P(k) \frac{sen\ (kr)}{kr} \; dk \text{ and } r \equiv |\vec{r}|.$$

In Practice, the two-point angular correlation function is estimated as,

$$dP(\theta) = n^2[1+\omega(\theta)]d\Omega_1 d\Omega_2,$$

where Ω_1 and Ω_2 are two solid angles separated by an angular distance θ.

Given a catalogue of positions, we can now estimate $\omega(\theta)$. The most common estimators are:

A- Infante et al. (1994)

$$\omega(\theta) = \frac{N_{dd} \; N_r}{B \; N_{dr}(N_{d-1})} - \frac{N_{rr} \; N_{r_1}}{N_{rr_1}(N_r - 1)},$$

where N_{dd} and N_{dr} are the numbers of data-data and data-random pairs respectively, N_d and N_r are the numbers of data and random points respectively; r_1 refers to a different set of random points. In this estimator the following corrections are considered:

-Edge effects $\quad \omega_{rd} = \frac{N_{rd}}{N_{rr}} - 1$

-Integral constraint $(B) \quad \int w(\theta) d\Omega_1 \; d\Omega_2 = 0$

-Unclustered objects. A constant to be determined after estimating the number of stars miss-classified as galaxies.

B- Landy & Szalay (1993)

$$w(\theta) = \frac{N_{dd} - 2N_{dr} + N_{rr}}{N_{rr}}.$$

2.2. *The co-moving Correlation Length*

We relate our $\omega(\theta)$ measurements to the spatial correlation functions through inversion. The spatial correlation function is assumed to be a power law weighted by the standard phenomenological evolutionary factor, $\xi(r,z) = \left(\frac{r}{r_0}\right)^{-\gamma} (1+z)^{-(3+\epsilon)}$, where r is the proper distance, r_0 is the proper correlation length, and ϵ is the clustering evolution index. $\epsilon = \gamma - 3$ corresponds to clustering fixed in comoving coordinates, while $\epsilon = 0$ represents stable clustering in physical coordinates (Phillips et al. 1978).

To obtain the correlation length r_0 we invert Limber's equation (Limber 1953). If $\omega(\theta) = A_\omega \theta^{(1-\gamma)}$, then r_0 is given by

$$r_0^\gamma = A_\omega^{-1} \; C \; \frac{\int_0^\infty g(z)(dN/dz)^2 dz}{[\int_0^\infty (dN/dz)dz]^2},$$

$$C = \pi^{1/2} \; \frac{\Gamma[(\gamma-1)/2]}{\Gamma(\gamma/2)},$$

$$g(z) = \left(\frac{dz}{dx}\right) x^{1-\gamma} F(x)(1+z)^{-(3+\epsilon-\gamma)};$$

where $x(z)$ is the coordinate distance and

$$F(x) = [1-(H_0 a_0 x/c)^2(\Omega_0 - 1)]^{1/2}.$$

The H_0 dependence of $g(z)$ is cancelled by that in the measured value of r_0. The strong dependence of A_ω on dN/dz is clear in this equation. Note in particular that in the above equations *there is no dependence on galaxy evolution, except in the calculation of the redshift distribution dN/dz.* (Peebles 1980, eqs. [56.7] and [56.13]). However Limber's equation, as presented here, does assume that the clustering is independent of luminosity; the high-redshift objects in our narrow magnitude slice are of higher luminosity than the low-redshift objects.

2.3. *Effective Volumes and Mean Separations*

The effective volume of the sample is,

$$V = \frac{\int_{z_{min}}^{z_{max}} (dN/dz)V(z)dz}{\int_{z_{min}}^{z_{max}} (dN/dz)dz},$$

which then yields the number density, $n = \frac{N_{systems}}{V}$, and mean separation, $d = \left(\frac{1}{n}\right)^{1/3}$.

3. CLUSTERING OF GALAXIES

Traditionally, the derivation of $\omega(\theta)$ or $\xi(r)$ has been approached in two different ways: (a) by means of catalogues of galaxies - sometimes with redshift information - which cover vast areas and (b) from deep

X Reunión Regional Latinoamericana de Astronomía Unión Astronómica Internacional (Córdoba, Argentina, septiembre 17 -21 2001)
Editores: Juan José Clariá, Diego García Lambas, y Hugo Levato

samples of galaxies, obtained with large telescopes which have a small field size on the sky. These two methods sample the clustering properties of the Universe to very different depths.

Given the limited space, a complete review of the subject is not possible. Nevertheless, I will focus on the most recent results from the Sloan Digital Sky Survey (SDSS) and from the Canadian Network for Observational Cosmology (CNOC2). The latter describe the clustering properties of galaxies in the nearby Universe, while the former is a description of how the galaxy clustering pattern evolve with redshift.

3.1. *Clustering of Galaxies at Low Redshifts*

The Sloan Digital Sky Survey is a photometric and spectroscopic survey of about 1/4 of the sky, above Galactic latitude of $\sim 30°$ (York et al. 2000). The photometric data are taken with a dedicated 2.5 m altitude-azimuth telescope at Apache Point, New Mexico, with a 2.5° wide distortion-free field and an imaging camera consisting of a mosaic of 30 imaging 2048×2048 SITe CCDs with $0.4''$ pixels (Gunn et al. 1998). The CCDs are arranged in six columns of five CCDs each, using five broad-band filters (u', g', r', i' and z'). The total integration time per filter is 54.1 seconds. Each column of CCD's observes a *scanline* on the sky roughly $13'$ wide; the six scanlines of a given observation make up a *strip*. The measured survey depth is 22.0, 22.2, 22.2, 21.3, and 20.5 magnitudes for the 5 filters, respectively. The SDSS photometric system is measured in the AB_ν system (Oke & Gunn 1983; Fukugita et al. 1996).

Current analysis use imaging data taken during the commissioning period of SDSS (on 21 and 22 March 1999), which together make up a strip 2.5° wide and 100° long centered on the Celestial Equator (runs 752 and 756); these data are included in the SDSS Early Data Release (Adelman et al. 2001). These runs extend over the area $7^h.7 <$ R.A. (2000) $< 16^h.8$ and $-1°.26 <$ Dec. (2000) $<$ 1°.26, although the two strips overlap for only the central part of this right ascension range. The seeing ranged from 1.2" to 2.5".

The main result is that $\omega(\theta)$ is a power-law with an exponent of ≈ -0.7 over 2 orders of magnitude, from scales between 1 arcmin to 0.5 degree. There appears to be a break in the power-law at scales 1 - 2 degrees. At $\theta < 1'$, the fit is not consistent with a power-law.

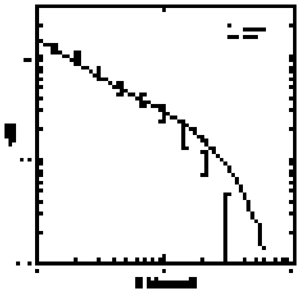

Fig. 1. The SDSS early galaxy data angular 2-point correlation function (filled circles), within the magnitude interval $18 < r^* < 19$. (This plot is Fig. 1 in Connolly et al. 2002). It is compared with the correlation function from the APM (Maddox et al. 1990) measured over the magnitude interval $18 < B_j < 20$ (solid line). The SDSS correlation function has been scaled to the depth of the APM data using Limber's equation.

3.2. *Evolution of Clustering*

To date, one of the best measures of the evolution of galaxy clustering is from the CNOC2 Redshift Survey (Shepherd et al. 2001). They determine the clustering evolution up to $z \sim 0.6$ for late and early type galaxies. Their survey covers a field of about 1.55 deg^2 and have redshifts for ~ 3000 galaxies with $R_c < 21.5$ ($M_R < -20$). They measure the projected correlation function ω_p in a volume limited sample, over the comoving projected separation range $0.04h^{-1}Mpc < r_p < 10h^{-1}Mpc$. Galaxies were classified as being early or late according to their Spectral Energy distributions, as determined from broad band $UBVR_CI_C$ photometry.

Early type galaxies were found to be more strongly clustered, with a larger power-law index ($r_0 = 5.45 \pm 0.28h^{-1}Mpc$ and $\gamma = 1.91 \pm 0.06$), than late type galaxies ($r_0 = 3.95 \pm 0.12h^{-1}Mpc$ and $\gamma = 1.59 \pm 0.08$). In terms of evolution, Shepherd et al. find that both type of objects have clustering amplitudes, in comoving coordinates, that decrease with time, $\epsilon = -3.9 \pm 1.0$ and $\epsilon = -7.7 \pm 1.3$ for early and late type galaxies respectively. However, the authors warn that the strong increase of clustering with redshift might be apparent; it might be caused by evolution of the galaxies themselves.

X Reunión Regional Latinoamericana de Astronomía Unión Astronómica Internacional (Córdoba, Argentina, septiembre 17 -21 2001)
Editores: Juan José Clariá, Diego García Lambas, y Hugo Levato

4. CLUSTERING; FROM SMALL GROUPS TO RICH CLUSTERS

We use a sample of 330,041 galaxies located on 278 square degrees of the Celestial Equator, with magnitude $18 \leq r^* \leq 20$, obtained from SDSS commissioning imaging data. We use these data to select isolated pairs of galaxies. We determine the angular correlation function of the galaxies and of the galaxy pairs. We find the following results: (1) Pairs of galaxies are more strongly clustered than single galaxies. The angular correlation amplitude of galaxy pairs is 2.9 ± 0.4 times larger than the amplitude of galaxies. (2) The slopes of both correlation functions are the same: 1.77 ± 0.04 (0.77 in 2D). (3) We measure $\omega(\theta)$ to just under 1 deg scales, corresponding to $\sim 9\ h^{-1}$ Mpc at the mean redshift of 0.22. No breaks are detected in the correlation functions. (4) Assuming a redshift distribution from the CNOC2 survey, we invert the angular correlations and determine a spatial correlation length, r_0. We find that pairs have a significantly larger correlation length than galaxies; $r_0 = 4.2 \pm 0.4\ h^{-1}$ Mpc for galaxies and $7.8 \pm 0.7 h^{-1}$ Mpc for pairs. (5) The mean separation between systems ($d = n^{-1/3}$) are $d = 3.7$ and $10.24\ h^{-1}\ Mpc$ for galaxies and pairs respectively. The results fall right on the global $r_0 - d$ relation observed for galaxy systems (Bahcall & West 1994).

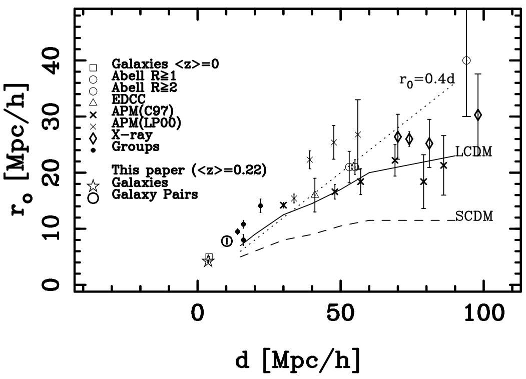

Fig. 2. Correlation length r_0 versus mean separation $d(= n^{-1/3})$ for galaxies and pairs, as well as groups and clusters of galaxies. Two cosmological models, LCDM and SCDM, are shown for comparison, as well as the original approximate relation of $r_0 \approx 0.4\ d$ for rich systems (Bahcall 1988).

5. CONCLUSIONS

Current results of the clustering properties of galaxy systems, ranging from galaxies to rich clusters, are examined. Techniques for estimating the two-point angular correlation function are discussed. We relate $\omega(\theta)$to the co-moving correlation length and estimate effective volumes by assuming a redshift distribution.

The first $\omega(\theta)$ results from the Sloan Digital Sky Survey are presented. The evolution of clustering to $z \sim 0.6$ is discussed. I take the CNOC2 survey as the most up to date example.

The work on galaxy pairs and triplets suggests a number of follow-up studies. We have defined pairs of galaxies from their position in projection, and we need to quantify what fraction of these objects are physical pairs. We are in the process of carrying out a redshift survey of a subset of the pairs sample, which will also be useful in tying down the dN/dz relation. We can also use photometric redshifts from the multi-band photometry of the SDSS to define a cleaner sample of galaxy pairs consistent with being at the same redshift. Finally, we have used just under 300 square degrees of SDSS data; the survey has now imaged over five times this much sky. Thus we are in the process of defining samples of richer (and thus rarer) systems from this larger sample, and measuring their correlations, to fill in the $r_0 - d$ relation between galaxies and groups.

LI acknowledges support from *Proyecto Puente PUC* and *Proyecto FONDAP "Centre for Astrophysical Research"*.

REFERENCES

Abadi, M.G., Lambas, D.G., & Muriel, H. 1998, ApJ, 507, 526
Adelman et al. 2001, submitted
Bahcall, N.A. 1988, ARA&A, 26, 631
Bahcall, N.A. & Cen, R. 1992, ApJ, 309, L81
Bahcall, N.A. & Soneira, R.M. 1983, ApJ, 270, 20
Colberg, J.M. et al. 2000, MNRAS, 319, 209
Connolly, A.J. et al. 2001, AJ, submitted (astro-ph/0107417)
Croft, R.A.C, Dalton, G.B., Efstathiou, G., Sutherland, W.J., & Maddox, S.J. 1997, MNRAS, 291, 305
Fukugita, M., Ichikawa, T., Gunn, J.E., Doi, M., Shimasaku, K., & Schneider, D.P. 1996, AJ, 111, 1748
Gunn, J.E., Carr, M.A., Rockosi, C., et al. 1998, AJ, 116, 3040
Groth, E.J. & Peebles, P.J.E, 1977, ApJ, 217, 385
Infante, L., Strauss, M.A., Bahcall, N.A., et al. 2002, ApJ, in press
Kaiser, N. 1984, ApJ, 284, L9
Landy, S.D. & Szalay, A.S. 1993, ApJ, 412, 64
Limber, D.N. 1953, ApJ, 117, 134
Limber, D.N. 1954, ApJ, 119, 655
Maddox, S.J., Sutherland, W.J., Efstathiou, G., Loveday, J., & Peterson, B.A. 1990, MNRAS, 247, 1P

Oke, J.B. & Gunn, J.E. 1983, ApJ, 266, 713
Peebles, P.J.E. 1980, "The Large Scale Structure in The Universe" (Princeton: Princeton University Press)
Phillips, S., Fong, R., Fall, R.S., Ellis, S.M., & MacGillivray, H.T. 1978, MNRAS, 182, 673
Postman, M., Huchra, J.P., & Geller, M.J. 1992, ApJ, 384, 404
Shepherd, C.W., Carlberg, R.G., Yee, H.K., Morris, S.L., Lin, H., Sawicki, M., Hall, P.B. & Patton, D.R., 2001, astro-ph/0106250
York, D.G. et al. 2000, AJ, 120, 1588

X Reunión Regional Latinoamericana de Astronomía Unión Astronómica Internacional (Córdoba, Argentina, septiembre 17 -21 2001)
Editores: Juan José Clariá, Diego García Lambas, y Hugo Levato