Chapter 1

Introduction

Imminent large-scale galaxy redshift surveys such as the Sloan Digital Sky Survey (SDSS; Gunn & Weinberg 1995) and the Two-degree Field (2dF; Colless 1998) will probe the galaxy density field of the universe with unprecedented precision. Were galaxies accurate tracers of the mass density field, the results of these surveys would put severe constraints on cosmological models. If the Cold Dark Matter (CDM) picture for the linear theory power spectrum is correct, these surveys can in fact measure cosmological parameters, such as the mass density $\Omega_0$, the vacuum density $\Lambda$, and the baryon density $\Omega_b$ (de Laix & Starkman 1998; Goldberg & Strauss 1998; Eisenstein, Hu, & Tegmark 1999; Wang, Spergel, & Strauss 1999). However, the visible matter in galaxies is only a small percentage of the baryons in the universe, which in turn is a small percentage of the mass in the universe (Fukugita, Hogan, & Peebles 1998; Cen & Ostriker 1999a). Moreover, the process of galaxy formation is complex and nonlinear, including a complicated interplay between gravitational fields, hydrodynamics, microphysics and star formation. Does this complicated process produce a population of galaxies whose number density field traces the mass density field perfectly? If not, does the clustering of galaxies in the universe then contain useful information on the processes that formed them? In the first half of this introductory section, I argue that observations already suggest that the relationship between the density fields of galaxies and mass is biased, scale dependent, and nonlinear, as well as dependent on morphological type. The second half of this introduction describes the approaches to answering the theoretical question which is the basis of this thesis: how is the galaxy density field related to that of the mass?

1. Observational Evidence for Complicated Bias

The argument given here for a complicated relationship between galaxies and mass is twofold. First, because different morphological types of galaxies have different density fields
Chapter 1: Introduction

(Hubble 1936; Oemler 1974), it is probable that none of these galaxy density fields trace the mass density field exactly. Second, it is clear that galaxies observed at high redshift are highly biased compared to the mass distribution. This indicates that the physics of galaxy formation affects the relative clustering of galaxies and mass in ways which may remain observable even for galaxies at \( z = 0 \).

Consider the observed differences between the clustering strengths of galaxies of different types. Various authors have compared elliptical and spiral galaxies, generally finding that the fluctuation amplitude of ellipticals is stronger than that of spirals by a factor of 1.3–1.5 (Davis & Geller 1976; Giovanelli, Haynes, & Chincarini 1986; Santiago & Strauss 1992; Loveday et al. 1996; Hermit et al. 1996; Guzzo et al. 1997). Similarly, a comparison of the galaxy distribution in the IRAS redshift survey (Strauss et al. 1992b) with those in the Center for Astrophysics Redshift Survey (CfA; Huchra et al. 1983) shows that optically-selected galaxies are clustered more strongly than infrared-selected galaxies by a similar factor (Davis et al. 1988; Babul & Postman 1990; Strauss et al. 1992a). A more recent comparison by Seaborne et al. (1999) of the IRAS Point Source Catalog redshift survey (PSCz; Saunders et al. 1997) to the Stromlo-APM survey (Loveday et al. 1992) finds similar results. These differences in the amplitude of clustering may be accounted for by invoking a deterministic linear bias prescription:

\[
\delta_g(r) = b \delta(r),
\]

where \( \delta_g(r) \equiv \rho_g(r)/\langle \rho_g \rangle - 1 \) is the galaxy overdensity and \( \delta(r) \equiv \rho(r)/\langle \rho \rangle - 1 \) is the mass overdensity, smoothed on some scale. To explain the observations, one must assume that different galaxy populations are “biased” by different factors \( b \); clearly all but one of these bias factors must differ from unity.

A relationship between galaxies and mass such as Equation (1-1) would be independent of scale, because of its linearity. On the other hand, there are differences between the shapes of correlation functions of galaxies of different types, at least at small scales. For instance, the ratio of the correlation functions of ellipticals and spirals found in the Optical Redshift Survey (Hermit et al. 1996) and in the Perseus-Pisces Redshift Survey (Guzzo et al. 1997) declines with scale over the range between 1 and 10 \( h^{-1} \) Mpc. This scale dependence cannot result from a deterministic linear bias. Therefore, there must exist a more complicated relation between the density fields of different morphological types, which involves either nonlinearity or stochasticity at some level.

An alternate measure of these differences is the density-morphology relation in clusters of galaxies, quantified by Dressler (1980) and Postman & Geller (1984). Similarly, Whitmore, Gilmore, & Jones (1993) demonstrate a radius-morphology relation in clusters. In the field, spirals comprise about 70% of all galaxies, and ellipticals and lenticulars comprise the rest; in the cores of rich clusters the situation is reversed, and ellipticals and
lenticulars account for 90% of all galaxies. The relationship between spiral density and the density of all galaxies is extremely nonlinear in the densest cluster regions.

Finally, the possibility exists that the density fields of different galaxy types, even on large scales, are not related in a deterministic way. That is, given a region with some overdensity of early-type galaxies $\delta_e$, it may not correspond uniquely to an overdensity of late-type galaxies $\delta_l$, even when the intrinsic noise in the measurement (such as Poisson noise) is taken into account. Tegmark & Bromley (1999) claim to measure this effect in the Las Campanas Redshift Survey (LCRS). They measure the correlation coefficient $r$ between counts-in-cells of different galaxy types (as defined by Bromley et al. 1998), finding that $r \sim 0.4$ between the most early-type galaxies and the latest-type galaxies. On the other hand, Seaborn et al. (1999) find that $r > 0.72$ at 95% confidence in their comparison of the PSCz survey to the Stromlo-APM survey. In an attempt to explore this issue further, I devoted Chapter 4 of this thesis to a more careful analysis of the LCRS to test the result of Tegmark & Bromley (1999); the analysis in fact casts doubt upon the existence of stochasticity in the relationship between galaxy types.

Whether or not stochasticity exists in the relationships between galaxy types, the differences among different morphological types which do exist suggest that the relationship between the density field of galaxies and the density field of mass is comparably complicated, possibly being nonlinear and scale dependent. If it does exist, stochasticity would indicate that there are variables other than the local mass density which determine the local galaxy density. It would be a coincidence if the overdensity field of the union of all types of galaxies exactly traced the full mass overdensity, despite the fact that the different morphologies have clearly formed in quite different ways, as evidenced by their different clustering properties. In any case, the selection effects of redshift surveys (color, surface brightness, luminosity, etc.) will cause any catalog to contain a mix of morphological types that differs from the mix in a purely volume-limited sample. Since the overdensity fields of different morphologies have different density fields, the results of every survey are “biased” to some degree.

Not only can one consider galaxies of different types, but one can consider galaxies at different stages of the history of the universe. Measurements of the clustering of Lyman-break objects (LBOs) at $z \sim 3$ (Steidel et al. 1998) indicate that these objects are highly biased with respect to the mass. In particular, measurements of the amplitude of counts-in-cells fluctuations (Adelberger et al. 1998) or the angular autocorrelation function (Giavalisco et al. 1998) suggest that galaxies were as strongly clustered in comoving coordinates at $z \sim 3$ as they are today. If the LBOs are unbiased tracers of the mass density field, these results contradict the widely accepted gravitational instability (GI) model for the formation of large-scale structure, unless one assumes an unacceptably low
value of $\Omega_0$ in order to prevent the growth of the clustering of galaxies between \( z = 3 \) and today. Therefore, the objects observed at high redshift are probably more highly "biased" tracers of the mass density field than are galaxies today. That is, they have a high value of \( b_g \equiv \sigma_g / \sigma \), where \( \sigma_g \equiv \langle \delta^2 \rangle^{1/2} \) is the rms galaxy density fluctuation and \( \sigma \equiv \langle \delta^2 \rangle^{1/2} \) is the rms mass fluctuation. The counts-in-cells analysis of Adelberger \textit{et al.} (1998) suggests that at \( z = 3 \) the bias of LBOs is \( b_g \sim 2b_0 \) (for \( \Omega_0 = 0.2, \Omega_\Lambda = 0 \)), \( b_g \sim 4b_0 \) (for \( \Omega_0 = 0.3, \Omega_\Lambda = 0.7 \)), or \( b_g \sim 6b_0 \) (for \( \Omega_0 = 1 \)), where \( b_0 \) is the bias of galaxies today. As Adelberger \textit{et al.} (1998) shows, this result is in rough agreement with a theoretical picture in which galaxies form at peaks of the density field. At least in this case, then, the process of galaxy formation is intimately connected with the clustering properties of the galaxies.

It is thus interesting to explore theoretically how the mass density in the universe might be related to the process of galaxy formation, how this relationship evolves and how it determines the density fields of galaxies of different types at \( z = 0 \).

2. \textbf{Theoretical Approaches to Bias}

Theoretical modeling of the relationship between galaxies and mass invariably boils down to modeling the formation of galaxies, whether the modelers explicitly state so or not. The first attempt at investigating the effects of a model of galaxy formation on large-scale structure was that of Davis \textit{et al.} (1985), who assumed that galaxies formed at peaks of the initial density field. Since then, the models have become somewhat more sophisticated. Workers in the field have taken, in a basic sense, three ways of approaching the problem: the first is a phenomenological approach which specifies a relationship between galaxies and mass and investigates the consequences of such an assumption; the second is to model the history of dark matter halos in the universe and apply simple models of star-formation to determine the nature of the galaxies forming in these halos; the third is to attempt a full hydrodynamic simulation of galaxy formation in a cosmological context, again using simple models of star-formation to track the cooling and collapse of gas below the code resolution.

The earliest and most common approach has been the phenomenological one, which consists of modelling the galaxy density as a function of the dark matter density (or other variables). Either the initial density field (Lagrangian bias) or the \( z = 0 \) density field (Eulerian bias) can be used; there are a number of papers describing the transformation of a Lagrangian bias scheme into the observed Eulerian bias, which in the linear regime is determined essentially by the continuity equation (Mo \& White 1996; Catelan \textit{et al.} 1998a,b; Sheth 1998). Peaks-biasing (Davis \textit{et al.} 1985; Bardeen \textit{et al.} 1986), which postulates that galaxies form at peaks of the initial density field, was the first suggestion for a Lagrangian biasing scheme; in fact, Davis \textit{et al.} (1985) used the peaks-biasing model.
to reconcile an $\Omega_0 = 1$ universe with observations. Most of the recent work involving
Lagrangian bias has been based on the extensions of Bond et al. (1991) of the Press &
Schechter (1974) formalism. An important issue on small scales (< 1 Mpc) is that the
contents of halos which are close together in Eulerian space in fact overlap in Lagrangian
space; Sheth & Lemson (1999) address this issue of halo “exclusion.”

The simplest versions of Eulerian bias are the deterministic linear bias of Equation
(1-1) and simple threshold biasing (Kaiser 1984), which places galaxies in regions which
are overdense in mass by some factor. Fry & Gaztaña (1993), Coles (1993), Mann,
Peacock, & Heavens (1998), and Narayanan, Berlind, & Weinberg (1998) have studied
how nonlinearity and other complications in an Eulerian bias scheme can affect large-scale
structure statistics. Scherrer & Weinberg (1998) show how adding scatter to an Eulerian
bias relation affects the correlation function and the power spectrum of galaxies, while
Pen (1998) discusses its effect on redshift space distortions. Dekel & Lahav (1999) have
developed a general formalism (which I adopt in this thesis) for expressing such a nonlinear
and stochastic relation. The goal of all of this work is to find a prescription which
makes consistent different determinations of statistics such as the density distribution
function, the correlation function, $M/L$ ratios of clusters, measured velocity statistics, and
redshift-space distortions. A novel technique along these lines introduced by Narayanan &
Weinberg (1999) is to use initial-condition reconstruction techniques. One can construct
many different initial conditions which produce the observed local universe by assuming
different cosmological models and prescriptions for bias. A goodness-of-fit test is provided
by running the simulations forward again and testing how well the simulated universe
reproduces the original observations. One can then test the relative quality of different
assumptions about the bias. These phenomenological models are useful for analyzing the
data and testing how different biasing schemes can be affecting observations, but they lack
a physical basis, which is provided more fully by the other two approaches to the problem
of bias discussed below.

These investigations have provided interesting results nonetheless. For instance, it was
Davis et al. (1985) who first noted that the bias of galaxies decreases substantially with
redshift, a prediction which may have been recently confirmed by Adelberger et al. (1998),
as discussed in the previous section. An important result of Scherrer & Weinberg (1998),
Narayanan, Berlind, & Weinberg (1998), and Coles, Melott, & Munshi (1999) is that bias
is scale-independent on large scales, if, on small scales, it depends only on the local mass
density, even if that dependence is nonlinear and stochastic. Finally, a number of people
have combined second-order gravitational perturbation theory with a phenomenological
bias relation which is nonlinear at second order to devise observational tests of the existence
of bias (Szapudi, Meiksin, & Nichol 1996; Matarrese, Verde, & Heavens 1997; Szapudi 1998;
Verde et al. 1998; Buchalter & Kamionkowski 1999; Frieman & Gaztaña 1999; Szapudi et
Chapter 1: Introduction

al. 1999). These tests essentially all consist of testing whether the deviations from Gaussianity of the density field are the sort one would expect from gravitational instability, or whether these deviations are being altered by some biasing effect. It is worth noting that the initial results from some of these investigations actually indicate that the galaxies trace the mass density field well (Szapudi, Melksin, & Nichol 1996; Frieman & Gaztañaga 1999).

The second approach is to use semi-analytic models to simulate galaxy formation (the earliest inklings being found in Ostriker & Rees 1977 and White & Rees 1978; practitioners of the modern approach being White et al. 1987; White & Frenk 1991; Cole et al. 1994; Kauffmann et al. 1993, 1999; Somerville, Primack & Faber 1998; Benson et al. 1999). These methods use either N-body simulations or the Bond et al. (1991) extension of the Press & Schechter (1974) formalism (an approach exemplified by Mo & White 1996) to follow dark matter halos and their merging histories. When halos merge, the galaxies are allowed to merge only on the timescale of dynamical friction, thus accounting for the overmerging in the extended Press-Schechter approach and due to finite resolution in N-body codes. Simple rules for star formation and feedback follow the evolution of the baryonic gas inside these halos. Kauffmann, Nusser, & Steinmetz (1997) and Kauffmann et al. (1999) used such models to study the relation between mass density and galaxy density. In Chapter 2, we compare these results with the results of hydrodynamic simulations. Similar analytic approaches, which account for the reduced accretion onto halos when the intergalactic medium is shock-heated, are being developed by Oh & Ostriker (1999), although these models cannot yet address the clustering of galaxies. All of these approaches have the virtue that they can quickly test many different assumptions about galaxy formation and cosmology. They have the drawback that they have many different assumptions about galaxy formation and gas cooling in halos which need to be tested.

The semi-analytic approach has already yielded some important insights into the nature of galaxy formation. It was in the development of the semi-analytic models that the "overcooling" phenomenon of the CDM model was first recognized; that is, the low-mass halos which form first yield far too much star-formation at early times, given the luminosity of stars we see in the universe at $z = 0$. White & Rees (1978) proposed what is now the standard solution to this dilemma, which is the suppression of star-formation in low velocity dispersion ($< 20-30$ km/s) halos due to gas heating by supernovae. One of the outstanding problems in galaxy formation theory which semi-analytic models have enunciated is the incompatibility of the Tully-Fisher relation with the luminosity function of galaxies (Kauffmann, White, & Guiderdoni 1993), without drastic assumptions about the effects of dust, supernova feedback, or surface-brightness selection effects on the luminosity function (Somerville, Primack & Faber 1998; Kauffmann et al. 1999; Dalcanton, Spergel, & Summers 1997). With the assumption that ellipticals form through the mergers of spirals, these models can nevertheless account for the morphological segregation of
ellipticals and spiral galaxies at about the level observed. The degree of bias relative
to mass depends somewhat on how the normalization of galaxy light to halo mass is
set (using the Tully-Fisher relation, the luminosity function, or fitting the Milky Way
galaxy), since clustering properties are tied intimately to halo mass in the Press-Schechter
formalism. Under the right assumptions about all of the effects, these methods produce
a bias which transforms a dark matter correlation function with a complicated shape to
a power-law function which matches the galaxy correlation function found by the APM
survey (Benson et al. 1999). Finally, the semi-analytic models naturally predict a bias
which increases with redshift to an extent which agrees with the observations of Adelberger
et al. (1998).

The third approach, which this thesis adopts, is to use hydrodynamic simulations
with heuristic models of galaxy formation built into them. Generically, these simulations
follow the radiative physics of the gas and use physically motivated prescriptions to
convert baryonic fluid into collisionless stellar particles, which is necessary due to our
ignorance of the details of the star-formation processes and the finite resolution of the
code. Carlberg & Couchman (1989), Evrard, Summers, & Davis (1994), Katz, Hernquist,
& Weinberg (1992, 1998), and Pearce et al. (1999) have all used the Lagrangian smoothed
particle hydrodynamic (SPH) method to address the question of galaxy formation; Gnedin
(1996a,b), Cen & Ostriker (1992, 1998a), and this thesis use mesh-based methods for the
same purpose. The main advantage of hydrodynamic simulations is that they handle the
interaction of gravitation, gas inflow, and heating and cooling processes self-consistently
and according to known physical laws. On the other hand, the physics of star-formation
still needs to be put in by hand, and the resolution of current simulations is still marginal,
though improving rapidly.

A notable success of the hydrodynamic simulations is that by assuming a reasonable
value of the $\Omega_b$, the cosmological density of baryons, they can reproduce the mass in stars
observed in the local universe to within a factor of two, despite having very few parameters
to tweak (namely, a critical overdensity $\eta \sim 5$, plus feedback efficiency parameters), none of
which affect this number greatly. The star-formation history in these simulations also is a
good fit to observations, in particular doing a good job of reproducing the sharp decrease
of star-formation since $z \sim 1$ (Nagamine, Cen & Ostriker 1999). Furthermore, Cen &
Ostriker (1993) realized early on that if one related the effective age of galaxies to their
morphology, one could reproduce the level of morphological segregation observed in the
real universe. Simulations have generally predicted a bias which increases on scales below
the nonlinear length scale and is independent of scale on large scales. In addition, these
simulations predict an increase of bias with redshift in agreement with both observations
and the semi-analytic models. I will describe more results along these lines in Chapters 2
and 3.
Chapter 1: Introduction

In this thesis, I approach the problem of bias and galaxy formation from theoretical and observational perspectives. In Chapters 2 and 3, I examine the results of hydrodynamic simulations, first looking at the properties of galaxy clustering at redshift zero and next considering the time evolution of the clustering of galaxy formation. The main conclusion of these chapters is that the simulations indicate that the local temperature of the gas around forming galaxies plays an important role in determining the clustering properties of the galaxies. In Chapter 4, I use the Las Campanas Redshift Survey to test one of the consequences of these models, namely, that there is scatter in the relationship between the density fields of different types of galaxies. In an Epilogue, I summarize my results and briefly describe future theoretical and observational prospects in this field.

REFERENCES

Chapter 1: Introduction

Frieman, J. A., & Gaztañaga, E. 1999, 521, L83
Chapter 1: Introduction

Chapter 1: Introduction


Chapter 1: Introduction