The Astrophysical Journal, 205:L121–L125, 1976 May 1

ON THE USE OF CORRELATION FUNCTIONS IN FINDING PHYSICAL ASSOCIATIONS OF GALAXIES

S. Michael Fall,* Margaret J. Geller,† Bernard J. T. Jones,* and Simon D. M. White*

Received 1976 January 14

ABSTRACT

We show that the flatness of the "singles" two-point correlation function found by Turner and Gott is an artifact of the selection criteria used to define a "single" galaxy. Thus their method cannot be used to distinguish a true field from a cluster population of galaxies.

Subject heading: galaxies: clusters of

I. INTRODUCTION

The description of galaxy clustering in terms of correlation functions has been discussed extensively by Peebles and his co-workers in a recent series of papers (see Peebles and Groth 1975, and references therein). From catalogs of galaxies they have determined the two- and three-point correlation functions which give the "excess" probabilities (above those expected for a Poisson distribution) for finding two- and three-galaxy configurations. Totsuji and Kihara (1969) and Peebles (1974) find that over a wide range of separations the two-point correlation function can best be fitted by a power law of the form

$$w(x) \approx Ax^{-\delta}, \quad \delta \approx 0.8, \tag{1}$$

where x is the angular separation of the two galaxies and A is a constant which depends on the limiting magnitude of the catalog. Peebles's interpretation of equation (1) is that clustering exists on all scales from $50h^{-1}$ kpc to $5h^{-1}$ Mpc with no preferred scale (h is Hubble's constant in units of 100 km s^{-1} Mpc^{-1}). Although $w(x)$ completely specifies the pairwise clustering of galaxies, a complete description of groups of galaxies requires the three-point and higher order correlation functions.

Several lists of groups of galaxies already exist (Burbidge and Burbidge 1961; Holmberg 1964; Karachentsev 1966; de Vaucouleurs 1976). The selection of group members is generally based on criteria like apparent proximity of galaxies on the sky, galaxy types, and similarity of magnitudes, diameters, and velocities. Turner and Gott (1975) suggest a scheme for picking out group members on the basis of purely objective statistical criteria. Using the Zwicky catalog to a limiting magnitude, 14.0 mag, they put each galaxy into one of two classes depending on the presence or absence of a neighboring galaxy within a chosen fiducial angular scale α. Galaxies which have one or more neighbors within α are called "associated"; those without a neighbor within α are called "singles." For the particular selection angle $\alpha = 45'$, Turner and Gott find that the correlation function w_s of the "singles" with all other galaxies in the sample is very nearly zero on all scales greater than α. Because Turner and Gott claim they have no *a priori* (mathematical) reason to expect such a result, they conclude that their class of "single" galaxies constitutes a true field population and that the "associated" galaxies may be identified as group members.

Our purpose is to show that (i) w_s can be uniquely expressed in terms of correlation functions for the overall sample and (ii) the flatness of w_s is almost certainly a necessary mathematical result of the selection criteria.

Our ignorance of the higher-order correlations prevents us from establishing the result (ii) above rigorously. However, we shall present two arguments to show that this result probably holds for the actual distribution of galaxies.

The results of Turner and Gott are not, therefore, inconsistent with the results of Peebles and co-workers who discuss only a single population of galaxies. The apparent lack of preferred clustering scales demonstrated by the overall two- and three-point correlation functions (Peebles and Groth 1975) suggests that statistical methods based on angular separation alone cannot be used to assign a particular galaxy to a physically distinct population.

II. ANALYSIS

a) Definitions and Basic Equations

In order to derive a relationship between $w_s(\boldsymbol{x})$ (the "singles" correlation function) and $w(\boldsymbol{x})$ (eq. [1]), we express these functions in terms of probability statements. We then use the calculus of probabilities to derive relationships among these statements (see, for example, Feller 1968). We refer to points on the sky by two-vectors relative to some origin $\boldsymbol{O}$. The quantity $x = |\boldsymbol{x}|$ is the angular separation of the points $\boldsymbol{O}$ and $\boldsymbol{x}$ measured along the great circle

* Institute of Astronomy, Madingley Road, Cambridge, England.

† Center for Astrophysics, Harvard College Observatory and Smithsonian Astrophysical Observatory, 60 Garden Street, Cambridge, Massachusetts.

passing through them. We shall restrict our attention to small enough areas of sky that all vectors may be regarded as Cartesian.

We define the propositions:

O: a galaxy at $\boldsymbol{O}$

O_s: a "single" galaxy at $\boldsymbol{O}$

$X, Y_1 \ldots Y_N$: a galaxy in $d^2\boldsymbol{x}$ at $\boldsymbol{x}$, $d^2\boldsymbol{y}_1$ at $\boldsymbol{y}_1 \ldots$ and $d^2\boldsymbol{y}_N$ at $\boldsymbol{y}_N$

Φ_α: no galaxies in the disk of radius α centered at $\boldsymbol{O}$.

In terms of the above propositions, we define the overall and "singles" two-point correlation functions:[1]

$$n[1 + w(\boldsymbol{x})]d^2\boldsymbol{x} = P\{X|O\}\,, \tag{2a}$$

$$n[1 + w_s(\boldsymbol{x})]d^2\boldsymbol{x} = P\{X|O_s\}\,, \tag{2b}$$

where n is the mean surface density of galaxies. We assume that the sample is statistically homogeneous; i.e., all correlation functions are assumed to be independent of the chosen origin $\boldsymbol{O}$. We shall also assume throughout that the magnitudes of all correlation functions decrease monotonically at large separations. From equations (2a) and (2b) we have the basic relation

$$\frac{1 + w_s(\boldsymbol{x})}{1 + w(\boldsymbol{x})} = \frac{P\{\Phi_\alpha|OX\}}{P\{\Phi_\alpha|O\}}\,, \tag{3}$$

since $O_s = O\Phi_\alpha$. In general, relation (3) involves all correlation functions including the unknown four-point and higher order functions.

b) *First Argument*

For $x \gg \alpha$ we may approximate equation (3) using the following scheme. First we note that $P\{\Phi_\alpha|OX\} \to P\{\Phi_\alpha|O\}$ as $x \to \infty$ because, in the limit of large separation, the presence of a galaxy at $\boldsymbol{x}$ cannot affect the probability of a configuration of galaxies occurring near $\boldsymbol{O}$. Second, we make the reasonable assumptions (a) that $P\{\Phi_\alpha|OX\}$ depends on $\boldsymbol{x}$ only through $w(\boldsymbol{x})$ for $x \gg \alpha$, and (b) that $P\{\Phi_\alpha|OX\}$ is analytic in $w(\boldsymbol{x})$. We may then write

$$\frac{P\{\Phi_\alpha|OX\}}{P\{\Phi_\alpha|O\}} \approx 1 + f(\alpha)w(\boldsymbol{x}) \text{ for } x \gg \alpha\,, \quad w(\boldsymbol{x}) \ll 1\,; \tag{4a}$$

$$f(\alpha) = \frac{\partial}{\partial w(\boldsymbol{x})} \ln P\{\Phi_\alpha|OX\}\,|_{w(\boldsymbol{x})=0}\,. \tag{4b}$$

(A justification for the approximations in eqs. [4] is discussed more fully below.) We note that f will in general be negative for some range of α because, for a positively correlated sample, the presence of a galaxy at $\boldsymbol{x}$ decreases the probability that the region around $\boldsymbol{O}$ is empty. Finally from equations (3) and (4) we have

$$w_s(\boldsymbol{x}) \approx [1 + f(\alpha)]w(\boldsymbol{x}) \qquad (x \gg \alpha) \tag{5}$$

for $w(\boldsymbol{x}) \ll 1$. Thus, $w_s(\boldsymbol{x})$ at large angular scales depends on the selection angle α. Regardless of the particular form of $w(\boldsymbol{x})$, the "singles" correlation functions is asymptotically flat [in the sense that $|w_s(\boldsymbol{x})/w(\boldsymbol{x})| \to 0$ as $|\boldsymbol{x}| \to \infty$] for any value of α such that $1 + f(\alpha) = 0$.

In order to justify the approximations in equations (4), we must be more specific about the dependence of $P\{\Phi_\alpha|O\}$ on higher order correlation functions. We make the ansatz:

$$P\{X|Y_1 \ldots Y_N\} \approx n[1 + Nw(\boldsymbol{x})]d^2\boldsymbol{x} \text{ for } y_1 \ldots y_N \ll x,\, w(\boldsymbol{x}) \ll 1\,. \tag{6}$$

Equation (6) is consistent with the asymptotic behavior (under the specified conditions) of the two- and three-point correlation functions found by Peebles and Groth (1975) and is the most natural generalization of their results given that the objects under consideration form a self-gravitating system. Our ansatz (6) is mathematically equivalent to the statement

$$P\{Y_1 \ldots Y_N|X; n\} \approx P\{Y_1 \ldots Y_N\,; n + \delta n\} \text{ for } y_1 \ldots y_N \ll x,\, w(\boldsymbol{x}) \ll 1\,, \tag{7}$$

where $\delta n = nw(\boldsymbol{x})$, and the functional dependence of probability statements on the density, n, is denoted explicitly by the expressions following the semicolons in equation (7). After some tedious manipulation it follows from equation (7) that

$$P\{\Phi_\alpha|OX; n\} \approx P\{\Phi_\alpha|O; n + \delta n\} \text{ for } x \gg \alpha,\, w(\boldsymbol{x}) \ll 1\,. \tag{8}$$

[1] The notation "$P\{X|O\}$" means "the probability that X occurs given that O occurs." The notation OX refers to the intersection (i.e., the conjunction) of propositions O and X; i.e., $O \cap X$.

Substituting equation (8) in equation (3) gives an expression of the form (4) with

$$f(\alpha) = \frac{\partial}{\partial \ln n} \ln P\{\Phi_\alpha | O; n\} . \tag{9}$$

In Figure 1 we plot the function $f(\alpha) + 1$ as given by equation (9) for

$$P\{\Phi_\alpha | O\} = \exp\left(-n \int_{\alpha > x} [1 + w(x)] d^2 x\right) . \tag{10}$$

Here $w(x)$ is given by equation (1). We use the density $n \approx 0.2$ per square degree and $A \approx 0.8$ appropriate for Turner and Gott's sample.

Equation (10) approaches the actual distribution when $w(\alpha) n \pi \alpha^2 \ll 1$. Although this inequality is only weakly satisfied for the parameters of interest, in view of the results shown in Figure 1, we expect that the zero of $f(\alpha) + 1$ required to make $w_s(x)$ flat (cf. eq. [5]) will be approximately the selection angle $\alpha = 45'$ used by Turner and Gott.

c) Second Argument

The first argument, although suggestive, might be criticized on the basis of our ansatz for the unknown higher order correlation functions. Thus, it is of some interest to explore the dependence of our conclusion (the flatness of the "singles" correlation function is a consequence of Turner and Gott's particular selection angle α) on the adopted form of these higher correlations. We now do this by expressing w_s exactly in terms of the known lower order correlation functions while simply ignoring the higher order correlations as follows.

First, we express $w_s(x)$ as an infinite series involving correlations of all orders. We divide the disk of radius α into a large number, N, of annuli with radii x_i ($<\alpha$). Then in the previous notation

$$P\{\Phi_\alpha | O\} = P\{\bar{X}_1, \bar{X}_2, \ldots \bar{X}_N | O\} , \tag{11}$$

where $\bar{X}_i$ is the negation of the proposition X_i that a galaxy be in the ith annulus. By considering various combinations of the propositions on the right-hand side of equation (11) in the usual way, (e.g., considering the propositions as sets in a Venn diagram and using the analogy between probabilities and areas), the probabilities may be expressed as

$$P\{\Phi_\alpha | O\} = 1 - \sum_i^N P\{X_i | O\} + \tfrac{1}{2} \sum_{i,j,i \neq j}^N P\{X_i X_j | O\} - \ldots . \tag{12}$$

Next, we pass to the continuous limit (large N) by replacing the discrete sums in equation (12) by integrals:

$$P\{\Phi_\alpha | O\} = 1 - n \int_{|x| < \alpha} [1 + w(x)] d^2 x$$
$$+ \tfrac{1}{2} n^2 \int_{|x|,} \int_{|y| < \alpha} [1 + w(x) + w(y) + w(x - y) + z(x, y)] d^2 x d^2 y - \ldots . \tag{13}$$

In the above expression, we have also replaced probability statements by their corresponding definitions as correlation functions; $z(x, y)$ being the three-point correlation function defined in analogy with $w(x)$. Similarly, one may express $P\{\Phi_\alpha | OX\}$ in a series which, when combined with $P\{\Phi_\alpha | O\}$ in the form (13) according to (3), gives the required exact series expression for $w_s(x)$ in terms of correlations of all orders.

Setting all correlations of fourth order and above to zero gives the "singles" correlation function shown in the solid curve of Figure 2. Here, we have used an overall pair-correlation function of the form (1) with $A = 0.8$ and $\delta = 1$ in

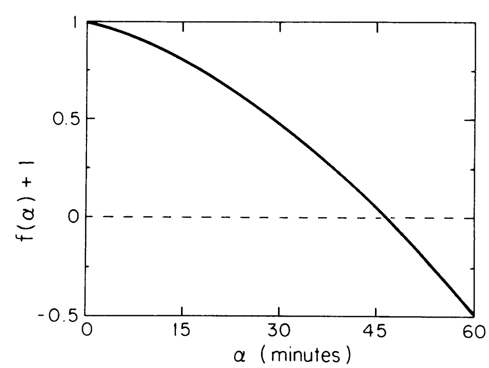

FIG. 1.—Behavior of $f(\alpha) + 1$ for $P\{\Phi_\alpha | O\}$ given by equation (10)

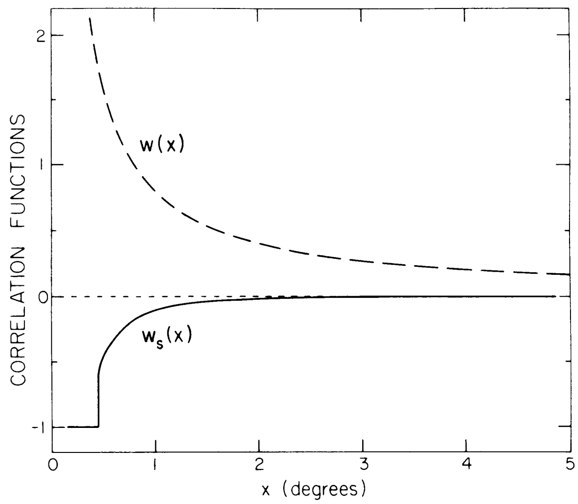

FIG. 2.—"Singles" correlation function $w_s(x)$ for galaxy distribution derived from the two-particle correlation function $w(x)$ displayed in the upper dashed curve (see eq. [1]) and the three-particle correlation function of the simple form suggested by Peebles and Groth (1975), the fourth and higher order correlations being set to zero. The $w(x)$ curve (*dashed*) shown here is the pair correlation function for *all* galaxies according to equation (1) with $A = 0.08$ and $\delta = 1$.

order to facilitate evaluation of the integrals, together with a three-point correlation function of the form

$$z(x,y) = w(x)w(y) + w(x)w(x - y) + w(y)w(x - y) . \tag{14}$$

This is the form found by Peebles and Groth (1975) to represent the actual distribution of galaxies. The selection angle α has been chosen to make the dominant terms in $w_s(x)$ vanish at large x. From Figure 2, it is apparent that $w_s(x)$ can be made very small in comparison with $w(x)$ over most of its range ($x > \alpha$). It is also apparent that the value of α is fairly close to the value 45′ used by Turner and Gott and suggested by our first argument.

Because the series expression for w_s alternates as correlations of higher order are included, we may hope that its convergence is fairly rapid. Inclusion of a positive fourth order correlation function would increase the value of α required for asymptotic cancellation. Of course, it is not expected that all correlation functions above some level are zero for the actual distribution of galaxies. Nevertheless, neglecting the fourth and higher order correlations, we have found a selection angle α which differs little from the one derived on the basis of our first argument.

III. CONCLUSIONS

a) The result $w_s \approx 0$ for $x > 45'$ found by Turner and Gott is probably a consequence of the particular selection angle chosen. Thus, in spite of the apparent separation of the galaxy population into two classes, their results are consistent with those of Peebles and his coworkers.

b) The "singles" population defined by Turner and Gott cannot necessarily be regarded as a true field component of the galaxy population: a similar separation could have been achieved given any underlying form of galaxy clustering with no preferred scale. In order that the "singles" population represent a physically unclustered component, one requires that it be uncorrelated with itself *and* uncorrelated with the "associated" component. The "singles" component selected by Turner and Gott will not in general satisfy these requirements (cf. eqs. [4]). The vanishing of w_s is an artifact of the selection process and does not necessarily imply the vanishing of these correlations. The autocorrelation of "singles" with themselves should be checked for the Turner-Gott sample.

c) We feel that statistical selection methods based on angular separations cannot be used by themselves to isolate galaxies into physically distinct field and cluster populations. Some supplementary criteria, based perhaps on magnitudes, redshifts, or diameters, must be used, though, no doubt, Turner and Gott's "associated" component contains a larger fraction of the galaxies which lie in physical groups than does the "singles" component.

We thank Marc Davis, Richard Gott, Jim Peebles, Paul Schechter, and Ray Soneira for helpful and entertaining discussions. This work was done during the Cambridge Workshop on Galaxy Formation sponsored by the Gravity Research Foundation.

REFERENCES

Burbidge, E. M., and Burbidge, G. R. 1961, *A.J.*, **66**, 541.
de Vaucouleurs, G. 1976, in *Stars and Stellar Systems*, Vol. **9**, ed. A. and M. Sandage (Chicago: University of Chicago Press).
Feller, W. 1968, *An Introduction to Probability Theory and its Applications*, Vol. **1** (New York: Wiley).
Holmberg, E. 1964, *Ark. f. Astr.*, **3**, 389.
Karachentsev, I. D. 1966, *Astrofizika*, **2**, 39.
Peebles, P. J. E. 1974, *Ap. J.* (*Letters*), **189**, L51.
Peebles, P. J. E., and Groth, E. J. 1975, *Ap. J.*, **196**, 1.
Totsuji, H., and Kihara, T. 1969, *Pub. Astr. Soc. Japan*, **21**, 221.
Turner, E. L., and Gott, J. R. 1975, *Ap. J.* (*Letters*), **197**, L89.

S. Michael Fall, Bernard J. T. Jones, and Simon D. M. White: Institute of Astronomy, Madingley Road, Cambridge, England

Margaret J. Geller: Harvard College Observatory and Smithsonian Astrophysical Observatory, 60 Garden Street, Cambridge, MA 02138