1

Introduction

Since Edwin Hubble showed in the 1930s that the Universe we live in is expanding, our knowledge about the formation of structure has also been expanding. Just like that of the Universe, the expansion of our body of knowledge seems to be accelerating, aided by the ever increasing power of telescopes, detectors, computers and software. This thesis deals with the formation of structure in the Universe, from the viewpoint of computer simulations. As large scale (>kpc) processes proceed on timescales many orders of magnitude longer than a human life (or the preparation of a PhD thesis), directly observing this evolution is impossible. Simulations are used to understand how objects evolve, while snapshots of the simulations may be compared to observations, which in essence are nothing more than snapshots of the real Universe.

1.1 Galaxy formation

1.1.1 The growth of structure in the Universe

About 13.7 Gyr ago, the Universe was born in a hot, dense and highly uniform state. The temperature and density of the hot plasma were almost completely uniform. Through the tight coupling between baryonic matter and radiation, the growth of density perturbations in the dark matter was hardly followed by the baryons. At the time of recombination, about 380.000 years later, the Gaussian deviations from the mean density were of order $\delta\rho/\rho \sim 10^{-5}$ and the wavelengths of the perturbations exhibit a spectrum that evolved from an initial power-law spectrum: $P(k) \sim k^{n_s}$, where $P(k)$ is the power at wave number k. The spectral index $n_s \simeq -1$. We know this, because the radiation emitted by the recombination of hydrogen atoms is redshifted by a factor ~1100 and observed as the Cosmic Microwave Background by e.g. the COBE and WMAP satellites (for a recent review, see Hu & Dodelson, 2002). The baryons decouple from the radiation and flow into the potential wells already in place (and still growing) in the dark matter.

Because the fluctuations are well within the linear regime, linear theory can be used to calculate the growth of the perturbations, until they get to the non-linear regime, where calculations with pen and paper will generally not suffice. The perturbations grow under gravity, making over-dense regions even denser in the course of time. Meanwhile the Universe expands, lowering the overall density.

The non-linear perturbations decouple from the expansion of the Universe and collapse into gravitationally bound, eventually virialised structures, that are generally named ‘haloes’. Within these haloes galaxies may form. This involves more than just gravity and needs to take full account of hydrodynamics, star formation, feedback effects and other ‘gastrophysical’ phenomena.

1.1.2 The evolution of baryons

The important difference between the formation of dark matter haloes and the formation of galaxies inside them, is the fact that gas is collisional, whereas dark matter is collisionless. Dark matter cannot cool and only acts upon the other (dark) matter through gravity. Gas can cool. Pressure gradients will drive gas flows from high to low pressure and there are many possible ways of injecting heat into gas, both as a result of shocks, where kinetic energy of gas is transformed into internal energy, but also by the absorption of radiation.

At a redshift of about nine (or somewhere between six and fifteen Komatsu et al., 2009, 2010) the first sources of light reionized the Universe. There is a roughly uniform UV background that is the result of young stars and accreting supermassive black holes (which of the two dominates is redshift dependent) that keeps intergalactic gas above a temperature floor of roughly 10^4 Kelvin (this temperature floor is density dependent, because of the interplay between the ionization by this background and recombination, and redshift dependent due to the adiabatic expansion of the Universe).

Initially, the baryons will just follow the dark matter and decouple from the expansion, to collapse into the dark matter haloes. In the center of these, where the densities are high, radiative cooling will become more efficient (atomic/ionic line cooling scales with the density squared). When gas cools, pressure support is lost and higher densities can be reached.

When galaxies start to form, a large variety of processes come into play that can influence the future evolution of the baryons. In the following two sections we will discuss some internal and external processes that could in principle be important for galaxy evolution. We will hereby focus on processes that are investigated in the remainder of this thesis.

1.1.3 Internal processes in galaxy formation

Gas in haloes is at higher density than gas in the intergalactic medium (IGM). Inside a halo the gas density follows a profile with densities higher in the center than in the outskirts. This density gradient corresponds to a gradient in cooling rate (in the simplest assumption that the temperature and metallicity are initially uniform). Pressure support is lost from the center where the cooling is very efficient, so more gas can fall in and an inward cooling flow establishes. In the center of the halo the gas cools down to roughly 10^4 K and settles in a disk, because the (specific) angular momentum is conserved.

In high density gas, radiative cooling through metal lines can become very efficient and dust column densities can become high enough such that clouds can become self-shielded from photo-dissociating and ionizing radiation. The gas can now become mostly molecular. The rotational and vibrational levels of molecules give rise to many new channels of cooling of the gas. In the gas in the disk, a multiphase interstellar medium (ISM) establishes, consisting of denser, colder, (partly) molecular clouds, embedded in warmer gas, the spatial distribution of which is fractal.

Eventually, when the gas is cold and dense enough, stars may form. The process of star formation itself is very complicated and many theses could be, and have been written on the formation of stars from a giant molecular cloud. For people who work on scales of galaxies and bigger, this star formation process is often heavily simplified, sweeping all details on scales smaller than $\sim 10^4 M_\odot$ under the carpet. Empirical relations between e.g. gas surface density and star formation rates are used. Star formation in galaxies is observed to follow a close relation between the gas surface density and the star formation rate surface density, a relation known as the Kennicutt-Schmidt (Kennicutt, 1998a) law: $\dot{\Sigma}_* = A\Sigma_g^n$, with $n \approx 1.4$ and A a normalization factor that depends on the stellar initial mass function (IMF), because star formation rate indicators are only sensitive to stars above some mass, because it is the effect of ionizing radiation that is measured.

After stars have formed, they evolve and eventually die. During their lifetimes, stars of different mass expel different chemical elements at different times, through stellar winds and/or explosions. These chemical yields are added to the interstellar medium (ISM) around the stars. This enrichment results in more efficient cooling, and in the possibility to form dust. The relative yields of different metals depends on the IMF, as different types of stars are the main producers of different elements.

The most massive stars already start exploding as Supernovae (SNe) after a few million years and inject about 10^{51} erg of kinetic energy per explosion into the surrounding gas. Part of the energy will be thermalized in shocks and radiated away. The remainder can stir up the surroundings of the star forming region (increasing the local turbulence), it can blow ‘super bubbles’ around complexes of star forming regions and might even blow large scale galactic winds. It is therefore obvious that SNe have a considerable impact on their host galaxy.

Most, if not all, galaxies that have a spheroidal component (elliptical galaxies, or disk galaxies with a bulge) also host a supermassive black hole (SMBH) in their centres (Kormendy & Richstone, 1995; Ferrarese & Merritt, 2000). These SMBHs accrete gas from an accretion disk, a process in which part of the rest mass energy that is accreted is not added to the mass of the black hole but radiated away. This radiation may heat and push surrounding gas. This AGN feedback comes in two flavours in nature: the relatively quiet accretion mode for low accretion rates (compared to their Eddington limit), which is called ‘radio mode’, as these systems are observed as radio galaxies. AGN with high accretion rates (comparable to the Eddington limit) have strong optical emission lines and the feedback corresponding to this mode is often called ‘quasar mode’. The energy output corresponding to black hole growth depends on the mass and accretion rates of the black hole, and is stronger for more massive black holes, which tend to live in more massive spheroids.

1.1.4 External processes in galaxy formation

Galaxies do not live alone in the Universe. Galaxies do have neighbouring galaxies, either within the same dark matter halo (for sufficiently massive haloes) or in haloes next to them.

One of the main drivers of galaxy evolution is the mass of the galaxy’s host dark matter halo. This host halo mass sets the gravitational potential well, and therefore affects the central density. Also, more massive haloes are in general older and therefore can start forming stars earlier. Most of the gas in the halo which is not already cold and inside the galaxies, is very tenuous and hot, at about the virial temperature of the halo. In order to form stars this gas needs to cool down, and the cooling time is a strong function of the temperature (the gas has to cool down further from higher temperatures, and the cooling rate is a complicated but in general a decreasing function of temperature in the temperature range $10^5 - 10^7$ K, see e.g. Wiersma et al., 2009a).

The rate at which, and the mode in which, gas accretes onto haloes is also a function of halo mass. Gas can accrete in two main modes: hot and cold (e.g. Kereš et al., 2005; Ocvirk et al., 2008). With hot accretion we indicate gas that flows in and shock heats near the virial radius to about the virial temperature of the halo. When gas accretes cold, streams (and clumps) of high density fall into the center, but the energy gained in the (smaller) shocks are efficiently radiated away and therefore do not add to the temperature. Gas that accretes cold does not have to cool down much before it can participate in star formation, whereas shock heated gas at the virial temperature (at least in massive haloes) has very long cooling times and therefore can delay star formation significantly. The transition from cold to hot accretion is not sharp (in many haloes a fraction of the gas accretes hot and a fraction accretes cold) and lies at around a halo mass of order $10^{12} M_{\odot}$, with more massive haloes accreting more gas in the hot mode (Dekel & Birnboim, 2006; Dekel et al., 2008).

Dark matter haloes are clustered. The amount of clustering is a function of mass, such that more massive haloes cluster more strongly (e.g. Kaiser, 1984; Cole & Kaiser, 1989; Mo & White, 1996). More massive haloes also have more subhaloes containing galaxies and the fraction of the mass of a DM halo that is in substructure is roughly constant with halo mass (Gao et al., 2004). If there is a minimum (sub-) halo mass for galaxy formation, then a more massive halo hosts more small (satellite) galaxies.

As revealed by the marvelous images of colliding galaxies, interactions between two systems are also of importance in the growth of galaxies. The tidal torques the two galaxies exert on each other drive gas flows inward, thereby fueling a central star burst, and possibly a quasar outburst of the central SMBHs (which eventually may merge too). For a few dynamical times, the galaxies will have an elevated star formation rate, and the end product of a major merger (mass ratio $\geq 1/3$) is often an elliptical galaxy, regardless of the Hubble types going in to the collision. For elliptical galaxies, which have hardly any cold gas to form stars from, the dominant growth mechanism is mergers, and the most massive ellipticals in the known Universe are thought to be the result of a series of major and minor mergers, deep inside the potential wells of massive DM haloes.

1.1.5 The interplay between internal and external processes in galaxy formation

In the previous paragraphs we listed many processes that are important in galaxy formation. To what extent the different physical processes are important, and how they act together to make up galaxy evolution, is the largest unknown in galaxy formation theory.

Whereas star formation is a necessary ingredient to form (optically) observable galaxies, it is not clear what sets the star formation rate in galaxies. Gas accretion is a necessary ingredient, and so is gas cooling. Feedback processes counteract cooling, and possibly also accretion, by blowing gas out of galaxies. It is expected that star formation in galaxies is to some extent self-regulated. If the cooling and accretion processes are dominant over feedback processes, stars will form while gas pressure support falls. Therefore, more gas will collapse and stars will form until feedback, e.g. in the form of SN explosions, is able to ‘counteract’ star formation. If, on the other hand, feedback is dominant over the cooling and accretion processes, star formation will cease, SN feedback will decrease and cooling and infall will result in more star formation, until the rate of star formation reaches some sort of quasi-equilibrium with the amount of feedback.

In galaxies of very different mass, the equilibrium between feedback and star formation may happen at very different scales. In more massive galaxies, the pressure in the ISM is higher, the amount of mass that need to be swept up by winds blown by SNe is larger, and the potential well from which the wind needs to escape deeper. With only an energy limit to the amount of feedback (the total amount of energy from SNe) it is not clear how this equilibrium settles in different environments. For example, at the same energy a lot of mass can be kicked at low velocity or vice versa. An upper limit for the energy input in winds is not necessarily related to the amount of energy available from SNe, if winds are driven by radiation pressure of the stellar population, rather than by the SN explosions themselves. An equilibrium between feedback and star formation may not always be possible. For example, if SN driven bubbles do not blow out of the galaxy and feedback is very inefficient. If that is the case, other feedback mechanisms like AGN feedback are required to suppress the star formation rate of galaxies.

The interaction between accretion flows bringing in new fuel for star formation and the outflows driven by star formation is a complicated non-linear process and requires accurate, high-resolution numerical simulations. Studies have not yet converged on how this interplay works, how the hot and cold accretion fractions depend on halo mass, redshift and feedback.

The problem of ‘nature versus nurture’ deals with the extent to which internal and external processes influence galaxy properties. A lot of observational work has been done in this field. Star forming properties of galaxies are found to be correlated with the stellar mass of galaxies, with their gas fractions and with surrounding galaxy densities (see references in Table 3.1 in Chapter 3). The colours and magnitudes of galaxies in turn depend largely on their stellar mass and recent star formation histories (e.g. Kauffmann et al., 2003; Blanton et al., 2005), and therefore also correlate with mass and environment. Because we know that the stellar mass (at least for the central galaxies of haloes) correlates with halo mass, and environmental density also correlates with halo mass Lemson & Kauffmann (1999), it is not yet clear what the main driving factor is, and whether there is more than one driving factor in galaxy evolution at all.

1.2 Numerical simulations

Although astronomy has always been, and will probably remain, an observationally driven science, a large part of our understanding of the evolution of structure in the Universe stems from simulations. In simulations, the evolution of a physical system is in principle completely determined by the code and the initial and boundary conditions. In some sense, you will ‘get out what you put in’. In practice it is, however, usually far from trivial to understand the outcome from the physics in the code and the initial and boundary conditions. Non-linear behavior of the system and the interplay between different ingredients of the simulation require a detailed investigation of the results in order to increase our understanding of the simulated objects.

In galaxy formation simulations, a lot of progress has been made over the past decades (for a somewhat dated review see Bertschinger, 1998). Although Eulerian mesh based codes (in which the volume is discretized) have also been used successfully for simulations of galaxies in a cosmological context, I will here focus on Lagrangian simulations, in which the mass in the Universe is discretized in particles, because those kind of simulations form the basis of a large part of this thesis.

1.2.1 Simulations of the dark matter component of the Universe

Until recently, cosmological simulations (simulations of a large, representative volume of the Universe with box sizes much larger than the objects of interest) were mainly N-body simulations in which only gravity is followed in a Universe in which the mass is discretized in point-like particles. These simulations predict the evolution of the dark matter component of the Universe, with the ‘details’ of the baryonic physics neglected. As following only gravity is relatively easy, our understanding of the large-scale structure of the Universe has rapidly increased due to simulation projects which followed only the dark matter component of the Universe (see e.g. Springel et al., 2005).

Galaxies in N-body simulations

With the evolution of dark matter alone, nothing can be said about the properties of the galaxies in such simulations. Semi-analytic models (SAMs) have been created in order to form a galaxy population on top of the dark matter simulations (Croton et al., 2006; Bower et al., 2006; De Lucia & Blaizot, 2007). These consist of analytic recipes, which depend on the merger history of the halo the galaxies are in. The recipes describe how gas flows into the haloes, cools, form stars, explode as SNe and how SMBHs form, grow and influence the gas in their haloes.

These SAMs generally come with a large number of free parameters (which are mostly motivated by the baryonic physics described in Section 1.1), so it is very well possible to form a galaxy population that is very representative of the galaxy population that is observed. The model parameters are usually tweaked to reproduce a few observables (principally the $z = 0$ galaxy luminosity function), and the model is then used to predict others. Given the large number of free parameters and functions, many of which may be poorly constrained or even unphysical, the predictive powers of these models may be questionable, but at least it is possible to create a galaxy population that matches a variety of observations.

Variations on N-body simulations

Galaxy formation models do not necessarily need N-body simulations in order to predict the behaviour of the dark matter component of the Universe. Several alternatives exist and are often used (they are usually less computationally expensive, but may lack small scale details and are less accurate). In extended Press-Schechter theory, for example, the dark matter halo merger histories can be obtained analytically. Another variation uses halo mass functions obtained from either analytic theory or $N-$body simulations and link the luminosities of observed galaxies to the dark matter haloes. In halo occupation distribution (HOD) models, the self-similarity of dark matter haloes is used, such that the number of galaxies, and their mass distribution is known as a function of the halo mass (e.g. Berlind & Weinberg, 2002). Using the resulting distributions of galaxies in haloes, galaxy luminosities from an observed galaxy luminosity function can be linked to these galaxies.

1.2.2 Smoothed Particle Hydrodynamics and sub-grid physics

Although the above methods can produce galaxy populations that satisfy observational constraints, they are not always physically well-motivated. Often, the different ingredients do not have a chance to interact, and in many cases the physical prescriptions for ingredients like gas cooling, star formation, feedback etc. are strongly simplified versions of reality. In order to gain physical insight into the interplay between baryonic processes in galaxy formation, one needs to follow the evolution of the baryons more self-consistently using simulations.

One way to do so, and this is the method used for most of this thesis, is to simulate the Universe with both a dark, and a baryonic component in hydrodynamical simulations. Gas particles discretize the mass in the simulated volume, and their hydrodynamical properties (e.g. density and pressure) are obtained by averaging over a kernel containing a fixed number of neighbours. These particles can then exert gravity and pressure on each other and evolve hydrodynamically, rather than just under the act of gravity. Such simulations are much more computationally challenging than the $N-$body simulations discussed above.

Although we take the smoothed particle hydrodynamics (SPH) approach (Monaghan, 1992) in order to follow the evolution of the gas (and the response of the dark matter to the presence and evolution of baryons), many of the processes shaping galaxies happen on scales below the resolution limit of cosmological simulations. As we will see below, typical particle masses in simulations of representative volumes of the Universe are limited (because the particle number is limited by computer memory and processor speed) to $m_p \gtrsim 10^5 M_\odot$ in the highest resolution simulations available, but more often one or more orders of magnitude higher. Gas cooling is a process on the scale of atoms, stellar evolution happens on scales of about 1 $M_\odot$ and the evolution of supernova remnants may require a similar resolution. These are just a few examples, in Chapter 2 we will go into many small-scale processes. Obviously, recipes have to be developed in order to describe the effect of small scale processes below the resolution scale of the simulation. These are called ‘sub-grid models’, and these are the ingredients that make different SPH simulations differ from each other (strongly).

To date, an extensive and fair comparison between the many different sub-grid models has not been made. A systematic comparison of sub-grid models requires a suite of simulations, run with the same code, on the same initial conditions, varying the sub-grid recipes one-by-one. That is exactly what the OverWhelmingly Large Simulations project set out to do.

1.3 The OverWhelmingly Large Simulations

A recent effort in simulating a representative volume of the Universe is conducted at Leiden Observatory, the Netherlands and is called 'The OverWhelmingly Large Simulations' (*OWLS*, Schaye et al., 2010), a project catalyzed by the temporary availability of the IBM BlueGene/L supercomputer 'Stella' which was built for the LOFAR collaboration in Groningen. The name of the project is not only supposed to tell you that the simulated volumes are large, but also that the number of variations in the sub-grid modeling is unprecedented for cosmological simulation projects. The philosophy of *OWLS* is to keep the simulation ingredients as simple as possible and to vary sub-grid models and/or parameters one by one (away from what we call the 'reference model').

By "keeping things simple", we mean that when we have to introduce a sub-grid model for an unresolved process, we keep this model simple, and do not introduce more parameters than necessary and/or justifiable.

The simulations are extensively described in Schaye et al. (2010) and an extensive summary is given in Chapter 2. Over 50 high resolution simulations have been carried out, totaling many tens of terabytes as a result of millions of CPU-hours of calculation. In this thesis we will focus on the populations of galaxies formed in the different *OWLS* runs. With such an extensive set of simulations, many studies are possible, and this thesis only contains a small subset of what *has been* done, let alone what *could be* done.

1.4 Thesis summary

In **Chapter 2** a summary is given of all *OWLS* runs used in this thesis. The influence of the physics and resolution of the simulation (in terms of mass as well as box size) on the resulting galaxy population at $z = 2$ are discussed. As the variation of sub-grid models is the unique feature of *OWLS*, we discuss the effect of all the sub-grid models in quite some detail. We look at the relation between properties of Friends-of-Friends (FoF) haloes in the high resolution simulations. In particular, the star formation rate (SFR), the build up of stellar mass, and gas, star and baryon mass fractions as a function of halo mass are used to assess the effectiveness of the various feedback models and we compare shortly to observations. Interesting conclusions from this chapter are that the star formation rate of a galaxy is self-regulated by gas accretion (set by halo mass and gas cooling) and feedback and that the star formation recipe regulates the amount of available fuel (i.e. the gas mass fraction) of the haloes, but not the star formation rate.

As the extent to which halo mass and environment sets galaxy properties is not yet clear, we investigate in **Chapter 3** how to disentangle the influence of halo mass and environment. It is well known that environmental density and halo mass correlate. In the literature, many different definitions of environmental density occur. We make use of the Millennium Simulation and semi-analytic models of galaxy formation in order to investigate the correlation between environmental parameters and halo mass on a galaxy population which matches observational constraints. We show how well popular environmental parameters correlate with halo mass, as a function of the scale on which environment is measured. We will show that if the minimum mass/luminosity of the neighbours used to characterize the environment is fixed relative to the mass/luminosity of the halo in question, and if the distance to these neighbours is scaled to a typical distance for the galaxy in question (e.g. the virial radius of its host halo) that then the measure of environmental density can be made to be independent of halo mass. If one wants to investigate the effects of halo mass (‘internal environment’) and ‘external environment’ separately, it is most useful to use one parameter that correlates very strongly with halo mass (e.g. the number of galaxies within roughly a virial radius) and one that is independent of halo mass.

In order to compare simulations and observations one can go two ways: determine physical properties from observables and compare these to the simulations, or extract observables from the simulation and compare these to observed galaxy properties. In **Chapter 4** we extract luminosity functions from the OWLS simulations and investigate how these depend on input physics, dust attenuation and galaxy selection. The dependence of the LF on input physics is very similar to the dependence of the stellar mass function on input physics, which was already shown in Chapter 2. Dust attenuation is hard to estimate in SPH simulations with particle masses exceeding the mass of absorbing clouds in the ISM of galaxies. We estimate it from the column density of metals, normalized to the extinction as a function of metal column in the solar neighbourhood. As the definition of galaxies used by simulators (gravitationally bound structures of particles) and observers (some region of an image that exceeds the background in intensity) are fundamentally different, we try to assess if the obtained luminosity function in simulations can be expected to be the same as the observed luminosity function of galaxies, under the assumption that the underlying galaxy populations are identical. To that end we project our star particles onto images, smear the images with a point spread function (PSF) and extract the galaxy luminosity function with the tools observers would use. We find that the LFs are in general very similar to the ones directly obtained from the simulations, but that PSFs which are large compared to the galaxies may flatten the faint end of the LF, which would alleviate a major tension between observed and simulated LFs.

Finally, in **Chapter 5** we investigate the stellar content, broadband photometry and metal enrichment of idealized galaxy models in the framework of the so-called ‘integrated galactic initial mass function’ (IGIMF, see Kroupa & Weidner, 2003; Weidner & Kroupa, 2004, 2005, 2006). Star formation occurs mostly in clusters (Lada & Lada, 2003; Piskunov et al., 2008). These clusters follow a mass distribution, that is quite similar in shape to the stellar IMF (e.g. Larsen, 2002; Bastian, 2008). The cluster mass function favours low mass objects, so if the power-law mass function were to extend all the way down to clusters of just a few solar masses, it is clear that the IMF summed up over all clusters must be deficient of high mass stars compared to the underlying IMF. A star can, after all, not be more massive than its host cluster. We will investigate the IGIMF under various assumptions for the method used to sample the stellar masses in the clusters and for different cluster mass functions. We use the IGIMFs as input IMFs for the GALEV population synthesis models (Bicker et al., 2004; Kotulla et al., 2009) to obtain broadband magnitudes and metallicities of closed box galaxy models. We find that the change in broadband colours from IMF to several versions of the IGIMF is smaller than the galaxy-to-galaxy scatter of colours. The O-star content of our Milky Way is significantly altered by the effects of clustered star formation, but the exact number that e.g. GAIA (Perryman et al., 2001) will observe depends on various other uncertain quantities. If the IGIMF indeed significantly deviates from the IMF (which depends on the unknown low mass behaviour of the star cluster mass function), then the metal content of galaxies is the most promising discriminator between (IG)IMFs.

1.5 The (near) future

The studies described in this thesis do answer some open questions in the field of galaxy formation, but are by no means final answers to the large open questions. Many of the simulated properties of galaxies do not correspond to observations and many of the physical processes in the simulations are highly (over-) simplified. In the near future much progress can be made on both the computational and the observational side of this topic. Whereas numerical models will become ever more sophisticated (due to the availability of more computing power and due to an improvement of software), observations with the new and upcoming observational facilities like JWST, ALMA, LOFAR, E-ELT and many others will shed new light on the state of galaxies and larger scale structures in the near and distant Universe.