Groups, clusters and superclusters of galaxies

Moore, Ben

Groups, Clusters and Superclusters of Galaxies

by

Ben Moore

March 1991

An account of work done at the Department of Physics, submitted to the University of Durham in accordance with the regulations for admission to the degree of Doctor of Philosophy.

PREFACE

The work described in this thesis was undertaken between 1987 and 1991, whilst the author was a research student under the supervision of Dr. C. S. Frenk, in the department of Physics at the University of Durham. This work has not been submitted for any other degree at the University of Durham or at any other University.

In addition to collaborating with Dr. Carlos Frenk, some of the work in Chapters 3 and 4 was carried out in collaboration with Dr. Simon White (presently at Cambridge University) and Dr. David Weinberg (presently at the University of California at Berkeley).

Abstract

Galaxies are observed in a diverse range of associations. Understanding the dynamical, statistical and clustering properties of aggregations of galaxies forms the main body of this thesis.

On the smallest scale, we use a model for the Local Group to study the formation of a typical galaxy system and to understand the mass distribution within the Local Group. Our model is a binary system excised from numerical simulation of a Universe dominated by cold dark matter which has similar radial velocity and separation as M31 and the Milky Way. We find that the timing argument provides a reliable method for placing a lower limit to the mass of the Local Group. The anisotropy parameter of the particle orbits within the dark halos of CDM are predicted to radially biased. To reconcile the mass of the Milky Way with the predictions from the timing argument and the mass of our model halo, the satellites of our Galaxy must be on circularly biased orbits. The asymptotic values of the rotation curves of the two halos match very closely those of M31 and the Milky Way, as the halos come closer together the curves become very distorted. A simple treatment of gas within our model extends the rotation curves into the central regions of the halos. In our model, M31 and the Milky Way collide 2.5 Gyrs from the present time, a fraction of the current crossing time.

Intermediate scales are probed using a statistical analysis of groups of galaxies identified using well defined selection criteria from the CfA redshift survey. The grouping algorithm is optimised using artificial galaxy catalogues constructed from N-body simulations which have similar low order correlations to the original survey. We develop a method of estimating the total luminosity of groups of galaxies identified within magnitude limited redshift surveys and use it to calculate the luminosity function of galactic systems. This statistic measures the abundance of gravitationally bound structures, independent of the detailed arrangement of the

luminous material within them. We find that this function has a smooth transition from single galaxies to rich clusters. The distribution of group velocity dispersions shows a discontinuity at the transition between groups and rich clusters. The correlation function of groups is found to depend on the mass range of the sample, luminous groups are more strongly correlated than faint groups. We compare these results with predictions from the CDM model and have extended to intermediate scales the previous success of the model on galactic and cluster scales.

On large scales we have used an all sky redshift survey of galaxies detected by IRAS to investigate the topology of the Universe to a depth of $200h^{-1}$ Mpc. Qualitatively, the distribution of galaxies out to this distance appears similar to a Gaussian density field with a sponge-like topology. High and low density regions are topologically similar and surfaces of constant density are interconnected. Quantitatively, we have used the genus-threshold density relation of Hamilton *et al.* to measure the *slope* of the power spectrum over a range of length scales between $10h^{-1}$ Mpc and $50h^{-1}$ Mpc. To constrain the slope of the power spectrum we used artificial "galaxy" catalogues constructed from N-body simulations and a variety of Monte-Carlo and bootstrap techniques. Our topological analysis is consistent with a spectrum of power-law form with $n \simeq -1$ (in $\delta k^2 \propto k^n$) over the range of scales considered. Values of $n < -1.8$ and $n > 0$ are strongly ruled out by our data. The inferred power spectrum of the distribution of IRAS galaxies is similar to the predicted mass spectrum in the standard cold dark matter model on scales $\lesssim 15\ h^{-1}$ Mpc, but falls off less steeply on larger scales. This discrepancy is significant at over 2σ and implies that structure identified by IRAS galaxies is coherent over scales larger than expected from the CDM model.

Acknowledgements

Firstly my thanks are to my supervisor Carlos Frenk, for his guidance, patience and wisdom throughout my postgraduate training. Carlos is a man of many talents; as a researcher he excels and I could not have wished for a more imaginative and thoughtful tutor.

My parents deserve my gratitude for making me eat my cabbage whilst I was knee high to an astronomer, my father especially deserves my gracious thanks for introducing cosmology to me in my early years.

All the students and post-docs in the astronomy department deserve my gratitude for creating a great atmosphere for working (and playing!). Ian Smail deserves special thanks for running my N-body code for several hundred hours on his Sun workstation. For many interesting debates and discussions on the basics of Cosmology I would particularly like to thank Richard Bower and Ioannis Georgantoupolus.

Gabrielle, my lover and my mentor, deserves more gratitude than I can put into words. Thank you for your continuous supply of love, affection and wonderful sustenance, and your incredible patience and support during February.

In the search for truth there are certain questions that are not important. Of what material is the Universe constructed? Is the Universe eternal? Are there limits or not to the Universe? If a man were to postpone his search and practice for Enlightenment until such questions were solved, he would die before he found the path.

Buddha

All men dream; but not equally. Those who dream by night in the dusty recesses of their minds wake in the day to find that it was vanity; but the dreamers of the day are dangerous men, for they may act out their dreams with open eyes to make it possible.

Seven Pillars of Wisdom T.E. Lawrence

Contents

1 INTRODUCTION

1.1 COSMOLOGICAL MODELS

1.1.1 Recent history

In the 1970's cosmology was dominated by two competing theories known as the adiabatic and isothermal models. Both were based on the evolution of small primordial random phase fluctuations in a Universe containing only baryonic matter. The adiabatic theory adopted the assumption that only the fastest growing adiabatic fluctuation mode would emerge from the early Universe. This led to a model in which all small scale fluctuations were heavily damped so that the first non-linear structures to form were large "pancakes" of the size of superclusters (Silk 1968). Galaxies then formed from the fragmentation and cooling of these structures (Zeldovich 1970). The isothermal model supposed that some unspecified process suppressed the growing adiabatic mode. Fluctuations at late times are then dominated by the slowly growing isothermal mode which is only weakly damped, leaving fluctuations on all scales (Peebles *et al.* 1968). In this model the characteristic masses of galaxies are a natural reflection of the cooling timescales of gas clouds, galaxy associations form by hierarchical clustering of pre-existing objects (Binney 1977, Silk 1977, White and Rees 1978).

Big Bang nucleosynthesis sets stringent limits to the amount of baryons in the Universe. The latest calculations described by Pagel and Schramm (1991)

show that the contribution to the cosmic density in the form of baryons is $0.01 \geq \Omega_b h^2 \leq 0.02$. Thus if $\Omega = 1$, most of the mass in the Universe must be non-baryonic. Observationally we find that visible galaxies contribute only $\Omega_{vis} \approx 0.05$ to the cosmic density. The inflationary model (Guth 1981) provided a physical mechanism for the origin of primordial seeds. Combining inflation with standard theories of cosmological nucleosynthesis gives a unique cosmogonic model in which $> 90\%$ of the closure density is made up non-baryonic dark matter. Initially, the most attractive candidates for the dark matter were neutrinos, since these are the only candidates known to exist and because of a claim that the mass of the electron neutrino is in an interesting range (Lyubimov *et al.* 1980). Particles with lower velocities than neutrinos, known collectively as cold dark matter (CDM), were also considered. Particle physics suggests several candidates for CDM and although none have been detected, experiments searching for specific particles are under way. In addition to these theories, many alternative models for the formation of structure have been proposed. Perturbations induced by cosmic strings rather than quantum fluctuations (Vilenkin 1986), or models in which structure results from recent hydrodynamic (Ostriker and Cowie 1981) or radiative processes (Hogan and Kaiser 1983). The immediate aims of these theories are to reproduce observations of the large scale distribution of galaxies. Predictions from these latter models tend to be qualitative, partly because they are less developed, but mainly because they are not as fully specified as the neutrino or CDM models. The ultimate aim of any cosmological model is to create a framework within which we can understand and predict all the observable characteristics of galaxies and their environments.

1.1.2 The linear regime

The linear density fluctuation field is usually specified in terms of $\delta(r) = \rho(r)/\overline{\rho} - 1$, or its Fourier transform δ_k, where k denotes spatial frequency. In the

absence of known physical processes which could have introduced phase correlations or preferred scales, the fluctuations are assumed to have random phases and a power spectrum given by

$$|\delta_k|^2 \propto k^n. \tag{1.1}$$

Fluctuations generated by quantum effects during inflation are random phase, and their power spectrum is a power law with index $n = 1$, corresponding to the Harrison-Zeldovich "constant curvature" spectrum. As the Universe expands, increasingly long wavelengths come within the horizon which allows causal pressure and diffusive effects to alter their subsequent evolution. The primordial power spectrum is modified from a pure power law depending on which damping mechanisms act on the particular mix of particles.

Hot dark matter (HDM), such as massive neutrinos, retain their thermal motions over an extended period of time. Whilst the temperature of the Universe is over ($\sim 30/K$) ev the neutrinos are relativistic and any fluctuations coming within the horizon at this time are damped away by free streaming. This damping mechanism wipes out wavelengths shorter than $\sim 13(\Omega h^2)^{-1}$ and leads to the characteristic spectrum in Figure 1.1 (Bond *et al.* 1980). The qualitative properties of the neutrino model appeared to agree with the apparent "cell like" distribution of galaxies (Doroshkevich *et al.* 1980), however quantitative studies have shown that sufficient gas to make galaxies is unlikely to cool and fragment in pancakes (Shapiro *et al.* 1983, Bond *et al.* 1984) and the natural scale of clustering is too large to be acceptable (White *et al.* 1983, White *et al.* 1984). Regions of space entirely empty of galaxies are predicted to fill most of the Universe which conflicts with galaxy redshift surveys (White 1986).

Cold dark matter particles are either very heavy or slow moving, thermal motions are never important and free streaming is negligible. The CDM spectrum gains its characteristic shape in the radiation dominated era when the dominant

photon-baryon fluid undergoes acoustic oscillations (Mezaros 1974). This effect bends the $n = 1$ inflationary spectrum at the characteristic scale $\lambda_c \sim 13(\Omega h^2)^{-1}$ towards $n = -3$ at short wavelengths. The cold dark matter model has proven to be the most successful candidate for describing the features which we observe in the Universe. Simple scaling laws and cooling arguments were initially used to show that the CDM spectrum could produce the characteristic properties of halos, groups and clusters of galaxies (Blumenthal *et al.* 1984).

The CDM power spectrum of Davis *et al.* (1985) can be represented by

$$P(k) \propto \frac{k}{(1 + 1.7Lk + 9.0L^{3/2} + L^2k^2)^2} \tag{1.2}$$

where $L = \Omega^{-1}h^{-2}\text{Mpc} = 100\Omega^{-1}h^{-1}\text{km s}^{-1}$and we have parameterised the Hubble constant as $H_o = 100h\text{km s}^{-1}\ \text{Mpc}^{-1}$. There are two free parameters in the power spectrum, the amplitude of equation (1.2) and the length scale $L \propto (\Omega H_o^2)^{-1}$. In the standard model, $\Omega = 1$ and $\Omega h = 0.5$.

Figure 1.1 shows the power spectra at late times (after recombination) for adiabatic, constant curvature initial fluctuations imposed on a Universe now dominated by collisionless particles. The value $k^3|\delta_k|^2$ is closely related to the rms amplitude of density fluctuations averaged over regions of size k^{-1}, and thus gives an indication of the order in which objects of different scale fragment out of the general expansion.

1.1.3 The non-linear evolution of density fluctuations

Various approaches have been used to follow the non-linear evolution of the density field. Press and Schechter (1977) used analytic methods to predict the number of non-linear clumps as a function of time. Bardeen (1986) and Kaiser (1986) have studied the statistics of peaks in a Gaussian random field and applied this to the cosmological density field. An alternative approach is to follow the non-linear evolution of the density field using the Zeldovich approximation and

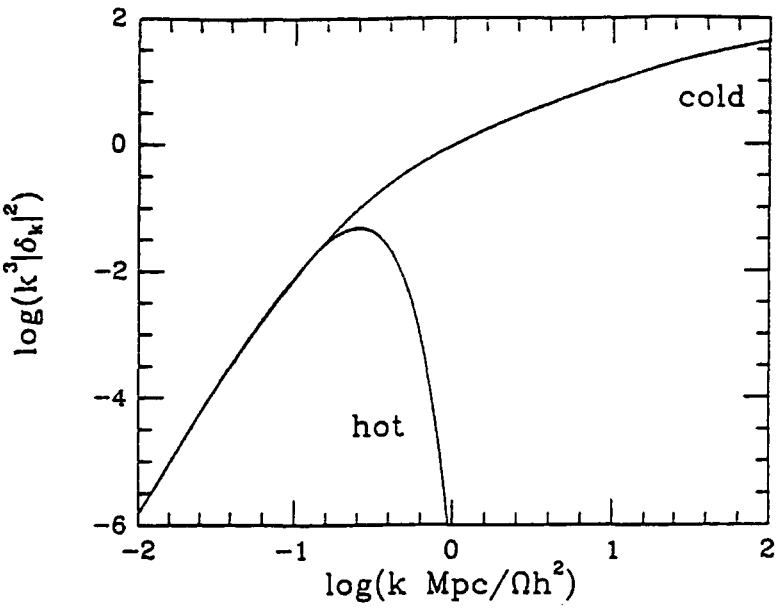


Figure 1.1 The power per decade as a function of spatial frequency for density fluctuations in Universes dominated by collisionless elementary particles. These are linear power spectra at late times evolved from the adiabatic, constant curvature fluctuations predicted by inflation. The curves are taken from results by Bond and Szalay (1983) and Bond and Efstathiou (1983).

Burgers equation (Gurvatov 1989). These analytic methods provide useful insight into the evolution of structure in the Universe, however they cannot predict the internal properties of bound structures. A more general treatment, which can be used for predictive purposes, requires numerical simulations.

Numerical simulations allow us to follow the evolution of $\sim 10^7$ particles by integrating the collisionless Boltzmann equation in an expanding Universe. This approach is only limited by the available computing facilities which constrain the resolution and dynamic range which can be followed. The CDM spectrum has power on all scales although it is expected that a range of mass structures will form almost simultaneously and aggregate hierarchically into larger systems. The "crosstalk" between different spatial frequencies questions the validity of the simple hierarchical clustering model. For example, the properties of small clumps depends largely on the large scale overdensities in which they are embedded. This provides

a possible explanation of the environmental properties of galaxies, although a large range of scales must be followed numerically in order to study these effects. A second feature of the crosstalk between different scales is a possible mechanism for natural biasing. The necessity of a biasing scheme arises in the CDM model due to the *a priori* assumption that $\Omega = 1$. Numerical models of a flat CDM Universe showed that the clustering length of the mass could only match the observations for unacceptable values of the Hubble constant. Correspondingly, the rms peculiar velocity of pairs of galaxies in the model were about a factor of three larger than the observed values of 400-700km s^{-1}(Davis *et al.* 1985). A flat Universe dominated by CDM can only be reconciled with these estimates if "galaxies" are biased with respect to the mass. A quasi-analytic model for this bias is to assume that galaxy formation takes place only within the *high peaks* of the initial density field. The currently adopted biasing procedure has two free parameters for turning a mass distribution into a galaxy distribution. One parameter for the scale of the smoothing filter applied to the linear density field and the other for the threshold height for identifying peaks of the density field with galaxies. Plausible values for these parameters lead to a "galaxy" distribution with two and three point correlation functions and pairwise velocities in agreement with observations. The justification for biasing is that the properties of the galaxy distribution are expected to reflect the environment within which they are situated, *i.e.* massive halos can form more easily in protoclusters than protovoids but they are also more easily disrupted within dense regions. Within the CDM model such an effect would lead to an over-representation of bright galaxies within dense regions, thus the clustering of the galaxies is enhanced relative to that of the dark matter.

The remaining free parameter, the amplitude of the fluctuation spectrum, is usually fixed by comparing the slope of the galaxy correlation function with observations. In Chapters 3 and 4 we use artificial galaxy catalogues constructed from the N-body simulations of Frenk *et al.* (1990). These galaxy catalogues are

constructed from the CDM density field using the analytic machinery of Bardeen *et al.* (1986; hereafter BBKS). BBKS showed that the local number density of peaks of a smoothed Gaussian random field can be related to the linear overdensity of the surrounding region. The high peak model is then used to predict the relative number of galaxies associated with each simulation particle from the linear initial conditions. The "galaxies" are given luminosities drawn from a Schechter function assumed to be independent of position with the overall normalisation adjusted to match the observed luminosity density for the Universe. The strength of the bias or "contrast" is determined by the threshold of the high peak model. We can define the biasing parameter b by $\sigma_{gal} = b\sigma_{\rho}$ where σ is the rms density fluctuation of the field filtered with a "top hat" window function. Observationally, $\sigma_{gal}(8h^{-1}\mathrm{Mpc}) \sim 1$. Once this procedure is completed then the predictions of the CDM model can be rigorously compared against a wide range of observations.

The CDM model has had a remarkable degree of success at predicting the observed distribution and dynamical properties of galactic size objects. On intermediate scales the model has again proven successful at predicting the characteristic properties of groups of galaxies (Nolthenius and White 1987) and the general "foamy" appearance of the galaxy distribution (White *et al.* 1987). On large scales the model is having some difficulties. The amplitude of the observed cluster-cluster correlation function (Bahcall and Soneira 1983) is $2 \sim 3\sigma$ higher than predicted by the model (White *et al.* 1987) and large scale streaming motions of galaxies are predicted to be small. Proponents of CDM claim that the Abell cluster catalogue (Abell 1958) is not suitable for a severe test of the model due to projection effects (Frenk 1989, Frenk *et al.* 1990, Sutherland 1988, Dekel *et al.* 1989) and the original claims for bulk flows in the galaxy distribution seem to have weakened as more observations are accrued (Kaiser *et al.* 1989). Two new galaxy catalogues have recently been completed which are posing more problems for the standard CDM model on large scales. The APM galaxy survey of $\sim$ two million galaxy magnitudes and positions (Maddox *et al.* 1990) and a survey of

sparsely sampled IRAS galaxies (Lawrence *et al.* 1989) show indications of large scale power somewhat in excess of the predictions of CDM (Efstathiou *et al.* 1990, Saunders *et al.* 1990, this thesis).

1.2 GALAXY ASSOCIATIONS

Galaxies are observed to exist in a wide range of environments, from small loose associations of a few galaxies, to compact groups with very high overdensities and to rich clusters containing many thousands of galaxies. Isolated galaxies are quite rare and most galaxies are observed to lie in small agglomerations or groups. Whereas the typical galaxy has luminosity $\sim L_*$, we shall see in Chapter 3 that a typical group has a total luminosity $\sim 6L_*$. This luminosity represents the characteristic scale on which bound systems of galaxies are currently observed in the clustering hierarchy. In Chapter 3 we demonstrate that groups of galaxies have typical mass to light ratios of $\Gamma \sim 200h\Gamma_\odot$ where $\Gamma_\odot = M_\odot/L_\odot$ (Table 1.1) and we have parameterised the Hubble constant as $H_o = 100h\text{km s}^{-1}\text{ Mpc}^{-1}$. If we express the critical density $\rho_c = 3H_o^2/(8\pi G)$ in terms of a mass to light ratio we find that $\Gamma \approx 1600h\Gamma_\odot$ is required for a flat Universe (where we have taken the luminosity density of the Universe to be the average of the estimates of Davis and Huchra 1982 and Kirshner *et al.* 1983). Extrapolating the mean density of groups to the Universe as a whole we find that most galaxies lie within systems which comprise about 12% of the cosmological critical density.

Rich clusters of galaxies, roughly defined as associations with total luminosity greater than $50L_*$ are very rare. Observers however, tend to focus on these very luminous aggregations for the simple reason that they are easy to identify in projection on the sky. Theoreticians also like rich clusters because rare or extreme objects provide effective tests of any theoretical paradigm and rich clusters probe the amplitude of the fluctuation spectra on large scales. Clusters of galaxies are often observed to lie in large superclusters which have not as yet collapsed out of the general Hubble flow. Zwicky (1938) was the first to suggest that the Local Group was part of a vast system of galaxies which we now know as the Local Supercluster.

1.2.1 Timescales

Various timescales distinguish between the systems we have described. The most fundamental timescale is the crossing time $t_c = 2r_h/\sigma\sqrt{3}$, where r_h is the median inter-particle separation (particles may be stars, galaxies or clusters) and σ is the velocity dispersion. A system of particles may not have had sufficient time to reach virial equilibrium if the crossing time is comparable to the Hubble time, $H_o^{-1} \approx 10^{10}$ yrs. The relaxation time of a system is the time required for random interactions to change a typical velocity vector of the order of itself. For a uniform distribution of particles the relaxation time $t_r = (0.1N/lnN)t_c$ where N is the number of particles. If the relaxation time of an object is much larger than its age then it evolves as a collisionless system in which the constituent particles move under the mean potential generated by all the other particles. The evolution of such a system can be followed with the collisionless Boltzmann equation. Another fundamental timescale is the collapse time t_{coll}. This is the time needed for a homogeneous sphere to collapse out of the Hubble flow and is within a factor of two equal to the crossing time (Gunn and Gott 1972).

For gas to collapse and form stars its cooling timescale must be significantly shorter than the collapse time. We can make a rough calculation of the cooling time of a protogalactic cloud as follows (Rees and Ostriker 1977). The virial temperature of an ideal gas cloud can be calculated by equating the mean square velocity of an ideal gas with the virial theorem. This yields $T_{vir} = 0.13GM\mu m_p/(k_B r_h)$, where μ, m_p and k_B are the mean molecular weight of the gas, the proton mass and Boltzmann's constant respectively. A typical protogalactic cloud at its virial temperature of $T_{vir} \sim 10^5$K will be fully ionised, therefore the main cooling mechanism is the emission of continuum photons by the bremsstrahlung process. Such clouds would be optically thin and would radiate energy per electron E at a rate proportional to the electron density, $\dot{E}_{cool} = \Lambda(T)n_e$, where n_e is the electron density and Λ is the cooling function for the particular mixture of primordial gas.

For temperatures in the range $5 \times 10^5 \mathrm{K} \lesssim T \lesssim 2 \times 10^7 \mathrm{K}$ then $\Lambda \sim 10^{-39} \mathrm{Js}^{-1}\mathrm{m}^3$ for a hydrogen-helium mixture. The energy per electron is $E = \frac{3}{2} x k_B T$ where x is the number of particles per electron and is ~ 2. Assuming no heat sources a protogalactic cloud at the virial temperature would have cooled on a timescale $t_{cool} \equiv (E/\dot{E}_{cool}) \sim 6.3 \times 10^5 (r_h/10\mathrm{kpc})\mathrm{yrs}$.

Object	N	r_h /kpc	σ /kms^{-1}	t_c /H_o^{-1}	t_r /t_c	t_{cool} /t_c	$\frac{M}{L}$ /$\frac{M_\odot}{L_\odot}$
Globular clusters	10^5	0.01	7	10^{-4}	900	10^{-7}	2
Galaxies	10^{11}	10	220	10^{-3}	10^8	0.02	12
Local Group	2	700	120	0.7	0.29	0.4	100
Groups	6	700	160	0.5	0.33	0.6	200
Clusters	1000	2000	800	0.3	15	8.7	400
Superclusters	100	20000	2000(!)	1.2	2.2	213	1600(?)

Table 1.1 Dynamical timescales and M/L ratios for systems of stars, galaxies and clusters.

Table 1.1 summarises these timescales for various systems of stars, galaxies and clusters of galaxies. The parameter values come from either work in this thesis or from Binney and Tremaine (1987). The crossing times of groups and clusters of galaxies are very similar and one would therefore expect little difference between their dynamical properties. The crossing times are comparable to the Hubble time which makes it uncertain as to whether these systems have had time to virialise within the age of the Universe. The relaxation times for groups of galaxies is $\sim 0.1 H_o^{-1}$, therefore gravitational encounters could play an important role in their evolution. Galaxies, however, have very small crossing times and extremely large relaxation times. These objects are therefore virialised but particle-particle encounters will never play an important part in their evolution, thus the dynamics of a galaxy will be described by the mean potential well, within which it is residing.

Hickson (1982) has compiled a catalogue of groups with extremely high over-density. The median separation of the galaxies within these groups is $\sim 40 h^{-1}$ kpc

with a velocity dispersion $\sigma \sim 200\text{km s}^{-1}$. The apparent dynamical timescales are of the order $2 \sim 4\%$ of the age of the Universe and we could expect the group to merge within a fraction of the Hubble time. This leads to a paradox as to their descendants and precursors. Barnes (1985) demonstrated that the apparent lifetime of a dense group is on average only a fraction of its real lifetime which on average was $\sim 20\%$ of the age of the Universe. This longer lifetime and the rarity of these objects make the paradox disappear.

There is a distinct dynamical difference between galaxies and galaxy associations. The fact that galaxies stand out in surface brightness above the background clustering testifies to the different, non-gravitational physics that must have played a role in their formation. The surface brightness of an elliptical galaxy within its half light radius lies well above the extrapolation of $\xi(r)$ from larger scales. Dissipational processes provide the physical means which distinguish between a galaxy and a group or cluster. For galactic size objects $T_{cool} \ll T_{coll}$, so dissipation can be expected to dominate their dynamics. Proto-systems larger than a small group have $T_{cool} \gtrsim T_{coll}$.

Spherical accretion of matter onto an accreting seed provides an approximate model for investigating the dynamics of protosystems in an expanding Universe. The turnaround radius is the radius which a perturbation of mass M reaches before collapsing (Peebles 1980):

$$r_{ta} = 0.343 M/{\rho_b}^{2/3} \tag{1.3}$$

where $\rho_b = 3H_o^2/8\pi G$ for a flat Universe. Groups, clusters and superclusters have a turnaround radius of $\approx 2r_h$. (Superclusters are certainly not virialised since they are still expanding with the Hubble flow.) For virialised systems this is about the reduction in radius which one expects in order that the potential energy liberated in the collapse be turned into the binding energy of the system. A galaxy of mass $10^{12}M_\odot$ has $r_{ta} \sim 500$ kpc, much larger than its present radius, which indicates

substantial shrinkage due to dissipation. A supercluster of total mass in the range $10^{17} \sim 10^{18} M_{\odot}$ has $r_{ta} \approx 25 \sim 50$ Mpc.

1.3 AIMS AND OBJECTIVES

The distinctions between galaxies and galaxy associations can be understood in terms of the relative cooling times and collapse times, but the distinction between galaxy groups and clusters of galaxies is not clear. One fundamental property of these systems is their relative abundances. Hierarchical clustering theories make definite predictions for the mass function or multiplicity function of "bound objects". Prior to the advent of large redshift surveys, two dimensional catalogues of galaxies were used to determine the relative numbers of projected density enhancements in the galaxy distribution (Turner and Gott 1976, Bahcall 1979). We now have complete catalogues of galaxies with quasi 3-dimensional information available which allows us to eliminate most of the projection effects intrinsic to the 2-dimensional catalogues. Our aim is to optimise an algorithm for selecting groups of galaxies from galaxy redshift surveys which can be used to create a catalogue of groups at some specified overdensity. We shall use artificial galaxy catalogues with full positional information to select the optimum parameters for minimising contamination of our group sample. Our objective is to study the continuity of properties between single galaxies and rich clusters.

A problem with groups of galaxies selected from a magnitude limited catalogue is that a group with total luminosity L at a distance s, will appear brighter than the same group at distance $2s$. The reason for this is that faint galaxies within the more distant group have fallen below the detection threshold of the catalogue. We shall attempt to correct the "hidden" luminosity in each group assuming a universal luminosity function (Schechter 1976, Felten 1977, Dressler, 1978, Tammann *et al.* 1979, Bingelli 1987), hopefully obtaining an effectively volume limited sample of groups. Adopting this procedure will allow us to calculate the correlation

properties between different luminosity ranges. Kaiser (1986) has suggested that rich clusters owe their apparent clustering properties to the strong spatial correlations of peaks in a Gaussian random field. Groups of galaxies should therefore show clustering properties inbetween to those of galaxies and rich clusters.

Galaxies have existed for most of their lifetime within small associations or groups. Before we can understand galaxy formation we must understand the environments within which galaxies are formed and have spent most of their subsequent time. The proximity of galaxies within the Local Group has allowed their true positions and radial peculiar motions to be measured. This allows us to compare the Local Group in detail with predictions from gravitational clustering theories. The kinematics of these galaxies have been used by many authors to constrain the mass distribution within the Local Group. We shall attempt to further our understanding of the mass distribution within the Local Group and to study in detail the formation of a small group of galaxies within the CDM model.

Our understanding of the Universe advances at about the same rate as the instrumentation with which we observe the Universe. Prior to the last decade, most of our understanding of the galaxy distribution came from analyses of the two-dimensional galaxy maps such as the Zwicky (Zwicky *et al.* 1961-1968) or Lick catalogues (Shane *et al.* 1967). With the advent of large telescopes, multifibre spectrographs and data storage systems capable of handling the increased amount of information, large complete redshift surveys have recently become accessible with many additional surveys currently in progress. Obtaining complete redshift surveys to depths $z \sim 0.1$ would still require many thousands of galaxy redshifts at considerable expense. The random sampling procedure, adopted by the Durham deep survey team in the early 1980's and proposed quantitatively by Kaiser (1984), allows maximum information on the large scale galaxy distribution at the expense of a loss of information on small scales. Lawrence *et al.* (1991) adopted this

procedure to create a complete, all sky redshift survey of IRAS galaxies (the "QDOT" survey).

Structure in the Universe is often described as filamentary, spongy, foamy etc. Observations of large scale structure sometimes reach the front pages of the newspapers. This has the unfortunate consequence of turning "Great Attractors" and "Great Walls" into standard measures with which we gauge the success or failure of a particular cosmogonic model. The visual appearance of the galaxy distribution helps us appreciate the fantastic range of structures that are present in the Universe, but it does little to convince most theoreticians that a particular model should or should not be discarded. The two point galaxy correlation function is the statistic most often used to quantify the galaxy distribution. Ignoring the complex problem of normalising the correlation function and weighting the galaxy density, interpreting the results can prove to be difficult. The pioneering work of Hamilton *et al.* (1986) and Gott *et al.* (1986) led to a method by which the overall appearance of the large scale galaxy distribution could be quantified and related to the initial conditions of the fluctuating density field. In Chapter 4 we shall study the topology of the QDOT redshift survey with the aims of measuring the initial mass fluctuation spectrum and testing the hypothesis that structure in the Universe grew from Gaussian initial fluctuations.

REFERENCES

Abell, G. 1958, *Ap.J. Suppl.*, **3**, 211.

Bahcall, N.A. 1979, *Ap.J.*, **232**, 689.

Bahcall, N. and Soneira, R. 1983, *Ap.J.*, **270**, 20.

Bardeen, J.M. 1986, *Inner space/Outer space. University of Chicago press.*

Bardeen, J.M., Bond, J.R., Kaiser, N., and Szalay, A.S., 1986, *Ap.J.*, **304**, 15.

Barnes, J. 1985, *Mon.Not.astr.Soc.*, **215**, 517.

Bingelli, B., Sandage, A. and Tammann, G.A. 1985, *A.J.*, **90**, 1681.

Binney, J. 1977, *Ap.J.*, **215**, 483.

Binney, J. and Tremaine, S. 1987, *Galactic Dynamics. Princeton University press.*

Bond, J.R., Efstathiou, G. and Silk, J.I. 1980, *Phys.Rev.lett.*, **45**, 1980.

Bond, J.R. and Efstathiou, G. 1983, *Ap.J.Lett.*, **285**, L45.

Bond, J.R. and Szalay, A.S., 1983, *Ap.J.*, **274**, 443.

Bond, J.R., Centrella, A.S., Szalay, A.S. and Wilson, J.R. 1984, *Mon.Not.R.astr.Soc.*, **210**, 515.

Blumenthal, G.R., Faber, S.M., Primack, J.R. and Rees, M.J. 1984, *Nature*, **311**, 517.

Davis, M. and Huchra, J. 1982, *Ap.J.*, **254**, 437.

Davis, M., Efstathiou,G., Frenk, C.S. and White, S.D.M. 1985, *Ap.J.*, **292**, 371 (DEFW).

Dekel, A., Blumenthal, G.R., Primack, J.R. and Olivier, S. 1989, *Ap.J letters*, **338**, L5.

Doroshkevich, A.G., Khlopov, M. Yu., Sunyaev, R.A., Szalay, A.S. and Zeldovich, Ya.B. 1980, *Ann.N.Y.acad.sci.*, **375**, 32.

Dressler, A. 1978, *Ap.J.*, **223**, 765.

Efstathiou, G., Kaiser, N., Saunders, W., Lawrence, A., Rowan-Robinson, M., Ellis, R.S.E and Frenk, C.S. 1990 *Mon.Not.R.astr.Soc.* **247**, 10p.

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

Frenk, C.S., White, S.D.M., Efstathiou, G. and Davis, M. 1988, *Ap.J.*, **327**, 507.

Frenk, C.S. 1989, in **The Epoch of Galaxy Formation**, eds. C.S. Frenk, R.S. Ellis, T. Shanks, A. Heavens and J. Peacock, (Dordrecht: Kluwer Acad. Publ.).

Frenk, C.S., White, S.D.M., Efstathiou, G. and Davis, M. 1990, *Ap.J.*, **391**, 2.

Gott, J.R., Mellot, A.L. and Dickinson, M. (GMD) 1986, *Ap.J.*, **306**, 341.

Gunn, J. and Gott, J. 1972, Ap.J., **176**, 1.

Gurvatov, S.N., and Saicher, A.I. and Shandarin, S.F. 1989, *Mon.Not.R.astr.Soc.*, **236**, 385.

Guth, A. 1981, *Phys.Rev.D*, **23**, 347.

Hamilton, A.J.S., Gott, J.R., and Weinberg, D. 1986, *Ap.J.*, **309**, 1.

Hickson, P. 1982, *Ap.J.*, **255**, 382.

Hogan, C.J. and Kaiser, N. 1983, *Ap.J.*, **274**, 7.

Kaiser, N. 1984, *Ap.J.*, **284**, L9.

Kaiser, N. 1986. *Mon.Not.R.astr.Soc.* **219**, 795.

Kaiser, N. and Lahav, O. 1989, *Mon.Not.R.astr.Soc.*, **237**, 129.

Kirshner, R.P., Oemler, A., Schechter, P.L. and Shectman, S.A. 1983, *A.J.*, **88**, 1285.

Lawrence *et al.* , 1991, in preparation.

Lyubimov, V.A., Novikov, E.G., Nozik, V.Z., Tretyakov, E.F. and Kozik, V.S. 1980, *Phy.Lett.*, **B94**, 266.

Maddox, S.J., Efstathiou, G., Sutherland, W. and Loveday, J. 1990, *Mon.Not.R.astr.Soc.*, **242**, 43p.

Meszaros, P. 1974, *Astr.Ap.*, **37**, 225.

Nolthenius, R. and White, S.D.M. 1987, *Mon.Not.R.astr.Soc.*, **225**, 505.

Ostriker, J.P. and Cowie, L.L. 1981, Ap.J.Lett., **243**, L127.

Peebles, P.J.E. and Dicke, R.H. 1968, *Ap.J.*, **154**, 891.

Peebles, P.J.E. 1980, *The Large Scale Structure of the Universe.*

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

Rees, M.J. and Ostriker, J.P. 1977, *Mon.Not.R.astr.Soc.*, **179**, 541.

Saunders, W., Frenk, C.S.F, Rowan-Robinson, M., Efstathiou, G., Lawrence, A., Kaiser, N., Ellis, R.S.E, Crawford, J., Xiao-Yang Xia and Ian Parry 1991, *Nature* **349**

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

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

Silk, J.I. 1968, *Ap.J.*, **151**, 459.

Silk, J.I. 1977, *Ap.J.*, **211**, 638.

Shapiro, P.R., Struck-Marcell, C. and Melott, A.L. 1983, *Ap.J.*, **275**, 413.

Sutherland, W. 1988, *Mon.Not.R.astr.Soc.*, **234**, 159.

Tammann, G.A., Yahil, A. and Sandage, A. 1979, *Ap.J.*, **224**, 775.

Turner, G. and Gott, J.R. 1976, *Ap.J.Supp.*, **32**, 409.

Vilenkin, A. 1986, *Inner space/Outer space. University of Chicago press.*

White, S.D.M. and Rees, M.J. 1978, *Mon.Not.R.astr.Soc.*, **183**, 341.

White, S.D.M., Frenk, C.S. and Davis, M. 1983, *Ap.J.Lett.*, **271**, L1.

White, S.D.M., Davis, M. and Frenk, C.S. 1984, *Mon.Not.R.astr.Soc.*, **209**, 27p.

White, S.D.M. 1986, *Inner space/Outer space. University of Chicago press.*

White, S.D.M., Frenk, C.S., Davis, M. and Efstathiou, G. 1987b, *Ap.J.*, **313**, 505 (WFDE).

Zeldovich, Ya.B. 1970, *A.A.*, **5**, 84.

Zwicky, F. 1938, *Pub.Astron.Soc.Pacific*, **50**, 218.

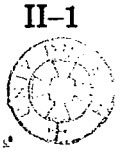
Zwicky, F., Herzog, E., Wild, P., Karpowicz. M. and Kowai, C.T. 1961-1968, *Catalogue of galaxies and clusters of galaxies. CIT*

2 THE LOCAL GROUP

2.1 INTRODUCTION

Evidence for the existence of a dark matter component has been available for over 50 years since Zwicky's study of the dynamics of the Coma cluster (Zwicky 1933). Since then the dynamical presence of dark matter has been observed on all scales from galaxies to galaxy clusters, although its nature, quantity and distribution remain a mystery. The distribution of dark matter and the dynamical role it plays have been inferred from a variety of techniques. Standard methods for constraining the amount of dark matter use the velocity dispersions of galaxies within groups and clusters and the infall of the Local Group towards the Virgo cluster (Davis and Peebles 1983). Dynamical studies of complete redshift surveys selected from the IRAS database have recently been used to probe the mass distribution on even larger scales by mapping the local peculiar gravity field and comparing the velocities predicted for the Local Group (Rowan-Robinson *et al.* 1990) and for a sample of 1000 galaxies (Kaiser *et al.* 1991) with observations. These methods give substantially higher values of Ω than the more "traditional" techniques and follows the trend of more dark matter on larger scales.

In this chapter we concentrate on the mass distribution around individual galaxies. This has often been studied using statistical studies of binary galaxies (Holmberg 1937, Turner 1976, White *et al.* 1983). The primary difficulties of this technique are our ignorance of the full orbital parameters and the selection of



an unbiased and uncontaminated sample of binaries. After adopting a model for the galactic masses and assuming some form for the anisotropy parameter, Jeans equation can be solved for the whole sample. The data are inconsistent with point mass potentials because no correlation between line of sight separation and velocity difference is observed. The potential of extended massive halos is consistent with the data although by adjusting the anisotropy parameter the halo masses can be reduced to zero. Assuming an isotropic velocity distribution the best available data give $v_c = 170\text{km s}^{-1}$, significantly smaller than the measured rotation curves of bright spiral galaxies $\sim 300\text{km s}^{-1}$ (White *et al.* 1983). Observations of hot X-ray gas surrounding elliptical galaxies provide additional evidence for large amounts of dark matter around these galaxies (Fabbiano 1989).

The rotation velocities of HI regions in spiral galaxies provide the strongest evidence that massive dark halos surround galaxies like our own. Reliable rotation curves have been obtained for over 70 spiral galaxies via radio measurements of the 21-cm emission line of neutral hydrogen (Faber *et al.* 1979). In almost all cases the rotation curve is flat or slowly rising out to the last measured point which is always well beyond the optical radius. *There is no example of a Keplerian fall off in any galaxy's rotation curve,* consequently there is no example of a spiral galaxy with a well determined total mass.

The first systematic attempt to determine the mass of the Milky Way was made by Kapteyn (1922). Kapteyn's model for the Universe had the Sun at the centre of an oblate spheroid of stars of total mass $\approx 10^{11} M_\odot$. Since Kapteyn our picture of the Galaxy and the Universe has changed somewhat, although coincidently some of the best estimates for the mass of the galaxy are not so different from Kapteyn's. The rotation velocity of the Milky Way is currently known out to twice the Sun's Galactocentric distance and is observed to have fairly constant amplitude of $\approx 220\text{km s}^{-1}$ (Fich *et al.* 1989). Independent mass estimates within this distance are obtained from studies of the kinematics of distant stars, the orbits

of globular clusters and the local stellar escape velocity. A comprehensive review of the techniques and results used to determine the mass of the inner Galaxy are given in Fich and Tremaine (1991).

Our situation within a small group of galaxies gives us a unique advantage in determining the mass distribution of the Galaxy's outer halo and in the Local Group itself, although the results have been the subject of some controversy since it was appreciated that there is more to this Universe than our Galaxy alone. The debate has been renewed recently with Little and Tremaine's (1987) analysis of the kinematics of the Galaxy's outer satellites. These authors found a very low mass for the Galaxy which disagreed with the value found from the timing argument (Kahn and Woltjer 1959). However, new observations of the radial velocities of several of these satellites gave a completely different result when the same analysis was performed (Zaritsky *et al.* 1989).

It would be a strange occurrence if the rotation curve of our Galaxy is dissimilar to that of M31 or any other spiral galaxy and does not continue with the same amplitude out to at least 30 kpc. Although there have been many attempts to determine the local distribution of mass using the kinematics of the galaxies within the Local Group, we still do not know the mass of our Galaxy or the mass of the Local Group to within a factor of order ten.

In this chapter we attempt to understand more fully the effects of different mass estimators and the implications of the assumptions that are implicit in their use. In Section 2.2 we briefly discuss the various techniques that have been used and the shortcomings of each method. Then we describe a model for the Local Group which was extracted from a numerical simulation of a Universe dominated by cold dark matter. The model serves several purposes: we can apply the standard mass estimators to the model in order to test their effectiveness at recovering the real mass of the system. We can also use the model to study the "natural"

formation of a small group of galaxies within the gravitational instability scenario. A comparison between the model and the observations provides a test of CDM on the smallest scales numerically accessible at present.

2.2 DETERMINING THE MASS DISTRIBUTION WITHIN THE LOCAL GROUP

2.2.1 The satellites of the Milky Way

The satellite dwarf galaxies and globular clusters of our Galaxy have been used by several authors to constrain the mass distribution of the outer galaxy (Hartwick and Sargent 1978, Frenk and White 1980, Frenk and White 1982, Lynden-Bell *et al.* 1983, Peterson 1985, Little and Tremaine 1987, Zaritsky *et al.* 1989). Table 2.1 lists all the known dwarf galaxies within 300kpc of our Galaxy and the globular clusters with distances greater than 50kpc (values taken from Zaritsky *et al*).

Object	r /kpc	v_{los}/kms^{-1}	μ /$(M_\odot/G)$
LMC+SMC	51	61	2.9
Ursa Minor	65	-88	7.8
Draco	75	-95	10.5
Sculptor	79	74	6.7
Carina	93	14	2.5
Fornax	140	-34	0.3
Leo II	220	16	0.9
Leo I	230	177	111.6
Pal 14	75	166	32.0
Eridanus	85	-138	25.1
Pal 3	95	-59	5.1
NGC 2419	98	-26	1.0
Pal 4	108	54	4.9
AM-1	117	-42	3.2

Table 2.1 Orbital parameters of satellite galaxies and globular clusters between 50 and 300 kpc from the Milky Way.

The main problems with any statistical analysis of these data are the small number of systems and the unknown anisotropy parameter β which describes the shape of the velocity dispersion tensor:

$$\beta(r) = \frac{\overline{v_r^2}(r)}{\overline{v_r^2}(r) + \overline{v_t^2}(r)}, \tag{2.1}$$

where $\overline{v_r^2}(r)$ is the mean square radial velocity at radius r and $\overline{v_t^2}(r) = \overline{v_\phi^2}(r) + \overline{v_\theta^2}(r)$ is the mean square velocity in the direction perpendicular to the radial vector. If all velocity vectors are equally probable then the orbits are described as isotropic and $\beta = 0.3$; satellites on radial orbits have $\beta = 1$ and circular orbits would have $\beta = 0$. N-body simulations of the formation of stellar systems through spherical collapse (Aguilar and Merrit 1990) suggest that β rises from zero near the centre, to unity in the outer parts. However, in more realistic scenarios of galaxy formation one expects large transverse motions to arise from gravitational torques during the formation process.

Various dynamical effects could easily bias the anisotropy parameter. There could be a process of "survival of the fittest" in which galaxies on highly eccentric orbits venture further into the potential well and become disrupted by strong tidal forces. The Magellanic Stream is a vivid example of tidal stripping due to the Milky Way. Murai and Fujimoto (1980) argue that the Stream could consist of gas torn from the Magellanic Clouds during the previous pericentric passage past the Galaxy at a distance $r_{min} \approx 50$ kpc. This process could leave a population of satellite galaxies on predominantly circular orbits (Lynden-Bell, Canon and Godwin 1983). Dynamical friction between the satellite galaxies and the halo provides another mechanism which would tend to decrease the anisotropy parameter (Bontekoe and van Albada 1987), although the process is less effective on low mass systems. A problem which has not been addressed in the literature is that the orbits of the most distant satellites of the Milky Way could be perturbed by M31.

We can illustrate some of these uncertainties via the techniques employed by Little and Tremaine (1987; hereafter LT). Assuming a model for the Galactic potential, LT evaluated the probability of observing the dimensionless quantity $\mu = rv_r^2/G$, where r is the satellite's Galactocentric distance, v_r is its radial velocity and G is the gravitational constant. This probability is expressed as $P(\mu|M_G)$ and can be calculated by applying Bayes theorem to the known quantities $P(M_G|\mu)$ and $P(M_G)$, where $P(M_G)$ is the *a priori* probability of finding a mass M_G. Once the eccentricity distribution of the satellites is specified this technique directly yields confidence intervals for the mass of the Galaxy. Standard models for the Galactic potential ϕ are an isothermal sphere characterised by its circular velocity v_c, $\phi = v_c^2 \ln(r)$, or a point mass potential $\phi = -GM/r$. For isotropic orbits around a point potential, LT found a mass for the Galaxy $M_G = 2.4 \times 10^{11} M_\odot$ with 90% confidence interval $[1.4 \times 10^{11}, 5.2 \times 10^{11}]$. The assumption of radial orbits decreases this mass estimate by about a factor of two. LT concluded that a Galactic potential modelled by an isothermal halo with $v_c = 220$km s^{-1} must fall off fairly sharply after 40kpc. Zaritsky *et al.* (1989) provided new observations of stellar velocities within the satellites Pal 14, Eridanus, Leo I and Leo II. The revised radial velocities of these satellites suggested a re-examination of LT's result, in particular, the original radial velocity of Leo I was over 100 km s^{-1} too low. Requiring Leo I to be bound to the Galaxy constrains the mass within 230 kpc to be at least $1.7 \times 10^{12} M_\odot$ which is well outside the original confidence interval. Zaritsky *et al.* performed the same analysis as LT with the new radial velocity data excluding Leo I and found a mass for M_G 100% larger than the original prediction and *adding Leo I to the data set further increased M_G by a factor of three.*

2.2.2 The Magellanic Stream

The Magellanic Stream is a trail of neutral hydrogen which stretches in a great circle across 110^o of arc, starting from the Magellanic Clouds and running

across the South Galactic Pole. The radial velocity of the Stream varies linearly with distance from the Clouds, reaching the very high value of -200km s^{-1} at its tip. The most plausible origin of the Stream is that tidal forces from the Milky Way stripped off the gas from the Clouds during a close passage past the Galaxy (Fujimoto and Sofue 1969). The beauty of the Magellanic Stream lies not only in its magnificent sweep across the night's sky, but in the way it can be used to probe the mass distribution in the outer halo of our Galaxy.

Murai and Fujimoto (1980) and Lin and Lynden-Bell (1982) modelled the formation of the Magellanic Stream using numerical simulations. The orbital parameters of the Clouds were adjusted until particles tidally stripped from the Clouds matched the run of radial velocity along the length of the Stream. The numerical experiments of these authors indicated that an extended halo out to at least 70 kpc with a circular velocity of 244 ± 12km s^{-1} is needed to fit the Stream kinematics (Lin and Lynden-Bell 1982). The simulations predict that the Magellanic clouds are presently at their pericentric distance from the Galaxy for the third time. The resulting transverse velocity of the clouds $V_t = 373$km s^{-1}, is within observational capabilities and if confirmed would provide some of the strongest evidence for an extensive dark halo surrounding our Galaxy. Measurements of V_t are underway using old plates of the Southern sky (Lin, private communication). These analyses assume that gravity is the only force acting on the Stream; there could be some hydrodynamic effects which are not included in these models.

2.2.3 Timing the Local Group

M31 and the Milky Way are falling towards each other with a relative velocity of 123km s^{-1} (Tully 1988; assuming a circular speed of 220km s^{-1} at the Sun's position). A plausible model for the evolution of the Local Group was proposed by Kahn and Woltjer (1959) and Peebles (1971) who approximated the Milky Way and M31 as an isolated system of two point masses with zero angular momentum. The

Hubble expansion of these two masses was slowed and eventually turned around by their mutual gravitational attraction. The equations which describe the orbit of a test particle emitted radially at $t = 0$ and attracted by a mass M at the origin were derived by Kahn and Woltjer and applied to the MW and M31 system in an attempt to measure the total mass of the Local Group (hereafter denoted M_{LG}). This technique is often called the "timing argument" since it is based on achieving the velocity reversal within a given time period. The constraint can be written in terms of the dimensionless parameters

$$\alpha = \frac{v_r t_o}{d}, \qquad \gamma = \frac{G M_{LG} t_o^2}{d^3}, \tag{2.2}$$

in the parametric form

$$\alpha = \frac{\sin\eta(\eta - \sin\eta)}{(1 - \cos\eta)^2}, \qquad \gamma = \frac{(\eta - \sin\eta)^2}{(1 - \cos\eta)^3}, \tag{2.3}$$

where η is the eccentric anomaly of the orbit, t_o is the age of the Universe, $d = 725 \pm 30$kpc is the separation (Pritchet and van den Bergh 1987) $v_r = -123$km s^{-1} is the radial velocity. Solving these equations for M_{LG} yields a range of masses between $5.6 \times 10^{12} M_\odot$ and $3.3 \times 10^{12} M_\odot$ for t_o in the range 10 to 20 Gyr.

Although the Kahn-Woltjer model is an idealised picture of the Local Group, any refinements of the model tend to increase the value obtained for M_{LG}. For example, it is unlikely that the two galaxies existed soon after the Big Bang as isolated objects with their present masses and that the radial velocities are purely radial. The analysis assumes there are no external torques on the Local Group from other nearby mass concentrations which could spin up the system. Let us consider each of these points in more detail.

A popular view of the formation process of groups of galaxies is through the gravitational instability picture whereby small perturbations in the early Universe accrete mass and merge hierarchically into larger systems. By relating the present

structure and kinematics of the Local Group to plausible models of the initial perturbations we can constrain the distribution of mass within it. Therefore, a more plausible approach based on the spherical infall model, is to assume that the initial mass distribution consists of two seed masses that grow at the rate expected for an isolated mass concentration in a critical density Universe, $M \propto t^{2/3}$. In the linear regime the comoving density surrounding each seed remains constant and the parameters α and γ can be calculated using linear perturbation theory giving a set of equations which can be solved for M_{LG} as a function of t_o, the time since the Big Bang (Peebles *et al.* 1989). The resulting M_{LG} is about 10% larger than the values given by the Kahn-Woltjer model. The estimates are similar since most of the peculiar velocity develops at late times, after the galaxies have grown to nearly their present masses. Peebles *et al.* (1989) followed numerically the growth of two seed masses within a flat Universe to check the analytic calculation described above. The simulation was evolved until the the parameter α equalled the observational value for the Galaxy and M31. The rotation velocity, $v_c^2 = GM/r$, of the halos agreed well with the observed values and the mass within the turnaround surface was similar to M_{LG} found using the timing argument.

The first conjecture concerning the transverse velocity of M31 was made by Einasto and Lynden-Bell (1982) who calculated that the motion of the Milky Way lies within $27^o \pm 15^o$ of the direction of M31. From this fact the total relative velocity of these galaxies was deduced to be 137 ± 20km s^{-1} and using the full orbital information the corresponding mass estimates for M_{LG} were raised by about 20%. Raychaudhury and Lynden-Bell (1989; hereafter RL) analysed the time dependence of the external torques on the barycentre of the Local Group using the observed distribution of galaxies within 12 Mpc. They conclude that the direction of M31 has swung through 19° since the Big Bang which implies a transverse velocity of $\sim$ 40km s^{-1}, slightly smaller than that predicted by Einasto and Lynden-Bell. Peebles (1989 and 1990) used a least action approach to trace

the orbits of the Local Group galaxies back in time. His best solution gave M31 a transverse velocity of $\sim$ 110 km s^{-1}.

It should be noted that these analyses would need to be modified if there are any large galaxies hidden behind the plane of our galaxy. McCall (1989) remeasured the Galactic extinction towards IC 342 and found that its revised luminosity made it nearly as bright as the Milky Way and calculated its distance to be 1.8 Mpc. RL found that IC 342 provided a large component of the tidal field, even at the distance they had adopted of 5.9 Mpc! IC 342 is at a galactic latitude 11° and at its new distance is only 1.1 Mpc behind M31 and 0.7 Mpc away from the giant elliptical Maffei 1. The "new" picture of the Local Group is essentially four large galaxies spaced fairly evenly along a line joining the Milky Way with M31. The close proximity of IC 342 to M31 questions one of the basic assumptions in the timing argument - that there are no nearby galaxies to perturb the system.

One final caveat in the Kahn-Woltjer model is the assumption that the mass of the Local Group lies essentially within two separate distributions surrounding the large spiral galaxies. For example, if the two galaxies are moving within a "sea" of dark matter or in a single large potential well, the required mass of the Local Group could be reduced by quite a large factor. One could argue that this scenario is unlikely when we observe small dwarf galaxies clumped around M31 and the Milky Way. From Table 2.1 we can see that 9 of the 14 known satellites of our Galaxy lie between 50 and 100kpc but only three lie between 100 and 200kpc, a volume eight times larger. However, one could always argue that the satellite galaxies could have originated in their present positions.

2.2.4 Timing Leo I

The latest observations of Leo I show that it has an extreme distance and velocity among the satellites of our Galaxy (see Table 2.1). Zaritsky *et al.* (1989)

applied the timing argument to Leo I to place a lower limit on the mass of the Milky Way. The adopted solution of equations (2.2) and (2.3) for the Milky Way and M31 have an eccentric anomaly which implies that the galaxies passed their maximum separation for the first time. Leo I however, has a large radial velocity away from our Galaxy, therefore a lower limit to M_G is found by assuming that the orbit is approaching maximum separation for the second time. The resulting mass that they find for the Milky Way is consistent with the Kahn-Woltjer model with approximately one third of M_{LG} residing in M_G. This analysis is strongly dependent on the assumptions that Leo I is bound to our galaxy, on a purely radial orbit and has only been under the gravitational influence of the Milky Way. If Leo I is unbound to the Local Group then it is most likely to have originated outside the Local Group, since at its observed radial velocity it would have travelled far beyond its current position in a Hubble time. An alternative to this scenario is to have Leo I ejected from the Local Group via a multi-body encounter with some other satellites of the Milky Way, although this would appear unlikely.

2.2.5 The local velocity field

A promising technique for evaluating the deceleration of the Hubble expansion in nearby field galaxies due to the presence of the gravitational potential well of the Local Group has been attempted by Sandage (1986 and 1987). For $H_o = 55, \Omega_o = 1$ and $t_o = 1.2 \times 10^{12}$yr he finds $M_{LG} = 3.5 \times 10^{12} M_\odot \pm (?)$. (Sandage does not quote errors on his estimate of M_{LG} due to the very large uncertainties in the distances to the nearby galaxies in his sample.)

2.2.6 Two models (at least!)

It is clear from these results that we do not have a clear cut picture of the mass distribution in the Local Group. At present there are two plausible models which can satisfy all the observational constraints.

(i) A high mass model: The lower limits to M_{LG} from the timing arguments indicate that the flat rotation curves of M31 and the Milky Way extend to at least 100 kpc. Additional support for this model comes from the simulations of the Magellanic Stream and the latest analysis of the Galaxy's satellites. If this model is accepted then at least 95% of the mass within the Local Group is dark.

(ii) A low mass model: One has to assume that the Local Group is not bound *i.e.* the high relative velocity of the Milky Way and M31 is due to a chance encounter (Burbidge 1972) and Leo I is just "passing through" the Group. The timing arguments do not apply and the analysis of the Galaxy's satellites without Leo I are consistent with low mass or truncated halos.

Support for the low mass model is fairly tenuous: including Leo I in the LT analysis gives a mass for M_G inconsistent with the analysis without Leo I. This could imply that the statistical analysis is dominated by small number statistics or that Leo I is indeed unbound to the Galaxy. Shuter (1989) managed to reproduce some aspects of the Magellanic Stream with numerical simulations of the Magellanic Clouds with hyperbolic orbits. In his model, the Galaxy and M31 are given low masses consistent with the LT analysis without Leo I, and a close passage of the Clouds past M31 is responsible for the Magellanic Stream.

2.3 A MODEL FOR THE LOCAL GROUP

In the previous section we have summarised the techniques that have been used to probe the mass distribution of the outer halo and to determine the mass of our Galaxy and the mass of the Local Group. The uncertainties and assumptions implicit in these analyses have been highlighted. In this section we describe a model for the Local Group which has been extracted from a numerical simulation of a CDM Universe. The evolution of the model Local Group is followed from an early time with a high resolution N-body code until a time when the orbital

parameters of the system resemble those of the Local Group. This is the time with which we identify the present. The simulation is then continued until a merger takes place between the halos with which we identify M31 and the Milky Way. The main purposes of the model were outlined in the introduction: with full orbital information available at any time during the evolution of the system we can compare the mass found using the estimators described previously, with the real mass distribution in the model. This provides an efficient test of the sensitivity of the different estimators and the effects of the assumptions that are implicit in their use. Most galaxies belong to small associations typically like the Local Group. By studying in detail the formation of a small group of galaxies we can attempt to further understand the environments where galaxies spend most of their time.

2.3.1 The initial conditions and orbital parameters

The Local Group is essentially a binary system dominated by the two massive spiral galaxies M31 and the Milky Way. Each of these galaxies is observed to have a cloud of satellites and globular clusters in its proximity. The nearest large cluster to the Local Group is Virgo at $10h^{-1}$ Mpc although several small groups such as the Coma-Sculptor cloud lie nearby. To locate a candidate model for the Local Group we looked for binary systems in the N-body simulations of Frenk *et al.* (1985; hereafter FWED). These simulations were performed to study the growth of galactic sized halos within a Universe dominated by cold dark matter by following the evolution of structure within a cubic region of present length 14 Mpc, from a redshift of 2.5 to the "present day". The present is identified in the P^3M simulation by matching the amplitude of the pairwise galaxy peculiar velocities to observations. This choice is equivalent to a normalisation of the CDM power spectrum, corresponding to a biasing parameter $b = 2.5$ (Frenk *et al.* 1990). Each of 5 simulations contained 32768 particles which represented a total mass of $1.9 \times 10^{14} M_{\odot}$ in each simulation. Within such a region the abundance of halos

was found to match the expected numbers of galaxies typically observed in such a volume.

Figure 2.1 shows the evolution of one of these simulations in a sequence of plots which show the projected particle positions in co-moving coordinates from a redshift $z = 2.5$ to the present. The most striking feature of these plots is the speed at which structures form within the model. A likely candidate for the model Local Group is indicated by an arrow on Figure 2.1 with projected centre of mass (0.15,0.35). The system is a pair of massive halos with several smaller subclumps nearby. We used a "friends of friends" program with a linking length 0.2 times the mean interparticle separation to extract the model particles at the final time. In Figure 2.1 we also plot the particle positions in physical coordinates of the candidate model in two different projections. (Hereafter the original simulation is denoted by P^3M.) The candidate model consists of a small system of halos containing 3008 particles which have almost merged by the present. The system is dominated by two large halos which are taken to represent M31 and the Milky Way, hereafter written in italics when representing the model.

A preliminary study of the candidate model from the P^3M run showed that at a time $t_o = 10.2 \times 10^9$yrs the infall velocity and separation of the two halos are similar to those of the Local Group. At this time $v_r = -99$km s^{-1} and $d = 770$kpc which yields $\alpha = -1.3$. This should be compared with the observations which give $\alpha = -1.8 \pm 0.2$, with errors on α due to the uncertainty in v_r and d. The turnaround radius for a mass $\sim 10^{13} M_\odot$ can be calculated from equation (1.1) which yields $r_{ta} = 1.8$ Mpc for $H_o = 50$h^{-1} Mpc km s^{-1}, whereas the actual value is ~ 1.2 Mpc. This order of magnitude agreement is all that can be expected since the turnaround surface of the Local Group is very elongated.

The region surrounding the binary system was excised from the P^3M simulation at several early epochs so that its evolution could be followed with a high

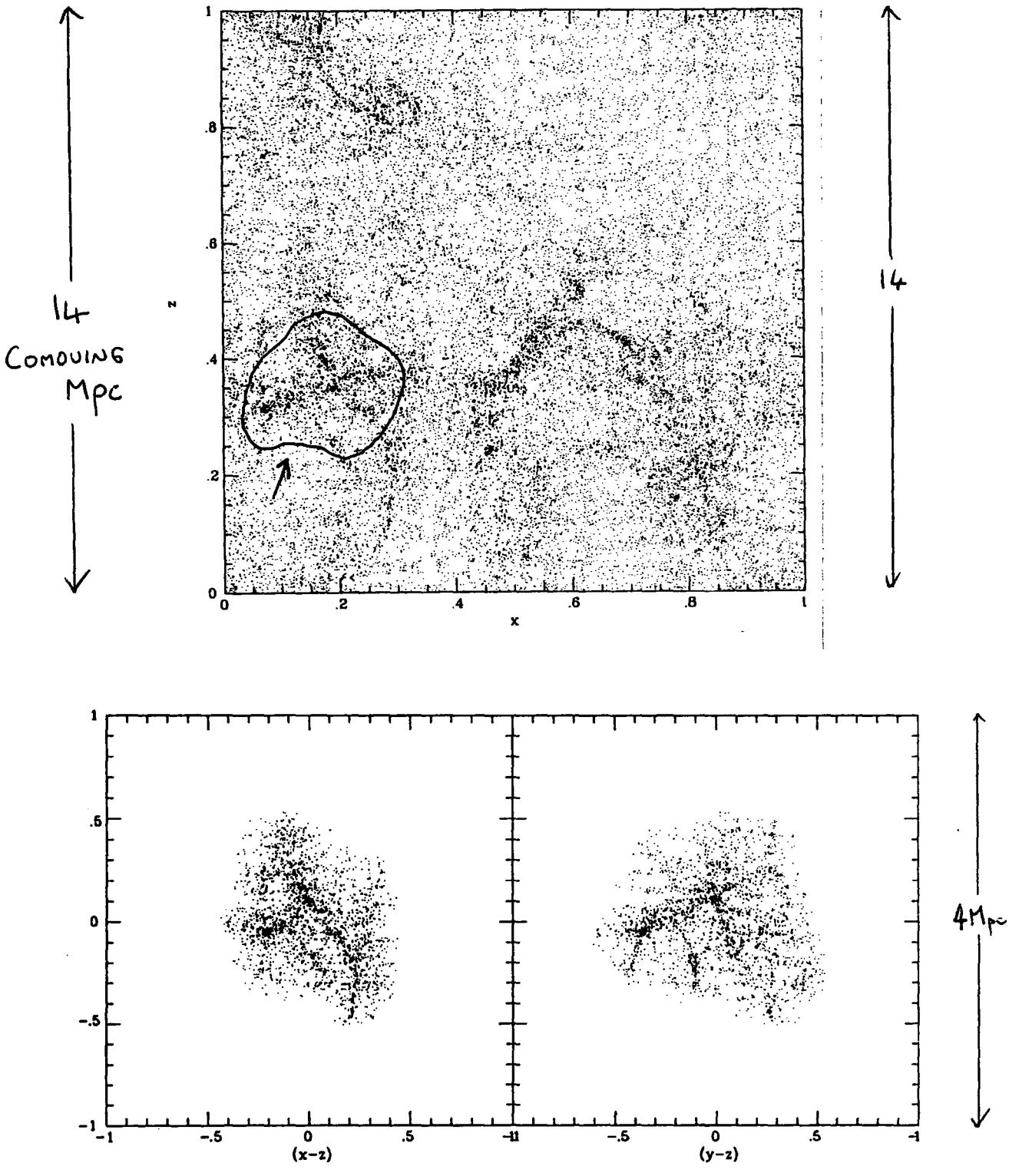
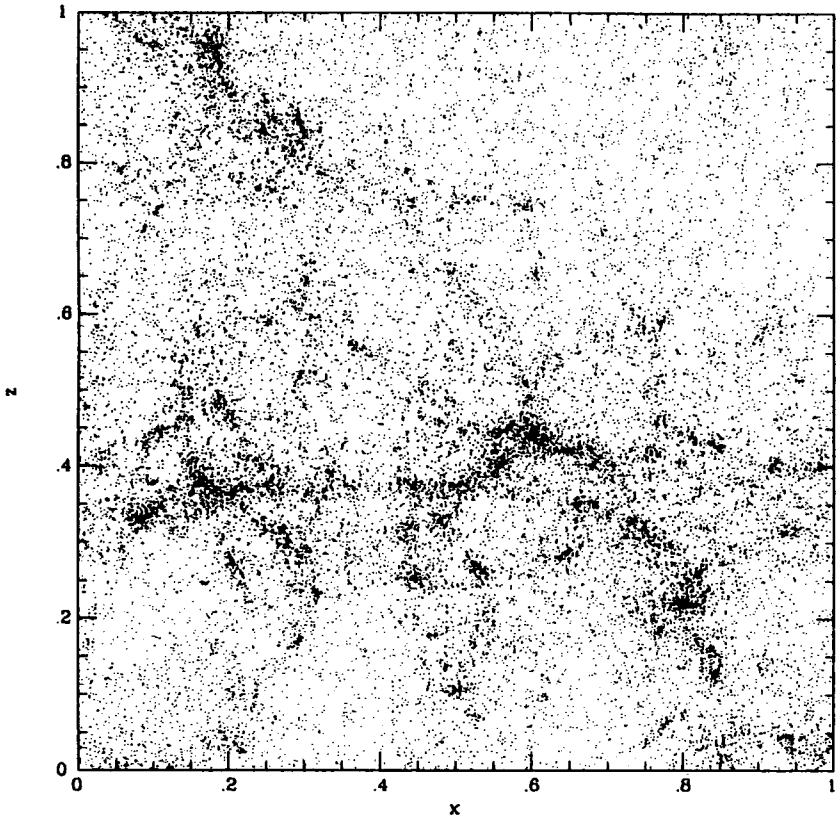
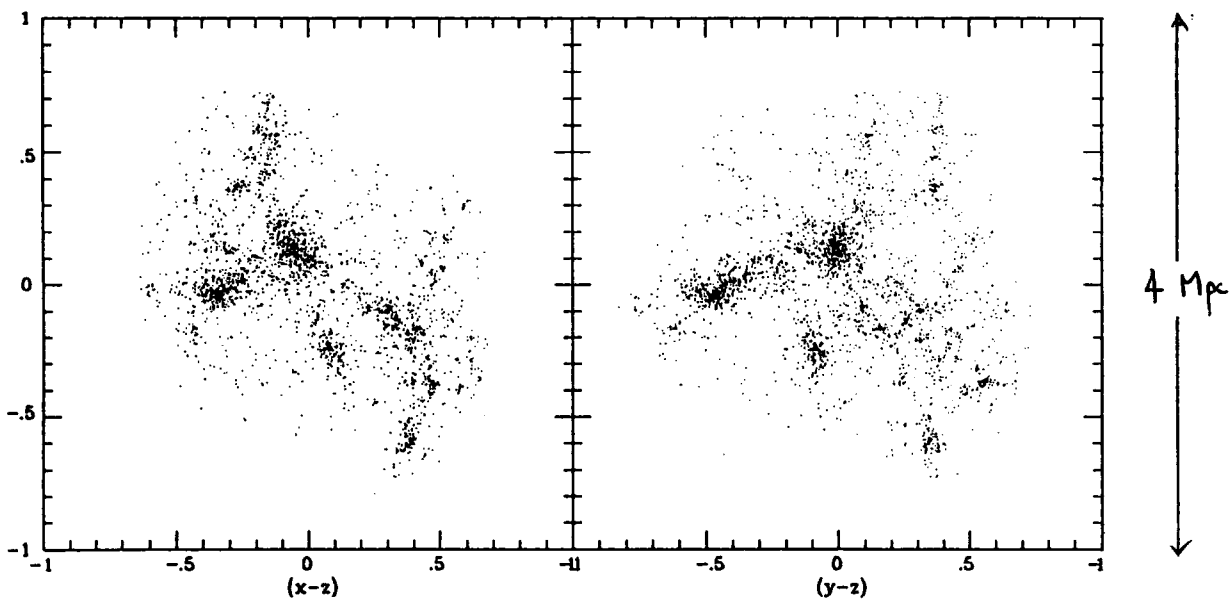
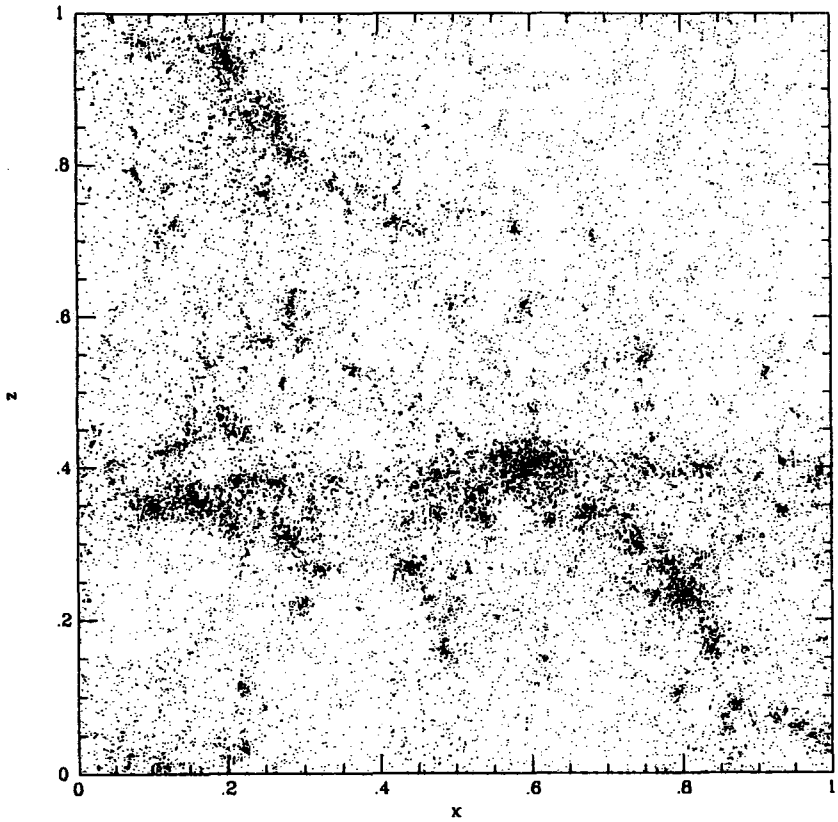
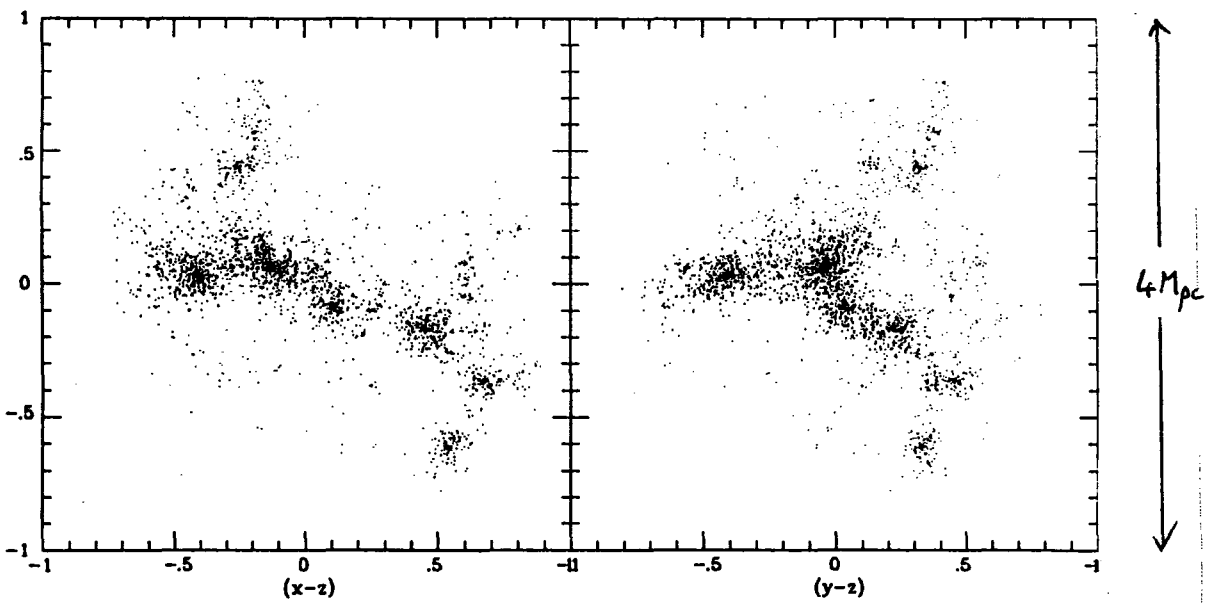


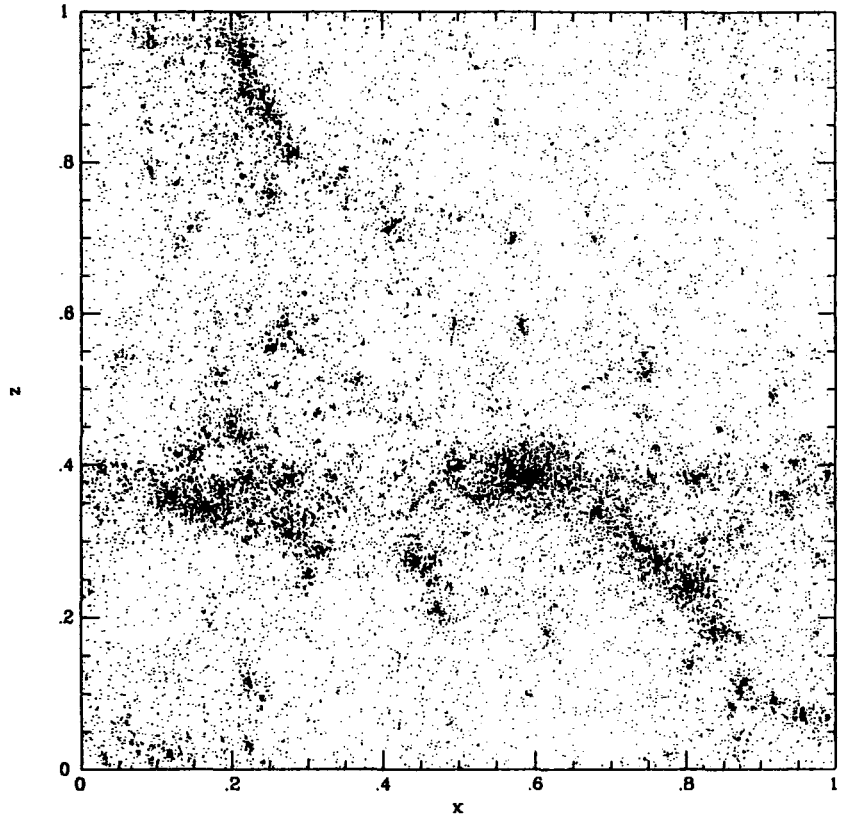
Figure 2.1 Projected comoving positions of the 32768 particles within the P^3M simulation at the expansion factors $a = 2.0, 3.5, 4.9, 5.9, 6.9$ (corresponding to redshifts $z = 2.5, 1.0, 0.4, 0.2, 0.0$). The lower panels show the candidate Local Group model plotted in physical coordinates at the same epoch.

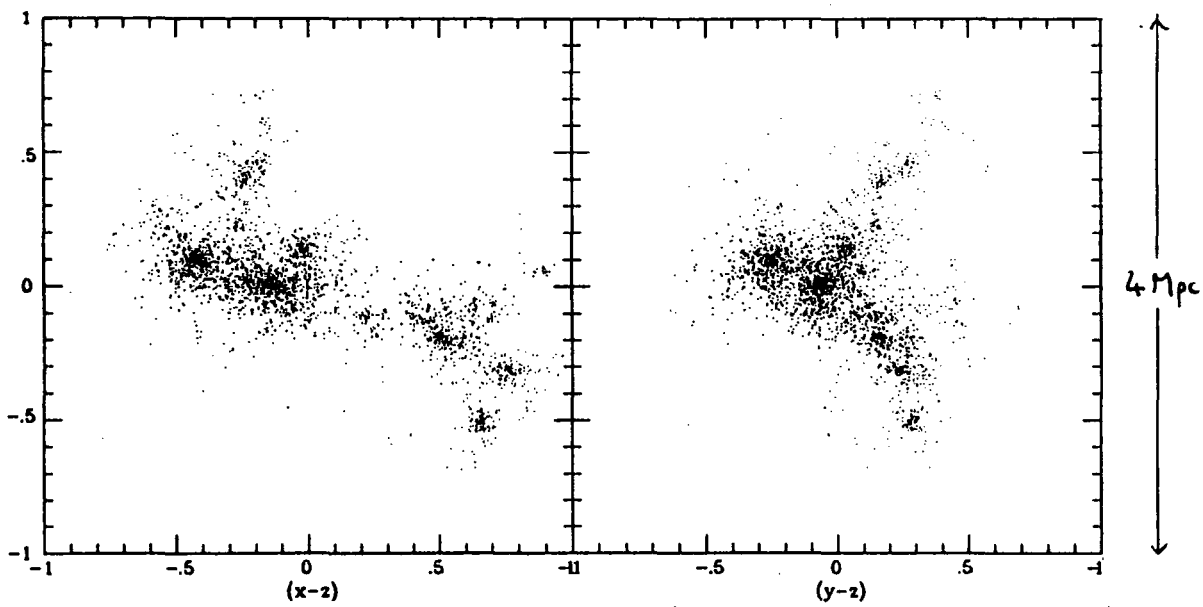

z
x
0
.2
.4
.6
.8
1

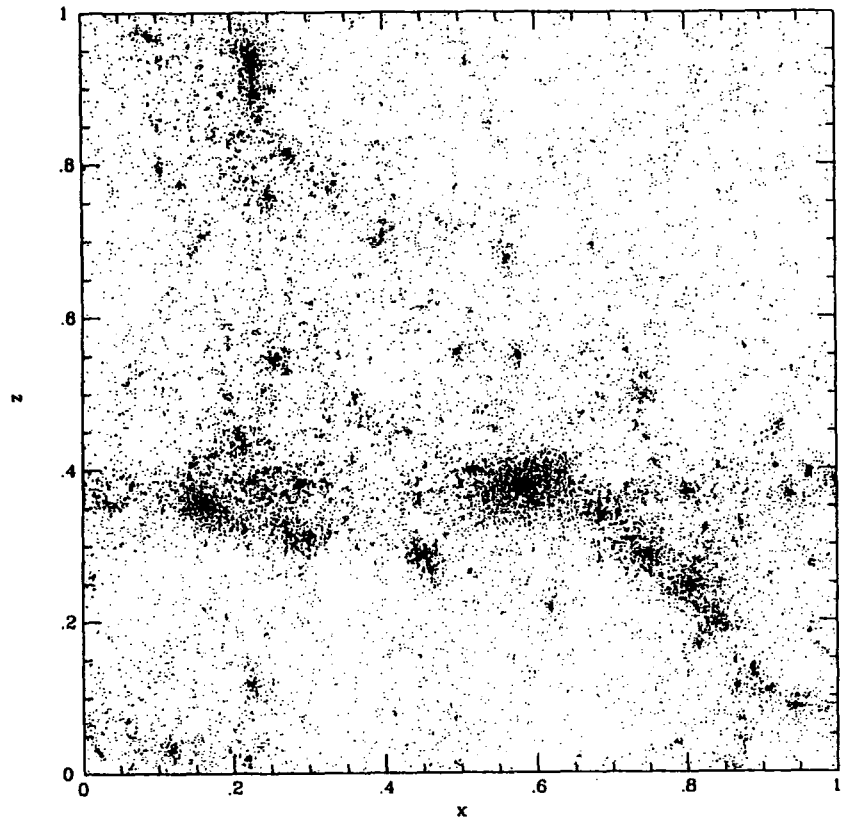

1
.5
0
-.5
-1
(x-z)
(y-z)
4 Mpc

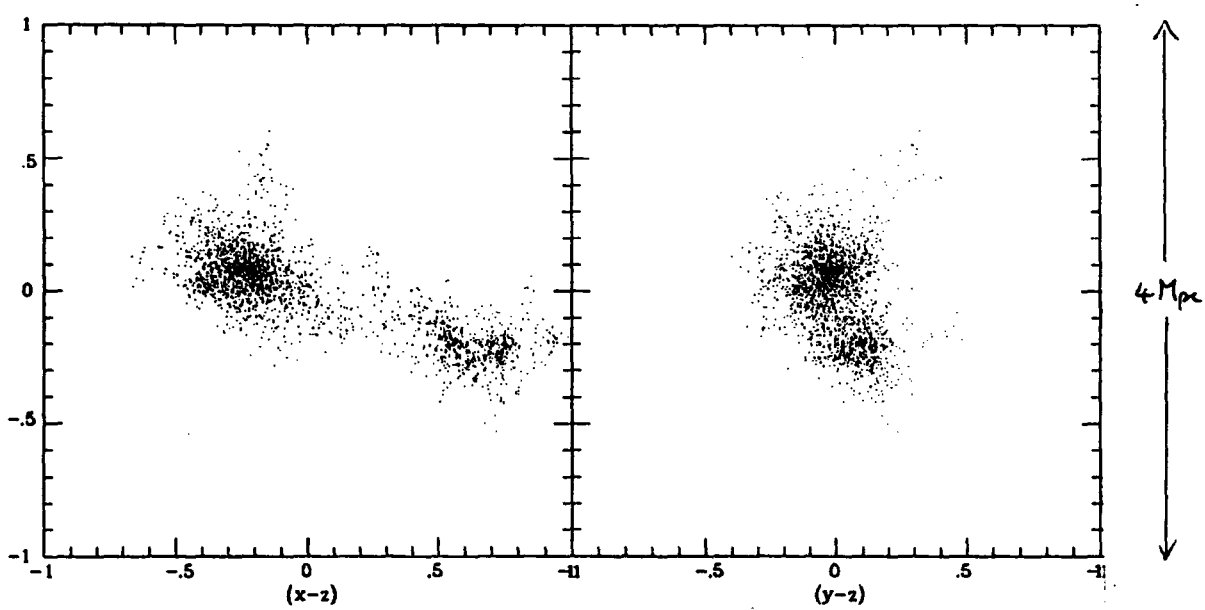

1
.8
.6
.4
.2
0
z
0
.2
.4
.6
.8
1
x


1
.5
0
-.5
-1
-1
-.5
0
.5
1
(x-z)
-1
-.5
0
.5
1
(y-z)
4Mpc


z
x
0
.2
.4
.6
.8
1


(x-z)
(y-z)
4 Mpc
1
.5
0
-.5
-1


z
x
0
.2
.4
.6
.8
1


(x-z)
(y-z)
4 Mpc

resolution N-body code. The purpose of this is twofold. The resolution in the P^3M run, determined from the force softening, was fixed in comoving coordinates to be 14 kpc in present day units. By evolving only the 3008 Local Group particles from an early time we can achieve higher resolution and with many output times the orbital and halo parameters can be studied in more detail.

The TREE code developed by Barnes (1986) provides the ideal N-body program for this type of calculation. The particles are embedded into a "tree structure" in which different levels of the tree represent the clustering hierarchy. The bottom level of the tree contains information on individual particles. Intermediate levels store information on larger volumes of space containing several particles until we reach the top level of the tree which contains the whole simulation. The force calculation for each particle proceeds as follows. The tree structure is searched through and the distances and masses are stored for all cells whose length/distance are within some specified criteria. Nearby cells may contain single particles but cells further away may contain a large number of particles.

The problem with extracting a cluster of particles at an early time from a larger simulation is that the external tidal field is suddenly turned off. By comparing the dynamics of the model with the results from the P^3M run we can assess how large this effect is. We ran several experiments with initial conditions taken from the first three output times of the P^3M run at $z = 2.5, z = 1.0$ and $z = 0.4$. The timestep was set to 10^6 years which conserved energy throughout the run to better than 0.1%. The particle positions and velocities were output at the same epochs as the P^3M run so that their distributions could be compared at the same times. The models excised at $z = 1$ and $z = 0.4$ looked very similar to the P^3M run at late times, however the model excised at $z = 2.5$ looked quite different from the P^3M output at the final time. The two main halos had less orbital angular momentum and merged rather more rapidly than in the other runs. This demonstrates that a large fraction of the angular momentum is generated

from the external tidal field at an early time. The results below correspond to the model Local Group excised from the P^3M run at $z = 1.0$ which gave a resolution of 7 kpc. Figure 2.2 shows the evolution of the model from this time until a after a merger has taken place between *M31* and the *Milky Way*.

We identify the *present* at a time $t_{P^3M} = 12.0 \times 10^9$yrs after the Big Bang when $\alpha = -1.7$, which agrees very well with the observed value. The particular system we have chosen best resembles the Local Group at a slightly earlier time than FWED identified with the present. This time corresponds to a redshift $z \approx 0.1$ and would imply a biasing parameter $b = 2.7$ rather than the value $b = 2.5$ advocated by FWED. This renormalisation is purely a matter of convenience; in a larger simulation we would expect to find fluctuations with similar dynamical properties, but slightly smaller amplitude than the system we have chosen to consider.

Table 2.2 shows the orbital properties of the two halos with which we identify M31 and the Milky Way at various epochs in their evolution. The radial and transverse velocities of *M31* listed in Table 2.2 are from a viewpoint centered on the *Milky Way*.

t_{BB} /10^9Gyrs	d /kpc	v_r /kms^{-1}	v_t /kms^{-1}	M_{LG} /$10^{12}M_\odot$ Using v_r only	M_{LG} /$10^{12}M_\odot$ Using v_r and v_t	M_{LG} /$10^{12}M_\odot$ True
2.3	912	[illegible]	20	....	....	6.9
5.2	1132	-15	37	....	....	8.3
7.0	1086	-50	44	11.4	12.0	9.1
8.8	969	-105	83	10.0	11.6	9.4
10.6	844	-110	156	6.4	10.0	9.6
13.0	735	-121	177	4.5	9.1	9.7
13.6	654	-130	196	3.9	8.8	9.8
15.1	304	-202	379	2.3	8.6	9.8
16.6	244	45	52	....	....	....

Table 2.2 Orbital parameters of the halo representing M31 and values of M_{LG} obtained from the timing argument.

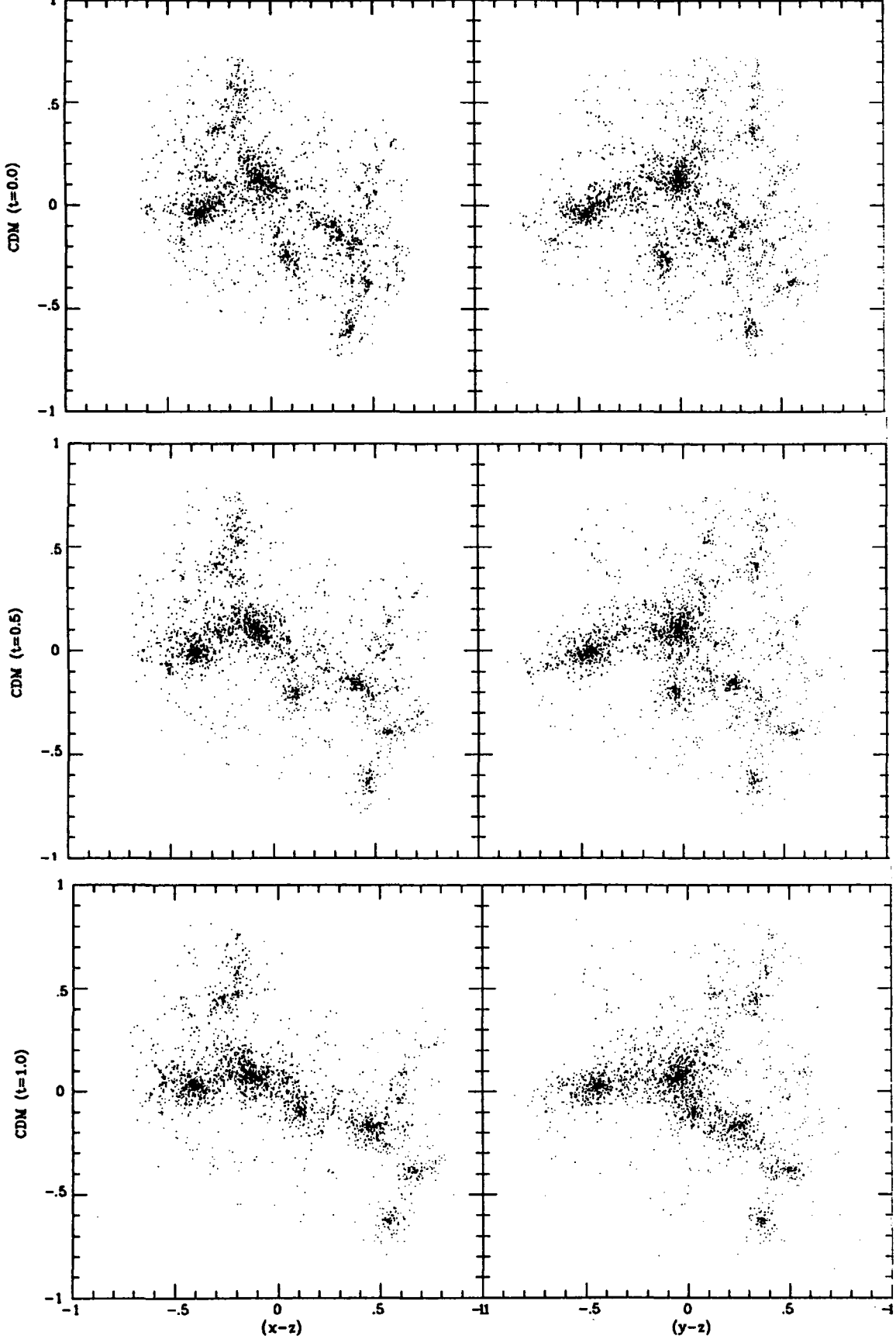
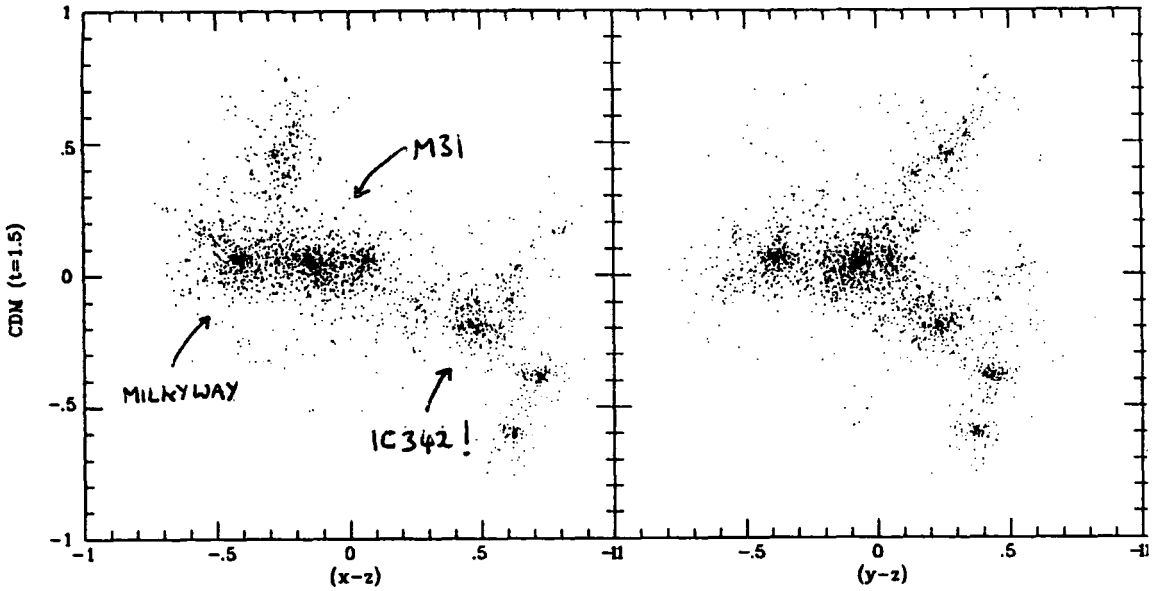
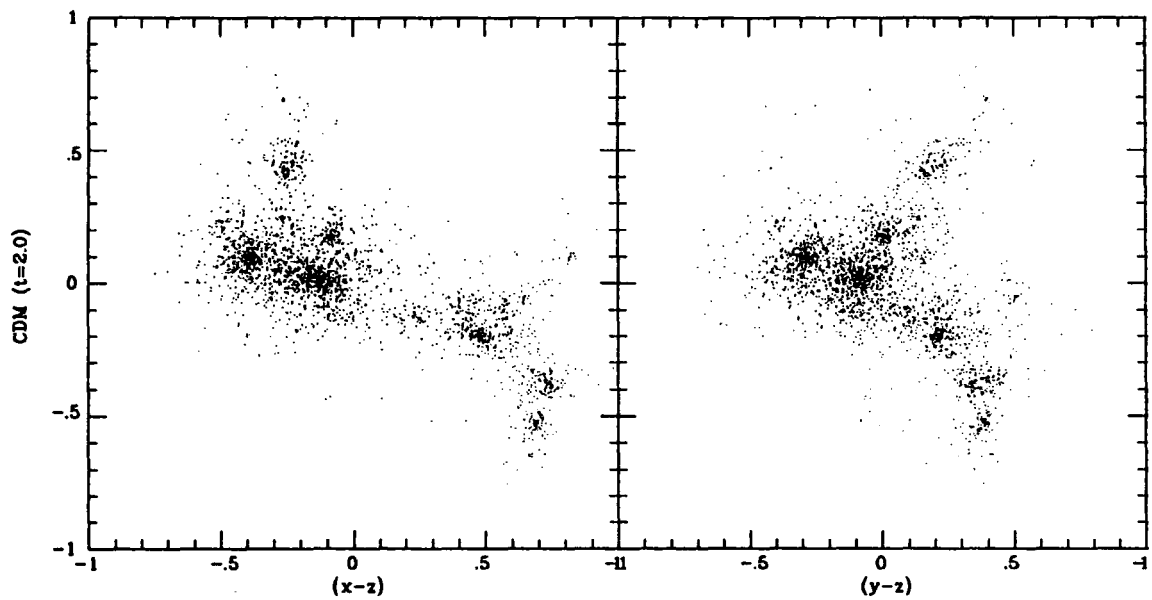


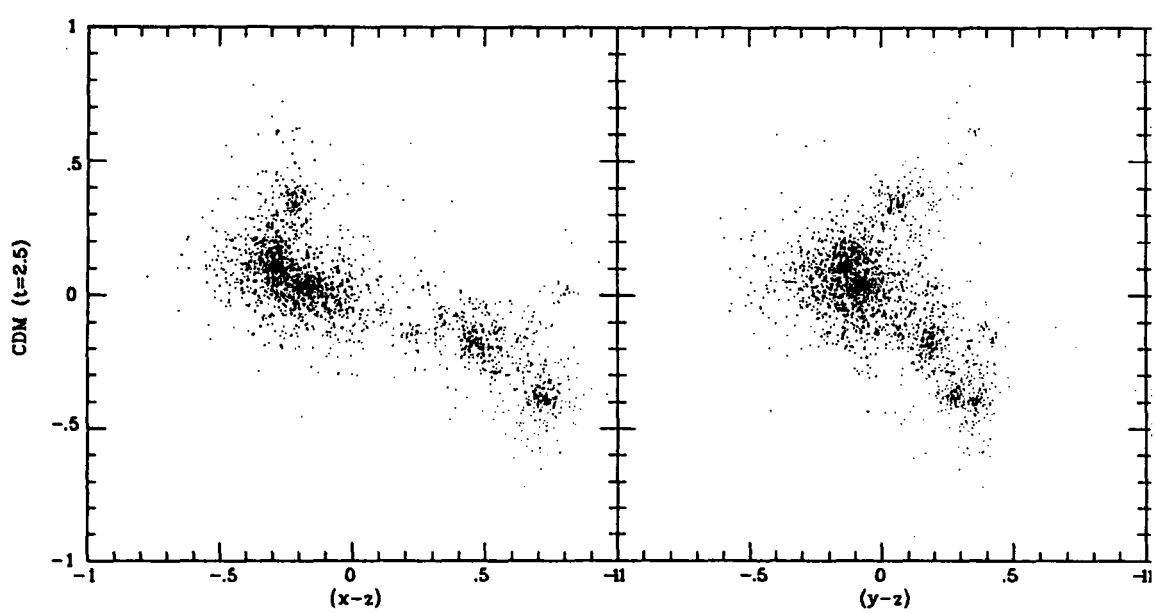
Figure 2.2 The particle positions from the Local Group in three projections in physical coordinates. The sequence starts at $z = 2.5$ and continues until after a merger has taken place between the *Milky Way* and *M31*. The box length in these plots is 4 Mpc.

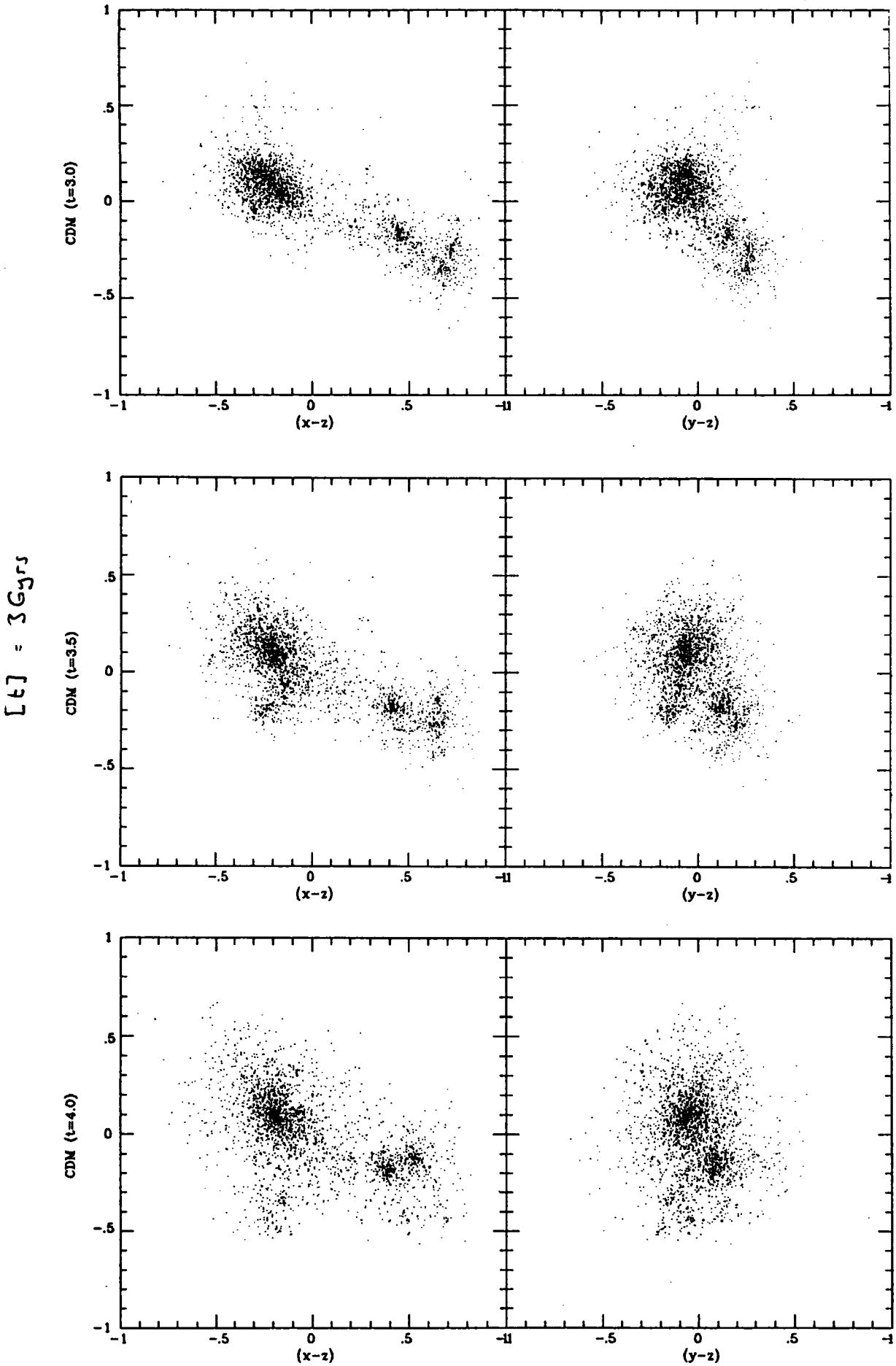
PRESENT



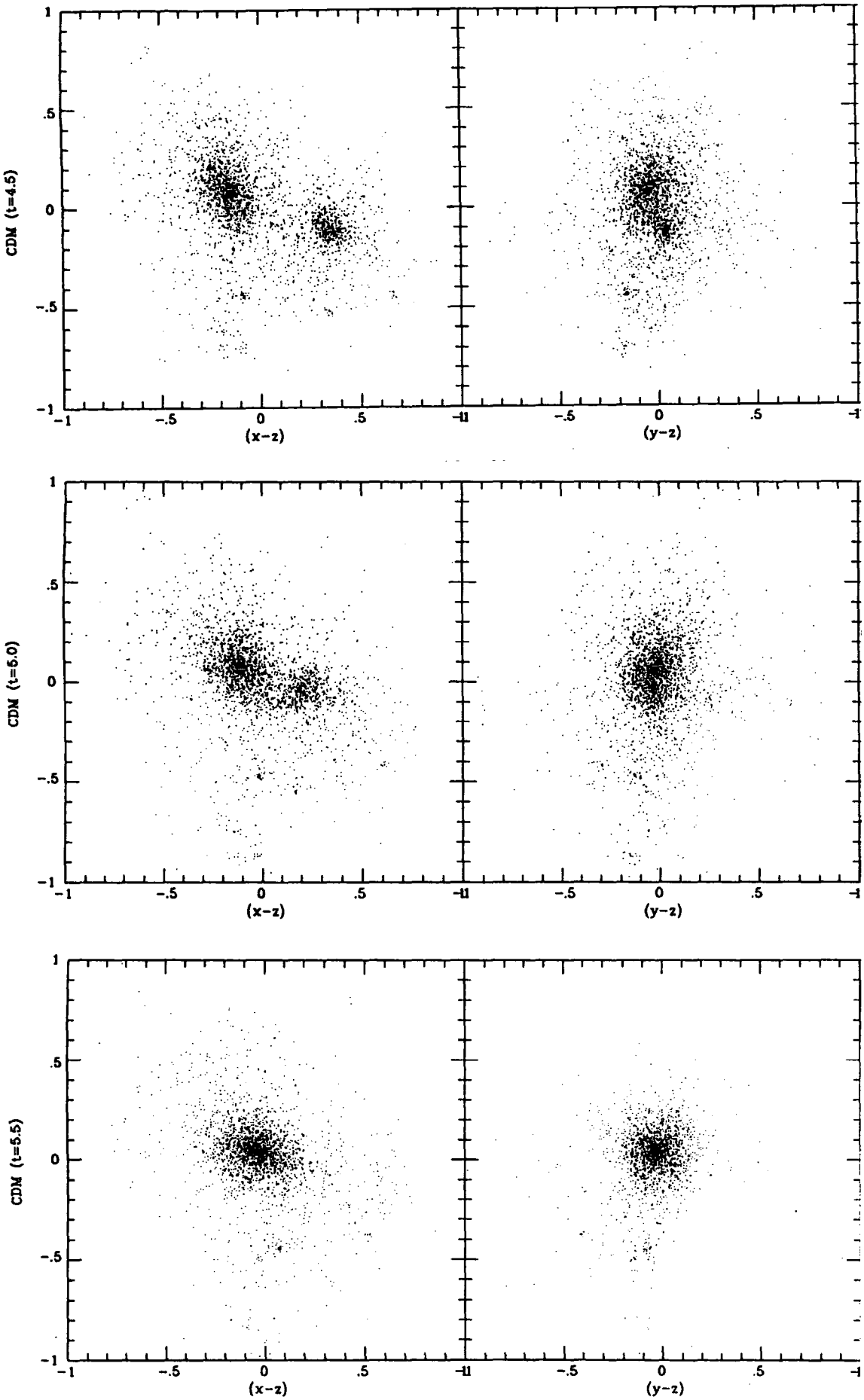
[t] = 3 Gyrs



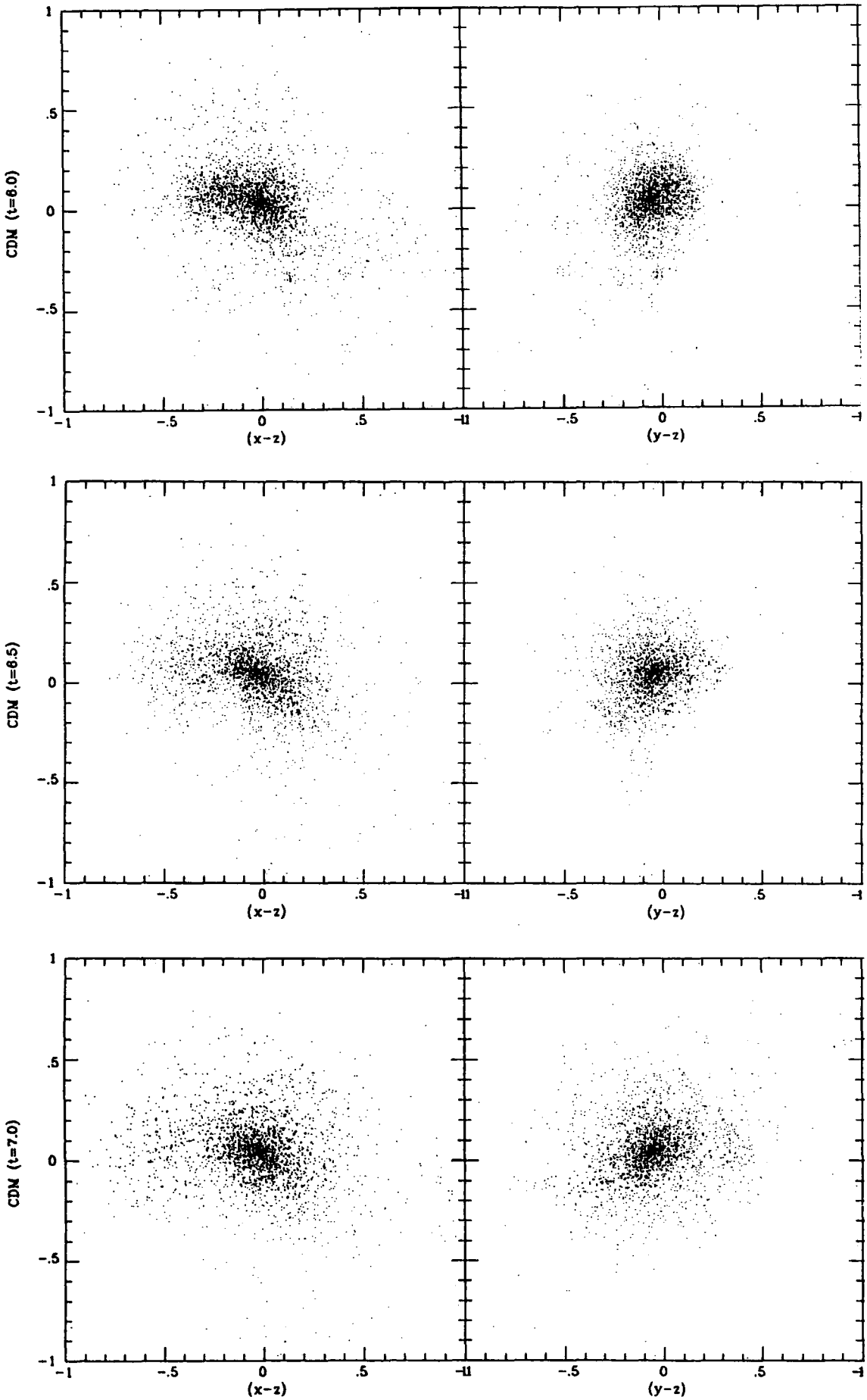




[t] = 3Gyrs



[t] = 3 Gyrs



[t] = 3 Gyrs

At a redshift $z = 2.5$, a time just 2 Gyrs from the Big Bang, the halos are still separating under the general expansion of the Universe. Four Gyrs after the Big Bang the halos have just reached their maximum separation of 1.2 Mpc, which agrees well with the maximum separation predicted using equations (2.3). The initial orbital angular momentum of the system is small, but from a redshift $z = 1$ the transverse velocity of *M31* increases rapidly. At the present the transverse velocity of the *M31* halo is much larger than that predicted for Andromeda by Einasto and Lynden-Bell (1982) or Raychaudhury and Lynden-Bell (1989). The origin of this large angular momentum is discussed in more detail later.

2.3.2 The rotation curves

The rotation curves are sensitive to the location of the "centre" of the halo. Defining the centre of the halo as the centre of mass can give a distorted rotation curve if the halos are not highly spherical. A more realistic centre can be defined as the position in which a galaxy would most likely be found *i.e.* the bottom of the potential well of the halo. The most bound particle within a halo should determine this position fairly accurately and is relatively simple to find. The distance between the centres of the halo defined in these ways is typically ~ 1 kpc. Figure 2.3 shows the rotation curves of both the massive halos in our model at six different stages of their evolution. (The rotation curves shown here give the velocity a particle would have in a circular orbit at distance r from the centre of the halo, $v_c^2 = GM/r$.) At a redshift $z = 2.5$ the rotation curve of *M31* rises rapidly to a peak then falls off slightly slower than Keplerian although the rotation curve of *Milky Way* rises fairly slowly to a constant value of 190 km s^{-1}. By a redshift $z = 1$ both the rotation curves of the *Milky Way* and *M31* are fairly flat. At the time we identify as the present the asymptotic values of the curves are remarkably similar to those of the Galaxy and M31 $\approx$ 220km s^{-1} and 260km s^{-1} respectively. Although the rotation curves do not attain their asymptotic values until a distance $\sim$ 40kpc, several authors have argued that the gravitational effects of the galaxy forming

at the centre will extend a flat rotation curve to the inner parts of the galaxy (Blumenthal *et al.* 1986, Ryden and Gunn 1987). This problem is deferred to a later section in which we make a crude attempt at modelling a gaseous component.

Continuing into the future; 2 Gyrs from now the halos are infalling at about twice their present speed and are only 300 kpc apart. The rotation curves are virtually overlapping and start to become distorted even in the inner regions. At $t_o = 15.6$ Gyrs the halos suffer a devastating collision. The orbital angular momentum is almost all lost during the first collision after which the halos emerge slightly battered. Figure 2.3d shows the rotation curves of the halos just after this time when they reach their second maximum separation of 250 kpc. The curves both rise rapidly and are extremely distorted. Similar rotation curves are observed for spiral galaxies which are in dense environments or interacting with other systems (Whitmore 1990). Very soon after the first collision the halos fall together very rapidly and merge completely with substructure becoming efficiently erased. The merger takes place surprisingly fast, within a third of the present day crossing time, and demonstrates the efficiency of dynamical friction. Much of the orbital angular momentum is transferred to the outer halo as the merger proceeds and the inner region of the remnant shows little rotation. This behaviour has been studied in detail by Frenk *et al.* (1985), Frenk (1987) and Efstathiou and Barnes (1987) and is typical of mergers between halos. The final plot in Figure 2.3 shows the rotation curve for the merger product which is again nearly flat in the outer parts and reaches the asymptotic value of 300 km s^{-1}.

The centres of the halos came within 20 kpc of each other before the halos separated again. It is not obvious what effect the first collision would have had on the galaxies residing at the centres of the halos. At this separation the tidal forces would have been very large and one could imagine gas and stars becoming torn from both galaxies.

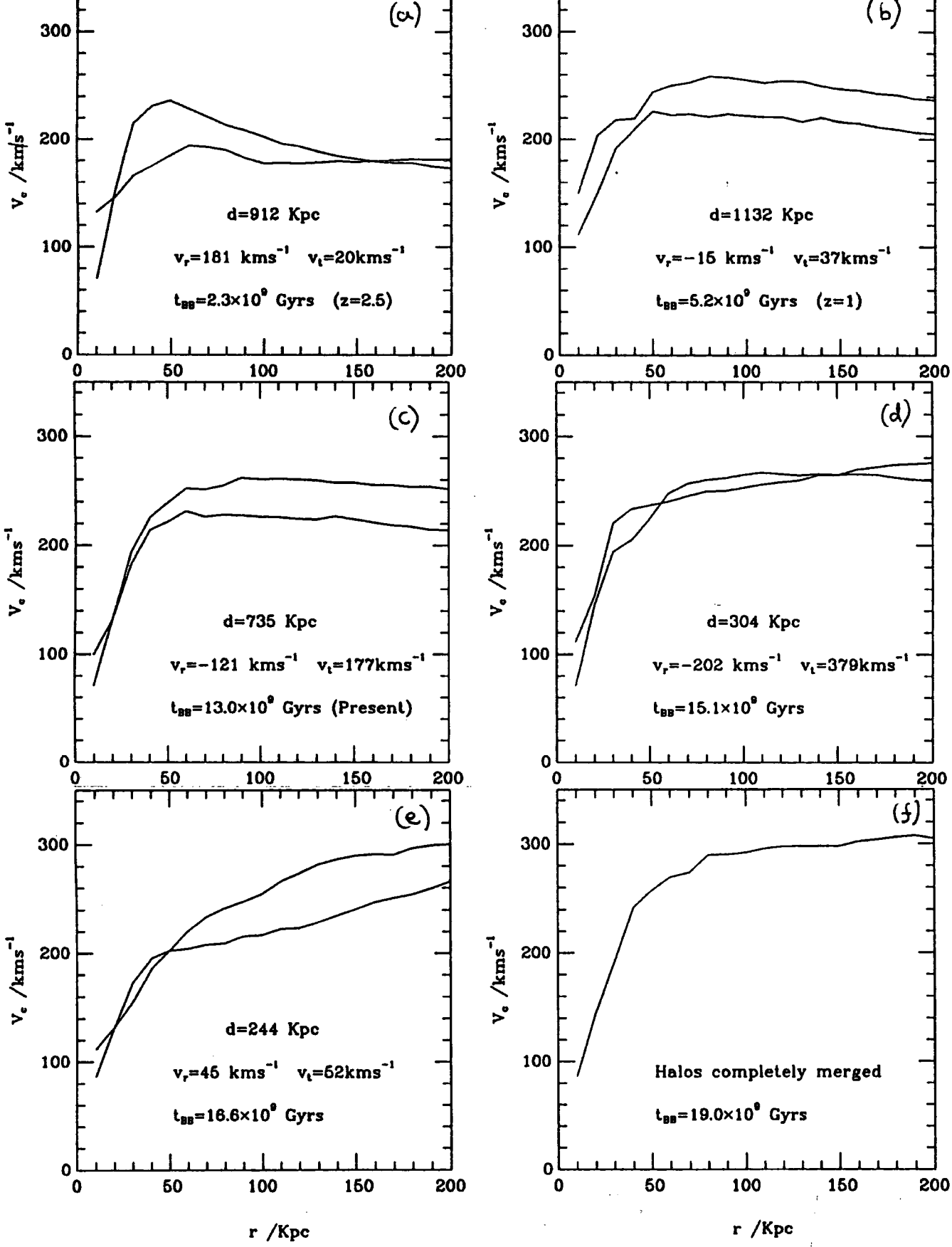


Figure 2.3 The circular velocities, $v_c = GM/r$, of the halos in our model which represent the Milky Way and M31 at different epochs in their evolution.

Peebles (1969) estimated the growth of angular momentum in linear theory and the subsequent non-linear evolution of the tidal field on the protosystem and found a value for the dimensionless spin parameter $\lambda \approx 0.08$.

$$\lambda \equiv \frac{J|E|^{1/2}}{GM^{5/2}}, \tag{2.4}$$

where J, E and M are the total angular momentum, energy and mass of the halo respectively. These analytic calculations and those of White (1984) and Doroshkevich (1970) have been confirmed by the numerical experiments of Efstathiou and Jones (1979), Efstathiou and Barnes (1983) and Barnes and Efstathiou (1987) who found the median value of $\lambda = 0.06$ with little correlation between initial conditions or group size. The halos in our model have values of λ typically in the range $0.05 \sim 0.09$ throughout their evolution with little sensitivity to the formation and merging processes.

2.3.3 The timing argument

Solving equations (2.2) and (2.3) for the mass of the model Local Group is straightforward. For $t_o = 2/3H_o^{-1}$, $v_r = -121$ and $d = 735$ we find a mass $M_{LG} = 4.5 \times 10^{12} M_\odot$. This is very close to the observational result which is not surprising since the model has very similar observational parameters to those of the real Local Group. The *real* mass of the model Local Group is however $9.7 \times 10^{12} M_\odot$. Here we have defined the mass of the Local Group as the sum of the masses within spheres centered on each galaxy extending out to the centre of mass of the two halos. Thus in this case, the timing argument has underestimated the real mass by a factor of two. The reason for this discrepancy is the significant transverse velocity of *M31* in our model; the eccentricity of the orbit at the present is 0.65. Using the full orbital information in the timing argument as Einasto and Lynden-Bell (1982), we find $M_{LG} = 9.1 \times 10^{12} M_\odot$, in good agreement with the actual mass. Table 2.2 lists M_{LG} calculated at various stages in the model's evolution. The mass estimates in columns 5 and 6 come from the timing argument. Those in

column 5 use only the radial velocity information whereas those in column 6 use all the orbital information. Column 7 gives the real mass of the model as defined in the last paragraph.

Zaritsky *et al.* applied the timing argument to Leo I in an attempt to set a lower limit to the mass of the Galaxy. To see how effective this approach is we isolated the particles between $r = 210$ kpc and $r = 250$ kpc and applied the timing argument to their radial velocity vectors. The crossing time of the halo representing the Milky Way is of order several Gyrs. Therefore to obtain a lower limit for the mass of the halo we assumed that the satellite particles were approaching either their first or second apocentric position, depending on whether the radial velocity was positive or negative. Using the 55 test particles lying within the annulus we find a median mass of $1.0 \times 10^{12} M_{\odot}$ compared to the real mass within 230kpc of $2.4 \times 10^{12} M_{\odot}$. This difference is again due to significant transverse motions of the particles which we now discuss in more detail.

2.3.4 The satellite distribution

The anisotropy parameter $\beta(r)$ (equation 2.1) is plotted in Figure 2.3 as a solid line for the *Milky Way* and a dashed line for orbits about the centre of *M31*. Radial orbits would have $\beta = 1$ and isotropic orbits, $\beta = 0.3$. The two distributions are similar and have $\beta \approx 0.4$, decreasing slightly with radius. The anisotropy parameter shows a similar behaviour at $z = 1$. This conflicts with the results of Aguilar and Merrit (1990) mentioned earlier and is probably due to complex tidal effects very early in the formation of the halos.

To compare the distribution of radial velocities within the model with the observed satellites of the Galaxy we adopt the same statistic used by Peebles *et al.* (1990). For each satellite in Table 2.1 and three others with distances less than 50

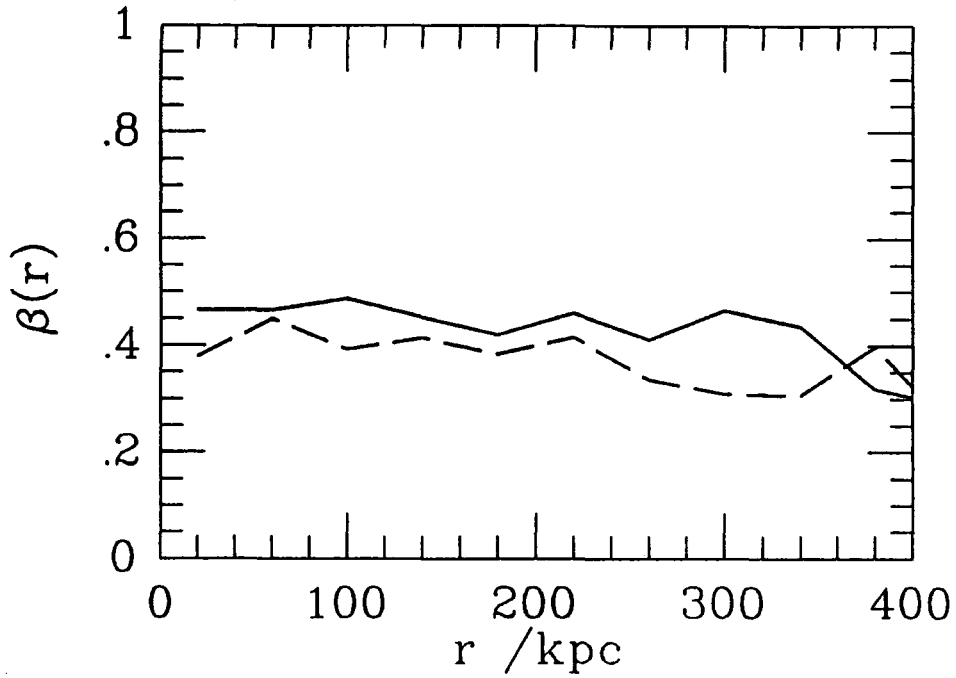


Figure 2.4 The velocity anisotropy parameter, β for the particles within the *Milky Way* (solid line) and *M31* (dashed line). $\beta(r)$ is defined in equation (2.1).

kpc taken from Zaritsky *et al.* (1989), we calculate the probability P_r of finding a radial velocity less than the observed value at the same distance as the satellite.

$$P_r = \frac{N_{v_r > v_{obs}}}{N_{v_r < v_{obs}} + N_{v_r > v_{obs}}} \tag{2.5}$$

where $N_{v_r > v_{obs}}$ is the number of particles with radial velocities greater than the observed satellite and $N_{v_r < v_{obs}}$ is the number of particles with radial velocities less than the observed value. The number of particles used for each calculation is ≈ 50 and is found by taking an annulus of width 40 kpc around the position of each satellite. A successful model should have a uniform distribution of probabilities between 0 and 1 with mean 0.5. The probabilities are plotted in Figure 2.5 as a function of distance from the centre of the *Milky Way*. For our model we find a mean of 0.49 whereas Peebles' *et al.* two seed model yielded a rather low mean

of 0.39. If the dispersion of radial velocities in the model were much larger than the observations then the mean probability would always tend to 0.5; however the probabilities shown in Figure 2.5 are fairly uniformly distributed between 0 and 1. The probability of observing a radial velocity as large as that of Leo I is ~ 0.01. At the distance of Leo I there is a small inflow of particles whereas Leo I is moving away from the Galaxy with a velocity of 177 km s^{-1}. If Leo I is bound to the galaxy then this low probability poses a large problem for the CDM model.

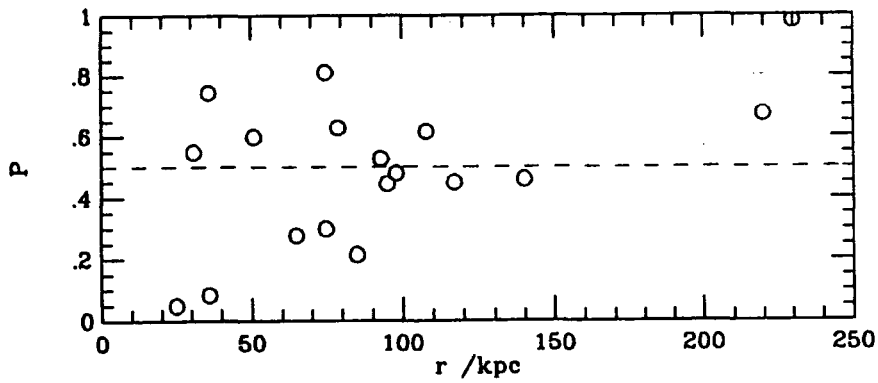


Figure 2.5 The open circles give the probability of observing a radial velocity in our model less than the radial velocity of a satellite at distance r. A successful model would give a uniform distribution of probabilities with a mean of 0.5.

Lin and Lynden-Bell (1982) concluded that the transverse velocity of the Magellanic Clouds was 373km s^{-1}. The probability of finding a transverse velocity as large as this at the distance of the Clouds in the model was 0.15. The

corresponding probability for the radial velocity is 0.6. If the Clouds are just passing pericentre as advocated by Lin and Lynden-Bell, the low probability of finding such a large transverse velocity in the model is expected because satellites spend most of their time near apocentre where they are moving their slowest.

Techniques which use the satellite distribution to determine the mass of the Galaxy are very sensitive to the choice of the anisotropy parameter. The difference between radial and isotropic orbits is a factor of 3 in the mass. For orbits about a point potential, Lynden-Bell *et al.* (1983) derive the mass estimator:

$$M_{est} = \left(\frac{2}{< e^2 >_s}\right)\mu, \tag{2.6}$$

where $\mu = rv_r^2/G$, e is the orbital eccentricity and the angle brackets with subscript s denote the sample average. Radial and isotropic velocity distributions give $< e^2 >_s = 1$ and $\frac{1}{3}$ respectively and are the values usually adopted by previous authors. LT argue that mass estimators like equation (2.6) are not self consistent and do not give reliable confidence intervals and hence they used the method described in Section 2.2.1. This technique and that of LT, assume that the value of μ averaged over the satellite orbits gives a correct picture of the Galactic potential. In Figure 2.6 we compare the mean values of μ for the model, with the real mass as a function of radius. The values for μ are found by averaging over radial shells in the halo representing the *Milky Way*. A typical error bar is plotted on the open circle at $r = 165$ kpc. This error is due to the 1σ dispersion of v_r^2 for particles at this distance and is $\sim \pm 50$km s^{-1}.

We can use the epicycle approximation to calculate the mean eccentricity of the particles within our halo (Chandrasekhar 1942). Using $v_c = 220$km s^{-1} we find that $< e^2 >^{1/2} = < v_r^2 >^{1/2} / v_c = 0.67$. Putting this value in equation (2.6) we overestimate the true mass by a factor of two. This error arises because the halo in the model is not a point potential. The values of μ for all the satellites in Table 2.1 and those within 50 kpc from Zaritsky *et al.* are plotted in Figure 2.6 as

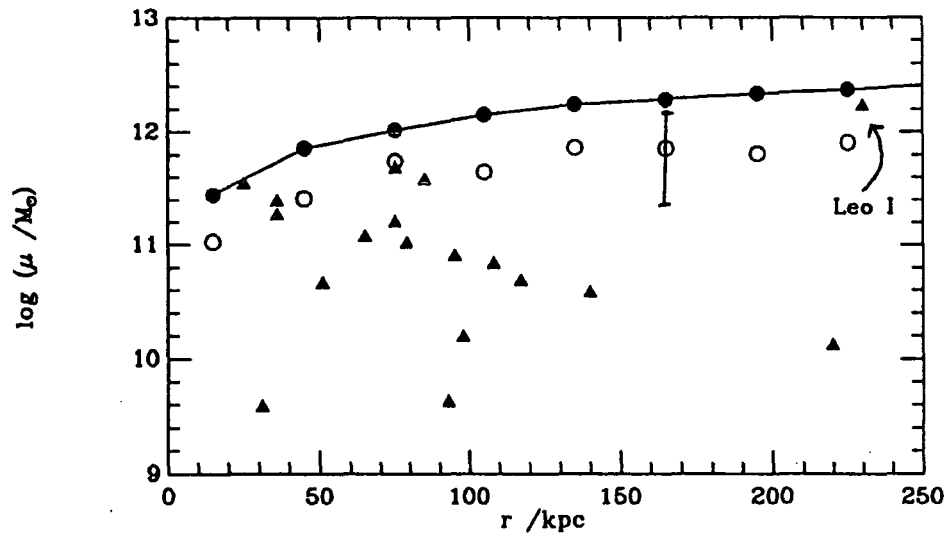


Figure 2.6 The values of $\mu = rv_r^2/G$ for the satellites listed in Table 2.1 (filled triangles) and the mean values of μ for the model (open circles). The joined filled points show the real mass of the halo representing the Milky Way

filled triangles. The model values of μ are substantially higher than the observed values. At first sight this result appears to contradict the probability distribution of radial velocities, however the dispersion of v_r in the model is larger than the observed dispersion of radial velocities, hence rv_r^2 is generally greater in the model.

Including Leo I in the LT analysis gives a mass for the Milky Way comparable to the mass of the halo in our model. The value of μ for Leo I is labelled in Figure 2.6 and lies distinctly above all the other satellites. We found that the probability of observing a radial velocity greater than that of Leo I in the model was 0.01.

2.3.5 The evolution of the tidal field

The total angular momentum of the Local Group can be calculated from $J \equiv \sum_{i=1}^{n}(m_i \mathbf{r_i} \times \mathbf{v_i})$, where $\mathbf{r_i}$ and $\mathbf{v_i}$ are the positions and velocities of particle i with respect to the group centre of mass. Linear perturbation theory predicts that J grows $\propto a^{3/2}$ (Peebles 1980). For the model Local Group we found that this relation is valid for redshifts $z > 0.4$. After this time the angular momentum grows rather erratically. By comparing the dynamics of the Local Group model excised at an earlier time from the P^3M run with the adopted initial conditions, we found that approximately 20% of the orbital angular momentum of *M31* is generated before a redshift $z = 1$ by external torques (*cf. Table 2.2*). After this time the subsequent torque comes from the smaller clumps within the region we identified as the model Local Group.

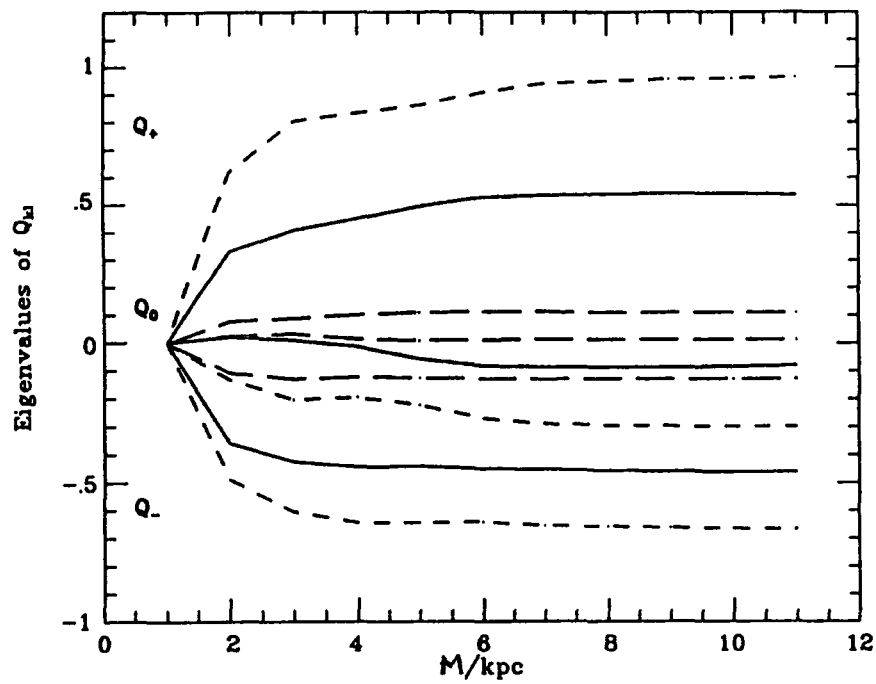


Figure 2.7 The eigenvalues Q_+, Q_o and Q_- of $\mathbf{Q_{kl}}$ for the model Local Group as a function of distance at the epochs $z = 2.5$ (long dashed lines), $z = 0.4$ (solid lines) and the present (short dashed lines).

The value of the quadrupolar component of the force from external masses at the barycentre of the Local Group can be estimated from:

$$\mathbf{Q_{kl}} = \sum_{\alpha=1}^{n} \frac{GM_\alpha}{|\mathbf{r}-\mathbf{r}_\alpha|^3}\left[\frac{3(\mathbf{r}-\mathbf{r}_\alpha)_k(\mathbf{r}-\mathbf{r}_\alpha)_l}{|\mathbf{r}-\mathbf{r}_\alpha|^2} - \delta_{kl}\right], \tag{2.7}$$

where $|\mathbf{r}-\mathbf{r}_\alpha|$ is the position vector to mass M_α (Raychaudhury and Lynden-Bell 1989; hereafter RL). RL used a sample of over 200 nearby galaxies to calculate the eigenvectors and eigenvalues of the 3×3 matrix $\mathbf{Q_{kl}}$. They found that the quadrupole components converge well before 12 Mpc and are mainly due to galaxies within 6 Mpc. The eigenvalues indicated both a compression and extension of the Local Group along two axis which implied a small transverse velocity for M31.

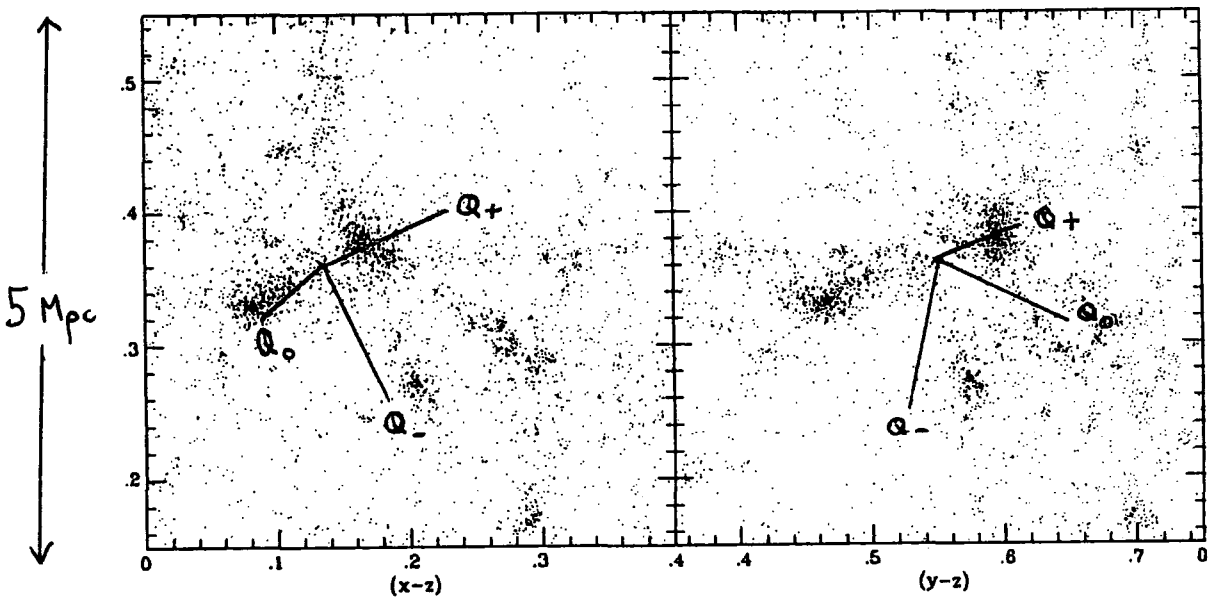


Figure 2.8 The projected particle positions of the Local Group model at the present epoch with the eigenvectors of $\mathbf{Q_{kl}}$.

To calculate $\mathbf{Q_{kl}}$ for our model we used the particle positions from the whole simulation, excluding those within 1 Mpc of the barycentre of the two massive halos, and taking into account the periodic box assumed by the P^3M run. Figure 2.7 shows the convergence of the eigenvalues of $\mathbf{Q_{kl}}$ about the barycentre of the two massive halos in our model at a redshift $z = 2.5$, $z = 1$ and the present. The quadrupole force from external clumps essentially converges within a few Mpc at all epochs from $z = 2.5$. Most of the angular momentum is generated by the nearby small clumps which are actually included within our model Local Group. The directions of the eigenvectors corresponding to the eigenvalues of $\mathbf{Q_{kl}}$ are shown on Figure 2.8 in order to visualise the direction of the tidal torque on the binary system.

2.4 MERGING AND ROTATION CURVES REVISITED

The N-body simulations previously described only follow the evolution of a dissipationless dark matter component. Big Bang nucleosynthesis sets stringent limits to the amount of baryons in the Universe. The latest calculations described by Pagel and Schramm (1991) show that the contribution to the cosmic density in the form of baryons is $0.01 \leq \Omega_b h^2 \leq 0.02$. Thus if $\Omega = 1$, most of the mass in the Universe must be non-baryonic. Observationally we find that visible galaxies contribute only $\Omega_{vis} \approx 0.05$ to the cosmic density.

Within the CDM model, dark matter is expected to be dynamically dominant over the gaseous and stellar components. Massive dark halos, like those studied in detail in the previous section, provide the sites for galaxy formation to take place. Gas is expected to dissipate energy and fall into the potential wells of the dark halos thus forming galaxies. The dynamical effect which infalling gas has upon the dark halos has been calculated using analytic methods by Blumenthal *et al.* (1986) Ryden *et al.* (1987). These authors find that this process should continue a flat rotation curve into the inner parts of the halo. The rotation curves shown in Figure 2.3 are remarkably flat over several hundred kpc although the curves do not attain their asymptotic values until $\approx$ 40kpc. Spiral galaxies however, show flat rotation curves a few kpc from their centres. The simulations presented here have a resolution comparable to the optical radius of spiral galaxies and therefore allow a comparison of the model rotation curves with the observations.

The rate at which halos merge is surprisingly fast with substructure soon becoming completely erased (Frenk *et al.* 1988). For example, the largest halo in Figure 2.1 is the end product of a complicated sequence of merger events between several small halos; the final remnant shows no clues of its violent past. Halos with a history like that described here have been proposed as the formation grounds for elliptical galaxies (Frenk 1986). The consequences of such rapid and

destructive mergers between clumps leads to a paradox concerning the formation of groups and clusters of galaxies within this picture. If small groups are transient phenomena then the origin of larger systems in the hierarchy is a problem. Several attempts have been made to model the evolution of galaxies within small groups by inserting pre-dissipated galaxies into the dark halos (Carlberg 1988). After a short dynamical timescale most of the galaxies have merged together. The shrinkage of these galaxies does not seem to be enough to maintain these objects as distinct units.

In this section a crude attempt is made at modelling a gaseous component so that the effects of dissipating gas within the CDM model can be examined. The questions we attempt to answer in this section are: Will the infalling gas contract the dark halos sufficiently to extend a flat rotation curve to within a few kpc of the centre? Could such a contraction break the merging hierarchy, allowing groups of galaxies to form and attain virial equilibrium?

2.4.1 Method

There are currently two techniques for modelling gas within N-body simulations. Smoothed Particle Hydrodynamics (SPH) uses a particle-mesh N-body code to follow the gas densities at grid points at each timestep. These densities can be related to shocks, winds, star formation, etc (Evrard 1988). The second method treats individual particles as gas clouds with size resolution set by the force softening. The clouds undergo inelastic collisions with some preset criteria for dissipating energy and turning gas into stars (Negroponte and White 1983). Although the first technique is more realistic we adopt the second method since our initial conditions only allow us to resolve scales of a few kpc *i.e.* that of small dwarf galaxy and this simple treatment should be sufficient for our purposes.

To examine the effects of dissipating gas upon the rotation curve we take the particles from the Local Group model at a redshift $z = 1$. Initially the gas

clouds are assumed to occupy the same phase space as the dark matter. Each dark matter particle is therefore given a neighbouring gas cloud with the same velocity vector. The total mass is kept constant with $m_{CDM} = 10 m_{GAS}$ and the initial kinetic and potential energy are therefore unchanged to within 0.1%. By the time we start our models at $z = 1$, we assume that the baryonic component has collapsed into relatively small clouds which we imagine to be similar to II Zwicky 40 (Brinks 1990). Such systems are almost purely HI gas clouds of primordial element abundance and are only visible due to recently induced star formation. The mass of gas in such systems is of the order of a few times $10^8 M_\odot$. Each gas cloud in our model has a mass $5.3 \times 10^8 M_\odot$ and "radius" set by the softening parameter of 3 kpc.

As the particles are embedded into the tree structure before each force calculation, we form a list of possible collision partners which lie within 30 kpc of each gas cloud. The list is sorted through to find the first collision partner which is heading towards the cloud and will pass within a prespecified distance r_{coll}. The collision is then treated in a similar fashion to Negroponte and White (1983). The relative radial velocity of the two clouds is reversed and reduced by a factor $\sqrt{\alpha}$ where α is a parameter expressing the elasticity of the collision and mimics the efficiency of star formation and the energy that is lost during such a collision. The collision treatment conserves momentum and angular momentum whilst a fraction between 0 and $(1 - \alpha)$ of the centre of mass kinetic energy is dissipated. We adopt $\alpha = 0.8$ and $r_{coll} = 7$ kpc. If a collision leaves two clouds bound (*i.e.* their total energy is negative) then we merge the particles, conserving momentum but dissipating a fraction of the kinetic energy. This solves the problem of tight binaries forming with repeated collisions taking place.

2.4.2 Results

Figure 2.9 shows the dark matter particles and gas clouds plotted separately

in a sequence of equally spaced times as in Figure 2.2. The evolution of the model with and without gas proceeds in a similar fashion, but the gas soon falls into the central regions of the nearest dark matter clump.

In Figure 2.10 we plot the circular velocity $v_c = \sqrt{GM/r}$, of the halos with which we identify the Milky Way and M31. Figure 2.10(a) shows the original rotation curves and Figure 2.10(b) shows the rotation curves including gas in the simulation. The effect of incorporating a dissipational component is to continue the flat rotation curve into the central regions of the halo. The number of collisions that a particular cloud undergoes depends upon r_{coll} and the local density. The effect of gas infalling into the halo at large distances has had the undesirable consequence of turning a flat rotation curve into a slightly falling one. This problem could be avoided by starting the model from an earlier time and choosing a smaller value for r_{coll} which would decrease the number of collisions in the outer halo.

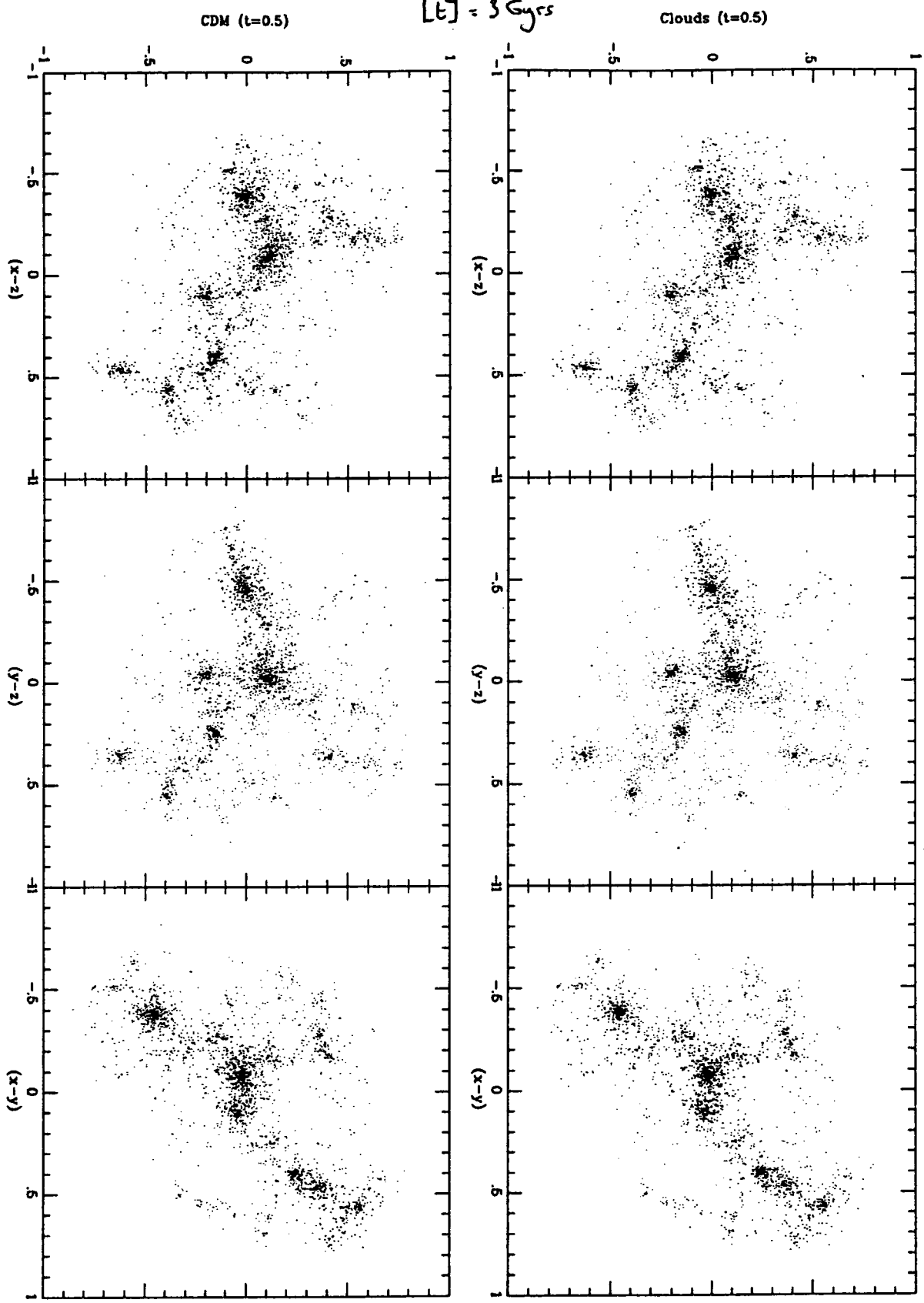
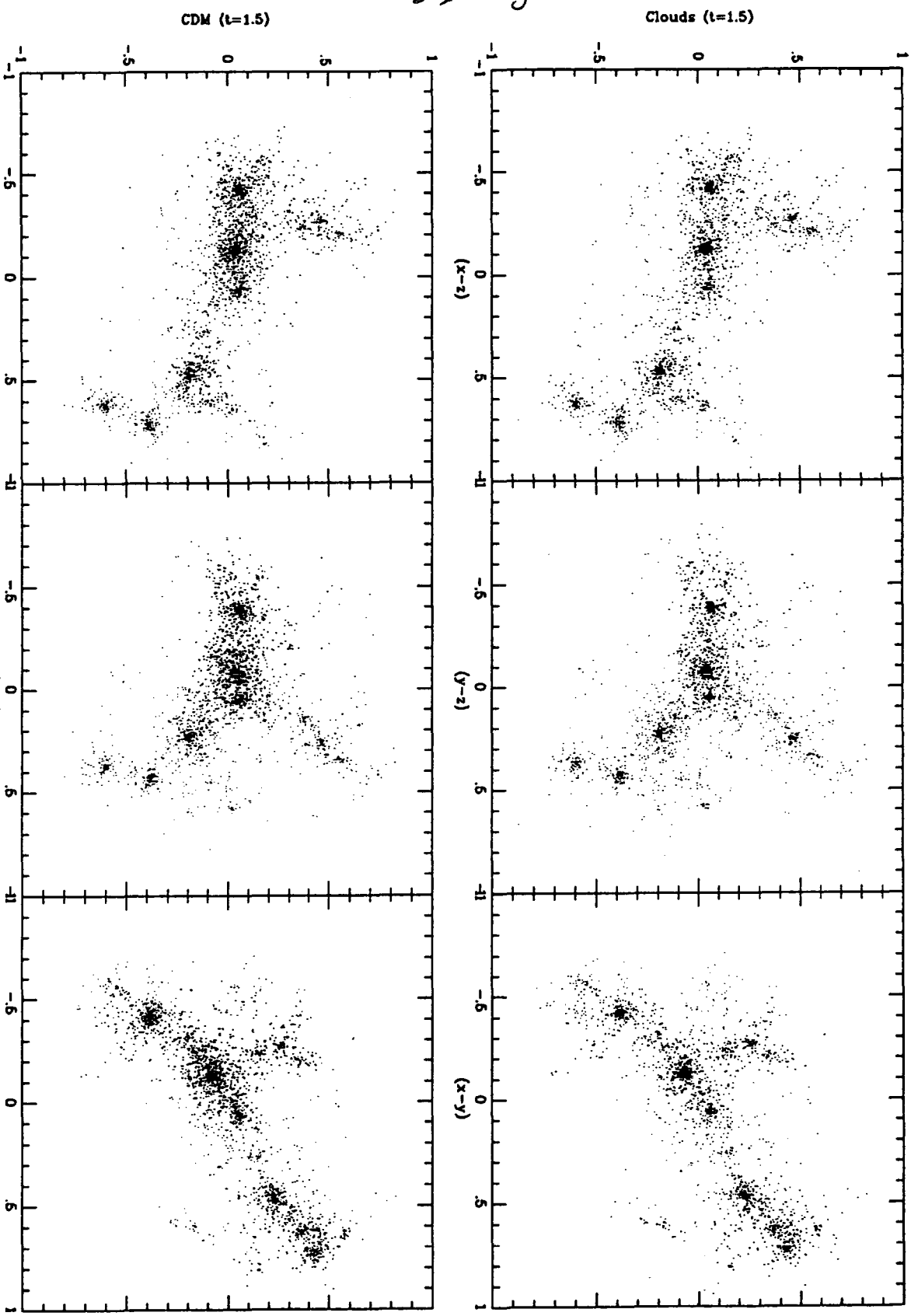


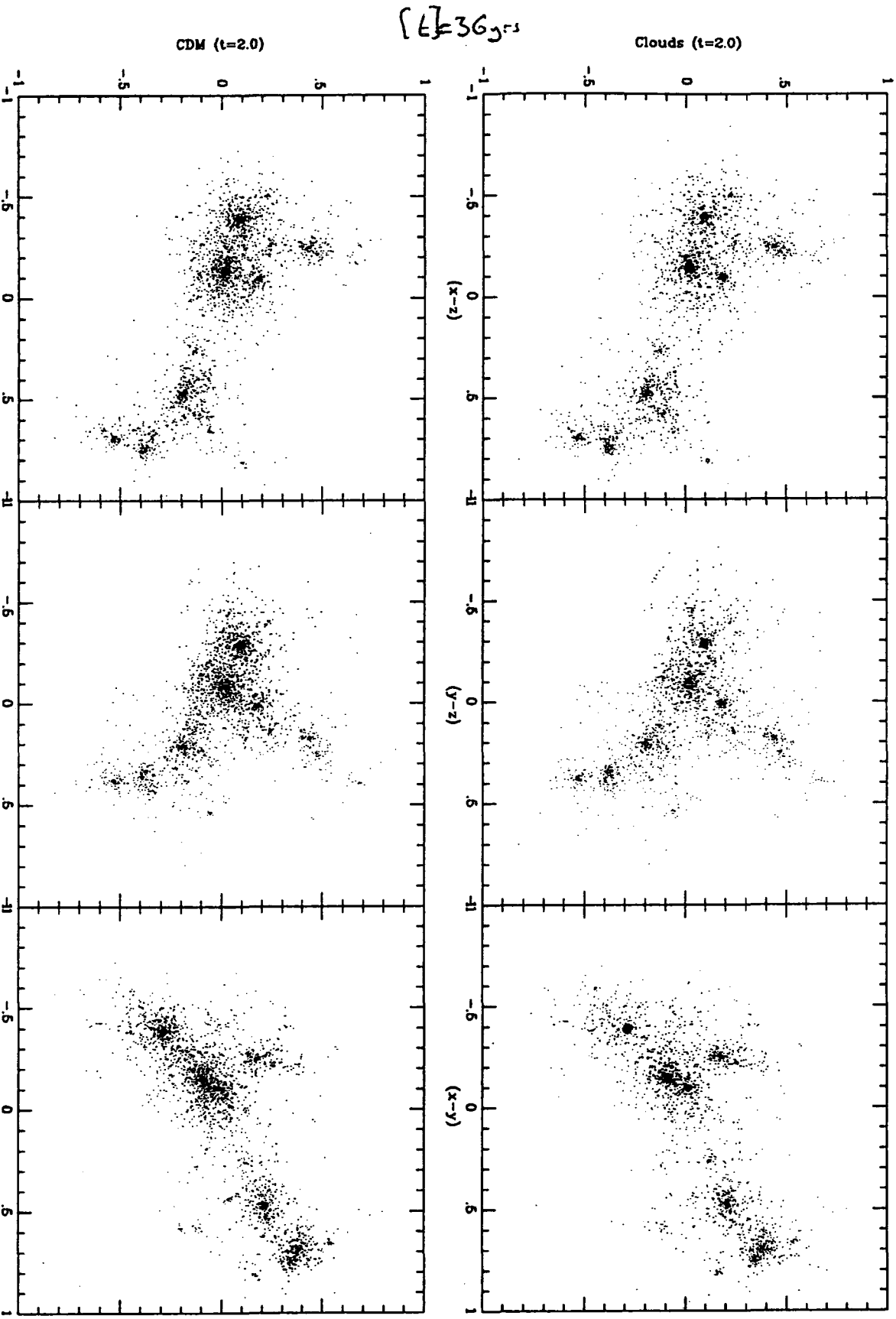
Figure 2.9 The separate panels show the dark matter particles and gas clouds from the Local Group model plotted in projection at the same times as Figure 2.2. Again we show physical coordinates in a box of length 4 Mpc.

[t] = 3Gyrs

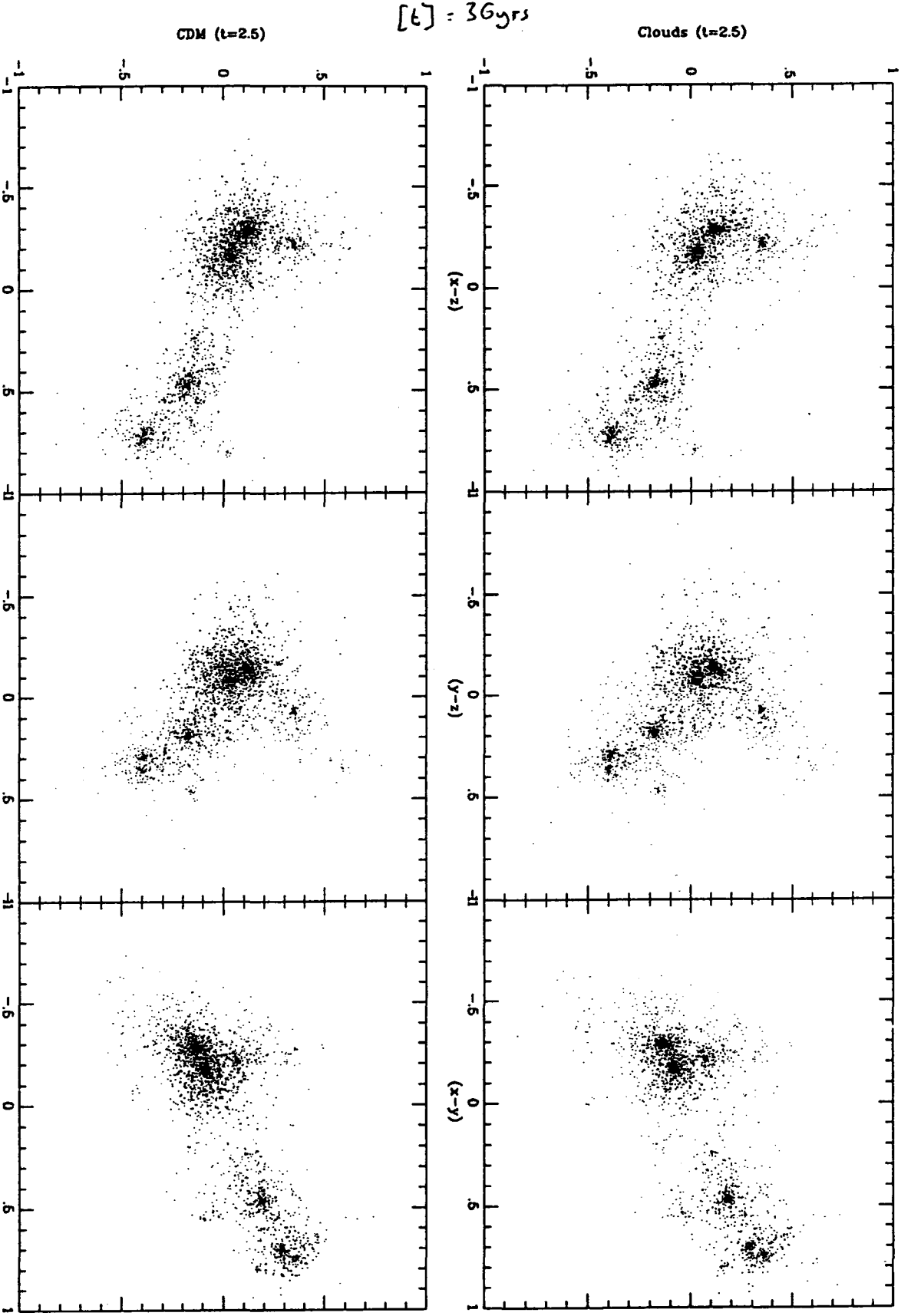


THE "PRESENT"

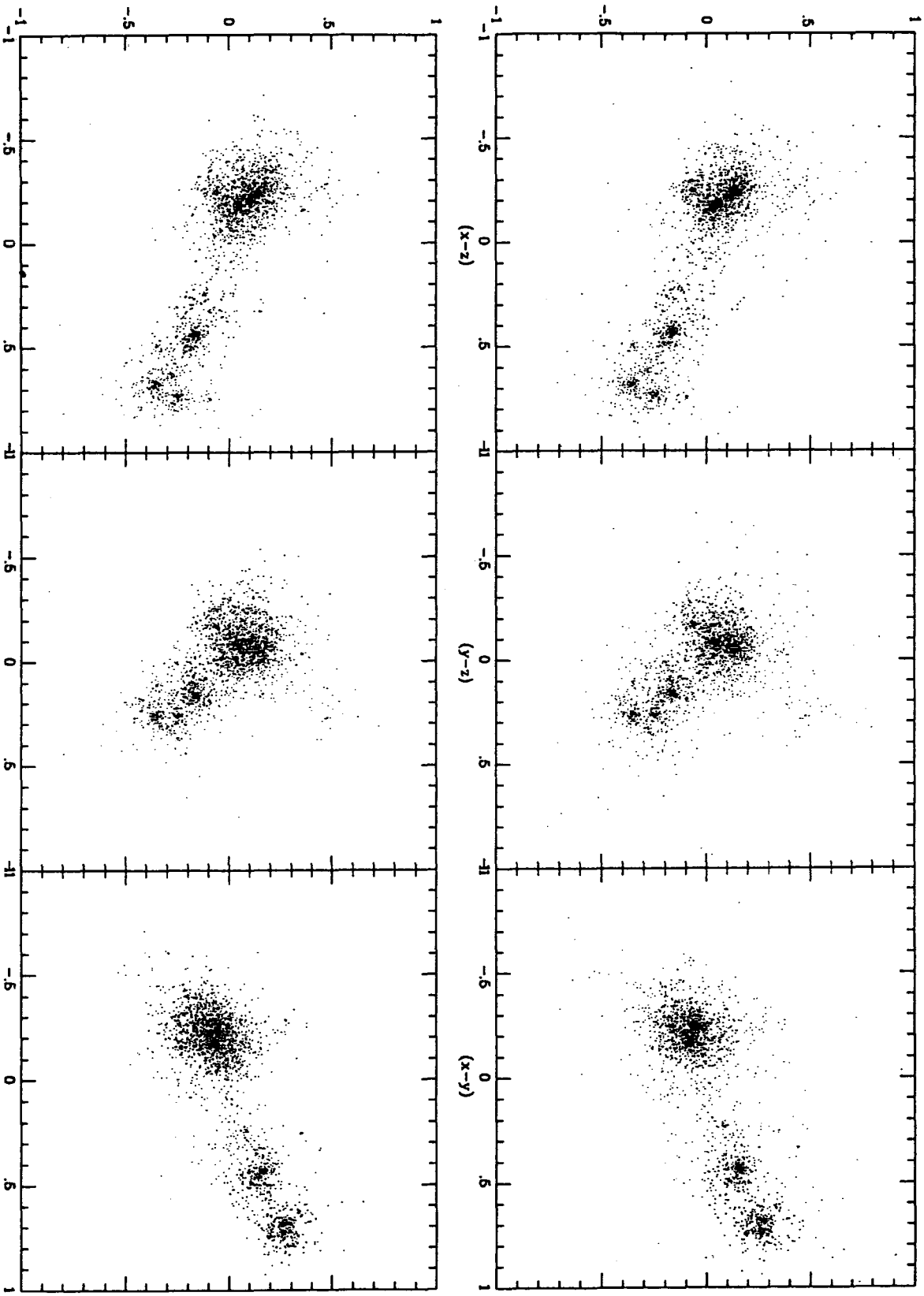
[t]=3Gyrs



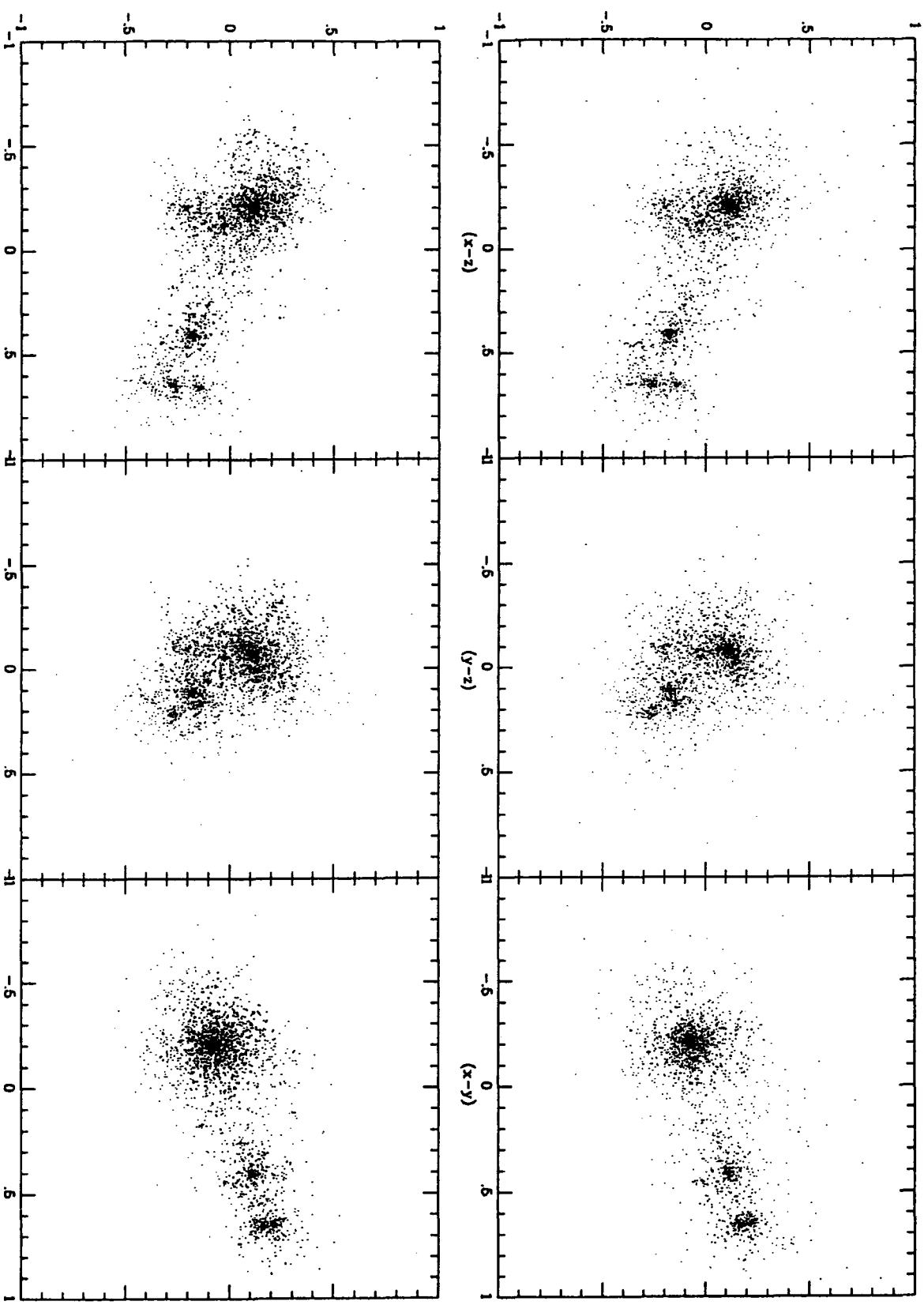
[t] = 3Gyrs



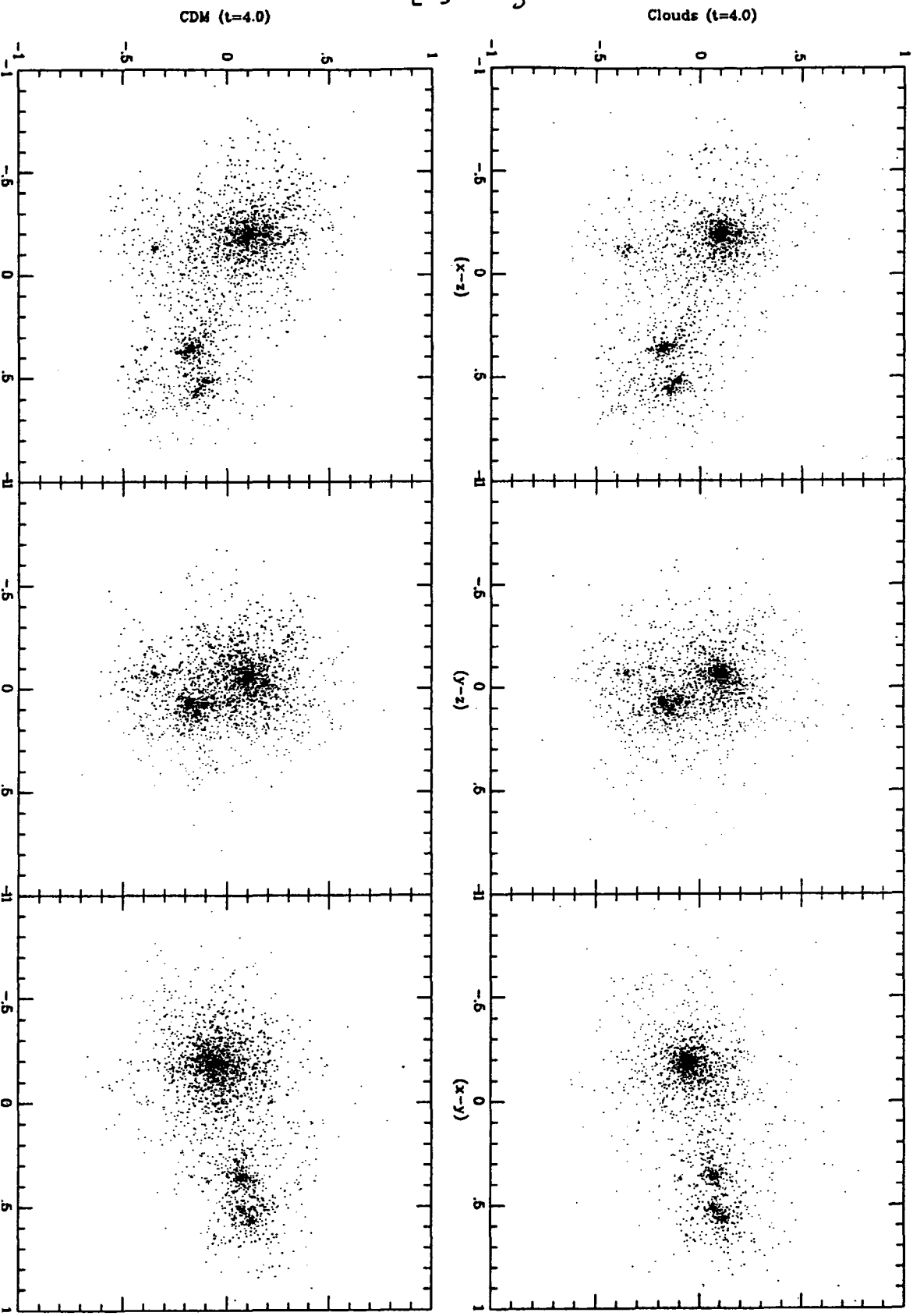
[t] = 3 Gyrs



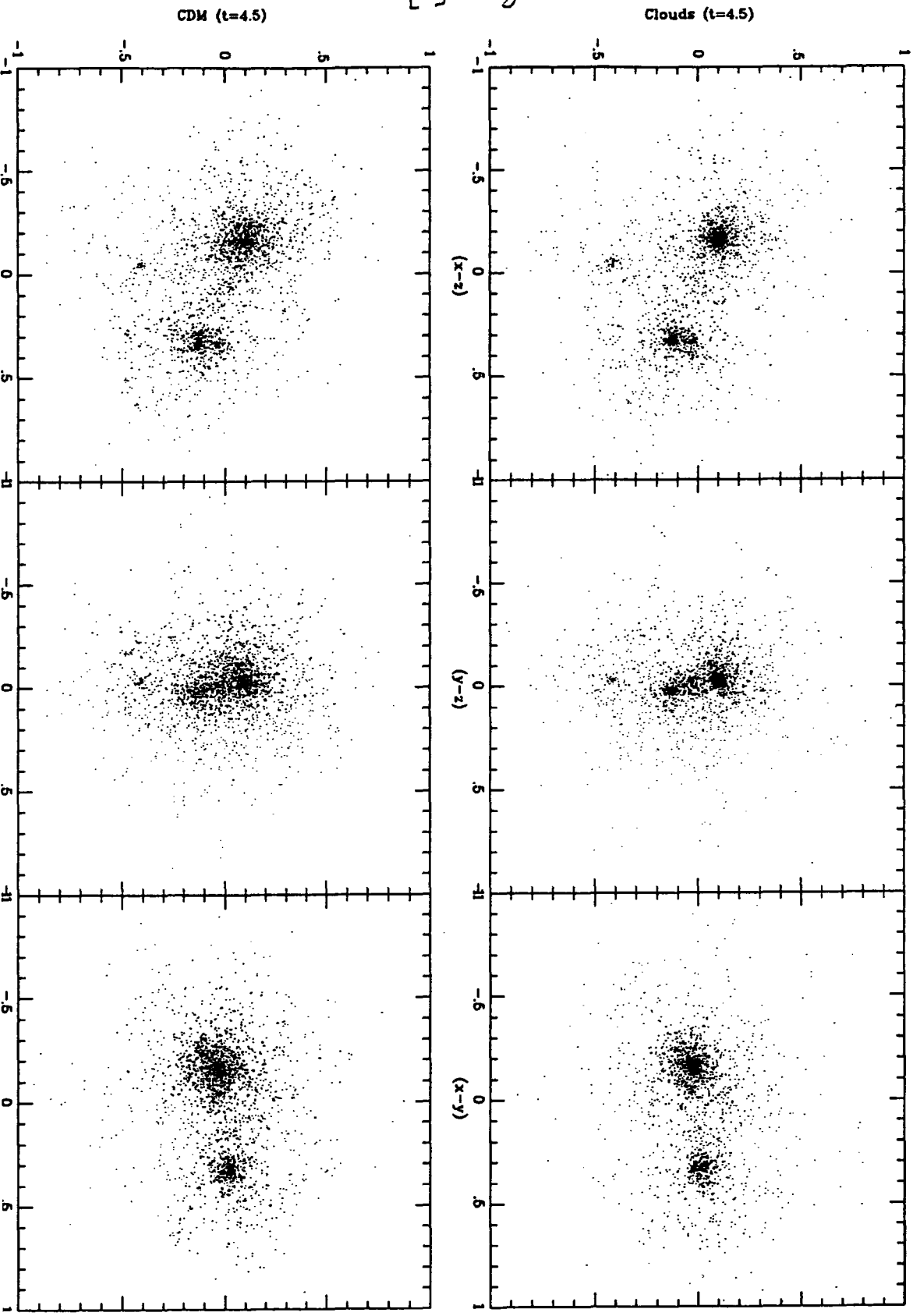
[t] = 3Gyrs



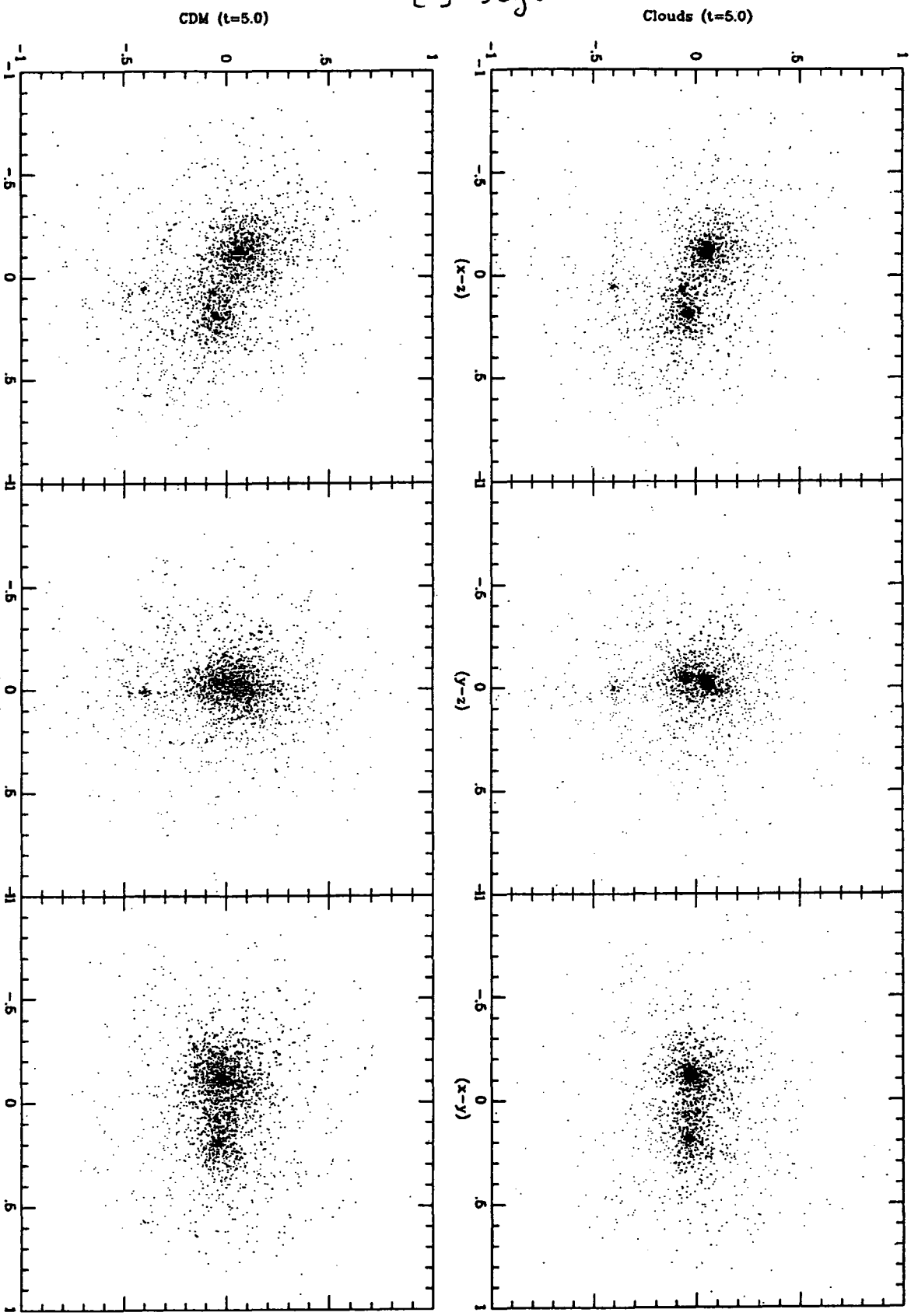
[t] = 3Gyrs



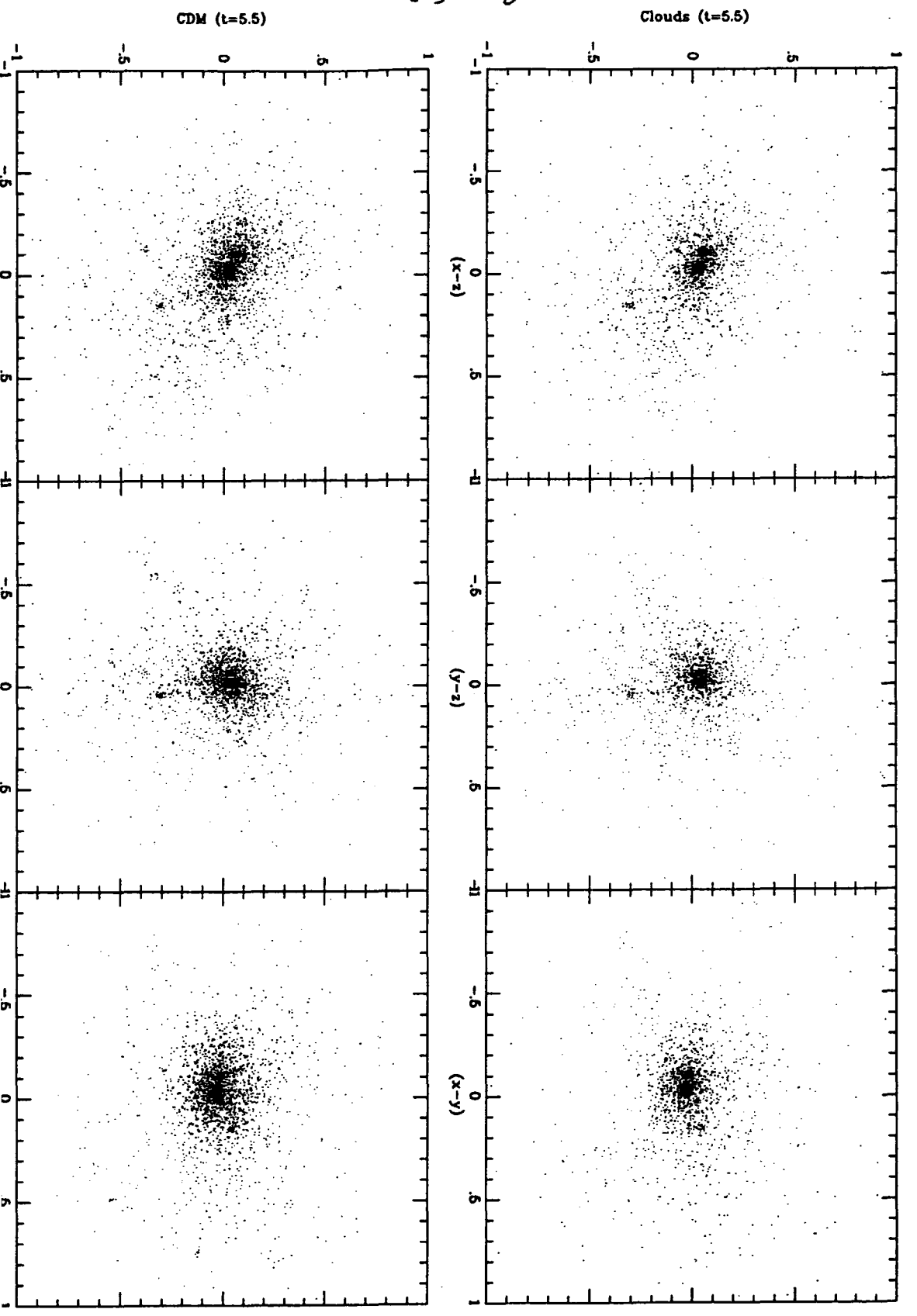
[t] = 3 Gyrs



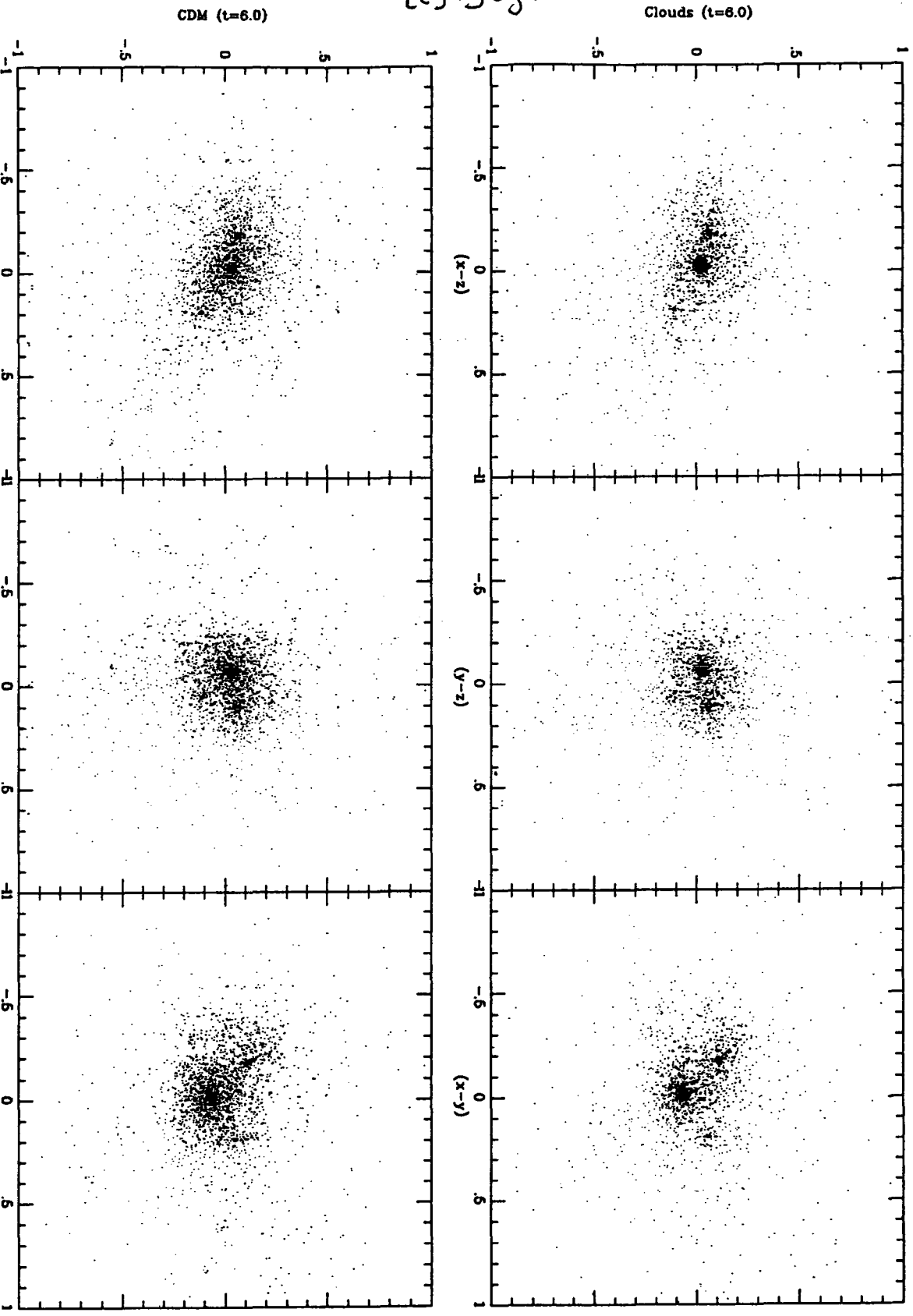
[t] = 3Gyrs



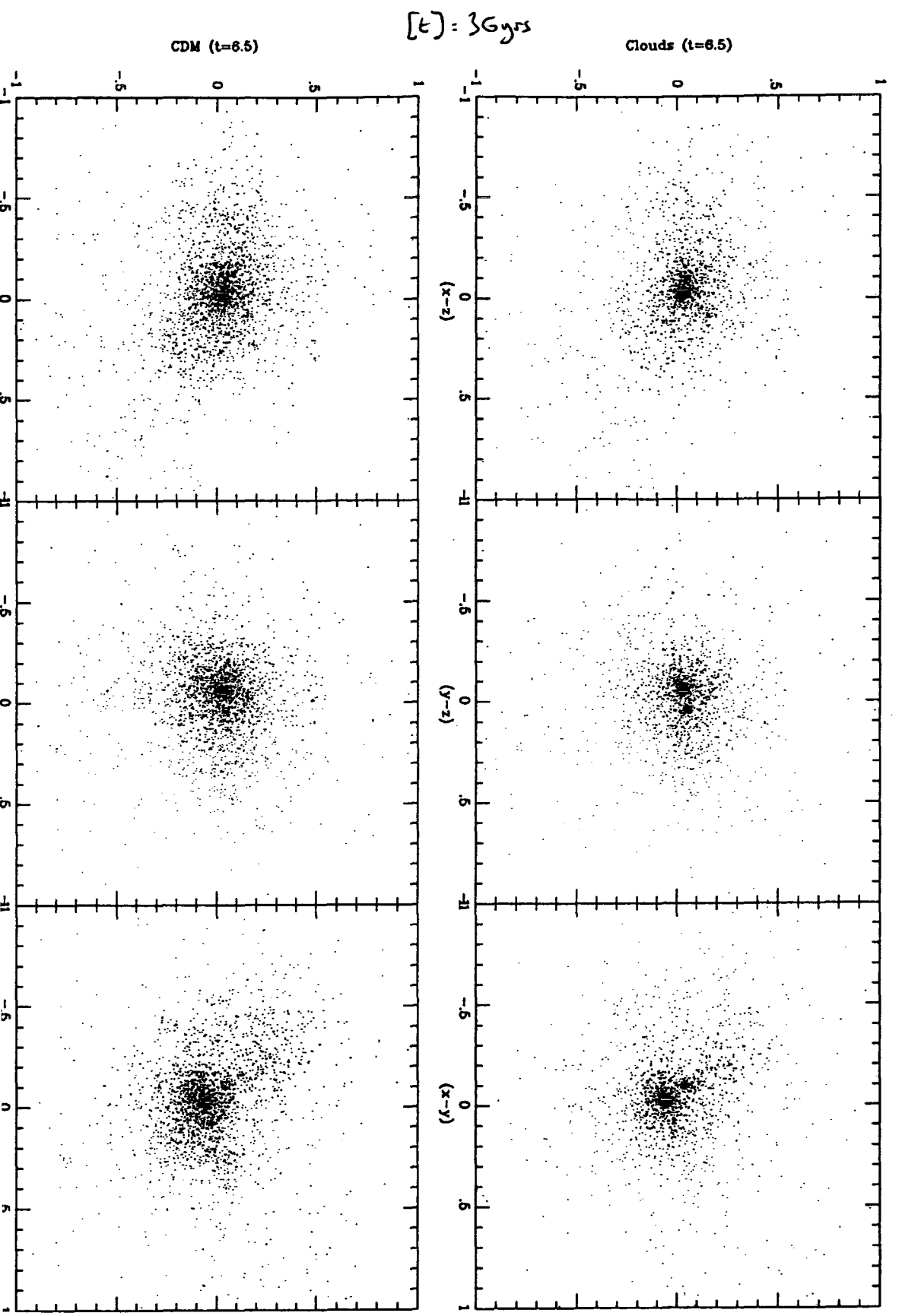
[t] = 3 Gyrs

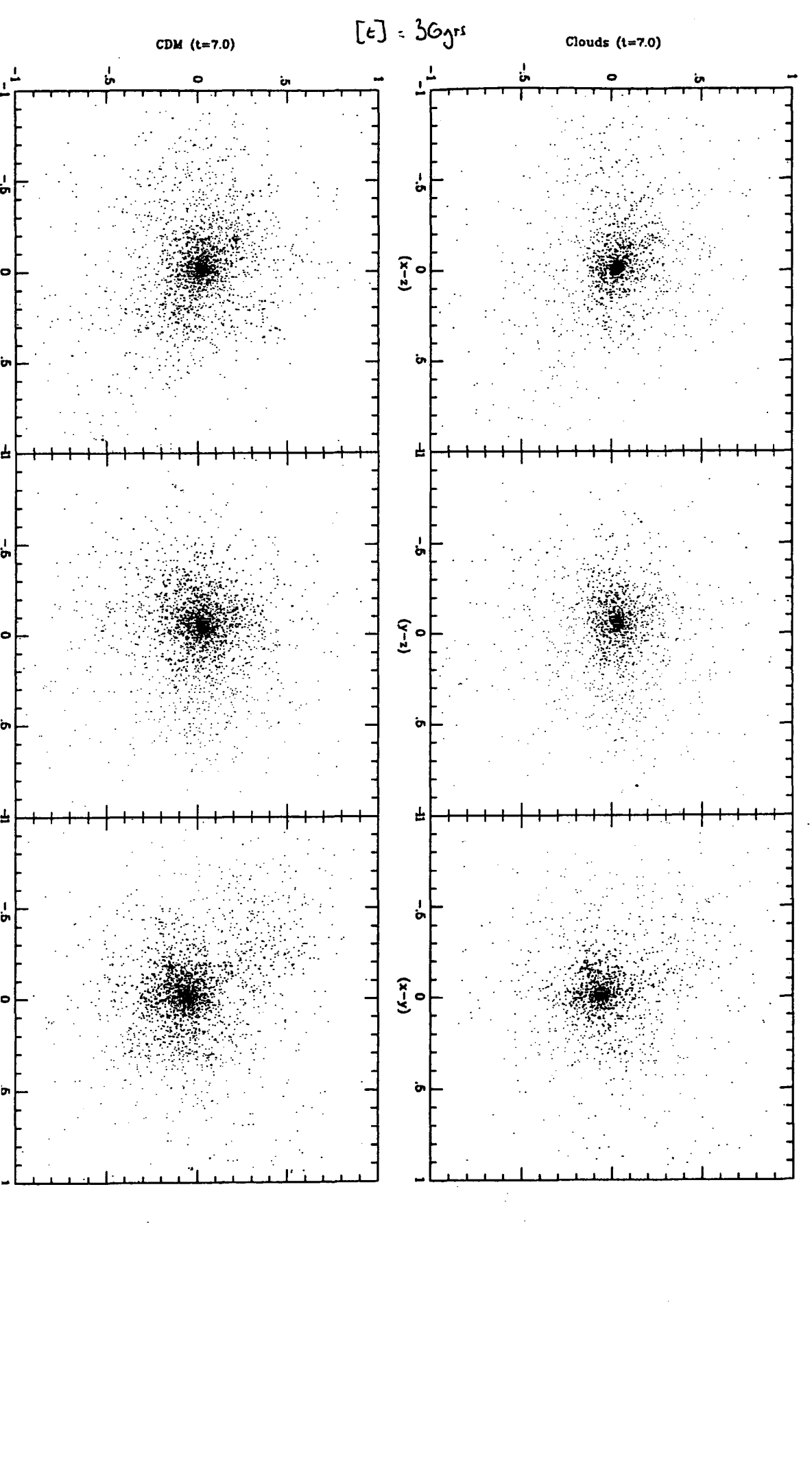


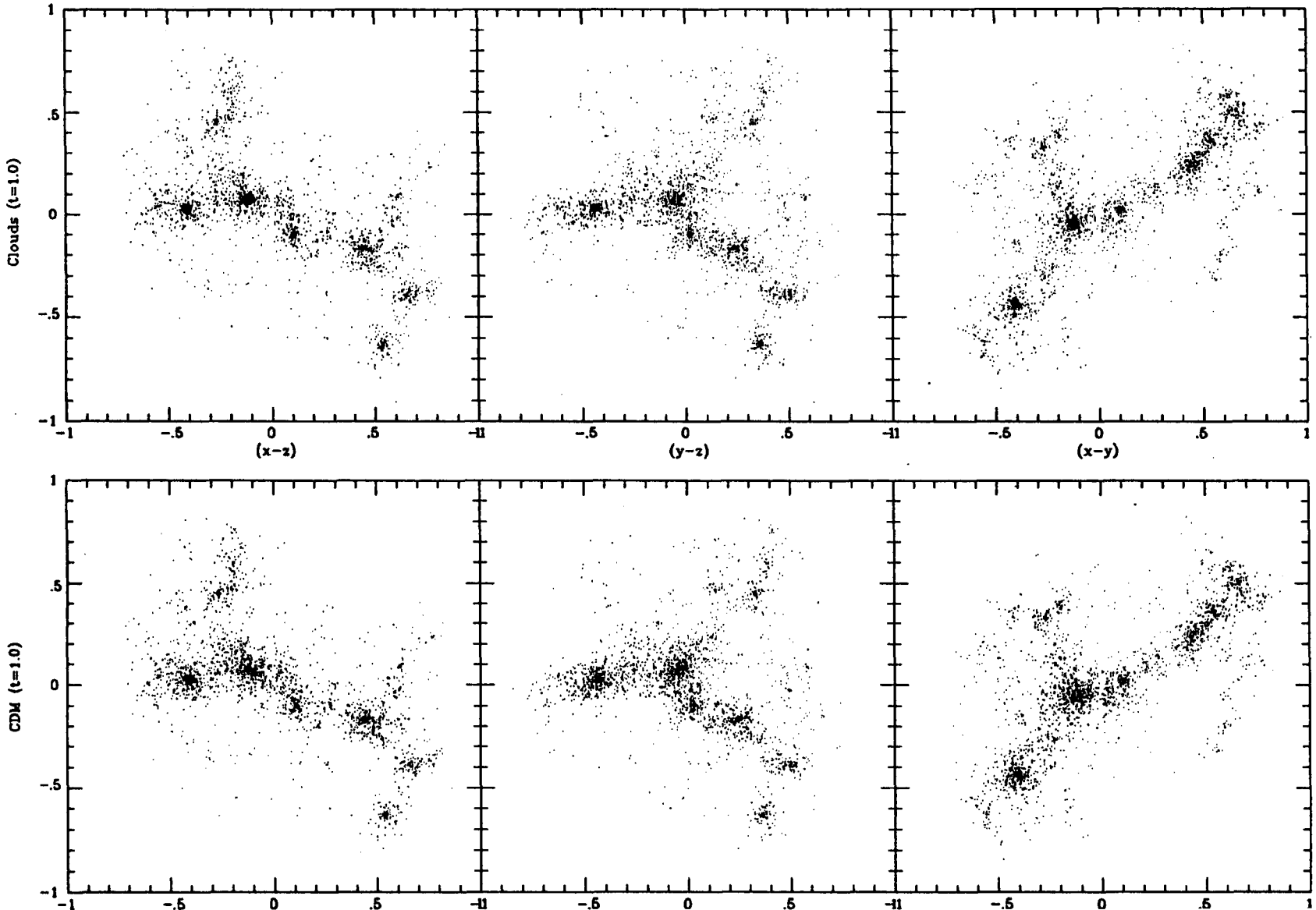
[t] = 3 Gyrs



[t] = 3Gyrs




[t] = 3Gyrs
CDM (t=7.0)
Clouds (t=7.0)
(z-x)
(y-z)
(x-y)


Clouds (t=1.0)
CDM (t=1.0)
(x-z)
(y-z)
(x-y)
1
.5
0
-.5
-1

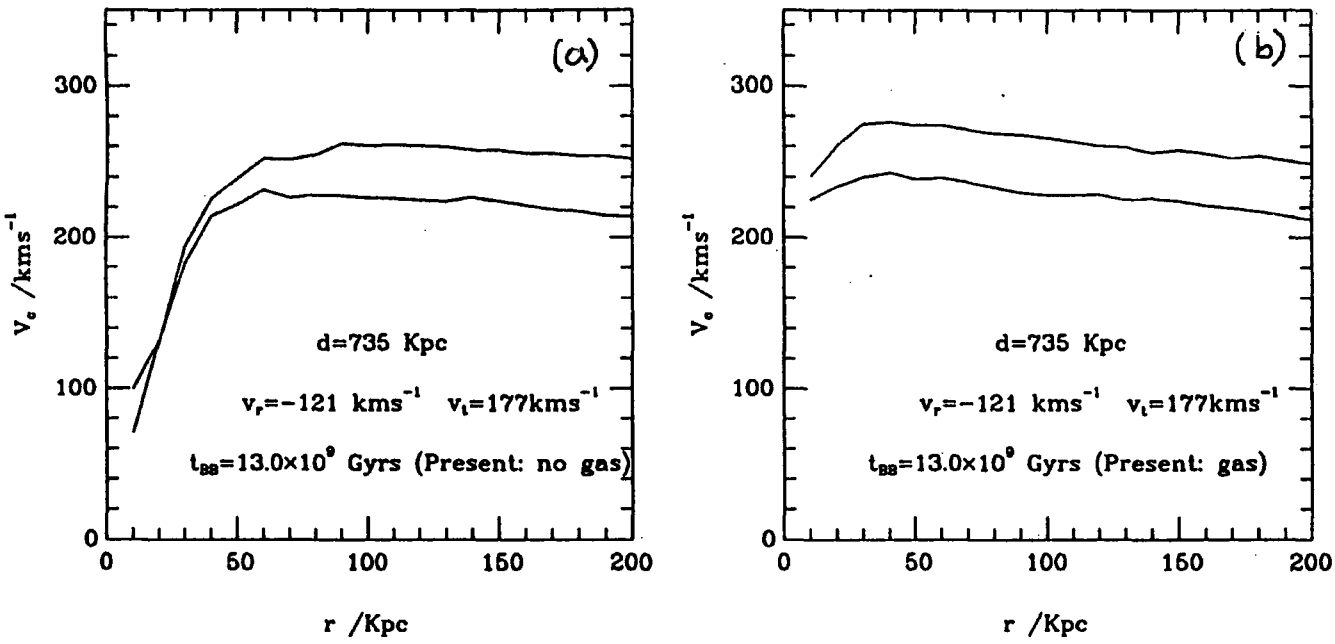


Figure 2.10 The rotation curves of the halos in our model Local Group at the present time, *(a)* without gas and *(b)* with gas.

We can see from Figures 2.9, 2.10 and 2.11 that the dissipating gas contracts within the halos thus increasing its binding energy. At present we have only a visual picture of the differences, but it is clear from a comparison of the models with and without gas that the merging timescales of some of the clumps are significantly longer when dissipation is included. Whether or not this effect is enough to break the merging hierarchy is not yet clear and further work is neccesary to fully tackle this problem (Katz and Gunn *preprint*, Barnes 1985).

2.5 DISCUSSION AND CONCLUSIONS

Determining the mass distribution within the Local Group has presented a challenge to observers and theoreticians throughout the past decades. In this chapter we have shown that almost all the observations are consistent with a high mass model for the Local Group in which massive halos extending for several hundred kpc surround the Milky Way and M31. These conclusions are based on a model for the Local Group which we have used to test the predictive capabilities of standard mass estimators. The model is a binary system of halos drawn from a numerical simulation of a Universe dominated by cold dark matter. The formation sequence of these halos is typical of the growth of fluctuations within the CDM model.

The rotation curves of the two large halos in the model attain the same asymptotic values as those of M31 and the Milky Way $\sim$ 220 and 260km s^{-1} respectively. As the merger proceeded dynamical friction was found to be extremely efficient at dragging the halos together in a fraction of the crossing time. As the halos come closer together the rotation curves become very distorted, displaying the same behaviour as observed in interacting spiral galaxies or spirals in dense environments.

We applied the timing argument to the "observed" orbital information and "full" orbital information of the massive halos representing M31 and the Milky Way to determine the mass of the model Local Group. The results support the use of the timing argument for estimating the mass of the Galaxy-M31 system, although we find that the real mass is underestimated by a factor of two due to the large transverse velocity component of *M31* in the model. We investigated the origin of the large angular momentum of the two massive halos in the model. The torque from external clumps converges within a few Mpc from the barycentre

of the model Local Group, with most of the angular momentum generated by the torque of a few small clumps which were included within the model.

The use of satellite galaxies to determine the Galactic potential has always been fraught with problems due to the small number of systems and the unknown anisotropy parameter. The best analysis of the satellite galaxies predicts a mass for the Galaxy similar to the mass of our model. This agreement occurs in spite of apparent differences in detail between the inferred velocity fields around the model halo and the Milky Way. The model successfully matched the observed distribution of radial velocities of the Galaxy's satellites and globular clusters. However, the dispersion of radial velocities in the model is larger than the observed dispersion *i.e.* In the model we find $\mu/GM_{\odot} \sim 50$ whereas the satellites in Table 2.1 have $\mu/GM_{\odot} \sim 16$ ($\mu/GM_{\odot} \sim 8$ without Leo I). The high radial velocity of Leo I remains a puzzle. Only one particle in a sample of 66 at the position of Leo I had a radial velocity as large as Leo I, whereas the probability of observing the radial velocities of the other satellites of our galaxy were large. If we reject Leo I then we are left with the conjecture that the satellites of the Milky Way are on predominantly circular orbits, perhaps due to tidal disruption of satellites on eccentric orbits. The particles within the model halos have an approximately constant anisotropy parameter $\beta \sim 0.4$ out to distances of 400 kpc. β shows the same form at $z = 1$ and the present. To reconcile the large mass inferred for our Galaxy from the timing argument with the kinematics of the satellites, we must assume that the satellites of our Galaxy have a fairly low anisotropy parameter.

Modelling a gaseous component within N-body codes is a fairly new area of research which will undoubtedly become routine within cosmological N-body simulations. We made a crude attempt at modelling gas clouds within our Local Group model. The infalling gas had the effect of contracting the halos, and with a reasonable choice of parameters for the gas clouds, the flat rotation curves could be extended into the inner regions of the halos. The contraction of the halos had

the added effect of prolonging the lifetimes of some of the merging clumps and substructure was not immediately erased as found in simulations containing only dark matter. This could provide a possible solution to the problem of creating long-lived groups and clusters of galaxies within hierarchical clustering models. The future of such simulations is promising and we can expect to learn a great deal about galaxy formation. The ultimate aim of the numerical simulations will be to explain fundamental observational results such as the Hubble sequence of galaxies, correlations between galaxy types and environment, merger rates as a function of time, etc; the list is long but the techniques are within reach.

REFERENCES

Aguilar, L.A., and Merritt, D. 1990, *Ap.J.*, **354**, 33.

Barnes, J. and Efstathiou, G. 1987, *Ap.J.*, **319**, 575.

Barnes, J., 1986, *Nature*, **324**, 446.

Barnes, J., 1988, *Dynamics and Interactions of Galaxies; Roland Wielen (Ed.) Heidelberg.*

Blumenthal, G.R., Faber, S.M., Flores, R. and Primack, J.R. 1986, *Ap.J.*, **301**, 27.

Bontekoe, T.R. and van Albada, T.S. 1987, *Mon.Not.R.astr.Soc.*, **224**, 349.

Brinks, E. 1990, *Dynamics and Interactions of Galaxies; Roland Wielen (Ed.) Heidelberg.*

Burbidge, G. 1975, *Ap.J.*, **196**, L7.

Carlberg, R.G. 1988, *Ap.J.*, **324**, 664.

Chandrasekhar, S. 1942, *Principles of Stellar Dynamics* (New York: Dover).

Davis, M. and Peebles, P.J.E. 1983, *Ap.J.*, **267**, 465.

Davis, M. and Peebles, P.J.E. 1983, *Ann.Rev.Astron.Astrophys.*

Doroshkevich, A.G. 1970, *Astrofizica*, **6**, 581.

Efstathiou, G. and Jones, B.J.T. 1979, *Mon.Not.R.astr.Soc.*, **186**, 133.

Efstathiou, G. and Barnes, J. 1983, in *Proc. 3d Moriond Astrophysics Meeting, Formation and Evolution of Galaxies and Large Scale Structure, page 361.*

Einasto, J. and Lynden-Bell, D. 1982, *Mon.Not.R.astr.Soc.*, **199**, 67.

Evrard, A. 1988, *Mon.Not.R.astr.Soc.*, **235**, 911.

Fabbiano, A. 1989, *Ann.rev.astr.astro.*

Fich, M., Blitz, L. and Stark, A.A. 1989, *Ap.J.*, **342**, 272.

Fich, M. and Tremaine, S. 1991, *Preprint.*

Frenk, C.S. and White, S.D.M. 1980, *Mon.Not.R.astr.Soc.*, **193**, 295.

Frenk, C.S. and White, S.D.M. 1982, *Mon.Not.R.astr.Soc.*, **198**, 173.

Frenk, C.S., White, S.D.M., Efstathiou, G. and Davis, M. 1985, *Nature*, **317**, 595.

Frenk, C.S. 1986, *Phil. Trans. R. Soc. Lond.*, **A 330**, 517.

Frenk, C.S. 1987, *Nearly normal galaxies.*

Frenk, C.S., White, S.D.M., Efstathiou, G. and Davis, M. 1990, *Ap.J.*, **391**, 2.

Fujimoto, M. and Sofue, Y. 1969, *A.A.*, **61**, 199.

Hartwick, F.D.A. and Sargent, W.L.W. 1978, *Ap.J.*, **221**, 512.

Holmberg, E. 1937, *Ann Lund Obs*, **6**.

Kaiser, N., Efstathiou, G., Ellis, R.S., Frenk, C.S., Lawrence, A., Rowan-Robinson, M. and Saunders, W. 1991, *in preparation.*

Kahn, F.D. and Woltjer, L. 1959, *Ap.J.*, **130**, 705.

Kapteyn, J.C. 1922, *Ap.J.*, **55**, 302.

Lin, D.N.C. and Lynden-Bell, D. 1982, *Mon.Not.R.astr.Soc.*, **198**, 707.

Little, B. and Tremaine, S. 1987, *Ap.J.*, **320**, 493.

Lynden-Bell, D., Cannon, R.D. and Godwin, P.J. 1983, *Mon.Not.R.astr.Soc.*, **204**, 87p.

McCall, M.L. 1989, *A.J.*, **97**, 1341.

Murai, T. and Fujimoto, M. 1980, *Publ.Astr.Soc.Japan*, **32**, 581.

Negroponte, J. and White, S.D.M. 1983, *Mon.Not.R.astr.Soc.* **205**, 1009.

Pagels and Schramm, 1991 *Physica Scripta, in press.*

Peebles, P.J.E. 1969, *Ap.J.*, **155**, 393.

Peebles, P.J.E. 1971, *Physical Cosmology.*

Peebles, P.J.E. 1980, *The Large Scale Structure of the Universe.*

Peebles, P.J.E. 1989, *Ap.J.*, **344**, L54.

Peebles, P.J.E., Melott, A.L., Holmes, M.R. and Jiang, R.L. 1989, *Ap.J.*, **345**, 108.

Peebles, P.J.E. 1990, *Ap.J.*, **362**, 1.

Peterson, R.C. 1985, *Ap.J.*, **297**, 309.

Pritchet, C.J. and van den Bergh, S. 1987, *Ap.J.*, **316**, 517.

Rowan-Robinson, M., Lawrence, A., Saunders, W., Crawford, J., Ellis, R., Frenk, C.S., Parry, I., Xiaoyang, X., Allington-Smith, J., Efstathiou, G. and Kaiser, N. 1990, *Mon.Not.R.astr.Soc*, **247**, 1.

Raychaudhury, S. and Lynden-Bell, D. 1989, *Mon.Not.R.astr.Soc.*, **240**, 195.

Rubin, V.C., Burstein, D., Ford, W.K. and Thonnard, N. 1985, *Ap.J.*, **289**, 81.

Ryden, B. and Gunn, J.E. 1987, *Ap.J.*, **318**, 15.

Sandage, A. 1986, *Ap.J.*, **307**, 1.

Sandage, A. 1987, *Ap.J.*, **317**, 557.

Shuter, W. 1989, *Dynamics and Interactions of Galaxies Roland Wielen (Ed.) Heidelberg.*

Tully, R.B. 1980, *nearby galaxy catalogue* (Cambridge: Cambridge University Press).

Turner, E.L. 1976, *Ap.J.*, **208**, 20.

White, S.D.M., Huchra, J., Latham, D. and Davis, M. 1983, *Mon.Not.R.astr.Soc.*, **203**, 710.

White, S.D.M. 1984, *Ap.J.*, **286**, 38.

Whitmore, B.C. 1990, *Dynamics and Interactions of Galaxies Roland Wielen (Ed.) Heidelberg.*

Zaritsky, D., Olszewski, E.W., Schommer, R.A., Peterson, R.C. and Aaronson, M. 1989, *Ap.J.*, **345**, 759.

Zwicky, F. 1933, *Helvetica Physica Acta*, **6**, 110.

3 GROUPS AND CLUSTERS

3.1 INTRODUCTION

Although the Universe is rich in structure over a very large range of scales, most cosmological studies focus either on bright galaxies or on rich clusters. There are practical reasons for this: magnitude limited surveys tend to pick out galaxies in a narrow range of luminosities while the identification of galaxy aggregates in projection requires the high contrast typical of the richest clusters. Only in recent years, have extensive redshift surveys over wide areas of sky made it possible to identify intrinsically poor groups of galaxies representing weak enhancements over the projected galaxy distribution. These groups are particularly interesting for cosmology since they probe scales intermediate between those of bright galaxies and those of rich clusters.

The variation of global properties along the sequence from individual galaxies to rich clusters has suggested a number of general arguments for or against particular cosmogonic models. For example, Peebles (1989) has invoked the apparent continuity between the mean distribution of mass around individual bright galaxies and the mean distribution of other galaxies around them to argue that, if galaxies formed as part of a dissipationless clustering hierarchy, then they must have collapsed by a modest factor at very early times. White (1990), on the other hand, has pointed out that while there does seem to be continuity in potential well depth from individual galaxies to rich clusters, there is a marked discontinuity in

the distribution of light (or radii) at galactic scales. This argues against formation within a dissipationless hierarchy and suggests instead substantial dissipative contraction of the material destined to become luminous at relatively recent epochs. Arguments of this sort are attractive because of their generality but must ultimately rely on detailed investigations of the observed properties of structures on a wide range of scales.

In hierarchical clustering theories, the mere existence of small groups of physically associated galaxies is a puzzle. N-body studies have demonstrated that substructure within virialized objects survives only for a very short period of time (Frenk *et al.* 1988, Efstathiou *et al.* 1988). If evolution is driven purely by gravity, bound structures at all levels of the hierarchy tend to become smooth and relaxed soon after they collapse. This result provides some evidence that the luminous component of galaxies must have undergone substantial dissipation if hierarchical clustering has anything to do with their formation. The properties of groups, in particular their observed abundance as a function of luminosity and velocity dispersion, provide important constraints on processes of this kind.

Another example of the interconnection between structures on different scales is Kaiser's (1984) suggestion that rich galaxy clusters owe their apparent clustering to the strong spatial correlations of peaks in Gaussian fields. He argues, in the context of hierarchical clustering models, that the linear theory autocorrelation function of rich clusters should be enhanced over that of galaxies by a constant factor which depends exclusively on the amplitude of the primordial fluctuation spectrum. A consequence of this model is that intermediate objects such as poor galaxy groups should exhibit intermediate clustering properties. Accurate measurements of correlations on various scales could, in principle, be used to infer the shape and amplitude of the primordial fluctuation spectrum.

In this chapter we study the properties of groups identified in the Center for Astrophysics (CfA) redshift survey (Davis *et al.* 1982, Huchra *et al.* 1982). One of our primary aims is to determine the abundance of bound structures as a function of their total luminosity, a sort of "universal" luminosity function of galactic systems which is independent of the detailed arrangement of luminous material within each structure. The motivation behind this is to provide an observational test of hierarchical gravitational clustering. Theoretical models, based either on N-body simulations or on analytic calculations such as those pioneered by Press and Schechter (1974), make fairly unambiguous predictions for the abundance of bound clumps on different scales, but not for the number of galaxies within them which is set by complicated dissipative phenomena. We also consider the distribution of group velocity dispesions, the luminosity-velocity dispersion relation and the group-group two-point correlation function. In hierarchical clustering models, the former depends primarily on the amplitude of primordial fluctuations and the latter on the shape of the fluctuation spectrum.

Although our approach is applicable in general to hierarchical clustering models, we focus exclusively on the comparison of our results with the cold dark matter (CDM) cosmogony, a particularly successful version of the hierarchical clustering hypothesis. The CDM model is based on the premises that the Universe contains a closure density of weakly interacting elementary particles and that primordial fluctuations are adiabatic and scale invariant. This cosmogony has been extensively investigated in recent years and has been shown to give good agreement with the observed internal properties and distribution of galaxies and galaxy clusters (Blumenthal *et al.* 1984, Davis *et al.* 1985, Frenk *et al.* 1985, 1988, 1989, White *et al.* 1987a, 1987b and other references therein). In this chapter we adopt the "standard CDM model" of Davis *et al.* 1985, in which $H_o = 50$ km s^{-1} Mpc $^{-1}$ and the biasing parameter (defined as the ratio of fluctuations in the galaxy and density distributions filtered with a "top hat" window function) is $b \sim 1.5 - 2.5$.

Several properties of groups in the CfA survey have been investigated in the past. Huchra and Geller (1982) and Geller and Huchra (1983) constructed the first catalogues of groups identified as number density enhancements over the local background. They concentrated on estimates of mass-to-light ratios. A similar analysis of two slices of the extended CfA survey (limited to $m_{B(0)} = 15.5$) was recently carried out by Ramella, Geller and Huchra (1989). They find that groups in this survey are brighter on average and have a higher median velocity dispersion than groups in the original CfA survey, even though both surveys span a comparable volume. The methods introduced by Geller and Huchra were refined and extended by Nolthenius and White (1987; hereafter NW) who used N-body simulations to help determine the optimal group selection procedure in the CfA survey. NW reconsidered the issue of mass-to-light ratios and discussed various other properties of groups both in the CfA survey and in N-body simulations of CDM Universes. Prior to the work of NW, estimates of the group luminosity function had been based on groups identified in projection (Gott and Turner 1977, Bhavsar 1978, 1980.) A previous attempt at constructing a "universal" luminosity function, similar in spirit but not in detail to our own, was carried out by Bahcall (1979) who combined the group catalogue of Turner and Gott with the rich cluster catalogue of Abell (1958).

The present study employs similar techniques to those of NW to identify groups; it amplifies and complements their analysis. To compare with CDM models we use the simulations described by Frenk *et al.* (1989) which are larger and, in some ways, more sophisticated than the simulations of Davis *et al.* (1985) that NW analysed. In Section 3.2 below we briefly review the technique for finding groups in the CfA survey and in the simulations. In Section 2.3 we introduce the concept of the luminosity function of galactic systems and estimate it for our data. In Sections 3.4 we discuss the distribution of group velocity dispersions, the luminosity-velocity dispersion relation and the group-group correlation function. Finally, in Section 3.5 we present a general discussion and summary of our main

results. We parameterise the Hubble constant as $H_o = 50h_{50}$ km s^{-1} Mpc $^{-1}$ and, unless otherwise stated, give our numerical results for $h_{50} = 1$.

3.2 GROUP SELECTION

A slice of declination $0^o < \delta < 35^o$ through the Northern portion of the CfA survey is illustrated in the "pie" diagram of Figure 3.1*a*, in which observed velocity v is plotted as the radial coordinate and right ascension α as the angular coordinate. Gravitationally bound associations show up as clumps which are elongated towards the observer. Rich clusters like Virgo ($v \sim 1000$ km s^{-1}, $\alpha \sim 13$ hrs) or Coma ($v \sim 7400$ km s^{-1}, $\alpha \sim 13$ hrs), give rise to prominent "fingers of God", but groups with smaller velocity dispersions are much harder to identify visually. The purpose of a group finding algorithm is to pick out these elongated structures taking into account the variation in galaxy number density with distance. The optimal algorithm is one that maximizes the identification of physically associated members while minimizing contamination due to chance alignments along the line-of-sight. It is clear from Figure 3.1*a* that constructing such an algorithm is not straightforward. However, artificial galaxy catalogues in which the distance to each galaxy is known, in addition to its redshift, provide a valuable benchmark against which the performance of a given algorithm can be assessed. In this chapter we rely heavily on catalogues constructed from N-body simulations evolved from cold dark matter initial conditions. Such catalogues have similar low order correlations to those of the CfA survey. They therefore serve the dual purpose of providing the desired benchmark while, at the same time, enabling a direct comparison of the theory with observations. Note that in their former role their utility is independent of the validity of the CDM model. Before describing our group finding procedure we briefly discuss the way in which our artificial "CfA" catalogues are generated.

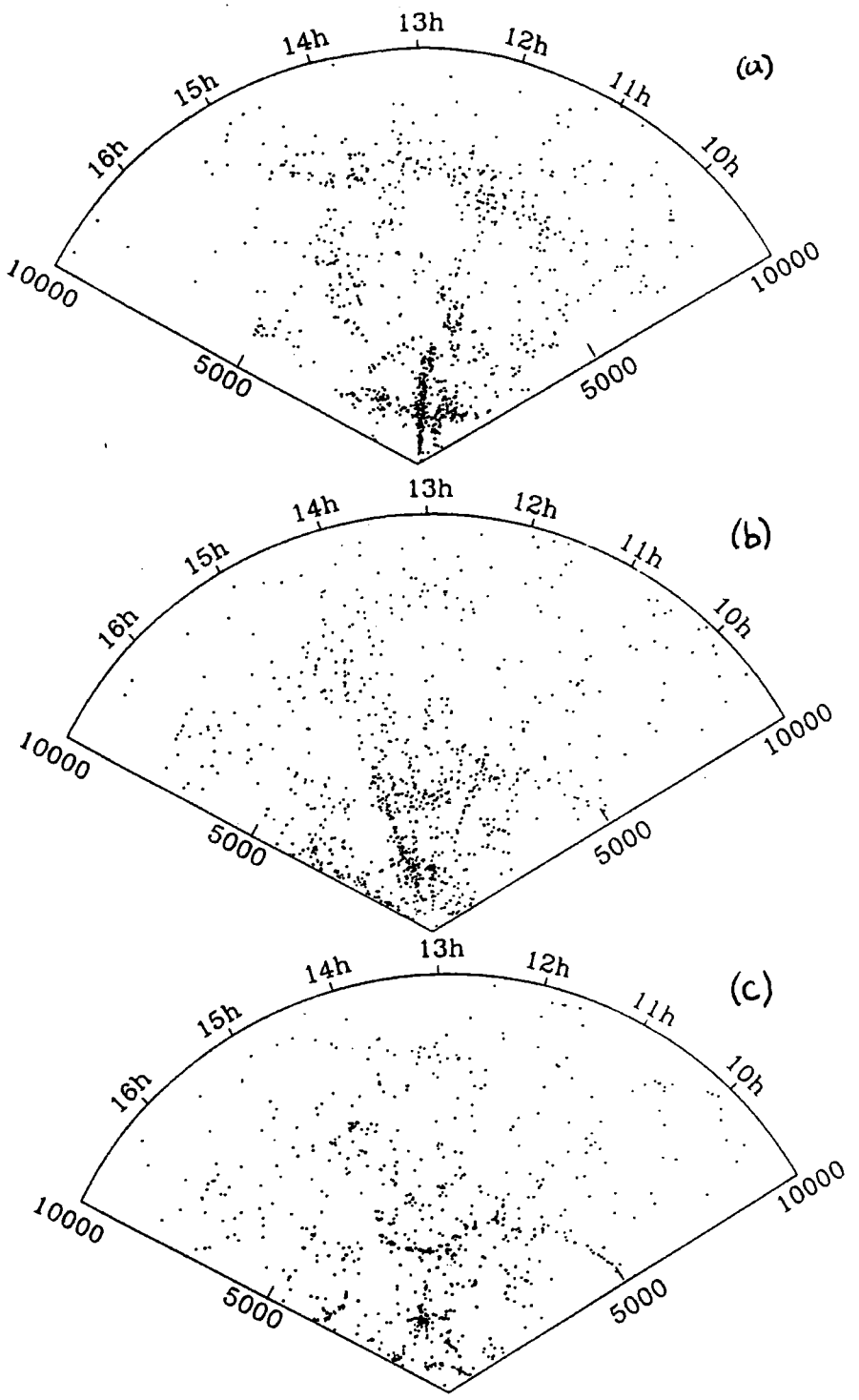


Figure 3.1 *(a)* A slice of declination $0^o < \delta < 35^o$ through the Northern portion of the CfA survey with observed velocity plotted as the radial coordinate and right ascension as the angular coordinate. *(b)* A similar slice through one of the $b = 2$ CDM simulations and *(c)* the same simulation with galaxies plotted using the true distance rather than velocity as the radial coordinate.

3.2.1 Artificial galaxy catalogues

We used the five CDM N-body simulations described by Frenk *et al.* (1989) to extract "galaxy" catalogues analogous to the CfA survey. Each simulation represents a comoving cube of a flat Universe of present size 360 Mpc containing 262144 particles. Results were analysed assuming three different values for the amplitude of the primordial fluctuation spectrum, corresponding to values of the biasing parameter $b = 1.6$, 2.0, and 2.5.

"Galaxies" were identified in the simulations using the "high peak" model as described and implemented by White *et al.* (1987b) and Frenk *et al.* (1989). Peaks of the initial density field lying above a global density threshold are identified as the sites of galaxy formation. A magnitude limited catalogue is then constructed by assuming a universal luminosity function. For most of our analysis we adopted the luminosity function estimated from the CfA survey itself by Davis and Huchra (1982). This has a Schechter form (*c.f.* equation (3.6) below), with $\alpha = -0.9$ and $L_* = 3.3 \times 10^{10}$ (for $H_o = 50$ km s^{-1} Mpc^{-1}). Catalogues with the luminosity function proposed by Felten (1985), ($\alpha = -1.25$ and $L_* = 4.5 \times 10^{10}$) were also generated as a test, but, as discussed below, our results are insensitive to this choice. Our catalogues include the effect of galactic obscuration, modeled according to the modified cosecant law of Fisher and Tully (1981). For most of the analysis our catalogues were limited at the same "apparent magnitude" as the CfA survey, 14.55 in the Zwicky system.

With our assumed luminosity functions, models with the value $b = 2.5$ adopted by Davis *et al.* (1985) are slightly affected by the undersampling problem described by White *et al.* (1987b). This causes $\sim 20\%$ of single particles, mostly near the centres of nearby clumps, to be associated with more than a single galaxy. Although this is not a serious problem for our analysis, we illustrate most of our

results using the $b = 2.0$ models for which only $\sim 10\%$ of the galaxies are undersampled. As we discuss below, many of the groups' properties we consider are quite insensitive to the choice of b.

Our simulated catalogues have similar power spectra, correlation functions, and other global properties to those of model E36 of White *et al.* (1987b), illustrated in Figures 8-10 of that paper. From each "all sky" catalogue we extracted all "galaxies" lying within the cone corresponding to the Northern region of the CfA survey. We generated 12 such surveys from the 5 different simulations for each of our adopted values of b. The mean number of galaxies in the artificial catalogues is very similar to that in the Northern CfA survey. For example, for $b = 2$ we find 1730 ± 160 "galaxies" compared to 1816 galaxies in the corresponding region of the real survey. Groups were identified and analysed in exactly the same way in our artificial and real catalogues. (In the latter case, we analyze both the Northern and Southern portions of the survey.)

Figure 3.1*b* shows a slice from one of our simulations (with $b = 2$) with the same geometry as the CfA slice in Figure 3.1*a*. The gross features of these two diagrams are clearly quite similar. A detailed analysis shows that the two and three point correlation functions are similar in the simulations and in the real data, although in the model the pairwise galaxy velocities are slightly too high. (DEFW). Figure 3.1*c* shows the same slice as Figure 3.1*b*, but with true distance, rather than velocity, as the radial coordinate. Comparison of these two plots shows the way in which groups are distorted in velocity space. The strategy we adopt below is to tune our group finder so that the groups found in velocity space resemble those found in real space as much as possible.

3.2.2 Groups in the CfA survey

Our procedure for selecting groups in the CfA survey is similar to that described by NW. Groups are defined as aggregates of three or more galaxies picked

out by a "friends-of-friends" or percolation grouping algorithm. This is roughly equivalent to picking objects whose density contrast relative to the local mean density is above some chosen minimum. This minimum determines the typical size of groups and their abundance. It should be high enough that most groups are single dynamical entities, but small enough that it produces a large sample of groups. The linking length in the "friends-of-friends" algorithm must allow for the facts that the galaxy sample is magnitude limited and that redshift separations rather than true distance separations are known. The former requires specifying a linking parameter which varies with redshift and the latter requires specifying separate linking parameters in angular distance and in redshift separation.

We can see how the procedure works as follows. Suppose first that true distance separations were known. Then, to obtain groups approximately bounded by isodensity contours, the linking parameter, r_L, should scale inversely with the one-third power of the local density:

$$r_L \quad \propto \quad \left(\int_{L_F}^{\infty} \Phi(L) dL \right)^{-1/3} \tag{3.1}$$

where $\Phi(L)$ is the galaxy luminosity function of the sample and L_F is the luminosity corresponding to the apparent magnitude limit at each redshift. When the depth information comes from the observed velocities, this scaling must be modified to take into account the smearing of positions along the line-of sight. Let us denote the linking cut-offs in angular separation and in line-of-sight velocity by D_L and V_L respectively. NW suggest that D_L should be taken to scale as the inverse square root of the group's surface density which leads to

$$D_L \quad = D_0 \left(\frac{\int_{L}^{\infty} \Phi(L) dL}{\int_{L_F}^{\infty} \Phi(L) dL} \right)^{-1/2} \left(\frac{V}{V_F} \right)^{-1/3} , \tag{3.2}$$

where V_F is a fiducial distance which we take to be 1250 km s^{-1}. They also argue that V_L should scale as the velocity dispersion of a typical group picked out at each distance since the luminosity and presumably, therefore, the velocity dispersion of the groups picked out increases with distance. The expected variation of group velocity dispersion with distance can be estimated by examining groups in artificial catalogues constructed in *real* space. To identify groups in real space we used the linking parameter of equation (3.1), with the constant of proportionality adjusted so that the median density contrast was 35. In agreement with NW, we find that this procedure leads to:

$$V_L = V_0 + 0.03(V - 5000)\mathrm{km\,s^{-1}} \tag{3.3}$$

where V is the mean velocity of the two galaxies to be linked. The scaling in equation (3.3) is fairly insensitive to the median density contrast of the groups in the artificial catalogues, but the value of V_0 does depend on it.

Equations (3.2) and (3.3) completely specify our grouping algorithm. There are two free parameters, D_0 and V_0. Most of the properties of interest here are fairly insensitive to D_0. We adopt $D_0 = 0.35$, a value which is effectively 15% smaller than that adopted by NW, but which avoids linking several small outlying groups to the main body of the Virgo cluster. The choice of V_0 is more difficult. As might be expected, it has a strong effect on the distribution of group velocity dispersions. Too small a value excludes galaxies with large relative velocities, artificially depressing the measured velocity dispersion. Too large a value, on the other hand, links non-members into the group, artificially increasing its velocity dispersion. We proceed by comparing artificial group catalogues in real and redshift space. We assume that the "true" velocity dispersion distribution is that obtained for groups (with three or more members) identified in real space. This is plotted in Figure 3.2 as solid triangles; it is fairly insensitive to the choice of proportionality constant in equation (3.1). Velocity dispersion distributions for

groups identified in redshift space in our artificial catalogues for various values of V_0 are compared to the true distribution in Figure 3.2. (These distributions do not depend sensitively on the adopted value of D_0.) The left hand panel gives ensemble averages over 12 catalogues with $b = 2.5$, whereas the right hand panel gives similar averages for $b = 1.6$. Each group is weighted by $1/\Gamma_{max}$, where Γ_{max} is the maximum volume out to which the third brightest galaxy in the group is above the magnitude limit of the catalogue. As V_0 is increased from 350 km s^{-1} to 950 km s^{-1} the distributions change systematically with the median velocity dispersion increasing by $\sim 60\%$. The total density of groups remains approximately constant due to the competing effects of merging previously distinct groups and identifying new associations.

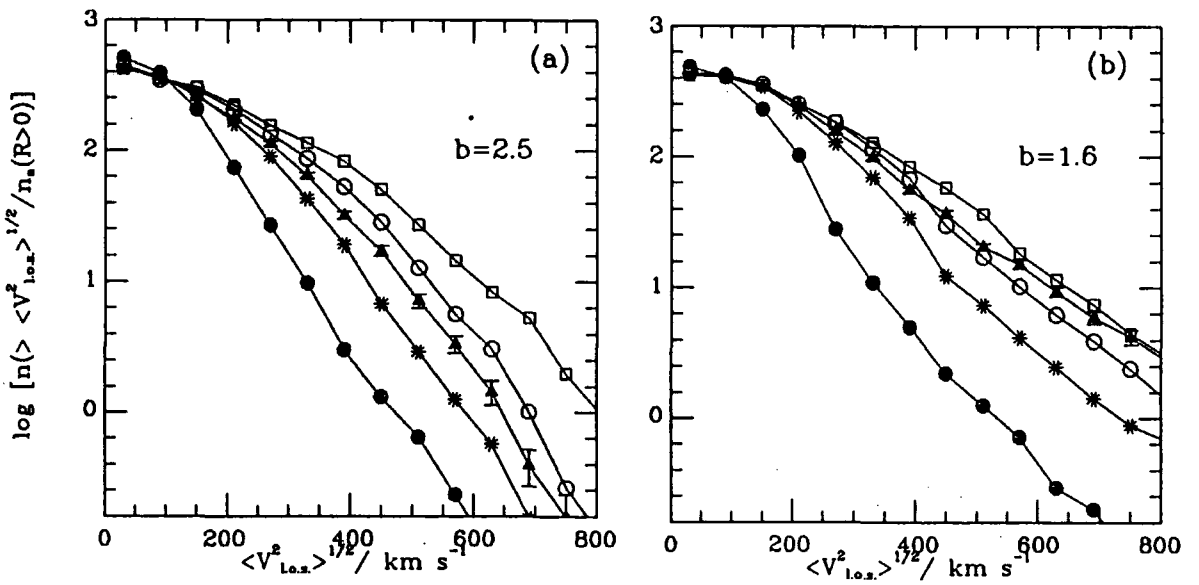


Figure 3.2 The cumulative distribution of velocity dispersions for groups found in projection with values of $V_0 = 350, 550, 750$ and 950kms^{-1} (closed circles, stars, open circles and open squares respectively). The closed triangles show the distributions for groups found in real space by applying equation (3.1) to the true galaxy positions. Each curve is from 12 averaged catalogues from *(a)* $b = 2.5$ and *(b)* $b = 1.6$ simulations.

The median velocity dispersion in our catalogues in real and redshift space and their dependence on the bias parameter b are summarised in table 3.1. As expected, the dispersions increase systematically with both V_0 and b. This table and Figure 3.2 suggest that the appropriate value of V_0 needed to match the real space data is somewhere between 550 and 950 km s^{-1}. In this range, the number of galaxies in common, grouped in the real and redshift space catalogues respectively, is approximately constant, at $\sim 75\%$. We adopt as our standard parameters the following values: $D_0 = 0.35$, $V_0 = 550$ km s^{-1}. Our choice of V_0 is somewhat larger than the value 350 km s^{-1} adopted by NW who carried out a similar comparison to ours but used unbiased CDM open models and focussed on somewhat different statistics than we have. Our adopted value of V_0 gives a reasonable match between the distributions in real and redshift space for $b \geq 2$, but underestimates the high dispersion tail in the $b = 1.6$ case. Figure 3.2, however, emphasizes this tail where the abundance is rapidly declining. In fact, with our adopted value of V_0 the true median dispersion for the $b = 1.6$ case is underestimated by only $\sim 10\%$.

Table 3.1 Median velocity dispersions

V_o/km s^{-1}	$\sigma_{b=2.5}$/km s^{-1}	$\sigma_{b=2.0}$/km s^{-1}	$\sigma_{b=1.6}$/km s^{-1}
350	166	170	174
550	224	244	265
750	262	287	308
3-d	230	265	303

We now illustrate graphically the performance and limitations of our procedure. In Figure 3.3*a* we show the same real space slice as in Figure 3.1*c* from one of our $b = 2$ CDM simulations. Groups identified in this slice using equation (3.1)

are traced by joining the galaxies linked together by our group finding algorithm. The groups are quite concentrated and aspherical, and show no preferred orientation. In Figure 3.3*b* we plot this slice in redshift space indicating the groups of Figure 3.3*a* (identified in *real* space). In redshift space, the groups become drawn out along the line of sight. It is these "fingers of God" which our group finder in redshift space is designed to pick out. In Figure 3.3*c* we show the groups actually found by our algorithm in redshift space, using our adopted parameter values. Comparison with Figure 3.3*b* illustrates how well our procedure works. Most of the "true" groups in Figure 3.3*b* are also identified in redshift space, although their detailed memberships are not identical. In Figure 3.3*d* we show the groups found in redshift space, but now plotted in real space. This should be compared with Figure 3.3*a*. The majority of the groups are found in both plots, but residual contamination by non-members is apparent in Figure 3.3*d*. Thus, some velocity dispersions will be slightly overestimated as a result of finding groups in redshift space, but this effect does not appear to be excessively strong. Figures 3.3*e* and 3.3*f* show the groups found in redshift space with our adopted value for D_0, but with $V_0 = 350$ km s^{-1}, plotted in redshift and real space respectively. Finally, Figures 3.3*g* and 3.3*h* are similar to Figures 3.3*e* and 3.3*f*, but with V_0 set to 750 km s^{-1}. Comparing Figures 3.3*e* – *h* with Figures 3.3*c* and 3.3*d* shows that the smaller value of V_0 excludes true members with large relative velocities and splits large groups into smaller ones while the larger value of V_0 gives rise to excessive contamination by non-members. Our adopted value $V_0 = 550$ km s^{-1} is a compromise between these two undesirable extremes. In our opinion, this is the best that can be done given the limitation of working with redshifts rather than distances. (Note that, as with Figure 3.2, these plots tends to emphasize the few groups in the high dispersion tail of the distribution and should, therefore, be interpreted in conjunction with the median values given in Table 3.1.)

Having used our N-body simulations to tune the group finding procedure, we now apply it to the CfA survey. Figure 3.4*a* shows the groups identified in the

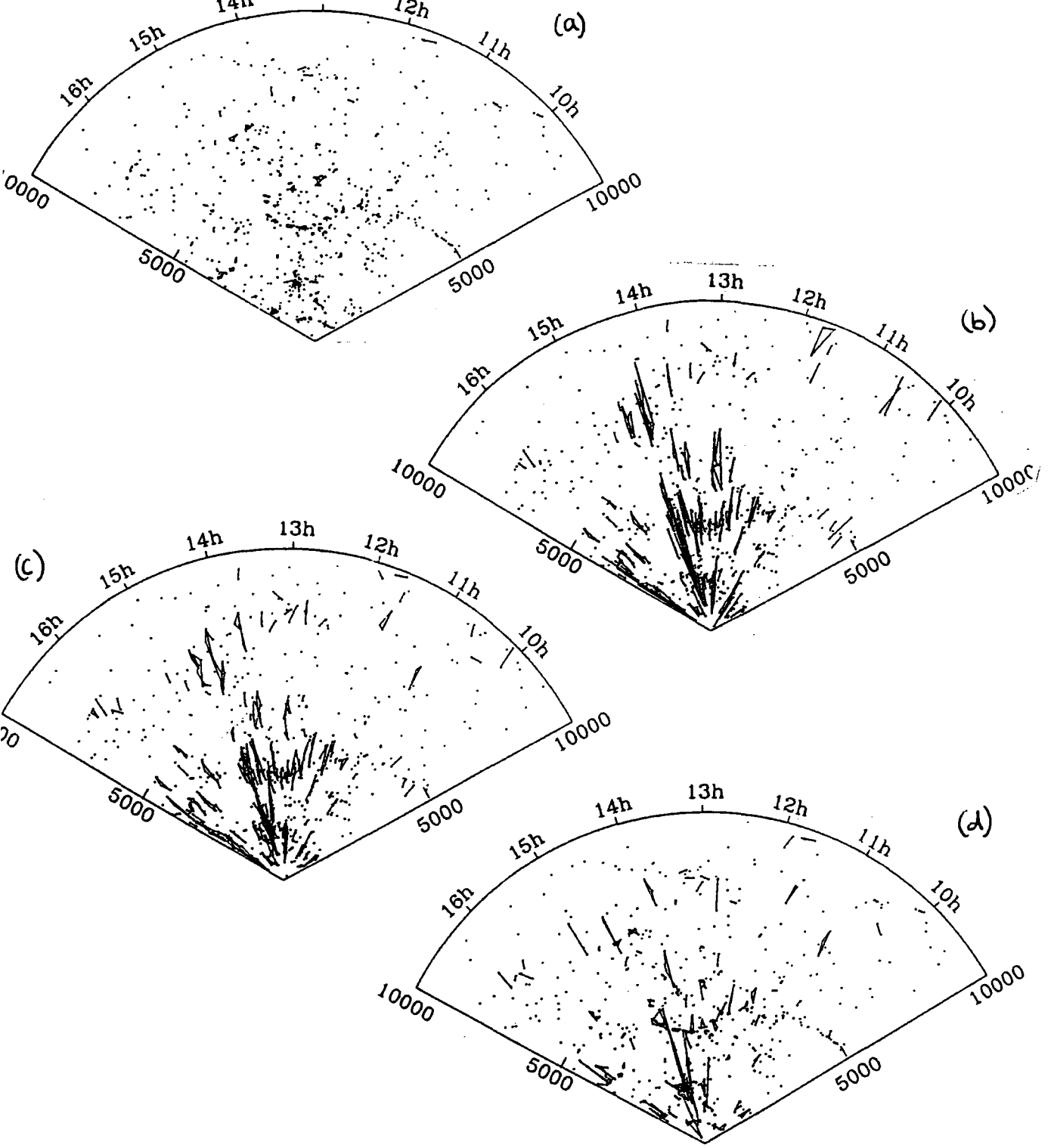


Figure 3.3 *(a)* Shows the same real space slice as Figure 3.1c. Groups have been identified in this slice using equation (3.1) and are traced by joining the galaxies linked together by our group finding algorithm. *(b)* The same groups as *(a)* but plotted in redshift space. *(c)* The groups identified in redshift space using equations (3.2) and (3.3) with our standard parameters are plotted in redshift space and *(d)* real space for comparison with *(b)* and *(a)* respectively.

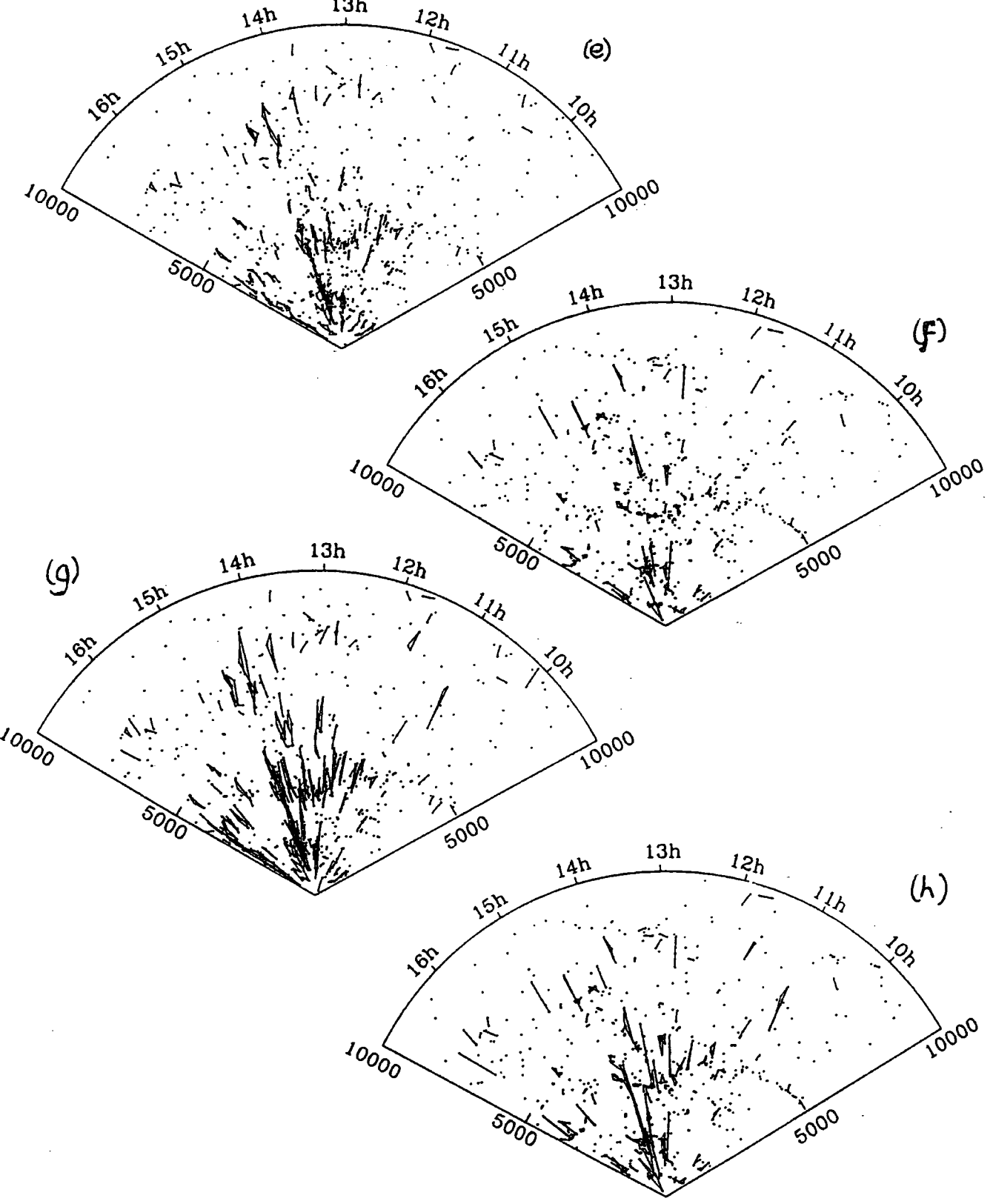


Figure 3.3 *(e)* Groups found in redshift space with $V_0 = 350\text{km s}^{-1}$ plotted in redshift space and *(f)* plotted in real space. *(g)* Groups found in redshift space with $V_0 = 750\text{km s}^{-1}$ plotted in redshift space and *(h)* plotted in real space.

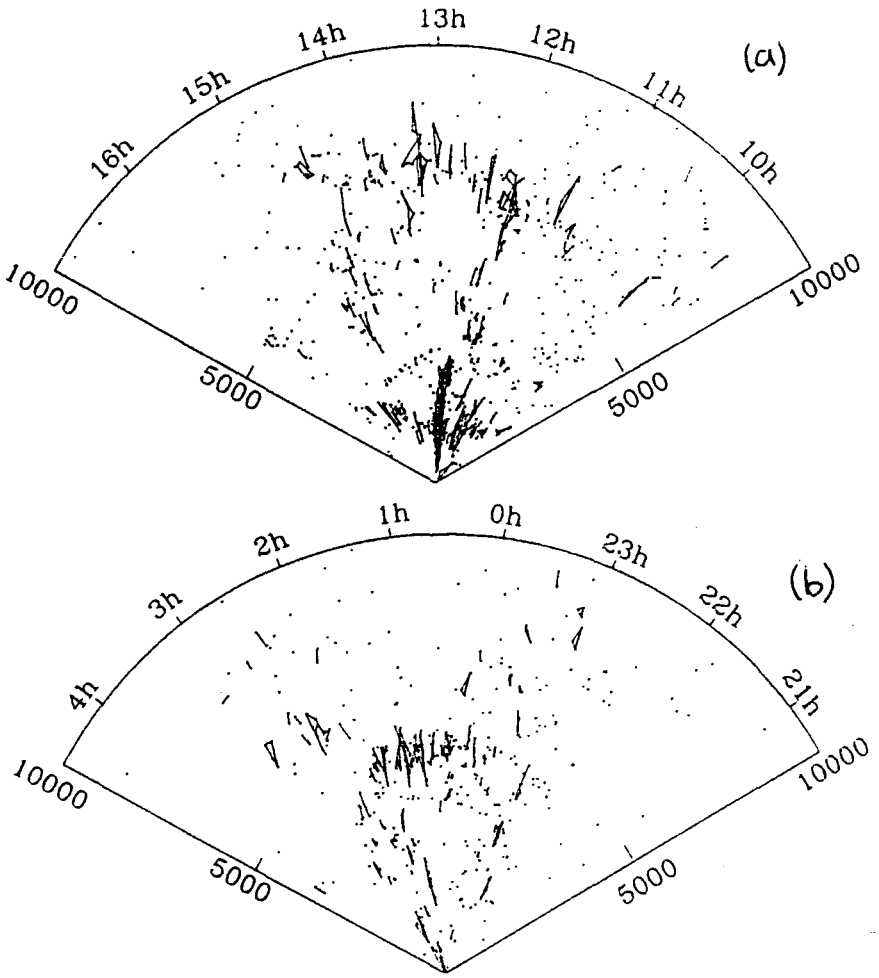


Figure 3.4 *(a)* The same slice as Figure 3.1a with groups identified in the Northern and *(b)* the Southern CfA survey using our adopted values for V_0 and D_0.

slice plotted in Figure 3.1*a* while Figure 3.4*b* shows a similar slice in the South. The most visible cluster in the North is Virgo which has about 200 members and extends out to 3000 km s^{-1}; Coma and Abell 1367 can be also be seen at $\sim$ 7000 km s^{-1}. There are fewer clusters in the South with several of them concentrated in a large supercluster at $\sim$ 5000 km s^{-1}. (The large structure in the South subtends an angle of $\sim 30^o$ on the sky and is part of the "Great Wall" discussed by Geller and Huchra, 1989.) Over the entire CfA survey we find 166 groups

with at least 3 members, which include 55% of all galaxies; 13% of galaxies lie in pairs and 32% are singles. The median density contrast of the groups is 50, the median line-of-sight velocity dispersion is 155 km s^{-1} and the median crossing time is $0.43H_o^{-1}$. (See NW for the exact definitions of these quantities and for a discussion of other related properties of the CfA groups.) Note that properties derived from the size and velocity dispersion of groups are strongly dependent on the selection cut-offs adopted. We have argued that our adopted value of V_0 provides a reasonable compromise between artificially truncating the velocity distribution of the richer groups and minimizing contamination by non-members. The angular linking parameter, D_0, is less strongly constrained. In Table 3.2 we show how the CfA group properties vary as this parameter is changed, so as to obtain group catalogues with density contrasts ranging between 5 and 644. The third column in this table gives the percentage of galaxies grouped in each case, the fourth column gives the number of groups with ≥ 3 members, and the fifth column gives median velocity dispersions. The remaining columns will be discussed below. Note that with our standard parameters, the median dispersion of the CfA groups is significantly smaller than the corresponding values in our artificial catalogues, even for $b = 2.5$. We discuss this discrepancy further in Section 3.4.

Table 3.2 Group properties

D_L	$d\rho/\rho$	% grouped	n_g	σ_{median}	M/L_z	β	γ	M_z^*	$\frac{\phi^*}{(10^{-4}\,Mpc^{-3})}$
0.65	5	77	147	178	199	0.90	1.2	-21.4	2.6
0.55	12	71	151	169	181	1.05	1.4	-22.0	2.2
0.45	30	64	169	161	176	1.12	1.6	-22.1	2.7
0.35	50	55	166	155	149	1.16	1.7	-22.1	3.0
0.25	154	49	145	141	123	1.36	2.0	-22.5	2.4
0.15	644	33	110	132	96	1.52	2.8	-22.8	1.9

3.3 THE LUMINOSITY FUNCTION OF GALACTIC SYSTEMS

3.3.1 Groups in the CfA survey

In this section we discuss our method for estimating the *total* luminosity, L_{tot}, of groups identified in magnitude limited surveys. We begin by assuming a universal galaxy luminosity function, $\Phi(L)$. The observed galaxies in each group are then a random sample of that portion of $\Phi(L)$ which lies above the local magnitude limit. We define L_{tot} as the sum of the luminosities of these galaxies plus the integrated luminosity of galaxies below the magnitude limit of the survey. Thus L_{tot} provides a measure of the absolute luminosity of groups independently of the magnitude limit of the survey in which they are identified. Note that even single galaxies must be regarded as the sole member above the apparent magnitude threshold of a group. We refer to the distribution of L_{tot} as the luminosity function of groups or "galactic systems." Let L_{obs} be the luminosity of galaxies included in the catalogue and L_{cor} the luminosity of galaxies fainter than the local absolute

magnitude limit of the catalogue, $M_{lim} = -5\log v + 5\log H_o - 10.45$. Then, for each group,

$$L_{tot} = L_{obs} + L_{cor}. \tag{3.4}$$

To estimate the contribution of galaxies present in the catalogue we take the N_{obs} members of a group to be at the mean redshift distance of the group. Then,

$$L_{obs} = \sum_{i=1}^{N_{obs}} L_i. \tag{3.5}$$

To estimate the contribution of galaxies below threshold we assume that group members are independently drawn from a group luminosity function with universal Schechter form,

$$\Phi(L)dL \propto L^{\alpha} exp(-\frac{L}{L^*})dL. \tag{3.6}$$

We normalise this function to give the observed number of galaxies in each group. The expected luminosity of faint group members is then

$$L_{cor} = N_{obs} \frac{\int_0^{L_{lim}} L\Phi(L)dL}{\int_{L_{lim}}^{\infty} \Phi(L)dL}, \tag{3.7}$$

where

$$L_{lim} = 10^{0.4 \times (M_{\odot} - M_{lim})}. \tag{3.8}$$

We take $M_{\odot} = 5.48$ and adopt the parameters for the luminosity function in equation (3.6) derived by Davis and Huchra (1982) from the CfA catalogue $\alpha = -0.9$ and $L_* = 3.3 \times 10^{10} \mathrm{L}_{\odot}$ ($h_{50} = 1$).

Out to the typical group distance of 5000 km s^{-1}, the mean add on luminosity L_{cor}, is relatively small, beyond that it increases rapidly. This is illustrated in Figure 3.5 which shows the add on luminosity for a single galaxy as a function of distance. At 5000 km s^{-1}the add on luminosity per galaxy is $\sim 0.5\, L_*$; at

9000 km $s^{-1}$the correction has increased to 7 L_*. Also plotted in Figure 3.5 is the mean group luminosity per galaxy, $\tilde{L}_{obs}$, contributed by galaxies brighter than the magnitude limit. At 7000 km s^{-1}, the correction is 100% and increases to 800% at 9000 km s^{-1}. For most of our analysis we shall consider only galaxies up to a distance of 7000 km s^{-1}, where the correction for a single galaxy, $\sim 1L_*$ is still reasonably small. This limit excludes the most distant 343 galaxies of the survey. We shall however, examine the effect of varying this cutoff in some of the tests described below.

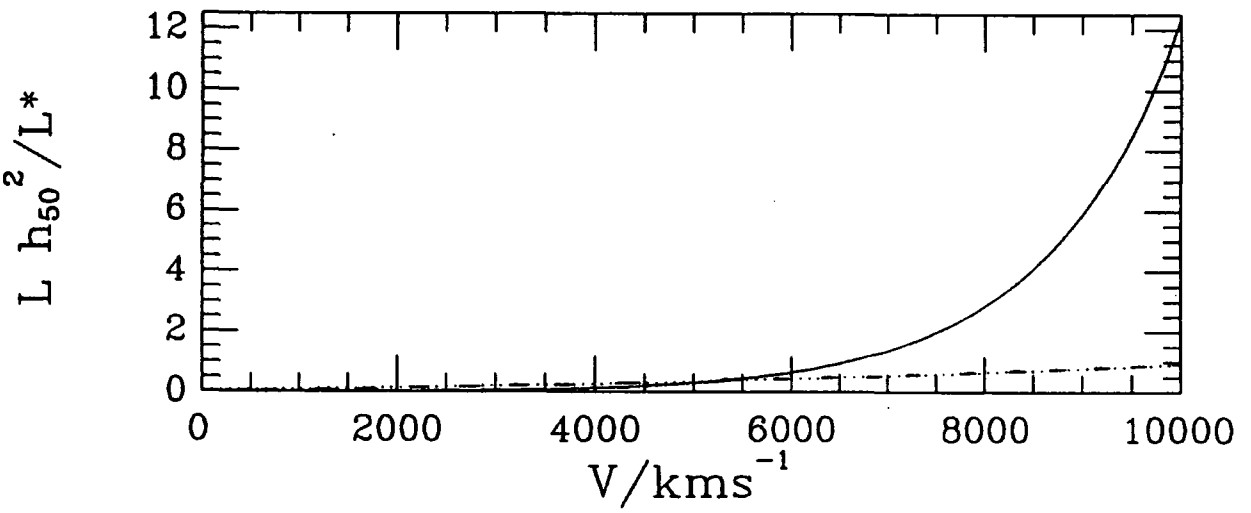


Figure 3.5 The additive luminosity correction as a function of distance to a single galaxy in units of L^*. At our velocity cutoff of 7000kms^{-1} the correction per galaxy is $\sim 1L_*$. The dashed line shows the mean luminosity per galaxy, contributed by galaxies brighter than the magnitude limit at distance v.

An immediate test of our procedure so far is to compare our values of L_{tot} for Virgo and Coma with direct estimates of their total luminosity from deep photometry. Figure 24 of Sandage *et al.* (1985) shows that the cumulative luminosity

of galaxies brighter than $B_T = 18$, within a 4×4^o region around the core of Virgo converges to $L_B = 2.4 \times 10^{12} L_\odot$, for an "effective" Hubble constant (for their assumed flow model) of 65 km s^{-1} Mpc^{-1}. Correcting to the Zwicky system (with the photometric transformation $B_T = m_z - 0.29$ given by Efstathiou *et al.* 1988) and to our adopted value of $H_o = 50$ km s^{-1} Mpc^{-1}, this corresponds to $L_z = 3.1 \times 10^{12} L_\odot$. Our grouping algorithm associates 212 galaxies with Virgo in the same 4×4^o region; our estimated total luminosity is $L_z = 3.4 \times 10^{12} L_\odot$. This good agreement is not surprising since we adjusted our angular linking parameter D_0, to avoid merging outlying groups with the main body of Virgo. Furthermore, the luminosity correction at the distance of Virgo is still quite small. A more stringent test of our grouping algorithm and of the validity of our luminosity correction is provided by Coma. This cluster lies at over 7000 km s^{-1}, where our correction is quite large. Our algorithm finds only 23 galaxies in Coma, from which we estimate $L_{tot} = 7.9 \times 10^{12} L_\odot$, 4 times the observed value. From Godwin's (1983) sample of 6724 galaxies in Coma complete to $m_b = 21.0$, we find a total luminosity of $L_z = 7.3 \times 10^{12} L_\odot$ in a 2×2^o square around the core of Coma. The excellent agreement between our inferred value and the direct estimate suggests that our correction procedure is working well and that our fiducial cutoff at 7000 km s^{-1}is quite conservative.

To estimate the group luminosity function, Φ_g, we obtain the space density of groups in the i^{th} absolute magnitude bin by weighting each group by $1/\Gamma_{max}$, where Γ_{max} is the maximum volume within our velocity cutoff out to which the luminosity of the brightest group member is above L_{lim}, *i.e.*

$$\Phi_g(L_i) = \sum_j 1/\Gamma_{max}, \tag{3.9}$$

where the sum extends over all groups with luminosity in the i^{th} bin and

$$\Gamma_{max} = \left(\frac{\omega}{3}\right)\left(V_{lim} H_o^{-1}\right)^3 \left[1 - \frac{3}{2}(1 + q_o)\frac{V_{lim}}{c}\right] \tag{3.10}$$

Here $V_{lim}H_o^{-1}$ is the smaller of 7000 km s^{-1}or the maximum recession velocity at which the brightest galaxy in the group would be brighter than the magnitude limit; ω is the solid angle of the catalogue, c is the speed of light, and q_o is the deceleration parameter which we set equal to 0.5.

Our estimate of the CfA group luminosity function is shown as a histogram in Figure 3.6*a*. In Figures 3.6*b* – *d*, we decompose this luminosity function into separate contributions from groups in which only single galaxies, pairs, or three or more members are brighter than the magnitude limit. The faint end of the group luminosity function is made up almost entirely of single galaxies; binaries link these to the steep bright tail of richer groups and clusters. There are three rich clusters of Abell richness $R \geq 1$ in the CfA survey, Virgo, Coma and A1367, but only Virgo and A1367 are included within our fiducial cutoff at 7000 km s^{-1}. The Northern and Southern surveys give consistent results, although there is a slight excess of faint nearby groups in the North associated with the Local Supercluster. (This is clearly visible in Figure 3.6.) The group luminosity function is reproduced in Figure 3.7 (filled circles), where we also show the result obtained without applying our luminosity correction (open circles). The correction shifts the bright end of the luminosity function by ~ 1 mag. For comparison, we show the field galaxy luminosity function for the CfA survey (dashed line). A Schechter function (equation (3.6) provides a good fit to the group luminosity function over a decade in absolute magnitude but is too steep at the bright end (dotted line). We find that a good fit over the entire range is provided by a "Marshall" (1987) double power law,

$$\phi(L)dL = \phi^{*'}[(\frac{L}{L_*})^{\beta} + (\frac{L}{L_*})^{\gamma}]^{-1}\frac{dL}{L_*}, \tag{3.11}$$

with best-fit parameters $\beta = 1.16 \pm 0.07$, $\gamma = 1.7 \pm 0.12$, $M^{*'} = -22.2 \pm 0.2$ and $\phi^{*'} = 3.0 \pm 0.5 \times 10^{-4}$ Mpc^{-3}. This is shown as a solid line in Figure 3.7.

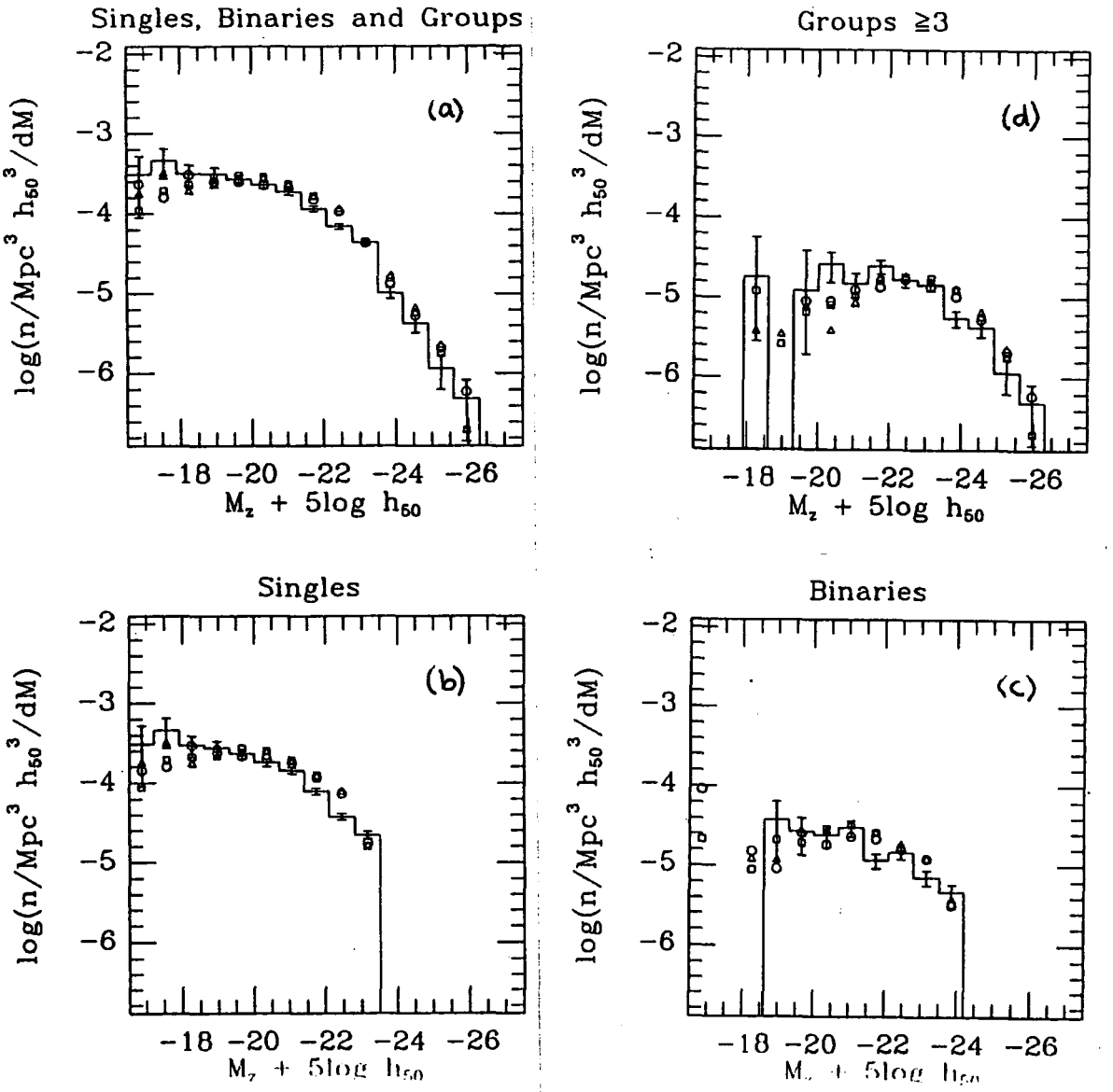


Figure 3.6 The CfA group luminosity function is shown as a histogram in *(a)*. The contributions from single galaxies, pairs and groups with 3 or more members are shown in *(b)*, *(c)* and *(d)* respectively. Open circles, squares and triangles represent the CDM simulations with values of the biasing parameter $b = 2.5, 2.0$ and 1.6 respectively.

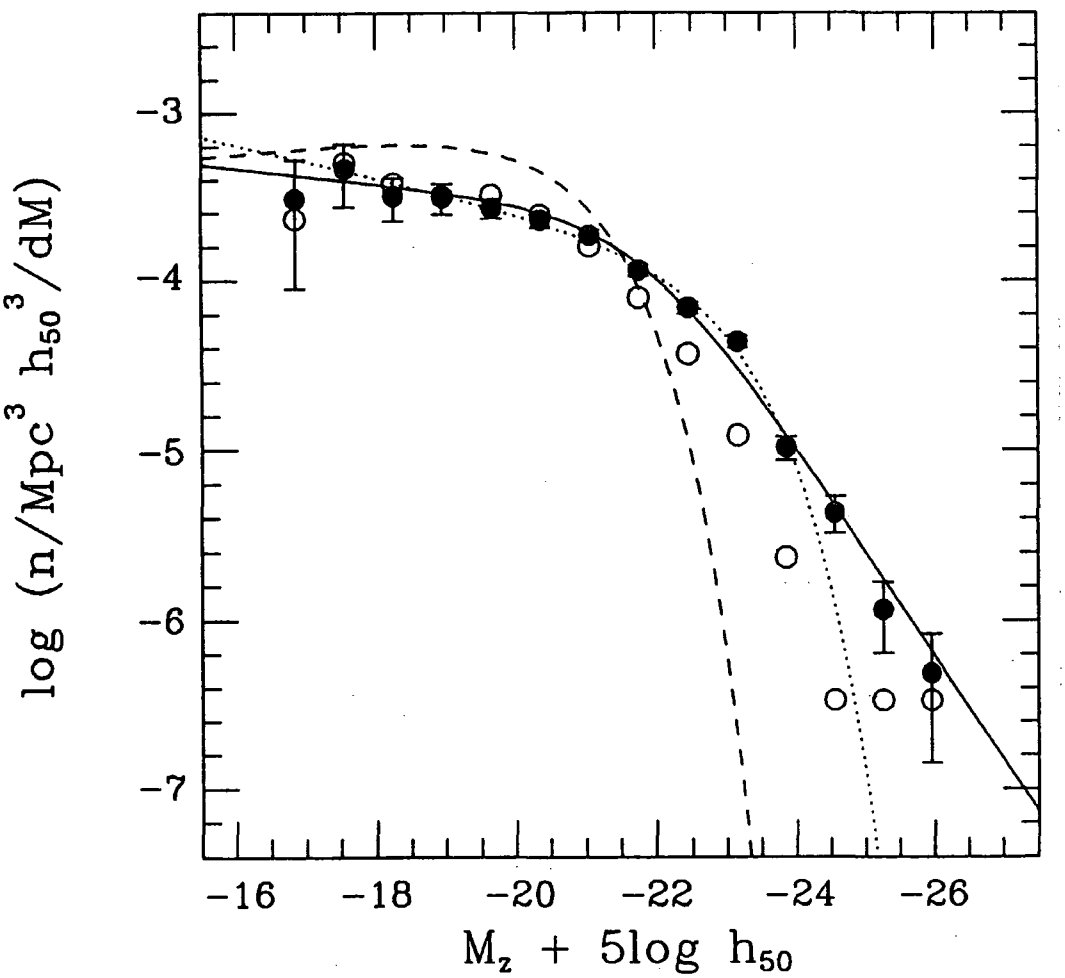


Figure 3.7 The open and filled circles show the CfA group luminosity function before and after the correction to their luminosity has been made. The solid line shows the best fitting double power law to the data. The best fitting Schechter function is plotted as a dotted line. The field galaxy luminosity function with our adopted parameters is shown for comparison. 1σ error bars are given.

The group luminosity function depends on the way in which groups are defined and identified in a galaxy catalogue. We have argued that our definition of groups as aggregates bounded by a surface of approximately constant density equal to a few tens the mean galaxy density is physically reasonable. Furthermore,

we have also argued that our procedure for identifying such groups in the CfA survey minimizes contamination by non-members and gives the best correspondence between simulated groups found using observational quantities and those found using full three-dimensional information. There is some freedom, however, in the choice of grouping parameters which satisfy these conditions. We have therefore obtained luminosity functions for group catalogues constructed at median density contrasts varying between 5 and several hundred by varying the parameter D_0 while keeping V_o and V_F fixed at 550 km s^{-1} and 1250 km s^{-1} respectively. The best-fit parameters for the double power-law luminosity function are listed in Table 3.2. The characteristic group luminosity, M_*, is fairly insensitive to the grouping parameters, but the slopes at the bright and faint ends vary systematically with the median density contrast and thus with the fraction of galaxies grouped. Also shown in Table 3.2 are median values of the virial mass-to-light ratio for our catalogues. Virial masses were obtained from equation (A5) of NW while total luminosities (in the Zwicky system) include the correction of equation (3.4). With our adopted transformation to the B_T, $(M/L)_{B_T} = (M/L)_{Zw}/1.31$. For groups with $(\delta\rho/\rho) \sim 50$, the median value of $(M/L)_{B_T}$ is $\sim 15\%$ smaller than the values found by NW.

3.3.2 Comparison with CDM and other tests

In this subsection we compare the results obtained so far with the predictions of the CDM cosmogony. We also use our N-body simulations to carry out further checks of our techniques. Group luminosity functions for CDM groups are shown by symbols in Figure 3.6. These groups were selected and analysed in exactly the same way as groups in the CfA survey. The different symbols in Figure 3.6 correspond to different values of the biasing parameter assumed in the CDM simulations; circles correspond to $b = 2.5$, squares to $b = 2.0$ and triangles to $b = 1.6$. Averages from all the simulations in each ensemble are shown. There is very little difference between results for the different values of b. All the CDM

group catalogues produce luminosity functions which are in reasonable agreement with that of the CfA groups. This agreement extends to the individual luminosity functions of the various types of systems into which we have decomposed the group luminosity function in Figure 3.6.

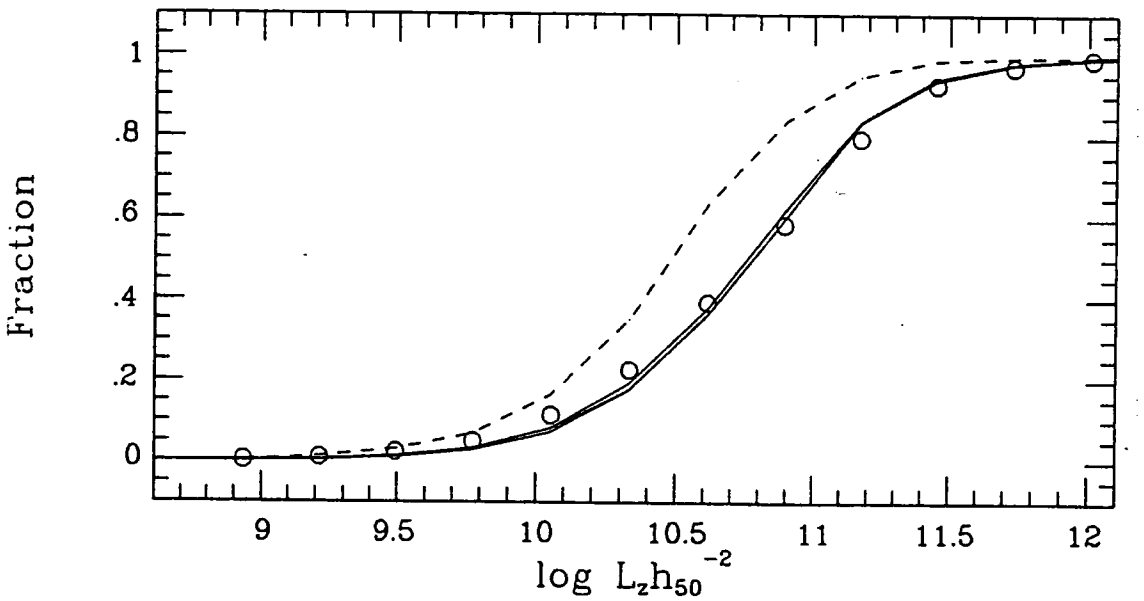


Figure 3.8 The cumulative fraction of groups less than a certain luminosity for the CfA, open circles, and CDM, solid lines. The dashed line shows the same distribution before the luminosity correction equation (3.4) is applied to the CfA groups.

We can compare our theoretical predictions with the data in a slightly different way by plotting cumulative, rather than differential distributions. In Figure 3.8 we plot the fraction of groups brighter than a given absolute magnitude for both the CfA and CDM group catalogues. The open circles represent the real

data and the solid lines the CDM predictions for our three values of b. This plot confirms the good agreement between the models and the data seen in Figure 3.6. Also shown as a dashed line are the CfA groups without the luminosity correction of equation (3.7). At the median point of the cumulative distribution, the correction to the total group luminosity is about one magnitude.

In addition to providing a test of the CDM theory, our simulated catalogues can be used to check the accuracy of our procedure for estimating total group luminosities. In particular we can verify that, as advertised, our results do not depend on the magnitude limit of the catalogue or on the adopted velocity cutoff. To test the first point we generated six artificial galaxy catalogues limited at $m_z = 15.5$ from our $b = 2.0$ ensemble of simulations. We compare results for these catalogues with those from our standard ones limited at $m_z = 14.55$ and with subcatalogues extracted from them at a magnitude limit of $m_z = 13.5$. In all cases, we consider only the region with $V < 4000$ km s^{-1}which contains the shallow catalogue. To identify groups at a similar density contrast (~ 50) in all cases, we varied the angular linking parameter in the deep and shallow catalogues by a small fraction. The group luminosity functions for the three limiting magnitudes are plotted in Figure 3.9*a*. There are no systematic differences amongst the different data sets indicating that our method for estimating L_{tot} is indeed independent of the magnitude limit. As a related check we compare in Figure 3.9*b* our standard CfA group luminosity function with that for groups of similar density contrast identified in the subset of the CfA survey limited at $m_z = 13.5$. Again, both catalogues are limited at 4000 km s^{-1}. Although the statistics are much poorer than in our ensemble of simulated catalogues, the result is the same: the group luminosity function is independent of the magnitude limit of the catalogue used to identify the groups.

As a separate test we compare group luminosity functions for catalogues limited at $m_z = 14.5$, but with different velocity cutoffs. We use both our ensemble

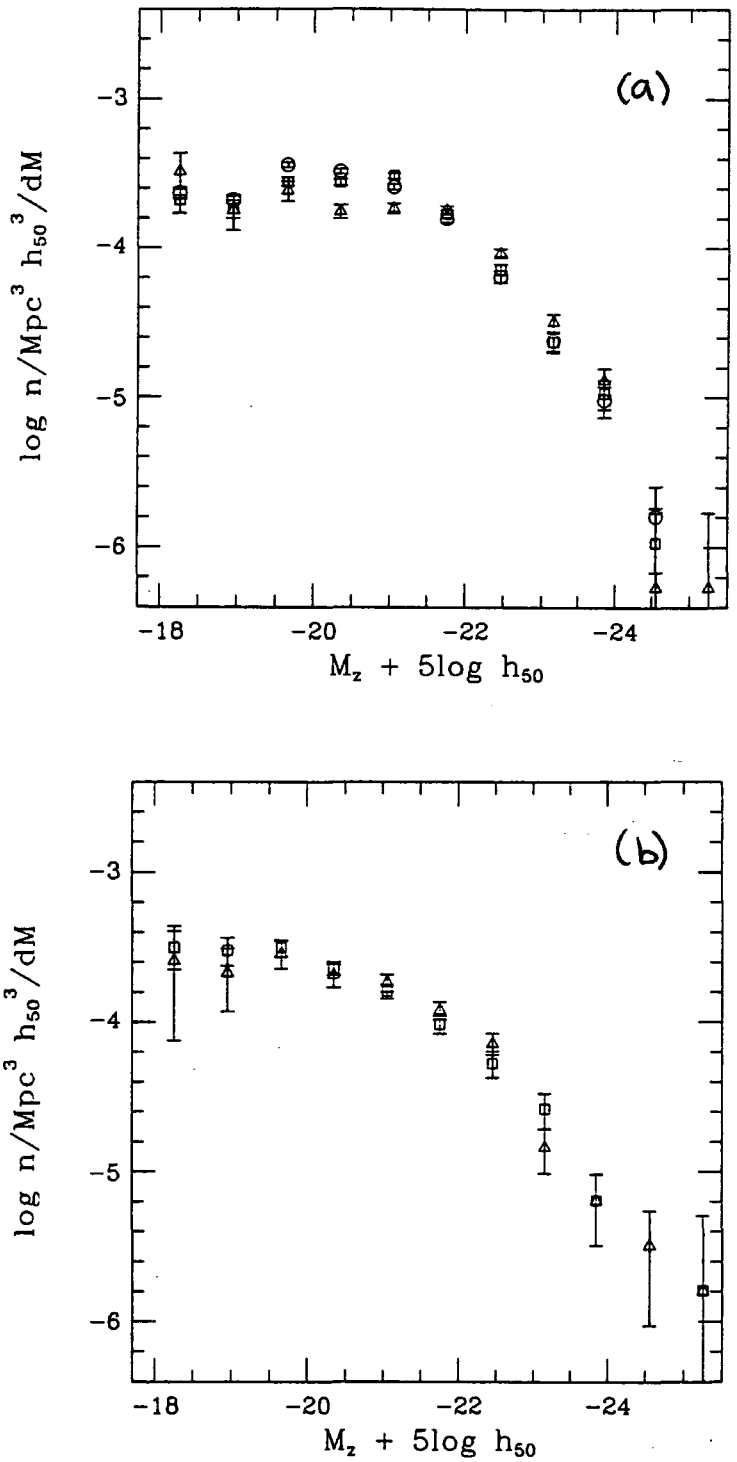


Figure 3.9 *(a)* The circles, squares and triangles show the group luminosity functions for the $m_z = 15.5, 14.55$ and 13.5 magnitude limited samples constructed from simulated galaxy catalogues with an apparent magnitude limit $m_z = 15.5$. it (b) The CfA group luminosity functions for $m_z = 14.5$ and $m_z = 13.5$ limited catalogues are plotted as open squares and circles.

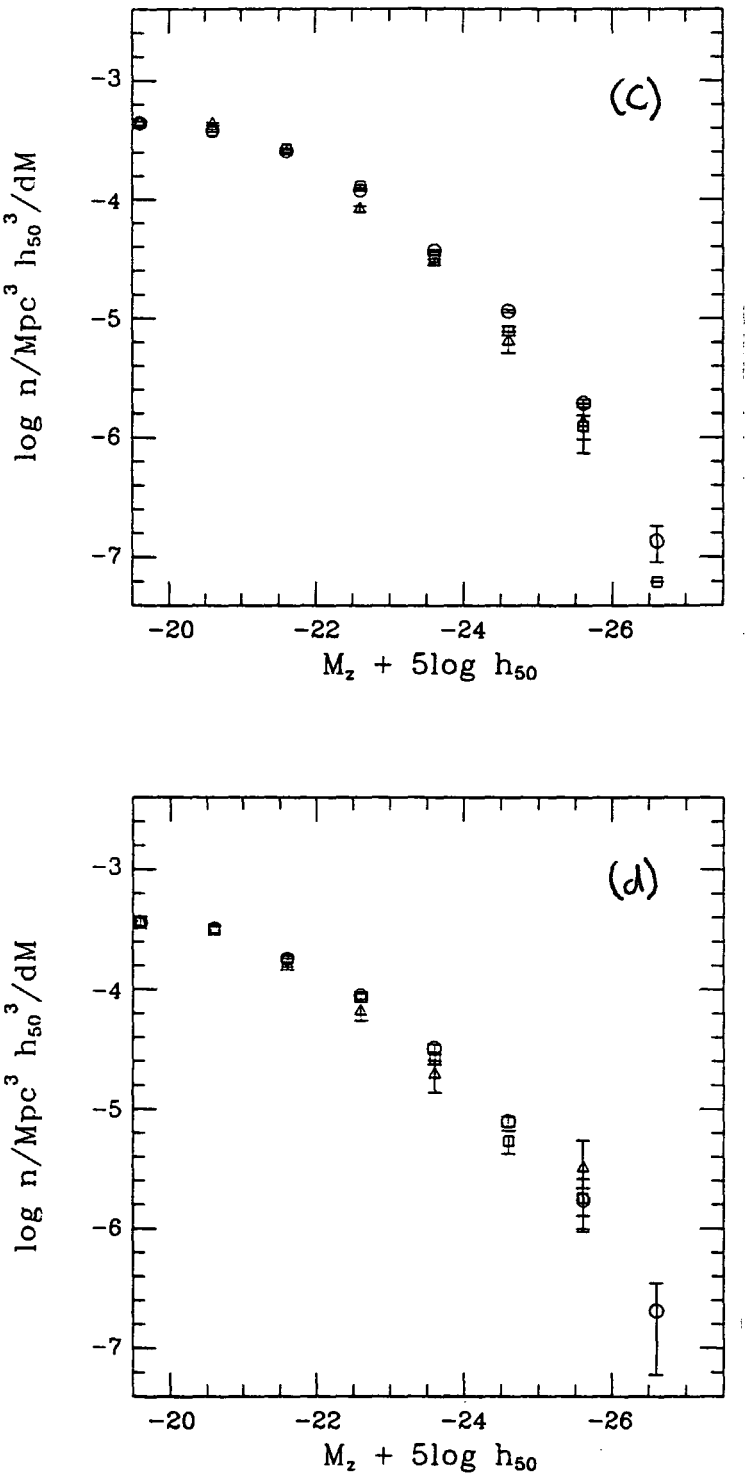


Figure 3.9 *(c)* The circles, triangles and squares show the group luminosity functions for the CDM groups and *(d)* the CfA groups within the volume limits of 10000, 7000 and 4000kms^{-1} respectively.

of $b = 2$ simulations and the CfA survey. We find groups in catalogues limited at 10000 km s^{-1}, 7000 km s^{-1}(our standard value) and 4000 km s^{-1}. The resulting group luminosity functions are plotted in Figures 3.9*c* and 3.9*d*. At the faint end, the groups are identified using the same galaxies in all cases, so this test is only sensitive for $M_z \gtrsim -20$. All the catalogues give consistent results, indicating that our procedure is reliable even in the region where the luminosity correction is large, well beyond our standard cutoff velocity of 7000 km s^{-1}.

Finally, we tested the sensitivity of the luminosity function to the accuracy of the group finding algorithm by obtaining the luminosity function of groups found in *real* space in our CDM catalogues at a similar density contrast to that used in redshift space. This luminosity function turned out to be practically indistinguishable from that found of our standard catalogues plotted in Figures 3.6 and 3.7.

3.4 OTHER GROUP PROPERTIES

3.4.1 The distribution of velocity dispersions

The cumulative distribution of line-of-sight velocity dispersion, $< V^2_{l.o.s} >^{1/2}$, for CfA groups is plotted in Figure 3.10*a*. To obtain a true spatial density, each group was weighted by $1/\Gamma_{max}$, with Γ_{max} given by equation (3.10) and in this case is the maximum volume out to which the third brightest group member could be observed above L_{lim}. The abundance is normalized to the mean density of Abell clusters of richness class $R > 0$, which, according to Bahcall and Soneira (1983), is $n_a = 7.5 \times 10^{-7}$ Mpc^{-3}. Also shown in the figure is the corresponding distribution for a sample of Abell clusters which Frenk *et al.* (1989) claim is fairly representative of the $R > 0$ cluster population. This sample contains 90 rich cluster velocity dispersions which we found from a literature search. Each cluster was required to have its velocity dispersion measured with at least 10 galaxies. The distributions of groups and clusters do not seem to match at all.

The group distribution falls steeply with velocity dispersion, whereas the rich cluster distribution is much flatter. At $< V^2_{l.o.s} >^{1/2} = 500$ km s^{-1}, the CfA groups are $\sim 2-3$ times more abundant than $R > 0$ Abell clusters. The disagreement at low velocity dispersions can be understood in terms of the definition of Abell clusters as rich dense systems. Many of the CfA groups with $< V^2_{l.o.s} >^{1/2} = 500$ km s^{-1} are not dense enough to satisfy Abell's criterion.

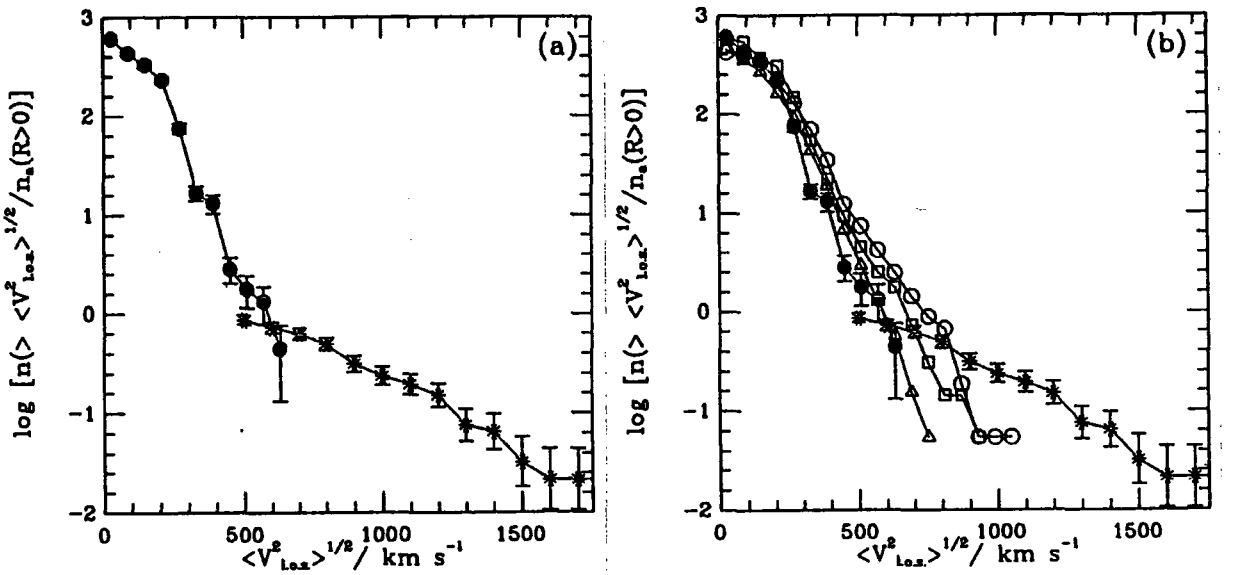


Figure 3.10 *(a)* The cumulative distribution of line-of-sight velocity dispersions of galaxy groups. The filled circles show the distribution for the CfA groups. The stars represent the fractional cumulative distribution of a sample of Abell clusters with good measurements of $< V^2_{l.o.s} >^{1/2}$. *(b)* The open triangles, squares and circles show the same distributions for groups found in the CDM simulations with biasing, $b = 2.5, 2.0, 1.6$ respectively. The group and cluster distributions are also shown for comparison. The abundances are normalised to 7.5×10^{-7} Mpc^{-3}, the observed mean density of Abell clusters of richness class $R > 0$.

The sharp discontinuity between the distributions for groups and clusters, on the other hand, is more difficult to interpret. Taken at face value it would seem to suggest that poor groups and rich clusters are dynamically different systems. We do not believe that this interpretation is correct. We suggest instead, that the flatness of the velocity dispersion distribution for Abell clusters is an artefact caused by spuriously large measurements, as would be obtained if projection effects were important. This suggestion is based on the results of Frenk *et al.* (1990). They showed that contamination by foreground galaxies can produce a spurious tail of apparently high dispersion clusters, similar to that in Figure 3.10, even if the true distribution contains no clusters with $< V^2_{l.o.s} >^{1/2} \gtrsim 1000$ km s^{-1}. The groups we have identified in the CfA survey do not suffer severely from projection effects, so their velocity dispersions reflect the true depth of the potential wells.

The predictions of our CDM models are compared with the CfA data in Figure 3.10*b*. At small values of $< V^2_{l.o.s} >^{1/2}$, the predicted distributions depend only weakly on the biasing parameter b, and are all slightly higher than the CfA data. However, the simulations do not take into account the internal degrees of freedom within individual galaxies. As a result group velocity dispersions comparable with typical internal velocity dispersions of individual galaxies are likely to be slightly overestimated in the simulations. This effect becomes negligible for $< V^2_{l.o.s} >^{1/2} \gtrsim 400$ km s^{-1}. In this regime, the predicted abundance is quite sensitive to the value of b. The model predictions with $b \sim 2 - 2.5$ are in good agreement with the CfA data although, given the uncertainties in the data, values of b around 1.5 cannot be strongly ruled out. The theoretical distributions shown in Figure 3.10*b* merge smoothly into the distributions predicted for rich clusters identified in three-dimensions in the simulations of Frenk *et al.* (1990) (*c.f.* their Figure 3). The agreement displayed in Figure 3.10*b* complements similar results for rich clusters obtained by Frenk *et al.* (1990). It is worth noting that group and cluster properties test the theory on different scales and are subject to different observational uncertainties.

In Figure 3.11 we plot luminosity against velocity dispersion for our CfA (circles) and $b = 2$ CDM (crosses) groups. Since these systems are defined by an overdensity criterion, we expect $< V^2_{l.o.s} >^{1/2} \propto L^{1/3}$ if the groups are in dynamical equilibrium and have similar mass-to-light ratios. This expectation is approximately born out by both sets of data, providing further evidence that our grouping algorithm identifies real dynamical systems which are not unduly contaminated by unrelated galaxies. For comparison, we also plot the Tully-Fisher relation for spiral galaxies (dashed line; from Aaronson and Mould 1983) and the Faber-Jackson relation for elliptical galaxies (dotted line; from Davis *et al.* 1983).

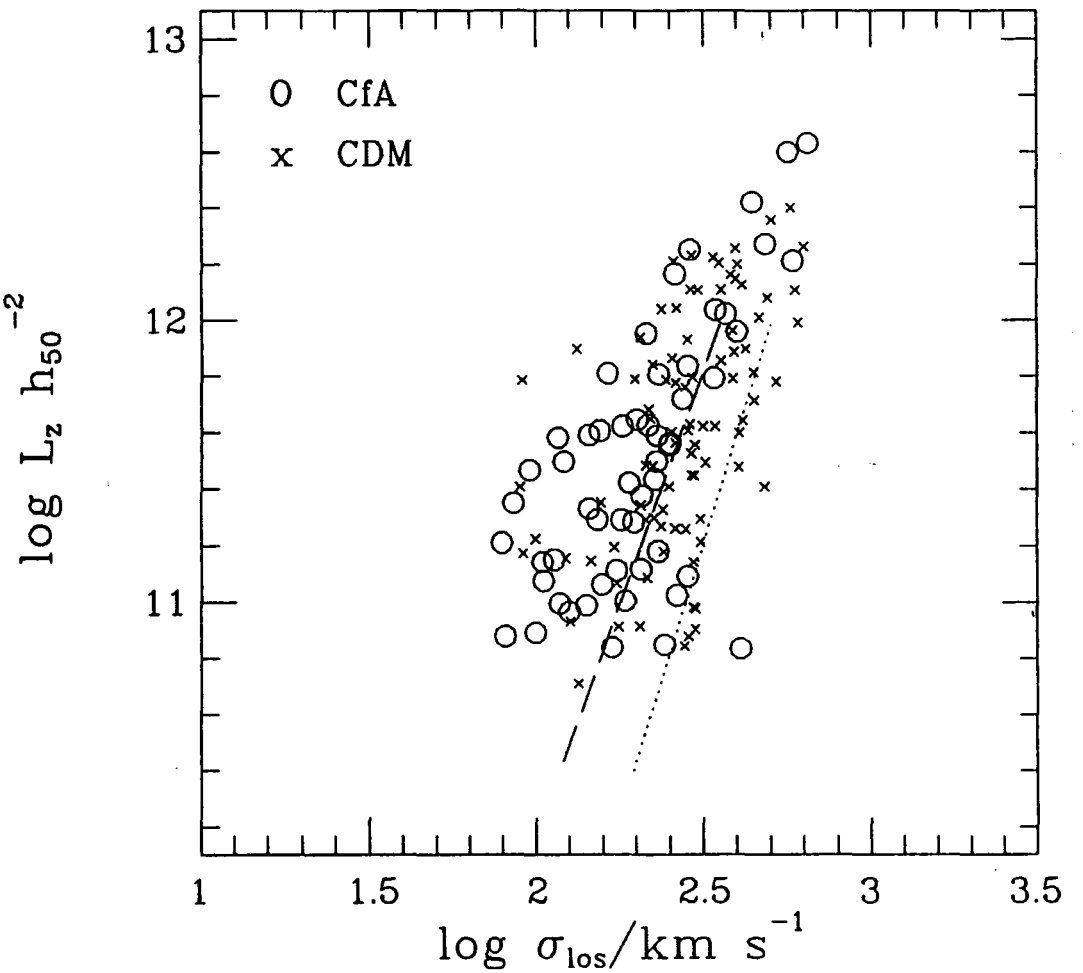


Figure 3.11 The luminosities of groups with ≥ 5 members are plotted against their velocity dispersions. The open circles are the CfA groups and the crosses represent groups from one of our $b = 2$ CDM catalogues. For comparison, we also plot the Tully-Fisher relation for spiral galaxies and the Faber-Jackson relation for elliptical galaxies.

3.4.2 The two-point correlation function of groups

The correlation statistics of rich clusters have given rise to much interesting debate in recent years. The strong correlations originally measured by Bahcall and Soneira (1983) and Klypin and Kopylov (1983) have been questioned by Sutherland (1988) and by Dekel *et al.* (1989). (See also Ling *et al.* 1986.) These authors argue that contamination and other systematic effects could introduce spurious correlations in Abell's cluster catalogue. Shectman (1986) analysed a different sample of relatively poor clusters and found them to have an intermediate clustering strength between that of galaxies and that of rich clusters. This sample, however, could be affected by the same kind of systematic effects discussed in connection with Abell clusters.

Our CfA group catalogue is, in principle, free from systematic effects due to projection or contamination since it is defined using quasi-three dimensional data. Its correlation statistics, however, cannot be directly compared with those of rich clusters since it seems likely that the amplitude of the two-point correlation function depends on cluster richness. This behaviour is predicted by Kaiser's (1984) model for the correlations of high peaks in Gaussian fields. Nevertheless, the group-group correlation function is a quantity of considerable interest. It probes the clustering pattern on intermediate scales, it can be reliably determined from data like ours and it can be compared directly with theoretical models such as the CDM cosmogony. Furthermore, this statistic is fairly insensitive to the details of the grouping procedure, at least on scales substantially larger than the mean group separation. It depends exclusively on the positions of the groups rather than on their exact membership.

To estimate the group-group correlation function, ξ_{gg}, we consider the brightest 100 "groups" in each of the Northern and Southern catalogues. Since the luminosities include the correction of Section 3.1 for members below the CfA survey

limit, a fraction of these groups correspond to single survey galaxies. Thus, in the Northern region our sample contains 46, 26 and 28 groups with $\geq$ 3, 2, and 1 CfA survey galaxies respectively. The Southern sample has a similar composition. The number and the number density of groups as a function of recession velocity are shown in Figures 3.12*a* and 3.12*b* respectively, with the North (solid lines) and the South (dashed lines) regions displayed separately. The Local Supercluster in the North and the large supercluster in the South at 4800 km s^{-1} stand out clearly in Figure 3.12*b*. This figure shows that the mean density of groups in our bright sample is approximately constant with distance, indicating that this sample is approximately volume limited.

To estimate ξ_{gg}, we take the median position of member galaxies as the group centre and use the estimator:

$$\xi_{gg} = \frac{N_o}{N_R} - 1, \tag{3.12}$$

where N_o is the observed number of group pairs in the catalogue and N_R is the number of pairs in a catalogue of points distributed at random within the same boundaries and with the same mean redshift distribution. N_R was estimated by averaging over 1000 random catalogues of the same size as the original group sample.

The resulting group-group correlation functions for the CfA catalogues is represented by filled circles with 1σ Poisson error bars in Figure 3.13. A least-squares fit to these data gives,

$$\xi_{gg}(r) = \left(\frac{r}{16 \pm 3 \ \mathrm{h_{50}^{-1}Mpc}} \right)^{-1.6 \pm 0.4.} \tag{3.13}$$

The correlation length, $r_o = 16 \pm 3 \ \mathrm{h_{50}^{-1}}$ Mpc, is only 36% of value estimated by Ling *et al.* (1986) for Abell clusters of distance class $D < 4$, but $\sim$ 60% of the value inferred by Sutherland (1988) for this same sample after correction for various

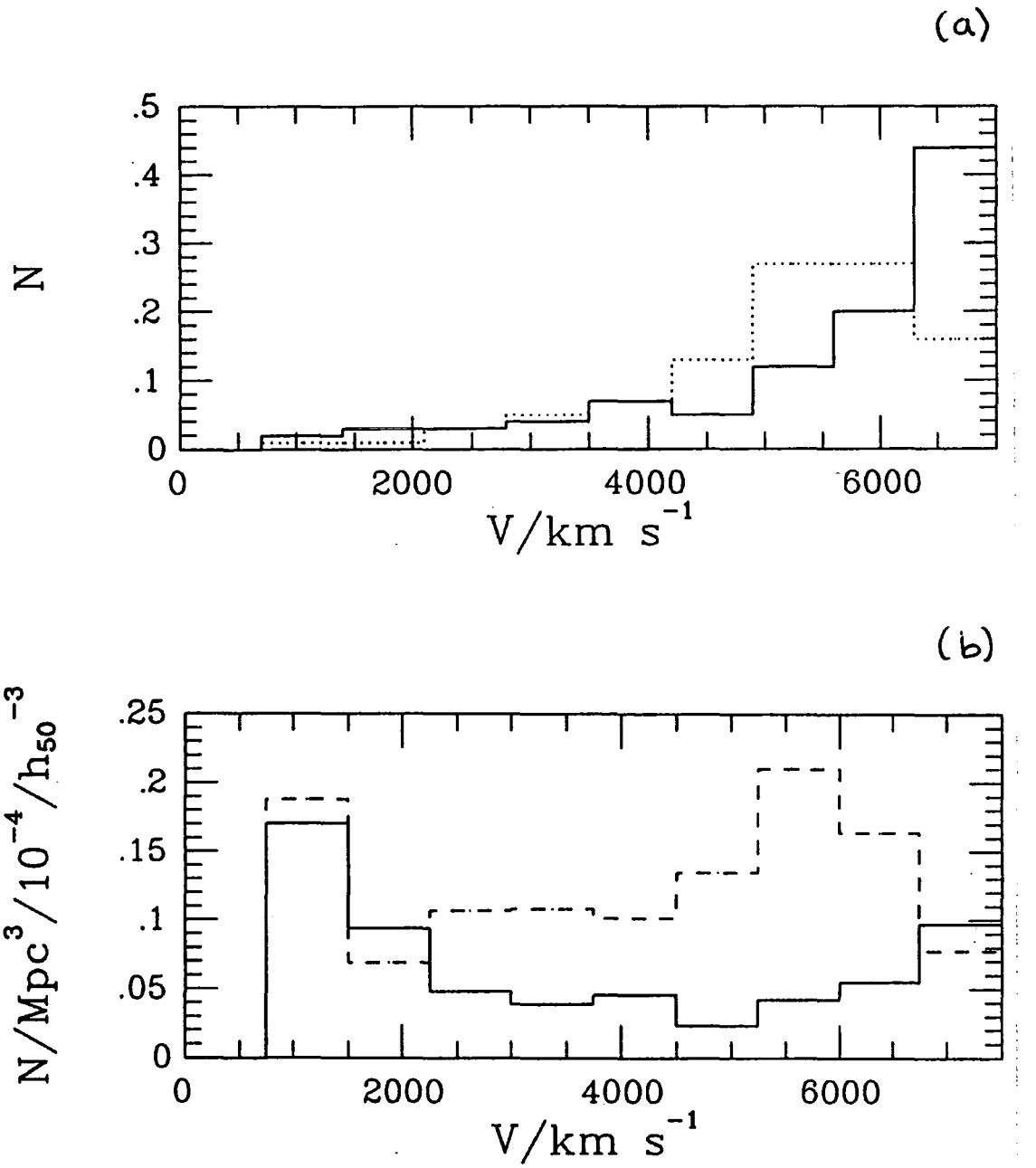


Figure 3.12 *(a)* The number and *(b)* the number density of groups in the Northern (thick line) and the Southern (dashed line) CfA survey as a function of recession velocity.

systematic effects. Increasing our sample to contain the brightest 400 "groups" leads to a correlation function with similar slope but with a clustering length $\sim 15\%$ smaller than equation (3.10) and similar to the galaxy two point correlation function. We have also calculated ξ_{gg} for the 166 CfA groups with ≥ 3 members picked out by our grouping algorithm and again find a similar clustering length as the galaxy correlation function. This sample, however, is not homogeneous because it is biased towards richer groups at large distance and very poor groups nearby.

The CDM predictions for ξ_{gg} for the brightest 100 groups are also shown in Figure 3.13, for $b = 2.5$, 2, and 1.6. There appears to be no trend for different biasing parameters. The most luminous 800 CDM groups has a clustering length less than that of the galaxies whilst, like the CfA groups, the most luminous 400 groups show similar clustering as the galaxies. In hierarchical clustering theories, ξ_{gg} is determined by the shape of the fluctuation spectrum; it is essentially independent of its amplitude and hence of b. This expectation is verified in our data. The model predictions are in good agreement with the CfA data. The group correlation length is a factor of 1.4 smaller than the correlation length obtained by White *et al.* (1987) for rich clusters identified in three dimensions in CDM N-body simulations. (The correlation functions for groups found in real space in our catalogues are virtually indistinguishable from those plotted in Figure 3.12.) This seems to confirm the trend, predicted in Kaiser's (1984) model, of an increase in clustering amplitude with cluster richness. It is interesting that the CDM spectrum seems to be consistent with the correlation strength of groups, but appears to be inconsistent with uncorrected estimates of Abell cluster correlations.

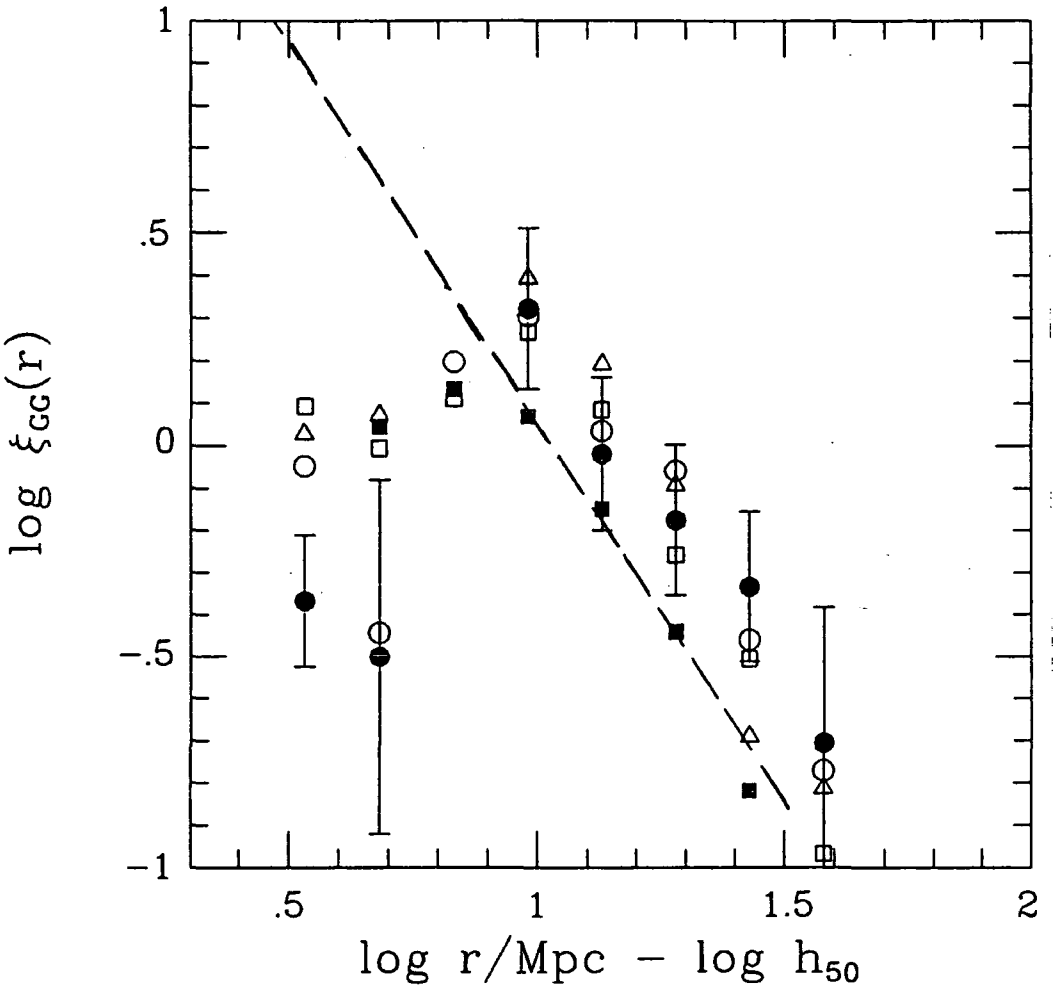


Figure 3.13 The group-group two point correlation function. The filled circles show the average correlation function of the brightest 100 groups from the Northern and Southern CfA survey. Error bars are taken from the scatter between 12 CDM realisations of the Northern CfA survey. The filled squares show the same correlation function but for the brightest 400 groups. Open circles, squares and triangles represent the biased CDM data 2.5, 2.0 and 1.6 respectively. The dashed line represents the galaxy two point correlation with amplitude 10 h_{50} Mpc and slope -1.8.

3.5 DISCUSSION AND CONCLUSIONS

We have identified groups of galaxies in the Center for Astrophysics (CfA) redshift survey using an algorithm similar to that proposed by Nolthenius and White (1987). We defined groups as galaxy aggregates approximately bounded by isodensity surfaces, with a mean density contrast of ~ 50. It is likely that they are close to virial equilibrium and have velocity dispersions in the the range $\sim 50 - 600$ km s^{-1}. These groups correspond to structures intermediate between bright galaxies and rich galaxy clusters. We have examined the continuity of properties over this range of scales and we have carried out a detailed comparison of the structure, abundance and clustering of groups with those predicted in the cold dark matter cosmogony. For this comparison we used artificial "galaxy" catalogues constructed from large N-body simulations and with the same geometry and magnitude limit as the CfA survey.

One of the aims of this chapter was to define and estimate the luminosity function of "galactic systems", that is the distribution of *total* luminosity in virialized systems on a wide range of scales. This is a useful diagnostic for theories which make firm predictions for the gravitational evolution of bound structures, but not for the dissipative physics which regulates the detailed distribution of luminous material within virialized clumps. Within specific cosmogonic models, predictions of this kind may be obtained through analytic approximations such as the Press-Schechter (1974) formalism or through N-body simulations such as those analysed here. A related property which we also consider here is the distribution of group velocity dispersions.

We defined the *total* luminosity of a group as the integral of a universal galaxy luminosity function normalised by the observed number of galaxies in the group. In this approach, even a single galaxy is regarded as the sole member of a group lying above the magnitude limit of the catalogue. With our definition and

estimation method we are able to predict correctly the observed total luminosity of Virgo and Coma from the relatively few galaxies above the apparent magnitude limit of the CfA survey. The luminosity function of galactic systems smoothly joins single bright galaxies at its faint end with luminous rich clusters at its bright end. It is well approximated by a double power law function with best-fit parameters $\beta = 0.16 \pm 0.07$, $\gamma = 1.7 \pm 0.12$, $M_z^{*'} = -22.1 \pm 0.2$ and mean density $\phi^{*'} = 3.0 \pm 0.5 \times 10^{-4}$ Mpc^{-3}. Using our artificial catalogues, we showed the luminosity function to be independent of the magnitude limit of the galaxy catalogue from which the groups are identified.

Identical procedures to those applied to the CfA data were implemented in artificial catalogues constructed from CDM simulations assuming three different values of the biasing parameter, $b = 2.5$, 2, and 1.6. The predicted group luminosity functions are essentially independent of b and are in remarkably good agreement with the CfA data. The distribution of group velocity dispersions, on the other hand, is sensitive to the normalization of the CDM fluctuation spectrum, measured by b. We find that models with $b \sim 2 - 2.5$ agree with the CfA data, although somewhat smaller values of b cannot be strongly ruled out by the present data. Previous studies had shown that the CDM cosmogony successfully reproduces the observed abundance of galaxy clusters (Frenk *et al.* 1985, 1988, 1989; White *et al.* 1987.) The present analysis extends this conclusion to intermediate structures.

The distribution of velocity dispersion for the CfA groups does not join smoothly into that of rich Abell clusters. The abundance of the richest groups we find is comparable to that of the poorest Abell clusters. However, the distribution of group dispersions falls steeply whereas the distribution for rich clusters exhibits an extended tail of high dispersion objects. We argued that this apparent discontinuity is most likely the result of spuriously large dispersion measurements for rich clusters, caused by contamination and projection effects.

The two-point correlation function of groups in the CfA survey is a a power-law of slope -1.6 ± 0.4. The correlation length, $r_o = 16 \pm 3 h_{50}^{-1}$ Mpc, is slightly larger than the correlation length of CfA galaxies but it is a factor of ~ 3 smaller than the value originally reported by Bahcall and Soneira (1983) for rich Abell clusters. The group correlation function we have estimated is not subject to the type of projection effects discussed by Sutherland (1988) and Dekel *et al.* (1989) in connection with Abell clusters. The correlation function of the brightest 400 groups is comparable to the galaxy correlation function. Similarly, the correlation function of the 166 groups with 3 or more members shows the same clustering strength as the galaxies, however, this sample is a "mixture" of poor groups nearby and rich groups at large distances.

As expected, the group-group correlation function in our models is independent of the biasing parameter, since this statistic reflects the *shape* of the fluctuation spectrum. The model results agree with the CfA data. This contrasts with the situation for rich clusters. The predicted correlations for clusters identified in three dimensions in CDM simulations are about a factor of 2 weaker than the correlations reported for Abell clusters (White *et al.* 1987b). Our study lends indirect support to the views of Sutherland (1988) and Dekel *et al.* (1989) who have argued that the estimates for Abell clusters are biased high by various systematic effects related to contamination and projection effects. It would seem surprising that the CDM spectrum can match the group data so well but fail by such a large factor on slightly larger scales.

One advantage of studying galaxy groups identified from redshift surveys is that properties such as velocity dispersion and clustering are not affected by the kind of projection effects which mar catalogues of clusters found in projection. The disadvantage is that the exact group membership and hence the measured velocity dispersions, depend sensitively on the details of the grouping algorithm. Here we have shown how N-body simulations can be used to optimize the grouping

procedure and circumvent this limitation. More extensive redshift surveys will enable these techniques to be extended to the study of rich clusters.

REFERENCES

Abell, G. 1958, *Ap.J. Suppl.*, **3**, 211.

Bahcall, N.A. 1979, *Ap.J.*, **232**, 689.

Bahcall, N. and Soneira, R. 1983, *Ap.J.*, **270**, 20.

Bhavsar, S.P. 1978, *Ap.J.*, **222.**, 412.

Bhavsar, S.P., Gott,J.R. and Aarseth, S.J. 1981, *A.J.*, **246**, 656.

Blumenthal, G.R., Faber, S.M., Primack, J.R. and Rees, M.J. 1984, *Nature*, **311**, 517.

Davis, M. and Huchra, J. 1982, *Ap.J.*, **254**, 437.

Davis, M. and Peebles, P.J.E. 1983, *Ap.J.*, **267**, 465.

Davis, M., Efstathiou,G., Frenk, C.S. and White, S.D.M. 1985, *Ap.J.*, **292**, 371 (DEFW).

Dekel, A., Blumenthal, G.R., Primack, J.R. and Olivier, S. 1989, *Ap.J letters*, **338**, L5.

Efstathiou, G., Davis, M. Frenk, C.S. and White, S.D.M. 1985, *Ap.J. Suppl.*, **57**, 241.

Efstathiou, G., Ellis, R.S. and Peterson, B. 1988a, *M.N.R.A.S.*, **232**, 431.

Efstathiou, G., Frenk, C.S., White, S.D.M. and Davis, M. 1988, *M.N.R.A.S.*, **235**, 715.

Felten, J.E. 1985, *Comm. Ap. Sp. Sci.*, **11**, 53.

Fisher, J.R. and Tully, R.B. 1981, *Ap.J.*, **47**, 139.

Frenk, C.S., White, S.D.M., Efstathiou, G. and Davis, M. 1985, *Nature*, **317**, 595.

Frenk, C.S., White, S.D.M., Efstathiou, G. and Davis, M. 1988, *Ap.J.*, **327**, 507.

Frenk, C.S. 1989, in **The Epoch of Galaxy Formation**, eds. C.S. Frenk, R.S. Ellis, T. Shanks, A. Heavens and J. Peacock, (Dordrecht: Kluwer Acad. Publ.).

Frenk, C.S., White, S.D.M., Efstathiou, G. and Davis, M. 1990, *Ap.J*, **391**, 2.

Geller, M.J. and Huchra, J.P. 1983, *Ap.J.Suppl.*, **52**, 61.

Geller, M.J. and Huchra, J.P. 1989, *Science*, **246**, 897.

Godwin, J.G., Metcalfe, N. and Peach, J.V. 1983, *M.N.R.A.S.*, **202**,113.

Gott, J.R. and Turner, E.L. 1977, *Ap.J.*, **216**, 357.

Huchra, J.P. and Geller, M.J. 1982, *Ap.J.*, **257**,423.

Kaiser,N. 1984, *Ap.J.*, **284**, L9.

Kaiser, N. and Lahav, O. 1989, *M.N.R.A.S.*, in press.

Klypin, A.A. and Kopylov, A.A. 1983, *Soviet Astr.letters*, **9**, 41.

Ling, E.N., Frenk, C.S. and Barrow, J. 1986, *M.N.R.A.S.*, **223**, 21.

Nolthenius, R. and White, S.D.M. 1987,*M.N.R.A.S*, **225**, 505.

Peebles, P.J.E., Daly, R.A. and Juszkiewicz, R. 1989, *Princeton preprint.*

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

Ramella, M., Geller, J.G. and Huchra, J.P. 1989, *Ap.J*, **344**, 57.

Sandage, A., Binggeli, B. and Tammann, G.A. 1985, *A.J.*, **90**, 1759.

Shectman, S.A. 1986, *Ap.J.Suppl.*, **57**, 77.

Sutherland, W. 1988, *M.N.R.A.S.*, **234**, 159.

White, S.D.M., Frenk, C.S., Davis, M. and Efstathiou, G. 1987a, *Nature*, **330**, 451.

White, S.D.M., Frenk, C.S., Davis, M. and Efstathiou, G. 1987b, *Ap.J.*, **313**, 505 (WFDE).

White, S.D.M. 1989, *To appear in Hubble Symposium proceedings, Berkeley.*

4 THE TOPOLOGY OF LARGE SCALE STRUCTURE

4.1 INTRODUCTION

Quantifying the large scale structure of the Universe is a difficult task and a variety of techniques have been developed to approach this problem. The first large data sets came with the 2-dimensional catalogue of Shapley-Ames (1932) and the Shane-Wirtanen (1967) galaxy counts. Large structures and clustering became apparent and statistics such as the angular correlation function were developed to quantify the deviations from a random distribution of galaxies. The advent of complete redshift surveys (Davis *et al.* 1982) brought to light the complex network of voids and filamentary structures that are present in the Universe. As each redshift survey is completed, structures approaching the size of the survey volume are suggested. This may be partly due to selection effects but only deeper surveys with well defined selection criteria will resolve this problem. The clustering of galaxies in 3-dimensions has proved to be one of the most important constraints on models of galaxy formation. A successful model must reproduce the observed clustering pattern of galaxies such as the two point correlation function (Peebles 1980), match the observed large scale velocity flows (Burstein *et al.* 1990), contain voids as large as Bootes (Kirshner *et al.* 1981) and have features as extended as the "Great Wall" (Geller and Huchra 1989).

In the standard inflationary models for galaxy formation the structures that we see today grew from small amplitude, random phase quantum noise in the

early Universe (Bardeen *et al.* 1983). The statistical properties of a random phase distribution are completely specified by its power spectrum. In the linear regime overdensities grow as $\delta\rho/\rho \propto a$, where a is the expansion factor. Therefore in co-moving coordinates, positive and negative fluctuations grow in place, increasing only in amplitude and contours of the density field selected by volume fraction do not change with time. When $\delta\rho/\rho \approx 1$ the high density regions begin to collapse rapidly and the initial power spectrum which was frozen into the density field becomes distorted. However we can still observe the initial conditions of the density field if we smooth the data on a scale much larger than the correlation length, as verified by Weinberg *et al.* (1987).

In 1850 Gauss discovered a definition of surface curvature which depended only on the intrinsic geometry of the surface, not on the way in which the surface was embedded in Euclidean three-space. Gauss became so excited by this result that he named it the *theorema egregrium* "the extraordinary theorem". In a letter to the Astronomer Hansen he wrote: "These investigations deeply affect many other things; I would go so far as to say they are involved in the metaphysics of the geometry of space." Until recently studies of the topology of the galaxy distribution were limited to qualitative remarks about the high and low density regions (Gott *et al.* 1986; hereafter GMD), although it was apparent from N-body simulations that the visual impression of the density field depends strongly upon the initial mass fluctuations. Hamilton *et al.* (1986) re-derived an expression relating the slope of the initial power spectrum to the curvature of the density field and developed an algorithm to calculate it. Gott *et al.* (1989) used a series of redshift surveys to sample the topology on a range of length scales between $3h^{-1}$ Mpc and $50h^{-1}$ Mpc. (we denote the Hubble constant by $H_o = 100h$km s^{-1} Mpc^{-1}) On scales less than $10h^{-1}$ Mpc Gott *et al.* used the the CfA survey (Huchra and Davis 1982), Tully's (1987) all sky catalogue of nearby galaxies and the Thuan and Schneider (1988) diameter limited survey of dwarf galaxies. On scales larger than $10h^{-1}$ Mpc , the Giovanelli and Haynes (1985; hereafter GH) and

Abell cluster catalogue (Abell 1958) were used. All these studies taken together provide a strong test for any theory of galaxy formation. Gott *et al.* compared the data with heavy neutrino models, the cold dark matter model (CDM) and bubble models and concluded that CDM provided the best fit to the data over the scales considered.

Although the GH and Abell cluster samples gave the highest resolution of the five surveys, they must be regarded with caution for the following reasons. The GH sample is biased towards a "baked bean" topology because it purposefully contains the massive filamentary Perseus-Pisces supercluster which dominates the survey. The Abell catalogue should also be used with caution because it has been selected by eye and may be affected by contamination and projection effects (Lucey 1983, Frenk *et al.* 1990). It should also be noted that only a few percent of galaxies actually lie in Abell clusters.

The Queen Mary and Westfield College, Durham, Oxford and Toronto (QDOT) "one-in-six" redshift survey of galaxies detected by the IRAS satellite (Lawrence *et al.* 1991) samples the density field in the Universe to a redshift $z \approx 0.1$, well beyond the Local Supercluster. There are 2163 galaxies in the survey randomly sampled from the IRAS point source catalogue with 60μ flux greater than 0.6Jy covering almost the whole sky at galactic latitudes $|b| > 10^o$ (See Saunders *et al.* 1991 for a detailed description of the survey). Regions of sky contaminated by interstellar dust (the infra red "cirrus") and a patch of sky not covered by the IRAS satellite are coded into a mask which excludes about 17% of the whole sky. The random sampling strategy allows the maximum information on large scales to be obtained from a given amount of telescope time (Kaiser 1986). The QDOT redshift survey is well suited to investigate structure on large scales because it provides near full sky coverage with galaxies selected in a uniform way. The distribution of recession velocities, $n(v)$, shown in Figure 4.1 has a median of 9000km s^{-1} and a long tail which samples a fairly large range of scales. A separate

survey of all IRAS galaxies brighter than 2Jy is being carried out by Strauss and Davis (1988). The expected redshift distribution for the 2Jy survey is shown for comparison in Figure 4.1.

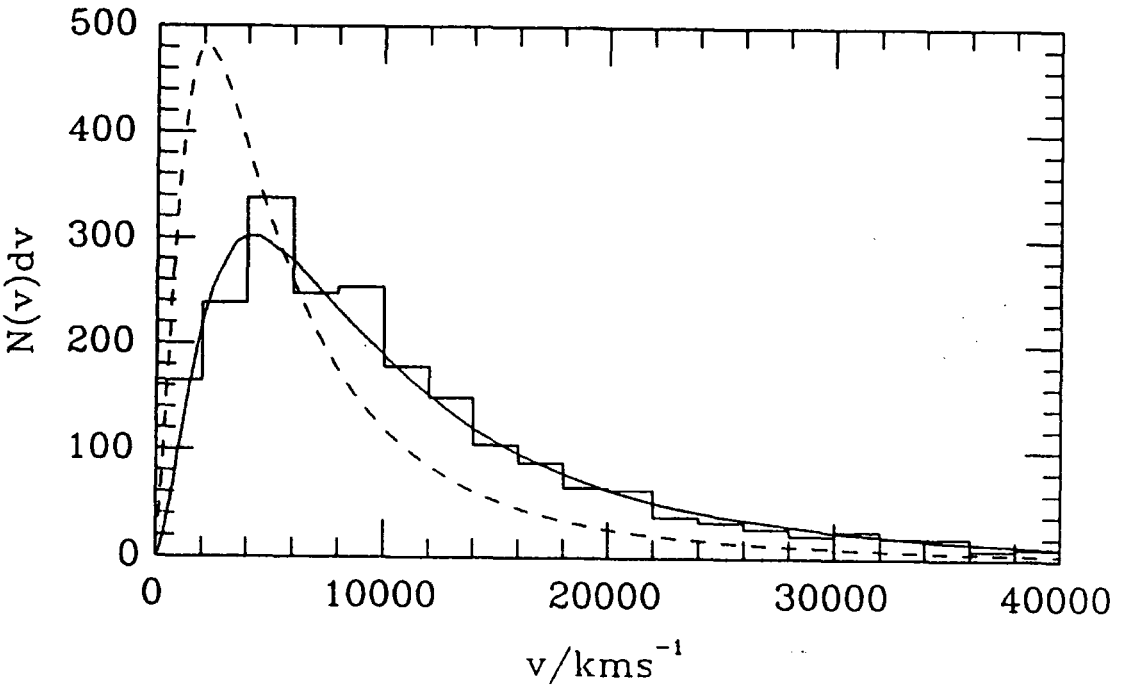


Figure 4.1 The redshift distribution of galaxies in the QDOT survey. The solid curve shows the distribution calculated by integrating the QDOT luminosity function (solution 19 in Table 3(a) of Saunders *et al.* 1990). The dashed curve shows the expected distribution of the 2Jy limited survey of IRAS galaxies by Strauss and Davis.

Several results have already emerged from analyses of the QDOT survey. The survey has been used to constrain the cosmological density parameter, Ω by mapping the local peculiar gravity field and comparing the velocity predicted for the Local Group (Rowan-Robinson *et al.* 1990) and for a sample of 1000

galaxies (Kaiser *et al.* 1991) with observations. Efstathiou *et al.* (1990) and Saunders *et al.* (1991) have estimated the variance of QDOT galaxy counts in cells and find fluctuations on large scales in excess of those predicted in the standard CDM model. Saunders *et al.* also obtained a marginal detection of skewness in the count distribution on large scales. The counts in cells analyses are sensitive to "hot spots", regions like the Hercules supercluster which may dominate the count variance. A topological analysis provides an alternative way of testing the Gaussian nature of the initial density field and measuring the slope of the power spectrum on similar length scales. This analysis is quite different to the counts in cells and it gives equal weight to all structures with density above some threshold value. The large depth coverage of the QDOT survey allows us to compare the results with similar analysis of the GH and Abell surveys on scales greater than $10h^{-1}$ Mpc.

As in the previous chapters in this thesis, we compare our results with theoretical predictions, particularly with those of the standard cold dark matter (SCDM) model. This is the best studied and in many ways the most successful cosmogony to date. Certainly on scales less than $\sim 10\ h^{-1}$ Mpc, the model predictions are roughly consistent with many observed properties of galaxies and their spatial distribution (see Frenk 1991 for a recent review). On larger scales the model predicts significantly less superclustering than has been measured in recent surveys, including the QDOT survey. It is perhaps worth reviewing the main tenets of the SCDM model and to be precise about which aspects of the theory are tested by our data. The SCDM model is based on four premises: *(i)* the dark matter consists of weakly interacting "cold" elementary particles; *(ii)* the Universe has critical density; *(iii)* the primordial fluctuations which seed the formation of structure are of the type predicted by the inflationary theory of the early Universe, *i.e.* Gaussian, adiabatic, and scale invariant; and *(iv)* the distribution of galaxies is biased relative to the distribution of mass according to the "linear

biasing model" whereby fluctuations in the galaxy distribution are assumed to be a constant multiple of the corresponding mass fluctuations.

The QDOT survey has already proved extremely valuable in testing directly some of the tenets of the SCDM model. The dynamical analyses of Rowan-Robinson *et al.* (1990) and Kaiser *et al.* (1991) provide support for assumption *(ii)* above. Since standard Big Bang nucleosynthesis limits the density of baryons to less than about 10% of the critical value (Olive *et al.* 1990), this result also provides indirect support for assumption *(i)*. The counts-in-cells analyses of Efstathiou *et al.* (1990) and Saunders *et al.* (1991) test a combination of model assumptions, particularly assumption *(iii)* concerning the shape of the primordial density fluctuation spectrum and assumption *(iv)* concerning the way in which these fluctuations are manifested in the galaxy distribution. The present analysis is especially sensitive to the shape of the fluctuation spectrum and to the assumption of random phases but is less sensitive than the counts-in-cells analyses to the assumed biasing prescription.

In this chapter we use a combination of analytic and numerical tools to compare the topology of the QDOT survey with predictions of the SCDM model. For the latter, we rely heavily on artificial "QDOT" surveys constructed from N-body simulations. Details of the simulations are given by Frenk *et al.* (1990) while the general procedure to construct flux limited "galaxy" catalogues from simulations is described by White *et al.* (1987). For the present analysis we matched this procedure to the particular circumstances of the QDOT survey. We assumed the luminosity function given by solution (19) of Saunders *et al.* (1990) (which includes a model for luminosity evolution and k-corrections). The artificial "QDOT catalogues" have the same luminosity function, redshift distribution and sky coverage of the real data and were masked in exactly the same way. They assume a biasing parameter $b = 2$, in the notation of Frenk *et al.* (1990).

In the next section we describe qualitatively the density field traced by IRAS galaxies, and make visual comparisons with artificial surveys constructed from the CDM model and power law fluctuation spectra. In Section 4.3 we lay down the groundwork for a topological analysis of galaxy redshift surveys, namely the genus-density relation. Section 4.4 describes the analysis of the QDOT data and a preliminary estimate of the errors on the genus curve using artificial "QDOT" surveys constructed from the N-body simulations. In Section 4.5 we calculate the slope of the power spectrum over a range of length scales. The errors in the slope are calculated using the results of Section 4.4 and Monte-Carlo simulations. In Section 4.6 we summarise our results on the shape of the power spectrum with the counts-in-cells analysis of the same data.

4.2 QUALITATIVE TOPOLOGY

Most of the galaxies within the QDOT survey are disk systems which avoid the cores of rich clusters but appear to trace the same large scale structure as elliptical galaxies (Babul and Postman 1990). The most luminous IRAS galaxies are predominantly interacting, starbursting disk systems (Joseph and Wright 1985, Leech *et al.* 1989). IRAS galaxies can be split roughly into two populations depending on T_{fir}, their far-infrared temperatures (Rowan-Robinson and Crawford 1989). Seyfert or star-bursting galaxies have $T_{fir} \geq 36K$ ("warm") and disk-dominated galaxies have $t_{fir} < 36K$ ("cool"). Saunders *et al.* (1991) find that the luminosity function of galaxies in the QDOT survey is dominated by cool galaxies below $\sim 5 \times 10^{10} h^{-2} L_{\odot}$ and by warm sources above this. Most of the galaxies in the survey are late type spiral galaxies similar to the Milky-Way.

Visual impressions of the three dimensional galaxy distribution of the QDOT survey have been given by Saunders *et al.* (1991) and Kaiser *et al.* (1991). These pictures portray the wealth of structure that can be identified in the catalogue.

Topological analysis uses contours of constant density and these provide an interesting visual representation of the structures that make up the high and low density field, in a way which complements the previous representations.

We obtain a smoothed density field by weighting the observed galaxy distribution with the inverse of the selection function and then smoothing with the Gaussian function

$$W(r,\lambda) = \frac{1}{\pi^{3/2}\lambda^3} e^{-r^2/\lambda^2}, \tag{4.1}$$

where r is distance and λ is the smoothing length. W is normalised so that its integral over all space is unity. The selection function, S(r), is defined as the expected number of galaxies in the survey at a distance r in an unclustered distribution,

$$S(r) = \int_{L_{lim}}^{\infty} \phi(L) dL, \tag{4.2}$$

where L_{lim} is the minimum observable luminosity at distance r and $\phi(L)$ is the galaxy luminosity function.

The solid line in Figure 4.2 shows the selection function obtained using the luminosity function given by Saunders *et al.* (1990). For comparison the dotted line in Figure 4.2 shows the selection function calculated directly from the data using

$$S\prime(r) = \frac{3}{\omega} \sum_{D_{max,i}>r} \frac{1}{D^3_{max,i}}, \tag{4.3}$$

where $D_{max,i}$ is the maximum distance out to which the ith galaxy can be detected and ω is the solid angle of the survey. Equation (4.3) overestimates the local density below $30h^{-1}$ Mpc but agrees fairly well with the estimate from equation (4.2) beyond that distance.

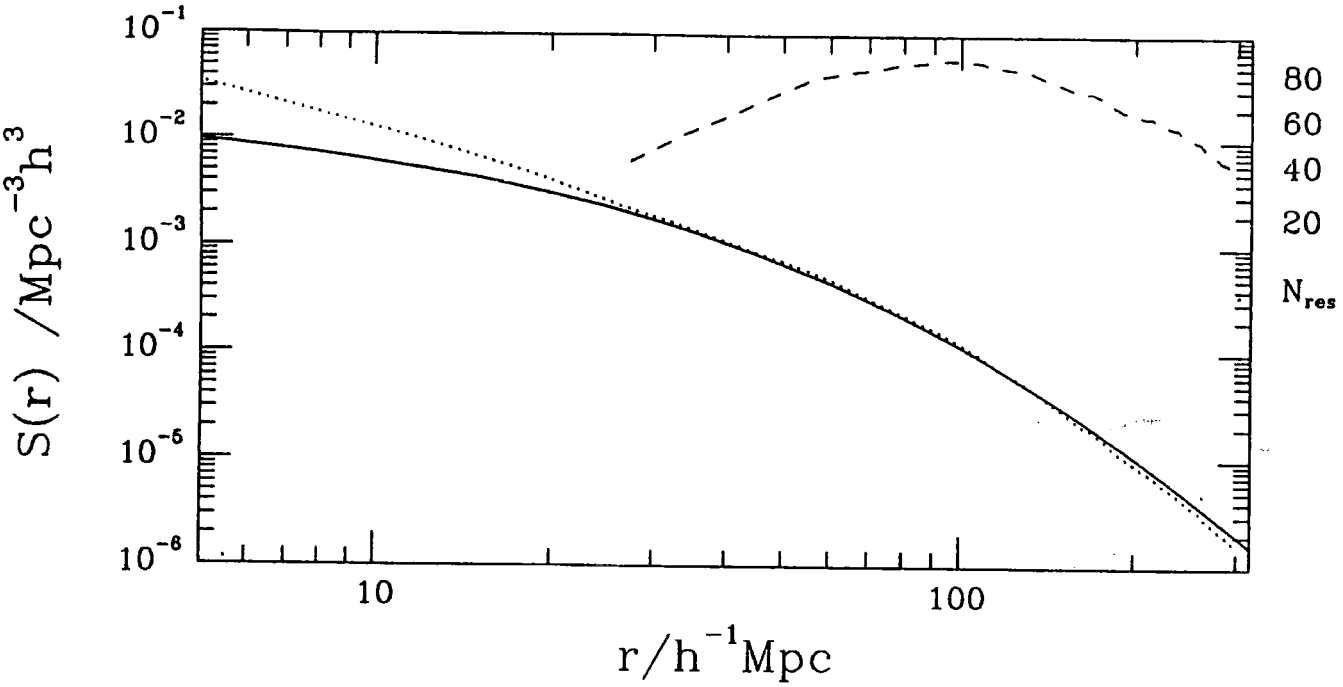


Figure 4.2 The selection function of the QDOT survey (solid line) calculated from the luminosity function of Saunders *et al.* (1990). The selection function calculated using equation (4.3) is shown as the dotted line. The dashed line shows the number of independent resolution elements (survey volume divided by the smoothing volume), it reaches a maximum of 80 at $r = 100h^{-1}$ Mpc.

We now make plots of the QDOT survey by carving out a sphere from the survey, tabulating the smoothed density field in a lattice and drawing isodensity contours above or below a given threshold. Our galaxy is located at the centre of the sphere with the galactic plane running horizontally across the plot. The galactic centre lies towards the right hand side in the positive x-axis direction.

In Figure 4.3a we plot the density field out to a distance $r_{max} = 75h^{-1}$ Mpc using a smoothing length $\lambda = 12h^{-1}$ Mpc. The surfaces in Figure 4.3a are drawn at the contour corresponding to the high density regions which encompass one third of the total volume of the sphere. The dominant feature in this plot is the "Great

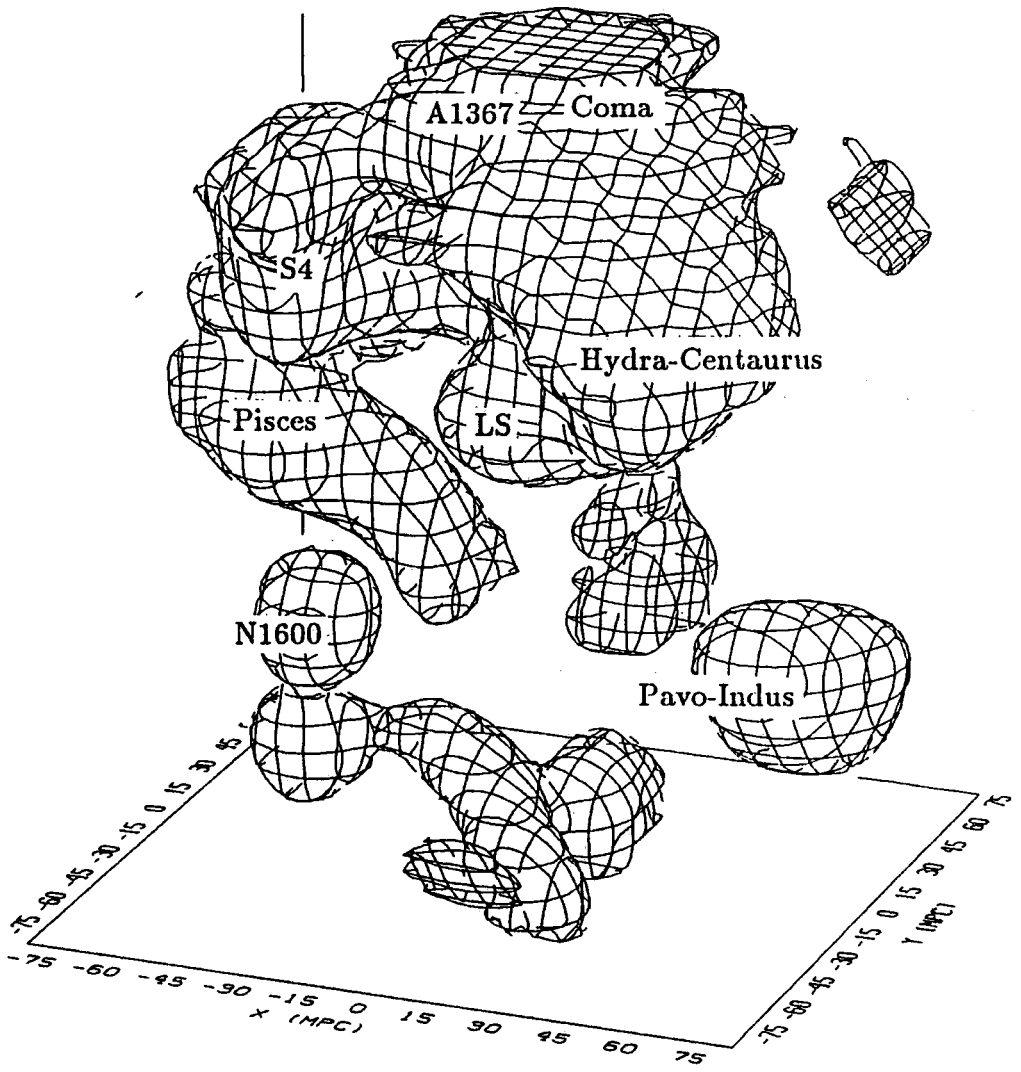


Figure 4.3 High and low density regions of the QDOT survey which enclose roughly one third of the total volume. The galactic centre points towards the right hand side of the sphere along the +ve x-axis with galactic longitude running anti-clockwise around the sphere. 13(a) The high density regions within a sphere of radius $75h^{-1}$ Mpc and smoothed with a Gaussian of width $\lambda = 12h^{-1}$ Mpc. The plot is dominated by the "Great Attractor" region (GA) and the Local-Supercluster (LS).

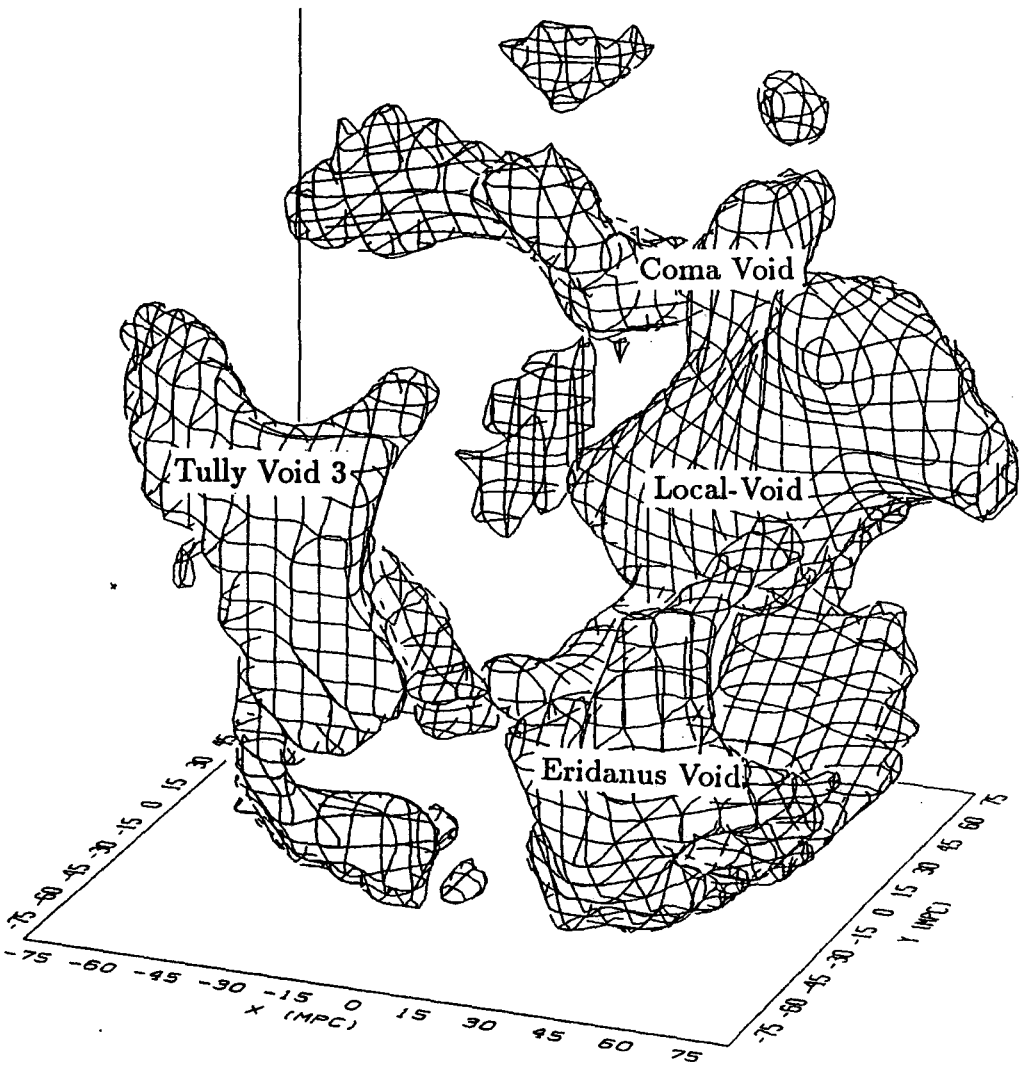


Figure 4.3 (*b*) The low density regions on the same scale and with the same smoothing length as (*a*).

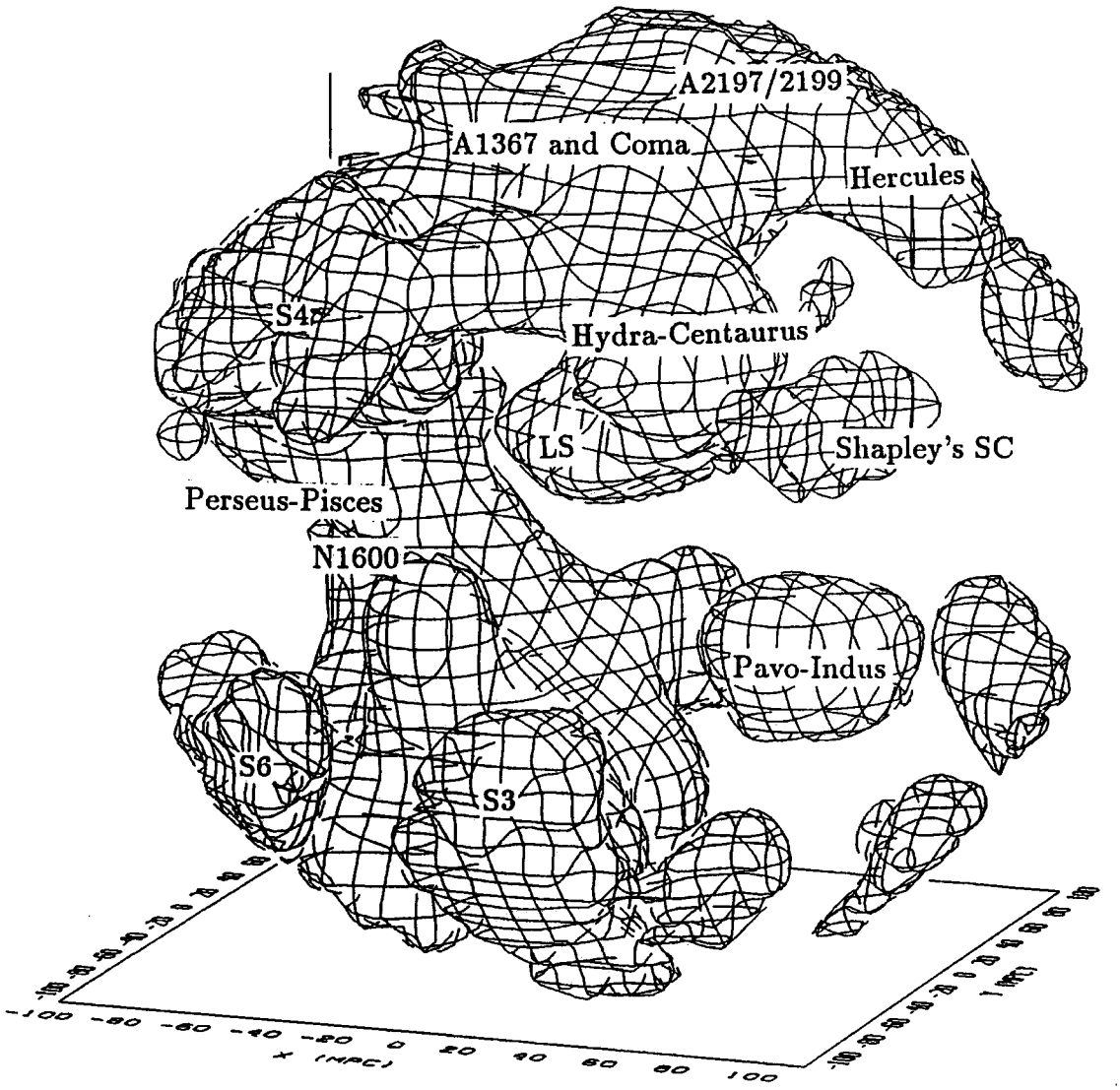


Figure 4.3 (*c*) The high density field to a depth of $100h^{-1}$ Mpc and smoothed on scale $\lambda = 15h^{-1}$ Mpc.

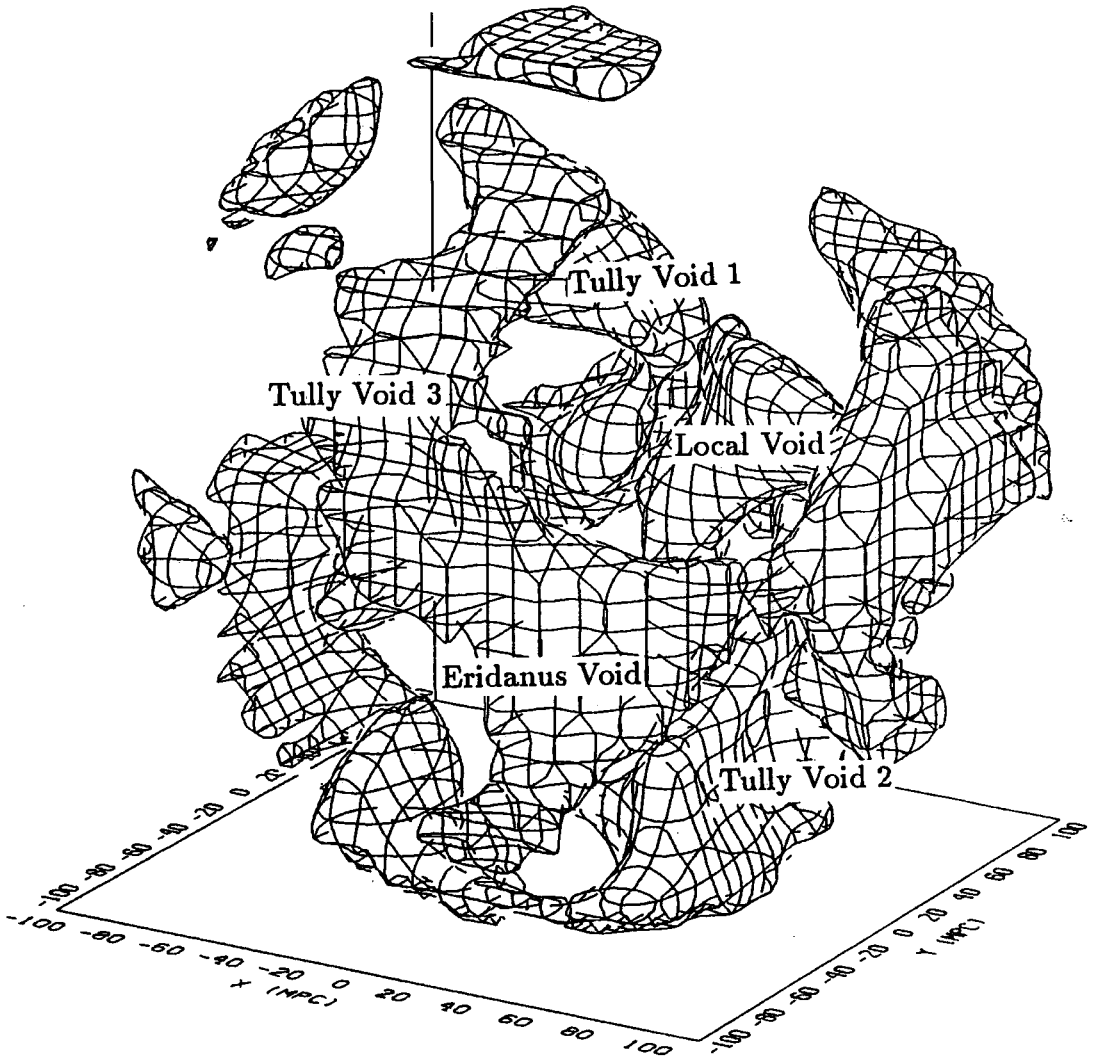


Figure 4.3 (*d*) The corresponding low density regions to (*c*) on the same scale and smoothing length.

Attractor" region (GA) (Burstein *et al.* 1990 and references within) consisting of the Hydra-Centaurus and Pavo-Indus superclusters. The Great Attractor joins smoothly to Coma and A1367 in the northern hemisphere. From this vantage point the large overdensity of the Local-Supercluster (LS) is hidden partially by the clusters forming the Great-Attractor. The surface in Figure 4.3b represents the low density regions which encompass one third of the total volume. This map is dominated by the Local Void which joins to the Eridanus Void in the Southern hemisphere. Tully Void 3 (Tully 1987) is also clearly seen.

Figures 4.3c and 4.3d show similar plots but on the larger scale $r_{max} = 100h^{-1}$ Mpc with $\lambda = 15h^{-1}$ Mpc. As before, the surfaces encompass roughly one third of the total volume. In Figure 4.3c the Local Supercluster now lies insignificantly at the centre of the picture with A1367, Coma, A2197 and Hercules joining together to form part of the Great Wall in the Northern hemisphere. Joined to this huge structure with a bridge of galaxies passing through the galactic plane is Perseus-Pisces and N1600, which make up the frightening Pisces-Cetus complex. The Hydra, Centaurus and Pavo-Indus clusters which delineate the Great Attractor region are still visible in this plot. The superclusters S2 and S6 were identified by Saunders *et al.* (1991). Figure 4.3d shows the corresponding low density regions on the same scale and the same smoothing. The Tully Voids are all present along with the Eridanus and Local Voids which join together smoothly to form a continuous surface. The overdensity of the structures shown in these pictures is about $\pm 1\sigma$ from the mean density.

In Figure 4.4a the threshold is such that the surface encompasses about a tenth of the total volume of the sphere. These high density regions correspond to roughly 1.7σ peaks of the density field. The inter-connected structures in Figure 4.3c have broken up and only isolated overdense super-clusters, such as Hercules and Pavo-Indus, remain visible. Figure 4.4b shows the high density field to $r_{max} = 150h^{-1}$ Mpc smoothed with $\lambda = 24h^{-1}$ Mpc. Hercules, the most dominant feature

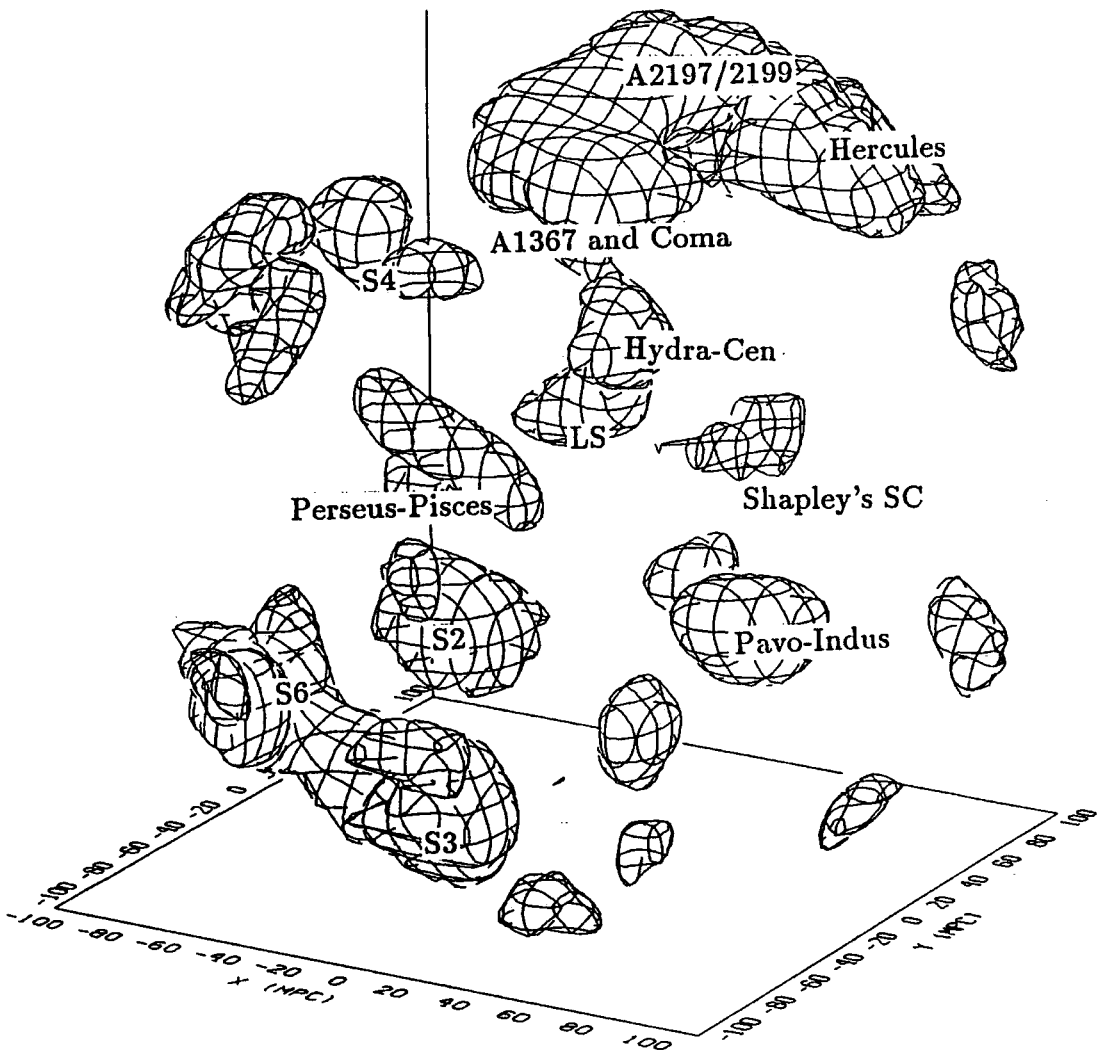


Figure 4.4 Contour plots of the QDOT survey with coordinates as in Figure 4.3. (*a*) The high density field of Figure 4.3c but at a higher threshold so that roughly only one tenth of the total volume is enclosed. The continuous structures break up into isolated clusters which can be easily identified with well known features.

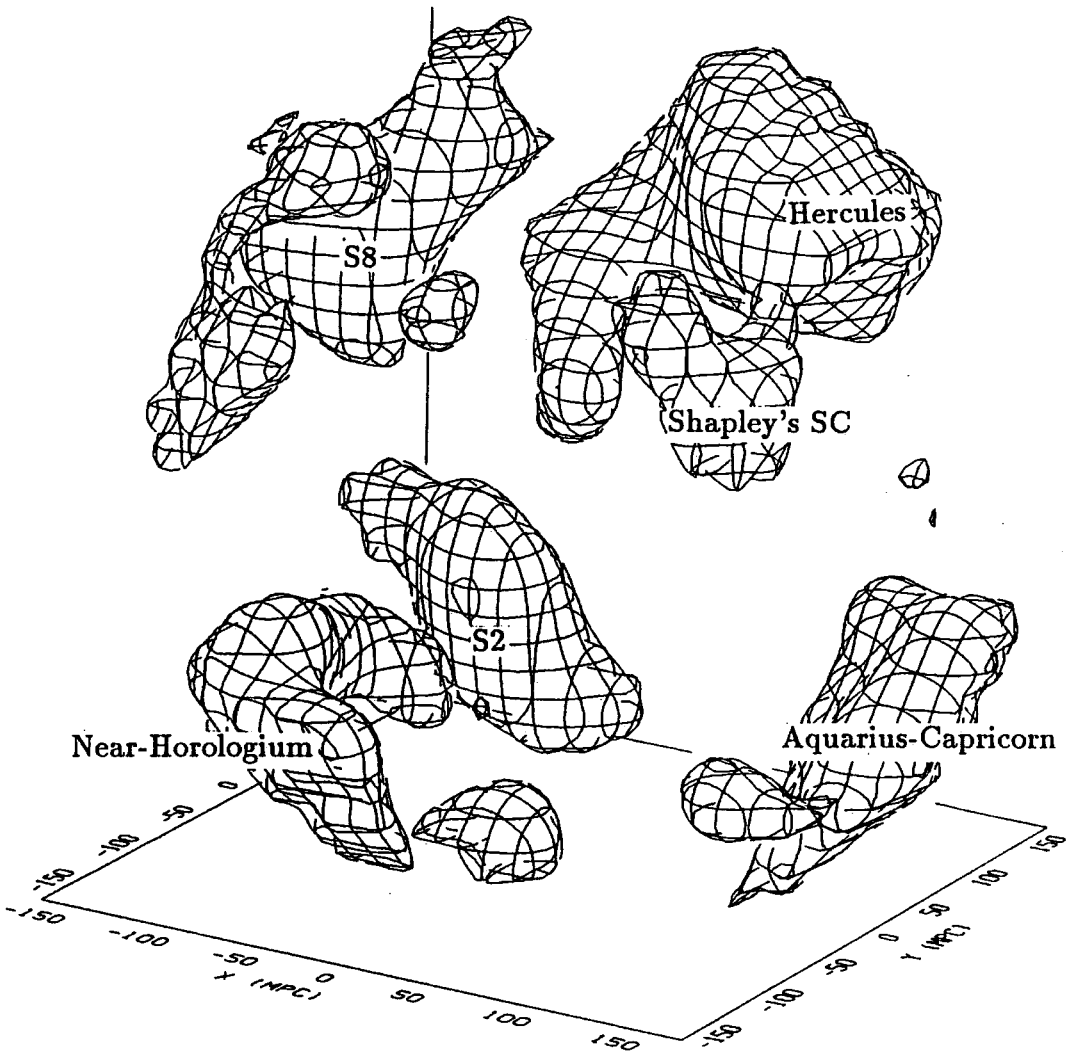


Figure 4.4 (*b*) The high density regions on larger scales; the radius of the sphere has been increased to $150h^{-1}$ Mpc and the data smoothed with $\lambda = 24h^{-1}$ Mpc.

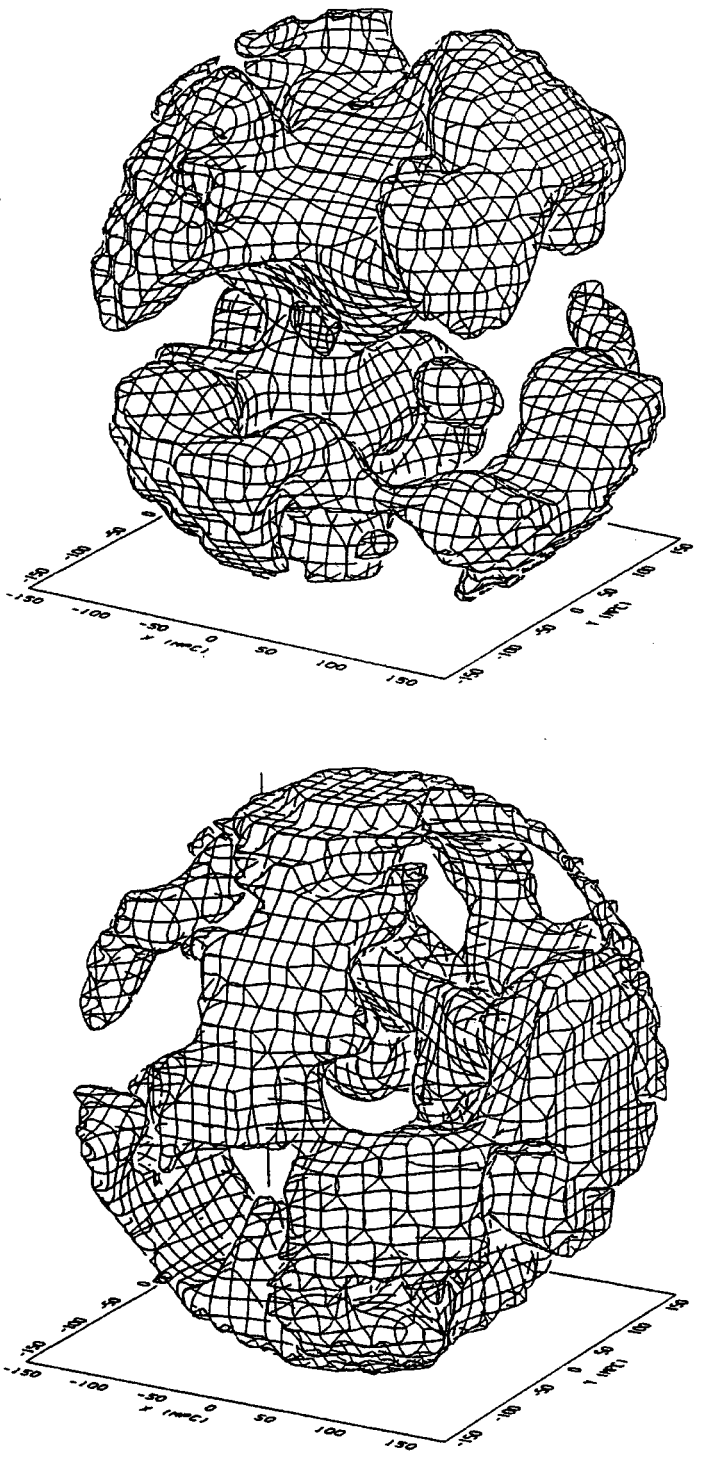


Figure 4.4 (*c*) The same density field as (*b*) but at the median density contour. This plot shows the high density regions which make up one half of the total volume. (*d*) The corresponding low density regions to (*c*) The structures are interlocking and sponge like, representative of a Gaussian random field.

in this plot and in the whole QDOT survey, extends from $100h^{-1}$ Mpc to beyond the scale of this diagram. The superclusters Aquarius-Capricorn, Near-Horologium and S2 dominate the Southern hemispheres. The structures shown in Figure 4.4b have a similar overdensity to those in Figure 4.3, approximately 1σ from the mean. Over 200 galaxies lie within these superclusters. If we raise the threshold so that only 2σ peaks remain visible on this scale then we only see the Hercules supercluster containing 27 galaxies. Finally, on the same scale and smoothing length as Figure 4.4b, Figures 4.4c and 4.4d show the interlocking median density contour of the high and low density regions. The similarity between the high and low density regions and their apparent connectedness is representative of a sponge or Gaussian random field topology.

For comparison with the QDOT data depicted in Figure 4.3c, Figure 4.5a shows a similar plot of the density field in an artificial catalogue constructed from a CDM N-body simulation as described in Section 1. Figure 4.5b shows the corresponding low density regions on the same scale and smoothing. The artificial catalogue's galaxy distribution is treated in an identical way to the real QDOT survey data. The topology of the high and low density fields of the model look quite similar to that of the QDOT survey. On this smoothing scale the slope of the CDM spectrum can be approximated by a power law of slope $n = -0.25$, where $P(k) \propto k^n$.

It is illuminating to compare the plots of the QDOT survey and CDM model with similar plots of distributions with different power spectra. Figure 4.6a shows high density regions from a smoothed density field laid down using a power law fluctuation spectrum with $n = 1$ on all scales. The absence of any large scale overdensities is characteristic of such a power spectrum. Similarly, Figure 4.6b shows a smoothed density field with $n = -2$. This spectrum has dominant fluctuations larger than the $100h^{-1}$ Mpc sphere we have shown here. The power law spectra are scale free density fields which we normalised by specifying $\lambda_{sf}/r_{sf} = \lambda_{QDOT}/r_o$

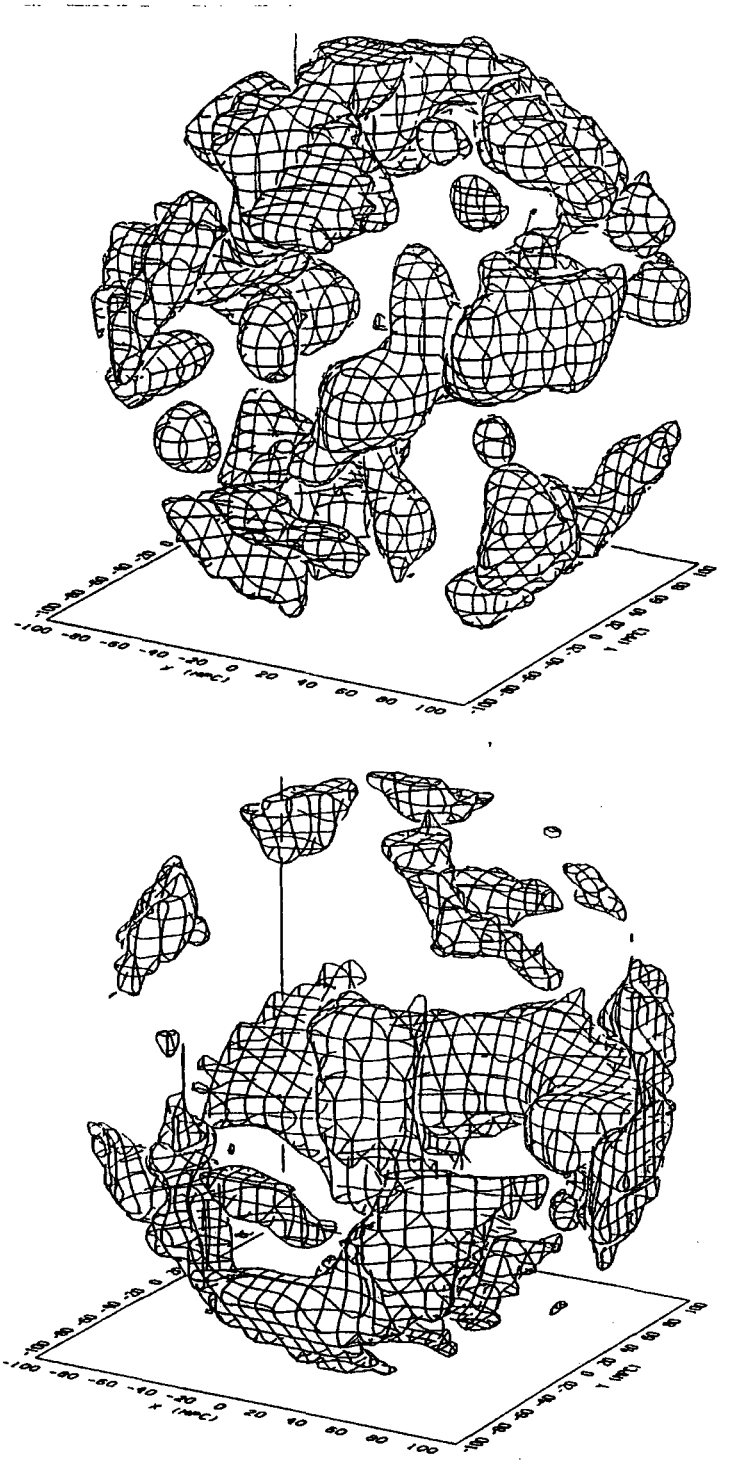


Figure 4.5 (*a*) The high density field of an artificial "QDOT" survey constructed from a CDM N-body simulation with the same selection function, flux limit and volume as the QDOT survey. (*b*) The low density field of the same artificial catalogue. The smoothing lengths and volumes of these plots are the same as those used to produce Figures 4.3c and 4.3d, with which these plots should be compared.

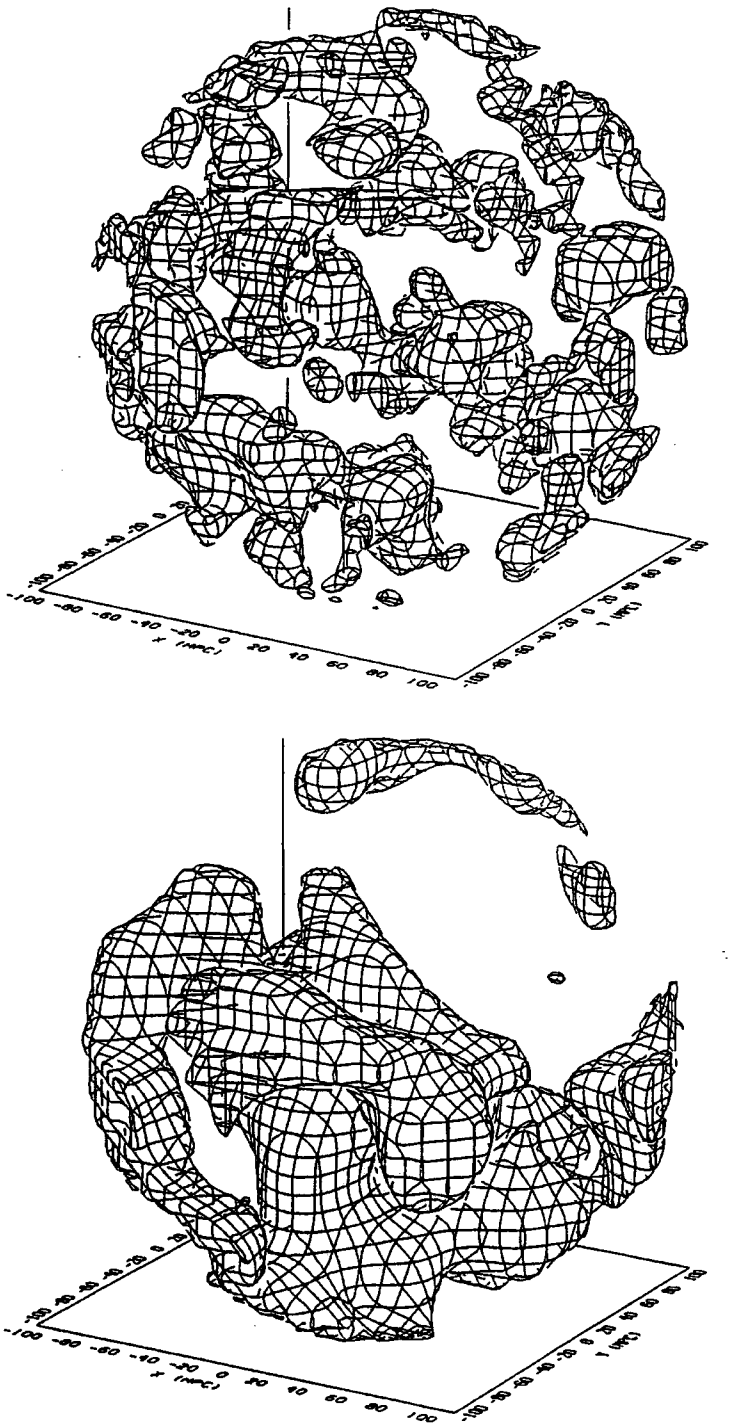


Figure 4.6 The high density regions of density fields with power law fluctuation spectra of spectral index $n = 1$ (a) and index $n = -2$ (b). The smoothing lengths and volumes of these plots are the same as those used to produce Figure 4.3c.

where λ_{sf} and λ_{QDOT} are the smoothing lengths used on the model and the data respectively and r_{sf} and r_o are their respective clustering lengths. We obtain r_{sf} directly by calculating the Fourier transform of the power spectrum (Peebles 1980), and we now discuss the calculation of r_o.

Figures 4.3 and 4.4 show the presence of very large structures in the survey. It is interesting to compare the scale of these structures with the characteristic clustering scale in the survey obtained from two-point correlation function $\xi(s)$. This statistic is complementary to the topological analysis; $\xi(s)$ depends on the amplitude and shape of the power spectrum of density fluctuations whereas the topology depends only on the slope of the power spectrum.

The clustering properties of galaxies can be measured in various ways. The standard estimator of galaxy clustering is the two point correlation function, $\xi(s)$, which describes the excess probability over random of finding a galaxy at a distance s from a chosen galaxy. The correlation function is usually approximated by a power law, $\xi(s) = (s/s_o)^{-\gamma}$; for optically selected galaxies the clustering length $s_o \approx 500\text{km s}^{-1}$ and $\gamma \approx 1.8$ (Davis *et al.* 1988). These values are based on the analysis of the CfA redshift survey (Huchra *et al.* 1983) and the Southern Sky Redshift Survey (da Costa *et al.* 1988). An independent redshift catalogue of 2278 IRAS galaxies complete to a flux limit of 2Jy by Strauss and Davis (1988) gives a correlation function with similar slope but a smaller clustering length $s_o \approx 450\text{km s}^{-1}$ (Davis *et al.* 1988).

We calculate $\xi(s)$ from pair counts in redshift space in the same way as Davis *et al.*

$$1 + \xi(s) = \frac{DD(s)}{DR(s)} \left(\frac{N_R}{N_D} \right) \tag{4.4}$$

where $DD(s)$ is the number of pair counts of galaxies at velocity separation s, and $DR(s)$ is the number of pairs consisting of a galaxy and a point chosen at random

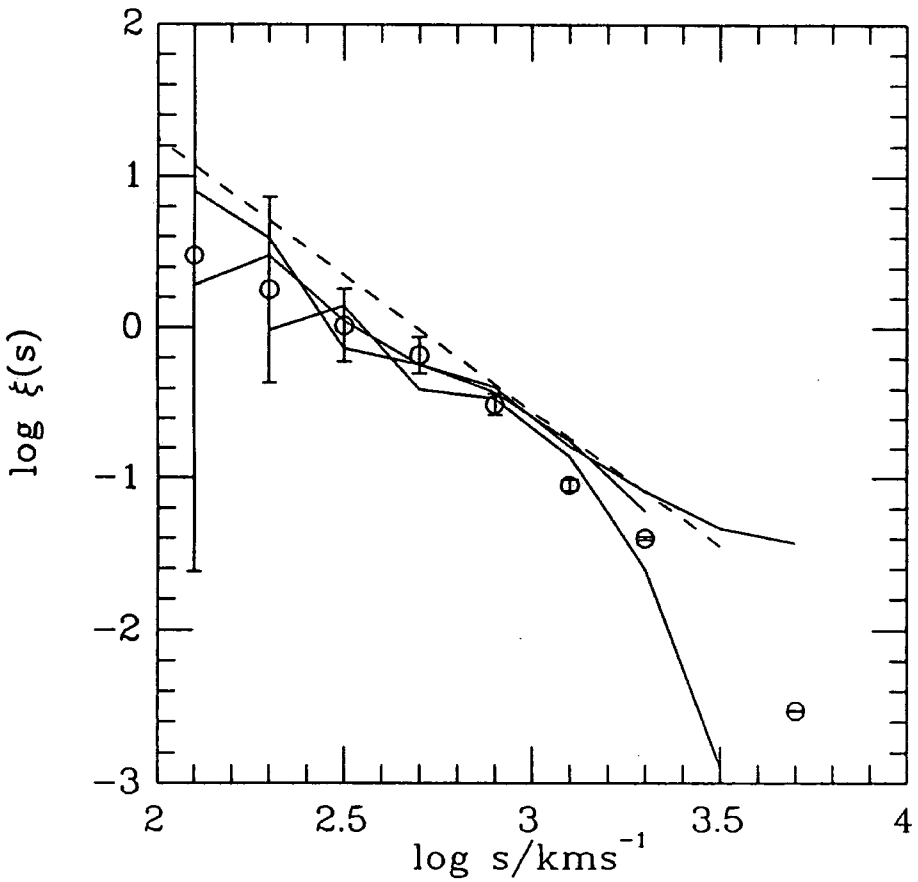


Figure 4.7 The two point correlation function of IRAS galaxies (open circles). The three lines show subsamples of the data which are volume limited at $s_{max} = 4000, 6000$ and 8000km s^{-1}. The dashed line has slope $\gamma = -1.8$ and $s_o = 500$km s^{-1}.

within the same catalogue window. The total number of points in the random and galaxy catalogues are denoted by N_R and N_D respectively. It is important to treat the IRAS mask correctly since it could introduce spurious large scale correlations. If a randomly distributed galaxy falls into a masked area then it is removed from the catalogue. To test the sensitivity of our our estimate to different weighting schemes and to check that our survey provides a "fair sample", we calculate $\xi(s)$

for the survey as a whole, and for 3 subsamples, volume limited at 4000, 6000 and 8000km s^{-1} .

Figure 4.7 shows the correlation function of the QDOT survey. From a least squares fit to the open symbols we find a slope for $\xi(s)$ of 1.4 ± 0.6 and a value of $s_o = 340 \pm 100$km s^{-1}. The errors are from the 1σ scatter of five simulated QDOT surveys constructed from the N-body simulations. The three lines show separate correlation functions for the volume limited subsamples. The small discrepancy in s_o between the QDOT survey and that of Strauss *et al.* could be due to the fact that the mean luminosity of galaxies within the Strauss *et al.* survey is about a factor of three brighter than those in the QDOT survey. The correlation function of the CDM artificial QDOT catalogues is similar to that of the real QDOT survey except that it falls towards zero at $\approx 20,000$km s^{-1}.

The plots in Figure 4.5 and 4.6 contrast the structures we would expect to observe depending on the form of the initial power spectrum. Quantifying the differences between these density fields is the subject of the next section.

4.3 QUANTITATIVE TOPOLOGY

The genus of a surface defined by some arbitrary threshold density is related to

$$\text{genus} \approx (\text{Number of holes}) - (\text{Number of isolated regions}) \qquad (4.5)$$

where "hole" means a hole like that in a doughnut and where an isolated (*i.e.* compact) region may be above or below the threshold density. For example, a torus has a genus of zero and an isolated sphere has a genus of minus one. GMD used the Gauss-Bonnet theorem to express the genus G_s, in terms of the integral over the contour surface of the Gaussian curvature K, defined as the reciprocal of the product of the two principal radii of curvature;

$$G_s = -\frac{1}{(4\pi)} \int_s K dA. \qquad (4.6)$$

As an example consider the simple case of a density contour which defines two isolated spherical clusters of radius r_s. The Gaussian curvature at all points on each sphere is $K = r_s^{-2}$ and the total area is $A = 2(4\pi {r_s}^2)$ which gives $G_s = -2$. This is as expected since we have two isolated regions and no holes. A contour threshold can be specified by its overdensity with respect to the mean $\overline{\rho}$, $\delta\rho/\rho = (\rho - \overline{\rho})/\rho$. The mean density contour of a Gaussian random field will have a positive genus, since its surface is all in one piece and it is multiply connected with many holes, *i.e.* the integrated curvature is negative since the two principal radii of curvature tend to point in opposite directions.

A formula for the genus per unit volume, g_s, was independently derived for Gaussian initial conditions by Doroshkevich (1970), Adler (1981), Bardeen *et al.* (1986; hereafter BBKS) and Hamilton *et al.* (1986):

$$g_s = N(1 - \nu^2)e^{-\nu^2/2} \tag{4.7}$$

where

$$N = \frac{1}{(2\pi)^2}\left(\int k^2 P'(k) d^3k \Big/ 3\int P'(k) d^3k\right)^{3/2}. \tag{4.8}$$

Here, ν is the number of standard deviations by which the contour threshold overdensity is above or below the average density and $P'(k)$ is the power spectrum smoothed using equation (4.1), $P'(k) = P(k)e^{-k^2\lambda^2/2}$. Smoothing a distribution introduces a short wavelength cut off in the power spectrum.

For a Gaussian random field positive and negative fluctuations are statistically indistinguishable; hence the genus curve is symmetric about the mean density. The transition from a positive to a negative genus occurs at $\nu = \pm 1$, which marks the change from multiply connected contours to isolated clusters or voids. The shape of $g_s(\nu)$ is independent of the amplitude of the power spectrum since this cancels out of equation (4.8). However the k^2 factor in equation (4.8) implies

that the normalisation of the curve depends primarily on the slope of the power spectrum at smoothing length λ.

For the simple case of a power-law spectrum, $P(k) \propto k^n$, the normalising factor can be calculated explicitly:

$$N = \frac{1}{(2\pi)^2 \lambda^3} \left(\frac{3+n}{3}\right)^{3/2} \qquad n \geq -3. \tag{4.9}$$

By studying the topology using different smoothing lengths, we constrain the shape of the power spectrum over a range of scales.

If the density fluctuations are non-Gaussian then the genus curve may be skewed. For example, if galaxies sat mainly in isolated clusters surrounded by thin voids (rather like a baked bean topology), the genus curve would be skewed to the left (Weinberg *et al.* 1987; hereafter WGM); one would have to look at a very small volume fraction of the data before the voids could be detected. On the other hand, if galaxies sat on the surfaces of large voids, the genus would be skewed to the right. In this case the characteristic asymmetry in the genus curve could be detected if the radii of the voids was less than about $\lambda/4$ (WGM). A symmetrical genus curve does not guarantee that the density field is Gaussian; for example, the square of a Gaussian field would have a symmetric genus curve. If, as in the standard CDM model, the galaxy distribution is biased relative to the mass in such a way that the mass to light ratio is a monotonic function of the total density, then contours of constant mass density are also contours of constant luminosity density. Provided that contours are identified by volume fraction, then contours in the biased galaxy distribution are very close to those in the unbiased mass distribution. This is a direct consequence of the fact that high density peaks of the mass distribution tend to map into high peaks of the galaxy distribution.

4.4 ANALYSIS

4.4.1 The data

The galaxies are binned in a 64^3 array and then smoothed with the Gaussian filter of equation (4.1) with width λ taken to be equal to the mean inter-galaxy separation at the edge of the survey, at distance r_{max}:

$$\lambda = [S(r_{max})]^{-1/3}, \tag{4.10}$$

where $S(r)$ is the selection function defined in Section 2. This choice of λ is made in order that the density field is not severely undersampled. A contour threshold is specified by means of the fractional volume f occupied by the high density regions:

$$f = (2\pi)^{-1/2} \int_{\nu}^{\infty} e^{-t^2/2} dt. \tag{4.11}$$

The curvature of a particular contour identified in the cubic lattice is concentrated at the vertices where square faces meet. (The faces have zero curvature.) The curvature at a vertex i on the surface is equal to its angle deficit $D_i = 2\pi - \sum A_i$ where A_i are the surrounding vertex angles (see GMD). The genus of the surface can then be calculated directly as $G_s = -\sum D_i/4\pi$. We used the program "CONTOUR" described by Weinberg (1988) which uses this technique to measure the curvature of density contours.

Smoothing in Fourier space requires a periodic box; to allow for periodic effects the region we are interested in is embedded into a larger box. The space outside r_{max} has to be treated carefully so as to minimise the effects of the boundary on the genus curve $g_s(\nu)$. If this space is left empty then smoothing brings down the mean density at the edge of the survey, skewing the genus curve towards the right. We tried several ways of dealing with this. *(i)* Include the galaxy distribution outside r_{max}; (This is possible because the $n(v)$ distribution extends well into high redshifts.) *(ii)* Set the density in all the cells with $r > r_{max}$ to

$S(r_{max})$ and then smooth the data. *(iii)* Smooth the data assuming zero density outside r_{max}, and divide this array by the smoothed selection function $S(r)$, also with zero outside r_{max}. We tested each of these methods on artificial "QDOT surveys" constructed from CDM simulations by comparing the genus curves of the data with the expected curves calculated using equation (4.8) and the CDM power spectrum. The least satisfactory method was *(ii)* which tended to overestimate the mean density at the edge of the array. Method *(i)* worked well for distances up to $100h^{-1}$ Mpc, but beyond that Poisson noise in the galaxy distribution outside r_{max} introduced noise into the results. We adopted method *(iii)* which worked well at all distances. The 64^3 array containing the galaxy selection function was multiplied by the volume of each cell visible to the observer, taking into account the volume hidden by the IRAS mask (which excludes about 8% of the sky above and below the galactic plane).

To obtain the maximum resolution, r_{max} should give the largest number of smoothing volumes, $V_{sm} = \pi^{3/2}\lambda^3$, per survey volume. The dashed line in Figure 4.2 shows how the number of resolution elements N_{res}, varies with distance.

$$N_{res} = \frac{\Omega_s r^3}{3\pi^{3/2}\lambda^3}. \tag{4.12}$$

The number of resolution elements reaches a maximum of 80 at $100h^{-1}$ Mpc where the mean galaxy separation is $20h^{-1}$ Mpc. This is larger than any previous redshift survey (The CfA1 survey, for example, has a maximum N_{res} of 54). We have examined the topology on a range of smoothing scales between $10h^{-1}$ Mpc and $50h^{-1}$ Mpc. Table 4.1 summarises the smoothing lengths we adopted, the distance r_{max} at which we cut the data off and the number of resolution elements per survey volume.

Figures 4.8a-d show the genus curves of the QDOT data for four different smoothing scales. The open squares show the genus values at 15 threshold over-densities between $\pm 2\sigma$ from the mean density contour. The error bars on the

QDOT data are 1σ and will be discussed in the next section. The smooth curves show the best fit theoretical curve to the data using equation (4.7). The one-in-six sampling strategy has resulted in a loss of information on small scales which is apparent from the large scatter in Figure 4.8a. However, we gain on larger scales and Figures 4.8b-d show that the topology of the QDOT survey is consistent with that expected from a Gaussian random field. There is a slight shift in the genus curves to the left towards an "isolated cluster" topology but not at a very significant level.

4.4.2 Error analysis using N-body simulations

There are two kinds of sampling errors which contribute to the uncertainties in our measurement of the genus curve, errors due to the one-in-six sampling and errors due to having surveyed only one region of space. We estimate these uncertainties for the artificial catalogues constructed from our CDM simulations. From five different simulations we generated 5 fully sampled galaxy catalogues each containing over 12000 galaxies with identical selection function as the QDOT survey. From each of these we then extracted six one-in-six randomly sampled sub-catalogues similar to the QDOT survey.

The filled circles in Figure 4.9 show the mean genus curve from the 5 fully sampled catalogues and the open circles show the average genus curve from the 30 randomly sampled catalogues with parameters $\lambda = 20\mathrm{h}^{-1}$ Mpc and $r_{max} = 100\mathrm{h}^{-1}$ Mpc. Comparing these two genus curves we can see that the random sampling procedure has not introduced any systematic biases and that we have successfully recovered the topology of the fully sampled simulations. The smooth curve in Figure 4.9a is the theoretical genus curve expected from the CDM power spectrum (equations (4.7) and (4.8)). The evolved simulations provide a good fit to this curve demonstrating that non-linear effects have not introduced a significant bias. As a further test of the effects of shot noise in our sparsely sampled catalogues

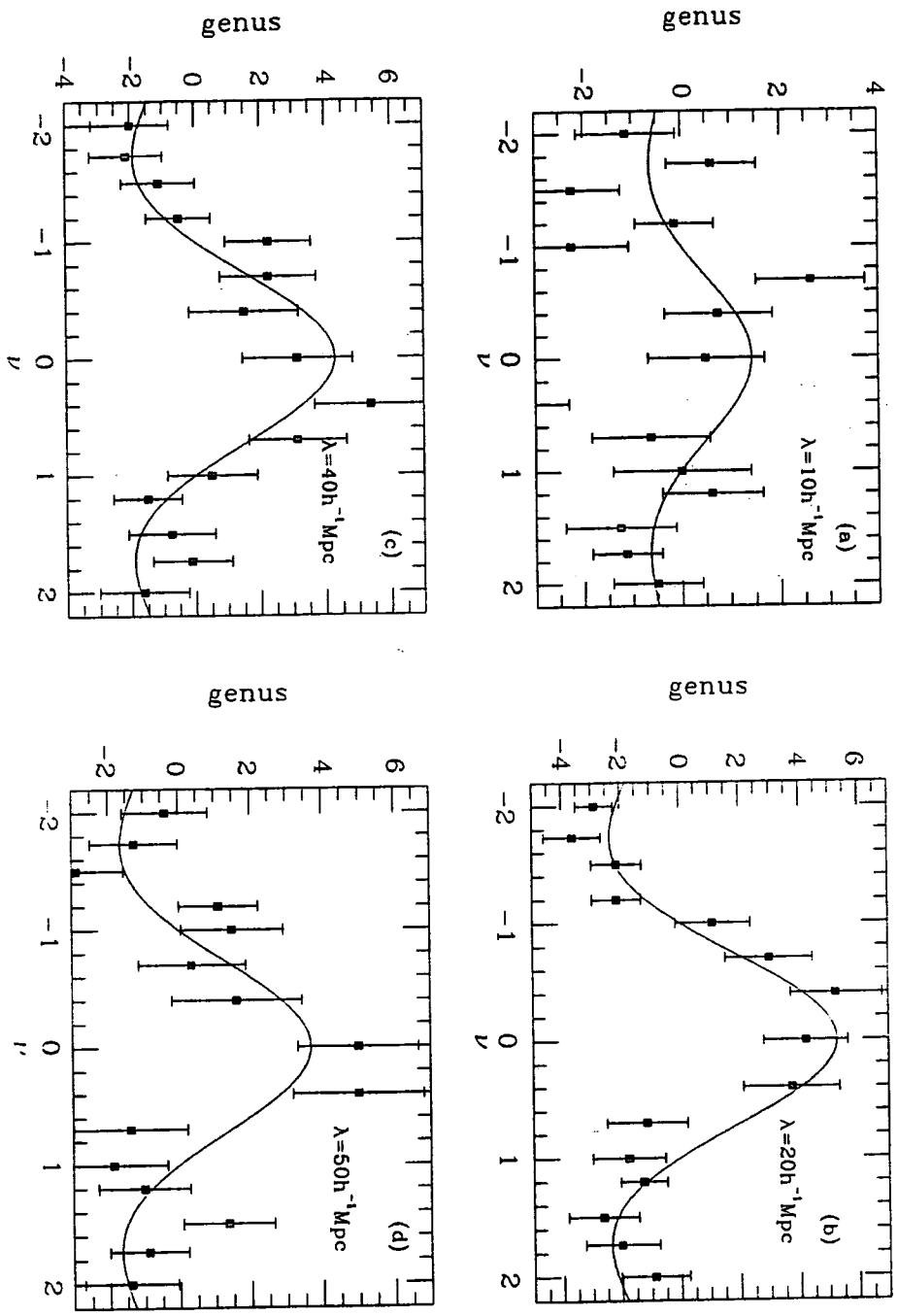


Figure 4.8 The genus curves of the QDOT survey on four different smoothing scales. On the horizontal axis, ν is the number of standard deviations from the mean density contour. The solid curves are the best fit curve of the form given by equation (4.7) to the data.

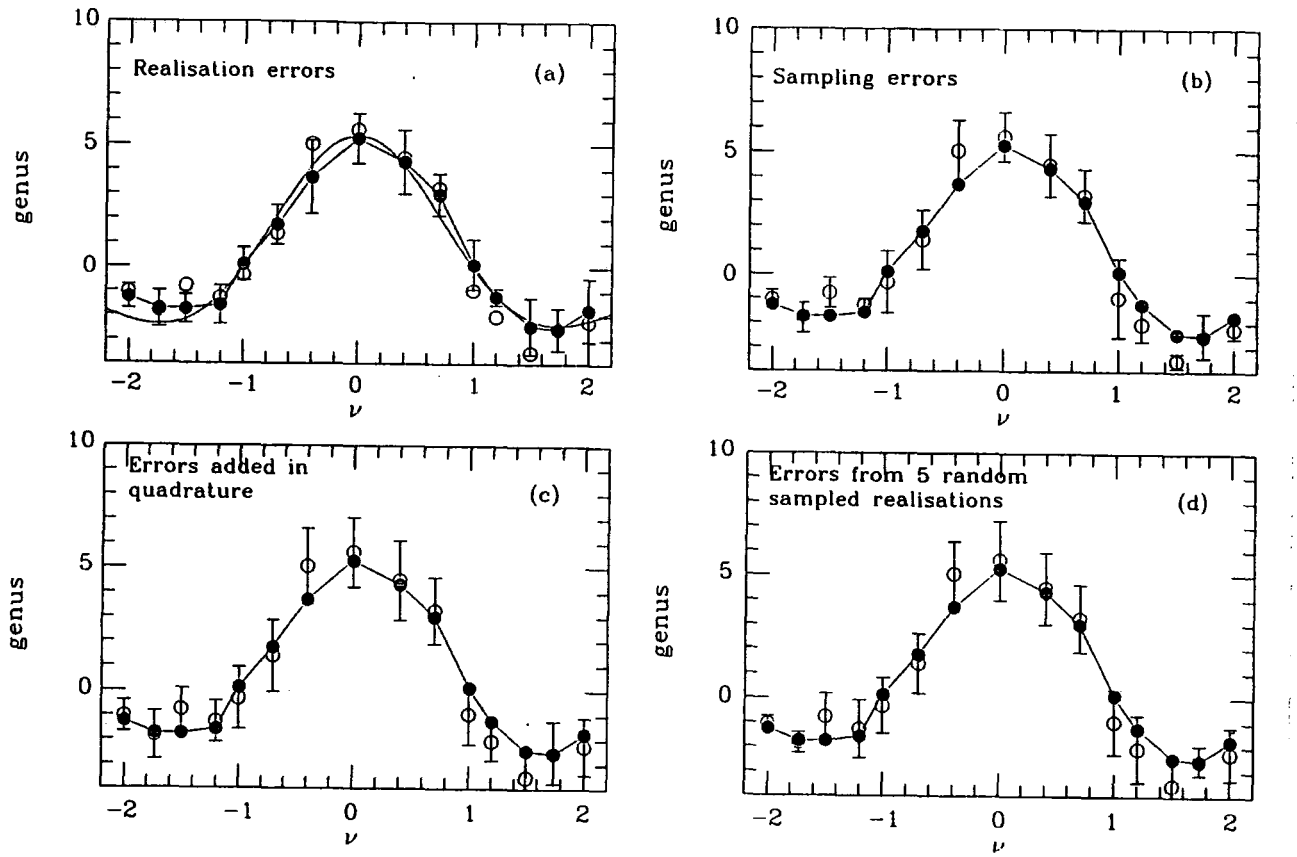


Figure 4.9 The mean genus curves for five fully sampled artificial catalogues made from CDM N-body simulations (filled circles) and the mean curve obtained by random sampling the galaxy distribution from each simulation six times (open circles). The parameters are $\lambda = 20h^{-1}$ Mpc and $r_{max} = 100h^{-1}$ Mpc. The expected genus curve for the CDM model is shown as the smooth curve in (*a*). The error bars are 1σ and arise from; (*a*) the scatter between 5 fully sampled artificial "QDOT" catalogues constructed from 5 separate CDM N-body simulations; (*b*) the scatter between 30 sparsely sampled artificial catalogues; (*c*) the errors from (*a*) and (*b*) added in quadrature; (*d*) the errors from 5 sparsely sampled artificial catalogues, each from separate simulations.

- on the scales we are considering the CDM spectrum is close to shot noise - we compare contour plots of fully sampled simulations with several one-in-six sampled subsets. Figure 4.10a shows an artificial "QDOT" survey constructed from a CDM simulation with $r_{max} = 100h^{-1}$ Mpc and $\lambda = 20h^{-1}$ Mpc. Figures 4.10b-d show three randomly sampled distributions drawn from the same artificial catalogue. The agreement between structures ranges from good to bad depending upon who is making the comparisons.

For a more quantitative test of shot noise, we compare the density per pixel in an artificial QDOT catalogue with the density field obtained from sub-catalogues sampled at a rate one-in-six, with smoothing lengths, $\lambda/2, 3\lambda/4, \lambda$ and $4\lambda/3$ where $\lambda = 20h^{-1}$ Mpc and $r_{max} = 100h^{-1}$ Mpc.

Figure 4.11 shows that when the data are smoothed on a scale smaller than the mean intergalaxy separation λ, at r_{max}, the agreement between the fully sampled and sparsely sampled density fields is poor, but the agreement becomes increasingly better as the smoothing length is increased and is quite acceptable when it is set equal to λ. We also compared the contour plots of the density field of the fully sampled and sparsely sampled distributions at these smoothing lengths and confirmed that λ should be at least as large as the mean inter-galaxy separation at the edge of the survey. As λ was reduced the numbers of spurious structures started to increase rapidly.

The error bars on the genus curves in Figure 4.9a show the scatter between the five fully sampled galaxy catalogues and represent the realisation to realisation differences. The error bars in Figure 4.9b represent the 1σ scatter between the 30 sparsely sampled simulations. Figures 4.9c and 4.9d compare these errors added in quadrature with the 1σ errors obtained from randomly sampling each of the 5 simulations once only. These two methods give very similar error bars for the model, roughly a mean error per genus measurement of ± 1.2. It is also

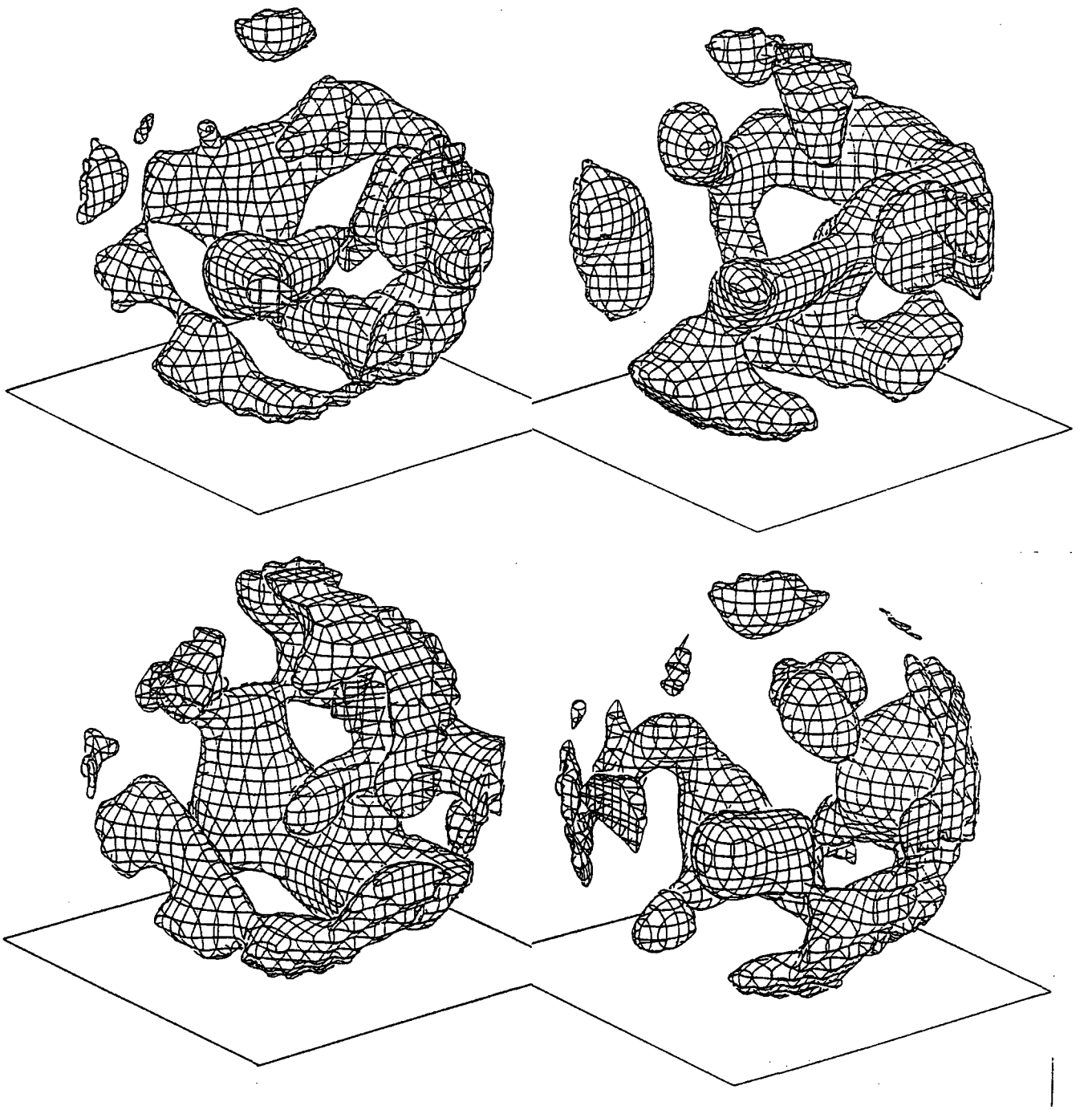


Figure 4.10 (*a*) High density regions of a fully sampled CDM N-body simulation with $\lambda = 20h^{-1}$ Mpc and $r_{max} = 100h^{-1}$ Mpc. (*b-d*) High density regions of catalogues constructed from the simulation in (*a*) by randomly sampling the density distribution at the rate one-in-six.

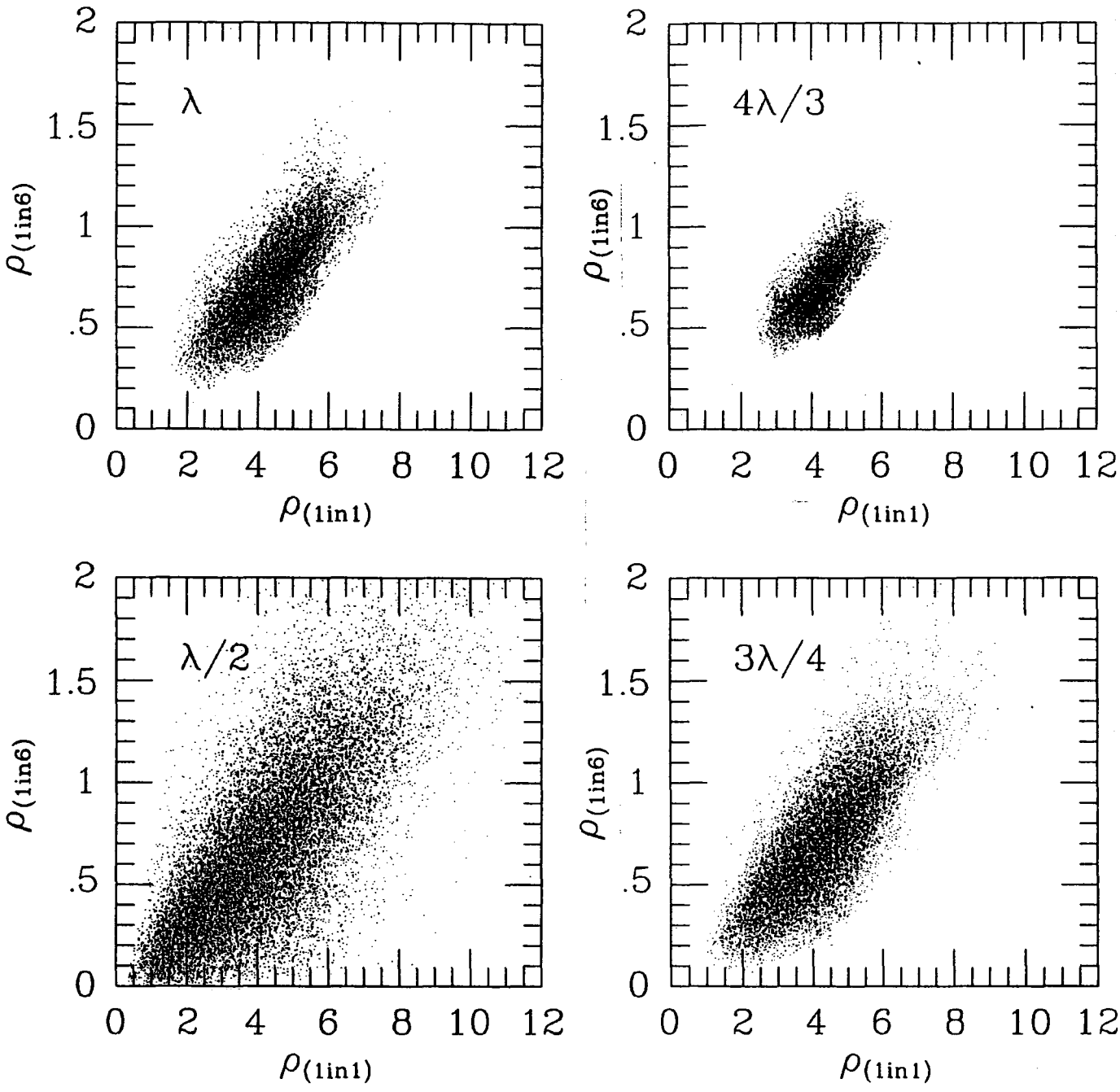


Figure 4.11 The density per pixel obtained from a fully sampled and sparsely sampled (at a rate one-in-six) catalogue. The separate panels show a range of smoothing lengths with $\lambda = 20h^{-1}$ Mpc and $r_{max} = 100h^{-1}$ Mpc.

clear that the scatter introduced by the one-in-six sampling is comparable to the scatter between realisations. Gott *et al.* (1989) have advocated a resampling procedure for estimating the errors in the genus curve. To perform this resampling we assign to each galaxy a position randomly chosen from the original list of galaxy positions. Some positions may have more than one galaxy and some may be left empty with probabilities given by the Poisson formulae $P(N) = (N!)^{-1}e^{-1}$. It is interesting to compare the errors obtained from the scatter between many resampled data sets with the errors estimated from the bootstrap resampling technique. Resampling artificial catalogues gave a mean error per genus of ± 1.35, similar to the error found from resampling the original QDOT survey. This is about 10% larger than the errors from the simulated surveys. It is interesting to note that had we measured the redshifts of all 12,000 galaxies of the fully sampled QDOT catalogue the simulations imply that the amplitude of the genus curve could have been determined only to ± 0.8.

4.5 THE SHAPE OF THE POWER SPECTRUM OF DENSITY FLUCTUATIONS

We showed in Section 4.1 that the genus curve for the QDOT survey has the characteristic shape expected for a Gaussian random field. In this section we use the genus curve to further investigate the density field. We begin by assuming that the galaxies "trace the mass" in the sense that the rms fluctuations in the distributions of galaxies and mass are the same apart from an arbitrary normalisation factor. We then use the genus curves of Section 4.1 and equations (4.7) and (4.9) to obtain the local slope of the power spectrum of density fluctuations on the scale of the smoothing length which we characterise by an "effective" power law index. We obtain this by fitting equation (4.7) to the genus curves to estimate the amplitude, $G_s(0)$, and substituting this into equation (4.9). To calculate the error on the slope one must be careful because points on the genus curve are not independent. As ν is increased the high peaks of the density field simply become

more prominent at the expense of the low density troughs so the contour surface at threshold $\nu + \delta\nu$ is similar to the contour surface at threshold ν. We can make a rough estimate of the value of $\delta\nu$ which causes the number of peaks to fall by a factor of two and gives a roughly independent measurement of the genus of threshold ν. For a Gaussian random field the number of peaks is given by

$$n(\lambda)d\lambda \propto exp(-\frac{\nu^2}{2\sigma^2})d\lambda, \tag{4.13}$$

where λ is the smoothing length and σ is the standard deviation of the peak widths (Bardeen *et al.* 1986). Requiring $n(\lambda)$ to fall by a factor of two gives us $\delta\nu > \sigma$. We have measured the genus over the range $\pm 2\sigma$ from the mean density contour which gives us about five independent points with which to measure just one parameter, the amplitude of the curve.

A rough estimate of the errors in n follows from simply taking the error in the amplitude of the genus curve to be the mean error per genus (computed using the simulations as described above). The errors vary with the volume and smoothing length under consideration so that the mean error per genus measurement must be recalculated for each scale (*i.e.* ± 1.2 for $\lambda = 20\mathrm{h}^{-1}$ Mpc). The values of n and their 1σ errors are listed in the fourth and fifth column of table 4.1. Our most accurate estimate of the slope of the power spectrum corresponds to the scale with the maximum number of resolution elements, $\lambda = 20\mathrm{h}^{-1}$ Mpc On this scale we find $n = -0.79 \pm 0.35$, with values in the range $0.2 < n < -1.8$ ruled out at over the 3σ level. When n is calculated directly from the bootstrapped data we find a significant bias towards lower values such that $n_{data} = n_{boot} - 0.15$.

The above results are summarised in Figure 4.12. The filled circles show our estimate of the effective power spectrum slope at a range of scales between 10 and $50\mathrm{h}^{-1}$ Mpc. Our data are consistent with a constant value $n \approx -1$ over this range. The solid line gives n_{eff} for the CDM spectrum calculated from equations (4.7) and (4.8). For comparison we also show n_{eff} for the Abell cluster sample

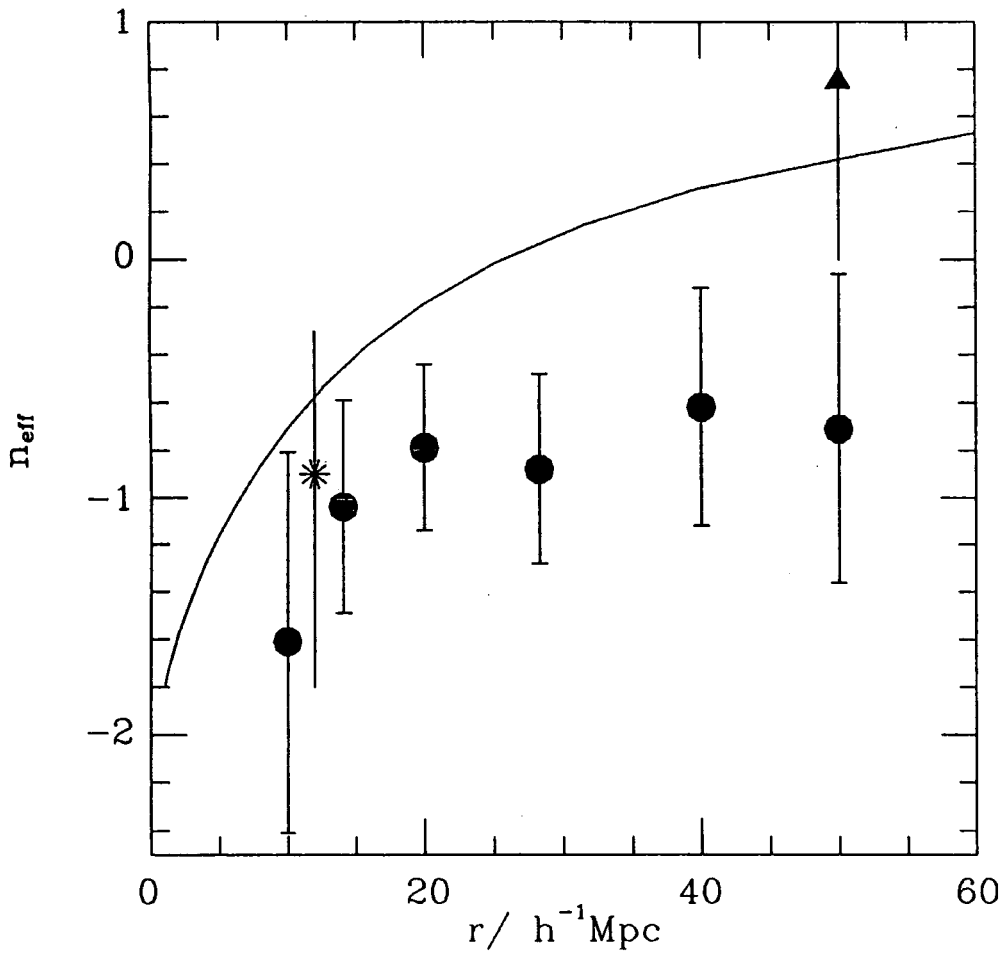


Figure 4.12 The effective spectral index, n_{eff}, for the CDM spectrum as a function of smoothing scale (solid curve). Filled circles show the results from the topology analysis. The star and filled triangle give n_{eff} for the GH redshift survey and the Abell cluster sample respectively.

(filled triangle) and Giovanelli-Haynes sample (star) from Gott *et al.* (1989). The topology of the QDOT survey agrees well with that of the GH survey but is marginally inconsistent with that of the Abell cluster sample. This discrepancy could be due to projection effects in the Abell cluster catalogue.

In the last column of table 4.1 we list the slope of the CDM spectrum on each scale we have studied. In brackets we give the significance level at which this slope is inconsistent with that of the QDOT survey. On our most accurate scale the slope of the CDM spectrum is $n = -0.25$, 1.6σ away from the QDOT result. The largest difference occurs on $28h^{-1}$ Mpc scales on which the CDM spectrum has $n \approx 0$ and the QDOT data has a spectrum of slope $n = -0.88 \pm 0.4$.

λ /h^{-1}Mpc	r_{max} /h^{-1}Mpc	$S(r)$ /h^3Mpc^{-3}	N_{res}	n (QDOT)	n_{eff} (CDM)
10	40	1.0×10^{-3}	48	-1.61 ± 0.80	-0.93(0.8σ)
14	70	3.0×10^{-4}	70	-1.04 ± 0.45	-0.59(1.0σ)
20	100	1.2×10^{-4}	80	-0.79 ± 0.35	-0.25(1.6σ)
28	135	4.2×10^{-5}	66	-0.88 ± 0.40	0.07(2.4σ)
40	175	1.6×10^{-5}	52	-0.62 ± 0.50	0.34(1.9σ)
50	210	8.1×10^{-6}	46	-0.71 ± 0.65	0.48(1.8σ)

Table 4.1. Characteristics of the QDOT survey and CDM predictions

To obtain a more accurate estimate of the uncertainties in our determination of the spectral power index, n, we carried out a series of Monte-Carlo simulations. For a particular value of n, we lay down a density field in a 64^3 grid with the specified power spectrum. We then divide the smoothed density field by the smoothed mask and assume the same value of λ/r_{max} as for the QDOT data. Repeating this 200 times we calculate the mean genus curve for this particular spectral index, $G(\nu)_{mean}$. Figure 4.13 shows $G(\nu)_{mean}$ for $n = 1, 0, -1, -2$ and the theoretical CDM spectrum for $\lambda = 20h^{-1}$ Mpc and $r_{max} = 100h^{-1}$ Mpc. We then calculate

the statistic

$$\Delta G = \sum_{\nu=1}^{15} \mathrm{Abs}(G(\nu)_{mean} - G(\nu)_{realisation}) \qquad (4.14)$$

where the summation is over the 15 threshold levels at which we calculate the genus.

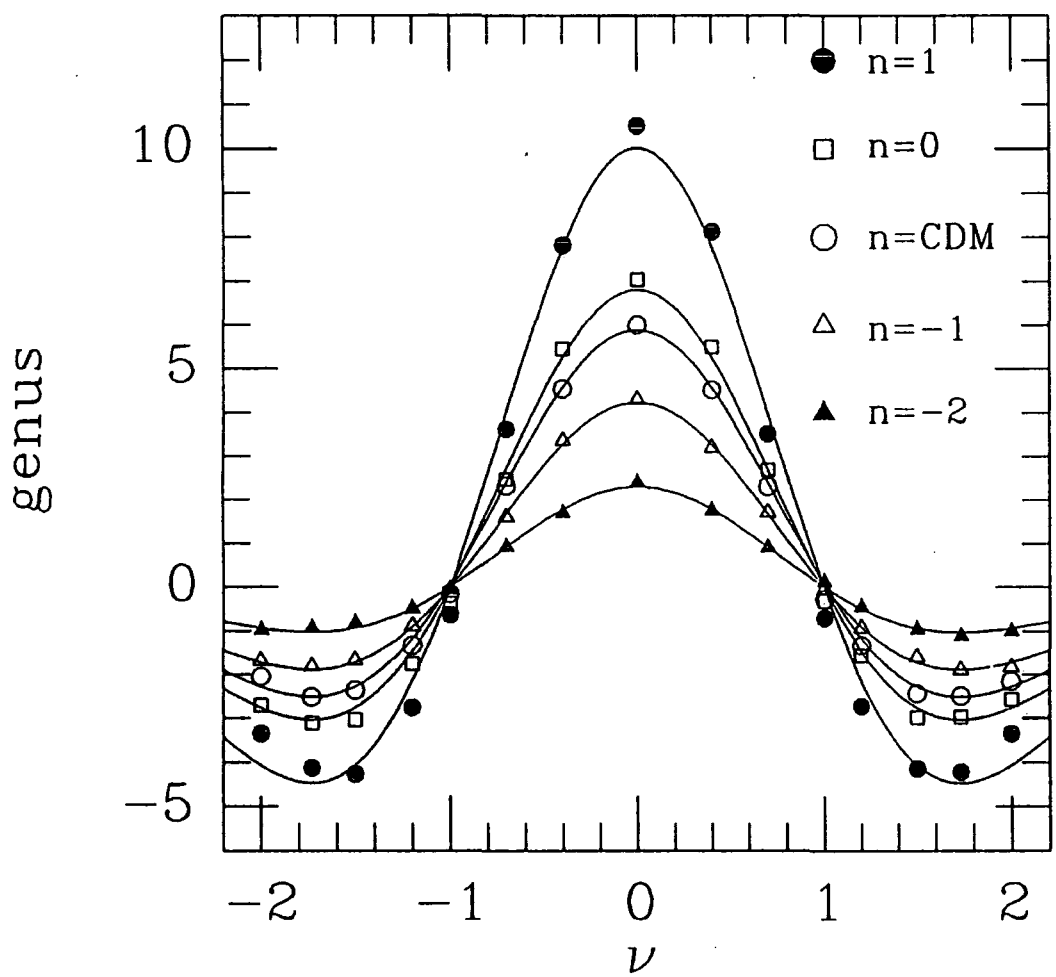


Figure 4.13 The mean genus curves for 200 Monte-Carlo simulations with power law fluctuation spectra of index n =+1, 0, -1, -2 and for the CDM spectrum. The smooth curves show the theoretical curve for each model calculated using equations (4.7), (4.8) and (4.9).

Figure 4.14 shows the distribution of ΔG for $n = 1, 0, -1, -2$ and for the CDM power spectrum. The distributions are symmetric about the mean and become broader for simulations with less power on large scales.

Our analysis of the artificial "QDOT" surveys in Section 4.2 showed that the scatter in the genus curve due to the one-in-six sampling is comparable to the scatter between different realisations of the survey. At least for the case of CDM the effect of the one-in-six sampling can be quantified by the distribution of the quantity

$$\Delta G_{1in6} = \sum_{\nu=1}^{15} \mathrm{Abs}(G(\nu)_{1in1} - G(\nu)_{1in6}), \tag{4.15}$$

which we calculate from six one-in-six artificial QDOT surveys constructed from each of the five CDM N-body simulations discussed in Section 4.2. The distribution of ΔG_{1in6} is shown in Figure 4.14c.

Assuming that the effect of random sampling depends only weakly on n, the convolution of the distributions of ΔG and ΔG_{1in6} gives the expected frequency distribution of the statistic

$$\Delta G_d = \sum_{\nu=1}^{15} \mathrm{Abs}(G(\nu)_{mean} - G(\nu)_{QDOT}), \tag{4.16}$$

where $G(\nu)_{mean}$ is the average genus curve from the 200 Monte-Carlo simulations for each model and $\Delta G(\nu)_{QDOT}$ is the genus curve of the QDOT survey. The distributions of ΔG_d are shown as solid curves in Figure 4.14. We can then make a straightforward calculation of the confidence levels for each assumed power spectrum. The whole procedure must be repeated for each set of smoothing lengths and volumes.

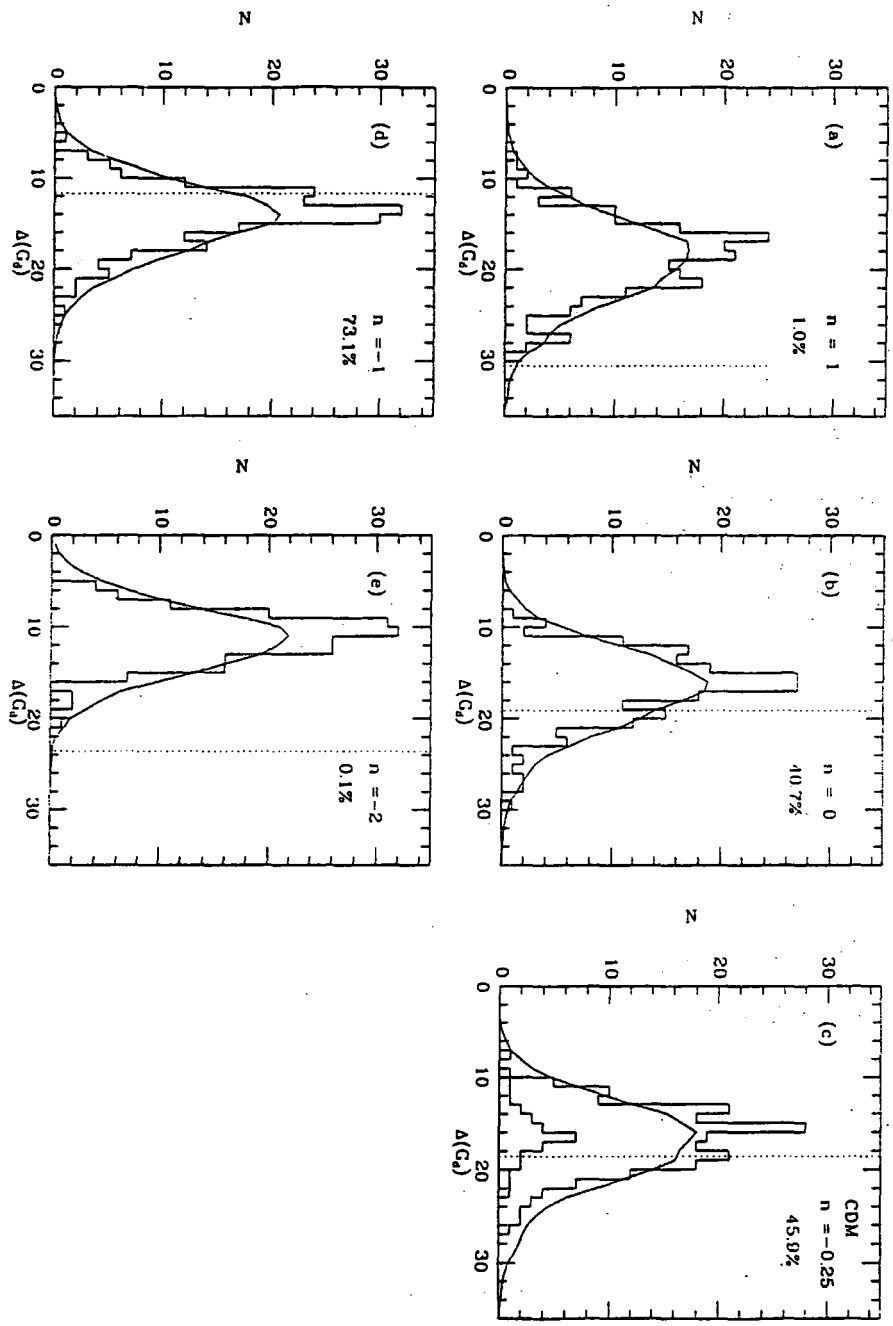


Figure 4.14 The distribution of ΔG for various fluctuation spectra. The solid curves in each diagram show the result of convolving the histogram of ΔG with the histogram of sampling errors ΔG_{linG} (shown as the shaded histogram in (c)).

Table 4.2 summarises the significance levels at which the various models can be rejected for three different scales $\lambda = 10, 20$ and $40 h^{-1}$ Mpc. The results from these Monte-Carlo simulations are roughly consistent with our previous, more naive estimates.

Model	$\lambda = 10h^{-1}$Mpc $r_{max} = 40h^{-1}$Mpc	$\lambda = 20h^{-1}$Mpc $r_{max} = 100h^{-1}$Mpc	$\lambda = 40h^{-1}$Mpc $r_{max} = 175h^{-1}$Mpc
CDM	73.6% ($\Delta G = 7.7$)	45.9% ($\Delta G = 18.6$)	56.2% ($\Delta G = 15.3$)
n = +1	8.9% ($\Delta G = 16.4$)	1.0% ($\Delta G = 30.5$)	38.8% ($\Delta G = 18.9$)
n = 0	62.8% ($\Delta G = 9.8$)	40.7% ($\Delta G = 19.1$)	97.5% ($\Delta G = 7.9$)
n = -1	85.4% ($\Delta G = 5.3$)	73.1% ($\Delta G = 11.7$)	93.8% ($\Delta G = 7.6$)
n = -2	75.8% ($\Delta G = 5.4$)	0.1% ($\Delta G = 23.6$)	13.7% ($\Delta G = 14.2$)

Table 4.2. Confidence levels and ΔG for CDM and power spectra models.

4.6 DISCUSSION AND CONCLUSIONS

The topological analysis discussed here complements previous studies of the galaxy distribution using the QDOT survey. In particular, from a counts-in-cells analysis, Efstathiou *et al.* (1990) and Saunders *et al.* (1991) have determined the variance $\Delta^2(r)$, of the density field smoothed on scale λ,

$$\Delta^2(r) = b^{-2} \int_0^\infty P(k) W^2(kr) d^3k \tag{4.17}$$

where $W^2(kr)$ is a window function and b is the biasing parameter. While the topology is sensitive to the shape of the power spectrum, the variance $\Delta^2(r)$ depends both on its shape and amplitude.

Figure 4.15 gives the variance as a function of smoothing scale as estimated by Efstathiou *et al.* (squares) and Saunders *et al.* (circles). In this plot we have taken into account the different smoothing functions used in these studies, cubical cells in the first and Gaussian spheres in the second. The relation between these depends slightly on the spectral index. We have used the conversion appropriate to a power law spectrum with $n = -1$. For a Gaussian field, the variance in Gaussian spheres of radius $r\mathrm{h}^{-1}$ Mpc is the same as that in top hat spheres of radius $1.8r\mathrm{h}^{-1}$ Mpc or cubes of side $2.9r\mathrm{h}^{-1}$ Mpc. (Note that our definition of a Gaussian window in equation (4.1) differs from the definition of Saunders *et al.* by a factor of $1/\sqrt{2}$.) The variances predicted in the CDM model calculated from the same N-body simulations used in this chapter are shown as open symbols in Figure 4.15. The solid line gives the theoretical variance calculated directly from the CDM spectrum with a biasing parameter $b = 1.4$, which is close to the value inferred from the dynamical studies of the QDOT survey (Kaiser *et al.* 1991).

The variance in top hat spheres of radius $8\mathrm{h}^{-1}$ Mpc is close to unity for optical glaxies (Davis and Peebles 1983). In our units this scale corresponds to $6.3\mathrm{h}^{-1}$ Mpc ; on this scale Figure 4.15 shows the variance of the IRAS galaxies is

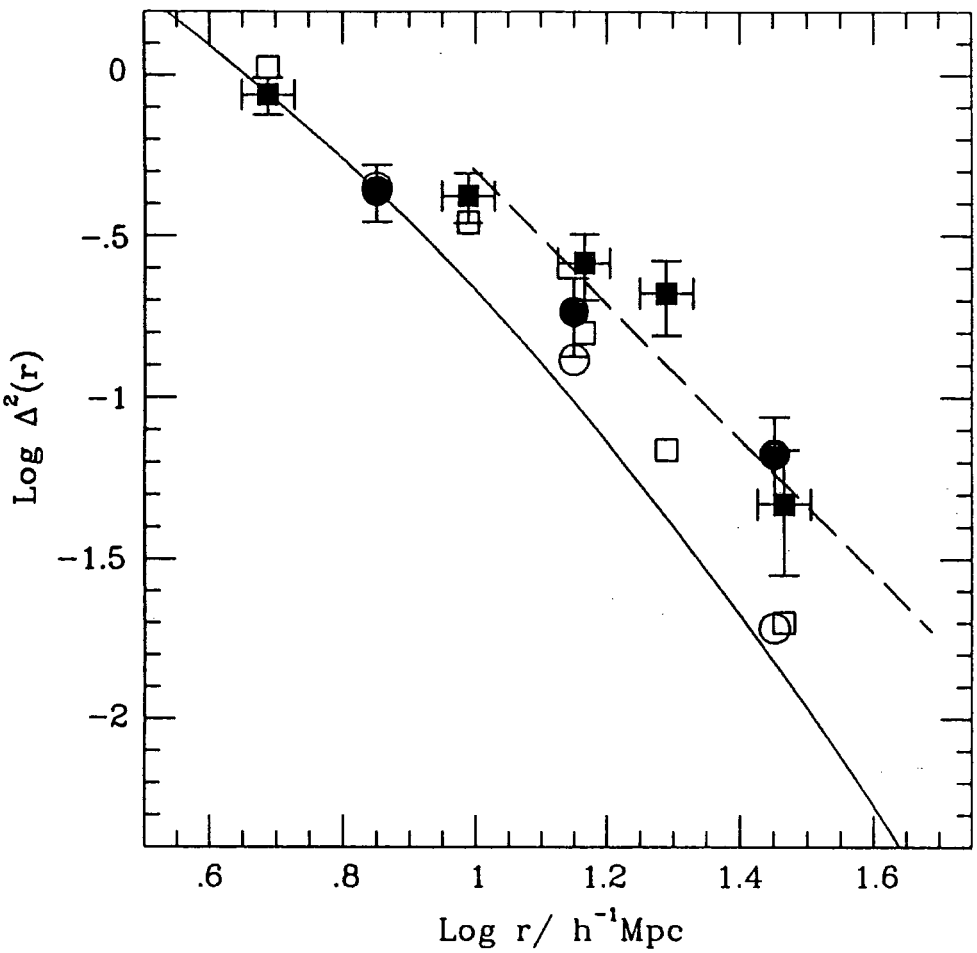


Figure 4.15 The variance of counts-in-cells in the QDOT survey measured by Efstathiou *et al.* (1990; solid squares) and Saunders *et al.* (1991; solid circles). The open squares and circles are their respective measurements of the variance in the N-body simulations of a CDM Universe used in this chapter. The vertical error bars are 1σ. The horizontal error bars on the squares stem from the uncertainty in converting between different smoothing functions. The variance calculated for the CDM spectrum using equation (4.17) and a biasing parameter $b = 1.4$ is represented by the solid curve. The approximate power law spectrum, $n = -0.9$, of the QDOT data found from the topology analysis is shown as a dashed line. This line was placed at a height that gives the best fit (by eye) to the results of the counts-in-cells analysis.

well below unity. A least squares fit to the first three data points in this Figure gives $\Delta(r) = 0.74 \pm 0.12$ on this scale. This is consistent with our correlation analysis in Section 2 which showed the IRAS galaxies to be somewhat less clustered than optically selected ones. The dashed line in Figure 4.15 has the slope $n = -0.9$ obtained in Section 5, over scales of 10 to $50h^{-1}$ Mpc. We have arbitrarily placed the line at the height at which it gives the best possible consistency (as judged by eye) with the results of the counts-in-cells analysis. The results from the different analysis are roughly consistent with one another although some significant discrepancies are evident.

To summarise, the QDOT redshift survey of IRAS galaxies has proved to be a useful data set with which to study the topology of the galaxy distribution on large scales. Although IRAS galaxies are less strongly clustered than optical galaxies, the QDOT survey contains a wealth of information on large scale structure. The survey provides useful information at least out to $200h^{-1}$ Mpc and many well known clusters, superclusters and voids can be identified within this distance. Within $75h^{-1}$ Mpc ,the most prominent feature is the "Great Attractor" complex. However, the Hercules supercluster extending between $100h^{-1}$ Mpc and $150h^{-1}$ Mpc dominates the far field. Smoothed maps of the galaxy distribution exhibit the topology characteristic of Gaussian density fields. High and low density regions show remarkable similarity in structure. Voids are not isolated empty spaces but rather regions of low density which connect together in a sponge like fashion. Clusters are joined together via walls and bridges of galaxies to form massive overdense regions.

We have used the genus-threshold density relation developed by Gott *et al.* (1986) and Hamilton *et al.* (1986) to investigate the topology of this survey on a range of length scales. Between $10h^{-1}$ Mpc and $50h^{-1}$ Mpc the genus curve has the shape expected for a Gaussian random field. The amplitude of the genus curve measures the slope of the initial fluctuation spectrum on the smoothing

scale. We used artificial "QDOT" surveys constructed from N-body simulations to calculate the errors introduced by sparsely sampling the galaxy distribution and realisation to realisation differences. We found that the power spectrum of the galaxy distribution in the QDOT survey is consistent with a power law of index $n = -0.9 \pm 0.4$. Pure power law fluctuation spectra with $0.2 < n < -1.8$ can be ruled out at over 3σ. In the standard cold dark matter model the shape of the power spectrum of the mass distribution is assumed to be the same as that of the galaxy distribution. Its effective index is $n \approx -1$ on scale $\sim 10h^{-1}$ Mpc, but the spectrum steepens gradually on larger scales. The largest discrepancy between the QDOT results and the standard model occurs on a scale of $\sim 30h^{-1}$ Mpc and is significant at the 2.4σ level.

References

Abell, G. 1958, *Ap.J. Suppl.*, **3**, 211.

Adler, R.J. 1981, *The geometry of Gaussian random fields* (Chichester: Wiley)

Babul, A., and Postman, M., 1989, *Ap.J* **359**, 280.

Bardeen, J.M., Steinhardt, P., and Turner, M. 1983, *Phys.Rev.D,* **28**, 679.

Bardeen, J.M., Bond, J.R., Kaiser, N., and Szalay, A.S., 1986, *Ap.J.*, **304**, 15.

Burstein, D., Faber, S., and Dressler, A. 1990. *Ap.J.*, **354** 18.

da Costa, L. *et al.* 1988, *Ap.J.*, **327**, 544.

Davis, M., Huchra, J., Latham, D.W. and Tonry, J. 1982, *Ap.J.*, **253**, 423.

Davis, M. and Peebles, P.J.E., 1983, *Ap.J.*, **267**, 465.

Davis, M., Efstathiou, G., Frenk, C.S. and White, S.D.M. 1985, *Ap.J.*, **292**, 465.

Davis, M., Meiksin, A., Strauss, M.A., da Costa, L. and Yahil, A., 1988, *Ap.J. (letters)*, **333**, L9.

Doroshkevich, A.G. 1970, *Astrophysics*, **6**, 320 (translated from *Astrofizika,* **6**, 581).

Efstathiou, G., Kaiser, N., Saunders, W., Lawrence, A., Rowan-Robinson, M., Ellis, R.S.E and Frenk, C.S. 1990 *Mon.Not.R.astr.Soc.* **247**, 10p.

Frenk, C.S., White, S.D.M., Efstathiou, G. and Davis, M. 1990, *Ap.J.*, **391**, 2.

Frenk, C.S. 1991, *Physica Scripta, in press.*

Geller, M.J. and Huchra, J.P. 1983, *Ap.J.Suppl.*, **52**, 61.

Geller, M.J. and Huchra, J.P. 1989, *Science*, **246**, 897.

Giovanelli, R. and Haynes, M.P. 1985, *A.J.*, **90**, 2445.

Gott, J.R., Mellot, A.L. and Dickinson, M. (GMD) 1986, *Ap.J.*, **306**, 341.

Gott, J.R, Millar, J., Thuan, T.X., Schneider, E., Weinberg, D.H, Gammie, C., Polk, K., Vogely, M., Jeffrey, S., Bhavsar, S.P., Mellot, A.L., Giovanelli, R., Tully, R.B. and Hamilton, A.J.S 1989, *Ap.J.*, **340**, 625.

Hamilton, A.J.S., Gott, J.R., and Weinberg, D. 1986, *Ap.J.*, **309**, 1.

Huchra, J.P. and Geller, M.J. 1982, *Ap.J.*, **257**, 423.

Kaiser, N. 1986. *Mon.Not.R.astr.Soc.* **219**, 795.

Kaiser, N., Efstathiou, G., Ellis, R.S., Frenk, C.S., Lawrence, A., Rowan-Robinson, M. and Saunders, W. 1991, *in preparation.*

Kirshner, R.P., Oemler, A., Schechter, P.L., and Schectman, S.A., 1981, *Ap.J. (letters)*, **248**, L57.

Joseph, R.D. and Wright, G.S. 1985, *Mon.Not.R.astr.Soc.* **214**, 87.

Lawrence *et al.* , 1991, in preparation

Leech, K.J., Penston, R.J., Terlevich, R.J., Lawrence, A. Rowan-Robinson, M. and Crawford, J. 1989, *Mon.Not.R.astr.Soc*, **240**, 349.

Lucey, J.R. 1983, *Mon.Not.R.astr.Soc*, **204**, 33.

Olive, K.A., Schramm, D.N, Steigman, G. and Walker, T.P., 1990 *Phys.Lett B.*, **236**, 434.

Peebles, P.J.E. 1980 *The Large Scale Structure of the Universe*

Rowan-Robinson, M. and Crawford, J., 1989, *Mon.Not.R.astr.Soc.* **238**, 523.

Rowan-Robinson, M., Lawrence, A., Saunders, W., Crawford, J., Ellis, R., Frenk, C.S., Parry, I., Xiaoyang, X., Allington-Smith, J., Efstathiou, G. and Kaiser, N., 1990 *Submitted to Mon.Not.R.astr.Soc.*

Saunders, W., Rowan-Robinson, M., Lawrence, A., Efstathiou, G., Kaiser, N., Ellis, R.S.E. and Frenk, C.S.F., 1990, *Mon.Not.R.astr.Soc.* **242**, 318.

Saunders, W., Frenk, C.S.F, Rowan-Robinson, M., Efstathiou, G., Lawrence, A., Kaiser, N., Ellis, R.S.E, Crawford, J., Xiao-Yang Xia and Ian Parry 1991, *Nature* **349** 32.

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

Shapley, H. and Ames, A., 1932. *A survey of the external galaxies Brighter than the Thirteenth Magnitude, Annals of the Harvard College Obs*, **88** no.2.

Strauss, M.A., Davis, M., Yahil, A. and Huchra, P.J. 1990 *Ap.J.* **361**, 49.

Thuan, T.X and Schneider, S., 1988, in preparation.

Tully, R.B. *nearby galaxy catalogue* (Cambridge University Press).

Tully, R.B., 1987, *Ap.J.* **323**, 1.

Weinberg, D.H., Gott, J.R. and Mellot, A.L. (WGM) 1987 *Ap.J.* **321**, 2.

Weinberg, D.H. 1988, *Pub.astr.Soc.Pacific*, **100**, 1373.

White, S.D.M. 1988, in *Minnesota Lectures on Clusters of Galaxies and Large-scale Structure,* ed J.M. Dickey, Astr. Soc. Pac. Conf. Series **5**, 197.

5 CONCLUSIONS

In this chapter we shall briefly review some of the main results of this thesis and discuss the consequences and prospects for theories and models for the evolution of structure in the Universe.

The location of the Milky Way within a group of ~ 2 galaxies is not special. We can learn more about our Universe by studying the commonplace objects rather than the rarer exotic objects which many sometimes focus on. The Local Group provides us with the ideal laboratory for studying the dynamics of a bound system of galaxies. Peebles (1990) used the existence of the Local Group and its infall towards the Virgo cluster to argue against its formation within a large pancake. The small relative velocities also suggest that the Local Group is collapsing for the first time and that the Local Group is in fact, much younger than our Galaxy.

We used a model for the Local Group to study the formation of a typical galaxy system and to further our understanding of the mass distribution within the Local Group. The model was excised from a numerical simulation of a Universe dominated by CDM and is a binary system which has a similar radial velocity and separation as M31 and the Milky Way. We find that the timing argument provides a reliable method for placing a lower limit to the mass of the Local Group. If we assume that the Local Group is bound then we can use the mass estimates from the timing argument to calculate its mass to light ratio. Taking the luminosity of the Milky Way in the V band as $1.4\times10^{10}L_{\odot}$ and M31 to be twice as luminous, the

corresponding mass to light ratio of the Local Group lies in the range $75 - 135\Gamma_{\odot}$ for an age of our Universe between 10 and 20 Gyr. If the transverse velocity of M31 is of the order of its radial velocity then the mass to light ratio could be as high as $215\Gamma_{\odot}$. These values are consistent with the median mass to light ratio of 200 which we find for our sample of groups of galaxies identified within the CfA survey. However, if the Local Group is indeed a quartet of galaxies and not a simple binary (McCall 1989), then the dynamics becomes much more complicated to analyse.

There are some inconsistencies between the model and the observations. The probability of observing a radial velocity in the model as large as Leo I was found to be very small. At the distance of Leo I the particles in the model halo tended to be infalling whereas Leo I is moving away from the Galaxy with a radial velocity of 177 km s^{-1}. A second incongruity is the orbits of the satellites of the Milky Way. The anisotropy parameter of the particle orbits within the CDM halos were found to be radially biased. The radial velocity dispersion in the model was found to be larger than the observed radial velocity dispersion of the satellites. To reconcile the mass of the Milky Way with the predictions from the timing argument and the mass of the model halo, the satellites of our Galaxy must be on circularly biased orbits. Simple analytic calculations are unclear as to whether tidal disruption would be sufficient to leave a population of satellite galaxies on circular orbits. We are currently performing numerical simulations to check these calculations.

The asymptotic values of the rotation curves of the two halos in the model match very closely those of M31 and the Milky Way. A simple treatment of gas within our model extended the flat rotation curves into the central regions of the halos. Galaxy rotation curves are one means by which the dark matter distribution within clusters of galaxies can be probed. The key question here is whether galaxies retain individual dark halos within clusters or whether they are orbiting within a common extended halo. As the halos within our model were

dragged closer together the rotation curves became very distorted. In our model, M31 and the Milky Way collide 2.5 Gyrs from the present time, a fraction of the current crossing time. The inclusion of gas within our simulations did not affect the rate at which these halos merged, although as a result of dissipation, the "galaxies" in their centres survived slightly longer.

We have studied the continuity of properties of galaxy associations from galactic to cluster scales in an attempt to understand the wide range of galaxy associations which we observe in the Universe. Intermediate scales were probed using a statistical analysis of groups of galaxies identified in the CfA redshift survey using well defined selection criteria. The grouping algorithm was optimised using artificial galaxy catalogues constructed from N-body simulations which have similar low order correlations to the original survey. With the deeper CfA2 catalogue due to be released soon, these techniques are likely to prove valuable in studying the properties of galaxies as a function of their environment.

We used artificial group catalogues to optimise our grouping algorithm. Comparing groups selected in real space with those selected in redshift space shows that the best which can be achieved is $\sim 75\%$ agreement between galaxies grouped in common. Projecting the galaxy distribution from an artificial catalogue and performing the same test, we could only obtain $\sim 40\%$ of grouped galaxies in common between groups identified in real space and groups identified as surface density enhancements in projection. Even with redshift information there is residual contamination which undoubtedly leads to some overly high group velocity dispersions and correspondingly overestimated mass to light ratios.

We developed a method of estimating the total luminosity of groups of galaxies identified within magnitude limited redshift surveys. Our technique was shown to give results which are independent of the magnitude limit and depth of the catalogue. (Note the importance of correcting the total luminosity of the groups

when calculating the mean cosmological density. The correction has the effect of lowering the median mass to light ratio by a factor ~ 0.75 from 270 to 200, which corresponds to a mean cosmological density of $\Omega \sim 0.12$.) We then constructed the luminosity function of galactic systems which measures the abundance of gravitationally bound structures, from single galaxies to rich clusters, independently of the detailed arrangement of the luminous material within them. The smooth transition between the abundances of galaxies, groups and clusters, suggests that these associations correspond to different levels in a clustering hierarchy.

It is interesting to compare our estimated abundance of "galactic systems" with Bahcall's (1979) attempt to calculate this distribution using groups and clusters of galaxies selected from 2-dimensional catalogues. Her results were quite different to those of the present work, reflecting the importance of calculating total group luminosities and the problems of using galaxy associations identified in projection.

We found no evidence for a dependence of mass to light ratio on either group mass or luminosity. This allows us to compare our distribution with theoretical predictions of the mass multiplicity function. The abundance of systems falls off rapidly at about $\sim 6L_*$. In the context of the Press-Schechter (PS) theory this scale is to be interpreted as the scale on which the mass fluctuations are currently collapsing out of the clustering hierarchy and turning non-linear. Given a normalised power spectrum of density fluctuations, the PS theory predicts the abundance of bound objects at any epoch. Using arbitrarily normalised power law spectra we fitted the PS predictions to our luminosity function of galaxy groups. We found that for reasonable values of the spectral index, $-2 < n < 1$, the PS theory successfully predicts the luminous end of the distribution, but greatly overestimates the abundance of low mass objects.

We found that the distribution of group velocity dispersions shows an apparent discontinuity at the transition between groups and rich clusters. Even by increasing V_o to unrealistically large values, we could not match the distribution of group velocity dispersions with those of rich clusters. The grouping algorithm was optimised using the velocity dispersions of groups found in the CDM artificial catalogues. We therefore have the ideal parameters for a comparison between CDM and the CfA survey, but if groups in the real Universe are much richer than their CDM counterparts, then our procedure would artificially depress the estimated velocity dispersions. However, the median velocity dispersion of the CfA groups is significantly lower than the CDM groups, which therefore indicates that the CfA groups have intrinsically low velocity dispersions. There are several reasons why we should be cautious about the interpretation of the discontinuity between the velocity dispersions of groups and rich clusters. Abell clusters of richness class R=0 were not included in our comparison because the sample would undoubtedly be incomplete. This could lead to an underestimation of low velocity dispersion clusters in our sample. Frenk *et al.* (1990) found that projection effects in clusters selected in 2 dimensions lead to an overestimate of the abundance of high velocity dispersion clusters. The break occurs at the conjunction of our catalogue of groups from the CfA survey and the clusters identified in the Abell catalogue. In Chapter 1 we demonstrated that the characteristic timescales of groups and clusters are very similar, so it would be surprising if a physical process could cause such a discontinuity.

The luminosity function of groups of galaxies is not very sensitive to a few contaminators. Neither is the correlation function of groups which depends mainly on the group positions. The correlation function of groups was found to depend upon the sample chosen. The most luminous groups are more strongly clustered than galaxies, groups with 3 members are only as strongly clustered as the galaxies, very faint groups are less strongly clustered than the galaxies. We find similar clustering properties for groups of galaxies in the CDM model. These results lend

support to Kaiser's (1984) idea that more luminous (and hence more massive) objects are more strongly clustered than fainter ones due to the Gaussian nature of primordial density fluctuations. These results are marginally significant and analyses of larger and deeper catalogues are needed to confirm these findings. As discussed in the conclusions of Chapter 3, it is surprising that there could be a large jump in clustering strength between our sample of groups and that of rich clusters, especially since our catalogue contains several rich clusters. A catalogue of clusters identified by ROSAT will provide a "projection free" correlation function, hopefully resolving this discrepancy.

On large scales we have used an all sky redshift survey of galaxies detected by IRAS to investigate the topology of the Universe to a depth of $200h^{-1}$ Mpc. This survey provides a remarkable picture of structure within the Universe. At the mean density contrast the high and low density regions merge into one another, structure is interlocking and sponge-like and isolated clusters or voids are not found. At high overdensities we observe only the prominent Hercules supercluster. A quantitative analysis of the data yielded two interesting results. The distribution of galaxies on smoothing scales between 1500km s^{-1} and 5000km s^{-1} is in good agreement with the hypothesis that structure formed from small perturbations within a Gaussian density field. The simplest inflationary model predicts that random phase fluctuations arise from quantum noise in the early Universe. Topological analyses of shallower redshift catalogues by Gott *et al.* (1989), showed that these data were also consistent with Gaussian fluctuations on smaller scales than we have studied. This is a pleasing result because it allows us to completely specify the linear density field in terms of its power spectrum. On large scales we found that the slope of the power spectrum can be represented by a power law of index $n \approx -1$. The amplitude of the power spectrum can be found from counts-in-cells analysis or the correlation function. (So could the slope of the power spectrum, but the weighting given to the galaxy distribution is different in all these methods. The topological approach gives all the galaxies a similar

weighting. Each method has its advantages and disadvantages, the power spectrum should be inferred from a study of all these statistics.) So we now know all there is to know about the galaxy distribution, not convinced?... you shouldn't be! The implicit assumption in this discussion is the inference of a mass distribution from a sample of infra-red selected galaxies. This is a big step and we must be careful not to over-simplify the Universe. Galaxy redshift catalogues are unlikely to probe deeper than the QDOT survey within the next decade, but the topology of large scale structure can be studied using clusters of galaxies such as might be selected by ROSAT or large two dimensional catalogues such as the APM survey.

The CDM N-body simulations of DEFW have served a dual purpose throughout this thesis. In Chapter 2 we used a model for the Local Group extracted from a numerical simulation of a CDM universe which helped us further our understanding of the mass distribution in the real Local Group. In Chapter 3 we used artificial redshift catalogues generated from CDM N-body simulations to help us evaluate the success of our grouping algorithm and our technique for correcting the luminosities of a magnitude limited sample of groups. In Chapter 4 we constructed fully sampled and one-in-six sampled redshift surveys similar to the QDOT survey to investigate the shot noise of the sparse sampling technique. Many of these applications were relatively independent of the assumed cosmological model. Our main concern was that the galaxy distribution in the model had similar correlations to the observed galaxy distribution.

Throughout this thesis we have compared the observational results with the predictions of the cold dark matter model. The agreement between the model and the observations of groups of galaxies has complemented previous work on galactic and cluster scales. This is a great achievement for such a fully specified model. In Chapter 4 we studied the topology of the galaxy distribution on scales larger than previous studies. We found that the standard CDM model fails to match the coherence of structure observed in the QDOT redshift survey. The significance of

this result and the counts-in-cells analysis of the same data is difficult to quantify although they appear to "rule out" the model at $\sim 3\sigma$. The main uncertainty in the standard CDM model is the technique for assigning the "light" to the "mass" which will undoubtedly become the subject of careful scrutiny. The fate of the cold dark matter model will probably be decided in the next few years, not by observations of large scale structure, but by either the detection or non-detection of fluctuations in the microwave background or the actual particles themselves. The nature, origin and distribution of the dark matter and its role in galaxy formation and dynamics are issues the resolution of which is likely to determine the direction of studies in cosmology for some years to come.

References

Bahcall, N.A. 1979, *Ap.J.*, **232**, 689.

Frenk, C.S., White, S.D.M., Efstathiou, G. and Davis, M. 1990, *Ap.J.*, **391**, 2.

Gott, J.R, Millar, J., Thuan, T.X., Schneider, E., Weinberg, D.H, Gammie, C., Polk, K., Vogely, M., Jeffrey, S., Bhavsar, S.P., Mellot, A.L., Giovanelli, R., Tully, R.B. and Hamilton, A.J.S 1989, *Ap.J.*, **340**, 625.

Kaiser,N. 1984, *Ap.J.*, **284**, L9.

McCall, M.L. 1989, *A.J.*, **97**, 1341.

Peebles, J.E. 1990, *Clusters of Galaxies. Baltimore STSI.*