USE OF AN AUTOMATED PHOTOGRAPHIC OBJECT DETECTION SYSTEM TO ANALYSE THE EFFECT OF MAGNITUDE ON THE ANGULAR CORRELATION FUNCTION OF GALAXIES

Thesis by

William L. Sebok

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California
1982

(Submitted May 18, 1981)

ACKNOWLEGEMENTS

There are many people who helped me in this project. I can therefore only name a few.

Centermost is Jim Gunn, who suggested this project, in a field with which I soon fell in love, and who supported it through the end.

Fred Harris provided electronics know-how and muscle that finally finished the Mann Machine. Drs. Gunn and Oke provided more know-how and support for this machine.

I thank the staff of Palomar Observatory for help during the period 1976 - 1978 when my observations were made.

I thank Greg Hungerford (guitar), Katrina Curdy (flute), Chris Joyce (bass), Lisa Loos (guitar), Sister Dorothy (guitar), Marieke Schuurs (singer), and all of the members of the guitar Mass folk groups for whom I have played my flute during this stay in Pasadena. They have provided me with an alternative to Astronomy and computers necessary for my sanity.

Peter Young, John Hoessel, Doug Rabin, Peter Wannier and Russell Redman provided me with helpful discussions --- especially during those late (or very early) hours of the day when I most needed to discuss things with someone.

Richard Green, by his charisma and warmth provided me with an ideal to follow.

Finally, I thank my family for providing support throughout the years of this thesis.

ABSTRACT

An automated system is created which scans photographic plates, locating, measuring, and identifying the objects on these plates. Four deep red sensitive plates from the Palomar 1.2 m. telescope are examined with this system. The objects classified as galaxies are divided into subcatalogs of different apparent magnitudes. The angular correlation function $w(\theta)$ (a measure of how galaxies cluster) is computed for each subcatalog for galaxy separations between 18" and 4'32". This enables an analysis of the behavior of the angular correlation function with changing apparent magnitude.

The slope of log correlation function versus log angle, for galaxies $16<m_r<20.5$, is measured to be -1.23 ± 0.18, with a hint of higher negative slopes at fainter magnitudes. This is inconsistent at the 2σ level with the standard slope of -0.8, but may be possible if the slope is a function of magnitude or an effect of the smaller angular size (~ 1 arc. min.) used here.

The slope of the log correlation function as a function of magnitude is most consistent with the slope of the log correlation function as a function of angle for bright galaxies being about -1.1, consistent with the above result.

If this slope is -0.8, then moderate to strong correlation function evolution is favored. If the slope of the log correlation function versus angle is higher then less correlation function evolution is needed.

If the slope is allowed to float to its preferred higher value, then amount of correlation function evolution cannot be distinguished.

No luminosity or color evolution is included in the models, but the models have no trouble fitting the data without it.

CONTENTS

Chapter 1 .... 1 Introduction

Chapter 2 .... 8 Star/Galaxy Classifier

Chapter 3 .... 19 Pattern Recognition System

Chapter 4 .... 29 Plate Scanner

Chapter 5 .... 49 Results

Appendix A .... 82 Edge Correction Derivation.

References 84

Figure Captions 87

1.0 INTRODUCTION

Two trends seem discernible in the direction of observational astronomy at the present time. One route is the development of more and more sensitive instruments and techniques which, however, view smaller and smaller sections of the sky. The other route is the development of techniques and instruments that apply the sensitivity previously available on small scales to larger and larger areas of the sky. This thesis follows the latter route.

Automatic detection, classification, and measurement of astronomical objects on photographic plates (or other detectors) is a marriage between astronomy and a much younger science, artificial intelligence. Artificial intelligence is a fully developed science, with its own literature and, unfortunately, its own jargon. A large fraction of the work spent on this thesis was spent on the artificial intelligence aspect.

The astronomical branch of artificial intelligence is no longer an infant field. It has been carried forward by the work of Herzog and Illingworth (1977), Kibblewhite et al. (1975), Kron(1980), Green (1977), Jarvis and Tyson (Jarvis and Tyson 1979, 1980, and 1981 Tyson and Jarvis 1979), Pratt et. al. (1975), Rheault and Hardy (1980) and others. In the analysis of the output the systems generated by these workers have produced, the major emphasis has tended to be on counts of various objects.

The field of angular correlation function analysis, or, in simple terms, the measurement of how galaxies clump, has been greatly advanced by the work of Peebles (references from 1973 to 1980). Earlier work in the field was done by Limber (1953 and 1954), Scott (1962), and Neyman (1962), among others. Since the mid 1970's the literature on the subject has become rather large. Some of those workers are Totsuji and Kihara (1969), Dautcourt (1977), Davis (Davis and Peebles 1977), Sharp (1979), Fry (Fry and Peebles 1980), Groth (Groth and Peebles 1977), and Lake and Tremaine (1980). For reviews of the field see Fall (1979) or Peebles (1980).

Correlation function analysis requires large catalogs of objects. Automatic scanning systems produce large catalogs of objects. Therefore, the union of the two things seems natural. In particular, one would hope that these catalogs would be so large that small subsets of them could be taken and still have enough objects to do a correlation function analysis of each catalog. Thus, catalogs can be constructed, each selected at a particular apparent magnitude, and the relationship between the correlation function and apparent magnitude can be explored.

In this thesis, a scaling law is derived to aid in the comparison of catalogs of different magnitude. Departures from this scaling law (which is valid for a Euclidean, non-evolving universe) are of interest. Chapter two, which appeared in the Astronomical Journal, (Sebok (1979)) derives the classifier used to distinguish faint stars from galaxies. Chapter three, which appeared in an SPIE conference report (Sebok (1980)) describes the computer system used in this project. In chapter four, telescopic material is described and the properties of the plate scanning system further are described. Finally in chapter five, the catalog of objects generated by the plate scanning system is used to construct object count versus magnitude diagrams and correlation function diagrams as a function of magnitude. Models are constructed for the data and the results are interpreted.

1.1 Scaling Law

Given the existence of a randomly chosen galaxy, the probability that another galaxy is seen in the sky within a solid angle $\delta\Omega$ at an angular distance θ from the chosen galaxy is:

$$\delta P = N[1+w(\theta)]\delta\Omega \tag{1-1}$$

N is the density of galaxies per unit solid angle. This equation defines the angular correlation function $w(\theta)$.

Peebles (1973) has shown that, in a Euclidian, non-evolving universe, given a catalog to some fixed limiting distance D, the correlation function calculated from this catalog scales as:

$$w(\theta,D) = Df(\theta D) \tag{1-2}$$

when compared to the correlation function calculation from a catalog selected to some other limiting distance. However, we will show that this scaling law is valid when comparing two catalogs selected at two different apparent magnitudes. If two catalogs are selected in apparent magnitude by selection functions identical except for a shift in magnitude, then correlation functions will differ only by the scaling law. The magnitude selection function can be anything that cuts off sufficiently rapidly at faint levels. In particular the selection function can be the delta function, which yields catalogs selected at a single apparent magnitude. Evolution of the correlation function with magnitude (i.e. departures from the scaling law) will then not be obscured by an integration over magnitude.

Given the existence of a randomly chosen galaxy, the probability that another galaxy is within a volume δv at a distance $\underline{r}$ from the first galaxy is:

$$\delta P = \rho[1+\xi(r)]\delta v \tag{1-3}$$

ρ is the density of galaxies per unit volume. This equation defines the spatial correlation function $\xi(r)$.

Also, define $\phi(r)$ as the probability that an object at a distance $\underline{r}$ will be included in the catalog. These quantities are connected by Limber's equation:

$w(\theta) =$

$$\int_0^\infty dx\ x^4\phi^2(x)\left[\int_{-\infty}^{\infty} dy\,\xi[(x^2\theta^2+y^2)^{1/2}]\right] \bigg/ [\int_0^\infty dx\ x^2\phi(x)]^2$$

(1-4)

Suppose that in a different catalog $\phi(x)$ is replaced by a new selection function $\phi'(x) = \phi(x/s)$. This new catalog will have a correlation function:

$w'(\theta) =$

$$\int_0^\infty dx\ x^4\phi^2(x/s)\left[\int_{-\infty}^{\infty} dy\,\xi[(x^2\theta^2+y^2)^{1/2}]\right] \bigg/ [\int_0^\infty dx\ x^2\phi(x/s)]^2$$

(1-5)

Let $u = x/s$

$w'(\theta) =$

$$\frac{1}{s}\int_0^\infty du\ u^4\phi^2(u)\left[\int_{-\infty}^{\infty} dy\,\xi[(u^2(s\theta)^2+y^2)^{1/2}]\right] \bigg/ [\int_0^\infty du\ u^2\phi(u)]^2$$

(1-6)

$$w'(\theta) = \frac{1}{s}\,w(s\theta)$$

(1-7)

An apparent magnitude selection function is a selection function of this form (in the Euclidian, non-evolving approximation that is used here). A selection function $\phi(r)$ yields a magnitude selection function $f(m)$

$$f(m) = \int_0^\infty r^2 dr\,\phi(r)\varphi(m-5\log(r)-25) \bigg/ \int_0^\infty r^2 dr\,\phi(r)$$

(1-8)

$\varphi(M)$ is the differential luminosity function. If in a different catalog if $\phi(r)$ is replaced by $\phi'(r) = \phi(r/s)$ then the new magnitude selection function $f'(m)$ is:

$$f'(m) = \int_0^\infty r^2 dr \phi(r/s)\varphi(m-5\log(r)-25) \Big/ \int_0^\infty r^2 dr \phi(r/s)$$

(1-9)

Let u = r/s

$$f'(m) = \int_0^\infty u^2 du \phi(u)\varphi[m-5\log(u)-25-5\log(s)] \Big/ \int_0^\infty u^2 du \phi(u)$$

(1-10)

$$f'(m) = f[m-5\log(s)]$$

(1-11)

Therefore f'(m) is the original magnitude selection function shifted in magnitude.

Putting these two things together, if two catalogs, catalog 1 and catalog 2, are compared, if the magnitude selection function of catalog 2 is the selection function of catalog 1 shifted by Δm, the relationship between the correlation function of catalog 1 and catalog 2 is:

$$w_2(\theta) = 10^{-\Delta m/5} w_1(10^{\Delta m/5}\theta)$$

(1-12)

CHAPTER 2

Reprinted from:

THE ASTRONOMICAL JOURNAL VOLUME 84, NUMBER 10 OCTOBER 1979

OPTIMAL CLASSIFICATION OF IMAGES INTO STARS OR GALAXIES—A BAYESIAN APPROACH

WILLIAM L. SEBOK

Hale Observatories,[a] California Institute of Technology, Pasadena, California 91125

Received 9 April 1979; revised 2 July 1979

ABSTRACT

Bayesian decision theory is applied to the star/galaxy classification problem and a classifier is derived. This classifier can be written in the form of a correlation with a single stored function. The classifier is then applied to images on 48-in Schmidt plates. There are no free parameters which need to be adjusted for each plate. Calibration of the classifier to a plate simply involves obtaining images of (brighter) objects known to be stars. Sources of error are discussed and a second classifier, which is insensitive to variations in the sky background, is derived and applied to the plate data. Finally a prediction of the magnitude limit is derived for both classifiers and compared to the observed magnitude limit. This observed magnitude limit is about one to one and a half magnitudes above the plate detection limit.

I. INTRODUCTION

The advent of panoramic detectors, as well as the increasing availability of high-speed microdensitometers to digitize photographs, has meant the appearance of a large amount of two-dimensional picture data in digital form. This has spurred interest in digital techniques to extract information from such pictures.

One such digital technique is the extraction of lists of various types of objects that are present in the picture. To accomplish this, one has to (1) determine the positions and extents of the objects, and (2) identify the objects. This paper is concerned with one facet of the second problem, the distinguishing of star images from galaxy images, or more precisely unresolved images from resolved or partially resolved images. Several papers have examined this problem. *Ad hoc* or nonoptimal classifiers are described in Pratt *et al.* (1975), Kibblewhite *et al.* (1975), Lorre *et al.* (1979) and Kron (Ph. D. thesis, 1978) (however, note that Kron's classifier is a form of the classifier derived here). A nonparametric classifier is given by Jarvis and Tyson (1979). The classifier derived here can be incorporated with other shape discriminators into a nonparametric classifier of the type discussed by Jarvis and Tyson. This may be useful as the classifier derived here is a linear algorithm that is only useful on nonsaturated portions of the image to be classified.

Star images have fairly simple properties. Basically, all star images consist of a single star profile multiplied by some scaling factor (and, of course, with the addition of noise). Galaxy images have quite a large variety of shapes. Even so, they are always larger than a star image of the same brightness.

To apply Bayesian pattern recognition theory to this problem, one takes these assumptions and constructs a statistical model of the probabilities of getting different types of images (1) if the image is a galaxy, or (2) if the image is a star. These probabilities are then used to decide which possibility is more likely to be correct for a particular image. A numerical functional of the image (called a classifier), whose value decides among the possible classifications of the image, is constructed from the probabilities. The hope is, of course, that this will produce a classifier simple enough to be useful. The other hope is that the effort spent in construction and understanding of the resultant optimal classifier will result in greater understanding of the objects being classified.

In the paper, the basic classifier is first derived. The classifier is then tested on a sample of real star and galaxy profiles, and the results discussed. Next, various modifications to the basic classifier are examined, and a form of the classifier insensitive to background variations is derived. This modified classifier is also applied to the profile samples. The last part contains a calculation of expected error rates.

II. DERIVATION OF CLASSIFIER

The notation here follows Duda and Hart (1973), which provides a good introduction to the ideas presented in this paper.

The basic idea is to calculate the probability that the unknown object is a star and compare it to the proba-

---

[a] Operated jointly by the California Institute of Technology and the Carnegie Institution of Washington.

bility that the object is a galaxy. Optimal classification is then achieved by picking whichever probability is greater. The difficulty, of course, is this calculation of these probabilities. Usually these probabilities are not completely known and must be calculated from a model. The "resultant" classifier is then only "optimal" within its model. However, if the model is a reasonable approximation to nature, then the classifier may still be quite good. Thus it is possible to generate several rather different looking classifiers by making different assumptions in the model of the probabilities. Each classifier is optimal within its own model. Criteria such as simplicity and speed of implementation, as well as insensitivity to effects ignored in the model, are used to pick which classifier is "best."

We start out with a list of numbers, the measurement vector y. Those numbers are a set of surface brightness measurements (in linear intensity units) over the object to be classified. The analysis in this paper used measurements arranged in a two-dimensional square grid, all of which lay within a circular aperture of a certain diameter. Of course, the theory itself is much more general than this, being valid for other sizes and shapes of apertures, one-dimensional rather than two-dimensional profiles, or even, with the proper initial assumptions, data which do not represent profiles of objects in a picture.

Suppose there are two classes of objects, stars S, and galaxies G. The Bayes Error Criterion is to classify y into one of these two classes such as to minimize the loss:

$$\text{Loss} = \begin{cases} \lambda_{s \rightarrow g}\, p(S \mid \mathbf{y}), & \text{if y is classified as a galaxy,} \\ \lambda_{g \rightarrow s}\, p(G \mid \mathbf{y}), & \text{if y is classified as a star,} \end{cases} \tag{1}$$

where $p(S \mid \mathbf{y})$ is the probability, given a particular set of measurements y, that the object is a star, and $p(G \mid \mathbf{y})$ is the probability, given a particular set of measurements y, that the object is a galaxy. $\lambda_{s \rightarrow g}$ is the loss associated with misclassifying a star as a galaxy, and $\lambda_{g \rightarrow s}$ is the loss associated with misclassifying a galaxy as a star.

Clearly, the optimal way to classify an object is to minimize this Loss. That is, to say that the object is a galaxy when

$$\lambda_{g \rightarrow s}\, p(G \mid \mathbf{y}) > \lambda_{s \rightarrow g}\, p(S \mid \mathbf{y}). \tag{2}$$

This is a generalization of the case in which we are simply interested in being right most of the time (called the unbiased case). For the unbiased case, $\lambda_{g \rightarrow s} = \lambda_{s \rightarrow g}$, and

$$p(G \mid \mathbf{y}) > p(S \mid \mathbf{y}) \text{ is the criterion.} \tag{3}$$

With this in mind, since the ratio of $\lambda_{s \rightarrow g}$ to $\lambda_{g \rightarrow s}$ enters as a bias factor, define

$$w = \lambda_{s \rightarrow g} / \lambda_{g \rightarrow s}. \tag{4}$$

How do we calculate these probabilities? $p(S \mid \mathbf{y})$, the probability that, given y, the object is a star, is not initially known. The classification problem would be trivial if it was known. However, $p(\mathbf{y} \mid S)$, the probability of getting measurements y given that the object is a star, is a quantity much easier to model. How do we "turn the probabilities around"? It is done by applying Bayes' theorem:

$$p(S \mid \mathbf{y}) = p(\mathbf{y} \mid S) p(S) / p(\mathbf{y}), \tag{5}$$

thus earning the title "Bayesian Classifier" for the resultant classifier. Here $p(S)$ is the probability that, knowing no measurements, that the object is a star (this is called the a priori probability that the object is a star). $p(\mathbf{y})$ is the total probability of getting the set of measurements y. Since the classifier works with the ratios of probabilities, $p(\mathbf{y})$ drops out and can be ignored. $p(S)$ and $p(G)$ (the *a priori* probability that the object is a galaxy) enter the classifier as another set of bias factors. They enter as a ratio $p(G)/p(S)$, the ratio of galaxy counts to star counts. If we do not know this ratio or, more importantly, do not wish to be biased by it, we may choose to ignore this factor.

Applying all of this to the classifier, we are left with

$$p(\mathbf{y} \mid G) / p(\mathbf{y} \mid S) > w p(S) / p(G), \tag{6}$$

as the condition for the object to be classified as a galaxy.

The ratio $p(S)/p(G)$ is the ratio of the number of stars to the number of galaxies. One can explicitly use some value for this ratio. This would be one of the possible "models" for a classifier. However, another possible model is one in which this ratio is set to one, not necessarily because one does not know the value of that ratio (although I may not) but because one wants the classifier to treat stars and galaxies evenhandedly and judge only on the properties of the profile under scrutiny. One may not want the decision to be influenced by which object is intrinsically more probable. Another way of looking at it is that one is free to set the bias factor w to whatever value will cancel $p(S)/p(G)$.

The problem then reduces to modeling the $p(\mathbf{y} \mid G)$ and $p(\mathbf{y} \mid S)$'s. Note again that within such a model the resultant classifier is the optimal classifier for that model. *One could not even in principle do better.*

The more obvious quantity to model is $p(\mathbf{y} \mid S)$. Let us make our assumptions explicit:

(1) y contains linear measurements of light intensity. The data could either come from a linear detector or from a photographic plate which has been linearized by some means, such as a calibration wedge.

(2) The intensity of the sky has been measured and subtracted.

(3) Noise in adjacent pixels is uncorrelated (the effect of noise correlations will be investigated later).

(4) The object to be classified has been (perfectly) centered within the aperture y.

(5) An ideal (noiseless) star image would have the

form $A\mathbf{s}$, where $\mathbf{s}$ is the star profile for that telescope and plate ($\mathbf{s}$ is a vector in the same vector space as $\mathbf{y}$) and A is a scaling constant. A has the property of a magnitude indicator. This assumption is eminently reasonable up to the point where it breaks down due to the saturation of the detector.

(6) The deviations of the image from ideal are Gaussian of standard deviation σ_y.

(7) One further assumption must be made about the magnitude indicator A. This assumption will be clearer if it is specified when needed.

With assumptions (1)–(6), we can immediately write

$$p(\mathbf{y}|S) = [\sigma_y(2\pi)^{n/2}]^{-1} \times \exp\{-[(\mathbf{y} - A\mathbf{s})\cdot(\mathbf{y} - A\mathbf{s})/(2\sigma_y^2)]\} \quad (7)$$

where n is the number of pixels in $\mathbf{y}$ and $\mathbf{s}$.

Here is where assumption (7) comes in. What is A? What we really have here is $p(\mathbf{y}|S,A)$, the probability of getting a set of measurements $\mathbf{y}$ given that the object is a star *and* that it has a magnitude indicator whose value is A. Assumption (7) says that for a particular $\mathbf{y}$, $p(\mathbf{y}|S)$ can be obtained by substituting into $p(\mathbf{y}|S,A)$ the value of A that maximizes $p(\mathbf{y}|S,A)$. That is, the maximum likelihood value of A is used. This is a reasonable assumption but not the only possible one. With a few assumptions as to the smoothness of the star luminosity function, it can be shown that integrating over all A's is equivalent to using the maximum likelihood value for A. This problem is further explored in Appendix A.

Maximizing $p(\mathbf{y}|S,A)$ with respect to A yields

$$A = \mathbf{y}\cdot\mathbf{s}/\mathbf{s}\cdot\mathbf{s}. \quad (8)$$

Substituting this back into Eq. (7) yields

$$p(\mathbf{y}|S) = [\sigma_y(2\pi)^{n/2}]^{-1}\exp(-1/2\sigma_y^2) \times [\mathbf{y}\cdot\mathbf{y} - (\mathbf{y}\cdot\mathbf{s})^2/\mathbf{s}\cdot\mathbf{s}]. \quad (9)$$

The case for stars was relatively straightforward. The case for galaxies is more complicated. Galaxy images occur in a remarkable variety of sizes and shapes. A realistic $p(\mathbf{y}|G)$ would have to involve some integral over all galaxy shapes (and would yield a classifier too cumbersome to be useful). To generate a useful classifier, I make the same (?!) assumption for galaxies as I made for stars, that is, in the absence of noise, that the profile would have the form $C\mathbf{g}$, where $\mathbf{g}$ is a "standard" galaxy model and C is a magnitude indicating constant. With that assumption, following the same steps as for the star case, we get

$$p(\mathbf{y}|G) = \frac{1}{\sigma_y(2\pi)^{n/2}} \times \exp\left[-\frac{1}{2\sigma_y^2}\left(\mathbf{y}\cdot\mathbf{y} - \frac{(\mathbf{y}\cdot\mathbf{g})^2}{\mathbf{g}\cdot\mathbf{g}}\right)\right]. \quad (10)$$

Near the plate limit, galaxy images are small. The details of their profile shapes are obscured by the seeing. Only the large-scale properties of these images remain, the most important one being the shape of the outer part of the profile. This can usually be approximated as a power law $f(r) \sim r^n$ where n is about -2. This is the justification for using a single galaxy template $\mathbf{g}$. Near the plate limit, galaxies are basically seeing discs with only this scale-independent power-law tail visible.

There is, in fact, one other factor that the seeing might not completely obliterate: the ellipticity of the galaxy. However, a galaxy with observable ellipticity looks even less like a star and, in fact, the classifier is even less likely to think that that galaxy is a star. Away from the plate limit, as galaxy images get bigger, they also look less like star images. A classifier constructed for objects near the plate limit should therefore have even less trouble with brighter galaxies.

Recalling inequality (6) and inserting $p(\mathbf{y}|G)$ and $p(\mathbf{y}|S)$ yields

$$\exp\left[+\frac{1}{2\sigma_y^2}\left(\frac{(\mathbf{y}\cdot\mathbf{g})^2}{\mathbf{g}\cdot\mathbf{g}} - \frac{(\mathbf{y}\cdot\mathbf{s})^2}{\mathbf{s}\cdot\mathbf{s}}\right)\right] > \frac{wp(S)}{p(G)}, \quad (11)$$

as the condition for classification of the object as a galaxy. Alternatively,

$$(\mathbf{y}\cdot\mathbf{g})^2/\mathbf{g}\cdot\mathbf{g} - (\mathbf{y}\cdot\mathbf{s})^2/\mathbf{s}\cdot\mathbf{s} > 2\sigma_y \ln[wp(S)/p(G)]. \quad (12)$$

Let us examine the unbiased case, that is, when the right hand side vanishes. The inequality becomes

$$(\mathbf{y}\cdot\mathbf{g})^2/\mathbf{g}\cdot\mathbf{g} > (\mathbf{y}\cdot\mathbf{s})^2/\mathbf{s}\cdot\mathbf{s}, \quad (13)$$

$$\mathbf{y}\cdot\mathbf{g}/\mathbf{y}\cdot\mathbf{s}\sqrt{\mathbf{s}\cdot\mathbf{s}/\mathbf{g}\cdot\mathbf{g}} > 1. \quad (14)$$

Thus we define the classifier

$$\phi = \frac{\mathbf{y}\cdot\mathbf{g}}{\mathbf{y}\cdot\mathbf{s}}\sqrt{\frac{\mathbf{s}\cdot\mathbf{s}}{\mathbf{g}\cdot\mathbf{g}}} = \frac{\sum yg}{\sum ys}\sqrt{\frac{\sum s^2}{\sum g^2}}. \quad (15)$$

The object is classified as a galaxy when $\phi > 1$.

This is the desired classifier. It is the ratio of the correlation of $\mathbf{y}$, the unknown, with two stored functions. In practice, a classifier boundary different than one often has to be used because of differences between the templates $\mathbf{g}$ and $\mathbf{s}$ and "nature."

Inequality (15) can be rewritten in the equivalent form

$$\mu = \sum y[g/\sqrt{\sum g^2} - s/\sqrt{\sum s^2}] > 0. \quad (16)$$

This form of the classifier, a correlation of the unknown with only one stored function, has the beauty of being very efficient and fast to implement on a computer.

What should one use for the templates $\mathbf{s}$ and $\mathbf{g}$? Clearly, for best results, the templates should resemble nature as closely as possible. I have found that, fortunately, there is a rather wide latitude in the templates that will still work, although a classifier boundary different from the theoretical value of $\phi = 1$ often has to be used. The comparison of the integrated magnitude within a narrow aperture to the integrated magnitude within a wide aperture, for example, is equivalent to this classifier, with crude step function star and galaxy templates.

Basically, for the ϕ classifier to work at some level, g has to be something wide and s, something narrow.

It is relatively easy to construct a good realistic model for s by simply averaging some star images together. They should be faint enough so that their profiles are nowhere saturated but otherwise as bright as possible to produce a profile relatively free from noise. To get the galaxy template, I constructed a galaxy model (typically a Hubble law $y = a(r/c + 1)^{-2}$ where a and c are constants and r is the radial distance). This was then convolved with the star profile s to simulate the smearing of the true galaxy shape by the seeing. The hope was, with both the star and galaxy templates equally influenced by the seeing, that the effect of the seeing would (at least to some order) "cancel out" of the classifier, leaving a classifier that was not systematically affected by the seeing. Of course, the classifier is expected to do worse in bad seeing, as a bigger seeing disk decreases the difference in appearance between a star and a galaxy, but at least the classifier would not be biased by changes in the seeing.

III. PREPARATION OF DATA

To check the performance of the classifier, small regions (frames) were digitized on several plates with the PDS Microphotometer at the Jet Propulsion Laboratory. These plates were all of the red 127-04 (now IIIa-F) emulsion, and were taken on the Palomar 1.2-m Schmidt telescope. PS21344 was lent by J. E. Gunn. PS23102 and PS24307 were taken by the author as part of a project to make a catalog of faint galaxies covering 600 square degrees of sky. The characteristics of each plate are summed up in Table I.

For each of these fields, there was also a plate taken with the 90-mm image tube on the Palomar 5-m telescope. These plates were used to check the performance of the classifier.

The 1.2-m telescope plates were digitized at a resolution of 10 μm, which at the plate scale of 67.2″ mm^{-1} corresponds to a pixel size of 0.672 arcsec. A 512 × 512 (or in one case a 1024 × 512) pixel frame was digitized around each field. In addition, PS24307, the plate from which field 3 was taken (see Table I), also contained another instance of field 2. The 512 × 512 frame, which was digitized around field 2 on that plate, was used for the photometry of the objects in that field. In addition, a 512 × 512 region was scanned on that plate around the image of the 15th magnitude elliptical galaxy IC 75. Surface photometry of this galaxy was used to provide the zero point for the photometry of fields 2 and 3. Calibration wedges were present on each plate (light incident on the plate varying exponentially with distance along the wedge). These wedges were digitized at the same resolution as the rest of the plate.

The calibration wedges were used to linearize the pixels in each frame to units of light intensity. Then the sky background was measured and subtracted, pixel by pixel. To measure the sky intensity in a frame, a histogram was first made of the intensities of the pixels in that frame. Then the histogram was displayed on the screen of a computer CRT terminal and a gaussian fitted through it, the boundaries for the fit being selected with a cursor. Generally, when selecting those boundaries, one favors the "dark" (lower intensity) side of the histogram. The "objects" in the frame (stars, galaxies, etc.) affect only the "bright" side, introducing a skew which mainly affects the bright side of the histogram. The position of the peak of the Gaussian is one estimator for the sky level and provides a consistency check for the method that follows. The width of the fitted Gaussian is noted and the derivative of a centered Gaussian of that width is computed and stored. The Gaussian derivative is then convolved back with the original histogram and the position of the zero-crossing point of the convolution is used as an estimator for the sky level. If the histogram were really a Gaussian of that width, then this would be the optimal method for finding the sky background (proof in Appendix B). In practice it is very good. It is rather insensitive to skew in the histogram.

After the sky background was measured in each of the frames, the zero points of their magnitude scales had to be fixed. The frame of IC 75 on plate PS24307 was compared to surface photometry obtained from the Direct Silicon Intensified Target camera on the Palomar 1.5-m telescope. Two "SIT" frames were available, one taken 30 September 1978, the other 2 October 1978. Each SIT frame, a 256 × 256 digital picture, was reduced to the red color in the magnitude system of Thuan and Gunn (1976). A ring with a 11.4″ inner diameter and a 27″ outer diameter was used for the intercomparison with the photographic frame (as inside the inner diameter the photographic frame was saturated). This zero point was then applied to field 3 and the PS24307 frame of field 2. In addition, as those two frames were rather near the corner of the plate, a vignetting correction was applied. The ratio of the sky backgrounds, which had been subtracted from those frames, to the sky background, which had been subtracted from the IC 75 frame, was used as this correction. Although crude, this

TABLE I. Characteristics of plate material.

Field	Plate number	Emulsion	Filter	Coordinates α 1950	Coordinates δ 1950	Region scanned	Position in plate
1...	PS21344	127-04	red plexiglas	14 09.5	52 26	512 × 512	center
2...	PS23102	127-04	red plexiglas	0 28.8	5 39	1024 × 512	corner
3...	PS24307	127-04	Wratten 25	0 26.3	5 39	512 × 512	corner, but farther in than 2

was judged to be better than no correction. The correction was about 0.15 mag.

Field 1 contains 3C295. Multichannel photometry of this galaxy was available in the paper by Gunn and Oke (1974). A red magnitude was synthesized for this galaxy from that photometry and the bandpass curves given in the Thuan and Gunn paper. The same thing was also done for the neighboring galaxy, called "foreground 3C 295" in the Gunn and Oke paper. The zero points obtained from those two galaxies were averaged together to fix the zero point of field 1.

Once the fields had been calibrated, a set of objects was selected from each field. Everything down to about 20.5 mag was included. The image tube plates from the 5-m telescope were carefully examined and each object was classified as a star or a galaxy. For the fainter objects there is a finite probability of mistakes in this classification. Coordinates of the objects were measured on each frame to the nearest pixel. Two centering algorithms were used: (1) a simple centroid finder, (2) an algorithm for galaxies which involved fitting a Lorentzian to the object and taking the position of the maximum of that Lorentzian. For stars the Lorentzian was replaced by a Gaussian.

Magnitudes for each object were measured within a 10″ circular aperture. Of course, for large and bright objects, this produces an underestimate of the brightness. When the classifier was applied to an object, the "object" was considered to be those pixels contained within this same 10″ (15 pixel) circular aperture.

IV. BEHAVIOR OF CLASSIFIER

The behavior of this classifier is shown in the plots of the classifier versus magnitude, Figs. 1–6. The star template in each case is the average of several star images from the same field. The star images chosen were generally from 17th to 18th magnitude. The galaxy template was this star template convolved with the galaxy models given in the figure captions. A large number of such models was tried, but due to space limitations, only a representative sample of two is shown here. The dashed lines indicate where the templates themselves fall on the graph. The stars and galaxies are indicated by +'s and O's respectively.

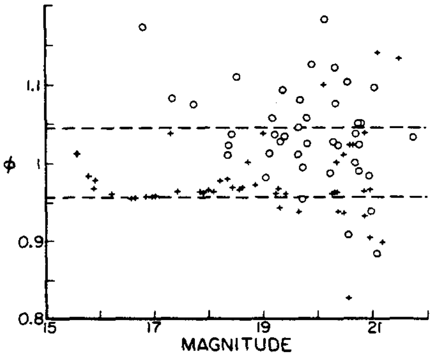

FIG. 2. Field 2, Galaxy model is Hubble Law, $(r + \frac{1}{3}'')^{-2}$, +'s are stars, O's are galaxies.

In some ranges of magnitude, the stars lie in a horizontal line of fairly tight dispersion. Towards lower magnitudes, this "star line" curves up. This is due to the pixels becoming nonlinear as the linearization of the plate densities breaks down. Toward high magnitudes, the star line spreads out, as one would expect. The galaxies lie in a cloud above the star line. There is an upward trend in ϕ as galaxies get brighter, nearer, and bigger. But, as claimed earlier, this is all for the better.

To use these graphs, one would first draw two vertical lines within which the classifier produces valid results. To the left the classifier breaks down because of image saturation, while to the right the classifier breaks down due to inadequate signal to noise. Between these vertical lines one draws a horizontal line to separate the stars from the galaxies.

If the templates really resembled nature, then this horizontal line would be at the theoretical value of $\phi = 1$. Conversely, a boundary close to $\phi = 1$ is a necessary (but not sufficient) condition that one's templates resemble nature. Figs. 1, 2, and 3, which use a Hubble law of core radius $\frac{1}{3}''$, are such a case. On the other hand, for Figs. 4, 5, and 6, which use a Hubble law of core radius

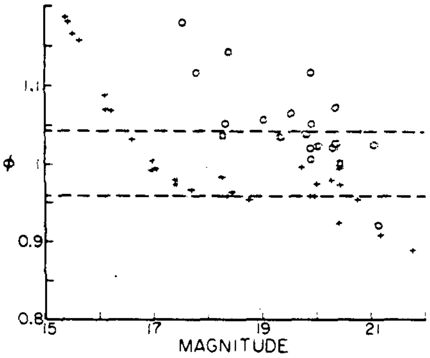

FIG. 1. Field 1, Galaxy model is Hubble Law, $(r + \frac{1}{3}'')^{-2}$, +'s are stars, O's are galaxies.

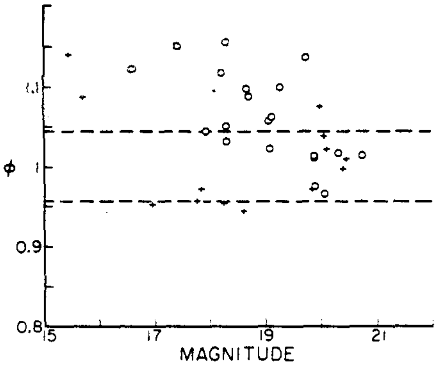

FIG. 3. Field 3, Galaxy model is Hubble Law, $(r + \frac{1}{3}'')^{-2}$, +'s are stars, O's are galaxies.

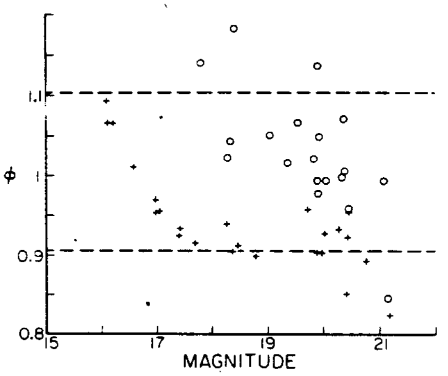

FIG. 4. Field 1, Galaxy model is Hubble Law, $(r + \frac{2}{3}'')^{-2}$.

$\frac{2}{3}''$, $\phi = 0.98$ would be more appropriate. However, note that for a particular galaxy model, *the same classifier boundary works for all three fields.* These fields were from different plates and different parts of the sky and were taken under different conditions. The only adjustment for seeing differences was due to using actual star images from the same field for the star template. These star images can be bright enough to be easily identified by inspection as stars. No detailed inspection of faint objects needs to be done on a plate-by-plate basis. *There are no free parameters that need to be adjusted for each plate.*

In all of these fields, the classifier seems to work down to about 20 mag or, to be safe, 19.5 mag. However, especially in field 2, there are stars present which lie far above where they should. On inspection, these objects turn out to be due to (1) double stars, (2) the presence of a nearby bright star or the diffraction spike from such a star, (3) the presence of a smudge on the plate, (4) errors in the centering of the object. Problem 1, overlapping objects, must be resolved by other tests. It can be reduced by keeping the size of the templates as small as possible (10″ templates were used here). Problems 2 and 3 are a form of local (small scale) sky background variations.

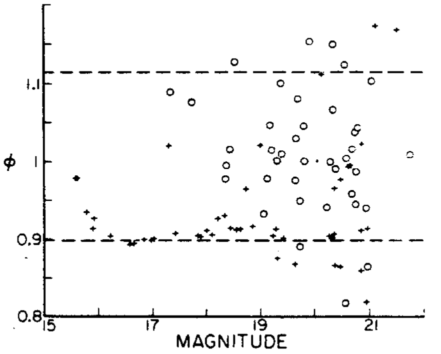

FIG. 5. Field 2, Galaxy model is Hubble Law, $(r + \frac{2}{3}'')^{-2}$.

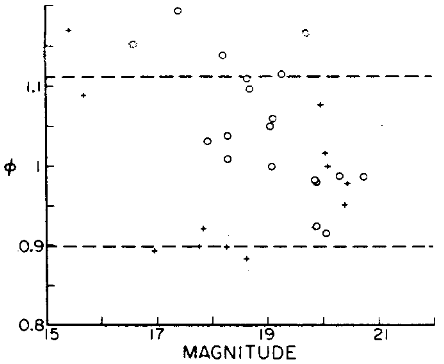

FIG. 6. Field 3, Galaxy model is Hubble Law, $(r + \frac{2}{3}'')^{-2}$.

All centering was done to within one pixel (no interpixel interpolation was done). The data analyzed here were scanned at a high spatial resolution to reduce the problem of objects which do not center exactly on a pixel. Centering the fainter objects is intrinsically a difficult problem (see Appendix B for an attempt in that direction).

Figure 7 shows the effect of local sky variations on the classifier. The templates are the same as for Fig. 3. The star template is scaled to the desired magnitude and a constant offset added. One unit of offset corresponds to a surface brightness of 27.56 mag $(\text{arcsec})^{-2}$, or a variation of 22.38 mag integrated over the whole 177 pixel aperture. As one can see, the effect is substantial. The curve for offset = 0 also bends down a bit. This is because of rounding-off errors due to truncation of the integer pixels.

Figure 8 shows the effect of centering errors on the classifier. Again, the templates are from Fig. 3. The classifier is plotted against shifts of one or more $\frac{2}{3}''$ pixels.

V. VARIATIONS ON THE CLASSIFIER

It was noted earlier that the classifier just derived assumed that there are no correlations between pixels, and that the dispersion of the pixel values is a constant. However, for photographic data, that is not necessarily a good assumption. How can the classifier be modified to take these things into account?

The answer is that every dot product, such as $\mathbf{s}\cdot\mathbf{s}$, $\mathbf{y}\cdot\mathbf{s}$,

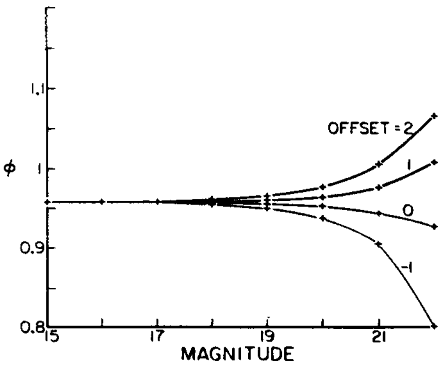

FIG. 7. Effect of Local Sky Variations on Classifier.

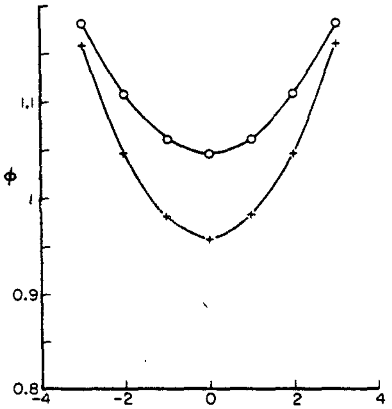

FIG. 8. Effect of Position Errors on Classifier.

and g•s is replaced by products of the form s $\mathbf{W}^{-1}$s, y $\mathbf{W}^{-1}$s, and g $\mathbf{W}^{-1}$s, where **W** is the covariance matrix of the pixels. The intersection of a particular row with a particular column of the matrix corresponds to the correlation of one particular pixel with another particular pixel.

If there are no correlations between pixels, the diagonal elements of **W** become equal to σ_y^2, the variances of the pixels, and the sums mentioned above become $\Sigma 1/\sigma_y^2 s^2$, $\Sigma 1/\sigma_y^2 ys$, and $\Sigma 1/\sigma_y^2 gs$. The σ_y's can be functions of y, that is, functions of the pixel values themselves. This is a more reasonable model of the situation in a linearized photographic plate. More generally, since $\sigma_y^2 \rightarrow \infty$ when the detector becomes saturated, the classifier can be modified to handle image saturation by simply throwing out the terms in *all* sums (including the Σs^2 and Σg^2 sums) that correspond to a value in y which is above a saturation threshold.

If the σ_y's are assumed to be constant for every pixel, then they can be factored out, leaving a matrix with 1's on the main diagonal. This matrix reduces to the identity matrix in the case of no correlation between the pixels.

The matrix is rather large—177 by 177 in the case of the templates used here. It presented some intractable problems of storage and computation on the minicomputer used for these calculations. Also, it was hard to get reliable measurements for elements farther from the main diagonal than those corresponding to a shift of one pixel. As the classifier seemed tolerant of fairly large variations in the form of the templates (and this correction is one more adjustment of the templates), it was thought that this correction could be ignored. Larger-scale correlations are clearly present but are hard to measure. They show up as local background variations. In fact, these larger-scale variations in the photographic emulsion (scale size about ½ mm = 50 pixels at this plate scale and resolution) seem to be one of the major sources of error for this classifier, possibly even dominating the pixel-to-pixel noise that is included in the model on which the classifier is based. Also, this classifier is sensitive to smudges and nearby objects (for instance, some of the misclassified stars lie over a diffraction spike of a bright star). These problems would perhaps be better handled by a "background-independent" classifier. It would be desirable to modify the theory (and the classifier) to make it insensitive to uniform changes in the background level. Let us attempt to do this.

Let **u** be the vector (in the same vector space as y, s, and g) whose components are all equal to 1. Then, for example, $\mathbf{y}\cdot\mathbf{u} = \Sigma y$. Equation (7) is then modified by replacing y with $y - B\mathbf{u}$, where B is a constant which reflects the background level.

$$p(\mathbf{y}|s) = 1/\sigma_y(2\pi)^{n/2} \exp\{-1/2\sigma_y^2 \times [(\mathbf{y} - A\mathbf{s} - B\mathbf{u})\cdot(\mathbf{y} - A\mathbf{s} - B\mathbf{u})]\}. \quad (17)$$

The obvious thing to do is to treat B the same way as A was treated, by inserting its maximum likelihood value. $p(\mathbf{y}|S)$ would be maximized simultaneously with respect to both A and B. This yields a classifier identical to the old one, except that y, s, and g are replaced by the same quantities with their mean values subtracted. In practice, that classifier does not seem to work very well. Too much of the difference between g and s is similar to the difference between u and s. That is, it is too hard to tell the effect of the difference between a star and a galaxy from the effect of a change in the background. A better background estimator is one that uses the pixels away from the center of the image, where both the components of s and g are small.

Suppose we carry B along, fixing its value with an estimator later. If the same B is used for $p(\mathbf{y}|G)$, as was used in $p(\mathbf{y}|S)$, this yields

$$\phi_{\text{new}} = \left\{\frac{\sum s^2}{\sum g^2}\right\}^{1/2} \left\{\frac{\sum yg - B\sum g}{\sum ys - B\sum s}\right\}. \quad (18)$$

For the estimator for B let us use

$$B = \sum ky/\sum k, \quad (19)$$

where **k** is a function weighed most strongly away from the center of the image. One more restriction needs to be imposed, that **k** be orthogonal to both s and g. That is,

$$\sum kg = \sum ks \equiv 0.$$

Then,

$$\phi_{\text{new}} = \left\{\frac{\sum s^2}{\sum g^2}\right\}^{1/2} \left\{\frac{\sum y(g - k\sum g/\sum k)}{\sum y(s - k\sum s/\sum k)}\right\}. \quad (20)$$

The condition of orthogonality is necessary to prevent the presence of a star or galaxy from affecting the background estimate. **k** has to be orthogonal to both s and g in order to permit the use of the same B in both $p(\mathbf{y}|G)$ and $p(\mathbf{y}|S)$. The classifier derived from $p(\mathbf{y}|G)$ and $p(\mathbf{y}|S)$'s with different B's cannot be reduced to the simple linear form of ϕ.

If **h** is some arbitrary function (again weighted mostly toward the outside), a **k** can be constructed from it of the form

$$\mathbf{k} = \mathbf{h} + X\mathbf{g} + Y\mathbf{s}. \tag{21}$$

Orthogonality of **k** with **s** and **g** fixes X and Y, and yields

$$\mathbf{k} = \mathbf{h} + \mathbf{g}[(\mathbf{h}\cdot\mathbf{g})(\mathbf{s}\cdot\mathbf{s}) - (\mathbf{h}\cdot\mathbf{s})(\mathbf{g}\cdot\mathbf{s})] + \mathbf{s}[(\mathbf{h}\cdot\mathbf{s})(\mathbf{g}\cdot\mathbf{g}) - (\mathbf{h}\cdot\mathbf{g})(\mathbf{g}\cdot\mathbf{s})]/(\mathbf{g}\cdot\mathbf{g})(\mathbf{s}\cdot\mathbf{s}) - (\mathbf{g}\cdot\mathbf{s})^2 \tag{22}$$

The resultant classifier (in its linear μ form), is

$$\mu_{\text{new}} = \sum y \left\{ \frac{g}{(\sum g^2)^{1/2}} - \frac{s}{(\sum s^2)^{1/2}} - \left[\frac{\sum g}{(\sum g^2)^{1/2}} - \frac{\sum s}{(\sum s^2)^{1/2}}\right] \times \left[\frac{h[\sum s^2 \sum g^2 - (\sum gs)^2] - g[\sum hg \sum s^2 - \sum hs \sum gs] - s[\sum hs \sum g^2 - \sum hg \sum gs]}{\sum h[\sum s^2 \sum g^2 - (\sum gs)^2] - \sum g[\sum hg \sum s^2 - \sum hs \sum gs] - \sum s[\sum hs \sum g^2 - \sum hg \sum gs]}\right]\right\}. \tag{23}$$

In spite of its complex appearance, it is still the correlation of the unknown **y** with a stored function. It has the property that the classifier behavior towards the original templates **g** and **s** is conserved over the unmodified classifier. Figures 9–11 show the application of this modified classifier on the same sample of objects used earlier. The **s** and **g** templates are the same as for Figs. 1–3. The background measuring template **h** was the set of all pixels in the 15 pixel square that lie outside the 15 pixel diameter circle.

Figures 12 and 13 compare the behavior of the background-independent classifier with the behavior of the background-dependent (original) classifier. Figure 12 shows the percentage of stars correctly classified and Fig. 13 does the same thing for galaxies. The sets of objects from all three were folded together and divided into magnitude bins one unit of magnitude in width. The points of the background-dependent classifier are labeled with squares and the points of the background-independent classifier are labeled with triangles.

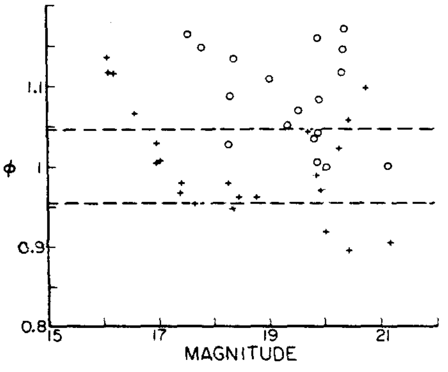

FIG. 9. Field 1, Floating background classifier, Galaxy model is Hubble Law with 1/3″ core radius.

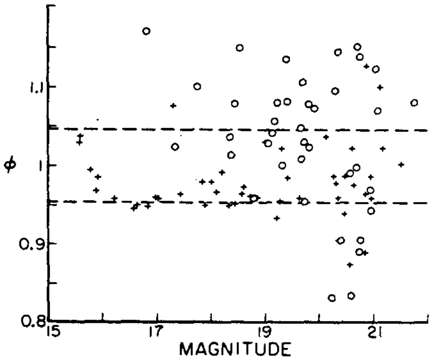

FIG. 10. Field 2, Floating background classifier, Galaxy model is Hubble Law with 1/3″ core radius.

Both classifiers show a drop in correct classifications for brighter stars. As mentioned previously, this is due to saturation of the images. The drop in correct classifications for fainter objects is also expected. In most cases, the background-independent classifier does less well than the background-dependent classifier. An examination of the individual graphs (Figs. 1–3 versus Figs. 9–11) reveals why. The background-independent clas-

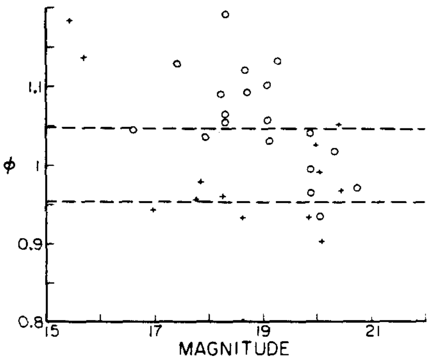

FIG. 11. Field 3, Floating background classifier. Galaxy model is Hubble Law with 1/3″ core radius.

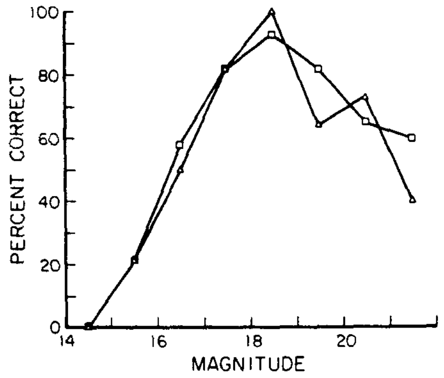

FIG. 12. Stars of all fields, percent of correct classifications. Galaxy model is Hubble Law with 1/3″ core radius. Background-dependent classifier points (data from Figs. 1–3) are □'s. Background-independent classifier points (data from Figs. 9–11) are △'s.

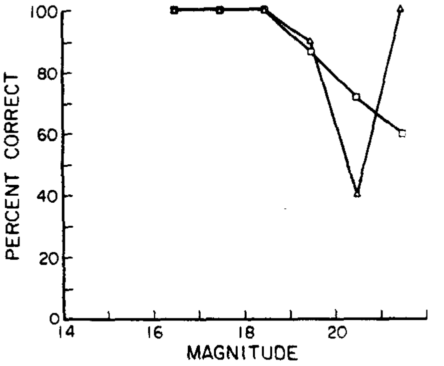

FIG. 13. Galaxies of all fields, percent of correct classifications. Galaxy model is Hubble Law with ⅓″ core radius. Background-dependent classifier points (data from Figs. 1–3) are □'s. Background-independent classifier points (data from Figs. 9–11) are △'s.

sifier has a higher dispersion than the background-dependent one. This is most easily seen in the "star line" on the graphs, which appears wide. Some discrepant stars, ones with bright objects nearby, come into line, but this gain is offset by errors due to the higher general dispersion. In conclusion, some of the information that separates stars from galaxies seems to be lost when the background is allowed to float. However, the concept of working with ease in the vicinity of bright stars and smudges still seems attractive.

VI. DERIVATION OF EXPECTED MAGNITUDE LIMIT

The last thing to do is to calculate the expected magnitude limit of the classifiers and compare it to the observed classifier magnitude limit and the plate detection limit.

The most straightforward quantity to calculate is $p(S \rightarrow G)$, the probability that a star will be misclassified as a galaxy. The probability of misclassifying a galaxy as a star $p(G \rightarrow S)$, suffers from the intrinsic dispersion of galaxies as a class.

To be general, suppose we have a quantity μ of the form $\mu = \mathbf{q} \cdot \mathbf{y}$, and $\mathbf{y}$ is distributed with Gaussian statistics of standard deviation σ_y around a mean of $A\mathbf{s}$ [as in Eq. (7)]. Then μ, as a linear function of $\mathbf{y}$, also has a Gaussian distribution,

$$p(\mu) = \frac{1}{\sigma_\mu (2\pi)^{1/2}} \exp\left(-\frac{(\mu - \mu_0)^2}{2\sigma_\mu^2}\right). \qquad (24)$$

It is only necessary to determine μ_0 and σ_μ. After a little algebra, these are

$$\mu_0 = A\sum qs \qquad (25)$$

$$\sigma_\mu^2 = \sigma_y^2 \sum q^2. \qquad (26)$$

$p(S \rightarrow G)$ can be obtained by integrating $p(\mu)$ from 0 to ∞ (as $\mu = 0$ to ∞ is the region where the object will be mistakenly classified as a galaxy). This yields

$$\begin{aligned} p(S \rightarrow G) &= \tfrac{1}{2}\,\mathrm{erfc}\{-\mu_0/(\sigma_\mu\sqrt{2})\} \\ &= \tfrac{1}{2}\,\mathrm{erfc}\{-A\textstyle\sum qs/(\sigma y\sqrt{2\sum q^2})\}. \end{aligned} \qquad (27)$$

Let I be the magnitude of the star $A\mathbf{s}$ in the ordinary astronomical sense (adding up pixels). That is, $I = A\sum s$. Then

$$p(S \rightarrow G) = \tfrac{1}{2}\,\mathrm{erfc}\{-I\textstyle\sum qs/[\sigma_y \sum s\sqrt{2\sum q^2}]\}. \qquad (28)$$

Inverting, this yields

$$I = (\sigma_y \textstyle\sum s\sqrt{\sum q}^2/\sum qs)\{\sqrt{2}\,\mathrm{erfc}^{-1}[2p(S \rightarrow G)]\} \qquad (29)$$

for the magnitude limit of the classifier. Call the function in braces $t[p(S \rightarrow G)]$. It is tabulated in Table II, both in intensity units and as a magnitude correction.

For an error rate of about 1 in 7 the magnitude correction vanishes. This will then be defined as the "magnitude limit" of the classifier.

Table III shows how the examples that have been displayed fared.

This compares to an observed classifier limit of about 19.5 or 20.0 mag. The "detection limit" is the magnitude of a star whose maximum pixel value is $2\sigma_y$ above the sky level. The predicted limit for the fixed background classifiers, Figs. 1–6, is, of course, a bit too optimistic. Neither background shifts or centering errors are taken into account by this formalism. The limits for Figs. 4–6 (which use a Hubble law of ⅔″ core radius) are a bit higher because the classifier boundary of $\phi = 1$, which is assumed here, was set a bit too high.

The floating background classifier, Figs. 9–11, has a lower predicted magnitude line. To some extent, this classifier does indeed have a larger dispersion. But it does treat plate variations (which are outside of the theory), more realistically. Only centering errors are not included in this magnitude limit. It agrees rather well with the observed limit.

TABLE II. Theoretical error rate as function of magnitude.

$p(S \rightarrow G)$	$t[p(S \rightarrow G)]$	$-2.5 \log\{t[p(S \rightarrow G)]\}$
0.000 . . .	∞	$-\infty$
0.005 . . .	2.576	−1.03
0.01	2.326	−0.92
0.02	2.053	−0.78
0.05	1.645	−0.54
0.1	1.281	−0.27
0.133 . . .	1.000	0.00
0.2	0.841	+0.19
0.3. . . .	0.269	+142
0.4	0.127	+2.24
0.5	0.000	$+\infty$

TABLE III. Magnitude limits for data presented in figures.

Field 1		Field 2		Field 3	
Fig. 1	21.06	Fig. 2	21.09	Fig. 3	21.04
Fig. 4	21.51	Fig. 5	21.62	Fig. 6	21.52
Fig. 9	20.00	Fig. 10	20.11	Fig. 11	20.00
detection limit	21.15	detection limit	21.23	detection limit	21.08

In conclusion, the classifier derived here is valid down to about a magnitude above the plate limit and is valid within the range that the picture data are linear. In either the fixed or floating background versions, the classifier is of a simple linear form, which can be quickly implemented on a computer. On a particular sample of data (plate or electronic detector frame), it is calibrated by brighter (but still unsaturated) star images. However, it requires no detailed examination of faint objects on a plate by plate basis. Within the model used to derive each classifier, that classifier is the optimal classifier. One could not even in principle do better.

On the debit side, the classifier is sensitive to things outside of its model, such as plate defects, double and overlapping objects, and centering errors. These must be resolved by other tests. Bright, saturated objects also require a different classifier.

I thank James Gunn, Michael Shantz, Peter Young, and John Hoessel for helpful discussions. I also thank Gilbert McCann for time on his Image Processing system, upon which the early stages of these ideas were explored, and Richard Green for help in the digitization of the plate material. This work was supported by a grant, No. AST 78-24842 from the National Science Foundation.

APPENDIX A

In this appendix, the choice of the maximum likelihood value for the magnitude indicator A is more carefully justified.

Repeating Eq. (7),

$$p(\mathrm{y}|S,A) = \frac{1}{\sigma_{\mathrm{y}}(2\pi)^{n/2}} \exp \left(-\frac{1}{2\sigma_{\mathrm{y}}^2}[(\mathrm{y} - A\mathrm{s})\cdot(\mathrm{y} - A\mathrm{s})]\right).$$

Now, $p(\mathrm{y},A|S)$, the joint probability of getting y and a particular value of A, given that the object is a star, is

$$p(\mathrm{y},A|S) = p(\mathrm{y}|S,A)p(A), \tag{A1}$$

where $p(A)$ is the total probability of getting a particular value of A. Essentially, $p(A)$ is a luminosity function for stars, normalized to relative units. Then

$$p(\mathrm{y}|S) = \int p(\mathrm{y},A|S)\, dA = \int p(\mathrm{y}|S,A)p(A)\, dA, \tag{A2}$$

where the integral is over all positive A.

Define α to be the maximum likelihood value of $A = (\mathrm{y{\cdot}s})/(\mathrm{s{\cdot}s})$. Adding and subtracting $\alpha\mathrm{s}$ from the $\mathrm{y} - A\mathrm{s}$ terms and rearranging

$$p(\mathrm{y}|S) = \frac{1}{\sigma_{\mathrm{y}}(2\pi)^{n/2}} \exp \left(\frac{1}{2\sigma_{\mathrm{y}}^2}[(\mathrm{y} - \alpha\mathrm{s})\cdot(\mathrm{y} - \alpha\mathrm{s})]\right) \times \int \exp \left(\frac{-1}{2\sigma_{\mathrm{y}}^2}[(A-\alpha)^2\mathrm{s}\cdot\mathrm{s} + 2(\mathrm{y}\cdot\mathrm{s} - \alpha\mathrm{s}\cdot\mathrm{s})]\right) \times p(A)\, dA. \tag{A3}$$

By the definition of α, the last term in the exponential under the integral vanishes. Inserting the definition of α,

$$p(\mathrm{y}|S) = \frac{1}{(2\pi)^{n/2}} \exp \left[-\frac{1}{2\sigma_{\mathrm{y}}^2}\left(\mathrm{y}\cdot\mathrm{y} - \frac{\mathrm{y}\cdot\mathrm{s}_2}{\mathrm{s}\cdot\mathrm{s}}\right)\right] \frac{1}{\sigma_{\mathrm{y}}} \times \int \exp \left(\frac{1}{2\sigma_{\mathrm{y}}^2}(A-\alpha)^2\mathrm{s}\cdot\mathrm{s}\right) p(A)\, dA. \tag{A4}$$

If (1), the probability $p(A)$ is slowly varying over changes in A of order $\sigma_y/(\mathrm{s{\cdot}s})^{1/2}$, and (2) the limits on the integration are at least of order $\sigma_y(\mathrm{s{\cdot}s})^{1/2}$ from the smallest α of interest (i.e., for the limit $A = 0$, that $\alpha > \sigma_y/(\mathrm{s{\cdot}s})^{1/2}$, then

$$1/\sigma_y \exp[-(A-\alpha)^2\mathrm{s}\cdot\mathrm{s}/2\sigma_y^2] \rightarrow 2\pi\mathrm{s}\cdot\mathrm{s}\delta(A-\alpha), \tag{A5}$$

and

$$p(\mathrm{y}|S) = (\sigma_y \sqrt{2\pi\mathrm{s}\cdot\mathrm{s}}\, p(A)|_{A=\alpha}) \frac{1}{\sigma_y(2\pi)^{n/2}} \times \exp \left[-\frac{1}{2\sigma_y}\left(\mathrm{y}\cdot\mathrm{y} - \frac{(\mathrm{y}\cdot\mathrm{s})^2}{\mathrm{s}\cdot\mathrm{s}}\right)\right]. \tag{A6}$$

For reasonable profile shapes, $A = \sigma_{\mathrm{y}}/(\mathrm{s{\cdot}s})^{1/2}$ can be shown to be equivalent to an object whose central components are of order (one sigma) above the background. Condition (1) then says that $p(A)$ should be slowly varying when the central components of the star image change intensity by one sigma.

Except for the factor in parentheses in front of the exponential the result is equivalent to Eq. (9). The presence of that factor is due to the difference in models. In Eq. (9), the brightness indicator A is assumed to *always* have the maximum likelihood value, while in Eq. (A6), A has a distribution of values. In the classifier, this term is absorbed into the (usually ignored) bias factor and essentially results in the renormalization of the units of the probabilities $p(S)$ and $p(G)$. Instead of the total ratio of galaxies to stars, the bias factor becomes the ratio of galaxies to stars at the particular magnitude of interest. As before, w (now a function of magnitude) can be chosen to cancel this factor.

APPENDIX B

The theory given earlier can be modified to allow for coordinate shifts in the templates. Assume that s is a function of some (shiftable) coordinate system **x**. Replacing s by s(**x**), Eq. (7) becomes

$$p(\mathrm{y}|\mathrm{S}) = 1/\sigma_{\mathrm{y}}(2\pi)^{n/2} \exp(-1/2\sigma_{\mathrm{y}}^2 [[\mathrm{y} - A\mathrm{s}(\mathbf{x})]]\cdot[[\mathrm{y} - A\mathrm{s}(\mathbf{x})]]). \tag{B1}$$

Simultaneously maximizing this with respect to **x** and A (i.e., $\nabla_{x}p(\mathrm{y}|S) = 0$ and

$$\frac{\partial}{\partial A}[p(\mathrm{y}|S)] = 0,$$

yields

$$\mathbf{y} \cdot [\nabla_x \mathbf{s} - \mathbf{s}(\mathbf{s} \cdot \nabla_x \mathbf{s})/(\mathbf{s} \cdot \mathbf{s})] = 0 \quad \text{(B2)}$$

as a condition on **x** for object centering. This is equivalent to finding a zero in the convolution of **y** with the function in brackets. At present it is my feeling that this algorithm is too slow to be used in star/galaxy classification. It has been my hope that the raw coordinates generated by the algorithm which scans the field for objects would be adequate. However, if more coordinate refinement turns out to be necessary, the above algorithm is as fast or faster than the other centering algorithms mentioned.

Equation (B2) can also be applied to the finding of the centroid of a histogram. In that case **y** is the histogram and **s** is the function used to model the histogram, in this case a Gaussian. As **s** is an even function, the righthand term vanishes, and the result is the same as the sky background finding algorithm mentioned earlier in the paper. Of course, the statistics of the histogram are not Gaussian as assumed here, but (for a finite size picture frame) binomial. However, near the peak of the histogram (which is what matters here) the statistics are close enough to Gaussian, for all practical purposes.

REFERENCES

Duda, R. O., and Hart, P. E. (1973). *Pattern Classification and Scene Analysis* (Wiley, New York), Chaps. 2 and 3.

Gunn, J. E., and Oke, J. B. (1974). Astrophys. J. **195**, 255.

Jarvis, J. F., and Tyson, J. A. (1979). Proc. SPIE (Instrumentation in Astronomy, III) **172**, 422.

Kibblewhite, E. J., Bridgeland, M. T., Hooley, T., and Horne, D. (1975). *Image Processing Techniques in Astronomy*, Proceedings (Reidel, Dordrecht), **54**, 245.

Kron, R. G. 1978, Ph. D. thesis, Univ. California, Berkeley, "Photometry of a Complete Sample of Faint Galaxies."

Lorre, J. J., Benton, W. D., and Elliot, D. A. (1979). Proc. SPIE (Instrumentation in Astronomy, III) **172**, 394.

Pratt, N. M., Martin, R., Alexander, L. W. G., Walker, G. S., and Williams, P. R. (1975). *Image Processing Techniques in Astronomy*, Proceedings (Reidel, Dordrecht), **43**, 217.

Thuan, T. X., and Gunn, J. E. (1976). Publ. Astron. Soc. Pac. **88**, 543.

CHAPTER 3

Reprinted from SPIE Vol. 264—Applications of Digital Image Processing to Astronomy

20

A faint galaxy counting system

William L. Sebok
Palomar Observatory, California Institute of Technology 105-24, Pasadena, California 91125

Abstract

A system for detecting and classifying faint objects on astronomical digital pictures is described. Algorithm and implementation considerations are given, and its performance is evaluated. The optimal "phi" classifier (briefly described) is used to separate stars from galaxies. The major emphasis is on speed. A 512 by 1536 pixel picture can be scanned and fully processed in about 5 minutes total time on a PDP 11/34 minicomputer. On plates taken with the Palomar 1.2 m Schmidt telescope, stars can be reliably separated from galaxies down to a limiting red magnitude of 19.6. Objects down to about 21st magnitude in red light can be detected. This system is part of a larger system which includes a highly modified David W. Mann two-dimensional scanning microdensitometer for image acquisition, and an extensive image processing system, all operating under compatible formats and conventions.

Introduction

Even in this age of digital detectors, the large Schmidt telescopes are still one of the resources of modern astronomy. A 14 by 14 inch plate taken on the 1.2 m (48") telescope at Mt. Palomar digitized at a typical resolution of 10 micron, produces a file about 30000 by 30000 pixels across, or about a billion pixels!! It will be a while before any electronic detector can match that field of view. This naturally leads to the question of what one is to do with these data. One of the more obvious task is to construct a list of all of the objects on the plate. Thanks to modern digital computers, this is now practical.

This idea is not quite so rare any more. Several groups have tried their hand at it (see references 1,2,3,4, and 5, for example). However, automatic collection of objects, and pattern recognition in general, are rather difficult problems. It is valuable at this stage of the art to have different approaches tried. My own approaches to this problem were influenced by:

1) an improved theretical understanding of the star/galaxy separation problem[6].
2) a desire to be able to handle some subset of a moderately large library of Schmidt plates (100 plates) in a reasonable amount of time. The emphasis, therefore, has been on "fast" algorithms.

Recently, a fast new plate scanning machine has come into operation at Caltech. It is basically a refurbished David W. Mann two-dimensional microdensitometer. The most drastic change in it is the replacement of the phototube by a diode array, a RETICON. The scanning rate for this machine is limited only by the speed of the computer and associated output devices. It can easily scan above 10 thousand pixels/second. In fact, the scanning speed which can be reached depends on the number of software options invoked. With the options used when scanning for the pattern recognition software, the speed is more like 3.7 thousand pixels/second. Even so, it is a strain for any pattern recognition system to keep up with the scanner. My goal was to build a pattern recognition system to run concurrently on the computer connected to the scanner, that would slow the scanner down by no more than a factor of 2.

Operating System and Environment

All software runs on a PDP 11/34 computer connected to the plate scanner, much of it executing while the scanner itself is running. The software is all in the FORTH language (with much recourse to FORTH's internal assembler for speed in critical routines). It runs under a stand-alone FORTH operating system, originally of FORTH Inc. vintage 1976 but heavily modified to support the needs of the plate scanner. Among the more important modifications affecting the plate scanner are:

1) a queuing disk handler that allows disk I/O to be issued from within (some other device's) interrupt handler. It also generally tightens the scheduling of disk accesses (the use of the disk being a major bottleneck in the system).
2) implementation of memory management, allowing programs larger than 56 K bytes of memory to run on the computer. The pattern recognition plus plate scanning software together take up 136 K bytes of memory.

There have been three major inter-compatible systems that I have written for this computer. They are:

1) An image processing system.
2) The plate scanner driver software, and associated setup software for the plate scanner (detector calibration, coordinate calibration, magnitude calibration, etc.).
3) The pattern recognition itself.

The image processing system has been a useful thing to have available. For instance, magnitude zero points of the Schmidt plates have been set by reference to surface photometry done with the Silicon Intensified Target (SIT) camera on the Palomar 60 inch telescope.[7] This photometry was reduced on the image processing system. Various random operations and tests on scanner data are made easy by invoking routines already available on the image processing system.

The SCANNER process

When scanning a plate, the strategy has been to break the area of the plate into bite-size chunks (call them "frames" or "pictures"). A frame size of 512 by 1535 pixels is typical, filling up a Digital RK05 disk with a bit of room left over for the other files that the pattern recognition routines require. First, the scanner process scans a frame into disk #1. When this is finished, it start to scan a frame into disk #2. At the same time the pattern recognition process starts to look at the frame on disk #1. The scanner process waits for the pattern recognition process to finish before starting to scan onto disk #1 again. Then the scanner process starts on disk #1 while the pattern recognition process starts on disk #2. Scanning and pattern recognition occur simultaneously, with the scanner process always one frame ahead.

Several things must be done to the "raw" output of the scanner's detector before it can be processed by the pattern recognition routines. The following things are done by the foreground process (which looks like an interrupt handler) that services the scanner:

1) The zero-light response of the detector is subtracted off (Dark counts are subtracted).
2) The various sensitivities of each detector are renormalized to one sensitivity (balance factors are applied).
3) A nonlinear transformation is applied to the data to convert the data (which are measurements of photographic transmission) to estimates of the intensity of light which originally fell on the plate.
4) The data from different regions of the frame (typically 128 by 128 pixels wide) are histogrammed by intensity. After the frame is scanned, the histograms are reduced to a file of sky backgrounds for each region of the frame. This file is passed, along with the frame itself, to the pattern recognition process, which subracts off the sky as the frame is read in. Bilinear interpolation between sky measurements is done there, so as not to produce sharp edges between regions. The sky estimate derived from the histograms is the zero of the convolution of the histogram with the derivative of a Gaussian. The Gaussian, computed earlier as part of the setup procedure, is intended to have the same width as the best-fit Gaussian for the histogram. If the histograms were really Gaussians of that width, this would be the optimal procedure for extracting the sky (see reference 8). The implicit assumption here is that, while the value of the sky changes over the plate, the width of the dispersion of the data around the sky does not. This is true to first order. The algorithm is fast. Generally, when the most likely values of the sky can be predicted, it requires evaluation of the convolution at only two points.

The FINDER process

The system maintains a "moving window" which travels through the picture. This "window" consists of a set (typically 15) of buffers, each of which contains a line from the picture. When the window is moved, a new line is read into a buffer previously occupied by an old line. Optionally, the previously determined sky background is subtracted from the new line and that line written back into the picture file. An extra line is read in look-ahead before it is needed. A great deal of speed is gained by having these two disk accesses scheduled asynchronously by the queuing disk handler. A "window control block" is used to expedite accesses to lines in the moving window.

As a starting point, local maxima in the data are considered to be objects. However, noise in the raw data tend to produce many spurious local maxima. Therefore, a three line moving window of smoothed data is generated from the raw data and local maxima are noted in this smoothed data. Here the seeing disk size sets the size of the boxcar. This is a 5 x 5 pixel box for a plate from the 1.2 meter telescope scanned at a resolution of 11 μm/pixel. Ideally, something optimal like a Weiner filter would be better. However, filtering and the object detection process are potentially one of the main bottlenecks in the system. Here, all of the data must be examined, while later steps only need examine portions. As there exists a fast algorith for boxcar smoothing which is independent of the size of the boxcar, boxcar smoothing is used.

The central line of the smoothed data is scanned from beginning to end, an entry being made whenever the smoothed data rises above a threshold THRESH (typically set about 5 times the standard deviation of the smoothed data). A different entry is made in the structure when the data goes below this threshold. The number of pixels counted above some high threshold (saturated pixels) is noted in this entry. After rising above THRESH the algorithm enters "uphill" mode. On a transition from "uphill" to "downhill" a test for a local maximum is made. The six pixels adjacent to the local maximum candidate on the other two lines of smoothed data are examined to determine whether the candidate is indeed a local maximum in two dimensions. If it passes this test, it is retained if it is also above a threshold 2THRESH (typically set about 1.5 times THRESH).

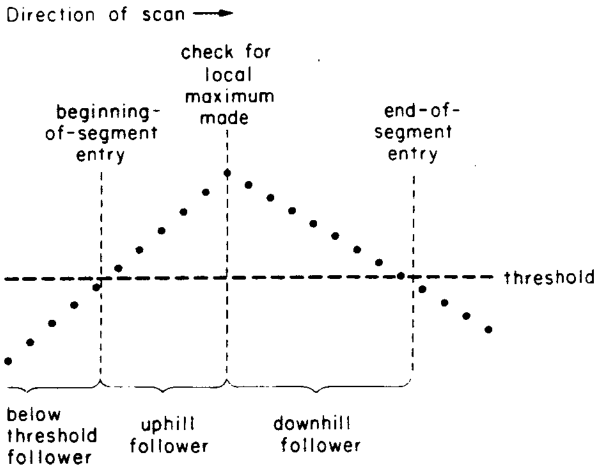

Figure 1. Local Maximum is detected on a transition from "uphill" to "downhill".

After a succesful local maximum detection, the raw (unsmoothed) data contained within the moving window are examined in a circular aperture centered on the position of the local maximum. Typically, the window size is 15 pixels in diameter. Various quantities are calculated on the data in this aperture. They are: (if the pixel values are labeled I_{ij}, where i and j are pixel indices on the two dimensions of the picture, (i,j)=(0,0) at the center of the aperture ≡ at the local maximum)

1) the quantities $\sum_{ij} I_{ij} s_{ij}$, $\sum_{ij} I_{ij} g_{ij}$ where s_{ij} and g_{ij} are star and galaxy models, respectively.

2) $\sum_{ij} I_{ij}$, the integrated brightness.

3) the moments, $\sum_{ij} i\, I_{ij}$, $\sum_{ij} j\, I_{ij}$, $\sum_{ij} i^2 I_{ij}$, $\sum_{ij} ij\, I_{ij}$, $\sum_{ij} j^2 I_{ij}$

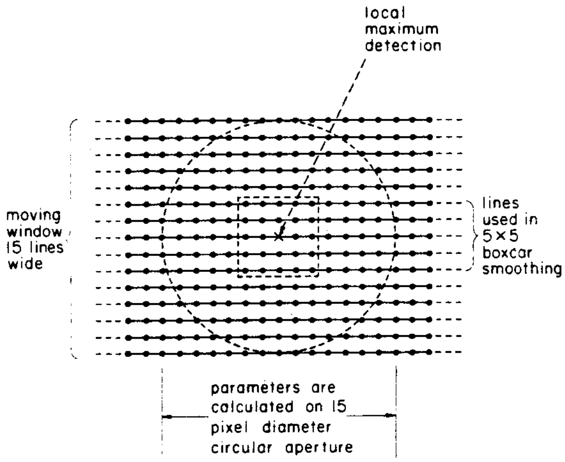

Figure 2. An aperture is drawn on the moving window when a local maximum is detected.

From the first set of quantities is calculated the classifier:

$$\phi \equiv \frac{\sum I_{ij} g_{ij}}{\sum I_{ij} s_{ij}} \left(\frac{\sum s_{ij}}{\sum g_{ij}}\right)^{\frac{1}{2}} \qquad 1)$$

If the templates s_{ij} and g_{ij} really resembled nature, if real stars and galaxies differed from their templates by only a scale factor, and if the statistics of noise were Gaussian, then the decision surface $\phi = 1$ would optimally classify objects into stars ($\phi<1$) or galaxies ($\phi>1$). The star template is constructed from an average of the profiles of about ten stars. The galaxy template is formed from the convolution of the star template with a galaxy model, typically a Hubble law of core radius about 1/3 second of arc. For a more detailed discussion of the properties and limitations of this classifier, see reference 6 .

From the sum of the pixels inside the aperture, the "Aperture Magnitude" is immediately computed:

$$m = m0 - 2.5 \log \sum I_{ij} \qquad 2)$$

Using moments, a paraboloid is fit to the data:

$$I = Ai^2 + Bij + Cj^2 + Di + Ej + F \qquad 3)$$

If N = number of pixels inside aperture, $Q = \sum i^2 = \sum j^2$, $P = \sum i^4 = \sum j^4$, $R = \sum i^2 j^2$, then $B = (\sum ijI)/R$, $D = (\sum iI)/Q$ and $E = (\sum jI)/Q$, and A,C, and F can be obtained by solving:

$$\begin{pmatrix} \sum I \\ \sum i^2 I \\ \sum j^2 I \end{pmatrix} = \begin{pmatrix} N & Q & Q \\ Q & P & R \\ Q & R & P \end{pmatrix} \begin{pmatrix} F \\ A \\ C \end{pmatrix} \qquad 4)$$

All of the coefficients A to F are linear combinations of the moments, with coefficients that, for any particular aperture, only need to be determined once. From these six coefficients, six quantities which are more useful can be derived.

$$\Delta = 4AC - B^2 \qquad 5)$$

$$\text{orientation} = \tfrac{1}{2}\arctan(B/(A-C)) \qquad 6)$$

$$\text{ellipticity} = 1 - \text{minor axis/major axis}$$

$$= \frac{2}{\Delta}\sqrt{(A+C)^2 - \Delta}\ \left((A+C) - \sqrt{(A+C)^2 - \Delta} \right) \qquad 7)$$

$$P = \text{peak value} = F + (BDE - CD^2 - AE^2)/\Delta \qquad 8)$$

$$\text{radius} = \sqrt{P/\sqrt{\Delta}} \qquad 9)$$

$$\Delta i = (EB-2DC)/\Delta = \text{Deviation of centroid from } i \text{ coordinate of Local Maximum.} \qquad 10)$$

$$\Delta j = (DB-2EA)/\Delta = \text{Deviation of centroid from } j \text{ coordinate of Local Maximum.} \qquad 11)$$

The innermost zone of the paraboloid can be matched with a Gaussian. The radius calculated here is defined to be equivalent to the standard deviation of that Gaussian.

Several tests are applied immediately. The local maximum is rejected if:

1) $\sum I < 0$ (unlikely)
2) $A + C \geq 0$ (Guarantees that the paraboloid is concave downward).
3) $\Delta \leq 0$ (local maximum is rejected if surface would be hyperbolic or degenerate.)
4) $|\Delta i| \geq \varepsilon$ or $|\Delta j| \geq \varepsilon$ (local maximum is rejected if centroid is too far from the local maximum's nominal position).

Condition #3 eliminates many spurious detections on the edges of bright stars. Spurious detections generally occur when the surface brightness is greater than THRESH and there is little gradient. Then, any little wiggle in the data will produce a detection. Most of the spurious local maxima on the edge of bright stars will be included in the envelope drawn around the star and not result in spurious objects. However, the surface brightness contours, and therefore the envelope, around a bright star are very convoluted, and it is possible for a piece of this envelope to split off and be detected as a separate object. The defense against this condition is the second threshold 2THRESH. One of these spurious objects is unlikely to have a local maximum much above THRESH, being close to the bright star's envelope where the level of the data is around THRESH. Therefore, almost all of these spurious objects will have local maxima which are rejected. The classification routines reject objects which contain no local maxima.

Condition #4 rejects many garbage objects, those which are so irregular that the centroid of the data within the aperture is not close to the local maximum. ε is typically 2.5 pixels.

At this stage, the information gathered consists of a list of segments describing envelopes around the objects on a frame, segments which as of yet are unconnected. Embedded in those segments are local maxima entries.

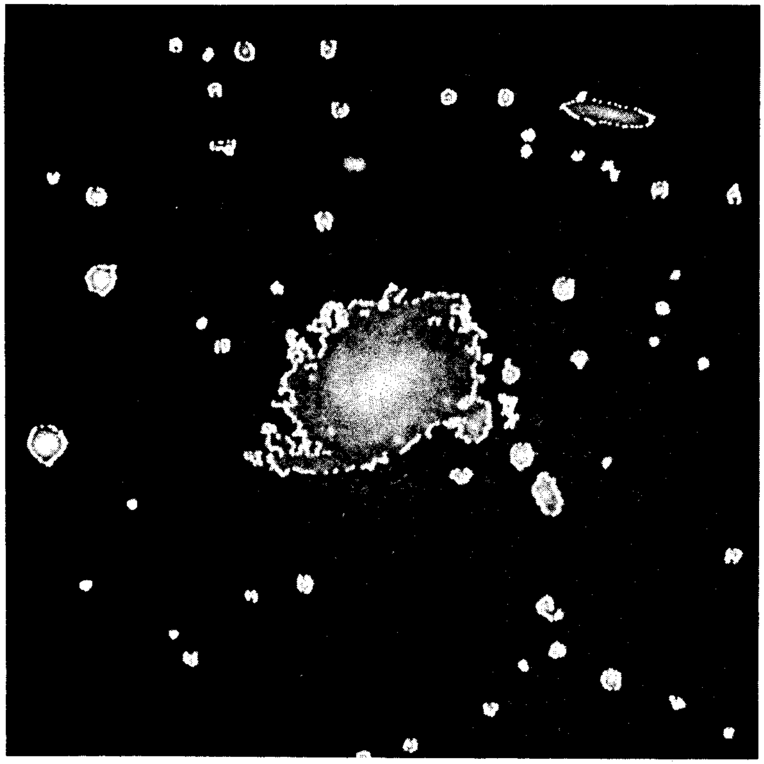

Figure 3. A 512 by 512 pixel area scanned by the Caltech plate scanner (digitally enhanced for this display). Envelopes and local maxima (white dots) of detected objects are shown.

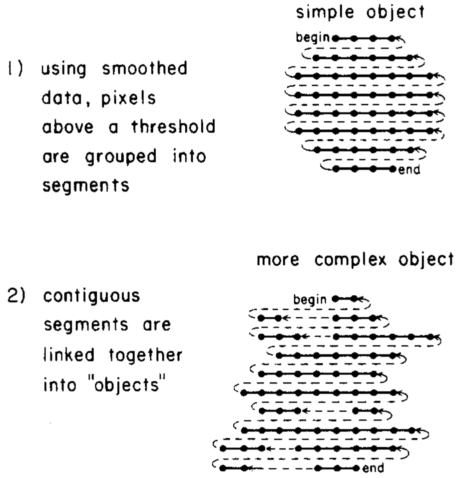

Figure 4. Segments are linked together to form Objects.

The CLASSIFIER process

After the whole frame is scanned, control passes to the classification routine. The classification routine reads in the list of segments created by the finder routine, and links together all segments which touch into a linked list structure. Among the complications handled by the routine are:

1) Branches, disconnected segments on a particular line of data which are discovered to be connected through a segment on some other line.
2) Holes, bifurcations which later reclose.

The outer envelopes of bright objects typically have ragged edges. It is necessary to handle the topology of this envelope correctly to avoid splitting off pieces of those objects, pieces, which would otherwise be treated as separate objects.

Each topologically distinct set of connected segments is considered to be an object. All of the information on each object is gathered together. The information on the N brightest local maxima (N=7 typically) is collected into a table. A pointer is set so th the local maxima parameters (the aperture magnitude, value of Phi, etc.) refer to the brightest (in aperture magnitude) local maximum.

In addition to the local maximum parameters, parameters are calculated on the envelope of each object. These are:

1) Number of pixels in the envelope (converted to a radius).
2) Number of saturated pixels within the envelope (also converted to a radius).
3) WIDTH of envelope (Length of largest segment in the object).
4) LENGTH of envelope (Distance between the midpoints of the first segment and of th last segment in the object).
5) FLAGS which indicate whether any of the margins around the picture were touched (margins are generally set one "aperture radius" away from the edges of the frame There is always enough room to compute the aperture parameters).

After all of the information on the object is collected, an attempt is made to classi the object.

If the envelope of the object touched the 1) left (first pixel in line), 2) right (last pixel), 3) top (first line in frame), or 4) bottom (last line) margin, then the object is output to one of four edge files (selected according to the above priority) and classification is postponed to the postprocessing stage (described below). At postprocessing the missing part of the object can be supplied from the adjacent frame.

The edge file consists of:
1) the collected envelope and local maxima information.
2) A formatted dump of the raw picture data in the object. A short header describing the size and position of each segment is followed by the raw pixels in that segment.

An object which touches more than one edge is only written to one edge file. However, every edge an object touches is noted in the FLAGS item for the object.

The number of local maxima is next examined. Objects with no local maxima are discarded. This eliminates objects which are either too faint to be properly handled, or which are spurious objects on the outer edges of bright stars.

If max(WIDTH/LENGTH , LENGTH/WIDTH) is greater than some threshold (typically 8), then the object is classified as a "streak" and written out to the object list. This test catches things like asteroids, meteor trails, and airplane trails. There are almost always several such trails running most of the way across each Schmidt plate.

Next, a major branch in the classification process takes place. The phi classifier separates stars from galaxies very well. However, it only is valid for linear data. For brighter objects, those with saturated pixels, a different classifier is used.

As part of the setup procedure, the number of saturated pixels is plotted against the number of pixels in the envelope. Most of the bright objects are stars and fall on a fairly well defined curve. An envelope is hand-drawn (using a cursor on a TV display) above and below this curve.

Back to the classification process: The number of saturated pixels is examined. If this number is greater than zero then the "bright object" classifier is used. The number of saturated pixels is compared to the number of pixels inside the object's envelope. If their relationship falls inside the hand-drawn curves, then the object is classified as a star and written out to the object list. Otherwise, the object is output to the "Eyeball" file for later human classification. The format for the entries in the Eyeball file is identical to the format of the Edge files. Besides bright galaxies, much garbage ends up in the Eyeball file where it can easily be rejected. Here was the point where it seemed better to rely on the still unexelled pattern-recognition abilities of the human brain, protected by the computer from the flood of ordinary objects.

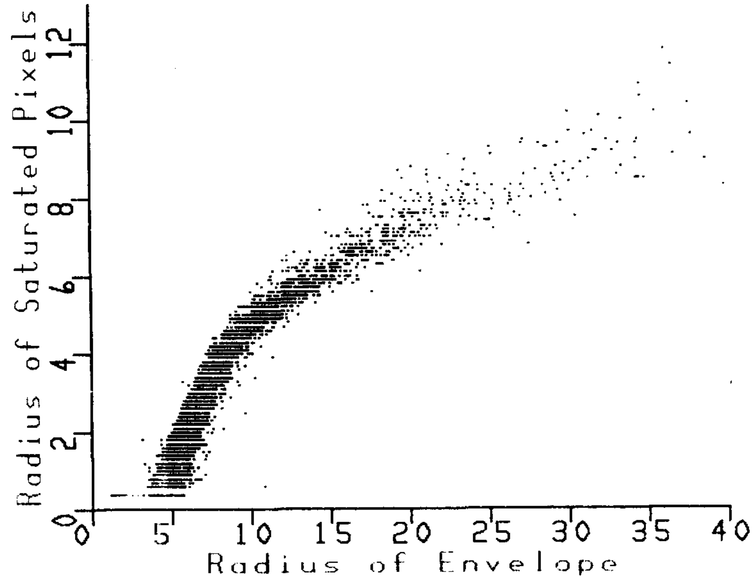

Figure 5. Plot of radius of envelope vs. radius of saturated part of image (both in seconds of arc) for all of the objects on one Schmidt plate which were classified as stars. An envelope drawn around this curve is used to distinguish bright stars from galaxies.

If there were no saturated pixels in the object, then the number of local maxima in the object is examined. If this number is greater than one, then the aperture magnitude is compared with a threshold to see whether it is worthwhile to output the object to the Eyeball file. Otherwise the object is classified "blend" and output to the object list.

Finally, the objects with no complicating factors (no saturated pixels, only one local maximum) are classified. Phi is compared to a threshold (typically 1). If it is greater than this threshold it is classified "galaxy" and output to the object list. If it is less than this threshold it is classified "star" and output to the object list.

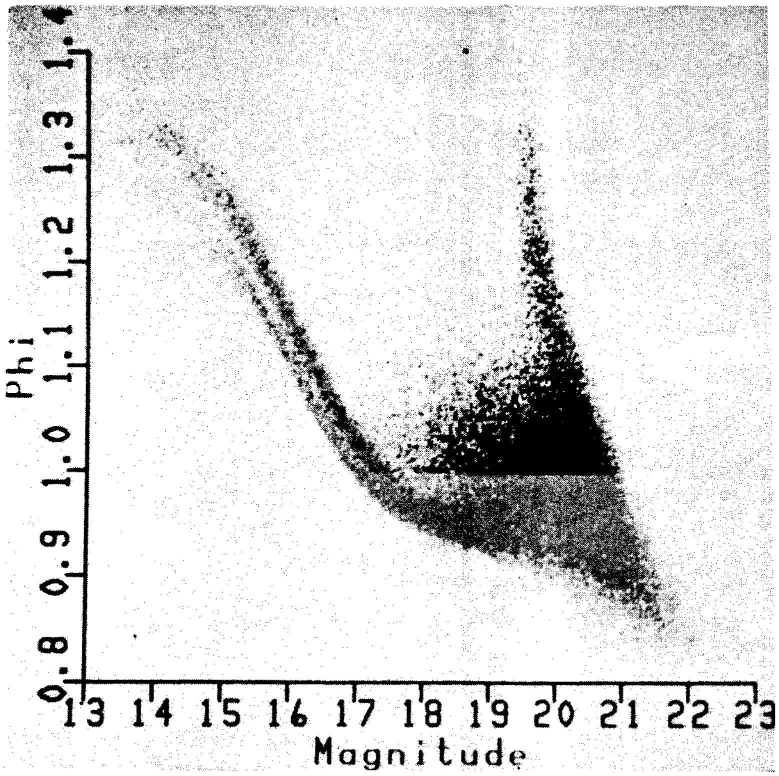

Figure 6.
Phi classifier versus brighter of aperture or isophotal magnitudes. All 130000 objects that were on one Schmid plate used. Light color points were classified as stars, dark color points were classified as galaxies.

Summary of Classification Decisions

```
If (at edge) then {save object in Edge file}

If (length>>width or width>>length) then {object is probably asteroid, meteor trail,
                                           or airplane trail, classify as streak}

If (saturated) then {decide if it is a star by comparing number of saturated pixels in
                     envelope to total number of pixels in envelope}

     If (starlike) then {classify as star}
                   else {save object in Eyeball file}

If (not saturated) then
     If (blended) then
          If (brighter than some threshold) then {save object in Eyeball file}
                                             else {classify as blended}

     If (not blended) then
          If ($\phi$> threshold) then {classify as galaxy}
                                 else {classify as star}
```

Output

For each classified object, a short entry is placed in the "classified object" list. This list contains information thought to be of later use (or useful as a check on marginally classified objects):

1) ID a unique ID number for every object.
2) RA 1950 right ascension
3) DEC 1950 declination
4) a flag indication the classification of the object.
5) Number of local maxima in the object.
6) Aperture Magnitude.
7) Isophotal Magnitude. This is computed within the same isophote as is used to trace out the envelope of the object. Isophotal Magnitude is only computed for the brighter objects. The criteria used is that the radius of the envelope of the object be larger than a threshold. This threshold is set slightly smaller than the

radius used for aperture magnitudes.

8) Ellipticity
9) Orientation
10) Core Radius
11) Peak Surface Brightness

quantities derived from fitted paraboloid.

12) Number of pixels in envelope, converted to a radius.
13) Number of saturated pixels in envelope, also converted to a radius.
14) Phi, the value of the star / galaxy classifier.

All of the quantities are saved in a plate independent form. For example, distances are expressed in seconds of arc rather than in pixels, and units of surface brightness rather than pixel intensities are used for the peak brightness. The Δi and Δj items derived from the fitted paraboloid are used to pin down the position of the object to a fraction of a pixel.

Postprocessing

After a plate is scanned, there is still a postprocessing step to do. The object lists for each separate frame are gathered together. The objects in the various "Eyeball" files are displayed on a television screen and the information gathered on that object is printed on a computer terminal. The operator then makes his choice as to the identity of the object. Pieces of objects which straddle the edge of a frame are merged with their pieces from the adjacent frame. The merged objects are classified by the same criteria used for the rest of the objects. If there are any objects in the Eyeball file produced by the processing of the merged file, then the Eyeball routine is entered immediately. The end result of the postprocessing is a master list of objects.

Timings and Behavior

The Plate scanner went into operation around the beginning of 1980. I have been using the pattern recognition routines on Schmidt plates since June 1980. Already (2nd week of August 1980), two plates have been scanned. There were about 130,000 objects on the first plate, down to the limit of 21st magnitude in red light. On many of the graphs of phi versus magnitude from different parts of the plate, an envelope can be drawn around the region occupied by starlike objects. This envelope crosses the decision boundary of $\phi=1$ at about 19.6 magnitude in red light. The conclusion is therefore that stars are adequately separated from galaxies down to 19.6 magnitude. There is some separation for another magnitude below this.

As mentioned earlier, the plate is broken up into 512 (width of RETICON) by 1536 (available disk space) pixel chunks. For plates from the Palomar 1.2 m Schmidt telescope, a pixel size of 10.8 microns = .722 seconds of arc is used, making the chunk size about 6.2 by 18.5 minutes of arc. A field of this size is scanned in about 3 min 30 sec, yielding an average counting rate of 3745 pixels/second. The pattern recognition process overlaps with the scanning process, the pattern recognition generally finishing about 40 seconds after the scanning process does. Average time for a scanning and processing cycle (overlap counted only once) is 4 min 14 sec, yielding an effective rate of 3096 pixels/second. The most likely time is actually more like 4 minutes even, the large amount of time needed to process the few bright stars skewing the distribution of processing times.

The PDP 11/34 computer running these programs is shared among many other users besides the plate scanner. A typical day's (or more exactly, night's) run on the scanner is something like 8 hours. A night's worth of work might consist of 108 of the previously mentioned frames. For each night's work there is about 2 hours of postprocessing to do. This lengthens the effective time to scan and process a frame to 5 min 6 seconds, yielding an effective scanning rate of 2570 pixels /second from plate to final list of objects. I think that I have met my goal that the pattern recognition processing slow down the scanner by less than a factor of two.

There are about 1000 of these frames usuable on a 14 inch wide plate from the 1.2 m Schmidt telescope. That corresponds to 9 nights work at 8 hours a night for the plate. With (unfortunately) one night's worth of setup time and two nights worth of postprocessing, plus an extra couple of nights lost for one reason or another, a plate can be processed from beginning to end in about two weeks.

Conclusion

The advent of a fast plate scanner at Caltech, plus the large amount of data available on plates from large Schmidt telescopes, have provided the challenge of designing automatic classification software that is as fast as the scanner. With a total effective rate from the scanning of the plate to the output of the final list of objects of 2570 pixels per second, this challenge has been met.

Acknowledgements

The author wishes above all to thank James E. Gunn for much help on all easpects of this project, part of my Ph.D thesis research. He, Fred Harris, and J.B. Oke have been extremely helpful during the long and difficult years of construction on the Mann Machine. Doug Rabin, Peter J. Young, and John Hoessel provided helpful comments, but most importantly the willingness to listen to me, helping me get things straight in my own mind. Finally, I thank the other astronomy graduate students for their tireless work in finding the bugs in my system, enabling me to correct them. This work was supported by grant No. AST 78-24842 from the National Science Foundation. Construction on the plate scanner was supported by NSF No. 75-22910.

References

1. Jarvis,J.F., and Tyson, J.A. (1979). Proc. SPIE (Instrumentation in Astronomy,III) **172**, 422.
2. Kibblewhite,E.J.,Bridgeland,M.T.,Hooley,T., and Horne, D. (1975). *Image Processing Techniques in Astronomy*, Proceedings (Reidel, Dordrecht), 54,245.
3. Kron,R.G. 1978, Ph.D. thesis, Univ. California, Berkeley, "Photometry of a Complete Sample of Faint Galaxies.".
4. Pratt,N.M.,Martin,R.,Alexander, L.W.G.,Walker,G.S., and Williams,P.R. (1975). *Image Processing Techniques in Astronomy*, Proceedings (Reidel,Dordrecht),43,217.
5. Shanks,T.,Fong,R.,Ellis,R.S., and MacGillvray,H.T. (1980) Mon.Not.R.Ast.Soc.,**192**,209.
6. Sebok,W.L. (1979). Astronomical Journal **84**,1536.
7. Kent,S.M. (1979) P.A.S.P. **91**,394.
8. Sebok,W.L. ibid. Appendix B

4.1 Observations

A total of four plates, taken on the Palomar 1.2 m. Schmidt telescope, were reduced to catalogs of galaxies. Their characteristics are listed in Table 4-1. These plates were taken to the sky limit with the red IIIaF emulsion and were arrayed in a two by two cluster on the sky. This pattern covered a total area about 10 degrees square, centered on $\alpha_{1950} = 10^h$ $\delta_{1950} = 47°$. There was a fair amount of overlap between adjacent plates. All of these plates were developed 8 minutes in a rocking bath with MWP2 developer. The plates were supplied with a log-linear calibration wedge, exposed in red light for about the same amount of time as the sky exposure.

Table 4-1 Plate Material

Field	Id	α_{1950}	δ_{1950}	position in cluster	seeing	date	exp.
111	PS23631	$9^h34^m37^s$	47°38'48"	NW	1.5"	Feb 16,1977	2^h
112	PS23591	$9^h37^m44^s$	43°14'54"	SW	1.5"	Jan 27,1977	3^h
113	PS23632	$9^h58^m13^s$	48°03'06"	NE	2.0"	Feb 16,1977	2^h
114	PS23634	$9^h57^m34^s$	43°33'41"	SE	1.0"	Feb 15,1977	2^h

More information

Field	l_{II}	b_{II}	density	fog	density above fog	Scanned on
111	170°.62	47°.05	1.20	0.49	0.71	June 2,1980
112	177°.50	48°.22	1.08	0.52	0.52	Sep 16,1980
113	169°.16	50°.80	0.99	0.38	0.61	Sep 6,1980
114	176°.19	51°.74	1.02	0.39	0.63	Jul 26,1980

In order to supply magnitude zero-points for the photometry of the Schmidt plates, surface photometry was obtained of two or three 13th, 14th, or 15th magnitude galaxies per plate. The most important criteria for these galaxies was that they have a well defined center point. Thus a position on a galaxy can be properly identified with a position on that galaxy's image on a plate. Calibrated digital pictures were obtained of each of the selected galaxies on the Direct SIT (silicon intensfied target) camera on the Palomar 1.5 m. telescope (see Kent, 1979). The red color in the photometric system of Thuan and Gunn (1976), was used for the calibration, as it is matched to the response of IIIaF plates and can be directly used to calibrate these plates.

4.2 Scanner

At the heart of this effort was the Caltech plate scanner. This scanner is a refurbished David W. Mann model 1032 two-dimensional micro-densitometer. The main refurbishment was the replacement of the light detector with a one-dimensional array of 512 silicon detectors, a Reticon. Light from a 250 W. quartz-halogen lamp is collimated and passed through an adjustable slit. An interference filter limits the bandpass to shortward of 700 nm. The slit is then focused onto a photographic plate which sits on a movable platform. The plate is then focused directly onto the reticon. This reticon sits on a thermo-electric cooler which cools the reticon to -40°C. A PDP 11/34 computer receives the data read from the reticon. This computer also controls the platform moving motors and reduces the data received from the reticon.

The spacing between pixels on the reticon is 1 mil. or 25.4 μm. The projected spacing of these pixels back onto the plate is selectable by a choice of 3 lenses mounted on a turret. For all of the scanning done in this project, the lens with the highest magnification was used. This lens gives a projected pixel spacing of 10.8 μm. At the 14.8

microns/arcsecond scale of the Schmidt plates, this yields a scale size of 0.722 arcseconds/pixel.

Two motors drive the 1 millimeter/turn screws that drive the scanning stage. These motors are of two types. The Y motor drives the stage parallel to the axis of the reticon. Since this motor is only used to position the stage it is a stepping motor. The X motor drives the stage perpendicular to the axis of the reticon. This motor is used to scan the photograph on the stage, forming a two-dimensional image out of lines from the one-dimensional reticon. With the emphasis here on smooth motion, this is a synchronous motor. Its speed is adjustable within factors of two by a gearbox. As the reticon is an integrating device, the pixel spacing along the direction of travel is equal to the product of the integration time and the motor speed. Since the motor speed can only be adjusted coarsely with the gearbox, any fine-tuning of the pixel size along the direction of travel must be done by adjusting the integration time. In particular, it is desirable to fine-tune the spacing of the pixels along the direction of travel in order to keep this spacing the same as the pixel spacing along the reticon, or, in other words, to keep the pixels square. The pixels were always kept square to at least a percent.

4.3 Plate Setup

A 14" by 14" plate does not quite fit on a 8" by 8" platen. The plate therefore had to be scanned in quarters. Fortunately, there were no obstructions preventing the plate from being laid on the platen. An advantage of the rather slow optics of the scanner (f2, slow for a plate scanner) was that there was little problem (except at the corners) of changes in focus caused by plate flexure. Before placing the plate on the platen the plate was carefully cleaned. The plate was placed on the platen with the edges aligned square with the edges of the platen. Some rather obvious landmark (usually the guide star) was selected and the coordinate system of the scanner was zeroed with the landmark at the origin. This process guaranteed that the coordinate system stayed roughly the same, at least closely enough so that some region of the plate could be selected as a "standard density" to tie down the photometric calibration of the scanner, and that this region could still be found even after the plate was moved.

With the plate in position, the scanner was moved to the lightest density likely to be encountered on the plate, and the intensity of light on the reticon adjusted (by a combination of neutral density filters and varying the slit width) so that the reticon was nowhere saturated. The alignment of the scanner was checked and readjusted if necessary and the scanner focused.

4.4 Scanning

When scanning a frame, one supplies the location of the center of the frame, the number of lines in the frame, the scan direction, a pixel size, and a motor speed. From the encoder coordinates of the center of the frame and from a coordinate system remembered by the computer, enough information is saved on the label of the frame to enable subsequent programs to compute the Right Ascension, Declination, and encoder coordinates of any pixel in the frame. From the center of the frame the computer calculates the location of the edge of the frame and moves the scanner two seconds travel time beyond that. Then the scanning motor is started. The computer monitors the encoders, starting the data-taking cycle when the scanner reaches the edge of the frame. When the desired number of lines is reached, a signal is sent for the motors to stop. Then, after a 100 ms. wait for the motors to stop and a wait for the subprocess previously created to finish, the Scanner process creates a subprocess to do pattern recognition on the frame which had just been measured. This subprocess (the Finder subprocess) executes at the same time that the Scanner subprocess proceeds to scan the next frame onto the other disk.

The data-taking cycle consists of:

1. Integrating light onto the the detector for a period of time.

2. reading out the detector.

3. Immediately reading out the detector again.

4. Sending the difference between the two readouts to the computer.

Before saving the reticon readout on the disk, the computer immediately does some processing on it:

1. Dark Current: A dark current estimate is subtracted from the readout.

2. Balance Factors: The readout is multiplied by balance factors.

3. HD Function: A nonlinear tranformation is done on the readout to convert the readout from estimates of plate transmission to estimates of the amount of light which originally fell on the plate.

4. Sky Tracking: Information is gathered from the readouts for use in a subsequent sky measurement and subtraction algorithm.

4.5 Calibration

Calibration of the plate scanner was a fairly lengthy process, typically requiring a full night to complete. Fortunately, the full calibration procedure did not have to be done every night or scanning would have been impractical.

4.5.1 Dark Count

To measure the dark count, the beam path of the scanner is blocked and about 500 to 1000 reticon readouts are averaged together. The dark count was measured rather frequently, at least at the start of every night. This was partly because it was easy to measure, but also because occasional changes were noticed in it at about the 2% level. These variations were only significant at the high densities found in the innermost portions of bright objects, regions in which scattered light was the dominant effect.

4.5.2 Flat Field

The scanner is moved to the previously decided "standard position" on the plate. Then, the scanner is started up and the previously decided "standard" number of reticon readouts are averaged together. The scanning motor is turned on during this process to average over any objects encountered on the plate. Averaging over a large region of the plate, even with the presence of objects, was found to be superior to averaging over a blank plate. Using two plates, it was found to be almost impossible to achieve a close enough match in the optical density and positions of the plates to measure the balance factors well enough to remove completely their effect from the data.

The main effect present in the balance factors was the imperfections of the slit. This completely dominated any change in sensitivity across the detector.

Once this measurement was made, its average across the reticon was noted. This gave an indication of the product of the current brightness of the lamp and the width of the slit. Renormalizing this average to a previously decided value then removed the effect of this day-to-day changes in sensitivity.

4.5.3 Coordinate Setup

Coordinate transformations on the scanner take place in two steps. The pixel row and column number (S,L) within a frame are converted to encoder coordinates (X,Y), which are then converted to Right Ascension and Declination. Each of these transformations must be separately calibrated.

4.5.3.1 Pixel To Encoder Coordinate Transformation.

The program to calibrate the (S,L) versus (X,Y) transformation first scans a frame at an arbitrary place on the plate. The user selects an object from this frame. The program then scans a pattern of frames on the plate, placing the object on different parts of the frame. An initial crude coordinate system is used to do this placing. Using one of the centroiding algorithms available, fitting a paraboloid to the object and noting the location of the maximum of the paraboloid, pixel coordinates of the object are measured in each of the frames. These coordinates, plus the encoder coordinates of the centers of each of the frames, form the raw material for a fit:

$$X = X_0 - A_{sx}S - A_{1x}L \tag{4-1}$$

$$Y = Y_0 - A_{sy}S \tag{4-2}$$

X_0, Y_0, A_{sx}, A_{sy}, and A_{1x} are adjustable for best fit. X_0 and Y_0 are estimates of the encoder coordinates of the fiducial object. A_{1x} is the scale along the direction of motor movement, while $(A_{sx}{}^2 + A_{sy}{}^2)^{1/2}$ is the scale along the reticon. $\tan^{-1}(A_{sx}/A_{sy})$ is a measure of the angle of skewness of the reticon with respect to the perpendicular of the direction of travel. This also includes the skewness due to the finite readout time of the reticon. The start the next integration essentially begins at the finish of the readout of the previous integration. As the pixels further along the reticon are read out later than the earlier pixels, a skewness is introduced. The position angle of the reticon is not adjustable. However, the pixels of the reticon are longer then they are wide. The slit which forms the illumination on the reticon is what defines the pixel extent along the direction of travel. The long pixels on the reticon can accommodate some rotation of the slit, thus simulating a rotation of the reticon. The slit angle is adjusted to zero the skewness in the coordinate system measured above. This can be done to about 5 minutes of arc.

This coordinate measurement is typically done with 10 to 20 measurements with the scanner running one direction, then an equal number of measurements with the scanner running in the other direction. The three scales A_{sx}, A_{sy}, and A_{1x} are averaged together. The difference between the two measurements of X_0 serves as an estimate of the total backlash in the system, and is used to set a software parameter to correct for this backlash.

All movements in the Y direction are handled by moving to the desired position plus 100 microns, then moving to the desired position. Thus, as the position is always approached from the same direction, there is no backlash. Also, because of this techique the Y motor reaches the desired position traveling at the same velocity. Since the Y motor will coast a velocity dependent distance after stopping this guarantees that a single coast-correction will be sufficient.

The residuals from the pixel coordinate to encoder coordinate measurement give an estimate of the local positional accuracy of the scanner. These residuals are typically on the order of two microns.

The set of pixel scales plus the encoder coordinates of the frame center are stored on the label of each frame, thus making it possible to compute the encoder coordinates for any pixel in that frame.

4.5.3.2 Encoder To Celestial Coordinate Transformation.

Transformation of encoder coordinates to Right Ascension and Declination is handled in two stages. In the first stage, the encoder coordinates (X,Y) are converted to a coordinate system (A,D) on a sphere whose equator and prime meridian are at a fiducial Right Ascension and Declination and whose axes are parallel to the local RA and Dec. axes. Two of these coordinate systems are used at different times: 1) the coordinate system for the plate, with fiducial RA and Dec. at the plate center, 2) the coordinate system for a single frame with fiducial RA and Dec. at the frame center. In the second stage the (A,D) coordinate system is rotated to the proper RA and Dec.

The plate coordinate system is used to transform the central encoder coordinates of a picture frame into RA and Dec. coordinates that can be used as the fiducial RA and Dec. for that frame's coordinate system.

To calibrate this plate coordinate system, position standard stars were gathered. These were stars out of the SAO catalog of $m > 8.0$, typically about 20 per plate.

To measure the standards, a small frame is scanned around the standard. The pixel coordinates of the standard are measured with the usual centroiding algorithm (with, however a rather large aperture of about 40 pixels, due to the large size of the images and their central saturation). The pixel coordinates are then converted to encoder coordinates with the previously determined transformation. The RA and Dec. of the standards are corrected for proper motion, and, using the plate center RA and Dec., rotated to the plate's (A,D) coordinate system. These quantities are then fit to:

$$A = T(X-X_0) + S(Y-Y_0) \tag{4-3}$$

$$D = -S(X-X_0) + T(Y-Y_0) \tag{4-4}$$

X_0, Y_0, T, and S are adjustable by the fit. A single scale size and a rotation are allowed in this fit. (X_0,Y_0) are determined to be the the encoder coordinates of the origin. This linear fit was found to be adequate for the Schmidt plates. Residuals from this fit were on the order of one or two arcseconds.

Whenever a frame is scanned, all of the information needed to convert location of any pixel to Right Ascension and Declination is stored on the label of that frame. As mentioned previously, the RA and Dec. of the frame center are calculated. Then a matrix of arcsecond/pixel scales, including all the scaling, stretching, rotation, and skewness of the underlying coordinate systems, is calculated by multiplying together the scaling matrices of the pixel --> encoder coordinate transformation and the encoder --> celestial coordinate transformation.

4.5.4 HD Calibration

Each of the Schmidt plates had on it a log-linear calibration wedge (intensity equal to the logarithm of the distance along the wedge). To reduce this wedge, its axis was carefully aligned with the scan direction of the scanner. The clear plate along the sides of the wedge (perpendicular to the scan direction) was blocked off to prevent light from scattering along the reticon onto the wedge. A frame was scanned along the wedge and displayed on the television screen. There the useful portion of the wedge (a trapezoidal region) could be selected with the cursor. The selected area was then averaged down to about 512 measurements along the wedge. These measurements were converted to densities (i.e. their logarithm was taken) and were fitted to a polynomial of about order 6 in the logarithm of the intensity. Chebyschev polynomials were used because of their greater stability. The quality of the fits was judged to be very adequate.

This function was chosen over other, more conventional (and complicated) functions because it did not require knowledge of the level of "clear plate". The light intensity of the scanner had been adjusted so that, using the area of lightest sky level on the plate, none of the pixels on the reticon, within the dispersion of the sky, were saturated. With this adjustment, clear plate was usually strongly saturated. For a typical sky density on a deep plate of about 1, it was judged a waste of the dynamic range of the scanner to be able to reach clear plate. When the transfer function is as densely sampled as it is when one uses a continuous calibration wedge, the form of the function fitted to it is not critical. It is only when the transfer function is coarsely sampled, as it often is with calibration spots or a step wedge, that the way the fitting function interpolates between measurements makes a difference.

When in use, the HD curve was stored as the array of numbers needed to convert the transmission measurements from the scanner to intensities. As this is a function which has a singularity at zero transmission (infinite intensity), the table was divided into 2 parts. The lower part was a straight one-for-one lookup table that reproduced the region around the singularity of the transfer function as closely as possible. The upper part of the table sampled the function more coarsely in the high-tranmission region where the curve was not changing very fast. This part of the table was used with linear interpolation. Even though the routine to do the lookup and interpolation was written as carefully as possible in machine code, the lion's share of the initial processing was spent in this conversion of the data from transmission to intensity.

4.5.5 Sky Tracking And Subtraction

As a frame is scanned, histograms of pixel intensity are created for subdivisions of the frame. Typically, these regions were 128 by 128 pixels across (about 1.5′ by 1.5′ on the scale in which the Schmidt plates were scanned) on a frame size of 512 by 1536 pixels (6.2′ by 18.5′). From each histogram, an estimate of the sky background is calculated. The array of sky backgrounds is passed as a file to the subsequent routine which finds and identifies the objects in the frame. This routine does a bilinear interpolation on the array of sky estimates to compute a sky estimate at the location of a particular pixel, and then subtracts this estimate. Doing at least a bilinear interpolation was necessary to avoid introducing sharp edges in the frame which could be later picked up as spurious galaxy correlations. However, no attempt was made to interpolate across the backgrounds of adjacent frames. Perhaps in a latter version that will be remedied. Doing anything higher order than a bilinear interpolation was judged too expensive in computer time. Any operation, like this one, which operates on every pixel in the frame, makes a large difference in the speed of the program and has to be carefully optimized.

To extract a sky background estimate from a histogram, the histogram was convolved with the derivative of a Gaussian. The sky estimate was taken to be the value of the zero crossing point of this convolution. If the histogram were really a Gaussian of that width, with the addition of Gaussian noise, then this would be the optimal way of estimating the sky from the histogram (see Appendix B of Chapter 2).

The assumption is made here that, although the sky background may vary across the plate, the dispersion of the pixels around the sky remains unchanged. This is only approximately true. Fortunately, this background finding algorithm is not very sensitive to the width of the Gaussian used. Also fortunately, the algorithm is not particularly sensitive to deviations of the histogram from Gaussian, such as those caused by the presence of stars or galaxies in the regions examined.

Because of computer memory limitations, the histogram could not look at the whole range of pixel intensities, but at only a narrow range around where the sky background was likely to be. Also, the convolution used by this algorithm was not computed completely, but only around where the background was expected to be found. A search algorithm took advantage of the fact that when the convolution was evaluated at some point, the sign of the result indicates the direction to search. Once the zero was bracketed its location was estimated by linear interpolation between the bracketing points.

The background finding algorithm needed two things to be initialized: an initial estimate of the value of the sky background, and a stored function --- the derivative of a Gaussian. Computation of that stored function required knowing the desired width (standard deviation) of the Gaussian. These things were measured by fitting a Gaussian to a histogram obtained from a test frame.

The initial sky estimate is used to set the location of an intensity window within which the histograms are measured when a frame is scanned. This width is typically 8 times the standard deviation of the sky (the width of the Gaussian used in the algorithm). The histograms measured in a row across the frame are examined together. If the sky background computed from any of these histograms is greater than the initial sky estimate by more than twice the standard deviation of the sky, then that sky estimate is replaced by the initial sky estimate. This may happen when there is a bright object in that region. Next, all of the (modified) estimates in the row are averaged and the average used as the initial sky estimate for the next row. The average from the last row on the frame is used as the initial sky estimate for the next frame, where it determines the window of intensities examined by the histograms of that frame. This number was printed out on the log for the scanner, where it served as a check on the stability of the algorithm and the scanner. For example, it could be used to determine which frames to throw out when the lamp on the scanner burned out. There was some worry that the algorithm could "get lost" when the sky changed so rapidly that the centroid of the intensities was no longer within the histogram. This effect was never actually observed to happen.

4.6 Subsequent Processing

At this stage, the catalog consisted of lists of objects from different places on each plate. There was much overlap in objects between these lists. The lists of objects for a particular plate were then merged into a master list of objects for that plate.

When merging lists, it is first necessary to sort the objects in each list into some order, in this case, increasing Right Ascension. The merging program operated on two of these sorted lists at a time, outputting a single sorted list. A moving window is maintained for objects in each of the two input files. Objects are read into a window from each file until they are more than some RA increment from the object of lowest RA in any window. Previously decided offsets are added to the coordinates and magnitudes of the second file. As an object enters a window it is compared to all of the objects in the other window. If the two objects are closer than a previously decided distance in RA and Dec. they are considered a match. When a match is found, both objects are removed from their windows, their parameters are averaged together, and result placed in a third window. Because of this, once an object is matched it will not take part in any further matches. Next, the object with lowest Right Ascension from any of the three windows is written to the output file and removed from its window. Using a small search box, say 3 by 3 arcseconds, the coordinate offsets between pieces are adjusted to maximize the number of common objects found between those pieces. The final merge between pieces of a plate is done with a 10″ by 10″ search box. The result is a separate catalog of all of the detected objects for each plate. A total of about half a million objects was counted. Of these, about a third were classified as galaxies.

5.1 Checks On Performance

5.1.1 ϕ Vs. Magnitude -- Completeness And Classifier Limit.

Two quantities of interest can be obtained immediately from examination of a plot of the classifier ϕ versus magnitude (refer to figure 6 of Chapter 2 of this thesis). For Schmidt plates, the distribution of objects tends to have a sharp magnitude cutoff, which tends toward brighter magnitudes as ϕ becomes larger, that is, as the objects become more diffuse. The lowest magnitude reached by this envelope for the most diffuse objects is about 20th. This is an estimate of the limiting magnitude to which the catalog is complete.

There is also information in the lower envelope of the ϕ vs. magnitude graph. The shape of the upper envelope of a distribution of stars should be roughly symmetrical to the shape of the lower envelope of the distribution. However the upper envelope of the stars is obscured by the galaxy distribution. Extending a mirror image of the lower envelope up to where it crosses the star/galaxy decision boundary gives an estimate of the magnitude at which the classifier breaks down to the misclassification of stars as galaxies. That magnitude was typically around 19.5. One would expect an earlier breakdown in reference to the misclassification of galaxies as stars. The distribution of galaxy shapes is much broader than the distribution of star shapes, making it much easier to mistake a galaxy (especially a compact one) as a star.

5.1.2 Image Saturation

Other quantities computed by the system enable the calculation of other essential parameters. An examination of a plot of peak surface brightness versus the faint object classifier, ϕ, gives the surface brightness at which saturation occurs. The ϕ classifier is rather sensitive to image saturation (which is why it is only useful for unsaturated, i.e. faint, images). On the ϕ versus central surface brightness graph, the stars will fall on a line of constant ϕ which begins to curve upward when the central pixels start to saturate. The beginning of surface brightness saturation is simply the surface brightness value at which the turn-up occurs.

This value was typically twice the sky intensity, or about a density of 2. The main reason for this rather poor performance was a scattered light problem in the plate scanner. In a multiplexing scanner like this one, with many detectors active at the same time, light that scatters once in the optics between the plate and the detectors may register on the wrong detector. However, in a scanner with a single detector, the most likely fate of light that scatters once is to miss the detector. Light has to scatter a second time to rejoin the beam and register on the detector. This difference between 1st and 2nd order scattering is a big problem in principle, made even worse by the fact that it takes very little scattered light to badly affect the dark central portions of the images of bright objects. Scattered light completely dominates digitization and amplifier noise, which begin to become noticeable around density 3.5.

Likewise, by examining the ø versus magnitude diagram for the turn-up of the distribution of stars, or by looking up the magnitude of objects whose central pixels are just becoming saturated, one can determine a magnitude at which the stars are becoming saturated. For stars brighter than this the magnitudes become unreliable. This occurs at 18th magnitude. For galaxies, whose lower density gradients are less susceptible to the scattered light problem, the magnitudes remain reliable for objects as bright as 17th magnitude.

5.1.3 Galaxy Counts.

A(m), the number of objects per square degree per unit magnitude, is plotted as a function of magnitude in Figures 1 through 5 for stars and galaxies. In Figure 1, A(m) is plotted for the stars and galaxies of all four plates. There is a fair agreement in the behavior of the plates. One plate (field PS23632) was scanned with a deeper threshold, as shown by the deeper magnitude cutoff.

Two plates from the Kitt peak 4 m. telescope were kindly loaned to me by J. Hoessel. The areas covered by these plates were within the region of sky covered by the Schmidt plates.

Table 5-1 Plates from KPNO 4 m. Telescope

plate	field	α_{1950}	δ_{1950}	Date	Exp.	Emulsion
MP3299	113	9 48 15	46 53 34	Mar 18,1980	75 m	IIIaF
MP2752	111	9 40 01	48 30 08	Feb 10,1980	90 m	IIIaF

These plates were scanned with roughly the same parameters that were used to the scan the Schmidt plates. The A(m) curves are plotted for these plates in Figs. 2 and 3, against A(m) curves for subfields of the Schmidt plates that cover the same area of the sky.

The average of all the Schmidt plates and all of the Kitt Peak 4 meter telescope plates is plotted in Figure 4. Finally, the total A(m) for galaxies from the Schmidt plates is plotted in Figure 5 together with some of the results of other groups. It is clear from examination that there are some systematic problems. These problems, however, are unlikely to affect the correlation function.

The star counts seem to have a plateau between $16^{m}_{\cdot}5$ and $18^{m}_{\cdot}5$. Also, at about $18^{m}_{\cdot}5$, the slope of the galaxy counts decreases. As the steep rise in star counts above $18^{m}_{\cdot}5$ is not recorded by other workers, the suspicion grew that these two things are related, both caused by an increase in the misclassification of galaxies as stars with increasing magnitude.

Pursuing this hypothesis, a line was fitted to the star logA(m) vs. magnitude function between $16^{\mathrm{m}}.5$ and $18^{\mathrm{m}}.3$. This line (slope 0.08) was then extrapolated to fainter magnitudes. Assuming that the excess stars above this line were really galaxies, the line was subtracted from the star counts and added to the galaxy counts. The result is the dotted line in Figure 5. An estimate for the rate at which stars are misclassified as galaxies follows from this exercise. As expected, the behavior here is poorer than the behavior for the misclassification of stars as galaxies. The adjusted galaxy counts agree well with the results of Kron (1980).

Table 5-2 Estimate of Stars Misclassified as Galaxies

Mag.	Stars/ deg^2/mag	Extra- polated	diff- erence	galaxies in star counts	galaxies/ deg^2/mag	corr- ected galaxies
16.50	262.4				15.8	
16.75	294.8				21.8	
17.00	307.1				27.5	
17.25	318.4				42.0	
17.50	328.8				56.3	
17.75	332.6				81.3	
18.00	355.7				112.8	
18.25	375.6	373.1	3.0	0.6%	151.7	154.7
18.50	425.6	390.1	35.6	8.4%	195.2	230.8
18.75	489.3	407.7	81.6	16.7%	251.9	333.5
19.00	566.8	426.2	140.5	24.8%	318.4	455.9
19.25	668.0	445.5	222.4	33.3%	407.1	629.4
19.50	818.0	465.7	352.3	43.1%	523.0	875.3
19.75	975.0	486.8	488.2	50.1%	680.2	1168.4
20.00	1180.9	508.9	672.1	56.9%	862.4	1534.5
20.25	1405.0	531.9	873.1	62.1%	947.1	1820.2
20.50	1636.1	556.1	1080.5	66.1%	700.1	1780.6

As the slope of the adjusted galaxy counts is rather steeper than the slope of the main galaxy count curve at the place that they meet, the extrapolation of the star count plateau probably underestimates the true star count. Thus, the behaviour of the classifier is probably better than the above numbers indicate. In retrospect, it is apparent that the star/galaxy decision boundary was set rather conservatively, favoring the classification of objects as stars. As the main interest was the computation of correlation functions, where an uncontaminated galaxy sample was desired, this decision was the correct one. At magnitudes faint enough so that many stars are misclassified as galaxies, the number of real galaxies has become so large that the contamination is not serious.

The Euclidian slope 0.6 line for bright galaxies from Kirshner, Oemler, and Shecter (1979, hereafter KOS), lies above the bright end of the objects on the A(m) graph. To check whether there is a local galaxy count deficiency in the region of this survey, A(m) estimates were computed from the <u>Catalog of Galaxies and Clusters of Galaxies</u> (Zwicky, et.al. 1961-1968, hereafter the Zwicky Catalog) for galaxies in the survey region.

To convert the magnitudes used in the Zwicky catalog to the magnitude system used here, the transformation of Felten (1977):

$$\underline{B}(0) = .886m_{zw} + 1.315 \qquad (5\text{-}1)$$

was used to transform to the $\underline{B}$ magnitude system. The $\underline{B}$ magnitudes were then transformed to $\underline{r}$ magnitudes with an average galaxy $\underline{B}$-$\underline{r}$ color (see below) of -1.009. The result is:

$$\underline{r} = .886m_{zw} + .306 \quad (5\text{-}2)$$

The resultant A(m) points, when converted to the $\underline{r}$ magnitude system, lie on the KOS line, indicating that there is no deficiency of bright galaxies in the survey region. With these galaxies in hand, the discrepancy could be investigated. All of the Zwicky galaxies brighter than $m_{zw}= 15.0$ (corresponding to $m_r= 13.6$) were examined in the survey. 13 out of total of 34 galaxies had been classified as stars. It is clear that a better bright object classifier must be found. This accounts for half of the discrepancy. The other half is due to a photographic saturation magnitude error of about 0.4, not surprising for photographic photometry of bright galaxies.

An attempt was made to use this information to correct the bright end of the magnitude scale. As mentioned earlier, the behavior of the classifier ϕ for stars as a function of magnitude indicated that there was no magnitude error due to saturation fainter than $18\overset{m}{.}0$. Using the simplest assumption, that the magnitude error grows linearly with magnitude, a line constrained to go through zero at $18\overset{m}{.}$ was fitted to the galaxy magnitude error as a function of magnitude between $m_{zw}= 14.0$ and $m_{zw}= 15.0$. The corrected magnitude scale is then:

$$m_{r'} = m_r - 0.0854(18-m_r) \qquad m_r < 18 \quad (5\text{-}3)$$

$$= m_r \qquad m_r > 18 \quad (5\text{-}4)$$

The value of this correction at the brightest magnitude used for correlation function analysis ($15^{m}.71$) was $0^{m}.19$.

As another check on the magnitude scale, isophotal magnitudes down to surface brightness $24^{m}.7$ were synthesized form the SIT profiles of the calibration galaxies. The results are in Table 5-3.

Table 5-3 Checks of Galaxy Photometry

	field	M_{zw}	$\underline{r}$(SIT)	$m_{scanner}$	Δ
ZwGal 09408+4254	112	15.0	13.63	13.69	0.06
ZwGal 09497+4305	112	14.5	13.03	13.51	0.48
NCG 3111	113	14.0	12.93	13.05	0.12
ZwGal 10002+4837	113	14.9	13.82	14.06	0.24
ZwGal 10010+4517	114	14.7	14.23	14.30	0.07
NGC 2857	111	14.3	12.67	12.74	-0.07
ZwGal 09372+4834	111	13.5	13.40	13.51	-0.11

For the last two galaxies on the list, the SIT profile did not reach $24^{m}.7$. Therefore $\underline{r}$(SIT) should have been brighter and the residual more positive for those two galaxies.

Subsections of the catalogs in a region of overlap between those catalogs were compared object by object as another consistency check. All of the objects detected in the region from $9^{h}44^{m}$ to $9^{h}45^{m}$ and from 44°00′ to 44°10′ were extracted from fields 112 (PS23591) and 114 (PS23634).

This was a sample of 186 objects. The objects were then cross-correlated between subcatalogs. The Right Ascension differed on the average by 2.4 seconds of arc, with a dispersion of 2.8 seconds of arc. Declination fared worse. The plates differed by an average of 3.5 seconds of arc with a dispersion of 11.4 seconds of arc. The behavior of the magnitude is given in table 5-4:

Table 5-4 Comparison of field in PS23591 and PS23634

Magnitude	Objects	difference	dispersion
15.50 - 16.00	4	-0.154	0.032
16.00 - 16.50	4	-0.160	0.021
16.50 - 17.00	5	-0.108	0.033
17.00 - 17.50	7	-0.091	0.020
17.50 - 18.00	9	-0.112	0.022
18.00 - 18.50	9	-0.090	0.072
18.50 - 19.00	6	-0.126	0.106
19.00 - 19.25	8	-0.163	0.102
19.25 - 19.50	16	-0.135	0.133
19.50 - 19.75	10	-0.028	0.187
19.75 - 20.00	12	-0.118	0.196
20.00 - 20.25	21	-0.144	0.245
20.25 - 20.50	22	+0.021	0.315
20.50 - 20.75	16	-0.054	0.557

All objects in PS23591 were found in PS23634 except for 1 object in the 20 to 20.5 magnitude bin, three in the $20\overset{m}{.}5$ to $20\overset{m}{.}75$ bin and 9 in fainter bins.

5.2 Correlation Function Analysis

5.2.1 Correlation Function Computation

The catalogs were used to compute the correlation function at small angular scales. To do this, the catalogs were divided into boxes about 30 arc minutes across. The distances between all pairs of objects was computed and binned into logarithmic bins, whose scale was fifteen bins per factor of 60 in angular distance, or a factor of 1.314 per bin. Using the total number of objects in each box and an analytic function for the number of pairs expected to be lost across the edges of the box (see appendix A), the number of expected pairs for an uncorrelated catalog was computed for each bin. The estimate for the correlation function for each bin is then:

$$\langle w \rangle_{bin} = (P_m/P_e) - 1 \qquad (5\text{-}5)$$

where P_m is the number of pairs measured and P_e is the number of pairs expected for an uncorrelated distribution.

The catalog in each box was inspected for imperfections, such as very bright stars producing a cluster of spurious objects (see below), or missing areas where a frame was lost due to a computer crash or disk file overflow. The boxes judged to be free from these imperfections were noted and their correlation functions averaged together. The averaging was done simply by summing the P_m and P_e values together.

Using the aperture magnitude scale (the magnitude calculated within an 11 arcsecond diameter aperture), the objects between $16^{m}.0$ and $20^{m}.5$ classified as galaxies were subdivided into subcatalogs half a magnitude wide. Then the small angle correlation functions were computed for each of these subcatalogs.

It became apparent that the "combined" magnitude scale, (the brighter of the aperture magnitude and the isophotal magnitude computed with a surface brightness isophote of $m_r = 24.7$) should have been used instead. The brighter of the bins already had a substantial aperture correction. The magnitudes were therefore corrected to the combined magnitude scale. Isophotal magnitudes for galaxies were plotted against aperture magnitudes (Figure 6) and a curve hand drawn through the distribution of galaxies on that graph. Using this curve, the magnitude of the bin centers was adjusted to the "combined magnitude" scale. The magnitudes were further adjusted for image saturation. The results are shown in Table 5-5.

Table 5-5 Correlation Function Magnitude Scale

Aperture Magnitude	Bin Center	Corrected to Combined Magnitude	Corrected for Saturation	Equivalent B Magnitude
16.0<m<16.5	16.25	15.71	15.52	16.63
16.5<m<17.0	16.75	16.40	16.26	17.27
17.0<m<17.5	17.25	17.08	17.00	18.01
17.5<m<18.0	17.75	17.60	17.57	18.58
18.0<m<18.5	18.25	18.20	18.20	19.21
18.5<m<19.0	18.75	18.74	18.74	19.75
19.0<m<19.5	19.25	19.25	19.25	20.26
19.5<m<20.0	19.75	19.75	19.75	20.76
20.0<m<20.5	20.25	20.25	20.25	21.26

A power law was fit to each correlation function between separations of 17.7" to 4'32", using a maximum likelihood routine which assumes that the number of pairs of objects in each angular bin follows Poisson statistics. A sample of these is shown in Figs 7-15. These results are shown in the left column of Table 5-6.

Table 5-6 Correlation Function Results

	All valid fields		Throwing out fields with SAO stars of $m_V<9.0$	
m_r	Log w(1')	Slope	Log w(1')	Slope
15.52	-0.04±0.60	0.32±1.06	no convergence	
16.26	0.18±0.15	-1.25±0.37	no convergence	
17.00	-0.11±0.12	-0.93±0.26	-0.25±0.20	-0.85±0.56
17.57	-0.11±0.06	-1.18±0.14	-0.03±0.07	-1.13±0.18
18.20	-0.55±0.07	-0.93±0.19	-0.56±0.10	-0.84±0.27
18.74	-0.64±0.05	-1.36±0.14	-0.64±0.07	-1.38±0.19
19.25	-0.83±0.05	-1.05±0.13	-0.81±0.06	-1.28±0.24
19.75	-0.57±0.02	-1.02±0.04	-1.05±0.07	-1.51±0.22
20.25	-0.86±0.02	-1.10±0.05	-1.02±0.04	-1.27±0.12

The amplitude of the fit at one arc minute is one of the parameters computed from the fit. As 1 minute of arc is near the midpoint of the range in logθ, its value should be nearly independent of the slope of the fit.

A regular least-squares fit (which assumes Gaussian statistics) can badly overestimate the amplitude of the correlation function. When there are fewer pairs of objects, statistics become less Gaussian and the overestimate of the correlation function increases. As the lower magnitude subcatalogs have fewer objects, the effect of this is to bias the slope of Log Correlation function versus magnitude graph to larger negative values.

The assumption of uncorrelated Poisson statistics for the number of pairs of objects in an angular bin is only an approximation, as the number of pairs in one bin is not independent of the number of pairs in other bins. Peebles (1973) showed that if the spherical harmonic transform of the correlation function u_1 is roughly constant over the range of angular separations included in the fit, then the standard deviation of the number of pairs in a bin is proportional to the Poisson value. The actual probabilities of these quantities has to be something resembling Poisson, having discrete values and being zero for a negative number of pairs of objects. Therefore it seems likely that those statistics resemble scaled Poisson statistics. The same fitting algorithm valid for Poisson statistics is also valid if the statistics are proportional to Poisson.

The sum of the pairs counted over all of the angular bins is equal to the total number of pairs n(n-1)/2. Likewise the sum over all bins of the number of expected pairs is n(n-1)/2. Since our estimate of the correlation function involves a comparison between two quantities that sum to the same value, it is clear that the estimate of the correlation function has to be negative somewhere. This is the explanation for why the correlation function turns down at the larger angles which are computed. Another way of viewing this is that the low spatial frequencies needed for estimates of larger angular scales have been filtered out by the process of binning the objects into boxes. This effect occurs even though a correction has been made for pairs lost over the edge of the box. The outermost point included in the power-law fits, at 4.5′ is at a much smaller angle than the box size of 30′, and the edge effects are negligible here.

Note the −0.26 jump in logW(1′) between $19^{m}.25$ and $19^{m}.75$. This jump is almost a factor of two and would be extremely interesting if real. However, the effect is almost certainly an artifact. In those automatic object detection systems that detect objects by looking for surface brightness enhancements above some threshold, it is always a problem that false objects can be detected on the outer edges of bright stars. In regions where the background hovers just below the threshold and the gradient of the background is low, any small wiggle in the background can be detected as an object. Also, any real object will have its brightness overestimated because of the ambient light from the nearby large object. The brighter a star, the lower the gradient of the background at the threshold. The result of this effect is a tightly correlated cluster of spurious galaxies.

I will now attempt to quantify the magnitude of the spurious component needed to produce this effect. To do this it is necessary to compute an addition rule for the correlation functions of two populations.

Assume two populations, one with correlation function $w_1(\theta)$, the other with correlation function $w_2(\theta)$. Assume that in a certain region there are n objects, n_1 objects of population 1 and n_2 objects of population 2. If the correlation function is estimated as in the above equation (5-5), then the correlation function estimate is (if n is large so that n-1 can be approximated by n):

$$\langle w(\theta) \rangle = P_m / P_e - 1 = P_m \Big/ \left[\tfrac{1}{2}A(\theta)n^2\right] \tag{5.6}$$

A(θ) is the analytic function computed in Appendix A, which corrects for the number of pairs lost off the edges of the field.

Broken down by population, with $P_{m_{11}}$ the pairs counted among population 1, $P_{m_{12}}$ the pairs counted between population 1 and 2 and $P_{m_{22}}$ the pairs counted among population 2 the relationships exist:

$$P_m = P_{m_{11}} + P_{m_{12}} + P_{m_{22}} \tag{5-7}$$

and

$$P_e = \tfrac{1}{2}A(\theta)n^2 = \tfrac{1}{2}A(\theta)(n_1+n_2)^2 \tag{5-8}$$

$$= \tfrac{1}{2}A(\theta)n_1^2 + A(\theta)n_1n_2 + \tfrac{1}{2}A(\theta)n_2^2 \tag{5-9}$$

$$= P_{e_{11}} + P_{e_{12}} + P_{e_{22}} \tag{5-10}$$

Note that in the middle term there is no factor $\frac{1}{2}$ correction for counting pairs twice, as there would be when counting pairs among a single collection of objects. The estimate for the correlation function is then:

$$\langle w \rangle = \frac{P_{m_{11}} + P_{m_{12}} + P_{m_{22}}}{[A(\theta)/2](n_1 + n_2)^2} - 1 \tag{5-11}$$

$$\langle w \rangle = \left[\langle w_1 \rangle n_1^2 + \langle w_{12} \rangle n_1 n_2 + \langle w_2 \rangle n_2^2 \right] \Big/ n^2 \tag{5-12}$$

If $\varepsilon = n_2/n_1$ is the ratio of objects in the two populations:

$$\langle w \rangle = \left[\langle w_1 \rangle + \langle w_{12} \rangle \varepsilon + \langle w_2 \rangle \varepsilon^2 \right] \Big/ (1+\varepsilon)^2 \tag{5-13}$$

The correlation functions add linearly by an amount proportional to the square of the densities.

The logarithm of the correlation function at $19^{m}.75$ goes from −1.05 to −0.57 when the fields suspected of contamination are added. This is a factor of 3 increase in the correlation function. If there were no cross correlation between the contaminating population (none is expected) and the true population of galaxies, and if the density of the contaminating population is, for example, one tenth of the density of the true population then the correlation function of the contaminating population at one minute of arc would have to be 23.7. If the density of the contaminating population is one hundredth the density of the true population then the correlation function of the contaminating population would have to be 1854. Finally, if the density of the contaminating population is equal to the density of the true population then the correlation function of the contamination population would only have to be 0.99. These values should bracket the true value. It does not take very much of a tightly correlated contaminating population to completely change the correlation function estimate of the true population.

There were tests and procedures in the pattern recognition software intended to handle this problem (see Sebok (1980), chapter 3 of this thesis), which, it is now apparent, were not enough. The main one of these procedures was that a region was grown from an object to include all of the pixels above the threshold. This is similar to the procedure of Jarvis and Tyson (1979). Thus, all of the pixels affected by a bright object should be included as part of that object. Carefully handling the holes and bifurcations which occur in the ragged envelope around a bright star should guarantee that no affected pixels are lost. However, if the sky gradient is low enough, it is still possible to split pieces off the envelope.

This problem is difficult to eliminate completely. To avoid it, the region around a bright star could be modeled very precisely and subtracted from the sky. That is rather hard however, and costly in computer time and computer memory. Also, a sky tracker which follows the sky level too quickly can cause light to be lost from the brighter galaxies (which may be all right if these objects are not of interest). More practically, a way of calculating the correlation function should be used that allows one to eliminate the affected areas more cleanly, without eliminating so much of the good areas.

Sufficiently many objects have been eliminated from the catalog that the statistics for bright galaxies are adversely affected. Between $18^{m}.20$ and $18^{m}.74$ the fitted values of slope and LogW(1') are essentially identical between the original (unculled) correlation function computation and the correlation function computation (culled) with the $m_V<9.0$ fields discarded. Therefore it appears that the contamination problem does not affect the catalog below $19^{m}.0$. The adopted results are then the unculled fits when $m<19.0$ and the culled fits when $m>19.0$. These are listed in table 5-7, and plotted against a range of models in Figures 17-19.

Table 5-7 Adopted Correlation Function Results

m_r	Log w(1')	Slope	objects
15.52	-0.04 ± 0.60	0.32 ± 1.06	83
16.26	0.18 ± 0.15	-1.25 ± 0.37	368
17.00	-0.11 ± 0.12	-0.93 ± 0.26	962
17.57	-0.11 ± 0.06	-1.18 ± 0.14	2057
18.20	-0.55 ± 0.07	-0.93 ± 0.19	3905
18.74	-0.64 ± 0.05	-1.36 ± 0.14	6506
19.25	-0.81 ± 0.06	-1.28 ± 0.24	1081
19.75	-1.05 ± 0.07	-1.51 ± 0.22	1974
20.25	-1.02 ± 0.04	-1.27 ± 0.12	3233

The main thing to note about this table is that the slopes tend to be higher than the canonical -0.8 (Peebles (1975)). The outermost point included in the fit, at 4.5 minutes of arc, is quite a bit smaller than where edge corrections are expected to be important.

If all measurements are included, a weighted average yields slope = -0.81 ± 0.19. This value is strongly affected by a single point. An inspection of Figure 7 reveals that the anomalous slope of the catalog at $15^{m}_{.}2$ suffers is due to small numbers statistics. Indeed moving the endpoints of the fit changes the slope to more normal values. This was not done because it would have introduced an arbitrary bias into the fit. With this point excluded the weighted average of the slopes becomes -1.23 ± 0.18. When a line of slope -0.8 is plotted on those graphs with the better statistics (those graphs of more numerous, fainter objects) it is clear that a -0.8 line is inconsistent with the data. There is a slight tendency to larger negative slopes at the fainter magnitudes.

Gott and Turner (1979), using measurements of 47 pairs of galaxies from the Zwicky catalog, measure the correlation function of the Zwicky catalog between 13" and 1'11". Their results are consistent with an extrapolation of Peebles' results (Peebles 1975). However, the power law slope derived from just their own data, -1.08±0.28, is consistent with the results measured here. Furthermore, an examination of the plot of these data reveals that if the measurement at the smallest angular separation is discarded (a measurement derived from only 2 pairs of galaxies) the slope becomes considerably steeper.

Lake and Tremaine (1979), using Holmberg's (1969) survey of counts of companion galaxies around 115 spiral galaxies of known redshift, derive a spatial correlation function slope of -1.52 ± 0.19 equivalent to an angular correlation function slope of -0.52 ± 0.19. This slope is inconsistent with the slope measured here.

A steeper slope of the correlation function could be due either to the small angular scale (~1 minute of arc.) or due to the generally fainter magnitudes --> greater distance of the catalogs analysed here. Of the models that various workers have contructed of the evolution of the correlation function with time, some steepen and other flatten (see Haggerty and Janin (1974), Press and Schecter (1974), Miyoshi and Kihara (1975), Peebles and Groth (1976), Gott, Turner and Aarseth (1979), and Fall (1978)). On the other hand, the steepening of the slope of the angular correlation function can be an effect of a steeping of the spatial correlation function at short scales. The theory of infall of galaxies around a single initial galaxy (Gunn and Gott (1972), Gott (1975)) leads to a spatial correlation function of $\xi \sim r^{-2.25}$ equivalent to $w(\theta) \sim \theta^{-1.25}$, which is very close to the result measured here. At small spatial distances, this type of clustering could dominate (Gott, Turner and Aarseth (1979)).

5.3 Models And Data Interpretation

5.3.1 K Corrections

The spectral scans of Wells (1972), for a variety of galaxy types, were compared to the transmission function of the *r* filter. Wells averaged his energy distributions F_λ, for 4 morphologies of galaxies: E and S0, Sa and Sb, Sbc, and Sdm and Irr. These scans are given as energy distributions for constant wavelength interval. If S_λ is defined as the filter transmission as a function of wavelength, then the K correction is given by (Oke and Sandage (1968)):

$$K(z) = 2.5\log(1+z) + 2.5\log\left[\int_0^\infty F_\lambda S_\lambda d\lambda \Big/ \int_0^\infty F_\lambda(\lambda/(1+z)) S_\lambda d\lambda\right] \tag{5-14}$$

where λ is the rest wavelength. These results are plotted in figure 16 and listed in table 5-8 out to a redshift of one. It is convenient that for a red filter, even at redshift one, no extrapolation into the ultraviolet is necessary. For the magnitudes reached here, the needed redshifts are quite a bit less than one, peaking at a redshift of about 0.2 for the faintest magnitudes.

Table 5-8 K correction for Gunn Red

z	E + S0	Sa + Sb	Sbc	Sdm + Irr
0.000	0.000	0.000	0.000	0.000
0.020	0.019	0.016	0.005	-0.008
0.040	0.038	0.031	0.011	-0.016
0.060	0.057	0.047	0.016	-0.022
0.080	0.073	0.062	0.021	-0.024
0.100	0.089	0.078	0.026	-0.026
0.120	0.106	0.096	0.032	-0.029
0.140	0.125	0.115	0.041	-0.029
0.160	0.150	0.139	0.051	-0.028
0.180	0.178	0.166	0.060	-0.029
0.200	0.206	0.192	0.069	-0.043
0.300	0.337	0.311	0.110	-0.121
0.400	0.499	0.438	0.171	-0.097
0.500	0.801	0.660	0.292	-0.029
0.600	1.237	0.969	0.477	0.037
0.700	1.640	1.289	0.717	0.150
0.800	1.916	1.537	0.976	0.324
0.900	2.045	1.717	1.229	0.514
1.000	2.151	1.889	1.486	0.639

5.3.2

Galaxy Colors and Transformations

Colors were synthesized from the energy distribution scans and filter transmissions for the different galaxy types listed above. This served to aid the comparison of the results obtained here with results in other color systems. The color differences between *r*, *g*, and the standard *B* and *V* colors are given in table 5-9.

Table 5-9 Galaxy Colors

	g-*r*	*B*-*V*	*V*-*r*	*B*-*r*	*g*-*V*	*B*-*g*
E and S0	0.443	0.934	0.204	1.138	0.179	0.695
Sa and Sb	0.404	0.824	0.190	1.017	0.154	0.610
Sbc	0.179	0.587	0.082	0.669	0.037	0.490
Sdm and Irr	-0.067	0.368	-0.027	0.578	-0.100	0.408
Effective Galaxy	0.365	0.827	0.169	1.009	0.134	0.632
BD +17 4708	0.000	0.430	-0.030	0.400	-0.030	0.400

BD +17 4708 is the primary standard of the Thuan-Gunn system (defined as $9^{m}.5$ in all colors) and was used here as the color standard. The effective galaxy color is an average color over the expected frequency of the different galaxy types selected at a particular magnitude. This is discussed further later.

5.3.3 Luminosity Functions

Luminosity functions were constructed for galaxies of the various morphologies. Columns from the tabulated differential luminosity functions in Christensen (1975) were summed to synthesize luminosity functions for the galaxy morphologies listed above. The columns labeled E, SA0-0/a, and SB0-0/a were added together into an "E" luminosity function, the columns labeled SAa-bc and SBa-bc added into an "Sa" luminosity function, the columns labeled SAc-m and SBc-m added into an "Sc" luminosity function, and the Irr column was used for an "Irr" luminosity function. The Sc luminosity function was identified with the Sbc energy distribution used earlier, as Wells' sample of galaxies had no Sc galaxies.

These luminosity functions were fit with a Shecter function (Shecter (1976)) written in the form (Felten (1977)):

$$\phi(M)dM = \frac{2}{5}\phi^*(\ln 10)\left[\mathrm{dex}\frac{2}{5}(M^*-M)\right]^{\alpha+1}\exp\left[-\mathrm{dex}\frac{2}{5}(M^*-M)\right] \qquad (5\text{-}15)$$

α was held constant at -1.2. M^* was converted from *B* magnitude to *r* magnitude, using the *B*-*r* color computed above for each galaxy type. The results are in table 5-10.

Table 5-10 Galaxy Luminosity Functions

	ϕ^* (objects Mpc.)	M^* (B)	M^* (r)	Magnitude Selected Frequency
E	.00153	-20.83	-21.96	45.2%
Sa	.00146	-20.67	-20.47	32.4%
Sc	.00526	-19.80	-20.47	20.7%
Irr	.005	-18.44	-19.41	4.7%

5.3.4 A(m) Model

A galaxy Number versus Magnitude (A(m)) model was computed from these luminosity function with the equation:

$$A(m) = \sum_i \int_0^\infty dz \; R_a{}^2 \frac{dR_m}{dz} (1+z)^3 \; \phi_i(m-K_i(z)-5\log R_L-25) \qquad (5\text{-}16)$$

Here *i* runs over the different galaxy types, K_i is the K correction for the galaxy types computed earlier,

$$R_m = \left[1+q_0+q_0 z+(q_0-1)(2q_0 z+1)^{1/2}\right] / \left[q_0{}^2(1+z)(H_0/c)\right] \qquad (5\text{-}17)$$

is the metric distance (in megaparsecs), $R_L = R_m/(1+z)$ is the luminosity distance, and $R_a = R_m(1+z)$ is the angular size distance.

A model was constructed from this equation for $H_0=50$ km/sec/Mpc., $q_0=0$, and no luminosity evolution, and is plotted on Figure 5, scaled to run through the KOS bright galaxy line. Note that it lies entirely above the data.

Because of the problems this A(m) graph has, I do not think it worthwhile to draw conclusions about it except that the model also lies above Kron's A(m) results. Apparently, some luminosity evolution is needed, in the sense that the average galaxy grows *brighter* with time.

The frequency of occurrence of the different galaxy morphologies (at zero redshift) was obtained by evaluating the integrands of equation (5-1)6 at zero redshift and dividing them by the sum of the integrands. Effective galaxy colors were computed by averaging galaxy colors together in these proportions.

5.3.5 Correlation Function Models

Models were also constructed for the behavior of the correlation function as a function of magnitude. Using the relativistic Limber's equation for a power law spatial correlation function ξ, the angular correlation function scales as (Fall(1979) with the notation used above):

$$S(m) = \frac{1}{A^2(m)}\int_0^\infty dz\ \frac{dRm}{dz}\ {R_a}^{5-\gamma}\ (1+z)^{3-\varepsilon}\ \sum_i \phi_i(m-K_i(z)-5\log R_L-25) \tag{5-18}$$

A(m) is the galaxy count model computed in equation (5-16). Again, it is assumed that there is no luminosity function (or K correction) evolution. ε is a correlation function evolution parameter. A simple model for the behavior of a power law correlation function with redshift is:

$$\xi(r,z) = (r/r_0)^{-\gamma}(1+z)^{-(3-\varepsilon)} \qquad (5\text{-}19)$$

An ε of 0 indicates BBGKY evolution for $\Omega = 1$ (Peebles (1975)), $\varepsilon = -3$ indicates no evolution in physical coordinates, $\varepsilon = \gamma-3$ indicates no evolution in comoving coordinates, and $\varepsilon = 3-\gamma$, as an extreme case, indicates clusters collapsing as fast as the universe is expanding.

5.3.6 Comparison Of Correlation Function Models With Data.

Several of these models have been plotted in Figures 17-19. The amplitude of these models was set by best fit to the data weighted by the error bars of the data.

The point at $13^{m}.60$, logW(1′)=1.267 is an extrapolation to one minute of arc of Peebles analysis of the Zwicky catalog (Peebles (1975)). The quoted magnitude limit of the analysis is at $m_{zw} = 15.0$. That magnitude, converted to the m_r system becomes $13^{m}.60$. The scaling law is not valid for comparisons between catalogs whose selection laws are of different types, such as between catalogs selected to a limiting magnitude and catalogs selected at a particular magnitude. However, for a power law spatial correlation function, it is still possible to find a relationship between the the amplitudes of the projected correlation functions measured from the two types of catalog. Substituting a power law spatial correlation function into the non-relativistic Limber's equation (equation 1-4), yields

$$w(\theta,m) = \left[5(6-\gamma) \big/ 9\ln 10\right] w(\theta, <m) \tag{5-20}$$

as the relationship between w(θ,m), the correlation function at angle θ for objects selected at magnitude m, and w(θ, <m), the correlation function at angle θ for objects selected down to magnitude m. The coefficient is of order unity. In fact, for $\gamma = 1.8$, which fits Peebles analysis, the coefficient is 1.02. The error in log W of 0.039 assigned to this point is inferred from the error quoted in Peebles (1975). One should be cautious about relying on this point, as extrapolations are risky. It would be good to eventually analyse a narrow magnitude range subset of the Zwicky catalog, or some other bright catalog, in order to do a proper comparison with these data.

It is clear that the dominant parameter in the models is γ, the negative of the slope of the spatial correlation function. As the slope of the bright end of the Log W vs. Magnitude relation is $-\gamma/5$ (derived from the scaling law), one would hope that this relation could serve as a check on that parameter. Three values of γ, 1.8, 2.1, and 2.4 are plotted in Figure 17, for models with the other parameters held at $q_0=0$, $H_0=50$, and $\varepsilon=0$. Chi^2 for 10 degrees of freedom is 47.4 for $\gamma=1.8$, 12.7 for $\gamma=2.1$, and 59.4 for $\gamma=2.4$. The data favors the $\gamma=2.1$ model. Of course, if γ varies with magnitude (as hinted earlier), then the simple models used here are inadequate.

Figure 18 explores the effect of the correlation function evolution parameter ε. The other parameters are held at $\gamma=1.8$, $H_0=50$, and $q_0=0$. Four values of ε are shown: $\varepsilon = -3$ (chi^2=33.7) no evolution in fixed coordinates, $\varepsilon = -1.2$ (chi^2=47.4) no evolution in comoving coordinates, $\varepsilon = 0$ (chi^2=66.0) BBGKY stable clustering evolution, and $\varepsilon = 1.2$ (chi^2=103.7) clusters collapsing as fast as the universe is expanding. If $\gamma=1.8$, the data seem to favor moderate to strong evolution of the correlation function. However, if γ is larger than 1.8, less correlation function evolution is needed to fit the data.

In figure 19 q_0 is set to 0 and 1, while the other parameters are held at $\gamma=1.8$, $\varepsilon=0$, and $H_0=50$. The effect of q_0 is so small compared to the effect of the other parameters that nothing can be said as to its preferred value.

Table 5-11 Best Fits for γ ($q_0=0$, $H_0=50$)

Model	γ	ε	chi^2	
1	2.02	0.98	13.2	clusters collapsing at rate Universe is expanding
2	2.08	0.00	12.4	"stable clustering"
3	2.13	-0.87	12.6	no evolution in comoving coordinates
4	2.24	-3	13.0	no evolution in physical coordinates

In table 5-11 the evolution parameter ε is varied while γ is adjusted for best fit. The models are essentially indistinguishable. The best fitting values of γ for these models are between 2.02 and 2.24, in agreement with the high value of gamma indicated by the slopes of the log W versus logθ graphs.

Evolution of the luminosity function and/or the K correction (i.e. color evolution) would add more parameters to the models, further entangling an already confused situation. There is no problem, however, finding models to fit the data if these things are ignored.

5.4 Conclusions

The slope of log correlation function versus log angle between 18″ and 4′32″, for galaxies $16<m_r<20.5$, is measured to be -1.23 ± 0.18, with a hint of higher negative slopes at fainter magnitudes. This is inconsistent at the 2σ level with the standard slope of -0.8, but may be possible if the slope steepens with increasing magnitude or the slope steepens with the smaller angular size (~ 1 arc. min.) used here. However at the fainter magnitudes, where $z \sim 0.2$ this size is about 200 Kpc., which is not a particularly small spatial size.

The power law slope for the projected correlation function as a function of angle is $1 - \gamma$, where $-\gamma$ is the power law slope of the spatial correlation function. From a scaling law, Equation 1-12, the power law slope of the angular correlation function as a function of magnitude is $-\gamma/5$ for a Euclidian, non-evolving universe. The correlation functions for the brighter galaxies should follow this relation. Models that include K corrections, evolution, or cosmology are still more sensitive to this parameter than to any other parameter. The data is most consistent with a γ between 2.02 and 2.24, with the higher gamma's associated with lower values of correlation function evolution. This is consistent with the above result.

If this γ is 1.8, then moderate to strong evolution of the spatial correlation function evolution is favored, evolution in the sense of decrease in the amplitude of the correlation function with increasing redshift. A higher value of γ would require less evolution of the correlation function. If γ is allowed to float toward the higher values it seems to prefer then the amount of evolution present cannot be distinguished.

No luminosity or color evolution is included in the models, but the models have no trouble fitting the data without it.

Appendix A

In this appendix is computed the number of pairs of uncorrelated objects expected in a bin which extends over some range in angular separations. Given a window, W(x,y), which contained N objects, how many pairs of objects P_e would be predicted if one object of each pair was selected from the window and the other object was selected from the intersection of the window W(u,v) and the binning function b(u-x,v-y).

$$P_e = \frac{N(N-1)}{2} \frac{\int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dy\int_{-\infty}^{\infty}du\int_{-\infty}^{\infty}dv\ W(x,y)\ W(u,v)\ b(u-x,v-y)}{\left[\int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dy\ W(x,y)\right]^2} \tag{A-1}$$

Let W(x,y) speciify a rectangular box of length a and width b, and let b(x,y) specify the area inside a circular region, i.e.:

$$\begin{aligned} W(x,y) &= 1 \quad 0 < x < a \quad 0 < y < b \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{A-2}$$

$$\begin{aligned} b(x,y) = b(r) = 1 \quad r = (x^2+y^2)^{1/2} &= 1 \text{ if } r < R \\ &= 0 \text{ otherwise} \end{aligned} \tag{A-3}$$

Let $b(u,v) = B[(u^2+v^2)^{1/2}]$ (i.e. some circularly symmetric function:

let $u - x = r\cos\phi \quad v - y = r\sin\phi$

$$P_e = \frac{N(N-1)}{2} \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}dxdy\int_{0}^{2\pi}d\phi\int_{0}^{\infty}rdrW(x,y)W(x+r\cos\phi,y+r\sin\phi)B(r)}{\left[\int_{-\infty}^{\infty}dx\int_{-\infty}^{\infty}dy\ W(x,y)\right]^2} \tag{A-4}$$

Remembering the definition of $\underline{w}$ and $\underline{B}$.

$$P_e = \frac{N(N-1)}{2a^2b^2}\int_0^{2\pi} d\phi \int_0^R r\,dr\; W(x+r\cos\phi, y+r\sin\phi) \tag{A-5}$$

$$\begin{aligned} P_e = \frac{N(N-1)}{2a^2b^2}\int_0^{2\pi} d\phi \int_0^R r\,dr \\ \times \quad \mathrm{Max}\Big[0, [\mathrm{Min}(a, a+r\cos\phi) - \mathrm{Max}(0, r\cos\phi)]\Big] \\ \times \quad \mathrm{Max}\Big[0, [\mathrm{Min}(b, a+r\sin\phi) - \mathrm{Max}(0, r\sin\phi)]\Big] \end{aligned} \tag{A-6}$$

$$P_e = \frac{N(N-1)}{2a^2b^2}\int_0^R r\,dr \int_0^{2\pi} d\phi\; \mathrm{Max}\Big[0, a-|\cos\phi|\Big]\mathrm{Max}\Big[0, b-|\sin\phi|\Big] \tag{A-7}$$

If $a > b > R > 0$ and, by symmetry, letting the ϕ integral run from 0 to $\pi/2$:

$$P_e = \frac{2N(N-1)}{a^2b^2}\int_0^R r\,dr \int_0^{\pi/2} (a-r\cos\phi)(b-r\sin\phi) \tag{A-8}$$

$$P_e = \frac{2N(N-1)}{a^2b^2}\int_0^R r\,dr \left[\frac{ab\pi}{2} - (a+b)r + \frac{r^2}{2}\right] \tag{A-9}$$

$$P_e = f(r) = \frac{N(N-1)}{2}\left[\frac{\pi R^2}{ab} - \frac{4R^3}{3a^2b} - \frac{4R^3}{3ab^2} + \frac{R^4}{2a^2b^2}\right] \tag{A-10}$$

For an annular bin between R_1 and R_2 with $R_1 < b < a$ and $R_2 < b < a$:

$$p_e\langle\text{annulus}\rangle = f(R_2) - f(R_1) \tag{A-11}$$

REFERENCES

Christensen,C.G. 1975,*A.J.*,**80**,282.

Dautcourt,G. 1977,*Astron.Nachr.*,**298**,253.

Davis,M.,Peebles,P.J.E. 1977,*Ap.J.Supp.*,**34**,425.

Fall,S.M. 1978,*M.N.R.A.S.*,**185**,165.

Fall,S.M. 1979,*Reviews of Modern Physics*,**51**,21.

Felten,J.E. 1977,*A.J.*,**82**,861.

Fry,J.N.,Peebles,P.J.E. 1980,*Ap.J.*,**236**,234.

Gott,J.R. 1975,*Ap.J.*,**201**,296.

Gott,J.R.,Turner,E.L. 1979,*Ap.J.*,**232**,L79.

Gott,J.R.,Turner,E.L.,Aarseth,S.J. 1979,*Ap.J.*,**234**,13.

Green,R.F. 1977,Ph.D Thesis "A Complete Sample of White Dwarfs, Hot Subdwarfs, and Quasars", (California Institute of Technology, Pasadena).

Groth,E.J.,Peebles,P.J.E. 1977,*Ap.J.*,**217**,385. VII.

Gunn,J.E.,Gott,J.R. 1972,*Ap.J.*,**176**,1.

Haggerty,M.J.,Janin,G. 1974,*A.A.*,**36**,415.

Herzog,A.D.,Illingworth,G. 1977,*Proc. Southwest Reg. Conf.* on "Direct Plate Automated Reduction Techniques", (Denton,Texas),**2**,147.

Holmberg,E. 1969,*Ark.für Ast.*,**5**,305.

Jarvis,J.F.,Tyson,J.A. 1979,*Proc. SPIE (Instrumentaion in Astronomy, III)*,**172**,422.

Jarvis,J.F.,Tyson,J.A. 1980,*Proc. SPIE (Applications of Digital Image Processing to Astronomy*,**264**,222.

Jarvis,J.F.,Tyson,J.A. 1981,*A.J.*,**86**,476.

Kent,S.M. 1979,*P.A.S.P.*,**91**,394.

Kibblewhite,E.J.,Bridgeland,M.T.,Hooley,T.,Horne,D. 1975

Image Processing Techniques in Astronomy,Proceedings (Reidel, Dordrecht),**54**,245.

Kirshner,R.P.,Oemler,A.,and Schecter,P.L. 1978,*A.J.*,**83**,1539.

Kirshner,R.P.,Oemler,A.,and Schecter,P.L.1979,*A.J.*,**84**,951.

Kron,R.G. 1980,*A.J.*,**43**,305.

Lake,G.,Tremaine,S. 1980,*Ap.J.*,**238**,L13.

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

Limber,D.N. 1954,*Ap.J*,**119**,655.

Matthews,T.A.,Sandage,A.R. 1963,*Ap.J.*,**138**,30.

Miyoshi,K.,Kihara,T. 1975,*P.A.S.Japan*,**27**,333.

Neyman,J. 1962,*Problems of Extragalactic Research*, Ed.G.C.McVittie,(New York:Macmillan),p.294.

Neyman,J.,Scott,E.L.,and Shane,C.D. 1954,*Ap.J.Supp.*,**1**,269.

Oke,J.B.,Sandage,A.R. 1968,*Ap.J.*,**154**,21.

Peebles,P.J.E. 1973,*Ap.J.*,**185**, 413 I.

Peebles,P.J.E.,Hauser,M.G.,1974 *Ap.J.Suppl.*,**253**,19 III.

Peebles,P.J.E.,Groth,E.J. 1975,*Ap.J.*,**196**,1 V.

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

Peebles,P.J.E.,Groth,E.J. 1976,*A.A.*,**53**,131.

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

Pence,W. 1976,*Ap.J.*,**203**,39.

Peterson,B.A.,Ellis,R.E.,Kibblewhite,E.J.,Bridgeland,M.T., Hooley,T.,Horne,D. 1979,*Ap.J.*,**233**,L109.

Pratt,N.M.,Martin,R.,Alexander,L.W.G.,Walker,G.S., Williams,P.R. 1975,*Image Processing Techniques in Astronomy*,Proceedings (Reidel,Dordrecht),**43**,217.

Press,W.H.,Shecter,P. 1974,*Ap.J.*,**187**,425.

Rheault,C.,Hardy,E. 1980,*SPIE Appl. Digital Image Processing to Astronomy*,**264**,200.

Scott,E.L. 1962,*Problems of Extragalactic Research*, ed.G.C.McVittie (New York:Macmillan).

Sebok,W.L. 1979,*A.J.*,**84**,1526 (chapter 2 in this thesis.).

Sebok,W.L. 1980,*SPIE Appl.Digital Image Processing to Astronomy*,**264**,213 (Chapter 3 in this thesis.).

Sharp,N.A. 1979,*A.A.*,**74**,308.

Shecter,P. 1976,*Ap.J.*,**203**,297.

Thuan,T.X.,Gunn,J.E., 1976,*P.A.S.P.*,**88**,543.

Totsuji,H.,Kihara,T. 1969,*Publ.Astron.Soc.Japan*,**21**,221.

Tyson,J.A.,Jarvis,J.F. 1979,*Ap.J.*,**230**,L153.

Wells,D.C. 1972,Ph.D.Thesis "Integrated Spectral Energy Distributions of Galaxies",University of Texas.

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

FIGURE CAPTIONS

1) Counts of Stars and Galaxies for the 4 Schmidt Plates.

2) Averaged Counts of Stars and Galaxies

_._. Galaxies from Schmidt Plates.

---- Galaxies from the 2 KPNO 4 m. Telescope Plates.

.... Stars from Schmidt Plates

solid line Stars from the 2 plates.

3) Comparison of KPNO 4m Plate MP 3299 with the Same Part of Sky on Schmidt Plate.

4) Comparison of KPNO 4m plate MP 2752 with the Same Part of Sky on Schmidt Plate.

5) Average Schmidt Galaxy Counts Compared with other Workers' Results.

6) Isophotal Magnitude vs. Aperture Magnitude for Plate PS23634.

7) Correlation Function

$16.0 < \text{Aperture Magnitude} < 16.5$

Solid Line = Fit

Dashed Line = Line of Slope -0.8.

8) Correlation Function

$16.5 < \text{Aperture Magnitude} < 17.0$

Solid Line = Fit

Dashed Line = Line of Slope -0.8.

9) Correlation Function

$17.0 < \text{Aperture Magnitude} < 17.5$

Solid Line = Fit

Dashed Line = Line of Slope -0.8.

10) Correlation Function

$17.5 < \text{Aperture Magnitude} < 18.0$

Solid Line = Fit

Dashed Line = Line of Slope -0.8.

11) Correlation Function

$18.0 < \text{Aperture Magnitude} < 18.5$

Solid Line = Fit

Dashed Line = Line of Slope -0.8.

12) Correlation Function

$18.5 < \text{Aperture Magnitude} < 19.0$

Solid Line = Fit

Dashed Line = Line of Slope -0.8.

13) Correlation Function

$19.0 < \text{Aperture Magnitude} < 19.5$

Solid Line = Fit

Dashed Line = Line of Slope -0.8.

14) Correlation Function

$19.5 < \text{Aperture Magnitude} < 20.0$

Solid Line = Fit

Dashed Line = Line of Slope -0.8.

15) Correlation Function

$20.0 < \text{Aperture Magnitude} < 20.5$

Solid Line = Fit

Dashed Line = Line of Slope -0.8.

16) K Correction for Gunn Red

17) Correlation Function at 1' vs Magnitude

Affect of γ on Models.

$H_0 = 50$, $q_0 = 0$ $\varepsilon = 0$

18) Correlation Function at 1' vs Magnitude

Affect of ε on Models.

$H_0 = 50$, $q_0 = 0$ $\gamma = 1.8$

19) Correlation Function at 1' vs Magnitude

Affect of q_0 on Models.

$H_0 = 50$, $\gamma = 1.8$ $\varepsilon = 0$

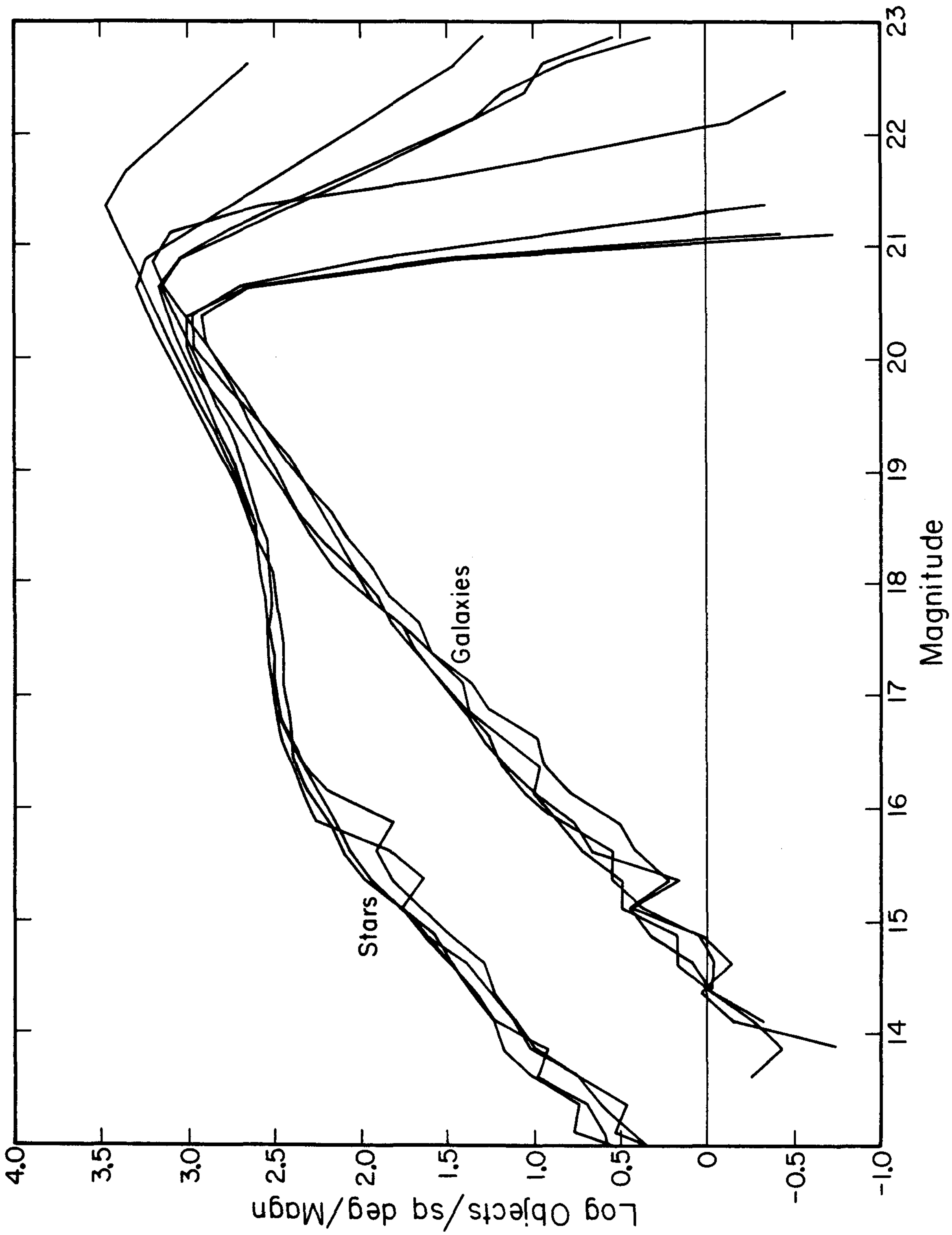
Log Objects/sq deg/Magn
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
-0.5
-1.0
Magnitude
14
15
16
17
18
19
20
21
22
23
Stars
Galaxies

Figure 1

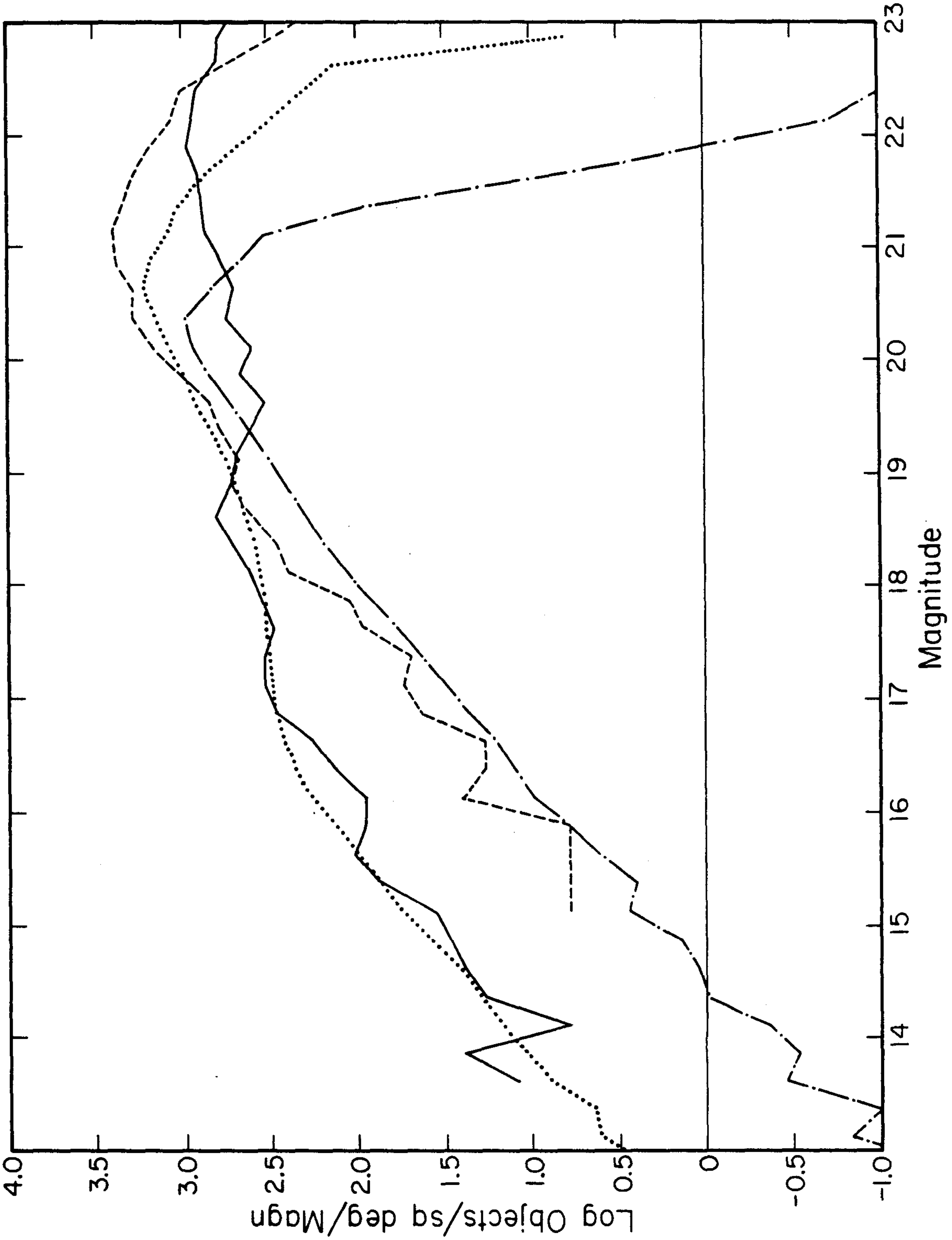
Log Objects/sq deg/Magn
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
-0.5
-1.0
Magnitude
14
15
16
17
18
19
20
21
22
23

Figure 2

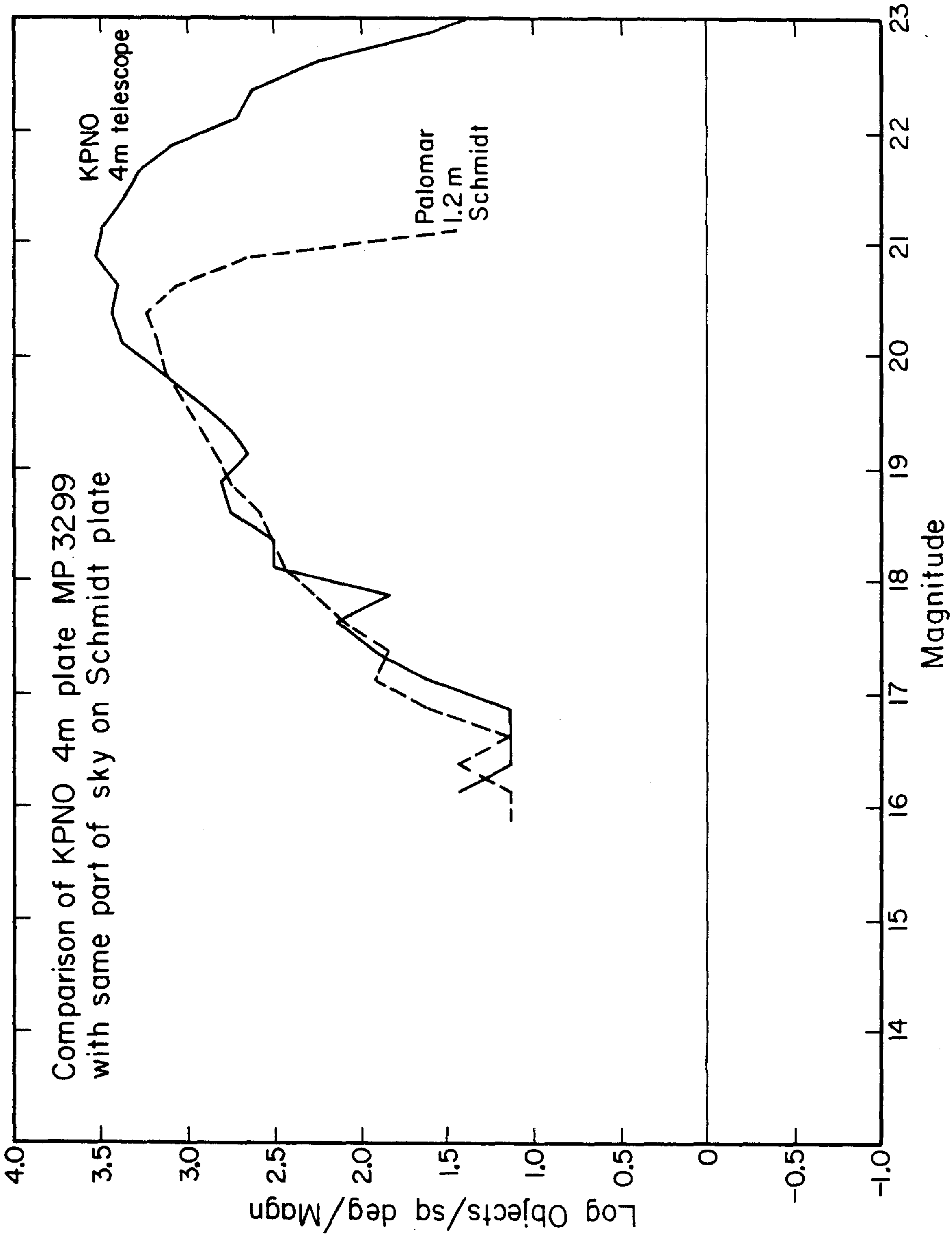
Comparison of KPNO 4m plate MP 3299
with same part of sky on Schmidt plate
KPNO
4m telescope
Palomar
1.2 m
Schmidt
Log Objects/sq deg/Magn
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
-0.5
-1.0
14
15
16
17
18
19
20
21
22
23
Magnitude

Figure 3

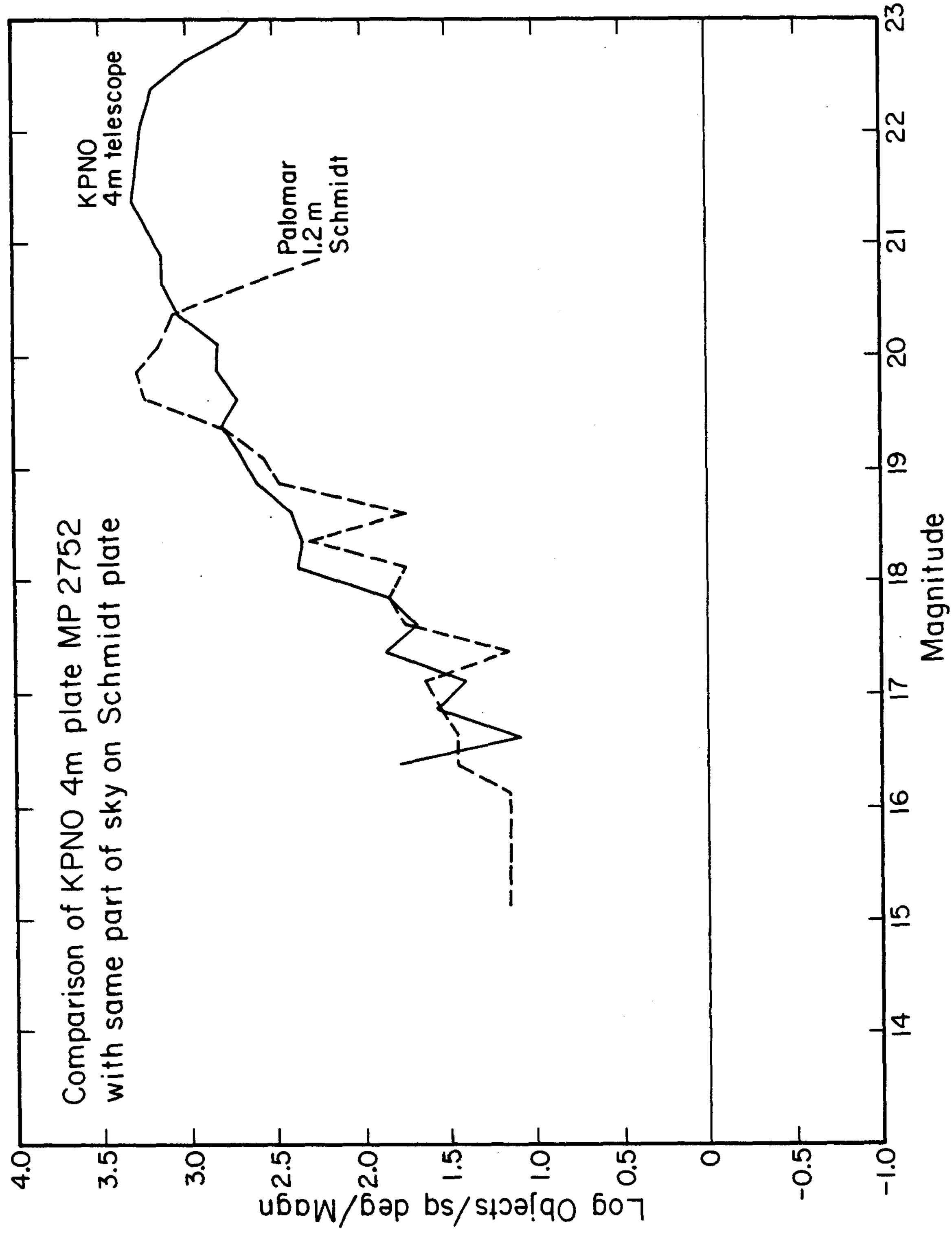
Comparison of KPNO 4m plate MP 2752
with same part of sky on Schmidt plate
KPNO
4m telescope
Palomar
1.2 m
Schmidt
Log Objects/sq deg/Magn
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
-0.5
-1.0
Magnitude
14
15
16
17
18
19
20
21
22
23

Figure 4

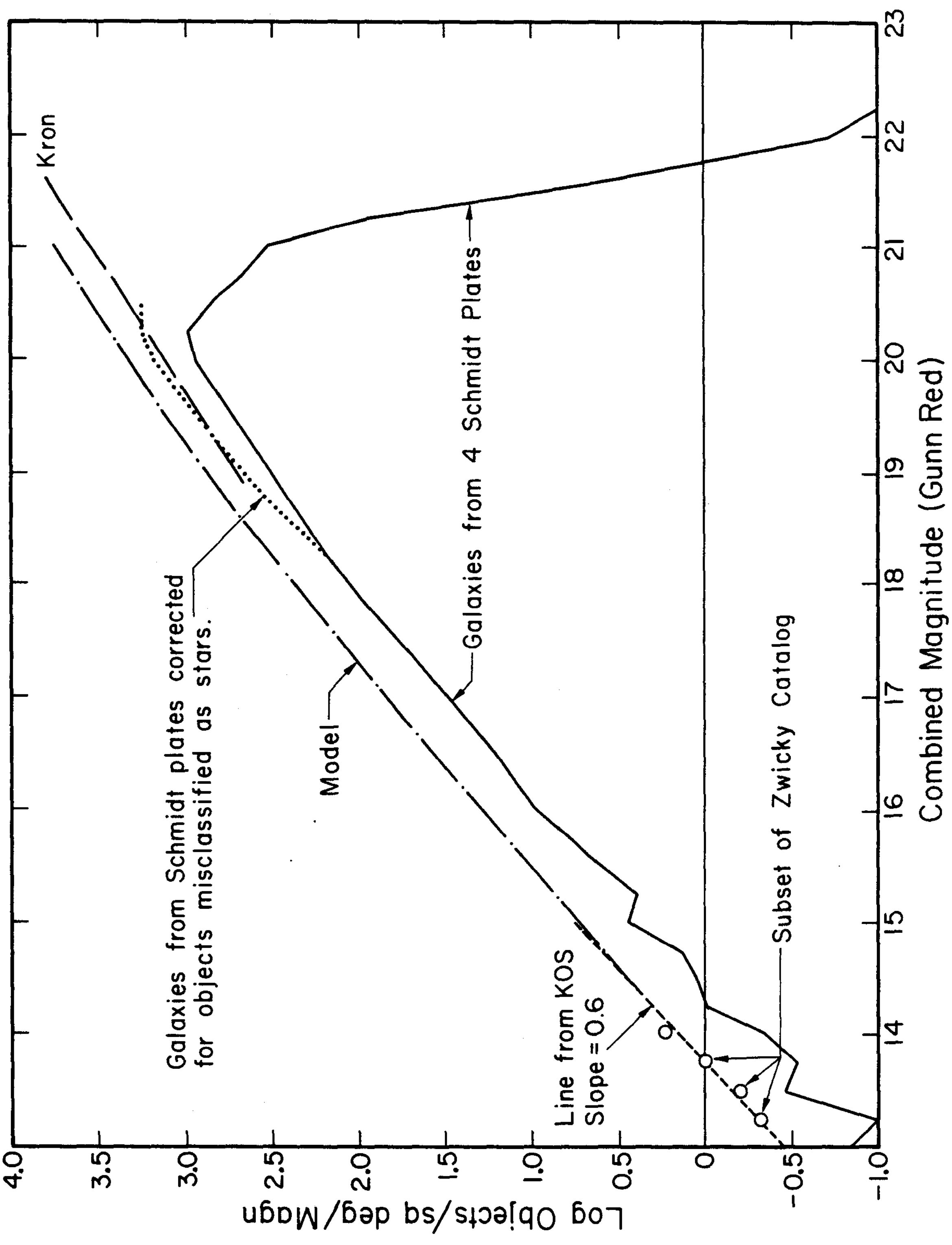
Kron
Galaxies from 4 Schmidt Plates
Galaxies from Schmidt plates corrected for objects misclassified as stars.
Model
Line from KOS
Slope = 0.6
Subset of Zwicky Catalog
Combined Magnitude (Gunn Red)
Log Objects/sq deg/Magn
14
15
16
17
18
19
20
21
22
23
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
-0.5
-1.0

Figure 5

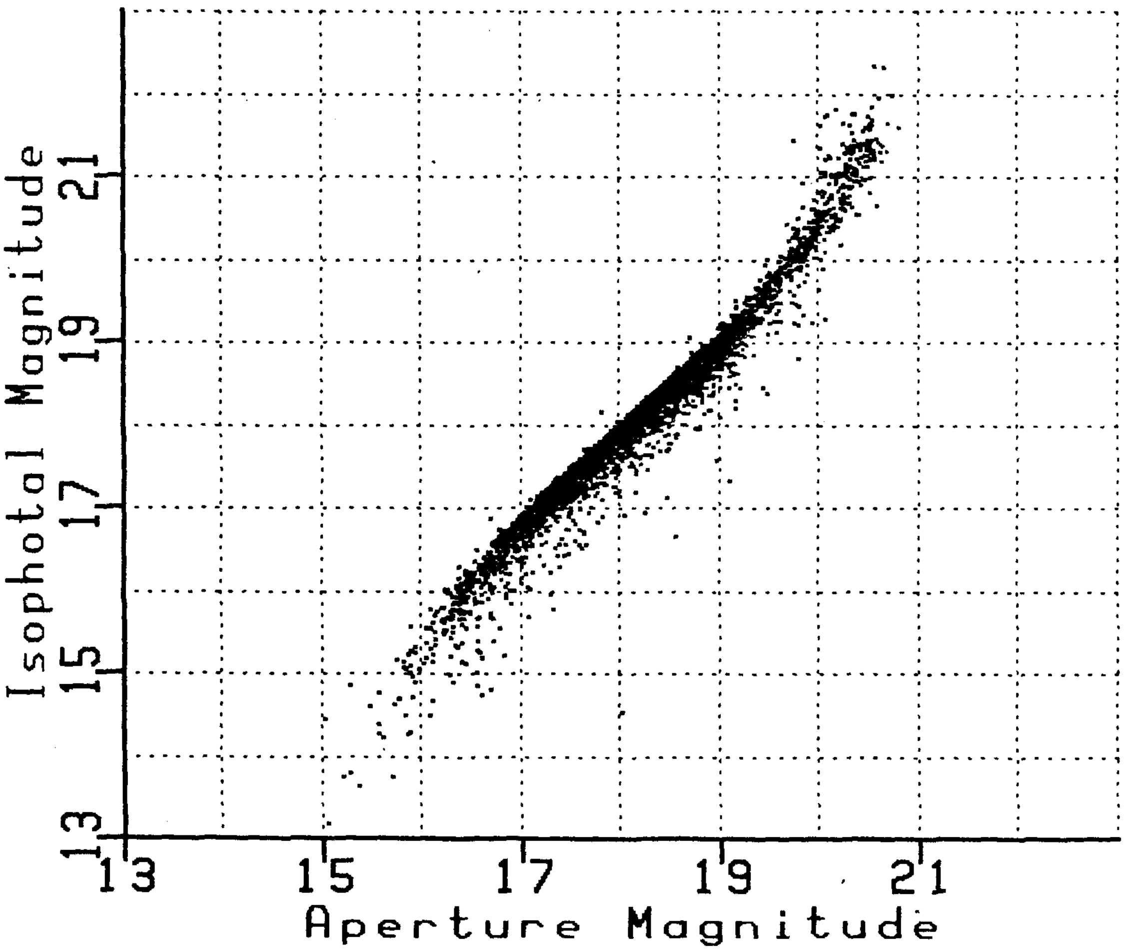
Isophotal Magnitude
13
15
17
19
21
13
15
17
19
21
Aperture Magnitude

Figure 6

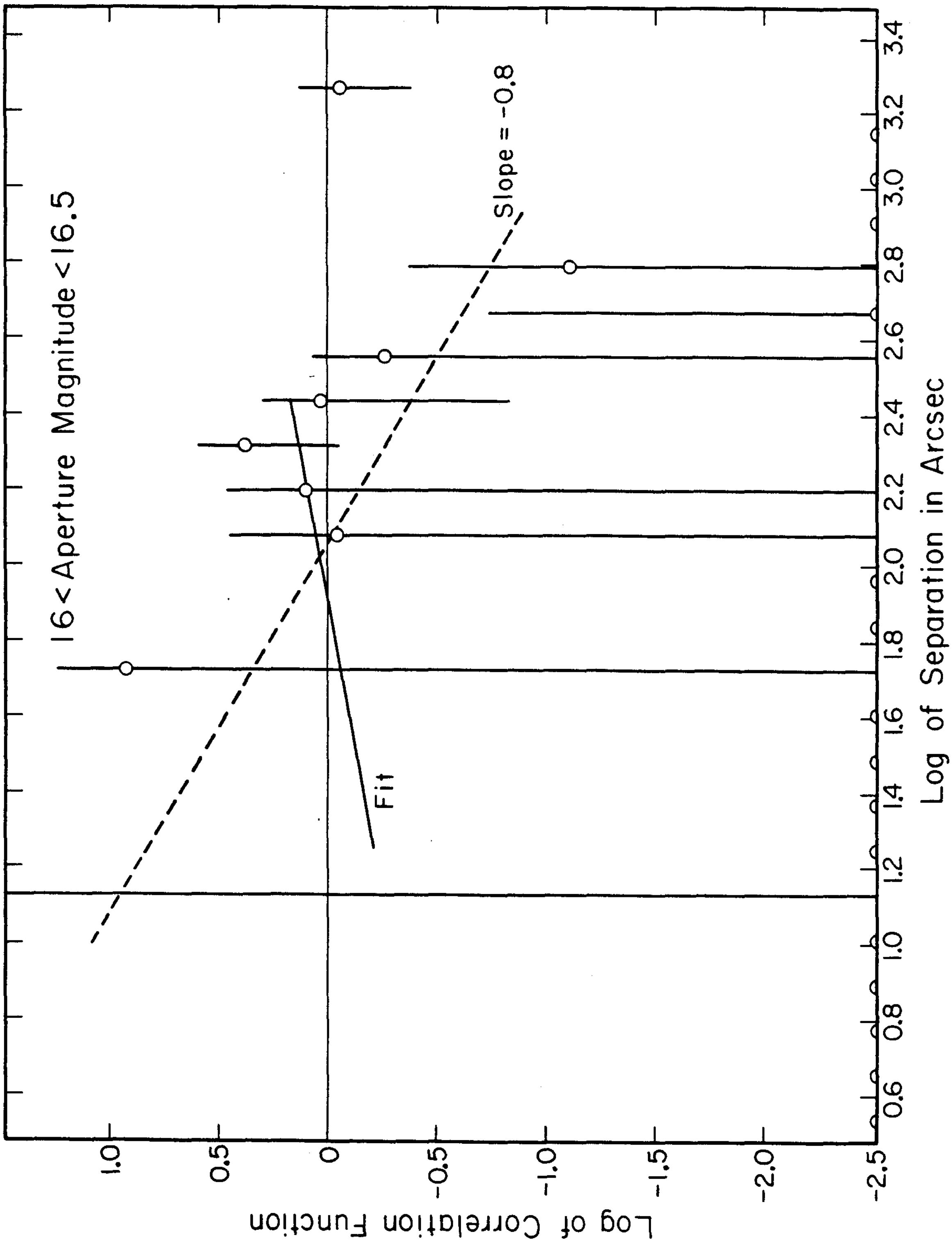
16 < Aperture Magnitude < 16.5
Slope = -0.8
Fit
Log of Correlation Function
1.0
0.5
0
-0.5
-1.0
-1.5
-2.0
-2.5
Log of Separation in Arcsec
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
3.2
3.4

Figure 7

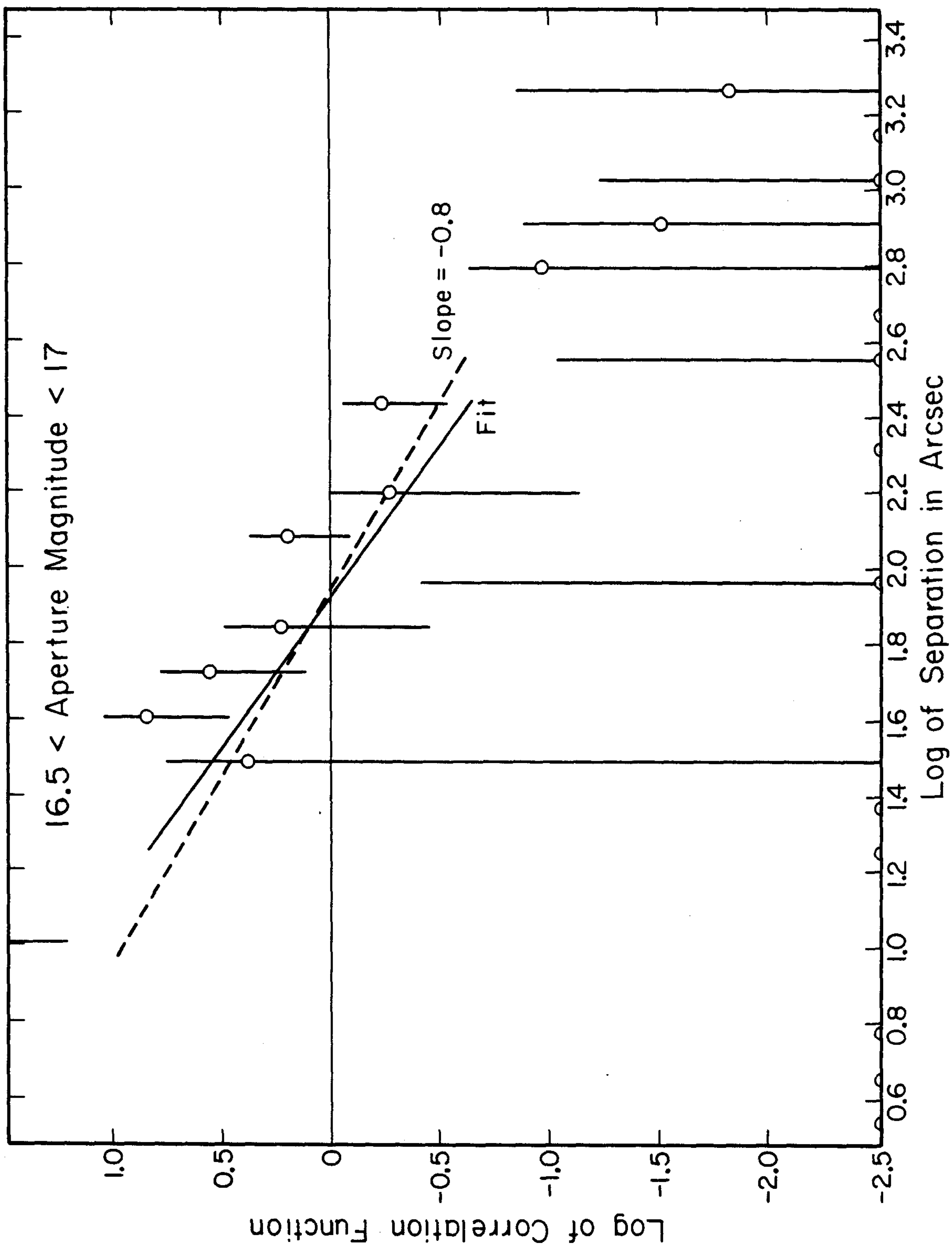
16.5 < Aperture Magnitude < 17
Slope = -0.8
Fit
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
3.2
3.4
Log of Separation in Arcsec
1.0
0.5
0
-0.5
-1.0
-1.5
-2.0
-2.5
Log of Correlation Function

Figure 8

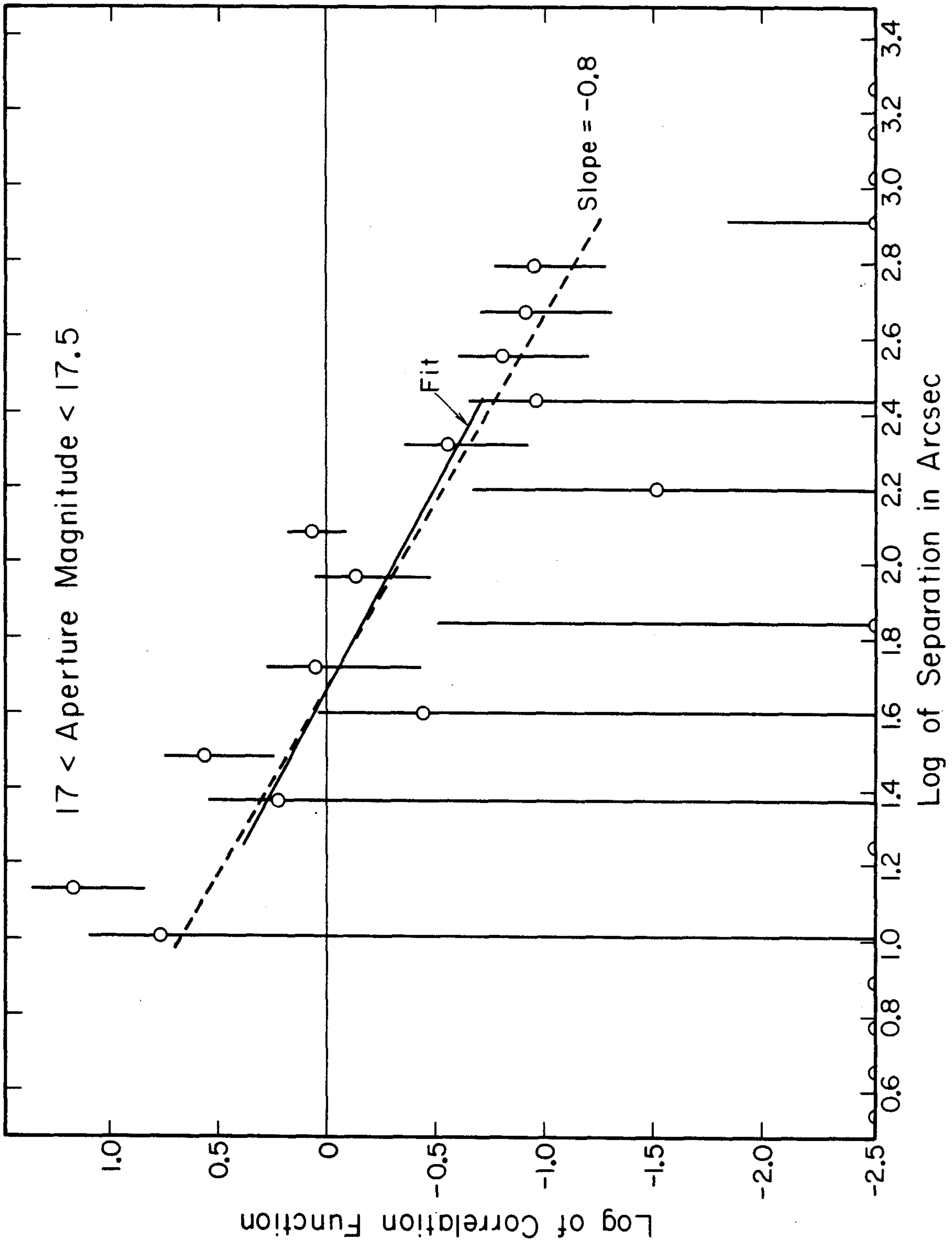
17 < Aperture Magnitude < 17.5
Fit
Slope = -0.8
Log of Separation in Arcsec
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
3.2
3.4
Log of Correlation Function
1.0
0.5
0
-0.5
-1.0
-1.5
-2.0
-2.5

Figure 9

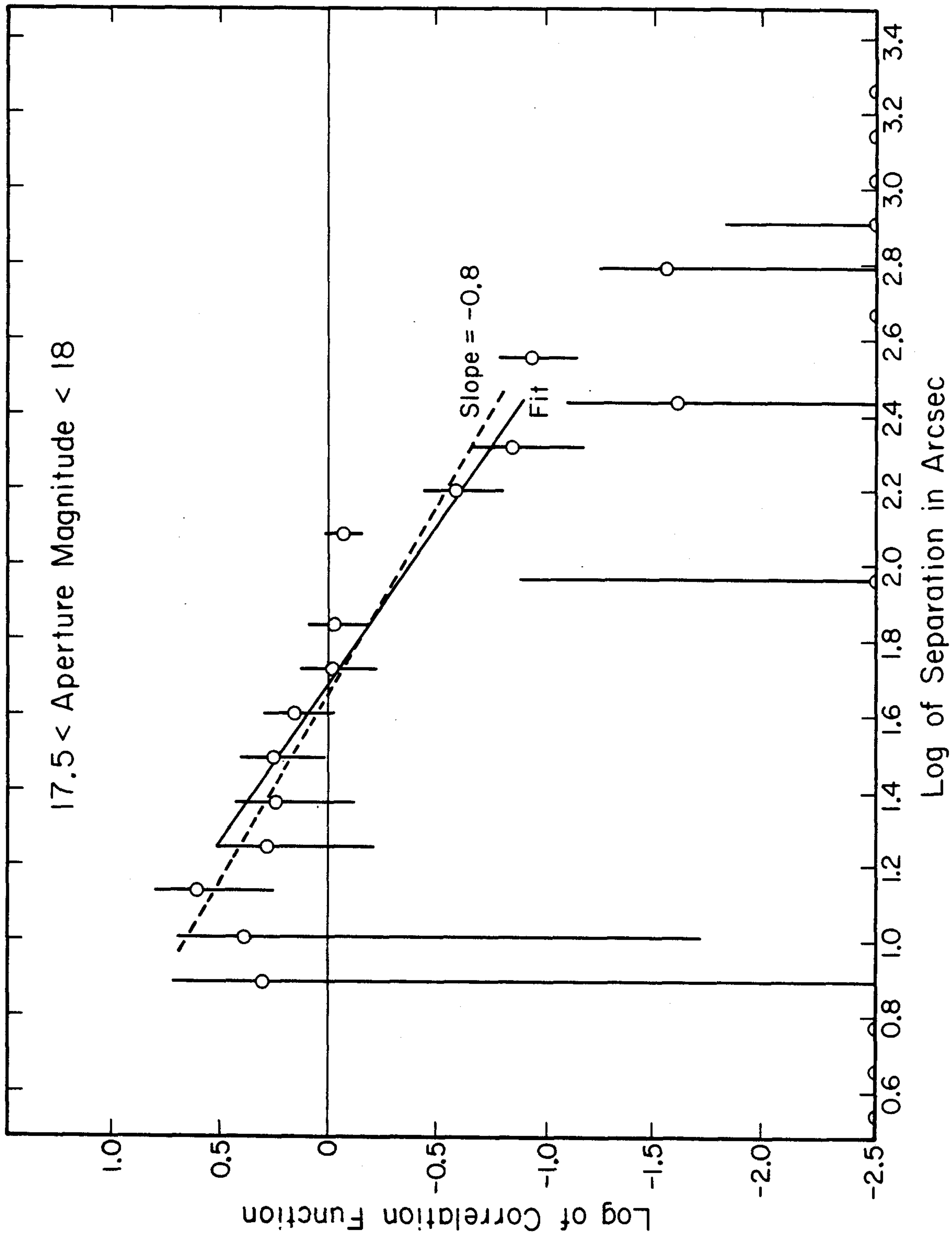
17.5 < Aperture Magnitude < 18
Slope = −0.8
Fit
Log of Separation in Arcsec
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
3.2
3.4
Log of Correlation Function
1.0
0.5
0
−0.5
−1.0
−1.5
−2.0
−2.5

Figure 10

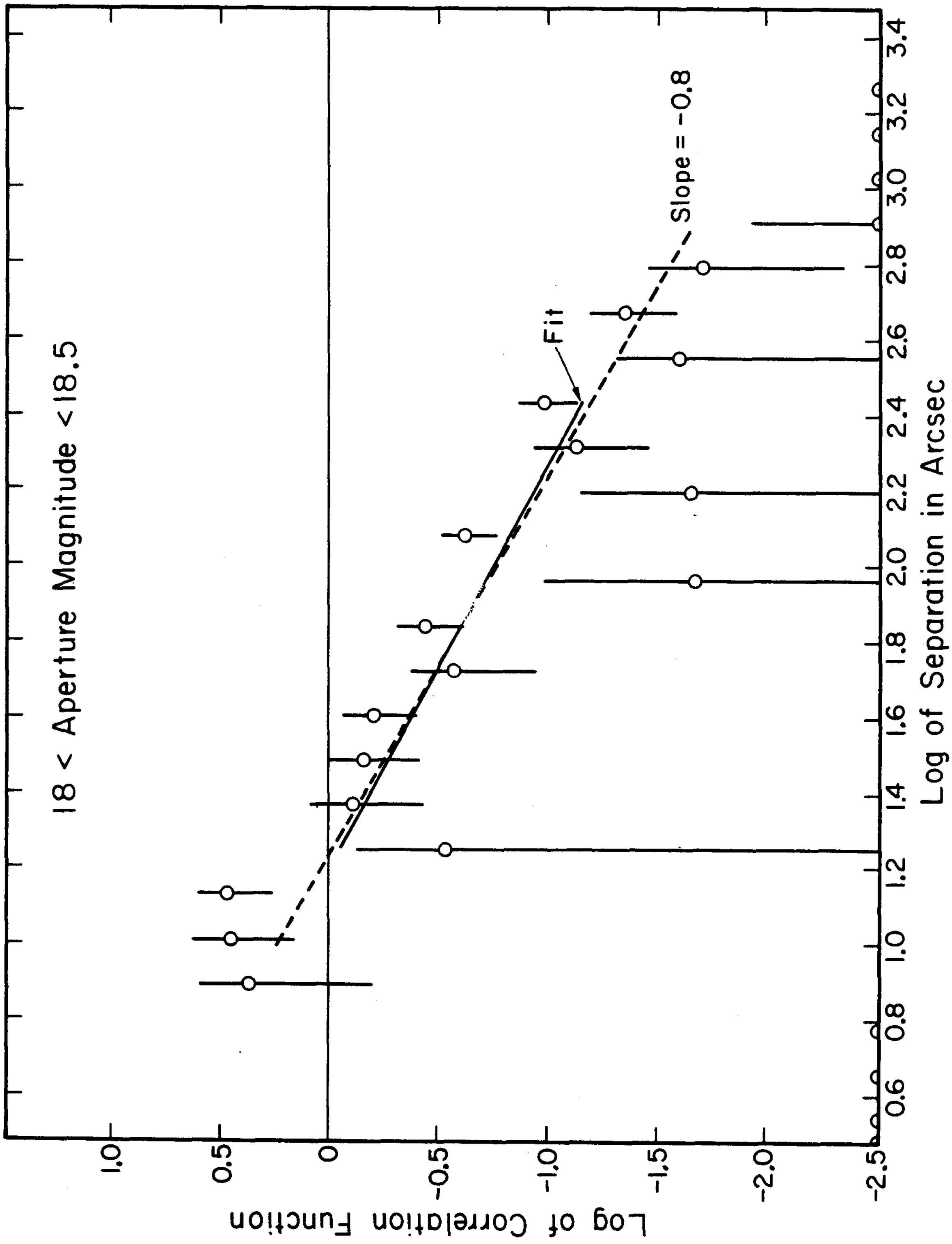
18 < Aperture Magnitude <18.5
Fit
Slope = −0.8
Log of Separation in Arcsec
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
3.2
3.4
Log of Correlation Function
1.0
0.5
0
−0.5
−1.0
−1.5
−2.0
−2.5

Figure 11

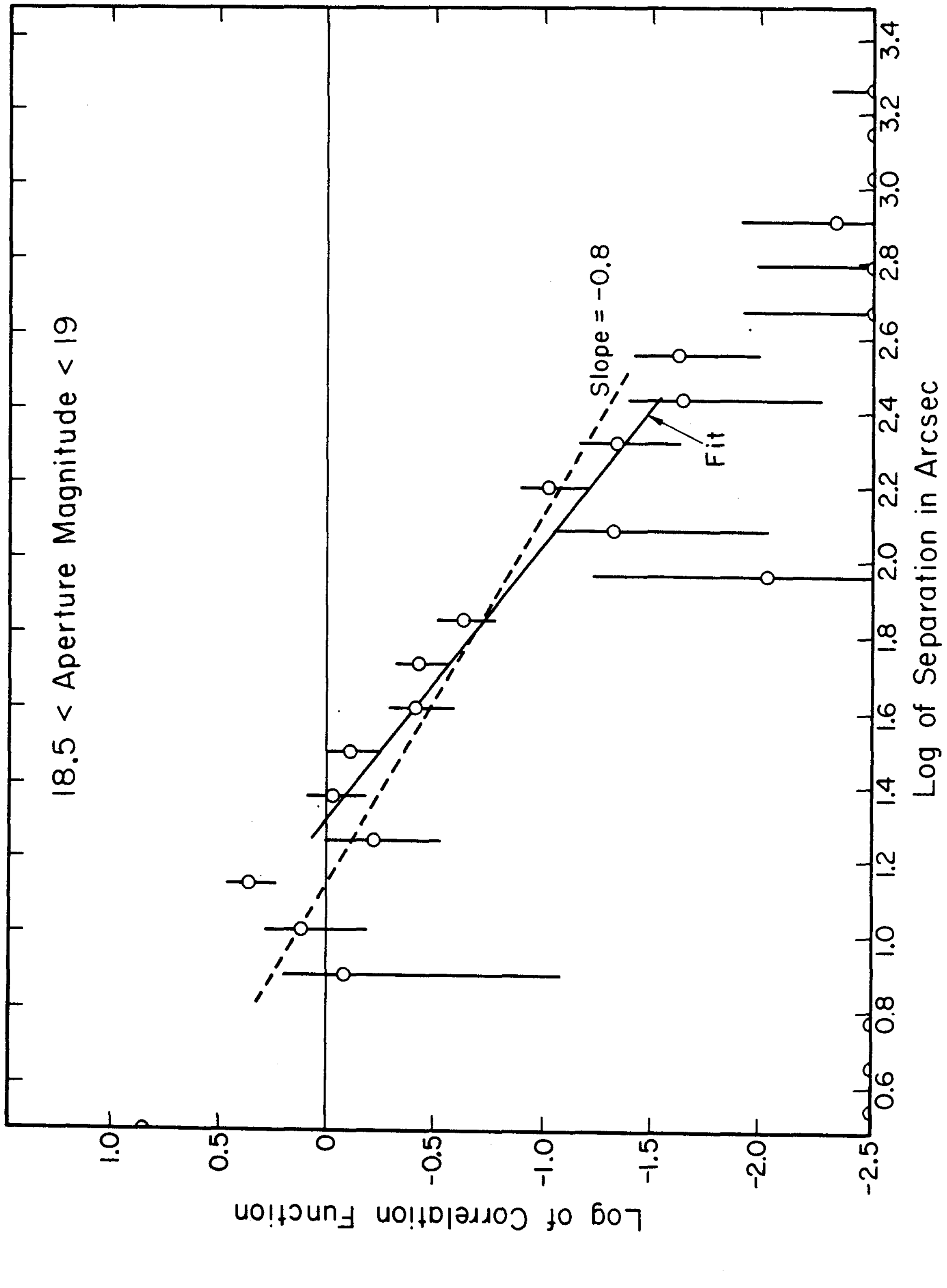
18.5 < Aperture Magnitude < 19
Slope = -0.8
Fit
Log of Separation in Arcsec
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
3.2
3.4
Log of Correlation Function
1.0
0.5
0
-0.5
-1.0
-1.5
-2.0
-2.5

Figure 12

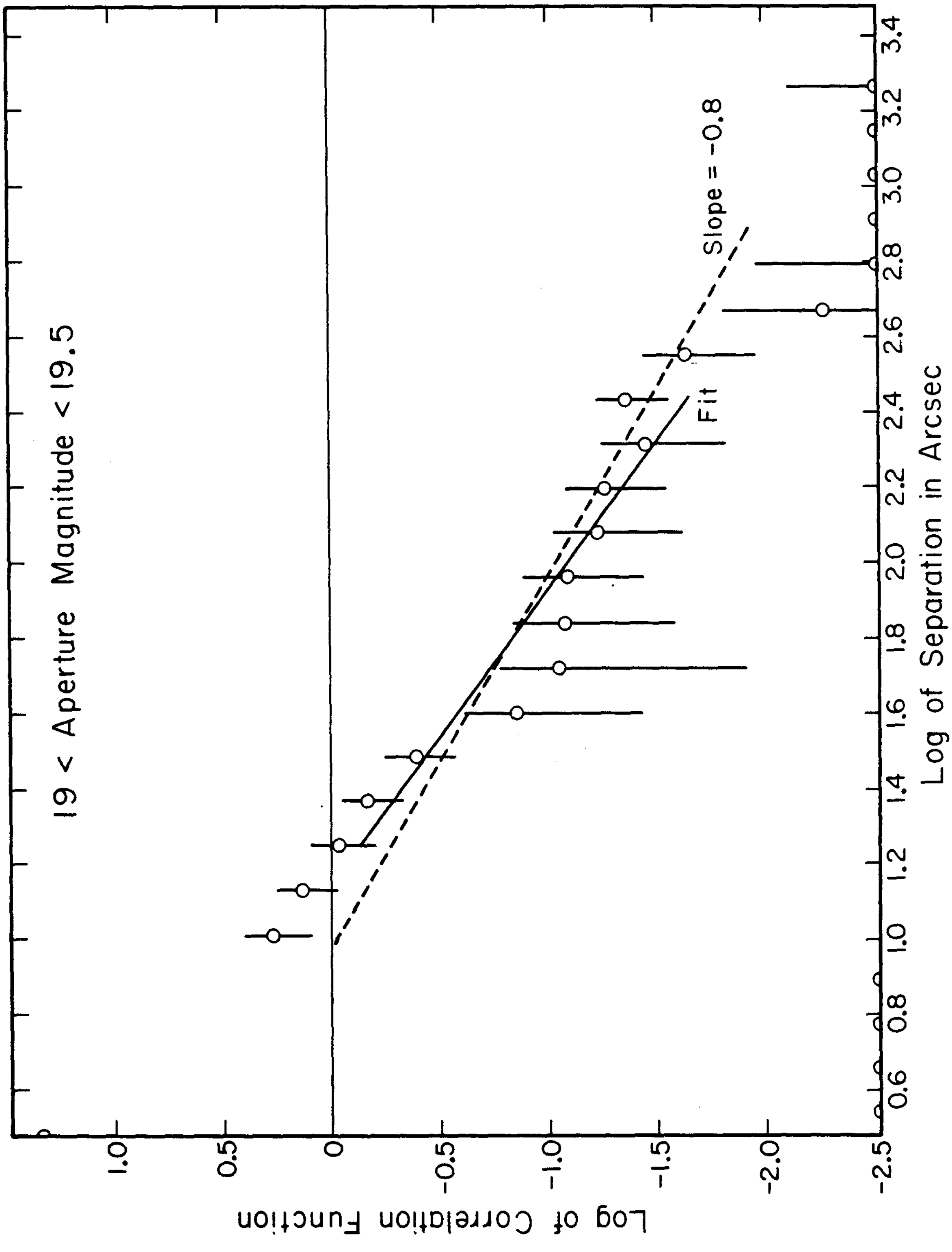
19 < Aperture Magnitude < 19.5
Slope = −0.8
Fit
Log of Correlation Function
1.0
0.5
0
−0.5
−1.0
−1.5
−2.0
−2.5
Log of Separation in Arcsec
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
3.2
3.4

Figure 13

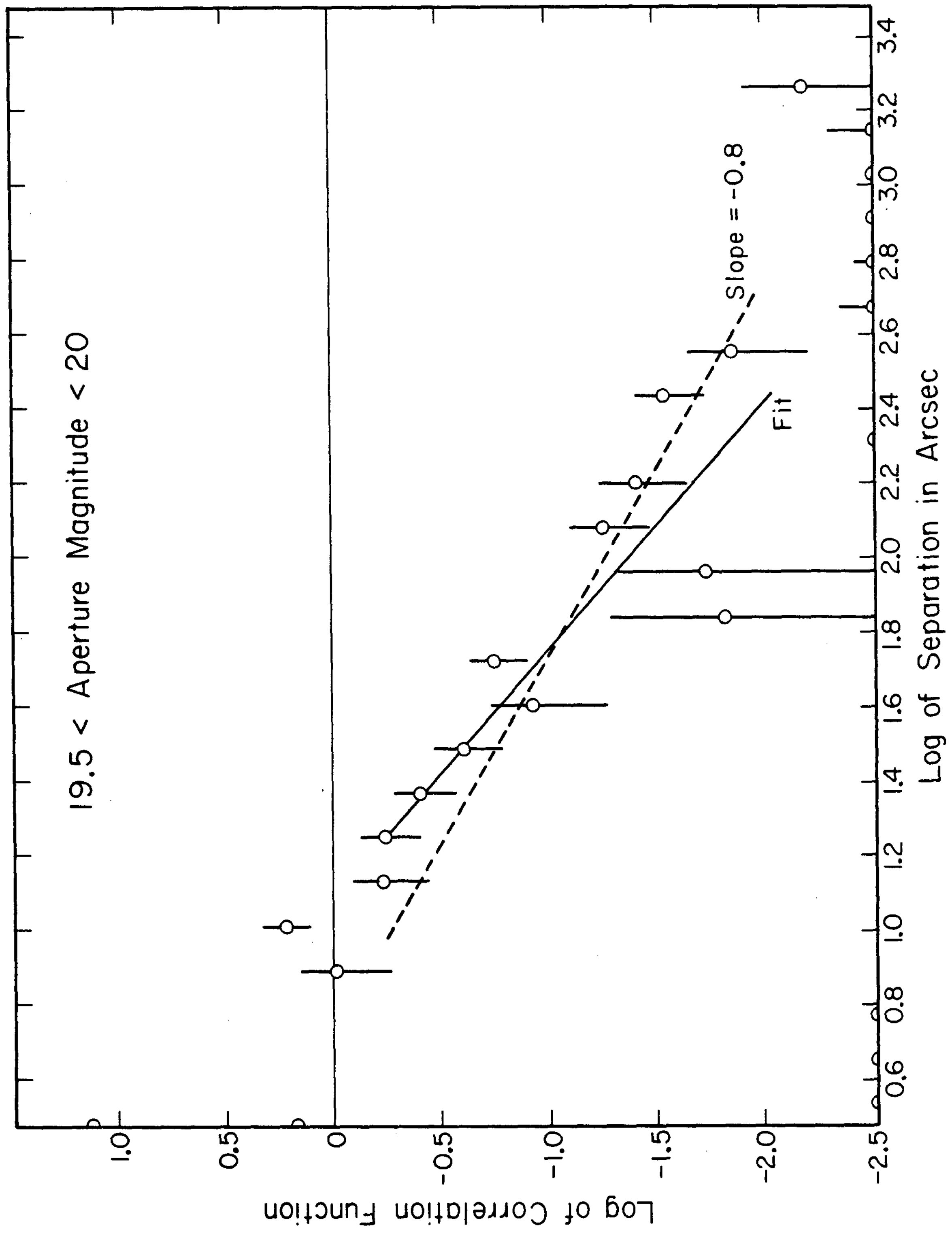
19.5 < Aperture Magnitude < 20
Slope = -0.8
Fit
Log of Separation in Arcsec
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
3.2
3.4
Log of Correlation Function
1.0
0.5
0
-0.5
-1.0
-1.5
-2.0
-2.5

Figure 14

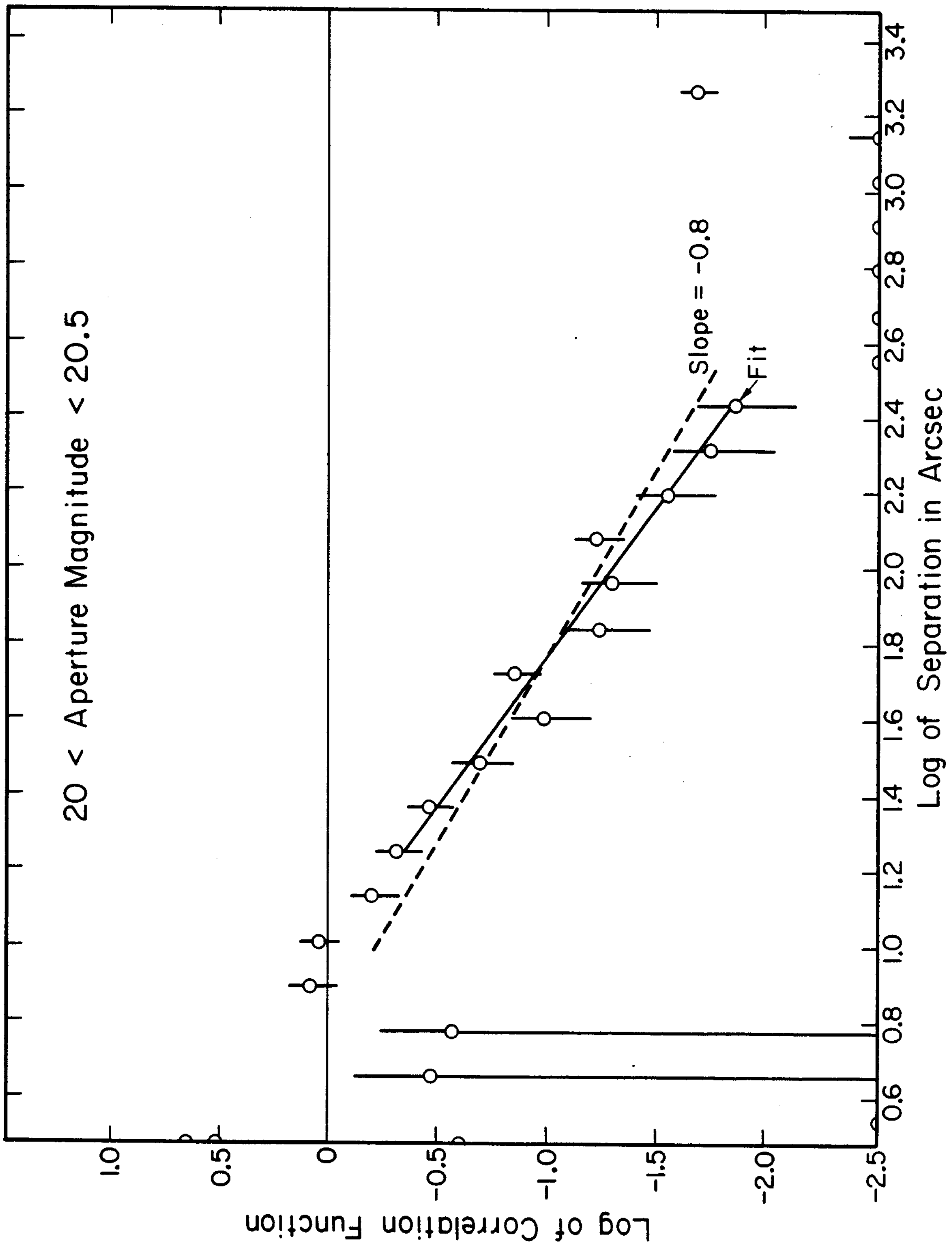
20 < Aperture Magnitude < 20.5
Slope = −0.8
Fit
Log of Separation in Arcsec
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
3.2
3.4
Log of Correlation Function
1.0
0.5
0
−0.5
−1.0
−1.5
−2.0
−2.5

Figure 15

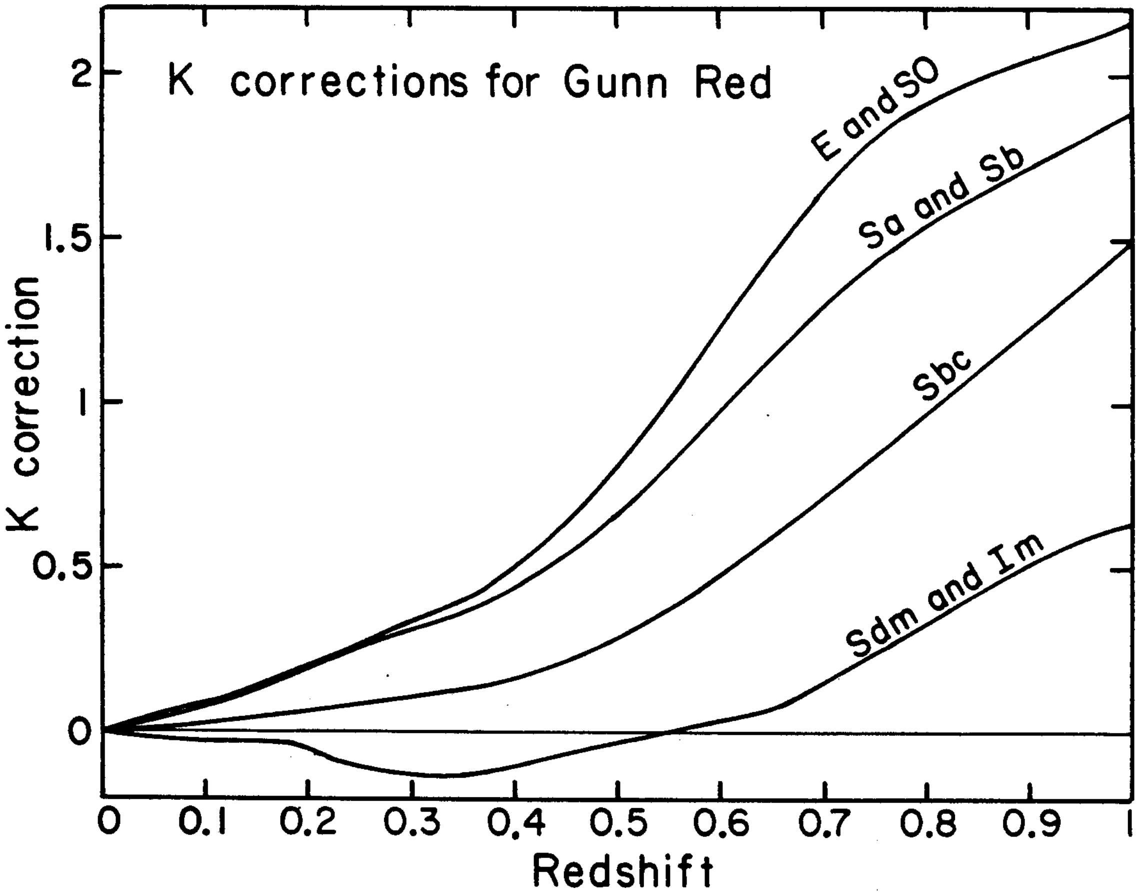
K corrections for Gunn Red
E and SO
Sa and Sb
Sbc
Sdm and Im
K correction
0
0.5
1
1.5
2
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Redshift

Figure 16

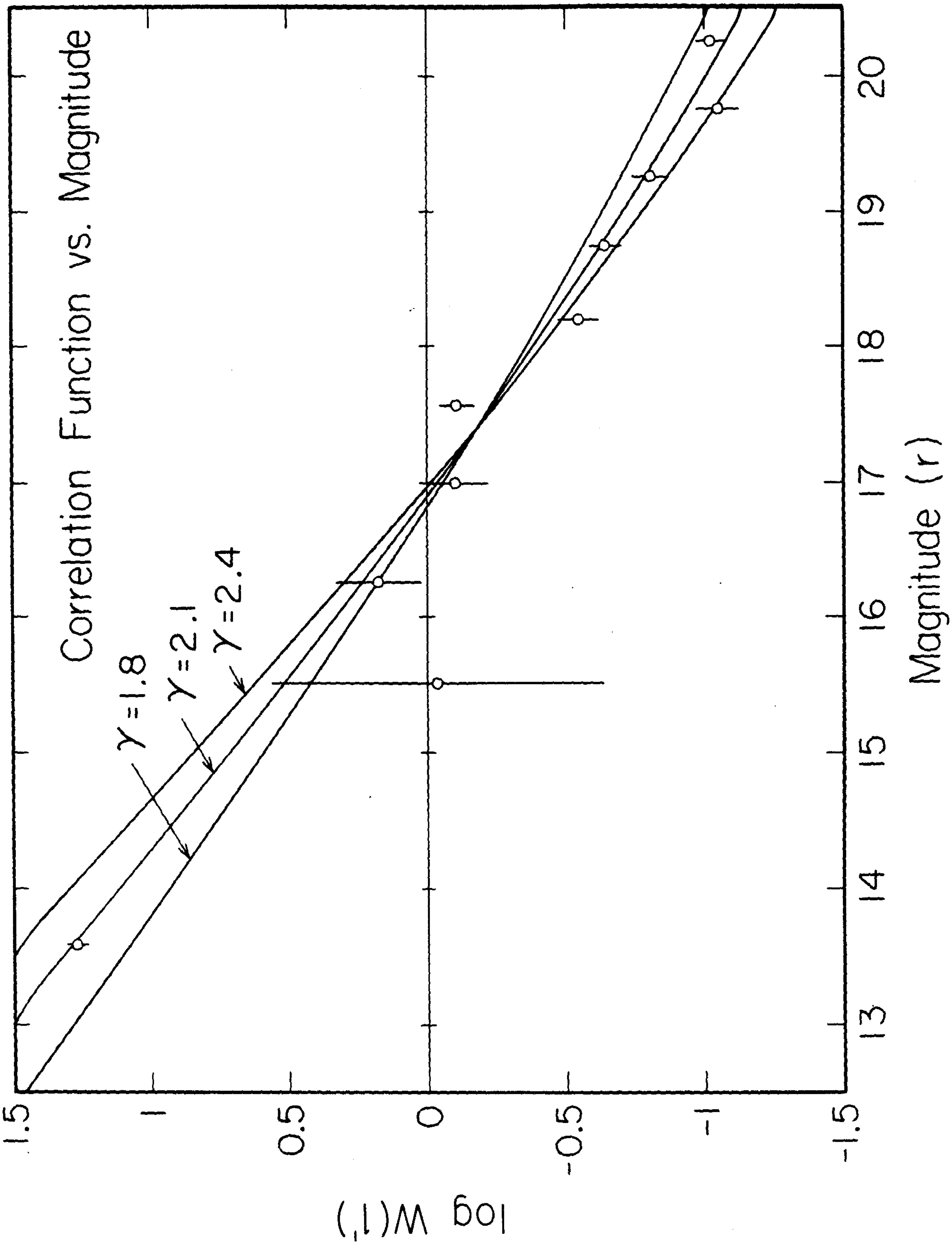
Correlation Function vs. Magnitude
$\gamma = 1.8$
$\gamma = 2.1$
$\gamma = 2.4$
1.5
1
0.5
0
-0.5
-1
-1.5
log W(1')
13
14
15
16
17
18
19
20
Magnitude (r)

Figure 17

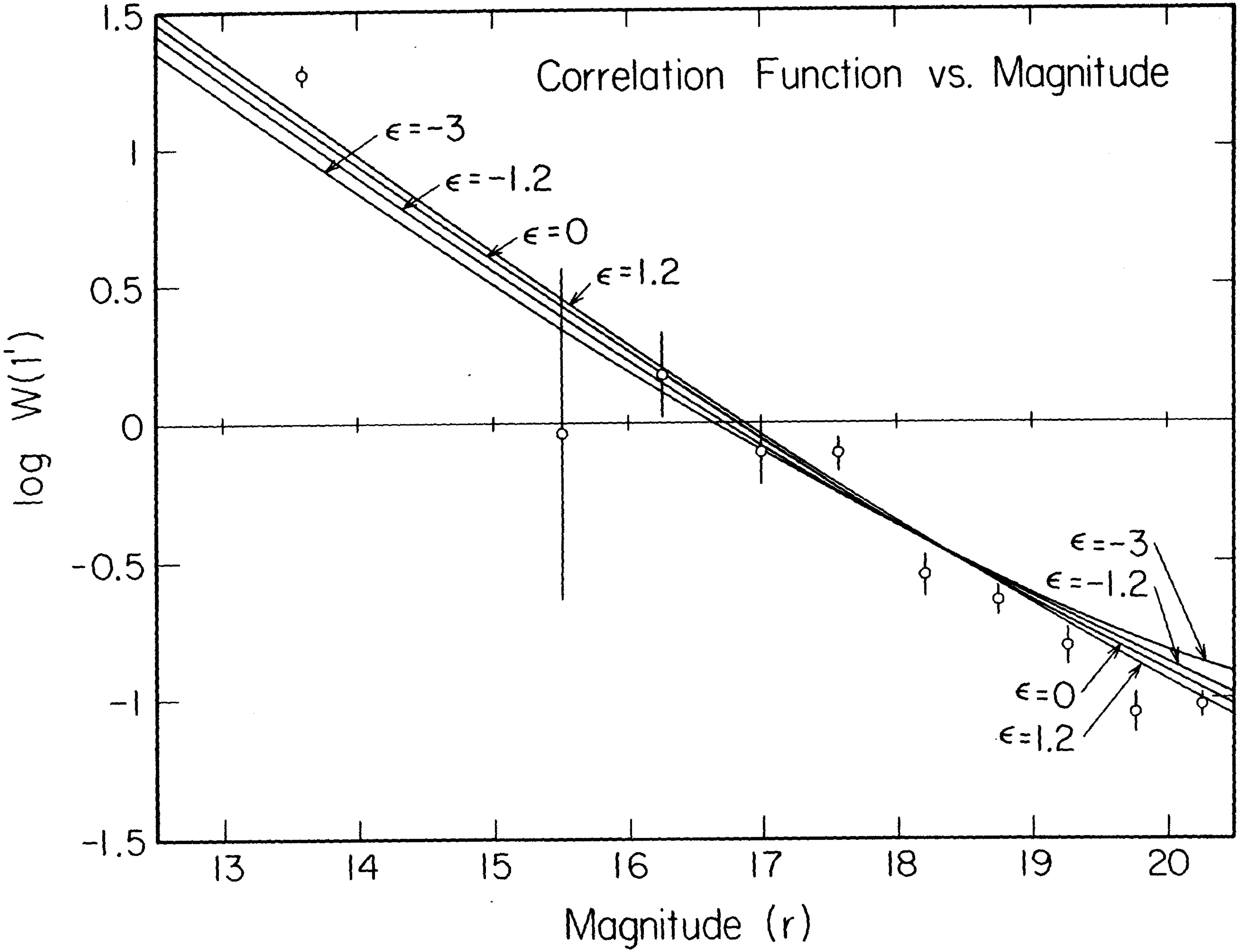
Correlation Function vs. Magnitude
$\epsilon=-3$
$\epsilon=-1.2$
$\epsilon=0$
$\epsilon=1.2$
$\epsilon=-3$
$\epsilon=-1.2$
$\epsilon=0$
$\epsilon=1.2$
log W(1')
1.5
1
0.5
0
-0.5
-1
-1.5
Magnitude (r)
13
14
15
16
17
18
19
20

Figure 18

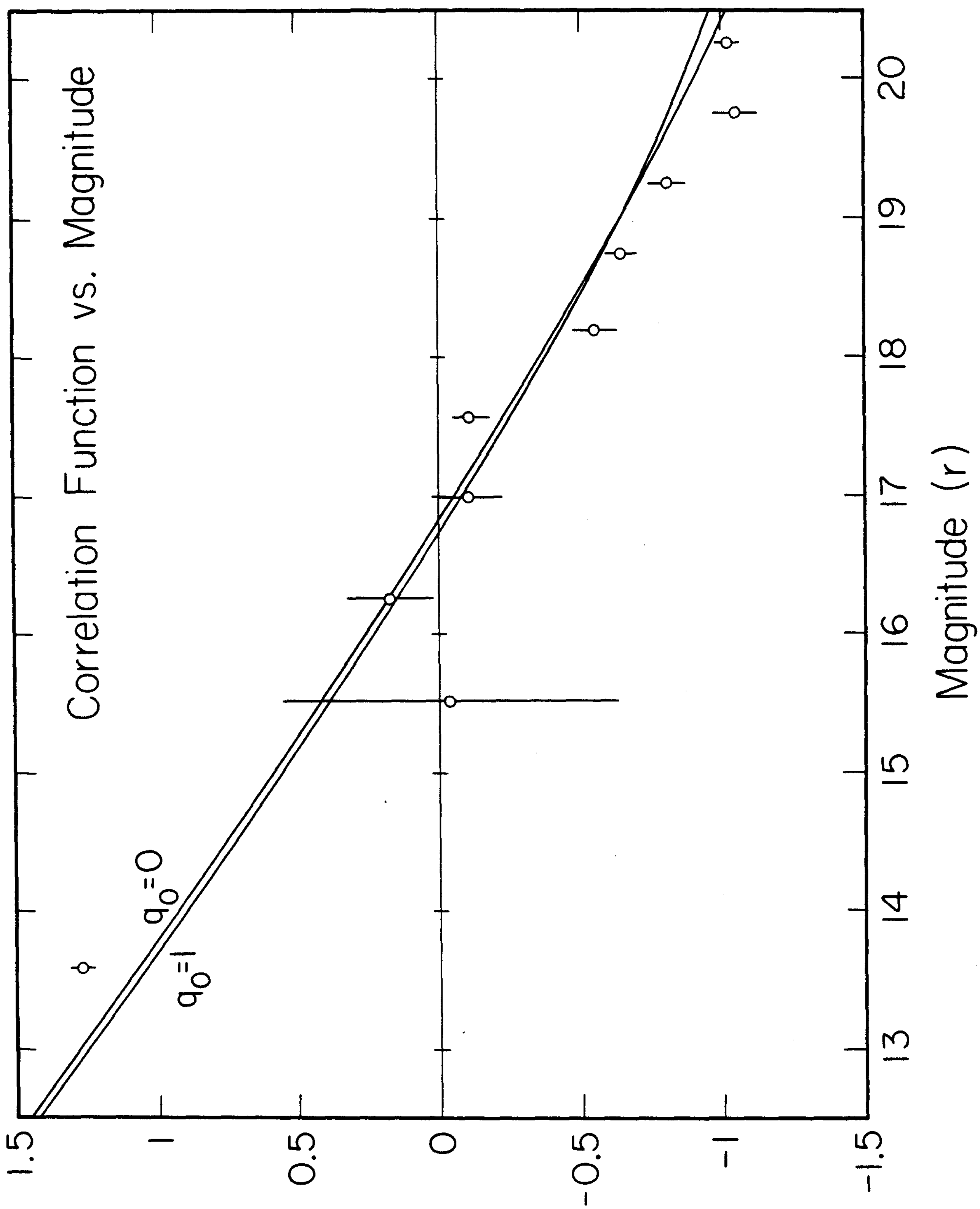
Correlation Function vs. Magnitude
$q_0=0$
$q_0=1$
log W(1')
1.5
1
0.5
0
-0.5
-1
-1.5
Magnitude (r)
13
14
15
16
17
18
19
20

Figure 19