ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

SCUOLA DI SCIENZE
Physics and Astronomy Department
Bachelor in Astronomy

Galaxy Formation

From Primordial Fluctuations to Structure Formation in The Universe.

Relator:
Prof.
Daniele Dallacasa

Presented by:
Mario Flores Herrera

Sessione III
Anno Accademico 2017 – 2018

November 28, 2017

Contents

1 Introduction

How everything was formed? Maybe is one of the biggest questions that can be out there and it certainly is one of the questions that led me to become a scientist.

An enormous progress has been made from the early years of cosmology until today, in both theoretical and observational; and this give us a solid enough framework to understand our universe in the more complete and general sense as possible, so now more that ever in the history of astrophysics we are not only able to ask ourselves the biggest questions, but to answer them.

Along this thesis we are going to treat the problem of structure formation in the universe, following the hierarchical model first proposed by **Press & Schechter** (1974) and then reinforced by **White & Ress** (1978). In the first part of this article, using the today-accepted ΛCDM cosmological model, we are going to briefly review the primordial fluctuations that led later to the formation of all structures, including galaxies. In the second part, we'll discuss the formation of dark matter halos, that will be of main importance for the understanding of galaxy formation. On the third and fourth part we are going to discuss the formation of luminous structure within the DM halos, and the causes of the different morphology observed in different galaxies, respectively. Following this approach, we'll compare some of the main theoretical results given in the historical papers already cited above with up-to-date observational data.

2 The Concordance Model

As we said it before, The so-called concordance model of cosmology proposes an isotropic, homogeneous and zero-curvature universe with a Hubble parameter at present time of $H_0 \approx 70$ $[Kms^{-1}Mpc^{-1}]$ and a composition of the universe as follows:

- 70% Dark Energy
- 20% Dark Matter
- 10% Baryonic Matter

Recent observational data that combined the results from WMAP satellite and observations of high-redshift type I_a supernova (Hinshaw et. al. 2009)[1] showed a Hubble parameter at present time of $H_0 \approx 70.5 \pm 1.3$ $[kms^{-1}Mpc^{-1}]$ and a composition of the universe as follows:

- 72% Dark Energy
- 23% Dark Matter
- 4.6% Baryonic Matter
- $<$ 1% Neutrinos

So, the ΛCDM model is strongly supported by observation. The discussion about the possible nature of both, dark matter and dark energy, although pretty interesting, exceeds completely the objective of this thesis.

[1] `https://arxiv.org/abs/0803.0732`

3 Thermal History of the Universe and Perturbation Growth

3.1 The Very Early Universe

As it was proposed by Guth in 1980 when the universe was only 10^{-36} seconds old, an rapid and exponential expansion occurred, lasting until $\approx 10^{-32}$ seconds after the Big Bang and making the universe grow by a factor of $\approx 10^{30}$. This period is known as inflation. The major importance of this phenomena for our analysis is that, during the inflation, the quantum fluctuations that arise naturally in quantum mechanics [2] grow to become real perturbations.

3.2 The Radiative Epoch

As the universe keeps expanding, the temperature keeps falling, allowing particles to form. At $t \approx 10^{-6}$ seconds, the first protons and neutrons were formed and at $t \approx 1$ second the first electrons were formed. Around this time, neutrinos decoupled from the other particles forming the CνB [3]

Regarding the structure formation, this era is governed by the Jeans-length

$$\lambda_J = c_s \sqrt{\frac{\pi}{G\rho}} \tag{3.1}$$

λ_J increases with the expansion of the universe and reaches a maximum when matter and radiation decouples. The grow of perturbations is not important during this epoch.

3.3 Linear Growth Epoch

The radiation dominance era ends and matter takes over. An important event during this epoch is the recombination [4] of electrons and protons into actual atoms about 379.000 yr after the Big Bang. Around this time, photons decoupled from matter forming the Cosmic Radiation Background (CMB) and baryonic perturbation can start to grow.

---

[2] This is a direct consequence of the Heisenberg uncertain principle, particularly as consequence of the fact that the number-particle operator does not commute with the Energy operator, i.e. $[\hat{N}_i, \hat{H}] \neq 0$

[3] Cosmic neutrino Background, observed by the very first time in 2005 by WMAP yielding strong support to the Big Bang model.

[4] Actually, this is a misnomer since there wasn't a previous *combination* before the so-called *recombination.*

Let's consider a volume V large enough that we can assume its homogeneity. Let's consider gravitationally unstable fluctuations in the density of baryonic and non-baryonic matter $\rho(\mathbf{x})$ where $\mathbf{x}$ is a comoving coordinate, so that the physical distance is actually the comoving distance $xa(t)$ where $a(t)$ is the scale factor. We can define the dimensionless overdensity as:

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x})}{\rho_0} - 1 \tag{3.2}$$

where ρ_0 is the average matter density over the volume V. As we have no idea about the values of $\delta(\mathbf{x})$ in the past, our better assumption is it to be a random variable, so the function δ defines a random field. Moreover, we require that $\delta(\mathbf{x})$ to be a Gaussian Random Field, so that the sum of large enough amount of independent random variables has a Gaussian distribution.[5] By consequence, any linear function of the values taken by the overdensity at several positions will have a Gaussian distribution too.

For linear regime $\delta(\mathbf{x}) \ll 1$, futhermore we consider the cosmic material to be a fluid with density ρ, pressure p and potential Φ governed by the following set of equations:

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\mathbf{v} = \frac{1}{\rho}\nabla p - \nabla\Phi \qquad \text{Euler's Equation}$$

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0 \qquad \text{Continuity Equation}$$

$$\triangle\Phi = 4\pi G\rho \qquad \text{Poisson Equation}$$

Now, we introduce a first order perturbation such as $\rho{=}\rho_0+\delta\rho$ in order to linearize the equations of motion and, for simplicity, we define the fractional density perturbation as $\delta = \frac{\delta\rho}{\rho_0}$

Changing our equations of motion into comoving coordinates

$$\mathbf{x}(t) = a(t)\mathbf{r}(t)$$
$$\delta\mathbf{v}(t) = a(t)\mathbf{u}(t)$$
$$\nabla_x = \frac{1}{a}\nabla_r$$

we now write the linearized equations for conservation of energy and momentum

$$\dot{\mathbf{u}} + 2\frac{\dot{a}}{a}\mathbf{u} = \frac{\nabla\delta\Phi}{a^2} - \frac{\nabla\delta p}{\rho_0}$$
$$\dot{\delta} = -\nabla\cdot\mathbf{u}$$

Closing the system, employing an equation of state for gravitational amplification of density perturbations we yield the Jeans equation[6]:

[5] See Central Limit Theorem.

[6] Although in the original Jeans paper it was considered a static fluid instead of an expanding one.

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\dot{\delta} = \left(4\pi G\rho_0 - \frac{c_s k^2}{a^2}\right)\delta \tag{3.3}$$

Where ρ_0 is the undisturbed value of the density whose disturbed values is described by δ. It can be observed that if the pressure term resists compression, i.e. if $\frac{c_s k^2}{a^2} > 4\pi G\rho_0$, we have oscillating solutions; these solutions take the following form

$$\delta = e^{i\mathbf{k}\cdot\mathbf{r}-\omega t}. \tag{3.4}$$

Where $\omega^2 = \frac{c_s k^2}{a^2} - 4\,\pi G\rho_0 > 0$. The oscillations are damped by a factor of $2\,\frac{\dot{a}}{a}\,\dot{\delta}$, so their amplitude decreases with time. This solution is a sub-Jeans scale solution, where perturbations do not grow [7] if the gravitational term dominates; instead, if $4\,\pi\,G\,\rho_0 > \frac{c_s k^2}{a^2}$, perturbations can grow. The solutions for equation (3.3), assuming that the universe is non-relativistic matter-dominated, are linear combinations of the growing and decaying modes. In time, only the growing modes survive and the solutions of Jeans equation behave as follows:

$$\delta \propto t^{2/3}. \tag{3.5}$$

The above hypothesized scenario holds for great part of the universe's history ($3100 \geq z \geq 0.5$) because vacuum energy gets an important role only recently in the universe's history.

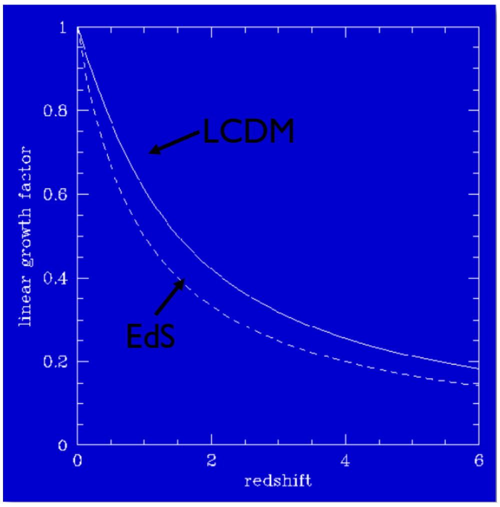

Figure 1 – The graph shows the time evolution of linear perturbation for an Einstein-deSitter universe and a ΛCDM universe. We can see that the later can be approximated fairly well as the former for much of the universe's history.

[7]We already encounter this kind of solution in the radiative epoch of the universe.

3.3.1 The Origin of Angular Momentum

In this section we are going to treat the origin of angular momentum of dark matter halos, using the kinematic approach used by Zel'dovich in 1970. Consider a displacement of a particle

$$\mathbf{x}(t) = \mathbf{q} - D(t)f(\mathbf{q}) \tag{3.6}$$

Where $\mathbf{x}$ is an Eulerian position, $\mathbf{q}$ is a Lagrangian position, $D(t)$ is a function scaling the time-independent displacement field $f(\mathbf{q})$.

To get the Eulerian density, use the Jacobian of the transformation between $\mathbf{x}$ and $\mathbf{q}$, in which frame ρ is constant:

$$\frac{\rho}{\rho_0} = \left[\Big(1 - D\alpha\Big)\Big(1 - D\beta\Big)\Big(1 - D\gamma\Big)\right]^{-1} \tag{3.7}$$

Where α, β and γ are the eigenvalues of the deformation tensor $T_{ij} = \frac{\partial f_i}{\partial \mathbf{q}_i}$. The morphology of the formed structure is determined by the sign of this eigenvalues. We can distinguish the following three cases:

$$\begin{cases} \text{If } (\alpha,\, \beta,\, \gamma) < 0 \text{ Void} \\ \text{If } \alpha > 0 \text{ Sheet} \\ \text{If } (\alpha,\, \beta) > 0 \text{ Filament} \\ \text{If } (\alpha,\, \beta,\, \gamma) > 0 \text{ Halo} \end{cases}$$

Since we assumed that the overdensity field $\delta(\mathbf{x})$ is a Gaussian random field, the values of $\delta > 0$ that enclose overdense regions are statistically indistinguishable for the negative values of δ in underdense regions. Let's imagine an almost spherical underdense region. The recession of matter from the center is slowed by gravity, in any case the underdense region expands more rapidly than the universe itself; this growth of underdense regions forms the **voids**. As time goes, different voids collide with each other shepherding matter into **sheets**. When the sheets surrounding three voids meet, a **filament** is formed. The points where more filaments intersect are call **nodes**.

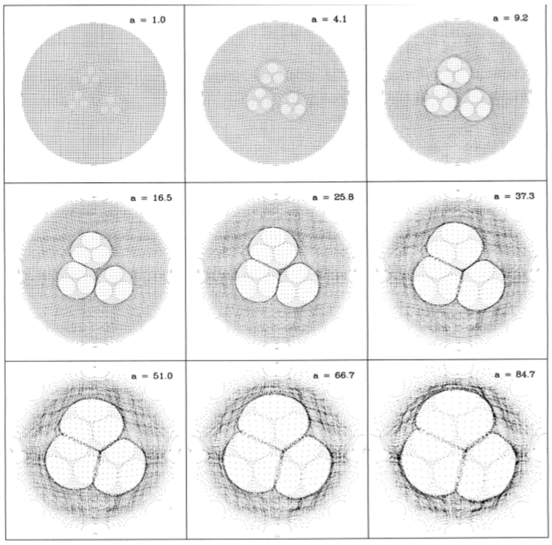

Figure 2 – Dubuski et al. 1993. In a background spherical expansion we place three regions of slightly reduced density and inside them three other regions with even lower density. the low density spheres expands faster that the backgrounds so they run into each other and merge. In the last frame (bottom right) we can see a web of high density regions that traces the original location of low density spheres.

3.3.2 The Tidal Torque Theory

The basic principle behind this formalism is that tidal torques exerted onto proto-halos by neighboring fluctuations.

The total angular momentum $\mathbf{L}\ (t)$ of a proto-halo with volume V at a certain time t is given by:

$$\mathbf{L}(t) = \int_V \rho(\mathbf{r}, t)\Big[\mathbf{r}(t) - \mathbf{r}_{com}(t)\Big] \times \mathbf{v}(t) dV \tag{3.8}$$

writing equation (3.8) in comoving coordinates

$$\mathbf{L}(t) = -\rho_0 a^5 \dot{D} \int_V (\mathbf{q} - \bar{\mathbf{q}}) \times \nabla\phi dV \tag{3.9}$$

The torque in a tidal field can be calculate as follows:

$$J_i = -a^2 \dot{D} \epsilon_{ijk} T_{jl} I_{lk} \tag{3.10}$$

Where T_{jl} is once again the Zel'dovich deformation tensor and I_{lk} is the inertia tensor. In our scenario, where the universe is dominated by matter $a^2\dot{D} \propto a^2\dot{a} \propto t$ angular momentum grows linearly with time, therefore most of the angular momentum of halos is acquired in the linear grow epoch of the universe. It is often to parametrized the angular momentum by the dimensionless spin parameter defined as follows:

$$\lambda = \frac{J\sqrt{E}}{GM^{5/2}} \tag{3.11}$$

Halos have an universal distribution, peaking at $\bar{\lambda} \propto 0.035$.

3.4 Non-Linear Growth Epoch

During this epoch perturbations have grown until $\delta \propto 1$. If we assume perfect spherical symmetry, the equation of motion for a matter-dominated universe is:

$$\frac{d^2\mathbf{r}}{d^2t} = \frac{-GM}{r^2} \tag{3.12}$$

The same equation of motion that would describe the motion of a projectile launched vertically from the surface of a spherical object of mass M.

We have to remind that the we are interested in the perturbation that collapses to eventually form a galaxy, so we have to assume that the projectile does not have enough energy budget to escape.

We can now define an important quantity, the **turnaround radius**, the radius at which the collapse starts.

$$r_{max} = 2a, \text{This occcure at } \eta = \pi \tag{3.13}$$

from this expression it can be derived the **turnaround time**

$$t_{max} = \pi\sqrt{\frac{a^3}{GM}} = \pi\Big(\frac{243}{2000}\Big)^{1/2}\frac{t_i}{\delta_i^{3/2}} = 1.095\frac{t_i}{\delta_i^{3/2}} \tag{3.14}$$

We can see from equation (3.14), that the larger the perturbations, the sooner the collapse will start. In our model, at $\eta = 2\pi$, that is at $t = 2\ t_{max}$, dark matter will settle into a stable configuration that we call a **halo**. This dark matter halo, when complete, satisfy the equilibrium condition, i.e. the viral theorem [8], thats why this process is sometimes called **viralization**.

We can use the virial theorem to define an useful quantity, the **virial radius** r_{200}, as the radius at which $\rho = 200\ \rho_{crit}$. [9]

Inside r_{200} the halo is considered to be in virial equilibrium and the mass inside r_{200} is considered to be the total mass of the halo. It's important to notice that DM halos have an almost universal density profile, called NFW profiles, independent of

[8] $V = -2K$.

[9] This is a rather arbitrary approximation. Upon viralization, the DM halo became ≈ 178 times denser than the background universe.

halo mass, initial density fluctuations spectrum or cosmological parameters, described pretty accurately by the following formula:

$$\rho(r) = \frac{\delta_c}{(r/r_s)^\alpha (1 + r/r_s)^{3+\alpha}} \rho_{crit} \tag{3.15}$$

Where δ_c is a characteristic density and r_s is the scale length and α is the slope parameter; recent research situates this parameter between 1 and 1.5.

The Press-Schechter theory predicts the following dark matter halo mass function:

$$\frac{dn}{dM}(t) = \sqrt{\frac{2}{\pi}} \frac{\rho}{M^2} \frac{\delta_c(t)}{\sigma(M)} \frac{d\log\sigma}{d\log M} e^{\frac{-\delta_c^2}{2\sigma^2(M)}} \tag{3.16}$$

Until 1990's most numerical simulations matched the **PS mass function** but when high-resolution simulations became available, it was clear that PS-theory overpredicts sub-characteristic mass halos $M < M^*$ and underpredicts $M > M^*$ halos.

DM halos grow hierarchically through mergers. One can follow the history of this mergers and form what is called a merger tree, as it is shown in figure 3. This is considered to be the backbone of the hierarchical formation theory.

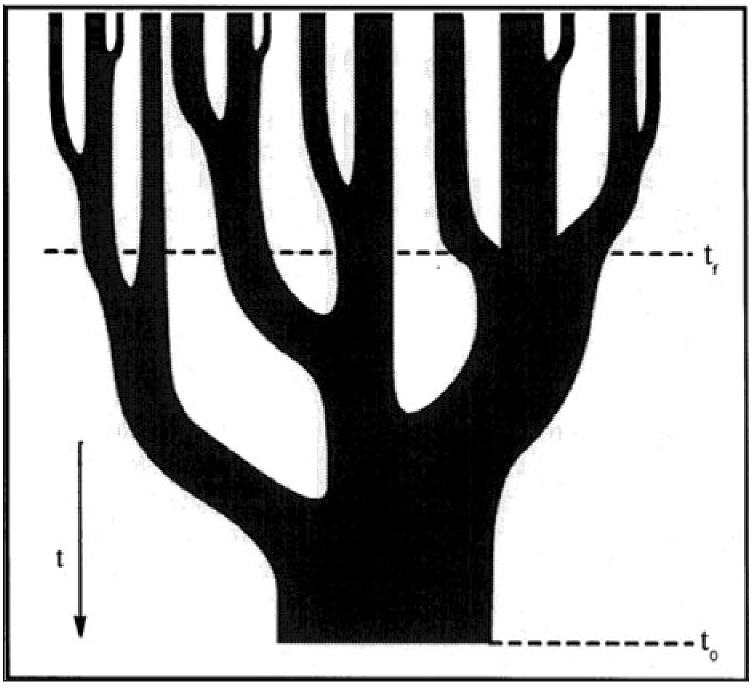

Figure 3 – Lancey & cole 1993. Time increases from top to bottom and the widths of the branches represents the masses of the halos. At t_f we see the progenitors of the final DM halo at t_0

3.5 Galaxy Formation Epoch

In the previous sections we've studied how dark matter halos form, now this already formed DM halos act as potential wells, for luminous matter to fall into, forming what is commonly known as a galaxy. During this epoch, the baryons have cool down enough to form luminous objects. This cooling process depends mainly on gas density, gas composition and of course temperature. There is a simple criteria that we can use to know when matter inside the DM halo has cooled down enough to start star formation. The cooling is mainly due to bremsstrahlung, recombination and collisionally-excited line emission. If we define the cooling rate due to all the factors above mentioned as $\Lambda(T)$, then the cooling time is :

$$t_{cool} = \frac{3kTm}{\rho\Lambda(T)} \tag{3.17}$$

Another important quantity in the process of star formation in galaxies is the dynamical time t_{dyn}, defined as the time that it would take to form a star if the pressure supporting it against gravity were suddenly remove.

If $t_{cool} < t_{dyn}$ and $t_{cool} \leq H^{-1}$ gas can fragment into stars.

In the White & Rees model, the minimum DM halo mass to host a luminous galaxy can be yield from the following expression:

$$\bar{M_b} = \frac{f_b M}{[1 + (2^{1/3} - 1)M_c/M]^3} \tag{3.18}$$

Where $\bar{M_b}$ is the mean baryonic mass, M is the halo mass, M_c is the characteristic mass below which a halo cannot host a luminous galaxy and $f_b = \frac{\Omega_b}{\Omega_0}$ is the cosmic baryonic fraction. Using expression (3.18) we can establish mass limits to DM halos. The minimum value of M_c in order to host a galaxy is $\approx 10^9\ M_\odot$. The upper limit for which a DM halo can host and individual galaxy is $\approx 10^{13}\ M_\odot$, DM halos with higher mass host groups and clusters of galaxies.

In the early universe heavy elements weren't present, so the cooling process was limited to H_2 formation, with this material, the first generation of stars were formed. Stars that were metal-free and lived shorts lives, generally ending in black hole or in SNe. It is believed that such kind of star population, called population III, were the seeds for the supermassive black holes in galactic centers and they started the reionization of the universe. This kind of galaxies, with early stars population are called **primeval galaxies**.

As population III stars die, some of them in a supernova event, they enrich the interstellar medium (ISM) making the next generation of stars more chemically complex. As old a galaxy is, as rich of heavy elements it will be. This richness of heavy elements or metals [10] in a star, is called metallicity, and is the total content of iron in a given star.

$$[Fe/H] = \log\left(\frac{N_{Fe}}{N_H}\right)_{Star} - \log\left(\frac{N_{Fe}}{N_H}\right)_{Sun} \tag{3.19}$$

The reason why we use the iron fraction, even if it isn't the most abundant metal in stars (oxygen is), is because it is easier to measure thought spectroscopy. As we already said, based on metallicity we can characterize star populations. Population III stars are extremely metal-poor , with mean iron abundances of 0.1%. Population II stars are metal-poor, with mean iron abundances of 1% - 1.6%, this are old stars that can be found mainly on the halo of the galaxy. Lastly we have population I stars, young stars with metallicity's from 2% up to 3%. Pop. I stars are likely to be found in the center and the disk of the galaxy. The distribution of stars populations in a galaxy is useful to characterize the chemical evolution[11] of the stellar system.

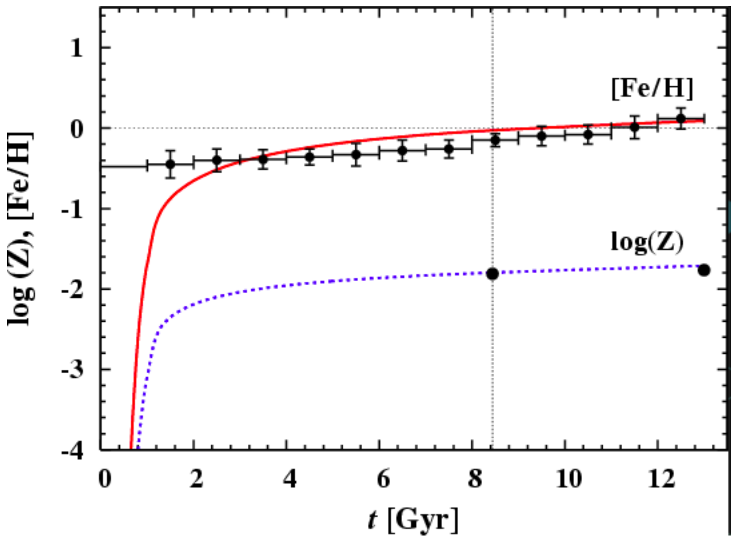

Figure 4 – This figure shows the chemical evolution of a typical galaxy on time

4 Morphology Of Galaxies

There are two type of galaxies if we categorize them based on its morphology, **ellipticals** and **spirals**[12]

[10]In astrophysics, a metal is any element heavier than helium

[11]This is a misnomer because the metal enrichment of the ISM is produced by nuclear reactions within stars and not by chemical reactions

[12]Of course, there are also irregular galaxies, but we are not going to treat them in this thesis

Even up-to-day it is not entirely clear how this morphologic differences arise between galaxies. In this section we are going to explore the state-of-the-art theories and experimental data that could lead us to a possible answer. It's worth noticing that the differences between ellipticals and spirals go further than the merely geometrical aspect, they differ in star velocity dispersion, star formation ratio (SFR) and age, just to name a few of the main physical differences between this two type of stellar systems.

4.1 Elliptical Galaxies

When it comes to ellipticals, historically, there are two schools of thought. At first it was believed that elliptical galaxies were formed by gas clouds that did not have a particularly strong amount of angular momentum, constructing a more round structure, without any axis of rotation .

The 1972 Toomre & Toomre paper was a turning point in the elliptical galaxy formation theory, stating that ellipticals were actually formed though mergers of spiral galaxies. A couple of very interesting facts support this theory, first of all, ellipticals are more likely to be found in rich galaxy clusters, where collisions are more frequent. Another interesting fact that should caught our attention is that the SFR in ellipticals today is very low, the star population founded in ellipticals are mainly Pop. II stars, i.e. red and old stars leading to the characteristic reddish color that this type of galaxies presents. Furthermore, ellipticals are gas-poor galaxies, this leaded us to the conclusion that there were an initial star formation burst due to the fact that collisions ignited almost all of the gas within the galaxy, turning it into stars.

4.2 Spiral Galaxies

When it comes to spirals, there are two main structures that we can identify, the bulge, and the spiral arms. The bulge is a central concentrate of Pop. II stars, very gas-poor for which SFR is very low in this part of the galaxy. In the other hand we have the spiral arms, sites with a high SFR, thus young blue Pop. I stars and consequently an important amount of gas and dust.

How this spiral disks formed is a rather complicate and ongoing field of research, however we are going to do our best to describe it briefly and accurately. Peter Goldreich and Donald Lynden-Bell from Cambridge and William Jullian and Alar Toomre from MIT independently discovered that a self-gravitating differentially rotating fluid disk can amplify initial perturbations by a factor of 10 or more, inducing a spiral-shaped wave in the stars of the disk. The conclusion is obvious, spiral patterns can be cause by a wide variety of phenomena like: tidal forces from companion or satellite galaxies, clumps of matter in the interstellar medium or sub-halo structure. The formalism that is used to study differentially rotating disks is called **density-wave theory**, inside this formalism, spiral pattern are density waves that, in the inner part of the galaxy, stars are moving faster than the spiral patter, so they overtake the density wave, in the outer part of the galaxy instead, the stars are moving slower that the spiral pattern, so the spiral arms overtake the stars. As gas clouds moves into the density wave, the local gas mass density increases, making more probable their collapse into stars, furthermore, as clouds get swept by spiral arms, they collide with each other, giving rise to shockwaves which heated the gas, making star-formation even more favorable.

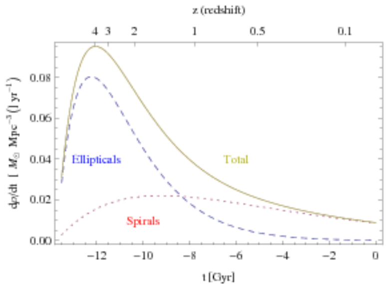

Figure 5 – The image shows the different star formation rate for elliptical galaxies and spiral galaxies

5 Conclusion and Open Questions

As it was our propose along this thesis, we studied the hierarchical model for structure formation in the universe, confronting some of the main theoretical results with recent experimental data such as observations from space-based and ground-base telescopes and high resolution simulations. today, more than ever we have a clearer picture of how structure and galaxies formed in the universe, nevertheless, at the same time, we know that there is a lot to be answered yet. What's the actual nature of dark energy? And of dark matter? Beyond our reasonable assumptions, what is the actual extension and shape of the DM halo? Why in ΛCDM simulations we cannot reproduce realistic disk galaxies? The further research that is necessary to bring enlightenment to this questions will be in the hands and minds of the next generation of physicists and astrophysicists, categories in which I'm glad to consider myself now.

References

[1] White & Rees, *Hierarchical structure formation. 1978*

[2] Press & Schechter, *Formation of Galaxies and Cluster of Galaxies By Self-Similar Gravitational Condensation. 1974*

[3] Efstathiou & Jones, *The rotation of galaxies: the numerical investigations of the tidal torque theory. 1978*

[4] Steinmetz & Bartelmann *On the spin parameter of dark-matter haloes. 1995.*

[5] Kamionkowski, *Possible Relics for new physics in the early universe: Inflation, the cosmic microwave background and particle dark matter. 1998*

[6] Hinshaw et. al, *Five-year Willkinson microwave Anisotropy Probe (WMAP) Observations: Sky maps and basic results. 2009*

[7] Carraro et. al, *On the galactic disk age-metallicity relation. 2008*

[8] James Binney & Scott Tremaine, *Galactic Dynamics (2nd Edition, Princeton University Press). 2008*

[9] Jonh A. Peacock, *Cosmological Physics. 1999*

[10] Houjun Mo, Frank van den Bosch & Simon White, *Galaxy formation and evolution (Cambridge University Press). 2010*