THREE-POINT CORRELATIONS OF GALAXY CLUSTERS

GÁBOR TÓTH AND JOSEPH HOLLÓSI
Space Telescope Science Institute and Roland Eötvös University

AND

ALEXANDER S. SZALAY
Roland Eötvös University and Johns Hopkins University

Received 1988 August 26; accepted 1989 January 6

ABSTRACT

We estimate the irreducible angular three-point correlation function of Abell clusters in distance classes 5 and 6 with Galactic latitude $|b| > 40^\circ$ from the Abell and Abell, Corwin, and Olowin catalog and of clusters identified in the Shane-Wirtanen catalog (Shectman). We find that the distribution of these clusters satisfies a relation between the two- and three-point correlation functions:

$$\zeta_{3} \approx Q(\xi_{r}, \xi_{s}, \xi_{u} + \xi_{r} + \xi_{u})$$

similar to galaxies. Higher order terms of $\xi$ are absent in $\zeta$, in contrast to predictions of biased galaxy formation theories. The value of $Q$ for the Abell clusters is found to be in the range

$$0.8 \leq Q \leq 1.1$$

the formal fits preferring lower values, and alternative weighting schemes pointing toward values close to 1.0. The less conservative bounds are best described by $Q = 1.0 \pm 0.1$. The SW clusters give a systematically lower value, $Q = 0.64 \pm 0.04$.

Subject heading: galaxies: clustering

1. INTRODUCTION

The statistical properties of the galaxy distribution were studied in detail during the last 20 years, the two-point, three-point, and even the four-point correlations being determined (for a summary see Peebles 1980). The two-point correlation function turned out to be very close to a power law: $\xi(r) = (r/r_0)^{-\gamma}$, with $\gamma \approx 1.8$ and $r_0 \approx 5h^{-1}$ Mpc. An interesting scaling relation was found between the spatial two- and three-point correlation functions $\xi$ and $\zeta$ (Groth and Peebles 1977):

$$\zeta(r, s, u) \approx Q\xi(\xi_{r}, \xi_{s}, \xi_{u} + \xi_{r} + \xi_{u})$$

where $Q = 0.8-1.3$ for the various catalogs. This scaling hierarchy seems to extend even further, and it is fairly well established for the four-point correlations (Fry and Peebles 1978). The scaling that seems to hold is

$$\xi_{n}(\ldots, \xi_{r}) = \lambda^{-n+1}\zeta(\ldots, \xi_{r})$$

This scaling relation is a stable closure of the BBKGY hierarchy (Davis and Peebles 1983; Guillen and Liboff 1988).

In the last few years both the angular and the spatial two point correlation function of the Abell clusters has been studied (Hauser and Peebles 1973; Bahcall and Soneira 1983; Klypin and Kopylov 1983; Couchman, McGill, and Olowin 1988). The Zwicky clusters, a much more numerous but statistically more uneven catalog was analyzed by Postman, Geller, and Huchra (1986). Surprisingly, the correlation function of clusters turned out to be very similar to the galaxy correlations, $\zeta(r) = (r/r_0)^{-\gamma}$, but for clusters $r_0 \approx 25-35h^{-1}$ Mpc. The similarity goes further: although at first glance the correlation length of the clusters is much larger than that of galaxies, it was pointed out that (at least on the two-point correlation level) pictures of galaxy and cluster catalogs would be indistinguishable from each other if we did not know in advance which is which, since $r_0$ in both cases is very close to the mean distance of objects in the catalog (Szalay and Schramm 1985). This is a sign of scale invariance, and the Shane-Wirtanen clusters (Shectman 1985), both in correlation length and mean separation in between the galaxies and the Abell clusters, satisfy the relation extremely well. The two-point correlation functions for all these samples are well represented by

$$\zeta(r) = 0.35((r/D_i)^{-\gamma}$$

where $D_i$ is the mean distance of objects in the catalog i. The fact that the correlation function of clusters is so much stronger and satisfies this scaling is somewhat puzzling and recently has provided one of the great challenges to galaxy formation theories. Here we try to investigate if this similarity of the clustering patterns is just accidental, or if it does extend to higher order correlations as well. Recently there has been some tendency to question the significance of Abell clusters because of the possibility that the visibility of a cluster is enhanced by the presence of an accidental foreground cluster (Sutherland 1988). Our results from angular cross-correlation of different distance classes (Szalay, Hollósi, and Tóth 1989), beside other arguments discussed in the literature, seem to contradict this theory. However, the Shane-Wirtanen clusters provide an independent check for our analysis.

In the CDM models one attempts to generate stronger correlations by biasing of density fluctuations, where we assume that from the primordial Gaussian fluctuations of the mass visible objects were only formed where the density exceeded a critical value (Kaiser...
1984; Politzer and Wise 1984; Bardeen et al. 1986). One could indeed get an amplification of the correlations, but at the price of making the objects rare. This biasing propagates to the distribution of the clusters as well, the large-scale part of the fluctuation spectrum modulating the small-scale part, forming galaxies in an exponential manner (Bardeen et al. 1986), and yielding the same biasing law as if the density smoothed on cluster scales had been clipped.

In the simple theories, the three-point correlations did not satisfy the scaling relation equation (1.1) exactly, and there is a term cubic in \( \xi \) present. One can still wonder if this is only an artifact of the particular biasing scheme used. It turned out that this is not the case. In a more general model (Szalay 1988) one can have an arbitrary nonlinear relation between luminosity and mass density. Assuming Gaussian fluctuations in the mass density, the luminosity fluctuations are strongly non-Gaussian. The three-point correlation function can be expanded in terms of \( \xi \), and the leading terms become

\[
\zeta(r, s, u) = Q(\xi, \xi_u + \xi_u \xi + \xi_u) + Q^2(\xi, \xi_u + \xi_u) + Q^3(\xi, \xi_u + \xi_u + \xi_u) + \cdots \cdots .
\]  

(1.4)

For the simple “clipping” discussed in the above references, \( Q = 1 \) and \( Q' = 0 \), and we recover the pure cubic term. But in any nonlinear biasing scheme the cubic term is always present in the three-point function with an amplitude \( Q^3 \), locked to the amplitude of the quadratic term. This can, of course, subsequently be altered by dynamical motion, which can be quite strong for galaxies, enough to wipe out the higher order terms and converge to the BBGKY similarity solution. Clusters are expected to have only small peculiar velocities, certainly not enough to move them substantially on 40 Mpc scales. One would thus expect some differences between these simple biasing scenarios and the scale-invariant hierarchy, which should be apparent in the distribution of the clusters.

In this paper we set out to determine the behavior of the angular three-point correlation function of clusters, with special emphasis on possible deviations from the quadratic scaling law (eq. [1.1]). In § II we describe the catalogs and the techniques we used to correct for the large-scale gradients. In § III we define our model for the three-point function and explain how that would project to the angular correlations. Section IV discusses the actual determination of the necessary bin counts and the preparation of random catalogs for comparison to the data. In § V we discuss the fitting procedure, error estimates, and the actual results of the various fits. In § VI we summarize our main results. Appendices A–C describe the details of the projection procedure and the probability distribution of random triplets.

II. DATA

We used several catalogs during this project. We used the original Abell catalog (Abell 1958) in part, and subsequently we obtained a floppy disk of the ACO catalog of 4073 rich clusters (Abell, Corwin, and Olowin 1988) from H. Corwin. This catalog contains complete southern data and some minor corrections to the “old” northern catalog. For the estimation of the selection function we also used the redshift compilation of Struble and Rood (1987a, b) containing redshifts for the “old” catalog.

The third catalog is the list of clusters identified in the Shane-Wirtanen catalog of galaxies (hereafter SW clusters) (Schectman et al. 1985). This survey contains 646 clusters, of which 40% are members of the “old” Abell catalog. The spatial density is ~6 times that of the \( R \geq 1, D \leq 4 \) sample. The redshift distributions of SW clusters and the \( D = 1–4 \) Abell clusters are similar. This is a good complement to the more distant \( D = 5 + 6 \) sample we use otherwise.

The extinction function is determined from the surface density distribution \( n(b^\parallel) \) of all clusters along the Galactic latitude. The clusters are counted in 50 equal area bins on both the north and south Galactic cap. The result is smoothed by a simple averaging process and is saved in a file to be used for generating random catalogs (Fig. 1). There are small differences between the north and south Galactic caps, between the “old” and “new” catalogs or the SW catalog. We compare our result to the well-known analytic formula for the extinction:

\[
E(b^\parallel) = N_{\text{obs}}/N_{\text{null}} = 10^{0.64(1 - \text{csc}(b^\parallel))}.
\]  

(2.1)

This expression does not necessarily hold for the cluster data since the cluster identification is a nonlinear process, but it is not far from the result of direct smoothing, either. For small separations the amplitude of the two-point correlation function depends on \( \langle n^2 \rangle \), the second power of the extinction function, and the amplitude of the three-point correlation function depends on \( \langle n^3 \rangle \), the third power of the function. These moments are not very sensitive to the details of the extinction function. We found that our extinction functions remove the large-scale gradients fairly well: the tails of the two-point correlation functions go to zero as expected (Fig. 2). In order to be safe, we restrict ourselves to high-latitude \( |b^\parallel| \geq 40^\circ \) clusters, where the extinction function is more uniform.

When estimating the three-point function, we used the \( D = 5 + 6 \) (distance classes 5 and 6 and richness class \( R \geq 1 \)) sample with Galactic latitudes \( |b^\parallel| \geq 40^\circ \) (HL for high latitude hereafter), since the \( D = 1–4 \) sample contained too few clusters to get a meaningful three-point correlation function. We split this into two subsamples (NC and SC) partly in order to estimate the errors and partly because the SC sample has been considerably revised by Abell et al. (1988) from the original catalog (see Table 1).

TABLE 1

<table>
<thead>
<tr>
<th>Sample</th>
<th>Galactic Latitude</th>
<th>Number of Clusters</th>
<th>Area (sr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High latitude (HL)</td>
<td>(</td>
<td>b^\parallel</td>
<td>\geq 40^\circ )</td>
</tr>
<tr>
<td>North cap (NC)</td>
<td>( b^\parallel \geq +40^\circ )</td>
<td>852</td>
<td>2.24</td>
</tr>
<tr>
<td>South cap (SC)</td>
<td>( b^\parallel \geq -40^\circ )</td>
<td>973</td>
<td>2.24</td>
</tr>
<tr>
<td>Shane-Wirtanen (SW)</td>
<td>(</td>
<td>b^\parallel</td>
<td>\geq 40^\circ )</td>
</tr>
</tbody>
</table>
Fig. 1.—The extinction function of the Abell catalog is shown for the NC and SC samples, respectively. Only clusters with $|b^\circ| \geq 40^\circ$ were used in our analysis.

SW sample is limited in declination as well, $\delta \geq -22^\circ.5$. Within these geometrical boundaries only Abell clusters of the “statistical sample” as defined by Abell (1958) can be found (the only exception is A915); thus, no further geometrical constraints are necessary to restrict ourselves to the statistical sample.

Since we calculate angular correlations, the redshift distribution does not have any effect on our results. Only when we relate the amplitudes of the angular correlations to spatial amplitudes does the selection function appear. As we shall see from the angular fits, we can conclude that higher order terms are absent from the three-point correlation function, and the amplitude of the quadratic term is rather insensitive to the actual shape of the selection function $P(x)$: the probability that at a given radial comoving distance a cluster will become a member of the Abell sample. For the sake of simplicity, we assign effective limits which are in reasonable agreement with those given by Seldner and Peebles (1977) but which have been updated somewhat by using the available sparse redshift information (Struble and Rood 1987a, b). We adopt $P(x) = 1$ within the limits 250–600 Mpc, and $P(x) = 0$ outside those limits as an effective selection function for the $D = 5 + 6$ sample.

III. THE MODEL FOR THE THREE-POINT CORRELATIONS

Let $\delta V_1, \delta V_2, \delta V_3$ be three volume elements with separations $r_{12}, r_{23},$ and $r_{31}$. The expected number of triplets in the volume elements is $\rho^3 \delta V_1 \delta V_2 \delta V_3$ for a uniform distribution, where $\rho$ is the mean density of clusters. In case of a correlated distribution there is an excess of triplets:

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

(3.1)

where $\xi$ and $\zeta$ are the spatial two- and three-point correlation functions, respectively.

In case of angular projections the definitions are similar, and the expected number of triplets becomes

$$n = n^3 \Omega_1 \Omega_2 \Omega_3 [1 + w(\theta_{12}) + w(\theta_{23}) + w(\theta_{31}) + z(\theta_{12}, \theta_{23}, \theta_{31})],$$

(3.2)

where $\Omega_1, \Omega_2,$ and $\Omega_3$ are three solid angle elements on the surface of the unit sphere with separations $\theta_{12}, \theta_{23},$ and $\theta_{31}, \eta$ is the mean surface density of clusters, and $w$ and $z$ are the angular two- and three-point correlation functions.

An easy way to estimate the spatial three-point correlation function $\zeta(r, s, u)$ would be to determine the number of triplets with given separations both in the data and random catalogs: $N_D(r, s, u)$ and $N_R(r, s, u)$. Assuming that we have already determined the two-point correlation function $\xi$, we could estimate $\zeta$:

$$\zeta(r, s, u) = \frac{N_D(r, s, u)}{N_R(r, s, u)} - \xi_r - \xi_s - \xi_u - 1.$$  (3.3)

Unfortunately only $\approx 300$ Abell clusters have measured redshifts in our sample, and even that sample is not statistically fair. Thus we cannot apply this method, since $N_D(r, s, u)$ would be too small, and we have to use angular correlations.

For two-point correlations the angular function $w(\theta)$ can be expressed from the spatial function $\xi(r)$ (Limber 1953, 1954), and vice versa (Fall and Tremaine 1977; Parry 1977). Thus, we can estimate $\xi(r)$ from $w(\theta)$ even if we do not know the individual redshifts, but only their statistical distribution. For three-point correlations we can express the angular function $z$ from the spatial function $\zeta$, but
the inversion is complicated (Regos 1988). In any case, all these inversions are rather unstable against statistical fluctuations, and usually one assumes a power-law form for the correlation functions for stability.

We are interested in the relation between the two-point and three-point functions of clusters. We prefer not to assume the existence of a quadratic scaling law, but rather to use a systematic expansion in terms of the various powers of the two-point functions. We consider all possible symmetric combinations in $\xi$ up to third order, which if present, should cause a signal strong enough to be detectable. Our model is using $\xi_\nu = \xi(r)$:

\[
\xi_M(r, s, u) = Q_0 + Q_1(\xi_s + \xi_u + \xi_s u) + Q_{11}(\xi_s \xi_u + \xi_s u + \xi_s u) + Q_2(\xi_s^2 + \xi_u^2 + \xi_s u^2 + \xi_s u^2) + Q_{111}(\xi_s \xi_u \xi_u + \xi_s \xi_u u + \xi_s u u) + \xi_{21}(\xi_s^2 \xi_u + \xi_s \xi_u^2 + \xi_s^2 u + \xi_s u^2) + \xi_{31}(\xi_s \xi_u^2 + \xi_s^2 \xi_u + \xi_s u^2 + \xi_s^2 u) + Q_3(\xi_s^3 + \xi_s^2 u + \xi_s u^2 + \xi_s^2 u) \tag{3.4}
\]

Using the notation $k \in \{0, 1, 11, 2, 111, 21, 3\}$, the expansion (eq. [3.4]) can be written in a shorter form:

\[
\xi_M(r, s, u) = \sum_k Q_k(\xi_k(r, s, u)) \tag{3.5}
\]

where $\xi_k(r, s, u) = \xi^{k_1}_{12} x^{k_2}_{23} x^{k_3}_{3} +$ symmetric terms.

Our goal is to determine the coefficients $Q_k$. Assuming that the selection function $P(r)$ is known, the expansion can be projected to the corresponding angular correlation functions with linear coefficients $A_k$:

\[
z_M(a, b, c) = \sum_k A_k w_k(a, b, c). \tag{3.6}
\]

The connection between $Q_k$ and its projection $A_k w_k$ is the following:

\[
A_k w_k(a, b, c) = \frac{\rho^3}{3} \int_0^\infty dr_1 r_1^3 P(r_1) \int_0^\infty dr_2 r_2^3 P(r_2) \int_0^\infty dr_3 r_3^3 P(r_3) \xi_k(r, s, u), \tag{3.7}
\]

where $r^2 = r_1^2 + r_2^2 - 2r_1 r_2 \cos a, s^2 = r_1^2 + r_3^2 - 2r_1 r_3 \cos b$ and $u^2 = r_2^2 + r_3^2 - 2r_2 r_3 \cos c$.

Accepting $\xi(r) = (r/r_0)^\gamma$, we can integrate analytically (except the $\xi_s, \xi_u, \xi_{su}$ term, where the $\gamma = 2$ assumption is necessary for analytical integration, which is a good approximation). The calculation is similar to the projection of $\xi(r)$ to $w(\theta)$, but it is more complicated (see Appendices A and B). The results are:

\[
\begin{align*}
L_0(a, b, c) &= 1, \\
L_1(a, b, c) &= w_a + w_b + w_c, \\
L_11(a, b, c) &= w_a w_b + w_a w_c + w_b w_c, \\
L_2(a, b, c) &= \left(\frac{w_a^2}{a} + \frac{w_b^2}{b} + \frac{w_c^2}{c}\right) \frac{H(2\gamma)}{H(\gamma)^2} \frac{180}{\pi}, \\
L_111(a, b, c) &= \left(\frac{w_a^2 w_b}{a b} + \frac{w_b^2 w_c}{b c} + \frac{w_c^2 w_a}{c a}\right) \frac{H(2\gamma)}{H(\gamma)^2} \frac{180}{\pi}, \\
L_3(a, b, c) &= \left(\frac{w_a^3}{a^3} + \frac{w_b^3}{b^3} + \frac{w_c^3}{c^3}\right) \frac{H(3\gamma)}{H(\gamma)^3} \frac{180}{\pi} \frac{2}{\gamma^2},
\end{align*}
\]

where $H(\gamma) = \int_1^\infty dx (1 + x^2)^{-\gamma/2}$.

The angular coefficients are proportional to the spatial coefficients, but the ratios are different for the different terms because of the various moments of $P(r)$ involved. Let us denote the ratios by $R_k = A_k/Q_k$. The numerical results for $\gamma = 2$ and for the selection function $P(r)$ (1 for 250 Mpc $\leq r \leq 600$ Mpc; 0 otherwise) are the following:

\[
\begin{align*}
R_0 &= 1, \\
R_1 &= 1, \\
R_11 &= \frac{M_1(3) M_3(9 - 2\gamma)}{M_2(6 - \gamma)^2} = 1.040, \\
R_2 &= \frac{M_1(3)^2 M_3(6 - 2\gamma)}{M_2(6 - \gamma)^2} = 0.672, \\
R_{111} &= \frac{M_3(3)^3 M_3(9 - 3\gamma)}{M_2(6 - 3\gamma)^2} = 0.642, \\
R_{21} &= \frac{M_1(3)^3 M_3(9 - 3\gamma)}{M_2(6 - 3\gamma)^3} = 0.642, \\
R_3 &= \frac{M_1(3)^4 M_3(6 - 3\gamma)}{M_2(6 - 3\gamma)^3} = 0.562,
\end{align*}
\]

where $M_k(\alpha) = \int_0^\alpha dr P(r) r^{k-1}$, and $a, b, and c$, are measured in degrees. The $R_k$ constants depend on the dimensionless ratios of the various $M_k(\alpha)$ moments; thus, they are not too sensitive to the shape of $P(x)$.

We need the angular two- and three-point correlation functions so as to determine the coefficients $A_k$ from the angular expansion, and then the spatial coefficients $Q_k = A_k/R_k$.

IV. THE DETERMINATION OF THE ANGULAR CORRELATIONS

a) Angular Two-Point Correlations

The angular two-point correlation function for the NC, SC, HL, and SW samples are estimated the following way. One hundred random catalogs are generated with the same boundaries and extinction functions and with the same number of clusters as in the data samples. The pairs in 12 linear bins are counted up to $5'$ separation for the data sample and the random catalogs as well. The estimator of the angular correlation for each bin is:

\[
w(\theta) = \frac{\langle DD \rangle_\theta}{\langle DR \rangle_\theta} - 1, \tag{4.1}
\]
Fig. 2.—Angular two-point correlation functions of the D = 5 + 6, R ≥ 1 Abell cluster samples with |b| ≥ 40°. Filled squares correspond to the high-latitude HL sample; empty symbols correspond to the south cap (SC) and the north cap (NC). The box in the upper right-hand corner shows w(θ) out to 60° separations.

where <DR>θ and <DD>θ are the number of data-data and data-random pairs in a bin of (θ − Δθ/2, θ + Δθ/2). To check how effectively the large-scale gradient is removed, the two-point correlation function is determined up to 60°. Figure 2 shows the behavior of w(θ), and at large angles the tail goes to zero as expected. Figure 3 shows the two-point correlation function on logarithmic scale; one can see how well the power-law fits.

b) Angular Three-Point Correlations

All triplets with separations up to 5° are counted and put into bins according to the length of the sides. We have decided to use these (a, b, c) instead of the usual (a, u, v) (Groth and Peebles 1977). Using a as the size parameter, one is averaging over u and v, triangles of very long sides, where the effect may be less apparent: it turns out that the three-point correlation function grows fastest in a direction orthogonal to a. Let n(a, b, c) be the number of triplets in a bin defined by

\[ V(a, b, c) = \left[ a - \frac{\Delta}{2}, a + \frac{\Delta}{2} \right] \times \left[ b - \frac{\Delta}{2}, b + \frac{\Delta}{2} \right] \times \left[ c - \frac{\Delta}{2}, c + \frac{\Delta}{2} \right], \]

where Δ is the size of the bin and a ≤ b ≤ c. The estimator for the angular three-point correlation function,

\[ \langle z(a, b, c) \rangle = \frac{\langle n_{DDD} \rangle - \langle n_{DDR} \rangle}{\langle n_{DRR} \rangle} + 2 = \frac{\langle n_{a, b, c} \rangle}{\langle n_{DRR(a, b, c)} \rangle}, \]

where n_{a, b, c} = n_{DDD} - n_{DDR} + 2n_{DRR} is the excess number of triplets due to the three-point correlation function in the V(a, b, c) bin. The quantity n_{DDD} is the total number of triplets in a given bin, n_{DDR} is cross-correlating the data catalog with itself and a random catalog (essentially a triplet weighted mean of the two-point correlation function), and n_{DRR} is the expected random triplet count for a uniform distribution. The latter two were calculated from 100 Monte Carlo catalogs with the proper extinction function (smoothed from the data). This method is more stable than subtracting the global average of the two-point correlation function, since it takes the edge effects better into account for a given triangle.

Equation (4.3) is the mean value of z(a, b, c) in the bin V(a, b, c), it becomes the true value of z in the center for infinitesimal Δ only. Reducing Δ is difficult because of the relatively small number of triplets (≈ 40,000 in the whole sample). As a compromise, we use Δ = (5/12)', yielding 239 bins, with the geometrical constraints for the sides: a ≤ b ≤ c ≤ a + b. Thus, having enough bins to determine z(a, b, c), we still have enough triplets within each bin to avoid a large statistical scatter, although Δ ≈ a, b, c does not hold for most of the bins.

Our goal is to determine the expansion of the three-point correlation function. In order to avoid the difficulties mentioned above, in the fitting process we use the integrated triplet counts rather than z itself. Let P(a, b, c) be the density distribution function of triplets:

\[ P(a, b, c) = \lim_{\Delta \to 0} \frac{n(a, b, c)}{\Delta^3}. \]
The estimator for \( z(a, b, c) \) in equation (4.3) becomes exact if we replace the bin counts by the density distributions:

\[
z(a, b, c) = \frac{P_{\text{DDD}}(a, b, c) - P_{\text{DDR}}(a, b, c)}{P_{\text{DRR}}(a, b, c)} + 2.
\] (4.5)

Let us rearrange this equation and integrate the volume of a bin, thus obtaining an exact equation for the triplet counts independently of the bin sizes, relating to the directly measurable quantities:

\[
\langle n_z(a, b, c) \rangle = \langle n_{\text{DDD}} - n_{\text{DDR}} + 2n_{\text{DRR}} \rangle = \iiint_{V(a, b, c)} da \, db \, dc \, P_{\text{DRR}}(a, b, c) z(a, b, c).
\] (4.6)

We model this integrated three-point part \( \langle n_z(a, b, c) \rangle \) of the data by the integral over the expansion in equation (3.6):

\[
n_M(a, b, c) = \sum_k A_k n_k(a, b, c),
\] (4.7)

with the integrated \( n_k \) functions

\[
n_k(a, b, c) = \iiint_{V(a, b, c)} da \, db \, dc \, P_{\text{DRR}}(a, b, c) w_k(a, b, c).
\] (4.8)

Having determined the \( n_k(a, b, c) \) functions from equation (4.8) we can fit \( n_M(a, b, c) \) to \( n_z(a, b, c) \) with the \( A_k \) parameters.

In order to obtain \( n_k \) we need the probability density of random triplets \( P_{\text{DRR}}(a, b, c) \). Since we use small separations, the nonlinear effects of the boundary and the density gradient are negligible, so the uniform distribution is a good approximation for the random catalogs. In this case the distribution of triplets can be calculated analytically taking \( P_{\text{DRR}}(a, b, c) = P_R(a, b, c) \) (Appendix C):

\[
P_R(a, b, c) = A \langle n^3 \rangle \frac{8\pi abc}{\sqrt{2(a^2 + b^2 + c^2)^3 - a^4 - b^4 - c^4}}.
\] (4.9)

The accurate coefficient in this equation for a finite area might be a bit smaller because of the presence of the boundaries, but we compensate for it.

Since \( P_R(a, b, c) \) is singular at \( a + b = c \), the terms in equation (4.8) must be integrated carefully. The values of \( P_R(a, b, c) \) are calculated at \( \approx 20,000 \) grid points in each bin and are multiplied by the various \( w_k \) terms (eq. [3.8]), estimated from the angular two-point correlations. Finally, the integrated values are multiplied by a normalization factor and are stored for the various samples, respectively.
We normalize these functions by comparing the total number of triplets expected for uniform distribution to the number of triplets in random catalogs which are taking into account the extinction and the boundaries. The triplets with separations up to $5^\circ$ are counted in 100 random catalogs generated with the proper extinction function and boundaries for each sample. Let us denote by $T_k$ the average of the counts in the 100 random catalogs and by $T_A$ the expected number of triplets calculated from the analytic formula:

$$T_K = \sum_{a,b,c} n_K(a, b, c),$$

$$T_A = \sum_{a,b,c} n_0(a, b, c) = \sum_{a,b,c} \int_{V(a, b, c)} da \, db \, dc \, P_R(a, b, c).$$

(4.10)

Assuming that the difference between $T_K$ and $T_A$ comes from linear effects of the large-scale gradient and the boundaries, the normalization factors are set to $T_K/T_A$ for each sample separately. To show that this linear correction is enough we tested $n_k(a, b, c)$ versus the corrected $n_0(a, b, c)$ function in every bin. The difference is reasonably small, considering that the statistical dispersion of triplet counts, even for random catalogs, can be higher than the Poisson error (Toth 1988).

V. FITTING

a) Two-Point Correlations

First, we determine the angular two-point correlation function of the clusters. We fit a power law to the data, out to about $20^\circ$. The best fit for the HL sample is, locking the slope to $-1$,

$$w(\theta) = \frac{0.71}{\theta}. \quad (5.1)$$

For the NC and the SC samples the amplitudes become $0.70$ and $0.72$, respectively (Fig. 3), in excellent agreement with both Bahcall and Soneira (1983) and Couchman, McGill, and Olowin (1988). We integrate the powers of these functions to obtain $n_k$. A higher resolution analysis of $w(\theta)$ shows that it becomes zero below $\theta_{\text{min}} = 0.225$, which is presumably due to the finite angular size of the clusters. In order to avoid excessive counts in the innermost bin, we truncated our fit the same way; otherwise we could have obtained an unreasonably large contribution from the small angles when we integrate in equation (4.8).

For the SW sample the angular two-point correlation function is not very well fitted by any power law, although it is consistent with one. Thus, we integrated the actual binned two-point correlation function in order to get the $n_k$.

b) Error Model

We determine $n_0(a, b, c)$ for each sample in 239 bins of size $\Delta = (5/12)^\circ$. The sides of the triplets are $a \leq b \leq c \leq 5^\circ$. In order to obtain the most likely values of the parameters $A_k$, we minimize

$$\chi^2 = \sum_{a,b,c} \frac{1}{\sigma^2(a, b, c)} \left[ n_0(a, b, c) - \sum_k A_k n_k(a, b, c) \right]^2,$$

(5.2)

where $\sigma(a, b, c)$ is the rms dispersion of the triplet count in the $V(a, b, c)$ bin. The whole fitting scheme depends on how we assign a value to this dispersion and, as a result, how the various parts of the $(a, b, c)$ space are weighted. The integral contribution to the reduced three-point function is

$$n_Z = n_{\text{DDD}} - n_{\text{DDR}} + 2n_{\text{DRR}} = n_{\text{DDD}} - n_{\text{DRR}} (1 + w_a + w_b + w_c). \quad (5.3)$$

There are two types of errors: statistical and systematic. The largest source of statistical error is in $n_{\text{DDD}}$; we take it to be Poissonian, not a bad approximation if the total number of the clusters is fixed in the sample. The statistical error in $n_{\text{DRR}}$ is minimal, since it was determined by averaging over 100 random catalogs. The statistical errors in the two-point subtraction are dominated by systematic errors which are due to large-scale gradients. The uncertainty in the estimate of the extinction function results in a deviation in the "tail" of $w(\theta)$. We take the systematic error of each $w$ value to be $\epsilon$, and we determine $\epsilon$ from the requirement that the reduced $\chi^2$ for the two-point correlation function fit should be about 1. This yields $\epsilon = 0.06$. Our expression for $\sigma(a, b, c)^2$ then becomes

$$\sigma(a, b, c)^2 = n_{\text{DDD}} + 3\epsilon^2 n_{\text{DRR}}. \quad (5.4)$$

It is comforting to see that this gives reasonable reduced $\chi^2$ values (slightly above 1) for our fits to the three-point function (see below).

c) Three-Point Correlations

All the functions with which we are fitting have a spike at the origin and tend to zero at large separations. The sample is dominated by triangles, where the values of all functions (and the data) is close to zero; therefore, the most interesting region will get a relatively small weight. On the other hand, in order to distinguish between different power laws one needs a wide dynamic range. The distribution of the $\langle z(a, b, c) \rangle$ values is shown in Figure 4 as a function of the smallest side $a$ and the mean size defined by $p = (a + b + c)/3$. The size of the hexagon indicates the value of $1 + \langle z \rangle$.

First, we fit to all higher order functions, and in order to have the maximum dynamic range we use all triangles in the HL sample. When we actually do the fit, we normalize the various $n_k$ to have the same norm: the total number of excess triplets in the 239 bins. The results are shown in Table 2, where $k$ identifies the function. The fitted coefficients for the normalized set are all consistent with
zero except for $k = 11$. When we scale to spatial functions, the relative amplitudes change (they are somewhat dependent on the shape of the selection function), but our conclusion does not change. The only noticeable thing is the relatively large uncertainty in the cubic ($k = 11$) component. At this stage we conclude that for the HL sample the amplitudes are

$$Q_{11} = 0.6937 \pm 0.0746, \quad Q_{22} = 0.0051 \pm 0.0068, \quad Q_{111} = 0.0025 \pm 0.1140,$$

$$Q_{21} = 0.0088 \pm 0.0079, \quad Q_3 = -0.0001 \pm 0.0001,$$

and $\chi^2 = 262.34$ for $N_{\text{df}} = 234$.

Next, we establish a fiducial point by fitting to a single function, the quadratic $\xi_{11} = \xi_{tt} + \xi_{tt} + \xi_{tt}$, and then consider relative changes in the value of $\chi^2$ from there. We obtain

$$Q_{11} = 0.82 \pm 0.04$$

using all 239 triangles, with $\chi^2 = 270.70$ for $N_{\text{df}} = 238$.

Now we investigate the effect of slightly modifying the relative weights associated with the small and large triangles by including triangles with their smallest side $a \leq a_{\text{max}}$. For orientation the distribution of bins as a function of $a$ is shown in Figure 4. We shrink the sample size down to $a = 1$, containing only 22 bins. These cuts have the advantage that besides confining the sample to the bins where $z$ is relatively high, we keep a large dynamic range. Cutting by a fixed value of $a$ instead of cutting in a ratio $a/p$ has the additional advantage that in the latter case the bins would still be heavily dominated by the outer ones, containing values close to zero. Such a cut depending on the geometry of the triangles is allowed: if we have a good fit, it should be good in any region. Given the absolute errors $\sigma(a, b, c)^2$, cutting in $a$ results in a region where the relative errors become smaller. It does not necessarily result in a smaller error in the value of the parameters since the sample becomes smaller, but we are considering a physically more

### Table 2

<table>
<thead>
<tr>
<th>$k$</th>
<th>Normalized</th>
<th>Angular</th>
<th>Spatial</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>0.7215 ± 0.0776</td>
<td>0.7215 ± 0.0076</td>
<td>0.6937 ± 0.0746</td>
</tr>
<tr>
<td>2</td>
<td>0.0848 ± 0.1133</td>
<td>0.0034 ± 0.0046</td>
<td>0.0051 ± 0.0068</td>
</tr>
<tr>
<td>111</td>
<td>0.0007 ± 0.0317</td>
<td>0.0016 ± 0.0731</td>
<td>0.0025 ± 0.1140</td>
</tr>
<tr>
<td>21</td>
<td>0.0792 ± 0.0714</td>
<td>0.0057 ± 0.0051</td>
<td>0.0088 ± 0.0079</td>
</tr>
<tr>
<td>3</td>
<td>-0.0800 ± 0.0640</td>
<td>0.0000 ± 0.0000</td>
<td>-0.0001 ± 0.0001</td>
</tr>
</tbody>
</table>

$N_{\text{df}} = 234; \chi^2 = 262.34$

**Note.**—Results of the fit with the five higher order functions to the angular three-point correlations in the HL sample; $k$ identifies the function. Three values of the coefficients and their errors are given: normalized, angular and spatial. The normalized values should be used to compare the relative contribution of the various functions. All but the quadratic term $Q_{11}$ are consistent with zero. The degrees of freedom and the resulting $\chi^2$ are given in the last row.
meaningful regime for the low values of \(a_{\text{max}}\). The results of applying a set of different cuts is shown in Table 3 and graphically in Figure 5.

We tested the same sample for the model expected from simple biasing (Szalay 1988), \(\xi = Q\xi_{11} + Q^3\xi_{111}\). The best value of \(Q\) and the corresponding \(\chi^2\) are shown in Table 3. The right-hand box in Figure 5 shows the change in \(\chi^2\) relative to the \(\xi_{11}\) fit: the typical value \(\Delta\chi^2 > 9.80\), corresponding to an acceptance probability less than 0.001, so the biasing prediction can be rejected.

From the results of the five-parameter fit we are still left with the possibility that there may be a small but nonnegligible contribution coming from the \(\xi_{111}\) term. We have tried a set of fits to test this (Table 4), changing \(a_{\text{max}}\) from 12 to 1 as in the previous cases. At \(a_{\text{max}} = 12\) we obtain

\[
Q_{11} = 0.76 \pm 0.04, \quad Q_{111} = 0.11 \pm 0.04, \tag{5.7}
\]

with \(\chi^2 = 264.64\) for \(N_{df} = 237\), barely different from the five-parameter fit, meaning that the other three parameters are indeed irrelevant. On the other hand, the difference from the canonical one-parameter fit seems to be somewhat significant: \(\Delta\chi^2\) is between \(-4\) and \(-6\) for a change of 1 in the effective degrees of freedom. However, as we decrease \(a_{\text{max}}\) this significance disappears, and at

![Figure 5](https://example.com/figure5.png)

**Fig. 5.—**The dependence of the fits to the HL sample on \(a_{\text{max}}\), the upper limit on the smallest side of the triangle. Smaller \(a\) values reduce the number of large triangles, changing the relative weighting. Filled squares are the value of \(Q_{11}\), a one-parameter fit; the star shows the amplitude of the combination \(Q\xi_{11} + Q^3\xi_{111}\). Open circles are the results of the two-parameter fit; the lower curve corresponds to \(\xi_{111}\). The right-hand side graph is the \(\chi^2\) value relative to the canonical case \(Q_{11}, \xi_{11}\) alone. Note that for small \(a_{\text{max}}\), the parameter \(Q_{111}\) approaches zero.
TABLE 4

<table>
<thead>
<tr>
<th>a_m</th>
<th>N_{et}</th>
<th>Q_{11}</th>
<th>Q_{111}</th>
<th>\chi^2</th>
<th>\Delta \chi^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>12...</td>
<td>237</td>
<td>0.76 ± 0.04</td>
<td>0.11 ± 0.04</td>
<td>264.64</td>
<td>-6.06</td>
</tr>
<tr>
<td>10...</td>
<td>233</td>
<td>0.76 ± 0.04</td>
<td>0.10 ± 0.04</td>
<td>259.16</td>
<td>-5.89</td>
</tr>
<tr>
<td>8</td>
<td>217</td>
<td>0.78 ± 0.05</td>
<td>0.10 ± 0.04</td>
<td>233.78</td>
<td>-5.03</td>
</tr>
<tr>
<td>6</td>
<td>181</td>
<td>0.78 ± 0.05</td>
<td>0.10 ± 0.04</td>
<td>194.14</td>
<td>-4.66</td>
</tr>
<tr>
<td>4</td>
<td>120</td>
<td>0.80 ± 0.05</td>
<td>0.09 ± 0.05</td>
<td>142.54</td>
<td>-4.81</td>
</tr>
<tr>
<td>3</td>
<td>85</td>
<td>0.80 ± 0.05</td>
<td>0.09 ± 0.05</td>
<td>113.78</td>
<td>-3.65</td>
</tr>
<tr>
<td>2</td>
<td>51</td>
<td>0.89 ± 0.07</td>
<td>0.05 ± 0.05</td>
<td>65.70</td>
<td>1.22</td>
</tr>
<tr>
<td>1</td>
<td>21</td>
<td>0.84 ± 0.10</td>
<td>0.03 ± 0.05</td>
<td>27.41</td>
<td>-0.37</td>
</tr>
</tbody>
</table>

Note.—Fit with two functions, \( \xi_{11} \) and \( \xi_{111} \), to determine whether the presence of \( \xi_{111} \) is significant. The dependence on the cuts in \( a_{\text{max}} \) is also shown. \( \Delta \chi^2 \) refers to the \( \chi^2 \) value in col (4) of Table 3, the one parameter fit with \( \xi_{11} \) alone. Note that the more we restrict the fit to the inner part, the larger \( Q_{11} \) and the smaller \( Q_{111} \) become. This is also shown in Fig. 6.

The same time the value of \( Q_{111} \) approaches zero. The quantity \( Q_{11} \) rises slightly toward higher values, from 0.76 to 0.84, as shown in Figures 5a–5b.

There is one other way of changing the weighting scheme to assign slightly higher weights to the inner part: changing the ratio of the systematic to statistical errors. This will not give any reasonable values for \( \chi^2 \) but may enable us to see the stability of the parameters with respect to various weighting strategies. We change the value of \( \epsilon \) from the "fair" value of 0.06 to about 1.0 and follow the behavior of the fits. We have done this for the \( \xi_{11} \) fit and the \( \xi_{11}, \xi_{111} \) combination, with \( a_{\text{max}} = 3 \) and 10. The curves are shown in Figure 6a. In all cases the tendency is the same: as \( \epsilon \) grows, \( Q_{11} \) approaches 1 and \( Q_{111} \) goes to zero.

d) The Other Samples

The SC subsample gives immediately a large value close to 1:

\[
Q_{11} = 0.93 \pm 0.04,
\]
with \( \chi^2 = 289.4 \). The tail of the NC sample has some troubles with the large-scale gradients; this shows up as a very low value

\[
Q_{11} = 0.48 \pm 0.05
\]

Fig. 6.—The dependence of the fit on the amount of systematic error \( \epsilon \). Increasing \( \epsilon \) has the effect of decreasing the weight of the outer bins, shifting weight closer to the origin. All fitted values of \( Q_{11} \) tend toward 1.0, whereas \( Q_{111} \) disappears. The left-hand box shows the results for the HL sample. Filled squares are the values for the fit with \( Q_{11} \), only and \( a_{\text{max}} = 3 \); the other two curves belong to the two-parameter fit with \( Q_{11} \) and \( Q_{111} \). Open circles show the value of \( a_{\text{max}} = 3 \), and asterisks show \( a_{\text{max}} = 12 \). The right-hand part shows the same behavior for the subsamples NC and SC.
in the formal fitting process. However, increasing $\epsilon$ results in a nice convergence between $0.9 \leq Q_{11} \leq 1.05$. This again confirms our belief in this weighting scheme (Fig. 6b).

The best fit to the SW clusters has a formal value

$$Q_{11} = 0.64 \pm 0.04 .$$

This value of $Q_{11}$ is stable with changes in $\epsilon$ or $a_{\text{max}}$. Since the SW sample is independent of the HL sample, it is interesting to examine the contribution of $Q_{11}$ in this sample as well:

$$Q_{11} = 0.75 \pm 0.05 , \quad Q_{111} = -0.13 \pm 0.05 ,$$

with $\chi^2 = 282.89$ for $N_{\text{of}} = 237$. The result is similar to equation (5.7). The SW sample is considerably smaller, so the statistical errors are bigger, and it is more affected by the finite size of the nearby clusters, which changed the behavior of the correlation function at small distances. Still, this is a reasonably good agreement, and with the future availability of the redshifts for this catalog it will be much easier to determine $Q_{11}$ more accurately.

VI. CONCLUSIONS

We reached the following conclusions regarding the three-point correlations of Abell and SW clusters:

1. We detected the presence of a strong irreducible three-point correlation function for the clusters.
2. The shape of the three-point correlation function is fully consistent with the quadratic scaling law found by Groth and Peebles (1977) for galaxies,

$$\zeta = Q(\xi_r \xi_s + \xi_s \xi_u + \xi_u \xi_r) = Q_{11} .$$

3. We searched for other higher order terms in the series expansion of $\zeta$, and all but the $\xi_{111} = \xi_r \xi_s \xi_u$ term can be rejected. The amplitude was zero with less than 1 $\sigma$.
4. The three-point correlation function is inconsistent with the expectations from biasing, where the minimal form becomes $\zeta = Q_{11} + Q_{111} \xi_{111}$. This form of $\zeta$ can be rejected at the 0.001 significance level.
5. Our best fit with two parameters for the combination $\zeta = Q_{11} \xi_{11} + Q_{111} \xi_{111}$ using the HL sample is given by

$$Q_{11} = 0.76 \pm 0.04 , \quad Q_{111} = 0.11 \pm 0.04 .$$

6. The best formal fit with one parameter is

$$Q_{11} = 0.82 \pm 0.04 .$$

7. Trying out various weighting schemes (without a strong control on the value of $\chi^2$) showed that when weights were shifted more toward the small separations, the value of $Q_{11}$ approaches unity, and the best estimate $Q_{111}$ becomes zero. Therefore we conclude that the acceptable range of $Q_{11}$ is

$$0.80 \leq Q_{11} \leq 1.10$$

with the most likely value close to 1, so we can set a less conservative bound as

$$Q_{11} = 1.0 \pm 0.1 .$$

8. The NC and SC catalogs are compatible with the above numbers, although NC has some problems with the removal of gradients. With increasing $\epsilon$, both samples converge to $Q_{11}$ between 0.9 and 1.05. The SW sample yields the best formal fit of $Q_{11} = 0.64 \pm 0.04$.

We would like to acknowledge useful discussions with N. Bahcall, R. Burg, and M. Fall. Electronic versions of the various catalogs were kindly provided to us by H. Corwin, R. Olowin, and S. Shectman. This research has been supported by NASA within the Graduate Student Program at STScI, by NSF and JHU in the US and OTKA in Hungary.

APPENDIX A

PROJECTION OF THE TWO-POINT CORRELATION FUNCTION

1. NOTATION

The connection between $\zeta$ and $w$ (Limber equation):

$$w(\theta) = \frac{\rho^2}{n^2} \int_0^\infty dr_1 r_1^2 P(r_1) \int_0^\infty dr_2 r_2^2 P(r_2) \zeta(r) ,$$

(A1)

where $r^2 = r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta$, $\rho$ is the spatial density of the clusters, and $n$ is the projected surface density.

Accepting $\zeta(r) = (r/r_0)^{-7}$ and a smooth selection function on the scale of the correlation length, we can integrate analytically for small separations in $\theta$. The following result was also derived by Peebles (1980), but we repeat it here because both the special techniques and the result will be used in our more complicated calculations subsequently.
Let us denote the various moments of $P(r)$ by $M_k(r) = \int_0^\infty dr r^k P(r)$. It is easy to see that the ratio of the surface and spatial density $\eta/\rho = M_1(r)$. Let us denote a frequent dimensionless integral by $H(n) = \int_{-\infty}^{\infty} dx (1 + x^2)^{-n/2}$. Some special values of $H(n)$ are $H(2) = \pi$, $H(4) = \pi/2$, and $H(6) = 3\pi/8$.

II. INTEGRATING THE LIMBER EQUATION

Let us use new variables and $x$:

$$r_1 = q, \quad r_2 = q(1 + \theta x), \quad r \approx qa\sqrt{1 + x^2},$$

and the Jacobian is $aq$. Considering the power-law form of $x$,

$$M_0 = M_1 \frac{q^2}{q} P(q) \int_{-\infty}^{\infty} dq q^2 P(q) \int_{-\infty}^{\infty} dx \theta q^2 (1 + \theta x)^2 P[q(1 + \theta x)] \left(\frac{q\theta \sqrt{1 + x^2}}{r_0}\right)^{-\gamma}. \quad (A3)$$

The separation $\theta$ is small ($\leq 5^\circ$), and at large $x$ values the integrand is negligible. Thus we can use the approximation $1 + \theta x \approx 1$, and we integrate from $-\infty$ to $\infty$ in the second integral:

$$w(\theta) = M_1(3)^{-2} \frac{q^2}{q} \int_{-\infty}^{\infty} dq q^2 P(q) \int_{-\infty}^{\infty} dx (1 + x^2)^{-\gamma/2}. \quad (A4)$$

The result is

$$w(\theta) = \frac{q^{3-\gamma} \theta^{1-\gamma} \int_{-\infty}^{\infty} dq q^{5-\gamma} P(q)^2 \int_{-\infty}^{\infty} dx (1 + x^2)^{-\gamma/2}}{M_1(3)^{2}} H(\gamma). \quad (A5)$$

APPENDIX B

PROJECTION OF HIGHER ORDER MOMENTS

I. NOTATION

For brevity we introduce the following notation: $\xi(r, s, u) = \xi_k + \text{symmetric terms}$, where $k \in \{0, 1, 11, 2, 111, 21, 3\}$. Now the spatial and the corresponding angular expansions are

$$\zeta_M(r, s, u) = \sum_k Q_k \xi_k(r, s, u), \quad z_M(a, b, c) = \sum_k A_k w_k(a, b, c). \quad (B1)$$

Our purpose is to determine the $w_k$ terms and their coefficients from the following relation:

$$A_k w_k(a, b, c) = \frac{\rho^3}{\eta} \int_0^\infty dr_1 r_1^2 P(r_1) \int_0^\infty dr_2 r_2^2 P(r_2) \int_0^\infty dr_3 r_3^2 P(r_3) Q_k \xi_k(r, s, u), \quad (B2)$$

where

$$r^2 = r_1^2 + r_2^2 - 2r_1 r_2 \cos a, \quad s^2 = r_1^2 + r_3^2 - 2r_1 r_3 \cos b, \quad u^2 = r_2^2 + r_3^2 - 2r_2 r_3 \cos c.$$ Accepting $\xi(r) = (r/r_0)^{-\gamma}$ and a smooth selection function, we can integrate analytically for small separations except the $\xi_r$ term, where the $\gamma = 2$ approximation is also necessary. Let us denote the ratio of the corresponding coefficients by $R_k = A_k/Q_k$. The arguments of $w$ and $\xi$ will be written as indices: $w_a, w_b, w_c, \xi_{a}, \xi_{b}$, and $\xi_{c}$.

II. TERMS OF THE FORM $\xi_k(r, s, u)$

There are three terms belonging to this category: $\xi_1$, $\xi_2$, and $\xi_3$. We have to integrate

$$R_k w_k(a, b, c) = \frac{\rho^3}{\eta} \int_0^\infty dr_1 r_1^2 P(r_1) \int_0^\infty dr_2 r_2^2 P(r_2) \int_0^\infty dr_3 r_3^2 P(r_3) (\xi_k + \xi_r + \xi_u). \quad (B3)$$

It is enough to integrate $\xi_k$ only, since the other terms are symmetric. Note that $\xi_r$ does not depend on $r_3$. Thus, the third integral can be evaluated, and it gives a factor of $M_1(3)$. Let us use new variables $q$ and $x$:

$$r_1 = q, \quad r_2 = q(1 + ax), \quad r \approx qa\sqrt{1 + x^2}, \quad (B4)$$

and the Jacobian is $aq$. Considering the explicit forms of $\xi$ and $P$,

$$R_k w_k(a, b, c) = M_1(3)^{-2} \int_0^\infty dq a q^2 P(q) \int_{-\infty}^{\infty} dx q^2 (1 + ax)^2 P[q(1 + ax)] \left(\frac{qa\sqrt{1 + x^2}}{r_0}\right)^{-\gamma} + \text{symmetric terms}. \quad (B5)$$

© American Astronomical Society • Provided by the NASA Astrophysics Data System
Taking $1 + ax \approx 1$ and integrating from $-\infty$ to $\infty$ in the second integral:

$$R_{i}w_{i}(a, b, c) = r_{0}^{3}a^{-1-k}M_{1}(3)\int_{0}^{\infty} dq \int_{-\infty}^{\infty} dx (1 + x^{2})^{-k/2} + \text{symmetric terms}.$$  \hspace{1cm} (B6)

Let us use $w_{a}$ to express $a^{-1-k}$ (eq. [A5]):

$$R_{i}w_{i}(a, b, c) = M_{1}(3)\frac{a^{2-k}}{M_{2}(6 - k)} \left( \frac{w_{a}^{2}}{a^{k+1}} + \text{symmetric terms} \right) H(k_{y})H(l_{y}) \cdot$$  \hspace{1cm} (B7)

For $k = 1$ we get the expected $w_{a} = w_{b} = w_{c}$ equation, but for $k = 2$ and $3$ the result is far from trivial.

III. TERMS OF THE FORM $\xi_{i}(r, s, u)$

There are two terms in this category: $\xi_{11}$ and $\xi_{21}$. We have to integrate

$$R_{i}w_{i}(a, b, c) = M_{1}(3)\int_{0}^{\infty} dq \int_{-\infty}^{\infty} dx (1 + x^{2})^{-k} \int_{-\infty}^{\infty} dy (1 + y^{2})^{-l/2} + \text{symmetric terms}.$$  \hspace{1cm} (B8)

Let us use new variables $q, x,$ and $y$:

$$r_{1} = q, \quad r_{2} = q(1 + ax), \quad r_{3} = q(1 + by), \quad r \approx qa\sqrt{1 + x^{2}}, \quad s \approx qb\sqrt{1 + y^{2}}, \quad$$  \hspace{1cm} (B9)

and the Jacobian is $aq^{2}$. Considering the explicit form of $\xi$:

$$R_{i}w_{i}(a, b, c) = M_{1}(3)\int_{0}^{\infty} dq \int_{-\infty}^{\infty} dx (1 + x^{2})^{-k} \int_{-\infty}^{\infty} dy (1 + y^{2})^{-l/2} + \text{symmetric terms}.$$  \hspace{1cm} (B10)

Again, with $1 + ax \approx 1, 1 + by \approx 1$, and integrating from $-\infty$ to $\infty$ in the second and third integrals,

$$R_{i}w_{i}(a, b, c) = r_{0}^{3}a^{-1-k}b^{-1-l}M_{1}(3)\int_{0}^{\infty} dq \int_{-\infty}^{\infty} dx (1 + x^{2})^{-k} \int_{-\infty}^{\infty} dy (1 + y^{2})^{-l/2} + \text{symmetric terms}.$$  \hspace{1cm} (B11)

Let us use $w_{a}$ and $w_{b}$ to express $a^{-1-k}$ and $b^{-1-l}$:

$$R_{i}w_{i}(a, b, c) = M_{1}(3)a^{2-k-l}b^{2-k-l}M_{2}(6 - k - l) \left( \frac{w_{a}^{2}w_{b}^{2}}{a^{k+1}b^{l+1}} + \text{symmetric terms} \right) H(k_{y})H(l_{y}) \cdot$$  \hspace{1cm} (B12)

IV. TERM $\xi_{i11}(r, s, u)$

We have to integrate equation (B2) for $k = 111$:

$$R_{i11}w_{i11}(a, b, c) = M_{1}(3)\int_{0}^{\infty} dq \int_{-\infty}^{\infty} dx (1 + x^{2})^{-k} \int_{-\infty}^{\infty} dy (1 + y^{2})^{-l/2} + \text{symmetric terms}.$$  \hspace{1cm} (B13)

Let us use new variables $q, x,$ and $y$:

$$r_{1} = q, \quad r_{2} = q(1 + ax), \quad r_{3} = q(1 + by), \quad r \approx qa\sqrt{1 + x^{2}}, \quad s \approx qb\sqrt{1 + y^{2}}, \quad u \approx qc\sqrt{1 + \left(\frac{a}{c} - \frac{b}{c} - y\right)^{2}}, \quad$$  \hspace{1cm} (B14)

and the Jacobian is $aq^{2}$. Let us denote the ratios of the separations by $u = a/c$ and $v = b/c$. Considering the power-law form of $\xi$:

$$R_{i11}w_{i11}(a, b, c) = M_{1}(3)a^{2-k-l}b^{2-k-l}M_{2}(6 - k - l) \left( \frac{w_{a}^{2}w_{b}^{2}}{a^{k+1}b^{l+1}} + \text{symmetric terms} \right) H(k_{y})H(l_{y}) \cdot$$  \hspace{1cm} (B15)

With $1 + ax \approx 1, 1 + by \approx 1$, and integrating from $-\infty$ to $\infty$ in the second and third integrals:

$$R_{i11}w_{i11}(a, b, c) = r_{0}^{3}a^{-1-k}b^{-1-l}c^{-1-\gamma}M_{1}(3)\int_{0}^{\infty} dq \int_{-\infty}^{\infty} dx (1 + x^{2})^{-\gamma/2} \int_{-\infty}^{\infty} dy (1 + y^{2})^{-\gamma/2} + \text{symmetric terms}.$$  \hspace{1cm} (B16)

© American Astronomical Society • Provided by the NASA Astrophysics Data System
Unfortunately the third integral can be evaluated only for special values of \( \gamma \); thus, the \( \gamma = 2 \) approximation is necessary. In this case the third integral can be decomposed into four partial fractions:

\[
\frac{1}{(1 + y^2)(1 + (\mu x - vy)^2)} = \frac{A + Cy}{1 + y^2} + \frac{B + Dv(\mu x - vy)}{1 + (\mu x - vy)^2},
\]

where \( A, B, C, \) and \( D \) are functions of \( \mu, v, \) and \( x \). The odd functions with the \( C \) and \( D \) coefficients give zero (i.e., the principal value of their improper integrals). The result of the third integral in equation (B18) is

\[
A\pi + B\frac{\pi}{v} = \pi \frac{1 + v}{(1 + v)^2 + (\mu x)^2}.
\]

The second integral can be decomposed into two partial fractions:

\[
\int_{-\infty}^{\infty} dx \frac{\pi(1 + v)}{(1 + x^2)((1 + v)^2 + (\mu x)^2)} = \pi(1 + v) \int_{-\infty}^{\infty} dx \left[ \frac{1}{1 + x^2} - \frac{v^2}{(1 + v)^2 + (\mu x)^2} \right] = \frac{\pi^2}{1 + \mu + v}.
\]

Expressing \( \mu \) and \( v \) by \( a, b, \) and \( c \) and evaluating the first integral, we obtain:

\[
R_{111}w_{111}(a, b, c) = \frac{r_0^2}{a^{1-\gamma}b^{1-\gamma}c^{1-\gamma}} \frac{M_g(9 - 3\gamma)}{M_1(\gamma)} \frac{\pi^2 c}{a + b + c}.
\]

Let us use \( w_a, w_b, \) and \( w_c \) to express \( a^{1-\gamma}b^{1-\gamma}c^{1-\gamma}:

\[
R_{111}w_{111}(a, b, c) = \frac{M_1(3)\frac{M_g(9 - 3\gamma)}{M_1(\gamma)}}{M_g(6 - \gamma)} \left( \frac{w_a w_b w_c}{a + b + c} \right) \frac{\pi^2}{H(\gamma)}.
\]

APPENDIX C

DISTRIBUTION OF THE RANDOM TRIPLETS

For a uniform distribution of points the distribution of triplets can be calculated analytically for small separations. The number of clusters in a sample is \( \lambda \lambda^3 \), where \( \lambda \) is the area of the sample and \( \lambda^3 \) is the mean surface density. Let us choose a cluster to be the first member of the triplet. The number of clusters with separations between \( b \) and \( b + \Delta b \) from this first cluster equals \( 2\pi b^2 \Delta b \frac{\lambda^3}{\lambda^2} \). Fixing the second member of the triplet, we can calculate the area of the region corresponding to points with separations between \( c \) and \( c + \Delta c \) from the second cluster. Considering that \( \Delta \) goes to zero, the result is \( \Delta^2 \sin \alpha \), where \( \alpha \) is the angle between the sides \( b \) and \( c \). Thus the number of triplets is

\[
P_{\text{R}}(a, b, c) \Delta^3 = A(\lambda^3)(2\pi b^2 \Delta b \frac{\lambda^3}{\lambda^2}) \frac{\Delta^2}{\sin \alpha}.
\]

Since \( \lambda \) is varying with the Galactic latitude, \( \langle \lambda^3 \rangle \) has to be corrected to \( \langle \lambda^3 \rangle \). This is a good correction for triplets with small separations. Expressing \( \sin \alpha \) from \( a, b, \) and \( c \), we get

\[
P_{\text{R}}(a, b, c) = A(\lambda^3) \frac{8\pi abc}{\sqrt{2(a^2 b^2 + b^2 c^2 + c^2 a^2) - a^4 - b^4 - c^4}}.
\]

The accurate coefficient in this equation for a finite area might be a bit smaller because of the presence of the boundaries.

REFERENCES


Joseph Hollósi, Alexandre S. Szalay, and Gábor Tóth: Department of Atomic Physics, Eötvös University, Puskin u. 5-7, H-1088 Budapest, Hungary

© American Astronomical Society • Provided by the NASA Astrophysics Data System