Galaxy Evolution
Observations, Analysis, and Theory

Jeremy Samuel Heyl

Institute of Astronomy
and
Fitzwilliam College

A thesis submitted to the University of Cambridge
in partial fulfillment of the requirements
for the degree of Master of Science

3 November 1994

Abstract

This thesis presents the coherent analysis of several redshift surveys and the results of a series of numerical simulation of galaxy formation. In the first part, the new `AUTOFIB` galaxy redshift is combined with several extant surveys (DARS, BES, LDSS-1, LDSS-2) to coherently and directly determine the evolution of the galaxy luminosity function. The construction of the luminosity function even at moderate redshift crucially depends on an understanding of the k-corrections of the galaxies within the survey. I introduce and implement a new method for determining (without human intervention) both an approximate spectroscopic classification and the k-correction. To measure the actual evolution of the luminosity function as a function of redshift and spectral type, I propose two new maximum-likelihood techniques. The resulting luminosity functions evolve strongly with redshift. Both the faint-end slope and normalisaton of the luminosity function increase with redshift while the cutoff luminosity remains nearly constant. This result appears to be independent of the method of derivation.

When the sample is divided by spectral type, the evolution both in terms of the spectra and the luminosity function is found to be strongest amongst late-type galaxies. The number of late-type galaxies increases markédly with redshift, especially at the faint end. On the other hand, the number of faint elliptical galaxies appears to decrease with redshift. That is, the vast majority of these starforming galaxies have faded below the limits of today's surveys since $z \sim 0.2$.

The second part of this thesis discusses the effect of cosmology on the observed properties of galaxies and their evolution. The "block" model of galaxy formation is used to explore galaxy formation in several cosmologies with structure forming in a framework either CDM or C+HDM. The simulated galaxy populations are compared with the observed luminosity functions in the B and K-bands, Tully-Fisher relation, $B-K$-colour distribution, number-magnitude relation (again in the B and K-bands), magnitude-limited redshift distributions, and evolution of the B-band luminosity function. I introduce a new method for deriving the properties of galaxies observed in magnitude-limited samples and use this method to calculate the redshift distributions and number counts directly from the simulations.

Each of the models has its own advantages. Low-density universes fit the observed Tully-Fisher relation best, and a universe with a low Hubble constant predicts colour distributions closest to those observed. A critical-density fiducial model with $H_0 = 60$ km sec^{-1}Mpc^{-1} and cold dark matter provides the best all around fit to the observed properties of galaxies, especially the bright-end cutoff of the luminosity function and its rapid evolution in the B-band. However, all the models overproduce faint galaxies relative to the local luminosity function. Several possible refinements to the "block" model are discussed: the inclusion of metallicity effects, non-local feedback, inhibited star formation in cooling flows, and initial mass function that varies in space and time.

Preface

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, with the exception of works cited in the text and the following.

The construction of the `AUTOFIB` survey has been a collaborative effort over many years of my supervisor, Professor Richard Ellis, Matthew Colless, Thomas Broadhurst and myself. My contribution to the survey itself, of course, came only during the latter stages of the observational campaign. The results of this survey (Chapters 2 and 5) will be submitted as Ellis, R. S., Colless, M. M., Broadhurst, T. J., Heyl, J. S. & Glazebrook, K., "The Evolution of the Galaxy Luminosity Function" to the *Monthly Notices of the Royal Astronomical Society*.

The simulations of galaxy formation (Chapters 7 and 9) were based on the "block" model developed by Shaun Cole and described in Cole et al. 1994a and papers referenced therein. This project was completed in collaboration with Shaun Cole, Carlos Frenk and Julio Navarro. The results of this investigation have been submitted to the *Monthly Notices of the Royal Astronomical Society* as Heyl, J. S., Cole, S., Frenk, C. S. & Navarro, J. F., "Galaxy Formation in a Variety of Hierarchical Models."

However, the major part of the research presented is my own work. This dissertation has not been submitted for any degree, diploma, or other qualification at any other University.

. . . и рука с рукоЮЮЮЮЮй доЮЮЮйдЮеїм мы оба.

Contents

List of Figures

List of Tables

Part I

The Observations

Chapter 1

Introduction

'It is a capital mistake to theorise before one has data.'
— Sir Arthur Conan-Doyle, *Scandal in Bohemia*

О, совсем бессмысленно, и всЮЮЮеї же неспроста.
— Georgi Ivanov

The study of the distribution of field galaxies with luminosity (the galaxy luminosity function) is a fundamental endeavour of extra-galactic astronomy. In the past decade, both observers and theorists have worked toward characterising the local luminosity function and its evolution with redshift. Although many redshift surveys have been completed in recent years (*e.g.* Peterson et al. 1986, Loveday et al. 1992), both the normalisation (Maddox et al. 1990b) and the faint-end slope (Davis, Summers & Schlegel 1992, McGaugh 1994) of the local luminosity function remain uncertain. The theoretical efforts have intensified the question. Although simulations of galaxy formation have enjoyed several successes (predicting the Tully-Fisher relation, the number counts *etc.*), they predict a faint-end slope markedly steeper than observed in these surveys (Kauffmann & White 1993, Cole et al. 1994a and Part II of this thesis). Is there an over-abundance of faint galaxies locally that the bright surveys miss? Or does our understanding of galaxy formation need revision?

These panoramic surveys have verified that the Schechter (1976) form is appropriate to describe the luminosity function and determine the cut-off luminosity (L_*) to high accuracy. However, extrapolating these results for the various morphological types to higher redshift fails to explain the observed steep slope of the number-magnitude relation $N(m)$ (Heydon-Dumbleton, Collins & MacGillivray 1989, Jones et al. 1991, Metcalfe et al. 1991) in the blue fainter than $B \sim 21$. On the other hand, Mobasher, Sharples & Ellis (1993) have measured the luminosity function in the K-band. This luminosity function predicts faint-number counts in the near-IR that agree with observations (Gardner, Cowie & Wainscoat 1993, Glazebrook, Peacock & Collins 1994). These observations indicate two

principal conclusions. Firstly, the local panoramic surveys do not probe the density of faint galaxies accurately. A galaxy with $M_B = -14 + 5 \log h$ could be detected to a distance of only 15 Mpc at a faint limiting magnitude of $B = 17$. A local population of faint blue galaxies would contribute to the B-band number counts with a nearly Euclidean slope (Driver 1994) – they suffer neither redshift dimming nor k-correction effects. However, in the IR, these galaxies would be too faint to affect the observed K-band number counts. Secondly, galaxies in the past may have been bluer and brighter than today, which again would not affect the K-band counts, but would explain the steep slope in the B-band. Even in the intermediate magnitude range of $17 < B < 21$, the number counts exceed the no-evolution predictions by a factor of two (Maddox et al. 1990b). This corresponds to a redshift ~ 0.1 for an L_* galaxy. It is difficult to accept that the galaxy population has evolved by such a large factor over the past $0.9\ h^{-1}$ Gyr. If the observations are correct, the remaining conclusion is that we live in a $150\ h^{-1}$ Mpc hole and the normalisation of the local luminosity function is not well determined.

Deep spectroscopic surveys directly explore the factors leading to the excess in the number counts and assess the various predictions of galaxy evolution. Spectroscopic surveys consisting of 100-300 galaxies in strict magnitude limited samples fainter than B=21 have been completed by Broadhurst, Ellis & Shanks (1988), Colless et al. (1990), Glazebrook et al. (1993), Lilly, Cowie & Gardner (1991) and Lilly (1993). These surveys have built up the picture that the density of galaxies is a function of apparent magnitude. The distribution of redshifts in these magnitude-limited slices exhibits neither a low nor a high redshift tail, indicative of the two possible explanation of the number counts excess mentioned in a preceding paragraph. Broadhurst, Ellis & Shanks (1988) noted that a luminosity function with an evolving faint-end slope could explain both the observed $N(z)$ and $N(m)$ distributions, but a non- evolving model with an increasing normalisation is also consistent with the data. Furthermore, they found that the median equivalent width of [OII] (an indicator of ongoing star formation) was higher in their fainter sample than in the DARS redshift survey (Peterson et al. 1986). However, Koo & Kron (1992) argue that the fainter surveys are biased toward bluer galaxies due to the k-correction, and therefore the underlying distributions of equivalent widths may be similar.

The deep spectral surveys, for reasons of observing efficiency, consist of samples restricted to lie within narrow apparent magnitude ranges, which make it impossible to strongly constrain the luminosity function, as found by Broadhurst, Ellis & Shanks (1988). Eales (1993) attempted to combine the various surveys to estimate directly the luminosity function as a function of redshift, however the inhomogeneity and limited size of the datasets then available precluded very reliable conclusions. A large homogeneous survey spanning a wide range of apparent magnitude is necessary to reliably estimate the evolution of the luminosity function.

The first part of this thesis describes the combination of the panoramic

DARS survey (Peterson et al. 1986), the multi-fibre Durham/Anglo-Australian Telescope faint galaxy redshift survey (Broadhurst, Ellis & Shanks 1988, BES), the multi-slit LDSS and LDSS-2 surveys (Colless et al. 1990, Glazebrook et al. 1993) with the recently completed `AUTOFIB` redshift survey. These surveys together probe a range of over 12 magnitudes, a factor of nearly 100,000 in flux. The same faint galaxy with $M_B = -14 + 5 \log h$ may be detected out to a redshift of 0.1 in the faintest LDSS-2 fields. Also, the photometry of these fainter surveys has been established using fainter isophotal limits; the surveys are more sensitive to low-surface brightness galaxies and could constrain their contribution to the galaxy population (McGaugh 1994). Moreover, galaxies near M_* may be observed to redshifts in excess of 0.5. The coherent combination of these surveys can constrain the number of faint galaxies over a large local volume as well as the evolution of the luminosity function to redshifts beyond 0.5.

The `AUTOFIB` survey contributes more than half of the galaxies in the combined survey and fills the undersampled region $17 < b_J < 21$. The `AUTOFIB` fibre aperture samples a comparable area to FOCAP (used by Broadhurst, Ellis & Shanks 1988), and the slits of LDSS and LDSS-2. All the faint spectra have been observed and analysed in a similar fashion, resulting in a catalogue of redshifts, magnitudes, equivalent widths and spectral types, a variety of tools to measure the evolution of the galaxy population. Chapter 2 describes some of the initial steps of this reduction and refers to more extensive explanations of the reduction. It also explores some of the major features of the combined survey. Chapter 3 explains the final stage of the reduction procedure, the spectral classification. Chapter 4 describes some estimators for the luminosity function, derives two new estimators and explores some of the statistical interrelationships amongst the estimators. The final chapters in this part (Chapters 5 and 6) describe the local luminosity function and the evolution observed in the combined survey.

Chapter 2

The AUTOFIB Redshift Survey

SUMMARY

The new AUTOFIB *redshift survey fills in the range in apparent magnitude from $b_J = 17$ to 21, between the brighter long-slit surveys and the fainter multi-slit surveys. By combining these new data with data from the brighter and fainter surveys, the luminosity function and its dependency upon redshift may be determined directly (as outlined in Chapters 4 and 5).*

2.1 Namesake

The principle goal of the new AUTOFIB survey, which forms the basis of the combined survey, is to extend the range of galaxy luminosities sampled at moderate z by completing the redshift – apparent magnitude plane in between the early B <17 surveys and the more recent 20< B <24 surveys. In this way it is hoped to provide a *direct* estimate of the shape of the luminosity function (LF) at various redshifts. The AUTOFIB fibre coupler at the AAT is particularly well suited to this task (Parry & Sharples 1988, Ellis & Parry 1988). AUTOFIB places 64 optical fibres with 2 arcsec cores across a 40 arc minute circular field at the f/8 Ritchey-Chretien focal plane of the AAT. During a typical observation, about ten fibres measure the sky. The remaining fifty fibres sample galaxy spectra. The number is well matched to the density of faint galaxies in the range of $17 < B < 21.5$. At the bright limit, a wide magnitude slice may be observed at a high sampling rate. While at the faint end, narrower magnitude slices provide better completeness, and the galaxies are sampled more sparsely. Nowhere in this range is one lacking galaxies. Furthermore, a 2-arcsec aperture corresponds to a physical distance of 2-3 h^{-1} kpc for redshifts between 0.1 and 1.1, well matched to the scale of galaxies.

An observing campaign begins with target selection from COSMOS scans of sky-limited UK-Schmidt plates or prime focus AAT plates. The magnitudes are converted to the colour-corrected photographic b_j system defined by Jones et al. (1991). $b_J \equiv$ Kodak IIIa-J plus GG395 where the limiting surface brightness of μ_J=26.5 arcsec^{-2} has been adopted (see Jones et al. 1991 for details and transformations). This places the survey on the same system as the Palomar Sky Survey. The magnitude zero-points and scale are set by comparing the measuring machine scans with reference to the Edinburgh-Durham Southern Galaxy Survey (Heydon-Dumbleton, Collins & MacGillivray 1989) using galaxies in the range $19 < b_J < 21$ and the CCD zero points used by Jones et al. (1991). These magnitudes were also correlated against those in the APM survey (Maddox et al. 1990a).

Star/galaxy separation of the tentative list of sources was performed using the COSMOS algorithm (MacGillivray, Beard & Dodd 1988). The COSMOS algorithm correctly identifies the classifications for 85 % of the faint sources, and undercounts the galaxies by no more than 5 %. Furthermore, the compact images which happen to be galaxies are neither bluer nor fainter than their extended counterparts. Therefore, the automated classification does not bias the target selection, unless there is a compact population at bright apparent magnitudes but not at faint ones. From the resulting list of galaxies, several target lists are generated, and each target is checked by eye, to verify astrometry and that the image of the plate is indeed a galaxy, not a plate defect.

The list of targets is passed along to an automated target selection program `CONFIGURE` (Sharples 1989). In addition to the target list, several astrometric fiducial stars and blank-sky positions must be provided. The sky-positions are verified as blank on either UK-Schmidt or AAT-Prime-Focus plates; the fiducial stars should be brighter than 15 magnitude to ensure that the field is easily acquired. `CONFIGURE` selects the optimum combination of targets to maximise the number of objects observed, without neglecting any of the mechanical constraints of the fibre positioner. The primary constraint is that a fibre head may not be placed within 33 arcsec of another fibre. Therefore, it is impossible to observe two objects closer that 33 arcsec at the same time. Cole et al. (1994b) found that the two-point correlation function of the `AUTOFIB` galaxies drops at scales less than 2 arcmin, pointing to an additional bias. Figure 2.1 schematically illustrates the `AUTOFIB` field. The galaxies in the centres of the various circles are selected for observation, while all other galaxies within the circles may not be observed or are less likely to be observed due to this mechanical constraint. Additionally, no object may be observed further than 19.15 arcmin from the centre of the field, nor within an annulus between 13.81 and 14.36 arcmin from the field centre. In Section 2.3 I discuss whether this may introduce a bias into the luminosity function determination.

During the observations, the exposures of the targets are alternated with sky observations (during which the telescope is moved 10 arcsec off-target) and arc wavelength calibrations. Bias and flat-field twilight sky

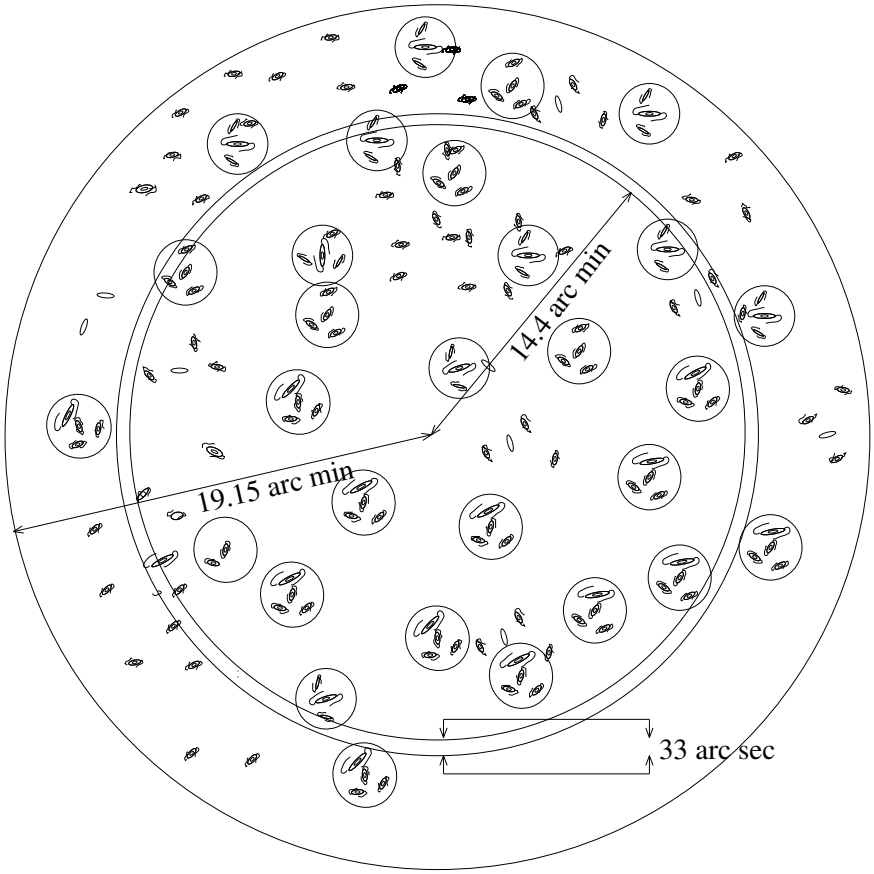

Figure 2.1: Schematic `AUTOFIB` Field. The large circle represents the extent of the `AUTOFIB` field, 38.30 arcmin. The smaller circles are centred upon a galaxy observed in the survey, and they circumscribe the exclusion region about each fibre. The construction of the positioner prevents any fibre button from being placed within 33 arcsec of another. Consequently, the two-point correlation function of galaxies within the survey becomes deficient at approximately 2 arcmin (Cole et al. 1994b). Additionally, the fibre plate has a step of 2 mm, 129 mm (14.4 arcmin) from the centre of the field to follow the curved focal plane. Fibres may not be placed within 5 mm (33 arcsec) inside of this step. This excluded region is represented by the narrow annulus.

Table 2.1: Characteristics of the combined survey

Survey	b_J	Area ($\Box^\circ$)	Fields	Galaxies	Complete
DARS	11.5–17.0	70.84	5	326	96%
Autofib bright	17.0–20.0	5.52	16	480	70%
Autofib faint	19.5–22.0	4.67	16	546	81%
BES	20.0–21.5	0.50	5	188	83%
LDSS-1	21.0–22.5	0.12	6	100	82%
LDSS-2	22.5–24.0	0.07	5	60	71%

frames are measured at the beginning and end of the night. Once the night's observations of the field are complete, the spectra are reduced in the standard fashion (described in Appendix B).

2.2 Survey strategy and coverage

The survey samples several independent pencil beams with a different sampling rates and magnitude limits in the range $17 < B < 21$, rather than a contiguous volume in a fixed apparent magnitude range. By sampling various areas of sky to various depths, the confusing effects of galaxy clustering may be reduced substantially. By selecting various magnitude limits within the larger range of $17 < B < 21$, all objects in a given field are measured to a uniform signal-to-noise level – optimising the exposure time for each field. Meanwhile, a larger range of magnitude is sampled when the various subsurveys are added coherently (Chapter 4), and the absolute magnitude *versus* redshift plane is filled in controlled manner as shown in Figure 2.2.

Table 2.1 summarises the salient characteristics of the combined survey. As well as the new data, we have included the brighter DARS survey (Peterson et al. 1986) and the fainter BES (Broadhurst, Ellis & Shanks 1988), LDSS-1 (Colless et al. 1990) and LDSS-2 surveys (Glazebrook et al. 1993). In total, the survey consists of 1733 redshifts in 53 fields. Within this survey, the new data reported here comprises 1026 redshifts in 32 fields taken primarily in 2 apparent magnitude ranges: $17 < b_J < 20.0$, and $19.5 < b_J < 22$.

Although there are a total of 53 fields, they fall into only ten independent regions. Most of the observations beyond $B \simeq 20$ were restricted to a few well-separated areas of sky with astrometric and photometric data from 4-m plates and CCD images. This information, as mentioned earlier, is needed for target selection and fibre placement. At the intermediate magnitudes, many of the fields are concentrated near the South-Galactic pole. This reflects the intensive study of the large-scale structure in this region by Broadhurst et al. (1990) and Szalay et al. (1991).

The LDSS-1 and LDSS-2 surveys (Colless et al. 1990, Glazebrook et al. 1993)

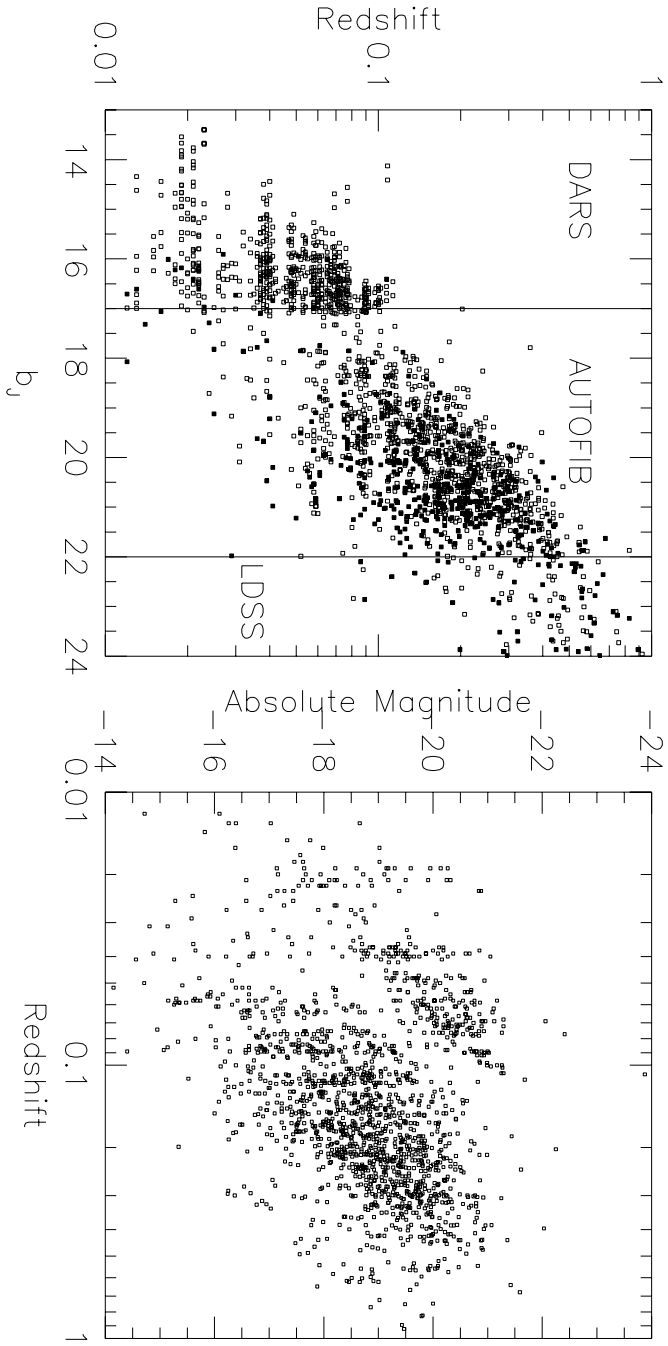

Figure 2.2: Apparent and Absolute Magnitudes versus Redshift. The upper pane shows the new `AUTOFIB` survey in relation to the bright DARS and faint LDSS surveys. The lower pane depicts the homogeneous sampling of the absolute magnitude versus redshift plane, helpful for the direct determination of the evolution of the luminosity function. The upper cloud is DARS. The large lower cloud is the new `AUTOFIB` survey combined with the BES survey. The LDSS surveys contribute a scattering of points at high redshift and faint absolute magnitude.

were selected from deep CCD data calibrated directly with reference to Jones et al and are automatically on this system. For the DARS data (Peterson et al. 1986), which was originally selected at a bright μ_J=24.0 arcsec^{-2} isophote, all of the apparent magnitudes and limits were shifted 0.28 magnitudes brighter. The spectra for a few of the subcatalogues were not available for the spectral classification described in Chapter 3. The DARS galaxies have morphological classifications which guided the assignment of k-corrections. The spectra for the entire `bes_mt` field could not be recovered and some spectra in the other catalogues were zeroed (apparently randomly) during reduction. These galaxies were given by default the k-corrections of Sbc galaxies. As it will be shown in Chapter 6, these uncertain galaxies make a minor contribution to the total luminosity function and therefore introduce little uncertainty in the results.

2.3 Selection biases

As mentioned earlier, the AUTOFIB instrument has a number of restrictions which could affect some of the survey characteristics. Figure 2.3 (or Figure 5 of Cole et al. (1994b)) shows the distribution of separations for pairs. The distribution drops dramatically at separations less than 2 arcmin (the fibre-fibre restriction). Also, it drops to a short dip between 30 and 34 arcmin due to the step in the fibre plate. Finally, there are no pairs separated by more than 38 arcmin as the edge of the field is reached.

The main effect of such selection constraints in our survey is likely to be in studies of the galaxy correlations rather than in estimating luminosity functions, unless there are luminosity-dependent correlations. Similar restrictions are also present in the LDSS-1 and LDSS-2 samples where multi-slit masks are optimally designed when sources are uniformly spaced. This restriction could bias the derived luminosity functions, if low luminosity satellite galaxies preferentially lie within 2 arcmin of brighter host galaxies. To check this, the relative magnitude distribution of pairs of which one was a galaxy selected for spectroscopic measurement and the other a source within a given separation was compared with the same distribution for all observed objects against all objects that could have been observed within the same field. No statistically significant difference was seen.

2.4 Incompleteness

Table 2.2 gives a more detailed summary of the survey including the statistics and incompleteness on a field by field basis. In general terms, incompleteness can arise in several ways and might, of course, seriously affect luminosity function estimation depending on whether it is systematic with *e.g.* redshift or spectral type. The most benign effect, which can be corrected, is incompleteness that arises purely from the increased difficulty of making redshift identifications because the spectra of the fainter galaxies

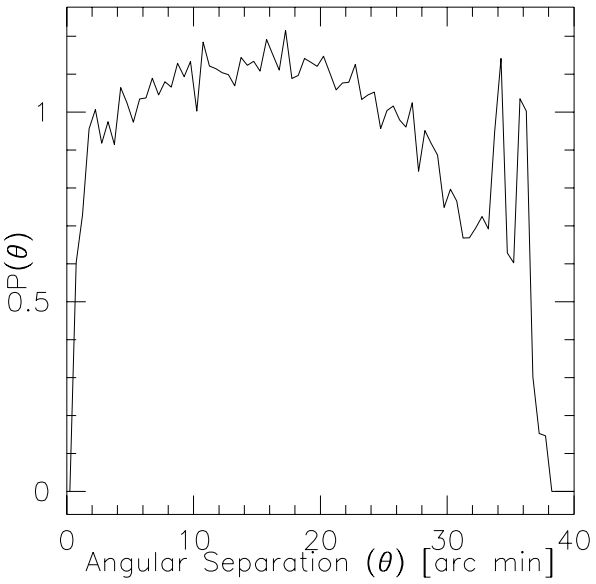

Figure 2.3: Normalised Ratio of Pair Counts. For each field, the distribution of separations for the observed galaxies is divided by the distribution of all galaxies in the sampling region. The mean of these ratios over all the fields is plotted. There is a deficit both at small and large separations.

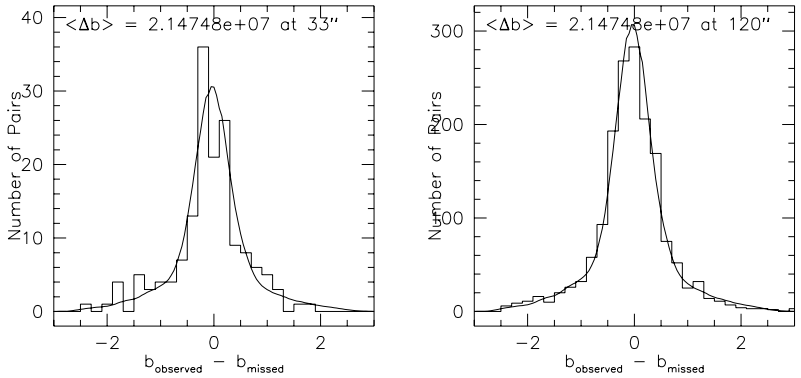

Figure 2.4: Distribution of Apparent Magnitude Differences for Pairs. The histogram in each pane shows the distributions of magnitude differences between objects included in the survey and nearby objects that were missed. The curves in each plot show the distribution for all pairs in a given field over the entire survey. Both the distribution of magnitudes differences for separations of 33″ and 120″ are equivalent to the total distribution with probabilities of 61 and 45 per cent respectively (using the Kolmogorov-Smirnov statistic).

Table 2.2: Field-by-field summary of the combined survey

Title	#	$B_{\rm min}$	$B_{\rm max}$	$N_{\rm target}$	$N_{\rm fibre}$	$N_{\rm z}$	$N_{\rm star}$	$N_{\rm dud}$	Area ($\Box'$)
10b	1	17.0	19.7	99	53	33	11	3	1260
10b2	2	17.0	20.0	83	53	34	8	0	1220
10f	3	20.3	20.8	144	54	36	10	1	1180
10m	4	19.7	20.3	99	54	42	8	1	1260
13b	5	17.0	19.7	77	50	39	3	3	1160
13b2	6	17.0	19.8	95	54	26	9	1	1260
13f	7	20.3	20.8	100	51	31	7	3	1210
13m	8	19.7	20.3	67	47	39	3	0	1220
197_b2	9	17.0	19.7	135	55	29	4	0	1260
197_b3	10	17.0	19.7	139	54	21	4	0	1170
22_b0	11	17.0	20.3	132	48	28	15	1	1260
22_f0	12	20.3	21.3	212	51	41	8	0	1820
22_f1	13	20.3	21.3	212	48	36	6	1	1820
22_f2	14	20.3	21.3	212	32	16	3	5	1820
03_f0	15	20.3	21.3	174	51	39	6	1	1340
03_f1	16	20.3	21.3	174	55	40	4	6	1340
411_b1	17	17.0	20.0	113	41	27	1	1	1200
411_f1	18	19.7	20.5	114	55	42	7	0	1260
412_b1	19	17.0	19.7	105	55	37	1	2	1260
412_b2	20	17.0	20.0	107	53	39	2	2	1260
412_f1	21	19.7	20.5	100	55	28	3	0	1260
474_b1	22	17.0	20.0	72	52	33	2	0	1260
474_b2	23	17.0	20.0	78	53	31	6	1	1260
475_f1	24	19.7	20.5	89	54	47	1	1	1260
475_f2	25	19.7	20.5	59	44	23	3	6	1260
529_b1	26	17.0	19.7	106	49	33	8	0	1260
529_b2	27	17.0	19.7	158	50	24	10	1	1260
529_b3	28	17.0	19.7	80	35	18	4	0	1260
bes_197	29	20.5	21.5	39	39	31	1	0	400
bes_419	30	20.0	21.0	114	38	32	2	0	502
bes_529	31	20.0	21.0	132	44	35	4	0	276
bes_mt	32	20.5	21.5	88	44	31	1	1	400
bes_sgp	33	20.5	21.5	71	71	59	1	0	217
ldss_002	34	20.9	22.7	150	49	21	7	14	120
ldss_102	35	21.0	22.5	89	26	17	4	2	65
ldss_104	36	21.0	22.5	91	36	20	10	5	65
ldss_132	37	21.0	22.4	84	29	18	7	1	65
ldss_134	38	21.0	22.4	84	23	14	5	1	65
ldss_135	39	21.0	22.4	79	21	10	6	0	65
13_b3	40	17.0	19.7	71	48	27	12	2	1260
13_f1	41	20.8	21.5	270	52	38	4	0	1235
13_xf	42	21.5	22.0	433	51	18	2	18	1235
mt_xf	43	21.5	22.0	329	50	32	4	3	1260

Table 2.3: Field-by-field summary of the combined survey, continued

Title	#	$B_{\rm min}$	$B_{\rm max}$	$N_{\rm target}$	$N_{\rm fibre}$	$N_{\rm z}$	$N_{\rm star}$	$N_{\rm dud}$	Area ($\Box'$)
l2_0331	44	22.5	24.0	504	30	19	2	0	67
l2_1021	45	22.5	24.0	110	20	14	1	0	45
l2_1022	46	22.5	24.0	110	19	14	1	0	45
l2_1321	47	22.5	23.3	77	13	9	0	0	51
l2_2231	48	22.5	23.0	49	6	4	0	0	46
dars_gsa	49	11.6	16.8	83	83	72	8	0	50904
dars_gsd	50	11.6	16.8	69	69	61	5	0	50220
dars_gsf	51	11.6	16.9	57	57	57	0	0	55188
dars_gna	52	11.6	17.1	78	78	69	2	0	50220
dars_gnb	53	11.6	16.9	70	76	69	7	0	48492
Totals				6666	2478	1703	253	87	299168

in each of the various magnitude ranges have inadequate signal relative to noise. *Provided* this magnitude-dependent incompleteness is independent of redshift or type, then it can be corrected by weighting each galaxy inversely with the survey success rate at that apparent magnitude. The completeness as a function of apparent magnitude for the categories defined in Table 2.1 is shown in Figure 2.5. All the surveys show some drop in completeness at the faint end of their magnitude range. The worst affected surveys are `AUTOFIB` bright and LDSS-2, while DARS is virtually complete. The relatively low completeness of the `AUTOFIB` bright survey arises from the agreed strategy of undertaking this portion of the survey whenever the observing conditions were too poor for fainter work. As a consequence, these spectra are often of poorer quality than those of fainter galaxies. We can estimate the effect of the observed incompleteness on our analyses by comparing the distributions of the $V/V_{\rm max}$ statistic for the various data subsets with and without the above correction for magnitude-dependent incompleteness. If the observed distribution of galaxies is unclustered and does not evolve then $V/V_{\rm max}$ should be uniformly distributed between 0 and 1. Actual clustering and evolution will cause departures from this expectation, but so will magnitude-dependent incompleteness even in their absence.

Fortunately the form of departure from uniformity of the $V/V_{\rm max}$ distribution is different for each of these cases: magnitude-dependent incompleteness will cause the sample to be deficient in the higher redshift galaxies of any given luminosity, and will therefore bias the $V/V_{\rm max}$ distribution to smaller values; clustering will cause peaks and troughs in the distribution at the values of $V/V_{\rm max}$ corresponding roughly to an L^* galaxy at the redshift of the relevant structure; evolution (at least if it takes the form of an increase in the number of galaxies of any given luminosity at higher red-

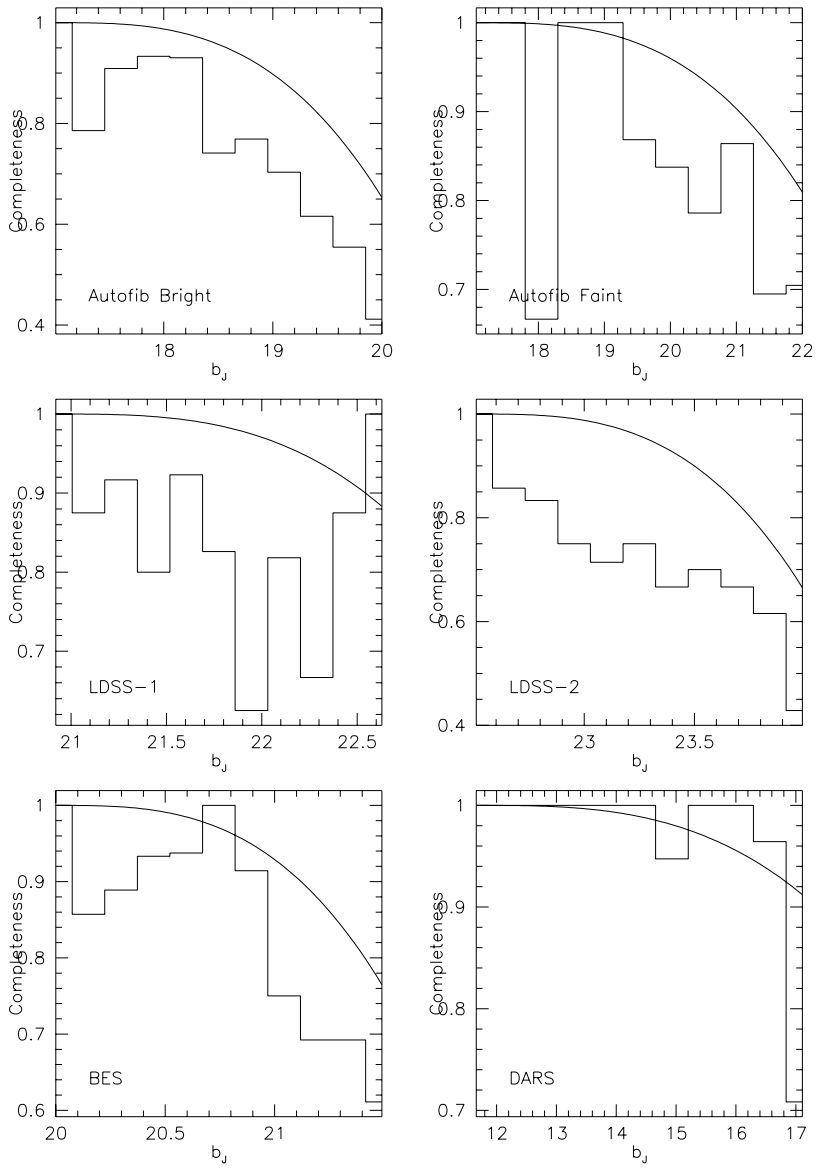

Figure 2.5: Completeness rate for the various subsurveys. The histogram in each pane follows the completeness as a function of apparent magnitude. To illustrate the completeness correction technique, the completeness is parameterised by $f(m-m_0) = 1-a\left[(m-m_0)/(m_1-m_0)\right]^3$ where m_0 and m_1 are the bright and faint limits of the survey. The best-fitting function for each survey is superposed.

shifts) will bias the distribution to larger values. Note that an important feature of our strategy of breaking our samples into several narrow apparent magnitude slices, is that we expect little relative evolution over any one subsample. Only by combining all the surveys and spanning a large range in apparent magnitude and redshift can we expect to see evidence for evolution. Thus the absence of any upward trend should *not* be taken as evidence against evolution.

To correct for magnitude-dependent completeness, the galaxies at the faint end of the survey receive higher weights than the brighter galaxies in inverse proportional to the completeness rate at the particular apparent magnitude within the field. For each field, the incompleteness rate is assumed to increase cubically from 0 at the bright limit of the field. It is assumed that no galaxies with W_λ[OII] greater than 20 Å are missed; we will see later that galaxies with high W_λ[OII] are a special subsample with the catalogue. In practice, the weighting function w also depends on the mean completeness rate C for the field:

$$w(m) = C - B\left(\frac{m - m_0}{m_1 - m_0}\right)^3 \tag{2.1}$$

where m_0 and m_1 are the bright and faint limits of the field, and B is selected such that the sum of the weights over a field is equal to the number of redshifts measured for that field. Thus, the sampling area for each field remains constant.

Figures 2.6 and 2.7 show the $V/V_{\rm max}$ distributions (with and without correction for incompleteness) for each of the surveys presented in Figure 2.2. In all cases the distribution is either flat or tends downward at large values, as expected for magnitude-dependent incompleteness. The completeness correction makes the distributions closer to uniform for all of the surveys, but overcorrects in the case of the LDSS-2 survey, which has a mean $V/V_{\rm max}$ of 0.49 before correction. Although dramatic on the $V/V_{\rm max}$ distributions, this completeness correction has little effect on the calculated luminosity functions as discussed in Section 5.2.2.

Unlike magnitude-dependent effects, incompleteness that is a function of galaxy redshift or spectral type can neither be directly quantified nor corrected. Furthermore, both sorts of incompleteness may be confused with the signal/noise-dependent losses, since both type and redshift are expected to correlate with apparent magnitude. However we can make tests to establish whether either of these problems is significant.

For type-dependent incompleteness we can again use the $V/V_{\rm max}$ statistic. In Chapter 3 I define a procedure to allocate a spectral type to each galaxy by correlating its spectrum with local templates. Anticipating this classification scheme, Figures 2.8 and 2.9 show $V/V_{\rm max}$ distributions for each spectral type, again with and without the correction for magnitude-dependent incompleteness. The greatest departures from uniformity are for galaxies of intermediate type, and their $V/V_{\rm max}$ distributions are only marginally improved by the correction for magnitude-dependent incom-

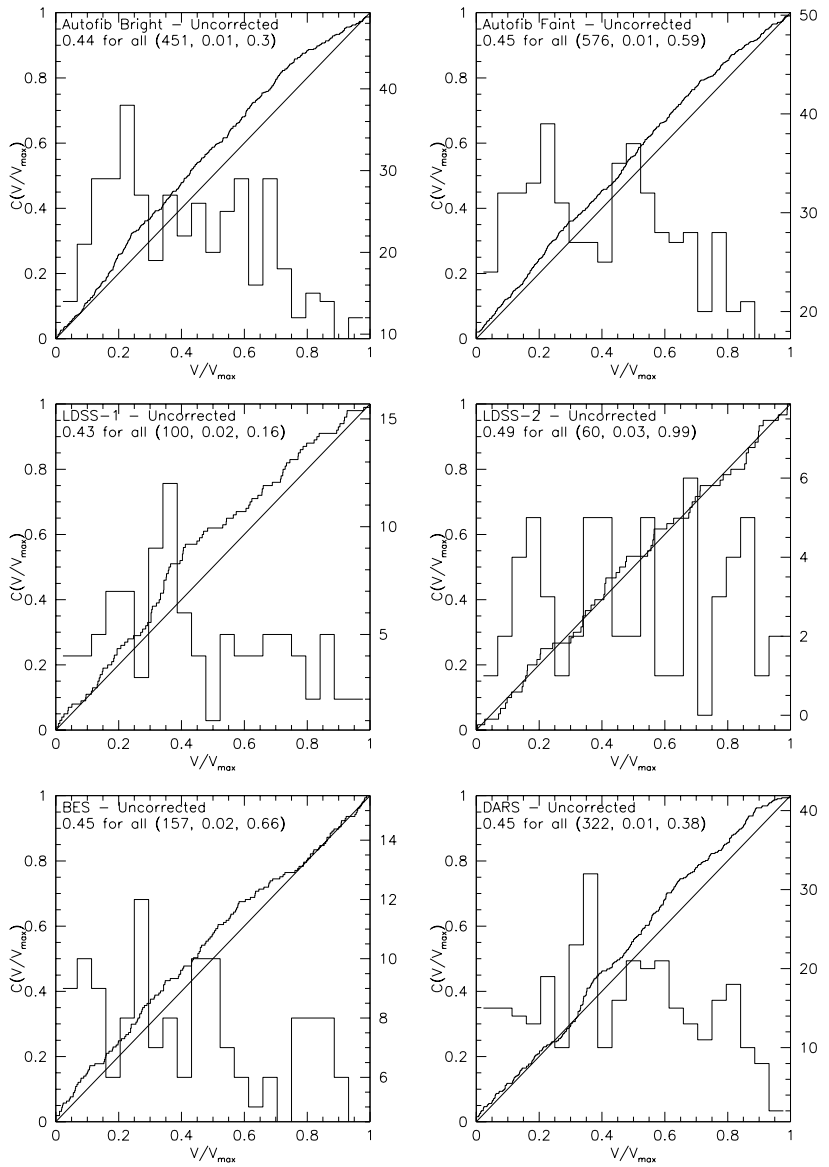

Figure 2.6: $V/V_{\rm max}$ distribution for the various surveys (without completeness correction). The histogram shows the distribution of $V/V_{\rm max}$ values over each subsample. The stepped curve traces the cumulative distribution of $V/V_{\rm max}$ and the diagonal is the cumulative distribution for a uniform sample. The legend gives the mean value of $V/V_{\rm max}$ over the subsample, the number of galaxies in the subsample, the excepted standard deviation for the mean value of $V/V_{\rm max}$ for a subsample of the given size. The final value is the Kolmogorov-Smirnov probability that the distribution of the $V/V_{\rm max}$ values could be drawn from a uniform disitrubtion.

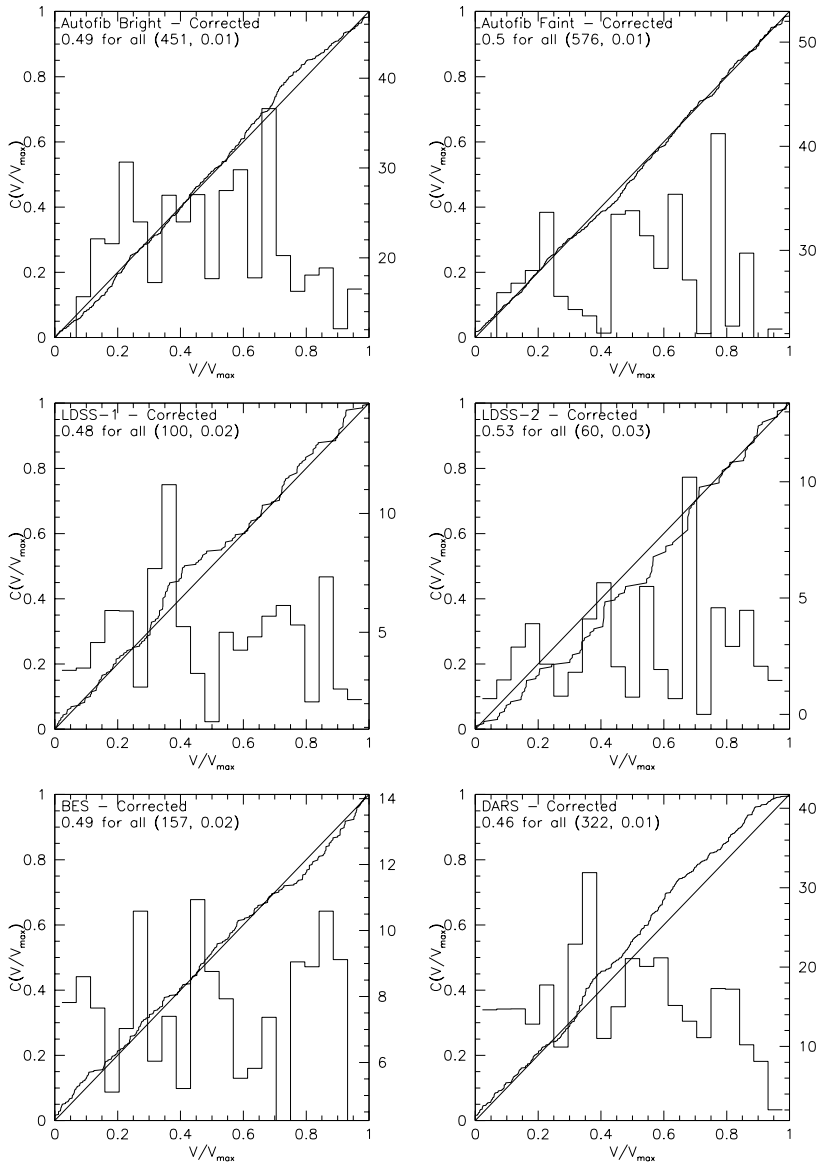

Figure 2.7: $V/V_{\max}$ distributions for the various surveys (with completeness correction). The curves are that same as in the preceding figure, and the legend shows the same information with the exception of the K-S probability.

pleteness. This is not unexpected, since early-type galaxies have strong absorption features and late types generally have strong emission lines, both of which render redshift determination more straightforward. In contrast, intermediate types have weaker absorption features and often no emission lines and so are generally harder to identify.

For redshift-dependent incompleteness the $V/V_{\rm max}$ statistic is worthless because V is a function of z. However we can check for redshift- dependent incompleteness by making use of the important fact that our combined sample is made up of sub-surveys with overlapping apparent magnitude ranges. By comparing the redshift distribution of the bright (high-completeness) end of a fainter survey with the faint (low-completeness) end of a brighter survey, within the limitations of clustering, we can check whether incompleteness distorts the redshift distributions.

Figure 2.10 shows the results of such comparisons—these can be summarised as follows:

1. DARS b_J=16.5–17 and Autofib bright b_J=17–17.5 agree well;
2. Autofib bright and Autofib faint disagree badly for b_J=19.5- -20, with Autofib faint having generally higher redshifts;
3. Autofib faint and BES, for b_J=21–21.5, agree well but LDSS-1 has generally higher redshifts;
4. LDSS-1 b_J=22–22.5 agrees well with LDSS-2 b_J=22.5–23.

Of course we cannot check in this way for redshift-dependent incompleteness in the LDSS-2 survey since we have no fainter survey with which to compare it. Glazebrook, Peacock & Collins (1994) describe the statistics of this deepest data set in some detail.

To summarise, there is significant incompleteness in all the surveys included in this work. However this incompleteness would appear to be dominated by the difficulty of identifying the fainter galaxies in each sample due to poorer S/N in their spectra. We can largely remove this effect by an appropriate magnitude-dependent completeness correction. Although residual systematic effects are doubtless present, particularly for intermediate spectral classes, even the dominant magnitude- dependent correction hardly changes the derived luminosity function results in Section 5.2.2.

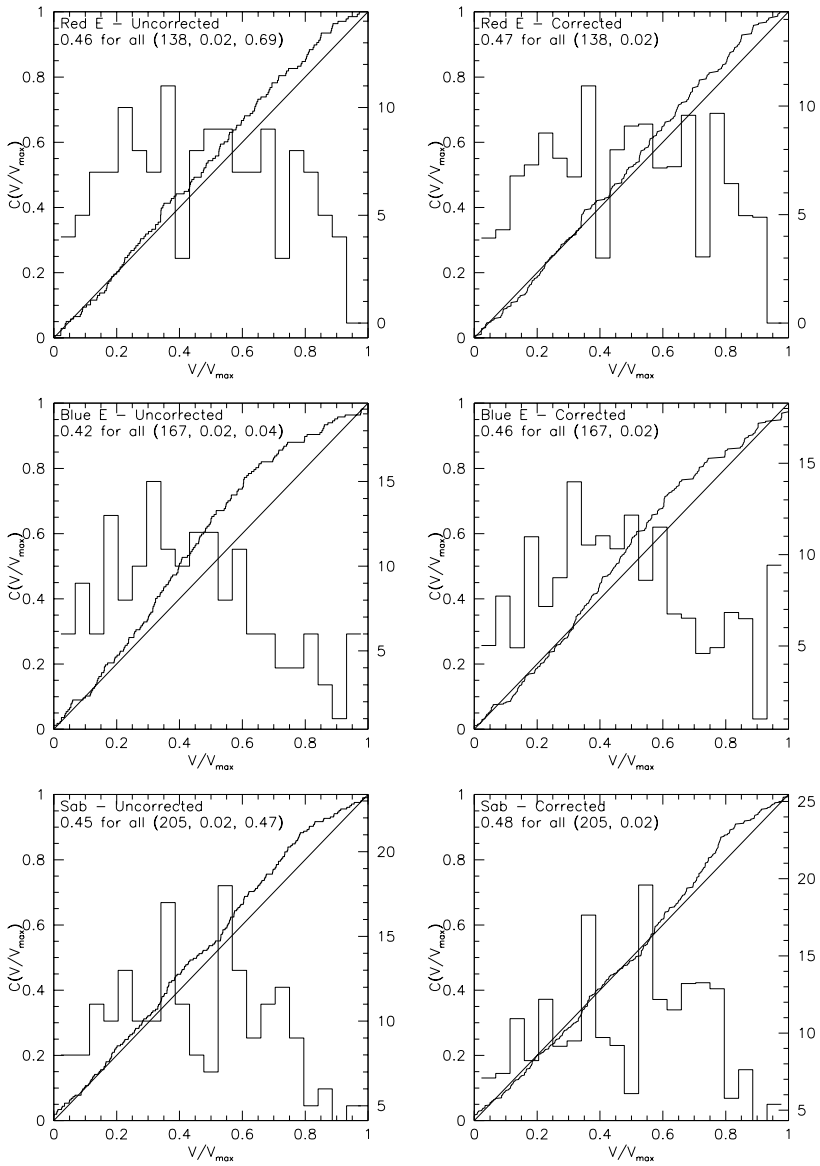

Figure 2.8: $V/V_{\max}$ distributions for the early spectral types

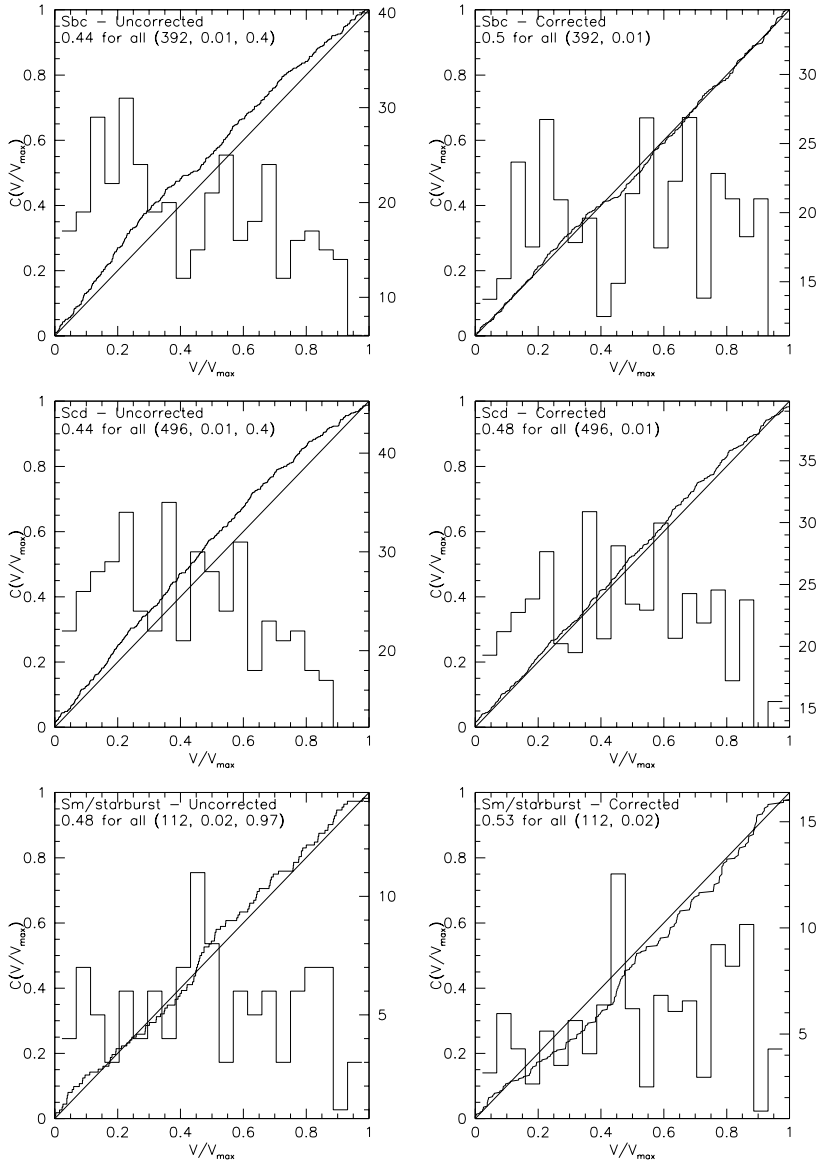

Figure 2.9: $V/V_{\max}$ distributions for the late spectral types

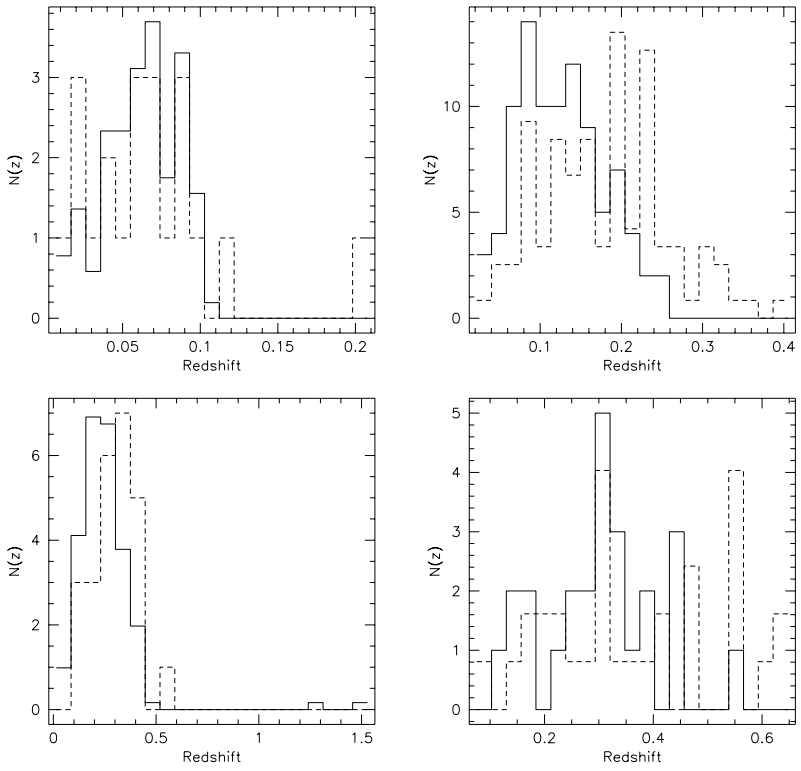

Figure 2.10: Redshift distributions in the overlap ranges. The panes depict (starting from top- left and proceeding clockwise) the overlaps of DARS with `AUTOFIB` bright, `AUTOFIB` bright with `AUTOFIB` faint, LDSS-1 with LDSS-2 and `AUTOFIB` faint and BES with LDSS-1.

Chapter 3

K-corrections

SUMMARY

A prerequiste in the determination of the luminosity function from a redshift survey is calculating the rest-frame magnitudes of the galaxies and the volume within which they could have been observed. This chapter discusses the obstacles to this process and introduces a cross-correlation technique to classify galaxy spectra, determining the k-corrections and accessible volumes. Finally, we present some tests of the algorithm and some sample spectra from the survey.

3.1 Introduction

An accurate estimate of the luminosity of the galaxies in a survey is the first and probably the most important step in determining the galaxy luminosity function. Once we have chosen a cosmological framework (*e.g.* $q_0 = 0.5$ and $H_0 = 100$ km/s/Mpc), the distance modulus for each galaxy is determined. However, in samples at moderate redshift, the k-correction is just as significant to the determination of the rest-frame luminosity of the object. Previous researchers have used a variety of approaches to this problem.

The most common method is to assume that the galaxies have k-corrections that increase linearly with redshift and each morphological type is assigned a different slope (*e.g.* Efstathiou, Ellis & Peterson 1988, Loveday et al. 1992). A next popular method is to assume some form for the galaxy spectral energy distributions (SED), and then use colour to determine the k-correction (*e.g.* Saunders et al. (1990) assume that the galaxies in their sample follow a single-temperature Planck function in the infrared). Eales (1993) takes a unique method which is to calculate the luminosities in a passband corresponding to the b_J band shifted blueward by the mean redshift of the sample. In this way, errors in the k-correction may be minimised, as the k-correction at the mean redshift of the sample is by definition zero. He then uses the morphological classifications (if they

exist) or assumes a classification for unclassified galaxies, and calculates the k-corrections from canonical SEDs.

In the survey, none of the above-mentioned methods can be directly applied with much success. Firstly, only a subset of the catalogue (the DARS galaxies) has been classified morphologically, and only a second subset has colours in $b_j - r_f$ (the LDSS and LDSS-2 galaxies). Applying a mean k-correction or defining a new passband probably would not be fruitful either as the galaxies range from $z \sim 0$ to nearly 1.0, so an extremely blue or red galaxy at high redshift would have a luminosity incorrect by over a magnitude and using a mean-redshift could lead to nearly as large errors. Furthermore, the volume weighting would be even more uncertain. Also, we would hope to derive luminosity functions as a function of class.

3.2 Testing for Spectral Evolution

In principle, a k-corrections could be "read" directly from each spectrum. In practice, the fluxings are uncertain if they exist at all, so this strategy would also fail. However, it points in the right direction. In absence of good fluxing, the spectra lack reliable information on large wavelength scales (*i.e.* colours), however on smaller scales (*i.e.* lines) the spectra are more reliable. After all, the redshifts may be reliably determined from the spectra. In this vein we developed two complementary techniques for classifying each spectra in the survey.

The first method follows in the spirit of Figure 6 in Broadhurst, Ellis & Shanks (1988, hereafter BES) which depicts the mean spectra over subsets of the BES survey. It shows that the galaxies in the survey are "normal" in the sense that one can find local counterparts to the distant galaxies in the survey. We took this technique further. First, we assume that the reddest galaxy observed within each subcatalogue is an elliptical. We also presume that the spectral response of the spectroscopic coupler does not change (except in normalisation) from galaxy to galaxy. Thus the galaxies in each subcatalogue may be crudely fluxed, using the supposed elliptical as a standard star.

After fluxing, each galaxy is classified according to the equivalent width of its [OII] line (if it exists) and the strength of the 4,000 $\mathring{A}$ break,

$$B_{4{,}000\mathring{A}} = \int_{4050\mathring{A}}^{4250\mathring{A}} f(\lambda)\lambda^2 d\lambda \Bigg/ \int_{3750\mathring{A}}^{3950\mathring{A}} f(\lambda)\lambda^2 d\lambda \qquad (3.1)$$

where the integrals are performed in the rest frame of the galaxy. We simply assign each galaxy one of four classes based on whether $B_{4{,}000\mathring{A}}$ is greater or less than 1.7 and whether the [OII] line was observed in the galaxy's spectra. As in BES, we then coadd the spectra in the rest-frame by selecting normalisation that minimise the mean scatter in flux over the spectra. We derived the median spectrum for each type and derived k-

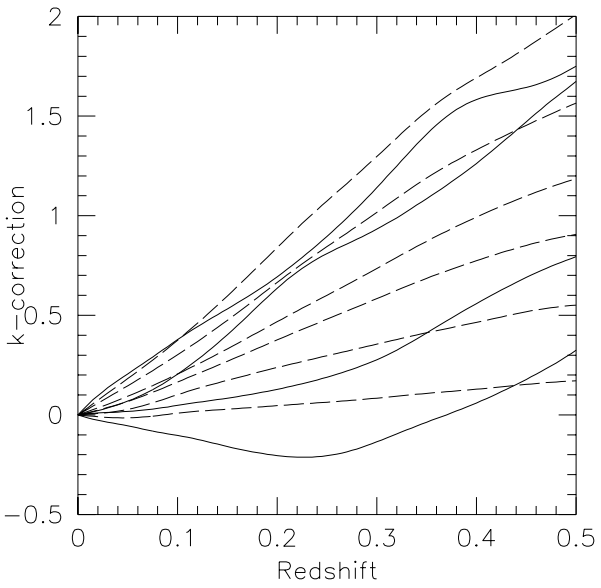

Figure 3.1: Pence and coadded k-corrections. The k-corrections for the various Pence classes (and the star-bursting galaxy NGC 4450) are traced with dashed lines, and those of the coadded spectra as solid lines. The k-corrections of the galaxies within the catalogue lie with the range of the Pence classes (with NGC 4449) included). Therefore, the Pence-and-starburst set is sufficient to calculate the k-corrections for the survey galaxies.

corrections from these coadded spectra as shown in Figure 3.1 along with the k-correction for various galaxy types given by Pence (1976).

Figure 3.1 shows that the coadded spectra have similar k-corrections to those of local galaxies and that the range in local type spans that of the survey (at least in the crude sense of the k-correction). Unfortunately, not every spectrum in the catalogue has the dynamic range for both the [OII] line and the break to be measured. Therefore, we cannot measure the k-corrections from these coadded spectra, and we must develop a method that can be consistently applied over the entire survey.

3.3 The Cross-Correlation Method

We decide to classify each galaxy not by colour as previous analyses, but by the spectral lines as in the four-class scheme described earlier. Each galaxy spectrum covers a different range in wavelength in the galaxy's rest frame, and the overlap in rest-frame wavelength is actually only few hundred $\AA$ from approximately 3,600 $\AA$ to 3,900 $\AA$. Therefore, the technique must not

rely on specific spectral information rather it must base its classification on varied, available data. We chose to cross-correlate the survey spectra against those of the Kennicutt (1992a, 1992b) spectral library. I thank the Astrophysics Data Center for providing the Kennicutt atlas. This library is highly appropriate for this task. It is important to use spectra integrated over a large portion of the galaxy rather than just spectra of the central region, as AUTOFIB and LDSS effectively admit light from the entire galaxy, and to match spectral resolution.

Initially, both the Kennicutt template spectra and the survey spectra are processed similarly. First, they are smoothed on a scale of 100 $\mathring{A}$ in the observer's frame and these smoothed versions are subtracted from the unsmoothed spectra, yielding continuum-subtracted spectra. Each of these spectra is rebinned to 2 $\mathring{A}$ per bin. Finally, each survey spectrum is cross-correlated against each Kennicutt spectrum. The survey spectrum is assigned the type of the template with which it most strongly cross-correlates. Thus, we have a two step mapping from survey spectrum to template and then using the published types of the Kennicutt spectra to a morphological classification. Using this classification and the King & Ellis (1985) k-corrections, we derive a k-correction for each galaxy.

3.4 Tests

3.4.1 Self-Testing

To verify this algorithm, we performed two series of tests. The first was to simulate the routine using fluxed spectra. First we selected one of the Kennicutt spectra randomly, normalised it to have a particular mean number of counts per bin and added a constant number of sky counts per bin. Next using the total number of counts in each bin, a random Gaussian deviate was chosen, and the noise was added to the value already in the bin. Finally the sky was subtracted again. Thus by repeating this process one hundred times, we created an ensemble of test spectra with a known signal-to-noise ratio. Each of these spectra was processed similarly to those in the survey, and the success rate for the "fluxed" spectra is calculated. As Table 3.1 shows, the routine's success rate is quite high, so we tested the algorithm again. This time we tried to simulate the unfluxed spectra in the catalogue and also the varying redshifts of the galaxies in the catalogue. The "test" spectra were generated in the same way as in the "fluxed" case except that each spectrum was multiplied by a response function. Our simulated spectrograph has zero sensitivity outside the band 3,600 to 7,200 $\mathring{A}$, and the response increases quadratically by a factor of two from the edge to the centre of this range. Before calculating the "observed" spectrum, the response function is blueshifted by a factor of 1 to 1.6; this is equivalent to redshifting the galaxy spectrum, multiplying it by the response function, and then blueshifting the resulting spectrum back into the galaxy's restframe.

Figure 3.2 depicts the results of the "unfluxed" simulations in more

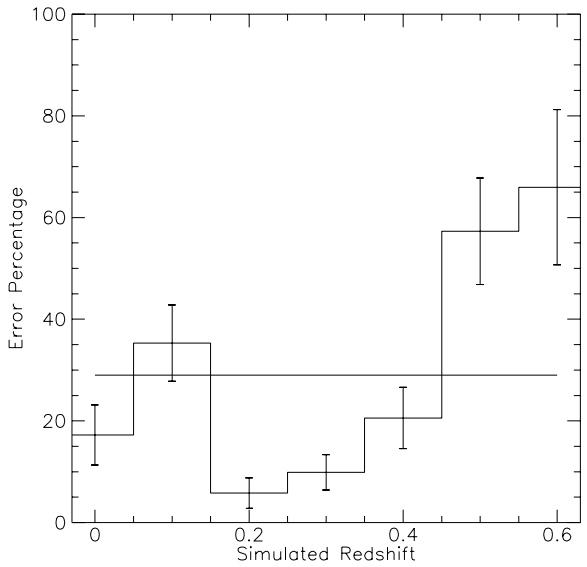

Figure 3.3: Classification error rate as a function of redshift. The error increases markedly with redshift, as the identifying features of the spectra are lost off the red end of our simulated spectrograph. The horizontal line is the mean error rate of the various tests: 29 %.

detail. The most striking feature of the distribution is the diagonal ridge line which traces the correct classifications. Classes #2 (red ellipticals) and #4 (early spirals) are sometimes confused with class #3 (blue ellipticals). Furthermore, class #5 (intermediate spirals) are sometimes classified as class #3 (blue ellipticals) and as class #7 (star-bursting spirals).

Each k-correction class corresponds to several Kennicutt spectra, each with varying strengths of spectral features; consequently, with the addition of noise, the spectra can be confused across a k-correction class. In the case of the intermediate spirals which have only weak features of both earlier and later classes, the classification may be off by more. Figure 3.5 shows an example intermediate-type spectrum from the bright section of the `AUTOFIB` survey – the features of both early and late-type galaxies appear weakly in this spectrum.

Additionally, we performed this blueshifting to simulate how observing galaxies at various redshifts through a fixed wavelength range would affect the success rate of the algorithm. Table 3.1 shows the routine performed well again. Furthermore, Figure 3.3 shows that the error rate increases with the simulated redshift of the galaxy. We observe only a portion of a galaxy's spectrum through the spectrograph, and this portion has a smaller range in wavelength in common with the Kennicutt templates as the redshift of the galaxy increases. Figure 3.7 depicts two galaxies from the faint sec-

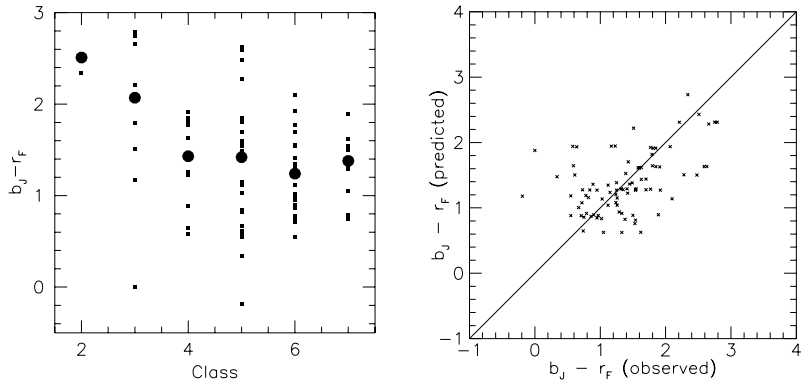

Figure 3.4: Results for LDSS-1 Spectra. The left panel depicts the spread in observed $b_J - r_F$ colour for the galaxies in the LDSS-1 (Colless et al. 1990) survey against the k-correction classification. Each small square represents a single galaxy. The large circles show the trend in the median colour for each class. The right panel shows colour predicted from the k-correction classification at the redshift of each galaxy against the observed colour of the galaxy. Each square represents a galaxy in the survey.

tion of the AUTOFIB survey, and their best- matching Kennicutt templates. The little overlap extends from just blueward of [OII] to slightly redward of the G-Band; little spectral information remains upon which to base a classification.

3.5 Results with observed spectra

3.5.1 Testing against LDSS

A subset of the catalogue have $b_J - r_F$ colours, which provide an independent test of the classification algorithm. The LDSS spectra (Colless et al. 1990) range from $B = 21$ to $B = 22.5$ and are amongst the faintest galaxies in the survey. They provide a stringent challenge for the method. We use a two-stage test: the correlation of observed colour against k-correction class and correlation of "predicted' colour against observed colour. We calculate the $b_J - r_F$ in the observer's frame by using the classification of each LDSS spectra from the algorithm and determining the colour of the appropriate SED at the redshift of the galaxy. Figure 3.4 shows the results of both these tests.

The trend of colour versus class exhibits a large spread in observed colour due to observing errors, colour corrections and partly misclassifications. The second pane shows the predicted colour against the observed colour, removing the effect of redshift upon colour. Although the trend is remarkably good; there is again a significant spread in the colour pre-

dictions. The mean-absolute error is approximately 0.4 mag. Budgeting the errors for this multistage process is difficult, as they arise from several places. The b_J and r_F magnitudes each have on average as uncertainty of 0.1 mag leading to a total error in the colour of 0.15 mag. Furthermore, a few galaxies have colours bluer than any of the k-correction SED; this is worrisome, but the effect is small and limited to only a few galaxies. Finally, the spread may be due to misclassifications. Although it depends on k-correction class, the error in the k-correction is approximately 1.5 times greater than the colour error, so if all the spread in this plot were due to misclassifications, the k-corrections for these galaxies would have an RMS error of 0.6 mag.

3.5.2 Sample spectra

The cross-correlation method uses features noticeable to the trained observer where available. Figures 3.5 through 3.7 illustrate the method in action. The lower curve is the observed spectrum in the composite survey, the middle curve is the Kennicutt spectrum selected by the algorithm as the best match for the lower observed spectrum. The continuum has been subtracted from both spectra as described earlier, and the observed spectrum has been smoothed further to accentuate the features. The upper curve is the product of the two spectra smoothed over 20 bins to show which features contribute most strongly to the total cross- correlation. As mentioned earlier, intermediate-type galaxies are the most challenging as seen in Figure 3.5; they have few strong features even these are ambiguous.

Looking at the bright end of the survey in Figure 3.6, one sees strong correspondence between the observed spectra and the best-match Kennicutt spectra. Again at the faint end (Figure 3.7, the algorithm performs well, however there is little overlap between the observed and template spectra for these moderate redshift galaxies. This is a stumbling block for the cross-correlation method. Presently there is no spectroscopic atlas like that of Kennicutt's to cover ultraviolet wavelengths; furthermore, there are few strong features blueward of the [OII] emission line that would enter the spectra range of the survey before a redshift of ~ 2. The simpler option is to give up spectral resolution in favour of wider spectral range, to find more familiar features in higher redshift galaxies. The only tradeoffs would be slightly less accurate redshifts and possibly a slightly higher incompleteness rate.

3.5.3 Testing with HST

Several galaxies from the LDSS-2 catalogue and one from the LDSS-1 catalogue were observed with the Hubble Space Telescope. The images were of sufficient quality to classify these galaxies morphologically, providing another independent test of the cross-correlation method.

Table 3.2 shows the results of the test. The cross- correlation method agrees with the morphological classification within one and a half classes

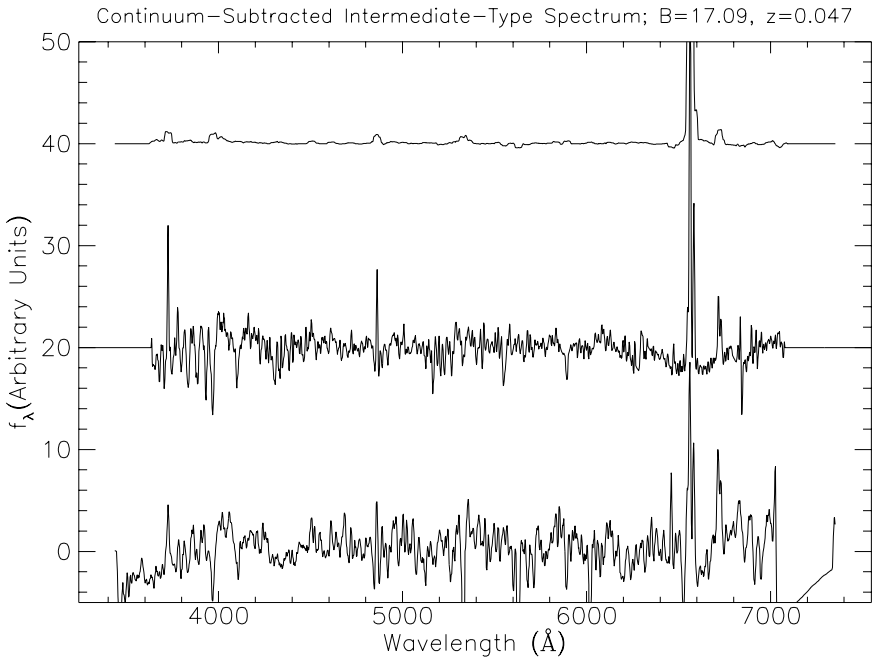

Figure 3.5: Intermediate sample spectrum from the bright section of the AUTOFIB survey ($B = 17.09, z = 0.047$). The spectrum has weak features of both early and late spiral galaxies. The lower curve traces the best-fitting spectrum from the Kennicutt atlas. The middle curve is the observed spectrum and the upper is the cross-correlation between the two curves.

Table 3.2: HST Morphologies of LDSS-1+2 Spectra. A total of seven galaxies were observed both by HST and the LDSS-1+2 spectrographs. Here is a comparison of the spectral and morphological classification of these galaxies.

ID	B	z	B-R	Spect. Class	Morph. Class
10.02.11	21.0	0.277	1.24	Scd	Sc?
10.21.15	22.7	0.177	1.10	Sbc	Sbc
10.21.17	23.7	0.492	2.92	Sm/starburst	Sdm
10.22.13	23.1	0.384	2.22	Sab	E
10.22.15	22.8	0.476	2.34	Scd	Sbc
10.22.16	23.7	0.436	2.22	Sbc	S0/a
10.22.19	23.1	0.724	1.08	Sm/starburst	Spec?

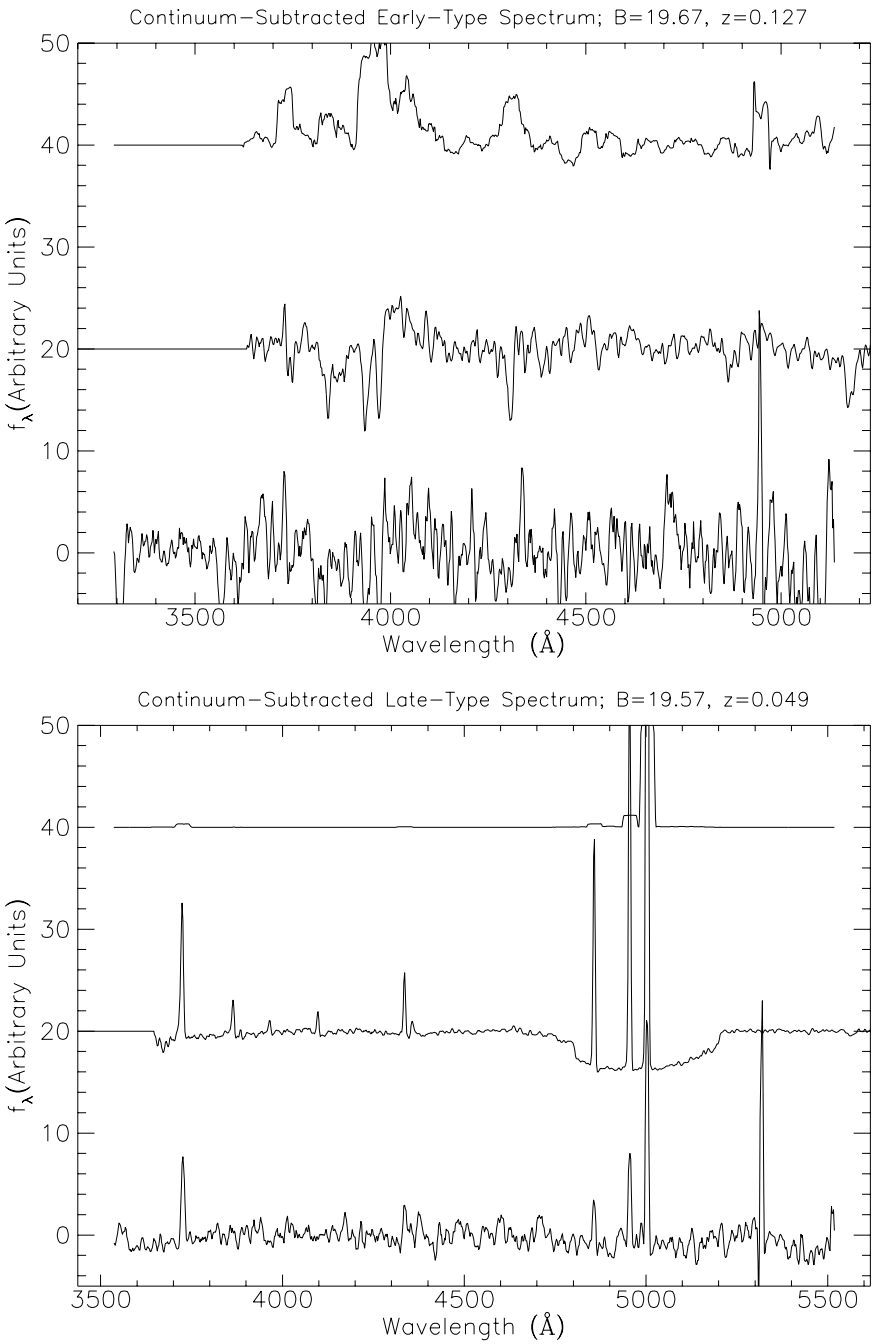

Figure 3.6: Bright sample spectra. These are two spectra selected from the bright section of the AUTOFIB survey. The upper panel is an early-type galaxy ($B = 19.67, z = 0.127$). The lower panel is a late-type galaxy ($B = 19.57, z = 0.049$). The three curves trace the same quantities as in the previous figure.

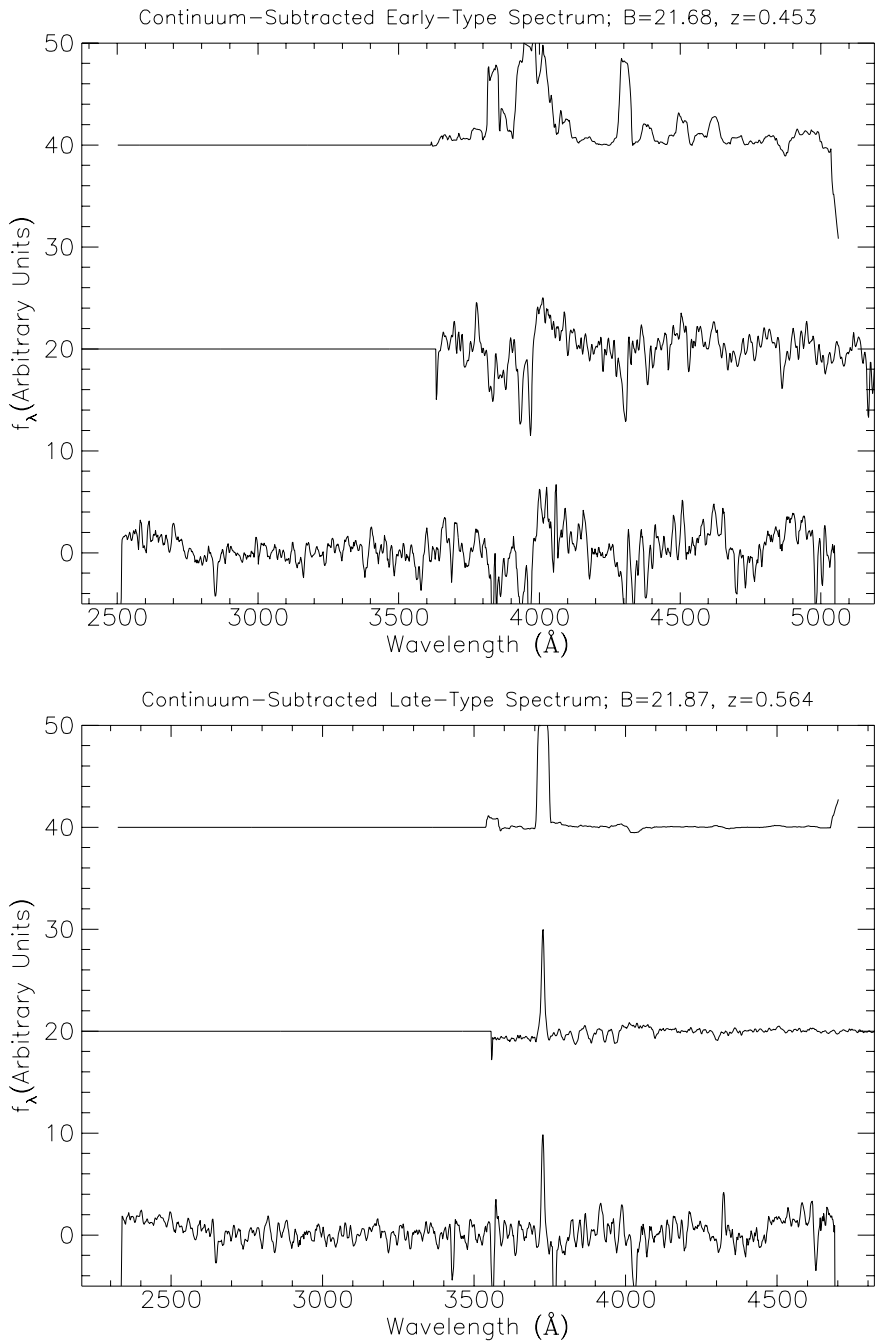

Figure 3.7: Faint sample spectra. These are two spectra selected from the faint section of the AUTOFIB survey. The upper panel is an early-type galaxy ($B = 21.87, z = 0.564$). The lower panel is a late-type galaxy ($B = 21.68, z = 0.453$). The three curves trace the same quantities as in the previous figures.

for all seven galaxies, within one class for six of seven; and for four of seven the two methods agree within the resolution of the cross-correlation method. This agreement is very encouraging and almost startling, as these galaxies are amongst the faintest to be observed spectroscopically and to be imaged with sufficient resolution for a morphological classification. The most significant discrepancy is 10.22.13, a galaxy with elliptical morphology and H, K, Hβ, and the G-band detected in the spectrum. Its k-correction at its redshift would be in error by about 0.5 magnitudes, and its value of $V_{\rm max}$ would be overestimated.

3.6 Conclusion

From simulations and testing with observed spectra, we find that the cross-correlation method performs well, assigning the correct k- correction class 80% of the time and to within one type more than 90% of the time. Furthermore, the types compare well with the observed colours of the LDSS-1 galaxies at the faint end of the composite survey. A final strong test compares the HST morphologies of a sample of LDSS-1 and LDSS-2 galaxies with the cross-correlation classification. The two methods agree excellently.

Chapter 4

Luminosity Function Estimators

SUMMARY

Over the years, the number of luminosity function estimators has multiplied nearly as quickly as the number of redshifts surveyed. Here, I describe several recent methods and introduce two more: modifications to the STY (Sandage, Tammann & Yahil 1979) and to the step-wise-maximum-likelihood method (Efstathiou, Ellis & Peterson 1988). This new method is compared with the most recent incantation of the traditional $1/V_{\rm max}$ *method (e.g. Eales 1993).*

4.1 Direct estimators

The canonical direct estimator of the luminosity function is the $1/V_{\rm max}$ method introduced for the study of quasar evolution (Schmidt 1968). The first step in this method is calculating the total volume within which the object could have been observed. In a single magnitude-limited survey:

$$V_{\rm max} = \frac{c}{H_0} A \int_{z_{\rm min}}^{z_{\rm max}} \frac{D_L^2 dz}{(1+z)^3 \sqrt{1+\Omega_0 z}} \tag{4.1}$$

where $z_{\rm min}$ and $z_{\rm max}$ are the minimum and maximum redshifts from which the object could have been observed in the survey considering the distance modulus and k-correction. A is the area of sky surveyed in sterradians. In its simplest form, the luminosity function is obtained by collecting the sources in bins of constant magnitude and summing the $(V_{\rm max})^{-1}$ values in each group.

Avni & Bahcall (1980) describe how to combine more than one sample coherently in a $V/V_{\rm max}$ analysis. The new variable, denoted V_a is simply the sum of $V_{\rm max}$ over all the surveys in which the object could have

been observed. Finally, Eales (1993) describes how to use this variable to construct the luminosity function as a function of redshift. The analysis proceeds according to Schmidt's method within the exception that one bins in redshift as well. Here $z_{\rm min}$ is the minimum redshift at which the object could have been observed in the magnitude-limited sample and be in the redshift range of interest. The result is similar for the maximum redshift.

Unfortunately, all these direct methods are sensitive to the presence of clusters in the field. Clustering tends to cause these direct estimators to overestimate the faint-end slope of the luminosity function.

4.2 Clustering-insensitive methods

Several authors have introduced estimators which although slightly biased reduce the effects of clustering on the resulting luminosity functions. All these techniques attempt to maximum the likelihood of observing the particular set of galaxies by varying parameters of the luminosity and completeness functions. How they differ is in the assumed probability of observing a particular galaxy.

Using Poisson statistics is one obvious choice. This was first applied to the analysis of quasar samples (Marshall et al. 1983) and later recast to analyse the CfA redshift survey (Chołoriewski 1986). The probability of observing k galaxies in the interval $dMdz$ in an area of sky $d\Omega$ is

$$P_k = e^{-\lambda}\frac{\lambda^k}{k!} \tag{4.2}$$

where

$$\lambda = \frac{1}{n}\phi(M)\rho(z)dMdzd\Omega \tag{4.3}$$

and n is the average number density of the survey. By binning $\phi(M)$ and $\rho(z)$, these two functions may be estimated without assuming particular forms. The only task that remains is to maximise the total probability of the survey (the product of all the P_k in each of the bins while varying the values of $\phi(M)$ and $\rho(z)$.

The C-method (Lynden-Bell 1971) takes a different approach. Looking at the plane of redshift *versus* absolute magnitude, this method uses the fact that the ratio of the number of galaxies observed between L and $L+dL$ to the the number brighter than L is proportional to the ratio of the number of galaxies actually in the field in this range to those brighter than L multiplied by a weighting factor to account for the differing volumes sampled. The method generates a cumulative luminosity function without normalisation. The differential luminosity function may be derived by fitting an appropriate model.

The STY method (Sandage, Tammann & Yahil 1979) twists this plane around. It examines the probability that a galaxy observed at a redshift z_0

is brighter than M:

$$P(M, z_0) = \frac{\int_{-\infty}^{M} \phi(M')D(z_0)f(m')dM'}{\int_{-\infty}^{\infty} \phi(M')D(z_0)f(m')dM'} \tag{4.4}$$

where $f(m)$ is the completeness of the survey at apparent magnitude m and $D(z)$ is the density of galaxies at redshift z divided by the mean density of galaxies ($\bar{\rho}$). Taking the derivative of this equation with respect to M yields the probability density for finding a galaxy with absolute magnitude M in a magnitude-limited survey. This probability is directly proportional to the density of galaxies with that that apparent magnitude and inversely proportional to that which could have been observed at that particular redshift,

$$p_k \propto \phi(M_k) \Bigg/ \int_{M_{\mathrm{faint}}(z_k)}^{M_{\mathrm{bright}}(z_k)} \phi(M')dM' \tag{4.5}$$

Here I have replaced the function $f(m')$ with a function that is zero outside the magnitude limits of the survey and one within (*i.e.* assuming that the survey is one-hundred percent complete). As the redshift of the galaxy is fixed, the discontinuities is $f(m')$ correspond with range of absolute magnitudes beyond which no galaxies at this redshift could have been observed. From Equation 4.5, the SWML technique (Efstathiou, Ellis & Peterson 1988) originated. The modification of the SWML method (SSWML) is a complement to Chołoniewski's method, but for a different statistical model.

4.2.1 Deriving the SSWML Method

The aforementioned clustering-insensitive methods by design probe the luminosity function as a function of luminosity only. To understand the evolution of the luminosity function with redshift, we must remove this restriction. We will derive two new methods without this restriction, generalisations of the STY method and SWML. These generalisations will be denoted with a prefixed 'S' for spatial. In the derivation, it becomes apparent that these generalisations reap additional rewards. They provide a straightforward prescription to combine various surveys coherently and to determine the absolute normalisation of the luminosity function.

The derivation of the SSWML method begins with Equation 4.4 of the STY method, but I will make two generalisations:

- $\rho(z, M) \neq \bar{\rho}D(z)\phi(M)$ (*i.e.* Luminosity and density evolution), and
- $f(m) = \Omega(m)$, where $\Omega(m)$ is the area of sky sampled at apparent magnitude m, accounting for sampling rate and mean completeness.

These two generalisations will allow the determination of the luminosity function as a function of redshift (as in Chołoriewski 1986) and the use of many surveys in a single coherent determination of the luminosity function.

The relevant probability is the probability that a galaxy in the survey is brighter than M and closer than z:

$$P(M,z) = \frac{\int_0^z \int_{-\infty}^M \rho(z',m')\Omega(m')\frac{dV}{dz}dM'dz'}{\int_0^\infty \int_{-\infty}^\infty \rho(z',M')\Omega(m')\frac{dV}{dz}dM'dz'} \tag{4.6}$$

Again, taking the derivative of this equation yields the probability density:

$$p_k = \frac{\partial^2 P(M,Z)}{\partial M \partial z} \propto \rho(z_k, M_k) \Big/ \int_0^\infty \int_{-\infty}^\infty \rho(z',M')\Omega(m')\frac{dV}{dz}dM'dz' \tag{4.7}$$

where $\Omega(m)$ again is the solid angle sampled at the apparent magnitude m, and m is the apparent magnitude corresponding to M and z considering the distance modulus and the k-correction; consequently m may best be written $m(z, M; c_k)$ where c_k is the k-correction class of galaxy k. This equation forms the basis of the generalised STY method or SSTY which will be used in Chapter 6 to derive the evolution of galaxies by spectral class.

Figure 4.1 compares Equations 4.7 and 4.5 on the absolute magnitude *versus* redshift plane. The lower double integral is simply the number of galaxies that one would expect to observe in a combined survey given a luminosity function. If the function $\Omega(m)$ is simply a series of steps (as in Figure 4.1), there is only one k-correction class, and the trial luminosity function is integrable at a given redshift (like an evolving Schechter function), the integral may be most efficiently calculated as

$$N_{\rm gal}^{\rm pred} = \int_0^\infty \sum_j \Omega_j \int_{B_{j,\rm min}}^{B_{j,\rm max}} \rho(z', m' - d_{\rm modulus}(z) - k(z))\frac{dV}{dz}dm'dz' \tag{4.8}$$

where the sum is over the subsurveys and the jth survey samples Ω_j sterradians in an apparent magnitude range from $B_{j,\rm min}$ to $B_{j,\rm max}$. Although the function $\rho(z, M)$ may not be separable, if the evolution of the luminosity function is posed as an evolution of the Schechter parameters, $\rho(z, M)$ takes the form of an integrable Schechter function at all redshifts.

Things have become much more complicated; however, if we reinstate the assumption that the density is separable and that we have a single magnitude-limited sample, this formula reduces to Equation 4.5:

$$p_k \propto n(z_k)\phi(M_k) \Big/ \int_0^\infty n(z)\frac{dV}{dz}\int_0^\infty \phi(M)\Omega(m)dM'dz' \tag{4.9}$$

We can calculate $\Omega(m)$ for z_k

$$\Omega(m) = \begin{cases} \Omega & \text{if } M_{\rm faint}(z_k) \le M \le M_{\rm bright}(z_k) \\ 0 & \text{otherwise} \end{cases} . \tag{4.10}$$

Since we are looking for an estimator for $\phi(M)$ which is assumed to be everywhere constant, we can chose to perform the latter integral at the

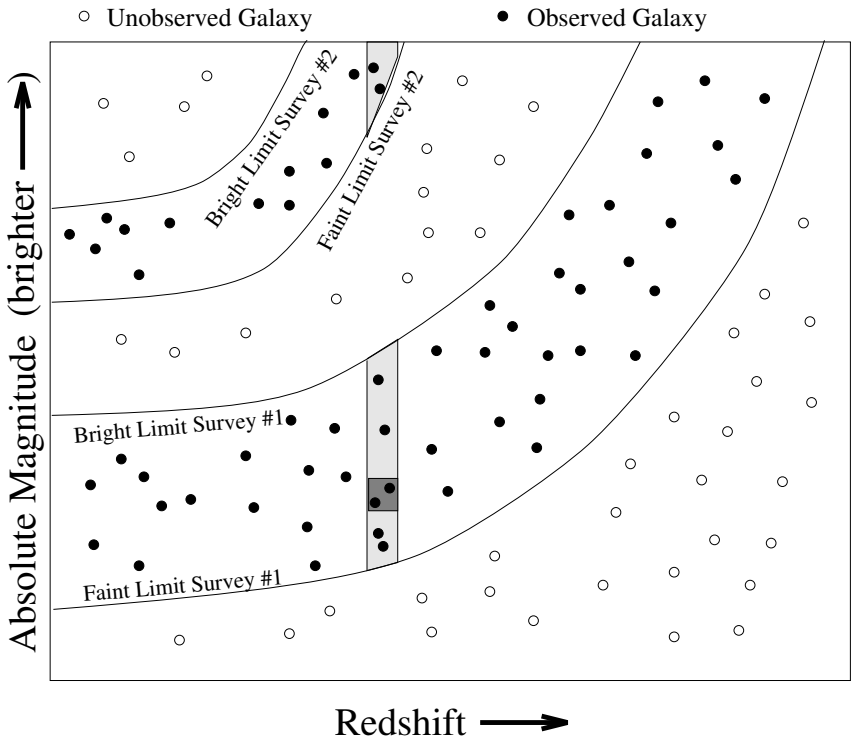

Figure 4.1: Geometric Comparison of the STY and SSTY Methods. In the STY method, the probability of observing a galaxy is the ratio of the number of galaxies at the galaxy's absolute magnitude (the dark-shaded region) to the total number of galaxies that could have been observed within the magnitude limits at the same redshift as the observed galaxy (the light-shaded region). The extension to the SSTY technique is straightforward. The probability of observing a galaxy at a particular redshift is the ratio of the density of galaxies at the observed redshift and absolute magnitude (again the dark-shaded region) to the total number of galaxies that could have been observed at any redshift (the region enclosed by the magnitude limits).

redshift of the observed galaxy (z_k), yielding

$$p_k \propto n(z_k)\phi(M_k) \Bigg/ \left(\int_0^\infty n(z)\Omega \frac{dV}{dz} dz \int_{M_{\rm faint}(z_k)}^{M_{\rm bright}(z_k)} \phi(M) dM \right) \quad (4.11)$$

which is equivalent to the STY result.

Now, let's proceed in the spirit of the SWML (Efstathiou, Ellis & Peterson 1988) method and assume that $\rho(z, M)$ is defined for a two-dimensional array of steps in both redshift and luminosity,

$$\rho(z, M) \equiv \sum_{ij} W(z - z_i, M - M_j)\rho_{ij} \quad (4.12)$$

where

$$W(z, M) = \begin{cases} 1 & \text{if } -\Delta z/2 \le z \le \Delta z/2 \text{ and } -\Delta M/2 \le M \le \Delta M/2 \\ 0 & \text{otherwise} \end{cases} . \quad (4.13)$$

Substituting this relation back into the formula for p_k gives

$$p_k \propto \sum_{ij} W(z_k - z_i, M_k - M_j)\rho_{ij} \Bigg/ \sum_{ij} \rho_{ij} \int_{z_i - \Delta z/2}^{z_i + \Delta z/2} \int_{M_j - \Delta M/2}^{M_j + \Delta M/2} \Omega(m) \frac{dV}{dz} dM dz \, . \quad (4.14)$$

Next we calculate the logarithm of p_k

$$\ln p_k = \sum_{ij} W(z_k - z_i, M_k - M_j) \ln \rho_{ij} - \ln I_k \quad (4.15)$$

where we have replaced the denominator with I_k. We have retained the subscript k because the value of the integral depends on the k- correction for the particular galaxy observed through $\Omega(m(z, M; c_k))$.

The likelihood of observing the entire survey is the product of the likelihoods for each galaxy

$$P_{\rm survey} = \prod_k p_k, \quad (4.16)$$

and taking the logarithm,

$$\ln P_{\rm survey} = \sum_k \ln p_k. \quad (4.17)$$

Because we would like to maximise the likelihood of the entire survey we can simply take the derivative with respect to ρ_{pq} and look for stationary points:

$$\frac{d \ln P_{\rm survey}}{d\rho_{pq}} = \sum_k \frac{d \ln p_k}{d\rho_{pq}} = 0. \quad (4.18)$$

Now we differentiate Equation 4.15,

$$\frac{d \ln p_k}{d\rho_{pq}} = \frac{W(z_k - z_p, M_k - M_q)}{\rho_{pq}} - \frac{1}{I_k} \frac{dI_k}{\rho_{pq}}. \quad (4.19)$$

Looking back to Equation 4.14, we see that

$$\frac{dI_k}{d\rho_{pq}} = \int_{z_p-\Delta z/2}^{z_p+\Delta z/2} \int_{M_q-\Delta M/2}^{M_q+\Delta M/2} \Omega(m) \frac{dV}{dz} dM\, dz \tag{4.20}$$

and

$$I_k = \sum_{pq} \rho_{pq} \frac{dI_k}{d\rho_{pq}} \tag{4.21}$$

Now we sum over the derivatives (Equation 4.19) which gives

$$\frac{d \ln P_{\rm survey}}{d\rho_{pq}} = \frac{1}{\rho_{pq}} \sum_k W(z_k - z_p, M_k - M_q) - \sum_k \frac{1}{I_k} \frac{dI_k}{d\rho_{pq}}. \tag{4.22}$$

if we set this equal to zero, we get

$$\rho_{pq} = \frac{\sum_k W(z_k - z_p, M_k - M_q)}{\sum_k \left\{ dI_k/d\rho_{pq} \Big/ \sum_{ij} \rho_{ij} dI_k/d\rho_{ij} \right\}} \tag{4.23}$$

which is similar in structure to the SWML result,

$$\phi_q \Delta M = \frac{\sum_k W(M_k - M_q)}{\sum_k \left\{ H[M_q - M_{\rm faint(z_k)}] \Big/ \sum_j \phi_j \Delta M H[M_j - M_{\rm faint(z_k)}] \right\}} \tag{4.24}$$

as we can identify I_k with the sum over j in the denominator and $dI/d\rho_{pq}$ with $H[M_q - M_{\rm faint(z_i)}]$ also in the denominator.

Calculating the values in the demonimator of Equation 4.23 is a bit less straightforward than in the SWML method, but looking at Equations 4.20 and 4.21 shows that this denominator is a function of the ρ_{pq}, the k-correction class of galaxy k, the cosmology (through dV/dz and the distance modulus) and the details of the survey (through $\Omega(m)$). Therefore we can calculate the values of the $dI_k/d\rho_{pq}$ for each k-correction class before beginning the iterative solution to Equation 4.23.

Furthermore, since the the integrals I_k are over volume, it is straightforward to calculate the total number of galaxies that one expects to observed given the current values of ρ_{ij},

$$n_{\rm gal}^{\rm pred} = \sum_{ijk} f_k \rho_{ij} \frac{dI_k}{d\rho_{ij}} \tag{4.25}$$

where f_k is the fraction of the galaxies observed in each k-correction class. Then one can normalise all of the ρ_{pq} so that Equation 4.25 predicts the number of galaxies observed in the catalogue. One can normalise either at each iteration by multiplying the right-hand side of Equation 4.23 by $n_{\rm gal}^{\rm observed}/n_{\rm gal}^{\rm pred}$, or multiplying by this ratio after the final iteration. The algorithm converges more quickly (two or three iterations) if the normalisation is performed at each step.

The errors for this analysis may be estimated using the following formula (following from Saunders et al. 1990):

$$\begin{aligned}\sigma(\log\rho_{pq}) &= (\ln 10)^{-1}\sigma(\ln\rho_{pq}) \\ &= (\ln 10)^{-1}\left(\frac{\partial^2 \ln P_{\rm survey}}{(\partial\ln\rho_{pq})^2}\right)^{-1/2} \\ &= (\ln 10)^{-1}\times \\ &\quad\left[\sum_k\left\{W(z_k-z_p,M_k-M_q)-\left(\rho_{pq}\frac{dI_k}{d\rho_{pq}}\Big/\sum_{ij}\rho_{ij}\frac{dI_k}{d\rho_{ij}}\right)^2\right\}\right]^{-1/2}.\end{aligned} \tag{4.26}$$

Additionally, we can easily estimate upper limits for bins in which we did not observe any galaxies,

$$\rho_{pq}^{\rm upper} = \frac{1/2}{\sum_k\left\{dI_k/d\rho_{pq}\Big/\sum_{ij}\rho_{ij}dI_k/d\rho_{ij}.\right\}} \tag{4.27}$$

In the derivation, the function $\Omega(m)$ is a rug under which to sweep a variety of problems. The simplest way to define $\Omega(m)$ is to calculate the area surveyed in each subcatalogue and then multiply this area by the sampling rate and completeness. In this way, we form a function $\Omega(m)$ which appears as a series of steps with a jump at the bright and faint limits of each subcatalogue. However, we may easily make this more general in two ways. Firstly, since we explicitly calculated $dI_k/d\rho_{ij}$ for each galaxy type, we can introduce completeness as a function of k-correction type as well by have various $\Omega_k(m)$. Secondly, we can calculate the completeness rate within a subcatalogue as a function of apparent magnitude and account for this in obtaining $\Omega(m)$ or $\Omega_k(m)$. These two simple generalisations make this technique an extremely versatile tool in analysing galaxy catalogues.

4.3 Tests and Comparisons

To test the $1/V_{\rm max}$ and SSWML methods in a coherent sample like the composite DARS, AUTOFIB and LDSS-1+2 surveys, I generated a random galaxy catalogue from a Schechter function (Loveday et al. 1992). Three-hundred galaxies were selected in each of the following magnitude ranges: $11.0-17.3$, $17.0-19.7$, $19.7-20.5$, $20.3-20.8$, $20.8-22.5$, and $22.5-24.0$. A Hubble constant of 100 km/s/Mpc and a deceleration parameter of 0.5 were assumed, and all the galaxies were assigned zero k-correction. Finally, the density of galaxies was doubled beyond a redshift of 0.2 to crudely simulate density evolution. A second catalogue with 3,000 galaxies in the above ranges was also generated to understand how the two algorithms converge to the true luminosity functions.

Figures 4.2 and 4.3 illustrate the results of these simulations. The two methods generate nearly identical luminosity functions, although the maximum likelihood method produces slightly smoother luminosity functions.

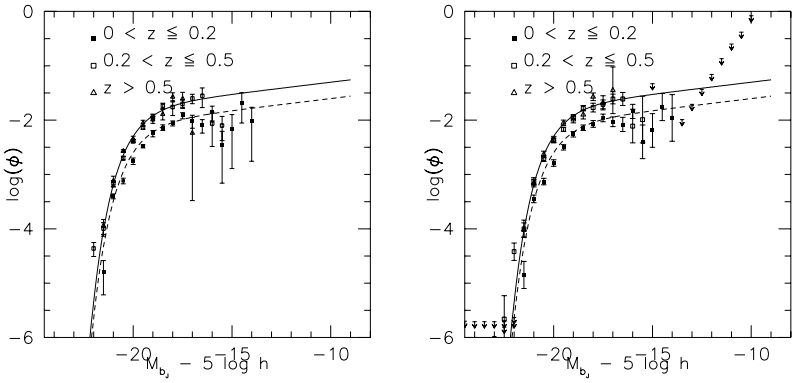

Figure 4.2: Test Catalogue of 1,800 Galaxies. The left panel depicts the results of the $1/V_{\max}$ analysis on the random galaxy catalogue and right panel shows the results of the SSWML method. In each pane, the lower curve is the nearby ($z < 0.2$) luminosity function, and the upper is the distant one. The errorbars for the $1/V_{\max}$ method are generated using a bootstrap technique, while the errors and upperlimits for the SSWML are determined as described in the text.

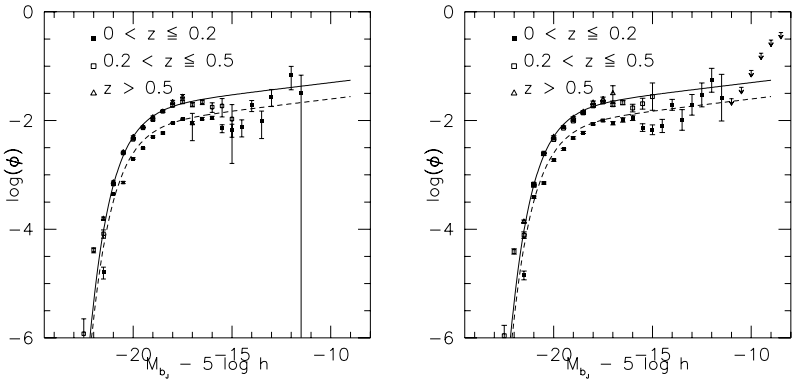

Figure 4.3: Test Catalogue of 18,000 Galaxies. The curves and symbols are as in Figure 4.2.

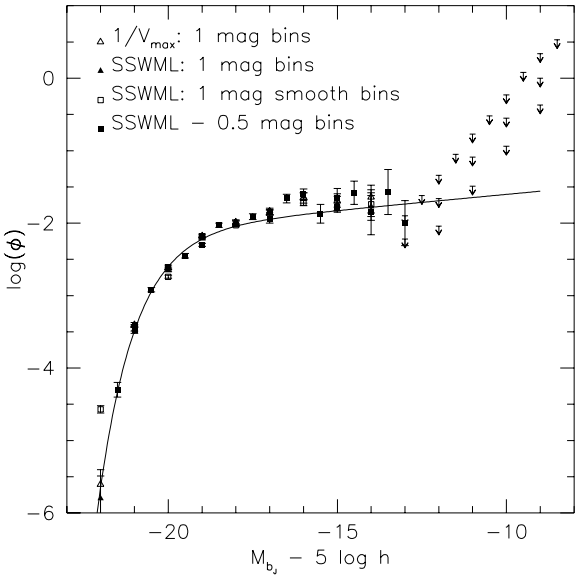

Figure 4.4: Clustered Test Catalogue of 1,800 Galaxies. The luminosity function of the test catalogue was determined using a variety of methods.

The bootstrap errors of the $1/V_{\rm max}$ method agree remarkably well with errors derived for the maximum likelihood technique. The SSWML technique has the additional advantage of determining upper limits, which further constrain the evolution of the faint-end slope.

To test the sensitivity of the algorithms to clustering, a second galaxy catalogue was generated with a "cluster" in one of the fields. The cluster is an overdensity of galaxies at a redshift of 0.05 in the field sampling from $17.0 - 19.7$. The luminosity function in the cluster is identical to that in the field. Figure 4.4 shows the results of this simulation for a variety of bin widths. The $1/V_{\rm max}$ and the narrowly binned SSWML methods overpredict the number of galaxies at the faint end. However, as the bin width of the SSWML method is increased, the overprediction decreases. The smooth binning technique uses two SSWML calculations with bins of two magnitudes, shifted by one magnitude relative to each other. This achieves the closest results to the input luminosity function.

4.3.1 $V_{\rm max}$ as a maximum likelihood estimator

Figure 4.4 reveals a interesting correspondence between the $1/V_{\rm max}$ and the maximum-likelihood technique. It appears that the smaller the SSWML bins are, the closer it approximates the $1/V_{\rm max}$ method. Looking back to Equation 4.23, the sum over the k objects in the survey in the numerator may be replaced with n_{pq} the number of objects in the appropriate bin.

Furthermore, we shall take the case where there is only one type of galaxy. This is not a restrictive assumption, as one could split the surveyed galaxies by type, calculate each type's luminosity function, and sum them up. These two alterations result in

$$\rho_{pq} = \frac{n_{pq}}{n_{\rm gal}^{\rm observed}} \frac{\sum_{ij} \rho_{ij} dI/d\rho_{ij}}{dI/d\rho_{pq}} \tag{4.28}$$

where the sum over k has been replaced by a multiplication. Looking back to Equation 4.25, the sum in the numerator may be replaced with $n_{\rm gal}^{\rm predicted}$, yielding

$$\rho_{pq} = \frac{n_{pq}}{n_{\rm gal}^{\rm observed}} \frac{n_{\rm gal}^{\rm pred}}{dI/d\rho_{pq}}. \tag{4.29}$$

It appears that the iterative process has entirely disappeared in this simplified case. The normalisation of ρ_{pq} must be determined by multiplying by $n_{\rm gal}^{\rm observed}/n_{\rm gal}^{\rm pred}$, yielding

$$\rho_{pq} = \frac{n_{pq}}{dI/d\rho_{pq}}. \tag{4.30}$$

The final connection is Equation 4.20 in the limit that ΔM is small, such that $\Omega(m)$ can be assumed to be constant across ΔM. A survey has magnitude limits, so this can only be an approximation. In this way, the inner integral may be approximated by a product:

$$\frac{dI}{d\rho_{pq}} \approx \int_{z_p-\Delta z/2}^{z_p+\Delta z/2} \Omega(m(z, M_q)) \Delta M \frac{dV}{dz} dz \approx V_{pq,\max} \Delta M. \tag{4.31}$$

Substituting Equation 4.31 into Equation 4.30 yields the familiar $1/V_{\rm max}$ equation,

$$\rho(z_p, M_q) = \rho_{pq} \approx \frac{n_{pq}}{V_{pq,\max} \Delta M} \tag{4.32}$$

and we have come full circle.

4.4 Conclusion

The $1/V_{\rm max}$ and a new maximum-likelihood technique (SSWML) are compared on random galaxy catalogues, and both give similar good results. The SSWML method has the additional advantage of being less sensitive to clustering and estimating upperlimits to the luminosity function for bins in which no galaxies were observed. Furthermore, the SSWML and the $1/V_{\rm max}$ are equivalent in the limit of small bins in absolute magnitude, showing that the $1/V_{\rm max}$ technique in a sense yields the survey that maximises the likehood of observing a redshift survey.

Chapter 5

The Evolving Galaxy Luminosity Function

SUMMARY

The luminosity function and its evolution are the fundamental features of galaxies as a population. Often these statistics are described by a simple parameterisation: the Schechter function, with luminosity or density evolution. In this section, the luminosity function at a variety of redshifts is derived from the `AUTOFIB`, `DARS`, `BES`, `LDSS-1` *and* `LDSS-2` *surveys, by means of two non-parametric methods. The first is the usual* $1/V_{\rm max}$ *method which is unbiased, and the second is a modification of the step-wise maximum-likelihood method described in Section 4.2.1 which is insensitive to galaxy clustering. Both of these methods avoid giving the answer ahead of time in the form of a specific parametrisation. The luminosity function is found to strongly evolve with redshift.*

5.1 The Local Luminosity Function

The volume density of intrinsically faint galaxies is a point of contention. Most hierarchical models (c.f. Kauffmann, White & Guiderdoni 1993, Cole et al. 1994a and Chapter 9) predict a steep faint-end slope to the luminosity function with $\alpha \simeq -2.0$. Looking to the Virgo cluster, Binggeli, Tammann & Sandage (1985) find a similar number of faint galaxies. Also the excess of blue galaxies at faint magnitudes is difficult to reconcile with a flat universe without a steep faint-end slope in the local luminosity function. However, observations of field-galaxy samples consistently find a flat faint-end slope, $\alpha = -1.0 \pm 0.1$ (Efstathiou, Ellis & Peterson 1988, Loveday et al. 1992, and Marzke, Huchra & Geller 1994) down to luminosities of of $M_B < -17 + 5 \log h$.

In their analysis of the CfA redshift survey, Marzke, Huchra & Geller (1994)

claim the first evidence of a possible upturn fainter than $M_{\rm Zwicky} = -16 + 5\log h$. Specifically, they observe 3 times as many low-luminosity objects in this category as would be expected from an extrapolation of the Schechter function fitted at brighter luminosities. Unfortunately, the uncertainties are still too great for Marzke et al's result to be considered definitive. A scale error in the photometric scale of the Zwicky catalogue could significantly reduce the excess and the effect is greatest in the northern cap where Binggeli et al's Virgo galaxies inevitably produce some of the signal. On the other hand, Bernstein (1994) has estimated the luminosity function of the Coma cluster, and found that it remains flat brighter than $M_R = -13 + 5\log h$ or $M_B \approx -11 + 5\log h$. With these various conflicting viewpoints, one must ascertain the uncertainties in the local luminosity function. With the exception of the Bernstein (1994) result, the surveys cannot constrain the luminosity function fainter than $M > -16 + 5\log h$, simply because of their bright magnitude limits. The faint work of Broadhurst, Ellis & Shanks (1988) to $b_J = 21.5$ showed that a significant contribution of low luminosity galaxies would shift the redshift distribution at these magnitudes to lower redshifts than they observed. The yet fainter LDSS and LDSS-2 surveys (Colless et al. 1990, Glazebrook et al. 1993) found redshift distributions consistent with no-evolution in the luminosity function and a flat faint-end slope. Additional surveys at faint magnitudes are necessary to further constrain the faint-end of the luminosity function as well as its evolution. Also, if these surveys are analysed coherently, the luminosity function as a function of redshift may be constructed directly.

A more poorly understood uncertainty is the selection effect introduced by the varying mean surface brightness of galaxies. The conjecture is that many field surveys may miss an entire population of *low surface brightness (LSB)* galaxies by virtue of selection effects inherent in common image detection algorithms (Disney 1976, Davies 1990). In general, a galaxy is included in a survey if its image is sufficiently large down to a given isophotal magnitude (isophotal-diameter selection), and its image is sufficiently bright within this isophote (isophotal-magnitude selection). Davies (1990) notes that this type of selection yields a peak in the observed surface-brightness distribution independent of the underlying distribution. Furthermore, this type of selection is biased toward spirals which have less concentrated brightness profiles than ellipticals. Disney (1976) proposed with this surface-brightness bias "galaxies are like icebergs and what is seen above the sky background may be no reliable measure of what lies underneath."

The only way to determine the location of these galaxies in the luminosity function is through observations. Impey, Bothun & Malin (1988) found that some LSBs are quite luminous, amongst the brightest galaxies observed. Like the largest living things, the most luminous galaxies may reveal themselves as insignificant fungi peeking through the obscuring humus of the night sky. However, Davies et al. (1994) have probed beneath the humus by cross-correlating exponential disk profiles against the noise of

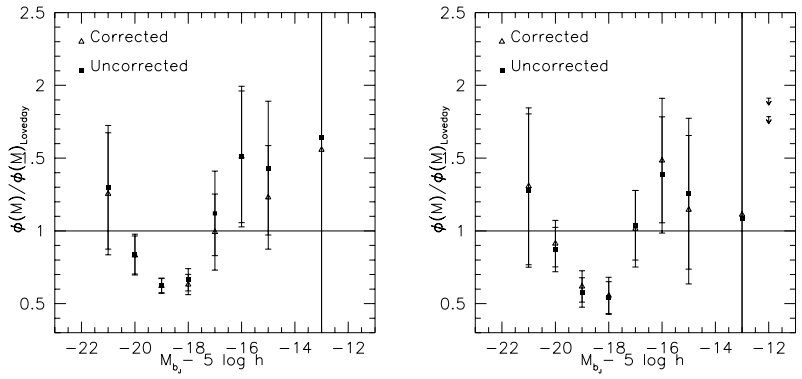

Figure 5.1: The Local Luminosity Function. The two panes depict the local luminosity function from the combined survey in the range $0.006 < z < 0.05$. The resulting luminosity functions have been divided by the Loveday et al. (1992) Schechter fit (without correction for Malmquist bias) to expand the dynamic range of the ordinate. The left shows the result of the $1/V_{\rm max}$ method and the right shows the result of the SSWML method. The results with and without completeness correction are shown.

"empty" sky. They found several extended sources but no hyperluminous objects like Malin 1.

They estimate that LSBs in the range $-22 < M_0 + 5\log h < -19$ and μ_0 about $26V$mag arc sec^{-1} are less common than their normal-surface-brightness counterparts by an order of magnitude or more. With the fainter isophotal limits, the AUTOFIB combined survey should uncover at least a portion of these and fainter low-surface-brightness galaxies.

Figure 5.1 compares the Loveday luminosity function with the luminosity function derived from the 205 galaxies in the combined survey with with $0.006 < z < 0.05$. The lower limit of $cz = 1800$km sec^{-1} avoids difficulties with peculiar motions and only excludes 12 galaxies. The upper redshift limit obviates the effect of the density enhancement at $z \sim 0.1$, shown in Figure 5.2. The luminosity functions are divided by the best fitting Schechter function of Loveday et al. (1992) to expand the useful range of the ordinate. The effects of including the correction for magnitude-dependent completeness and varying the algorithm are also shown.

Even with the large errorbars, AUTOFIB constrains the luminosity function more tightly than the brighter surveys. The survey measures the luminosity function fainter than -16, by sampling a larger volume to fainter limits. Furthermore, the AUTOFIB data is sampled to a fainter limiting isophote, approximately μ_{b_J}=26.5 arcsec^{-2} (Jones et al. 1991). Therefore, the survey would detect low-surface-brightness galaxies such as those found in Virgo. Although the luminosity function presented in Figure 5.1 differs from the Loveday et al. (1992) result, it is not consistent with the Marzke,

Huchra & Geller (1994) finding that the number of galaxies in the range $-16 \leq M_Z \leq -13$ exceeds the extrapolation of their best-fit Schechter function at the bright-end by a factor of 3.1 ± 0.54. The results here exhibit no such excess at the faint end. This result is independent of the luminosity function estimator and completeness correction. The excess observed in the Marzke, Huchra & Geller (1994) northern sample is consistent with Binggeli, Tammann & Sandage (1985) luminosity function for the Virgo cluster which contributes about half of the galaxies in the northern sample at faint absolute magnitudes. Furthermore, a scale error in the Zwicky magnitudes of 0.2 mag mag^{-1} would reduce the excess in the south to a more modest 1.8±0.3, marginally consistent with the results here. Moreover, observations of the Coma cluster indicate a flat faint-end slope to $M_R \approx -13 + 5 \log h$ (Bernstein 1994, private communication), beyond the results in Figure 5.1.

Probing deeper into the local universe reveals more structures. To study the local luminosity function, I used the SSWML method to measure the density of galaxies in two bins in absolute magnitude and twenty bins in redshift from 0 to 0.2. The luminosity is binned coarsely to reduce the errors in the luminosity function determined at a given redshift. The bright absolute magnitude bin extends from -21.73 to -17.73, centred on the Loveday et al. (1992) determination of $M^* = -19.73 \pm 0.13$. The value of the luminosity function in this bin will be denoted $\langle\phi(M^*)\rangle$. The faint bin ($\langle\phi(M^*+4)\rangle$) probes the slope of the luminosity function, and extends from -17.73 to -13.73 and is centred on M^*+4. Table 5.1 shows how each of these values change with varying α. The bright bin is sensitive to changes in the normalisation, and the ratio of the values in the two bins is strongly sensitive to changes in the faint end slope. Figure 5.2 depicts the values of $\langle\phi(M^*)\rangle$ and the ratio of $\langle\phi(M^*+4)\rangle$ to $\langle\phi(M^*)\rangle$ as a function of redshift. The density of galaxies in the bright bin varies by nearly a factor of three. The ratio between the two bins varies little, indicating that the faint-end slope is almost constant. It is straightforward to verify that the total normalisation changes more drastically than parameters of the shape. Although an increase in bright-end cutoff luminosity also would result in an increase in $\langle\phi(M^*)\rangle$, it would also reduce the ratio of the two densities. Increasing the value of α to approximately -2 would also increase $\langle\phi(M^*)\rangle$ by nearly a factor of three, but the ratio of the densities in the two bins would be more than 60. Therefore, the density of galaxies increases without great changes in the shape of the luminosity function.

At a redshift of 0.075, an L^* galaxy would have a apparent magnitude of approximately 17. At this apparent magnitude, the number counts begin to dramatically exceed the no-evolution prediction (*e.g.* Maddox et al. 1990b). Postulating such rapid evolution over the last 0.6 h^{-1} Gyr is difficult. However, if some of the fields in which the counts have been measured pass through the density enhancement depicted in Figure 5.2, the counts excess at these bright magnitudes may be explained. To test this hypothesis, the density enhancement is fit with two exponentials which meet at $z = 0.085$

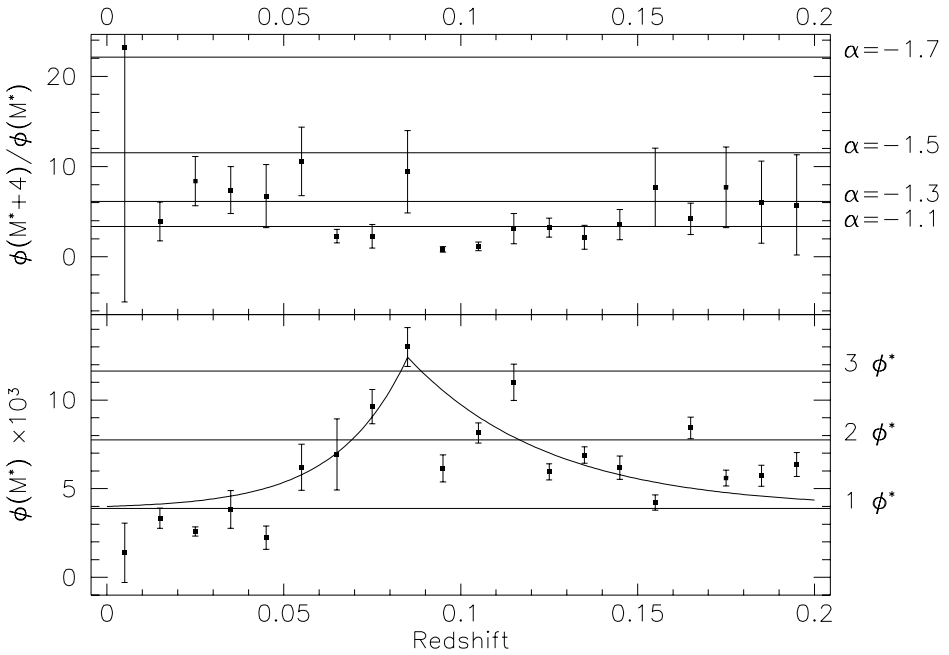

Figure 5.2: Local Density Fluctuations. The lower panel shows the values of the luminosity function in a bin from -21.73 to -17.73 as determined by the SSWML method. The horizontal lines show the value of this parameter for a Schechter with the Loveday et al. (1992) parameters for various values of ϕ^*, the normalisation. The curve shows a constrained fit to the changing normalisation. The upper pane depicts the measured values of the ratio of the two bins in the luminosity function. The value of this ratio for several different faint-end slopes (α) are illustrated with horizontal lines.

Table 5.1: Properties of a Schechter Function. The second and fourth columns correspond to the values depicted in Figure 5.2. The mean value of $\phi(M)$ near M^* changes little with α while the ratio changes quickly. Therefore, if M^* is assumed to change little, these two quantities can constrain both α and ϕ^*. For simplicity, the Schechter function has been normalised so that $\langle\phi(M^*)\rangle \equiv 1$ for $\alpha = -1.1$.

α	$\langle\phi(M^*)\rangle$	$\langle\phi(M^*+4)\rangle$	Ratio
-1.0	0.92	2.30	2.50
-1.1	1.00	3.35	3.35
-1.2	1.09	4.94	4.52
-1.3	1.20	7.37	6.15
-1.4	1.32	11.1	8.40
-1.5	1.46	16.9	11.5
-1.6	1.63	26.0	15.9
-1.7	1.82	40.2	22.1
-1.8	2.03	62.8	30.9
-1.9	2.29	98.9	43.3
-2.0	2.58	157.0	60.8

and decay at different rates to higher and lower redshift reaching the local density as measured by Loveday et al. (1992) in the limit,

$$\rho(z,L) = \begin{cases} (1+2.3e^{(z-0.085)/0.02})\phi(L), & z < 0.085 \\ (1+2.3e^{(z-0.085)/0.04})\phi(L), & z \geq 0.085 \end{cases} \tag{5.1}$$

where $\phi(L)$ is the luminosity function measured by Loveday et al. I assumed that all galaxies have a k-correction of $2z$. The resulting number-magnitude relation is not very sensitive to the k-correction at these low redshifts. Figure 5.3 illustrates the predicted number counts with and without this density enhancement. The upper curve with the density enhancement exceeds the observations, while the lower "no-evolution" curve falls short. The third middle curve is a linear combination of the two possibilities, assuming that 40% of the observed fields exhibit the density fluctuation in Figure 5.2 while the other fields follow the local Loveday luminosity function.

Although the model indicated by Figure 5.2 exceeds the observed number counts, the intermediate model fits the data to nearly $b_J = 20$ or $z \approx 0.2$. This redshift corresponds to a lookback time of 1.6 h^{-1} Gyr, a more realistic timescale for the galaxy population to evolve dramatically.

5.2 Evolution over Half a Hubble Time

The combined survey by probing to the faintest magnitudes so far observed spectroscopically reveals the properties on the galaxy population nearly to

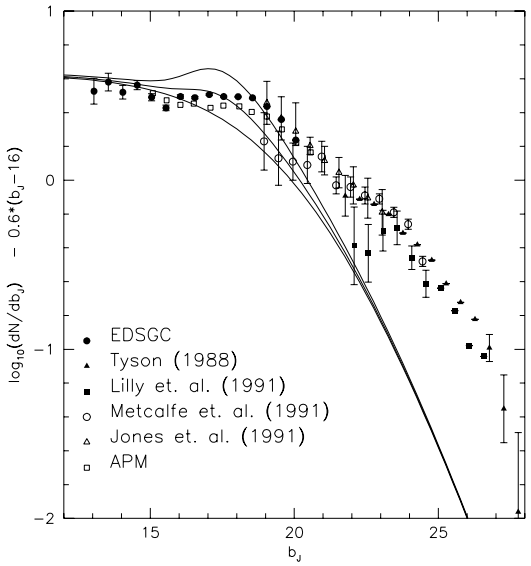

Figure 5.3: Number Counts Predictions with Density Fluctuations. The upper curve follows the model with density enhancement, and the lower curve traces a simple "no-evolution" model without the enhancement. The middle curve is a weighted average of the two possibilities which fits the observed counts nearly to $b_J = 20$. The data are taken from Maddox et al. (1990b), Jones et al. (1991), Metcalfe et al. (1991), Lilly, Cowie & Gardner (1991), Tyson (1988) and Heydon-Dumbleton *et al.* (1989; EDSGC). Where necessary Johnson B magnitudes have been converted to b_J assuming $b_J = B - 0.2$.

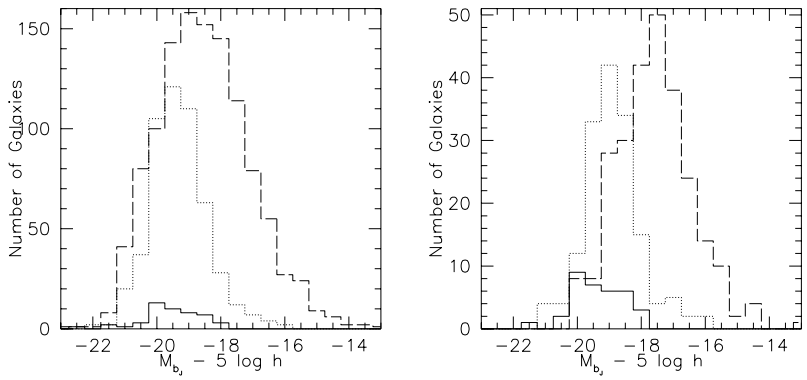

Figure 5.4: Coverage in Luminosity and Redshift. The left pane shows the number of galaxies in each of the redshift bins as a function of absolute magnitude. the right pane shows the same thing but for only the galaxies with W_λ[OII] greater than 20 Å. The dashed curve shows the number in the bin $0 < z \leq 0.2$, the dotted curve in the range $0.2 < z \leq 0.5$, and the solid curve for galaxies with $z > 0.5$.

a redshift of one, over more than half the age of the universe. Figure 5.4 shows the number of galaxies in the survey as a function of redshift and apparent magnitude. The right pane gives the number of galaxies with W_λ[OII] greater than 20Å which trace the star-forming population. Both the low and intermediate redshift bins are well-filled, and although the high redshift bin contains fewer galaxies, even here the survey constrains the luminosity function.

Figures 5.5 and 5.5 depict the evolution of the luminosity functions determined by both the $1/V_{\rm max}$ and the SSWML methods. For both methods, the low-redshift bins differs from the Loveday et al. (1992) fit at the 99.99% level. The intermediate bin differs from the low-redshift bin at the 99.9% and 99% levels for the $1/V_{\rm max}$ and the SSWML methods, respectively. Finally, the high-redshift bin differs from the intermediate bin at the 99% and 85% for the two methods. The SSWML method estimates slightly larger errors than the bootstrapped errors of the $1/V_{\rm max}$ method. It also yields slightly smoother luminosity functions which are better fit by Schechter functions.

Table 5.2 lists the parameters of the best-fitting Schechter functions with one-σ errors. The luminosity functions were fit over the entire range of absolute magnitude covered in each redshift bin. The errors among the three parameters are highly correlated; therefore, the simple one-parameter error values overestimate the uncertainties in the Schechter fits. The "Prob" column gives the χ^2 probability of realising a worse fit to the Schechter function than the data. A very low value indicates that a Schechter function does not provide a good fit to the data or that the errors have been

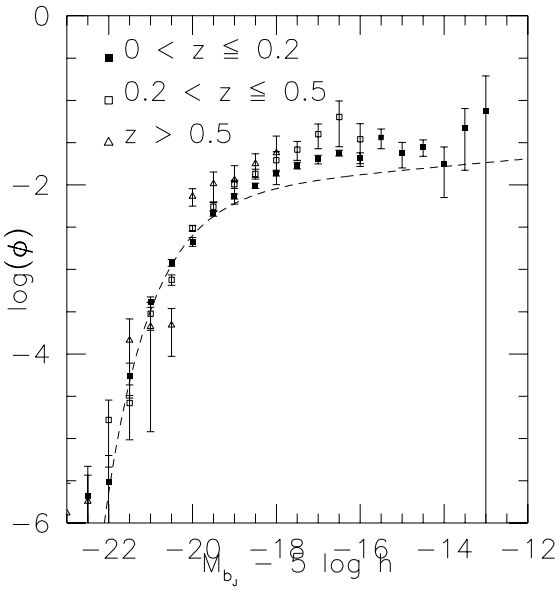

Figure 5.5: Evolution of the Luminosity Function - $1/V_{\rm max}$. The luminosity function at several redshifts is depicted by the symbols in the legend with one-sigma errorbars. The long-dashed curve is the best-fit Schetcher determined by Loveday et al. (1992) without correction for Malmquist bias. All the luminosity function figures that follow will have a similar format.

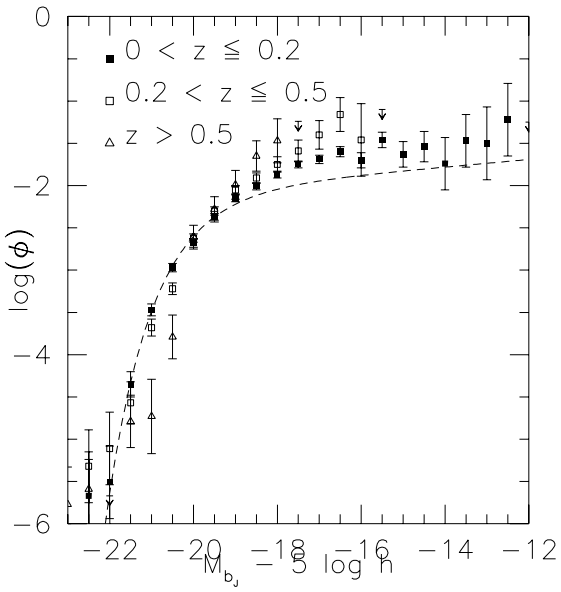

Figure 5.6: Evolution of the Luminosity Function - SSWML

underestimated. A value very near to one may imply that the errors have been overestimated.

The low-redshift bin has a steeper faint-end slope function than the local luminosity function described in the preceding section. The intermediate bin has a yet steeper faint-end slope, a slightly higher normalisation, and a fainter cutoff luminosity. The highest redshift bin appears to be poorly constrained by the data. However, more detail may be gleaned from the error ellipsoids shown in Figures 5.7 and 5.8. Only the ellipsoids for the SSWML method are depicted for simplicity. The successively larger ellipsoids represent higher redshift bins. The intermediate-redshift ellipsoid lies toward the upper-left-back corner *vis a vis* the low-redshift ellipsoid, corresponding to a combination of a fainter M^*, steeper faint-end slope and higher normalisation. The high-redshift ellipsoid intersects the intermediate-redshift ellipsoid only slightly and gives a still higher normalisation.

It is convenient to describe the evolution of the luminosity function in terms of the Schechter parameters; however, this evolution is also apparent in the binned luminosity functions themselves (Figures 5.5 and 5.6). These luminosity functions are not strongly dependent on the algorithm used to derive them, but the k-corrections and completeness corrections could have dramatic effects on the resulting luminosity functions.

Table 5.2: Schechter Fits to the Evolving Luminosity Functions

		$1/V_{\max}$		
Redshift	$\phi^*(10^{-2}\mathrm{Mpc}^{-3})$	α	M^*	Prob
$0 < z \leq 0.2$	1.08±0.41	-1.26±0.12	-19.77±0.29	0.11
$0.2 < z \leq 0.5$	1.82±1.30	-1.32±0.43	-19.46±0.49	0.46
$z > 0.5$	3.32±1.67	1.09±3.51	-18.09±0.84	0.12

		SSWML		
Redshift	$\phi^*(10^{-2}\mathrm{Mpc}^{-3})$	α	M^*	Prob
$0 < z \leq 0.2$	1.05±0.51	-1.29±0.17	-19.74±0.37	0.72
$0.2 < z \leq 0.5$	1.25±1.00	-1.50±0.58	-19.59±0.67	0.89
$z > 0.5$	5.03±5.00	-1.05±3.49	-18.7±1.57	0.92

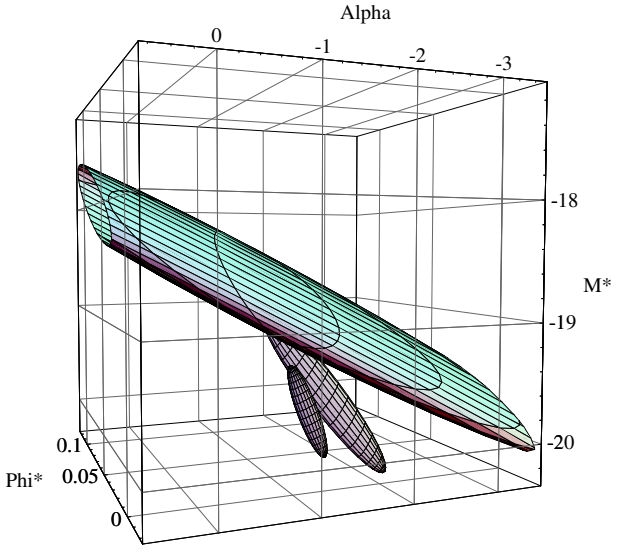

Figure 5.7: Error Ellipsoids for the Schechter Fits

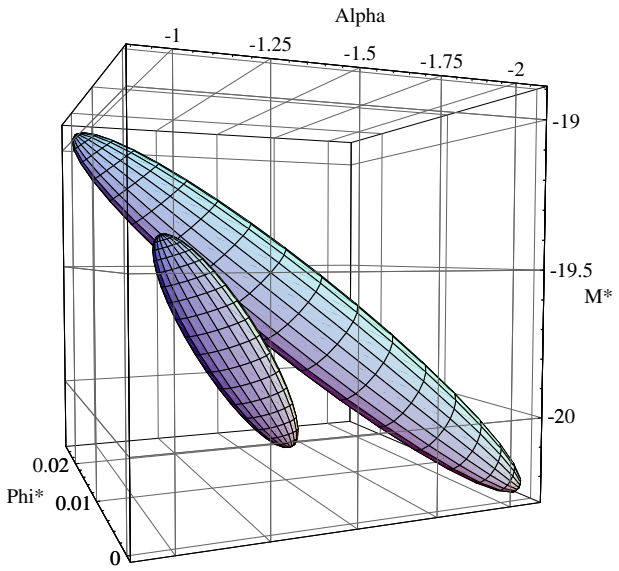

Figure 5.8: Error Ellipsoids for the Schechter Fits (Close-up)

5.2.1 K-corrections

The analysis in Chapter 3 shows that the cross-correlation technique may misclassify approximately one-fifth of the spectra in the survey by one k-correction class, blueward or redward. To examine this effect, 20% of the galaxies has been reclassified by one class blueward or redward, the sampling volumes and absolute magnitudes recalculated and the luminosity functions redetermined. Figure 5.9 presents the results of this analysis. Within the errorbars, the luminosity functions agree with those presented in Figures 5.5 and 5.6, showing that the luminosity functions derived are not strongly dependent on possible misclassifications.

As a second test, the luminosity functions in the observer's frame are calculated. Here, the k-corrections determined for the surveyed galaxies play no role in the derivation of the luminosity functions. However, a specific mix of galaxies with varying k-corrections has been assumed in constructing the "no-evolution" predictions (kindly provided by Tom Broadhurst). As expected as the galaxies suffer larger and larger k-corrections, the cutoff luminosity becomes markédly fainter with redshift in both the observations and the model. However, the model predicts fewer faint galaxies than are observed, even in the observer's frame luminosity functions. This effect is analogous to the increase in the faint-end slope in the rest-frame luminosity functions. Independent of possible errors in the k-corrections, the number of faint galaxies increases with look-back time.

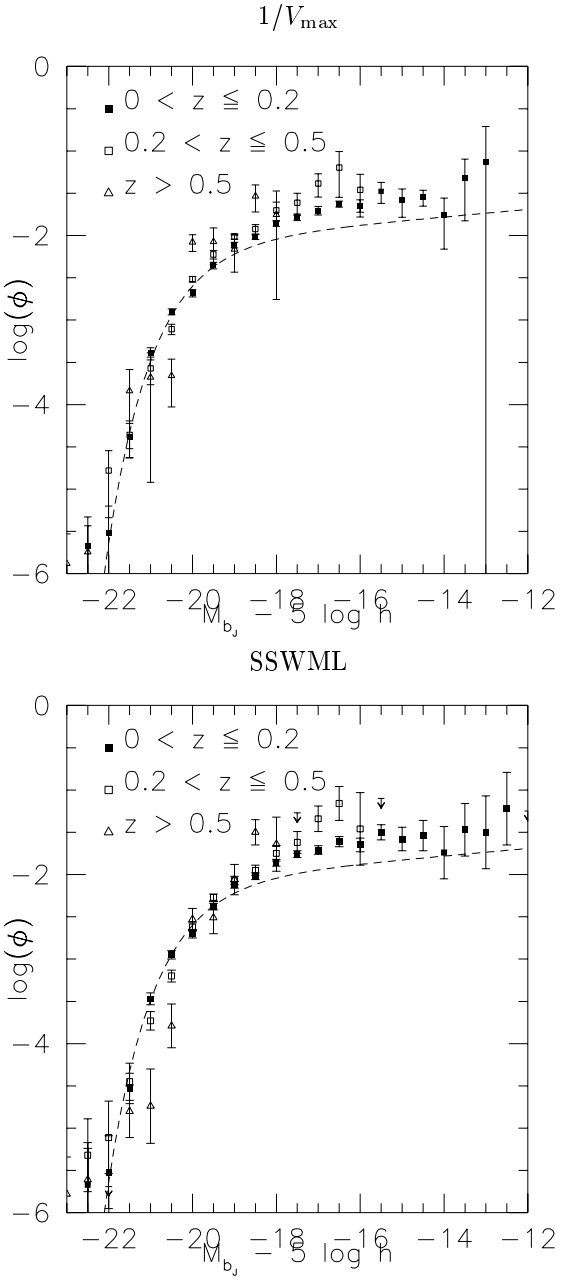

Figure 5.9: Evolution of the Luminosity Function with Random Classifications.

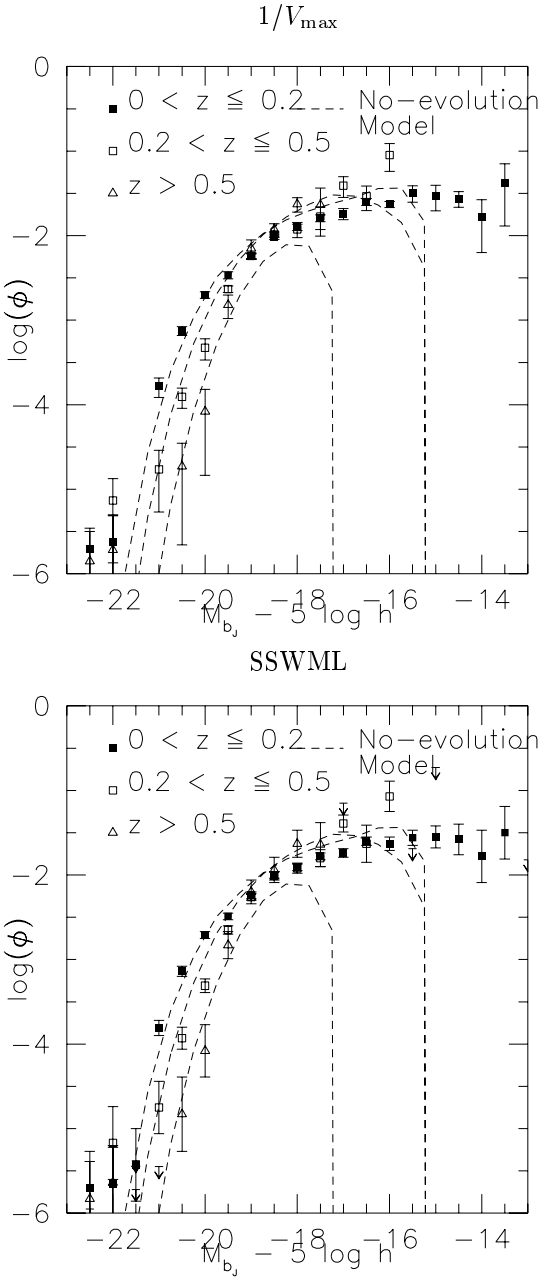

Figure 5.10: Observer's Frame Luminosity Functions

5.2.2 Completeness Correction

Section 2.4 describes the incompleteness in the AUTOFIB combined survey, and presents an algorithm to correct for incompleteness, which results in nearly uniform $V/V_{\rm max}$ distributions both on a survey-by-survey and a class-by-class basis. The AUTOFIB survey has significant incompleteness; one would expect this completeness correction to have as dramatic an effect on the luminosity functions as it has on the $V/V_{\rm max}$ distributions. Comparing Figure 5.11 with Figures 5.5 and 5.6 shows that the completeness correction does not have a statistically significant effect on the luminosity functions derived. The fields of the AUTOFIB combined survey (with the exception of the nearly complete DARS survey) are so narrow that both the complete bright-end and the incomplete faint-end probe similar luminosity functions, so reweighting the survey (*i.e.* the completeness correction) toward the faint-end of each survey has little effect on the derived luminosity functions. Furthermore, as the comparison of redshift distributions in Figure 2.5 implied, galaxies were probably missed simply because the signal-to-noise ratio was too low which would not introduce a systematic bias.

5.3 Implications

The measurement of the luminosity function is complicated by selection effects and intricate analysis. The results here are insensitive to errors in the k-corrections and incompleteness. The AUTOFIB combined survey probes fainter magnitudes and isophotal limits than previous surveys; therefore, it can uncover galaxies that may have been missed by the brighter surveys. Locally, the survey extends the flat faint-end slope found at brighter absolute magnitudes to $M_{b_J} \approx -15 - 5\log h$, in agreement with the Bernstein (1994) results for the Coma cluster. The shape of the luminosity function remains constant to $z \sim 0.15$. The excess in the number counts at bright magnitudes ($17 < b_J < 20$) may be explained by an increase in the normalisation of the luminosity function at $z \approx 0.09$; the density of galaxies returns to the local value by $z \approx 0.12$.

The luminosity function begins to steepen at a redshift of approximately 0.15. Past $z \sim 0.3$, it continues to steepen, and the normalisation increases. Letting the clock run forward, at about a redshift of 0.5, there was an excess of galaxies at all luminosities relative to today. By a redshift of 0.3 or so, the excess bright galaxies ($\sim L^*$) had disappeared or faded. The faint galaxies had also begun to vanish by this time and continued to vanish until a redshift of 0.15, yielding the flat faint-end slope of the local luminosity function. This evolution is explained by an interplay between several galaxy populations and Chapter 6.

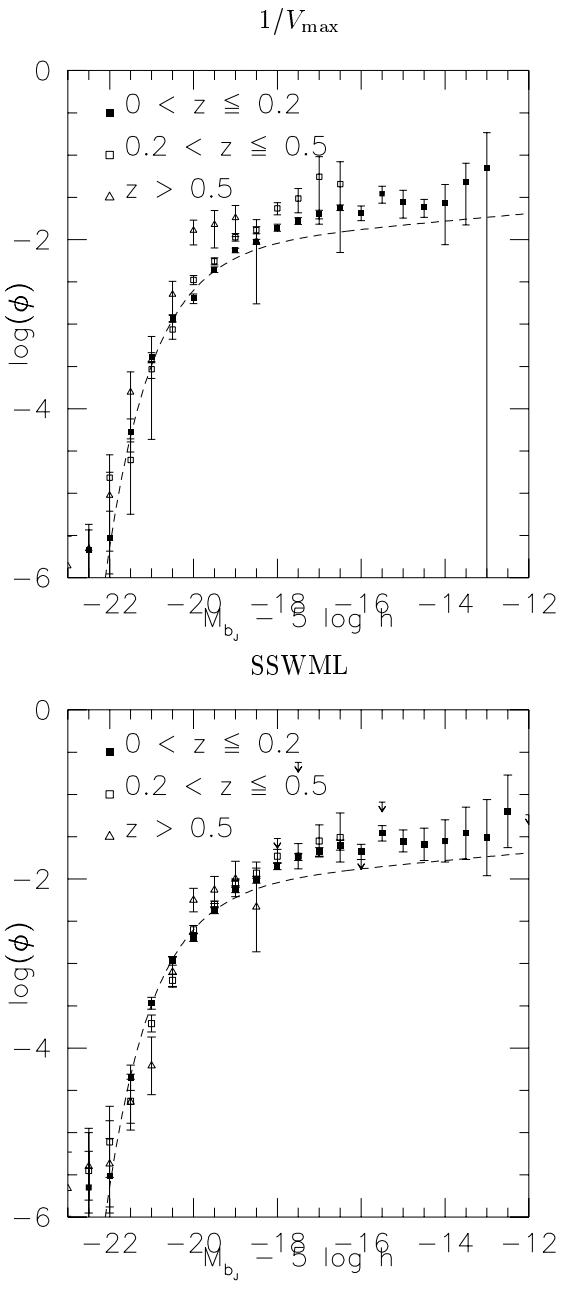

Figure 5.11: Completeness Corrected Luminosity Functions

Chapter 6

Evolution by Galaxy Type

SUMMARY

To understand the evolution of the galaxy population further, one must study how individual subpopulations of galaxies evolve. Subdividing the survey by spectral type is straightforward. Although it increases the noise in the various determinations, this examination is illustrative. The luminosity function for "star-forming" galaxies evolves strongly with redshift and to depend on the equivalent width (EW) of the forbidden line of singly-ionised oxygen (W_λ[OII] 3727). Furthermore, the evolution is strongly sensitive on galaxy type. The late-type galaxies were more numerous in the past especially at the faint end, while the early-type galaxies exhibit negative evolution.

6.1 Star-forming Galaxies

It is often found that as fainter samples are observed, stronger tracers of star formation are uncovered. Broadhurst, Ellis & Shanks (1988) found that distribution of W_λ[OII] 3727 was much wider for galaxies in the range $20.5 < b_J 21.5$ than for the DARS sample with $b_J < 17$. Colless et al. (1990) found a larger excess at yet fainter magnitudes. The `AUTOFIB` survey and LDSS-2 data augments these data. All the measured equivalent widths in the combined `AUTOFIB` survey are plotted against apparent magnitude in Figure 6.1. The median equivalent width for bins of one-half magnitude are plotted as large squares. There are not large jumps in the median equivalent width between the long-slit DARS data and fibre `AUTOFIB` data at $b_J \approx 17$. And the median equivalent width increases steadily across the fibre-multislit threshold at $b_J \approx 22$. This indicates that aperture effects are not very strong at these faint magnitudes.

This evolution of the equivalent-width distribution should be reflected in the evolution of the luminosity function of high-EW galaxies. This first cut at evolution by galaxy type is to divide the sample according to W_λ[OII]

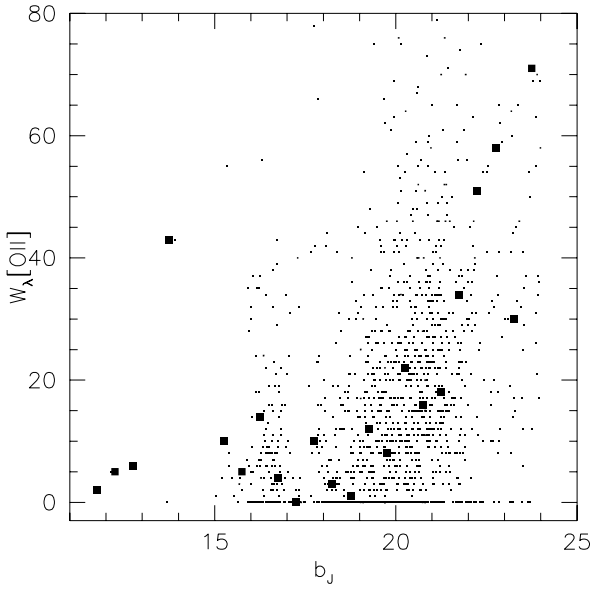

Figure 6.1: W_λ[OII] against apparent magnitude. The median EW is plotted as a larger square every 0.5 magnitudes in b_J. The boundary between DARS and `AUTOFIB` is $b_J \approx 17$. `AUTOFIB` and BES meet LDSS and LDSS-2 at $b_J \approx 22$.

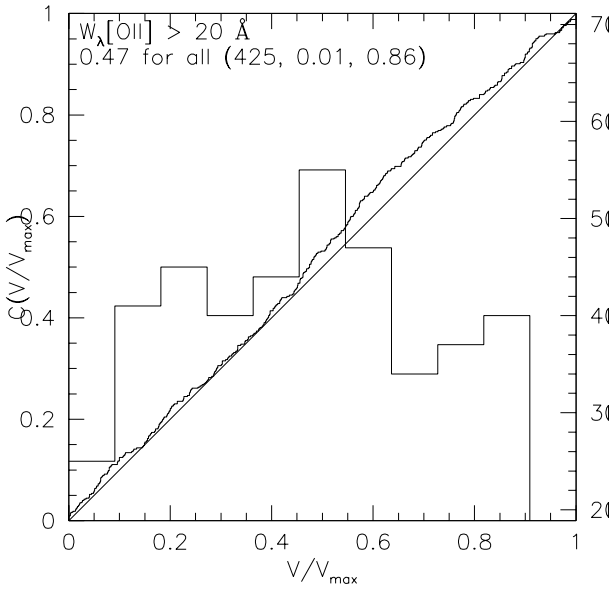

Figure 6.2: $V/V_{\rm max}$ Distribution for the High-equivalent-width Galaxies

3727. Although previous studies have used a more involved prescription (Cole et al. 1994b), I decided to simply to identify galaxies with W_λ[OII] greater than 20Å as "star-forming". Although the strength of the [OII] line depends on both the excitation and the oxygen abundance (*i.e.* it is *not* directly proportional to the star-formation rate), it provides a simple tracer of on-going star formation. Since the DARS survey is incomplete in regards to W_λ[OII] measurements, it is excluded from the luminosity-function analysis.

The galaxies with strong emission of [OII] 3727 are a nearly complete subsample of the `AUTOFIB` combined survey. Their mean $V/V_{\rm max}$ is nearly 0.5, the distribution is almost uniform (Figure 6.2). This isn't surprising as a galaxy with strong [OII] emission is easy to identify. Therefore, conclusions about this subset are not sensitive to systematic incompleteness within the survey.

From Figure 6.3, it is apparent that star-forming galaxies have become both fainter and less numerous from redshifts of $z \sim 0.5$ to the present. The luminosity functions show no bright galaxies with high EWs in the local luminosity function. This is simply because the EWs of the bright DARS catalogue were excluded from the luminosity-function analysis.

This luminosity-function analysis implies that the trend of median EW as a function of apparent magnitude may result from an underlying trend in the median EW as a function of redshift and absolute magnitude. Here, it is possible to include the DARS measurements in the calculation of the median EW, if one assumes that whether or not an equivalent width was

measured did not depend on the galaxy's EW, which appears to be the case.

The distribution of EWs changes markédly with redshift and absolute magnitude. Although the DARS EW measurements are included in this analysis, there is some concern about aperture effects. Since the DARS galaxies are much closer than the AUTOFIB BES, and LDSS data, the physical aperture is much smaller, centred on the nucleus. In the fainter data, the fibre or slit effectively integrates over much of the galaxy. Therefore, the DARS equivalent widths may be systematically underestimated (Figure 6.1 and a comparison with the Kennicutt (1992b) sample show that this is not a strong effect). Regardless, the DARS data contributes mainly to the bright nearby galaxies – the extreme front-left corner of Figure 6.4.

To analyse the EW distribution, the galaxies are binned by absolute magnitude (in one-magnitude bins) and by redshift (in bins of one tenth). Within each bin the median EW is determined, reducing the influence of galaxies with anomalously high EWs. Between the bin centres, the median EW is interpolated bilinearly, resulting in Figure 6.4.

The expected equivalent width of a galaxy depends strongly on its redshift and absolute magnitude. At all redshifts, fainter galaxies tend to have higher EWs. This effect mirrors the trend in W_λ[OIII] 5007 noted by Koo & Kron (1992). Furthermore, at a given luminosity, the median equivalent width increases with redshift (also noted by Koo & Kron (1992)). These two effects conspire to give a strong dependence of the median EW observed for a sample on the magnitude limit. At a redshift of 0.5, even L^* galaxies have [OII] emission nearly as strong as found in today's dwarf galaxies. Does this increase in median EW occur because more late-type galaxies are included in the high-redshift samples or do the EWs within each class of galaxies evolve?

6.2 Analysis by Spectral Classification

Dividing the survey by spectral classification can help us to understand this effect further. Since the number of galaxies in each subsample is much smaller than in the composite survey, I reduce the amount of data required from each subsample. The sample will be divided into elliptical, early-spiral and late-spiral galaxies. For each subsample, the mean coadded spectra, median EW, and luminosity function as a function of redshift will be calculated. However, for these subsurveys, the sampling of the luminosity function will be much coarser. A parallel analysis of the luminosity function will be to determine the evolution of the Schechter parameters with redshift using the generalised STY method derived in Equation 4.7. To summarise the evolution of the luminosity function, I found the evolving model that maximised the likelihood of observing to each subsample:

$$\begin{aligned} \phi^*(z) &= \phi_0^*(1+z)^{\phi_z^*} \\ M^*(z) &= M_0^* - 2.5L_z^*\log_{10}(1+z) \end{aligned} \tag{6.1}$$

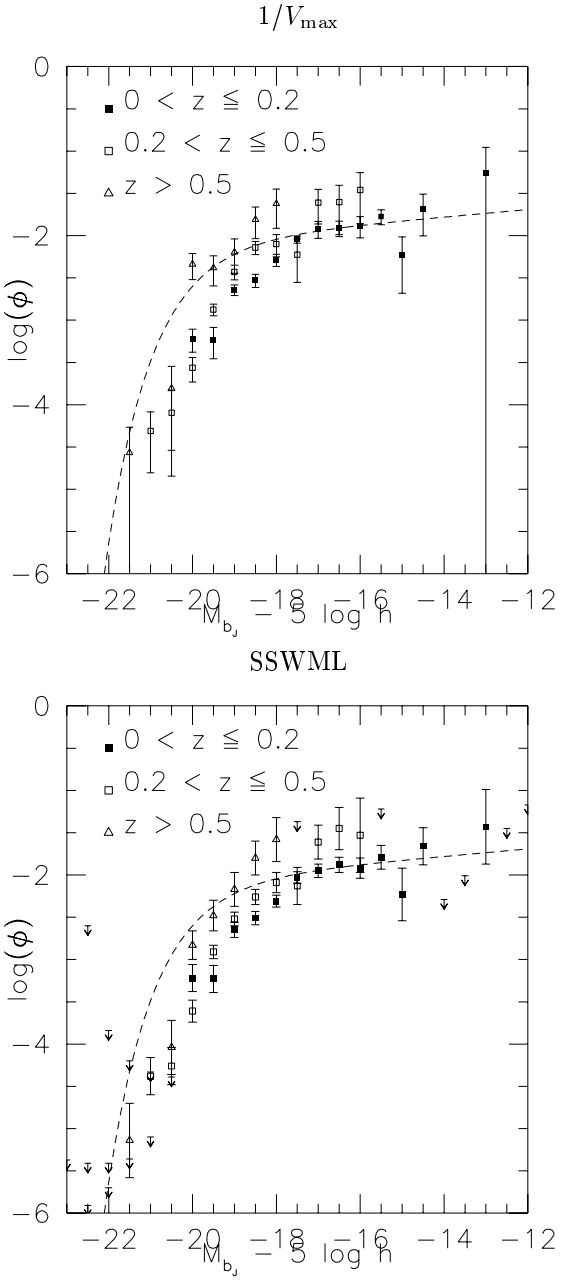

Figure 6.3: Star-forming Luminosity Functions

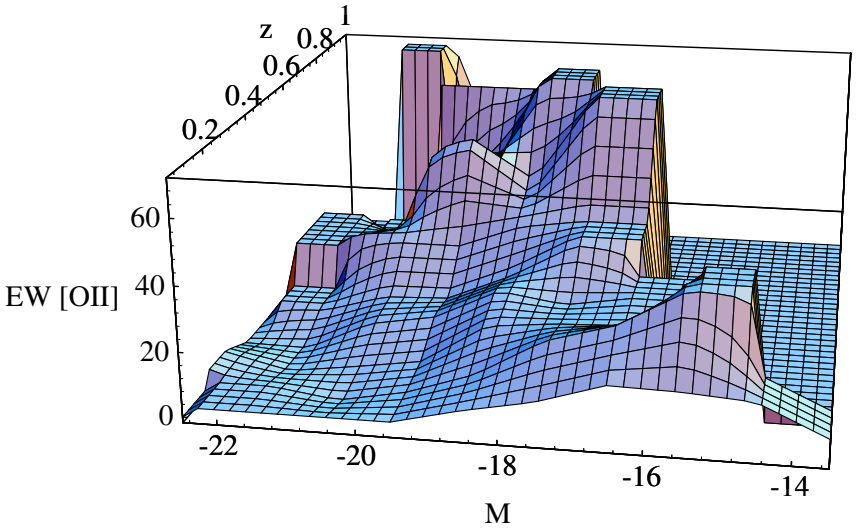

Figure 6.4: Evolution of the Median W_λ[OII].

$$\alpha(z) \quad = \quad \alpha_0 + \alpha_z z.$$

Both the normalisation and cutoff luminosity evolve as a power of cosmic time: $\phi^* \propto t^{-1.5\phi_z^*}$ and $L^* \propto t^{-1.5L_z^*}$. Meanwhile, the faint-end slope evolves linearly with redshift. As with a non-evolving Schechter function, the parameters are highly correlated. The density of galaxies is given by

$$\phi(L,z) = \phi_0^*(1+z)^{\phi_z^* - L_z^*(\alpha_0+\alpha_z z)} \left(\frac{L}{L_0^*}\right)^{\alpha_0+\alpha_z z} \exp\left(\frac{-L}{L_0^*(1+z)^{L_z^*}}\right). \quad (6.2)$$

For example, if the faint-end slope does not change with redshift ($\alpha_z = 0$), ϕ_0^* and the product $\alpha_0 L_0^*$ jointly determine the density of faint galaxies as a function of redshift. The six parameters are highly correlate making the errors difficult to interpret (imagine a six-dimensional error ellipsoid); therefore, the evolving parameters derived will be taken a guideposts elucidating the evolution apparent in the SSWML and $1/V_{\max}$ luminosity functions, not as definitive models for galaxy evolution. Although the parameters fit maximise the likihood of observing the survey, several sets of these parameters will result in nearly identical evolution of the luminosity function within the absolute magnitude and redshift range probed.

For this analysis, the survey is divided by k-correction class. Therefore, each of the subsamples described earlier will have two sets of these parameters. The two luminosity functions may be added to find the evolution of the entire subsample, which may be compared with the directly derived luminosity functions. Within each subsample, the spectra were coadded in the rest-frame (using the `squish` utility described in Section A.1.3) with

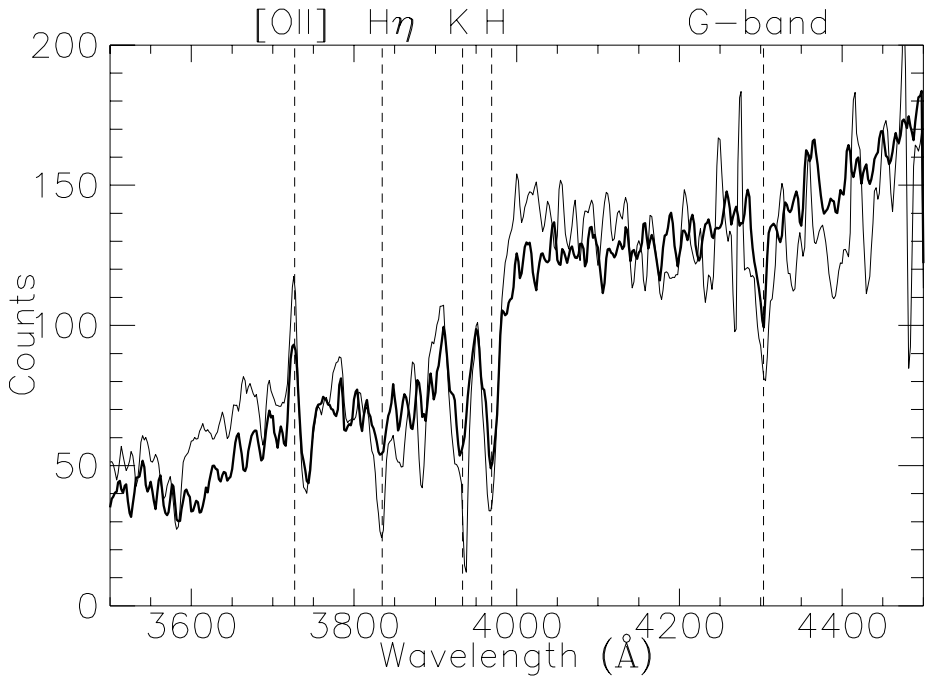

Figure 6.5: Coadded Spectra for Elliptical Galaxies. All the galaxies with elliptical spectral classifications have been coadded in two groups. The bold curve is the coaddition of all ellipticals with redshifts less than or equal to 0.2. The light curve is of all ellipticals with $0.2 < z \leq 0.5$. Both spectra have been smoothed on a scale of 10Å.

normalisations that were allowed to vary so as to minimise the mean dispersion between the various spectra. Since each spectrum covers a slightly different wavelength, this is not exactly equivalent to normalising the total flux in each spectrum, but this difference only subtly affects the coadded spectra in the range of interest.

6.2.1 Elliptical Galaxies

The elliptical galaxies are a small subset of the combined survey (312 of 1585 classified galaxies) and contribute approximately one-tenth of the volume density of intrinsically faint galaxies and one-half of the galaxies with $L \sim L^*$. Figure 6.5 compares the mean spectra of the elliptical galaxies with $0 < z \leq 0.2$ with those $0.2 < z \leq 0.5$. The two spectra appear nearly identical, and compare well with their local exemplars in the Kennicutt atlas (Kennicutt 1992a), indicating that the spectral classification algorithm performed well.

There are subtle differences between the two spectra, which probably arise from instrumental effects. The equivalent widths of the absorption

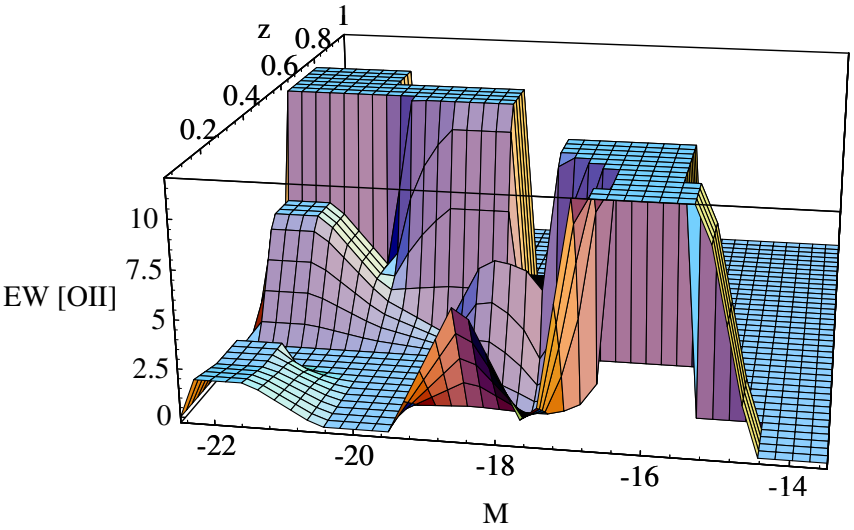

Figure 6.6: Evolution of the Median W_λ[OII] for Ellipticals.

Table 6.1: The Evolution of the Schechter Function: Elliptical Types

Type	$\phi_0^*(10^{-3}\mathrm{Mpc}^{-3})$	ϕ_z^*	M_0^*	L_z^*	α_0	α_z
Red Elliptical	3.38	-17.02	-20.0	5.41	-0.43	-4.82
Blue Elliptical	2.17	0.19	-19.5	0.19	-0.95	0.94

lines are consistent larger in the high-redshift coadded spectrum. This may be explained if the sky in this wavelength range was consistently overestimated when observing these faint galaxies. This reduces the strength of the continuum and increases the ratio between its and the absorption lines.

Figure 6.6 examines the strength of the [OII] feature in the elliptical-class galaxies further. The median equivalent width does increase slightly with redshift but not enough to explain the strong evolution show in Figure 6.4. This increase probably results from a combination of the sky-subtraction errors and increased fraction of elliptical galaxies with emission lines (E+A galaxies) at moderate redshift (*i.e.* evolution).

The luminosity function of elliptical galaxies with the survey also changes little with redshift. The luminosity function for $0.2 < z \leq 0.5$ is identical within the errors to the low-redshift result. So few elliptical galaxies are included in the survey at $z > 0.5$ that it is impossible constrain the luminosity function at high redshift.

The generalised STY (SSTY) method finds that the red and blue ellipticals evolve differently; however, few ellipticals constrain the luminosity function at high redshift. Although the two elliptical types at zero red-

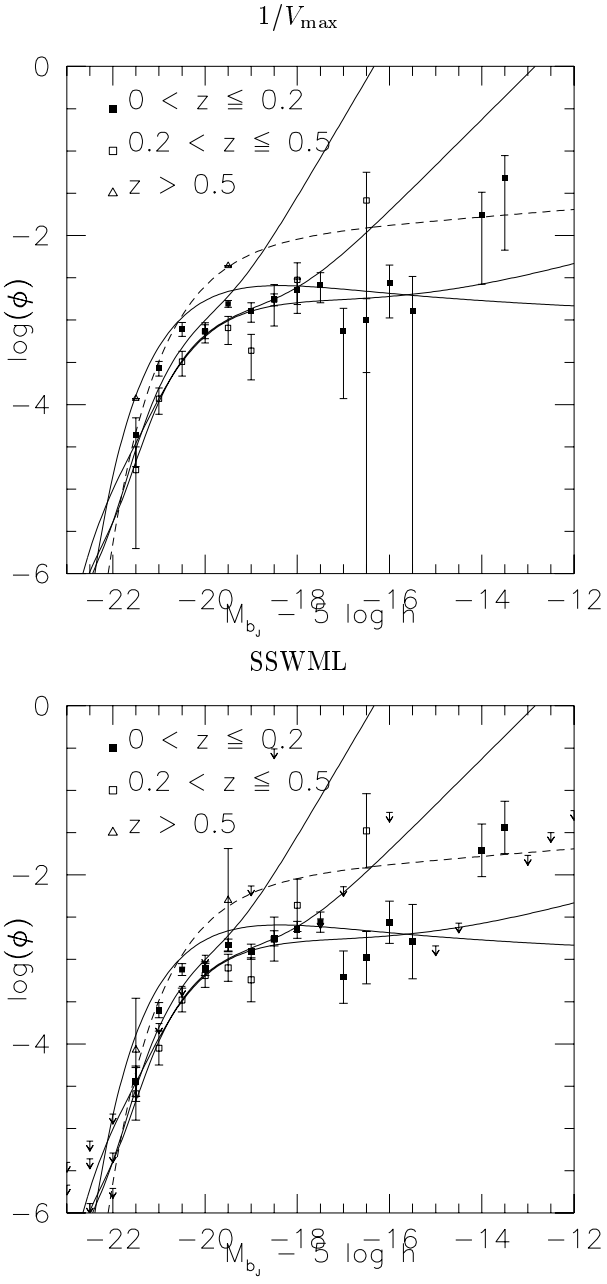

Figure 6.7: Luminosity Function of Elliptical Galaxies. The results of the generalised STY fitting for the elliptical types are superimposed for $z = 0.1$, 0.3 and 0.5.

Table 6.2: The Evolution of the Schetcher Function: Early-Spiral Types

Type	$\phi_0^*(10^{-3}\mathrm{Mpc}^{-3})$	ϕ_z^*	M_0^*	L_z^*	α_0	α_z
Sab	1.71	0.70	-19.9	-1.07	-1.19	0.79
Sbc	1.12	6.39	-20.0	-2.12	-1.42	-0.02

shift both have luminosity functions shallower than the total luminosity function (in agreement with the Loveday et al. 1992 results), the LF for the red ellipticals increases in magnitude and slope, becoming effectively a power-law by $z \sim 0.4$, while the slope of second elliptical type decreases with redshift with a nearly constant normalisation and cutoff luminosity. The red elliptical luminosity function performs contortions to fit the three elliptical galaxies with $z > 0.3$ which are essentially outliers in the galaxy distribution. Therefore, the decreasing luminosity function exhibited by the blue ellipticals probably provides the best estimate for the evolution of the elliptical luminosity function. Although the number of L^* elliptical galaxies has been constant over the past few billion years, there were far fewer faint ellipticals in the past – negative evolution.

6.2.2 Early Spiral Galaxies

Early spirals contribute nearly 40% of the classified galaxies (616 of 1585). They contribute about one-half the local density of galaxies for luminosities probed by the combined survey. The analysis proceeds as for the elliptical galaxies. Figure 6.8 compares the coadded spectrum for the early spirals with $z \leq 0.2$ and $0.2 < z \leq 0.5$. A similar effect to that found in the ellipticals is evident here. The absorption lines are weaker relative with respect to the continuum in the high-redshift bin than locally. The strength of [OII] 3727 changes little between the two spectra; the equivalent width actually decreases slightly with redshift which is reflected in Figure 6.9. The difference between the slopes of the two spectra at the blue end results from the response curve of the spectrograph which decreases quickly blueward of 3600 Å– affecting the low- redshift galaxies only.

Like the coadded spectra, the median W_λ[OII] for the early spiral galaxies evolves little with redshift. The median strength of [OII] decreases slightly with redshift before increasing to a peak value of approximately 30 Å, falling far short of the increase in the entire survey. The luminosity function for the early spiral galaxies appears nearly as constant as the luminosity function for elliptical galaxies.

The evolving Schechter parameters for the early sprials tell two stories. The Sab galaxies evolve similarly to the elliptical galaxies. The faint-end slope slowly decreases, while the normalisation slowly increases with redshift. As with the ellipticals, the number of L^* Sab galaxies has remained constant over the past few billion years. The later Sbc spirals exhibit an

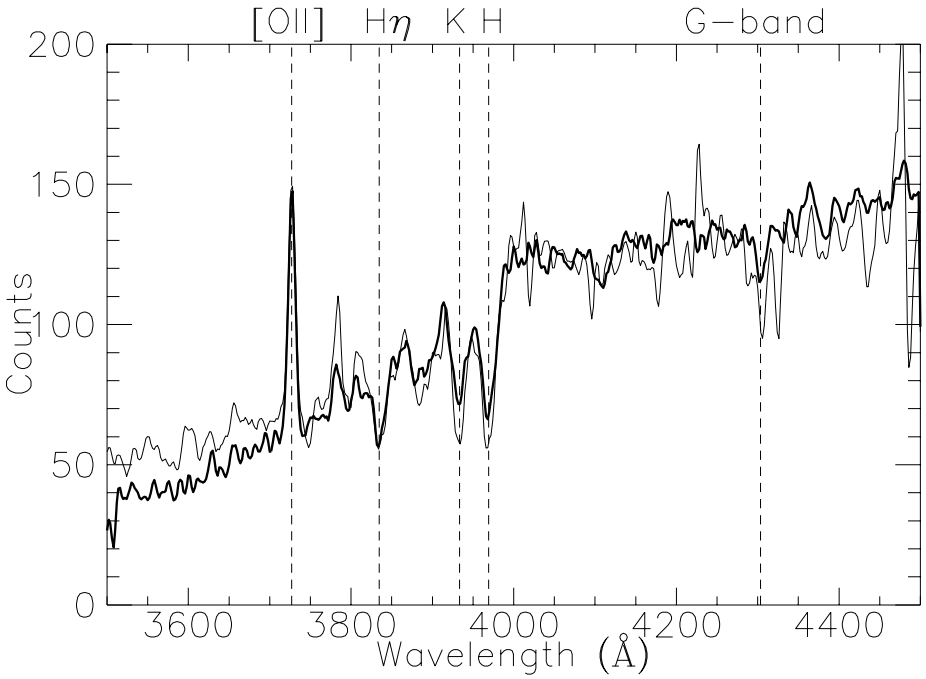

Figure 6.8: Coadded Spectra for Early Spiral Galaxies

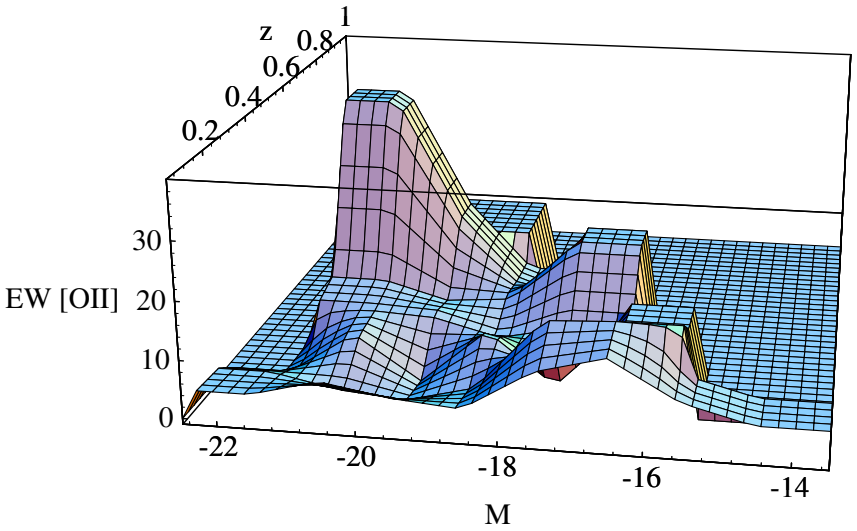

Figure 6.9: Evolution of the Median W_λ[OII] for Early Spirals.

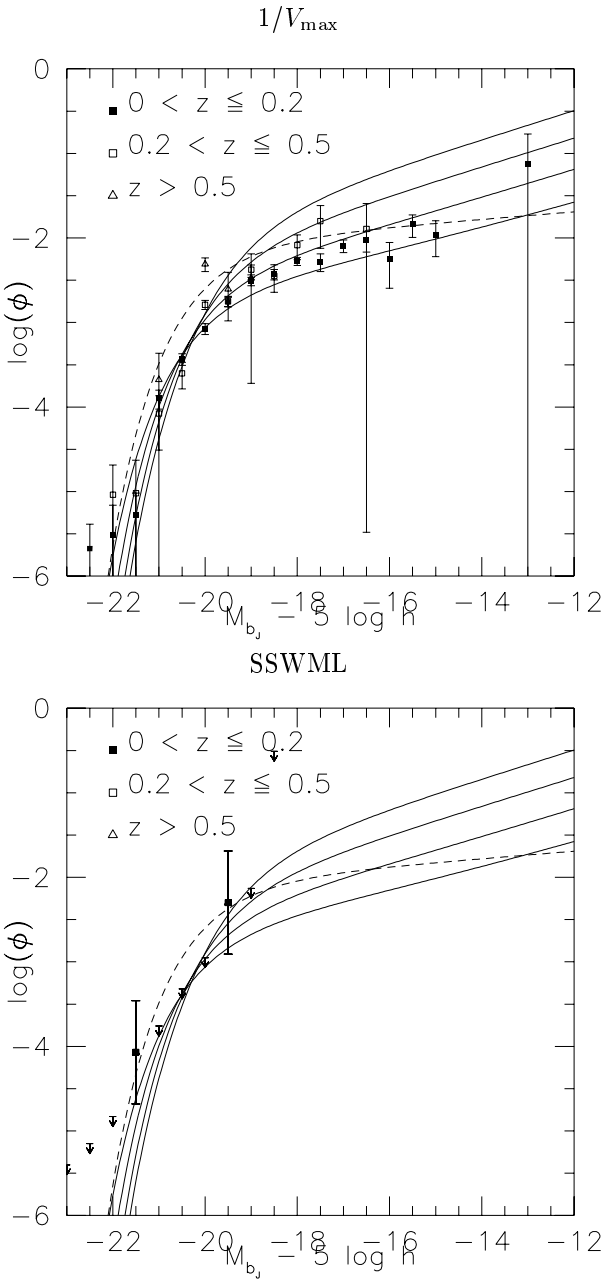

Figure 6.10: Luminosity Function of Early Spiral Galaxies. The results of the generalised STY fittng for the elliptical types are superimposed for $z = 0.1$, 0.3 and 0.5.

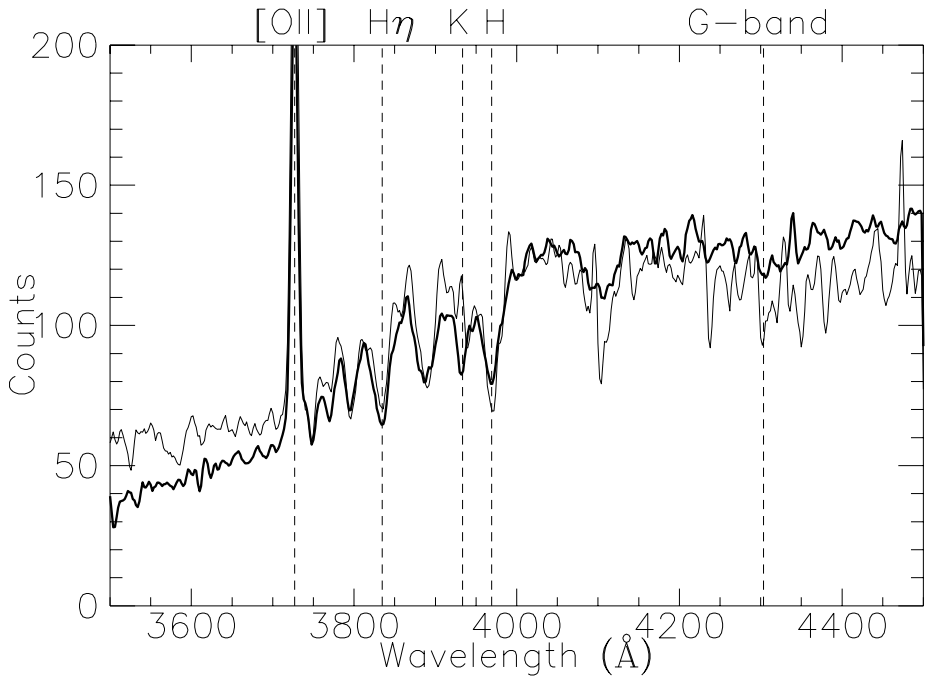

Figure 6.11: Coadded Spectra for Late Spiral Galaxies

opposing effect. Their density increases with redshift quickly without much change in the faint-end slope. The cutoff luminosity decreases quickly with redshift. As with the Sab galaxies, the number of galaxies with $M_{b_J} < -20$ remains more or less constant. The interplay between these two luminosity functions, yields a more constant (although increasing) combined luminosity for the early spirals. Perhaps as the Sbc galaxies age, they begin to appear more like the earlier Sab spirals, so the number of Sab galaxies increases at the expense of their later counterparts (or progenitors).

6.2.3 Late Spiral Galaxies

The evolution of the Sbc galaxies hints at the evolution of the late spiral galaxies. The two late-spiral spectral classifications also contribute about 40 % of the classified galaxies in the survey (657 of 1685). By process of elimination and extrapolation from the Sbc galaxies, these late galaxies must reflect the evolution of the total luminosity function, discussed in the previous chapter. The coadded spectra change markédly with redshift. W_λ[OII] increases whilst the absoption lines nearly disappear entirely, resulting in an emission-line galaxy by $z \sim 0.3$.

Figure 6.12 shows that the median equivalent width of the blue, late-type galaxies increases with redshift. Analysing these blue galaxies on their

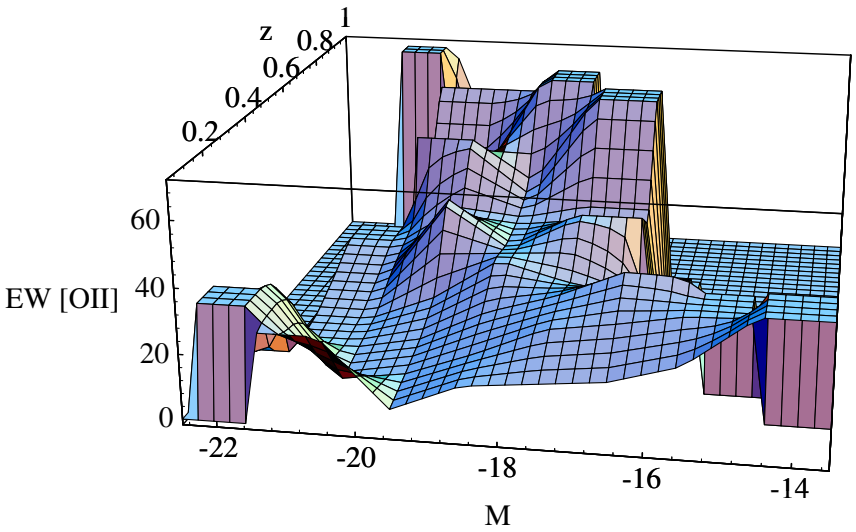

Figure 6.12: Evolution of the Median W_λ[OII] for Late Spirals.

Table 6.3: The Evolution of the Schechter Function: Late-Spiral Types

Type	$\phi_0^*(10^{-3}\mathrm{Mpc}^{-3})$	ϕ_z^*	M_0^*	L_z^*	α_0	α_z
Scd	2.82	2.45	-19.3	0.35	-1.33	-1.33
Sm/Starburst	2.88	-2.00	-17.9	3.14	-0.93	-1.15

own avoids the bias that fainter samples tends look at bluer populations because of the k-correction (Koo & Kron 1992). The blue galaxies themselves were forming stars more fiercely in the recent past than today.

The evolution of the Scd galaxies swamps the changes in the other types (see Figure 6.14). Their density, cutoff luminosity and faint-end slope all increase with redshift, so that they contribute a large fraction of the galaxies at all luminosities at moderate redshifts, and dominate the abundance of intrinsically faint galaxies. Although they were more common in the past (the increase in cutoff luminosity and faint-end slope overcomes the decrease in normalisation), the Sm/starburst-type galaxies appear always to have been fainter than normal galaxies and relatively rare. The prime actors in the evolution of the galaxy luminosity function are the late-spiral galaxies.

6.3 Discussion

To summarise the results for the six spectral types, Figure 6.14 depicts the evolution in the luminosity functions for each spectral type, and as a final

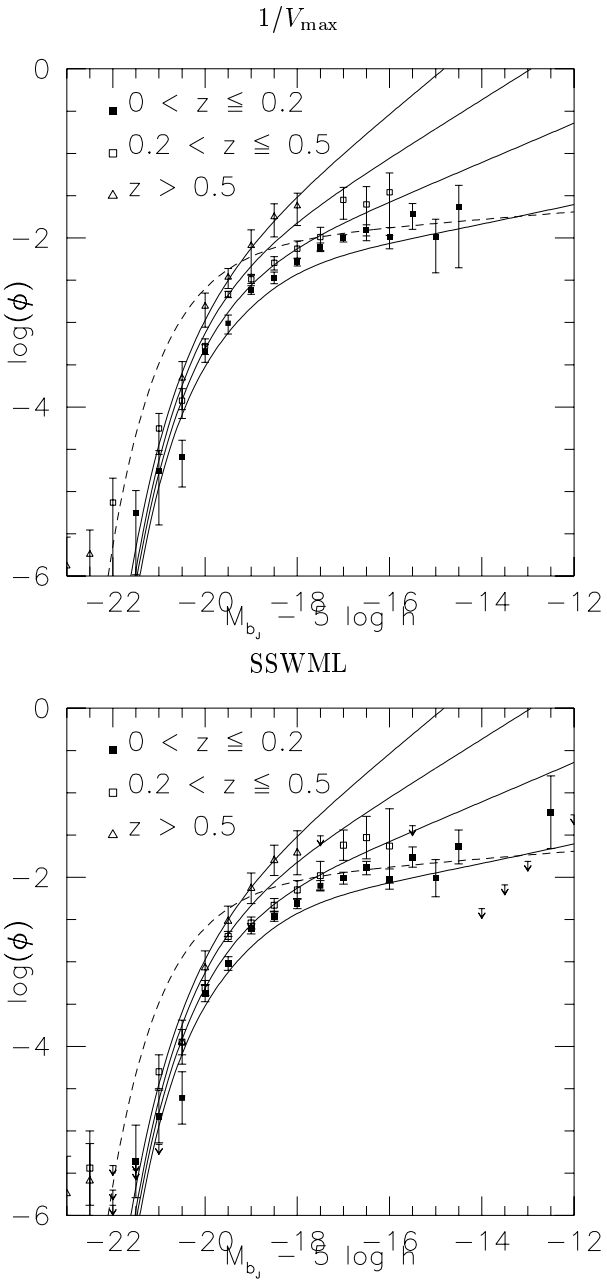

Figure 6.13: Luminosity Function of Late Spiral Galaxies. The results of the generalised STY fitting for the elliptical types are superimposed for $z = 0.1$, 0.3 and 0.5.

Table 6.4: The Evolution of the Schechter Function: Unclassified Galaxies

$\phi_0^*(10^{-3}\mathrm{Mpc}^{-3})$	ϕ_z^*	M_0^*	L_z^*	α_0	α_z
1.93	1.02	-18.4	3.49	-1.00	-1.11

check, these results are added together with the evolution inferred for the few unclassified galaxies (whose spectra were not available) to yield the total luminosity function illustrated in Figure 6.15. The SSWML results, first presented in Figure 5.6, are shown here again.

The properties of the unclassified galaxies are summarised in Table 6.4. Their density is a small fraction of that of all galaxies. Furthermore, their evolution is negligible, as would be expected if they were a random sampling of the classified types, rather than biased toward a particular galaxy type. The presence of these unclassified galaxies introduces little uncertainty in the conclusions as shown by Figure 6.15.

Many authors have postulated late-type galaxies as the prime movers in the observed evolution of the luminosity function, manifest in the number-magnitude relation. Their blue colours mean that they are affected least by k-corrections and can be observed to the largest distances for a given absolute magnitude. Broadhurst, Ellis & Shanks (1988) found that a luminosity function with a constant cutoff luminosity and a faint-end slope which increases with redshift is consistent with both the observed number counts and redshift distribution. Lacey & Silk (1991) and Treyer & Silk (1993) both concur with this finding, and add that this increase is driven by blue, late-type galaxies which have simply disappeared since a redshift of about 0.2. Lacey & Silk (1991) proposes several avenues for their disappearance. These excess galaxies could have produced more high-mass stars proportionally than modern galaxies. This bias toward high-mass stars would make them brighter while stars were forming, but the galaxies would quickly fade after the high-mass stars went supernova, leaving few stars to be observed today. These star-forming galaxies could have subsequently destroyed themselves as violent supernovae blew out their gas reservoirs and unbound the stars. A final explanation is that these galaxies have merged into the galaxies that we observe today. White (1990) and Efstathiou (1990) argue that such an evolutionary history is a natural consequence of the turnaround of larger and larger mass scale with time in a hierarchical universe. However, large amounts of merging would leave tracers on today's galaxies (*e.g.* Toth & Ostriker 1992). Furthermore, these excess galaxies are unlikely to be the progenitors of today's galaxies as they are more weakly clustered and more dense than galaxies today (Babul & Rees 1991).

From Figure 6.14 it is difficult to decide how the late-type galaxies disappeared so quickly. As the number of late spirals decreases toward the present day, the number of elliptical galaxies have increased. But could so

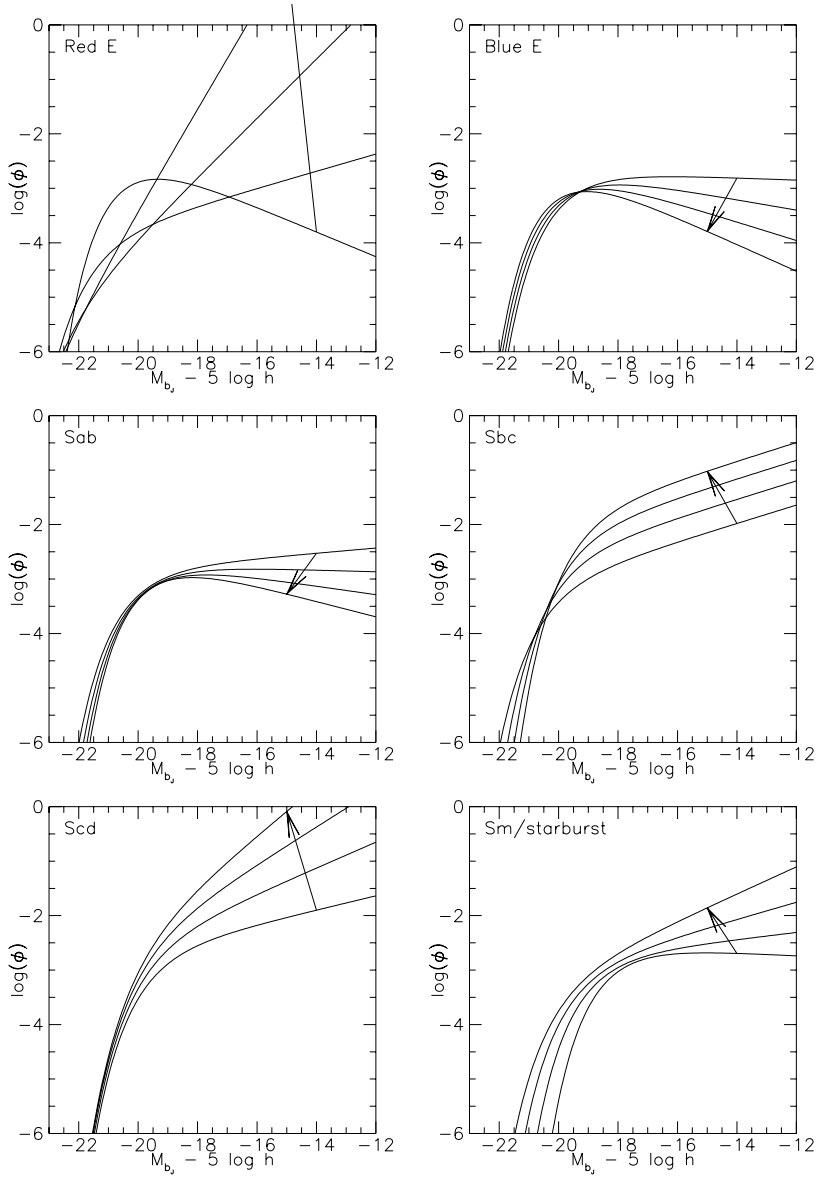

Figure 6.14: The Evolution of the Luminosity Function by Spectral Type. The curves trace the luminosity functions at $z = 0.1$, 0.3 and 0.5. The arrow points toward increasing redshift.

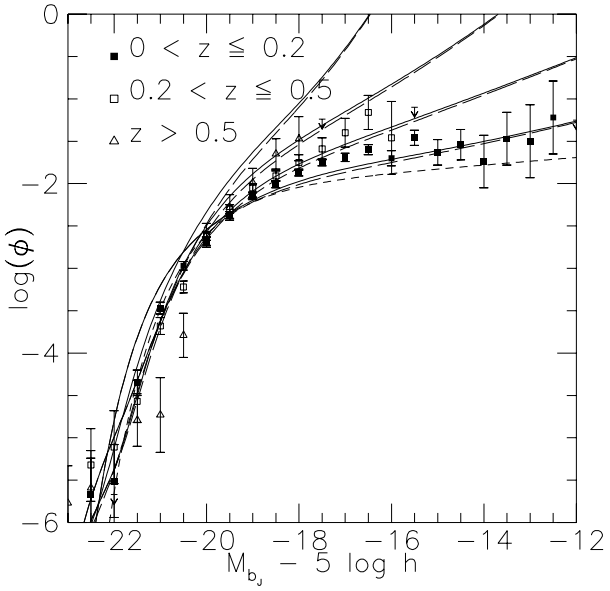

Figure 6.15: Comparison of the SSTY and SSWML Luminosity Functions. The sum of the luminosity functions for the six classifications and the unclassified galaxies (the solid curves) is compared with the luminosity function for all the galaxies in the survey as determined by the SSWML method and the Loveday result (short-dashed curve). Just below the solid curves are long-dashed curves, which show the evolution of the luminosity function with unclassified galaxies excluded from the analysis. The SSTY results are given for $z = 0.1$, 0.3 and 0.5.

many faint late-type galaxies hide in so few present-day ellipticals? Babul & Rees (1991) propose an alternative explanation (elucidated further in Efstathiou 1992). Although the small mass haloes ($\sim 10^9\,\mathrm{M}_\odot$) collapse and virialise before the nascent haloes of L^* galaxies, the UV flux produced by quasars may keep the gas in small haloes ionised until $z \sim 1$. Only then can these small galaxies begin to form stars. The stars would form quickly as in a starburst galaxy. During the starburst these galaxies would appear irregular, and their spectra would be most similar to those of present-day late-type spirals. Supernovae would blow out the gas slowing star formation, and the galaxy would begin to fade (more quickly in the B-band than in the K-band). Babul & Rees (1991) also proposed that in low-pressure environments the gas may escape the galaxy entirely, while in intermediate and high-pressure regions, some of the gas would return to the galaxy, possibly fueling further starbursts. The few galaxies that could still be observed today would be in the high-pressure regions clustered near more luminous galaxies. Gigayears of phase mixing could transform these irregular starburst galaxies into today's population of dwarf elliptical galaxies. However, the vast majority of these dwarf ellipticals would fade below detection limits. This explanation appear consistent both with the increasing abundance of early-type galaxies and the decreasing numbers of late-type galaxies over the past few billion years.

Part II

Galaxy Formation Simulations

Chapter 7

Introduction to the Block Model

'listen: there's a hell
of a good universe next door; let's go'
— e.e. cummings, *pity this busy monster, manunkind*

'Had I been present at the Creation, I would have given some useful hints for the better ordering of the universe.'
— Alfonso the Wise, King of Castile

SUMMARY

We predict observable properties of the galaxy population in several hierarchical models, using the "block" model of structure formation. We contrast the standard CDM cosmogony with CDM models with a low value of H_0, a low value of Ω and a low-density, flat universe ($\Omega + \Lambda = 1$). Furthermore, we compare galaxy formation in these CDM universes with the CHDM "mixed dark matter" model. To determine the numerous cosmological parameters in these models, we look to recent observations of large-scale structure, light-element abundances, and globular cluster ages, and to fix the astrophysical arguments of the "block" model, we attempt to find the best-fit to the present-day observed B-Band luminosity function. The result is several independent diagnostics of these models: the K-band luminosity function, the infrared Tully-Fisher relation, $B-K$ colours, number counts and redshift distributions.

7.1 Introduction

Studies of galaxy formation have progressed steadily over the past few years on three broad fronts: observations, numerical simulations and semianalytic treatments. Photometric and spectroscopic data on faint galaxies and

quasars suggest that the process of galaxy formation may be accessible to observation with existing techniques. Intense protogalactic activity seems to be occurring at redshifts $z = 1 - 3$. At this epoch, the amount of neutral hydrogen present in damped Lyman-α systems is comparable to the present mass density in stars, suggesting that these clouds might contain the raw material for most of the stars seen in galaxies today (Lanzetta, Wolfe & Turnshek 1994) The abundance of quasars peaks in this redshift interval, signalling strong evolutionary processes (Green 1989, Boyle et al. 1990, Hewitt, Foltz & Charree 1993). The total flux from faint blue galaxies in deep CCD counts implies intense star formation activity which may account for the production of a substantial fraction of the heavy element content of galaxies and perhaps also for a similar fraction of their stellar content. Although the redshift range at which this flux is emitted is still undetermined, it is quite possibly near $z = 1$. Even at redshifts less than 1, there appear to be symptoms of ongoing galaxy formation, manifest in the seemingly rapid evolution in the luminosity function and the colours of galaxies, as well as their mix of morphological types (this work, Cowie et al. 1988, Lilly 1993, Ellis et al. 1994; Butcher & Oemler 1978).

Theoretical studies of galaxy formation have also progressed at a rapid rate, as semianalytic models and numerical simulations of the relevant gravitational, gas dynamical and radiative processes in the universe become increasingly sophisticated. Although N-body/gas dynamical simulations of large cosmological volumes do not yet have sufficient resolution to follow galaxy formation in detail, simulations of small volumes and of individual objects have shown that in hierarchical clustering theories, galaxies are assembled through mergers of dark matter halos within which gas rapidly cools and condenses, in the manner envisaged by White & Rees (White & Rees 1978, Katz, Hernquist & Weinberg 1992, Cen & Ostriker 1993, Katz, Hernquist & Weinberg 1992, Evrard, Summers & Davis 1994, Navarro & White 1993, Navarro, Frenk & White 1994). The detailed mode and epoch of galaxy formation depends on the nature of the assumed primordial fluctuations. For example, in the standard cold dark matter (CDM) cosmogony, the paradigm of hierarchical clustering, galaxy formation activity peaks at relatively recent epochs, $z \lesssim 2$ (Davis et al. 1985, Frenk et al. 1988).

Semianalytic models provide a powerful tool to explore the validity of various physical assumptions and simplifications, and have been considerably extended in several recent papers (Cole 1991, White & Frenk 1991, Lacey & Silk 1991, Kauffmann, Guiderdoni & White 1994, Cole et al. 1994a). The goal is to construct "ab initio" models in which the growth of structure arising from an assumed spectrum of primordial density perturbations is represented by a set of simple rules which encapsulate our current understanding of gravitational clustering, radiative hydrodynamics, star formation and feedback, and the ageing of stellar populations. The outcome is a set of predictions for the observable properties of the galaxy population – abundances, luminosities, colours, circular velocities – as a function of time. Although the detailed implementation of the rules differs somewhat

in different studies, there seems to be general agreement regarding the difficulties that models such as standard CDM face in order to reproduce the observations.

The successes and failures of models based on the standard CDM cosmogony are summarised in Cole et al. (1994a). They showed that in order for a CDM model to be successful, star formation must be strongly regulated by feedback from supernovae and evolving stars and galaxy mergers must play a central role. With these ingredients, it is possible to construct a "fiducial model" that predicts B-band and K-band luminosity functions in general agreement with observations (although the faint-end slopes are slightly steeper than observed in the field); acceptable mass-to-light ratios; a wide range of galaxy colours and a colour-luminosity relation with the correct sign; star formation rates similar to those observed; and faint number counts and associated redshift distributions in excellent agreement with observations. However, this model also suffers from two major shortcomings:

- it does not produce galaxies as red as many observed ellipticals (by about 0.3 magnitudes in $B-K$) and
- the zero-point of the $I-$band "Tully-Fisher" relation – the correlation between the I-magnitude and the circular velocity of disk galaxies is predicted to be about 2 magnitudes too faint.

The first problem can be traced to the fact that standard populations synthesis models require more time to generate sufficiently bright red stars than is available in the fiducial CDM model. The second problem arises because a CDM universe contains an excessive number of dark galactic halos of size comparable to that of the Milky Way.

The inability to produce a fully successful model of galaxy formation may be due to an incorrect choice of cosmological parameters or to inadequacies in the modelling of the physics of galaxy formation. The aim of this exercise is to explore the first of these possibilities. Thus, we retain the basic astrophysical framework of Cole et al. (1994a) (although we reserve the freedom to adjust parameters appropriately), and apply it to a variety of alternative cosmological models. In an attempt to remedy the colour problem, we consider models with a longer timespan than CDM (by lowering the values of H_0 and Ω_0) and, in an attempt to remedy the Tully-Fisher discrepancy, we consider models that produce a lower abundance of dark galactic halos (by lowering Ω_0 or by assuming a mixture of cold and hot dark matter). Our main result is that none of these alternatives provides a satisfactory resolution to the problems affecting the fiducial model and, in many cases, they do not even share some of its successes. We are therefore led to the conclusion that some of the astrophysical processes included in our model require revision. The remainder of this part is organised as follows. The rest of this chapter describes the recipe for galaxy formation developed in Cole et al. (1994a) and introduces the cosmological models studied. Chapter 8 describes a new method for deriving number counts,

redshift distributions and several other observational data from numerical simulations, and the final chapter of this part describes the results of the calculations and their implications.

7.2 The Method

The factors that affect the formation of galaxies and the appearance of the population of galaxies today split into two areas: cosmology and astrophysics. The cosmology determines the sites in which galaxies can grow and the duration of their growth until we observe them today. The astrophysics of star formation, stellar evolution, gas dynamics and galaxy mergers (amongst other processes) affects the evolution of the galaxies within their nascent dark-matter halos.

7.2.1 Astrophysical Parameters: Mergers and Star Formation

The method we use to model the formation and evolution of galaxies is laid out in detail in Section 2 of Cole et al. (1994a). Here we summarise the main features of this modelling procedure and define the parameters that specify our model of star formation and the merging of galaxies within a common dark matter halo.

We follow the dynamical evolution of the population of dark matter halos using the block model of Cole & Kaiser (1988; see also Cole 1991). This is an approximate Monte Carlo implementation of the analytic description of halo merging based on the extension of the Press-Schechter theory and developed by Bond et al. (1991), Bower (1991) and Lacey & Cole (1993). The only input to the block model is the linear power spectrum normalised to the present epoch and a density threshold, δ_c, calculated from the collapse of a uniform spherical overdense region (δ_c increases with redshift and is dependent on both Ω and Λ). This analytic description has recently been shown to be in good agreement with the evolution found in large fully nonlinear N-body simulations (Lacey & Cole 1994, Kauffmann & White 1993). The basic Press-Schechter formalism can be applied to models with $\Omega < 1$ as detailed in Lacey & Cole (1993) and to models with $\Lambda \neq 0$ in an entirely analogous manner. However, it is not directly applicable when a mixture of hot and cold dark matter (CHDM) is present, for relativistic neutrinos do not cluster along with the CDM on scales less than their Jeans mass. Therefore, for this model we use a constant threshold of $\delta_c = 1.686$ and adopt the evolving CHDM power spectrum as parameterised by Klypin et al. (1993). As shown by these authors, the Press-Schechter mass function determined in this manner provides a reasonable approximation to the mass function of dark halos found in N-body simulations.

Individual halos are modelled as isothermal spheres in which any diffuse gas present when the halo forms is shock heated to the virial temperature of the halo and has initially the same $\rho \propto r^{-2}$ density profile as the dark

matter. This allows us to compute the fraction of gas that can cool during the halo's lifetime by computing the radius at which the cooling time, calculated assuming primordial abundances, equals the halo lifetime. The lifetime of a halo is defined as the time elapsed since the formation of a halo until its merger with a larger structure. The gas that cools is assumed to settle on a galaxy at the centre of the halo where it can then begin forming stars. In our model, this galaxy can experience more than one episode of star formation, as further episodes may be triggered by galaxy mergers. The transformation of the cooled gas into stars is, in our model, a self-regulating process. Star formation rates are moderated by supernovae and evolving stars, which inject thermal and kinetic energy into the gas. This energy feedback may expel gas from the galaxy, and return it to the hot diffuse phase. The efficiency of this process is assumed to depend sensitively on the depth of the potential well in which the galaxy resides. Thus, the cool gas reservoir is continuously depleted by both the transformation of gas into stars and the reheating of gas by supernovae. We assume that the star formation rate, $\dot{m}_\star(V_c,t)$, is proportional to the current mass of cool gas, $m_c(t,V_c)$,

$$\begin{aligned}\dot{m}_\star(t,V_c) &= m_c(t,V_c)/\tau_\star(V_c) \\ &= [m_c(0,V_c) - m_\star(t,V_c) - m_{\rm hot}(t,V_c)]/\tau_\star(V_c), \end{aligned} \tag{7.1}$$

where $m_{\rm hot}(t,V_c)$ is the mass of cooled gas reheated by the energy released from supernovae that is returned to the hot phase and $m_\star(t,V_c)$ is the mass of stars formed at time t after the onset of this episode of star formation. ($m_c(0,V_c)$ is the total amount of gas that can cool in the lifetime of the halo.) We further assume that the mass of gas reheated is proportional to the mass of stars formed

$$\dot{m}_{\rm hot}(t,V_c) = \beta(V_c)\dot{m}_\star(t,V_c) \tag{7.2}$$

The time scale, $\tau_\star$, and the ratio of the mass gas returned to the hot phase to the mass of stars formed, β, are both assumed to depend only on V_c. Hence

$$m_\star(t,V_c) = \frac{m_c(0,V_c)}{1+\beta}\,[\,1-\exp(-(1+\beta)\,t/\tau_\star)\,]\,, \tag{7.3}$$

We parameterise $\tau_\star(V_c)$ and $\beta(V_c)$ as simple power laws;

$$\tau_\star(V_c) = \tau_\star^0 \left(\frac{V_c}{300\,{\rm km\,s^{-1}}}\right)^{\alpha_\star} \tag{7.4}$$

$$\beta(V_c) = (V_c/V_{\rm hot})^{-\alpha_{\rm hot}}. \tag{7.5}$$

The four parameters $\alpha_{\rm hot}$, $V_{\rm hot}$, $\alpha_\star$, and $\tau_\star^0$ then specify completely our description of star formation. The simulations of Navarro & White (1993) suggest that the values of all these parameters depend only on the strength of the feedback as parameterised by their variable f_v (see their Table 2). The dependences of $\tau_\star(V_c)$ and $\beta(V_c)$ on f_v and V_c and the corresponding

Table 7.1: Astrophysical Parameters

Model	Υ	$\tau^0_{\rm mrg}/\tau_{\rm dyn}$	$\alpha_{\rm mrg}$	$\tau^0_*/{\rm Gyr}$	α_*	$V_{\rm hot}/\,{\rm km\,s^{-1}}$	α_h
Fiducial	2.7	0.5	0.25	2.0	-1.5	140.0	5.5
low-H_0	2.0	0.5	0.25	2.0	-1.5	140.0	5.5
low-Ω_0	3.0	2.0	0.25	2.0	-1.5	140.0	5.5
$\Omega_0 + \Lambda_0$	2.5	2.0	0.25	2.0	-1.5	140.0	5.5
CHDM	1.0	3.0	0.25	2.0	-1.5	140.0	5.5

values of $\alpha_{\rm hot}$, $V_{\rm hot}$, $\alpha_\star$ required to fit these dependences can be found in Figure 2 and Table 1 of Cole et al. (1994a).

The star formation histories computed for each galaxy according to the above prescription are converted into luminosities and colours using the stellar population synthesis model of Bruzual & Charlot (1993). Here we adopt the Scalo (1986) IMF for luminous stars with masses $0.1 < M/\,{\rm M}_\odot < 125$. The mass in non-luminous brown dwarfs with masses $M < 0.1\,{\rm M}_\odot$ is characterised by a further parameter Υ, defined to be the ratio of the total mass in stars to that in luminous stars.

The fate of galaxies whose halos merge is determined by a merger timescale $\tau_{\rm mrg}$. If $\tau_{\rm mrg}$ is shorter than the lifetime of the newly formed common halo then we merge the two galaxies, whereas if $\tau_{\rm mrg}$ is longer than the halo lifetime the galaxies remain distinct as either a dominant galaxy and a satellite or simply as members of a cluster or group of galaxies. Galaxy mergers within hierarchically growing halos have been studied by Navarro, Frenk & White (1994). They find that the probability of a merger, and hence the appropriate value of $\tau_{\rm mrg}$, depends sensitively on the angular momentum of the galaxy's orbit, but also increases with increasing galaxy mass as expected from simple consideration of dynamical friction. Hence we parameterise this merger timescale as

$$\tau_{\rm mrg} = \tau^0_{\rm mrg}(M_{\rm halo}/M_{\rm sat})^{\alpha_{\rm mrg}} \tag{7.6}$$

where $M_{\rm halo}$ is the mass of the newly formed common halo, $M_{\rm sat}$ the mass of the halo the satellite galaxy had prior to the halo merger and $\alpha_{\rm mrg} < 1$.

7.2.2 Cosmological Background

Once the astrophysical parameters presented above have been chosen, our model of galaxy formation is fully specified by the choice of a cosmological model, as this specifies the age and density of the universe, the initial spectrum of density fluctuations, and their growth rate. For our purposes, a cosmological model is specified by six parameters: the Hubble constant, $H_0 \equiv 100h\,{\rm km\,s^{-1}Mpc^{-1}}$; the total present mass density of the universe as well as those in baryons and relativistic particles (Ω, $\Omega_{\rm b}$, Ω_ν,respectively),

Table 7.2: Cosmological Parameters

	Cosmological Parameters					Constrained Quantities			
Model	Ω_0	Λ_0	h	σ_8	Ω_b	Γ	$\sigma_8\Omega^{0.6}$	$\Omega_b h^2$	$t_{\rm age}$/Gyr
Fiducial	1.00	0.00	0.50	0.67	0.06	0.5	0.67	0.015	13
low-H_0	1.00	0.00	0.25	0.67	0.20	0.25	0.67	0.0125	26
low-Ω_0	0.30	0.00	0.60	1.0	0.04	0.3	0.48	0.0144	13
$\Omega_0 + \Lambda_0$	0.30	0.70	0.60	1.0	0.04	0.3	0.48	0.0144	16
CHDM	1.00	0.00	0.50	0.67	0.06	–	0.67	0.015	13

all in units of the critical density; the cosmological constant Λ, in units of $3H_0^2$ (so that for a flat universe $\Omega+\Lambda = 1$); and the present linear amplitude of mass fluctuations in spheres of radius $8\,h^{-1}$Mpc, σ_8.

These parameters determine the properties of the cosmological model in a variety of ways. The *shape* of the power spectrum of linear density perturbations is determined by Ω, h and Ω_ν. If $\Omega_\nu = 0$ and the initial spectrum is that of Harrison-Zel'dovich, the shape of the post-recombination power spectrum is fully specified by the shape parameter $\Gamma = \Omega h$. The *growth rate* of perturbations depends mainly on Ω and Λ, although if $\Omega_\nu > 0$ the growth of fluctuations on small scales will be retarded. In an $\Omega < 1$ universe structure ceases to grow after a redshift $z \lesssim \Omega^{-1}$. This transition is similar, but more abrupt, when $\Lambda > 0$. Consequently, models with the same value of σ_8 (and hence the same present amplitude of fluctuations) will form galactic mass halos at higher redshift for low-Ω than for $\Omega = 1$. The spatial number density of these halos is also proportional to Ω, as for the same value of σ_8 these halos will contain some fixed fraction of the total mass. The *age of the universe*, $t_{\rm age}$, is proportional to H_0^{-1}, while for a given H_0 the age increases with decreasing Ω and increasing Λ. A modest increase in Ω_b can cause a large increase in the mass of stars formed in our models as the baryon fraction controls both the total amount of baryonic material available to form stars as well as the cooling of this material inside dark halos.

It is not feasible to present a thorough exploration of this wide parameter space. Instead, we have chosen to apply our galaxy formation framework to four new models, which we contrast with each other and with the fiducial model of Cole et al. (1994a). Three of the new models are variants of the CDM models in which H_0, Ω and Λ have been varied and the fourth is the CHDM "mixed dark matter" model advocated by Klypin et al. (1993) and Davis, Summers & Schlegel (1992). These four new models span the range of currently favoured cosmological models and serve to illustrate the effects of varying each of the cosmological parameters.

The parameters of these four new models are fixed by the following observational constraints:

1. The comparison of the galaxy peculiar velocities with the density field traced by IRAS galaxies implies $\Omega^{0.6}/b_{\rm IRAS} = 0.86 \pm 0.15$ (Kaiser et al. 1991), where $b_{\rm IRAS}$ is the bias parameter relating fluctuations in the density of IRAS galaxies to fluctuations in the underlying mass distribution. The correlation function of IRAS galaxies indicates that $b_{\rm IRAS}\sigma_8 = 0.58 \pm 0.14$ (*e.g.* Moore et al. 1994). Assuming that the bias parameter is independent of scale these combine to yield $\sigma_8\Omega^{0.6} = 0.5 \pm 0.15$. A very similar constraint is provided by the abundance of rich clusters which for spatially flat universes requires $\sigma_8\Omega^{0.56} = 0.57 \pm 0.05$ (White, Efstathiou & Frenk 1993).

2. Galaxy clustering on large scales as measured in the APM and IRAS galaxy surveys favour a spectrum with more large scale power than standard CDM, $\Gamma = \Omega h = 0.2$–0.3 (Maddox et al. 1990a, Efstathiou et al. 1990, Saunders et al. 1991, Feldman, Kaiser & Peacock 1994, Fisher et al. 1993).

3. Big bang nucleosynthesis (BBNS) limits on primordial light element abundances require $\Omega_{\rm b}h^2 = 0.0125 \pm 0.0025$ (Walker et al. 1991).

4. Recent estimates of the age of globular clusters require $t_{\rm age} \geq 13$ Gyr (Renzini 1986; Sandage 1993).

The parameters of the fiducial model and the four new models together with the values of these constrained quantities are shown in Table 7.2. The fiducial model uses the same cosmological parameters as standard CDM and therefore fails to satisfy the constraint on Γ. With the normalisation adopted here it also predicts cosmic microwave background fluctuations that are approximately 50% smaller in amplitude than those measured by COBE (Smoot et al. 1992). With the exception of low-Ω_0 the normalisation of all our new models are consistent with the COBE measurements.

With the cosmological and astrophysical parameters determined, the stage is set for the numerical simulations, but before presenting the results of these simulations and their implications, we will describe a new method of analysing numerical simulations.

Chapter 8

Bootstrapping a Galaxy Catalogue

SUMMARY

The generation of a redshift survey from numerical simulations provides an important link between theory and observations. I describe how a bootstrap resampling may be applied to this problem to generate number-redshift distributions, number-magnitude distributions and redshift surveys that mimic the errors and biases of the observations. The method integrates over the luminosity functions predicted by the simulations in their full detail using a Monte-Carlo technique; this adds no additional approximations, as the "block" model simulations (and N-body simulations in general) integrate the characteristic equations using the Monte-Carlo approximation. The test case of a non-evolving model without k-corrections effects is discussed and compared with the integrated solution.

8.1 The Problem

The statistical bootstrap, the cousin of the Quenouille-Tukey jackknife, was introduced by Efron in 1977 (Efron 1979). It most often makes its appearance in astronomy when estimating errors and biases in small samples especially those in the study of large-scale structure (Nobelis 1990; Bhavsar 1990); recently, bootstrap techniques have been applied to studies of Lyman alpha clouds (Press, Rybicki & Schneider 1993) and to the study of galaxy merger remnants (Heyl, Hernquist & Spergel 1994). Here, the resampling is used not to find errors, but to integrate over the curve threading through the distribution of galaxies in redshift and luminosity in a magnitude-limited sample.

The inputs for the technique are lists of the absolute magnitudes (in the observer's frame) of the galaxies realised in the simulations at a variety

of redshifts. The method resamples (hence the bootstrap) these lists with replacement to integrate the counts or generate a redshift survey. When deriving the number counts, the actual luminosities selected are not in the end important. They are simply the points used to evaluate the counts integral:

$$N(m) = \int_0^\infty \sum_i \rho_i(z, m - d(z) - k_i(z)) \frac{dV}{dz} dz \tag{8.1}$$

where $d(z)$ is the distance modulus and $k(z)$ is the k-correction. The sum is over the various types of galaxies to be included, each with its own luminosity function and k-correction curve. Here, the function $\rho(z, M_{\text{observer's}})$ is not known analytically but as a specific realisation. Furthermore, both $d(z)$ and dV/dz with a non-zero cosmological constant are given by additional integrals. One could fit an analytic function to the observer's band luminosity functions at the various redshifts; the k-correction term and the sum over galaxies are now unnecessary. However, the fitted function is both less general and less flexible than the original list – observers don't study smooth fitting functions, they count individual galaxies which may be missed or miscounted because of low surface brightnesses or other effects not included in the simple integral form (Equation 8.1). Including these additional effects is possible into an analytic treatment is possible but difficult and prone to error.

8.2 The Method

Here, I will describe a technique to include observational biases and to use the output of the galaxy formation simulation directly to calculate number counts and simulated redshift catalogues. The method is simply to select a galaxy from the distribution of galaxies in space and luminosity and calculate its apparent magnitude (considering the k-correction and surface brightness effects) and repeat. Since the luminosity function evolves with redshift, the first step is to choose the redshift of the galaxy and then once the redshift is known to determine from which redshift realisation of the luminosity function to select a galaxy.

The first step in determining the galaxy redshift is calculating the cumulative distribution of galaxies with redshift over the entire sky:

$$C(z) = \int_0^z n(z) \frac{dV}{dz} dz \tag{8.2}$$

where $n(z)$ is the total number density of galaxies at redshift z as predicted by the simulations. As the simulation only produces galaxies at discrete redshifts, this function must be interpolated. However, it is often quite smooth, and this interpolation introduces little error. This function is calculated only out to the maximum redshift that the brightest galaxy in the realisations could have been detected given the magnitude limits of the counts or redshift survey and is tabulated logarithmically in redshift.

Using the result that any distribution $f(z)$ may be translated to a uniform distribution

$$f(C(z)) = \text{constant} \tag{8.3}$$

by means of the cumulative distribution, a redshift is selected by selecting a uniform deviate from zero to the total number of galaxies out to the maximum redshift tabulated ($z_{\max}$), and again by interpolation, the cumulative distribution is inverted to give a redshift ($z_{\rm gal}$).

Once the redshift is known the luminosity of the galaxy must be found by selecting a galaxy from amongst the realisations. The two realisations that straddle the selected redshift are chosen, and the luminosity function in the observer's frame is assumed at the selected redshift to be a linear interpolation of the luminosity function at the two redshifts realised (z_a, z_b),

$$\phi(L, z_{\rm gal}) = \frac{z_b - z_{\rm gal}}{z_b - z_a}\phi(L, z_a) + \frac{z_{\rm gal} - z_a}{z_b - z_a}\phi(L, z_b). \tag{8.4}$$

Since neither of the luminosity functions are known analytically, we resort to a trick. We select a uniform deviate in redshift in the range $[z_a, z_b]$. If this number is greater than $z_{\rm gal}$, we use the realisation at redshift z_a, otherwise we use the other realisation.

Finally, to find the galaxy in question, we select randomly one galaxy from the appropriate realisation. Because this galaxy was simulated in a some sense, we have access to a variety of information possibly including circular velocity, rest-frame absolute magnitude, surface brightness, *etcetera*. Any of these parameters may be used to determine whether the galaxy ends up in the sample. Regardless of whether the galaxy is counted, we note that one galaxy has been selected, incrementing $n_{\rm selected}$ by one.

To determine whether the galaxy makes it into the catalogue, we calculate its apparent magnitude,

$$m = M + d_{\rm modulus}(z) + \Delta M(\Sigma, r_e, z...). \tag{8.5}$$

ΔM is a magnitude correction term to translate from total magnitude to isophotal magnitude or to add any other biases thought to be present in the observations. If m is in the apparent magnitude range of interest, the galaxy becomes part of the survey. In repeating this process, we build a magnitude-limited redshift survey from the simulations.

Many galaxy redshifts are selected from the cumulative distribution but few become part of the survey. However, the count of the number of redshifts selected gives the area of sky sampled in the survey:

$$\Omega_{\rm sampled} = 4\pi\frac{n_{\rm selected}}{C(z_{\max})} \tag{8.6}$$

Deriving the number counts proceeds similarly. As the error in the number counts is proportional to the square root of the number of galaxies counted, selecting a fixed number of galaxies from several narrow magnitude bins and dividing by the area sampled will predict a number-magnitude relation with constant relative errors.

As each galaxy is selected from simulations, a wealth of information is available about each object in the simulated survey. Galaxy formation simulations predict various correlations amongst galaxy properties (*e.g.* the Tully-Fisher relation). They give these relations for a volume-limited sample. Using this technique, these relations may be predicted for magnitude-limited samples. Although this resampling moves the predictions further from the underlying physical processes, the simulations now predict using a process similar to observation.

8.3 The Tests

A simple and quite rigorous test is to compare the number-magnitude relation predicted by this method with a similar relation derived by integrating directly,

$$N(m) = \int_0^\infty \phi(m - d_{\rm modulus}(z)) \frac{dV}{dz} dz. \tag{8.7}$$

First, we generate a list of 200,000 absolute magnitudes from a Schechter luminosity function,

$$\phi(L)dL = \phi_* \left(\frac{L}{L_*}\right)^\alpha e^{-L/L_*} d\left(\frac{L}{L_*}\right) \tag{8.8}$$

and

$$\phi(M)dM = 0.4 \ln 10 \phi^* 10^{-0.4(\alpha+1)(M-M^*)} \exp(0.4(M - M_*))dM \tag{8.9}$$

where as given by Loveday et al. (1992), $\phi^* = 1.01 \times 10^{-2} {\rm Mpc}^{-3}$, $\alpha = -1.11$, and $M^* = -19.73$. Next, we find the brightest and faintest galaxies in the realisation, and integrate the Schechter function between these limits to find the total density of galaxies in the luminosity range. Dividing 200,000 by this density gives the total volume of the realisation. This volume fixes the normalisation of the counts. Furthermore, since the realised luminosity function falls to zero outside this range, the integrand of Equation 8.7 must be slightly altered. This has little effect on the resulting number counts. Figure 8.1 shows the resulting number counts by numerical and bootstrap integration. For the bootstrap integration, 200 galaxies were sampled from 36 bins, each one-half magnitude wide, yielding a Poissonian one-sigma counting error of 7%. From the figure, it is apparent that the two methods agree within the Poisson error over the apparent magnitude range of 10 to 28. Higher accuracy may be achieved by increasing the number of galaxies per bin.

A second important test is deriving the $N(z)$ distribution for a magnitude-limited survey,

$$N(z) = \int_{m_{\rm min}}^{m_{\rm max}} \phi(m + d_{\rm modulus}(z)) \frac{dV}{dz} dm. \tag{8.10}$$

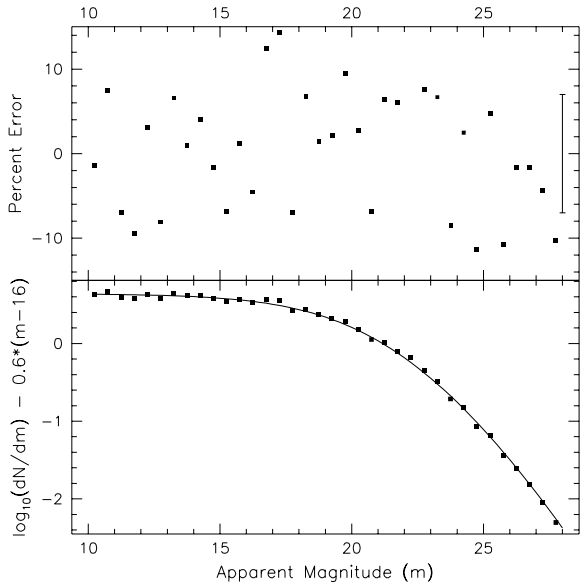

Figure 8.1: Counts test for the bootstrap integration. In the lower panel, the curve traces the counts predicted by the integration of Equation 8.7, and the points follow the bootstrapped results. The upper panel illustrates the error rate of the bootstrap integrator. The errorbar shows the one-sigma Poisson error of 7 %, corresponding to 200 galaxies per bin.

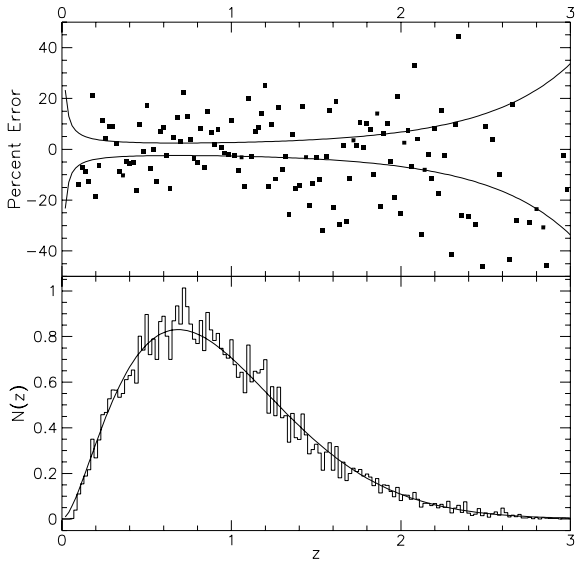

Figure 8.2: $N(z)$ test for the bootstrap integration. In the lower panel, the curve traces the $N(z)$ distribution for a magnitude-limited sample between 23 and 24 as predicted by the integration of Equation 8.10, and the histogram follows the bootstrapped results (a total of 100,000 galaxies were "surveyed"). The upper panel illustrates the error rate of the bootstrap integrator. The curves show the one-sigma Poisson errors as predicted by the analytically integrated $N(z)$ distribution.

For this comparison, it is important to integrate over a range in apparent magnitude, as the simulated survey is constructed. Figure 8.2 shows that the two methods agree well, although the error distribution may be a bit wider than the Poissonian prediction. This effect is due to rebinning effects. The bootstrap integrator bins the cumulative distribution of galaxies in 300 bins starting with a redshift of 0.001 and increasing logarithmically until the maximum redshift possible in the sample, in this case 4.5; consequently at a redshift of 1.0 these bins are about 0.02 wide, approximately the width of the bins used to plot the two distributions. Therefore, the errors should be slightly greater than Poissonian.

8.4 Conclusion

The statistical bootstrap is a Monte-Carlo method for generating magnitude-limited samples from numerical simulations. For a Schechter luminosity function, the method agrees well with the analytical forms for the number counts and redshift distribution. However, this bootstrap method is

much more flexible than the analytic method. It allows for the luminosity function to be arbitrarily complex and to agree exactly with the luminosity function realised in simulations. Furthermore, this method allows for the consideration of observation biases such as surface brightness effects, the translation from total magnitudes to observed isophotal magnitudes, and magnitude errors. Beyond redshift surveys, one can construct a magnitude-limited sample of luminosities and circular velocities, or any other galaxy properties calculated the simulations. This method observes a numerical simulation and calculates observable quantities with observational biases from the simulations. In the following chapter, it is applied to measure the number counts in the B and K-bands and the redshift distribution from a series of numerical simulations.

Chapter 9

Simulation Results

SUMMARY

Here the results of the numerical simulations and analysis are presented. We find that although the models have some success in remedying the shortcomings of the standard CDM cosmogony, none of these new models agree as broadly or as well with the observations as standard CDM. Although the low-Ω and $\Omega + \Lambda = 1$ models improve the agreement between the predicted and observed Tully-Fisher relations (the main weakness of galaxy formation in standard CDM), these models predict an inverted colour-magnitude relation and weak bright-end cutoff in the galaxy luminosity function. All of the models predict recent star formation in the majority of galaxies and exhibit galaxy colours bluer than observed; the CHDM model predicts colours two magnitudes too blue in $B - K$. We discuss several potential refinements to the "Block" model: the inclusion of metallicity effects, non-local feedback, inhibited star formation in cooling flows and an initial mass function that varies in time and location.

9.1 Main Results

As in Cole et al. (1994a), we choose to assess the various cosmological models described in the previous section with a host of diagnostics. We proceed as follows. The "astrophysical" parameters of Table 7.1 are varied until an acceptable fit to the present-day B-band luminosity function is found for each cosmological model. Typically, this involves choosing the appropriate value of the stellar mass-to-light ratio parameter Υ (to match the knee of the B-band LF); the merger rate parameters $\tau^0_{\rm mrg}$ and $\alpha_{\rm mrg}$ (which affect mainly the bright-end of the LF and are selected to suppress the formation of ultraluminous galaxies); and the parameters characterising the star formation rates and feedback, τ^0_*, α_*, $\alpha_{\rm hot}$, and $V_{\rm hot}$ (all of which have an appreciable effect on the faint-end slope of the LF). Although we did explore departures from the values outlined Cole et al. (1994a), we

choose to retain the same values as in the fiducial model for all parameters except Υ and $\tau^0_{\rm mrg}$. Varying the other parameters generally has little effect or results in an unacceptable B-band luminosity function. The parameters used in all the models are shown in Table 7.1.

Once these parameters have been specified, each cosmological model is fully determined. The good agreement or otherwise of each model with our additional diagnostics (the K-band LF, the infrared Tully-Fisher relation, the $B-K$ colours, the B and K-number counts, the $n(z)$ distributions, and the evolution of the luminosity function) should therefore be regarded as real successes or failures of that particular cosmogony. The first three diagnostics deal with the properties of the galaxy population at $z=0$, while the last four probe the evolutionary properties of galaxies. In some cases, and within the context of our modelling, it proved impossible to find an adequate fit to the B-band LF without violating one or more of the "cosmological constraints" mentioned in the previous section. When this occurs, we have explored how these constraints may be relaxed in order to improve the agreement of the model with observations. We shall comment on this in each individual case.

9.1.1 The B-band and K-band Luminosity Functions

Figure 9.1 presents the luminosity functions obtained for each model. The fiducial model (*i.e.* that presented in Cole et al. (1994a)) is a reasonable fit to both the B- and K-band data. The faint-end slope seems to be slightly steeper than the Loveday *et al.* and Mobasher *et al.* data for field galaxies, but the discrepancy is not dramatic, especially noting that LFs derived from different samples (*e.g.* the CfA redshift survey or the LF in clusters, de Lapparent, Geller & Huchra 1989, Colless 1989) tend to give steeper slopes than the data used for this comparison. The faint-end slope of the fiducial model is actually much shallower than the slope of the mass function of dark halos in this model, an effect due largely to the strong suppression of star formation in low-mass halos. This point is especially important, for it solves a well-known problem for hierarchical clustering theories to produce galaxy luminosity functions as shallow as observed. The price to pay is a dramatic steepening of the Tully-Fisher relation at the low-mass end. We discuss this point in more detail below.

The good agreement at the bright-end is due partly to our moderate choice for merger rates but also to the relatively late formation of massive halos in this model. The short lifetimes of very massive halos prevent large amounts of gas from cooling to form ultraluminous galaxies at the centre of these halos. Finally, good agreement at the knee of the LF is obtained by choosing $\Upsilon = 2.7$, which indicates that a fair amount of mass should be in the form of "dark stars". The stellar mass-to-light ratios implied by choosing Υ in this manner and other properties of the stellar populations of the models are summarised in (Table 9.1).

The low-H_0 model is also a moderately good fit to the luminosity func-

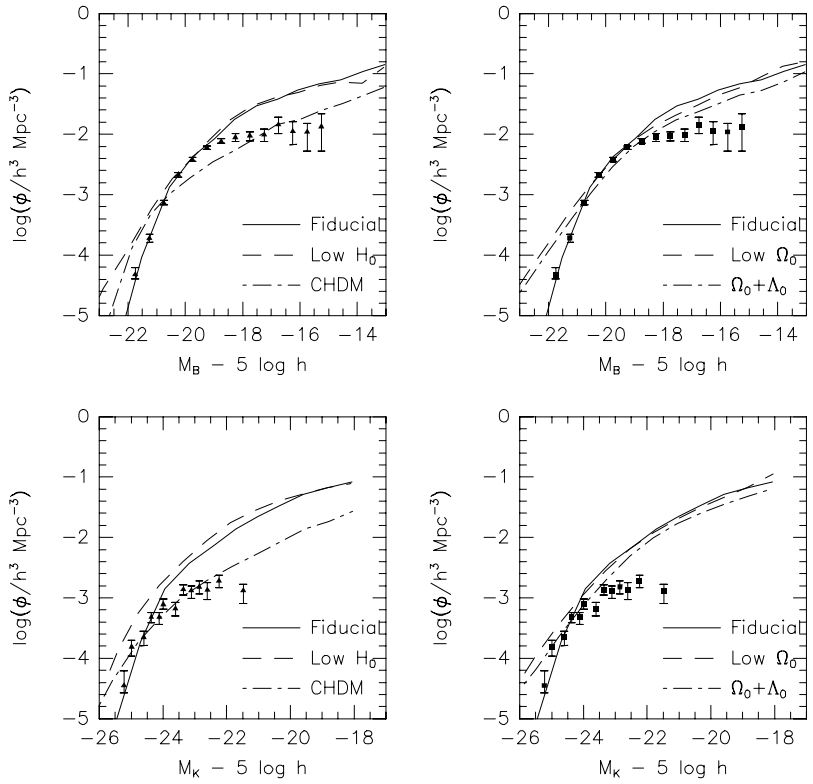

Figure 9.1: Luminosity Functions: The four panes show the luminosity functions at $z = 0$ in the B-band and K-band for the five models. The points with errorbars show the Loveday et al. (1992) results and Mobasher, Sharples & Ellis (1993) results. The solid line traces the luminosity function in the Fiducial model in both panes.

Table 9.1: Properties of the Stellar Populations: The first two columns are the median stellar mass-to-light ratios and current starformation rates of galaxies brighter than $M_B = -19.5$ The second and third columns give the redshift before which half the stars in the model formed and the corresponding time elapsed since this point.

Model	$(M_*/L_*)/(hM_\odot/L_\odot)$	$\dot{M}_*/M_\odot \mathrm{yr}^{-1}$	z_*	t_*/Gyr
Fiducial	17	4.3	0.87	8.0
low-H_0	34	7.2	0.71	14.5
low-Ω_0	8.9	4.7	1.17	8.3
$\Omega_0 + \Lambda_0$	7.7	7.7	0.86	8.3
CHDM	2.3	4.2	0.23	3.5

tion data, albeit for a slightly different choice of astrophysical parameters. However, there are more stars in this model (because of the higher Ω_b), and they are proportionally much older than the stars in the fiducial model (because the age of the universe has doubled). These two effects result in very high stellar mass-to-light ratios for typical galaxies; $(M_*/L_*) \sim 34h(M_\odot/L_\odot)$ for galaxies brighter than $M_B = -19.5$, compared to the observed $\sim 10 - 20hM_\odot/L_\odot$ in ellipticals (Lauer 1985) and $\sim 5M_\odot/L_\odot$ in the solar neighbourhood (Bahcall 1984). This we regard as a serious shortcoming of the low-H_0 model. Reducing the value of Ω_b to less than half that prescribed by primordial nucleosynthesis can reduce the stellar mass-to-light ratios to within the observational uncertainties. However, even with this rather ad-hoc modification the model can not account for the zero-point in the Tully-Fisher relation or for the observed colours of galaxies, as will be shown in the following subsections.

The CHDM model has the opposite difficulties. The general feature of this model is that halos of galactic size form so late that they have not had time to form enough stars by $z = 0$. Matching the knee of the B-band luminosity function, or equivalently the luminosity density of the universe, actually requires values of $\Upsilon < 1$, which are of course unacceptable. (Υ must be larger than unity because it represents the ratio of the total mass in stars formed in a star formation burst to the mass of "visible" stars, *i.e.* excluding brown dwarfs.) Removing the feedback from star formation altogether allows more stars to form and to form earlier, but the knee in the predicted luminosity function nearly disappears as the faint-end slope markédly increases and the luminosity density of the universe still falls short of that observed. Only increasing Ω_b to 0.12, twice the value allowed by big-bang nucleosynthesis, can provide more fuel for star formation and can make it possible to match the knee of the luminosity function. We will explore this high-Ω_b further through the evolution of its B-band luminosity function.

The stellar mass-to-light ratios do not seem to pose problems to the other two cosmological models. However, they seem to produce too many bright galaxies to be consistent with the data. Reducing the efficiency of merging has no significant effect on these galaxies, whose large luminosities are related to the fact that halos in this model have in general collapsed much earlier than in the fiducial model. Cooling has therefore had more time to act in massive halos, leading to the formation of overluminous galaxies. Bringing these models in agreement with observation would require adopting a star formation cutoff in very massive systems. The same problem was noticed by Kauffmann, Guiderdoni & White (1994), who decided to neglect star formation in halos with circular velocities larger than about $500\,\mathrm{km\,s^{-1}}$. Adopting a similar prescription here would reconcile the model with the bright end of the observed LF.

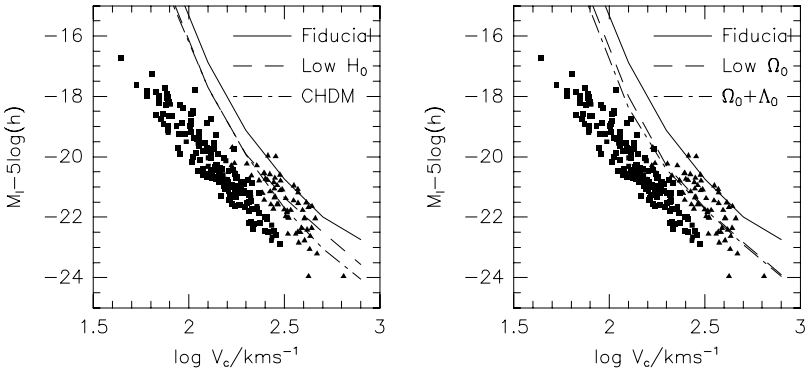

Figure 9.2: I-Band Tully-Fisher Relation: The various models are plotted as curves which trace the mean luminosity of the galaxies at a given circular velocity. The fiducial model is plotted in both panes as a solid line. The open squares are a sample of spirals compiled from new and published cluster data (Young et al. 1994, in preparation) and the triangles a sample of ellipticals from the Coma cluster (Lucey et al. 1991) which have been placed on this plane by defining an effective circular velocity in terms of the observed velocity dispersion, $V_c = \sqrt{3}\sigma_{1D}/1.1$

.

9.1.2 The Tully-Fisher relation

Comparing the results of our models with the observed Tully-Fisher relation requires that we assign a rotational velocity to the "galaxies" formed in our block model. There is no unique way of doing this as the rotational velocities of disks are affected by the spatial distribution of the baryonic component at the centre of the dark halos, an effect that is not taken into account in our model. The simplest procedure seems to be to assign to each galaxy a rotational velocity equal to the circular velocity of the halo in which it was formed. As mentioned in Section 7.1, in the case of the fiducial model this identification results in a zero-point for the Tully Fisher relation which is about two magnitudes fainter than observed (Figure 9.2).

In principle, we could change the value of the stellar mass-to-light ratio parameter Υ to bring the model into better agreement with the observed Tully-Fisher relation, but this would would result in a large overestimate of the luminosity density of the universe or, equivalently, in a significant disagreement with the observed luminosity function. The problem seems to be due to an overabundance of halos with circular velocities typical of galaxies, a problem that has also been noted by Lacey et al. (1993) and Kauffmann, Guiderdoni & White (1994).

The low-H_0 model has the same number density of halos as the fiducial model (per $(h\text{Mpc})^3$), so fitting simultaneously the galaxy luminosity function *and* the Tully-Fisher relation does not seem possible. The CHDM

model does not seem to do well either, despite the fact that this model does not fit the present-day luminosity function, and that fewer galaxy-sized halos have collapsed by $z = 0$ than in the fiducial mode. Both effects tend to make galaxies brighter at a given V_c, and to improve the agreement with the observed Tully-Fisher relation. It is also disappointing that increasing the value of the baryon density in order to improve the galaxy luminosity function has little effect on the Tully-Fisher relation. The number density of galaxy-sized halos is lower than the fiducial model in the cases with low- Ω_0 and $\Omega_0 + \Lambda_0$ and therefore their Tully-Fisher zero point is in better agreement with observations. However, only values of Ω_0 much lower that the one used here would produce a zero point in agreement with observations.

It is also important to note that in all models the discrepancy with the observed Tully-Fisher relation becomes more pronounced in low-mass halos. Indeed, the slope of the relation steepens below $V_{\rm hot} = 140\,{\rm km\,s^{-1}}$, due to strong suppression of star formation in these systems. It is in fact this steepening that is mainly responsible for the shallow faint end slope at the faint end of the luminosity unction and it is, therefore, a general prediction of our models. In other words, luminosity functions with faint end slopes shallower than the halo mass function *can* be obtained, but only at the expense of steepening the Tully-Fisher relation below $V_{\rm hot}$. Can such a steepening be ruled out by present data? Selection effects may cause such a steepening of the Tully-Fisher relation to go undetected, as would be the case if only the brightest galaxies have been used to define the Tully-Fisher relation at low V_c.

From this discussion it seems that, despite the wide range of parameters tried, none of our cosmological models can reproduce simultaneously the galaxy luminosity function and the Tully-Fisher relation. Do we need to conclude from this that all these models are fatally flawed? This would be perhaps premature. The weakest link between observations and our model predictions is certainly the assumption that the rotational velocity of a galaxy is the same as the circular velocity of its surrounding halo, and there are many ways in which this identification can go wrong. For example, if dark halos are not well represented by singular isothermal spheres but instead possess sizeable core radii, the rotational velocity of the galaxy's disk may not be a good indicator of its surrounding halo's V_c. This indeed seems to be the case in galaxy clusters, where the velocity dispersion of the central galaxy is generally several times lower than that of the cluster itself. If a similar effect were at work in galaxy halos it would mean that the rotational velocities assigned to the model galaxies are too large. In particular, if disk galaxies inhabit halos with mean circular velocities about twice the disk's circular speed, then the Tully-Fisher problem would be solved. Since detailed analysis of disk rotation curves and the dynamics of satellite systems strongly suggest that galactic halos are *not* isothermal spheres, this suggestion may not be as extravagant as it appears (Persic & Salucci 1992, Ashman 1992, Zaritsky et al. 1993, Flores et al. 1993).

9.1.3 Colours

Observed broad-band colours indicate that galaxies of different magnitudes have undergone a wide variety of star formation histories. The brightest galaxies tend to be very red ($B-K>4$), while fainter galaxies are noticeably bluer. At all magnitudes, the scatter in colours is quite large, about one magnitude in $B-K$. This is shown in Figure 9.1.3, where we plot the data from Mobasher, Ellis & Sharples (1986) as a histogram, after dividing the sample in two magnitude bins. The fiducial model fails this comparison in two counts; neither has is galaxies as red as the brightest ellipticals in Mobasher *et al.* 's sample nor do the galaxies in each magnitude bin span as wide a range in colours as observed. However, the trend is correct, as brighter galaxies tend to be redder than the rest. This in itself is a success for a hierarchical model in which larger systems collapse later, and comes about because stars in large galaxies today formed preferentially in smaller clumps that collapsed early and were only recently assembled into one single massive object.

It would be expected that the low-H_0 model fared better in this respect because of the increased age of the universe over that of the fiducial model. Indeed, galaxies are slightly redder in this model, but not nearly as red as observed. Although the universe in the low-H_0 model is 13 Gyr older than in the fiducial model, feedback prevents a large number of stars from forming in low-mass halos at high redshift. Star formation begins in earnest only when halos with $V_c \sim V_{\rm hot}$ collapse, and therefore stars form on average only about 5-6 Gyr earlier. The colours predicted by stellar population synthesis evolve only very slowly as stars age from $\sim$ 7-8 Gyr to $\sim$ 12-14 Gyr and no major improvement in the colours result. Similarly, no significant improvements are obtained in the low-Ω_0 and $\Omega_0+\Lambda_0$ models, where, if anything, the colour-magnitude trend seems to have been reversed. This is due to the effects of cooling and late star formation on the largest halos, as discussed in Section 9.1.1. Not surprisingly, the CHDM model performs poorly. Although it has the same age as the fiducial model, galaxy-sized halos collapse much later and their stellar populations have not had enough time yet to evolve to colours comparable to those of present-day galaxies.

9.1.4 Number Counts and Redshift Distribution

Magnitude-limited number counts and redshift distributions depend on the galaxy luminosity function and its evolution and therefore probe the evolutionary properties of our models. The observed B-band number counts point toward significant evolution of the luminosity function with lookback time while the K-band counts and B-band redshift distributions are consistent with little or no evolution. Figure 9.4 depicts the B- and K-band number counts predicted by the models as well as recent observational results. The fiducial model's predictions are consistent with the observations of the faint number counts *and* the redshift distributions. However, it overpredicts the number of galaxies at bright apparent magnitudes.

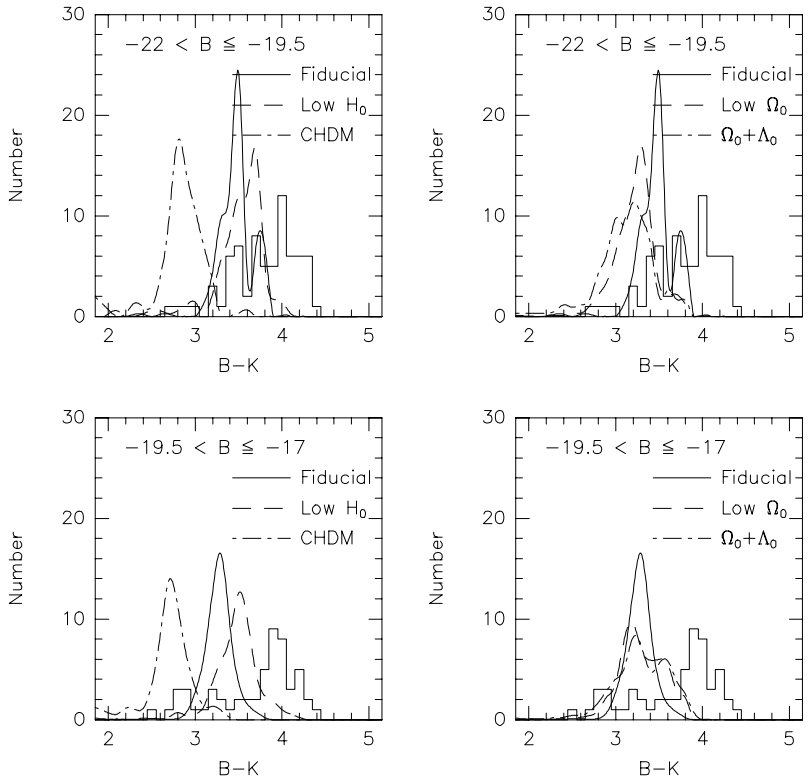

Figure 9.3: $B-K$ Colours: Each pane shows the observed colour distribution of galaxies (Mobasher, Ellis & Sharples 1986) as a solid histogram and the distribution predicted by the fiducial model as a solid line. The upper panes show the distribution for bright galaxies with $-22 < M_B < -19.5$. The lower panes show those galaxies with $-19.5 < M_B < -17$. In all four panes the theoretical distributions have been normalised to have the same area as the observed distributions.

In comparison, the other models do not fare nearly as well. Regarding the number counts, the low-H_0 model predicts fewer galaxies at faint apparent magnitudes and slightly more bright ones. This effect is more pronounced for the low-Ω_0 and $\Omega_0 + \Lambda_0$ models. The CHDM model fails even more dramatically. Galaxy formation occurs very late (75% of all stars have formed later than $z = 0.4$), and results in a surplus of very bright, blue galaxies. As a result, it underestimates the number of faint blue galaxies and predicts too few galaxies at all K-magnitudes.

The redshift distributions also reflect the luminosity functions produced by the models. Figure 9.1.4 shows that the fiducial model and low-H_0 model both predict $N(z)$ distributions consistent with the observations. The low-Ω_0 and the $\Omega_0 + \Lambda_0$ models exhibit a small excess of high redshift galaxies in both the $B = 22$ and the $B = 24$ distributions. These tails are symptoms of the overproduction of luminous galaxies in these two models which is apparent in their luminosity functions. On the other hand, since few stars in the CHDM model form at high redshift, it predicts a redshift distribution that peaks too early and has hardly any galaxies beyond $z = 0.8$, in disagreement with observations.

9.1.5 The Evolution of Galaxy Luminosity Function

The galaxy luminosity function and its evolution are a fundamental statistic of the galaxy population; the number-magnitude relation and the redshift distribution are simple convolutions of this evolution within the cosmological framework. Figure 9.6 summarises the evolution of the rest-frame B-band luminosity function for four of the five models. The critical density fiducial and low-H_0 model evolve smoothly until the present day. The faint-slope gradually but continually becomes shallower. The present-day slope, as mentioned in Section 9.1.1, is still much steeper than observed. Both the low-Ω_0 model and the $\Omega_0 + \Lambda_0$ model (only the $\Omega_0 + \Lambda_0$ model is depicted) evolve little since a redshift of one. From $z \approx 1.5$ to $z \approx 0.5$, the luminosity function does become slightly shallower, but changes little since that time. In low density universes, structure in the dark matter grows ceases to grow at $z \sim \Omega^{-1}$ or 3 in these models; the galaxies lag behind by a few billions years. The evolution in the CHDM model is most striking. In CHDM universes, the structure only begins to form quite recently; the galaxy population evolves quickly to reach its present-day state. The galaxies are fainter and less numerous even in the recent past, and beyond a redshift of one they are nearly nonexistent.

Which of these models compares most favourably with the observation evolution of the luminosity function presented in Chapters 5 and 6? Although the fiducial model begins with a steeper faint-end slope than observed, the observed slope in the luminosity function quickly catches up with that of the fiducial model. And if it is possible to extrapolate the earlier results, the observed slope at $z \sim 1$ may even exceed the slope predicted by the fiducial model. The low-H_0 model also evolves quickly, and it

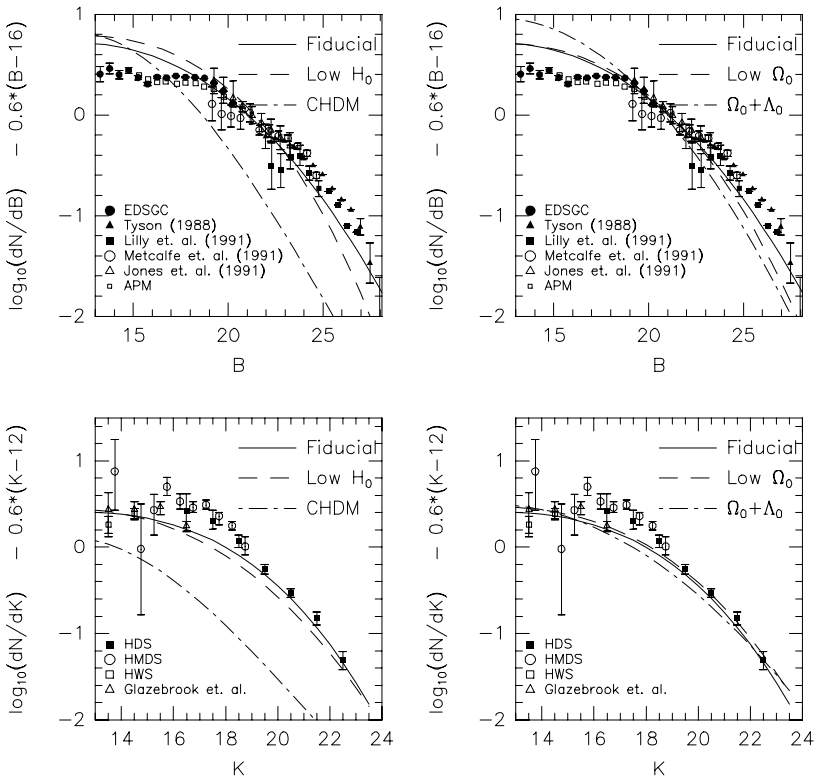

Figure 9.4: Number Counts: The upper show the B- band number counts for the five models, and the lower panes show the K-band counts. The various polygons are the observational data, from the sources given in the key. The raw counts have been divided by a pure power law with slope 0.6, so as to expand the useful dynamic range of the figure. Thus, the Euclidean number counts would appear as a horizontal line. The B-band data are taken from Maddox et al. (1990b), Jones et al. (1991), Metcalfe et al. (1991), Lilly, Cowie & Gardner (1991), Tyson (1988) and Heydon-Dumbleton *et al.* (1989; EDSGC). Where necessary b_j magnitudes have been converted to Johnson B assuming $B = b_J + 0.2$. The $K-$band data are taken from Glazebrook, Peacock & Collins (1994), the Hawaii Wide Survey (HWS), the Hawaii Medium Deep Survey (HMDS) and the Hawaii Deep Survey (HDS) as reported by Gardner, Cowie & Wainscoat (1993).

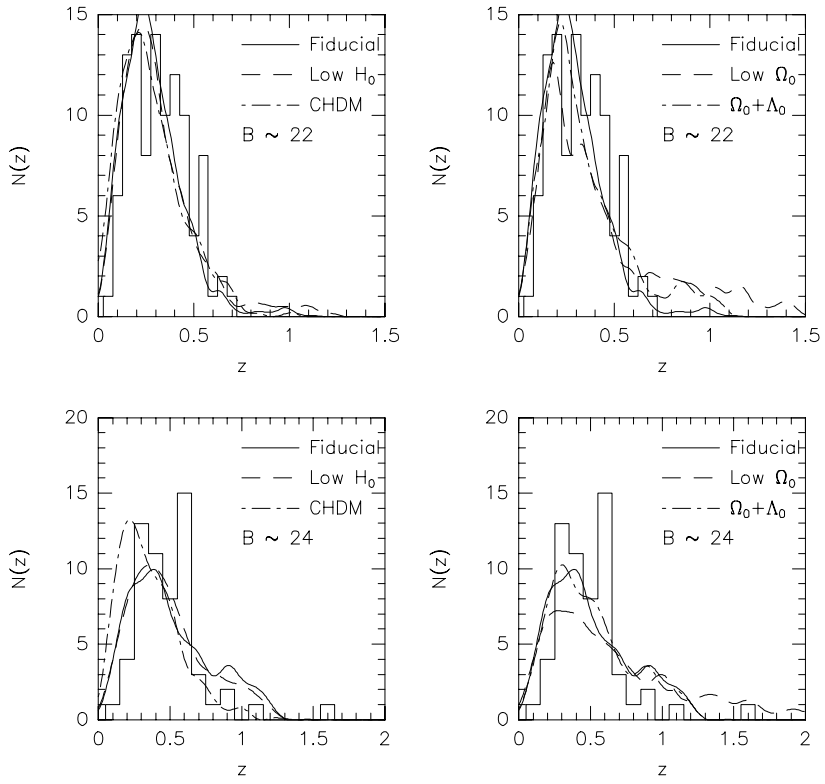

Figure 9.5: N(z): The upper panes show the distribution of redshifts for a magnitude- limited sample from $B = 21$ to $B = 22.5$. For comparison, the LDSS data (Colless et al. 1993) is plotted as a histogram. The lower panes show the N(z) distribution from $B = 22.5$ to $B = 24$. We have plotted the LDSS-2 data (Glazebrook et al. 1993) for comparison. In all four panes the theoretical distributions have been normalised to have the same area as the observed distributions.

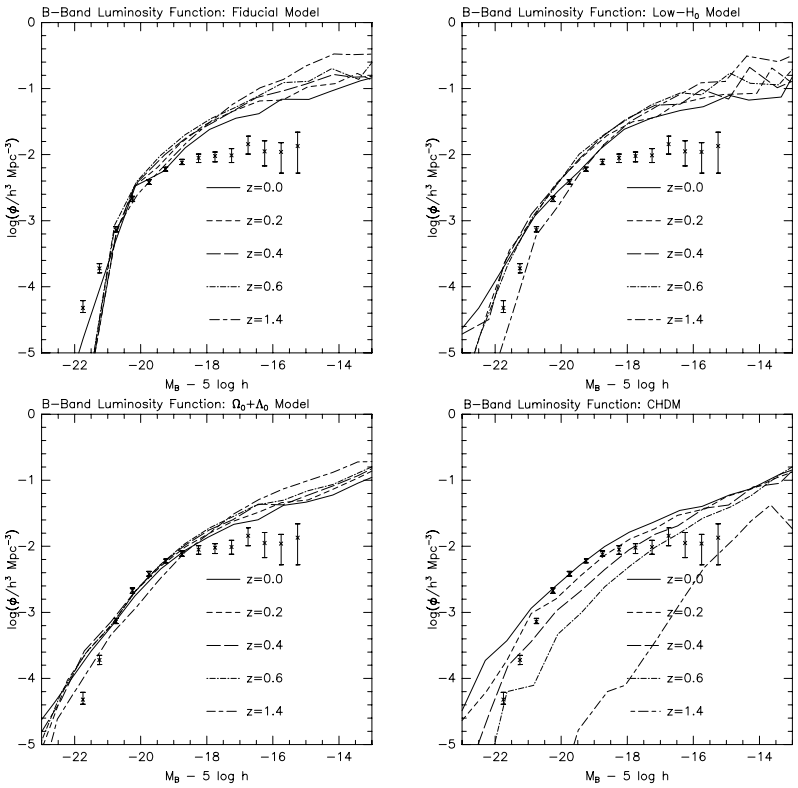

Figure 9.6: Evolution of the Luminosity Function. The evolution of the luminosity function for four out of the five models is presented. For brevity, the low-Ω_0 model whose evolution is similar to that of the $\Omega_0 + \Lambda_0$ model has been omitted. Furthermore, for fairness an alternative CHDM model is presented. It fits the knee of the local luminosity function but violates BBN constraints on the density of baryons.

begins with a slightly shallower slope. It provides an equally good fit to the observed evolution of luminosity function. Both the $\Omega_0 + \Lambda_0$ and low-Ω_0 models evolve too slowly; the CHDM model evolves too quickly and in the wrong direction. The poor fit of these models to the number-magnitude relation is simply a manifestation of their poorly performance in predicting the evolving luminosity function. It is difficult to reconcile a low-density universe in which structure evolves too early and a CHDM universe in which structure forms too late with the rapid and recent evolution of the luminosity function described in Part I of this thesis.

9.2 Discussion

The successes and failures of the fiducial CDM model of Cole et al. (1994a) were summarised in Section 7.1. We now assess, in turn, the pros and cons of each of the alternative models. These models were selected specifically to find out if the deficiencies of the fiducial model could be remedied within our general scheme for galaxy formation merely by changing the underlying cosmological assumptions. The parameters of these models (listed in Table 7.2 were chosen for consistency with recent data on galaxy clustering and peculiar velocities, Big Bang nucleosynthesis calculations, and main sequence determinations of the age of galactic globular clusters. Our strategy was to adjust the free astrophysical parameters in our scheme until the best possible agreement with the observed galaxy luminosity function was obtained.

Low-H_0 CDM. If H_0 is low, Big Bang nucleosynthesis requires a large baryon density, $\Omega_b = 0.2$ for a model with $H_0 = 25$ km s^{-1} Mpc^{-1}. Such a large value gives rise to very efficient star formation at early times which is not significantly suppressed even when feedback effects are as strong as we have assumed. As a result, the predicted stellar mass-to-light ratio of bright galaxies turns out to be unacceptably large. This difficulty may be circumvented by violating the nucleosynthesis constraint but, if Ω_b is reduced much below 0.1, the stellar populations become too old and too faint to account for the observed abundance of bright galaxies. The best model of this kind has $\Omega_b = 0.1$, strong feedback, and a moderate amount of galaxy merging.

The resulting luminosity function is similar to that of the fiducial model. However, the model does not fully resolve the problem which motivated it in the first place: the need to produce bright galaxies as red as the reddest field ellipticals. Although the age of the universe is in this case 26 Gyrs, feedback effects – required to prevent an excessively large abundance of dwarf galaxies – delay the onset of star formation until relatively low redshifts and results in a paucity of very red, bright systems. Indeed, the most extreme galaxies in the model have $B-K \simeq 4$, somewhat redder than those in the fiducial model but still about 0.5 mag bluer than the reddest field ellipticals.

The second main problem of the fiducial model, *i.e.* the incorrect zero point in the Tully-Fisher relation, is not resolved by lowering H_0. Although galaxies with a given circular velocity are about 1 mag brighter in the low H_0 model than in the fiducial model, they are still over 1.5 magnitude too faint. Overall, the low H_0 model appears rather unattractive, especially considering the growing observational evidence in favour of a large value of H_0 (see *e.g.* Jacobi et al. 1992 and references therein).

Low-Ω_0 CDM. Our main motivation for examining this model was the expectation that the lower abundance of galactic halos that form in this case would be enough to bring the predicted Tully-Fisher relation into agreement with observations. This expectation was only partially fulfilled. As in the fiducial model, the predicted Tully-Fisher relation has about the observed slope for $V_c > 100\,\mathrm{km\,s^{-1}}$, but the zero point is still about one magnitude too faint at $V_C \sim 200\,\mathrm{km\,s^{-1}}$. Although this represents a considerable improvement over the fiducial model, it cannot be claimed as a significant success. The low-Ω_0 model performs worse than the fiducial model on two counts: its luminosity function rolls over gently at the bright end, rather than cutting off exponentially, and the colour distribution of bright galaxies is shifted even further to the blue. The sign of the colour-magnitude relation – a notable success of the fiducial model – is inverted with brighter galaxies being bluer than fainter ones. These shortcomings can be traced back to excessive cooling of gas onto large dark matter halos which form much earlier in this model than in one with a flat geometry. As noted by Kauffmann, Guiderdoni & White (1994), they may be circumvented by postulating that cooling flows in large galaxies do not produce visible stars, as seems to be the case in the cooling flows inferred in the cores of rich clusters (eg Fabian 1991). The counts of faint galaxies in the low-Ω_0 model are as in good agreement with observations as those in the fiducial model, but the excess population of bright galaxies gives rise to a significant tail of high redshift galaxies which may well be inconsistent with existing data.

$\Omega_0 + \Lambda_0$ CDM. Adding a non-zero cosmological constant to the low-Ω_0 model has only a minor effect, although some of the small differences that there are seem to be in the right direction. The problem at the bright end of the luminosity function is slightly reduced, but the cutoff is still not as sharp as observed. The predicted Tully-Fisher relation and colour distributions change very little, but the B-band counts of faint galaxies drop by about a factor 2 below the fiducial model –which provides an excellent match to observations. This difference arises because the faint end slope of the luminosity function is flatter and evolves more slowly in this than in the fiducial model. This potential difficulty may not be too serious since, as shown in Cole et al. (1994a), the faint counts are rather sensitive to the assumed stellar initial mass function and to the details for the feedback prescription.

CHDM. Like the two previous cases, a model with a mixture of cold (70%) and hot (30%) dark matter was considered in the expectation that the Tully-Fisher discrepancy of the fiducial model might be resolved. With CHDM a lower abundance of galactic halos is produced because, for a given amplitude on large scales, the power spectrum has relatively less small scale power than with CDM alone. Better agreement with the Tully-Fisher relation is indeed obtained in the CHDM model, but the zero-point discrepancy is not fully removed. In fact, the Tully-Fisher relation in this model is virtually identical to those in the low-Ω_0 and $\Omega_0 + \Lambda_0$ models.

The reduced spectral power on galactic scales has an undesirable side-effect which makes the CHDM model rather unattractive: bright galaxies form much too late to be consistent with observations. With $H_0 = 60$ km s^{-1} Mpc^{-1}, the baryon density required by Big Bang nucleosynthesis constraints is too low to form enough bright galaxies to match the knee of the luminosity function. Even if we disregard the BBNS constraints and arbitrarily set $\Omega_b = 0.1$, the resulting luminosity function does not show the characteristic break at high luminosities. Perhaps more damning are the extremely blue galaxy colours predicted at the present epoch which, in the mean, are about 2 mag bluer than observed. The reddest objects in the model have $B - K \simeq 3.5$, one magnitude short of the reddest observed ellipticals.

These difficulties are also manifest in the counts of faint galaxies, which are a factor of 10 lower in the K-band than observed and in their redshift distribution which is strongly biased towards low redshift, in strong disagreement with observations. The problem of late galaxy formation in the CHDM model seems unsurmountable. This conclusion is virtually independent of the details of our galaxy formation model. Even if we switch off the feedback altogether and adopt a very short star formation timescale (which produces a completely unacceptable luminosity function) bright galaxies are still much too blue.

In summary, none of the models we have considered are completely satisfactory. Overall, the most successful ones are the fiducial CDM model of Cole et al. (1994a) and the $\Omega_0 + \Lambda_0$ model. This model only partly solves the Tully-Fisher discrepancy that afflicts the fiducial model, but does so at the expense of a rather poor fit to the observed luminosity function and an even worse colour problem than in the fiducial model. There is also a potential difficulty explaining the faint counts in the low-$\Omega_0 + \Lambda$ model.

Our results are quite consistent with those of Kauffmann *et al.* (1993, 1994). This agreement strengthens our conclusions since our two approaches, although similar in spirit, differ significantly in many astrophysical details. Our failure to find a fully consistent picture of galaxy formation within currently popular cosmologies, suggests that we should look carefully at the astrophysical inputs that go into our modelling procedure.

The colour problem, common to all the cases we have examined (including the long-lived low-H_0 model), is particularly puzzling. The stellar population synthesis model which we use produces acceptable fits to the in-

tegrated spectral energy distributions of present day galaxies of all spectral types. However, the more realistic star formation laws in our models invariably produce intermediate age stellar populations in bright galaxies from the late infall of gas expelled from halos in the lowest level of the clustering hierarchy. Regardless of the detailed prescription for feedback, gas must be prevented from forming stars profusely in these low-mass halos; otherwise virtually all the baryons would be turned into stars well before the present, and an unacceptably large abundance of dwarf galaxies would result. It is possible that a star formation rate more strongly biased towards high redshift than our models generically predict might circumvent these problems. Another possibility is that current stellar population synthesis models are predicting colours which are inaccurate at the 0.3 mag level in $B-K$. Such inaccuracies might arise from the treatment of the poorly understood late stages of stellar evolution (particularly the asymptotic and post asymptotic giant branch) or from the neglect of chemical evolution.

We have argued that some of form of feedback is an essential requirement in any hierarchical clustering theory of galaxy formation. The ejection of gas (and metals) observed in bright ellipticals, sometimes in the form of highly energetic superwinds (David, Forman & Jones 1991; Heckman, Armus & Miley 1990), provides an example of the sort of process which may be required. Nevertheless, there is no direct observational guidance for assuming any particular form of feedback in the highly specific conditions prevailing at high redshift. The feedback mechanism implemented in our scheme and in most other related ones is a local process where star formation is regulated *in situ*. Non-local processes such as photoionisation (Efstathiou 1992) or bulk gas motions could be important and it is not inconceivable that they could depend on the large-scale environment or act selectively, allowing early formation in some halos and delaying it or suppressing it altogether in others. Processes of this sort might alleviate the colour discrepancy discussed above and could even give rise to "naked halos", dark matter objects in which no visible galaxy ever forms. The Tully-Fisher discrepancy in the standard CDM model and probably also in the alternative models which we have considered, could be resolved if a substantial fraction of dark halos do not harbour bright galaxies.

A further source of uncertainty in our scheme for galaxy formation in general, and in the stellar population synthesis models in particular, is the stellar initial mass function (IMF). The universality of the locally determined IMF has been a longstanding matter of much debate. Perhaps the strongest argument for a non-universal IMF comes from studies of the metallicity of the intracluster gas which seems to require a bimodal IMF in the metal-producing galaxies (Arnaud et al. 1992). It is not difficult to speculate on the many outcomes possible with a variable IMF. For example, an IMF biased towards large masses in metal-poor systems, might alleviate the excess dwarf problem if their stellar populations have faded by the present day.

It should be clear from the above discussion that, subject to the ob-

servational constraints of large scale structure, Big-Bang nucleosynthesis, and globular cluster ages, our semianalytic recipe for galaxy formation fails to produce a fully acceptable model. Our results are quite consistent with those of Kauffmann *et al.* (1993, 1994). This agreement strengthens our conclusions since the two approaches, although similar in spirit, differ significantly in many astrophysical details. It is difficult to see how, without revision of our scheme or dramatic changes in the interpretation of observations, hierarchical models of the kind can successfully account for the observed properties of the galaxy population. This is illustrative of the potential of the semianalytic methods; they enable us to test a wide variety of models and assumptions as well as to isolate the root causes of disagreement between observations and specific cosmogonies. This, in itself, should be regarded as a success of our modelling technique, as it highlights the obstacles to be dealt with by future attempts at unravelling the process of galaxy formation in a hierarchical universe.

Part III

Finishing Touches

Chapter 10

Conclusion

We may approach the question of galaxy evolution from several angles. However, observational results must be comparable to theory and *vice versa*. We must find some middle ground. From an observational point of view, a good starting point is a large and deep redshift survey. A survey alone provides few theoretical constraints (*e.g.* a given redshift distribution may be consistent with a wide range of evolving luminosity functions). Using detailed analysis and exploiting information often overlooked, the galaxy evolution may be constrained.

The first step is accurately determining the k-corrections of the galaxies. The k-corrections define the restframe luminosities of the galaxies and the volume that the survey samples. The thesis has introduced and used a new technique, cross-correlating the observed spectra against templates with known morphologies, and mapping this morphological data onto k-corrections. Integrated observations extending the Kennicutt spectra atlas (Kennicutt 1992a,1992b) into the ultraviolet would remove this mapping step and reduce the uncertainties in the k-corrections. With the k-corrections in hand, one can take several avenues to determine the evolution of the luminosity function. This thesis has introduced two more, one parametric and one nonparametric (the SSTY and SSWML methods) and shown how these techniques relate to contemporary procedures. Both are insensitive to galaxy clustering.

It is important to meet halfway. Most theoretical models predict volume-limited samples while most observations probe magnitude-limited samples. Although it is possible to model the mapping from a volume-limited to a magnitude-limited sample, it is far from straightforward to simulate all the possible observational biases. The thesis introduces a technique akin to the statistical bootstrap to quickly generating a magnitude-limited sample from a volume-limited sample that can include any observational bias; *provided* the biasing is well-defined for a single galaxy. Depending on the information available to the theorist, this may include the translation from total to isophotal magnitudes and diameters as a function of surface brightness. The

technique is demonstrated by predicting the number counts and redshift distributions for several numerical models of galaxy formation. Many other applications abound. For example, it is possible to generate a magnitude-limited sample to study the Tully-Fisher relation in a variety of hierarchical universes.

Even with these new techniques, we are far from constraining the processes of galaxy formation from observations, but several results appear clear. In the B-band, the number of faint galaxies in a given magnitude range has been decreasing since a redshift of 0.5. Meanwhile, bright galaxies have evolved little. This result is independent of the k-corrections assumed for the galaxies, the method of deriving the luminosity function and incompleteness. Scrutinising the observations further reveals dramatic evolution is constrained to galaxies with late-type spectral identifications – blue galaxies. At the faint end of the luminosity function, these galaxies were several times more numerous at $z \sim 0.2$ than today and an order of magnitude more numerous at $z \sim 0.5$ than today. Furthermore, these galaxies, regardless of luminosity, were much more fiercely forming stars in the recent past, as revealed by the distribution of W_λ[OII] as a function of redshift and absolute magnitude. Meanwhile, galaxies with earlier spectral types have evolved little since $z \sim 0.5$. The number density of faint elliptical galaxies and early spirals actually appears to have increased over the past few billion years. And the tracers of star formation were no more common in the past than today in these galaxies.

These findings indicate that the faint galaxies were much more abundant in the recent past, but since that time they have disappeared below our detection limits. From a theoretical point of view, several models may explain the disappearance. These models often indicate that a galaxy's environment may be the driving force in its evolution. An archetype is the proposal of Babul & Rees (1991). They posit that the formation of dwarf elliptical galaxies is suppressed by the UV background radiation from quasars until $z \sim 1$, when their gas begins to collapse, and stars begin to form quickly. Suddenly these galaxies begin to appear in the surveys with late-type spectra. However, as supernovae begin to blow out the gas, star formation ceases. If the galaxy is in a low pressure environment, the gas will simply return to the inter-galactic medium, and the galaxy winks out. The gas will return to the few galaxies in high pressure environments fueling further star formation episodes, until the gas is exhausted. Many more stars form over these repeated episodes, and the galaxies do not fade into obscurity but into ubiquity. Babul & Rees (1991) argue that today these galaxies would be identified as dwarf ellipticals which are often found in clusters and near larger galaxies.

The effect of galaxies on the evolution of their neighbours has only recently been included in the modelling of galaxy formation. The "block" model explored in the second part of the thesis does include the effects of galaxy merging, and predicts many of the features of the present-day galaxy population. Some of the discrepancies between its predictions and

the observations such as Tully-Fisher zero-point fainter than observed, may be alleviated by varying the cosmology – in a low-density universe, the "block" model is closer to predicting the observed zero-point. However, regardless of how the model is pushed, some problems remain. The "block" model predicts many more faint galaxies than are observed. And if one were to insist on a model that predicts the slope of the Tully-Fisher relation at small V_c, it would predict an even larger excess of faint galaxies. Both the shallower faint-end slope of the luminosity function and the steeper faint-end slope of the Tully-Fisher relation result from a form of local feedback which suppresses star formation in galaxies with small circular velocities. In this framework, the tradeoff is inevitable. To lower the faint-end slope of the luminosity function, the haloes of a given circular velocity must be assigned a wide range of luminosities, increasing the slope and scatter of the predicted Tully-Fisher relation.

It is possible to quench star formation in dwarf galaxies without regard to their internal properties by including the effect of environment on the galaxy evolution – non-local feedback. A promising source of non-local feedback is the ionising background produced by quasars before $z \sim 1$, advocated by Efstathiou (1992) and Babul & Rees (1991). Lacey et al. (1993) propose an alternative method in which the tides of induce star formation in nearby dwarf haloes. This would have similar results to the photoionisation model. Dwarf galaxies would form only near larger galaxies, and only after their shared halo collapsed, which would delay star formation. Both of these attractive alternatives will be investigated in the next rendition of the "block" model.

From an observational point of view, the study of the effect of environment on galaxy evolution will advance briskly with new large redshift surveys such as the Sloan Digital Sky Survey with a dedicated telescope at Apache Point, New Mexico and the fainter 2dF survey at the AAT. Although the former survey will provide needed constraints on the properties of galaxies locally, it can probe only bright galaxies out to even moderate redshift. The 2dF and its faint extension will constrain evolution to fainter absolute magnitudes and more distant redshifts. However, although the faint multislit work such as LDSS and LDSS-2 is more costly, it is more rewarding. The evolution of galaxies is driven at the faint end of the luminosity function. Probing spectroscopically yet fainter apparent magnitudes is almost exponentially more expensive, but understanding the distribution of galaxies at yet higher redshifts is the only direct way to constrain the evolution of the population.

Only during the past few years with the advent of multiplexing spectrographs and semianalytic models for galaxy formation has it been possible to so directly compare galaxy formation theory and observations. Not only does this comparison deepen our understanding of the physics of galaxy formation, it may be possible to unravel the effects of galaxy evolution and cosmology to determine the fundamental features of our universe: its age, its fate and its composition. The next few years will be a most exciting

time in observational and theoretical cosmology as newer instruments (such as the 2dF on the AAT), more clever analysis, faster computers, and more ingenious modeling begin to reap rewards.

Acknowledgements

I thank my supervisor, Professor Richard Ellis, for entrusting the AUTOFIB catalogue and spectra, and their analysis to me, and for inspiring this thesis in a discussion more than two years ago in the visitor's office of Peyton Hall. I thank Matthew Colless and Thomas Broadhurst for many valuable discussions all over the world.

I thank Carlos Frenk for suggesting that I explore galaxy formation as a function of cosmology. I thank Shaun Cole for developing the "block" model and helping me through the intricacies of the software. I also thank Julio Navarro for direction during the completion of this project, and Steve Zepf and Alfonso Aragón-Salamanca for their help in defining the rules and observational tests of the "block" model. I thank all of the CAFNZ gang for many involved and enlightening conversations.

I thank the Marshall Aid Commemoration Commission which provided me financial support during the past two years and Diane Balestri who suggested the prospect of studying in Great Britain.

I thank the Shincliffers (Rafael, Francisco, Chris, Jason, Steve, Amy, Ramón, Kathryn, Guy, Richard and Rebecca ...) and the Durmites (Ghadir, Isabelle, Vicki, Roger, Jane ...) — for listening to my crazy ideas and a great year in the Northeast. And in Cambers, I thank Kester, George, Julia, Francisco again ... for fun on the Fens.

I thank Jenny for helping during what could have been the most difficult year.

Bibliography

Arnaud, M., Rothenflug, R., Boulade, O., Vigroux, L. & Vangioni-Flam, E. 1992, *A&A,* **254**, 49.

Ashman, K. M. 1992, *PASP,* **104**, 109.

Avni, Y. & Bahcall, J. N. 1980, *ApJ,* **235**, 694.

Babul, A. & Rees, M. J. 1991, *MNRAS,* **255**, 346.

Bahcall, J. N. 1984, *ApJ,* **276**, 169.

Bernstein, G. 1994, private communication

Bhavsar, S. P. 1990, in C. Jaschek & F. Murtagh (eds.), *Errors, Bias and Uncertainties in Astronomy*, p. 107, Cambridge Univ. Press, Cambridge

Binggeli, B., Tammann, G. A. & Sandage, A. 1985, *AJ,* **90**, 1681.

Bond, J. R., Cole, S., Efstathiou, G. & Kaiser, N. 1991, *ApJ,* **379**, 440.

Bower, R. G. 1991, *MNRAS,* **248**, 332.

Boyle, B. J., Fong, R., Shanks, T. & Peterson, B. A. 1990, *MNRAS,* **243**, 1.

Broadhurst, T. J., Ellis, R. S., Koo, D. C. & Szalay, A. S. 1990, *Nature,* **343**, 726.

Broadhurst, T. J., Ellis, R. S. & Shanks, T. 1988, *MNRAS,* **235**, 827.

Bruzual, G. & Charlot, S. 1993, *ApJ,* **405**, 538.

Butcher, H. & Oemler, A. 1978, *ApJ,* **226**, 559.

Cen, R. & Ostriker, J. P. 1993, *ApJ,* **414**, 407.

Chołoriewski, J. 1986, *MNRAS,* **223**, 1.

Cole, S. 1991, *ApJ,* **367**, 45.

Cole, S. & Kaiser, N. 1988, *MNRAS,* **223**, 637.

Cole, S. M., Aragón-Salamanca, A., Frenk, C. S., Navarro, J. F. & Zepf, S. 1994a, *A Recipe for Galaxy Formation*, Durham Preprint

Cole, S. M., Ellis, R. S., Broadhurst, T. J. & Colless, M. 1994b, *MNRAS,* **267**, 541.

Colless, M. 1989, *MNRAS,* **237**, 799.

Colless, M., Ellis, R. S., Broadhurst, T. J., Taylor, K. & Peterson, B. 1993, *MNRAS,* **261**, 19.

Colless, M., Ellis, R. S., Taylor, K. & Hook, R. N. 1990, *MNRAS,* **244**, 408.

Cowie, L. L., Lilly, S. J., Gardner, J. & McLean, I. S. 1988, *ApJL,* **332**, L29.

David, L. P., Forman, W. & Jones, C. 1991, *ApJ,* **380**, 39.

Davies, J. I. 1990, *MNRAS,* **244**, 8.

Davies, J. I., Disney, M. J., Phillips, S., Boyle, B. J. & Couch, W. J. 1994, *MNRAS,* **269**, 349.

Davis, M., Efstathiou, G. P., Frenk, C. S. & White, S. D. M. 1985, *ApJ,* **292**, 371.

Davis, M., Summers, F. J. & Schlegel, D. 1992, *Nature,* **359**, 393.

de Lapparent, V., Geller, M. J. & Huchra, J. P. 1989, *ApJ,* **343**, 1.

Disney, M. J. 1976, *Nature,* **263**, 573.

Driver, S. P. 1994, *The Non-Evolving Universe and Faint-Galaxy Number Counts*, Cardiff Preprint

Eales, S. 1993, *ApJ,* **404**, 51.

Efron, B. 1979, *Ann. Stat.,* **7**, 1.

Efstathiou, G. 1990, in R. Wielen (ed.), *Dynamics and Interactions of Galaxies*, p. 2, Springer, Berlin

Efstathiou, G. 1992, *MNRAS,* **256**, 43P.

Efstathiou, G., Ellis, R. S. & Peterson, B. A. 1988, *MNRAS,* **232**, 431.

Efstathiou, G., Kaiser, N., Saunders, W., Lawrence, A., Rowan-Robinson, M., Ellis, R. S. & Frenk, C. S. 1990, *MNRAS,* **247**, 10P.

Ellis, R. S., Colless, M., Broadhurst, T. J., Heyl, J. S. & Glazebrook, K. 1994, *The Evolution of the Galaxy Luminosity Function*, in preparation

Ellis, R. S. & Parry, I. R. 1988, in L. B. Robinson (ed.), *Instrumentation for Ground-Based Optical Astronomy*, Vol. 9 of *Santa Cruz Summer Workshop in Astronomy and Astrophysics*, p. 192, Springer-Verlag, New York

Evrard, A. E., Summers, F. J. & Davis, M. 1994, *ApJ,* **422**, 11.

Fabian, A. C. 1991, *MNRAS,* **253**, 29.

Feldman, H. A., Kaiser, N. & Peacock, J. A. 1994, *ApJ,* **426**, 23.

Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A. & Huchra, J. P. 1993, *ApJ,* **402**, 42.

Flores, R. A., Primack, J. P., Blumenthal, G. R. & Faber, S. M. 1993, *ApJ,* **412**, 443.

Frenk, C. S., White, S. D. M., Davis, M. & Efstathiou, G. 1988, *ApJ,* **327**, 507.

Gardner, J. P., Cowie, L. L. & Wainscoat, R. J. 1993, *ApJL,* **415**, L9.

Glazebrook, K., Ellis, R. S., Colless, M., Broadhurst, T. J., Allington-Smith, J., Tanvir, N. R. & Taylor, K. 1993, *A Faint Galaxy Redshift Survey to B=24*, Preprint

Glazebrook, K., Peacock, J. A. & Collins, C. A. 1994, *MNRAS,* **266**, 65.

Green, R. 1989, in C. S. Frenk *et al.* (eds.), *The Epoch of Galaxy Formation*, p. 121, Kluwer, Dordrecht

Heckman, T. M., Armus, L. & Miley, G. K. 1990, *ApJS,* **74**, 833.

Hewitt, P. P., Foltz, C. B. & Charree, F. H. 1993, *ApJL,* **406**, L43.

Heydon-Dumbleton, N. H., Collins, C. A. & MacGillivray, H. T. 1989, *MNRAS,* **238**, 379.

Heyl, J. S., Hernquist, L. & Spergel, D. N. 1994, *ApJ,* **427**, 165.

Impey, C., Bothun, G. & Malin, D. 1988, *ApJ,* **330**, 634.

Jacobi, G. H. *et al.* 1992, *PASP,* **104**, 599.

Jones, L. R., Fong, R., Ellis, R. S. & Peterson, B. A. 1991, *MNRAS,* **249**, 481.

Kaiser, N., Efstathiou, G., Ellis, R. S., Lawrence, A., Rowan-Robinson, M. & Saunders, W. 1991, *MNRAS,* **252**, 1.

Katz, N., Hernquist, L. & Weinberg, D. H. 1992, *ApJ,* **399**, 109.

Kauffmann, G., Guiderdoni, B. & White, S. D. M. 1994, *MNRAS,*, submitted

Kauffmann, G. & White, S. D. M. 1993, *MNRAS,* **261**, 921.

Kauffmann, G., White, S. D. M. & Guiderdoni, B. 1993, *MNRAS,* **264**, 201.

Kennicutt, R. 1992a, *ApJS,* **79**, 255.

Kennicutt, R. 1992b, *ApJ,* **388**, 410.

King, C. R. & Ellis, R. S. 1985, *ApJ,* **288**, 456.

Klypin, A., Holtzman, J., Primack, J. & Regos, E. 1993, *ApJ,* **416**, 1.

Koo, D. C. & Kron, R. G. 1992, *ARAA,* **30**, 613.

Lacey, C. & Cole, S. 1993, *MNRAS,* **262**, 627.

Lacey, C. & Cole, S. 1994, *MNRAS,*, in press

Lacey, C. G., Guiderdoni, B., Rocca-Volerange, B. & Silk, J. 1993, *ApJ,* **402**, 15.

Lacey, C. G. & Silk, J. 1991, *ApJ,* **381**, 14.

Lanzetta, K. M., Wolfe, A. M. & Turnshek, D. A. 1994, *ApJ,*, in press

Lauer, T. R. 1985, *ApJ,* **292**, 104.

Lilly, S. J. 1993, *ApJ,* **411**, 501.

Lilly, S. J., Cowie, L. L. & Gardner, J. P. 1991, *ApJ,* **369**, 79.

Loveday, J., Peterson, B. A., Efstathiou, G. & Maddox, S. J. 1992, *ApJ,* **390**, 338.

Lucey, J. R., Guzman, R., Carter, D. & Terlevich, R. J. 1991, *MNRAS,* **253**, 584.

Lynden-Bell, D. 1971, *MNRAS,* **155**, 95.

MacGillivray, H. T., Beard, S. M. & Dodd, R. J. 1988, in ? (ed.), *Astronomy from large databases: Scientific objectives and methodological approaches; Proceedings of the Conference, Garching, Federal Republic of Germany*, p. 389, European Southern Observatory

Maddox, S. J., Efstathiou, G., Sutherland, W. J. & Loveday, J. 1990a, *MNRAS,* **242**, 43P.

Maddox, S. J., Sutherland, W. J., Efstathiou, G., Loveday, J. & Peterson, B. A. 1990b, *MNRAS,* **247**, 1P.

Marshall, H. L., Avni, Y., Tananbaum, N. & Zamorani, G. 1983, *ApJ,* **269**, 35.

Marzke, R. O., Huchra, J. P. & Geller, M. J. 1994, *ApJ,* **428**, 43.

McGaugh, S. S. 1994, *The Number Density of Low Surface Brightness Galaxies*, IoA Preprint

Metcalfe, N., Shanks, T., Fong, R. & Jones, L. R. 1991, *MNRAS,* **249**, 498.

Mobasher, B., Ellis, R. S. & Sharples, R. M. 1986, *MNRAS,* **223**, 11.

Mobasher, B., Sharples, R. M. & Ellis, R. S. 1993, *MNRAS,* **263**, 560.

Moore, B., Frenk, C. S., Efstathiou, G. P. & Saunders, W. 1994, *MNRAS,*, in press

Navarro, J. F., Frenk, C. S. & White, S. D. M. 1994, *Blah*, in press

Navarro, J. F. & White, S. D. M. 1993, *MNRAS,* **265**, 271.

Nobelis, P. P. 1990, in C. Jaschek & F. Murtagh (eds.), *Errors, Bias and Uncertainties in Astronomy*, p. 89, Cambridge Univ. Press, Cambridge

Parry, I. R. & Sharples, R. M. 1988, in S. C. Barden (ed.), *Fiber Optics in Astronomy*, Vol. 3 of *Astronomical Society of the Pacific Conference Series*, p. 93, Astronomical Society of the Pacific, San Francisco

Pence, W. 1976, *ApJ,* **203**, 39.

Persic, M. & Salucci, P. 1992, *ApJ,* **360**, 68.

Peterson, B. A., Ellis, R. S., Efstathiou, G., Shanks, T., Bean, A. J. & Fong, R. 1986, *MNRAS,* **221**, 233.

Press, W. H., Rybicki, G. B. & Schneider, D. P. 1993, *ApJ,* **414**, 64.

Renzini, A. 1986, in B. F. Madore & R. B. Tully (eds.), *Galaxy distances and deviations from universal expansion; Proceedings of the NATO Advanced Research Workshop, Kona, HI, Jan. 13-17, 1986*, p. 177, D. Reidel Publishing Co.

Sandage, A. 1993, *AJ,* **106**, 719.

Sandage, A., Tammann, G. A. & Yahil, A. 1979, *ApJ,* **232**, 352.

Saunders, W., Frenk, C., Rowan-Robinson, M., Lawrence, A. & Efstathiou, G. 1991, *Nature,* **349**, 32.

Saunders, W., Rowan-Robinson, M., Lawrence, A., Efstathiou, G., Kaiser, N., Ellis, R. S. & Frenk, C. S. 1990, *MNRAS,* **242**, 318.

Scalo, J. M. 1986, *Fundam. Cosmic Physics,* **11**, 1.

Schechter, P. 1976, *ApJ,* **203**, 297.

Schmidt, M. 1968, *ApJ,* **151**, 393.

Sharples, R. 1989, *UM 26.1: Autofib User Manual*, Anglo-Australian Observatory

Smoot, G. *et al.* 1992, *ApJ,* **396**, 1.

Szalay, A. S., Ellis, R. S., Koo, D. C. & Broadhurst, T. J. 1991, in *Primordial nucleosynthesis and evolution of early universe; Proceedings of the International Conference, Tokyo, Japan, Sept. 4-8, 1990*, p. 435

Toth, G. & Ostriker, J. P. 1992, *ApJ,* **389**, 5.

Treyer, M. A. & Silk, J. 1993, *The Faint End of the Galaxy Luminosity Function*, Preprint CfPA-TH-93-26

Tyson, J. A. 1988, *AJ,* **96**, 1.

Walker, T. P., Steigman, G., Schramm, D. N., Olive, K. A. & Kang, H.-S. 1991, *ApJ,* **376**, 51.

White, S. D. M. 1990, in R. Wielen (ed.), *Dynamics and Interactions of Galaxies*, p. 380, Springer, Berlin

White, S. D. M., Efstathiou, G. & Frenk, C. S. 1993, *MNRAS,* **262**, 1023.

White, S. D. M. & Frenk, C. S. 1991, *ApJ,* **379**, 52.

White, S. D. M. & Rees, M. J. 1978, *MNRAS,* **183**, 341.

Young, P. *et al.* 1994, in preparation

Zaritsky, D., Smith, R., Frenk, C. S. & White, S. D. M. 1993, *ApJ,* **405**, 464.

Appendix A

The Software Library

A.1 Observational and Analysis Software

A.1.1 Figaro Applications

The Figaro application are located in the directory:
`/home/jsheyl/figaro`.

- `bin2dst` *binary-file dst-file*

  converts a spectral file in Fortran binary format into Figaro Format

- `calcolor` *spect-dst calib-dst output*

  calculates the ratio of integrated fluxes in two bands, centered on 4,250 Å and 5,050 Å, given the series of spectra in *spect-dst* and the calibration curve in *calib-dst*. The list of colours is output to *output*.

- `coadd` *input-dst add-dst*

  takes the mean of two spectra in "coadded" format, taking into account the weighting of each pixel. It overwrites *input-dst*.

- `dst2bin` *dst-file binary-file*

  converts a spectral file in Figaro format into Fortran binary format.

- `fake` *output-dst norm-1000 power*

  creates a fake power-law spectrum in Figaro format. The value of the flux at 1,000 Å is given in *norm-1000* and the power-law slope is *power*.

- `getbreak` *spect-dst red-file output*

  calculates the 4,000 Å break of the spectra in *spect-dst*. The user must supply the redshifts of the galaxies in *red-file*, so the program can shift the filter into the rest frame, giving the results in *output*

- `jspflux` *spect-dst calib-dst output-dst* fluxes a 2-D array of spectra, given a calibration spectrum.

- `jvachel` *input-dst redshift spectrum-number*

  shifts the spectrum into the rest frame and creates a new spectrum suitable for coadding. The spectral range of the resulting spectrum is 2,000 Å to 8,000 ÅThe spectrum is contained in the first row of the Figaro file, and the number of pixels used to calculate a single pixel is contained in the second row (*i.e.* the weight of each pixel).

The non-Figaro applications are located in the directory `/home/jsheyl/lumin/src`.

A.1.2 K-corrections and filters

- `dogetkc` *spectrum-files*

  calculates the k-correction for all of the spectrum files from $z = 0$ to $z = 0.75$ with 0.005 steps in redshift using the b_J filter.

- `getkc` *spectrum-file spectrum-number z-max step filter*

  calculates the k-correction of the given spectrum at a variety of redshifts from 0 to 0.5 and fits a line to this function. The given spectrum must already be in the restframe.

A.1.3 Classification by features

Technique This is the old-style analysis, using the spectra themselve to provide the k-corrections in a sense. The first step is to calibrate all of the DST files to be classified using the command `docal`. Next calculate the break strengths using `dobreak`. The penultimate step is to add the information about the break strengths with equivolent width data with `makeclass`. Finally, `makeflis.awk` will construct a `.lis` file from these data. An example series of commands is:

```
set file=spectra
docal
dobreak
makeclass spectra.cat spectra_bre spectra.cla
makeflis.awk spectra.cat spectra.cla > spectra.lis
```

which has produced a LIS-file, `spectra.lis`. To use this LIS-file with the luminosity function software, one must use the software `lis2eales` and `sswml` directly and provide the correct information for the bootstraped k-correction files: `/home/jsheyl/lumin/class/class`.

Programs

- `classspec`

  classifies all the spectra in the catalog given by the environment variable `file`, shifts each spectrum into its rest frame, and places it into a subdirectory containing only galaxies of its class.

- `dobreak`

  calculates $D_{4000\mathring{A}}$ for all the calibrated DST files in the current directory.

- `docal`

  calibrates the spectra in the DST file given by the environment variable `file`. This script assumes that the reddest galaxy in the DST file is an elliptical and then use this information to construct a calibration curve.

- `editcal`

  edits interactively the calibration curve for the DST file given by the `file` environment variable.

- `makeclass` *cat-file break-file class-file*

  given a *cat-file* and a list of the break strengths (as provided by *get-break*) of the galaxies in the *cat-file*, creates a list of classes for the galaxies in the *class-file*, in fiber number order. The file contains the following data:

  - Fiber number
  - The value of $D_{4000\mathring{A}}$
  - The integral over the red section of the break
  - The integral over the blue section of the break
  - The redshift of the galaxy
  - The equivalent width of the 3,727 Åline
  - The number of the class (0 - 4), with four meaning that the galaxy cannot be classified for some reason

- `makeflis.awk` *cat-file class-file*

  given a *cat-file* and a *class-file* this will create a `.lis` file with the absolute magnitudes of the galaxies with k-corrections.

- `prespec`

  prepares all the spectra in the current directory to be coadded. That is, it converts them from Figaro format into Fortran binary format.

- `squish` *do-norm output-file spectrum-files* coadds a group of spectrum files in Fortran binary format, producing a new spectrum with the following data:
  - The mean of the pixel values (for a given wavelength) from the original spectrum files
  - The deviation of the mean (at a given wavelength)
  - The mean of the middle 80% (by weight) of the pixels
  - The mean of the 10% to 30% range (by weight)
  - The mean of the 70% to 90% range
  - The total weight of a given pixel

  If *do-norm* is equal to one, `squish` will renormalise the spectra so as to minimise the mean dispersion in the pixel values over wavelength.

A.1.4 Classification by cross-correlation

Technique The standard method as discussed in Chapter 3 starts with a DST file and a CAT file for each subcatalogue. First, run the command `splitdstline blah.cat` to split the DST file into rest-frame spectra of each galaxy. Second, run `makelisk blah.cat` which will classify each of the spectra against the Kennicutt templates, convert from Kennicutt type to k-correction class and output the results in a file called, in this case, `blah.lisk`. The script `makelisk` uses the programs `findobj` and `makelis.awk`.

There are variations to this technique which use the spectrum with the continuum and attempt to find the best fit fluxing curve and the best-fitting galaxy templates simultaneously. To use this method, run the command `splitdst blah.cat` and `makelis blah.cat`. This by default will use the Pence spectra as templates. The script `makelis` uses the programs `findgal`, `findflux`, `findobj` and `makelis.awk`.

Programs

- `findflux` *N-Flux-Param Tolerance Input-File*

  finds the best fitting fluxing curve to match the spectrum and class files given in *Input-File*. *Input-File* must be in the form as the outputted by `findobj` or `findbest`.

- `findgal` *Class-Files - Spectrum-Files*

  iteratively finds the best fluxing curve and the best template spectrum from amongst the *Class-Files* for the *Spectrum-Files*. It assumed that the fluxing curve is a seven-term polynomial and that all of the spectrum files share the same fluxing curve.

- `findobj` *N-fit Fit-Param-File N-Skip Class-Files - Spectrum-Files*

  finds the best fitting templates from the *Class-Files*, assuming that the *Spectrum-Files* all have the same fluxing curve (a polynomial with *N-fit* terms given in `Fit-Param-File`. Sends the best fitting results to standard output.

- `ken2pen.awk` *res2k-file*

  converts a `res2k` file to a `res2p` file by substituting the Kennicutt templates with their appropriated k-correction classes.

- `makelis` *cat-files* generates from a `CAT` file and a series of spectrum files in binary format a `LIS` containing the classifications and k-corrections of all the galaxies, using a cross- correlation against the product of the Pence spectra and an arbitray fluxing curve, which is determined iteratively.

- `makelis.awk` *cat-file res-file*

  generates a file in the format of a `LIS` file from a `CAT` file and `res` file (the output of either `findgal` or `ken2pen.awk`).

- `makelisl` *cat-files*

  generates from a `CAT` file and a series of spectrum files in binary format a `LIS` containing the classifications and k-corrections of all the galaxies, using a direct cross-correlation against the Pence line spectra.

- `makelisk` *cat-files*

  generates from a `CAT` file and a series of spectrum files in binary format a `LIS` containing the classifications and k-corrections of all the galaxies, using a direct cross-correlation against the Kennicutt line spectra. The results of this analysis are converted to a k-correction class and a `LIS` file is created.

- `makelisnc` *cat-files*

  generates a bunch of `LIS` files from a series of `CAT` files without re-classifying the galaxies.

- `splitdst` *cat-files*

  split a single `DST` into a series of rest-frame `BIN` files for each of the galaxies in the *cat-files*.

- `splitdstline` *cat-files*

  split a single `DST` into a series of continuum-substracted rest-frame `BIN` files for each of the galaxies in the *cat-files*.

A.1.5 Number counts

Technique Once the galaxies have been observed and the sampling rates determined, calculating the number counts from the enitre coherent catalogue is straightforward. If the sampling data is in `samples.inf` and the catalogue in `cats.lis`, the number counts can be calculated using the command:

```
lis2counts.awk cats.lis samples.inf | countshist.awk > counts
```

The errors may be determined by constructing an intermediate file (*e.g.* `cats.counts`) and using `bootstrap.awk` to construct several bootstrap resampling of this file, and then pass all the files along to `countshist.awk`. The deviation in the various results is the error estimate (*c.f.* `lfeales` below).

Programs

- `countshist.awk` *counts-file*

  converts the results of `lis2counts.awk` to a number counts histgram.

- `lis2counts.awk` *lis-file inf-file*

  calculates the total area from which a galaxy in the *lis-file* could have been sampled from the sampling data in the `inf-file`. The result is a list of apparent magnitude and sampling areas, which may be summed in bins to get the number counts (see `countshist.awk` above.)

A.1.6 Luminosity function

Technique The input for the luminosity function routines is a LIS file containing all the samples to be added coherently and an INF file containing information about the samples. The script `makelf` will calculate the luminosity function with errors in four redshift bins: $0-2$, $0-0.2$, $0.2-0.5$ and $0.5-1$ using both the SSWML method and the $1/V_{\max}$ methods with the Pence k-correction curves.

Programs

- `eales2lf.awk` *eales-file M-min M-max Nm*

  sums the inverse of the volumes listed in the *eales-files*, in bins of absolute magnitude, yielding an estimate of the luminosity function.

- `lis2eales` *inf-file k-c-header n-kc lis-file z-min z-max*

  calculates the total volume from which a galaxy in the *lis-file* could have been sampled from the sampling data in the `inf-file` within the given redshift range. The k-correction files are numbered from 0 to *n-kc* minus one (in C-style). For example, if *k-c-header* is `/home/fred/class` and *n-kc* is `4`, the k- correction files are:

```
/home/fred/class0.kc
/home/fred/class1.kc
/home/fred/class2.kc
/home/fred/class3.kc
```

The result is a list of absolute magnitude and sampling volumes, which may be summed in bins to get the number counts (see `eales2lf.awk` above). Compare with `lis2counts.awk`.

- `lfeales` *eales-file boot-count M-min M-max nm*

  calculates the luminosity function from a single *eales-file* placing the output in a file ending with `.lf` It will also generate a series of `boot-count` realisations of the catalogue and calculate the luminosity function for each of these – these files named `??_n.lf`, where `n` is the number of the bootstrap may be used to estimate the errors in the luminosity function determination. Finally, it will sum over the bootstraps, yielding an `.elf` file containing the magnitude bins in absolute magnitude, $\phi(M)$ and the error in $\phi(M)$.

- `lumfunk` *tolerance ilf-file lis-files*

  calculates the luminosity function using the stepwise maximum likelihood technique or the STY method. An initial trial luminosity function, a group of LIS files, and a group of files with the lower and upper magnitude limits for each LIS file must be provided. The magnitude files should have the same names as the LIS files with the extension "inf".

  If the parameter *tolerance* is less or equal to zero, `lumfunk` will find the most likely Schetcher function from the data with the value of *tolerance* now being an estimate of the magnitude errors in the survey.

- `makelf` *lis-file inf-file*

  calculates the luminosity function in four redshift bins: $0-2$, $0-0.2$, $0.2-0.5$ and $0.5-1$ using both the SSWML method and the $1/V_{\rm max}$ methods. This routine automatically uses the Pence k-correction files in `/home/jsheyl/lumin/galaxies/pence?.kc`.

- `schfit` *ELF-file*

  calculates the best-fitting function to a binned luminosity function with errors. The first column of the *ELF-file* must contain the luminosity or absolute magnitude of the bin, the second column must have the density of galaxies within the bin ($\phi(L)$ or $\phi(M)$) and the third column must have the error on this density.

- `ssty` *inf-file kc-file lis-file [n-output]*

calculates the most-likely evolving luminosity function as defined by

$$\begin{aligned} \phi^*(z) &= \phi_0^*(1+z)^{\phi_z^*} \\ M^*(z) &= M_0^* - 2.5L_z^*\log_{10}(1+z) \\ \alpha(z) &= \alpha_0 + \alpha_z z. \end{aligned}$$

using the generalised STY method (Equation 4.7). Compare this with `lumfunk` in STY mode. This program, by design, uses the same k-correction curve for all the galaxies in the survey, so the *lis-file* should contain only galaxies of a particular k-correction class and the analysis repeated for each galaxy type. *n-output* is an optional parameter which tells `ssty` how often to output the current values of the evolving luminosity function. The programs output is to standard output, and the most-likely luminosity functions are given as follows

```
\phi^*_{0,L} = 0.00769459
\phi^*_z = -2.90301
M^*_0 = -20.5772
L^*_z = 1.75626
\alpha_0 = -0.951837
\alpha_z = 1.69353
```

where the values of ϕ_0^* and α_0 are for the function $\phi(L)$, although the fitting takes place in magnitude space.

- `sswml` *inf-file k-c-header n-kc lis-file out-file z-min z-max nz M-min M-max nm*

  calculates the SSWML luminosity function from the *inf-file* and *lis-file* with errors and upper-limits, over *nz* bins in redshift and *nm* bins in apparent magnitude. The k-correction files are given by *k-c-header* and *n-kc* as in `lis2eales`. The resulting *out-file* contains rows for each absolute magnitude bin and three columns for each redshift bin. The first column of the file contains the central absolute magnitude of each bin. The next three columns contain the logarithm of the density ($\phi(M)$) in the bin, the number of galaxies in the bin, and the error in the logarithm of the density. If no galaxies were observed in the bin, the density column has a value of -100.00 and the error column contains an upper- limit in the logarithm of the density for the given bin. This upper-limit will be 100.00 if the survey has no sensitivity in the bin.

A.1.7 Fakes

- `fakelf` *ilf-file seed number error inf-files*

  generates a faked `.lis` file with the *number* galaxies following the luminosity function given in *ILF-file* (L and $\phi(L)$). The *inf-file* contains

the bright and faint magnitude limits for the sample (in that order). If the second limit is less than the first, it will generate a volume-limited numerical realisation of the luminosity function. Magnitude errors with σ =*error* may be included. `seed` is the random number seed for the simulated `.lis` file. The area of sky sampled in the survey is given at the end of the `.lis` file.

- `fakespec` *N-fakes Seed Signal-Level Sky-Level Class-Files*

  generates a series of *N-fakes* spectra selecting randomly from the *Class- files*, normalising the spectra to a mean *Signal-Level* counts/bin and adding *Sky-Level* counts per bin of sky. Poisson noise is added to the product of the spectum and a moving responce function. The responce function is divided out yielding a fluxed spectrum with noise increase toward the edges. The response function moves with the red-shift of the simulate spectrum which is selected randomly from 0 to 0.6.

A.1.8 Astrometry and configuring

- `fieldcalc.awk` *fld-file*

  counts the number of objects within 40 arc minutes of the centre of the *fld-file*.

- `fieldextract.awk` *fld-file*

  extracts only those objects in the *fld-file* within 40 arc minutes of the centre of the field.

- `fld2ast` *field-name suffix*

  converts the given field file into an `.ast`, suitable to be used by the `ASTROM` program.

- `fld2dat` *field-name suffix*

  converts the given field file to a `.dat` file, for the `CHART` program.

- `log2ast` *field-name suffix*

  converts the given `.log` file into an `.ast`, suitable to be used by the `ASTROM` program.

- `log2dat` *field-name suffix*

  converts the given `.dat` file to a `.dat` file, for the `CHART` program.

- `newfld` *field-name suffix*

  takes the log file created by CONFIGURE and creates a new FLD file from the old one without the already observed galaxies

- `newfld2` *field-name suffix*

  takes the log file created by CONFIGURE and creates a new FLD file from the old one with only the already observed galaxies

- `readcos` *cos-file output-file*

  coverts a `COSMOS` file to ASCII format.

A.1.9 Matching

- `cat2ang.awk` *cat-file*

  creates a file with eight columns containing the R.A. and declination in degrees, the redshift, apparent magnitude, the subcatalogue number, and the x, y, z positions of the galaxy in redshift space.

- `cat2can.awk` *cat-file*

  creates a file with three columns containing the R.A. and declination in radians, and the apparent magnitude of each object

- `domatch` *cat-file distance*

  finds all objects in the `.fld` file assocated with *cat-file* with *distance* radians of each object in the `cat-file`. It produces a file `.mat` which contains the position and magnitude within each file of the matches.

- `fld2fan.awk` *field-name suffix*

  creates a file with three columns containing the R.A. and declination in radians, and the apparent magnitude of each object, for use with matchpos.

- `matchpos` *primary-file x1 y1 m1 secondary-file x1 y1 m2 tolerance*

  finds all objects in the *secondary-file* within *tolerance* radians of objects in the primary file. *x1*, *y1* and *m1* give the columns which contain the R.A., declination and a third column to compare on output (usually the apparent magnitude) in the first file. The second set give the columns in the second file.

- `pairmatch` *[cat-file or fld-file] distance*

  matches the objects in the `.cat` or `.fld` file against itself to generate a histogram the distance between all the pairs in the file in a `.cpair` or `.fpair` file.

A.1.10 Utility Software

- `addclass.awk` *cat-file lis-file*

  adds a column to the *cat-file* containing the classification from the *lis-file*. It outputs to standard output.

- `bootstrap.awk` *file-name random-seed*

  generates a bootstap resampling of the lines (without hash-marks) of *file-name* to the standard output.

- `calceffarea.awk` *inf-file option* calculates the effective area column of the *inf-file* using one of two methods. If *option* is equal to `2`, it will calculate the effective area by scaling to the observed number counts (using a fitting formula). If *option* is not equal to `2` or if it is absent, it will calculate the effective area using the sampling rate and completeness.

- `calcmeanz.awk` *lis-file*

  calculates the mean redshift of the galaxies in *lis-file* in three redshift bins ($0-0.1$, $0.1-0.3$ and $0.3-1.0$) and for the catalogue as a whole.

- `cat2lis.awk` *cat-file*

  converts a *cat-file* to a `.lis` file without applying k-corrections. Output to standard output.

- `catc2lis.awk` *cat-file*

  converts a *cat-file* to a `.lis` file. It applies k-corrections based on the classes given in the final column of the *cat-file*, as added by `addclass.awk` above. Output to standard output.

- `checkcats.awk` *check-file inf-file*

  checks the values *inf-file* against those in the *check-file* which is produced by `sumcat.awk` below. For example, to check the values in `cats.inf` against the observed catalogue `all.cat`, use the following two commands:

  ```
  sumcat.awk all.cat | checkcats.awk - cats.inf
  ```

  and a report of the errors will appear.

- `checklis.awk` *lis-file cat-file*

  checks that all of the objects with observed redshifts less than two in the `cat-file` are also in the *lis-file*. It also checks that no object with an observed redshift less than two appears more than once in either file.

- `compweight.awk` *inf-file lis-file* performs the magnitude-dependent completeness correction on the *lis-file* with output to standard output, given the completeness rates in the *inf-file*.

- `cutlis.awk` *lis-file B-min B-max*

  removes all objects in the *lis-file* in outside the given magnitude range, updates the columns `Bmin`, `Bmax`, `zmin` and `zmax` in the file. Output to standard output.

- `gammq` *Degrees-Of-Freedom Chi-Squared*

  given χ^2 and the degrees of freedom, calculates the probability that two samples are from the same parent distribution.

- `getsubcat.awk` *lis-file include-file*

  extracts from *lis-file* only those the objects in subcatalogues listed in the *include-file*, one per line.

- `lissummary.awk` *lis-file*

  calculates the mean redshift, apparent and absolute magnitudes of the galaxies in *lis-file* by k-correction class.

- `randclass.awk` *lis-file*

  reclassify randomly 20 % of the galaxies in the *lis-file* by one k-correction class redward or blueward. Results to standard output.

- `reckc.awk` *lis-file*

  recalculates the k-correction for all the galaxies in *lis-file*, as well as the minimum and maximum redshifts. Results to standard output. To be used if one of the k-correction files has changed.

- `setkc.awk` *lis-file constant*

  sets the k-correction of all the galaxies in *lis-file* to the product of *constant* and the redshift, and recalculates the k-corrections and minimum and maximum redshifts. Results to standard output.

- `setclass.awk` *lis-file class*

  sets the k-correction class of all the galaxies in *lis-file* to *class* and recalculates the k-corrections and minimum and maximum redshifts. Results to standard output.

- `shiftlis.awk` *lis-file Delta-m* shifts the apparent, absolute, and magnitude limits of the *lis-file* by *Delta- m*. Results to standard output.

- `split.awk` *cat-file inf-file*

  splits the *cat-file* into subcatalogues using the files listed in the first column of the *inf-file*.

- `subcatsum.awk`

  calculates the mean redshift, apparent and absolute magnitudes and equivolent width the galaxies in *lis-file* by subcatalogue.

- `sumcat.awk` *cat-file*

  counts the total number of fibres, duds, stars, redshifts and unknowns in each subcatalogue within the *cat-file*. Also, it gives the completeness rate, and bright and faint limits.

- `uniform` *Number-Of-Samples Sample-Size Random-Number-Seed*

  calculates the mean of a sample of *sample-size* of uniform deviates. It repeats this *Number-Of-Samples* times. It is use to find the deviation of the mean of a uniform distribution from 0.5.

- `vmax.awk` *lis-file*

  calculates the $V/V_{\max}$ statistic for every object in the *lis-file* and send these results to standard output. It uses the columns $z_{\min}$ and $z_{\max}$ in the *lis-file*.

A.1.11 Location

The DST and CAT files (along with calibration curves *etc.*) are contained in the directories below `/home/jsheyl/scratch/lumin`. A given DST file (when available) may be found by the date of the run or "ldss" or "bes".

```
/home/jsheyl/scratch/lumin/bes
/home/jsheyl/scratch/lumin/dars
/home/jsheyl/scratch/lumin/ldss
/home/jsheyl/scratch/lumin/ldss2
/home/jsheyl/scratch/lumin/90Sept
/home/jsheyl/scratch/lumin/91Sept
/home/jsheyl/scratch/lumin/92April
/home/jsheyl/scratch/lumin/92Sept
/home/jsheyl/scratch/lumin/93May
```

A.2 Galaxy Formation Analysis Software

- `lis2lb.exe` reads a list of 200,000 luminosities from standard input and converts it to two identical (except for the header) `lb` files (`lb0` and `lb1`). The first file is at $z = 0$ and the second at $z = 100$. By renaming these files and setting up appropriate symbolic links, the files may be used by `makecat` as a test case.

- `makecat` *Sim-Dir N-Sim N-Red Upsilon N-Gal B-min B-max*

  and additional optional parameters,

  B-min B-max N-Bin N-gal

  for number-counts calculation.

  The first seven parameters are required and give the directory of the simulation to be analysed, the number of realisations per redshift, the number of redshifts, the value of Υ for the simulation, the number of galaxies in the survey (may be zero) and the bright and faint limits for the survey. The results of the survey will be sent to standard output. If the number of redshifts is less or equal to zero, the program with

assume that all galaxies have $M_B = -19.73$ and a constant density (this is a debugging feature).

If the number of simulations is less than zero, the program will use the K-band data instead of the B-band.

If any of the final four parameters are given, the program will require all four parameters. They give the bright and faint limits for the counts calculation, the number of bins in apparent magnitude, and the number of galaxies in each bin. The program will produce a file named either `bcounts.data` or `kcounts.data` in the current directory.

The software requires all the `lb` and `lbr`, or `lk` files for the simulation to be in the directory specified. Also, this directory must contain a file called `cosmo.info` which gives the cosmological information for the simulation. This file is produced by `processinc.awk` from `galaxies.inc`.

- `processinc.awk` *INCfile*

  parses the `INCfile` into a list of cards and values suitable to be a `cosmo.info` file. It outputs to standard output.

- `testcosmo` *Omega Lambda h*

  calculates the number counts for a Loveday et al. (1992) Schechter function for an arbitray universe where `h` is the Hubble constant divided by 100 km/s/Mpc. It sends the results to standard output.

Appendix B

AUTOFIB Spectra Reduction

The reduction of AUTOFIB spectra splits into five stages:

Cleaning Subtracting off the bias and removing cosmic rays.

Clipping Cutting the CCD frames to the proper size, making sure that wavelength increases from left to right, and rotating them so that the fibres run horizontally.

Fibre Extraction Converting the raw CCD frames into an array of spectra.

Wavelength Calibration Calibrating the arc frames and copying wavelengths onto the object frames.

Sky Subtraction Subtracting the sky from each object frame and copying the wavelength calibration onto the final DST file.

B.1 Cleaning

The cleaning stage consists of two parts: removing the bias and removing cosmic rays. Before reaching this stage, you should verify that the bias is uniform across the CCD by checking the short-exposure bias frames. Once you've done this, you look for the overscan of the CCD which should be near one of the edges beyond the last fibre. You can do this by using the EXTRACT program to build a "spectrum" across the fibres for one sky frame and one object frame from each field. Focus in on the region beyond the last fibre and find the range of constant counts, past the light scattered from the last fibre. Then subtract the bias from each frame:

```
ISTAT 20MAY0019 MIN MAX 440 460
ICSUB 20MAY0019 'stat_mean' 20MAY0019DB
```

Instead of 440 460, you should give the range of the overscan that you discovered. Next remove the cosmic rays. The programs, CRSPOT and CRSKY, perform this task for object frames and sky frames respectively.

They both compare two frames and look for large differences between them to find the cosmic rays. If they find one, they simply copy the analogous region from the unblemished frame on top of the cosmic ray. Each program expects two input files and two output files, and a value of the comparison threshold: 250 is good for object frames and 100 for sky frames.

B.2 Clipping

The two CCD windows, `TEK_FIBRES` and `TEK_OFFSETSKY`, have slightly different dimensions, so the object and arc frames must first be clipped, using the `ISUBSET` program (*i.e.* `ISUBSET A MIN MAX 5 459 B`). Then the object, arc, and sky frames must be reversed in the X direction and rotated, using `IREVX` and `ROTATE`.

B.3 Fibre Extraction

There are two methods available in `FIGARO` for extracting fibres from the CCD frame. Only one may be available at your site. The first method was developed specifically for AUTOFIB, while the second for echelle spectra. Although the first is the "standard" method, the second is better documented (*e.g.* in the AAO document *The UCL Echelle Spectrograph*) and does an equally good job.

First Method To extract the fibres you will use three programs: `FINDSP` to find the best-fit polynomials, `OVERPF` to check the fits, and `POLEXT` to actually extract the fibres. Before doing the extraction you should look at the two "spectrum" files that you created to find the overscan. Plot these files again but now focus in on the first few fibres. Note the position of the centre of the first fibre and the separation between fibres. Sum (`ISUM`) up all of the sky frames and all the object frames. As the fibres shift little between exposures, this will increase the signal substantially when you're trying to fit the fibres. Run `FINDSP` on the total spectrum to define the best-fit fibre centroids. Fibres 7, 15, 17 and 23 are duds and must be deleted from the fits. Check the fits especially near the corners. Extract the fibres from all the object and arc frames. Repeat this process for the sky frames.

Alternative Method The echelle reduction software comes in three programs:

`SDIST` Determines the best-fit lines for a few spectra.

`CDIST` Straightens out the entire frame so that the spectra end up in horizontal rows.

`ECHSELECT` Sums up sets of rows on the CCD to produce a 2D spectra with one fibre per row.

The first step of the process must be repeated for each CCD window you have (*i.e.* once for the skies and once for the arcs and objects). Display (using `IMAGE`) an image showing the fibres clearlyg, either a flat field for the appropriate window or a sum of a few CCD frames. Use the `ICUR` program, select the centres of a few spectra throughout the image. Be sure that you have deleted or renamed the previous SDIST.DAT. Then run `SDIST`. Here's a sample:

```
$ SDIST
(IMage) Image containing distortion spectra [flat] -
(COLumns) Number of cols to average to get peaks [10] -
(TRace) Trace mode - G(aussian) C(OG) or E(dge) [g] -
(WIdth) Half width of spectra - in pixels [2] -
(MAXdeg) Maximum degree polynomial to use [10] - 2
(DIsplay) Use image display to show fit results? [YES] -
(SOft) Display results on graphics device? [NO] -
```

Next run `CDIST` to straighten out the spectra. Next you could run `ECHSELECT` directly which requires you to select the spectra with the mouse. Or run `YSTRACT` to get a fibre cross section of the image. Next run `FINDFIB` on this cross section which will simply give you a list of peaks in the cross section, their column positions, spacing and height. Edit this file removing spurious peaks, adding the dud fibres, abdinserting a line at the beginning containing the half width of the fibre (2 is a good choice). Run this list through `lis2ech.awk` to create `echselect.lis`:

```
# The awk line convert from a list of fibres to a file appropriate for
# ECHSELECT.  The first line should contain the half-width of each fibre
# and the following lines the fibre centres.
(NR==1) { width=$1
     printf("*\n*    Order selections from file: %s\n*\n",filename);
     printf("*    Created by LIS2ECH.AWK: HW=%.1f\n*\n",width); }
(NR>1) {  centre=$1
          fibno=NR-1
          for (colno=centre-width;colno<=centre+width;colno++) {
               printf("%9d%10d\n",colno,fibno);
          }
     }
```

Finally run `ECHSELECT` to see if everything has turn out right. Backup the current `sdist.dat` and `echselect.lis`.

Now you're ready to extract the spectra. Run each spectrum through `CDIST`:

```
cdist mt_xf_a5 ys=min ye=max out=mt_xf_a5x maxdegy=5
```

and then `ECHSELECT` to create the spectra.

This may be entirely automatic (although you will have to type q, enter the output file, and press return) if you want it to be, or you can check the success of the fits with `ECHSELECT` as it runs.

You must repeat this process entirely for the other CCD windows used.

B.4 Wavelength Calibration

Find the atlas for argon arc spectra before you think about continuing. This first arc is the hardest (if you have already calibrated an arc from the run, use it and its accompanying ARLINES.LIS as a starting point). Extract one fibre from the middle of the first arc frame (Fibre #31 is a good choice). Run ARC on this spectrum. If you already have an ARLINES.LIS file, use it. Otherwise identify the arc lines with the atlas and get a good fit. After completing the ARC program, rename the newly created ARLINES.LIS to ARC1.LIS. Repeat this process for the other arc frames, using ARC1.LIS as the starting point and renaming the ARLINES.LIS files successively.

Before calibrating each of the fibres, you must take care of the dud fibres. An easy way is to copy the adjacent fibres into the dud fibres:

```
extract mt_xf_a5x 6 6 mt_xf_a5f6
growx sp=mt_xf_a5f6 ys=7 ye=7 im=mt_xf_a5x nonew
extract mt_xf_a5x 14 14 mt_xf_a5f14
growx sp=mt_xf_a5f14 ys=15 ye=15 im=mt_xf_a5x nonew
extract mt_xf_a5x 16 16 mt_xf_a5f16
growx sp=mt_xf_a5f16 ys=17 ye=17 im=mt_xf_a5x nonew
extract mt_xf_a5x 22 22 mt_xf_a5f22
growx sp=mt_xf_a5f22 ys=23 ye=23 im=mt_xf_a5x nonew
```

Now copy the appropriate arc-list file (*e.g.* ARC1.LIS) to ARLINES.LIS and run IARC on the extracted arc frame (with the duds filled in). Repeat this process for each of the arc frames.

Now you must choose which arcs go with which object frames. Then you run ISCRUNCH to calibrate each of the object frames:

```
ISCRUNCH MT_XF_1X ARC1.IAR 3600 7700 1024 MT_XF_1S NOLOG NODEN
QUAD
```

B.5 Sky Subtraction

Before performing the sky subtraction the sky frames must be expanded to the size of the object frames using ISUPER and ISTRETCH. For example,

```
isuper mt_xf_s1x 1024 64 1 1 mt_xf_s1xt
istretch mt_xf_s1xt 16 1 0 0 1 1 mt_xf_s1xt
```

Now create a fibre information file for the program FRAME_ADD. The first line should contain the number of object frames followed by the number of sky frames that you want to use and the number of sky fibres. The following lines contain the range in object frames that you wish to use for each of the sky frames. Next give the numbers of each sky fibre, followed the numbers of each dud. Here's an example:

```
6 5 10
1 2
3 3
4 4
5 5
6 6
1
9
12
22
33
37
48
51
60
63
7
15
17
23
```

Now run FRAME_ADD. First give it the total number of frames (in our example 11), then the names of the object frames following by those of the sky frames (you can use the arrow key to use the previous entry as the basis for the current entry – a major time-saver). Finally give the name of the output file. FRAME_ADD doesn't copy the wavelength information from the input files into the output file. So you must run XCOPY on the output file before extracting the individual spectra using EXTRACT.

B.6 FIGARO Programs

1. ARC
2. CRSPOT
3. CRSKY
4. EXTRACT
5. FINDSP
6. FRAME_ADD
7. GROWX
8. ICSUB
9. IREVX
10. ISCRUNCH
11. ISTAT
12. ISUBSET

13. ISUPER
14. OVERPF
15. POLEXT
16. ROTATE
17. SPLOT
18. XCOPY
19. YSTRACT

As an alternative to FINDSP, OVERPF, and POLEXT – if they are not available – the following group of programs does the same thing:

1. CDIST
2. ECHSELECT
3. FINDFIB
4. SDIST