Dissertation

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Rupertus Carola University of
Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by
Diplom-Physicist: **Stefanie Phleps**
born in: **Groß-Gerau**

Oral examination: 7^{th} **of November, 2001**

The evolution of the large scale structure of the universe since $z = 1$

Referees: Priv.-Doz. Dr. Klaus Meisenheimer
Prof. Dr. Immo Appenzeller

In dieser Arbeit wird die Entwicklung der großskaligen Struktur im Universum ab einer Rotverschiebung von $z \sim 1.1$ bis heute untersucht. Als Datensatz dienen ~ 4000 Galaxien aus vier Feldern des Calar Alto Deep Imaging Surveys (CADIS). Diese Galaxien haben Helligkeiten von $I \leq 23^{mag}$ und Rotverschiebungsfehler von $\sigma_z \lesssim 0.02$. Zur Bestimmung der Amplitude der dreidimensionalen Korrelationsfunktion $\xi(r)$ dient die Deprojektion zweidimensionaler Korrelationsfunktionen. Am Las Campanas Redshift Survey (LCRS) wird die Verläßlichkeit der Deprojektion der Winkel- bzw. projizierten Korrelationsfunktion getestet. Beide, Winkel- sowie projizierte Korrelationsfunktion, werden in verschiedenen Rotverschiebungsintervallen berechnet und mit dem lokalen Meßwert aus dem LCRS verglichen. Um den direkten Vergleich zwischen CADIS und LCRS zu ermöglichen, muß der Einfluß der Rotverschiebungsungenauigkeit auf die projizierte Korrelationsfunktion berücksichtigt werden. Für die Entwicklung der Struktur wird der Ansatz $\xi(r_{com}, z) \propto (1+z)^q$ verwendet. Für die Gesamtheit aller Galaxien ergibt sich ein Entwicklungsparameter $q \approx -1.9$, entsprechend der Vorhersage durch lineare Störungstheorie. Eine formal gefundene Abhängigkeit vom kosmologischen Modell ist wohl auf die ungenügende Zahl der beobachteten Felder zurückzuführen. Das gemessene Anwachsen der Struktur ist jedoch eindeutig vom Hubbletyp abhängig. Für frühe Galaxientypen ist die Struktur bereits bei $z = 1$ deutlich stärker ausgeprägt, so daß schon ein Anwachsen mit $q \simeq -1$ ausreicht, um die heutige Amplitude der Korrelationsfunktion zu erreichen.

The large scale structure of the universe since $z \sim 1$

In this thesis, the evolution of galaxy clustering from a redshift of $z \sim 1$ to the present epoch is investigated. The data used for this analysis were ~ 4000 galaxies in four fields of the Calar Alto Deep Imaging Survey (CADIS). The galaxies have luminosities brighter than $I \leq 23^{mag}$, and redshifts with an error of $\sigma_z = 0.017$. The amplitude of the three-dimensional correlation function is estimated by deprojecting two-dimensional correlation functions. The reliability of the deprojection methods of the angular and projected correlation function is tested on the Las Campanas Redshift Survey (LCRS). Both angular and projected correlation function are calculated for different redshift bins, as local measurement the LCRS data is used. To facilitate the direct comparison of the two surveys, the influence of the redshift errors on the projected correlation function have to be taken into account. For evolution of the clustering strength the ansatz $\xi(r_{com}, z) \propto (1+z)^q$ is used. For the galaxies as a whole the evolution parameter turns out to be $q \approx -1.9$, according to the prediction of linear theory. A formal dependency on the cosmology is presumably due to the small number of fields observed. However, the measured clustering growth clearly depends on Hubble type. At $z \sim 1$ early type galaxies are already much stronger clustered, an increase with $q \simeq -1$ is sufficient to explain the present day amplitude of the correlation function.

Contents

1 Introduction **1**

2 CADIS – the Calar Alto Deep Imaging Survey **7**
2.1 Object detection 8
2.2 Classification of the objects 9
2.2.1 Multicolor classification 10
2.3 The data under consideration in this work 12

3 The statistical description of structure **15**
3.1 How to measure the two-point correlation function 16
3.2 Real space correlations and redshift errors 16
3.2.1 The angular correlation function 17
3.2.2 The projected correlation function 19
3.2.3 The special case of multicolor data 23

4 The LCRS - local sample and test **25**
4.1 The data 25
4.2 Consistency Checks – from 2d to 3d 27
4.2.1 The three dimensional correlation function 27
4.2.2 The angular correlation function 30
4.2.3 The projected correlation function 31
4.2.4 Influence of the redshift errors on the correlation function 35

5 The evolution of galaxy clustering **39**
5.1 The angular correlation function 40
5.2 The projected correlation function 44
5.2.1 Connection to the "local universe" 48
5.2.2 The evolution of the correlation function for different Hubble types . . 51
5.2.3 The bright and the faint ones 58
5.2.4 The cross-correlation between the faint and the bright galaxies 66
5.3 Summary 69

6 Discussion and conclusions **71**
6.1 Comparison with other observations 71
6.1.1 The CFRS determination 71
6.1.2 The CNOC determination 73
6.1.3 COMBO 17 73

6.2 Comparison with therory 77
6.2.1 Evolution of dark matter in different cosmologies 77
6.2.2 Biased galaxy formation 79
6.2.3 Semi-analytic models of biased galaxy formation 81
6.3 Conclusion 81

7 Outlook **85**
7.1 The merger rate at intermediate redshifts 85
7.2 The emission line galaxies in CADIS 89
7.3 COMBO 17 89

A The relation between $\xi(r)$ and the power spectrum of the density fluctuations **91**

B Cosmological Distances **93**

$$H_0 = h\ 100\,\mathrm{km\ s^{-1}/Mpc}$$

Chapter 1

Introduction

The universe is structured on all scales, from the molecular and crystal level to large clusters and superclusters of galaxies. Luminous matter is condensed into stars and planets, the stars are constituents of galaxies, which appear in a large variety of different types. Local galaxies are highly clustered. They are organised into a network of sheets and filaments which surround large underdense regions, usually referred to as voids. At the present epoch the universe seems to satisfy the *Cosmological Principle* (homogeneity and isotropy of the universe) only on the very largest scales.

Even before it became clear that the so-called *nebulae* were extragalactic objects, it was noticed by Hubble (1934) that their number found in different fields on the sky varies in a non-Poissonian way – they are obviously not homogeneously distributed. Shapley, who believed in the Galactic origin of those nebulae, thought of the non-uniformity of their distribution as a remnant of the formation history of the Milky Way (Shapley, 1933).

First systematic analyses of the distribution of galaxies and clusters did not occur before galaxy catalogues with large numbers of objects were drawn up – the first analyses of the clustering properties of galaxies were based on the Shane-Wirtanen, the Zwicky catalogue, and the catalogue of Abell clusters, and the results are outlined in a number of fundamental papers by Peebles (and co-workers) (Peebles, 1973; Hauser & Peebles, 1973; Peebles & Hauser, 1974; Peebles, 1974; Peebles & Groth, 1975; Peebles, 1975).

Only a few years later the CfA survey was completed (Davis et al., 1982; Davis & Peebles, 1983), and the analysis brought a distribution to light, which was amazingly inhomogeneous – filaments, sheets, walls and large voids emerged, and it became clear that the local universe is in fact far from homogeneous. Figure (1.1) shows a map of galaxies with $B \leq 15.5^{mag}$, in six contiguous 6° slices in the northern galactic cap (see Geller & Huchra, 1989). The structure running all the way across between 8^h and 17^h in right ascension and $cz = 5000$ and 10000 km s^{-1} is called the "Great Wall", perhaps the largest single structure yet detected in any redshift survey.

How did the structure we see today form and develop? How could the universe evolve from its smooth, homogeneous state immediately after the big bang[1] into the highly structured

[1] Observations of the cosmic microwave background show very little fluctuations.

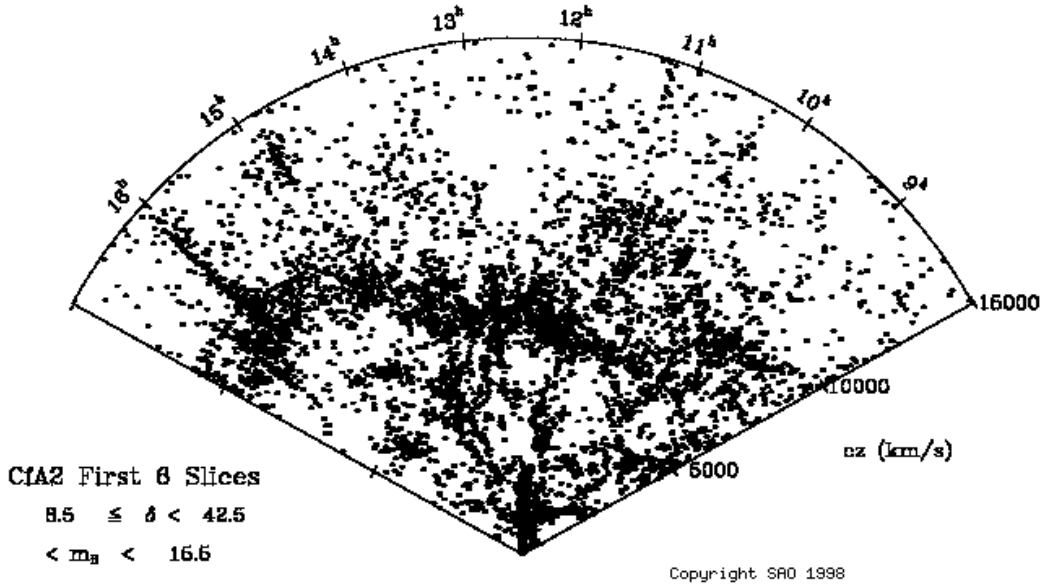

Figure 1.1: This initial map was really surprising, showing that the distribution of galaxies in space was anything but random, with galaxies actually appearing to be distributed on surfaces, almost bubble like, surrounding large empty regions, or "voids". The elongated structures along the line of sight are called *Fingers of God* (picture from Geller & Huchra (1989)).

world we live in today? Still this question is one of the most challenging ones in the field of cosmology. The opportunities for solving them them have emerged only recently with the advent of deep extragalactic surveys, which provide sufficiently accurate redshift estimates for a large numbers of galaxies.

Until now, there have been some attempts to explore this subject by observation, but only few possibilities to study the clustering of galaxies at high redshifts in detail. Not only the shallow depths of most surveys, but also missing redshift information or too small number statistics have limited the possibilities of analysing the data with regard to structure formation.

In general, two different types of surveys have to be distinguished – large angle surveys, which are limited to relatively bright apparent magnitudes, and pencil-beam surveys with small, but very deep fields. Furthermore one can make a distinction between surveys which contain

only a small number of galaxies, but with very accurate redshift information (deduced from spectroscopy), and surveys which provide huge catalogues of galaxies, but without or with extremely insecure redshift information. The **C**alar **A**lto **D**eep **I**maging **S**urvey (CADIS), see Meisenheimer et al. (in preparation), is a deep pencil-beam survey, the output of which at present is a catalogue of $\sim$ 6000 classified objects of up to $I \leq 23^{mag}$. Around 4000 of these are galaxies with relatively secure redshift information inferred by means of multi-color methods. This unique data base provides the possibility to investigate the evolution of galaxy clustering from a redshift of $z \approx 1.1$ to the present epoch. This is the aim of this thesis.

In this context the question how to quantify *structure* arises – the eye is an expert in finding structures, but if we want to describe the properties of a non-random distribution, we have to use a mathematical expression, which includes some kind of measure for the strength of the deviation from random. We also need this quantity to facilitate a comparison between the clustering properties at different redshifts, and the investigation of its development into what we see today. Usually, structure is described in terms of *n-point correlation functions*, the simplest of which is the *two-point correlation function*. In practice, computing the three-dimensional real-space two-point correlation function requires very accurate distances. Peculiar velocities as well as redshift errors distort the redshift-space relation, and, by making the distribution more Poisson-like, increase the noise. Different methods have been developed to overcome these problems. If no redshifts are available at all, it is possible to obtain information about the three dimensional distribution of galaxies by deprojecting the two-dimensional *angular correlation function*. If peculiar velocities are not negligible, or the data suffer from large redshift errors, one can use the deprojection of the *projected correlation function* to deduce the clustering strength of the three-dimensional distribution. Both methods are employed in this work.

Understanding the evolution of the large-scale structure of the universe is helpful for learning something about the history and nature of the universe, and the formation and evolution of galaxies. According to the standard theoretical paradigm, the structures observed today were formed by the gravitational amplification of small perturbations in an initially Gaussian dark matter density field. Those perturbations are believed to be remnants of the quantum fluctuations in the vacuum, and have undergone a tremendous growth during the epoch of inflation.

On the large scale dark matter provides the dominant mass in galaxies and systems of galaxies. The mass-to-luminosity ratios of galaxies and clusters of galaxies and the application of different variants of the Cosmic Virial Theorem on larger scales indicate that allmost all mass in the universe is in some dark form. Primordial nucleosynthesis of the light elements provides a firm upper limit to the value of the baryonic density parameter of $\Omega_B \leq 3.6 \cdot 10^{-2} h^{-2}$ (Longair, 1998). Thus, even adopting a value of $h = 0.5$, the upper limit to the baryonic mass density would barely be sufficient to account for values of the overall density parameter of the order $\Omega_0 \sim 0.25 - 0.6$, the values favoured by recent observations of the fluctuations of the microwave background (Hu et al., 2001; Griffiths et al., 2001).

One can think of the dark matter in different forms, but regardless of what sort of strange, yet unknown particles make up the contribution of non-baryonic dark matter, basically two different scenarios for the evolution of the large-scale structure of the universe are possible. In the *hot* dark matter (HDM) scenario (where "hot" means that the particles had relativistic

velocities at the time of decoupeling) the largest scales form first, and smaller structures such as galaxies and their contents form by fragmentation and instabilities ("top-down" model). The HDM model produces too much structure on large scales, and can not account for the observed two-point correlation function of the galaxies in the local universe. Another problem with the HDM model is that galaxies can only form once the large-scale structures have collapsed. Therefore it is inevitable that galaxies form rather late in the universe. This also is in contradiction to many observations of the early universe, for example the existance of quasars at $z > 4$, which can only be understood if structure formation was already well advanced at $z \sim 5$ (Turner, 1991), or the extremely massive, red elliptical galaxies which are observed in the centers of clusters of galaxies must have formed at redshifts $z > 6$ (Bruzual A. & Charlot, 1993).

The *cold* dark matter (CDM) scenario is much more efficient in creating power on small scales. The cold dark matter particles decoupled early in the universe, after they had already become non-relativistic. Small-scale densities were the first to collapse, and the resulting objects subsequently merged under the influence of gravity to form larger structures such as clusters and superclusters. One of the most attractive features of this "bottom-up" scenario is that the observed form of the galaxy correlation function can be reproduced (Davis et al., 1985).

In this hierarchical formation picture the clustering of the dark matter, as measured by the amplitude of the matter correlation function, increases monotonically with time (Hamilton et al., 1991; Peacock & Dodds, 1994; Jain et al., 1995; Jenkins et al., 1998). The only observable, however, is the clustering of galaxies, which *possibly* trace the underlying density field.

The evolution of the distribution of the dark matter in different world models can be investigated in large numerical N-body simulations. The development of large, powerful supercomputers has made simulations with $\sim 17 \cdot 10^6$ particles feasible (Pearce et al., 1999), in which the evolution of the dark matter density perturbations can be tracked from a redshift of $z = 50$ to the present epoch. These numerical computer experiments confirm what is predicted analytically by linear theory: In universes with low matter density structure forms early and evolves little between a redshift of $z = 1$ and today, whereas in high-density models structure forms much later, and a rapid growth of the clustering strength is observed between $z = 1$ and the present epoch.

The evolution of galaxy clustering in the different scenarios can be evaluated with semi-analytic models of galaxy evolution, which take simplified models of starformation, cooling, and feedback into account. The method of combining large N-body simulations with semi-analytic models was introduced by Kauffmann et al. (1999b). The evolution of galaxy clustering to $z = 0$ was investigated in two following papers (Kauffmann et al. (1999a), and Diaferio et al. (1999)).

Their results show that the measured clustering growth depends on the sample selection, and the clustering evolution of the dark matter density field can not directly be inferred from the clustering evolution of galaxies. The explanation for this fact can be given in the context of *biased galaxy formation*. One can imagine many reasons why galaxies should *not* trace the underlying density field, and the generic term for this phenomenon is *biasing*, meaning the preferential formation of galaxies in certain regions of space rather than others. Part of

the motivation behind the introduction of biasing was to improve the agreement between the predictions of the CDM scenario and the observed distribution of galaxies.

In this picture galaxies formed within dense haloes of dark matter, where gas was able to reach high enough overdensities to cool, condense and form stars. If we require the density perturbation to exceed some critical value which allows structures to form, galaxy formation would be biased towards the highest density perturbations over the mean background density. Also, the astrophysics which play a pivotal role in the process of galaxy formation and evolution might also change the relation of galaxy clustering and the clustering of the dark matter.

This thesis is structured as follows: The **C**alar **A**lto **D**eep **I**maging **S**urvey and the data used for the analysis are described in Chapter 2. An introduction into the fundamental principles of three- and two-dimensional correlation functions and the deprojection methods which used of in this work is given in Chapter 3. In Chapter 4 we test the reliability of the deprojection methods, using the Las Campanas Redshift Survey (Shectman et al., 1996). It is shown that the projected correlation function also suffers from the influence of redshift errors, although it was invented to overcome redshift-space distortions. Therefore we introduce some improvements to the procedure, which facilitate a direct comparison of the CADIS data with the Las Campanas Redshift Survey (which we use as a "local sample"). In Chapter 5 we investigate the evolution of the galaxy clustering from a redshift of $z \sim 1.1$ to today, using the methods introduced in the previous Chapters. In Chapter 6 the results are summarized and discussed. Further improvements are outlined and an outlook is given in Chapter 7.

Chapter 2

CADIS – the Calar Alto Deep Imaging Survey

The **C**alar **A**lto **D**eep **I**maging **S**urvey has been established in 1996 as the extragalactic key project of the Max-Planck Institut für Astronomie. It combines a very deep emission line survey carried out with an imaging Fabry-Perot interferometer with a deep multicolour survey using three broad-band optical to NIR filters and up to thirteen medium-band filters when fully completed (see Figure (2.1)). The combination of different observing strategies facilitates not only the detection of emission line objects but also the derivation of photometric spectra of all objects in the fields without performing time consuming slit spectroscopy.

CADIS provides the data base to address different aspects of extragalactic and even Galactic astronomy:

- The search for primeval galaxies at high redshifts
- Evolution of the luminosity function of field galaxies at intermediate redshifts
- Luminosity function of emission line galaxies at intermediate redshifts, and the deduction of the starformation rate
- Number counts and the luminosity function of quasars at $0 < z < 6$
- Number counts of EROs (**E**xtremly **R**ed **O**bjects)
- Search for very late-type M dwarfs and brown dwarfs
- Galactic structure and the stellar luminosity function
- Clustering properties of field galaxies with strong emission lines between $z \approx 0.24$ and $z \approx 1.2$
- Evolution of the clustering of field galaxies between $0.2 < z \lesssim 1.1$.

The last topic is subject of this thesis.

Details of the survey and its calibration will be given in Meisenheimer et al. (in preparation).

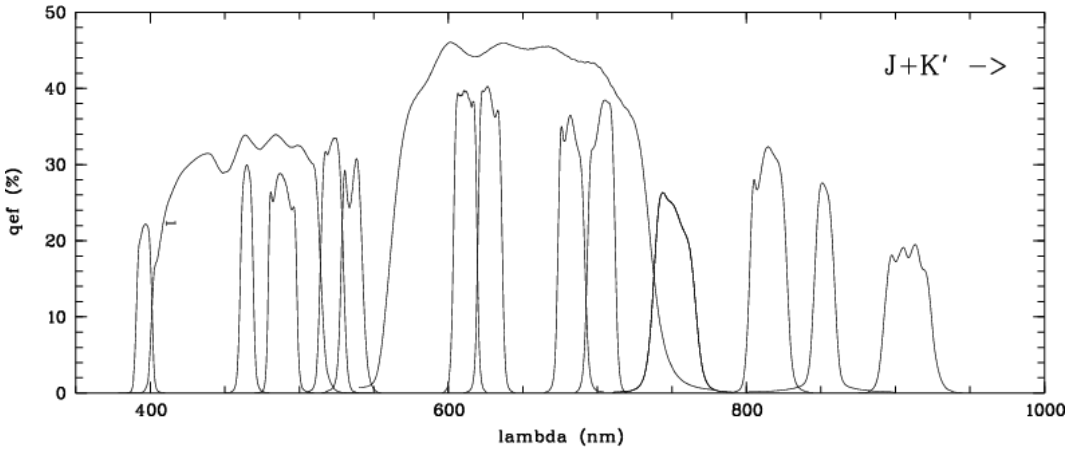

Figure 2.1: Transmission curves of the CADIS filter set. The Fabry-Perot settings, the K' filter at 2.2μ, and the J filter at 1.1μ are not indicated in this plot.

The seven CADIS fields measure $\approx 1/30\ \square^\circ$ each and are located at high Galactic latitude to avoid dust absorption and reddening. In all fields the total flux on the IRAS 100 μm maps is less than 2 MJy/sr which corresponds to $E_{B-V} < 0.07$. Therefore we do not have to apply any color corrections. As a second selection criterium the the fields should not contain any star brighter than $\approx 16^{mag}$ in the CADIS R band. In fact the brightest star in the four fields under consideration in this thesis has an R magnitude of 15.42^{mag}. Furthermore, the fields are chosen such that there should be at least one field with an altitude of at least 45° above the horizon of Calar Alto being observable each time throughout the year. Among the CADIS fields three equatorial fields allow follow-up observations with the VLT.

All observations were performed on Calar Alto, Spain, in the optical wavelength region with the focal reducers CAFOS (Calar Alto Faint Object Spectrograph) at the 2.2 m telescope and MOSCA (Multi Object Spectrograph for Calar Alto) at the 3.5 m telescope. For the NIR observations the Omega Prime camera was used.

In each filter, a set of 5 to 15 individual exposures was taken. The images of one set were then bias subtracted, flatfielded and cosmic corrected, and then coadded to one deep sumframe. This basic data reduction steps were done with the MIDAS software package in combination with the data reduction and photometry package MPIAPHOT (by Meisenheimer & Röser).

2.1 Object detection

Depending on their spectrum objects are detected in different filters with different signal-to-noise ratio. Therefore, object search is done on the sumframe of each filter using the source extractor software **SE**xtractor (Bertin & Arnouts, 1996). The filter-specific object lists are then merged into a master list containing all objects exceeding a minimum S/N ratio on any of the bands. For merging, all objects are considered identical which fall into a common error circle of $1''$ radius, while the typical seeing is $1\farcs5$. The positions of all detections in the different color bands are then averaged into a final position.

Photometry is done using the program *Evaluate*, which has been developed by Meisenheimer & Röser (see Meisenheimer & Röser (1986) and Röser & Meisenheimer (1991). Variations in seeing between individual exposures are taken into account, in order to get accurate colors. An optimum signal-to-noise ratio is achieved by integrating the photons over an aperture with a Gaussian weight distribution. In each image the aperture is located at the same position on the sky and its size and weight distribution is adapted to the seeing of the frame. Every image gets a weight aperture that simulates a Gaussian smoothing to a common seeing before the photon counting, in order to make sure that always the same fraction of an objects intrinsic light distribution is probed. The fluxes of extended sources have to be corrected for aperture losses. These corrections, which are dependent on the morphology of the galaxy, were derived from photometry on simulated images where the true magnitudes are known. Because the photometry is performed on individual frames rather than sumframes, realistic estimates of the photometric errors can be derived straightforwardly.

The measured counts are translated into physical fluxes outside the terrestrial atmosphere by using a set of "tertiary" spectrophotometric standard stars which were established in the CADIS fields, and which are calibrated with secondary standard stars (Oke, 1990; Walsh, 1995) in photometric nights.

From the physical fluxes magnitudes and color indices (an object's brightness ratio in any two filters, usually given in units of magnitudes) can be calculated. There are several definitions for the zeropoint of the magnitude scale, an astronomical definition handed down from history and an increasing number of physical definitions. We use a special magnitude scale, the *CD magnitude*, defined as

$$CD\mathrm{mag} = -2.5 \log F_{\mathrm{phot}} + 20\overset{\mathrm{m}}{.}01 \ , \tag{2.1}$$

where F_{phot} is the flux in photons $\mathrm{m}^{-2}\mathrm{s}^{-1}\mathrm{nm}^{-1}$. The zero point is chosen such that at 550 nm an object (observed through a quite narrow filter) will have the same magnitude in all distinct magnitude systems which are commonly used.

The corresponding color indices are then defined by

$$m_1 - m_2 = -2.5 \log \frac{F_{\mathrm{phot},1}}{F_{\mathrm{phot},2}} \ . \tag{2.2}$$

2.2 Classification of the objects

The CADIS data base essentially consists of two different subsets. One is a sample of galaxies with strong emission lines, which are detected on deep Fabry-Perot exposures. The other subset is a sample of field galaxies, which have been detected on the broad- and narrow filter frames, in the following *multicolor galaxies*. Since in this work only the multicolor galaxies are investigated, we will restrict ourselves to the discussion of the multicolor classification scheme.

2.2.1 Multicolor classification

With a typical seeing of $1\farcs5$ a morphological star-galaxy separation starts failing at $R \gtrsim 21$ where already many galaxies appear compact. Quasars have point-like appearance as well as stars, and thus can not be distinguished. Therefore a classification scheme was developed, which is based on template spectral energy distributions (*SED*s) (Wolf et al., 2001c,b). The classification algorithm basically compares the observed colors of each object with a color library of known objects. This color library is assembled from observed spectra by synthetic photometry performed on an accurate representation of the instrumental characteristics used by CADIS.

The stellar library is taken from Pickles (1998). The QSO spectral library is based on high signal-to-noise template spectra (Francis et al., 1991), but also includes different continuum slopes and line widths; the final QSO spectral library contains templates with redshifts up to $z \leq 6$. The spectral library for galaxies is derived from the mean averaged spectra of Kinney et al. (1996). From these, a grid of redshifted spectra was formed covering redshifts from $z = 0$ to $z = 2$ in steps of $\delta z = 0.01$ and 100 different intrinsic *SED*s, from old populations to starbursts ($SED = 1$ corresponds to an E0 galaxy, whereas $SED = 100$ is a starburst galaxy). Our final library of color indices (colors are obtained from synthetic photometry performed on our CADIS filterset) contains entries for 131 star-, 45150 QSO-, and 20100 galaxy templates.

Using the minimum variance estimator (for details see Wolf et al. (2001c)), each object is assigned a type (star – QSO – galaxy), a redshift (if it is not classified as star), and an *SED*. The formal errors in this process depend on magnitude and type of the object. For the faintest galaxies ($I < 22$) they are of the order of $\sigma_z = 0.017$, and $\sigma_{SED} = 2$, respectively.

Note that we do not apply any morphological star/galaxy separation or use other criteria. The classification is purely spectrophotometric.

The quality of the classification and redshift estimation was checked by spectroscopy of a subsample of 162 arbitrarily chosen objects, which are representative for the object catalogue as a whole. This subsample can further be divided into 103 bright ($17 < R < 22$) and 59 faint objects ($R > 22$, including 11 objects with $R > 24$). The bright sample contains only two misclassifications, which translates into $\sim 98\%$ correct classifications. The mistakes are Seyfert-1 galaxies (i.e. quasars) found by chance among the compact galaxies. The faint sample contains $\sim 25\%$ misclassifications and 10% unclassified objects, with most of them being galaxies. The others are one L star and one Seyfert-1 galaxy. This means that until $R < 22$ the classification procedure works almost ideally for all classes, but at fainter levels the abundant galaxies start contaminating the star class and the quasar class. More details about the performance and reliability of the classification are given in Wolf et al. (2001b).

The quality of the multicolor redshift estimation is shown in Figure (2.2). For the 162 objects, multicolor redshift is plotted against the spectroscopic redshift, the difference between both ($\Delta z = z_{\mathrm{multicolor}} - z_{\mathrm{spectroscopic}}$) against the spectroscopic redshift, and Δz against their R magnitudes. Half of the galaxies are estimated within an error margin of ± 0.017, thus we conclude that the redshift determination works properly until $R \lesssim 24$.

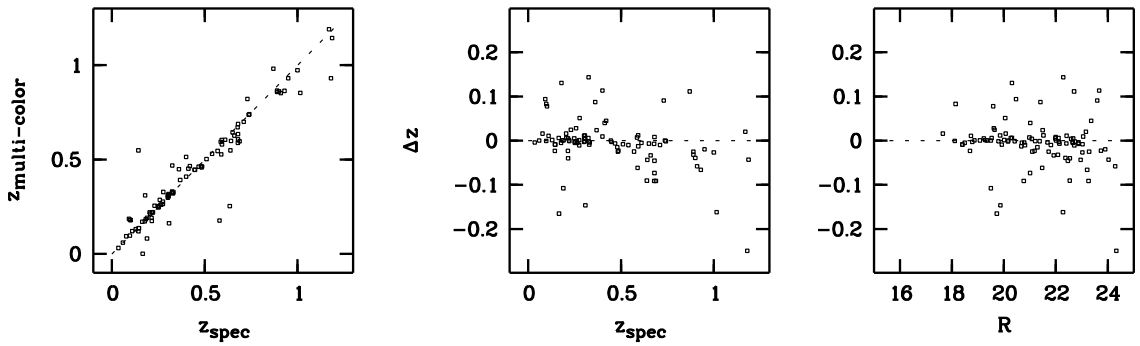

Figure 2.2: The quality of the multicolor redshift estimation: left panel: multicolor redshift versus spectroscopic redshift with the highest redshift galaxies residing at $z \approx 1.2$. center panel: the error of the estimate ($\Delta z = z_{\mathrm{multicolor}} - z_{\mathrm{spectroscopic}}$) versus redshift; right panel: Δz versus the R magnitude.

We also compared the true redshift errors Δz with the errors estimated by the multicolor technique itself on the basis of photometric errors and the galaxy distribution in the color space. The ratio $\Delta z/\sigma_z$ evaluates the error consistency of our redshift estimate. If the estimated errors were representative of true errors, this ratio should have a Gaussian distribution with an rms of 1.0. In fact, it turns out that for 30% of the galaxies this inconsistency is larger than 3σ, while the remaining $\sim 70\%$ show a more or less Gaussian distribution with an rms scatter of 1.2 (see Figure (2.3)). This result implies that for one third of the spectroscopic galaxy sample, the redshift estimation process considers itself too accurate, supposedly a consequence of cosmic variance that changes the galaxy SEDs and their estimated redshifts while not changing the photometric errors.

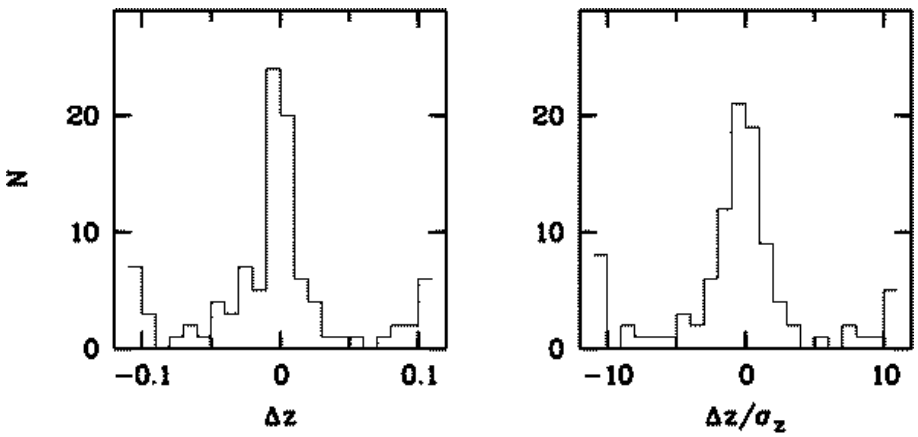

Figure 2.3: Most galaxy redshifts are estimated with a Δz error variance of ~ 0.04, but $\sim 10\%$ of the galaxies receive completely wrong redshift assignments with $\Delta z > 0.1$ (left). For 70% of the galaxies the true error distribution matches up with the one expected from the multicolor errors, but 30% of the objects have true errors larger than the estimated 3σ-errors (right), which are mostly starburst galaxies. The reason for the increased scatter in general is, that the observed SEDs are not perfectly matched by the library SEDs.

Table 2.1: The coordinates of the four fields investigated in this thesis, and the number of galaxies per field, $I \leq 23$ and $0.2 \leq z \leq 1.07$.

CADIS field	α_{2000}	δ_{2000}	$N(I \leq 23, 0.2 \leq z \leq 1.07)$
1 h	$1^{\rm h}47^{\rm m}33\overset{\rm s}{.}3$	$2°19'55''$	898
9 h	$9^{\rm h}13^{\rm m}47\overset{\rm s}{.}5$	$46°14'20''$	916
16 h	$16^{\rm h}24^{\rm m}32\overset{\rm s}{.}3$	$55°44'32''$	971
23 h	$23^{\rm h}15^{\rm m}46\overset{\rm s}{.}9$	$11°27'00''$	841

2.3 The data under consideration in this work

Four CADIS fields have been fully analysed so far (for coordinates see Table (2.1)). Although the fields are rectangular, only objects within a circle of $400''$ radius around the field center are included in the catalogue, to avoid spurious effects from the field edges, where object detection and classification becomes increasingly difficult due to the influence of the instrumental properties (like distortion of the optics, decreasing illumination of the field).

In the selected area we identified 4540 galaxies with $I \leq 23$. Out of them 3626 are located in the redshift range $0.2 \leq z \leq 1.07$, where we have analysed their clustering properties. The number of galaxies per field is given in Table (2.1)).

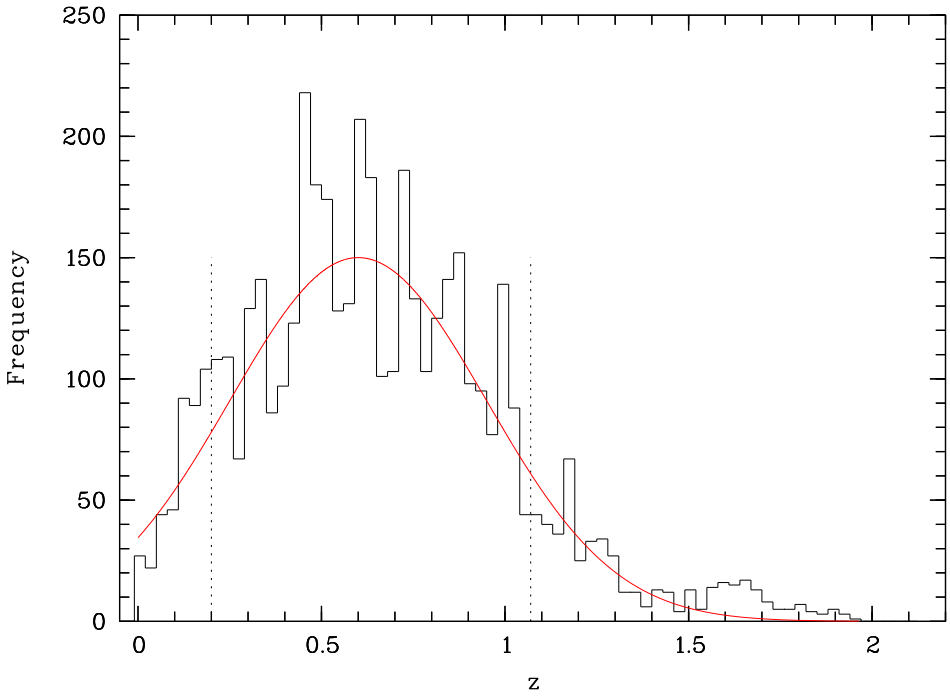

Figure 2.4: Redshift distribution galaxies with $I < 23$ in the four CADIS fields. The "smoothed" redshift distribution can be described by a Gaussian. The dotted lines indicate the redshift range, in which the clustering properties of the galaxies are analysed.

The redshift distribution (see Figure (2.4)) has a maximum at $z \approx 0.6$. The large scatter from bin to bin, which is discernably more than what is expected from a Poisson statistics, reflects the large scale structure of the galaxies. With four fields, the accumulation of galaxies at certain redshifts does not cancel out completely.

From the classification algorithm, the redshift for all galaxies in the sample is known with an accuracy of $\sigma_z = \pm 0.017$. For the first time a sample this deep, with large number statistics *and* redshift information is available. However, the redshift accuracy is not sufficiently high enough to calculate the three-dimensional two-point correlation function directly. Nevertheless, as we will demonstrate later, we can use the redshifts for our analysis. First of all it enables us to divide the catalogue into distinct redshift bins and we can analyse the clustering of the galaxies in each of them up to $z \approx 1.1$, and secondly, we are independent of the inversion of the angular correlation function: we can use the *projected* correlation function, which makes use of redshifts, and is less sensitive to redshift errors than the three-dimensional calculation.

We can further divide our sample into different subsamples and investigate the clustering properties of galaxies with different Hubble types (since the classification scheme provides the *SED*s of the galaxies), or different absolute restframe B magnitudes. We take only galaxies with redshifts $0.2 \leq z \leq 1.07$ into account for this analysis. In this range the number statistics is large enough, and the redshift estimation secure enough. Figure (2.5) shows the redshift versus B magnitude for galaxies with redshifts $0.2 \leq z \leq 1.07$.

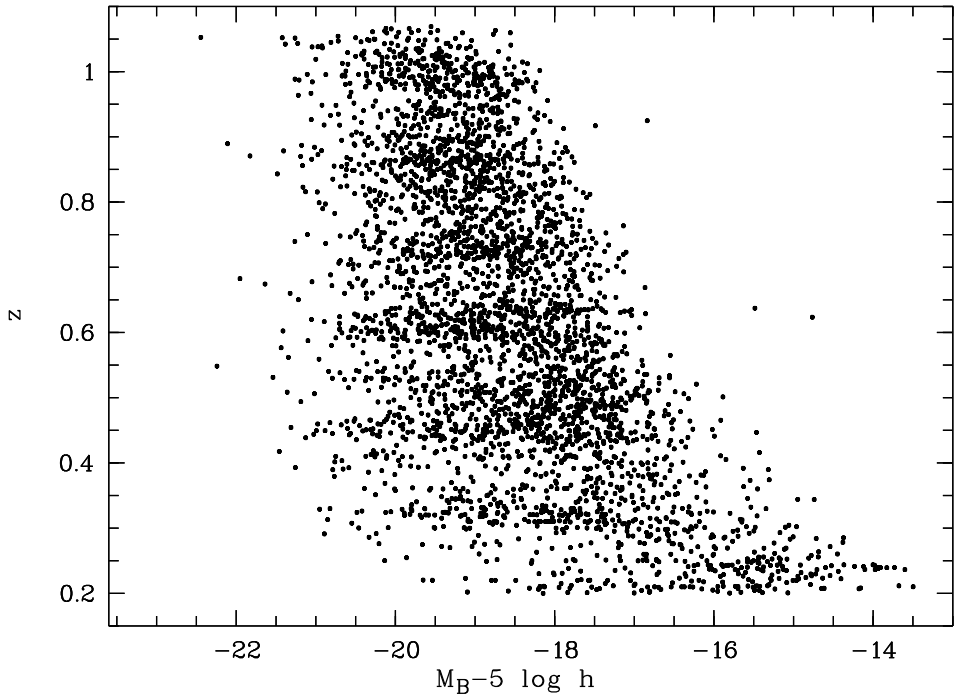

Figure 2.5: Redshift versus absolute magnitude for the 3626 galaxies with $I_{815} \leq 23$, in the redshift range under consideration in this work, for $\Omega_0 = 1$, $\Omega_\Lambda = 0$.

Chapter 3

The statistical description of structure

We have seen in Chapter 1 that galaxies are obviously not randomly distributed (see for example Figure (1.1)). In fact, much of the obvious clumping, the holes and the stringy structures are real features of the three-dimensional distribution of galaxies. To take away the arbitrariness in deciding whether any structure is real or not, and to quantify the strength of the clustering, one has to find a formalim, that describes the deviation from a uniform distribution. The simplest statistic commonly used to gain insight into the structure of point processes is the *two-point correlation function.* It describes the excess probability of finding a second galaxy at a distance r from any galaxy selected at random over the probability expected in a uniform, random distribution of galaxy locations. Thus the number of galaxies $\mathrm{d}N$ found in the volume element $\mathrm{d}V$ at distance r is given in the form:

$$\mathrm{d}N(r) = N_0[1 + \xi(r)]\mathrm{d}V \ , \tag{3.1}$$

where N_0 is a suitably-defined average background number density of galaxies. The correlation function, $\xi(r)$, can also be written in terms of the number of *pairs* of galaxies separated by distance r in the form

$$\mathrm{d}N_{pair} = N_0^2[1 + \xi(r)]\mathrm{d}V_1 \ \mathrm{d}V_2 \ . \tag{3.2}$$

It turned out that in the correlation function of galaxies does not appear any preferred scale. The natural description for a behaviour like this is the parametrisation with a power law, as first proposed by Totsuji & Kihara (1969):

$$\xi(r) = \left(\frac{r}{r_0}\right)^{-\gamma} \ . \tag{3.3}$$

The *correlation length* r_0 is only a "fit parameter", without any physical meaning. It can nevertheless be understood as a measure for the clustering strength, and is useful for comparing correlation functions with each other. One can also think of r_0 as the length, at which the clumping gets really significant – the probability of finding a galaxy separated by the distance $r = r_0$ from another galaxy is twice as large as it would be if the distribution of galaxies was random, see equation (3.1). Still one has to keep in mind that there is no deeper meaning hidden in the value of r_0, and the parametrisation with a single power law is nothing but a

useful description, expressing the fact that there are no preferred scales.

The two-point correlation function is the Fourier transform of the power spectrum, which is often referred to in the literature. The derivation of this relation is given in Appendix A.

3.1 How to measure the two-point correlation function

The simplest way to measure the deviation from a Poisson distribution is the following: one counts the number of pairs of real galaxies in a histogram of logarithmic distance bins ($\langle DD \rangle$ in the following). This histogram is normalised to the total number of pairs, $n(n-1)/2$, where n is the number of galaxies in the sample under consideration. An artificial catalogue has to be created, with randomly distributed data points, that exhibit the same area on the sky as the real data, and suffers from the same selection effects as the real data. Thus it serves as a *catalogue window*, representing the geometrical properties of the survey. In this random catalogue, we count the pairs of galaxies in the same way as in the real data, this histogram (in the following $\langle RR \rangle$) is normalised to the total number of artificial galaxies $n_r(n_r - 1)$. Thus the most obvious estimator for the correlation function (the **excess probability** of finding pairs of galaxies at distance r over a random sample) is:

$$\xi_{esti} = \frac{\langle DD \rangle}{\langle RR \rangle} - 1 \; . \tag{3.4}$$

One can also use $\langle RD \rangle$ (the distances between real and random data) instead of $\langle RR \rangle$ (Davis & Peebles, 1983). This very simple estimator has the defect that its uncertainty on large scales is limited by the uncertainty in the mean density. This is not the case for the estimator proposed by Hamilton (1993):

$$\xi_H \;\; = \;\; \frac{\langle DD \rangle \langle RR \rangle}{\langle RD \rangle^2} - 1 \; . \tag{3.5}$$

Nevertheless, use of the sample mean for the normalisation biases the estimate of the correlation function by missing out variances on scales comparable to the scale of the sample (Landy & Szalay, 1993). The statistic proposed by Landy & Szalay (1993) allows missing large-scale variance to be restored:

$$\xi_{LS} \;\; = \;\; \frac{\langle DD \rangle - 2\langle RD \rangle + \langle RR \rangle}{\langle RR \rangle} \; . \tag{3.6}$$

Therefore we decided to use this one in all the following investigations.

3.2 Real space correlations and redshift errors

The previous section is a bit deceptive in giving the impression that one "simply has to count distances". But exactly at this point the first difficulties appear: the coordinate system we are forced to use for the calculation of the distances is not homogeneous. The coordinates we are able to measure are the angular positions on the sky, α and δ, and the redshift z of each galaxy. The angular positions can be measured with extremely high accuracy, and

without time-consuming and difficult observation techniques. In contrast, the redshift, from which we have to infer the distance along the line of sight, can only be deduced with high precision from spectroscopic observations. But even if redshifts are known with very small errors, this third spatial coordinate can not be regarded equal to the two others: peculiar velocities add to the general Hubble flow, and the relation between measured redshift and distance is distorted. Structures tend to appear elongated along the line of sight. A typical example are the so-called "Fingers of God" (see Figure (1.1)), where the random motion in compact clusters lead to a spread in the real distances when displayed in three-dimensional galaxy maps.

Since it is extremely time-consuming to measure redshifts, most large galaxy catalogues consist only of the angular positions on the sky. If redshifts have been deduced by applying spectro-photometric (multicolor) methods, they have large errors, which has essentially the same effect on the estimation of the three-dimensional correlation function as have peculiar velocities.

The conclusion is that in reality the real space two-point correlation function can only be calculated directly, if redshift information is available with very high precision, and if peculiar velocities are negligible. If only angular coordinates are available, or the redshifts have huge errors, one has to deal with the two-dimensional distribution of the galaxies on the sky, measure the *angular correlation function*, and try to derive the parameters of the three-dimensional correlation function from the deprojection.

If the errors in the redshift measurement are not excessively large, but still too large to facilitate a direct computation, one can use the *projected correlation function* to derive the correlation length r_0 and the slope γ. Although it makes use of the redshift, it much less sensitive to redshift error, which we will demonstrate in chapter (4).

In the following we will describe the two different approaches to derive the three-dimensional correlation function by deprojection – of the angular correlation function, or the projected correlation function.

3.2.1 The angular correlation function

The angular correlation function $w(\theta)$, in analogy to equation (3.2) defined as

$$\mathrm{d}P = N^2(1 + w(\theta))\mathrm{d}\Omega_1\mathrm{d}\Omega_2 \ , \tag{3.7}$$

is related to the three dimensional correlation function $\xi(r)$ by Limber's equation (Limber, 1954).

Limber's equation is true for $z \ll 1$, but for deep samples, which cover a large range of redshifts, the expansion of the universe, curvature effects, and the possible evolution of the clustering have to be included. The redshift-dependent version of Limber's equation has been derived by Groth & Peebles (1977), and Phillipps et al. (1978). The general distribution of redshifts ($\mathrm{d}N/\mathrm{d}z$) has to be known or calculated from the selection function, which assembles

all selection effects due to observation and data reduction. The three-dimensional correlation function is conventionally parameterised as

$$\xi(r,z) = \frac{(r/r_0)^{-\gamma}}{(1+z)^{3+\epsilon}} \tag{3.8}$$

$$r_0(z) = r_0(0)(1+z)^{-(3+\epsilon)/\gamma} \tag{3.9}$$

(r_0 in *proper* distance), where the parameter ϵ takes some kind of evolution of the clustering properties with redshift into account. The evolution parameter ϵ becomes unimportant, if one investigates the angular correlation function only in small z-bins (presumably there is *no* evolution within a small redshift interval).

It should be noted that the correlation length $r_0(z)$ is here the correlation length (in physical units), that would be measured by a local observer at the epoch in question. Thus the correlation length will evolve (as $(1+z)^{-1}$), even if the clustering pattern is fixed in comoving space. In the case of a clustering pattern fixed in comoving coordinates, clustering does not grow with time, and $\epsilon = -1.3$ for $\gamma = 1.7$. When gravitationally bound units keep a fixed physical size, the clustering growth is the result of the increasingly diluted galaxy background (it is effectively the voids that are growing), and $\epsilon = 0$. For a standard CDM cosmology, the mass clustering should grow in the linear regime with $\epsilon = 0.7$.

Then from Limber's equation

$$w(\theta) = A_w \theta^{1-\gamma} \ , \tag{3.10}$$

with

$$A_w = CBr_0^{\gamma} \ , \tag{3.11}$$

where C is a numerical factor:

$$C = \sqrt{\pi}\frac{\Gamma((\gamma-1)/2)}{\Gamma(\gamma/2)} \ , \tag{3.12}$$

and

$$B = \int_0^\infty \frac{d_A(z)^{1-\gamma}(1+z)^{-(3+\epsilon)}}{g(z)} \left(\frac{\mathrm{d}N}{\mathrm{d}z}\right)^2 \mathrm{d}z \cdot \left[\int_0^\infty \frac{\mathrm{d}N}{\mathrm{d}z}\ \mathrm{d}z\right]^{-2} \ . \tag{3.13}$$

Here d_A is the angular diameter distance and $g(z)$ is the differential comoving distance element multiplied by the scale factor. $\mathrm{d}N/\mathrm{z}$ is the smoothed redshift distribution, see for example Figure (2.4)).

For $\Omega_\Lambda = 0$

$$d_A = \frac{c}{H_0}\frac{q_0 z + (q_0-1)(-1+\sqrt{1+2q_0 z})}{q_0^2(1+z)^2} \tag{3.14}$$

$$g(z) = \left(\frac{c}{H_0}\right)\left[(1+z)^2\sqrt{1+\Omega_0 z}\right]^{-1} \tag{3.15}$$

Thus from equation (3.11) the correlation length r_0 of the three-dimensional distribution can be calculated. On the other hand, if the correlation length at $z = 0$ is known, the evolution of the clustering with redshift can be deduced, namely the evolutionary parameter ϵ (see equation (3.9)).

3.2.2 The projected correlation function

If peculiar velocities are not negligible, or the available redshifts have large errors, one can use the projected correlation function $w(r_p)$ to derive the parameters of the real space correlation function (Davis & Peebles, 1983). The projected distance r_p between galaxies i and j (the distance perpendicular to the line of sight) can be interpreted as

$$r_p = [d_A(i) + d_A(j)] \tan(\theta_{ij}/2) \ , \tag{3.16}$$

where d_A is the angular diameter distance (equation (3.14)) and θ_{ij} is the apparent separation between galaxy i and galaxy j.

$$\frac{\mathrm{d}\pi}{\mathrm{d}z} = \frac{c}{H_0(1+z)^2\sqrt{1+2q_0z}} \tag{3.17}$$

is the increment in *physical* distance. For small angles $r^2 = r_p^2 + \pi^2$, see Figure (3.1).

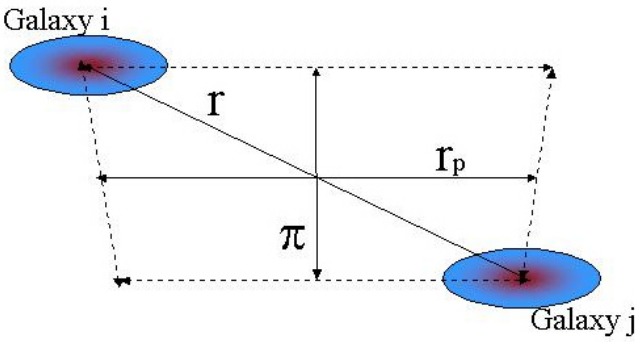

Figure 3.1: Schematic drawing of the geometry.

The projected correlation function is defined as

$$\begin{aligned} w(r_p) &= 2\int_0^\infty \xi\left[(r_p^2+\pi^2)^{1/2}\right]\,\mathrm{d}\pi \\ &= 2\int_{r_p}^\infty \xi(r)(r^2-r_p^2)^{-1/2} r\,\mathrm{d}r \ . \end{aligned} \tag{3.18}$$

The inverse is the Abel integral

$$\xi(r) = -\frac{1}{\pi}\int_r^\infty w'(r_p)(r_p^2 - r^2)^{-1/2}\mathrm{d}r_p \ . \tag{3.19}$$

If $\xi(r) = (r/r_0)^{-\gamma}$, then equation (3.18) yields

$$w(r_p) = C r_0^\gamma r_p^{1-\gamma} \tag{3.20}$$

with the factor C from equation (3.12). Thus computing $w(r_p)$ provides a measurement of the parameters of the three-dimensional correlation function, namely r_0 and γ.

The evolutionary parameter ϵ is not a very intuitive description of the evolution of the clustering strength. Instead of using ϵ, we can calculate the amplitude of the three-dimensional correlation function at a comoving separation of $r_{com} = 1h^{-1}\mathrm{Mpc}$, and introduce a more descriptive parameter q:

$$\xi(r_{com} = 1h^{-1}\mathrm{Mpc}) = \xi_0(1+z)^q \tag{3.21}$$

$$r_{0,com}^\gamma = r_{0,0,com}^\gamma (1+z)^q \ . \tag{3.22}$$

For a clustering pattern that is expanding with the universe (fixed in comoving coordinates), the clustering strength is simply a function of $(1+z)$, and $q = 0$. Thus q gives directly the deviation of the evolution from the global Hubble flow. Of course it can be related to ϵ: With equation (3.9) we can write

$$\xi(r = 1h^{-1}\mathrm{Mpc}, z) = \xi(r = 1h^{-1}\mathrm{Mpc}, 0)(1+z)^{-(3+\epsilon)} \ , \tag{3.23}$$

for r in proper coordinates. For $\xi(r = 1h^{-1}\mathrm{Mpc}) = r_0^\gamma$,

$$r_0^\gamma = r_{0,0}^\gamma (1+z)^{-(3+\epsilon)} \ . \tag{3.24}$$

For r_{com} $(r_0^\gamma \to r_0^\gamma(1+z)^\gamma)$

$$\begin{aligned} r_{0,com}^\gamma &= r_{0,0,com}^\gamma (1+z)^{-(3+\epsilon)}(1+z)^\gamma \\ &= r_{0,0,com}^\gamma (1+z)^{-(3+\epsilon-\gamma)} \ . \end{aligned} \tag{3.25}$$

Comparison with equation (3.21) yields

$$q = -3 - \epsilon + \gamma \ . \tag{3.26}$$

So for the canonical slope $\gamma = 1.8$, $q = 0$ corresponds to $\epsilon = -1.2$, $q = -1.2$ corresponds to $\epsilon = 0$, and $q = -2$ to $\epsilon = 0.8$ (linear growth).

Method of estimation

Following Davis & Peebles (1983) one can calculate the projected correlation function from

$$w(r_p) = \int_{-\delta\pi}^{+\delta\pi} \xi(r_p, \pi)\ \mathrm{d}\pi \ . \tag{3.27}$$

Since the three-dimensional two-point correlation function has the form of a power law, it converges rapidly to zero with increasing pair separation. Therefore the integration limits do

not have to be $\pm\infty$, they only have to be large enough to include all correlated pairs. Since the observable in the first place is the redshift z instead of the physical separation, we make a coordinate transform (Le Fevre et al., 1996):

$$w(r_p) = \int_{-\delta z}^{+\delta z} \xi(r_p, \pi) \frac{c}{H_0(1+z)^2\sqrt{1+2q_0 z}} \mathrm{d}z \ , \tag{3.28}$$

for $\Omega_\Lambda = 0$.

The way to estimate $w(r_p)$ in practice is to count the projected distances between pairs of galaxies that are separated in redshift space by not more than δz, in appropriate projected distance bins ($\langle DD(r_p)\rangle|_{d_z \leq \delta z}$). We use the estimator by Landy & Szalay (1993), which we will call $\zeta_{esti}(r_p)$ in the following, thus we have to do the same in a suitably generated random data set, and find the cross-counts between real and random data:

$$\zeta_{esti}(r_p) = \frac{\langle DD(r_p)\rangle|_{d_z \leq \delta z} - 2\langle RD(r_p)\rangle|_{d_z \leq \delta z} + \langle RR(r_p)\rangle|_{d_z \leq \delta z}}{\langle RR(r_p)\rangle|_{d_z \leq \delta z}} \ . \tag{3.29}$$

To derive $w(r_p)$, $\zeta_{esti}(r_p)$ has to be multiplied with the "effective depth" $\Delta r_\parallel$ in which galaxies are taken into account, thus

$$\begin{aligned} w(r_p) &= \zeta_{esti}(r_p) \cdot \Delta r_\parallel \\ &= \zeta_{esti} \cdot \int_{-\delta z}^{+\delta z} \frac{c \ \mathrm{d}z}{H_0(1+z)^2\sqrt{1+2q_0 z}} \end{aligned} \tag{3.30}$$

for $\Omega_\Lambda = 0$.

The effective depth

In reality, one has to cope with a selection function of some kind or another, and not with a top-hat function (of probability unity to find a galaxy within the borders of the survey and zero otherwise). The importance of this was first noticed in this work: in the context of the projected correlation function, a complication arises with the computation of the depth $\Delta r_\parallel$. At low redshifts, the effective depth is diminished by the geometry, since the physical area in which galaxies can be found grows with redshift. In a magnitude limited sample the number of galaxies per redshift bin decreases with increasing redshift (the fainter a galaxy, the smaller its limiting distance to which it can be detected). The varying probability to find pairs of galaxies separated by a redshift δz within the survey has to be taken into account. It can be included in the calculation by multiplying the integrand in equation (3.30) with the probability to find pairs of galaxies separated by a distance δz. The selection function of finding *pairs* of galaxies (ϱ_{pair}) can be expressed in terms of the squared redshift distribution ($\mathrm{d}N/\mathrm{d}z$), normalised to unity at it's maximum:

$$\varrho_{pair} = \left[\frac{1}{N_z}\frac{\mathrm{d}N}{\mathrm{d}z}\right]^2 \ , \tag{3.31}$$

where N_z is the normalisation constant. Then the integral (3.30) changes to

$$\Delta r_\parallel = \int_{-\delta z}^{+\delta z} \frac{c}{H_0(1+z)^2\sqrt{1+2q_0 z}} \varrho_{pair} \mathrm{d}z \ . \tag{3.32}$$

With this correction for the selection function, the "depth" converges to a fixed value and does not grow anymore, even if the integration limits cover more than the total depth of the survey. For comparison with results of other authors it has to be kept in mind that in no previous estimation of the projected correlation function (see for example Le Fevre et al. (1996) or Carlberg et al. (2000)) it was payed any attention to the influence of the survey selection function.

The choice of the integration limit δ_z in equation (3.27) and equation (3.30) is somewhat arbitrary. To capture the bulk of the correlation signal, the integration limits should be significantly larger than the correlation length r_0 and the length corresponding to the pairwise velocity dispersion, and of course they have to be larger than the errors.

Very large values of δ_z might more completely integrate the correlation signal, but they do so at the considerable cost of increased noise, for two reasons – first, the bigger the separation of two galaxies along the line of sight, the more meaningless the projected separation perpendicular to the line of sight gets. Second, if a pair of galaxies is separated by a very large physical distance along the line of sight, it is certainly not correlated at all, since the correlation function decreases very fast with distance. Nevertheless it can display a very small projected separation and will therefore be regarded as strongly correlated.

Robustness of the inversion

With equation (3.18), the inversion of the Abel integral, one should in principle be able to derive the correlation length r_0 of the real space distribution from the measurement of the projected correlation function. Implicitly, this includes the assumption, that the correlation function decreases monotonously (the slope of the power law stays constant), even for very large distances. This is obviously not the case, see Peacock (1997). The power spectrum (the Fourier transform of the correlation function), changes its slope at a wavenumber of $k \approx 0.1h\,\mathrm{Mpc}^{-1}$.

In the range $0.01 < k < 10$, the real space power spectrum can be described with a break between two power laws (Peacock, 1999):

$$\Delta^2(k) = \frac{(k/k_0)^\alpha}{1 + (k/k_1)^{\alpha-\beta}} \,, \tag{3.33}$$

with

$$\begin{aligned} k_0 &= 0.03h\mathrm{Mpc}^{-1} \\ k_1 &= 0.05h\mathrm{Mpc}^{-1} \\ \alpha &= 0.8 \\ \beta &= 4.0 \end{aligned}$$

This break in the power law has at least two implications: The trivial one is, that the amplitude of the correlation function (defined by $\xi(r_0) = 1$, or $w(r_p) = 1$, respectively) one derives by fitting the data with a single power law, depends on the range of distances that are observed in the sample. This means that the range of projected distances has to be compatible

with the range of the three dimensional distances.

The other implication is less trivial. If a survey is much deeper than the scale at which the break in the power law occurs, equation 3.20 is no longer a valid solution for the inverse of the Abel integral, because this solution implies $\xi(r) = (r/r_0)^{-\gamma}$. The integral over a power law with a break at $\sim 60h^{-1}$ Mpc yields a value for the correlation length r_0 that is slightly smaller than the value expected from a single power law. In order to simplify the verification, we did not Fourier-transform equation (3.33), but integrated the simple power law to $r = 60h^{-1}$ Mpc. Beyond this point where the break occurs, it decreases very strongly, so the contribution to the integral can be neglected. For $r_0 = 5.0h^{-1}$ Mpc, $\gamma = 1.8$, we find for the amplitude of the projected correlation function (for $r_p = 1h^{-1}$ Mpc).

$$\int_{r_p}^{\infty} \frac{\xi(r)}{\sqrt{r^2 - r_p^2}} r \, \mathrm{d}r = 62.57 \tag{3.34}$$

With equation (3.20), we recover $r_0 = 5.0$. But

$$\int_{r_p}^{60} \frac{\xi(r)}{\sqrt{r^2 - r_p^2}} r \, \mathrm{d}r = 60.15 \ , \tag{3.35}$$

and by applying equation (3.20), we find only $r_0 = 4.88$.

So the effect is of Order $\mathcal{O} = 10^{-2}$, and therefore negligible for our purposes, since the errors of the measurements are larger in any case.

3.2.3 The special case of multicolor data

The redshifts of the galaxies in our catalogue have errors of size $\sigma_z = \pm 0.017$ (see chapter (2)), so it is clearly impossible to compute the three-dimensional correlation function directly. However, the errors are still small enough to facilitate the analysis of the projected correlation function, as we will show in the next chapter. Since the number of pairs of galaxies that contribute to the histogram $\langle DD \rangle$ in the estimation of the projected correlation function is much smaller than for the angular correlation function (for the projected correlation function only galaxies separated by less than the integration limits in equation (3.28) are counted), we can use the larger number statistics of the angular correlation function to corroborate the results qualitatively. Therefore we calculate both angular *and* projected correlation function, and compare the results.

Usually, a straight line is fitted to the logarithmic data points, where γ is the slope, and the amplitude is the y axis intercept (of course in the case of the three-dimensional correlation function $\xi(r)$, the y axis intercept gives directly r_0). We have to make a very careful choice in the range where we fit the data: since the CADIS fields measure $11' \times 11'$ each, the comoving distance at the mean redshift $\langle z \rangle \approx 0.6$ is only $\sim 4.5h^{-1}$ Mpc (for $\Omega_0 = 0.2$, $\Omega_\Lambda = 0$). This is even less than the canonical correlation length of $r_0 = 5h^{-1}$ Mpc, and the major part of the correlation signal in our data is on very small scales. Therefore, we restrict the range in which the data are fitted to this region: for the projected correlation function $-1.7 \leq \log r_p \leq -0.3$,

and for the angular correlation function to $-0.9 \leq \log\theta \leq 0.5$ (θ in arcminutes). To minimize the errors of the measured amplitudes (which are clearly dominated by the uncertainty in the fitted slope, if one interpolates far outside the measured data points), we do not fit the amplitude at the y axis intercept, but at some point within the measurement, preferably somewhere in the middle. Since we want to compare our measurements with the local one (for which we use the Las Campanas Redshift Survey), we furthermore have to make sure that we compare similar comoving scales in the two different surveys. Thus we restricted the fit of the projected correlation function of the LCRS to $-1.15 \leq \log r_p \leq 0.3$, and the angular correlation function to $0 \leq \theta \leq 1.5$, to achieve a compromise between maximum overlap and high signal-to-noise of the correlation signal of both surveys. We then fitted the amplitude of the projected correlation function of CADIS at $\log r_p = -0.5$ ($r_p \simeq 316h^{-1}\,\mathrm{kpc}$), and for the LCRS at $\log r_p = -0.3$ ($r_p \simeq 500h^{-1}\,\mathrm{kpc}$) – in comoving coordinates this corresponds for CADIS ($\langle z\rangle = 0.6$) to $r_{p,com} \simeq 506h^{-1}$ kpc, for the LCRS ($\langle z\rangle = 0.1$) to $r_{p,com} \simeq 550h^{-1}$ kpc. In the case of the angular correlation function we fit the amplitude of $w(\theta)$ for CADIS at $\theta = 1'$ (at $z = 0.6$ this corresponds to $r_p \simeq 200h^{-1}\,\mathrm{kpc}$), and for the LCRS at $\theta = 0\overset{\circ}{.}1$ (at $z = 0.1$ this corresponds to $r_p \simeq 400h^{-1}\,\mathrm{kpc}$). From the values of the amplitudes, which are all measured at comparable scales, we can now deduce the parameter of the three-dimensional correlation function.

Chapter 4

The LCRS - local sample and test

The CADIS data provides the possibility to analyse the clustering of galaxies at relatively high redshifts, but since it is our intention to investigate the evolution of galaxy clustering from $z \sim 1$ to the present epoch and the CADIS data is only suited for the analysis down to a redshift of 0.2, we need a local sample to facilitate the connection to $z = 0$. A suitable catalogue, which can serve as a local sample is the Las Campanas Redshift Survey, in the following **LCRS**. On the other hand it is necessary to verify the validity of the deprojection methods described in the previous chapter. Before proceeding to derive the two-point correlation function of the multicolor galaxies in CADIS, one can use the LCRS to check all the above assumptions.

4.1 The data

The LCRS is described in detail by Shectman et al. (1996). The survey consists of ~ 26000 galaxies in six sky strips, three strips in the northern Galactic cap region and three in the southern region, see Figure (4.1). Each strip runs approximately 80° across the sky in right ascension and has a width in declination of $\sim 1.5^\circ$. The mean strip declinations are -3°, -6°, -12° in the northern sample and -39°, -42°, -45° in the southern sample.

Each strip was subdivided into 50 or so fields of square or nearly square dimensions. Galaxies in each field were selected on the basis of Kron-Cousins R-band magnitudes; a subset of these galaxies was chosen randomly for spectroscopic study using multiobject fiber spectrometers of either 50 or 112 fibers. Apparent magnitude limits vary from field to field, with typical isophotal limits of $16.0 \leq m_R \leq 17.3$ and $15.0 \leq m_R \leq 17.7$ for the 50 fiber and 112 fiber fields, respectively. Additional limits were imposed on the basis of the "central surface brightness" of the galaxies, corresponding approximately to the flux entering a fixed fiber aperture of $3''.5$; the limiting central magnitude is in the range of $m_c = 18 - 19$, depending on the isophotal magnitude.

The survey has a median redshift of $\langle z \rangle \sim 0.1$, and therefore can be regarded as "local"; the mean error in redshift is $\sigma_z \approx 2.24 \cdot 10^{-4}$, that is $\sigma_{cz} = 67.2\,\mathrm{km\ s^{-1}}$.

Different aspects of large scale structure in the LCRS are treated in a variety of papers, e.g Doroshkevich et al. (1996), Lin et al. (1996), Landy et al. (1996), Tucker et al. (1997), Colley (1997), Shandarin & Yess (1998), Bharadwaj et al. (2000), Best (2000), and Müller et al.

(2000)).

Tucker et al. (1997) found a comoving correlation length of $r_0 = 6.28 \pm 0.27h^{-1}$ Mpc and $\gamma = 1.52 \pm 0.03$, in the distance range $2.0 < r < 16.4h^{-1}$ Mpc, whereas Jing et al. (1998) found $r_0 = 5.06 \pm 0.12h^{-1}$ Mpc and $\gamma = 1.862 \pm 0.0.034$ for $r_p < 28h^{-1}$ Mpc. The latter value was estimated by means of the projected two-point correlation function.

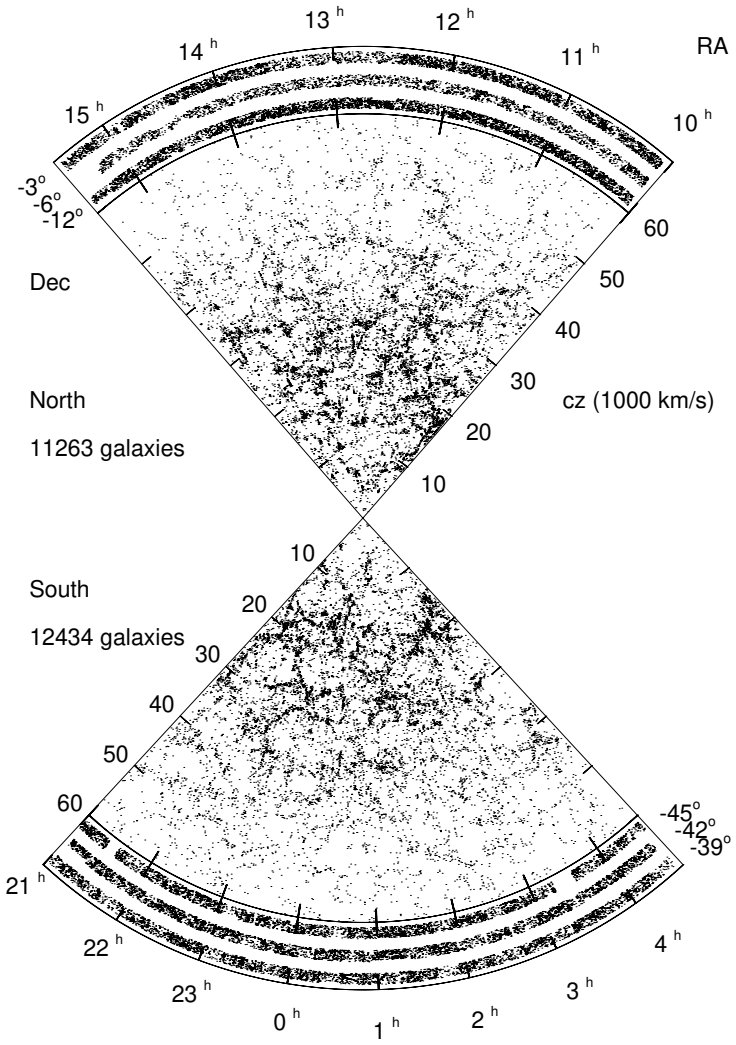

Figure 4.1: The Las Campanas Redshift Survey.

Table 4.1: Geometries and number of galaxies for the different sectors of the LCRS which are used in our study.

Name	x_1 /°	x_2 /°	Δx /°	y_1 /°	y_2 /°	Δy /°	N_{gal}
37-38	-38.829	-38.337	0.492	8.091	57.711	49.620	539
35-41	-41.993	-41.283	0.710	2.508	67.246	64.738	1055
190-3	-2.904	-2.278	0.627	156.583	230.366	73.783	1479
193-11	-11.947	-11.358	0.589	156.228	220.015	63.787	1457

4.2 Consistency Checks – from 2d to 3d

To check whether the integral inversions (be it Limber's equation or the Abel integral) work properly with *real* data, we calculated the three dimensional correlation function for four different sectors of the survey, and alternatively for the same sectors the projected correlation function and the angular correlation function. From those we derived the correlation length r_0 and the exponent γ. The main aim of this exercise is to check the validity of deprojection methods for the local sample.

Table (4.1) lists the geometries and the number of galaxies in the different parts of the LCRS we were using, right ascension and declination are both given in degrees.

4.2.1 The three dimensional correlation function

For each sector a catalogue of randomly distributed galaxies is generated, within the corresponding borders in right ascension and declination. Their redshift distribution (that follows the mean redshift distribution of *all* galaxies in the whole survey, see also Figure (4.2)) can well be approximated by a Maxwell function:

$$\frac{\mathrm{d}N}{\mathrm{d}z} = 1.27 \cdot 10^6 z^2 \exp\left[-0.5 * \left(\frac{z - 0.0074}{0.063}\right)^2\right] \tag{4.1}$$

We calculated the three dimensional correlation function for $0.04 \leq z \leq 0.16$, using the estimator of Landy & Szalay (1993) (equation (3.6)).

The distance $d_{i,j}$ between a pair of galaxies at redshifts z_i and z_j separated by an angle θ can be approximated by using Pythagoras' law:

$$d_{i,j} = \sqrt{(D_A(\bar{z}) \cdot \theta)^2 + \left(\frac{D(\bar{z})}{(1+\bar{z})}\right)^2}\,, \tag{4.2}$$

where D_A is the angular distance, D the comoving distance, which is projected to the physical distance at the mean redshift $\bar{z} = (z_i + z_j)/2$ by dividing by $(1 + \bar{z}$. For definitions and derivations see appendix (B).

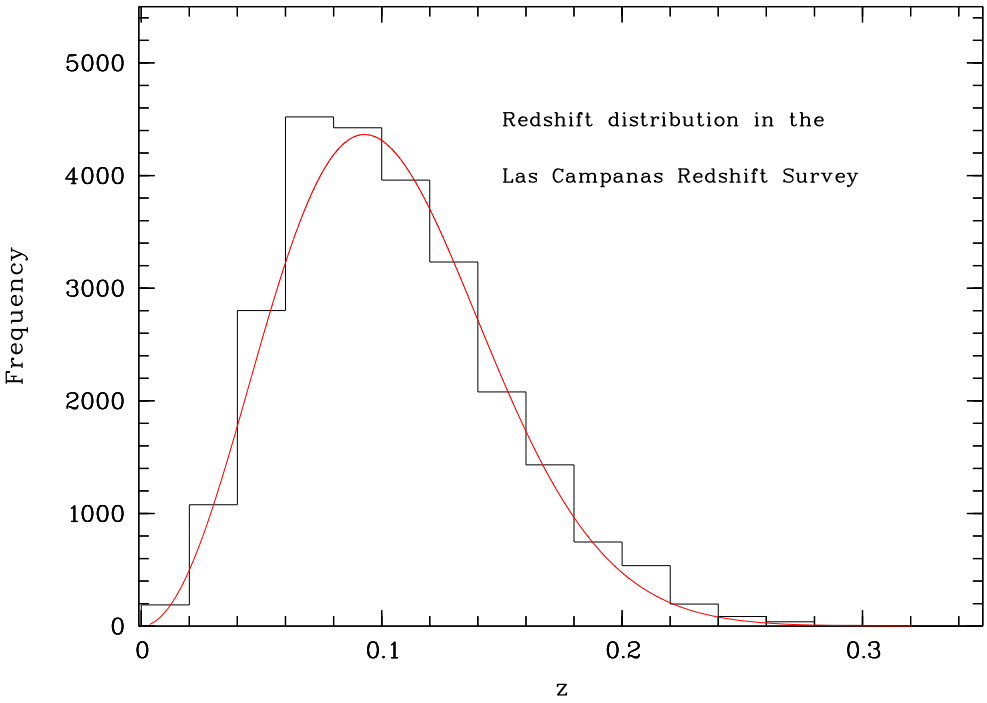

Figure 4.2: The redshift distribution of all galaxies in the LCRS.

Figure (4.3) shows the run of $\xi(r)$ for the above mentioned sectors of the LCRS, for $H_0 = 100$, $\Omega_0 = 1.0$, and $\Omega_\Lambda = 0$. Errors are calculated from Poisson statistics:

$$\sigma_\xi = 2 \cdot \sqrt{\frac{(1 + \xi(r))}{[DD]}} \,, \tag{4.3}$$

where $[DD]$ is the unnormalised histogram of the number of pairs of galaxies in logarithmic distance bins. We found that the errors calculated in this way are comparable with the errors using bootstrap resampling methods (Barrow et al., 1984). The individual errors of the data points are sometimes slightly smaller and sometimes slightly larger than bootstrap errors, but there is no systematic trend towards either case.

The correlation functions in the different sectors of the LCRS show a relatively large spread at the large scale end, thus the interval where the correlation function was fitted, was restricted to $-1.0 \leq \log r \leq 0.5$, as indicated in figure (4.3). Furthermore we did not fit r_0 directly, but the amplitude of $\xi(r)$ at $r = 0.5h^{-1}$ Mpc (from which we then are able to calculate r_0), because if one leaves the slope as a free parameter and determines the amplitude somewhere at the margin, or even outside the measured range, the errors of the slope dominate the errors of the amplitude, and the errors of the amplitude get unnecessarily large. The value $r = 0.5h^{-1}$ Mpc is more or less in the middle of our fitted range of data points, so it is better to determine the amplitude of $\xi(r)$ at this point. Also, as outlined in Chapter 3, it allows the direct comparison with the CADIS data.

Table (4.2) shows the results for the different sectors of the LCRS. In the first columns the

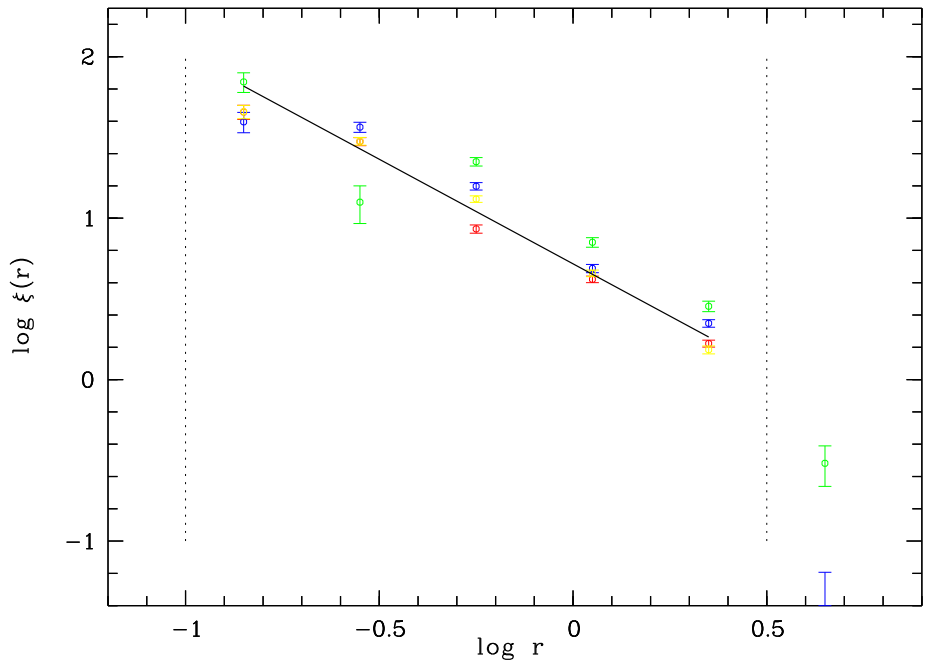

Figure 4.3: The two-point correlation function $\xi(r)$ in the different parts of the LCRS. Different colors refer to the different sectors; red: 193-11, green: 37-38, blue: 35-43, yellow: 190-3. The dotted lines indicate the interval, in which the correlation function is fitted. The black line is the mean fit.

amplitude of $\xi(r)$ at $r = 0.5h^{-1}\,\mathrm{Mpc}$ and the slope γ is given, respectively, the last column contains the value r_0 deduced from this.

The weighted mean r_0 in our calculations is 3.579 ± 0.092, the weighted mean of $\gamma = 1.295 \pm 0.016$. The mean fit is also plotted in figure (4.3). This mean slope is much flatter than the canonical slope of $\gamma = 1.8$, and flatter than what has been found by other authors (see above). The scatter between the values of $\xi(r = 0.5/h^{-1}\,\mathrm{Mpc})$ and r_0 in the different sectors is quite large, the values are not equal within the errors.

Table 4.2: Results for r_0 for the different field geometries, deduced from the amplitude of $\xi(r)$ at $r = 0.5h^{-1}\,\mathrm{Mpc}$.

Sector	$\xi(r = 0.5/h^{-1}\,\mathrm{Mpc})$	γ	$r_0/h^{-1}\,\mathrm{Mpc}$
37-38	19.991 ± 0.290	1.224 ± 0.046	5.777 ± 0.536
35-41	15.577 ± 0.183	1.285 ± 0.034	4.236 ± 0.243
190-3	13.116 ± 0.131	1.335 ± 0.029	3.438 ± 0.146
193-11	11.574 ± 0.131	1.290 ± 0.028	3.337 ± 0.140

4.2.2 The angular correlation function

For comparing three- and two-dimensional correlations, we first computed the **angular correlation function** ($w(\theta)$), and deduced the correlation length r_0. Again we used the estimator by Landy & Szalay (1993) for the estimation. For the random sample we used the same mock catalogues as for the analysis of the three-dimensional correlation function. Figure (4.4) shows $w(\theta)$ for the different field geometries (see Table (4.1)).

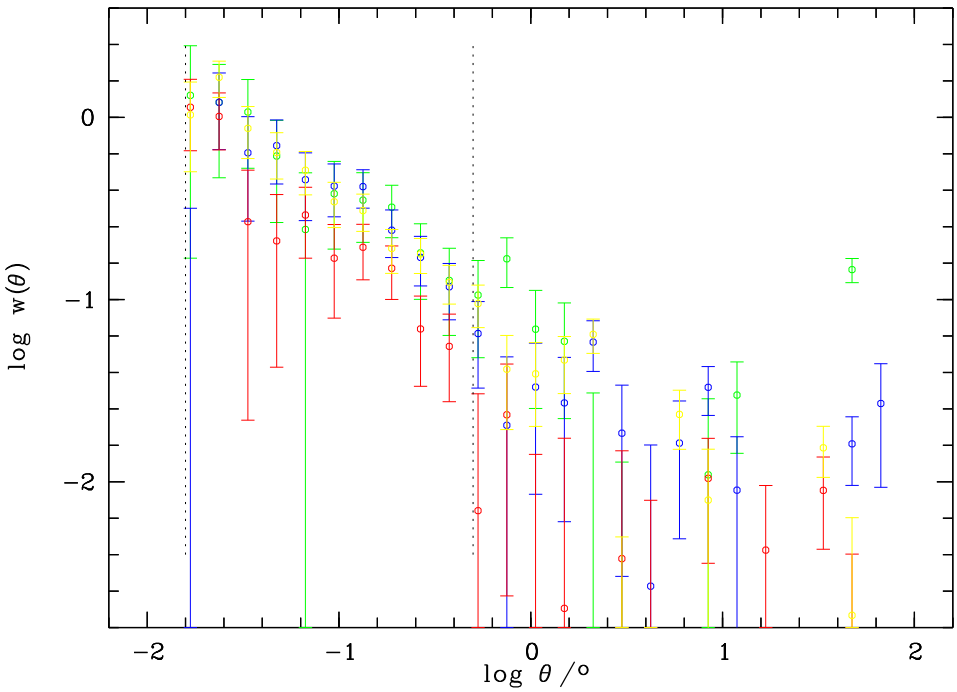

Figure 4.4: The angular two-point correlation function $w(\theta)$ in the different parts of the LCRS. Different colors refer to the different sectors (same as in Figure (4.3)). The dotted lines indicate the interval, in which the angular correlation is fitted.

To make sure that we compare to some extend the same distance ranges (for the three-dimensional correlation function the range was $0.14 \lesssim r \lesssim 2.24h^{-1}\,\mathrm{Mpc}$), we restrict the interval to be fitted fit to $0\overset{\circ}{.}0167 < \theta < 0\overset{\circ}{.}5$. At the mean depth of the survey ($\langle z \rangle = 0.1$), this corresponds to $0.09 \lesssim r \lesssim 2.62h^{-1}\,\mathrm{Mpc}$). In this range the signal-to-noise is high, and the distance ranges overlap. The borders are also indicated in figure (4.4). We fit the amplitude of $w(\theta)$ at $\theta = 0\overset{\circ}{.}1$ ($\simeq 400h^{-1}$ kpc).

We can now use Limber's equation to derive the connection of the amplitude of the angular correlation function at $\theta = 0\overset{\circ}{.}1$ with the amplitude of the three-dimensional correlation function at $r = 0.5h^{-1}$ Mpc. With $\xi(r = 0.5h^{-1}\mathrm{Mpc}) = (0.5/r_0)^{\gamma} = 0.5^{-\gamma} r_0^{\gamma}$, equation (3.10) and equation (3.11) we find:

$$A_w \cdot 0.5^{\gamma} = CB\xi(r = 0.5h^{-1}\mathrm{Mpc}) \tag{4.4}$$

$$0.5^{\gamma} w(\theta = 0\overset{\circ}{.}1) = CB\xi(r = 0.5\ h^{-1}\mathrm{Mpc}) \cdot 0\overset{\circ}{.}1^{1-\gamma}\ , \tag{4.5}$$

Table 4.3: Amplitudes of the angular correlation function at $\theta = 0.^\circ 1$ in different parts of the LCRS, and the amplitude of the three-dimensional correlation function at $r = 0.5\ h^{-1}$Mpc, and the correlation length r_0 calculated from that.

Sector	$A_w(\theta = 0.^\circ 1)$	γ	$\Rightarrow \xi(r = 0.5\ h^{-1}\mathrm{Mpc})$	r_0/h^{-1} Mpc
37-38	0.403 ± 0.057	1.726 ± 0.334	61.982 ± 0.701	5.462 ± 1.527
35-41	0.393 ± 0.042	1.770 ± 0.282	62.422 ± 0.583	5.167 ± 1.293
190-3	0.373 ± 0.028	1.848 ± 0.168	62.289 ± 0.337	4.677 ± 0.951
193-11	0.218 ± 0.016	1.927 ± 0.241	38.012 ± 0.286	3.302 ± 0.780

and

$$\xi(r = 0.5\ h^{-1}\mathrm{Mpc}) = \frac{0.5^\gamma\, w(\theta = 0.^\circ 1)}{CB\ 0.^\circ 1^{1-\gamma}}\ . \tag{4.6}$$

In this calculation ϵ was kept fixed at zero – the choice of the evolution parameter is not really important in the case of the LCRS, because at the shallow depth of the survey, it does not change the result in a significant way. We also chose zero because the effective evolution between a redshift of 0.1 and zero is negligible anyway.

Table (4.3) shows the amplitudes of the angular correlation function at $\theta = 0.^\circ 1$ in different parts of the LCRS, the amplitude of the three-dimensional correlation function at $r = 0.5\ h^{-1}$Mpc), and the correlation length r_0 calculated from that. Everything was computed for $\Omega_0 = 1$ and $\Omega_\Lambda = 0$.

The weighted mean is $\langle \xi(r = 0.5\ h^{-1}\mathrm{Mpc}) \rangle = 50.868 \pm 0.196$, with $\langle \gamma \rangle = 1.838 \pm 0.116$. This corresponds to a mean correlation length $r_0 = 4.240 \pm 0.572$. Although the amplitudes at $r = 0.5\ h^{-1}$Mpc) are much larger than the ones deduced from the three-dimensional estimation, the values of r_0 are equal within the errors. This is due to the steeper slope of the angular correlation function, which compensates for the large amplitudes.

4.2.3 The projected correlation function

Before we give results for the **projected correlation function** ($w(r_p)$), we have to make a choice for the appropriate integration limits (see equation (3.28)). Therefore we calculated the projected correlation function for increasing integration limits, and fitted the amplitude at $r_p = 0.5h^{-1}$Mpc) (the fit was done in the range $-1.15 \leq \log r_p \leq 0.3$). The result is shown in figure (4.5), the colors denote the different sectors of the LCRS, as before. The black line is the weighted mean of the four sectors, as in the case of the three-dimensional correlation function, the field-to-field variation is relatively large.

The amplitude of the projected correlation function rises very steeply with increasing integration limits, reaches a maximum when peculiar velocities become unimportant and the undistorted correlation signal is sampled, and then slowly declines, because the noise increases. The maximum lies around $\pm\delta cz \approx 3000\mathrm{kms}^{-1}$.

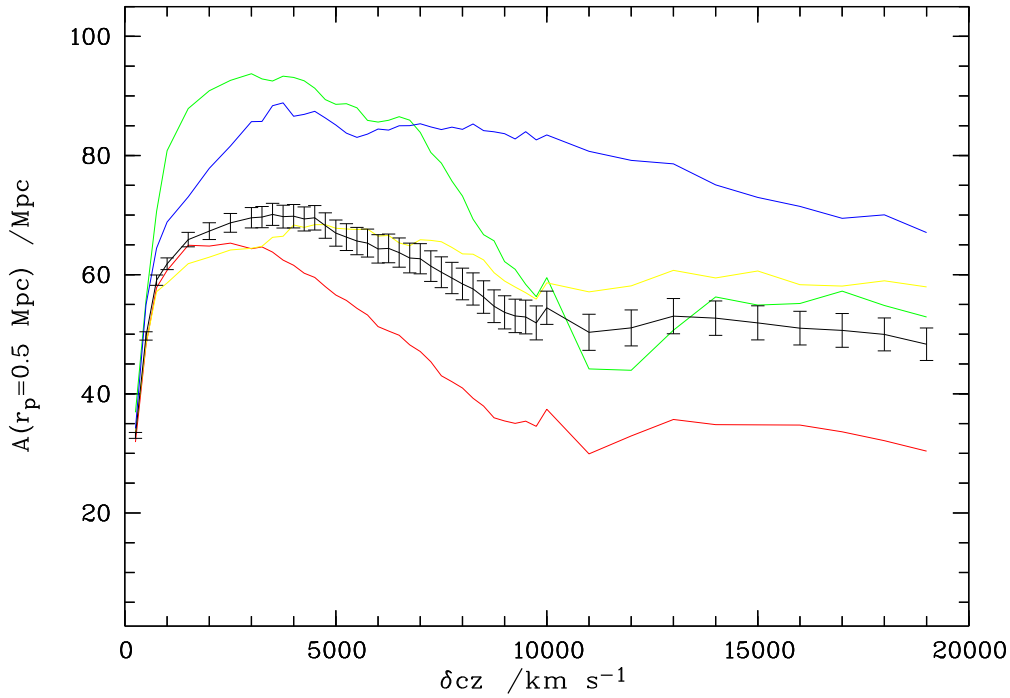

Figure 4.5: The amplitude of the projected correlation function at $r_p = 0.5h^{-1}$Mpc), versus integration limits ($\pm\delta cz$), for the different parts of the LCRS. The colors denote the same sectors as in figure (4.3). The black line is the weighted mean of the four sectors.

The projected correlation function is calculated by multiplying the estimator ζ_{esti}, with the effective depth $\Delta r_{\|}$ (see Chapter 3), $w(r_p) = \zeta_{esti} \cdot \Delta r_{\|}$. ζ_{esti} essentially relates the counted projected distances to those which can be found in a random distribution, and $\Delta r_{\|}$ is the effective depth in which those pairs of galaxies can be counted. Figure (4.6) shows the dependence of ζ_{esti} (fitted at $r_p = 0.5h^{-1}$Mpc) and $\Delta r_{\|}$ on the integration limits. At small integration limits, ζ_{esti} is relatively high and then decreases fastly, because the same kind of scaling relation applies as in the case of the angular correlation function (Peebles, 1980): the correlation signal is diminished when the sample depth increases, because at a certain angle one samples an increasing number of pairs of galaxies with large physical speparations, which are more or less uncorrelated. Then the noise starts to dominate, and the amplitude decreases further. For small integration limits the distance $\Delta r_{\|}$ increases linearly with increasing limits, which is not surprising, because at low redshifts and small redshift intervals, the redshift distance relation can be approximated with the linear Hubble relation $r = cz/H$. But since the selection function of the survey (in form of the redshift distribution) is taken into account in the calculation of $\Delta r_{\|}$, it converges to a constant value for $\delta cz \to \infty$.

Both effects together lead to the observed course of the amplitude of the projected correlation function. The errors of the redshift determinations are very small and thus have no influence on the amplitude, but peculiar velocities (the galaxies in a cluster have a velocity dispersion of typically $\Delta v \approx 2500\,\mathrm{km\ s^{-1}}$) do play a role and diminish the amplitude.

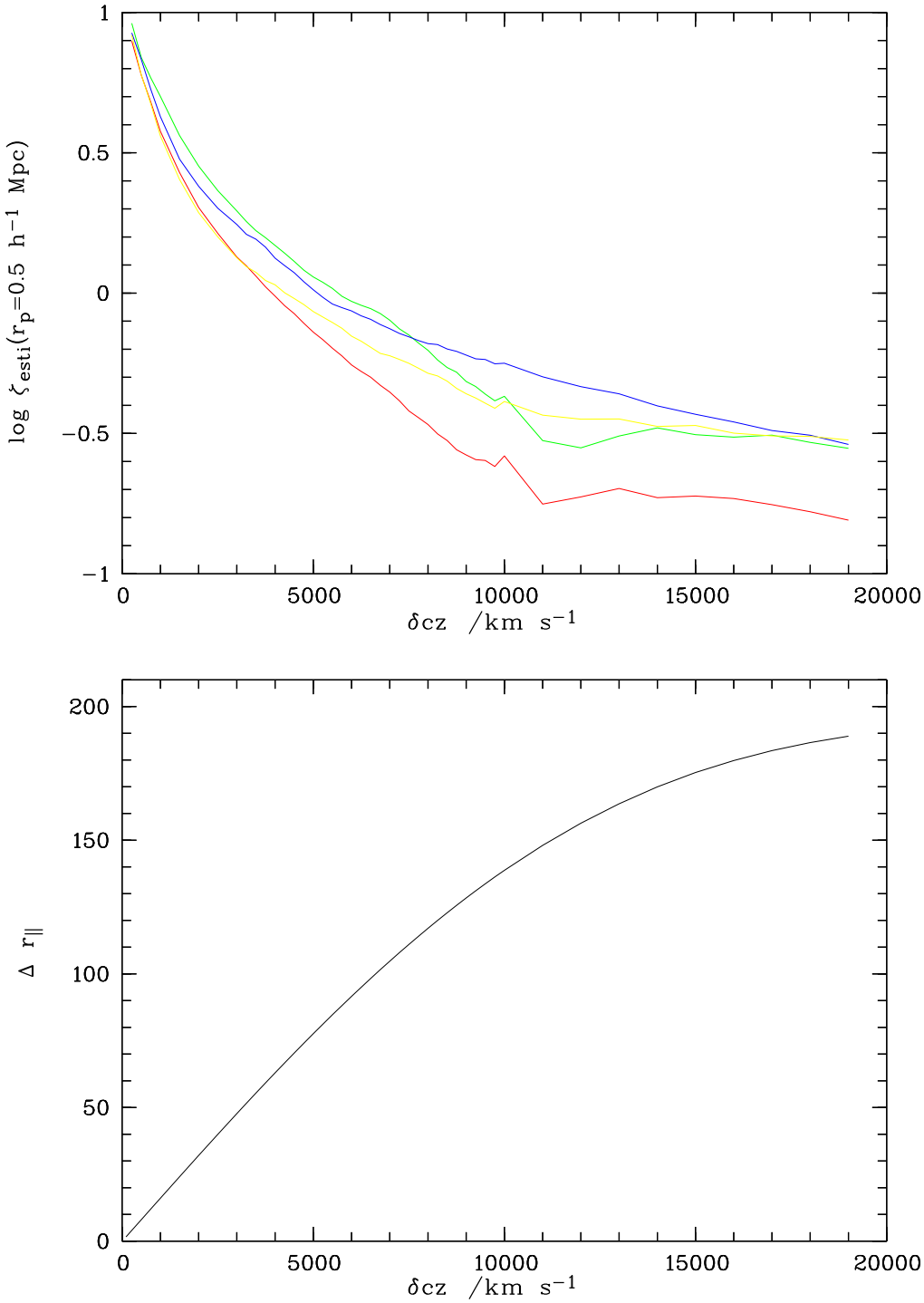

Figure 4.6: Upper plot: the amplitude of ζ_{esti} (see equation (3.30)), at $r_p = 0.5h^{-1}$Mpc). The colors denote the same sectors as in Figure (4.3), lower plot: the depth $\Delta r_{\|}$ in dependence of the integration limits.

Table 4.4: The amplitude of the three-dimensional correlation function at $r = 0.5h^{-1}$Mpc, deduced from the projected correlation function, and the correlation length r_0 which was calculated from that.

Sector	$w(r_p = 0.5h^{-1}\text{Mpc})\,/h^{-1}\,\text{Mpc}$	γ	$\Rightarrow \xi(r = 0.5h^{-1}\text{Mpc})$	$r_0/h^{-1}\,\text{Mpc}$
37-38	95.931 ± 1.193	1.739 ± 0.150	49.224 ± 0.732	4.699 ± 0.909
35-41	85.683 ± 0.877	1.788 ± 0.126	46.072 ± 0.549	4.259 ± 0.643
190-3	64.427 ± 0.691	1.933 ± 0.087	39.078 ± 0.278	3.331 ± 0.284
193-11	64.351 ± 0.633	1.789 ± 0.097	34.633 ± 0.422	3.627 ± 0.390

Table 4.5: Comparison of the values of $\xi(r = 0.5h^{-1}\text{Mpc})$ derived in the three different ways.

Sector	$\xi(r = 0.5\ h^{-1}\text{Mpc})$	$\xi(r = 0.5\ h^{-1}\text{Mpc})$ from $w(\theta)$	$\xi(r = 0.5\ h^{-1}\text{Mpc})$ from $w(r_p)$
37-38	19.991 ± 0.290	61.982 ± 0.701	49.224 ± 0.732
35-41	15.577 ± 0.183	62.422 ± 0.583	46.072 ± 0.549
190-3	13.116 ± 0.131	62.289 ± 0.337	39.078 ± 0.278
193-11	11.574 ± 0.131	38.012 ± 0.286	34.633 ± 0.422

As we have shown, the amplitude of the projected correlation function reaches it's maximum when the integration limits have the same size as the velocity dispersion. To capture the bulk of the correlation signal with a minimum of noise, the appropriate choice of the integration limits is given by the onset of the maximum – thus we chose $\delta cz = 3000\,\text{km s}^{-1}$ for the calculation of the projected correlation function in the LCRS.

Using equation (3.20) and $\xi(r = 0.5h^{-1}\text{Mpc}) = 0.5^{-\gamma} r_0^{\gamma}$, we can now calculate the amplitude of the three-dimensional correlation function at $r_p = 0.5h^{-1}$Mpc from the fitted amplitude of the projected correlation function:

$$\xi(r = 0.5h^{-1}\text{Mpc}) = \frac{0.5^{-\gamma} w(r_p)}{C r_p^{1-\gamma}} \ . \tag{4.7}$$

We also give the correlation length r_0 which was calculated from the amplitude at $r_p = 0.5h^{-1}$Mpc). Table (4.4) shows the results.

The weighted mean is $\langle \xi(r = 0.5\ h^{-1}\text{Mpc}) \rangle = 39.802 \pm 0.205$, with $\langle \gamma \rangle = 1.837 \pm 0.054$. This corresponds to a correlation length of $r_0 = 3.713 \pm 0.218$.

A comparison of all the values for $\xi(r = 0.5\ h^{-1}\text{Mpc})$, derived in the three different ways is given in Table (4.5).

Those values differ significantly from each other, but if one takes into account the different slopes, and calculates the correlation length r_0 from the amplitude at $r = 0.5\ h^{-1}\,\text{Mpc}$ (as

summarised in Table (4.6)), the values of r_0 are the same within their errors. A flat slope can compensate for a low amplitude, if one extrapolates to larger scales.

Table 4.6: Comparison of the correlation length r_0, derived in the three different ways.

Sector	r_0	r_0 from $w(\theta)$	r_0 from $w(r_p)$
37-38	5.777 ± 0.536	5.462 ± 1.527	4.699 ± 0.909
35-41	4.236 ± 0.243	5.167 ± 1.293	4.259 ± 0.643
190-3	4.236 ± 0.243	4.677 ± 0.951	3.331 ± 0.284
193-11	3.337 ± 0.140	3.302 ± 0.780	3.627 ± 0.390

4.2.4 Influence of the redshift errors on the correlation function

Errors in the redshift measurement basically lead to increasing noise in the correlation signal. If the clustering in redshift space is more and more washed out (the redshift distribution becomes more and more Poisson-like), the amplitude decreases, especially at small scales.

To prove this assumption and to estimate the size of the effect, we assigned artificial errors to the measured redshifts of the galaxies in the catalogue, picked randomly from a Gaussian error distribution:

$$\hat{z} = z + \Delta z \ , \tag{4.8}$$

where Δ_z is randomly drawn from the distribution

$$p(\Delta z) = \frac{1}{\sqrt{2\pi}\ \sigma} \exp\left[-0.5\left(\frac{\Delta z}{\sigma_z}\right)^2\right] \ . \tag{4.9}$$

With these artificial errors we computed $\xi(r)$ and $w(r_p)$. As expected, the three-dimensional real space correlation function $\xi(r)$ suffers extremely from the increase of noise and the effect of the smeard-out correlation signal along the line of sight. Figure (4.7) shows the correlation function of the galaxies in field 37-38, one calculation with $\sigma_z = 0.007$, and one calculation with $\sigma_z = 0.017$. In velocity space this corresponds to a full width at half maximum (FWHM) of $\Delta cz = 2\sqrt{2\ln 2} \cdot c\sigma_z = 5000\,\mathrm{km\ s^{-1}}$ and $\Delta cz = 12000\ \mathrm{km\ s^{-1}}$, respectively. In both cases the noise is too large (there are only three data points greater than zero) to recover the correlation signal. Effectively the measurement breaks down completely even for $\sigma_z = 0.007$.

To evaluate the effect of the redshift errors on the projected correlation function, we repeated the calculation of $w(r_p)$ for increasing integration limits δcz, one time with an simulated redshift error of $\Delta cz = 5000\,\mathrm{km\ s^{-1}}$, and a second time with $\Delta cz = 12000\,\mathrm{km\ s^{-1}}$, which corresponds to the size of the CADIS redshift errors. Figure (4.8) shows the weighted mean of the fitted amplitudes of the projected correlation function in the four sectors, fitted at $r_p = 0.5h^{-1}$ Mpc, for increasing integration limits. The maximum of the curve is shifted towards larger integration limits, for the calculation with $\Delta cz = 5000\,\mathrm{km\ s^{-1}}$, it can be located at around $\delta cz = 6000\,\mathrm{km\ s^{-1}}$. At this point, when the integration limits become clearly larger than the error, one starts to sample the correlation signal. Nevertheless, the

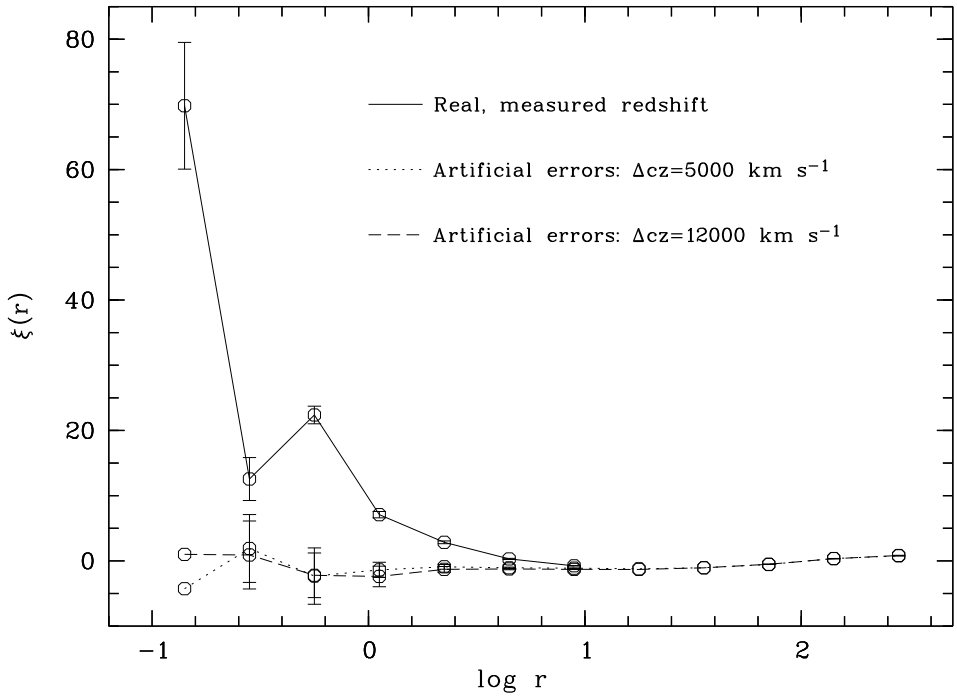

Figure 4.7: The influence of redshift errors on the real space two-point correlation function $\xi(r)$

maximum is not only shifted, but also the amplitude is smaller than in the case without or nearly without redshift errors. The noise, which makes the distribution more poisson-like, diminishes the amplitude, although the projected correlation function reacts not as sensitive to redshift errors as the three-dimensional correlation function. In the case of the calculation for $\Delta v = 12000\,\mathrm{km\ s^{-1}}$ (which corresponds to a depth that is almost as deep as the survey itself), the maximum obviously lies in the range where the correlation signal decreases again due to the scaling relation, as explained above.

The important result of this investigation is that:

- the integration limits have to be chosen at least as large as the full width at half maximum of the redshift error distribution to sample the complete correlation signal, and

- one has to take into account that the measured amplitude becomes smaller, if redshift errors exceed the noise due to peculiar velocities.

If we want to calculate the projected correlation function for the CADIS data, the appropriate choice of the integration limits is $\pm\delta cz = 15000\,\mathrm{km\ s^{-1}}$.

In principle there are two possibilities to take the influence of the redshift errors on the projected correlation function into account. To facilitate the direct comparison of CADIS with the LCRS, we simulate the influence of the redshift errors, and compare CADIS with the

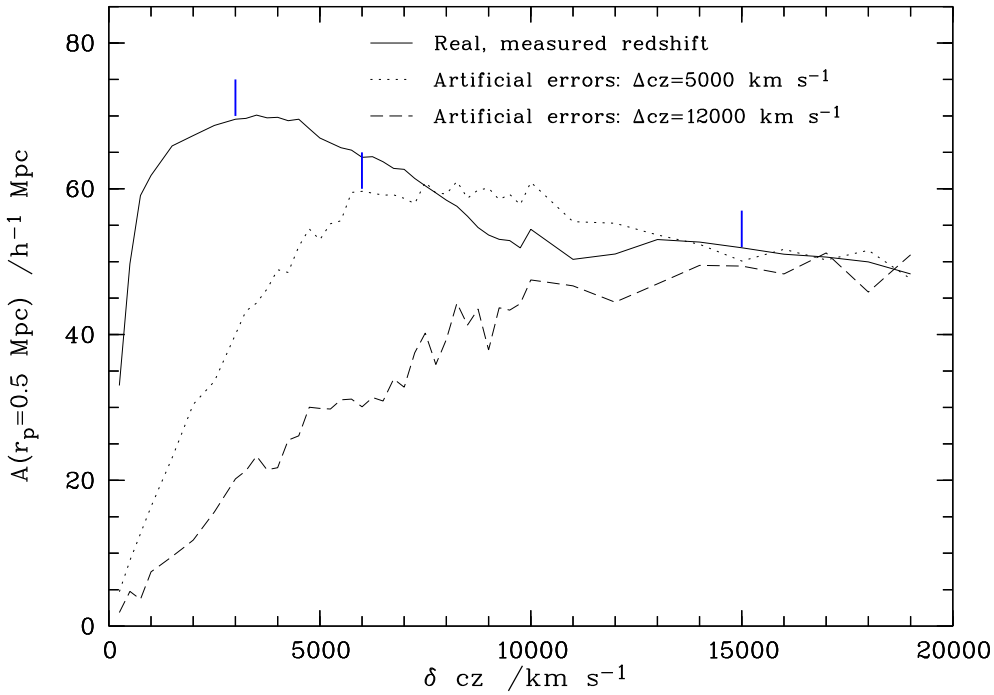

Figure 4.8: The influence of redshift errors on the projected two-point correlation function, shown is the weighted mean of the four sectors. $w(r_p)$ for increasing integration limits is calculated for an artificial redshift error of $\sigma_z = 0.007$ (which corresponds to a FWHM of $\Delta cz = 5000\,\mathrm{km\ s^{-1}}$), dotted line, and for $\sigma_z = 0.017$ ($\equiv 12000\,\mathrm{km\ s^{-1}}$, dashed line. The errors of the fits are not plotted here to avoid confusion. The blue marks indicate the integration limits which have to be chosen for the calculation of $w(r_p)$.

simulated LCRS data – which includes the same choice of the integration limits. For a comparison with results published by of other authors, we can correct the measured amplitudes of the CADIS data for the effect: the maximum amplitude deduced from the unchanged LCRS data is larger by a factor of ~ 1.4 than the maximum amplitude of the data with simulated redshift errors of the size of the CADIS redshift errors. Multiplying our measured amplitudes with 1.4 will yield values which can be compared with the literature.

Chapter 5

The evolution of galaxy clustering

In this chapter we will investigate the evolution of the large scale clustering of field galaxies. The data set on which this investigation was carried out, is the CADIS multicolor sample, as described in Chapter 2. For the comparison with the clustering properties of galaxies in the local universe we used the Las Campanas Redshift Survey. For the direct comparison with CADIS, we introduce artificial redshift errors in the LCRS data (see Chapter 4).

With redshift errors of $\sigma_z = 0.017$, as is the accuracy of the redshift determination for the faintest galaxies ($I > 22$), which was achieved using our special multicolor method (Wolf et al., 2001c)), it is obviously not possible to derive the correlation length directly from the three dimensional correlation function $\xi(r)$ (see Chapter 4). However, it is possible to derive it either from the angular correlation function $w(\theta)$, or the projected correlation function $w(r_p)$.

For each of the four CADIS fields under consideration we generate a catalogue of randomly distributed galaxies, with the same properties as the real data - namely the same borders in x and y coordinates, and with redshifts that follow a Gaussian distribution:

$$\frac{\mathrm{d}N}{\mathrm{d}z} = 6.6 \cdot 10^2 \exp\left[-0.5\left(\frac{z-0.6}{0.35}\right)^2\right] , \tag{5.1}$$

the smoothed mean redshift distribution of the galaxies in the four fields in the redshift range $0 \leq z \lesssim 1.1$ (see Figure (2.4)).

"Same properties" does not only mean the same field geometry and smoothed redshift distribution, but also to simulate the influence of bright stars in the field – they can cover galaxies, that therefore can not be accounted for in the calculation of the correlation function.

Bright stars ($R \leq 15.5$) are masked out, nine in the 16 h field, nine in the 23 h field, one in the 1 h field, and none in the 9 h field. For each star the radius of the circles for the mask is chosen such that the image profile has decreased to the flat minimum overall countlevel.

5.1 The angular correlation function

Instead of simply calculating the angular correlation function for the whole sample, we can make use of the redshift information available, and calculate $w(\theta)$ in distinct redshift bins. Then we can compare the amplitudes with model calculations for different evolution scenarios.

We computed the angular correlation function $w(\theta)$ for each of the four CADIS fields separately, and then calculated the weighted mean:

$$\begin{aligned} < w(r_p) > &= \sum_{fields} \left[\frac{w(r_p)_{field}}{\sigma^2_{w(r_p)_{field}}} \right] \cdot \left[\sum_{fields} \frac{1}{\sigma^2_{w(r_p)_{field}}} \right]^{-1} \\ &= \frac{\left[\frac{w(r_p)_{1\mathrm{h}}}{\sigma^2_{w(r_p)_{1\mathrm{h}}}} + \frac{w(r_p)_{9\mathrm{h}}}{\sigma^2_{w(r_p)_{9\mathrm{h}}}} + \frac{w(r_p)_{16\mathrm{h}}}{\sigma^2_{w(r_p)_{16\mathrm{h}}}} + \frac{w(r_p)_{23\mathrm{h}}}{\sigma^2_{w(r_p)_{23\mathrm{h}}}} \right]}{\left[\frac{1}{\sigma^2_{w(r_p)_{1\mathrm{h}}}} + \frac{1}{\sigma^2_{w(r_p)_{9\mathrm{h}}}} + \frac{1}{\sigma^2_{w(r_p)_{16\mathrm{h}}}} + \frac{1}{\sigma^2_{w(r_p)_{23\mathrm{h}}}} \right]} , \end{aligned} \tag{5.2}$$

and with errors

$$\begin{aligned} \sigma_{<w(r_p)>} &= \frac{1}{\sqrt{\sum_{fields} \frac{1}{\sigma^2_{w(r_p)_{field}}}}} \\ &= \frac{1}{\sqrt{\frac{1}{\sigma^2_{w(r_p)_{1\mathrm{h}}}} + \frac{1}{\sigma^2_{w(r_p)_{9\mathrm{h}}}} + \frac{1}{\sigma^2_{w(r_p)_{16\mathrm{h}}}} + \frac{1}{\sigma^2_{w(r_p)_{23\mathrm{h}}}}}} . \end{aligned} \tag{5.3}$$

Figure (5.1) shows the angular correlation function in four redshift intervals at $0.2 \leq z < 0.4$, $0.4 \leq z < 0.6$, $0.6 \leq z < 0.8$, and $0.8 \leq z \leq 1.0$, between angular scales of $1'' \lesssim \theta \lesssim 400''$. The mean angular correlation of the LCRS is also shown for comparison.

The data was fitted in the range $-0.9 \lesssim \log\theta \lesssim 0.5$ (θ in arcmin). The amplitude was fitted at $\theta = 1'$, Table (5.2) lists the values found in the different redshift bins. We also fitted the amplitude at $\theta = 1'$, keeping the slope fixed at 1.964, the mean value of γ.

In Table (5.1) we list the physical separation which corresponds to $\theta = 1'$ at the central redshifts of the four bins, for three different cosmologies ($\Omega_0 = 1$, $\Omega_\Lambda = 0$, $\Omega_0 = 0.2$, $\Omega_\Lambda = 0$, and $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$, respectively).

Table 5.1: The physical separations in $h^{-1}\,\mathrm{Mpc}$ which correspond to $\theta = 1'$, in different cosmologies.

$\langle z \rangle$	$\Omega_0 = 1$, $\Omega_\Lambda = 0$	$\Omega_0 = 0.2$, $\Omega_\Lambda = 0$	$\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$
0.3	0.165	0.175	0.187
0.5	0.213	0.235	0.256
0.7	0.239	0.273	0.300
0.9	0.252	0.298	0.327

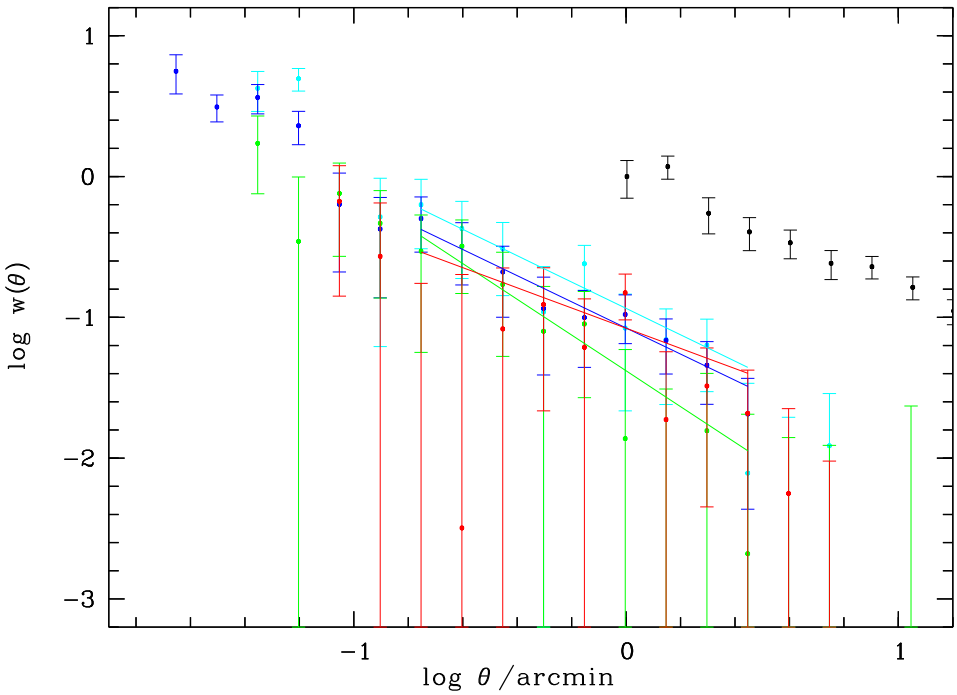

Figure 5.1: The mean angular correlation function of the CADIS multicolor galaxies. Different colors refer to different redshift intervals: light blue: $0.2 \leq z < 0.4$, blue: $0.4 \leq z < 0.6$, green: $0.6 \leq z < 0.8$, red: $0.8 \leq z < 1.8$. For comparison, the mean angular correlation function of the LCRS is also plotted (black). The straight lines are the best fits to the data points, fitted in the range $-0.9 \lesssim \log\theta \lesssim 0.5$.

Table 5.2: The amplitude of the CADIS correlation function at $\theta = 1'$, in different redshift intervals.

z interval	$A_w(1')$	γ	$A_w(1')$ for $\gamma = 1.965$
$0.2 \leq z < 0.4$	0.116 ± 0.010	1.936 ± 0.361	0.114 ± 0.018
$0.4 \leq z < 0.6$	0.084 ± 0.050	1.928 ± 0.285	0.083 ± 0.011
$0.6 \leq z < 0.8$	0.042 ± 0.005	2.270 ± 0.509	0.058 ± 0.014
$0.8 \leq z < 1.0$	0.084 ± 0.007	1.717 ± 0.702	0.081 ± 0.019

To examine how the clustering evolves with redshift, we use equation (3.13), from which we can calculate the theoretically expected amplitudes at $\theta = 1'$ in the four redshift bins for different values of ϵ, and compare the theoretical amplitudes with the measurement. For the correlation length r_0 at $z = 0$, we compute the comoving value from the angular correlation function of the LCRS sample (at $\bar{z} = 0.1$) – $r_0 = 4.664 \pm 0.629 h^{-1}$ Mpc.

For the flat cosmology with non-zero cosmological constant ($\Omega_\Lambda = 0.7$) we have to use approximations for the expressions for $g(z)$ and d_A. We compute equation (3.13), following Longair (1998) and Cabanac et al. (2000):

$$g(z) = \left(\frac{c}{H_0}\right) \frac{1}{(1+z)\sqrt{\Omega_0(1+z)^3 - \Omega_0 + 1}} \tag{5.4}$$

$$d_A = \left(\frac{c}{H_0}\right) \left[z + \frac{z^2}{2}(1+q_0) + \frac{z^3}{6}(3 + 6q_0 + 3q_0^2 - 3\Omega_0) \right] , \tag{5.5}$$

where

$$q_0 = \frac{\Omega_0}{2} - \Omega_\Lambda \ . \tag{5.6}$$

Equation (5.5) is a third-order approximation. Figure (5.2) shows the measured amplitudes and the theoretical "evolution tracks" for the three different cosmologies and for $0 \leq \epsilon \leq 5$.

A formal χ^2 estimation of the best fitting ϵ yields $\epsilon = 1.590 \pm 0.132$ for $\Omega_0 = 1$, $\epsilon = 1.107 \pm 0.134$ for $\Omega_0 = 0.2$, and $\epsilon = 0.209 \pm 0.133$ for $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$. The three corresponding evolutionary tracks are also plotted in Figure (5.2). The results are consistent with linear clustering, where the growth rate is highest for the closed high-density case, and relatively low in the flat model with non-zero cosmological constant.

We will now compute the projected correlation function in different redshift intervals, examine the evolution of the clustering, and then compare the results.

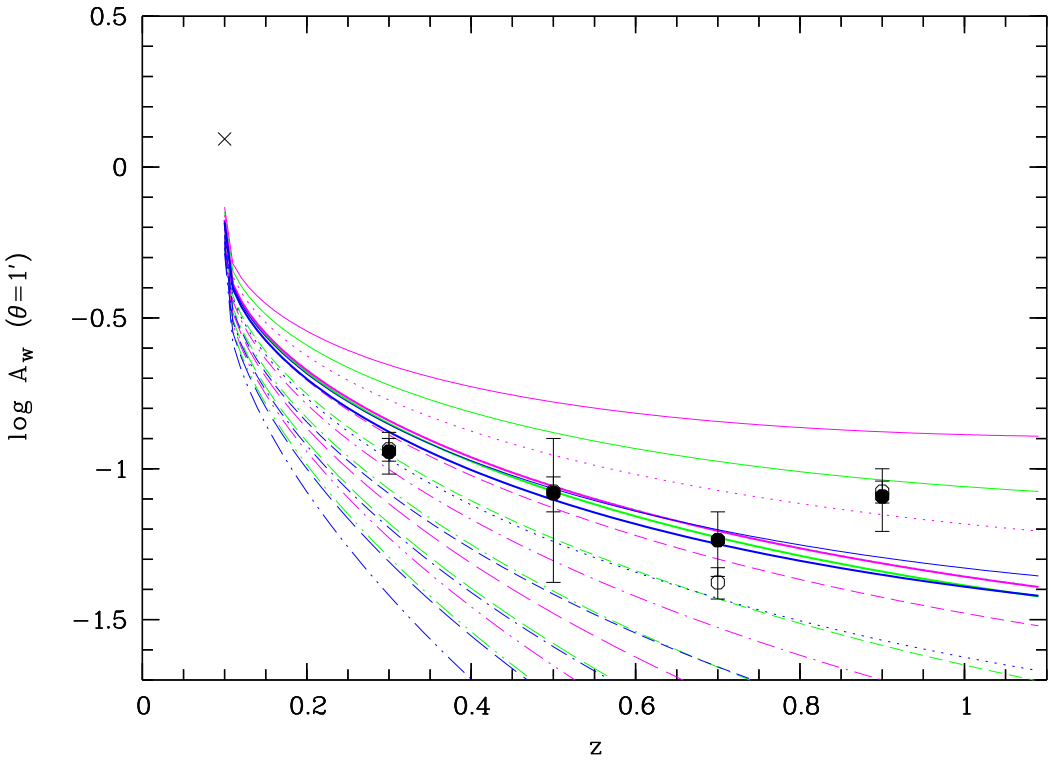

Figure 5.2: The fitted amplitudes of the angular correlation function in four redshift bins; open circles are the amplitudes fitted for γ as a free parameter, filled circles are amplitudes fitted with $\gamma = 1.965$ fixed. The cross is the amplitude of the angular correlation function of the LCRS. Expected evolution of the amplitudes ($A_w(\theta = 1', z)$) for different values of ϵ: green curves are for $\Omega_0 = 0.2$, magenta curves are for $\Omega_0 = 1$, and blue curves are for a flat model with non-zero cosmological constant ($\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$). In all cases the uppermost curve is for $\epsilon_0 = 0$, each succeeding curve is for $\epsilon_n = n$, $n \in N$. The thick lines are the supposed "best-fits".

5.2 The projected correlation function

To investigate the evolution of the clustering of field galaxies between $z = 0$ and $z \approx 1$, we use the inversion of the projected correlation function $w(r_p)$. This approach *does* make use of the redshift information, although an integration is carried out along the line of sight.

Before proceeding any further, we have to make a choice for the integration limits in equation (3.27). In Chapter 4 it was shown how errors in the redshift measurement influence the measurement of the projected correlation function. Therefore we have to take the integration limits slightly larger than the size of the errors ($\sigma_z = 0.017$, which corresponds to a FWHM of $\Delta cz = 12000\,\mathrm{km\ s^{-1}}$). We calculated the projected correlation function for the galaxies in the redshift range $0.4 \leq z \leq 0.8$, for increasing integration limits, see Figure (5.3), and fitted the amplitude at $r_p = 316h^{-1}$ kpc with γ as a free parameter, between $\sim 20h^{-1}$ kpc and $\sim 500h^{-1}$ kpc. The solid line is the weighted mean of the different CADIS fields (other lines).

The different fields behave quite differently, with the scatter getting larger with increasing integration limits. To capture the bulk of the correlation signal with a minimum of noise, we chose $\delta cz = 15000\,\mathrm{km\ s^{-1}}$ for the appropriate integration limits of the calculation of $w(r_p)$.

We divided the sample into three redshift bins, $0.2 \leq z < 0.5$, $0.5 \leq z < 0.75$, and $0.75 \leq z < 1.07$, and computed $w(r_p)$ for each of these bins. For each distinct bin with mean redshift $\bar{z}$, we have to take the selection function ϱ_{pair} of finding pairs of galaxies into account, see equation (3.32). Since the redshift distribution of a certain redshift bin does not "know" what lies outside, the normalisation can not be taken outside – e.g. at the maximum of the total distribution – it has to be chosen for each single redshift bin. Instead of taking the squared redshift distribution (normalised to unity at the mean redshift $\bar{z}$ in that bin), we multiply the redshift distribution of the whole sample with its reflection (reflected at $z = \bar{z}$), and normalise this product to unity at $z = \bar{z}$:

$$\varrho_{pair} = \exp\left[-0.5\left[\frac{z-0.6}{0.35}\right]^2\right] \cdot \exp\left[-0.5\left[\frac{2\bar{z}-z-0.6}{0.35}\right]^2\right] \cdot \left[\exp\left[-0.5\left[\frac{\bar{z}-0.6}{0.35}\right]^2\right]\right]^{-2} \quad (5.7)$$

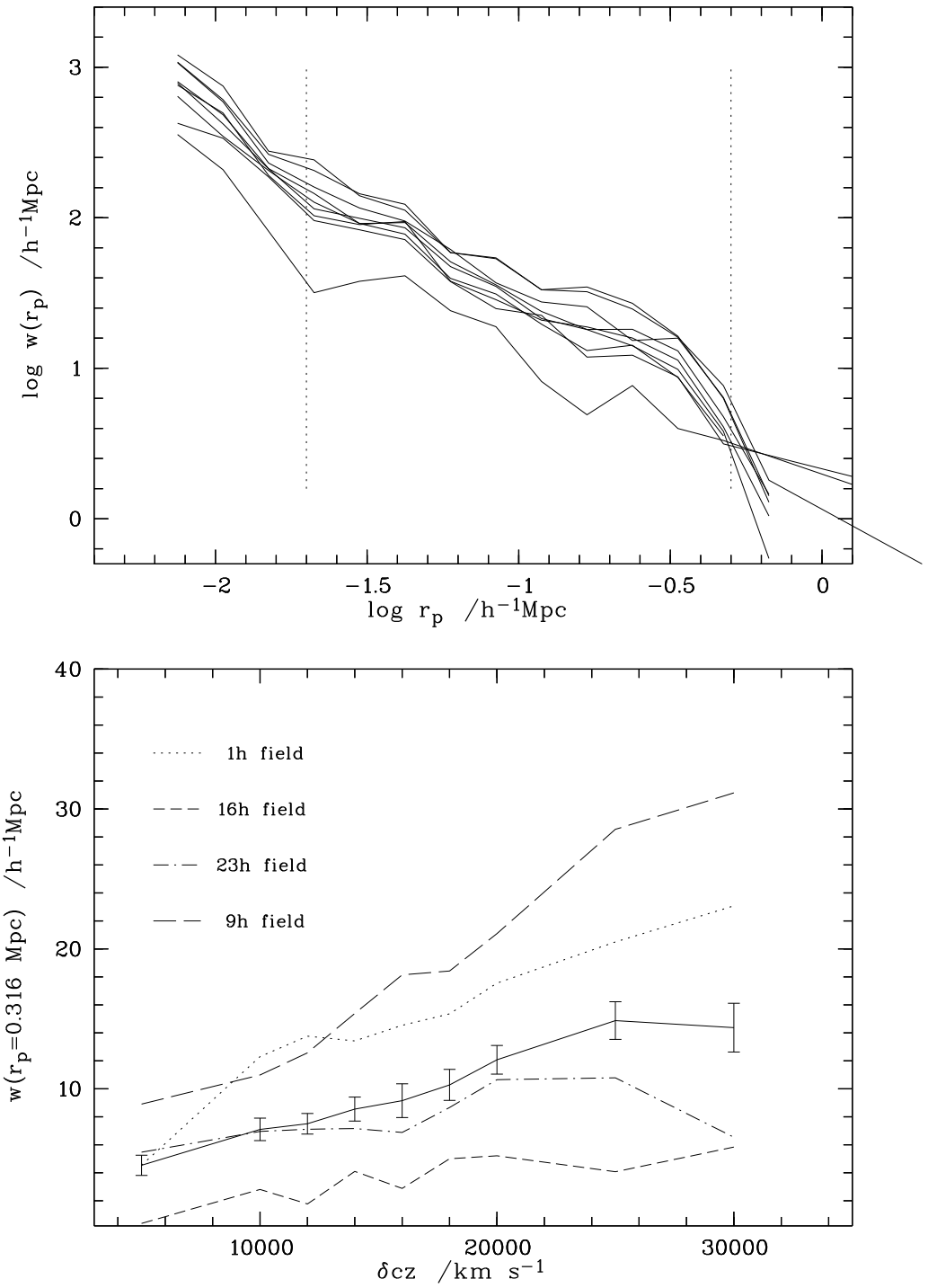

Figure 5.3: Top: Projected correlation function of the galaxies between $z = 0.4$ and $z = 0.8$, for integration limits δcz between 5000 and 30000. The dotted lines indicate the distance range, in which the amplitudes are fitted. Below: fitted amplitudes at $r_p = 316h^{-1}$ kpc – the solid line is the weighted mean of the different CADIS fields (dotted line: 1 h, short-dashed: 16 h, dot-dashed: 23 h, and long-dashed: 9 h).

The projected correlation function for different world models

Up to this point we have computed $w(r_p)$ for a flat universe with no cosmological constant ($\Omega_0 = 1.$, $\Omega_\Lambda = 0$). In the case of the LCRS, this was justified by the modest depth of the fields. A non-zero cosmological constant would not change the results in a significant way. For the CADIS data, where the redshift distribution extends well beyond $z = 1$, we have to take different world models into account. We computed $w(r_p)$ for three different cases: The flat universe with $\Omega_\Lambda = 0$, an open model with $\Omega_0 = 0.2$, and the model favoured by recent observations of supernovae of type Ia (Perlmutter et al., 1999), with $\Omega_0 = 0.3$, and $\Omega_\Lambda = 0.7$.

Using the data from the LCRS, we can connect the CADIS measurements at intermediate redshifts to the "local" universe. The LCRS extends to $z \sim 0.16$, with a median of $\bar{z} = 0.1$. Since the redshifts of the LCRS data are spectroscopically determined, the errors are extremely small ($\sigma_{cz} = 66.7\,\mathrm{km\ s^{-1}}$, and the results of the calculation of the projected correlation function can not directly be compared to the CADIS measurement. The influence of the errors on the projected correlation function was tested in Chapter 4 – the amplitude decreases, even if the integration limits are chosen such that the corresponding distance is larger than the errors.

To facilitate the direct comparison, we can incorporate artificial redshift errors (of the same size as the errors of the CADIS sample, $\Delta cz = 12000\,\mathrm{km\ s^{-1}}$) to the LCRS sample (see section (4.2.4)), and compare the resulting amplitudes of the projected correlation function. Therefore we computed $w(r_p)$ for the three cosmologies under consideration, with integration limits $\delta_{cz} = 15000\,\mathrm{km\ s^{-1}}$.

Figure (5.4) shows the projected correlation function for the flat standard model, an open model, and for the flat model with non-zero cosmological constant. In all three cases the "local" $w(r_p)$ (black) is the weighted mean of all sectors of the LCRS (see Chapter 4).

It is now of little use to estimate the correlation length r_0 from the measurement, since it is in any case far outside the measured range of distances. The CADIS data was fitted in the range $-1.7 \leq \log r_p \leq -0.3$, and the LCRS data in the range $-1.15 \leq \log r_p \leq 0.3$. The amplitude was fitted at $r_p \approx 316h^{-1}$ kpc for CADIS, and for $r_p \approx 500h^{-1}$ kpc for the LCRS, respectively, to make sure we compare the same comoving scales ($\sim 500h^{-1}$ kpc). The best fits are plotted in Figure (5.4) as well.

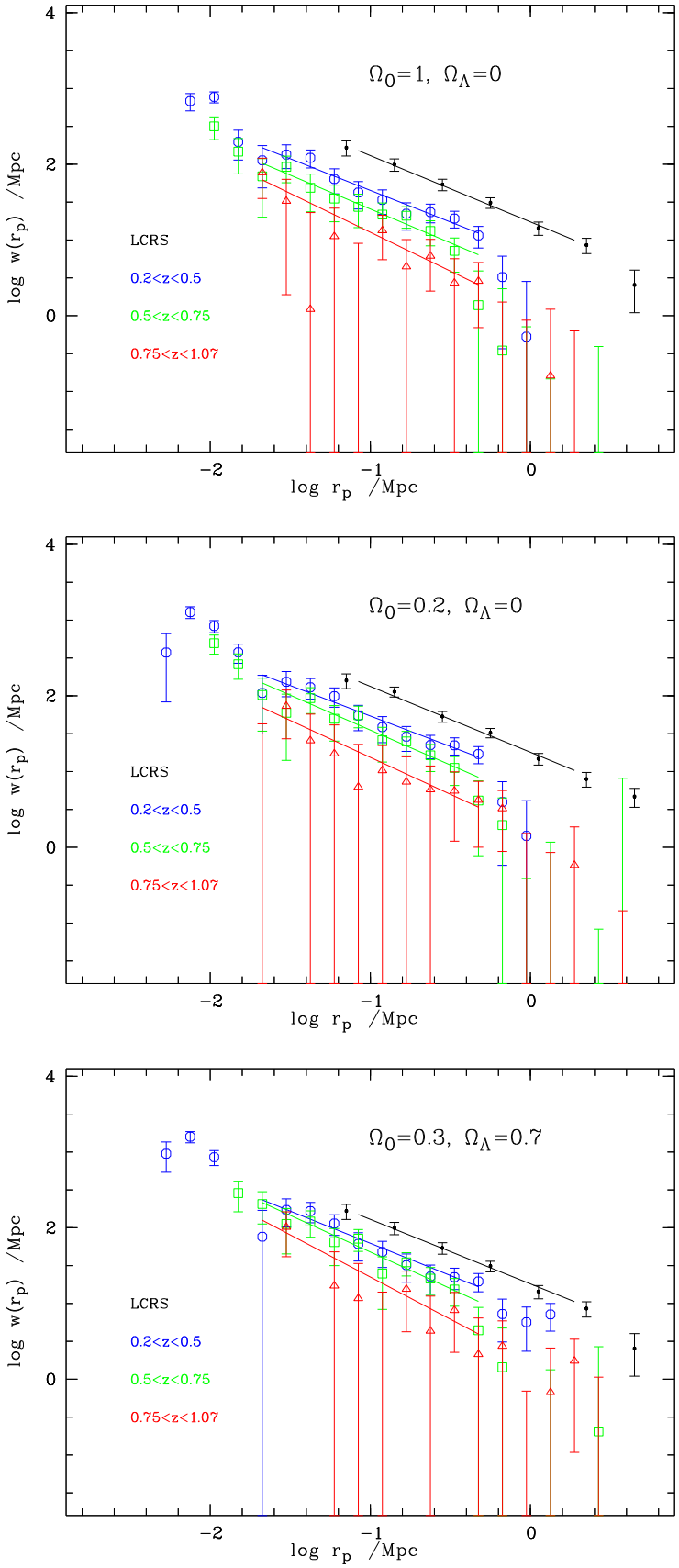

Figure 5.4: Projected correlation function in three redshift bins, for the flat standard model (top), an open model (middle), and for a flat universe with a non-zero cosmological constant (below).

5.2.1 Connection to the "local universe"

Since the purpose of this evaluation is to compare the clustering strength at different cosmic epochs, we can calculate the amplitude of the correlation function at $r = 1h^{-1}$ Mpc from $w(r_p)$, instead of r_0. With equation (3.20), and $\xi(r) = (r/r_0)^{-\gamma}$ we find:

$$\xi(r = 1\mathrm{Mpc}) = r_0^{\gamma} \tag{5.8}$$

$$w(r_p) = C r_p^{1-\gamma} r_0^{\gamma} = C r_p^{1-\gamma} \xi(r = 1\mathrm{Mpc}) \tag{5.9}$$

To facilitate a direct comparison of the possible changes, we compute the amplitude for a *comoving* distance $r_{com} = 1\,\mathrm{Mpc}$ ($r_0^{\gamma} \rightarrow (r_0(1+z))^{\gamma}$):

$$\xi(r_{com} = 1\mathrm{Mpc}) = \frac{w(r_p)(1+z)^{\gamma}}{C r_p^{1-\gamma}} \tag{5.10}$$

$$= \frac{w(r_p)(1+z)^{\gamma}}{\left(\frac{\Gamma(1/2)\Gamma[(\gamma-1)/2]}{\Gamma(\gamma/2)}\right) r_p^{1-\gamma}} \,. \tag{5.11}$$

The errors for the correlation length are calculated following Gaussian error propagation:

$$\Delta\xi(r_{com} = 1\mathrm{Mpc}) = \sqrt{\left(\frac{\partial\xi(r_{com} = 1\mathrm{Mpc})}{\partial w(r_p)}\Delta w(r_p)\right)^2 + \left(\frac{\partial\xi(r_{com} = 1\mathrm{Mpc})}{\partial\gamma}\Delta\gamma\right)^2}$$

where

$$\frac{\partial\xi(r_{com} = 1\mathrm{Mpc})}{\partial w(r_p)} = \frac{(1+z)^{\gamma}\Gamma\left[\frac{\gamma}{2}\right]}{\sqrt{\pi}\Gamma\left[\frac{\gamma}{2} - \frac{1}{2}\right] r_p^{1-\gamma}} \,, \tag{5.12}$$

$$\frac{\partial\xi(r_{com} = 1\mathrm{Mpc})}{\partial\gamma} = \frac{w(r_p)(z+1)^{\gamma}\Gamma\left[\frac{\gamma}{2}\right]\left[2(\ln r_p \ln(z+1)) - \psi^{(0)}\left(\frac{\gamma}{2} - \frac{1}{2}\right) + \psi^{(0)}\left(\frac{\gamma}{2}\right)\right]}{2\sqrt{\pi}\Gamma\left[\frac{\gamma}{2} - \frac{1}{2}\right] r_p^{1-\gamma}} \tag{5.13}$$

$\psi^{(0)}$, the *digamma function*, is the logarithmic derivative of the gamma function.

The fitted amplitudes A at $r_p \approx 316h^{-1}$ kpc for CADIS and at $r_p \approx 500h^{-1}$ kpc for the LCRS, respectively, and the amplitude of the three-dimension correlation function at $r_{com} = 1\,\mathrm{Mpc}$ derived from them, are listed in Table (5.3) (for γ as a free parameter).

The differences between the values in different world models are small, however, $\xi(r_{com} = 1\,\mathrm{Mpc})$ is at all redshifts smallest for a flat high-density cosmology, and largest in the flat case with non-zero cosmological constant. The projected separation r_p corresponding to the same angle θ is larger for smaller values of Ω_0, and even larger if $\Omega_\Lambda \neq 0$ (see Table (5.1)). This means that the physical distances are stretched (or compressed, respectively) corresponding to the cosmology, and the projected correlation function is shifted along the r_p-axis. Since those differences are increasing with increasing redshift, the size of this "shift" depends on redshift, and we expect to measure a slightly different evolution of the correlation function.

Table 5.3: The amplitude of the projected correlation function in different redshift intervals, at $r_p \approx 316h^{-1}$ kpc for CADIS and at $r_p \approx 500h^{-1}$ kpc for the LCRS (first lines), respectively, and $\xi(r_{com} = 1h^{-1}$ Mpc), for different world models. Numbers in italic indicate that this values are inferred from the LCRS.

Model	z interval	A	γ	$\xi(r_{com} = 1\,\mathrm{Mpc}$
	0.04≤ z ≤ 0.16	*32.05±0.796*	*1.876±0.180*	*5.999±0.124*
$\Omega_0 = 1.0$	0.2 ≤ z < 0.5	17.271±0.770	1.835 ± 0.203	3.210 ± 0.143
$\Omega_\Lambda = 0$	0.5 ≤ z < 0.75	9.160±0.822	1.895 ± 0.295	2.414 ± 0.234
	0.75 ≤ z < 1.07	3.881±0.831	2.026 ± 0.359	1.431 ± 0.318
	0.04≤ z ≤ 0.1	*33.297±0.732*	*1.886±0.163*	*6.198±0.215*
$\Omega_0 = 0.2$	0.2 ≤ z < 0.5	21.441±0.857	1.805 ± 0.207	3.981 ± 0.163
$\Omega_\Lambda = 0$	0.5 ≤ z < 0.75	12.202±0.925	1.922 ± 0.289	3.226 ± 0.261
	0.75 ≤ z < 1.07	5.013±0.957	1.975 ± 0.446	1.831 ± 0.389
	0.04≤ z ≤ 0.16	*33.678±0.786*	*1.882±0.183*	*6.148±0.334*
$\Omega_0 = 0.3$	0.2 ≤ z < 0.5	23.404±1.012	1.850 ± 0.221	4.348 ± 0.188
$\Omega_\Lambda = 0.7$	0.5 ≤ z < 0.75	15.721±1.082	1.966 ± 0.261	4.170 ± 0.293
	0.75 ≤ z < 1.07	6.195±1.260	2.113 ± 0.485	2.311 ± 0.485

Quantification of the evolution

To quantify the evolution of the clustering strength, we can now estimate the parameter q (see equation (3.21)) from the data. Figure (5.5) shows the logarithm of the amplitudes of the real space correlation function at $1h^{-1}$ Mpc, versus $\log(1+z)$. Then q is simply the slope of the straight line fitted into the data.

For the closed model we find $q = -2.68 \pm 0.16$, for the open model we find -1.92 ± 0.17, and for the model with non-zero cosmological constant $q = -1.23 \pm 0.20$. q gives directly the deviation from the global Hubble flow (see Chapter 3): large negative q-values indicate a rapid growth of the clustering strength between a redshift of $z \sim 1$ and today, $q = -2$ is expected for linear growth of structure, a value of $q \approx -1.3$ means that at $z \sim 1$ there exists already the same kind of structure as we see today, and the clustering growth is due to the dilution of the galaxy background. Although the individual values of $\xi(r_{com} = 1\,\mathrm{Mpc})$ are rather consistent for the different cosmologies under consideration, the values of q do not agree within their errors. This formal dependency on the cosmology adopted for the calculation is much larger than expected from the difference in physical scale (see Table (5.1)). We regard this as an indication that the peculiarities in our pencil beams have not averaged out properly yet. Presumably high statistics and more fields are required to reach a robust determination of q. Note that the mean of the above values ($q \approx -1.9$) agrees well with the value derived in COMBO 17 (see Chapter 6).

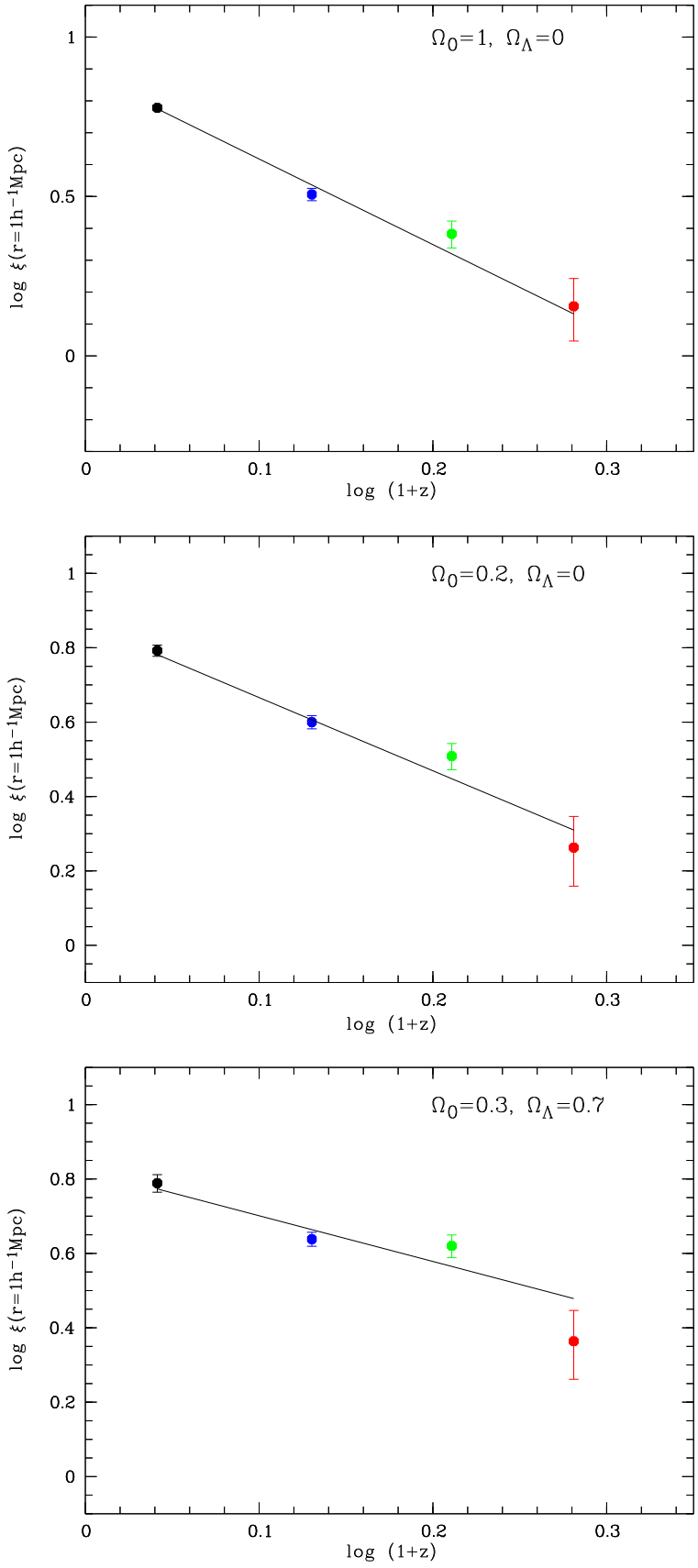

Figure 5.5: The evolution of the clustering strength (at $1h^{-1}$ Mpc) with redshift. The line is the fit to the data points, the first data point is the weighted mean of the LCRS sectors, the three other ones are CADIS data.

Table 5.4: Comparison between the values of $\xi(r = 1h^{-1}$ Mpc) calculated by using the results from the angular correlation function, with the values deduced from the projected correlation function (amplitudes have been multiplied with 1.4 to correct for the influence of redshift errors).

Model	z interval	$\xi(r_{com} = 1h^{-1}$ Mpc) from $w(\theta)$	$\xi(r_{com} = 1h^{-1}$ Mpc)from $w(r_p)$
$\Omega_0 = 1$ $\Omega_\Lambda = 0$	$0.2 \leq$ z < 0.5	4.868 ± 0.223	4.494 ± 0.200
	$0.5 \leq$ z < 0.75	2.956 ± 0.143	3.379 ± 0.328
	$0.75 \leq$ z < 1.07	1.914 ± 0.977	2.003 ± 0.445
$\Omega_0 = 0.2$ $\Omega_\Lambda = 0$	$0.2 \leq$ z < 0.5	5.627 ± 0.257	5.573 ± 0.228
	$0.5 \leq$ z < 0.75	3.738 ± 0.181	4.516 ± 0.365
	$0.75 \leq$ z < 1.07	2.616 ± 0.134	2.563 ± 0.545
$\Omega_0 = 0.3$ $\Omega_\Lambda = 0.7$	$0.2 \leq$ z < 0.5	6.497 ± 0.289	6.088 ± 0.264
	$0.5 \leq$ z < 0.75	5.097 ± 0.241	5.838 ± 0.411
	$0.75 \leq$ z < 1.07	4.125 ± 0.206	3.235 ± 0.678

Comparison with $w(\theta)$

We can now compare this result with the result found from the examination of the angular correlation function. We found $\epsilon = 1.590 \pm 0.132$ for $\Omega_0 = 1$, $\epsilon = 1.107 \pm 0.134$ for $\Omega = 0.2$, and $\epsilon = 0.209 \pm 0.133$ in the $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$ case. With $\langle\gamma\rangle = 1.964 \pm 0.195$, $r_0(z) = r_{0,z=0}(1+z)^{-(\epsilon+3/\gamma)}$ (in this relation r_0 is in proper coordinates, as would be measured by an observer in the epoch under consideration), and $r_{0,z=0} = 4.664 \pm 0.629h^{-1}$ Mpc (deduced from the angular correlation function of the LCRS data), we can calculate the expected values of $\xi(r_{com} = 1h^{-1}$ Mpc), and compare them with the values deduced from the projected correlation function. In the case of the projected correlation function we have to take the influence of redshift errors into account. The measured amplitudes are decreased by a factor of 1.4 (see Chapter 4), so we can correct for the influence of the redshift errors by multiplying the measured amplitudes by 1.4, and then with equation (5.8) calculate $\xi(r_{com} = 1h^{-1}$ Mpc) from the corrected amplitudes. Table (5.4) compares the values.

The amplitudes of $\xi(r_{com})$ at $1h^{-1}$ Mpc are equal within the errors. Thus we conclude, that both methods lead to the same result.

5.2.2 The evolution of the correlation function for different Hubble types

Different local surveys found the red galaxies to be clustered much stronger than the blue ones (Davis & Geller, 1976). Here we show that this is indeed the case, not only for $z = 0$, and investigate the evolution of the clustering for different Hubble types.

The galaxy library used for the multicolor classification resembles regular grids in redshift and **SED**, see Chapter 2, thus the Hubble type can also be estimated from the observations. This enables us to investigate the evolution of the clustering of different populations of galaxies in the field.

Table 5.5: The amplitude of the CADIS correlation function at $\theta = 1'$ in different redshift intervals, for galaxies with $SED \leq 60$, and $SED > 60$.

z interval	$A_w(SED \leq 60)$	γ	$A_w(SED > 60)$	γ
$0.2 \leq z < 0.4$	0.519 ± 0.088	1.904 ± 0.289	0.048 ± 0.010	2.427 ± 0.929
$0.4 \leq z < 0.6$	0.208 ± 0.022	1.727 ± 0.541	0.065 ± 0.007	2.175 ± 0.504
$0.6 \leq z < 0.8$	0.091 ± 0.014	2.222 ± 0.774	0.060 ± 0.011	1.974 ± 0.970
$0.8 \leq z < 1.0$	0.159 ± 0.035	2.651 ± 0.824	0.073 ± 0.009	2.061 ± 1.110

We devided the sample into two SED bins, with $SED \leq 60$, and $SED > 60$, respectively. For each SED bin we have to generate a sample of randomly distributed data points within the same field geometry. The redshift distribution for the red galaxies ($SED \leq 60$) can be modeled with a Gaussian with its maximum at $\bar{z} = 0.55$, and width $\sigma = 0.3$, the blue galaxies with a modified Gaussian:

$$\frac{\mathrm{d}N}{\mathrm{d}z}_{blue} = 2.95 \cdot 10^2 \exp\left[-0.5\left(\frac{z-0.6}{0.4}\right)^4\right] , \tag{5.14}$$

see Figure (5.6).

We can now calculate both angular and projected correlation function, and analyse the results.

Angular correlation function

For the angular correlation function, we divided the two samples into the same redshift bins as before, $0.2 \leq z < 0.4$, $0.4 \leq z < 0.6$, $0.6 \leq z < 0.8$, and $0.8 \leq z < 1.0$, respectively. Figure (5.7) shows the comparison between angular correlation function of the red galaxies and the blue galaxies, and the fitted amplitudes at $\theta = 1'$ in the four redshift bins for the two SED bins, for the range $-0.9 \leq \log\theta \leq 0.5$. Table (5.5) lists the fitted amplitudes and exponents.

It is obvious that the early type galaxies are stronger clustered than the late type ones. The calculation of the evolutionary parameter ϵ from $w(\theta)$ depends strongly on the value of $r_0(z = 0)$, which we do not know for the subsamples under consideration. We can not use the LCRS for this, and there exists no suitable catalogue of local galaxies which consists of exactly the same population mix as our subsamples. The values measured for r_0 at $z = 0$ certainly depend on the SED cut as well as the values at higher redshifts. To quantify the evolution of the clustering strength, we also computed the projected correlation function, with $\delta cz = 15000\,\mathrm{km\ s^{-1}}$, in the same redshift bins as we did in section (5.2).

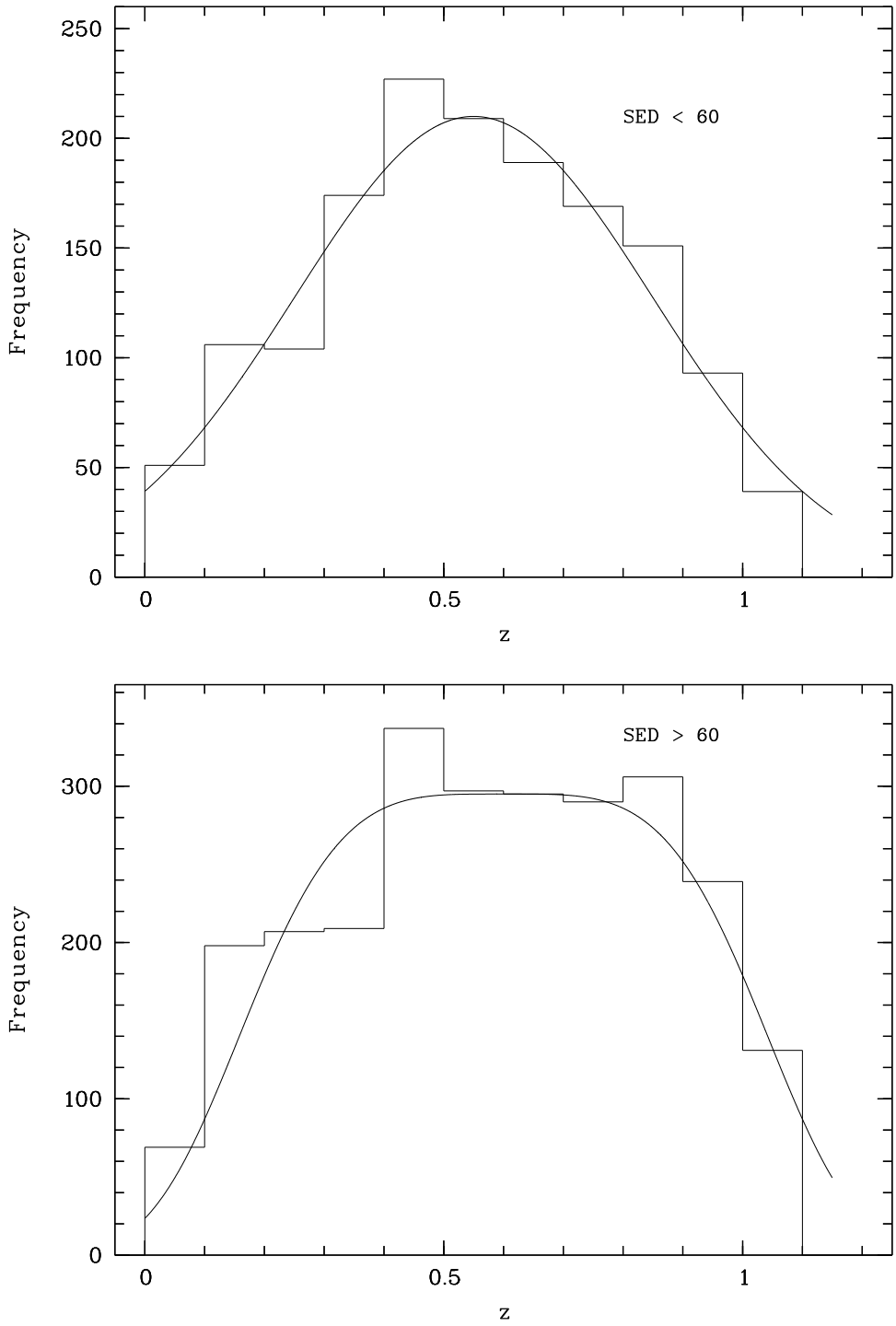

Figure 5.6: Redshift distribution of the galaxies in two redshift bins ($SED \leq 60$, upper figure, and $SED > 60$, lower figure, respectively).

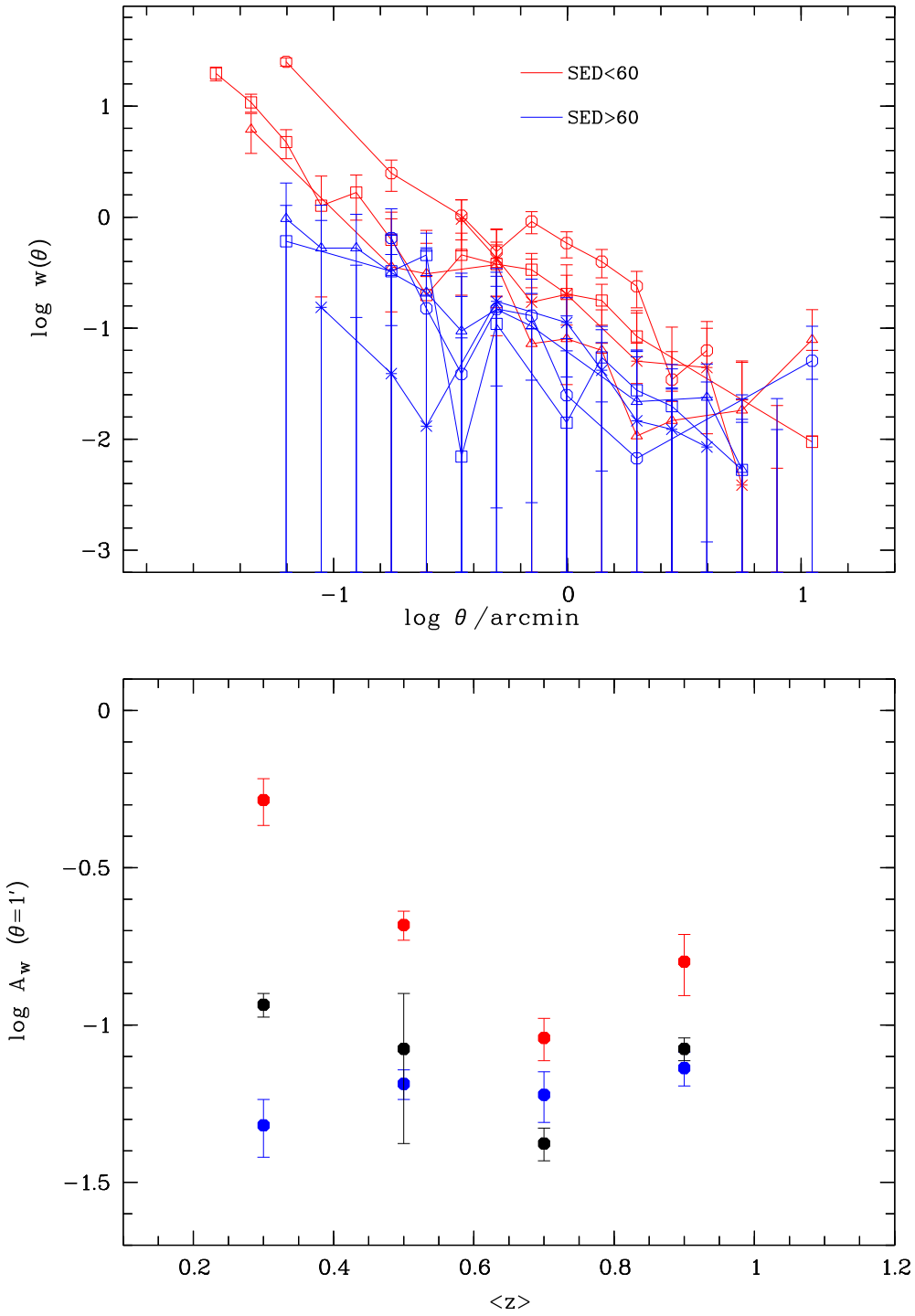

Figure 5.7: Upper plot: The angular correlation function of CADIS galaxies with $SED < 60$ (red), and with $SED > 60$ (blue). The different symbols refer to the different redshift bins: open circles – $0.2 \leq z < 0.4$, open squares – $0.4 \leq z < 0.6$, asterisks: $0.6 \leq z < 0.8$, triangles: $0.8 \leq z \leq 1.0$. Lower plot: amplitudes of the angular correlation function at $\theta = 1'$ in the two SED bins. Blue: $SED > 60$, red: $SED < 60$, black: all galaxies.

Table 5.6: The amplitude of the projected correlation function at $r_p = 316h^{-1}$ kpc, in different redshift intervals for two SED bins, for different world models.

Model	z interval	$A(SED \leq 60)$	γ	$A(SED > 60)$	γ
$\Omega_0 = 1.0$ $\Omega_\Lambda = 0$	$0.2 \leq$ z < 0.5	46.398 ± 1.808	1.906 ± 0.239	7.310 ± 1.476	2.140 ± 0.462
	$0.5 \leq$ z < 0.75	24.229 ± 1.539	2.149 ± 0.279	7.097 ± 1.742	2.053 ± 0.650
	$0.75 \leq$ z < 1.07	9.904 ± 3.524	2.527 ± 1.355	7.398 ± 1.158	1.941 ± 0.425
$\Omega_0 = 0.2$ $\Omega_\Lambda = 0$	$0.2 \leq$ z < 0.5	50.085 ± 1.791	1.998 ± 0.204	8.147 ± 1.587	2.120 ± 0.504
	$0.5 \leq$ z < 0.75	27.117 ± 1.601	2.012 ± 0.337	5.502 ± 1.710	2.317 ± 0.703
	$0.75 \leq$ z < 1.07	14.241 ± 3.336	2.585 ± 1.247	6.280 ± 1.286	1.302 ± 2.512
$\Omega_0 = 0.3$ $\Omega_\Lambda = 0.7$	$0.2 \leq$ z < 0.5	63.926 ± 2.100	1.998 ± 0.231	6.128 ± 2.092	2.474 ± 0.602
	$0.5 \leq$ z < 0.75	39.224 ± 1.801	2.072 ± 0.216	11.910 ± 2.879	2.066 ± 0.941
	$0.75 \leq$ z < 1.07	19.531 ± 4.179	2.157 ± 1.196	9.785 ± 1.833	1.996 ± 0.565

Table 5.7: The amplitudes of the three-dimensional correlation function $\xi(r)$ at a comoving distance of $r = 1\,\mathrm{Mpc}$, for red ($SED \leq 60$) and blue ($SED > 60$) galaxies.

Model	z interval	$\xi(r_{com} = 1\,\mathrm{Mpc},\ SED \leq 60)$	$\xi(r_{com} = 1\,\mathrm{Mpc},\ SED > 60)$
$\Omega_0 = 1.0$ $\Omega_\Lambda = 0$	$0.2 \leq$ z < 0.5	8.598 ± 0.367	1.301 ± 0.305
	$0.5 \leq$ z < 0.75	6.408 ± 0.431	1.886 ± 0.463
	$0.75 \leq$ z < 1.07	3.693 ± 1.381	2.681 ± 0.501
$\Omega_0 = 0.2$ $\Omega_\Lambda = 0$	$0.2 \leq$ z < 0.5	9.184 ± 0.441	1.458 ± 0.336
	$0.5 \leq$ z < 0.75	7.204 ± 0.427	1.425 ± 0.473
	$0.75 \leq$ z < 1.07	5.282 ± 1.410	1.301 ± 7.541
$\Omega_0 = 0.3$ $\Omega_\Lambda = 0.7$	$0.2 \leq$ z < 0.5	11.722 ± 0.574	0.973 ± 0.411
	$0.5 \leq$ z < 0.75	10.717 ± 0.517	3.163 ± 0.768
	$0.75 \leq$ z < 1.07	7.315 ± 1.708	3.588 ± 0.776

Projected correlation function

Figure (5.8) shows $w(r_p)$ for the different redshift bins, for a flat closed, hyperbolic open, and a flat $\Omega_\Lambda = 0.7$ model, respectively. Table (5.6) lists the amplitudes at $r_p = 316h^{-1}$ kpc, fitted between $\sim 20h^{-1}$ kpc and $\sim 500h^{-1}$ kpc, just as in section (5.2).
The corresponding amplitudes of the three-dimensional correlation function $\xi(r)$ at a comoving distance of $r_{com} = 1h^{-1}\,\mathrm{Mpc}$ are given in Table (5.7).

Figure (5.9) shows the amplitudes of the three-dimensional correlation function $\xi(r)$ at a comoving distance of $r = 1\,\mathrm{Mpc}$, for red ($SED \leq 60$) and blue ($SED > 60$) galaxies, for the different world models under consideration. For comparison the data points for the whole sample are also plotted. The lines are the fits for the q-parameter, the data point from the LCRS is not included in the fits for the early type galaxies.

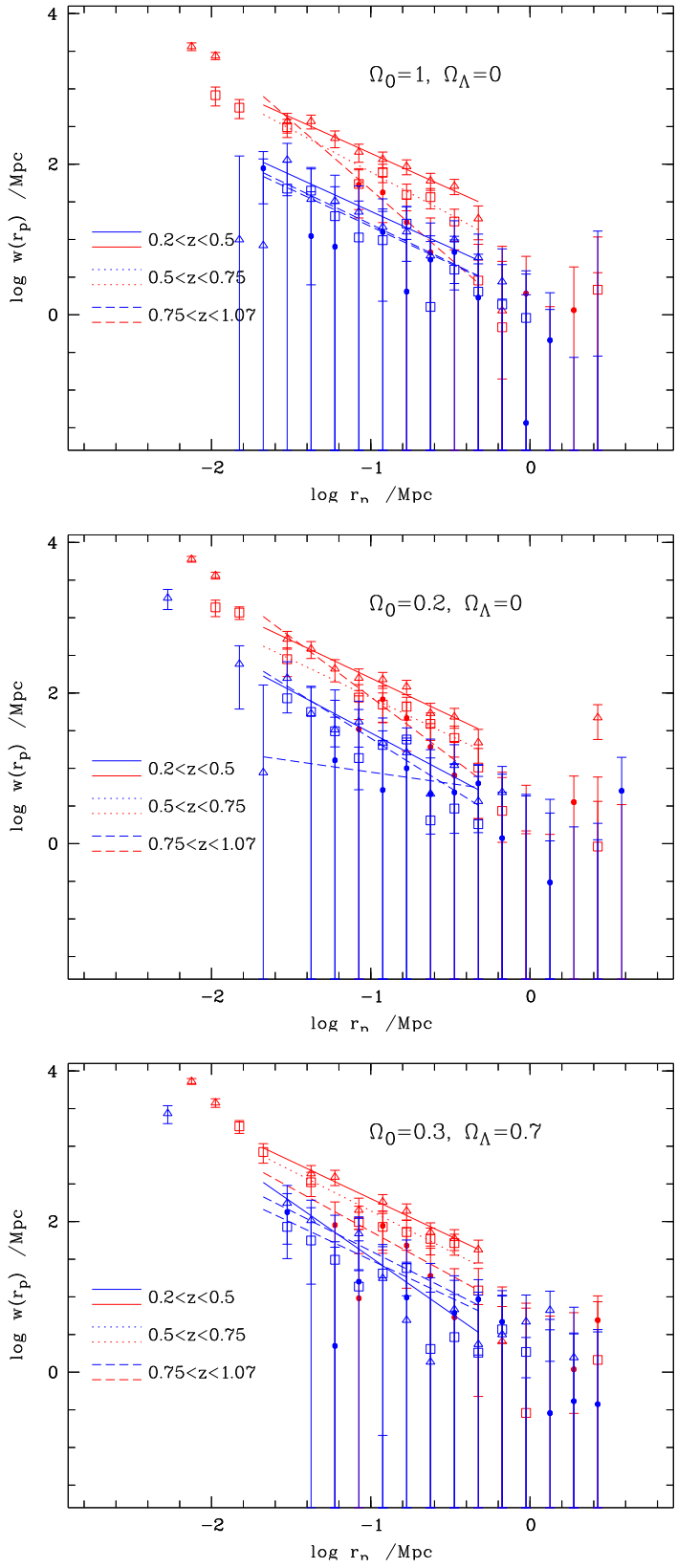

Figure 5.8: The projected correlation function in three different redshift bins, upper figure: flat closed model ($\Omega_0 = 1$, $\Omega_\Lambda = 0$); middle: hyperbolic open model ($\Omega_0 = 0.2$, $\Omega_\Lambda = 0$), lower figure: flat Ω model ($\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$). Red is for early type, blue for late type galaxies.

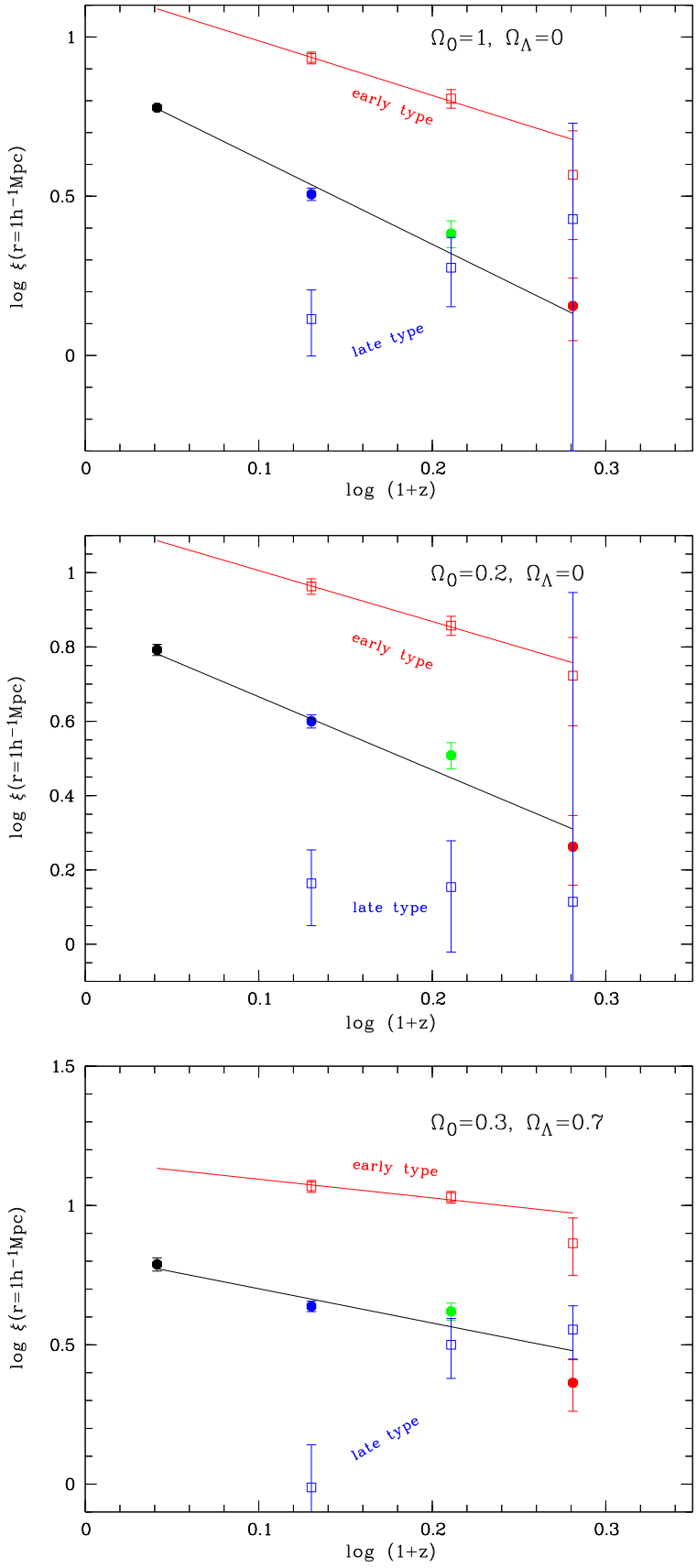

Figure 5.9: The evolution of the clustering strength (at 1 Mpc) with redshift, for early type ($SED \leq 60$) and late type ($SED > 60$) galaxies. The data for the complete sample is plotted for comparison. The first data point (the weighted mean of the LCRS sectors), is not included in the fit for q.

Table 5.8: The parameter q for early type ($SED \leq 60$) and late type ($SED > 60$) galaxies, for the different world models under consideration. The evolution of the complete sample is also shown for comparison. The data point from the LCRS is not included in the fit for early and late type galaxies.

Cosmology	$q(SED \leq 60)$	q_{all}
$\Omega_0 = 1.0$, $\Omega_\Lambda = 0.0$	-1.71 ± 0.39	-2.68 ± 0.16
$\Omega_0 = 0.2$, $\Omega_\Lambda = 0.0$	-1.37 ± 0.36	-1.92 ± 0.17
$\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$	-0.67 ± 0.33	-1.23 ± 0.20

As could already be seen in the case of the angular correlation function, the early type galaxies are significantly stronger clustered than the late type ones, and their evolution is different from the evolution of the whole sample. The evolution of the late type galaxies can not be quantified, because the number of galaxies with $SED > 60$ is not enough to calculate the amplitude of the three-dimensional correlation function with sufficiently high accuracy. The clustering amplitude is very low in any case, but q can not be determined. For the early type galaxies, q is found to be significantly smaller than the value for the whole sample. Table (5.8) lists q for the late type sample, and compares them to the values found for the complete sample.

The clustering evolution of early type galaxies is obviously much slower than for the whole sample, whose clustering behaviour is roughly consistent with linear clustering. This does not inevitably imply that we have to deal with a population of galaxies which act as "test-particles", that simply trace the expansion of the universe. A more plausible explanation for this behaviour arises in the context of biased galaxy formation (Bardeen et al., 1986), where the galaxies form at high redshift in the high-density peaks of the dark matter distribution, and their clustering evolves different from the clustering of the dark matter, see Chapter 6.

5.2.3 The bright and the faint ones

If massive, bright galaxies form in the high density peaks of the dark matter distribution (which are strongly clustered (Kaiser, 1984)), one expects the brighter galaxies to be more strongly clustered than the fainter ones. We divided our sample at $M_B = -18.$, which is approximately the brightness of the Large Magellanic Cloud. With this magnitude cut, we only have enough faint galaxies to calculate the correlation function at relatively low redshifts. Therefore we restrict the analysis of the clustering properties of bright and faint galaxies to the redshift range $0.3 \leq z \leq 0.6$.

Since the absolute magnitude depends on the cosmology chosen for the computation of the luminosity distance (see appendix (B)), the redshift distribution of (absolute) faint and bright galaxies is slightly different for different world models. This is shown in Figure (5.10).

For all cosmologies and $z \leq 1.1$, the smoothed redshift distribution of the faint galaxies can

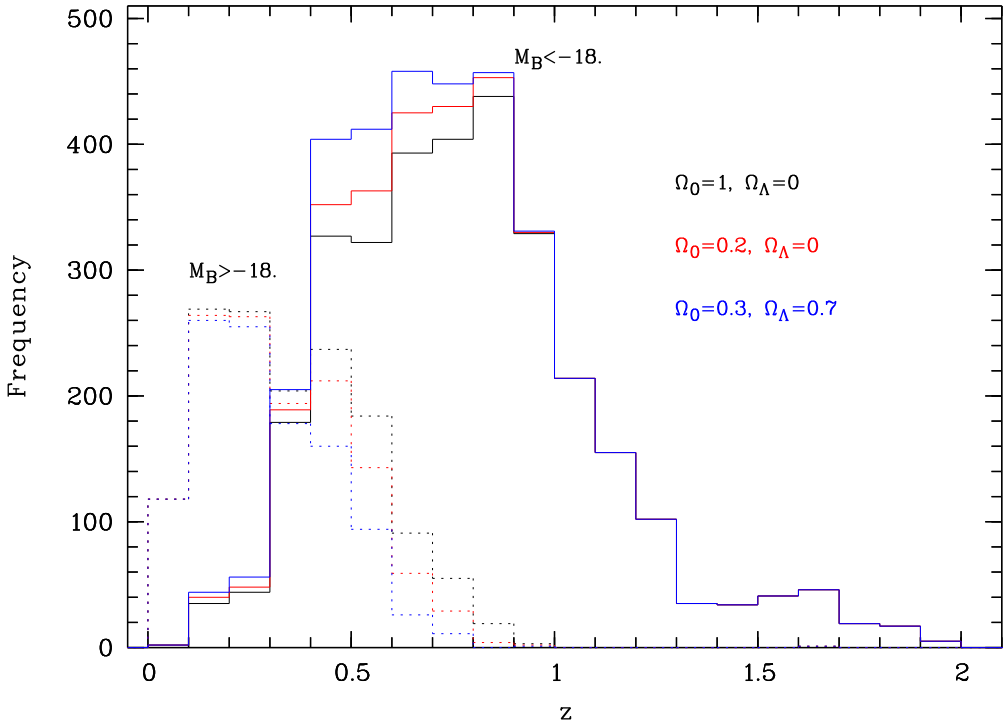

Figure 5.10: Redshift distribution of faint ($M_B > -18$) and bright ($M_B \leq -18$) galaxies, for the different cosmologies under consideration in this work.

be described with a modified Maxwellian:

$$\frac{\mathrm{d}N}{\mathrm{d}z} = n_0 \cdot (z + \hat{z})^2 \exp\left[-0.5 \left(\frac{z}{\hat{\sigma}_z}\right)^2\right] , \tag{5.15}$$

and for the bright galaxies with a modified Gaussian:

$$\frac{\mathrm{d}N}{\mathrm{d}z} = n_0 \cdot \exp\left[-0.5 \left(\frac{z - \bar{z}}{\bar{\sigma_z}}\right)^4\right] . \tag{5.16}$$

Table (5.9) gives $\bar{z}$, $\hat{z}$, $\bar{\sigma_z}$ and σ_z for for bright and faint galaxies, for the different cosmologies.

Table 5.9: Parameter for the description of the redshift distribution of faint and bright galaxies in different world models.

Model	$\hat{z}$	$\hat{\sigma}_z$	$\bar{z}$	$\bar{\sigma_z}$
$\Omega_0 = 1.0,\ \Omega_\Lambda = 0.0$	0.2	0.29	0.72	0.32
$\Omega_0 = 0.2,\ \Omega_\Lambda = 0.0$	0.18	0.27	0.7	0.32
$\Omega_0 = 0.3,\ \Omega_\Lambda = 0.7$	0.2	0.25	0.7	0.32

Table 5.10: Amplitudes of the angular correlation function at $\theta = 1'$, for bright ($M_B < -18$) and faint $M_B \geq -18$ galaxies.

Cosmology	$A_w(M_B \geq -18)$	γ	$A_w(M_B < -18)$	γ
$\Omega_0 = 1.0$, $\Omega_\Lambda = 0.0$	0.098 ± 0.009	1.971 ± 0.393	0.125 ± 0.008	1.617 ± 0.437
$\Omega_0 = 0.2$, $\Omega_\Lambda = 0.0$	0.096 ± 0.009	2.066 ± 0.426	0.128 ± 0.007	1.723 ± 0.332
$\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$	0.085 ± 0.008	1.948 ± 0.447	0.115 ± 0.006	1.744 ± 0.342

Angular correlation function

Again we computed the angular correlation function to get a qualitative insight with better number statistics, and the projected correlation function, to derive a quantitative description of the clustering of bright and faint galaxies.

Figure (5.11) shows the angular correlation function of bright and faint galaxies between $0.3 \leq z \leq 0.6$, in comparison with the angular correlation function of all galaxies, for the three different cosmologies, respectively. Table (5.10) gives the amplitudes at $\theta = 1'$.

The amplitude of the bright galaxies is clearly higher than that of the faint ones.

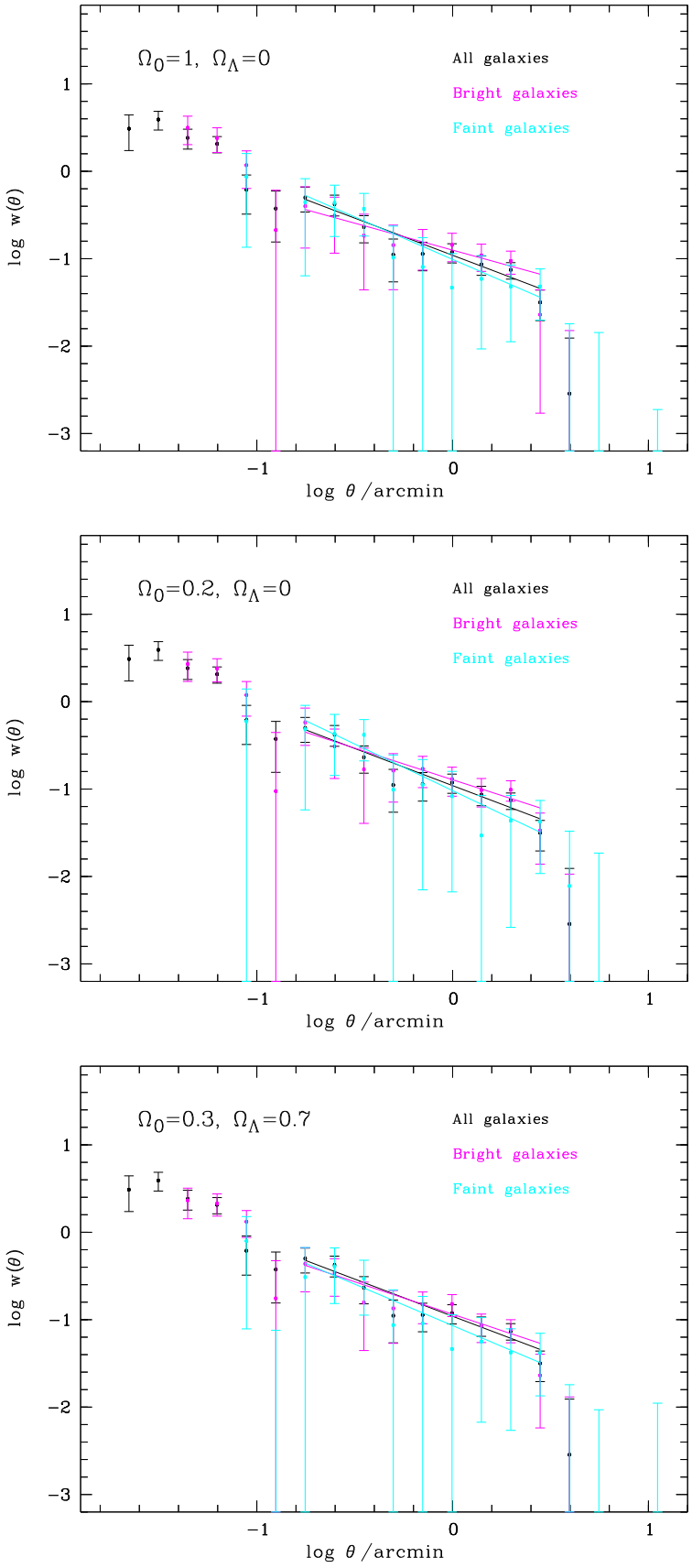

Figure 5.11: The angular correlation function of bright and faint galaxies between $0.3 \leq z \leq 0.6$, in comparison with the angular correlation function of all galaxies, for the three different cosmologies.

Table 5.11: Amplitudes of the projected correlation function at $r_p \approx 316\,\mathrm{kpc}$, for bright ($M_B < -18$) and faint $M_B \geq -18$ galaxies.

Cosmology	$A(M_B \geq -18)$	γ	$A(M_B < -18)$	γ
$\Omega_0 = 1.0$, $\Omega_\Lambda = 0.0$	6.756 ± 0.907	2.086 ± 0.381	15.126 ± 1.068	1.913 ± 0.292
$\Omega_0 = 0.2$, $\Omega_\Lambda = 0.0$	8.681 ± 0.939	2.136 ± 0.331	19.997 ± 1.193	1.772 ± 0.322
$\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$	12.444 ± 1.138	2.003 ± 0.367	22.237 ± 1.430	1.853 ± 0.321

Table 5.12: Amplitudes of the three-dimensional correlation function at $r_{com} = 1h^{-1}$ Mpc, for bright ($M_B < -18$) and faint $M_B \geq -18$ galaxies.

Cosmology	$\xi(r_{com} = 1h^{-1}\,\mathrm{Mpc}),(M_B \geq -18)$	$\xi(r_{com} = 1h^{-1}\,\mathrm{Mpc}),(M_B < -18)$
$\Omega_0 = 1.0$, $\Omega_\Lambda = 0.0$	1.414 ± 0.207	3.213 ± 0.227
$\Omega_0 = 0.2$, $\Omega_\Lambda = 0.0$	1.802 ± 0.224	4.204 ± 0.333
$\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$	2.631 ± 0.256	4.716 ± 0.315

Projected correlation function

We computed the projected correlation function for bright and faint galaxies with $\delta cz = 15000\,\mathrm{km\ s^{-1}}$ in the same redshift interval. For the calculation of the effective depth we used the above mentioned redshift distribution, see equation (5.16), and (5.15), respectively, and Table (5.9). The results for the different cosmologies under consideration in this thesis are shown in Figure (5.12).

The amplitudes of the projected correlation function at $r_p \approx 316\,\mathrm{kpc}$ are given in Table (5.11), Table (5.12) lists the corresponding amplitudes of the three-dimensional correlation function at $r_{com} = 1h^{-1}$ Mpc, which have been deduced by means of equation (5.10).

We find that the amplitude of the three-dimensional correlation function of the bright galaxies is larger by a factor of ~ 2 than the amplitude of correlation function of the faint galaxies. Figure (5.13) shows the values of bright and the faint ones in comparison with the amplitudes of the whole sample. Additionally plotted are the amplitudes of the early type and late type galaxies, as in Figure (5.9).

The amplitude of the bright galaxies is a bit larger than the total amplitude at that redshift, but not as large as the amplitude of the early type galaxies. The faint galaxies are much lower clustered, their amplitude can be found in the same region as the amplitudes of the late type galaxies. Obviously the difference in the clustering strength for subsamples of different Hubble types is larger than for subsample with different rest-frame B band luminosities.

This result corroborates the hypothesis that the differences in the evolution of clustering of galaxies are not only determined by the locus of their formation (large, bright galaxies form in the rare high density peaks and are thus more strongly clustered), but also on their evolution.

The faint sample is more dominated by late type galaxies than the bright sample, the ratio of late type ($SED \geq 60$) to early type ($SED < 60$) galaxies is ≈ 1.7, whereas for the bright sample it is ≈ 1.3. This might be the reason why the amplitude of the faint sample has more or less the same value than the starburst galaxies, and the amplitude of the bright sample, which contains bright, younger galaxies as well as the older population, is only slightly higher than the total amplitude.

If faint (late type) galaxies are less strongly clustered than the bright ones at all redshifts, then if we would see the same population mix of galaxies also in the highest redshift bin, the clustering evolution would be even steeper than we have measured. However, the faint galaxies are not seen at higher redshifts (see Figure (5.10)), and the correlation function in the highest redshift bin between $z = 0.75$ and 1.07 is completely dominated by the bright population.

Also the comoving number density of starburst galaxies, which are always less clustered than the rest, increases with increasing redshifts, and the space density of the highly clustered very early type (E-Sa) galaxies decreases by a factor of ~ 1.6 from $z = 0$ to $z = 1$ (Fried et al., 2001).

Therefore we expect the measured growth for a deeper sample (deeper than $I \leq 23$) to be even stronger.

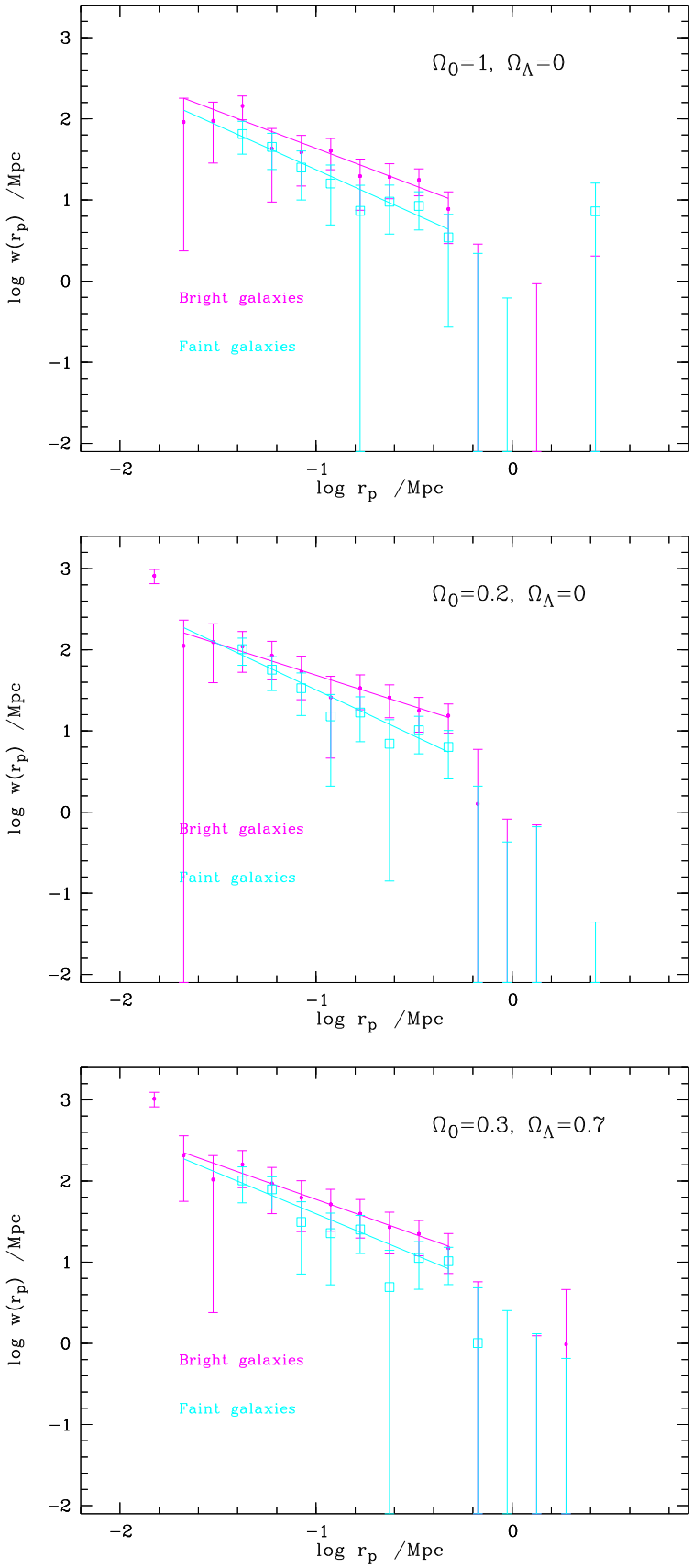

Figure 5.12: The projected correlation function of bright and faint galaxies between $0.3 \leq z \leq 0.6$, for the three different cosmologies.

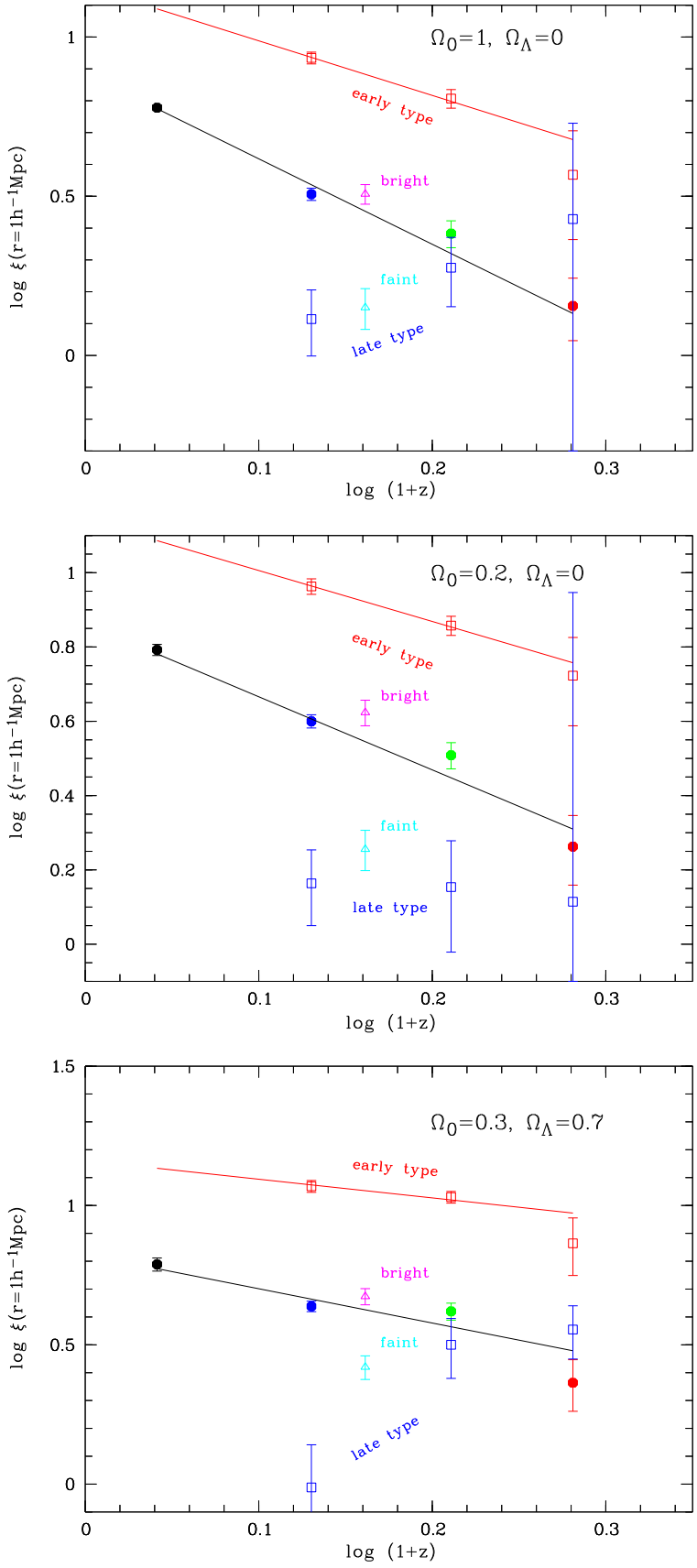

Figure 5.13: The amplitudes of the correlation function at $r_{com} = 1h^{-1}$ Mpc of bright and faint galaxies, in comparison with the whole sample, and the early and late type galaxies.

5.2.4 The cross-correlation between the faint and the bright galaxies

We also computed the correlation between bright and faint galaxies, to check whether the correlation signal on the very smallest scales is dominated by the distances *between* bright primaries and their faint satellites. We calculated this cross-correlation in the same redshift bin as we did for the correlation of bright and faint galaxies ($0.3 \leq z \leq 0.6$), by comparing the distances between the galaxies in the two samples ($\langle D_f D_b \rangle$), with the distances between two randomly distributed samples of galaxies ($\langle R_f R_b \rangle$) which follow the smoothed redshift distribution of bright and faint galaxies, respectively (see equation (5.15), and (5.16)). The estimator for the correlation function between faint and bright galaxies is

$$w_{esti} = \frac{\langle D_f D_b \rangle}{\langle R_f R_b \rangle} - 1 \ . \tag{5.17}$$

The angular correlation between the bright and the faint sample is shown in Figure (5.14), in comparison with the data of all galaxies, and the correlation found in the bright and the faint sample, respectively. There are no evident features which might serve as a diagnostics for the contribution of satellite galaxies to the faint galaxy sample.

Despite its smaller number statistics, the projected correlation function can give a more detailed insight, because projected distances are only counted if the galaxies are separated by a distance $\pm\delta z$, thus the contribution of randomly projected, intrinsically uncorrelated pairs is suppressed.

For the calculation of the effective distance in $w(r_p)$, we took the the redshift distribution of all galaxies, which represents the selection function for the whole sample.

Figure (5.15) shows the projected correlation function between the bright and the faint sample, in comparison with the correlation function of the whole sample, and the correlation among themselves (as shown in the last section).

The projected correlation function of the bright and the faint galaxies contains some information: it shows a bump at the small scale end, and this may be used as a diagnostic for the contribution of satellite galaxies to the faint sample. If we could extend our analysis to higher redshifts (which will be possible in future surveys), we could give an estimate of the change of the percentage of satellite galaxies with redshift, and thus place constraints of the merger rate (at least for minor mergers).

In contrast to the angular correlation function, the projected correlation function can be used for the investigation of the evolution of the comoving number of satellite galaxies.

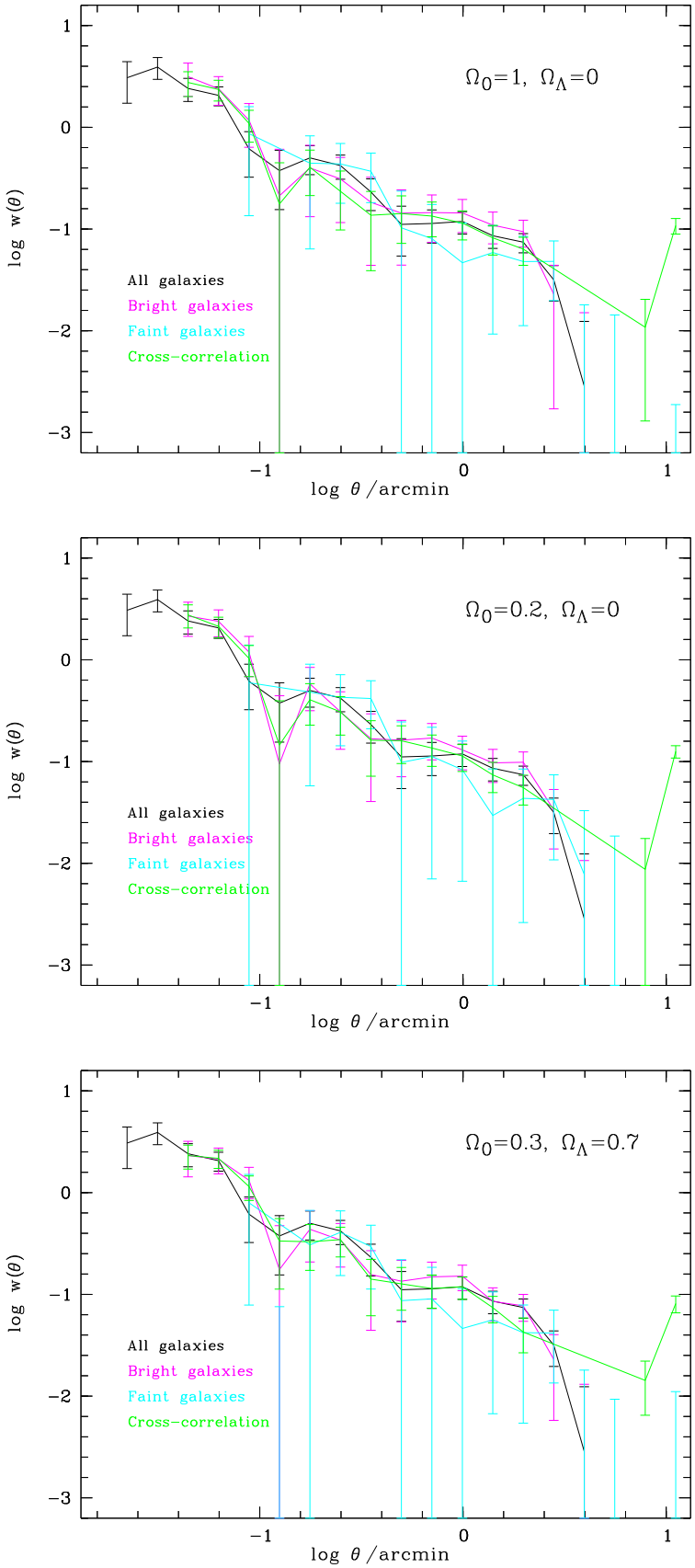

Figure 5.14: The angular cross-correlation between the bright and the faint sample (green) in the range $0.3 \leq z \leq 0.6$, in comparison with the data of all galaxies (black), and the correlation found in the bright (magenta) and the faint (light blue) sample, respectively.

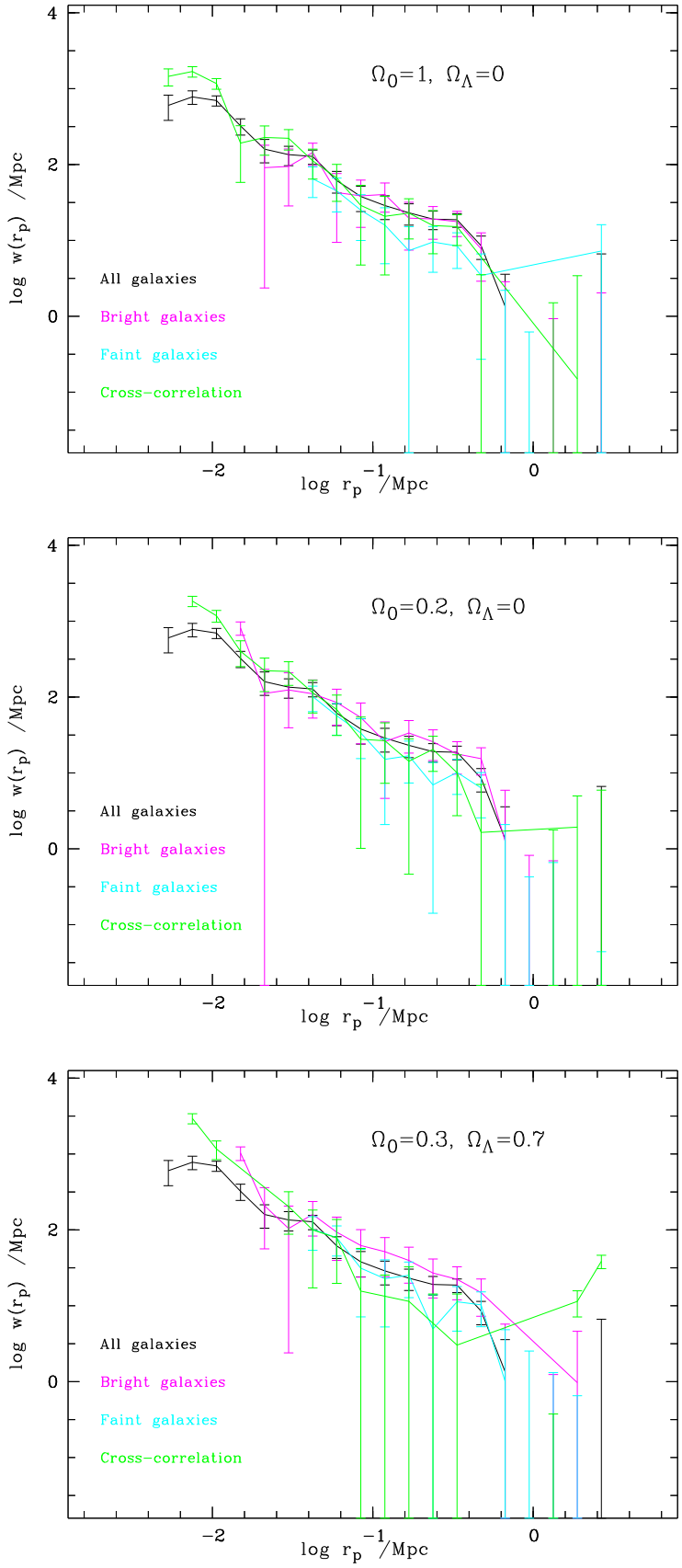

Figure 5.15: The projected correlation between the bright and the faint sample (green) in the range $0.2 \leq z < 0.5$, in comparison with the data of all galaxies (black), and the correlation found in the bright (magenta) and the faint (lightblue) sample, respectively.

5.3 Summary

We have carried out an analysis of the evolution of large scale clustering of field galaxies between a redshift $z \sim 1.1$ and the present epoch, using multicolor data from the **C**alar **A**lto **D**eep **I**maging **S**urvey **CADIS**. Since the redshift errors are too large to calculate the three-dimensional correlation function directly, we used both angular and projected correlation function to infer the amplitudes of $\xi(r)$ at different redshifts.

To facilitate a direct comparison with the LCRS (as a "local" sample), redshift errors of the size of the CADIS errors had to be added to the LCRS, and the projected correlation function had to be calculated with the modified redshifts.

We calculated the angular correlation function in four redshift intervals with $\Delta z = 0.2$. From comparison with model amplitudes calculated for different evolution scenarios, we deduced the evolution parameter ϵ for different world models (a flat, high-density model, an open model with $\Omega_0 = 0.2$, and a flat model with non-zero cosmological constant ($\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$). For the same cosmologies, we calculated the projected correlation function, in three redshift bins centered at $\langle z \rangle = 0.35$, $\langle z \rangle = 0.625$, and $\langle z \rangle = 0.91$. From the amplitude of the projected correlation function in the different redshift bins we calculated the amplitude of the three-dimensional correlation function at a comoving separation of $r_{com} = 1h^{-1}\,\mathrm{Mpc}$, and introduced a new parameter q, which describes the deviation of the clustering growth from the global Hubble flow. q can be estimated by fitting a straight line into the $\xi(r_{com} = 1h^{-1}\,\mathrm{Mpc})$ versus $\log(1+z)$ plot, q is then simply the slope. It can be related to ϵ: $\epsilon = -(q+3-\gamma)$. Both methods, deprojection of the angular and the projected correlation function are equivalent and lead essentially to the same results. The estimation of the new parameter q is a bit more straightforward, and its meaning is possibly more comprehensible.

Depending slightly on the cosmology adopted for the estimation, we find a more or less rapid growth of the clustering strength between a redshift of $z \sim 1.1$ and today. For the closed world model we find $\epsilon = 1.590 \pm 0.132$, $q = -2.68 \pm 0.16$. The values found in the open case are $\epsilon = 1.107 \pm 0.134$, $q = -1.92 \pm 0.17$, and for the flat model with non-zero cosmological constant we found $\epsilon = 0.209 \pm 0.133$, $q = -1.23 \pm 0.20$. The cosmology should not play *such* an important role between a redshift of $z = 1$ and today (see Table (5.1)). Thus we conclude that the apparent dependence on the world model is mainly due to low statistics, and with only four fields of $1/30\square^\circ$ each we can not make a definite quantitative statement about the clustering strength. However, all values of ϵ are positive, and $q \approx -1.9$, which is consistent with linear clustering.

We also calculated the angular and the projected correlation functions for late and early type galaxies. The different subsamples of early and late type galaxies exhibit a different clustering evolution: The early type galaxies are much stronger clustered than the late type ones, and the amplitude of their correlation function is also much bigger than the amplitude of the complete sample. This is the case at *all* redshifts, but the difference is larger at higher redshifts; the values seem to converge to a common r_0. The evolution of the clustering of the late type galaxies could not be quantified, because the errors of the deduced amplitudes of the three-dimensional correlation function at $r_{com} = 1h^{-1}$ Mpc are too large to enable a reliable fit. The amplitude of the correlation function of early type galaxies changes much

more slowly with redshift, we found $q = -1.71 \pm 0.39$ for $\Omega_0 = 1.0$, $\Omega_\Lambda = 0.0$, $q = -1.37 \pm 0.36$ for $\Omega_0 = 0.2$, $\Omega_\Lambda = 0.0$, and $q = -0.67 \pm 0.33$ for $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$.

Combining redshift information and SED, it is possible to calculate rest-frame B band luminosity, so we could further investigate the different clustering properties of bright and faint galaxies. With reasonable number statistics in the faint sample, this was only possible in the redshift range $0.3 \leq z \leq 0.6$. The clustering amplitude of the bright galaxies is larger than for the whole sample, and much larger than the amplitude of the faint galaxies, but not as strong as for the early type ones.

Chapter 6

Discussion and conclusions

In this chapter we will first compare our results with determinations of other authors from both the observational and theoretical point of view and then end with our conclusions.

6.1 Comparison with other observations

In the literature there are essentially only two investigations of the evolution of galaxy clustering, the results of which can be compared with this work: one analysis by Le Fevre et al. (1996) which has been carried out in the framework of the **C**anada **F**rance **R**edshift **S**urvey (in the following CFRS), and one by Carlberg et al. (2000), done on the CNOC sample (**C**anadian **N**etwork for **O**bservational **C**osmology). Furthermore we applied our methods to the COMBO 17 survey (Wolf et al., 2001a), and compared the results.

6.1.1 The CFRS determination

Le Fevre et al. (1996) used the projected correlation function to investigate the spatial clustering of 591 galaxies between $0.2 \leq z \lesssim 1.1$, in five CFRS fields (for a description of the survey see Lilly et al. (1995) and Schade et al. (1995), respectively) of approximately the same size as our CADIS fields. The objects are primarily located in three parallel strips for each of the five fields, within which almost 100% spectroscopic sampling was obtained, separated by regions where few spectroscopic observations were carried out. The galaxies have spectroscopic redshift determinations, and $I \leq 23^{mag}$. They computed the projected correlation function in three redshift bins between $0.2 \leq z \leq 0.5$, $0.5 \leq z \leq 0.75$, and $0.75 \leq z \leq 1.0$, with integration limits of $\delta_z = \pm 0.0075$. For the connection to $z = 0$ they took values of $r_0(z = 0)$ from Loveday et al. (1995) and Hudon & Lilly (1996).

Figure (6.1) shows the amplitude of the three dimensional correlation function at $r_{com} = 1h^{-1}$ Mpc, deduced from the projected correlation function, in comparison with our own data. For the direct comparison we have to multiply our measured amplitudes of the projected correlation function by 1.4 to correct for the influence of large redshift errors (see Chapter 4).

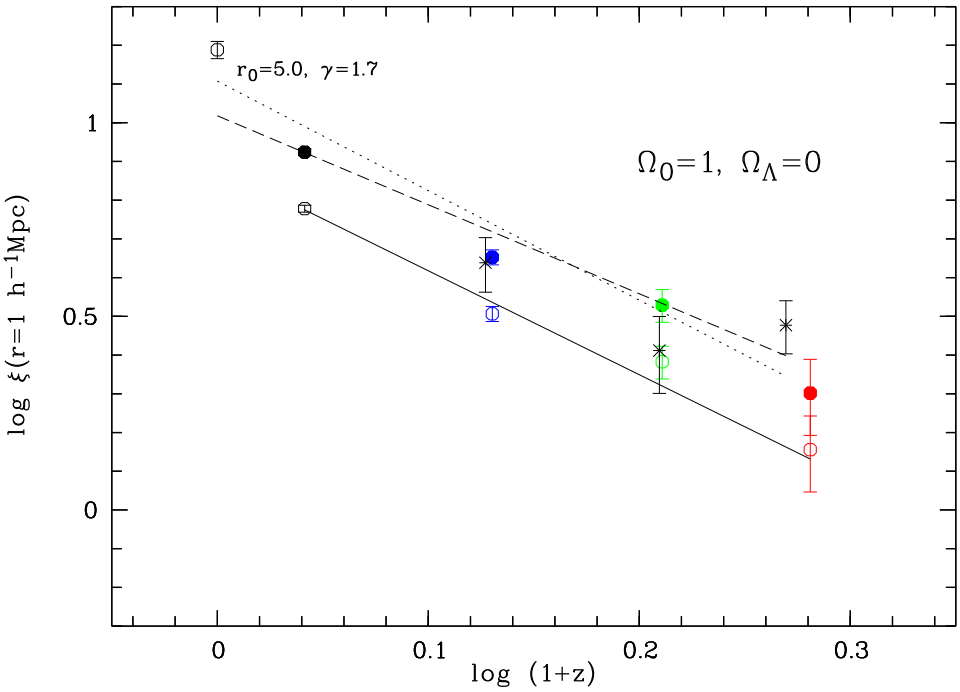

Figure 6.1: The amplitudes of the three dimensional correlation function at $r_{com} = 1h^{-1}$ Mpc, deduced from the projected correlation function of the CFRS data (crosses), in comparison with our own data (open symbols). The filled symbols are our data points corrected for the influence of the redshift errors on the amplitude of the projected correlation function. The dotted line is the fit of the CFRS data points including the value of $r_0(z = 0)$ from Loveday et al. (1995), the dashed line is the fit using the corrected LCRS point instead.

With this correction, the CFRS data points are consistent with our own measurement, although with large errors. Le Fevre et al. (1996) claim that if $r_0(z = 0) = 5h^{-1}$ Mpc, $0 < \epsilon \lesssim 2$. The fit of their data points, including the connection to $z = 0$, yields $q = -3.043 \pm 0.213$ ($\cong \epsilon = 1.8$). If the connection to $z = 0$ is disregarded, we find $q = -1.184 \pm 0.634$. The fit including our redshift error corrected LCRS point instead of the Loveday et al. (1995) point yields $q = -2.298 \pm 0.238$. This is even a bit less than our own measurement ($q = -2.68 \pm 0.16$), but nevertheless equal within the errors. This exercise shows that the value of q depends on the connection to the present epoch. First of all, as we have seen, the measured amplitude of the correlation function depends strongly on the Hubble type, so the adopted value of $r_0 = 5h^{-1}$ Mpc might not fit to the population mix of the CFRS sample and thus simply be too large. Second, Le Fevre et al. (1996) did not take the influence of the redshift errors into account. Although they are relatively small, it is probably not admissible to directly compare the measured amplitudes with a value which was inferred from a direct estimation of the three-dimensional correlation function.

The conclusion is that their *measurement* is consistent with ours, but the evolution is not as fast as they claim, because they did not treat the local measurement selfconsistently.

6.1.2 The CNOC determination

The survey itself is described in detail in Yee et al. (1996). The analysis of the clustering evolution (Carlberg et al., 2000) was carried out on a sample of 2300 bright galaxies (k-corrected and evolution-compensated R luminosities $M_R^{k,e}$ brighter than -20^{mag} ($H_0 = 100\,\mathrm{km\ s^{-1}}$)). The redshift distribution extends to $z = 0.65$, and for comparison with $z = 0$, they selected a comparable sample from the LCRS.

From the calculation of the projected correlation function in seven redshift bins and in the LCRS sample, they deduced the evolution parameter ϵ, and found a strong dependence on the cosmology adopted for the computation: $\epsilon = 0.80 \pm 0.22$ for $\Omega_0 = 1$, $\Omega_\Lambda = 0$, $\epsilon = -0.17 \pm 0.18$ for $\Omega_0 = 0.2$, $\Omega_\Lambda = 0$, and $\epsilon = -0.81 \pm 0.19$ for $\Omega_0 = 0.2$, $\Omega_\Lambda = 0.8$. The trend seen here is the same as in our case – strongest evolution for the high density model, slower evolution for the low-density cosmology, and the least evolution for the flat model with non-zero cosmological constant. In our analysis the differences in the value of ϵ are not so strong, we always find a positive value of ϵ, in all three cosmologies.

From a comparison of their results for different cosmologies with different theoretical models for clustering evolution (linear clustering: Peebles (1980) and Efstathiou et al. (1985), biasing models: Mo & White (1996) and Jing (1998), N-body experiments: Colin et al. (1997) and Kravtsov & Klypin (1999)) they conclude that the open and $\Omega_0 = 1$ models are consistent with linear growth, but the flat Λ model is marginally excluded. Biasing is marginally excluded in both the open and the high density cosmology, the low-density flat model is acceptable under all biasing models.

Between a redshift of $z = 0.65$ and today the cosmology adopted for the calculation is even less important than in our case (see Chapter 5), and therefore we conclude that their statistics was far too low to permit these detailed conclusions.

6.1.3 COMBO 17

All the methods developed in this thesis, especially the calculation of the projected correlation function, are universally valid and not restricted to a certain survey geometry or redshift accuracy. Larger, wide angle deep surveys have only recently become available, and one of them is the **COMBO 17** survey [1] (Wolf et al., 2001a), in some respect the successor of CADIS. All observations have been carried out on La Silla, Chile, with the WFI (Wild Field Imager) at the ESO-2.2 m telescope. Each of the four fields has a size of $1/4\,\square^\circ$. 17 filters have been observed, which facilitates a secure multicolor classification and redshift determination ($\delta z 0.017$) down to $R = 24$. The complete catalogue will include ~ 40000 galaxies with $I \leq 23$, in $1\ \square^\circ$, with SED and morphological information.

We already took a first glance on the data available so far, namely on one field centered at $\alpha_{2000} = 3^h32^m25^s$, $\delta_{2000} = -27^\circ48'50''$. The catalogue includes 6602 galaxies between $0.2 \leq z \leq 1.07$ and $I \leq 23$, with the same redshift accuracy as the CADIS galaxies. We masked out bright stars, calculated the projected correlation function in the same redshift intervals as for the CADIS data, and fitted the amplitudes in the same way as before.

[1] COMBO = **C**lassifying **O**bjects by **M**edium-**B**and **O**bservations in **17** filters

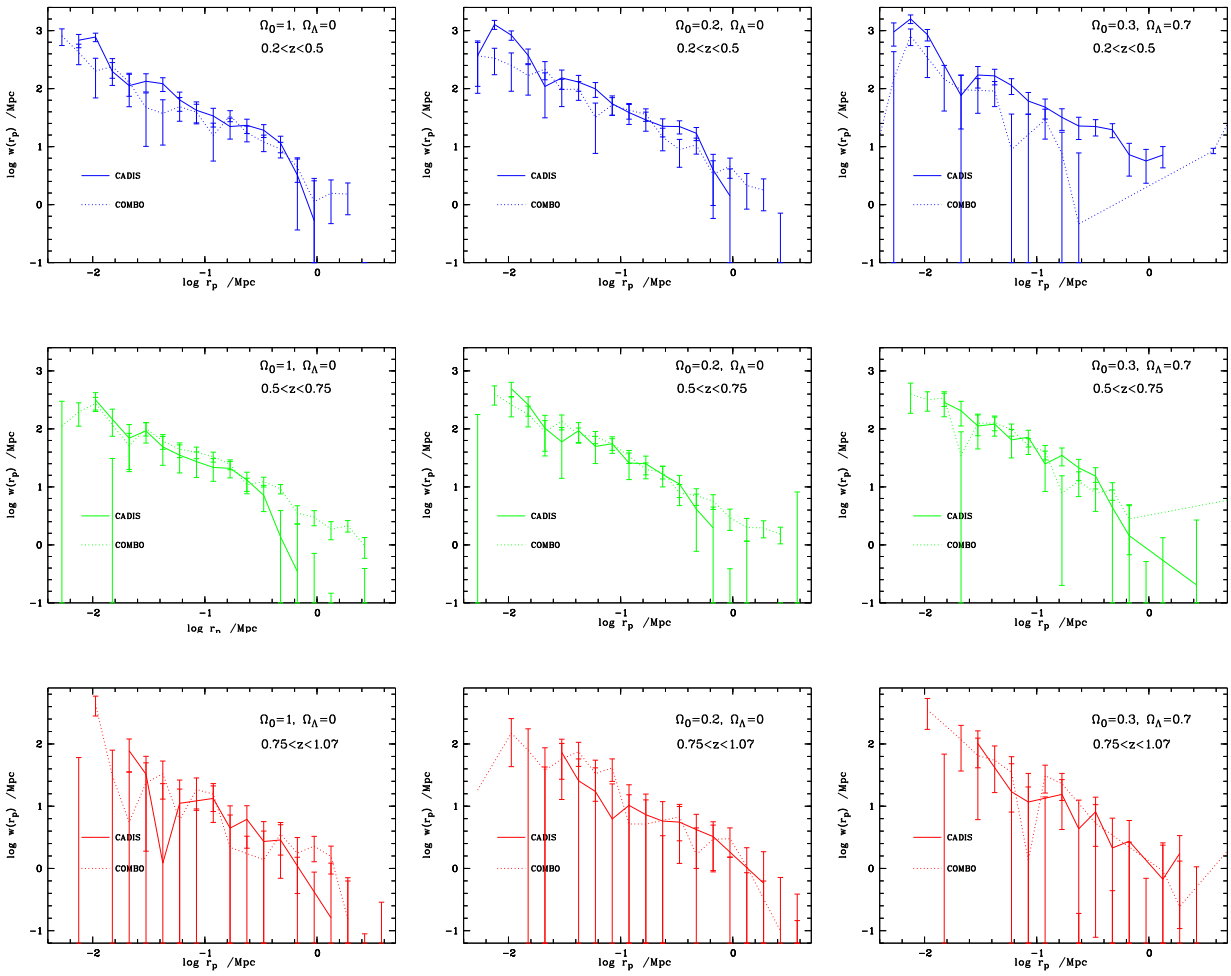

Figure 6.2: The projected correlation function of the COMBO data (dotted lines) in comparison with CADIS (solid lines). Upper panel: $0.2 \leq z < 0.5$, middle: $0.5 \leq z < 0.75$, lower panel: $0.75 \leq z \leq 1.07$.

A comparison of the projected correlation function of the COMBO 17 and the CADIS data is shown in Figure (6.2). For each cosmology and each redshift interval the projected correlation function is shown separately.

From the projected correlation function we calculated the amplitude of the three-dimensional correlation function at $r_{com} = 1\,\mathrm{h}^{-1}$ Mpc. The comparison with the CADIS data is shown in Figure (6.3).

As one can see, the results are only in few cases fully consistent with the CADIS determinations. For $\Omega_0 = 1$, $\Omega_\Lambda = 0$ and $\Omega_0 = 0.2$, $\Omega_\Lambda = 0$, the values for $\langle z \rangle = 0.91$ are equal within their errors, for the open cosmology the data at $\langle z \rangle = 0.625$ are almost identical. The amplitudes at the lowest redshift are different in all cases. It seems as if field-to-field variations still play a role, even if the number of galaxies per field is large.

From the fit of the data we find $q = -1.87 \pm 0.13$ for $\Omega_0 = 1$, $\Omega_\Lambda = 0$, $q = -1.94 \pm 0.12$ for $\Omega_0 = 0.2$, $\Omega_\Lambda = 0$, and $q = -1.54 \pm 0.20$ for $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$.

In the case of the closed model the parameter is significantly less negative than the one deduced from the CADIS data ($q_{CADIS} = -2.68 \pm 0.16$), whereas in the open case the value is almost exactly the same as for CADIS ($q_{CADIS} = -1.92 \pm 0.17$). In the flat model with non-zero cosmological constant, the values are equal within the errors ($q_{CADIS} = -1.23 \pm 0.20$). The differences between the cosmologies are not as significant as in the analysis of the CADIS data.

This result corroborates our claim that field-to-field variations and low statistic can feign an apparent dependency on the cosmology adopted for the calculation. We expect that for more fields and more galaxies q will settle to a robust value of $\simeq 2$ with the expected spread of about 20% due to the difference in angular distance in the different cosmological models.

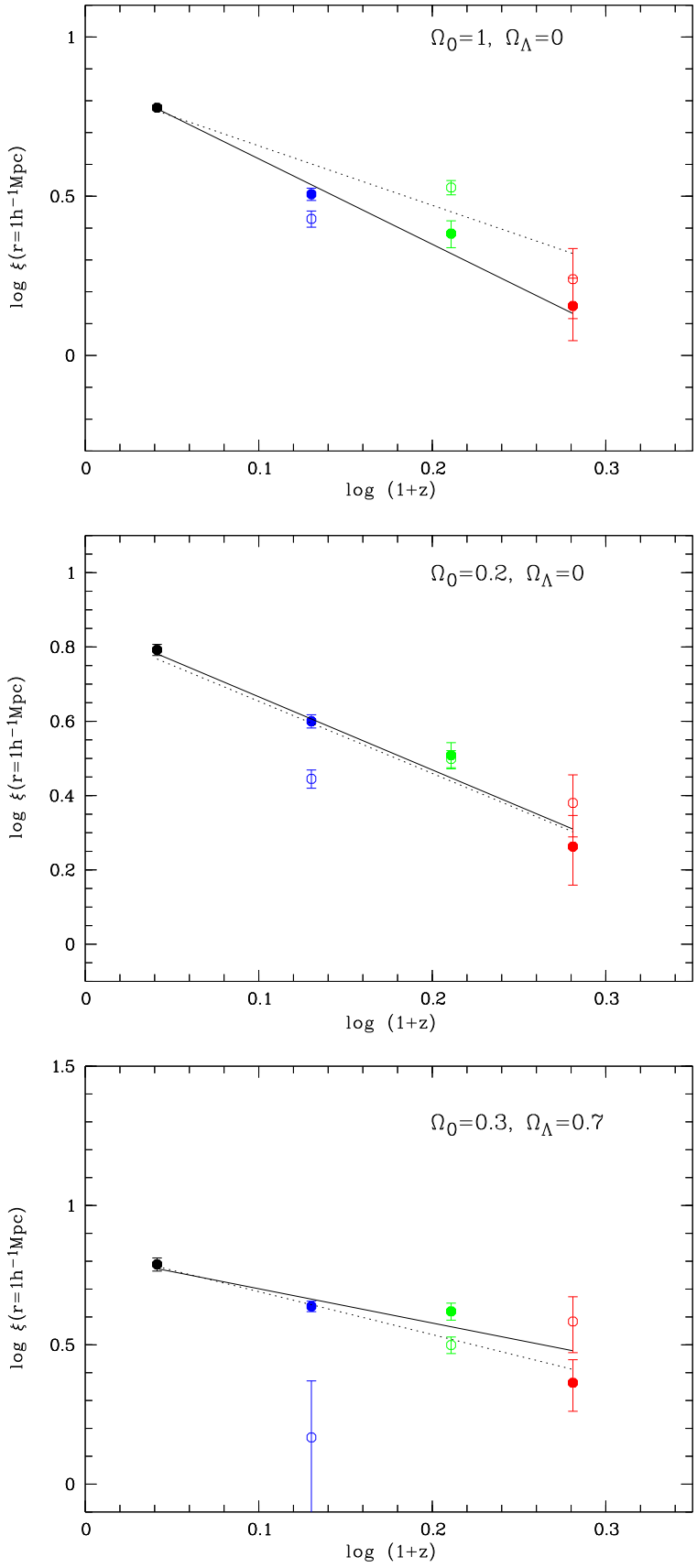

Figure 6.3: The amplitude of the three-dimensional correlation function at $r_{com} = 1\,\mathrm{h}^{-1}$ Mpc. Filled points are CADIS data, open circles are COMBO data. The dotted line is the fit for the COMBO data.

6.2 Comparison with therory

The evolution of the matter correlation function ξ_m with redshift has been studied extensively using both N-body simulations (Jenkins et al., 1998) and analytic models (Hamilton et al., 1991; Peacock & Dodds, 1994; Jain et al., 1995). Those analyses show, how an initial density field evolves in different cosmologies. The evolution of the galaxy clustering, however, need not necessarily follow that of the collisionless component of the mass density field. Galaxies have been subject to external phenomena such as tidal interactions, satellite accretion, merger, etc., as well as to internal phenomena such as galactic winds and supernovae. How galaxies form and evolve in the underlying mass field can partly be investigated by combining large high-resolution N-body simulations with semi-analytic models of galaxy formation (Kauffmann et al., 1999b,a; Somerville et al., 2001).

6.2.1 Evolution of dark matter in different cosmologies

Since the pioneering work of Davis et al. (1985) larger and larger N-body experiments have been carried out in order to investigate the evolution of the dark matter. The development of powerful supercomputers and sophisticated numerical methods has made extremely high-resolution simulations possible. The best ones available at the moment are the ones carried out by the Virgo consortium (Pearce et al., 1999). The evolution of structure was analysed for four different cosmologies in a simulation with $17 \cdot 10^6$ particles, running from $z = 50$ to $z = 0$. The world models include a flat high density model with $\Omega_0 = 1$, $\Omega_\Lambda = 0$ (SCDM), an open model with $\Omega_0 = 0.3$, $\Omega_\Lambda = 0$ (OCDM), a flat low-density model with non-zero cosmological constant (ΛCDM, $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$). Note that the parameters of these models are almost identical to those we have used for analysing the correlation functions in our data. Additionally a flat $\Omega_0 = 1$ τCDM model has been simulated, in which a massive neutrino (the τ neutrino) was present during the very early evolution of the universe and came to dominate the energy density for a short period. It then decayed into lighter neutrinos which are still relativistic, thus delaying the epoch of matter-radiation equality – it is obvious that structure starts to form later in this model. Figure (6.4) shows three "snapshots" of the universe in the four models, at $z = 3$, $z = 1$ and $z = 0$. The boxsize is $239.5\,h^{-1}$ Mpc. For the ΛCDM and the OCDM $H_0 = 70\,\mathrm{km\ s^{-1}/Mpc}$ and $H_0 = 50\,\mathrm{km\ s^{-1}/Mpc}$ for the SCDM and the τCDM models. The normalisation of the primordial power spectrum was chosen such that at $z = 0$ the models match the observed abundance of galaxy clusters.

In the two low-density models structure forms early, whereas in the τCDM cosmology and above all in the SCDM model, structure forms much later. At the earliest epoch shown in Figure (6.4), $z = 3$, the SCDM model is very smooth, with only little fine structure. The τCDM model has some embryonic large-scale structure but is even more featureless than the SCDM on the finest scales. By contrast, structure in the low-density models, particularly the OCDM is already well developed by $z = 3$.

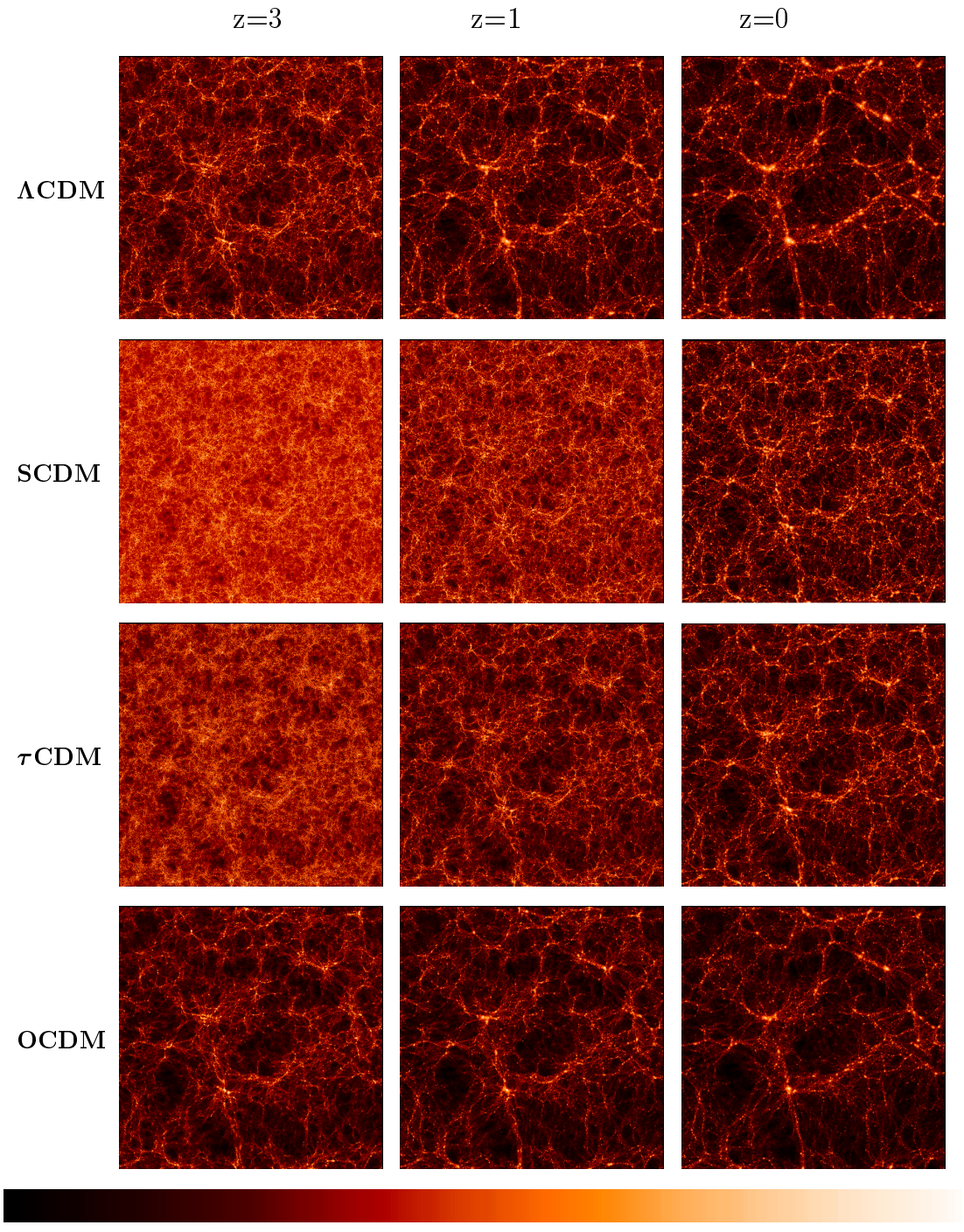

Figure 6.4: The evolution of dark matter in four different cosmologies.

The large-scale differences among the models are also apparent at $z = 1$. There is substantially more evolution for high-density than for low-density models. OCDM has the most developed large-scale structure at $z = 1$, while ΛCDM is intermediate between this and the two high-density models.

At $z = 0$, the general appearance of all the models is similar because, by construction, the phases of the initial fluctuations are the same. But non of the models is able to exactly reproduce the observed two-point correlation function as measured in the APM survey (Baugh, 1996). Thus, for any of these models to provide an acceptable representation of the present reality, the distribution of galaxies would need to be biased relative to the mass in a non-trivial, scale-dependent fashion.

This analysis does not only make clear that the dark matter density field is clustered in a different way than the galaxies, but also that structure developes very differently in different cosmologies. Colin et al. (1997) give values for the evolution parameter ϵ for dark matter haloes followed in an N-body simulation of ~ 2 million particles, between $z = 1.3$ and $z = 0$. In their simulation, $\epsilon = 1.5$ for $\Omega_0 = 1$, $\Omega_\Lambda = 0$ (corresponding to the evolution parameter $q = -2.7$ if $\gamma = 1.8$); $\epsilon = -0.3$, $q = -0.9$ for $\Omega_0 = 0.2$, $\Omega_\Lambda = 0$; $\epsilon = -0.6$, $q = -0.6$ for $\Omega_0 = 0.2$, $\Omega_\Lambda = 0.8$.

The degree to which those results constrain the mean density of the universe depends on how well the evolution of galaxy clustering is traced by the evolution of the mass density field or the halo population, respectively. Since we have seen in our analysis that galaxies do not trace matter, we can *not* conclude that our result, a more or less rapid growth of the structure, is a hint at a high-density universe. More detailed studies of the clustering evolution of galaxies of different Hubble types, and the comparison with *semi-analytic models* of galaxy evolution are required before we can understand biasing, and before structure evolution can result in statements about the mass density of the universe.

6.2.2 Biased galaxy formation

The evolution of the clustering of *dark matter* is a monotonous process, which depends on cosmological quantities such as Ω_0, Ω_Λ, and the initial power spectrum, which determine how collapsed structures – dark matter haloes – form and evolve. In those hierarchical cosmologies the evolution of *galaxy clustering* depends also on the physical processes – cooling, star formation, radiative and hydrodynamic feedback – which drive the formation of galaxies within these merging haloes. The relation of the clustering of the dark matter and of the galaxies is described by the biasing parameter b, which can be different for different types of galaxies.

Kaiser (1984) showed that, inherent in the notion of the power-spectrum of the perturbations (for the interrelation between correlation function and power-spectrum see appendix (A)) is the fact that the perturbations have a Gaussian distribution of amplitudes about the root mean squared value $\bar{\Delta}$ with variance $\bar{\Delta}^2$, so that the probability of encountering a density contrast Δ at some point in space is proportional to $\exp(-\Delta^2/\bar{\Delta})$. Galaxies form in the rare, highest density peaks of the dark matter distribution, when gas cools, and condenses to form stars. If we require the density perturbation to exceed some value Δ_{krit} in order that structures form, galaxy formation would be biased to the highest peaks over the mean background

density.

Consider a large-scale (very much greater than galaxy-scale), positive density enhancement in a given region of space, and that superimposed on this density enhancement are numerous bumps and wiggles of galaxy scale (see Figure (6.5)). The very first galaxies to collapse will be those at the very highest peaks in the density field. These objects correspond to the small-scale bumps and wiggles that reside along the highest "ridge" of the density enhancement. Therefore, one might expect that the earliest galaxy-sized objects to form would do so in a strongly clustered state [2] (Brainerd & Villumsen, 1994). An observation which fits into this picture is the strong clustering of Lyman break galaxies at a redshift of $z \sim 3$ (Steidel et al., 1998; Giavalisco et al., 1998). The Lyman break galaxies are believed to be the predecessors of today's very massive elliptical galaxies.

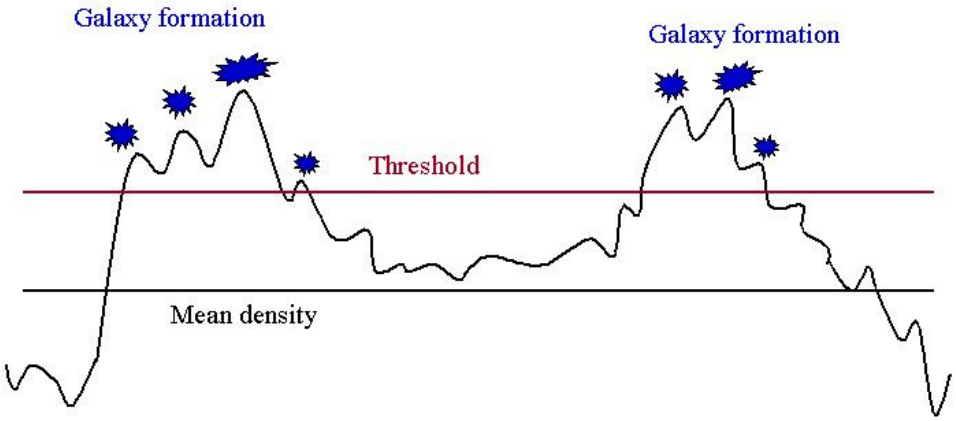

Figure 6.5: The first galaxies form in the bumps and wiggles superimposed on rare, large density enhancements and are therefore highly clustered.

As time goes by, smaller and smaller peaks that are farther and farther out in the wings of the large-scale density enhancement will collapse to form galaxies, resulting in a galaxy distribution that is less clustered than the distribution of the first generation of galaxies.

[2] The reason why the highest mountains in the world are in the Himalayas is because they are superimposed on the large-scale plateau, or long wavelength perturbation, caused by the plate supporting the Indian subcontinent crashing into the Asian plate.

6.2.3 Semi-analytic models of biased galaxy formation

A new technique for following the formation and evolution of *galaxies* in cosmological N-body simulations was introduced by Kauffmann et al. (1999b). Dissipationless simulation are used to track the formation and merging of dark matter haloes as a function of redshift. Simple prescriptions, taken directly from semi-analytic models of galaxy formation, are adopted for gas cooling, star formation, supernova feedback and the merging of galaxies within the haloes.

In this kind of simulations, the dependence of the measured evolution of the correlation function on sample selection can be investigated. Kauffmann et al. (1999a) showed, that a 'dip' in the amplitude of the correlation function between $z = 1$ and $z = 0$ can be a diagnostic for the process of galaxy formation. Such a dip occurs in low-density models, because structure forms early, and galaxies of $\sim L_*$ are unbiased tracers of the dark matter over this redshift range; their clustering amplitude then evolves similar to that of the dark matter. At higher redshifts, bright galaxies are strongly biased, and the amplitude increases again. In high-density models structure forms late, and bias evolves much more rapidly. As a result, the clustering amplitude of L_* galaxies remains constant from $z = 1$ to $z = 0$ (the 'dip' occurs earlier). The strength of this effect is sensitive to sample selection. The dip becomes weaker for galaxies with lower star formation rates, redder colors, higher luminosities and earlier morphological types, for which the comoving clustering strength does not change at all.

The simulation shows, that for a given cosmology, the clustering amplitude predicted for a sample of galaxies depends on the masses of the dark matter haloes they inhabit. The evolution of clustering depends on how the mass distribution of these haloes changes with redshift, and on the variation of the abundance of galaxies in the sample.

A sample of galaxies with a fixed star formation rate is expected to show a stronger dip (and connected with this, a steeper rise of the amplitude of the correlation function with redshift) than a sample of galaxies that tracked haloes of the same mass at all redshifts: galaxies with fixed star formation rate are found in smaller haloes at high redshift than at the present time, and also the abundance of starforming galaxies is expected to increase more strongly with redshift, because there are many more small haloes than large ones. Early type galaxies are found primarily in massive haloes at all redshifts. These galaxies do not exhibit any dip in clustering, and their abundance decreases strongly with redhsift, because massive haloes are rare objects at early times.

The results of this computer experiment fit very well to our own findings – a relatively strong growth of the clustering strength between $z = 1$ and today found for all the galaxies included in the CADIS sample, while the galaxies with early type SED (i.e. low star formation rate) show a much slower increase of the amplitude of the correlation function.

6.3 Conclusion

In this work, we have seen that older galaxies show a much slower evolution of their clustering than younger or even starburst galaxies. Light does *not* simply trace matter: the simple assumption of a direct correlation between galaxies and the dark matter distribution fails to explain why galaxies of different Hubble types show a different evolution of their clustering

with redshift. The rate of clustering growth one measures is dependent on sample selection.

A possible explanation can be found in the context of *biased galaxy formation*. The first galaxies are born in a highly clustered state, because they form in the bumps and wiggles which are superimposed on the very large-scale density enhancements.

In the most simple model of bias evolution of early type galaxies, the *galaxy conservation scenario* (Fry, 1996; Tegmark & Peebles, 1998), the galaxy population is conserved over cosmic time (i.e. no new elliptical forms and no one disappears). This scenario implies the assumption that all the early type galaxies (mainly ellipticals) were formed at high redshift and simply follow the growth of the perturbations without the additional non-linear effects such as virial collapse and merging. In this case the positive evolution of the bias which increases with redshift is more than compensated by the decline of linear growth, and the amplitude of the correlation function is expected to rise towards $z = 0$.

Obviously this model is an oversimplification. Galaxy formation is not supposed to be a burst-like event, but a continuous process with a spread in redshift from $z \gtrsim 3$ to $z \approx 1$. It is already well established that disk galaxies evolve to the present day: Late type disk galaxies have present-day star formation rates comparable to those at earlier cosmic times, early type disks have formed most stars at earlier times (Kennicutt et al., 1994). Galaxies may disappear from one sample and show up in another, when the stars age and the SED changes. Fried et al. (2001) found the evolution of the B band luminosity function of the CADIS galaxies to be clearly differential, the normalisation Φ^* of the early type E-Sa galaxy luminosity function and the integrated comoving space density decreases with increasing redshift. The normalisation for the Sa-Sc galaxy luminosity function increases with redshift as well as the space density. The luminosity function of the starburst galaxies steepens towards the faint end, and their comoving space density increases with redshift. The density evolution of the early and the late type galaxy population apparent in our data is suggestive of merging.

If we assume a substantial evolution of the individual galaxies (including merging) between a redshift the growth of the clustering strength well be slower: The next generations of galaxies form later in the wings of the large-scale enhancements, and are therefore less and less clustered. While the universe expands, the galaxies evolve, age, and eventually merge to form larger, brighter galaxies and ellipticals, and generally add to the population of earlier type galaxies, while new galaxies form at later times in less and less clustered environments.

The oldest galaxies have formed in a much more clustered state than their successors, and at lower redshifts, the population of old galaxies consists of galaxies of different ages, which have added to the "old galaxy popultion" at different times, and in more and more lower clustered states. Elliptical galaxies might not only form by monolithic collapse and passive evolution, but also by merging (Toomre & Toomre, 1972; Naab et al., 1999). Merging amplifies the effect, because first of all galaxies are formed, which "suddenly" add to the old population. Second, a merger event reduces not only the number of galaxies, but also the number of small pair separations in a sample, which reduces the probability of finding pairs of galaxies at small distances – and thus supresses the amplitude correlation function. Although the clustering strength of the underlying dark matter density field increases with redshift, the biasing decreases. The net effect is a very slowly rising clustering amplitude.

In our case, we see some amount of evolution in the early type sample ($0 \leq SED \leq 60$), but this sample also includes galaxies of later types than Sa ($SED = 30$). The comoving number density of certainly more weakly clustered Sa to Sc galaxies increases towards higher redshifts (Fried et al., 2001), so in our highest redshift bin the clustering signal is probably dominated by a more weakly clustered population of galaxies. If we could calculate the correlation function for a sample excluding galaxies of earlier type than Sa, we would expect the evolution parameter q to be even closer to zero than for our present "early type" sample which includes galaxies with $0 \leq SED < 60$.

The overall evolution is an increase of structure between a redshift of $z \sim 1.1$ and the present epoch, which might be even stronger than we have measured, since at the highest redshifts we miss the faint, less clustered galaxies.

Our results can be explained if we assume that

- the first galaxies form in a highly clustered state,
- their successors form in an increasingly lower clustered environment,
- the epoch of galaxy formation is not a burst-like event but is spread over a large range of redshifts, and
- there is a substantial amount of galaxy evolution (including merging) between a redshift of $z = 1$ and today.

Chapter 7

Outlook

We have seen that two-point correlation functions are not only suited to simply describe the presence or absence of structure in a distribution of galaxies, but also provide a powerful tool to investigate the formation and evolution of galaxies within the underlying dark matter density field. Not all galaxies trace mass (i.e. older galaxies are much more strongly clustered than young, starforming galaxies). The exploration of the processes which lead to the different evolution of the clustering of galaxies of different Hubble types can help us to understand the interaction between the pure structure growth of the dissipationless dark matter component and the development of the baryonic matter into stars and galaxies. In principle the evolution of the correlation function can also place constraints on the cosmological parameters which determine the geometry and dynamics of our universe, but more observations are needed to disentangle dark matter clustering growth and the evolution of the bias.

7.1 The merger rate at intermediate redshifts

An interesting question in the context of galaxy evolution is the evolution of the merger rate with redshift. In a hierarchical clustering scenario, merging plays an important role, for example for the mass function (and therefore the luminosity function of galaxies), or the formation of massive ellipticals. It also can be a possible explanation for the slow evolution of the clustering of early type galaxies, which can better be explained by biasing models which take merging into account, than by galaxy conservation models (Daddi et al., 2001). However, it is very difficult to deduce the merger rate directly – not each close pair (~ 5 to $20\,\mathrm{kpc}$) is in the process of merging, if their peculiar velocities are large ($\Delta v > 500\,\mathrm{km\ s^{-1}}$) they move on hyperbolic orbits and will not merge. Also there is an ambiguity in the relation between redshift space and real space – a redshift difference of $\Delta v = 500\,\mathrm{km\ s^{-1}}$ can correspond either to two galaxies at a small physical separation with a large infall velocity, or to two galaxies at a separation of $\sim 5h^{-1}\,\mathrm{Mpc}$, with no peculiar velocity. Thus not only their spatial separation must be known, but also their peculiar velocities. Although it is therefore not possible to deduce the merger rate from a correlation function, we can nevertheless at least try to estimate the *close pair fraction* of galaxies at different redshifts by means of correlation functions. This quantity is an upper limit of the merger rate. The present day close pair fraction for galaxies with $-21 \leq M_B \leq -18$ has been estimated by Patton et al. (2000) to be $N_c = 0.0226 \pm 0.0052$ at $z = 0.015$. This implies that $\sim 2.3\%$ of the galaxies have companions within a projected

physical separation of $5h^{-1} \leq r_p \leq 20h^{-1}\,\mathrm{kpc}$ and $\Delta v = 500\,\mathrm{km\ s^{-1}}$. In hierarchical galaxy evolution scenarios, the merger rate of field galaxies is expected to increase with increasing redshift, and probably also with fainter luminosities.

The signature of very close pairs is expected to be a bumb at the very small scale end, as we have seen in the case of the cross-correlation function between bright and faint galaxies (see Chapter 5). We already have carried out some simulations to quantify this effect. We took the CADIS data, and added a certain percentage of artificial neighbours to randomly chosen galaxies. Those neighbours are located in a spherical shell with radius $r + \mathrm{d}r = 15 + 5\,\mathrm{kpc}$ around the catalogue galaxy. We calculated both angular and projected correlation function in two different redshift bins ($0.2 \leq z \leq 0.5$ and $0.5 \leq z \leq 0.75$), for 1%, 10%, and 30% close pairs, respectively. Figure (7.1) and (7.2) show the results for $\Omega_0 = 1$, $\Omega_\Lambda = 0$.

In all figures the black line is the fit for the unchanged data, fitted as usually in the range $0.88 \leq \log\theta \leq 2.28$ for the angular, and $-1.7 \leq \log r_p \leq -0.3$ for the projected correlation function. Depending on the percentage of close pairs introduced into the catalogue, there is a more or less pronounced bump at the small-scale end of *both* angular and projected correlation function. This bump with respect to the fit is even visible in the unchanged data, which indicates that a certain amount of close pairs is in fact present and leads to a change of the correlation function for small pair separations.

The height of this bump could be used as an estimate of the close pair fraction at different redshifts. From a first comparison with the modified data we can conclude that the close pair fraction is certainly smaller than 10%, and the bump can more easily be found in the redshift bin $0.2 \leq z \leq 0.5$ than for the higher redshift interval ($0.5 \leq z \leq 0.75$), where we are hardly able to resolve pairs with a separation of $< 20\,\mathrm{kpc}$ on groundbased data.

We need to follow up this possibility and find a quantification for the effect.

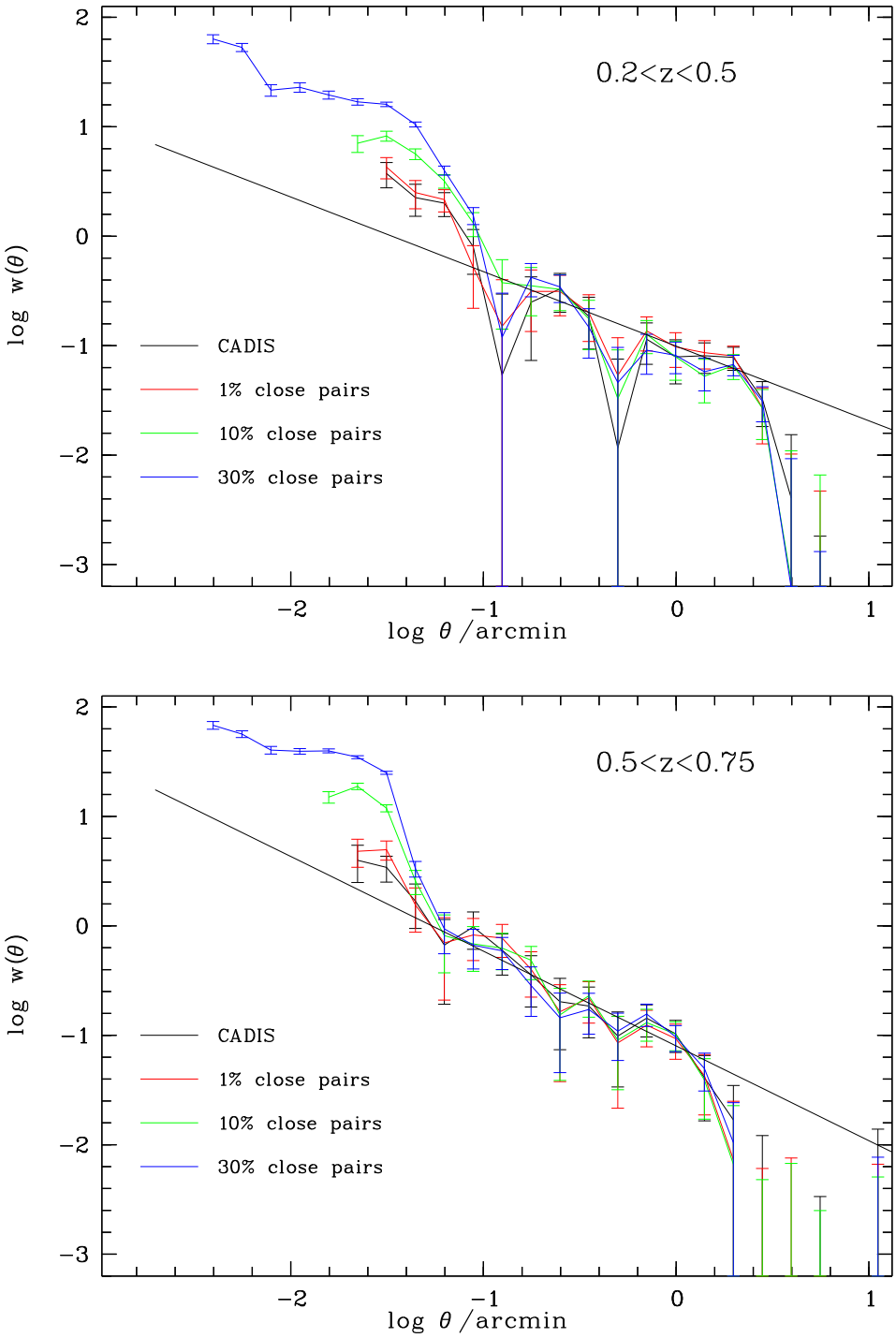

Figure 7.1: The influence of close pairs on the angular correlation function.

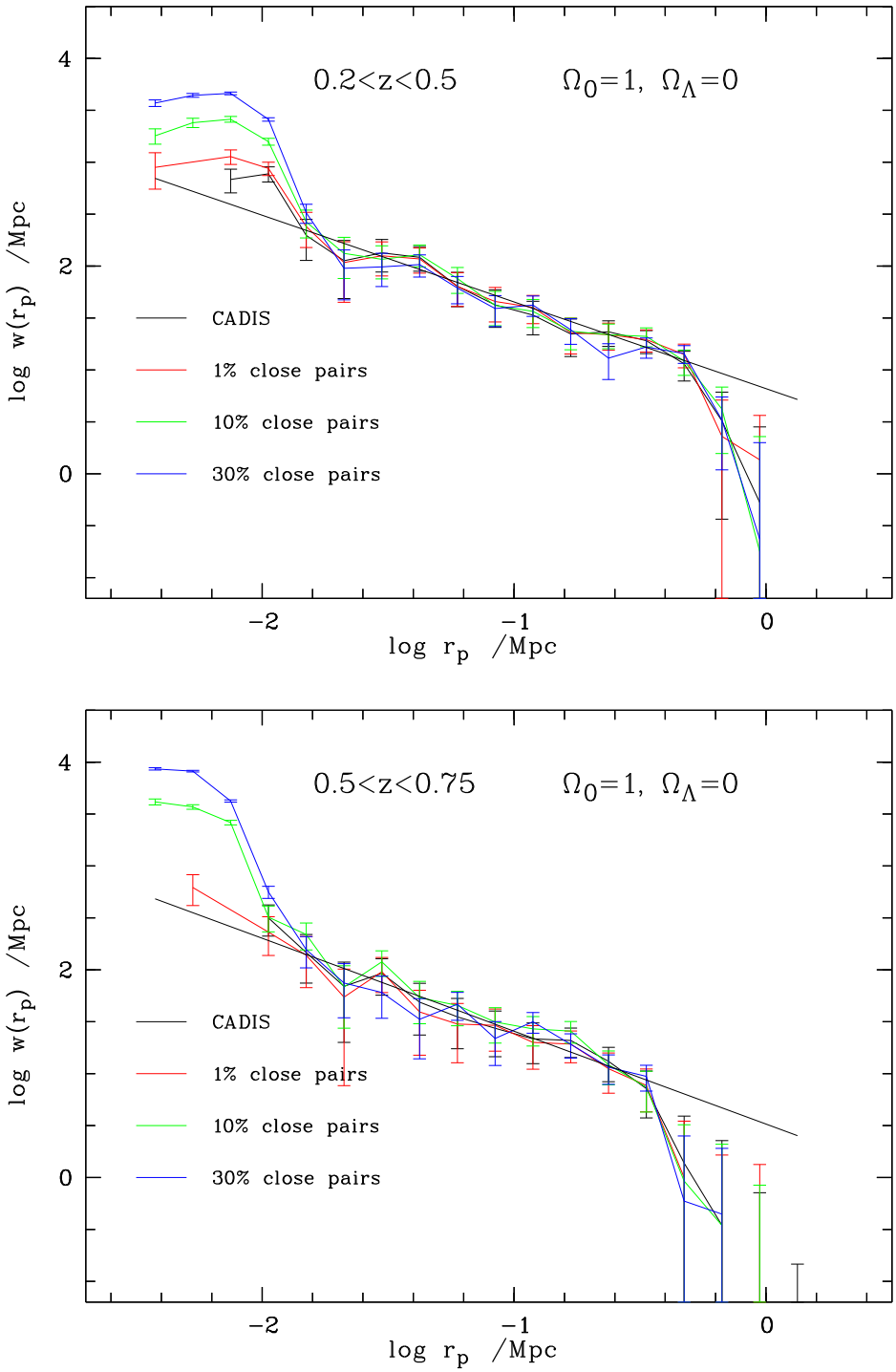

Figure 7.2: The influence of close pairs on the projected correlation function.

7.2 The emission line galaxies in CADIS

In this work only one part of the whole data base of CADIS was analysed, namely the multicolor galaxy sample. But CADIS assembles essentially two different survey strategies, the multicolor survey, and an emission-line survey using an imaging Fabry-Perot interferometer, to probe emission line galaxies down to a limiting line flux of $\sim 3 \cdot 10^{-20}$ W m^{-2}. These galaxies, which have been detected and classified by their emission lines, have redshifts with an accuracy of 120 km s^{-1} – good enough to calculate the three-dimensional correlation function directly. The special observing technique samples galaxies in distinct narrow redshift bins, which allows for the investigation of the evolution of the clustering properties of emission line galaxies between a redshift of $z \approx 1.4$ and $z \approx 0.24$. Galaxies which show strong emission lines are galaxies which form, or have recently formed stars. Therefore they make up a valuable sample for comparing the different clustering properties of early type (quiet) and actively star forming galaxies. We expect their clustering amplitude to be very low, and, following Kauffmann et al. (1999a), to be strongly evolving with redshift.

However, since their number is small ($N \approx 1000$ for the full CADIS), it is impossible to carry out this analysis before the full CADIS emission line survey has been completed.

7.3 COMBO 17

This amazing data base (see Chapter 6) can be used for various investigations, using either the projected or the angular correlation function. First of all the higher statistic allows for a more detailed analysis of the evolution of the clustering, a repetition of the analysis carried out on the CADIS data will not only corroborate the results, but might also show a possible dependency of the growth rate on redshift (a change of the parameter q with redshift). It also should be possible to investigate, whether the slope γ of the correlation function changes with redshift, or is different for different Hubble types. The clustering evolution of different Hubble types itself can also be investigated in smaller SED bins, and the difference of the amplitudes of the correlation function of bright and faint galaxies can be followed to larger redshifts.

With a deeper survey like COMBO 17, we can furthermore calculate the projected cross-correlation function between bright and faint galaxies for higher redshifts as we did before (see section (5.2.4)). As we have seen, the signature of satellite galaxies as well is a bump at the small-scale end, so this analysis will help place constraints on the merger rate. If a quantification of the close pair fraction can be found, with the COMBO 17 data it is probably also possible to investigate the dependency of the merger rate, or close pair fraction, respectively, on the Hubble type, or the absolute rest-frame B-band luminosity.

The investigation of the clustering properties of various subsamples from the COMBO 17 survey can be carried out rather easily with the methods and algorithms developed in this thesis. As soon as the classification and redshift determination for more fields is completed, this analysis will immediately yield results with a unique accuracy and significance.

Appendix A

The relation between $\xi(r)$ and the power spectrum of the density fluctuations

The correlation function is directly related to the density contrast $\Delta(x) = \delta\varrho/\varrho$. With $\varrho = \varrho_0[1 + \Delta(x)]$, equation (3.2) can be written as

$$\mathrm{d}N_{pair}(\mathbf{r}) = \varrho(\mathbf{x})\mathrm{d}V_1\varrho(\mathbf{x}+\mathbf{r})\mathrm{d}V_2 \ . \tag{A.1}$$

Therefore

$$\mathrm{d}N_{pair}(\mathbf{r}) = \varrho_0^2[1 + \Delta(\mathbf{x})][1 + \Delta(\mathbf{x}+\mathbf{r})]\mathrm{d}V_1 \ \mathrm{d}V_2 \ . \tag{A.2}$$

The average value of Δ is zero by definition and therefore the two-point correlation function is just

$$\mathrm{d}N_{pair}(r) = \varrho_0^2[1 + \langle\Delta(\mathbf{x})\Delta(\mathbf{x}+\mathbf{r})\rangle]\mathrm{d}V_1 \ \mathrm{d}V_2 \ . \tag{A.3}$$

This shows explicitly the relation between the density contrast on different scales r and the two-point correlation function:

$$\xi(r) = \langle\Delta(\mathbf{x})\Delta(\mathbf{x}+\mathbf{r})\rangle \tag{A.4}$$

We can now relate the spectrum of the fluctuations (in terms of the spatial Fourier transforms of $\Delta(\mathbf{r})$) to the two-point correlation function, which is, by definition, spherically symmetric about each point.

First we define the Fourier transform pair for $\Delta(\mathbf{r})$

$$\begin{aligned} \Delta(\mathbf{r}) &= \frac{V}{(2\pi)^3}\int \Delta_{\mathbf{k}}\mathrm{e}^{-i\mathbf{kr}} \ \mathrm{d}^3k \\ \Delta_{\mathbf{k}} &= \frac{1}{V}\int \Delta(\mathbf{r})\mathrm{e}^{-i\mathbf{kr}}\mathrm{d}^3x \ . \end{aligned} \tag{A.5}$$

With use of Parseval's theorem to relate the integrals of the squares of $\Delta(\mathbf{r})$ and its Fourier transform $\Delta_{\mathbf{k}}$, one gets

$$\frac{1}{V}\int \Delta^2(\mathbf{r})\mathrm{d}^3x = \frac{V}{(2\pi)^3}\int |\Delta_{\mathbf{k}}|^2\mathrm{d}^3k \ . \tag{A.6}$$

The quantity on the left-hand side of equation (A.6) is the mean square amplitude of the fluctuations per unit volume, and $|\Delta_{\mathbf{k}}|^2$ is the *power spectrum* of the fluctuations, $P(k)$. Therefore we can write

$$\langle\Delta^2\rangle = \frac{V}{(2\pi)^3}\int |\Delta_{\mathbf{k}}|^2 \mathrm{d}^3k = \frac{V}{(2\pi)^3}\int P(k)\mathrm{d}^3k \ . \tag{A.7}$$

Since the two-point correlation function is spherically symmetric, the element of k-space can be written $\mathrm{d}^3k = 4\pi k^2\mathrm{d}k$ and so

$$\langle\Delta^2\rangle = \frac{V}{2\pi^2}\int |\Delta_{\mathbf{k}}|^2 k^2\mathrm{d}k = \frac{V}{2\pi^2}\int P(k)k^2\mathrm{d}k \tag{A.8}$$

With equation (A.4) we can relate $\langle\Delta^2\rangle$ to the two-point correlation function. $\Delta(\mathbf{x})$ can be written as a Fourier series:

$$\Delta(\mathbf{x}) = \sum_{\mathbf{k}} \Delta_{\mathbf{k}} \mathrm{e}^{-i\mathbf{kx}} \ . \tag{A.9}$$

$\Delta(\mathbf{x})$ is a real function and therefore we can find $|\Delta(\mathbf{r})|^2$ by writing $|\Delta(\mathbf{r})|^2 = |\Delta(\mathbf{r})\Delta^*(\mathbf{r})$, where $\Delta^*(\mathbf{r})$ is the complex conjugate of $\Delta(\mathbf{r})$. Taking the average value of the product of $\Delta(\mathbf{x})$ and $\Delta(\mathbf{x}+\mathbf{r})$ in the same way, one finds

$$\xi(r) = \langle\sum_{\mathbf{k}}\sum_{\mathbf{k}'} \Delta_{\mathbf{k}}\Delta^*_{\mathbf{k}'}\mathrm{e}^{-i(\mathbf{k}-\mathbf{k}')\mathbf{x}}\mathrm{e}^{-i\mathbf{k}'\mathbf{r}}\rangle \ . \tag{A.10}$$

All cross terms vanish except those for which $\mathbf{k} = \mathbf{k}'$. Therefore

$$\xi(r) = \sum |\Delta_{\mathbf{k}}|^2 \mathrm{e}^{-i\mathbf{kr}} \ . \tag{A.11}$$

Conversion into a Fourier integral gives

$$\xi(r) = \frac{V}{(2\pi)^3}\int |\Delta_{\mathbf{k}}|^2 \mathrm{e}^{-i\mathbf{kr}}\mathrm{d}^3k \ . \tag{A.12}$$

$\xi(r)$ is a real function, so we can take only the real part of $\mathrm{e}^{-i\mathbf{kr}}$, and because of the spherical symmetry of the two-point correlation function, we integrate over an isotropic distribution of angles θ. Thus we obtain

$$\xi(r) = \frac{V}{2\pi^2}\int |\Delta_k|^2 \frac{\sin kr}{kr}k^2\mathrm{d}k = \frac{V}{2\pi^2}\int P(k)\frac{\sin kr}{kr}k^2\mathrm{d}k \ . \tag{A.13}$$

The function $\sin kr/kr$ acts as a window function, which allows only wavenumbers $k < r^{-1}$ to contribute to the amplitude of the fluctuations on the scale r. Fluctuations with larger wavenumbers, corresponding to smaller scales, average out to zero on the scale r.

Appendix B

Cosmological Distances

The calculation of the three-dimensional correlation function raises the question how to compute the distances between two galaxies with different redshifts z_i and z_j, which are separated by an angle θ. The only measurement one has for the radial distances of the galaxies is the redshift, so we have to find a way to estimate the distances from the redshifts. This is not straightforward, and so the most important points are listed here.

The starting point and basis for all further calculations is the General Theory of Relativity developed by A. Einstein, together with the **Cosmological Principle**. The Cosmological Principle states that the Universe is spacially homogenious and isotropic. This assumption results in a simlification of Einsteins field equations which reduce to the pair of independent equations:

$$\dot{R}^2(t) = \frac{8\pi G\varrho(t)R^2(t)}{3} + \frac{\Lambda R^2(t)}{3} - kc^2 \ , \tag{B.1}$$

and

$$\frac{\ddot{R}(t)}{R(t)} = -\frac{4\pi G\varrho(t)}{3} + \frac{\Lambda}{3} \ . \tag{B.2}$$

G is the gravitational constant, $\varrho(t)$ the matter density (throughout this work negligible pressure is assumed), Λ the cosmological constant and the sign of k determines the curvature of the three-dimensional space. R is the scale factor, which has the dimension of a length, c is the speed of light, and t the cosmic time as measured by a *fundamental observer*. The concept of fundamental observers can be introduced according to *Weyl's postulate*, which says that the worldlines of all points in the universe do not intersect, except at one point at the very beginning – the big bang. Once this postulate is adopted, it becomes possible to assign a notional observer to each worldline – those are the fundamental observers. The time measured by the fundamental observer with a standard clock is called cosmic time. There is a very simple way to synchronize all the clocks: all the fundamental observers set their clocks to one distinct time in that very moment when the temperature of the microwave background drops to a certain value.

The only isotropic curved spaces are those in which the curvature is constant throughout the whole space, and can take positive, zero or negative values.

The general form of the line element in a homogeniously expanding, isotropic curved space, which satisfies the field equations, is described by the *Robertson-Walker-Metric*, in the following RWM:

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - R^2(t)\left[\frac{\mathrm{d}\sigma^2}{(1-k\sigma^2)} + \sigma^2\mathrm{d}\theta^2 + \sigma^2\sin^2\theta\mathrm{d}\phi^2\right] , \tag{B.3}$$

where s is the four-dimensional distance, θ and ϕ are angular coordinates, and σ is the dimensionless radial coordinate. σ is also called *comoving coordinate*, because it can be understood as a kind of label fixed to a certain galaxy, which never changes it's value. The varying scale factor $R(t)$ is taking account of the expansion. With

$$\begin{aligned} H &= \frac{\dot{R}}{R} , \\ \Omega &= \frac{8\pi G\varrho}{3H^2} , \\ \Omega_\Lambda &= \frac{\Lambda}{3H^2} , \end{aligned}$$

H being the Hubble parameter, Ω the density parameter, and Ω_Λ the normalised cosmological constant, we can calculate

$$kc^2 = R^2H^2(\Omega + \Omega_\Lambda - 1) , \tag{B.4}$$

so that

$$k = sign(\Omega + \Omega_\Lambda - 1) . \tag{B.5}$$

Since $R > 0$ we can write

$$R = \frac{c}{H}\frac{1}{\sqrt{|\Omega + \Omega_\Lambda - 1|}} ; \tag{B.6}$$

this is the radius of curvature of the three-dimensional space at time t. For $k = 0$ it is convenient to *define* the scale factor R to be c/H.

As can be seen from equation (B.1), and (B.3), the constant k determines the geometry of the universe.

- If the curvature parameter $k = -1$, then the space is negatively curved (the sum of the angles of a triangle is less than 180°)
- For $k = 0$ the universe is flat (the sum of the angles of a triangle equals 180°)
- If the curvature parameter $k = +1$, then the space is positively curved (the sum of the angles of a triangle is more than 180°)

For the special case of $\Omega_\Lambda = 0$, k determines not only the curvature, but also the dynamical properties of the universe:

- If the curvature parameter $k = -1$ and $\Omega_\Lambda = 0$, the universe is open, and the expansion will contiue forever.

- For $k = 0$ and $\Omega_\Lambda = 0$ the universe is neither open nor closed, coming to a halt only only as $t \to \infty$.
- If the curvature parameter $k = +1$ and $\Omega_\Lambda = 0$, then the expansion will someday halt and reverse itself

With the cosmological constant present, the dynamic properties and the geometric properties are independent.

In the following the index 0 will be used to denote the present value of a given quantity, fixed at the time t_0 of observation. The explicit dependence on t will be dropped for brevity. Taking matter conservation into account and using the present day values, we have $\varrho R^3 = \varrho_0 R_0$, and so

$$\dot{R}^2 = H_0^2 R_0^2 \left(\frac{\Omega_0 R_0}{R} + \frac{\Omega_\Lambda R^2}{R_0^2} - (\Omega_0 + \Omega_\Lambda - 1) \right) . \tag{B.7}$$

With

$$z = \frac{R_0}{R} - 1 \tag{B.8}$$

we get from equation (B.7)

$$\mathrm{d}z = \frac{\mathrm{d}z}{\mathrm{d}R} \dot{R} \mathrm{d}t = -H_0 (1+z) \sqrt{Q(z)} \mathrm{d}t \ , \tag{B.9}$$

where

$$Q(z) = \Omega_0 (1+z)^3 - (\Omega_0 + \Omega_\Lambda - 1)(1+z)^2 - \Omega_\Lambda \ . \tag{B.10}$$

The coordinate distance σ of the RWM can be represented as

$$\sigma = F(\chi) = \begin{cases} \sinh\chi & : \quad for\ k = -1 \\ \chi & : \quad for\ k = 0 \\ \sin\chi & : \quad for\ k = +1 \end{cases} \tag{B.11}$$

with

$$\chi = \frac{c}{H_0 R_0} \int_{z_1}^{z_2} \frac{\mathrm{d}z}{\sqrt{Q(z)}} \tag{B.12}$$

In a static Euclidean space, one can define a variety of distances according to the method of measurement, which are all equivalent. In an expanding universe with non-zero curvature one can define the same set of distances in a way such that they look as Euclidean as possible, and they all can be related to each other.

The dimensionless comoving coordinate can be transformed into a distance by multiplying σ with the present day scale factor R:

$$D = R_0 \sigma \ . \tag{B.13}$$

This distance has the meaning of the proper distance that an object has today. The proper distance at some other time can be obtained by deviding equation (B.13) by $(z+1)$, where z is the redshift at the corresponding time. For the special case of $\Omega_\Lambda = 0$, an analytic formula for this distance could be derived (Mattig, 1958):

$$D = \frac{2c}{H_0}\frac{1}{\Omega_0^2}\left[\Omega_0 z - (2-\Omega_0)(\sqrt{1+\Omega_0 z}-1)\right]\frac{1}{1+z} \tag{B.14}$$

The angular diameter distance is a distance defined such that the relation between proper length d of an object and its angular projection on the sky (θ) looks as Euclidian as possible:

$$d = D_A \cdot \theta \ , \tag{B.15}$$

$$D_A = \frac{R_0 \sigma}{(1+z)} \ . \tag{B.16}$$

In the same way one can define a luminosity distance:

$$D_L^2 = \frac{L}{4\pi F} \ , \tag{B.17}$$

where L is the luminosity of the source, and F is the flux measured by the observer. Let a source of light be located at the origin of a comoving coordinate system. The source emmits photons that arrive at a spherical surface around the origin, which (at the present time) has the area $4\pi D_C^2$, so the radiant flux will be diminish as $1/D_C^2$. Two effects, in addition to the inverse square law, act to reduce the value of the radiant flux measured at this sphere. The energy of each photon is reduced by a factor of $1+z$. Also, cosmological time dilation affects the average time intervalls between photons emitted by the source. This means that the rate at which the photons arrive at the sphere is less than the rate at which they leave the source by another factor of $1+z$. Combining these effects, the flux at the sphere's surface is

$$F = \frac{L}{4\pi (R_0\sigma)^2 (1+z)^2} \ , \tag{B.18}$$

so for the luminosity distance one finds

$$D_L = R_0 \sigma (1+z) \ . \tag{B.19}$$

For the comparison of the correlation function at different times in the evolution of the large scale structure we need to know the physical distances between the galaxies in the epoch under consideration. This is a major problem, because the physical distance between two galaxies at two different redshifts z_i and z_j is not properly defined - since two different redshifts mean two different epochs, it is only possible to give a physical distance of the two objects at the present epoch.

For the angular distance the *reciprocity theorem* applies:

$$D_A^{i,j} = D_A^{j,i}\frac{1+z_i}{1+z_j} \ . \tag{B.20}$$

The symmetry properties of the angular distance are discussed in detail in Kayser et al. (1997).

A good approximation for $\Delta z = z_i - z_j \ll 1$ is to first calculate the distance the two galaxies have today, and then project it to the epoch under consideration, by deviding by $(1+\bar{z})$, where $\bar{z} = (z_i + z_j)/2$. For small radial intervals where curvature is negligible, we can also make a further simplification: instead of integrating the line element of the RWM with $\mathrm{d}\phi \neq 0$ and $\mathrm{d}\theta \neq 0$, we use the law of Pythagoras, as if it were static, Euclidean space. Thus

$$\begin{aligned} d_{i,j} &= \sqrt{(D_A(\bar{z}) \cdot \theta)^2 + \left(\frac{D(\bar{z})}{(1+\bar{z})}\right)^2}\,, \\ &= \sqrt{\left(\frac{R_0 \sigma_{0,\bar{z}}}{\bar{z}} \cdot \theta\right)^2 + \left(\frac{R_0 \sigma_{z_i,z_j}}{\bar{z}}\right)^2}\,. \end{aligned} \tag{B.21}$$

The subscripts in σ indicate that in the case of D_A the integration in χ (see equation(B.12)) is carried out from 0 to $\bar{z}$, for D it is carried out from z_i to z_j.

Bibliography

Bardeen, J. M., Bond, J. R., Kaiser, N., Szalay, A. S.: 1986. The statistics of peaks of Gaussian random fields. *Astrophysical Journal*, **304**, 15.

Barrow, J. D., Sonoda, D. H., Bhavsar, S. P.: 1984. A bootstrap resampling analysis of galaxy clustering. *Monthly Notices of the Royal Astronomical Society*, **210**, 19P.

Baugh, C. M.: 1996. The real-space correlation function measured from the APM Galaxy Survey. *Monthly Notices of the Royal Astronomical Society*, **280**, 267.

Bertin, E. Arnouts, S.: 1996. SExtractor: Software for source extraction. *Astronomy and Astrophysics Supplement Series*, **117**, 393.

Best, J. S.: 2000. An examination of the large-scale clustering of the Las Campanas Redshift Survey. *Astrophysical Journal*, **541**, 519.

Bharadwaj, S., Sahni, V., Sathyaprakash, B. S., Shandarin, S. F., Yess, C.: 2000. Evidence for Filamentarity in the Las Campanas Redshift Survey. *Astrophysical Journal*, **528**, 21.

Brainerd, T. G. Villumsen, J. V.: 1994. On the evolution of clustering and the biased galaxy formation scenario. *Astrophysical Journal*, **431**, 477.

Bruzual A., G. Charlot, S.: 1993. Spectral evolution of stellar populations using isochrone synthesis. *Astrophysical Journal*, **405**, 538.

Cabanac, R. A., de Lapparent, V., Hickson, P.: 2000. Evolution of faint galaxy clustering. The 2-point angular correlation function of 20,000 galaxies to V<23.5 and I<22.5. *Astronomy and Astrophysics*, **364**, 349.

Carlberg, R. G., Yee, H. K. C., Morris, S. L., Lin, H., Hall, P. B., Patton, D., Sawicki, M., Shepherd, C. W.: 2000. Galaxy clustering evolution in the CNOC2 High-Luminosity Sample. *Astrophysical Journal*, **542**, 57.

Colin, P., Carlberg, R. G., Couchman, H. M. P.: 1997. The Omega 0 Dependence of the Evolution of $\xi(r)$. *Astrophysical Journal*, **490**, 1.

Colley, W. N.: 1997. Two-dimensional Topology of Large-Scale Structure in the Las Campanas Redshift Survey. *Astrophysical Journal*, **489**, 471.

Daddi, E., Broadhurst, T., Zamorani, G., Cimatti, A., Röttgering, H., Renzini, A.: 2001. The Spatial Clustering od Distant, $z \sim 1$, Early-Type Galaxies. *astro-ph/0107340*.

Davis, M., Efstathiou, G., Frenk, C. S., White, S. D. M.: 1985. The evolution of large-scale structure in a universe dominated by cold dark matter. *Astrophysical Journal*, **292**, 371.

Davis, M. Geller, M. J.: 1976. Galaxy Correlations as a Function of Morphological Type. *Astrophysical Journal*, **208**, 13.

Davis, M., Huchra, J., Latham, D. W., Tonry, J.: 1982. A survey of galaxy redshifts. II - the large scale space distribution. *Astrophysical Journal*, **253**, 423.

Davis, M. Peebles, P. J. E.: 1983. A survey of galaxy redshifts. V - The two-point position and velocity correlations. *Astrophysical Journal*, **267**, 465.

Diaferio, A., Kauffmann, G., Colberg, J. . M., White, S. D. M.: 1999. Clustering of galaxies in a hierarchical universe - III. Mock redshift surveys. *Monthly Notices of the Royal Astronomical Society*, **307**, 537.

Doroshkevich, A. G., Tucker, D. L., Oemler, A. J., Kirshner, R. P., Lin, H., Shectman, S. A., Landy, S. D., Fong, R.: 1996. Large- and superlarge-scale structure in the Las Campanas Redshift Survey. *Monthly Notices of the Royal Astronomical Society*, **283**, 1281.

Efstathiou, G., Davis, M., White, S. D. M., Frenk, C. S.: 1985. Numerical techniques for large cosmological N-body simulations. *Astrophysical Journal Supplement Series*, **57**, 240.

Francis, P. J., Hewett, P. C., Foltz, C. B., Chaffee, F. H., Weymann, R. J., Morris, S. L.: 1991. A high signal-to-noise ratio composite quasar spectrum. *Astrophysical Journal*, **373**, 465.

Fried, J. W., von Kuhlmann, B., Meisenheimer, K., Rix, H.-W., Wolf, C., Hippelein, H. H., Kümmel, M., Phleps, S., Röser, H. J., Thierring, I., Maier, C.: 2001. The luminosity function of field galaxies and its evolution since z=1. *Astronomy and Astrophysics*, **367**, 788.

Fry, J. N.: 1996. The Evolution of Bias. *Astrophysical Journal Letters*, **461**, L65.

Geller, M. J. Huchra, J. P.: 1989. Mapping the universe. *Science*, **246**, 897.

Giavalisco, M., Steidel, C. C., Adelberger, K. L., Dickinson, M. E., Pettini, M., Kellogg, M.: 1998. The Angular Clustering of Lyman-Break Galaxies at Redshift Z approximately 3. *Astrophysical Journal*, **503**, 543.

Griffiths, L. M., Melchiorri, A., Silk, J.: 2001. Cosmic Microwave Background Constraints on a Baryonic Dark Matter-dominated Universe. *Astrophysical Journal Letters*, **553**, L5.

Groth, E. J. Peebles, P. J. E.: 1977. Statistical analysis of catalogs of extragalactic objects. VII - Two- and three-point correlation functions for the high-resolution Shane-Wirtanen catalog of galaxies. *Astrophysical Journal*, **217**, 385.

Hamilton, A. J. S.: 1993. Toward better ways to measure the galaxy correlation function. *Astrophysical Journal*, **417**, 19.

Hamilton, A. J. S., Matthews, A., Kumar, P., Lu, E.: 1991. Reconstructing the primordial spectrum of fluctuations of the universe from the observed nonlinear clustering of galaxies. *Astrophysical Journal Letters*, **374**, L1.

Hauser, M. G. Peebles, P. J. E.: 1973. Statistical Analysis of Catalogs of Extragalactic Objects. II. the Abell Catalog of Rich Clusters. *Astrophysical Journal*, **185**, 757.

Hu, W., Fukugita, M., Zaldarriaga, M., Tegmark, M.: 2001. Cosmic Microwave Background Observables and Their Cosmological Implications. *Astrophysical Journal*, **549**, 669.

Hubble, E.: 1934. The Distribution of Extra-Galactic Nebulae. *Astrophysical Journal*, **79**, 8.

Hudon, J. D. Lilly, S. J.: 1996. The Clustering of Faint Galaxies and the Evolution of $\xi(r)$. *Astrophysical Journal*, **469**, 519.

Jain, B., Mo, H. J., White, S. D. M.: 1995. The evolution of correlation functions and power spectra in gravitational clustering. *Monthly Notices of the Royal Astronomical Society*, **276**, L25.

Jenkins, A., Frenk, C. S., Pearce, F. R., Thomas, P. A., Colberg, J. M., White, S. D. M., Couchman, H. M. P., Peacock, J. A., Efstathiou, G., Nelson, A. H.: 1998. Evolution of structure in cold dark matter universes. *Astrophysical Journal*, **499**, 20.

Jing, Y. P.: 1998. Accurate fitting formula for the two-point correlation function of dark matter halos. *Astrophysical Journal Letters*, **503**, L9.

Jing, Y. P., Mo, H. J., Boerner, G.: 1998. Spatial correlation function and pairwise velocity dispersion of galaxies: Cold dark matter models versus the Las Campanas Survey. *Astrophysical Journal*, **494**, 1.

Kaiser, N.: 1984. On the spatial correlations of Abell clusters. *Astrophysical Journal Letters*, **284**, L9.

Kauffmann, G., Colberg, J. . M., Diaferio, A., White, S. D. M.: 1999a. Clustering of galaxies in a hierarchical universe - II. Evolution to high redshift. *Monthly Notices of the Royal Astronomical Society*, **307**, 529.

Kauffmann, G., Colberg, J. M., Diaferio, A., White, S. D. M.: 1999b. Clustering of galaxies in a hierarchical universe - I. Methods and results at z=0. *Monthly Notices of the Royal Astronomical Society*, **303**, 188.

Kayser, R., Helbig, P., Schramm, T.: 1997. A general and practical method for calculating cosmological distances. *Astronomy and Astrophysics*, **318**, 680.

Kennicutt, R. C., Tamblyn, P., Congdon, C. E.: 1994. Past and future star formation in disk galaxies. *Astrophysical Journal*, **435**, 22.

Kinney, A. L., Calzetti, D., Bohlin, R. C., McQuade, K., Storchi-Bergmann, T., Schmitt, H. R.: 1996. Template Ultraviolet to Near-Infrared Spectra of Star-forming Galaxies and Their Application to K-Corrections. *Astrophysical Journal*, **467**, 38.

Kravtsov, A. V. Klypin, A. A.: 1999. The Origin and Evolution of Halo Bias in Linear and Nonlinear Regimes. *Astrophysical Journal*, **520**, 437.

Landy, S. D., Shectman, S. A., Lin, H., Kirshner, R. P., Oemler, A. A., Tucker, D.: 1996. The two-dimensional power spectrum of the Las Campanas Redshift Survey: Detection of Excess Power on 100 H -1 Mpc Scales. *Astrophysical Journal Letters*, **456**, L1.

Landy, S. D. Szalay, A. S.: 1993. Bias and variance of angular correlation functions. *Astrophysical Journal*, **412**, 64.

Le Fevre, O., Hudon, D., Lilly, S. J., Crampton, D., Hammer, F., Tresse, L.: 1996. The Canada-France Redshift Survey. VIII. Evolution of the Clustering of Galaxies from z to approximately 1. *Astrophysical Journal*, **461**, 534.

Lilly, S. J., Le Fevre, O., Crampton, D., Hammer, F., Tresse, L.: 1995. The Canada-France Redshift Survey. I. Introduction to the Survey, Photometric Catalogs, and Surface Brightness Selection Effects. *Astrophysical Journal*, **455**, 50.

Limber, D. N.: 1954. The analysis of counts of the extragalactic nebulae in terms of a fluctuating density field. II. *Astrophysical Journal*, **119**, 655.

Lin, H., Kirshner, R. P., Shectman, S. A., Landy, S. D., Oemler, A., Tucker, D. L., Schechter, P. L.: 1996. The power spectrum of galaxy clustering in the Las Campanas Redshift Survey. *Astrophysical Journal*, **471**, 617.

Longair, M. S., editor: 1998. *Galaxy formation.*

Loveday, J., Maddox, S. J., Efstathiou, G., Peterson, B. A.: 1995. The Stromlo-APM redshift survey. 2: Variation of galaxy clustering with morphology and luminosity. *Astrophysical Journal*, **442**, 457.

Madau, P., Pozzetti, L., Dickinson, M.: 1998. The Star Formation History of Field Galaxies. *Astrophysical Journal*, **498**, 106.

Mattig, W.: 1958. Uber den Zusammenhang zwischen Rotverschiebung und scheinbarer Helligkeit. *Astronomische Nachrichten*, **284**, 109.

Meisenheimer, K. Röser, H.-J.: 1986. In *Use of CCD Detectors in Astronomy, Baluteau J.-P. and D'Odorico S., (eds.)*, Seite 227.

Mo, H. J. White, S. D. M.: 1996. An analytic model for the spatial clustering of dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, **282**, 347.

Müller, V., Arbabi-Bidgoli, S., Einasto, J., Tucker, D.: 2000. Voids in the Las Campanas Redshift Survey versus cold dark matter models. *Monthly Notices of the Royal Astronomical Society*, **318**, 280.

Naab, T., Burkert, A., Hernquist, L.: 1999. On the Formation of Boxy and Disky Elliptical Galaxies. *Astrophysical Journal Letters*, **523**, L133.

Oke, J. B.: 1990. Faint spectrophotometric standard stars. *Astronomical Journal*, **99**, 1621.

Patton, D. R., Carlberg, R. G., Marzke, R. O., Pritchet, C. J., da Costa, L. N., Pellegrini, P. S.: 2000. New Techniques for Relating Dynamically Close Galaxy Pairs to Merger and Accretion Rates: Application to the Second Southern Sky Redshift Survey. *Astrophysical Journal*, **536**, 153.

Peacock, J. A.: 1997. The evolution of galaxy clustering. *Monthly Notices of the Royal Astronomical Society*, **284**, 885.

Peacock, J. A.: 1999. *Cosmological physics.* Cosmological physics. Publisher: Cambridge, UK: Cambridge University Press, 1999. ISBN: 0521422701.

Peacock, J. A. Dodds, S. J.: 1994. Reconstructing the Linear Power Spectrum of Cosmological Mass Fluctuations. *Monthly Notices of the Royal Astronomical Society*, **267**, 1020.

Pearce, F. R., Jenkins, A., Frenk, C. S., Colberg, J. M., White, S. D. M., Thomas, P. A., Couchman, H. M. P., Peacock, J. A., Efstathiou, G., The Virgo Consortium: 1999. A Simulation of Galaxy Formation and Clustering. *Astrophysical Journal Letters*, **521**, L992.

Peebles, P. J. E.: 1973. Statistical Analysis of Catalogs of Extragalactic Objects. I. Theory. *Astrophysical Journal*, **185**, 413.

Peebles, P. J. E.: 1974. Statistical Analysis of Catalogs of Extragalactic Objects.IV. Cross-Correlation of the Abell and Shane-Wirtanen Catalogs. *Astrophysical Journal Supplement Series*, **28**, 37.

Peebles, P. J. E.: 1975. Statistical Analysis of Catalogs of Extragalactic Objects. VI - The galaxy distribution in the Jagellonian field. *Astrophysical Journal*, **196**, 647.

Peebles, P. J. E.: 1980. *The large-scale structure of the universe.* Research supported by the National Science Foundation. Princeton, N.J., Princeton University Press, 1980. 435 p.

Peebles, P. J. E. Groth, E. J.: 1975. Statistical Analysis of Catalogs of Extragalactic Objects. V - Three-point correlation function for the galaxy distribution in the Zwicky Catalog. *Astrophysical Journal*, **196**, 1.

Peebles, P. J. E. Hauser, M. G.: 1974. Statistical Analysis of Catalogs of Extragalactic Objects. III. The Shane-Wirtanen and Zwicky catalogs. *Astrophysical Journal Supplement Series*, **28**, 19.

Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., Castro, P. G., Deustua, S., Fabbro, S., Goobar, A., Groom, D. E., Hook, I. M., Kim, A. G., Kim, M. Y., Lee, J. C., Nunes, N. J., Pain, R., Pennypacker, C. R., Quimby, R., Lidman, C., Ellis, R. S., Irwin, M., McMahon, R. G., Ruiz-Lapuente, P., Walton, N., Schaefer, B., Boyle, B. J., Filippenko, A. V., Matheson, T., Fruchter, A. S., Panagia, N., Newberg, H. J. M., Couch, W. J., The

Supernova Cosmology Project: 1999. Measurements of Omega and Lambda from 42 High-Redshift Supernovae. *Astrophysical Journal*, **517**, 565.

Phillipps, S., Fong, R., Fall, R. S. E. S. M., MacGillivray, H. T.: 1978. Correlation analysis deep galaxy samples - I. Techniques with applications to a two-colour sample. *Monthly Notices of the Royal Astronomical Society*, **182**, 673.

Pickles, A. J.: 1998. A stellar spectral flux library: 1150-25000 å. *Publications of the Astronomical Society of the Pacific*, **110**, 863.

Röser, H. . Meisenheimer, K.: 1991. The synchrotron light from the jet of 3C 273. *Astronomy and Astrophysics*, **252**, 458.

Schade, D., Lilly, S. J., Crampton, D., Hammer, F., Le Fevre, O., Tresse, L.: 1995. Canada-France Redshift Survey: Hubble Space Telescope Imaging of High-Redshift Field Galaxies. *Astrophysical Journal Letters*, **451**, L1.

Shandarin, S. F. Yess, C.: 1998. Detection of network structure in the Las Campanas Redshift Survey. *Astrophysical Journal*, **505**, 12.

Shapley, H.: 1933. *Harvard Bull.*, **890**, 1.

Shectman, S. A., Landy, S. D., Oemler, A., Tucker, D. L., Lin, H., Kirshner, R. P., Schechter, P. L.: 1996. The Las Campanas Redshift Survey. *Astrophysical Journal*, **470**, 172.

Somerville, R. S., Lemson, G., Sigad, Y., Dekel, A., Kauffmann, G., White, S. D. M.: 2001. Non-linear stochastic galaxy biasing in cosmological simulations. *Monthly Notices of the Royal Astronomical Society*, **320**, 289.

Steidel, C. C., Adelberger, K. L., Dickinson, M., Giavalisco, M., Pettini, M., Kellogg, M.: 1998. A Large Structure of Galaxies at Redshift Z approximately 3 and Its Cosmological Implications. *Astrophysical Journal*, **492**, 428.

Tegmark, M. Peebles, P. J. E.: 1998. The Time Evolution of Bias. *Astrophysical Journal Letters*, **500**, L79.

Toomre, A. Toomre, J.: 1972. Galactic Bridges and Tails. *Astrophysical Journal*, **178**, 623–666.

Totsuji, H. Kihara, T.: 1969. The Correlation Function for the Distribution of Galaxies. *Publications of the Astronomical Society of Japan*, **21**, 221.

Tucker, D. L., Oemler, A., Kirshner, R. P., Lin, H., Shectman, S. A., Landy, S. D., Schechter, P. L., Muller, V., Gottlober, S., Einasto, J.: 1997. The Las Campanas Redshift Survey galaxy-galaxy autocorrelation function. *Monthly Notices of the Royal Astronomical Society*, **285**, L5.

Turner, E. L.: 1991. Quasars and galaxy formation. I - The z greater than 4 objects. *Astronomical Journal*, **101**, 5.

Walsh, J.: 1995. Optical and uv spectrophotometric standard stars. *http://www.eso.org/observing/standards/spectra*.

Wolf, C., Dye, S., Kleinheinrich, M., Meisenheimer, K., Rix, H.-W., Wisotzki, L.: 2001a. Deep BVR photometry of the Chandra Deep Field South from the COMBO-17 survey. *astro-ph/0012474*.

Wolf, C., Meisenheimer, K., Röser, H. .: 2001b. Object classification in astronomical multi-color surveys. *Astronomy and Astrophysics*, **365**, 660.

Wolf, C., Meisenheimer, K., Röser, H. ., Beckwith, S. V. W., Chaffee, F. H., Fried, J., Hippelein, H., Huang, J. ., Kümmel, M., von Kuhlmann, B., Maier, C., Phleps, S., Rix, H. ., Thommes, E., Thompson, D.: 2001c. Multi-color classification in the Calar Alto Deep Imaging Survey. *Astronomy and Astrophysics*, **365**, 681.

Yee, H. K. C., Ellingson, E., Carlberg, R. G.: 1996. The CNOC Cluster Redshift Survey Catalogs. I. Observational Strategy and Data Reduction Techniques. *Astrophysical Journal Supplement Series*, **102**, 269.

Danksagung

Ich danke...

an erster Stelle meinem Betreuer Dr. Klaus Meisenheimer, für die spannende Aufgabenstellung, seine volle Unterstützung und die Geduld, die er so oft für mich aufgebracht hat

Herrn Prof. Dr. Immo Appenzeller und Herrn Prof. Dr. Hans-Walter Reinhard Rix, die mir die Promotion am Max-Planck-Institut für Astronomie ermöglichten

Herrn Prof. Dr. Immo Appenzeller für die Anfertigung des Zweitgutachtens

den Korrekturlesern, Natalie Winn, Dr. Thorsten Naab, Christopher Popp und Sebastian Jester

allen am MPIA für die entspannte Arbeitsatmosphäre

allen auf dem Calar Alto für ihre Hilfe und Unterstützung beim Beobachten, und für die schöne Zeit, die ich dabei so oft in Spanien hatte – ganz besonders danke ich Manolo Alises, Luzma Montoya, Alberto Aguirre, Felipe Hoyo sowie Uli Thiele, Jens Helmling und Enrique De Guindos

Thorsten, Sebastian, Olaf, Michael, Helmut, Markus, Markus, Sabine und Bernd für die Ausflüge zum Turm

meinem Spanischlehrer Miguel Lizaso, der mir in den letzten drei Jahren geholfen hat, meinen Horizont zu erweitern

meiner Mitbewohnerin Claudia, für den Spaß und die vielen langen Abende, an denen wir mal wieder vom Hundertsten ins Tausendste gekommen sind

Sebastian, Kris, Meike, Lennon, David & Hella – manche Dinge ändern sich nie

Natalie und Bruno für ihre wunderbare Freundschaft

Manuel, für das Öffnen verborgener Türen

Besonderer Dank gilt meinen Eltern, für ihre immerwährende Unterstützung in meinem Vorhaben, Astronom zu werden, ihr Vertrauen und ihr großes Verständnis für so Vieles