The distribution of faint galaxies near the South Galactic Pole — II

P. C. Hewett  Department of Astronomy, University of Edinburgh, 
Blackford Hill, Edinburgh EH9 3HJ, Scotland

H. T. MacGillivray and R. D. Dodd  Royal Observatory, 
Blackford Hill, Edinburgh EH9 3HJ, Scotland

Received 1980 October 1; in original form 1980 June 4

Summary. The clustering of galaxies in the MacGillivray & Dodd (1980b) sample, obtained from COSMOS machine measures, has been investigated using the method of Mead’s analysis and dispersion—subdivision curves. The results show two characteristic clustering scale-lengths in the data, with angular sizes ~0°.12 and ~1°.00. At the redshift z = 0.5 (assuming q0 = +1), the corresponding linear dimensions of the features are ~3h⁻¹ and ~25h⁻¹ Mpc respectively. This is interpreted as providing evidence for the second-order clustering of galaxies. Comparison with computer simulations of galaxy fields indicates that this second-order clustering is consistent with a cellular model for superclustering.

1 Introduction

In a recent paper (MacGillivray & Dodd 1980b), hereafter referred to as Paper I, the distribution of faint galaxies in a field of 15 square degrees to a limit of B ≈ 22.0 was examined from COSMOS machine measures on a deep UK Schmidt Telescope (UKST) plate. The field was centred on the ESO/SRC southern survey field number 413, approximately 10° from the South Galactic Pole. The smallest images on the plate were 2 arcsec in diameter. From this study it was found that, on small angular scales, the galaxies were distributed in small groups or clusters with a mean ‘multiplicity’ of 3.6, in agreement with similar studies of the Lick and Jagellonian samples (Fesenko & Pi’ev 1975; Fesenko & Fesenko 1978).

In this paper, we analyse the distribution of galaxies in the MacGillivray & Dodd sample by other means, viz. the method of Mead’s analysis and dispersion—subdivision curves. To support our conclusions, comparison is made with computer simulations of galaxy fields.

2 Observational material

The data have already been described at length in MacGillivray & Dodd (1980a) and in Paper I, together with details of the photometry and star/galaxy separation technique (performed down to B = 22.0). The number—magnitude count for galaxies (MacGillivray & Dodd 1980a)
was found to be in good agreement with the results of other investigators at both North and South Galactic Poles, apart from a deficiency of galaxies with $B \sim 16.0$. A possible systematic error in the zero point provided by the night sky photometer of $\pm 0.2$ mag may be present. Other work (Corwin, in preparation) confirms the accuracy attainable with the night sky photometer on the UKST (in its original position, i.e. pointing to the South Celestial Pole) and the linearity of the galaxy magnitudes from COSMOS data (at least for galaxies in the range $13.0 < B < 17.5$). Checks of the COSMOS magnitudes for galaxies have been published in Dodd et al. (1979).

3 Results

3.1 Mead's Analysis

A description of the method of Mead's analysis (Mead 1974; Besag & Diggle 1977) as applied to the statistical clustering of galaxies has been given by Shanks (1979) and references therein. Of special note is the ability of the technique to identify a number of clustering scales due to the use of the local galaxy density rather than the average density. This means that effects of clustering will not be overwhelmed by large-scale density gradients as with the dispersion—subdivision technique (Section 3.2). The drawback of both methods is the small number of discrete scales that may be examined.

The method was applied by Shanks to computer simulations of galaxies in the Catalogue of Galaxies and Clusters of Galaxies (Zwicky et al. 1961–1968) and to a number of samples including the Jagellonian and Dodd et al. (1975) sample. More recently (Shanks et al. 1980), the statistic has been used in a study of a sample of galaxies to a limiting magnitude of $J \sim 21.5$ from COSMOS measures on UKST plates at the South Galactic Pole. The scale of clustering detected by Shanks et al. using both Mead's analysis and covariance analysis (Peebles 1973), is somewhat smaller ($3 h^{-1}$ Mpc) than that obtained from covariance analyses on the brighter surveys ($9 h^{-1}$ Mpc) (Groth & Peebles 1976).

To apply the method to the data of Paper I, the raw data were rebinned at higher resolution ($0.06 \times 0.06$ cells) giving a $64 \times 64$ element array. The Mead's statistic was calculated using the approximation of Shanks (1979). We used 100 random fields with the same number of galaxies to determine $T_{\text{mean}}$ — the mean value of the 'raw statistic' (cf. Shanks 1979) for the random distributions — and $S_A$ its standard error. The final statistic $S$ was calculated as

$$S = \frac{T_{\text{obs}} - T_{\text{mean}}}{S_A}$$

for each scale examined. The results are shown in Fig. 1(a) (filled circles), where two peaks are evident at $0.12$ and $0.96$, providing evidence that two scales of clustering are present in the data.

Note that our results are quantitatively lower (by a factor of 5) than the results of Shanks. This is due to the different methods of 'randomizing' the data. However, a comparison with our results on the Dodd et al. field confirms that while the results are quantitatively lower, the features are none the less identical.

3.2 Dispersion—Subdivision Curves

The method of dispersion—subdivision curves has been described at length by Zwicky (1957) and de Vaucouleurs (1971, 1977). The significant features of the technique (which is related to the cellular $\chi^2$ method, e.g. Abell 1961) are:
Figure 1. (a) Results of the Mead's analysis on the data of Paper I (rebinned with a resolution of $0^\circ.06$). (b) Comparison of the Mead's analysis for the present data, limited at $B = 22.0$, with that for the Jagellonian, Dodd et al. and Ellis et al. red samples (from Shanks 1979). To allow intercomparison of the shape of the curves, a zero point shift in the $y$-direction has been performed bringing the first peaks into approximate coincidence.

(a) The use of the mean galaxy density (averaged over the field) as the standard from which the cell variances are calculated.

(b) The comparison of observed variance with that of the Poisson distribution with the same mean number of galaxies/cell, for square cell sizes which are exact divisions of the total area under investigation.
Use of the mean galaxy density over the whole field and the comparison with Poisson statistics results in apparently significant values of the index of clumpiness \( k \) (at small angular sizes) that are not necessarily due to intrinsic properties of the galaxy distribution. At large angular cell sizes, the presence of large-scale density gradients in the galaxy distribution introduces high values of \( k \) also.

The technique as used by other workers has produced differing results. Zwicky (1957) found a feature at \( 3^\circ \) cell size in the Coma cluster, whereas analyses of the Corona (Zwicky 1957) and Tautenburg M3 (Dautcourt et al. 1978) fields showed only a steadily increasing value of \( k \) over the range of cell sizes examined. The strong feature in the region of the Coma cluster is not unexpected. A continuous rise in the value of \( k \) with cell size in the latter cases could be due primarily to factors not intrinsic to the galaxy distribution (e.g. plate non-uniformity and/or physiological effects). As Section 5 indicates, very strong features are necessary to produce discrete 'peaks' in the characteristically rising curve of this method of analysis.

Results from the data of Paper I are presented in Fig. 2 (filled circles), where the original resolution of Paper I has been retained. De Vaucouleurs (1971) suggests that values greater than 3 for the index of clumpiness \( k \) usually indicate the presence of clustering. For galaxies down to \( B = 22.0 \), values of \( k \) greater than 3 are obtained on cell sizes greater than \( 0^\circ 48 \). The dispersion–subdivision curve for these galaxies shows a peak at the cell size of \( 0^\circ 96 \). We interpret this peak as providing further evidence for a preferred clustering scale-length at angular size of \( 1^\circ \) in our sample. The final point at \( 1^\circ 92 \) shows a continuing rise in the value of \( k \) for larger cell sizes. It is tempting to speculate that an even larger scale structure exists beyond \( 2^\circ \) angular size. However, a reliable investigation of this can only be made with a larger sample of data.

3.3 Variation with limiting magnitude

Although the data of Paper I are in the form of counts to a single limiting magnitude, the original galaxy data on magnetic tape contain the positional and photometric parameters for each image, allowing the counting of galaxies to different limiting magnitudes.

In Fig. 1(a) are also shown the Mead's analysis for the data to three other limits, viz. \( B = 19.0, 20.0 \) and \( 21.0 \). There are two points worthy of note. First, the peak at small
angular sizes remains peaked and moves to larger angular size with brighter limiting magnitude, indicating that the smaller clustering scale-length is discrete and scales with limiting magnitude. Secondly, the peak at larger angular size disappears for brighter magnitudes. This is probably due to the fact that this large-scale feature is itself defining the background for the bright galaxies and is therefore not resolved in the data.

The corresponding dispersion—subdivision curves are shown in Fig. 2. The flattening of the curve with magnitude is also consistent with the above interpretation.

4 Possible explanations for the large-scale feature

We cannot be absolutely certain that the large-scale feature detected in the data of Paper I has not been produced by some means other than an intrinsic property of the galaxy distribution, although the fact that we can reproduce the feature from simulations is an argument in favour. Before investigating the structure of the observed feature further, it is worthwhile to discuss other effects that might give rise to a feature of this nature in the data. There are four possible causes:

(a) The peak is due entirely to some artefact of the plate, the measuring machine or the software techniques (e.g. the background representation algorithm or method for star/galaxy separation). This is always a possibility in measurements of this type. However, such variations occur either at a very much higher or a very much lower frequency. For example, the stability of the COSMOS machine remains unchanged over periods of days; the background following algorithm is allowed to vary on scales of 0°.02 following variations in the background plate transmission; the star/galaxy separation criteria were constant over the whole area studied (MacGillivray & Dodd 1980a). Furthermore, an examination of Fig. 1 of Paper I reveals no effect on scales of 1° that might be attributable to machine or software. The fact that the Jagellonian data produces a similar shaped curve (Fig. 1b) when the plate material, counting procedure, star/galaxy separation were completely different is also strong evidence against this possibility.

(b) The peak is an artefact of the statistical methods. We argue against this suggestion on the grounds that the two techniques give similar results in that a feature is present at 1° angular size.

(c) The peak is real and is due entirely to patchy galactic obscuration, a suggestion similar to that of Fesenko (1978). Work done by Hewett et al. (in preparation) and the simulations of Section 5 argue against the hypothesis that a feature of this magnitude in the galaxy distribution at scales of 1° can be produced by patchy obscuration. This is discussed further in Section 5.

(d) The peak is a statistical fluctuation present only in this field. Analysis of the Jagellonian data (Shanks 1979) produces a curve with the same features as in the present sample (Fig. 1b). Two smaller data sets, the Dodd et al. (1975) sample and Ellis, Fong & Phillipps (1977) red sample, also show the small scale ~ 0°.2 peak (Fig. 1b). The angular extent of the latter two samples is smaller than the Jagellonian or Paper I samples, but the rise evident at the largest angular scale in the two smaller samples is consistent with a large-scale feature at 1° being present. The number of galaxies present in a sample affects the absolute values of the Mead statistic, S. Similar clustering in a large sample gives higher values than in a small sample. Consequently, in Fig. 1(b) we have shifted the curves in the y direction by a zero-point to produce approximate coincidence for the first peaks. The important point to note is the shape of the curves which show remarkable agreement.

Two other data sets, the Ellis et al. (1977) J sample and the Shanks et al. (1980) sample, produce different Mead’s curves. It is known that the Ellis et al. J sample is affected by
emulsion variations; the Mead's curve which indicates a strong overall increase with no pronounced peak at $0^\circ.1 \sim 0^\circ.2$ is entirely consistent with this. The Shanks et al. data agrees well with the small scale peak at $0^\circ.2$ as expected, clustering on scales up to several Mpc being well established. However, the results at larger scales are markedly different; they observe a steep drop-off in $S$ beyond $\sim 0^\circ.3$. Because of the size of structures producing the secondary peak ($\sim$ some tens of Mpc, equivalent to $2^\circ$ at $B \sim 22.0$) it is possible that in a $4^\circ \times 4^\circ$ sample from a UKST plate, the presence or absence of a few of these structures can markedly affect the large-scale results. Otherwise, some other explanation, possibly in the data reduction techniques, must be sought to explain the differences. It is noted that one of the Shanks et al. fields was filtered to remove effects of an area apparently deficient in galaxies.

Our confidence in the presence of a large-scale feature present in our sample is further strengthened by the evident scaling of the curves which behave as expected with change in limiting magnitude (Fig. 1a).

5 Comparison with computer simulations

Model fields of galaxies have been generated using a Monte Carlo technique which has been described in other papers (Dodd 1977; MacGillivray et al. 1980a, b). The volume of space modelled is a pyramid with the observer at the apex and the base being defined by the measured area on the photograph. Galaxies are generated to populate this space with distances in the range $100 < r < 3000$ Mpc. The highest-order clustered objects (e.g. clusters, superclusters or super-superclusters) are randomly distributed in this space. Clusters whose centres lie outside the space but which have a possibility of contributing individual galaxies within the space are also generated. In the final analysis, only galaxies whose projected coordinates fall within the field of the measured area are included.

All angular separations are corrected for cosmological effects (Mattig 1958; Sandage 1961). We have assumed throughout that $q_0 = +1$ and $H_0 = 75$ km s$^{-1}$ Mpc$^{-1}$. Galaxy magnitudes are generated by pseudo-random pair selection of points on intersecting lines (Dodd 1977), biased by the luminosity function of Arakelyan & Kalloglyan (1970). The spatial frequency of galaxy types generated is taken from Pence (1976). The spectral energy data of Pence are also used to generate $K$ corrections for all galaxy types with a resolution in $z$ of 0.01. The galaxy luminosities are increased for luminosity evolution (Brown & Tinsley 1974) and reduced for effects of distance, cosmology, $K$-dimming, patchy galactic obscuration and telescope vignetting.

A full account of the modelling of patchy galactic obscuration and the effects on the observed distribution of faint galaxies is the subject of a separate paper (Hewett et al. in preparation). However, a brief outline is appropriate at this point.

Computer models of the interstellar medium may be produced, allowing simulation of the effects of extinction over entire UKST and AAT (Anglo-Australian Telescope) plates. These models are based on the theoretical work of McKee & Ostriker (1977) and Spitzer (1978), on observational work of Bohlin, Savage & Drake (1978), Kalberla, Mebold & Reich (1980) and Heiles & Cleary (1979) – H$^1$ emission and absorption profiles – and on the stellar photometry of Hilditch, Hill & Barnes (1976). The model parameters used are consistent with both the theoretical and observational constraints.

We find that small variations of $\sim \pm 0.03$ in $E_B - V$ on scales of $0^\circ.2$ to $3^\circ.0$ can produce considerable effects in the galaxy distribution near the limiting magnitude of the sample. However, we have been unable to reproduce features of the magnitude evident in the sample of Paper I (Fig. 1a) or in the Jagellonian sample (Shanks 1979 Fig. 12, and Fig. 1(b) of this paper).
To determine a suitable multiplicity function for the simulations, i.e. the frequency of clusters with a given number of galaxies, we have adapted data from Gott & Turner (1977), using simple assumptions to derive the number of clusters with $n$ members instead of the number of clusters of a given integrated luminosity. From these data, we derive that the frequency of clusters with $n$ members is $f_1 = 0.545$ for $n = 1$ and $f_n = 1.097 \ n^{-7/3}$ for $2 < n < 700$

(with $\sum_{n=1}^{700} f_n = 1$).

This multiplicity function gives a mean cluster membership

$\langle n \rangle = \sum_{n=1}^{700} f_n \ n$

of 3.0.

The radial distribution of galaxies in a cluster is generated with a power law bias, with index $\gamma = -1.77$, and where the cluster radius defines the cut-off value. Cluster and supercluster radii may be generated at random or with a preselected probability distribution (e.g. delta or Gaussian functions). The ‘core’ radius is defined as the radius containing 50 per cent of the galaxies. We have used a value of 1/3 for the ratio of core to cluster radii. This is purely arbitrary and the results are not significantly altered by changing this quantity.

The magnitude of each generated galaxy (after corrections as described above) is compared with the observed plate limit and the galaxy rejected if fainter. Galaxies are generated until the number brighter than the plate limit is equal to the number in the real field. The galaxies retained are projected on to the plane of the sky and their apparent distribution analysed by the same techniques (Mead’s analysis and dispersion–subdivision curves) as were applied to the real data.

Figs 3 and 4 show the results for five models; the points plotted are in fact the mean of 10 consecutive simulations for each model, thus smoothing fluctuations due to different initial random numbers. In model (a) cluster sizes were restricted to $3 \ h^{-1} \text{Mpc}$ and no superclustering was included; in model (b) a continuous range of cluster sizes in the range $0.75 \ h^{-1}$

![Figure 3. The Mead's analysis for the computer simulations as described in the text. Each model is the mean of 10 simulations. The error bars denote the $2\sigma$ difference in the simulations of model (d).](https://example.com/figure3.png)
to $30 \, h^{-1}\text{Mpc}$ was allowed, the sizes being generated at random within this range; in model (c) clusters with sizes of $3 \, h^{-1}\text{Mpc}$ were distributed with a Gaussian bias in superclusters of maximum size $30 \pm 2 \, h^{-1}\text{Mpc}$ and with $10 \pm 5$ clusters per supercluster (Gaussian bias); model (d) was similar to model (c) but with the clusters situated at the supercluster radii, i.e. in a bubble-type manner; model (e) was the same as for model (a) but with patchy galactic obscuration included. The error bars denote the $2\sigma$ difference in the results of the 10 simulations for model (d). For the other models the error bars are similar.

It is seen from Figs 3 and 4 that the only model to produce a peak similar to that in the real case is model (d), i.e. the model in which the $3 \, h^{-1}\text{Mpc}$ clusters are distributed in a cellular manner. These models are, of course, very simple and the real case is obviously more complex, e.g. the clusters being distributed in sheets or filaments. The results for the Mead's analysis are also quantitatively higher than the real case, a fact which could be explained by the adopted multiplicity function underestimating the contribution from isolated galaxies. However, the fact remains that the features observed in the real data are better reproduced by a model in which the clusters are distributed in a cellular manner rather than in centrally concentrated superclusters.

A qualitative examination of the isoplethal map of Fig. 1 in Paper I supports the conclusion of a cellular-type distribution, since the high-density peaks of the galaxy distribution appear to form rings. A similar effect is noticeable in Fig. 1(a) of MacGillivray & Dodd (1979). In the latter paper we commented that the clusters appeared to be linked together by 'bridges' of galaxies. This observation is consistent with the cellular picture for the organization of galaxies and clusters.

6 Discussion and conclusions

The results from the COSMOS data indicate that two preferred scales of clustering are present in the distribution of galaxies in the sample of Paper I, viz. at $\sim 0.12$ and $\sim 0.96$. Using $M^* = -21.0$, corresponding to the knee in the luminosity function (Christensen 1975), $q_0 = +1$ and $z^* = 0.5$ (i.e. the redshift at which a galaxy of intermediate type with $M^* = -21.0$ would be seen at the limiting magnitude), then the corresponding linear dimensions of the features are $\sim 3 \, h^{-1}\text{Mpc}$ and $\sim 25 \, h^{-1}\text{Mpc}$ respectively. Comparison with computer simulations of galaxy fields suggests that the second-order organization is in the form of clusters being distributed in a bubble or cellular manner.
Studies of galaxy distributions from nearby samples of galaxies and clusters (e.g. Groth & Peebles 1977) have primarily utilized the two-point and higher-order correlation functions. From these analyses a picture of clustering on all scales (up to $9\ h^{-1}\ Mpc$) has emerged, consistent with the simple power law representation of the covariance functions. There has been much debate concerning the ability of the covariance analysis to discriminate between different clustering schemes (Jones 1976; Soneira & Peebles 1978; Shanks 1979; Peebles 1979). It would appear that the observed form of the two-point covariance function may be well reproduced by a variety of such schemes (Shanks 1979).

Amplitudes in the covariance functions due to large-scale features (i.e. $> 9\ h^{-1}\ Mpc$) may be so low that detection of such is difficult (for nearby samples, where large linear scale corresponds to large angular scale) in the presence of effects of selection and/or galactic extinction. For this reason, the presence of clustering on larger preferred scales is not ruled out. Thus our detection of a second-order preferred scale is not at variance with the results so far obtained. The covariance function is an inappropriate method to use on our sample to study the large-scale structure, i.e. out to $\sim 2^\circ$, because this is a significant fraction of the sample size. Peebles (1974a) has discussed this and for the Jagellonian sample only calculated the covariance function out to $1^\circ$ (Peebles 1975) for this reason. For samples of small angular size, cellular methods of the Dispersion—Subdivision and Mead's type are to be preferred because all the data are used to calculate the value of the statistic at each angular scale, in contrast to any correlation technique which has to allow for edge effects.

The form of the Mead's analysis curves observed in the Jagellonian, Dodd et al., Ellis et al. red and the present sample are consistent with the cellular supercluster structure proposed by Jöeveer, Einasto & Tago (1978) from observations of galaxies and clusters in the southern hemisphere. The conclusions of Jöeveer et al. are supported by recent observations of the Coma/Abell 1367 region (Gregory & Thompson 1978), the Hercules supercluster region (Tarenghi et al. 1979, 1980) and the Indus supercluster (Corwin, in preparation), the distribution of bright Sc galaxies (Rubin et al., 1976) and redshift surveys (Chincarini & Rood 1979 and references therein). The redshift observations reveal a remarkably consistent picture; 'holes' of $\sim 100\ Mpc$ diameter and the concentration of visible matter (galaxies) in sheets or shells being the main features.

An important difference between the cellular model and the continuous clustering model is that preferred scales will exist if the 'holes' and sheets have particular sizes. The wide range of structures studied in the above references and the uniformity of the scale estimates of 'holes' and sheets suggests that this may be the case.

Our results from the computer simulations have shown that input parameters, derived from nearby redshift observations, can reproduce the features in the deep distributions of the Jagellonian and Paper I samples. We note that the actual sizes of the simulated structures are larger than the detected size. This implies that the interpretation of observed clustering 'scale-lengths' is not straightforward. The cellular model outlines the spatial relationship between isolated galaxies and small and large groups, together with rich Abell-type clusters, distributing them with some form of large-scale ordering. That this scheme should be able to reproduce the features of the Mead's statistic in the deeper samples lends support to the model.

We emphasize that the large angular scale of these structures $\sim 1^\circ$ to $2^\circ$ at the typical limiting magnitude ($B \sim 22.0$) and the typical small sample size (i.e. $4^\circ \times 4^\circ$) may combine to produce variations in the large-scale Mead's statistic; such effects have been reproduced in our simulations. Statistical variations from field to field are to be expected and the error bars on the simulated curves are thus overestimates. In contrast, none of the more standard models (i.e. models a, b, c and e) used in the simulation produced a Mead's curve which was
similar to the observed Mead’s curve in any one simulation, i.e. the statistical fluctuations to be expected in the other models are not large enough to produce the observed Mead’s curve.

The above observations have important consequences for our understanding of the formation of superclusters, implying the presence of a turbulence mechanism (Doroshkevich, Sunyaev & Zel’dovich 1974) rather than a gravitational instability mechanism (Peebles 1974b; Aarseth, Gott & Turner 1979). Further deep samples together with more discriminating tests of galaxy clustering are now necessary to confirm these conclusions. A prediction of the turbulence model is the formation of elongated clusters (‘pancakes’). This has gained some observational support (MacGillivray et al. 1976; Carter & Metcalfe 1980). More evidence for the turbulence mechanism may be provided by further such studies of the shapes of clusters.

We are currently studying much larger samples of galaxies with COSMOS, covering fields some tens of degrees in diameter to look for confirmation of our results in other areas. A search for higher order clustering of galaxies on scales of hundreds of Mpc might also prove possible with such larger samples.

Acknowledgments

Thanks are due to Dr D. Emerson, Mr H. G. Corwin Jr., Dr N. M. Pratt and Dr B. Guthrie for useful comments on an original draft of this paper. Analysis of the COSMOS data was carried out on the IBM 360 computer at the Rutherford Laboratory. The galaxy simulations were performed on the ‘Starlink’ VAX 11/780 computer at the Royal Observatory Edinburgh. We are grateful to Mr R. Martin for assistance in transferring the simulation program to the VAX. PCH acknowledges the support of an SRC studentship while this work was being carried out. The comments of an anonymous referee are also acknowledged.

References

Galaxies near the South Galactic Pole – II

1981MNRAS.195..613H