The Muenster Redshift Project. Automated Analysis of Galaxy Clustering on Schmidt Plates

H. Horstmann
Astronomisches Institut
Westfälische Wilhelms-Universität
Münster, F.R. Germany

Abstract

The Muenster Redshift Project (MRSP) collects information about the three-dimensional distribution of galaxies by processing pairs of direct and objective prism Schmidt plates. This contribution describes the role of direct plates in the survey. Methods of analysis and their errors are discussed. Maps of two-dimensional galaxy distributions are presented, including the distributions of different types of galaxies (using ellipticity as a coarse morphological criterion). We present preliminary results for a region near the South Galactic Pole, containing several clusters of galaxies, of which five form a physical group at redshift $z = 0.11$.

1 Introduction

The basic goal of the MRSP is to obtain information on the three-dimensional distribution of galaxies and the properties of matter and space which can be derived from it. The basic data are obtained by processing automatically pairs of direct and objective prism Schmidt plates. The material consists of ESO/SRC atlas plates (film copies of IIIa-J plates) and film copies of IIIa-J objective prism plates with a dispersion of $246\,\mathrm{nm\,mm^{-1}}$ at Hγ. Both kinds of plates are taken with the UK Schmidt telescope.

The faint limiting magnitudes ($21\overset{\mathrm{m}}{.}5$ on the direct and $20\overset{\mathrm{m}}{.}5$ on the prism plates) and the large area of the sky covered, even by a single plate, enables us to study extended and relatively distant structures in the distribution of galaxies. The large numbers of objects detected, typically 150 000 on the direct and 50 000 on the objective prism plates at high galactic latitudes require fully automatic reduction procedures on the level of single objects.

The processing of objective prism plates, in particular automatic redshift measurements, are described in a paper by Schuecker (1988). The present contribution describes the reduction of the direct plates and, in particular, the study of internal properties of clusters of galaxies based on these data.

2 The role of direct plates in the MRSP

2.1 Service functions of the direct plates

The primary function of the direct plates is to supply information which can be obtained more easily and/or with higher precision from direct images. For example, the **segmentation** (search for objects) can cause substantial problems when executed on the objective prism plate, especially in regions of high object density (clusters of galaxies), because the relatively long spectra often overlap. Even if these overlaps are very slight, it is difficult to recognize them properly on the objective prism plates. Other examples are the object positions in R.A. and Decl., star/galaxy separation, and measurement of apparent magnitude.

The transport of information from the direct to the objective prism plate is possible after coordinate transformations have been applied, using 12 to 15 reference stars. Object positions on the objective prism plate are predicted with an accuracy of $1''$ to $2''$ over the whole plate, which is sufficiently precise for object identification.

A more sophisticated use of the direct plate is made in the determination of wavelength zero points for the objective prism spectra. A widely used procedure takes the emulsion cutoff as wavelength reference. This, however, introduces large systematic errors due to the brightness and colour dependence of the cut-off position. In the MRSP a high accuracy plate transformation is employed, using approximately 1 000 reference stars and a higher order transformation model. This procedure is described in detail in the paper of Tucholke (1988).

2.2 Deeper surveys from direct plates

There is, of course, more information available on the direct plates than is needed for processing the objective prism plates: the direct plates reach fainter objects and supply (coarse) morphological parameters of the galaxies. When deep direct plate studies are applied to clusters of galaxies – whose distances are measured on the objective prism plates – shapes, orientations, density profiles, morphological contents and luminosity functions of the clusters can by derived, while the redshift data help to disentangle clusters superimposed on each other along the line of sight.

3 Plate digitization and segmentation

3.1 Digitization

The plates are scanned with the PDS 2020 GMplus microdensitometer at the Astronomical Institute Muenster (AIM). The central $300 \times 300\,\mathrm{mm}^2$ of the plates are digitized with a step size of $15\,\mu m$ or about $1''$. This corresponds to an area of $5.5 \times 5.5\,\mathrm{deg}^2$ on the sky. The scan time for a whole plate is presently about 20 hours, if the maximum mechanical speed of the machine is used. The density which is measured by the PDS ranges from density 0.0 to 4.5, with a resolution of 0.00125 over the whole measurable range.

The search for objects is performed online. An *object* is defined as a group of connected pixels (6 pixels in the case of the ESO/SRC atlas plates) which exceed a minimum density threshold above sky background. The latter is determined by applying a strong median filter to a scanline (or, in practice, to each 10th line, since the sky background is varies very slowly) the significance level is defined at $3\,\sigma$ above sky background. The **search** process leads to typically 150 000 objects in a high galactic latitude field, varying between 125 000 and 195 000 for the 9 plates processed so far. During the online search, only a few *object parameters* are computed: object position, obtained by density weighted object moments, area covered by the object (in pixels), extent of the object in x and y, local plate density and local standard deviation of the plate density.

3.2 Segmentation

Together with the object parameters, the images themselves are stored permanently in the form of quadratic scan segments, $21 \times 21\,\mathrm{arcsec}^2$, called **picture frames.** The objects found on one plate can be saved on one to several high density tapes. The permanent storage permits direct access to an individual star or galaxy image (not only to a list of image parameters). This makes subsequent processing possible without going back to the plate. All classification processes become very fast by simply using the digitized images.

The following image analysis, including the computation of effective radius, central intensity, ellipticity, orientation, apparent magnitude and other parameters, uses these picture frames. It is obvious that the limited size of the frames leads to systematic errors for the brighter objects whose outer parts are lost. For galaxies these errors arise at about $16^{\mathrm{m}}.5$ or brighter, depending also on morphological type. The most important error is the magnitude error. Because of this, bright objects must either be excluded from the survey or looked at with larger picture frames.

Generally, the restrictions can be tolerated, mainly for two reasons. The first is that galaxies brighter than about 17^{m} are rare objects compared with the total number of galaxies found in deep surveys. The second reason is, that we cannot measure redshifts for these extended objects with the slitless technique. The bright dominant galaxies in the cluster centers are treated interactively during investigation of the cluster morphology (see Sect. 6).

4 Object classification

4.1 Star/Galaxy separation process

In the following, a brief description of the **star/galaxy classification** algorithm is given. For information about complications due to instabilities of the plate parameters see also Horstmann (1988).

For star/galaxy separation the difference in the *peak density/effective radius* relations for stellar and non-stellar images are used. *Fig. 1* shows a diagram with several thousand objects detected on an ESO/SRC atlas plate. The classification *stellar/non-*

stellar image is performed by defining a discriminating function along the right hand side of the stellar image strip. This curve (spline interpolation) is established automatically by fitting Gauss curves to the stellar density distributions at several positions of the stellar curve (see *Fig. 2*).

4.1.1 Field dependent star/galaxy separation

The curve separating stars from galaxies cannot be assumed constant over the whole plate. The vertical as well as the horizontal part of the stellar curve in *Fig. 1* changes slightly with position on the plate. Reasons for this are primarily the variations of the plate background density which in turn is caused mainly by desensitization during plate exposure (Malin 1983, Dawe *et al.* 1984). Some maps of the background density are shown in Horstmann (1988). Other reasons may be coordinate-dependent

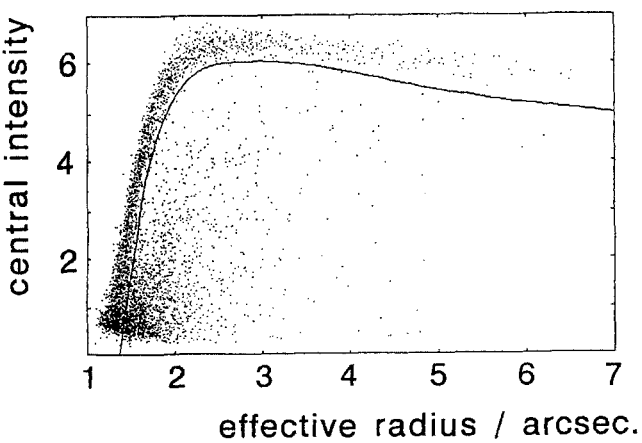

Fig. 1. Central intensity vs. effective radius for several thousand objects from the ESO/SRC field No. 411. The solid line separates stars from galaxies.

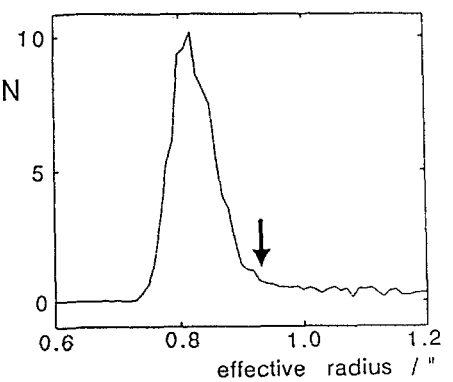

Fig. 2. Tracing through the stellar distribution in *Fig. 1* at about half the maximal intensity. The arrow indicates the 3σ point which defines the location of the star/galaxy separating function.

defocussing during the exposure of the plates and/or during scanning. We correct for this effect by dividing the plate into 25 fields. For each of these fields the discriminating function is determined separately with the methods described above. For each position between the centers of these reference fields, the functions are interpolated.

4.1.2 The discriminating curve

Figure 2 shows a tracing across the stellar curve at about half maximum intensity. The location for the fitting points is chosen at the 3σ level obtained from the Gauss fit and is indicated by an arrow in *Fig. 2.* If the distribution is Gaussian, this position leads to a misclassification rate for stars of about 0.1 %. As the plot shows, the tracing is not strictly Gaussian, in particular, the object density on the right side is significantly larger, making the stellar distribution asymmetric. Nevertheless, from inspection by eye we have obtained an error rate for objects brighter than about 20^{m} of no more than 2 %.

For fainter objects star/galaxy separation becomes more difficult. Stellar and galaxy regions overlap and it is hard to establish a discriminating function which optimally separates stars from galaxies. A reference sample cannot be obtained from classification by eye, because the personal errors become very large at faint magnitudes and strongly influence the result. The discriminating function is thus extrapolated into the region of the faintest objects. We expect that the classification error up to the completeness limit of a given plate, typically 21^{m}, is never larger than a few %. This is also supported by the two point correlation functions for the star and galaxy distributions (Sect. 4.2).

4.1.3 Recognition of double stars

After distinguishing between stellar and non-stellar images, all *non-galaxies* among the non-stellar images must be eliminated. They are blends of star and galaxy images and plate flaws, dust *etc.* For a non-disturbed galaxy we expect, in the majority of cases, that the position of the density center does not depend on the density- or intensity-threshold used in image analysis, and that the maximum density is reached at the position of the density center. By testing the rejection parameters interactively, we have obtained $0''.75$ as a suitable maximum for the permitted deviation between object centers at various density levels. For a undisturbed galaxy image we expect that the maximum image intensity does not exceed the central intensity by more than 15 %.

The rejection rate as a function of apparent brightness is shown in *Fig. 3.* The increase of the rejection rate with decreasing object brightness follows from the fact that an object can be influenced only significantly by objects of comparable brightness, and that therefore the probability of being disturbed by a neighbour increases for faint images. For objects fainter than 22^{m} (most of which are nothing but plate noise), the rejection rate increases steeply to 100 %.

It is apparent from the description of the classification algorithm, that it cannot distinguish between stars and galaxies when the images are saturated. The apparent

magnitude at which this occurs depends for galaxies on the morphology of the object; but normally only galaxies brighter than 15^m to 16^m will reach maximum plate density. For these objects we have developed a method which makes use of the fact that on Schmidt plates bright stellar images show spikes, while galaxy images do not. For the data described in the following, this classifier was not employed, since the images were already excluded because they exceed the limit of the picture frames.

4.2 Results from the star/galaxy separation

As one result of the star/galaxy separation process, *Fig. 4* shows the two-point angular correlation functions $w(\theta)$ for stars and galaxies up to the limiting magnitude 21^m. We have counted stars and galaxies in cells of $3'.3 \times 3'.3$ and used the estimator (Peebles 1975, Hewett 1982)

$$w(\theta) = \frac{\langle n_1 n_2 \rangle}{\langle n_1 n_2 \rangle} - 1 \,, \tag{1}$$

where the n_1, n_2 are cell counts and $\langle\ \rangle$ denotes averaging over all pairs n_1 n_2 with cell separation θ. The functions indicate the expected null correlation for a purely stellar sample and the expected correlation on the scale of 1° for the galaxies. All data are taken from field No. 411.

5 Morphological classification of galaxies

Since the work of Dressler (1980), Tully and Shaya (1984) and others, the distribution of different morphological types of galaxies has become of great interest. The MRSP data seem well suited for such studies, because they are complete up to faint magnitudes (at least to $20^m.5$) and because large areas of the sky are covered. Additionally, the automated classification has the advantage of being free of any personal bias. On the other hand the problem arises of how to obtain morphological information at apparent magnitudes around 20^m. Classification using object colours is possible, but it requires the scanning of additional direct plates in different spectral regions. They

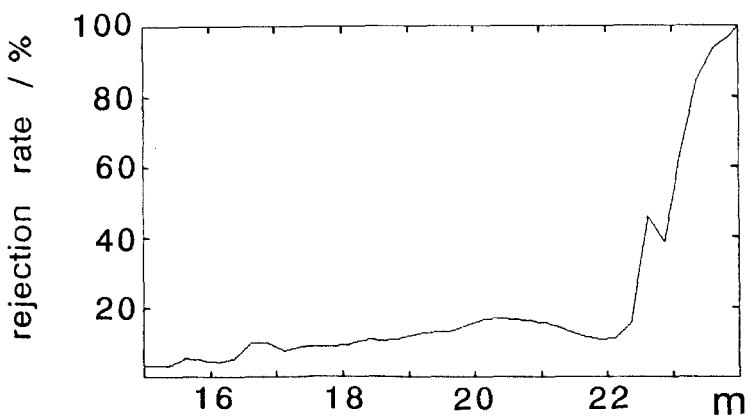

Fig. 3. Object rejection rate as a function of apparent magnitude. See text for details.

are not available at the moment (even the ESO red survey is still very incomplete). Colour measurements from the objective prism spectra cannot be used because the blue parts of the spectra are missing at these faint magnitudes. We have attempted to perform a classification using intensity profiles, but it appeared that these tracings, with a length of typically four or five pixels, do not contain enough information.

The quantity finally used for the classification is the *apparent ellipticity* of the galaxy images. The basic idea is that the images of elliptical galaxies seldom show extreme axial ratios. Ellipticities larger than about 0.5 occur for much less than 10% of all E type galaxies (Sandage *et al.* 1970, Schechter 1987 and references therein). The images of spiral galaxies, on the other hand, show high axial ratios when seen more or less edge-on. Ellipticities larger than 0.5 should occur for about 30% of all images of S type galaxies. Thus, the content of objects with large ellipticities should be an indicatior of the morphological content of a given cluster of galaxies or of a given region on the sky.

Because the classification is statistical, one cannot decide for a single object whether it is an E- or S-type galaxy. Additionally, some problems arise because fainter objects appear to have systematically lower ellipticities, as is shown in *Fig. 5*. This is due to seeing effects, which are more severe for faint objects where they tend to reduce the ellipticity. Also, experience shows that objects consisting of very few pixels tend to be 'rounder' than brighter and larger images, partially due to aperture effects from the scans.

Ellipticities are computed as

$$E = \frac{\sqrt{(M_{2,0} - M_{0,2})^2 + 4M_{1,1}^2}}{M_{0,2} + M_{2,0}} \tag{2}$$

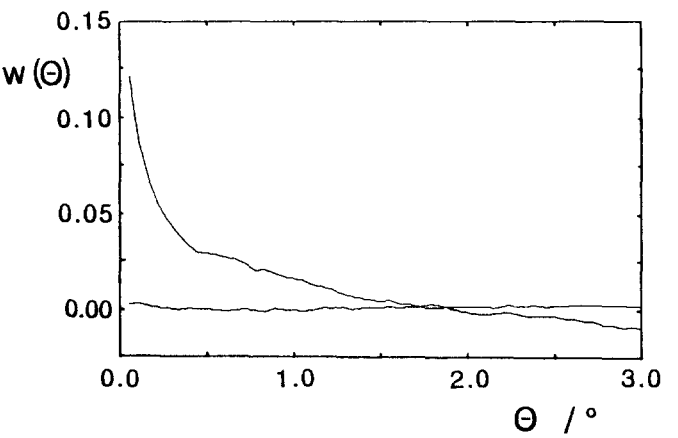

Fig. 4. Two-point angular correlation function for all galaxies (upper curve) and stars up to 21^{m} from field No. 411. Note that the stellar images do not show any significant clustering, as is expected in the direction of the galactic pole.

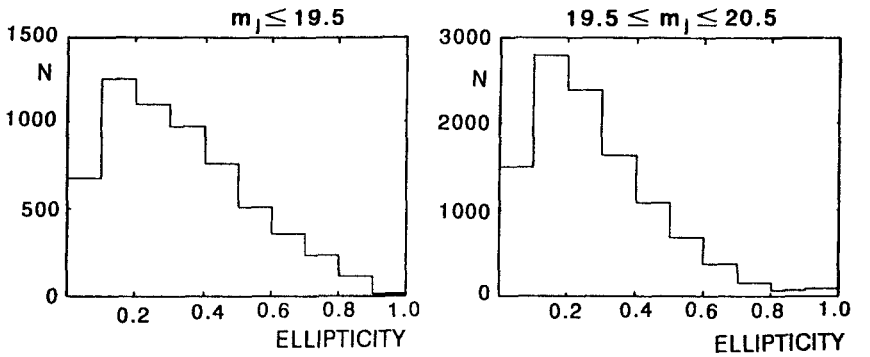

Fig. 5. Apparent ellipticity histograms for galaxies in field No. 411. Left: Galaxies brighter than $19^{\mathrm{m}}.5$. Right: Galaxies between $19^{\mathrm{m}}.5$ and $20^{\mathrm{m}}.5$. The 'rounder' images at fainter magnitudes are due to seeing and digitization effects.

with the moments $M_{i,j}$ defined as

$$M_{i,j} = \sum_{A} (x - x_c)^i \, (y - y_c)^j \, I(x, y) \tag{3}$$

where A is the summation over all pixels of a given object, $I(x, y)$ the intensity above the local sky background intensity, and x_c and y_c are the coordinates of the object center.

The equations from which the ellipticity is derived include contributions from all intensity levels; the axial ratios are not determined at a fixed intensity level only. There is a noticeable lack of objects in the interval 0.0 to 0.1 in *Fig. 5*, which may be an artefact, although this has also been found by other authors (*e.g.* Benacchio and Galetta 1980), and, recently, by Davies *et al.* (1988), who find a similar deficit of 'round' systems for dwarf ellipticals in the Fornax Cluster region.

6 Results

We have applied the methods described above to a region near the south galactic pole. In the following, first results are given.

Figure 6 shows a map of the galaxy distribution in three adjacent ESO/SRC fields for all galaxies brighter than $20^{\mathrm{m}}.5$. The total number of galaxies is about 60 000. The most striking feature on this map is the group of 5 rich clusters in the central field, marked A, B, C, D and E. From our redshift measurements (Schuecker 1988) we have found a distance of $z = 0.11$ for all five clusters, indicating that they are physically connected, probably forming the dense nucleus of a supercluster.

Histograms of the ellipticities of galaxies in the five clusters are given in *Fig. 7*. Each histogram contains the counts of all galaxies brighter than $20^{\mathrm{m}}.0$ within a radius of 15' around the center of the corresponding cluster. As the histograms show, the counts at higher ellipticity levels are significantly lower than one would expect from the histograms of all galaxies from the whole plate. With our interpretation of ellipticities

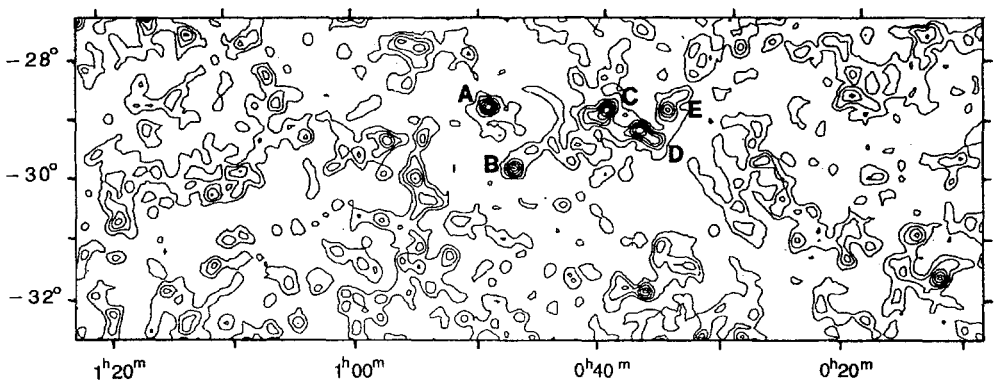

Fig. 6. Maps of the galaxy distribution obtained from three fields, Nos. 410, 411 and 412 (from right to left), near the South Galactic Pole. The isopleths correspond to 1.1, 1.4, 1.7... times the mean galaxy backround density. The five clusters A to E are at the same distance and probably form the nucleus of a galaxy supercluster. See the text for details.

this means that compared to the field *spiral galaxies are underrepresented in rich clusters*, in agreement with the results of Dressler (1980) and others.

Our result is also supported by the correlation functions for objects in different ellipticity intervals. *Fig. 8* shows plots of the two-point angular correlation functions for the three fields of *Fig. 6*, for galaxies with ellipticities below and above 0.4 separately. For fields 412 and 410, both showing few prominent clusters, the two curves are nearly identical. For the central field 411, containing the five large clusters, the low ellipticity galaxies have significantly stronger correlations than those of higher ellipticity.

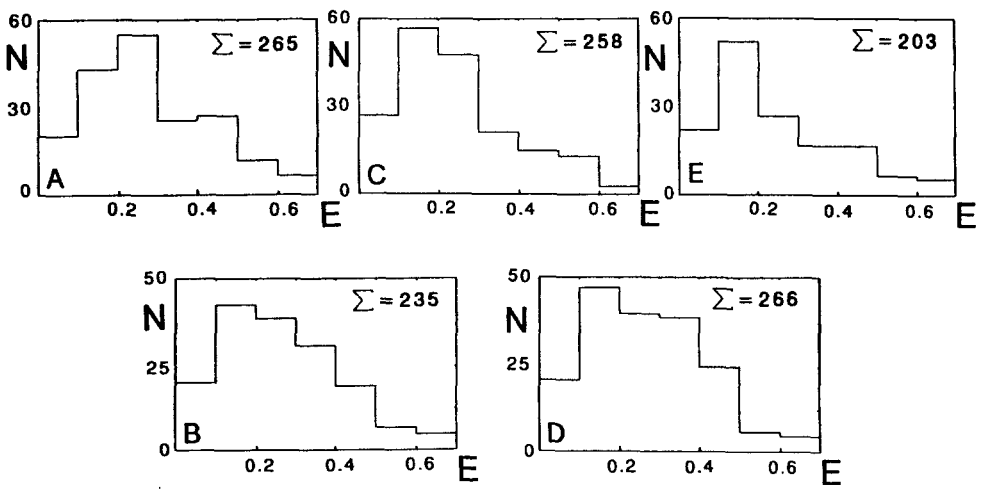

Fig. 7. Histograms of galaxy ellipticities for the five clusters marked in *Fig. 6*.

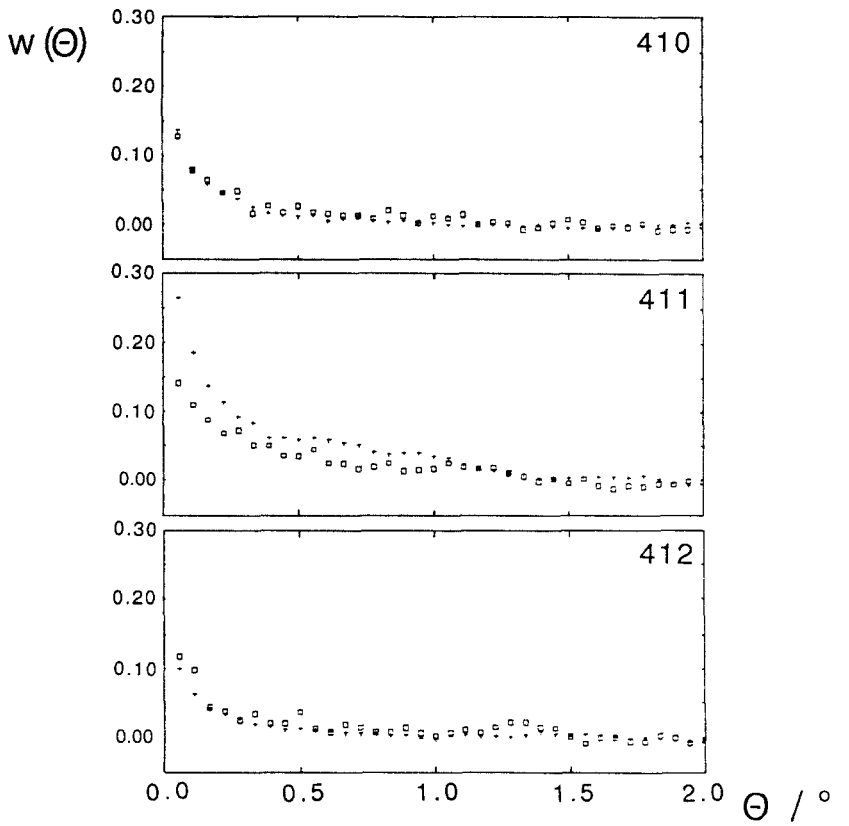

Fig. 8. Two-point angular correlation functions for galaxies brighter than $20^{\mathrm{m}}.0$ from ESO/SRC fields 410, 411 and 412. In field 411, containing five prominent clusters, galaxies with apparent ellipticities less than 0.4 (crosses) show stronger clustering than those of higher ellipticity (rectangles). In the two other fields, containing few prominent clusters, the correlation functions for both populations show no significant differences.

References

Benacchio, L., Galetta, G., 1980. *Mon. Not. R. astr. Soc.*, **193**, 885.

Davies, J.I., Phillips, S., Cawson, M.G.M., Disney, M.J., Kibblewhite, E.J., 1988. *Mon. Not. R. astr. Soc.*, **232**, 239.

Dawe, J.A., Coyte, E., Metcalfe, N., 1984. In *Astronomical Photography*, Occasional Reports Royal Obs. Edinb., **14**, eds. Sim, M.E., Ishida, K., p. 59.

Dressler, A., 1980. *Astrophys. J.*, **236**, 351.

Hewett, P.C., 1982. *Mon. Not. R. astr. Soc.*, **201**, 867.

Horstmann, H., 1988. In *Proc. IAU Workshop Astrophotography*, ed. Marx, S., Springer, Berlin, p. 186.

Malin, D.F., 1983. In *IAU Coll. No. 78, Astronomy with Schmidt–Type Telescopes*, ed. Capaccioli, M., Reidel, Dordrecht, p. 57.
Peebles, P.J.E., 1975. *Astrophys. J.*, **196**, 647.
Sandage, A., Freeman, K.C., Stokes, N.R., 1970. *Astrophys. J.*, **160**, 831.
Schechter, P.L., 1987. In *IAU Symposium No. 127, Structure and Dynamics of Elliptical Galaxies*, ed. de Zeeuw, T., Reidel, Dordrecht, p. 217.
Schuecker, P., 1988. *These proceedings*, p. 142.
Tucholke, H.J., 1988. *These proceedings*, p. 136.