THE APM GALAXY SURVEY

S.J. Maddox, G. Efstathiou and J. Loveday
Institute of Astronomy,
Cambridge, England.

1. INTRODUCTION

Wide-field surveys of galaxies and clusters are an indispensable tool for studying large scale structure in the universe. The Abell catalogue (Abell 1958), Zwicky catalogue (Zwicky *et. al.* 1961-1968), and the Lick survey (Shane and Wirtanen 1967, Seldner *et. al.* 1977) have provided many statistical results of key importance to our understanding of galaxy formation and clustering (see *e.g.* Peebles 1980). However, these surveys were constructed more than 20 years ago. Since then, there have been major technological developments in photographic emulsions, automatic scanning machines and computers. It is therefore possible to improve significantly on earlier surveys by generating deep galaxy catalogues with high photometric precision and uniformity over wide areas of sky. Over the last four years, we have taken advantage of these developments to construct a new survey of several million galaxies.

The photographic material that we have used is based on the $6° \times 6°$ SERC IIIa-J plates taken by the UK Schmidt Telescope Unit (UKSTU). The UKSTU plates cover the entire southern sky south of $\delta = -20°$. The limiting magnitude corresponds to $b_J \approx 21$, more than 2 magnitudes deeper than the Lick survey. The European Southern Observatory (ESO) has supplied us with high quality glass copies of the original plates (West 1978) which we have scanned in Cambridge.

The SERC Automatic Plate Measuring (APM) machine in Cambridge is a high speed automated densitometer with on-line sky subtraction and image analysis (Kibblewhite *et. al.* 1984). It can scan the central $5.8° \times 5.8°$ of a UKSTU plate in about 14 hours giving accurate measurements of the position, magnitude and shape parameters for each image above a fixed detection threshold.

We have used the APM machine to scan 176 of the 190 UKSTU fields within the area $b < -40°$ and $\delta < -20°$. The remaining 14 plates have yet to be supplied by ESO. Over the 4400 square degrees covered by our survey, we have detected about 3.6 million galaxies at a magnitude limit of $b_J = 21$. Figure 1 shows the field boundaries in an equal area projection centred on $\alpha = 15°, \delta = -45°$. For statistical studies of the large-scale structure, it is important that the galaxy selection function is uniform over the whole survey. We have therefore used a combination of photometric CCD calibrations and comparisons between plate overlaps to ensure that the magnitude limit and star-galaxy classification are uniform. The positions of the calibration sequences are shown by the dots in Figure 1.

151

Figure 1. Equal area projection of the APM survey fields. The UKSTU field numbers are indicated. The dots show the positions of our CCD sequences.

2. THE APM MEASUREMENTS

A typical APM scan records about $200,000$ images on each plate. The machine measures an isophotal intensity for every image, with the threshold set to be $\approx 8\%$ of the sky background. The threshold corresponds to $\approx 25\,\mathrm{b_J}$ mags arcsec^{-2}. Multiple scans of the same plate show that the measured magnitudes are repeatable to 5%, even for images close to the plate limit. However, comparisons between different plates of the same field show that variations in sensitivity and grain noise lead to large magnitude errors for images fainter than $\mathrm{b_J} \approx 21$. For the final catalogue we have therefore aimed at uniformity for objects brighter than $\mathrm{b_J} = 20.5$. At this limit we find about 17000 galaxies per plate.

The star-galaxy separation technique locates the stellar loci in several parameters and uses maximum likelihood to give an estimate of the classification for each object. Figure 2 shows the parameters that we have used as measured on a typical plate. The scatter plots are of magnitude against area, magnitude against mean size, peak surface brightness against mean size and (area/size) against magnitude. The first three plots provide different measures of surface brightness. The stellar loci are well defined, with the galaxies forming a distinguishable group of images. The fourth plot is designed to pick out merged pairs of objects for which the area should be small in comparison to the mean radius. Recently we have developed a more accurate method of star-galaxy classification which uses the surface brightness profiles in conjunction with the moments of each image, but this has not yet been applied to every field in the survey.

We have compared the automatic classifications with our own visual checks from the plates. The star-galaxy separation algorithm agrees with our visual classifications for $\approx 90\%$ of the galaxies in the range $17.5 < \mathrm{b_J} < 20.5$. For the fainter images there is not enough information on the plates for either us or the machine to distinguish reliably between stars and galaxies. The classification algorithm becomes inaccurate for bright objects because stellar halos and the saturation of the IIIa-J emulsion lead

to similar image parameters for bright stars and galaxies. To overcome this problem we are examining bright objects visually with the aim of achieving a 100% complete sample of galaxies. During the checking we are also noting a rough morphological classification. The distribution of galaxies brighter than $b_J = 16$ is shown in Figure 3. This is the current limit of our eyeballed sample, set by the shallowest plate in the survey. For most fields, the bright sample is now complete to $b_J = 16.5$.
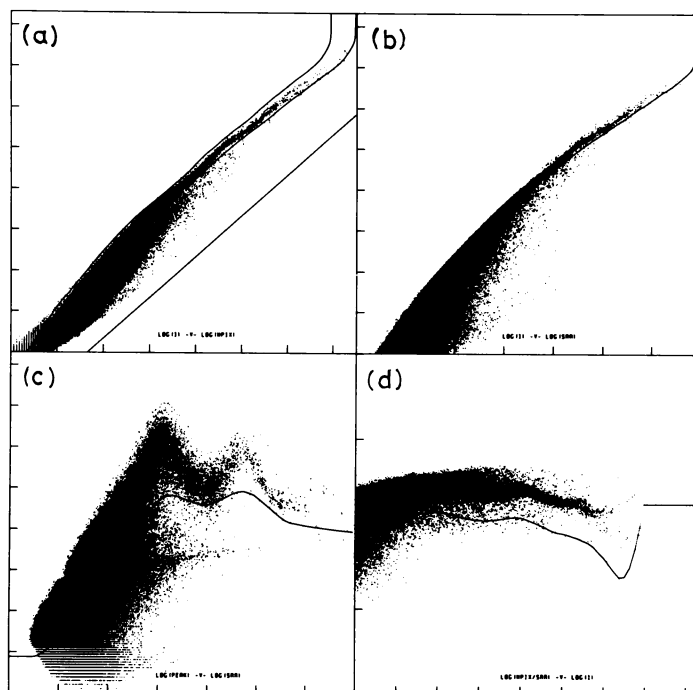

Figure 2. Parameters used in separating stars from galaxies. (a) shows area v magnitude, (b) mean size v magnitude, (c) size v peak surface brightness, (d) magnitude v area/size. Points between the curved lines in (a) and above the curved lines in (b) and (c) are classified as stars. Points below the line in (d) are classified as merged images. The straight line in (a) shows the minimum possible area for a given magnitude.

The *rms* positional errors of objects in the plate overlaps are only $\approx 0.5''$. We also find small distortions over the $5.8° \times 5.8°$ field which give rise to systematic errors of $\approx 1.5'' - 2''$. For studies of galaxy clustering these positional errors are insignificant. The galaxy positions are well within the tolerances required for multi-object fibre spectroscopy at large telescopes. For example, Colless and Hewett (1987) have used the APM survey to generate fibre masks at the Anglo-Australian Telescope for their study of the dynamics of rich clusters.

Figure 4 shows an isopleth map in the same projection as Figure 1 for about $4,000,000$ galaxies before making any corrections for uniformity. The magnitude limits and star-galaxy separation accuracy therefore vary between plates reflecting differences in plate quality and detection threshold. These variations can be seen as changes in the number density of galaxies across the field boundaries. However,
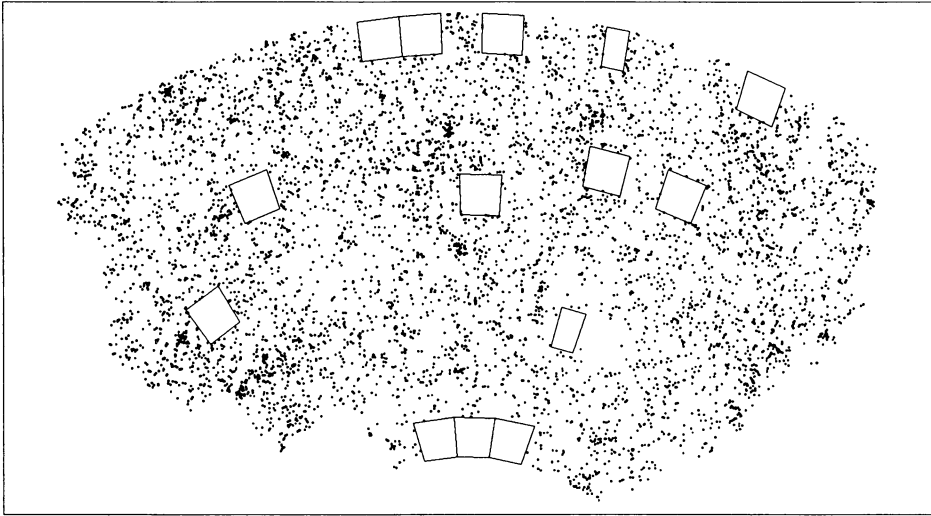
Figure 3. Distribution of bright galaxies ($b_J \leq 16$) in the equal area projection of Figure 1. The boxes show the boundaries of areas which have not yet been processed.
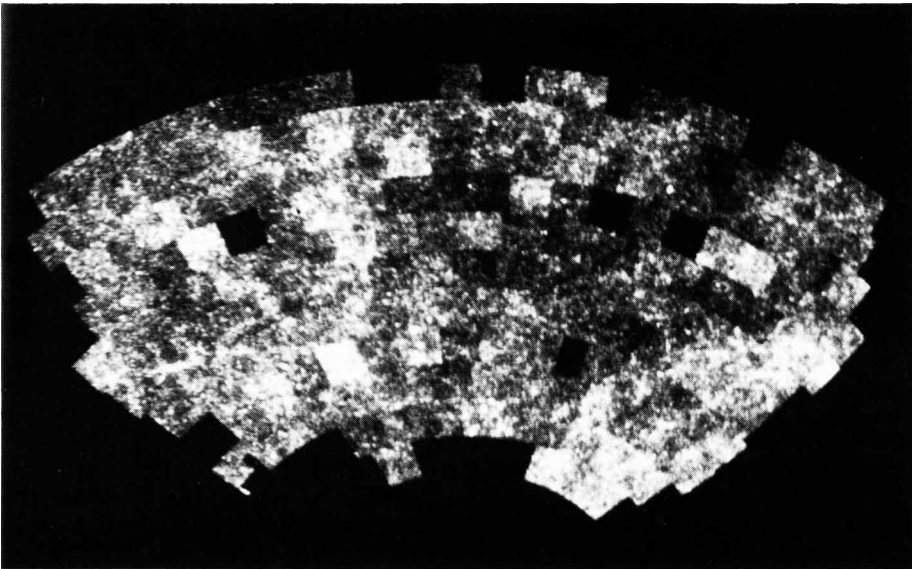
Figure 4. Equal area projection showing the density of galaxies in the unmatched APM survey scans at a magnitude limit of $b_J \approx 21$.

even with no corrections interesting features in the galaxy distribution may be traced across plate boundaries.

3. PLATE OVERLAPS AND MATCHING

The field centres are separated by 5° so there is a generous overlap area of about 6° × 1° between each plate. The parameters of matched pairs of images in the overlaps provide our primary means of ensuring uniformity in the selection function over the survey area.

A good demonstration of our photometric accuracy and the extent to which we can remove systematic field effects is provided by plate pairs near $\alpha = 0$. Unlike the majority of plates in the survey, some of the plates at $\alpha = 0$ have large overlaps of up to 3° × 6°. Thus we can compare measurements made at one plate centre with those made at another plate edge. Any systematic errors in the magnitudes on each plate will show up as a positional variation in the mean difference between the two sets of measurements. Figure 5 shows a contour map of the average difference between the magnitudes of images measured from one plate at $\alpha = 0$ and its neighbour. The differences over most of the area are less than 0.04 magnitudes peak-to-peak.

Figure 5. Contour map illustrating magnitude differences between two plates with a large overlap (fields 292 and 241, see Figure 1). The x and y scales are marked in degrees. The countours are spaced at 0.02 magnitude intervals according to the following key:

——————— 0.02
——————— 0.00
· · · · · · · · · −0.02
··· ············ −0.04
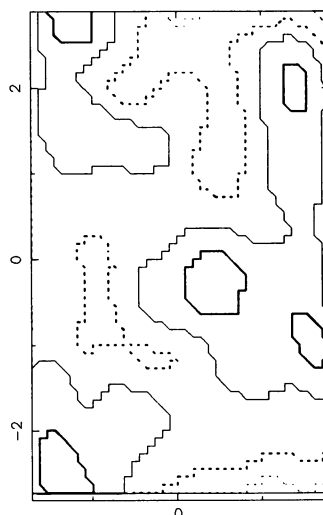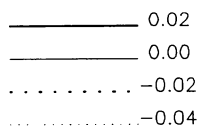
We use an algorithm similar to that applied to the Lick counts by Seldner *et. al.* (1977) to determine corrections to the magnitudes from the plate overlaps. A polynomial fit to the magnitude-magnitude plot for each overlap is used to give a conversion between the magnitudes from each plate to those from its neighbours. We then apply an iterative algorithm to find the set of field corrections which is the most consistent with the measured overlap conversions. Matching the magnitudes in this way gives a residual scatter of 0.02 magnitudes in the zero point of each plate.

The edge matching procedure is sensitive to residual field effects which introduce small systematic drifts in the plate zero points. Without any additional checks, these would cause large-scale gradients in the survey. To prevent this evidently undesirable effect, we have constructed 66 faint photometric galaxy sequences using CCD frames in B and V taken with the 1.0m telescope at the South African Astronomical Observatory. These are used as tie points in the plate matching algorithm.

156

At present, the main residual source of error in our matched maps is caused by variations in star-galaxy separation. Over the survey area the stellar density varies by a factor of ~ 3 from the galactic pole to low galactic latitudes. Most of the residual variation is caused by an increase in stellar contamination in the directions of the galactic centre and anticentre. We have not yet applied an algorithm to maintain uniformity over the whole survey, though this will be done shortly. Nevertheless, the star-galaxy separation is quite accurately uniform over large areas. An example is given in Figure 6 which shows an area of about $26° \times 26°$. Notice that the boundaries between fields are not visible in this map (*cf.* Figure 4).
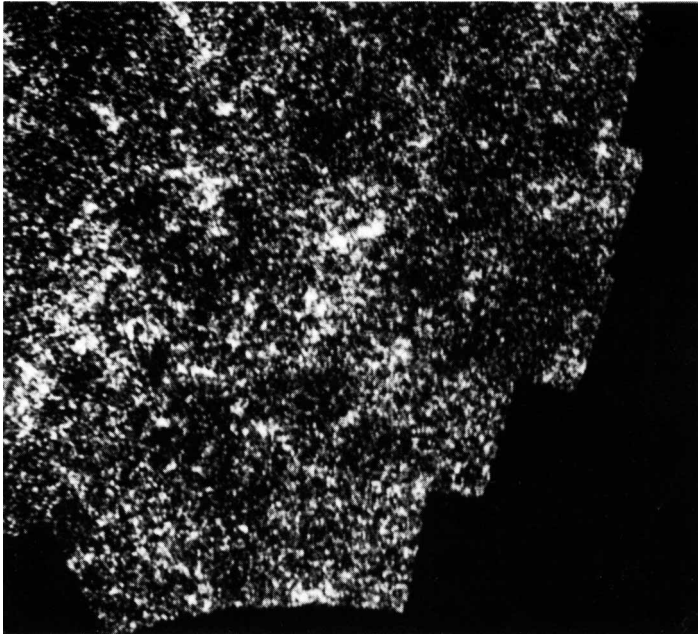
Figure 6. The galaxy distribution at $b_J = 20$ in a $26° \times 26°$ area of the APM survey. The plates have been matched as described in Section 3.

4. THE ANGULAR TWO-POINT CORRELATION FUNCTION

One of our main aims in generating the APM survey has been to determine the two-point galaxy correlation function $\xi(r)$ at large-scales. This may be accomplished by measuring the angular two-point correlation function $w(\vartheta)$ which is related to the spatial correlation function by Limber's formula (Limber 1953, Peebles 1973). Angular catalogues such as the Lick and APM surveys contain vast numbers of galaxies, well in excess of the size of any conceivable redshift survey. Even though the amplitude of $w(\vartheta)$ is small, the statistical power of angular catalogues may be used to trace correlations to large spatial scales provided that systematic and random errors can be kept to acceptable levels.

Groth and Peebles (1977) have measured $w(\vartheta)$ for the Lick sample and find

a break from the power law $w(\vartheta) \propto \vartheta^{-0.74}$ at an angular scale of $\vartheta \approx 3°$. The feature corresponds to a spatial scale of $\approx 10h^{-1}\text{Mpc}$, where $\xi(r) \approx 0.3$. As Groth and Peebles remark, the break from a power law occurs interestingly close to the expected transition between non-linear and linear density fluctuations. The shape of $\xi(r)$ beyond the Groth and Peebles break point could therefore set important constraints on theories of primordial density irregularities.

The reality of this feature has been challenged by Geller *et. al.* (1984) and de Lapparent *et. al.* (1986). It has been vigorously defended by Groth and Peebles (1986a,b). In our view, the the excellent agreement that Groth and Peebles find between intra-plate and inter-plate estimates of $w(\vartheta)$ represents the most significant piece of evidence in support of their result. Intra- and inter-plate estimates are subject to different types of error; for example, random plate matching errors would produce a positive offset in the intra-plate estimates (corrected for by Groth and Peebles, 1977), but would average out for inter-plate estimates over the survey area. It is possible to imagine correlated errors which could produce a break in the inter-plate w, but it is difficult to see how these could lower the intra-plate estimates. However, the plate-matching algorithm applied by Seldner *et. al.* (1977) to the Lick counts produces large scale gradients in the corrected counts. The position of the break, and the slope of $w(\vartheta)$ beyond the break point depends on the filtering applied to remove these large scale features. This is probably the largest source of systematic error in the analysis of the Lick catalogue.

In this Section we present the first results on large scale clustering in the APM survey. The conclusions are preliminary since we have not yet analysed the entire contiguous region shown in Figure 1. A more detailed discussion, together with an analysis of our systematic errors, will be published elsewhere.

We estimate the angular function $w(\vartheta)$ using two methods. On small scales, $\vartheta < 0.5°$, we apply the estimator

$$w(\vartheta) = F\frac{\text{DD}}{\text{DR}} - 1, \qquad (1)$$

where DD is the number of data pairs with separations in the range $\vartheta \pm \delta\vartheta$. This count is normalized by the term DR which is the pair count determined by cross-correlating the data points with a set of points distributed at random within the survey boundary with F times the mean density of the galaxies. This corrects in a straightforward way for the complex shape of the boundaries and for areas which we have excluded around bright stars *etc.*. At large angular scales, we group the data into cells of count n_i and use the estimator

$$w(\vartheta) = \frac{\langle n_1 n_2 \rangle}{\langle n_1 \rangle \langle n_2 \rangle} - 1 \qquad (2)$$

(Peebles 1975, Hewett, 1982).

Estimates of $w(\vartheta)$ at a magnitude limit of $b_J = 20$ for our survey are shown in Figure 7. The crosses in Figure 7a show the average of $w(\vartheta)$ determined from the central $5° \times 5°$ region on each of 166 fields. The error bars show the variance in the mean deduced from the scatter between plates. The results for each plate have been fitted to the power law $w(\vartheta) \propto \vartheta^{1-\gamma}$ by least-squares. The mean and standard deviation of γ over the 166 fields are

$$\gamma = 1.660 \pm 0.014, \quad \text{for } 0.01° \le \vartheta \le 0.1°. \qquad (3)$$

A least-squares power law fit to the mean $w(\vartheta)$ shown in Figure 7a gives

$$w(\vartheta) = 2.89 \times 10^{-2}/\vartheta^{0.676}, \quad \text{for } 0.01° \le \vartheta \le 0.1°. \qquad (4)$$

All of our estimates of $w(\vartheta)$ have been multiplied by a factor of 1.29 to correct for residual stellar contamination. At small angular scales ($\lesssim 20''$) the amplitude of $w(\vartheta)$ is low because close pairs of images are sometimes merged in our scans. The slopes in (3) and (4) are somewhat shallower than the value $\gamma = 1.77 \pm 0.06$ inferred by Peebles (1975) from the Zwicky, Lick and Jagellonian catalogues. The effective depth of our survey at $b_J = 20$ is $D^*_{eff} \approx 570 h^{-1} Mpc$ ($H_o = 100 h\, km/s/Mpc$), thus the above estimates of γ refer to spatial scales $0.1 h^{-1} Mpc \lesssim r \lesssim 1 h^{-1} Mpc$. Only the Zwicky catalogue provides a substantial overlap over this range of scales and it is our impression that the correlation functions are consistent within the errors.
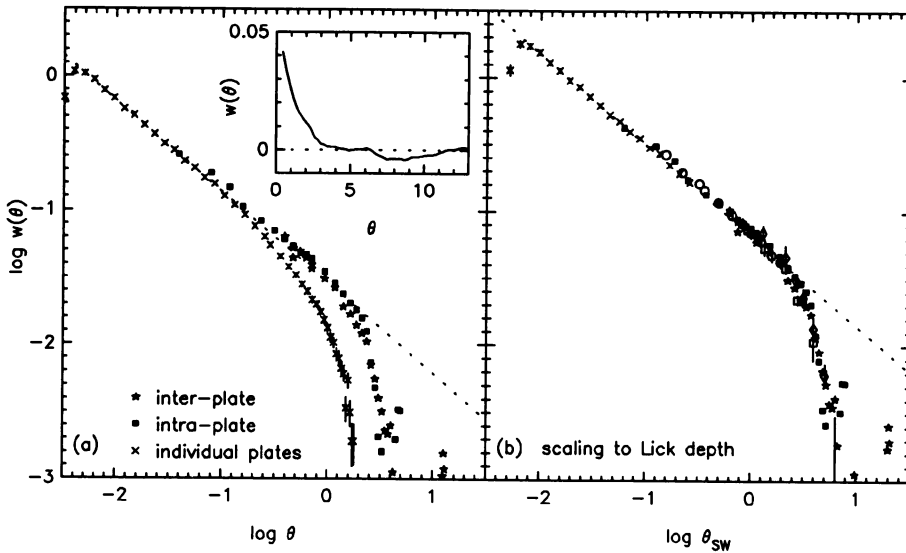
Figure 7. Estimates of the angular two-point correlation function from the APM survey at $b_J = 20$. Crosses show the mean value of $w(\vartheta)$ derived from single plates. The filled symbols show intra- and inter-plate estimates for the large contiguous region of Figure 6. The inset in (a) shows the combined intra- and inter-plate estimates on a linear scale. These results have been scaled to the depth of the Shane-Wirtanen catalogue in (b). Open symbols show the correlation function for the Lick sample taken from Figure 5 of Groth and Peebles (1977). In (b) we have plotted the crosses in (a) at $\vartheta < 0.2°$ to which we have added an offset of 0.0098 to account for the integral constraint described below. The power law fit of equation (4) is plotted as the dotted line in (a) and has been scaled to the Lick depth in (b).

On scales $\vartheta > 0.5°$, the crosses in Figure 7a consistently fall below a power law extrapolation. However, this should not be directly interpreted as a real feature of the galaxy distribution. Estimates of w from individual plates are biased low because the mean density is set to that observed on each field. The estimated $w(\vartheta)$ is therefore subject to a constraint of the form

$$\int w(\vartheta_{12})\, d\Omega_1 d\Omega_2 \approx 0, \qquad (5)$$

(Groth and Peebles, 1977) where the integral extends over the area of a single plate. If Groth and Peebles' (1977) estimates of w were correct, then at $b_J = 20$ we would expect estimates of w from individual $5° \times 5°$ fields to be biased low by ≈ 0.0098.

Adding this constant to the crosses shown in Figure 7a would not eliminate the break, but would cause it to move out to $\vartheta \approx 1.5°$ bringing it roughly into line with the scaled Lick break (see below). We have verified this conclusion numerically using monte-carlo simulations based on the Soneira and Peebles (1978) model. We have mentioned this point in some detail because our interpretation differs from that of other authors who have analysed clustering on single UKSTU fields (Shanks *et. al.* 1980, Stevenson *et. al.* 1985).

Clearly, an analysis of w over a contiguous area containing many plates is required to check the reality of features on scales $\sim 1°$. So far, we have analysed only the large region illustrated in Figure 6. The solid symbols in Figure 7a show the intra- and inter-plate estimates of w for this region determined with the estimator (2). No smoothing has been applied to these estimates. The angular correlation function falls below the power law at about 1.5°. The break is seen in both the intra- and inter-plate estimates of w, though the intra-plate estimates are systematically higher by ≈ 0.003 at $\vartheta \lesssim 2°$. This offset agrees with our estimate based on the residual errors in the galaxy counts on single plates. The inset in Figure 7a shows $w(\vartheta)$ on a linear plot at large angular scales. At 8.6°, $w(\vartheta)$ reaches a minimum of ≈ -0.0041. This relatively large value probably indicates that our map contains large scale gradients that are not attributable to the true galaxy distribution. Adding 0.0041 to the solid symbols in Figure 7 would not eliminate the break in $w(\vartheta)$.

In Figure 7b, we have scaled these estimates to the Lick depth using the relativistic Limber equation. We have adopted $q_o = 0.5$, a Schechter luminosity function with parameters in the b_J system as determined by Efstathiou *et. al.* (1987), and a k-correction of $3z$. We have ignored luminosity evolution and assumed that the clustering is stable ($\epsilon = 0$ in the notation of Groth and Peebles 1977). The estimates of w from individual plates on scales $< 0.2°$ are also plotted in this Figure with an additive offset of 0.0098 to correct for the integral constraint. The amplitude of w on small scales and the position of the break agree extremely well with the results from the Lick catalogue. The slope beyond the break also agrees with the Lick data, but this must be partly fortuitous because our estimates of systematic errors would lead us to question the reliability of values of $w(\vartheta)$ below ~ 0.004. The integral constraint (5) implies that the solid points may be underestimated by $\lesssim 0.001$ if we adopt the Groth and Peebles model for $w(\theta)$. This is a negligible correction, thus we conclude that our estimates of $w(\vartheta)$ are consistent.

The accuracy of the scaling test illustrated in Figure 7b provides further evidence that the break in the correlation function is intrinsic to the galaxy distribution. However, the example is intended primarily as an illustration of the potential of the APM survey. The contiguous region shown in Figure 6 represents only $\approx 1/7$th of the total area of our survey and was selected because of the known uniformity of the star-galaxy classification. The results presented here clearly need to be confirmed from an analysis of the whole survey. When the APM survey is in final form, it will be possible to apply internal scaling checks over a wide magnitude range.

The APM survey will be applicable to many other aspects of galaxy clustering, *e.g.* deep redshift surveys, the spatial distribution and morphologies of rich clusters of galaxies. Preliminary work has begun in these directions and we hope to present results in due course.

REFERENCES

Abell, G.O., 1958, *Astrophys. J. Suppl.*, **3**, 211.

Cannon, D., 1984, in *Astronomy with Schmidt Type Telescopes*, ed. M. Capaccioli, IAU Coll. **78**, Reidel, Dordrecht, p25.

Colless, M. and Hewett, P., 1987, *Mon. Not. R. astr. Soc.*, **224**, 453.

de Lapparent, V., Kurtz, M.J. and Geller, M.J., 1986 *Astrophys. J.*, **304**, 585.

Efstathiou, G., Ellis, R.S. and Peterson, B.A., 1987 *preprint.*

Geller, M.J., de Lapparent, V. and Kurtz, M.J., 1984 *Astrophys. J.*, **287**, L55.

Groth, E.J. and Peebles, P.J.E., 1977, *Astrophys. J.*, **217**, 385.

Groth, E.J. and Peebles, P.J.E., 1986a, *Astrophys. J.*, **310**, 499.

Groth, E.J. and Peebles, P.J.E., 1986b, *Astrophys. J.*, **310**, 507.

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

Kibblewhite, E.J., Bridgeland, M.T., Bunclark, P. and Irwin, M.J., 1984, in *Astronomical Microdensitometry Conference*, ed. D.A. Klinglesmith, NASA Conf. Publ. **2317**, p277.

Limber, D.N., 1953, *Astrophys. J*, **117**, 134.

Peebles, P.J.E., 1973, *Astrophys. J.*, **185**, 413.

Peebles, P.J.E., 1975, *Astrophys. J.*, **196**, 647.

Peebles, P.J.E., 1980, *The Large-Scale Structure of the Universe*, Princeton University Press, Princeton.

Seldner, M., Siebers, B., Groth, E.J., and Peebles,P.J.E, 1977, *Astron. J.*, **82**, 249.

Shane, C.D., and Wirtanen, C.A., 1967 *Pub. Lick Obs.*, **22**, Part 1.

Shanks, T., Fong, R., Ellis, R.S. and MacGillivray, H.T., 1980, *Mon. Not. R. astr. Soc.*, **192**, 209.

Soneira, R.M. and Peebles, P.J.E., 1978. *Astron. J.*, **83**, 845.

Stevenson, P.R.F., Shanks, T., Fong, R. and MacGillivray, H.T., 1985, *Mon. Not. R. astr. Soc.*, **213**, 953.

West, R.M., 1978. in *Modern Tecnhiques in Astronomical Photography*, eds R.M. West and J.L. Heudier, ESO Publ., Geneva, p193.

Zwicky, F., Herzog, E., Wild, P., Karpowicz, M. and Kowal, C.T., 1961-1968, *Catalogue of Galaxies and Clusters of Galaxies v1-6*, California Institute of Technology, Pasadena.