the Universe was younger. Therefore, we may hope to infer how galaxies form and evolve by comparing their properties, in a statistical sense, at different epochs. In addition, at each epoch we can try to identify regularities and correspondences among the galaxy population. Although galaxies span a wide range in masses, sizes, and morphologies, to the extent that no two galaxies are alike, the structural parameters of galaxies also obey various scaling relations, some of which are remarkably tight. These relations must hold important information regarding the physical processes that underlie them, and any successful theory of galaxy formation has to be able to explain their origin.

Galaxies are not only interesting in their own right, they also play a pivotal role in our study of the structure and evolution of the Universe. They are bright, long lived and abundant, and so can be observed in large numbers over cosmological distances and time scales. This makes them unique tracers of the evolution of the Universe as a whole, and detailed studies of their large scale distribution can provide important constraints on cosmological parameters. In this book we therefore also describe the large scale distribution of galaxies, and discuss how it can be used to test cosmological models.

In Chapter 2 we start by describing the observational properties of stars, galaxies and the large scale structure of the Universe as a whole. Chapters 3 through 10 describe the various physical ingredients needed for a self-consistent model of galaxy formation, ranging from the cosmological framework to the formation and evolution of individual stars. Finally, in Chapters 11–16 we combine these physical ingredients to examine how galaxies form and evolve in a cosmological context, using the observational data as constraints.

The purpose of this introductory chapter is to sketch our current ideas about galaxies and their formation process, without going into any detail. After a brief overview of some observed properties of galaxies, we list the various physical processes that play a role in galaxy formation and outline how they are connected. We also give a brief historical overview of how our current views of galaxy formation have been shaped.

1.1 The Diversity of the Galaxy Population

Galaxies are a diverse class of objects. This means that a large number of parameters is required in order to characterize any given galaxy. One of the main goals of any theory of galaxy formation is to explain the full probability distribution function of all these parameters. In particular, as we will see in Chapter 2, many of these parameters are correlated with each other, a fact which any successful theory of galaxy formation should also be able to reproduce.

Here we list briefly the most salient parameters that characterize a galaxy. This overview is necessarily brief and certainly not complete. However, it serves to stress the diversity of the galaxy population, and to highlight some of the most important observational aspects that galaxy formation theories need to address. A more thorough description of the observational properties of galaxies is given in Chapter 2.

(a) Morphology One of the most noticeable properties of the galaxy population is the existence of two basic galaxy types: spirals and ellipticals. Elliptical galaxies are mildly flattened, ellipsoidal systems that are mainly supported by the random motions of their stars. Spiral galaxies, on the other hand, have highly flattened disks that are mainly supported by rotation. Consequently, they are also often referred to as disk galaxies. The name 'spiral' comes from the fact that the gas and stars in the disk often reveal a clear spiral pattern. Finally, for historical reasons, ellipticals and spirals are also called early- and late-type galaxies, respectively.

Most galaxies, however, are neither a perfect ellipsoid nor a perfect disk, but rather a combination of both. When the disk is the dominant component, its ellipsoidal component is generally called the bulge. In the opposite case, of a large ellipsoidal system with a small disk, one typically talks about a disky elliptical. One of the earliest classification schemes for galaxies, which is still heavily used, is the Hubble sequence. Roughly speaking, the Hubble sequence is a sequence in the admixture of the disk and ellipsoidal components in a galaxy, which ranges from early-type ellipticals that are pure ellipsoids to late-type spirals that are pure disks. As we will see in Chapter 2, the important aspect of the Hubble sequence is that many intrinsic properties of galaxies, such as luminosity, color, and gas content, change systematically along this sequence. In addition, disks and ellipsoids most likely have very different formation mechanisms. Therefore, the morphology of a galaxy, or its location along the Hubble sequence, is directly related to its formation history.

For completeness, we stress that not all galaxies fall in this spiral vs. elliptical classification. The faintest galaxies, called dwarf galaxies, typically do not fall on the Hubble sequence. Dwarf galaxies with significant amounts of gas and ongoing star formation typically have a very irregular structure, and are consequently called (dwarf) irregulars. Dwarf galaxies without gas and young stars are often very diffuse, and are called dwarf spheroidals. In addition to these dwarf galaxies, there is also a class of brighter galaxies whose morphology neither resembles a disk nor a smooth ellipsoid. These are called peculiar galaxies and include, among others, galaxies with double or multiple subcomponents linked by filamentary structure and highly distorted galaxies with extended tails. As we will see, they are usually associated with recent mergers or tidal interactions. Although peculiar galaxies only constitute a small fraction of the entire galaxy population, their existence conveys important information about how galaxies may have changed their morphologies during their evolutionary history.

(b) Luminosity and Stellar Mass Galaxies span a wide range in luminosity. The brightest galaxies have luminosities of $\sim 10^{12}\,L_\odot$, where $L_\odot$ indicates the luminosity of the Sun. The exact lower limit of the luminosity distribution is less well defined, and is subject to regular changes, as fainter and fainter galaxies are constantly being discovered. In 2007 the faintest galaxy known was a newly discovered dwarf spheroidal Willman I, with a total luminosity somewhat below $1000\,L_\odot$.

Obviously, the total luminosity of a galaxy is related to its total number of stars, and thus to its total stellar mass. However, the relation between luminosity and stellar mass reveals a significant amount of scatter, because different galaxies have different stellar populations. As we will see in Chapter 10, galaxies with a younger stellar population have a higher luminosity per unit stellar mass than galaxies with an older stellar population.

An important statistic of the galaxy population is its luminosity probability distribution function, also known as the luminosity function. As we will see in Chapter 2, there are many more faint galaxies than bright galaxies, so that the faint ones clearly dominate the number density. However, in terms of the contribution to the total luminosity density, neither the faintest nor the brightest galaxies dominate. Instead, it is the galaxies with a characteristic luminosity similar to that of our Milky Way that contribute most to the total luminosity density in the present-day Universe. This indicates that there is a characteristic scale in galaxy formation, which is accentuated by the fact that most galaxies that are brighter than this characteristic scale are ellipticals, while those that are fainter are mainly spirals (at the very faint end dwarf irregulars and dwarf spheroidals dominate). Understanding the physical origin of this characteristic scale has turned out to be one of the most challenging problems in contemporary galaxy formation modeling.

(c) Size and Surface Brightness As we will see in Chapter 2, galaxies do not have well-defined boundaries. Consequently, several different definitions for the size of a galaxy can be found in the literature. One measure often used is the radius enclosing a certain fraction (e.g. half) of the total luminosity. In general, as one might expect, brighter galaxies are bigger. However, even for a fixed luminosity, there is a considerable scatter in sizes, or in surface brightness, defined as the luminosity per unit area.

The size of a galaxy has an important physical meaning. In disk galaxies, which are rotation supported, the sizes are a measure of their specific angular momenta (see Chapter 11). In the case of elliptical galaxies, which are supported by random motions, the sizes are a measure of the amount of dissipation during their formation (see Chapter 13). Therefore, the observed distribution of galaxy sizes is an important constraint for galaxy formation models.

(d) Gas Mass Fraction Another useful parameter to describe galaxies is their cold gas mass fraction, defined as $f_{\rm gas} = M_{\rm cold}/[M_{\rm cold} + M_\star]$, with $M_{\rm cold}$ and $M_\star$ the masses of cold gas and stars, respectively. This ratio expresses the efficiency with which cold gas has been turned into stars. Typically, the gas mass fractions of ellipticals are negligibly small, while those of disk galaxies increase systematically with decreasing surface brightness. Indeed, the lowest surface brightness disk galaxies can have gas mass fractions in excess of 90 percent, in contrast to our Milky Way which has $f_{\rm gas} \sim 0.1$.

(e) Color Galaxies also come in different colors. The color of a galaxy reflects the ratio of its luminosity in two photometric passbands. A galaxy is said to be red if its luminosity in the redder passband is relatively high compared to that in the bluer passband. Ellipticals and dwarf spheroidals generally have redder colors than spirals and dwarf irregulars. As we will see in Chapter 10, the color of a galaxy is related to the characteristic age and metallicity of its stellar population. In general, redder galaxies are either older or more metal rich (or both). Therefore, the color of a galaxy holds important information regarding its stellar population. However, extinction by dust, either in the galaxy itself, or along the line-of-sight between the source and the observer, also tends to make a galaxy appear red. As we will see, separating age, metallicity and dust effects is one of the most daunting tasks in observational astronomy.

(f) Environment As we will see in §§2.5–2.7, galaxies are not randomly distributed throughout space, but show a variety of structures. Some galaxies are located in high-density clusters containing several hundreds of galaxies, some in smaller groups containing a few to tens of galaxies, while yet others are distributed in low-density filamentary or sheet-like structures. Many of these structures are gravitationally bound, and may have played an important role in the formation and evolution of the galaxies. This is evident from the fact that elliptical galaxies seem to prefer cluster environments, whereas spiral galaxies are mainly found in relative isolation (sometimes called the field). As briefly discussed in §1.2.8 below, it is believed that this morphology–density relation reflects enhanced dynamical interaction in denser environments, although we still lack a detailed understanding of its origin.

(g) Nuclear Activity For the majority of galaxies, the observed light is consistent with what we expect from a collection of stars and gas. However, a small fraction of all galaxies, called active galaxies, show an additional non-stellar component in their spectral energy distribution. As we will see in Chapter 14, this emission originates from a small region in the centers of these galaxies, called the active galactic nucleus (AGN), and is associated with matter accretion onto a supermassive black hole. According to the relative importance of such non-stellar emission, one can separate active galaxies from normal (or non-active) galaxies.

(h) Redshift Because of the expansion of the Universe, an object that is farther away will have a larger receding velocity, and thus a larger redshift. Since the light from high-redshift galaxies was emitted when the Universe was younger, we can study galaxy evolution by observing the galaxy population at different redshifts. In fact, in a statistical sense the high-redshift galaxies are the progenitors of present-day galaxies, and any changes in the number density or intrinsic properties of galaxies with redshift give us a direct window on the formation and evolution of the galaxy