Study of stellar populations in nearby galaxies using integrated spectroscopy

Vera Alexandra Fonseca Patrício

Thesis to obtain the Master of Science Degree in

Engineering Physics

Supervisors

Supervisor:	Prof.ª Ana Maria Vergueiro Monteiro Cidade Mourão
Co-supervisor:	Dr.ª Myriam Arnal Rodrigues

Examination Committee

Chairperson:	Prof.ª Maria Teresa Haderer de la Peña Stadler
Members of the Committee:	Prof.ª Ana Maria Vergueiro Monteiro Cidade Mourão
	Dr. Hector Alejandro Flores Figueroa

February 2014

Resumo

O estudo do conteúdo estelar das galáxias pode fornecer pistas importantes para a completa compreensão da evolução galáctica. Exceptuando algumas dezenas de galáxias próximas, nenhuma galáxia pode ser resolvida em estrelas individuais e apenas é possível estimar o seu conteúdo estelar através da sua luz integrada. Um método que permite fazer isto é o ajuste ao espectral com modelos de populações estelares, método este que é frequentemente usado em estudos extragalácticos. No entanto, a sua validade ainda não foi convenientemente estabelecida para galáxias, que são muito mais complexas em termos de conteúdo, história e dinâmica que os objectos - aglomerados de estrelas e espectros sintéticos - usados até à data neste tipo de testes. Este trabalho apresenta os primeiros resultados do projecto Starfish - STellar Population From Integrated Spectrum - cujo objectivo é estabelecer a validade deste método usando pela primeira vez espectros reais de galáxias. Neste trabalho foi investigada a exactidão das populações estelares simples derivadas num amostra de 5 galáxias anãs. O ajuste foi feito com o algoritmo STARLIGHT, que encontra a combinação de espectros de populações estelares que melhor reproduzem um espectro observado. Foi possível encontrar populações estelares com menos de 10 Myr e as populações mais velhas encontradas, signiticativas em termos de massa total, situavam-se por volta dos 1-2 Gyr. Conclui-se também que este método subestima a massa criada entre 1 e 10 Gyr e sobrestima no intervalo 300 Myr a 1 Gyr.

Orientado por Myriam Rodrigues, ESO fellow, e Ana Mourão, investigadora do CENTRA.

Palavras-chave: Galáxias, Espectro integrado, Populações estelares, Ajuste espectral completo

Abstract

The study of the stellar content of galaxies may provide important clues to fully understand galaxy evolution. Except for a few dozens of nearby galaxies, galaxies are not resolved into individual stars and it is only possible to probe their stellar content through their integrated light. One of the methods to do this is to fit the optical spectrum with stellar population models, a method is commonly used in extragalactic studies. Despite this, this method has still not been conveniently validated for galaxies studies, that are far more complex in terms of content, history and dynamic than test objects - star clusters and synthetic spectra - used so far. This work presents the first results of the Starfish - STellar Population From Integrated Spectrum - project which has the goal of determining the accuracy of the derived physical properties using for the first time real galaxies' spectra. I investigated the accuracy of the simple stellar populations retrieved with this method in a sample of 5 nearby dwarf galaxies. Full spectral fitting was performed with STARLIGHT code, finding the best combination of spectra of simple stellar populations that reproduce the observed spectra. Known stellar populations with less then $\sim$10 Myr were recovered and the oldest mass significant populations were typically around 1-2 Gyr old. The method underestimated the mass contribution in the 1 - 10 Gyr interval and overestimated the created mass between 300 Myr - 1 Gyr.

Under supervision of Myriam Rogrigues (ESO-Chile) and Ana Mourão (CENTRA).

Keywords: Galaxies, Full spectral fitting, Integrated spectroscopy, Stellar populations

Contents

List of Tables

List of Figures

Glossary

HR, Hertzsprung-Russel diagram.

CMD, Colour-Magnitude Diagram. Equivalent to Colour-Magnitude diagram.

MS, Main Sequence. The stage of a star life when it is fulled by hydrogen burning.

Z, Metallicity. Mass fraction of elements heavier than Helium.

SSP, Simple Stellar Population. Ensamble of stars with the same age and composition.

SFH, Star Formation History. Star formation rate throughout time.

IMF, Initial Mass Function. Mass distribution of stars created in a burst.

HST, Hubble Space Telescope.

Starfish, STellar PopulAtion From Integrated Spectra.

GMOS, Gemini Multi-Object Spectrograph.

ANGST, ACS Nearby Galaxy Survey Treasury. Team in collaborarion in the Starfish project (responsible for the CMD studies).

Chapter 1

Introduction

1.1 Context: Galaxy Formation

Galaxies are the visible imprints of the large scale structure of the Universe and the harbors of its small inhabitants, the stars. The study of these cosmic islands is fairly recent: it was only in 1924 that Edwin Hubble was able to determine that Andromeda is 2 million light years away, placing it outside the Milky Way. With further observations of several other nebulae, that for the first time were recognized as other galaxies, it became possible to create a classification scheme based on their morphology. Galaxies were grouped into three major types: *ellipticals*, with a regular ellipsoid shape, *spirals*, arranged in a disk with arms and an inner bulge, and *irregulars*. Hubble also proposed a classification scheme based on the complexity of the observed morphology. Simpler galaxies - elliptical - were classified as early type galaxies, and more complex ones (spirals) as late type, proposing diagram known as the Hubble Fork (see figure 1.1).

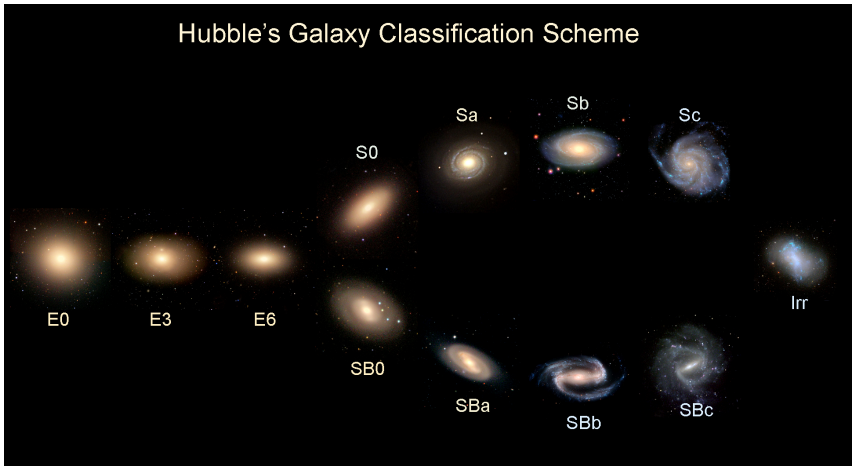


Figure 1.1: Hubble Turning Fork constructed with Sloan Digital Sky Survey pictures - John Kormendy. *E - Eliptical Galaxies; S - Spirals; SB - Spirals with a bar; Irr - Irregulars.* The associated number classify the galaxies in subclasses, related with the eccentricity or the tightness of the arms.

The modern understanding of galaxy formation and evolution comes both from theoretical predic-

tions and observational facts. Cosmological models describe an early Universe with small mass density fluctuations in an otherwise uniform distribution and where cold dark matter dominated over visible matter. This view is backed up by observations of the cosmic microwave background, a relic from the recombination era, that shows a high uniformity in the matter distribution with some small anisotropies. During the expansion of the Universe denser regions grew through gravitational accretion, a process that originated dark matter halos, that are thought to be the seeds of galaxies [Mo et al., 2010].

The actual challenge is to understand how dark matter halos were populated with visible matter. There are currently two alternative theories on how baryonic matter clumped in the halos: the *hierarchical* and the *secular* scenarios. In the hierarchical view the accumulation of visible matter started in small and relatively low density halos that continued their growth by merging with other halos, dragging along the visible matter that also formed bigger structures. The galactic zoo that is visible today would therefore be explained by each particular merging history. On the other side, the secular view argues that baryonic matter piled up preferentially in the higher density halos that continued to grow though accretion of dark matter and consequently visible matter too. The morphological differences in galaxies would be explained by the initial size of the gas cloud as well as by the peculiarities of the accretion process.

It is known both from observations and simulations that the formation of elliptical galaxies is consistent with the merging scenario [Toomre and Toomre, 1972; Toomre, 1977]. But in what disk galaxies are concerned, the most common type of intermediate mass galaxies accounting for roughly 70% of the total galaxy population in the local universe, there is still no final answer. The process of forming a disk is well explained by the secular scenario: given an initial cloud of gas, with enough angular momentum (that can be inherited from the dark halo), a disk will be formed and will grow through accretion. However, recent improvements in merging simulations that also include gas influence, demonstrate that mergers with a high gaseous content can also give origin to disk galaxies [Hammer et al., 2005]. It is still unknown whether it is necessary for an initial secular phase in Universe evolution to explain the abundance of disk galaxies or if the merging scenario is sufficient.

Since 1924 until now the study of galaxies has evolved considerably but there are still some fundamental questions that remain unanswered, such as when and how did galaxies assemble their stellar mass, how is the morphological type of a galaxy determined and what triggers star formation during the evolution of galaxies. In this work, I focus my study on the stellar populations of galaxies, that provide valuable clues to these questions.

1.2 The building blocks of galaxies

In 1899, J. Scheiner compared an Andromeda's spectrum with a solar spectrum and found a *'surprising agreement of the two'*. This first spectroscopy study of a galaxy confirmed the ongoing hypothesis that these objects were in fact composed by stars.

1.2.1 Stars

A few years earlier, in the late 19th century, the study of individual stars was already being developed. In 1897 Antonia Maury organized stars in increasing strenght of a number of absorption lines resulting in the O B A F G K M sequence, that later on Cecilia Payne proved to correspond to a sequence of decreasing surface temperature. The classification eventually extended into star subtypes, from 0 to 9, in order of decreasing temperature and to luminosity strength, from type I, the supergiants, to type VI, sub-dwarf stars. Spectra of different types of stars, following that classification, are displayed in figure 1.2.

The spectrum of stars consists of by a continuum, well described by black-body radiation, with absorption lines arising from electronic transitions in the elements in the surface of the star. The star's surface temperature can be inferred from the continuum shape using Wien law. A star's colour and temperature are then connected: hotter stars are bluer and colder stars redder, as it can be seen in figure 1.2. A description of the spectroscopic characteristics as a function of temperature in the different types of stars is given bellow.

- **Type O (T > 30 000 K)** Temperature is so high that all hydrogen is ionized and transitions are not likelly to occur. As a consequence, hydrogen absoption lines are very week and the spectra is dominated by He II (single ionized helium) lines. Most of the flux is being emitted in the blue wavelengths.

- **Type A (T < 11 000 K)** Temperature is still sufficiently high to ionize most of the metals, whose lines are not clearly visible. On the other hand, the star is cool enough to allow the presence of atomic hydrogen but there is still sufficient thermal energy for ionization to take place, resulting on clearly visible and strong hydrogen lines.

- **Type G (T < 6000 K)** Temperature has drooped to a point where hydrogen ionization is not very likely to occur. Neutral atoms are present and lines from elements such as Ca II (from which lines K,H and G band) appear.

- **Type K (T < 5000 K)** Atoms are now mainly in their neutral form. The cumulative effect of absorption lines of several metals before the 4000 Å decreases most of the stars' flux, which causes a feature called the 4000 Å break, that is also seen in evolved galaxies.

- **Type M (T < 4000 K)** The temperature on these stars atmosphere is cold enough to allow the presence of not only of neutral atoms but also molecules such as TiO and VO, which lines form the most prominent features. Most of the flux is emitted in the redder region of the spectra.

Although different metals - defined here as all elements heavier than helium - have distinct spectral features, they are generally treated as a single quantity, the metallicity (Z). Metallicity is defined as the mass fraction of elements heavier then helium. Because it is difficult to measure each element contribution, mettalicity is estimated by the quantity of some specific metals (Fe or O, for example). Usually metallicities are given as fractions of the solar metal content.

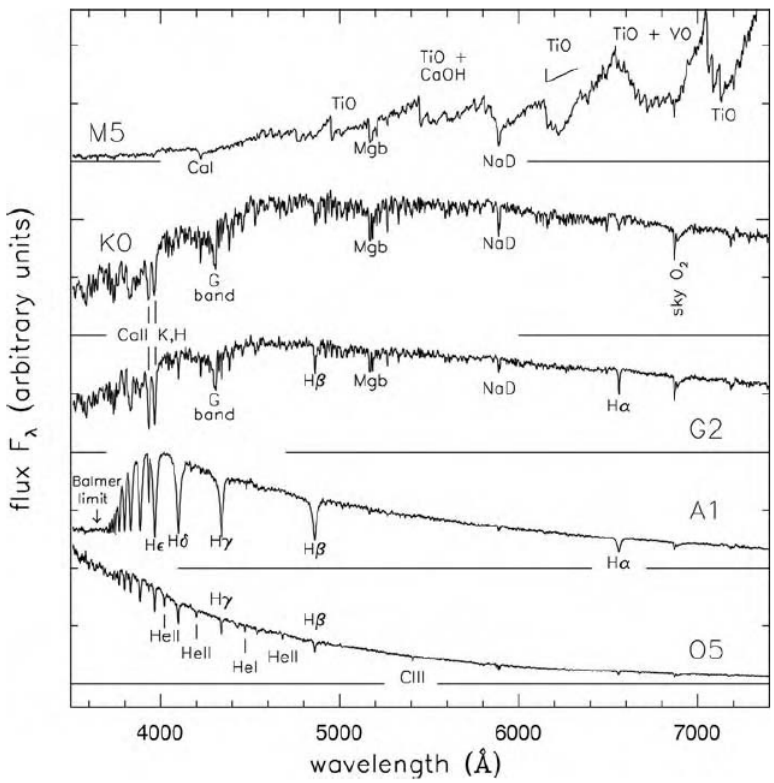


Figure 1.2: Examples of spectra of main sequence stars (increasing surface temperature from top to bottom) - from *G. Jacoby et al., spectral library*

Much of a star's properties can also be derived from photometry, from its colour, that is determined by its surface temperature, and its luminosity. Figure 1.3 shows a HR diagram of nearby stars, where absolute luminosity is ploted against the temperature (or, equivalentely, colour) of a star [Sparke and Gallagher, 2007]. Observing this diagram for a random sample of stars it is obvious that they do not span all possible temperatures and luminosities. The reason is that indeed not all combinations are physically possible. What is more, there are some regions in the HR diagram where the stars number density is higher: a main diagonal line that crosses all diagram, called the main-sequence, upper branches where bright but cold stars can be found, and a lower branch where relatively hot and dim stars are placed. This distribution is not particular to nearby stars and is a reflection of the physical mechanisms taking place inside stars. Stellar atmosphere physics describes these mechanisms and allows the prediction of a star's luminosity and surface temperature throughout its life, given its initial mass. This evolution can be plotted in an colour-magnitude diagram (CMD) for a single star, forming a path in the diagram that is called an evolutionary track. Figure 1.4 shows a few evolutionary tracks predicted by stellar models for main sequence stars with different masses.

A star is born from a dense molecular gas cloud sufficiently dense to ignite hydrogen fusion at its core. During most of its life a star will be fueled by the nuclear fusion of hydrogen into helium, that releases sufficient radiative energy to contra balance the gravitational collapse of the star. This stable and long period corresponds to the main sequence (MS) on the HR diagrams. The time that a star spends in the MS, and its evolutionary path after it, depends mainly on its initial mass. More massive stars will

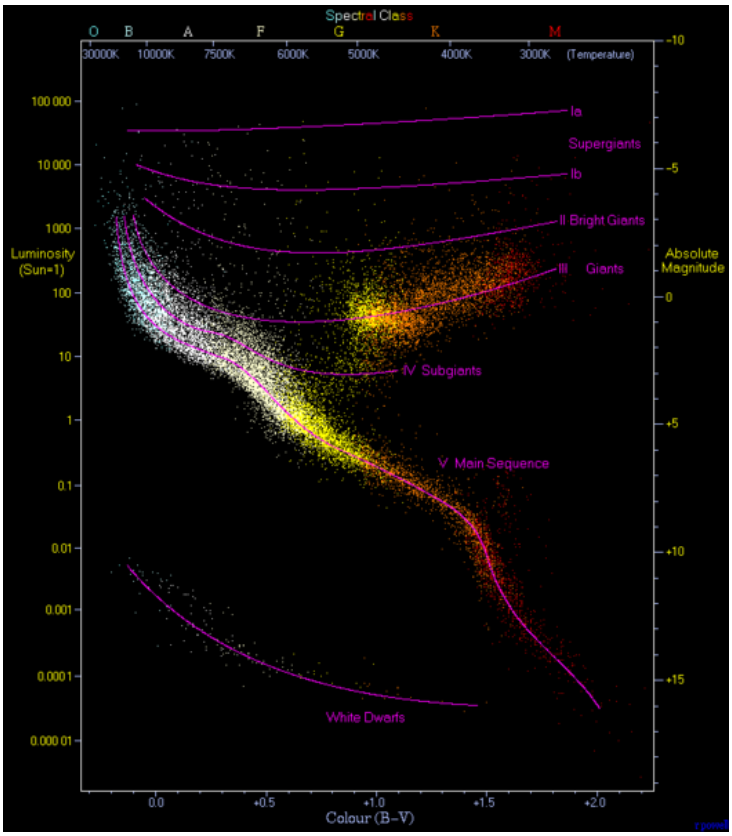


Figure 1.3: HR diagram with 22,000 stars plotted from the Hipparcos catalog and 1,000 from the Gliese catalog of nearby stars - *Galaxy Formation and Evolution*, Mo, van den Bosch and White

have shorter lifes, since their cores are hotter due to their higher gravitational pressure, which increases the rate at which hydrogen is converted into helium. Metals in the star play a secondary role, mainly increasing the opacity of the stellar atmosphere, preventing photons produced in the core of escaping to the surface as easily as in metal poor stars. The small increase of radiation pressure caused by this opacity allows a slower hydrogen burning, increasing the lifetime of stars [Sparke and Gallagher, 2007].

When a star has converted around 10% of its hydrogen into helium it reaches the end of the main sequence. At this point, it becomes necessary to distinguish between stars of different masses. Low mass stars ($M < 8M\odot$) become red giants and form the Giant Branch visible in the HR diagram. After the end of the MS, their cores collapse, increasing core temperature and pressure, while the outer layer expands up to a hundred times the initial radius, becoming cold and red. The nucleosynthesis proceeds in the core, burning helium into heavier elements but at a much faster rate. Each time the radiative energy produced in the nucleus diminishes due to the decrease of the burning element and becomes insufficient to contra balance the gravitational pressure, the star contracts. These contractions increase nuclear temperature and pressure, allowing the nucleosysthesis of successive heavier elements. However, for low mass stars, this process can only sustain a star until carbon and oxygen are created. At this point, the star's mass is not sufficiently high to initiate the next element synthesis. The outer layers have been blown away from the core and the star ends its life as a white dwarf. These are the stars situated in the lower branch of the HR diagram (see figure 1.3). More massive stars ($M > 8M\odot$) are ruled by the same processes but

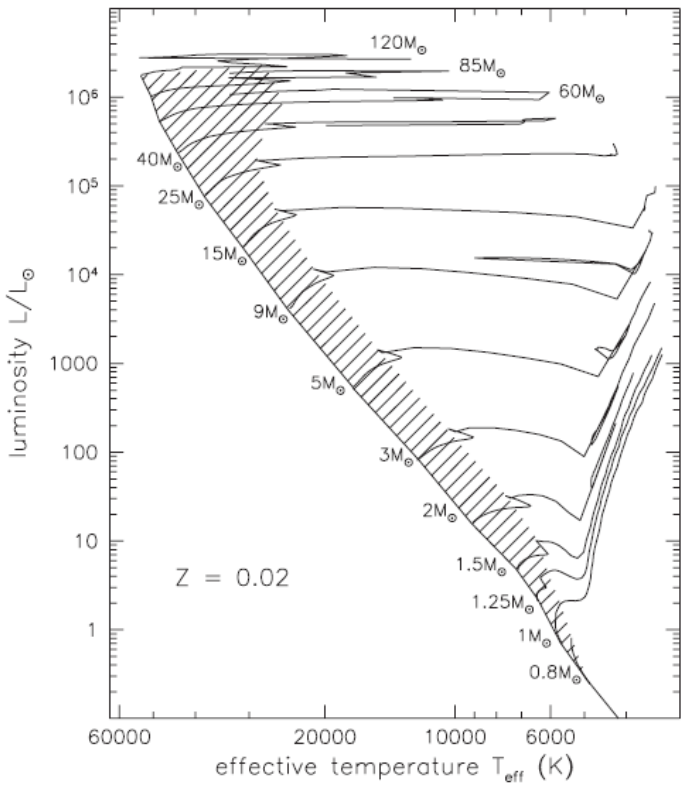


Figure 1.4: Main sequence stars stellar tracks (hatched region corresponds to hydrogen core burning)- Geneva Observatory tracks

their higher masses allow the burning of heavier elements, up to iron. This is the most stable element in nature and thus the synthesis of heavier elements is not energetically favourable. This means that once the nucleosynthesis has reached iron the star will collapse, until the pressure is so high that the core becomes proton degenerate and explodes as a nova or supernova, ending its life as a neutron star or a black hole.

But stars in the Universe are not born alone. Instead, when a dense gas cloud has the right conditions to trigger star formation several thousands of stars will be born simultaneously.

1.2.2 Simple Stellar Population

Stars born at the same time and from the same gas (with the same chemical composition) are what is defined as a Simple Stellar Population (SSP). In nature, star clusters are the most similar objects to this theoretical concept. Figure 1.5 shows a color-magnitude diagram of such a group, the Pleiades, a near open cluster, visible with naked eye.

Stars in this diagram are also distributed along a diagonal line, but the main sequence is shorter and not all the branches of the HR diagram of nearby stars (figure 1.3) are present. This happens because nearby stars can have very different ages, whereas in clusters we are looking at stars with the same age. Their distribution is described by an isochrone - a path in the HR diagram of stars with different masses but equal age and metallicity - which is a 'snapshot' of the position of all the stars in the cluster at a

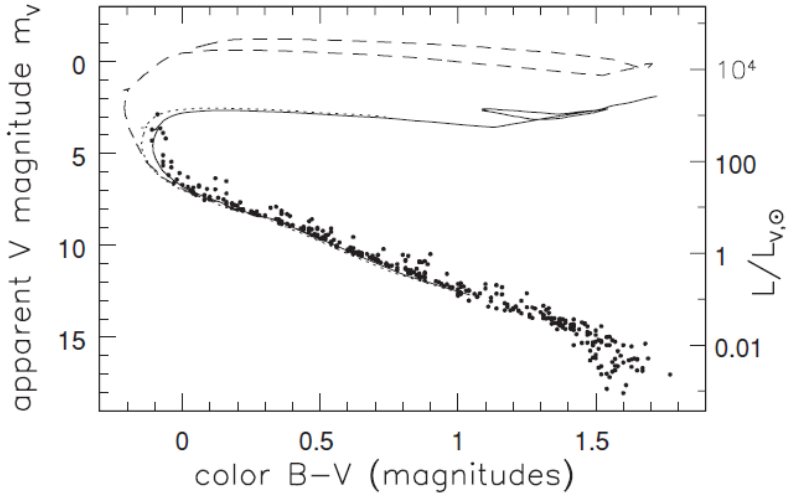


Figure 1.5: HR diagram of the Pleides open cluster. *Solid line*: 100 Myr isochrone. *Dotted line:* 100 Myr isochrone corrected for dust reddening. *dashed line*: isochrone with 16 Myr - J.C.Mermilliod

certain time of their evolutionary tracks, such as the ones shown in figure 1.4. In the diagram an isochrone for stars with 100 Myr is visible in solid line, which is the age of the cluster itself.

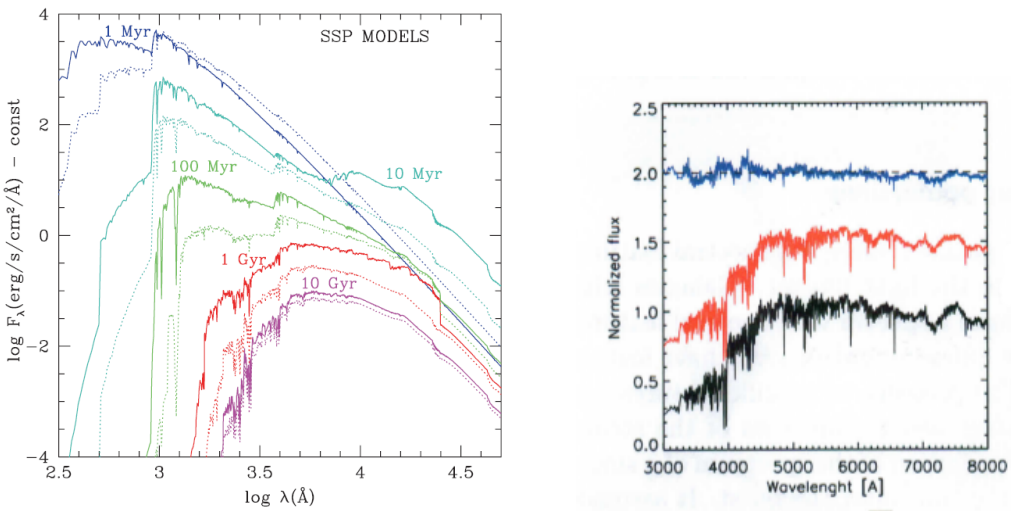


Figure 1.6: *Left:* SSP spectra with solar metallicity from 1 Myr to 13 Gyr - from *Stellar Population, a user guide from low to high redshifts* Laura Greggio and Alvio Renzini. *Right:* Example of age-metallicity degeneracy. In red: Spectrum of a SSP with 2.5 Gyr and solar metallicity. In black: Spectrum of a SSP of 10 Gyr and Z=0.4 Z⊙. In blue: ratio between the two spectra. - by Myriam Rodrigues

The spectrum of a SSP is a mixture of the spectra of the stars within it. In figure 1.6 examples of SSP spectra for populations with the same metallicity and different ages can be seen. Some of the features of individual stars are imprinted in the spectra. Young SSP will have a higher flux in the blue part of the spectra, since hot blue giant stars are still in the main sequence. As time goes by, the flux distribution is flattened because the major contributors to the light of an evolved population are low mass red stars, that survive for several Gyr. Nevertheless, mettalicity also plays a part in determining a star's continuum

shape. Stars that have a high quantity of metals in their atmosphere will have a redder spectrum than the ones with lower metal content, because of the oppacity increase caused by those elements. This effect will be visible in the SSP spectra and a young stellar population that is metal rich may be mistaken by an older population with low metallicity. In figure 1.6 an example of this age-metallicity degeneracy can be seen. This makes the determination of the age of a SSP from its spectrum a complex task, since the age effect on spectra - the continuum redenning and spectral line modifications - is much similar to the effects caused by the increase of metallicity. Specifically, stellar populations that have the same $age \times Z^{3/2}$ combination will have nearly indistinguishable spectra [Worthey, 1999]. The example of figure 1.6 also reveals that SSP may be differentiated by a detailed analysis of a high resolution spectrum, particularly the blue spectral range, where the difference is greater. Yet, predicting the spectrum of an SSP with a determined age and composition is in itself a complex and difficult task.

Evolutionary Population Synthesis

To model a Simple Stellar Population spectrum three main inputs are needed: the Initial Mass Function (IMF), which is the mass distribution of the stars created at the burst, the isochrones, that describe the evolution of stars, and a spectral library, to attribute a spectrum to a star with a determined set of parameters [Conroy, 2013]. Figure 1.7 gives a schematic view of the process.

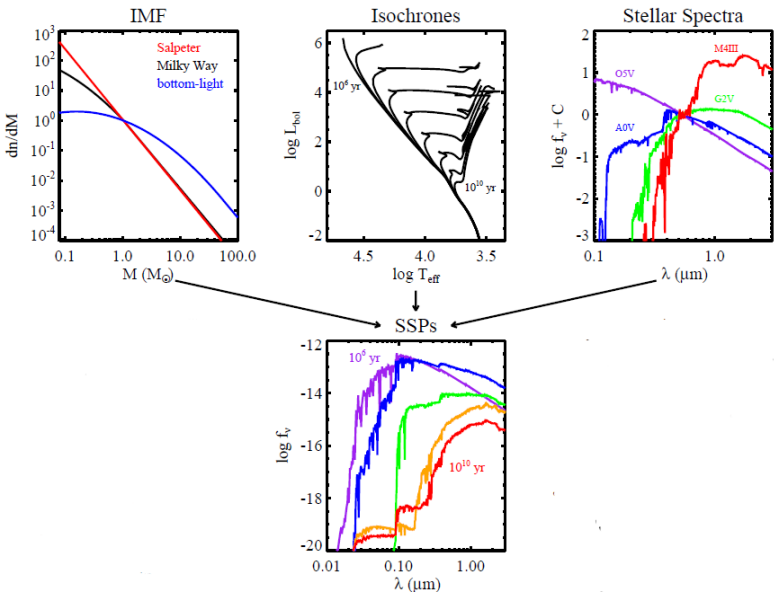


Figure 1.7: Spectra synthesis overview - adapted from [Conroy, 2013]

Initial Mass Function

Firstly it is necessary to define the metallicity of the SSP. Most models use a solar abundance pattern, this is, the relative amount of each metal follows the distribution observed in the sun, although the total quantity of metals is varied. It is also necessary to define the mass distribution of young stars that were created in the burst. This fraction is given by the initial mass function. The mass function was defined formally by Salpeter as:

$$\epsilon(log\ m) = \frac{dN/V}{d\ log\ m} \tag{1.1}$$

where N is the number of stars in the volume V and m the mass. The functional form of this distribution is a current subject of research and there are several proposed models, such as the Salpeter (1995), the Kroupa (2001) or the Chabrier (2003). Figure 1.8 displays some of these models.

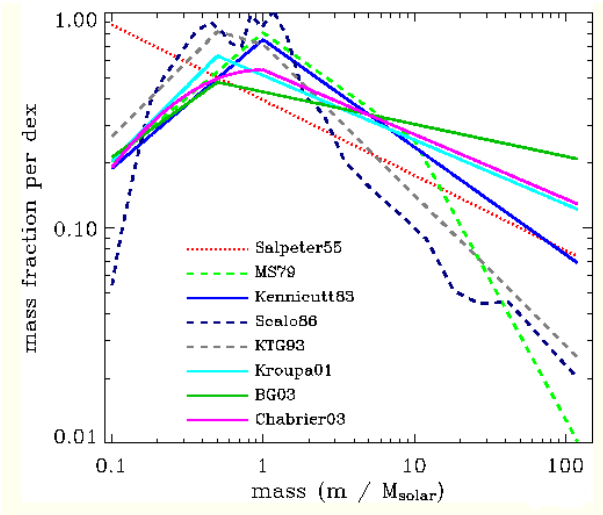


Figure 1.8: Initial Mass Function Models - by Ivan Baldry, 2006

It is quite challenging to derive a mass function mainly because there is no possibility to directly measure it [Chabrier, 2003]. Stellar masses can only be directly determined making use of gravitational interaction, which limits it to near binary systems. So the mass function has to be indirectly derived from the luminosity function - distribution of stars with luminosity L per volume, an analogous to the mass function but for star luminosities - if representative samples are to be used. To convert luminosity into mass functions stellar atmosphere models are needed, which will introduce errors due to the refereed uncertainties. Moreover, we have only access to the current distribution of masses and it is necessary to interpolate the initial distribution.

The several IMF models available (see figure 1.8) are based on empirical observations. The models mostly differ in the massive stars regime ($> 10M\odot$). These stars are short-lived and less numerous than lower mass stars, so the number of observed stars is low and therefore its estimative not accurate. Because these are precisely the stars that most contribute to the integrated spectra, these uncertainties could have a major impact in the results derived from full spectral fitting.

Evolutionary Models

After the ensemble of initial stars has been defined through the choice of an IMF (left upper panel of the Figure 1.7), they are evolved accordingly to their individual evolutionary paths (middle upper panel), making it possible to know in what stage of its evolution a star is and the corresponding luminosity and surface temperature.

Stellar libraries

From the ensemble of evolved stars, a certain number is chosen and matched with a correspondent

spectra taken from a stellar library - a group of spectra with the same resolution and wavelength coverage - that can be composed of empirical or synthetic spectra, or even a mixture of the two

There is a considerable number of empirical star spectra libraries covering different spectral ranges and resolution and with different sampling criteria, from which a desired base can be constructed. As previously mentioned, the available sample of stars is biased. Hot main sequence stars, specially at low metallicities, are rare in the solar neighborhood, and rapid phases of the stellar evolution, such as Thermally-Pulsating Asymptotic Giant Branch (TP-AGB) or Wolf-Rayet stars, are difficult to probe. This constitutes a problem, because although these stars do not contribute much for the total mass, their contribution to the total light is not negligible as they are very luminous.

There are also several synthetic spectral libraries options. Calculating these synthetic spectra is a research field in itself and only a very brief description is made. The ingredients needed to compute a stellar spectra are a model of the stellar atmosphere for the desired star and a list of atomic and molecular lines that come from laboratory measurements, when possible, or from theoretical calculations [Coelho et al., 2005]. In the modelling of stellar atmosphere models numerous approximations, such as calculating 1D or 2D stellar atmospheres models, and assuming local thermal equilibrium, have to be made, which introduces uncertainties in the final product. A major problem of current stellar atmosphere models is how to deal with convection, rotation and mass loss, that is an important factor in giant luminous stars [Conroy, 2013]. Overall, the short lived and bright stars - massive stars, TP-AGB stars, extreme horizontal branch (EHB) stars and blue stragglers - are the major modeling challenge, as in empirical libraries. Other problems also arise from the line lists, that are still not complete, which means that not all observed lines are reproduced. This situation affects more the cool stars, where molecular lines, that carry the largest uncertainties, are more important.

Each type of library has its own advantages and disadvantages. Empirical spectra have the advantage of being physically accurate since they are provenient from real objects. However, there are observational uncertainties, such as the atmospheric corrections and flux calibration, that impose a certain degree of error in empirical spectra. They also have a limited wavelength coverage and spectral resolution, that depend on the characteristics of the instrument used to observe. This is a problem for the construction of spectral libraries with a wide wavelength coverage and constant resolution, such as it would be desirable to perform full spectral fitting, because it is necessary to use data from different instruments that do not have homogeneous characteristics. This is precisely one of the advantages of synthetic spectra: it is possible to produce spectra with constant and high resolution throughout wide wavelength windows. But the major disadvantage of empirical libraries is actually their limited coverage of age and metallicity. Objects that are suitable to be observed with sufficient spectral quality to be included in a base belong to the Milky Way, particularly to the solar neighborhood, and besides not spanning all possible age and metallicity combinations, tend to be biased towards solar abundances. As a result the HR diagram is not homogeneously sampled and interpolation is needed to complete the coverage. This is also a problem that synthetic libraries can overcome. However, a great number of theoretical assumptions is needed to build these synthetic spectra, introducing error in the final spectra. Finally, empirical spectra are not completely model independent. Their classification in age and metallicity, needed to interpret the model

that results from the full spectral fitting, requires some theoretical hypothesis to be made. Despite this, the degree of model's dependency is obviously much smaller than in theoretical spectra and the main advantage of empirical spectra is its continuum accuracy and spectral line coverage.

Final expression

The stellar spectra correspondent to the evolved stars are then superimposed resulting on a synthetic integrated spectra with a given age. The final SSP spectrum is computed in the following way [Walcher et al., 2011]:

$$F_{SSP}(t, Z) = \int_{m_{min}}^{m_{max}} F_\star(m, Z)\phi(m, t)dm \tag{1.2}$$

where $F_{SSP}(t, Z)$ is the spectrum of a SSP of age t and metallicity Z; $F_\star(m, Z)$ is the spectrum of a star with mass m and metallicity Z, and $\phi(m, t)$ is the stellar mass function, that describes the fraction of stars of mass m that are still contributing to the SSP spectra at time t, derived from the IMF and the isochrones.

Simple Stellar Populations are a fundamental concept to understand how galaxies have assembled their mass.

1.3 Estimating the stellar population of a Galaxy

Knowing when a galaxy formed its stars is a vital key to reconstructing its history. Nowadays it is possible to deduce the star formation history (SFH) - the rate at which mass was converted from gas into stars throughout time - of a galaxy by studying their current stellar populations. But estimating the stellar population of a galaxy remains a tough challenge and the techniques used to do it depend mainly on the characteristics of the available observational data. This can be divided in major two classes: resolved galaxies, where individual stars can be distinguished, and observations of unresolved galaxies.

1.3.1 Resolved Galaxies

In the nearest galaxies, individual stars can be resolved, making it possible to perform photometry on them. This technique consists of observing the galaxy with different filters and measuring the flux of each star in each of these filters. This provides the information needed to construct an observational colour-magnitude diagram (CMD) which is comprised of the colour of the stars, calculated from the relative flux of one band to another, and the luminosity deduced by the total observed flux.

This analysis was first performed on Milky Way star clusters. The visual inspection of a cluster's CMD allows some assumptions to be made about its age, from the presence of some of the more evolved branches that give a lower limit for the cluster age. Nevertheless, for a more precise age determination the main sequence has to be inspected. Stars in the upper part of the main sequence, the more massive stars, will disappear progressively with time, changing the 'turnoff' to the red giant branch. To accuratelly determine the age of the cluster, the MS must be observed up to its turning point, where populations with different ages can be distinguished. The age is then found by fitting an isochrone to the stellar distribution.

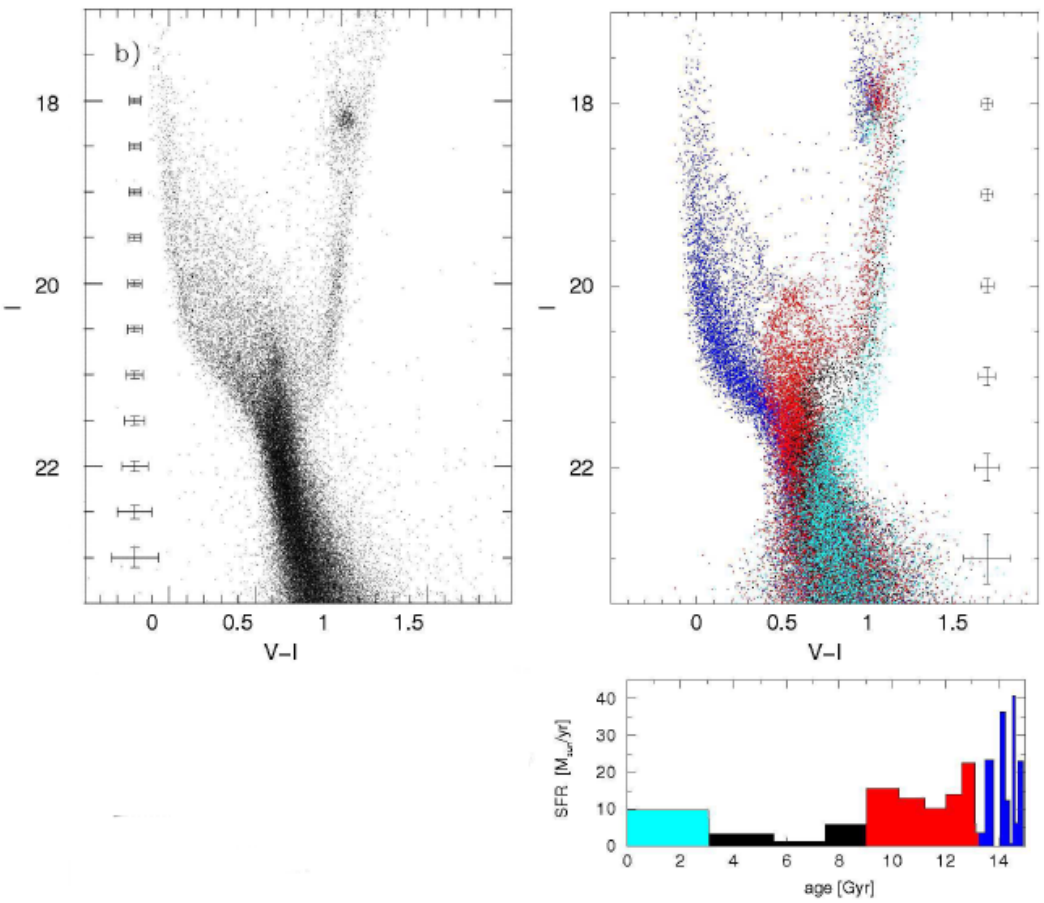


Figure 1.9: *Left:* CMD of the Large Magellanic Cloud observed with the HST instrument WFPC2 (with error bars on the left). Axis in magnitude units - from (Smecker-Hane et al.,2002). *Right:* Synthetic CMD for the same field and respective SFH. Stellar populations originating from 4 different phases of star formation are coded in different colors - from (Lilly and Alvensleben, 2005)

Galaxies however are much more complex systems, since they are composed of more than one stellar population and their star formation history composed by a continuous sequence of bursts and not only one. What is more, SSP with different ages can occupy the same area of the diagram (see figure 1.9) and consequently different SSP will appear superimposed, which makes the fitting of several isochrones extremely challenging and brings up the need of using other methods.

The technique used to tackle this problem compares the observed CMD with an artificial one. To built this synthetic CMD it is necessary to make a first educated guess on the general star formation history (SFH) of the galaxy and to produce the several isochrones that would describe the stars originated by that SFH. From these isochrones, stars (points in the distribution) are picked using a Monte Carlo approach and used to construct a CMD. This synthetic diagram is statistically compared to the observed one, typically using a form of likelihood analysis, and the proposed SFH is modified until the two diagrams are satisfactory alike [Tolstoy et al., 2009]. After this, using the information from the inputed isochrones (age, metallicity, number of stars) it is possible to deduce the true the star formation history of the galaxy. Figure 1.9 compares an observed CMD of the Large Magellanic cloud and a synthetic model with

its respective SFH.

1.3.2 Unresolved galaxies

Even with the latest generation of telescopes the vast majority of galaxies we see in the Universe are unresolved. In this case only their integrated light, the sum of the light of all their luminous components attenuated by dust, is observable. Its study can yield informations about all the components that contribute to the integrated light - stars, gas and dust - using photometric or spectroscopic methods. The goal of this work is to explore what can be learned from the integrated spectra of galaxies.

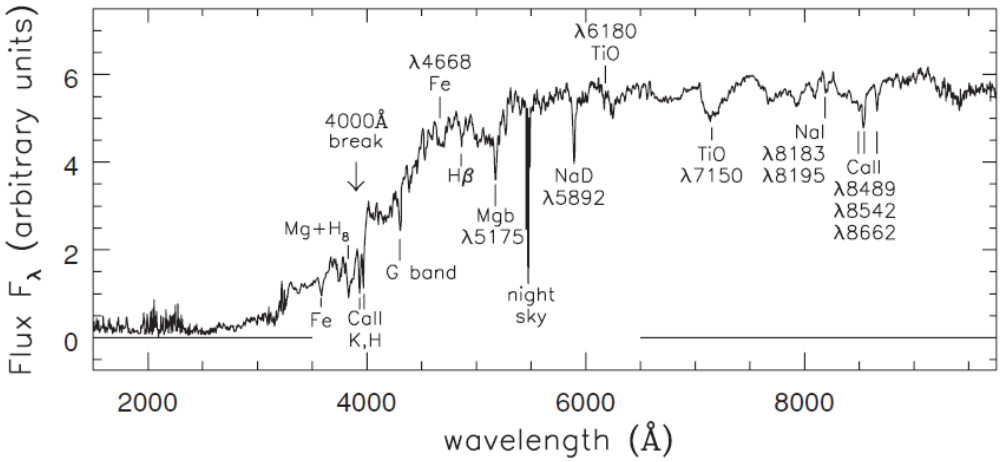


Figure 1.10: Spectrum of an elliptical galaxy - by A. Kinney

In Figure 1.10 an example of an elliptical galaxies' spectrum from 2000 to 9000 Å is shown. Like the spectra of stars, it is made up by a continuum overlayed with absorption lines, but it also contains emission lines, due to the contribution of gas. A brief description of these features in the optical is given bellow:

- **Continuum** The continuum is mainly shaped by its stellar content. Galaxies with young stellar populations still have very luminous massive blue stars (Type A and O) which means that they will emmit most energy in the blue region of the spectrum. On the contrary, the stellar content of galaxies with no recent period of star formation will mainly be of low mass red stars, which translates in a redder spectrum. Gas heated by stars also emmits in the continuum contributing to its general shape. Dust affects the shape of the continuum, by reemission and absorption of the starlight, although emission is only relevant in the far infrared. Dust attenuation causes the diminishing of the total light by scattering and absorption of starlight. Because photons with lower wavelengths are more likely to be scattered and absorbed by the small particles that form interstellar dust, the bluer part of the spectra will be more attenuated than the red part, causing the reddening.

- **Breaks** Breaks are spectral discontinuities, wavelengths at which the continuum experiences an abrupt change. As previously referred, type K and cooler stars have strong metal absorption lines

that absorb most of the light below 4000 Å.

- **Absorption Lines** These are originated in the stars photosphere by electronic transitions to higher energy quantum levels. For this process to occur the atom must absorb energy, in the form of a photon, using it to promote an electron to a higher level. The atom becomes exited or, if the electron is promoted to such a high level that is no longer under the nucleus influence, it becomes ionized and its charge is no longer neutral. Because the energy levels are quantified, the energy difference between the ending and starting levels is also well defined for each transition and for each type of atom. This allows to make the correspondence between an absortion line, the sum of inumerous atomic absorptions taking place in a star's photosphere, to the a chemical element in a specific exitation state. The strength - the flux intensity - of a line depends on the abundance of the element that originated it and can therefore be used to measure the metal content of a star. Despite beeing created by the same electronic transition, which implies that the absorbed photons have the same energy and consequently the same wavelength, the lines do not have an infinitesimal width. The thermal motion of the elements on the star causes individual Doppler shifts that summed broaden the lines. This broadening can be connected to the temperature of the star, with hotter stars having greater Doppler broadening. However, in a galaxy's spectrum these lines have a voight profile that is not only explained by this thermal broadening. It is a consequence of the sum of the multiple individual star spectra in the integrated spectrum, each of them having their particular thermal broadening and Doppler shift due to the star's relative motion.

- **Emission Lines** They have their origin in the desexitation of atoms or molecules in the interstellar gas. In opposition to process described above, when an ion captures an electron or an excited atom relaxes to a lower energy state, energy is released in the form of photons. Just like in the case of absorption lines, the energy of the emitted photon corresponds to the energy gap between the excited and lower energy level, and will have a specific value for each transition. This will also allow the identification the components of the gas and their excitation state. For this process occur it is necessary to ionize the gas. This can be done via photoionization, when an electron absorbs the energy necessary for the transition from an photon, or via collisions between electrons, when temperature is high enough to provide a considerable amount of electrons with the required energy. For the hydrogen, the main constituent of the interstellar gas (and of the Universe), the ionization energy if of 13.6 eV. Only type A or O stars are sufficiently hot to radiate photons capable of ionizing the hydrogen gas. Around these stars, regions with high quantities of ionized gas, the H II regions, are formed. Because of this coupling between gas and stars necessary to create the emission, the strenght of the emission lines indicates how powerfull is the ionization flux. Other elements can also be ionized and suffer the same desextitation process, emitting photons with specific wavelengths. Gas metallicity can therefore be inferred from the presence of these emission lines. In a galaxy's spectrum, because there are elements that are abundant both in stars and gas (such as the hydrogen) the emission and absorption lines will be superimposed. This causes the attenuation of the lines and is the origin of the characteristic 'deep' visible on the sides of some

emission lines, that are the effects of the flux subtraction caused by the correspondent absorption lines.

Spectroscopy uses the spectrum of an astronomical object to understand its physics, by matching the spectral features with their physical origin. The advantage of this technique regarding photometry in the study of unresolved galaxies relies on the possibility of breaking the age-metallicity degenerency, that is further complicated by the dust presence, that also reddens the spectra. It is hoped that spectroscopic analysis allows to differenciate age, metallicity and dust effects, since narrow spectral features are relatively immune to the overall reddening.

Spectroscopic Analysis

One strategy to disentangle the contributions of stars, gas and dust to the integrated light is to construct synthetic spectra that reproduce the observed spectra. Due to the complexity of these spectra, I will only focus on absorption lines and the continuum, and in what they can yield about the stellar content of a galaxy.

One approach to describe the stellar component of a galaxy's integrated spectrum is to consider it as a sum of the individual spectra ($F^\star$) of the stars that compose it. The model spectrum (M) is then described as:

$$M = \sum^{N} F^\star_{o,j} \tag{1.3}$$

where N is the total number of stars in the galaxy and $F^\star_{O,j}$ the observed spectrum of the star j. The observed spectra corresponds to the intrinsic spectra with the flux attenuated by the interstellar dust and the spectral lines blue or red shifted by the star's proper motion. The model can then be modified to:

$$M = \sum_{j}^{N} F^\star_j r_j(\lambda) f(v_{\star,j}, \lambda) \tag{1.4}$$

where the $F^\star_j$ are the intrinsic spectra of the stars, $r_j(\lambda)$ is the dust attenuation law and $f(v_{\star,j}, \lambda)$ is the wavelength shift given by the Doppler effect.

The attenuation law describes the intrinsic attenuation suffered by a galaxy, this is, the effects of the dust within the observed object have in its own light. Its has the following general form:

$$r_{\lambda,j} = 10^{-0.4E_s(B-V)k'(\lambda)} \tag{1.5}$$

where B is the flux in the blue band, V in the visible and $E_s(B-V) = (B-V)_{observed} - (B-V)_{intrinsic}$. The $k'(\lambda)$ factor is the selective attenuation of the stellar continuum, that describes the dust-light interaction. This interaction is quite complex and, so far, no theoretical model has conveniently described it. Intrinsic attenuation depends not only on the chemical properties of the dust (composition and temperature), but also on its geometrical distribution. Dust is not uniformly distributed, and so some stars are likely to be more strongly attenuated than others. Bearing in mind that the surrounding

environment of a star evolves with its life (young stars are enveloped in gas whether old stars are more often in a low density environment), the geometrical distribution may also depend on the stellar population age. Moreover, there is also scattering not only out of the line-of-sight, but also into it, which further complicates the scenario. For these reasons, the physical interpretation of the factors on $k'(\lambda)$ is quite difficult. Figure 1.11 displays some of the empirically determined selective attenuations.

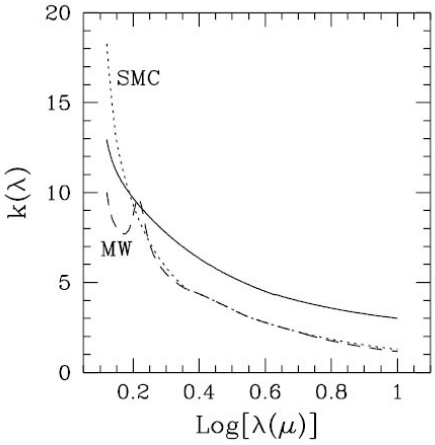


Figure 1.11: $k(\lambda)$ factors. *Continuous line*: Calzetti et al.(1994) derived law. *Dashed line* Determined for the Small Magellanic Cloud. *Dotted line* Determined for the Milky Way

It is also possible to describe the integrated spectra as a sum of Single Stellar Populations (SSP) spectra. In this case, the kinematic part of the model will depend on the mean velocity of the stars composing the SSP (v_{SSP}) and their velocity dispersion (σ_{SSP}). The mean velocity is responsible for the wavelength shift that the SSP spectrum suffers, similarly to what happen to stars spectra. The dispersion gives rise to the broadening of the spectral lines, caused by the different proper motions of stars within the same SSP. As this effect will depend on the specific kinematics of each SSP, the general model can be formulated as:

$$M = \sum_{j}^{N} F_j^{SSP} r_j(\lambda) \otimes f(v_{SSP,j}, \sigma_{SSP,j}) \tag{1.6}$$

where $f(v_{SSP}, \sigma_{SSP})$ represents a general law that will translate v_{SSP} and σ_{SSP} into Doppler shifts and line broadening.

Algebraic description

The problem of finding the best model for a specific observed spectrum can be addressed as an algebraic inversion problem, treating this model as a linear combination of spectra. In this methodology, a set of spectra of stars or SSP, either theoretical or empirical, is used as a base and the linear contribution of each element is estimated. This immediately posses a limitation to the method, since there is a limited number of available spectra to be used as base elements. This means that the sum will not be in all galaxy elements (N) but in a more modest number (n), the size of the base. The implemented model is

also discretized - $M = \sum_\lambda M_\lambda$ - which permits different wavelength intervals to be used, depending on the anaysisis being done. The final model formulated in an algebraic language is:

$$M_\lambda = \sum_j^n c_{j,\lambda} L_{j,\lambda} r_{j,\lambda} \otimes f(v_j, \sigma_j) \tag{1.7}$$

where where M_λ is the model in a wavelength interval, $c_{j,\lambda}$ is the contribution of the base element j to the integrated spectrum and $L_{j,\lambda}$ is the flux of the j element in the considered wavelength interval. This model is very general and it may be applied to different analysis techniques - colour, spectral indices and full spectra fitting - mainly depending on the spectral resolution of the data to be fitted.

Analysis techniques

Colour analysis is mostly used with photometric data but can also be performed in low resolution spectra, integrating the flux in a relatively broad and well defined band. In terms of final data these two options are equivalent, but photometry has the advantage of consuming less telescope time. Moreover, it is also possible to perform photometry in the ultra violet (UV), that is better for estimating young stars contribution, and near infrared (NIR) bands, best for old stars, and not only optical as its usual with most spectrograph. In colour analysis, the wavelength interval considered in the model is quite broad and only a few points - that correspond to the available photometric bands - are fitted. The basis elements are either empirical colours or the star and SSP spectral models can be integrated equivalently to the observed band. Figure 1.12 is an example of this type of fitting for photometric data from the far UV to the far IR.

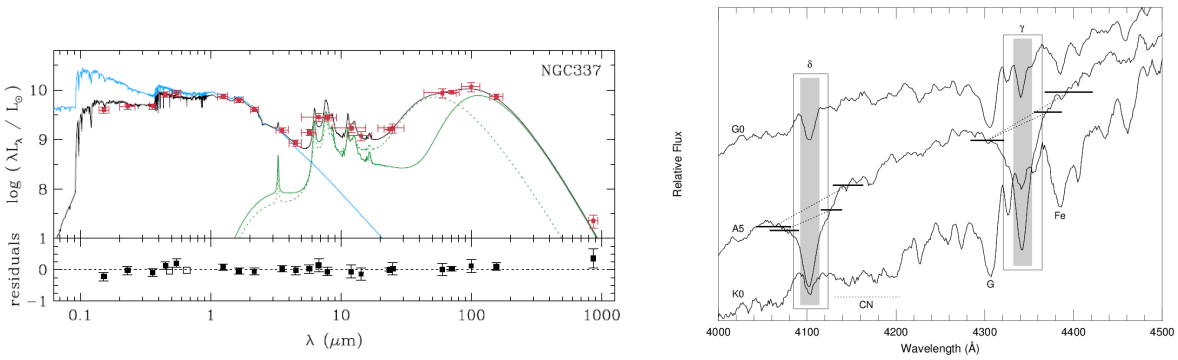


Figure 1.12: Analysis techniques. *Left: Colour fitting example.* In red: observed data. In black: best fit model. In blue: dust unattenuated model. In green: dust emission - by E da Cunha. *Right: Spectral line fitting example* The rectangles (in grey and white) sign two possible definitions for the central band. The lateral lines correspond to the mean of the corresponding side bands from which the I_c is calculated (dotted line) - [Worthey et al., 1994]

With low resolution spectra it is possible to use the absorption lines to get a better description of the stellar content of a galaxy, diminishing the wavelength interval of the model to a few dozens of Å. This is done using spectral indices, that are calculated defining a central band ($\delta\lambda_0$) around an emission line, with intensity (I_c), and two side bands, a red and a blue one, for which the corresponding intensity (I_s) is the mean of both bands. From these, the equivalent width, $W = (1 - I_c/I_s)\Delta\lambda_0$, is calculated [Worthey

et al., 1994]. An illustration of these measurements can be seen in figure 1.12. A collection of spectral indices, one of the most common being the Lick indices catalog, is used as basis for the above model.

When good resolution spectra are available it is possible to perform full spectral fitting, that will be described with more detail in the next section.

1.4 Full Spectral fitting

1.4.1 Technique

Full spectral fitting is a technique that attempts to derive the model described in equation 1.7 in a wide wavelength range with a high spectral resolution, fitting continuum and absorption lines at the same time. Figure 1.13 shows a schematic of the spectrum model components.

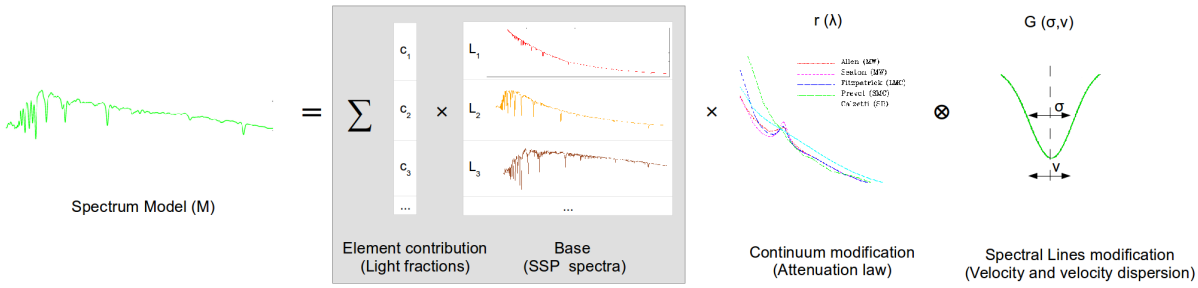


Figure 1.13: STARLIGHT model spectrum components

In this work, an algebraic description will be used, estimating the linear contribution of indivual spectra to the final model.These can be either stars or SSP spectra, empirical or theoretical. The only requirements are that each element covers at least the considered spectral range, usually the optic, and that it has a resolution comparable to the one of the data to be fitted. Figure 1.14 shows an example of the use of this technique.

1.4.2 Elements of the base

The ensamble of individual spectra that can contribute to the final spectrum model forms the algebraic base used in this description. As mentioned, the base can be composed by stellar or SSP spectra. Stars spectra involve less theoretical treatment or, in the case of observational spectra, more age- metallicity combinations are available, and are therefore more acurate. However, using stars as base elements will return the stellar content of the galaxy, but no information about its star formation history. For the purpose of studying galaxy evolution, it is more usefull to use SSP spectra as base elements. Observational spectra of star clusters suitable for full spectral fitting is limited to Milky Way clusters. This has the drawback of providing a limited basis, since the Milky Way clusters are either metal poor globular clusters, with ages older than 12 Gyr, or open clusters of younger age, the great majority with less than 1 Gyr [Sparke and Gallagher, 2007]. On the other hand, the construction of synthetic Simple Stellar Population spectra requires a great deal of inputs, as explained in section 1.2.2. All the referred uncertainties of

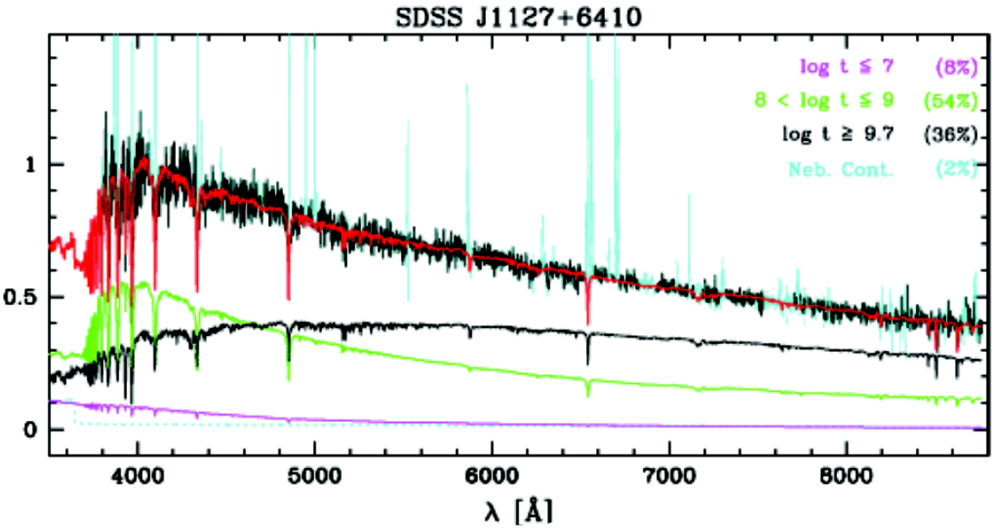


Figure 1.14: Full Spectral Fitting example in optical wavelengths: *In black*: observed spectrum *In red*: best fit model *In purple, dark blue and green* individual basis elements contribution *In light blue:* estimated gas contribution - Corbin et al. 2006

stellar spectra are carried into the SSP spectra: the imperfections on stellar physics of the theoretical spectra or the inaccuracies that arise from the interpolation of empirical spectra needed to cover all necessary age-metallicty combinations. More frequently, a mixture of both theoretical and observational uncertainties is propagated into the synthetic SSP, since a mixture of theoretical and observational stellar ingredients are used.

1.4.3 State of the art of full spectral fitting

Efforts have been made to verify the consistence of different fitting techniques, the impact of the different input physics and internal errors. These tests found a good agreement between the different fitting algorithms available today. However, estimating the accuracy of the method is still a work in progress.

Star Clusters

Star clusters (SC) are the first natural candidates to test the reliability of full spectral fitting since they are real SSP. In Wolf et al. [2007] a systematic validation of the method was done using 101 globular clusters spectra from 3200 to 10000 Å and varying spectral resolution (from 6 to 23 Å.). These clusters covered a range age of 4 Myr $< t <$ 20 Gyr and metal content of $1.6 < [Fe/H] < 0.3$ and their individual values were known from CMD and individual stars spectroscopic analysis. The authors found that full spectral fitting (performed with the [Bruzual and Charlot, 2003] SSP models) overestimated the ages of many clusters younger than 1 Gyr, with an average offset of 0.69 dex (0.54 Gyr) and a dispersion of 0.9 dex (3.2 Gyr), a result that they atributed to extra redenning of the spectra due to TP-AGB stars. This young clusters also revealed to be problematic in metallicity determination, and no correlation with literature values was found. For older clusters an agreement of 0.16 dex in age and 0.12 dex in metallicity

was found. González Delgado and Cid Fernandes [2010] continued in this line of investigation, using a sample of hight quality spectra in the 3650 to 4600 Å range of 27 young SC from the Magellanic Clouds. Several bases were tested and the results reveal a consistency of 0.17 dex in the ages estimated by the several bases. Metalicity and extinction results were more problematic, with a calculated precision of 0.5 dex for metallicity values between models. Comparing the results with literature values, the method yielded an accuracy of 0.1 dex in age estimation and 0.3 dex in metallicity.

Galaxies

The Large Magellanic Cloud (LMC) bar was used as a testing object by Lilly and Alvensleben [2006], following the methodology outlined by Alloin et al. [2002]. In a preparatory study for this analysis, six synthetic galaxy spectra with distinct star formation histories (SFH), but same metallicity, were used to access to what degree the SFH could be reconstructed from integrated spectra [Lilly and Fritze-v. Alvensleben, 2005]. The results showed that very different SFH (see figure 1.15) could result in very similar integrated spectra properties. From this investigation, the authors concluded that the optimal approach to this problem was to divide the look back time in several long periods, with the earlier ones having a bigger weight. Using this methodology, the SFH derived from CMD analysis was smoothed to longer periods and a set of variations of that original SFH were produced. The respective synthetic spectra was calculated and compared with the observed one, reaching the conclusion that stellar populations could only be reasonably well estimated up to 1 Gyr after the last star burst (4 Gyr in the best scenario), since blue stars luminosity greatly difficult the identification of the faint spectral features of older red stars. It is worth mentioning that in these studies the model spectra were derived assuming a SFH (much similarly to the process discussed to derive SFH from colour-magnitude diagrams) and not from fitting.

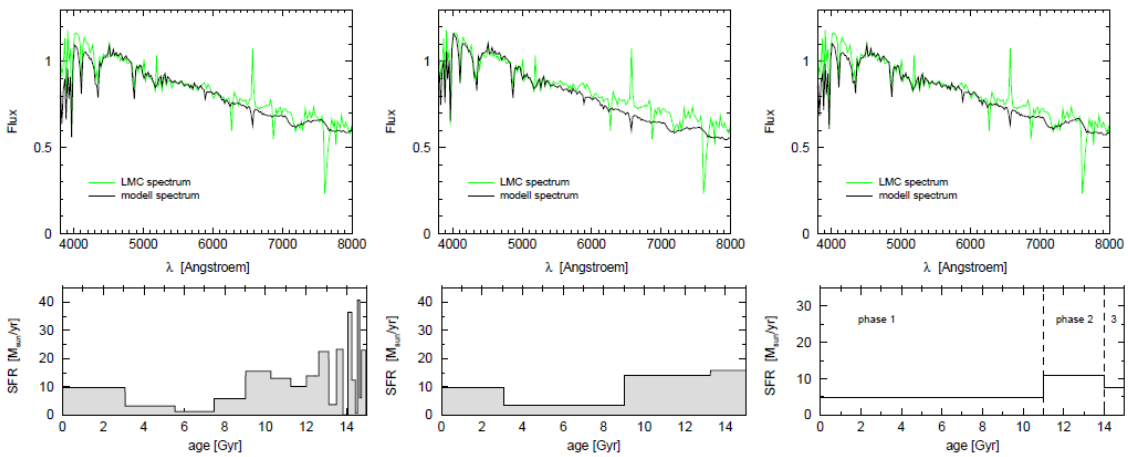


Figure 1.15: Comparison of integrated spectra from different SFH. *Upper panels*: Model spectra (in black) and observed spectrum (in green). *Lower panel*: Star formation histories. Left: SFH from Smecker-Hane et al. [2002] CMD analysis. Middle: smoothed Smecker-Hane SFH. Right: 3-phase SFH - from [Lilly and Alvensleben, 2006]

Milky Way dwarf satellite galaxies are the other potential targets for these validation tests. A first study was done using the dwarf galaxy KDG 64 by Makarova et al. [2011]. Full spectral fitting was used,

but in a limited way, only allowing for two components to be recovered: a young one with less than 2.4 Gyr and an older with a fixed age of 10 Gyr. The results were consistent with the previously published.

Limitations of the current validation tests

Although these previous studies point to the potentialities of full spectral fitting, its validity in the study of the stellar populations of galaxies is still not fully addressed.

Star clusters are composed by a single stellar population, which limits their SFH to one burst of stellar formation. In some major and more complex clusters up to three SSP can be found, but this is still far from the complexity of a galaxy SFH. Star clusters are then inefficient to validate the accuracy of the star formation history. Synthetic spectra can avoid this problem, by simulating more complex SFH. However, these synthetic spectra are built using the current knowledge of stellar physics, that is also used to built the bases of the model being tested. This tests can, in some extent, only demonstrate that the method is self consistent and not that is accurate. What is more, synthetic spectra still do not include dust and gas effects, since models are rather crude, factors that have a major impact on real integrated spectra of galaxies.

It is then important to estimate the uncertainties of full spectral fitting when applied to real galaxies. With this in mind, the Starfish project was started, led by Myriam Rodrigues, European Southern Observatory fellow hosted by the Observatoire de Paris. This project will use a sample of nearby galaxies to verify if a galaxy spectrum can be accurately described as an inversion problem, if the current models already contemplate all the ingredients needed to describe it, what physical properties can we hope to determine and to how far in the past can the SFH be probed with this method. In my thesis I present the first results of the Starfish project. The next chapter is dedicated to the description of this project.

Chapter 2

Starfish: STellar PopulAtion From Integrated Spectra

2.1 Goals and Objectives

The goal of the Starfish - STellar populAtion From Integrated Spectra - project is to test the accuracy of the stellar population parameters (age, metallicity, mass and star formation history) derived from integrated spectra through full spectral fitting. It aims to validate this method for the first time by comparing real integrated spectra with Color Magnitude Diagrams of the resolved stellar populations within the same galaxy. The main objective is to determine the systematic errors and underlying uncertainties of full spectral fitting. The chosen methodology will also provide two valuable resources to the scientific community. The first, is a calibration of stellar population models used to infer the properties of unresolved galaxies. This is particularly important for high redshift galaxies that share some of the intrinsic characteristics of the selected sample used in Starfish (high gas content, low metallicity and young ages). The second, is a library of 38 integrated spectra of nearby galaxies with star formation histories (SFH) independently determined by the CMD method. This library will constitute an ideal testing sample for stellar synthesis models and fitting methods for all unresolved galaxies.

2.2 Methodology

The validation of the full spectral fitting technique will be done comparing the derived SFH of a sample of galaxies with the SFH determined via CMD analysis, that so far is the most accurate technique available for this type of study. The methodology to be followed in this project is briefly outlined in figure 2.1.

The integrated spectra of the selected galaxies was carefully reduced and calibrated before being fitted with one of the available inversion algorithms, from which the SFH will be obtained (lower panel of figure 2.1). The exact spacial area where the integration of light took place in each galaxy is provided to the ANGST (ACS Nearby Galaxy Survey Treasury) team [Weisz et al., 2011] that has already performed a detailed CMD analysis of these galaxies but in a wider area than the spectroscopically integrated one.

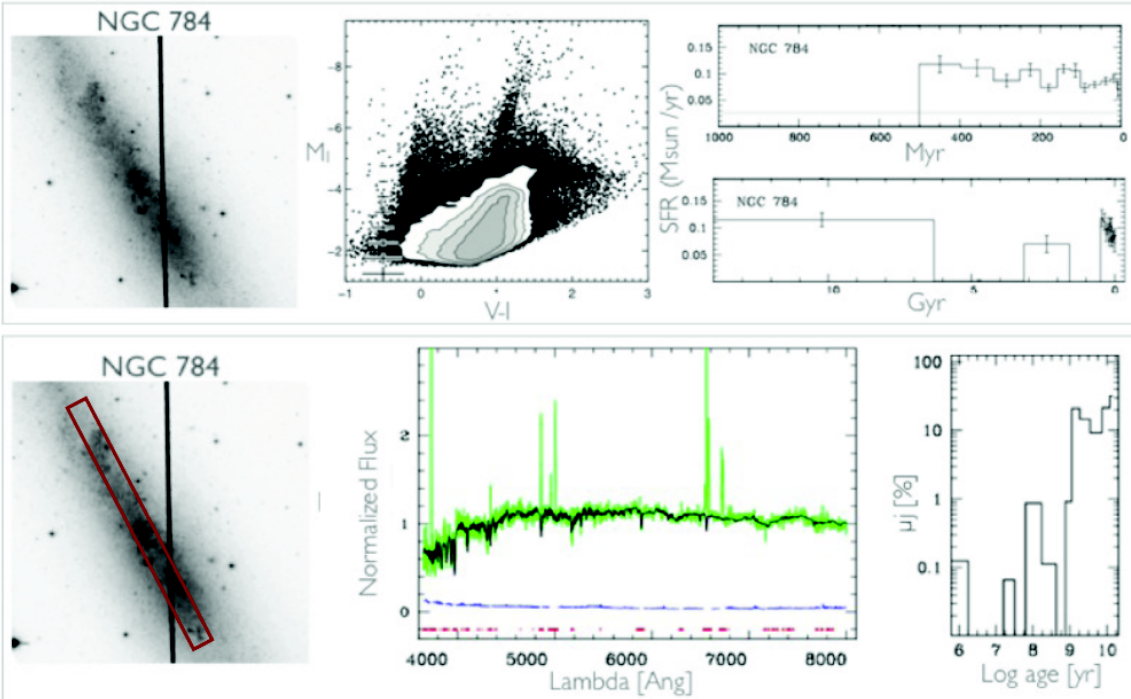


Figure 2.1: Methodology of the Starfish project (from the GMOS observing proposal) . Example with NGC 784 *Upper panel: Hubble Space Telescope image, observational CMD and SFH derived from the CMD method - Mc Quinn et al. 2010. Lower panel: SDSS image, full spectral fitting performed in a galaxy's integrated spectrum (observed spectrum in black and fit in green) and SFH derived from the model - Cid Fernandes et al 2005*

The ANGST team recalculates the SFH from the CMD of the given area (upper panel) and it is this SFH and corresponding stellar populations that is used for full spectral fitting and error analysis.

2.3 Sample Selection

This project requires photometric observations deep enough to measure low luminosity and low mass stars, that contribute significantly to the total stellar mass. The magnitude limit was set so that stars at least 1.5 magnitudes below the tip of the red giant branch could be resolved, which allows for estimates of the SFH up to 7 Gyr in look-back time to be made from the CMD. The photometric images must also have a good spacial resolution, in order to measure individual stars to construct the observational CDM. These conditions dictate the need for Hubble Space Telescope (HST) data. At optical wavelengths the HST has spacial resolution superior to any groud base telescope. However, even HST has finite resolving power, and the galaxies for this combination of requirements were fulfilled restricted the sample to Milky Way dwarf galaxies, 1 to 5 Mpc away. This selection was made for the ANGST project and figure 2.2 shows the distribution in distance and magnitude of the 60 suitable dwarf galaxies.

Further selection criteria had to be defined for the spectroscopic observations. The first was a lower limit of 24 mag/arcsec^2 for the galaxies' surface brightness, to keep the integration period of the observation within reasonable limits. The targets need to have a small angular size in order to cover a representative area of the galaxy with the field of view of the instrument. This narrowed the sample to

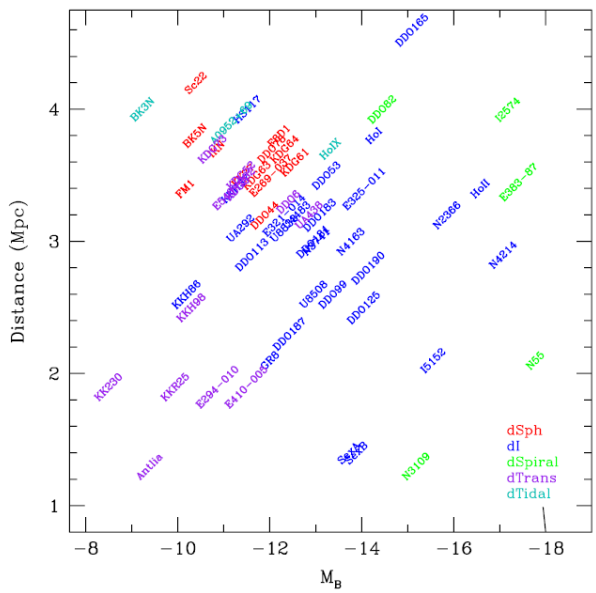


Figure 2.2: Distribution of the ANGST sample of dwarf galaxies in distance and M_B. Galaxies are color-coded by morphological type - [Weisz et al., 2011]

38 dwarf galaxies: 4 transition dwarves, 2 spheroidal, 4 dwarf spirals, 10 irregulars and 18 starbursts.

2.4 Observing strategy

To fulfil the project goals, the observed spectra were required to be observed in a wide vavelegth range within the optical band (3200 to 6100 Å) and to have resolution ($\lambda/\Delta\lambda$) of at least 800, to be comparable with those used in typical spectral fittings.

Due to their low surface brightness the targets were observed with 8 m class telescopes. In the southern hemisphere the observations were made in Chile with the VIsible MultiObject Spectrograph (VIMOS), at the Very Large Telescope, and with the Gemini Multi-Object Spectrograph (GMOS) on the Gemini South Observatory. Observations in the northern hemisphere were also performed GMOS, on the Gemini North Observatory, on Mauna Kea in Hawaii. To date, 24 of the 38 galaxies in the sample have been observed.

A different strategy was used for each of the instruments. For VIMOS, the Integral Field Unit (IFU) in high resolution mode was used (see figure 2.3). In this mode the field of view has an area of 27″ x 27″ and so multiple pointings were needed to cover an entire galaxy. For that reason only galaxies with angular size smaller than 7′ were observed with VIMOS. For both GMOS instruments (GMOS-N and GMOS- S are identical) long-slit spectroscopy was performed (see figure 2.3). An 1 arcsec slit was used to observe the galaxies, as a compromise among spectral resolution, signal-to-noise and spatial coverage of the targets within a reasonable amount of time. Targets were observed with the G5307 grating in the optical band - 3600 to 7000 Å - with a long-slit of 1″ aperture to achieve a spectral resolution of $\sim$ 800 and a dispersion of 0.050 nm/pixel. Observations with consecutive offsets, of variable number depending on the target morphology, from a central position were summed in order to obtain the spatially integrated

spectra. This technique has the advantage of allowing to remove foreground stars that contaminate any of the acquisitions, since it preserves the spacial correspondence of each observed spectrum. For each long-slit position an exposure of 60 minutes yielding a S/N of 14 per exposure and a predicted S/N of 35 for the summed exposures.

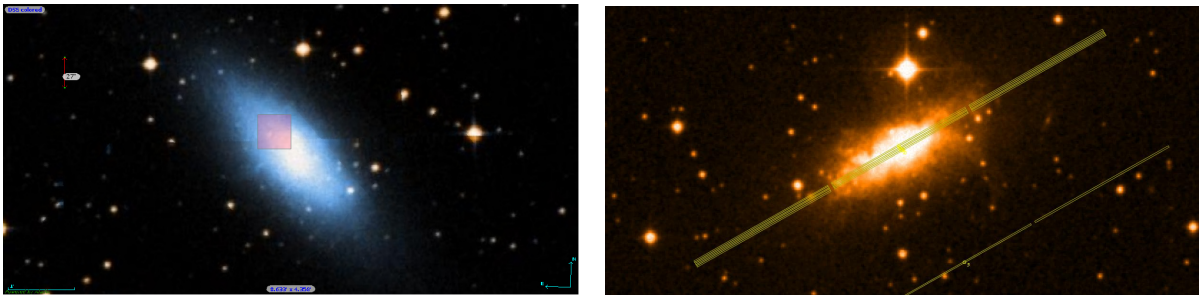


Figure 2.3: Observational setup. *Left*: VIMOS in Integral Field Unit (IFU) mode. Target NGC 5253 overlayed with field of view in the high resolution mode (only one observed position) - Generated with the Aladin tool *Right*: GMOS in long-slit mode. Target NGC 1569 overlayed with long-slit masks (all positions to be observed) - Generated with GMOS observing tool

Chapter 3

Data Reduction

In my master thesis I reduced and calibrated GMOS data: 5 galaxies observed in the GMOS-N observatory and 2 from GMOS-S. For the standard reducing steps, I have used the IRAF pipeline developed by the Gemini staff. For the combination of the several observations I developed a IDL script that allowed the inspection and removal of foreground stars in the final result. I also produced a second script to extract and flux calibrate the final spectra to be used in the analysis.

3.1 GMOS Data

Astronomical data is commonly stored in *Flexible Image Transport System* (FITS) files. These contain two extensions: an image with the scientific data and a header with information about that data, such as the identification of the observed object, date, origin, etc. In GMOS, a single acquisition produces six independent FITS files, corresponding to the data read from each of the 6 chips that compose the detector, that are saved as a single *Multi-Extension FITS* (MEF) that also contains its own header, the *Primary Header Unit* (PHU), where information concerning the observation is saved. MEF files are very flexible and can contain several different extensions besides raw or reduced data, such as *Mask Definition Files* (MDF), with information about the layout used to do the cutting of images identifying the gaps and overscan region, data quality files (DQ), images where saturated pixels are flagged, and the variance of the data (VAR). These extensions are generated and used during the reduction process.

Figure 3.1 is a typical raw data image of a GMOS long-slit observation. Besides the spectrum of a galaxy, sky lines and cosmic rays, the two vertical gaps between the detectors are visible (in black), as well as the two shadows cast by the bridges of the long-slit supporting structure (in light grey). The image, or frame, is a 2D spectra with the vertical axis corresponding to the spatial direction, the horizontal axis to the spectral direction and the flux, in electron counts, being plotted with a color gradient from back, the lowest value, to white. If a cut is made along the rows of the array (horizontally in the image) a spectrum is obtained, whereas the spatial profile can be obtained along the lines.

GMOS detector consists of three CCDs, each divided into two 1024 x 4609 pixel chips that are read individually, in an array of 6144 x 4609 total pixels. In this work, a frame is composed by the six

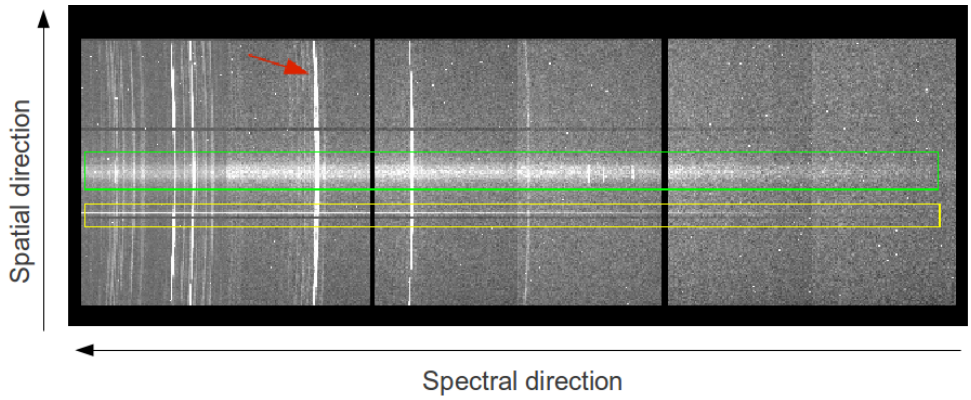


Figure 3.1: A typical GMOS-N long slit image. *In green:* galaxy spectrum. *In yellow:* spectrum of a near object. *In red:* a sky line. The two vertical gaps correspond to the detector gaps and the two grey horizontal stripes to the slit structure. Wavelength axis runs from right (blue) to left (red).

individual data images from the six detector chips, but for simplicity and since the Gemini pipeline is built to deal with the all FITS files simultaneously, all frames shown and referred to in the following sections will be the ensemble of the individual data images.

3.2 Data reduction process

To reduce the data, a combination of Gemini pipeline tasks in the Image Reduction and Analysis Facility (IRAF) environment and Interactive Data Language (IDL) scripts was used. A schematic overview of the complete reduction and calibration process can be seen in Figure 3.2.

The first step to perform the reduction is to process the raw calibration files (with the *gbias* and *gsflat* IRAF tasks). In the case of the bias, this corresponds to the combined median of the multiple exposures creating a master bias. Multiple flat exposures were also median combined, if more than one were available. The master bias and flat fields were then used to remove the instrument signature from the raw images (a), by subtracting the master bias and dividing by the flat field. These pre-reduced (b) frames are then wavelength calibrated using a dispersion solution (λ solution) calculated from an exposure of a known source, usually an arc lamp, corresponding to Xe,Ar,Ne or Th. At this point it is possible to subtract the night sky backgroud flux and emission lines. Up to this point, the IRAF Gemini pipeline is used to process the files. The final files of these processed, sky subtracted images (d), are then visually and manually inspected, using an IDL script specially developed for the task. In this step individual mask files are produced that flag the pixels in each frame that are contaminated and will be rejected from the inclusion in the final integrated spectra. After this, using a second IDL script, the processed files are summed and the final spectra is extracted (e) from the selected regions. The coordinates of the final region are then converted in spatial coordinates, to be used by the ANGST team to calculate the CDMs for the same patial region. Finally, with the sensitivity function derived for a standard star reduced in the same way as the science files, the spectra are flux calibrated and corrected for redshift and galatic extinction. A detailed description of these steps is given bellow.

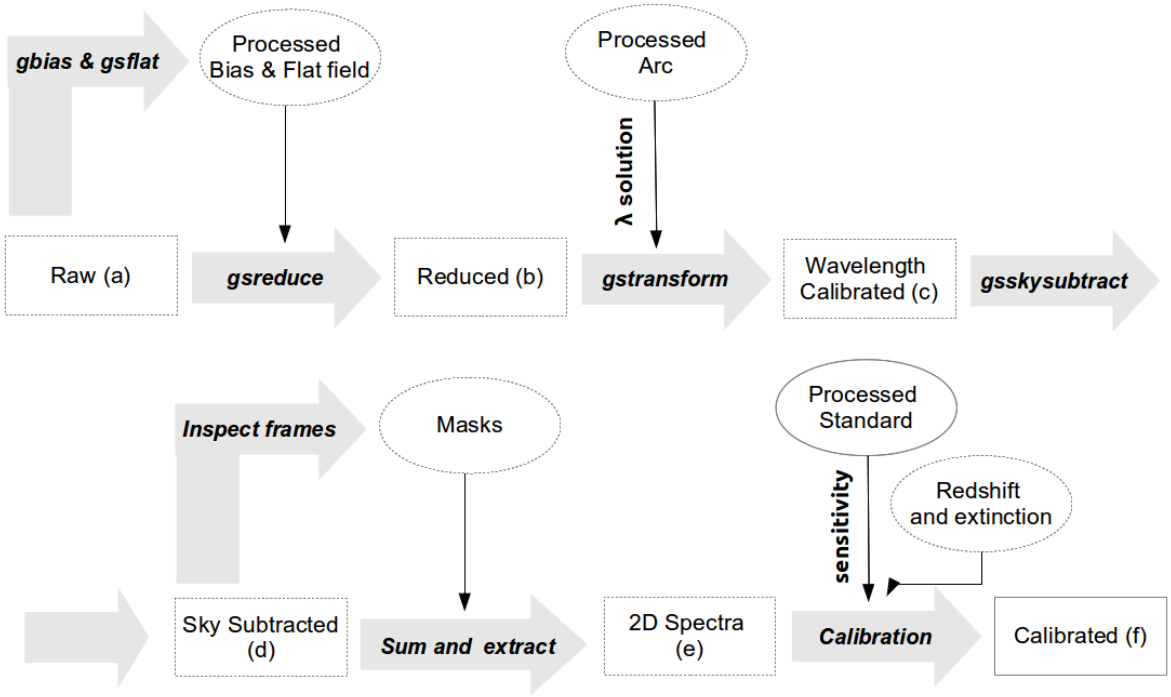


Figure 3.2: Reduction process. The top tasks were performed with the IRAF gemini pipeline and the bottom in IDL environment

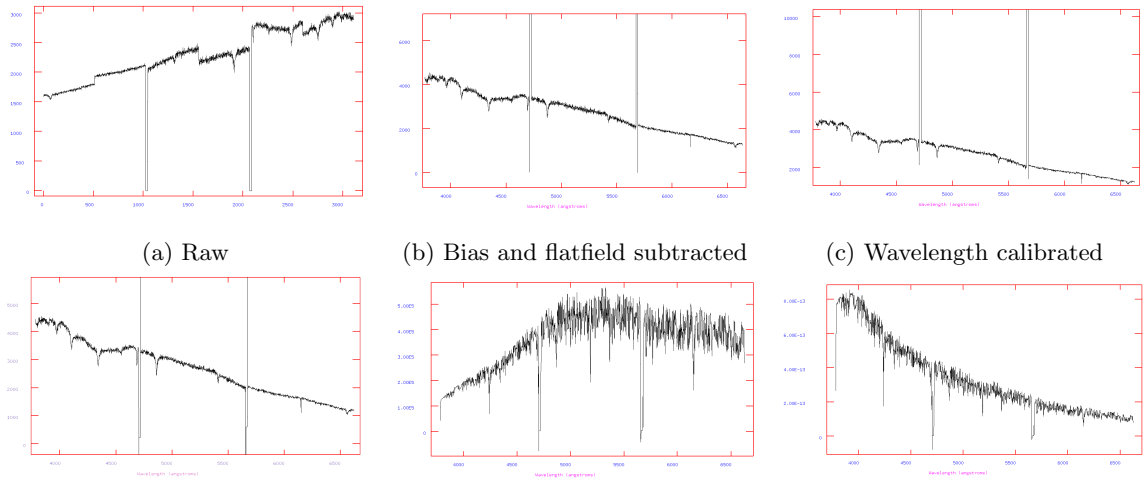
3.2.1 Bias subtraction and flatfielding

During the CCD readout a number of electrons are artificially added to each pixel to ensure that there are no negative values being read. Although this number is meant constant, in some CCD's it can differ from readout to readout and even from pixel to pixel. The first step in data reduction is the subtraction of these extra, or bias, electrons. This can be done in two ways: using the overscan region or bias frames[Massey, 1997].

The overscan region is an array of virtual pixels that are not read in the detector but generated during the readout with the number of extra electrons added in that particular exposure. In GMOS this region is composed of 32 virtual columns added to the edges of each chip. To perform bias subtraction using the overscan, its columns are averaged into one single column and a low order function is fitted along the rows. The fitted function is then subtracted row by row in the entire array.

Overscan subtraction has the advantage of being recorded in each acquisition, but it only corrects variations of the number of added electrons row by row. For a complete spacial correction bias frames are used. These are CCD readouts with zero exposure time and no illumination on the detector that are made during the instruments regular calibrations. Several of these frames are obtained and averaged, if compatible between themselves, and a main bias frame is produced and subtracted to the data pixel by pixel. In either case the overscan is trimmed off before proceeding with the reduction.

In this reduction, the bias images did not show any spacial structure and the mean values of the bias on the master bias frames were in agreement with the mean value of the corresponding overscan region for all acquisitions. Bias frame subtraction was preferred to the overscan region for technical reasons,



(a) Raw (b) Bias and flatfield subtracted (c) Wavelength calibrated

(d) Skysubtracted (e) Extracted (f) Flux Calibrated

Figure 3.3: Reduction example of a single frame (N20120711S0865.fits) - *(a) to (d)* : data of the central line of the frame. *(e) and (f)*: data of the sum of all the lines of the frame. - generated with IRAF

since the Gemini pipeline needed a bias frame to create and propagate the variance (see section 3.3.8).

The second step in any data reduction is the correction for pixel sensitivity, that is not uniform across the CCD. This is done with flat field images, obtained with an ideally uniform light source - a lamp, the illuminated telescope dome or the twilight - and using the same instrument setup as in the science acquisitions. Once more, several of these images are combined and the final frame is normalized. This is mainly done to remove the lamp, dome or sun's spectral signature of the flat field, since it is almost impossible to use a source that has constant flux in all (and any) spectral window considered. This normalization is performed row by row, in the spectral profile, by fitting the acquired flux with an high order polynomial function and dividing that same row by the fitted value, assuring that only variations in the pixel sensitivity are kept in the master flat field, and not the spectral signature of the illumination used. The adequate order of the function to be fitted can only be determined comparing the final result of the reduction with a standard spectra, so this process may involve several tries. Once the final normalized master flat field is obtained, all science data are divided by it, normalizing the number of electron counts along the CCD.

Finally, cosmic ray removal is also performed at this first stage. This is done by fitting the data with an extremely high order cubic spline function (the order depends on the resolution and size of the detector). The dispersion (σ) of the fitted function is determined and points of the data that are more then a determined number of σ away from the fitted function are removed and replaced by the average of the surrounding pixels. A 10 σ threshold was used in this reduction. This value was determined by trial and error, assuring that was high enough not to remove emission lines and low enough to remove a reasonable amount of cosmic rays.

Gemini pipeline tasks used:

gbias - combination of bias frames;

gsflat- combinationa and normalization of flatfield frames;
gsreduce - bias subtraction, flat division and cosmic ray removal (subroutine: gscrreject).

3.2.2 Wavelength Calibration

After the first reduction the 2D frames are wavelength calibrated. For this, a spectrum with known emission lines is taken, with the instrument in the same configuration as the frames to be calibrated. The lines are identified and dispersion function is calculated, matching columns with the respective wavelength.

Gas lamps, in GMOS case an ArCu lamp, are used to obtain the calibration frames. These are only bias subtracted, since it is the spacing between emission lines that is considered to derive the wavelength solution and not their flux. The dispersion function is calculated and applied row by row, since the dispersion of light is not perfectly homogeneous in the CCD. For this, a complex and multiphase pattern recognition algorithm, that uses as inputs the lamp spectra and a list of the corresponding emission lines, is used of witch a brief description is made bellow:

1. Selection of a number of local maximum in the lamp spectra. The algorithm guarantees that these maximum are searched for thoughtout all the horizontal axis.
2. Selection of twice as much lines in the list of emission lines.
3. Calculate the spacings between neighboring lines and the rations between consecutive spacings. This is done for both for the emission lines on the spectra and the listed ones.
4. Both sets of ratios are matched (assuming an error in the spectra). This is done in several iterations where only the strongest matches are kept.
5. For spacings with two common pixels a linear dispersion function is calculated. Once more only the best are kept.
6. Using these local functions, the pixel values are converted in wavelength. Finally a polynomial of second or fourth order is fit to the data, that is taken to be the final wavelength solution.

The wavelength solution yielded a 0.46 Å resolution per pixel. The corrected frames have the wavelength axis increasing from left to right (horizontally flipped in relation to the uncorrected ones). Moreover, this transformation is visible in the straitening of the sky lines, since at this point each column corresponds a single wavelength, which allows the next step to take place.

Gemini pipeline tasks used:
gswavelenght - determines the detector wavelenght response;
gstransform - aplies the calibration.

3.2.3 Sky Subtraction

Sky line subtraction is done by fitting a low order function to the sky background in the science frames. This fit is done spacially, column by column, away from the central part of the frame, where the spectra is, which should in principle provide a good night sky sample. The fitted value, or the interpolated flux in the spectra are, is subtracted in all frame. Broad and saturated lines are hard to remove completely because, despite wavelenght correction, lines are not perfectelly straight as it is required by this process. For galaxy NGC2366 this process was not used, since the galaxy extended through all the long-slit length, not leaving enough space to evaluate the night sky flux. Instead, frames with only night sky spectrum were obtained during the observation, offsetting the slit enough so that the galaxy would be out of the field of view. This frame was directly subtracted in the images to be corrected.

Gemini pipeline tasks used: *gsskysubtract*

3.2.4 Combination and extraction

Although the full reduction and calibration process can be performed in the Gemini pipeline and with IRAF tasks, from this point on IDL scripts developed for this work were used. This allowed a better control and understanding of the data and its associated errors. Particularly, the combination process is crucial to the project, since it is the sum of the frames correspondent to different slit offsets that will create the integrated spectrum.

Before the combination of the frames, the extracted spectrum of individual frames was inspected in order to identify possible foreground star contamination. To correct for these contaminations, mask files flagging the areas of the frames containing undesirable data are created during the individual inspection of the frames. A final frame is produced, rejecting the masked data. A 1D spectra is extracted from this summed frame, by summing the flux it along the columns. The selection of the rows to be included in this sum is also critical, since it defines witch part of the galaxy is being included in the spectra, what will be essential to the CDM comparison.

To perform these tasks two interactive scripts were created. The first script creates the mask files from manual inspection of the frames. The second script takes the reduced frames and performs a simple sum of the unmasked pixels. The extraction area is manually selected and its rows summed, producing the final 1D spectrum and its correspondent error.

The next steps are calibrations and corrections applied to the extracted 1D spectra that were also included in the scripts.

3.2.5 Flux calibration

A precise flux calibration is essential to the Starfish project, since the shape of the continuum will be one of the elements to be analyzed. The instrument is not equally efficient in converting incoming flux into electron counts for all considered wavelengths. This means that besides correcting for total flux efficiency

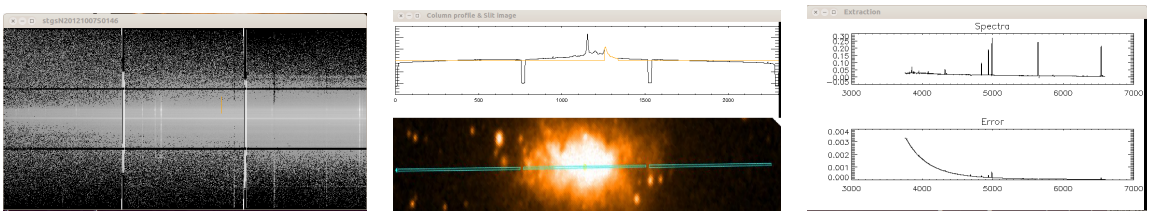


Figure 3.4: Interface for the inspection and extraction of the data. *Left:* Individual frame where the extraction region is defined. *Middle:* Galaxy spactial profile and physical image with slit position (extracted region in yellow). *Right:* Extracted spectra and corresponding error

(which is done with the flat fields) it is also necessary to correct for the wavelength sensitivity of the instrument. This calibration is done with standard stars, for which the flux is well established, observed in the same conditions as the science to be calibrated. The standard stars are reduced the same way as the science frames, differing only on the combination process. Since all images are acquired with the same telescope orientation and correspond to the same physical object, all available frames are averaged, not summed. Pixels that deviate more than 3 sigma from the mean of that particular pixel are not included in the average, increasing the rejection of cosmic rays. For the standard stars the overscan area was used instead of a bias frame, since there was none available in the Gemini archives that would be suitable.

At the end of the reduction, the extracted spectrum of the standard star is binned to match the tabulated flux and the calibration factor is calculated point to point:

$$C = 2.5 \log \left(\frac{observed\ counts}{exposure\ time \times bin\ width \times tabulated\ flux} \right) \tag{3.1}$$

This calibration factor is then fitted with a smooth function, which will be used as the sensitivity function of the detector. This function is interpolated in order to match the observed spectra resolution and used to calibrate in flux that input spectra acoordingly to the formula:

$$F_{calib} = 0.4 * 10^{F_{obs}\ *exposure\ time\ *\ C^{-1}} \tag{3.2}$$

This is not an absolute flux calibration, but only a 'normalization' calibration, that corrects the shape of the continuum but does not provide an absolute energy scale.

The five galaxies that are part of the STARLIGHT observation program in GMOS were observed with the same grating but at two different central wavelengths - 5200 and 5250 Å - and so two flux calibrations were needed. In the science program two standards were available: EG 131 and BD+28 4211.

EG 131 (5200 Å)

EG131 standard star, observed with a central wavelenght of 5200 Å was included in the observation program and reduced as described above. The magnitude values used in the calibration process were available at Bessell [2007]. The final result can be seen in Figure 3.5.

The reduced spectra deviated significantly from the tabulated standard flux, probably due to the poor quality of the acquisition, that had a low number of counts. To perform an accurate flux calibration a

good signal to noise is needed, so this could explain the flux deviation. Because a precise flux calibration is needed to perform the data analysis, this calibration was not used.

BD+28 4211 (5200 Å)

Observations of the standard BD+28 4211 from three different months were used for the calibration of the galaxies with 5200 Å central wavelenght. The stars were treated individually and three different sensitivity functions were calculated. The individual calibration was used to calibrate each star (see Figure 3.5) with good results. As all sensitivity functions were similar a final average function was produced and used to calibrate the science data (see Figure 3.7).

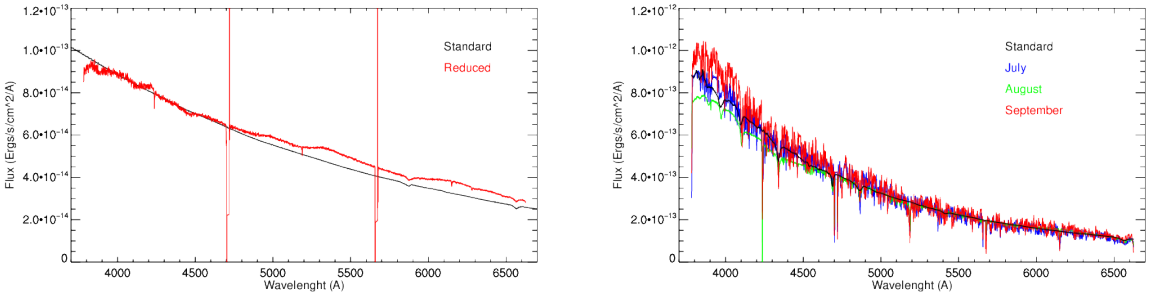


Figure 3.5: Flux comparison of 5200 Å central wavelength standard stars: *Left* - EG131. *Right* - BD28 + 4211 (observed in July, August and September). Observed fluxes were normalized to the standard flux.

Hiltner 600 (5250 Å)

For the 5250 Å central wavelength there was no standard available on the science program. Searching in the Gemini Science Archive an observation of the standard star Hiltner 600 was found that matched the instrument setup used in the project. The procedure followed was the same as the one mentioned above. In Figure 3.6 the normalized flux of both the reduced star and the tabulated flux are plotted. The reduced standard was smoothed to match the resolution of the tabulated spectra from [Hamuy et al., 1992], that has a much lower resolution.

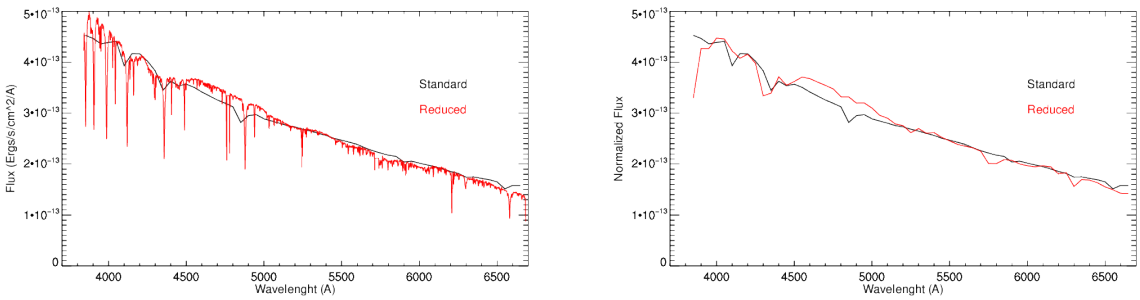


Figure 3.6: *Left* - Observed and stantadard Hiltner 600 flux. *Right* - Normalized and smoothed

In Figure 3.7 both final sensitivity functions used to calibrate the galaxies spectra are plotted.

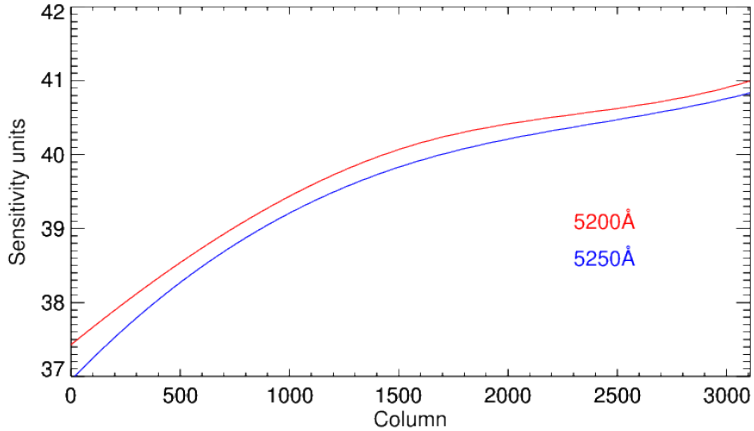


Figure 3.7: Sensitivity functions

Estimating flux calibration error

The systematic error was accessed by calculating the difference between the flux obtained with the calibration and the tabulated values. This difference was normalized by the tabulated flux value for each wavelenght value and smoothed in order to interpret any visible trend. The result can be seen in figure 3.8 and it is clear that there is not a systematic desviation in either calibration.

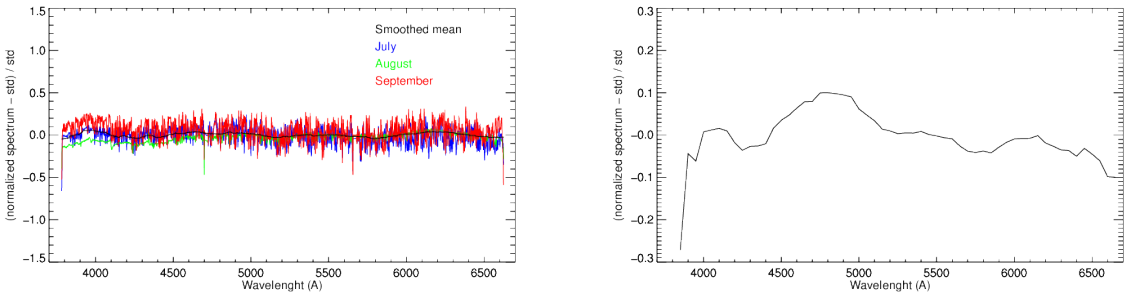


Figure 3.8: Normalized deviation to the tabulated flux. *Left* - BD28 (5200 Å) *Right* -Hiltner 600 (5250 Å)

The uncertainty was only possible to access for the BD28 standard star, since only for this star there were three measurements available. To estimate the uncertaities, the difference spectra and its mean (on the left panel of figure 3.8) were smoothed and the distance of each of the three differences to the mean was calculated. The maximum and minimum values for each wavelenght of these distances were taken as the upper and lower uncertaity values of the flux calibration (see figure 3.9). The acuracy of the calibration corresponds to the mean of these values. Flux calibration is accurate to less than 5% in the redder wavelengths (λ >5000Å) and to about 15% in the bluer part ($\lambda <$ 5000 Å).

3.2.6 Foreground galatic extinction correction

After being flux calibrated, the spectra need to be corrected for the galactic reddening caused by the dust present in the interstellar medium (ISM) of our own galaxy. The general form of this correction is:

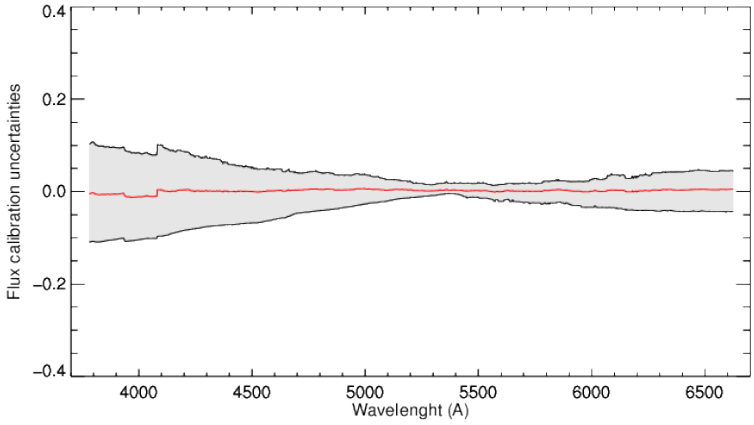


Figure 3.9: *In Grey*: Uncertainty *In Red*: Accuracy

$$F_{dered} = F_{obs} 10^{0.4A(\lambda)} \tag{3.3}$$

where F_{dered} is the corrected flux, F_{obs} is the reddened spectra to be corrected and $A(\lambda)$ the attenuation factor, that depends on the dust composition and size. Interstellar dust is composed by the heavy elements that stars have synthesized and are released to the ISM at the final stages of their lives. When these remains cool down, they form dust grains that will scatter and absorb passing light, with an intensity that is wavelength dependent. The attenuation factor is often stated as:

$$A(\lambda) = E(B-V)[k(\lambda - V) + R(V)] \tag{3.4}$$

where $E(B-V)$ is the colour excess between the observed (B-V) and the intrinsic (B-V) (the magnitude difference between the blue (B) and visible (V) band), $k(\lambda - V)$ is the normalized extinction curve and $R(V) = E(B-V)/A(V)$. It is this extinction curve $k(\lambda)$ that describes the physics of the dust interaction with the light. However there is still no theoretical model that can reazonably describe it from physical principles. Instead, empirical laws are used to correct for the galactic reddening. In this work, the Fitzpatrick and Massa [1990] parametrization was used.

This reddening correction assumes that the dust attenuation curve is independent of the location of the target on the sky. In a way, it assumes that the dust properties are homogeneous within our galaxy. The different amounts of reddening that targets at different lines-of-sight suffer is given by the extinction, the $E(B-V)$ parameter. In this attenuation description, the extinction corresponds to the amount of dust between us and the source. This quantity has to be empirically measured for each area of the sky. The $E(B-V)$ values for each galaxy were taken from Schlafly and Finkbeiner [2011].

3.2.7 Redshift correction

Finally, wavelength axis is corrected for the doppler shift (to the red or blue part of the spectra) caused by the relative motion of the galaxies. Literature values for the redshift were used to perform this correction.

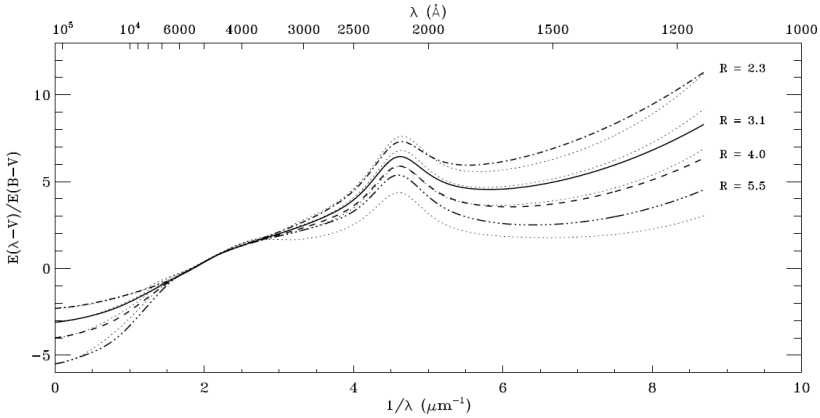


Figure 3.10: Examples of the R-dependent far-IR through UV extinction curves. *Dashed lines*: from Fitzpatrick 2008. *Dotted lines*: from Cardelli, Clayton and Mathis (1988)

$$\lambda_{corrected} = \lambda_{observed}/(1+z) \tag{3.5}$$

3.2.8 Error Propagation

During the first part of the reduction, when dealing with 2D spectra images, errors were calculated pixel by pixel, and a variance image (the VAR extensions referred in section 3.2) was produced and propagated throughout all steps.

The intrinsic error of the initial science frames is the sum of the read out noise plus poison noise. The poisson noise is the square root of the number of electron counts - the ones corresponding to incoming photons of the source - and is estimated to be the square root of the number of electron counts in the raw image minus the bias electrons added. The initial variance (σ^2) is then:

$$\sigma^2 = ron^2 + (Sci - Bias) \tag{3.6}$$

where ron is the read-out-noise, Sci the number of electron counts in the science frame and $Bias$ the counts in the bias frame. From this starting point the variance image is propagated, through usual error progagation. The following table summarizes the errors added to the original variance image throughout the data reduction:

Transformation	Associated Variance
Bias subtraction	σ^2_{Bias}
Flatfield division	$(\frac{Sci-Bias}{Flat^2})^2\sigma^2_{Flat}$
Sky Subtraction	$(rms_{fit})^2$
Frame Sum	$\sum_i \sigma_i^2$
1D Extraction	Extraction in the same area

When the extraction is reached, the variance image is extracted exactly as the science frame and is also transformed in a 1D spectra. After the extraction, the spectra are flux calibrated and the errors are propagated:

$$\sigma^2_{calib} = \left|\frac{1}{total\ exposure\ time\ *C}\right|^2 \sigma^2_{extracted} + \left|\frac{F_{extracted}}{(total\ exposure\ time\ *C)^2}\right|^2 \sigma^2_C \tag{3.7}$$

where σ^2_{calib} is the error of the flux calibrated spectrum, $\sigma^2_{extracted}$ the error of the extracted spectrum (last step in the previous table) and σ^2_C is the error of the sensibility function. Once this last error was not available, the uncertainties added by the flux cablibrations were intruduced via the estimation done in section 3.3.5. The previous equation was then substituted by:

$$\sigma^2_{calib} = \left|\frac{1}{total\ exposure\ time\ *C}\right|^2 \sigma^2_{extracted} + F_{calib} * f_{uncertainty} \tag{3.8}$$

where F_{calib} is the flux calibrated spectra, the result of the extraction, and $f_{uncertainty}$ is the difference between the maximum and minimum deviation, that corresponds to the height of the grey area in figure 3.9.

As for the reddening of the spectra, the usual error propagation could be used:

$$\sigma^2_{dered} = \left|\frac{F_{dered}}{F_{calib}}\right|^2 \sigma^2_{calib} \tag{3.9}$$

where σ^2_{dered} is the error of the final derenened spectrum.

3.3 Extracted Spectra

To perform the wavelength calibration and the galactic extinction correction, literature values for the redshift and galactic extinction in the V band were used (see Table 3.1). The signal-to-noise ratio (S/N) was estimated in a blue wavelength interval, aproximatelly from 4100 to 4200 Å and in a red interval, aproximatelly from 6100 to 6200 Å. The regions were chosen in order not to contain absorption or emission lines (except in the NGC2366 case). The lower value of the Red S/N for NGC2366 is due to the sky emission lines present in the interval. With the exception of NGC2366, where it was not possible to conveniently remove sky emission lines, all targets have a better S/N in the red window than in the blue.

Galaxy	Classification	Redshift	A(V)	Distance (Mpc)	Blue S/N	Red S/N
NGC 1569	Irregular	-0.00347	1.903	2.896	41	56
NGC 6789	Blue Compact	-0.000470	0.187	3.660	33	20
NGC 2366	Irregular	0.000267	0.100	3.567	22	38
UGC 4483	Irregular	0.000519	0.093	3.440	13	31
UGC 4459	Irregular	0.000067	0.103	3.182	7	20

Table 3.1: Columns 2 to 5 are literature parameters used in the reduction process. S/N was estimated in the final spectra. Blue window: 4100 to 4200 Å Red window: 6100 to 6200 Å

In the UGC4459 case, the first extractions revealed a spectrum typical of an older (redder) galaxy, unlike the previous ones. After a more careful analysis of the individual frames, a foreground star was found to be contaminating the spectra. Figure 3.11 shows both spectra.

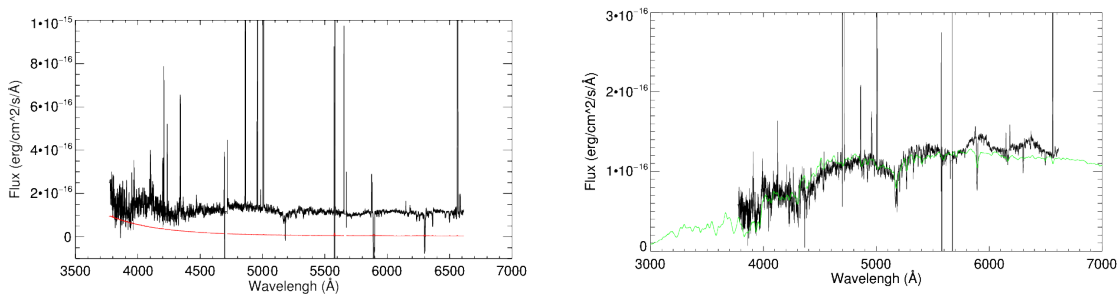


Figure 3.11: UGC4459: *Left* - Extraction with contamination; *Right* - Contaminating star (black) and K2V template spectrum (green). Template from Pickles Atlas

Comparing the extracted star spectrum with templates for the several types of stars, we have have concluded that it is most probably a type K star. Visually comparing the extracted spectrum with star subtype templates, the star appears to be a K2V, a red dwarf. The spectrum of the star was masked in the individual frames and a new extraction was made.

The reduced spectra and the corresponding extraction sky areas, that will be used by the ANGST team to calculate the specific star formation histories for the presented spectra, are displayed from figure 3.12 to 3.16. These areas were calculated with the GMOS-N observing coordinates, the observational offset and the extraction limits in the summed frames. A brief description of the known relevant properties of the galaxies is also made.

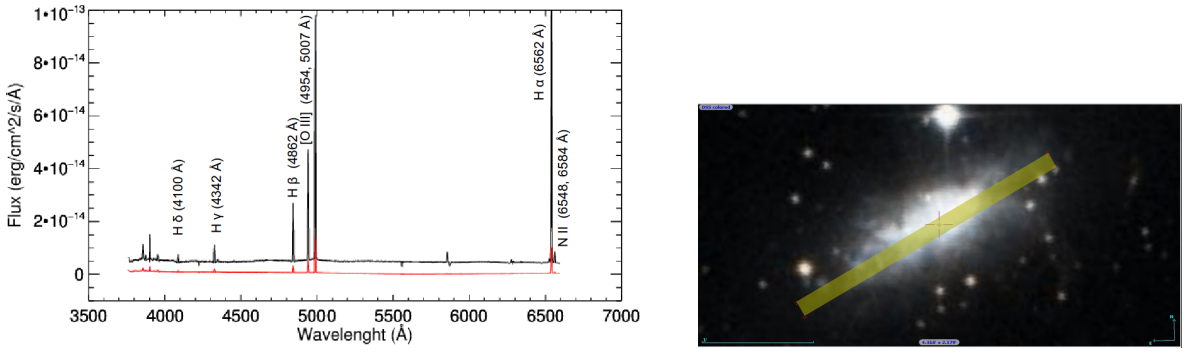


Figure 3.12: NGC1569: *Left* - Calibrated extracted spectrum and error (black and red, respectivelly); *Right* - Extraction area superimposed on a Digitized Sky Survey (DSS) image.

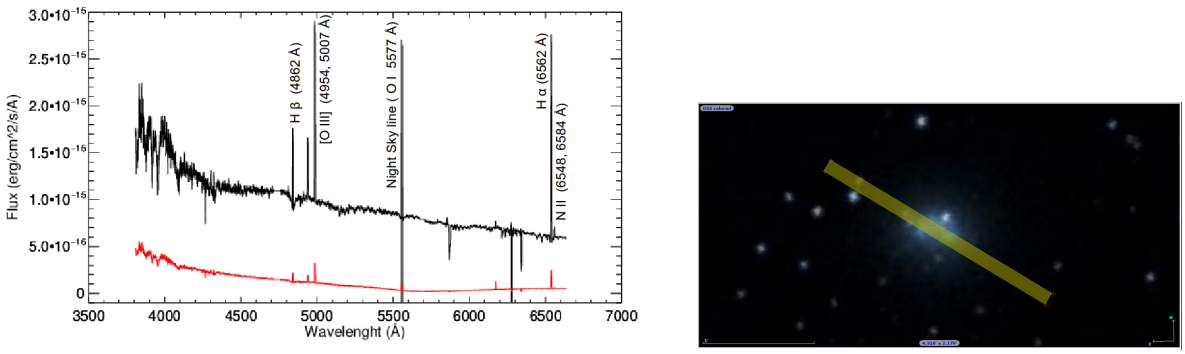


Figure 3.13: NGC6789: *Left* - Calibrated extracted spectrum and error (black and red, respectivelly); *Right* - Extraction area superimposed on a Digitized Sky Survey (DSS) image.

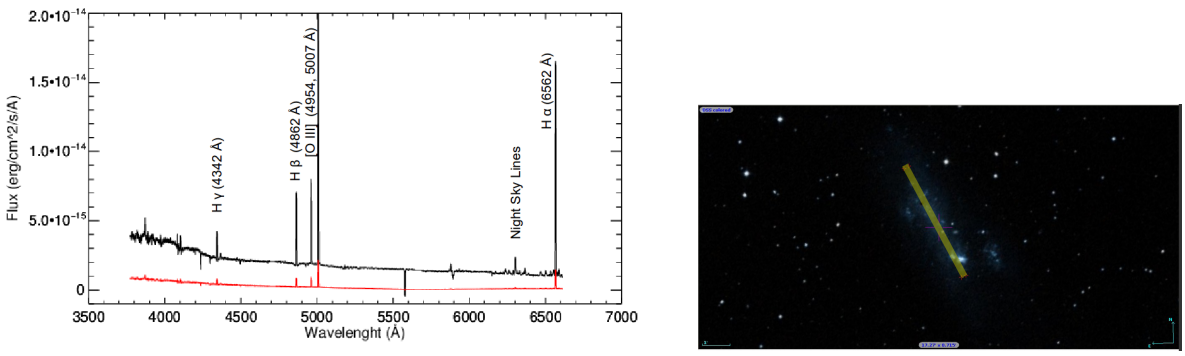


Figure 3.14: NGC2366: *Left* - Calibrated extracted spectrum and error (black and red, respectivelly); *Right* - Extraction area superimposed on a Digitized Sky Survey (DSS) image.

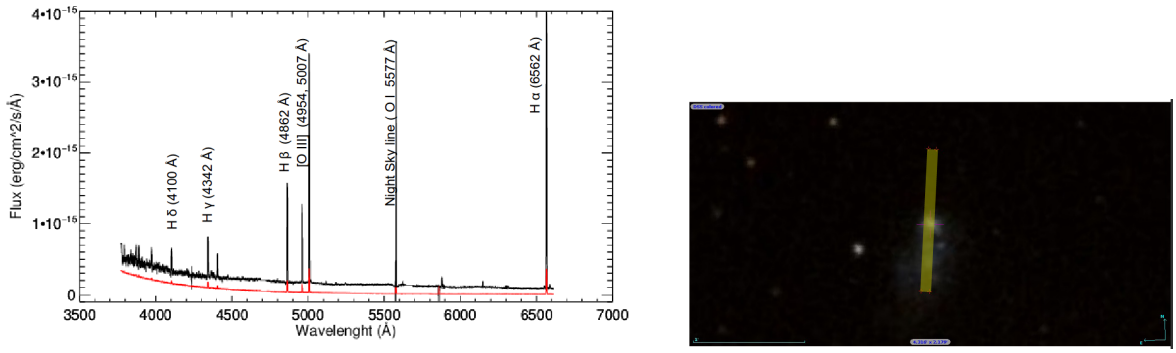


Figure 3.15: UGC4483 Extraction *Left* - Calibrated extracted spectrum and error (black and red, respectivelly) *Right* - Extraction area superimposed on a Digitized Sky Survey (DSS) image.

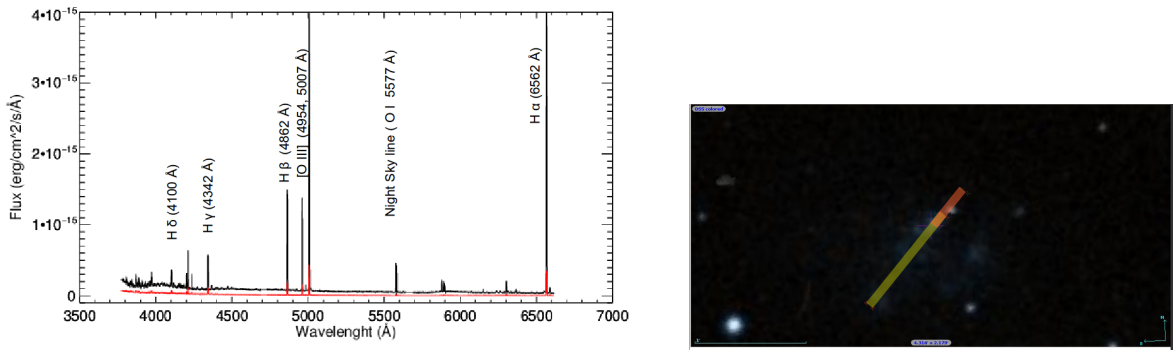


Figure 3.16: UGC4459: *Left* - Calibrated extracted spectrum and error (black and red, respectivelly); *Right* - Extraction area and mask (in red) superimposed on a Digitized Sky Survey (DSS) image.

Chapter 4

STARLIGHT code

The next step is to analyze the processed data by fitting the observed spectrum with one or more models to infer the overall properties. There are several available codes thar perform full spectral fitting, such as the ULySS [Koleva et al, 2009], the VESPA [Tojeiro et al. 2007] and the MOPED [Heavens et al. 2000]. The software STARLIGHT (Fernandes et al. 2005) was selected for this task. In this section the method by which the reduced spectra were analyzed is described and the results presented

4.1 STARLIGHT spectrum model

STARLIGHT is an algorithm that performs full spectral fitting, finding the best spectrum model that fits an observed spectrum. It follows an algebraic approach to the problem, where the model is a linear combination of individual spectra (L_λ - luminosity per wavelenght unit λ), attenuated by galactic dust and convolved with a kinematic factor. The implemented model is fundamentally the one described in equation 1.7 (section 1.3.2) with some technical modifications.

In the general model described, the linear combination of spectra was given by the product $c_j F_j$ (plus the dust and velocity factors). In STARLIGHT, normalized base spectra (b_λ) are used:

$$b_{\lambda,j} = \frac{L_{\lambda,j}}{L_{\lambda_0,j}} \tag{4.1}$$

where $L_{\lambda_0,j}$ is the luminosity of the spectra j at wavelenght λ_0 (that is defined by the user). The model is then defined as:

$$M_\lambda = \sum_j^n b_{\lambda,j} L_{\lambda_0,j} r_{\lambda,j} \otimes f(v_j, \sigma_j) \tag{4.2}$$

where r_λ is the extinction law and $f(v_{SSP,j}, \sigma_{SSP,j})$ the kinematic factor. The normalization is also done in the spectra to be fitted in a small window around λ_0. The objective of this normalization is to avoid dealing with absolute flux calibration, that are hard to recover from the observations.

Kinematics

To simplify the fitting process, the kinematics are simplified to a global behavior, where it is assumed that the stars' velocity can be reasonably described by a Gaussian distribution, defined by the mean velocity of all stars in the galaxy and its dispersion. These considerations result on the following mathematical model, for SSP basis:

$$M_\lambda = \sum_j^n b_{\lambda,j} L_{\lambda_0,j} r_\lambda \otimes G(v,\sigma) \tag{4.3}$$

The convolution with the gaussian filter $G(v,\sigma)$ is done in the following way:

$$I(\lambda) = \int F\left(\lambda' = \frac{\lambda}{1+\upsilon/c}\right) G(\upsilon; v, \sigma) d\upsilon \tag{4.4}$$

where $I(\lambda)$ is the convolved spectra and $G(\upsilon; v, \sigma)$ is the gaussian function of mean v and dispersion σ.

If the models used as base elements have a different spectral resolution than the input spectra, the velocity dispersion given by the fit do not correspond to the physical velocity dispersion on the galaxy, since the broadening will also account for the resolution difference. It is possible to correct this value, but since the focus of this work is on the SSP, the velocities will not be studied.

Dust treatment

There are several reddening-law options available in STARLIGHT such as the Cardelly, Clayton and Mathis (1998), Calzetti et al. (1994), Gordon et al (2003) and Allen (1976), among others. For this part of the analysis, the Calzetti law [Calzetti et al., 1994] was chosen, since it was empirically determined from low-redshift and metal-poor starburst galaxies, which are characteristic the sample we are analysing. The attenuation is introduced as:

$$r_{\lambda,j} = 10^{-0.4E_s(B-V)k'(\lambda)} \tag{4.5}$$

where B is the flux in the blue band, V in the visible and $E_s(B-V)$ is the color excess in the stellar continuum given by $(B-V)_{observed} - (B-V)_{intrinsic}$. The $k'(\lambda)$ factor is given by:

$$k'(\lambda) = 2.659(-1.857 + 1.040/\lambda) + 4.05\, for 0.63\mu m \leq \lambda \leq 2.20\mu m \tag{4.6}$$

$$k'(\lambda) = 2.659(-2.156 + 1.509/\lambda - 0.198/\lambda^2 + 0.011/\lambda^3) + 4.05\, for 0.12\mu m \leq \lambda \leq 0.63\mu m \tag{4.7}$$

In STARLIGHT, the intrinsic extinction law is written as:

$$r_{\lambda,j} = 10^{A_V q_{\lambda_0} + A_V(q_\lambda - q_{\lambda_0})} \tag{4.8}$$

where the $q(\lambda)$ factors are the extinction curve $k(\lambda)$ and $A(V)$ the attenuation factors, that are the equal for all stellar populations. This is a disadvantage, but it was still a technical challenge to allow such a great number of free parameters to be determined. With this extinction law, the model spectrum is written as:

$$M_\lambda = \sum_j^n [L_{\lambda_0,j} 10^{-0.4A_{\lambda,j} q_{\lambda_0}}] b_{\lambda,j} 10^{-0.4A_{\lambda,j}(q_\lambda - q_{\lambda_0})} \otimes G(v,\sigma) \tag{4.9}$$

The term in brackets corresponds to flux of j basis spectra attenuated by dust at the normalization wavelength. It corresponds to the light fraction contribution of the element j to the total spectra. Defining x_j as:

$$x_j \equiv L^0_{\lambda_0,j} 10^{-0.4A_{\lambda,j}(q_{\lambda_0})} \tag{4.10}$$

we arrive to the final model that will be fitted:

$$M_\lambda = \sum_j^n x_j b_{\lambda,j} 10^{-0.4A_{\lambda,j}(q_\lambda - q_{\lambda_0})} \otimes G(v,\sigma) \tag{4.11}$$

Turning once again to the general model stated by equation 1.7 the parallelism with the $c_j F_j$ elements can be made. In STARLIGHT, these correspond to $b_{\lambda,j} x_j$, the j base element and its contribution to the total flux, respectively.

4.2 Numerical Details

The fitted parameters are the extinction $A(V)$, the velocity v, the velocity dispersion σ, and the x_j spectral models. The data to be fitted is an observed spectra, reduced and flux calibrated. The fitting algorithm is a mixture of several techniques with the goal of minimizing the difference between the observed spectra and a model:

$$\chi^2 = \sum [(O_\lambda - M_\lambda) e_\lambda^{-1}]^2 \tag{4.12}$$

where O_λ is the observed spectrum to be fitted, M_λ is the model described in the previous section and e_λ^{-1} the weight on the total χ^2 of this difference at λ. The e_λ^{-1} factor translates how significant a particular spectral region is for the total fit. The weight factor is included because even within one single spectrum data quality varies with wavelength. In this work, the inverse of the error propagated during the reduction process was used as e_λ^{-1}.

STARLIGHT's fitting method is a hybrid of Metropolis and Markov Chain algorithms with the goal of approximating a function (in this case M_λ) of the desired distribution (O_λ) by iteratively updating the parameters that determine the function (in this case x_j, A_V, v and σ). A Metropolis technique is used to fully probe the parameter space, by randomly modifying the parameters' values. The new values are kept if the obtained χ^2 is better than the previous one. This makes the process converge, despite the random choice of the new parameter update. The size of the difference between a new parameter's value and the previous one varies from iteration to iteration. The fitting process starts with big differences, to avoid that the system converges at a χ^2 local minimum, and diminishes as the fitting progresses, to achieve convergence. The Metropolis algorithm has the disadvantage of converging very slowly, especially given the number of parameters to be fitted. For this reason, Markov chains are also used, where the new parameters state are only influenced by the previous one. The parameters are then probed in a

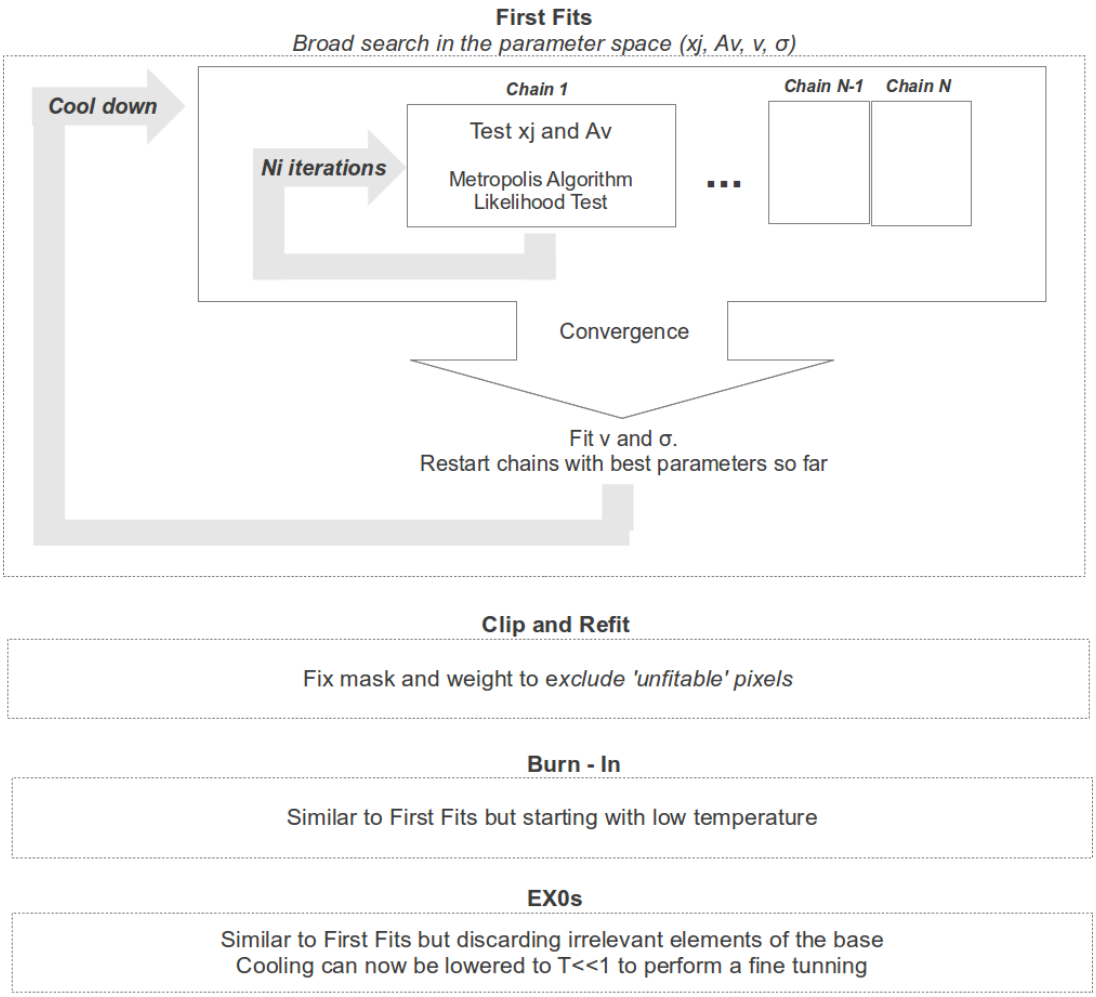


Figure 4.1: Starlight's simplified fitting process

Metropolis random order but in a region of the parameter space determined by the previous state of the system (Markov Chain technique). To further tune the fitting, there are n independent Markov Chains involved in the process. The final convergence criteria depends on the comparison of the typical variance of a parameter within each of the chains with the variance of the parameter between all chains. These variances must be alike for convergence to be reached.

A simplified version of the entire process can be seen in figure 4.1. In the first step of the fit, the *First Fits*, a set of n chains is set to evolve in Ni iterations (7 chains were used in this work). Because convolution with the kinematic kernel is quite time consuming, only the x_j and A_V parameters are modified in these chains. Once the Ni iterations are finished, the v and σ values are fitted with the best model. The previous step is then repeated, starting with the best values of all parameters from the last iteration. In this new run of iterations the size of the random steps are diminished, by reducing the factor by which the weights are multiplied.

Once this first chain iterations are done, the system move to the *Clip and Refit* phase. In this stage, points that could not be appropriately fit in the *First Fits*, that significantly deviate from the best M_λ found, are rejected.

The next stage, the *Burn-In*, is much similar to the first step, with the difference that the raw value of the errors are used to calculate the χ^2, so the convergence criteria is tighten than in the *First Fit* phase. The best values achieved in the *First Fits* are used as the starting point.

Finally, the *EX0s* phase - fits with a condensed base - is reached. Once more the fitting process is much similar to the one of the *First Fits*, but with different temperature and number of iterations. The difference in this stage is that it is possible to discard irrelevant base elements. There are two ways of defining an irrelevant base element. The first is if the light contribution (the x_j) of that element is lower then a defined threshold, that element is not going to be used. The second is to exclude the elements with the smaller light faction, until its total contributions sums up to the defined threshold. In this work the last option was used. After this, the a final fit is done and the spectrum model produced.

4.3 Model spectrum components

STARLIGHT is quite flexible in the choice of the stellar population models for base elements. Any spectrum can be included in the base, as long as the data file is written in the conventionalized way. For this analysis, SSP spectral models from Bruzual and Charlot [2003] were used as base elements. The base was consisted in 45 different base elements, divided into 3 metallicities and 15 ages. The age range covers SSP from 1 Myr to 13 Gyr and three metallicity values: subsolar (0.004), solar (0.020) and above solar (0.050). All spectra have a resolution of 3 Å across the whole spectral range.

As described in the introduction (section 1.4.3) the three main ingredients to construct a population synthesis code are an initial mass function (IMF), an evolutionary model and a spectral library. The IMF adopted to compute the base elements was the Chabrier (2003b). To diminish the uncertainties introduced by stellar evolution models, several were used. Together, they cover the main sequence up until the asymptotic giant branch (AGB) for low and intermediate mass stars. Evolutionary tracks for the late AGB phase were also included in the model. All models assume solar abundance distribution of the elements. The stellar spectra used come from a mixture of theoretical and empirical libraries - BaSel, STELIB and Pickles libraries - were used.

4.4 Light to Mass fractions

Besides light fractions, STARLIGHT also computes the mass fractions (μ_j) of each SSP population. If absolute flux calibration is available, it is possible to convert the light contributuion into mass contribution, trough the mass-to-light ratios ($M = L_{observed} * (M/L)$). The base elements used to perform the fitting were produced in $L\odot$ Å $^{-1} M\odot^{-1}$ units, which means that they are L/M ratios and can be used to calculate the total stellar mass. STARLIGHT calculates the total mass that was converted into stars in all bursts in the units of the input spectra. To convert the this mass into true mass the following relation is used:

$$M_{\star\ ini} = M_{SL} * 4\pi d^2 * L\odot^{-1} \tag{4.13}$$

where M_{SL} is the mass given by the fit, d is the distance of the galaxy (in cm), and $L\odot$ is the sun's luminosity in erg/s. When stars die, they lose mass to the intertellar medium, so the total mass of a SSP diminished with time. The current stellar mass of a SSP is then given by:

$$M_{\star\ cur} = M_{\star\ ini} * f_{\star} \tag{4.14}$$

where $f_{\star}$ is the mass fraction of stars that is still in the SSP and was not yet converted again into interstellar gas. This fraction is calculated in the stellar evolution models and its values for all base elements are available. The initial an current masses of the galaxies were calculated following the above equations. However, these masses do not correspond to the galaxies true mass. Firstly because the integrated spectra only cover a fraction of the galaxies, and not to the entire galaxy. Secondly, because the absolute flux calibration acuracy is still to be determined, so the used values might not be correct.

Chapter 5

Results

The five reduced spectra were fitted with STARLIGHT and the obtained model spectra are presented in this section. Emission lines and the spectral features caused by GMOS gaps were masked in the input spectra. The plotted residuals were calculated as the difference *observed spectrum − model spectrum* , normalized to the observed spectrum. Both observed and model spectra were severely smoothed to calculate this residual, so that it depicted the continuum and not spectral lines differences. Some of the fitted quantities are also plotted in the graphics: the χ^2, as defined in equation 4.12, the *adev*, which is the flux deviation of the model to the data in percentage, and the fitted attenuation $A(V)$, velocity dispersion σ (the uncorrected value) and velocity v of the model. The Calzetti attenuation law was used allowing the factor $A(V)$ vary between 4 and -1 magnitudes, not imposing any physical constrains in order to freely test the method. The SSP light fractions (the x_j quantities) with the respective age and metallicity are also presented (visualization purposes only light fractions greater than 1% are displayed). Once more, the metallicities values will also not be studied in detail, since this is out of the scope of this work.

The mass fractions derived from the STARLIGHT results are also compared with the ones available in the literature. This comparison is still preliminary, since the literature values used are relative to the entire galaxies and not only to the area where the spectra was integrated. The star formation regions in these galaxy are quite strong and localized and consequently the star formation history of the galaxy may vary significantly from region to region. Moreover, the slit was typically centered in the central or most luminous part of the galaxies, and it is natural that more young stellar populations are found that what would be expected for the entire galaxy.

5.1 Derived masses

The galaxy masses calculated from the results and the number of recovered SSP can be found in table 5.3. Usually the current mass is two times lower than the initial mass of a galaxy ($2M_{\star\ cur} \sim M_{\star\ ini}$) [Asari et al., 2007]. In this work, an average mass relation of $M_{cur} \sim 1.4M_{ini}$, was found, which suggests that less stellar mass that was expected has returned to the ISM. This can be a consequence of a high

Galaxy	$M_{ini}(M\odot)$	$M_{cur}(M\odot)$	SSP	SSP $x_j < 1\%$
NGC1569	1.70×10^7	1.36×10^3	16	3
NGC2366	8.41×10^5	4.99×10^5	7	2
NGC6789	7.26×10^5	4.59×10^5	7	0
UGC4459	3.02×10^4	2.07×10^4	8	5
UGC4483	1.02×10^4	1.98×10^3	5	0
Mean	3.73×10^6	1.98×10^3	8.6	2

Table 5.1: Column 2: Initial mass from STARLIGHT results; Column 3: Current Mass from STARLIGHT results; Column 4: Number of SSP spectra used in the final spectrum model of the galaxy. Column 5: Number of SSP spectra that contributed less than 1% to the total light.

number of young populations, that have not yet returned much of its mass to the interstellar medium. The only exception to this scenario is NGC1569 that has a much lower current mass that what would be expected. Not surprisingly, the recovered masses were one to four orders of magnitude lower to the ones calculated from CMD analysis. Excluding once more NCG1569, for which a higher number of SSP were recovered, the mean number of 6.75 base elements were used in the model.

Two criteria were defined to examine the accuracy the results: if the number of recovered bursts - periods where a considerable high fraction of mass was created - and their ages matched the CMD values, and if the relative mass contribution in each time bin was comparable with the CDM results. The relative mass fractions created in each time bin were calculated for the STARLIGHT results and estimated from the star formation histories of the CDM analysis. To get a first notion of the method accuracy only three age bins were defined: a young age, with SSP up to 300 Myr, an intermediate age from 300 Myr to 1 Gyr, and an old one, with SSP with more than 1 Gyr old. The results of these measures are in table 5.2.

For NCG1569, NGC2366 and UGC4483 the expected bursts were recovered by full spectral fitting, although with some age discrepancies (0.5 Gyr in young ages and 3 Gyr in older populations). For the remaining two galaxies, older bursts with more than 1 Gyr were not found. As for the relative created mass in the 3 defined bins, only NGC2366 shows a good agreement. The other galaxies show in general a bad agreement, typically underestimating the mass contribution in the old age interval and overestimating the created mass in the intermediate age bin.

It is worth noticing that the colour-magnitude diagrams results were estimated using a single slope power law IMF, that has negligible differences from the Kroupa 2001 IMF for these galaxies [Weisz et al., 2011]. From figure 1.8, it can be seen that this IMF predicts less stars with more than 10 $M\odot$ than the used Chabrier IMF. It is then possible that some of the discrepancies between this work relative masses and literature values are due to this difference. Particularly, our results should predict a higher percentage of high mass stars, which might be related with the overestimation of star formation rates for younger populations.

Galaxy	Method	Bursts (Gyr)	Relative Mass Created		
			1 Myr - 300 Myr	300 Myr - 1 Gyr	1 Gyr - 14 Gyr
NGC1569	CDM	0; 0.5; 2; 10	0.106	0.055	0.839
	FSF	0; 1; 2.5; 10	0.003	0.213	0.784
NGC2366	CDM	1; 2; 5; 10	0.040	0.094	0.866
	FSF	0; 4; 5; 10	0.030	0.148	0.822
NGC6789	CDM	0.4; 0.5; 2.5; 10	0.020	0.050	0.929
	FSF	0; 0.5; 1	0.005	0.947	0.069
UGC4483	CDM	0; 1; 8	0.012	0.029	0.959
	FSF	0; 1; 5	0.157	0.153	0.691
UGC4459	CDM	1; 5; 10	0.015	0.036	0.949
	FSF	0	0.348	0.652	0.000

Table 5.2: Results estimated from the SFR of the colour magnitude diagram (CMD) and full spectral fitting (FSF) analysis. *Bursts* Aproximate age in Gyr of the burst in the SFH. *Relative Mass* Relative mass created in each age bin.

5.2 Galaxies individual results

5.2.1 NGC1569

NGC1569 is an irregular dwarf galaxy with low metallicity and abundant H II regions. It underwent a recent episode of strong star formation, between 10-100 Myr ago [Greggio et al, 1998], that may possibly explain its complex and dynamic morphology. A well known morphological feature are two superstar clusters in its central region [Arp and Sandage, 1985]. These type of objects are thought to be the progenitors of globular clusters and are dense gas regions where young massive stars can be found. These stars ionize the surrounding gas forming a HII region, that is in turn surrounded by dust that strongly attenuates the starlight.

Gas outflows are also present in the galaxy. These are composed of large-scale gas superbubles expanding in the outer part of the galaxy [Martin 1998]. There are indications that the superstar clusters may provide enough energy, through type II supernova explosions - very energetic explosions of high mass stars that have exhausted its fuel and unable to maintain the equilibrium between pressure and radiation - to expel the gas from the central part of the galaxy forming the outflow [Martin et al., 2002]. In a spectroscopic study of the Hα line of NGC1569 near the central superstar clusters, Westmoquette et al. [2007] have described a line profile composed of a strong narrow component with an underlying broad component. The physical origin of the broad component is still not fully understood. There are two main theories that try to explain this phenomenon. The broadening may arise from the superimposition of the emission lines of unresolved expanding shells of the outflow. Other possibility is that it may be originated in the turbulent mixing layers between the hot outflow gas of the galaxy and the cool interstellar medium

[Westmoquette et al., 2007]. This last explanation was preferred by the authors.

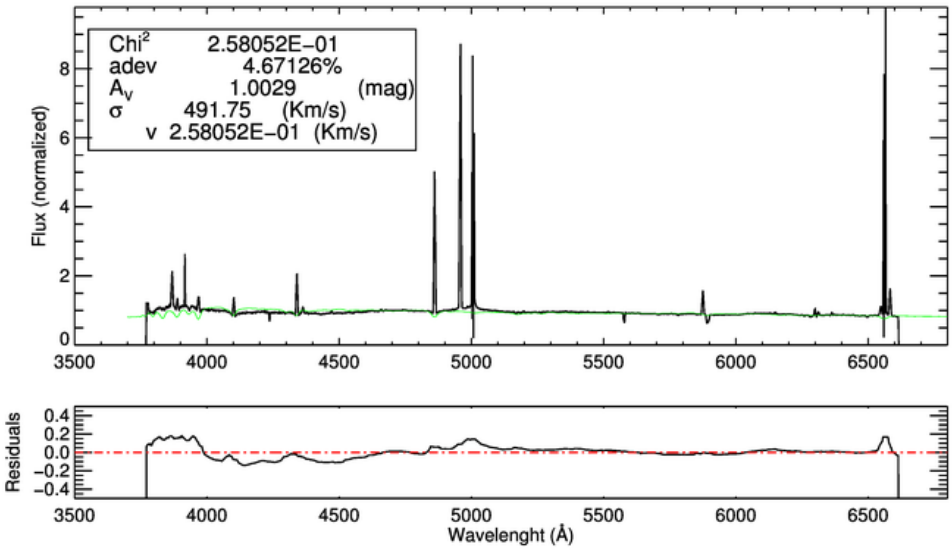


Figure 5.1: NGC1569: Input spectrum in black, model in green and residual (in the lower panel)

The observed $H\alpha$ emission line (see figure 5.1) is extremely strong and other hydrogen lines, up to the $H\delta$, are visible which is in concordance with the existence of a high content of ionized gas. The spectrum has the characteristic continuum shape of a starburst galaxy, that is not accurately reproduced by the model spectrum. Specially in the bluer part of the spectrum (< 4000 Å) the model has $\sim 20\%$ less flux than the observed spectrum. In the regions near the $H\alpha$ and [OIII] lines the model also deviate considerably from the observed spectrum. Analyzing the continuum shape of the input spectrum more closely, a broadening of the lower part of these emission lines are visible, with a profile much in agreement with the work of Westmoquette et al. [2007]. This can explain why the model could not reproduce the spectrum in these regions, since the excess of observed flux is most probably due to gas outflows.

This galaxy was included in the starburst sample of Calzetti et al. [1994] work on intrinsic extinction. Its estimated $A(V)$ was of 0.217 mag. However, this extinction was calculated based on the ratio between hydrogen balmer lines ($H\alpha/H\beta$ or $H\gamma/H\beta$), lines that are produced in the H II regions of the galaxy. As discussed before, this regions are usually more dusty than others in the galaxy, so the extinction values obtained from full spectral fitting may differ from these. Asari et al. [2007] demonstrated that the $A(V)$ values obtained from SSP are usually half of the estimated by emission lines. This would give an estimative of 0.1 mag for the intrinsic extinction of NGC1569. The results from STARLIGHT yield $A(V) = 1.0029mag$, ten times higher. It is quite intriguing why the attenuation estimated value is so high, specially given that the model lacks flux in the bluer part (that is more severely attenuated) when compared with the observed spectrum. This might indicate that there is a problem in the continuum shape, and that no combinations of base elements could conveniently described the observed shape without a strong 'distortion' introduced in this case by the attenuation law.

The model's velocity v found was of 0.23 Km/s, a low value that confirms the accuracy of the wavelength calibration. As for the SSP results (see figure 5.1), only relatively young populations, between

1 to 10 Myr, were found, that do not match the recent star formation period, between 10 and 100 Myr reported by Greggio et al., [1998]. An older populations, with more than 1 Gyr, were also retrieved.

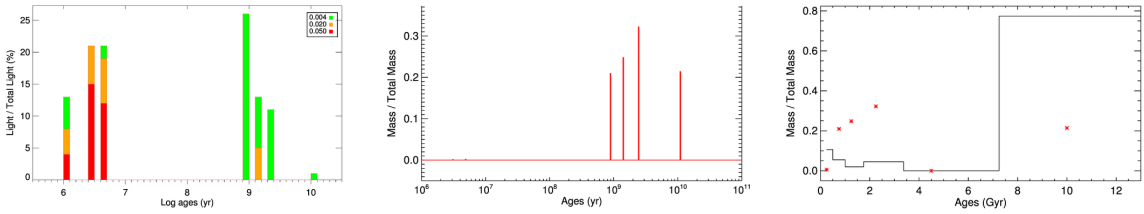


Figure 5.2: NGC1569: *Left* - Light fractions of the simple stellar populations that contributed to the model spectrum (logaritmic time scale). Metallicities are represented in different colours; *Middle* - Stellar mass fractions from STARLIGHT; *Left* - Comparision between the derived stellar mass fractions, from the two methods, in the age intervals 0-0.5, 0.5-1, 1-1.5, 1.5-3, 3-6, 6-14 Gyr. In red stars: FSF; In black: CMD from [McQuinn et al., 2010]

NGC1569 is a good example of the difficulty that deriving the stellar mass created from integrated spectra poses. The SSP found could be divided into two groups - a young one ($\leq$ 10 Myr) and an older one ($\geq$ 1 Gyr) - that have approximately the same contribution to the total light. Observing the mass fraction distribution (middle panel in figure 5.2) the picture changes, and the contribution of the younger population is estimated to less than 1% of the total mass.

Comparing the obtained masses fractions with the results from CMD (left panel of figure 5.2) the relative contributions to the total mass of each period do not match. The mass created in recent ages ($<$ 4 Gyr) was overestimated, whereas in older periods ($>$ 6 Gyr) was underestimated. This suggests that although the method is capable of recovering old stellar populations even when recent bursts are present, it does not correctly estimate its contribution.

5.2.2 NGC2366

NGC2366 is an irregular dwarf galaxy with an interesting structure characterized by a bright small region and a dimmer elongated body, much like a comet [Thuan and Izotov, 2005]. The brightest component is a star forming region, currently undergoing a period of star formation, one of the most powerful in the local universe [Drissen et al., 2000].

The extracted spectrum has the expected starburst profile. As referred in the data reduction chapter, due to the galaxy's large angular size, sky subtraction was performed with an extra frame observed with a large offset from the galaxy center. However, even with this technique the sky subtraction is not as accurate as in the rest of the sample, and sky emission lines are still visible from 6200 Å until the end of the observed wavelength range.

In this galaxy, the model was also not able to fully reproduce the continuum shape, lacking up to 20 % of flux for wavelengths shorter than 4100 Å. The strong emission lines confirm the presence of ionized gas in the galaxy.

The Calzetti et al. [1994] work on starbursts calculated an upper limit for the attenuation of 2.883 *mag*. Correcting the overestimation of this value for an SSP aproach following [Asari et al., 2007], an

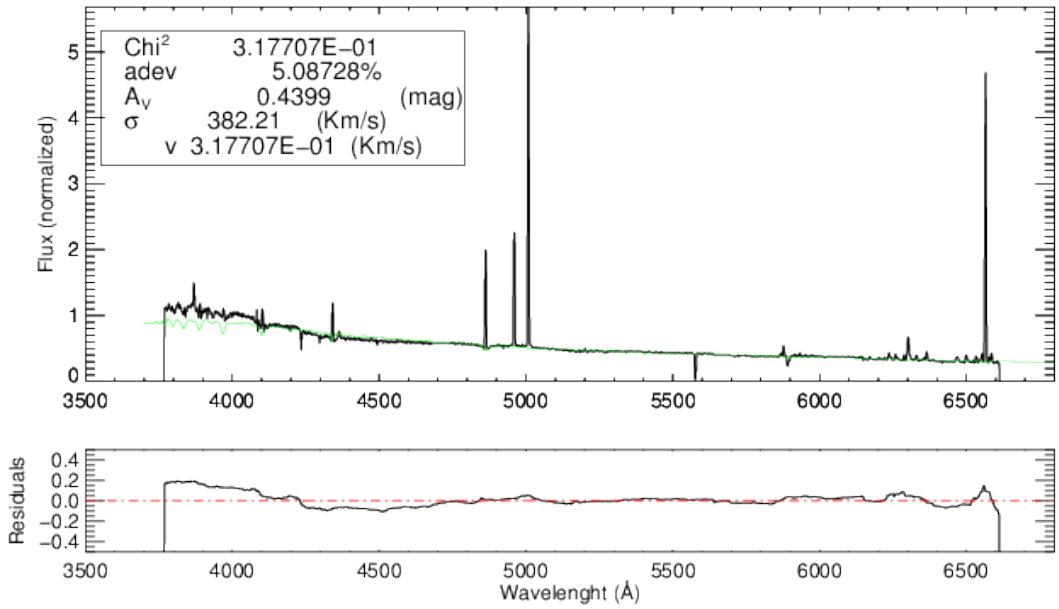


Figure 5.3: NGC2366: Input spectrum in black, model in green and residuals (in the lower panel).

upper value of 1.44 *mag* for the $A(V)$ is estimated. NGC2366 fitted attenuation is 0.4399 *mag* which is safely below that upper limit. The velocity was small, as it was expected.

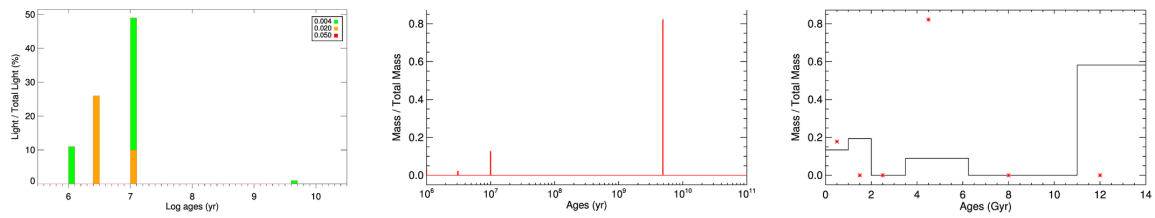


Figure 5.4: NGC2366: *Left* - Light fractions of the simple stellar populations that contributed to the model spectrum (logaritmic time scale). Metallicities are represented in different colours; *Middle* - Stellar mass fractions from STARLIGHT; *Left* - Comparision between the derived stellar mass fractions, from the two methods, in the age intervals 0-1, 1-2, 2-3, 3-6, 6-10, 10-14 Gyr. In red stars: FSF; In black: CMD from [Weisz et al., 2011]

Four young SSPs were recovered (left panel in figure 5.3) with ages between 1 to 10 Myr, that may correspond to the reported star forming central region. A very low light contribution ($\sim$ 1.5 %) from an old population ($\sim$ 5 Gyr) was found.

NGC2366 recovered relative masses (right panel in figure 5.3) do not agree with much precision with the predicted ones. The mass created in recent periods, that CDM results place throughout the last 2 Gyr, has only been found by full spectral fitting in the last Gyr . Star formation around the 5 Gyr period, was overestimated, acconting for nearly 80% of the galaxy's mass, whereas a contribution of $\sim$ 10% was expected. The intermediate age stellar populations (1-2 Gyr) as well as the extremelly old (more than 10 Gyr) present in the CMD derived relative masses were not recovered.

5.2.3 NGC6789

NGC6789 is a blue compact dwarf galaxy, characterized by an irregular central part with large HII regions and bright blue stars. Its core and halo regions have different morphologies and stellar populations [Drozdovsky and Tikhonov, 1999]. The first is characterized by a high surface brightness, with large H II regions, and it is composed by a very young stellar population of type O stars, whereas the latter has low surface brightness and a stellar population with age between 20 and 100 Myr.

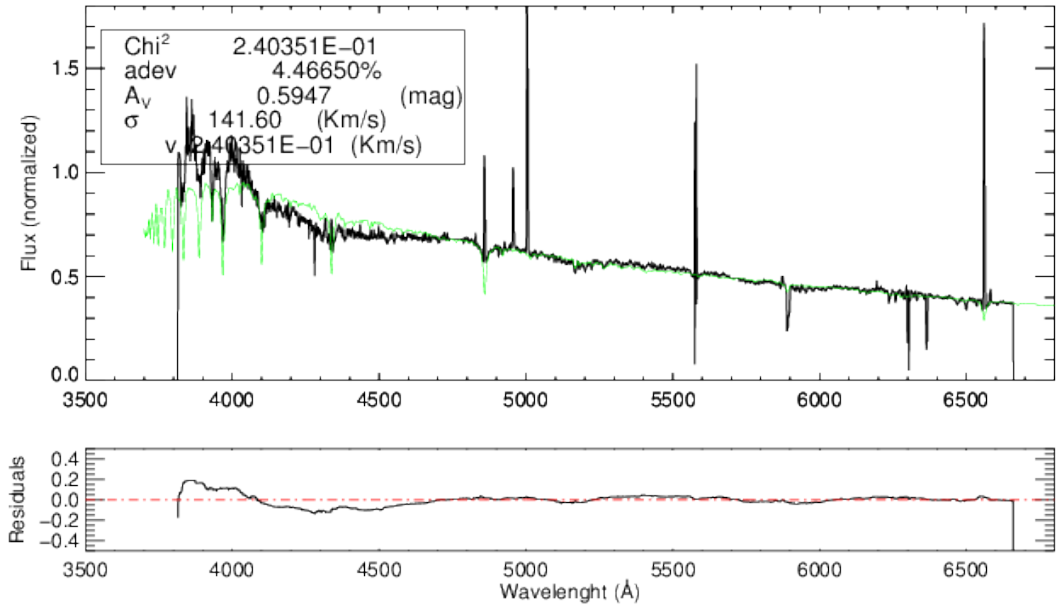


Figure 5.5: NGC6789: Input spectrum in black, model in green and residuals (in the lower panel).

Once more, the bluer part of the spectra was not correctly predicted by the model (see figure 5.5). The fitted $A(V)$ value (0.5947 mag) does not exceed the referred maximum for starburst galaxies. The mean velocity v is also small.

Four young SSP were found with ages between 3 to 8 Myr. These may correspond to the populations of the central part. An intermediate age stellar population was also found, with a lower age limit around 200 Myr. This might correspond to the mentioned stellar population of age 20-100 Myr, demonstrating that the method accuracy in these intermediate ages is low.

The mass fractions reveal two recent starburts, one from stellar populations with a few Myr ($\sim$ 3-5 Myr) and other from stellar populations with hundreds of Myr (see figure 5.6). The colour-magnitude diagram results also show star formation in these ages, although their relative contribution is much less, around 5%, than the estimated 80%. Full spectral fitting results overestimated mass creation in young ages (in the order of Myr) and has found no traces of old stellar populations.

5.2.4 UGC4483

UGC4483 is an irregular dwarf galaxy. It has a young star cluster [Karachentseva et al., 1985], with an attributed age of $\sim$ 10 - 15 Myr [Dolphin et al.,1994]. Metal rich AGB stars have been identified in CMD

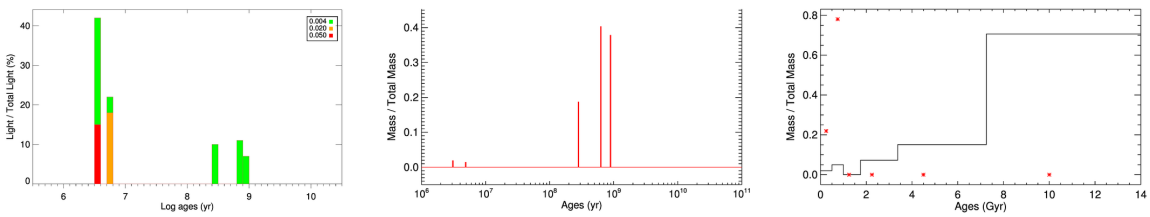


Figure 5.6: NGC6789: *Left* -Light fractions of the simple stellar populations that contributed to the model spectrum (logaritmic time scale). Metallicities are represented in different colours; *Middle* - Stellar mass fractions from STARLIGHT; *Left* - Comparision between the derived stellar mass fractions, from the two methods, in the age intervals 0-0.5, 0.5-1, 1-1.5, 1.5-3, 3-6, 6-14 Gyr. In red stars: FSF; In black: CMD from [McQuinn et al., 2010]

studies.

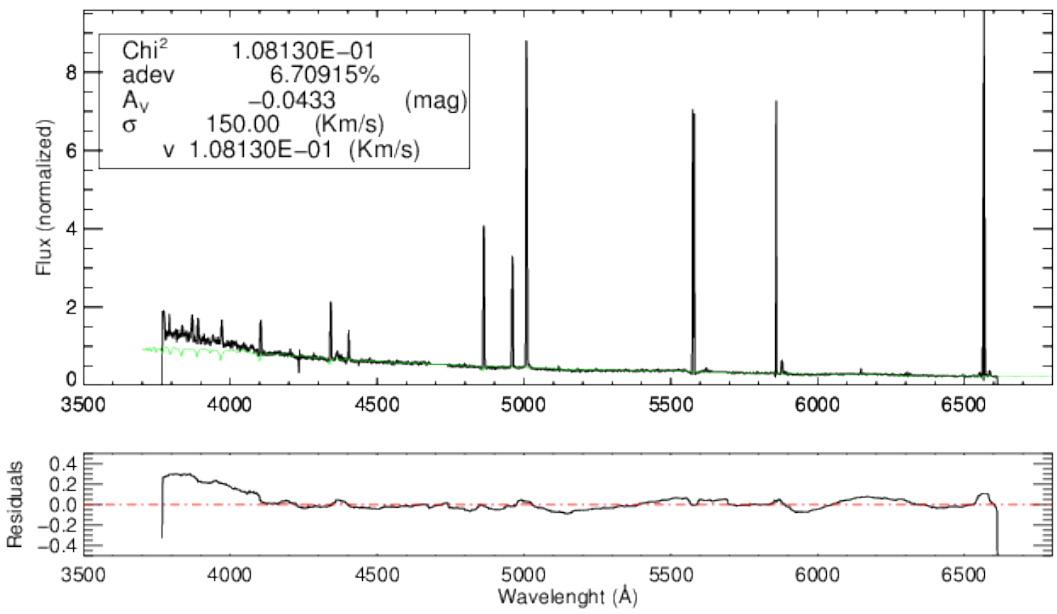


Figure 5.7: UGC4483: Input spectrum in black, model in green and residuals (in the lower panel).

The observed continuum has a starburst shape and the strong emission lines point to the presence of gas and young stars, as it is typical of these galaxies. Once more the model spectrum could not accurately describe the flux in for wavelengths lower than 4200 Å, predicting up to $\sim$ 30% less flux. The $A(V)$ value in this case was negative, which indicates that the model spectrum had actually to be dereddened. Physically, this value may only indicate that the galaxy is very poor in dust, since the difficulty in estimating the dust content results in high $A(V)$ uncertainty. This results rules out the hypothesis that the flux discrepancy between model and spectrum arose from artificially high values of attenuation, since that for a wide range of $A(V)$ values (from 1.0029 to -0.0433 mag) the problem persisted.

A 1 Myr old SSP was found, indicating a high current star formation rate. Also, a population with 10 Myr old was found, that matches the age of the referred star cluster of the galaxy.

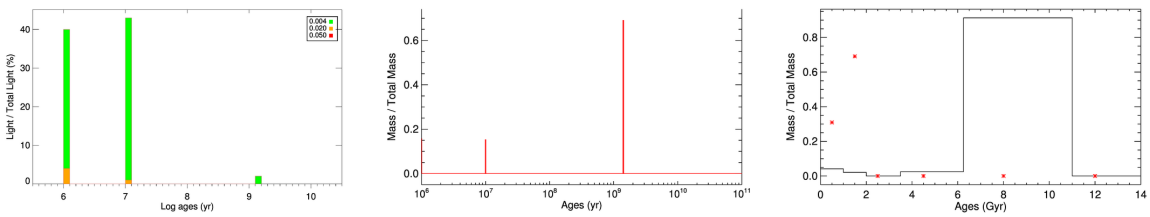


Figure 5.8: UGC4483: *Left* - Light fractions of the simple stellar populations that contributed to the model spectrum (logaritmic time scale). Metallicities are represented in different colours; *Middle* - Stellar mass fractions from STARLIGHT; *Left* - Comparision between the derived stellar mass fractions, from the two methods, in the age intervals 0-1, 1-2, 2-3, 3-6, 6-10, 10-14 Gyr. In red stars: FSF; In black: CMD from [Weisz et al., 2011]

UGC4483 relative masses shows similar results to the previous galaxies: the recent period (< 2 Gyr) of star formation was found but overestimated, specially the very young stellar populations (see figure 5.8). The highest mass created period for this galaxy (from CDM analysis) is actually situated between the 6 and 10 Gyr. However, full spectral fitting results show no signs of this star formation period.

5.2.5 UGC4459

UGC4459 is an irregular dwarf galaxy with a compact HI distribution and hight $H\alpha$ emission [Dicaire et al., 2008]

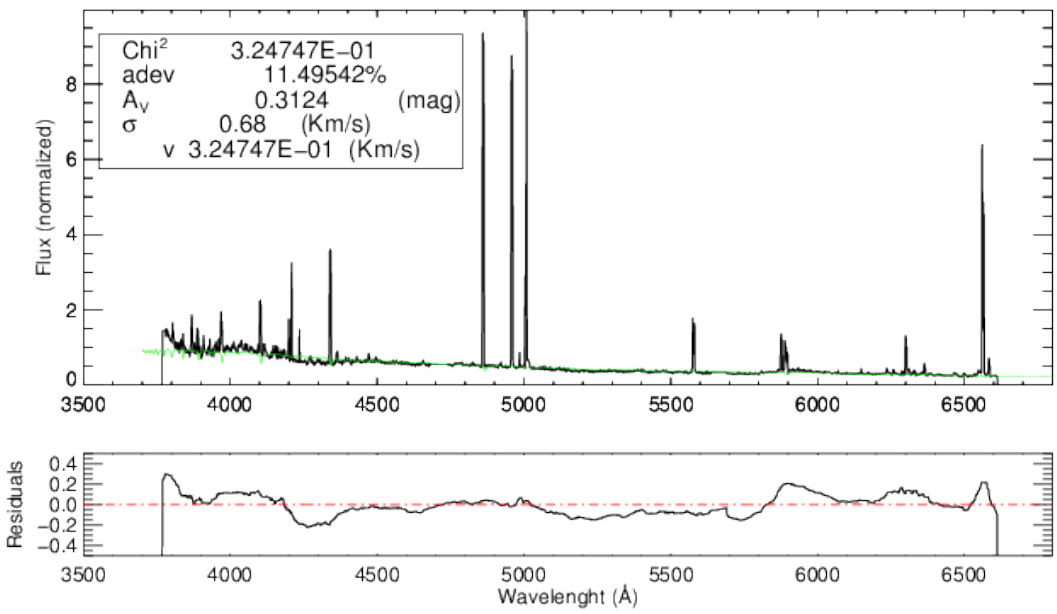


Figure 5.9: UGC4459: Input spectrum in black, model in green and residuals (in the lower panel).

As discussed in the reduction chapeter, this galaxy had a foreground K type star (or similar) contaminating the spectrum. The star contribution was masked during the extraction, but is possible that this correction was not perfect and some contamination still remains. This could also explain the poor match between observed and model spectrum observed in this galaxy when compared with the previous fits.

Nevertheless, the results for the $A(V)$ and v are within the expected ones. The velocity dispersion in this case is very small, only a few Km/s, and in this case no physical velocity dispertion can be derived [Cid Fernandes et al., 2005]. Was the spectrum still so contaminated that the spectral lines from the single star dominated and a consequent very low dispertion was needed to broaden the lines? Further quality tests are necessary to check the degree of contamination, to allow to confirm this hypothesis.

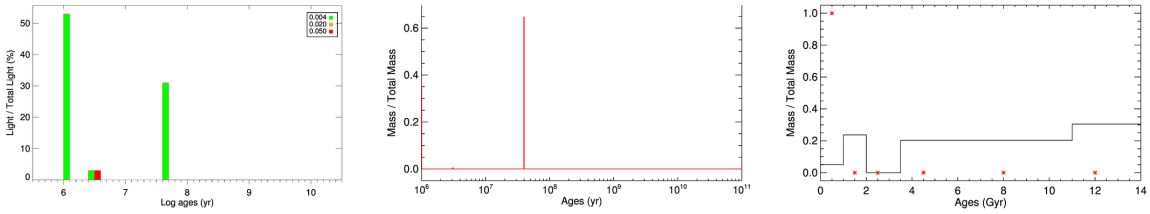


Figure 5.10: UGC4459: textitLeft - Light fractions of the simple stellar populations that contributed to the model spectrum (logaritmic time scale). Metallicities are represented in different colours; *Middle* - Stellar mass fractions from STARLIGHT; *Left* - Comparision between the derived stellar mass fractions, from the two methods, in the age intervals 0-1, 1-2, 2-3, 3-6, 6-10, 10-14 Gyr. In red stars: FSF; In black: CMD from [Weisz et al., 2011]

Only relativelly young stellar populations were found, with less than 50 Myr. This resulted on a very poor match between the recovered relative masses and the predictions.

5.3 Discussion of the quality of the fit

Visually inspecting the residuals, a 20% difference between the obtained model and the observed spectrum for $\lambda \leq 4200$ Å can be observed in all targets. This error can either be from the extracted spectra, and caused by the reducing steps or instrumental limitations, or from the proposed model.

Extracted Spectra

Sky subtraction Sky subtraction influences the shape of the continuum of the extracted spectra. It is important to remember that NGC2366 had a slightly different sky subtraction, with an exposure of night sky used to correct this effect. This means that problems arriving from this step came from the subtraction itself, an not from galaxy contamination in the frame area where the sky was estimated. In an attempt to further investigate the sky influence, a sky spectrum was extracted, flux calibrated and added to the base to fit the corresponding (sky unsubtracted) galaxy. The emission lines were not masked in this run. Interestingly, the quality of the fits did not decrease and in some cases the flux discrepancy in the blue diminished. What is more, only in the NGC6789 case did the sky spectrum contributed to the model (3% in this case). Figure 5.11 depicts two examples of these analysis: NGC1569, where the previous trend in the blue flux is not present, and NGC5678, where the fit quality is much similar to the one properly sky subtracted.

As a test, the extracted sky spectra from the different targets were used as input for a STARLIGHT run and fitted following the exact same process The sky spectra that were being fitted were added to the base elements. Two examples of the results can be seen on figure 5.12. Surprisingly, only for the fit of

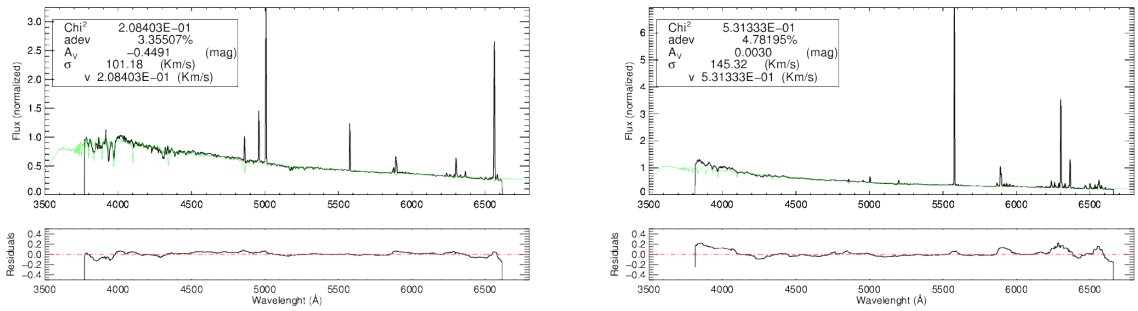


Figure 5.11: Fits of sky unsubtracted galaxies with sky extracted spectrum of each galaxy as base element *Upper panel* - NGC1569 (0% of sky contribution to the fit) *Lower panel* - NGC6789 (3% of sky contribution to the fit)

NCG6789 sky spectra did the sky ifself contribute to the final model (a contribution of 3%). This does not explain the discrepancy between model and observations from the previous section, but confirms that further investigation on this subject is needed, such as testing more complex sky subtraction methods.

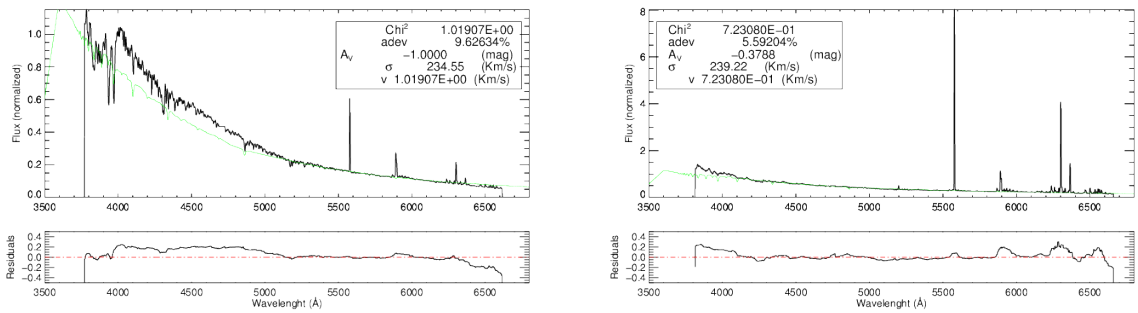


Figure 5.12: Sky fits with BC150 base plus the sky extracted spectrum of each traget *Upper panel* - NGC1569 sky (0% of sky contribution to the fit) *Lower panel* - NGC6789 sky (3% of sky contribution to the fit)

Continuum shape Flux calibration is, obviously, one of the probable sources of flux errors. Two sensitivity functions were used in the reduction, one for NGC6789 and other for the rest of the sample, so if there is any error it must be induced by the calibration method or from the instrument itself. GMOS-N CCD's have in fact a small efficiency in the $\leq$ 4200 Å spectral range (see figure5.13). This can explain the higher uncertainty is this spectral part and eventually the observed difference between model and observations. Barry Rotheberg from Postdam, specialist of HST imagery, will measure the total fluxes inside the areas where the spectra were extracted. The integrated flux of the spectra will be compared with these fluxes to access the relative spectra photometry quality.

Flux uncertainties The uncertainties propagated through the reduction process also play a role in the fitting algorithm, giving a weight to each spectral range. Indeed, the uncertainties were higher in the problematic wavelengths than in the rest of the spectrum. To test the influence of these uncertainties, fictitious errors were created, one of 1% of the flux of the spectrum and other of a small constant value. These errors were associated with the exact same spectra that were once more fitted with STARLIGHT.

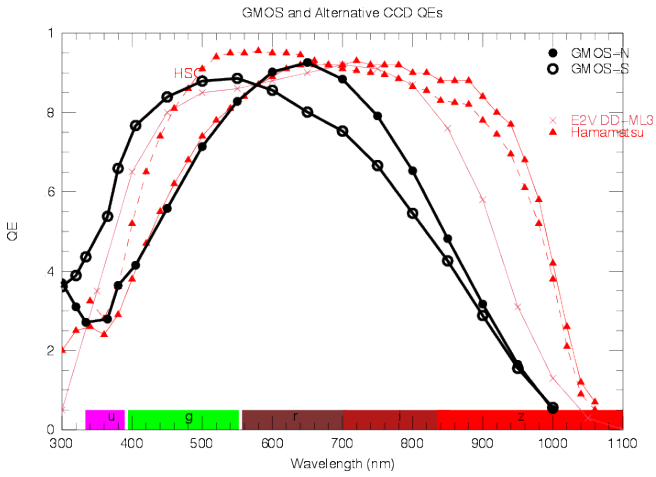


Figure 5.13: GMOS detectors quantum efficiency. Efficiency of the detectors used in the observations is shown in pink (E2V DD ML3) - credit: Gemini Observatory/AURA

The problem persisted in the results with both types of errors.

Model

There is also the possibility that the obtained fluxes are correct and the observed difference arise from the fitting technique or the model.

Dust One of the most problematic ingredients in this technique is to correctly account for the intrinsic attenuation. The Cardelli, Clayton and Mathis attenuation law (derived for our galaxy) was also tried, but no improvements in the fit were visible. If the error is fact caused by a bad description of the dust, it is nevertheless strange that the model predicts less flux in the blue. Previous studies have pointed to a higher attenuation in young stars, so one would expect that the model predicted more flux in the blue (where this young stars irradiate most) and not less.

SSP model Finally, this difference can be a reflection of the imperfection of the SSP models used. These may have its origin in the omission of some more exotic blue stars or on the IMF. Indeed this galaxies probe the most uncertain part of the IMF, since they contain a high number of massive stars that, as seen in the introduction, whose numbers are hard to estimate with accuracy. Assuming that the reduced spectra are correct, the difference could indicate that the used IMF is underestimating the number of massive stars produced.

5.4 Base effects

To investigate the possible base influence in the recovered model, the same spectra were fitted with other base. This was also composed of SSP spectra models from Bruzual and Charlot [2003], but this time

Galaxy	$M_{ini}(M\odot)$	$M_{cur}(M\odot)$	Nb of SSP	SSP $x_j < 1\%$
NGC1569	4.67×10^7	2.26×10^7	28	10
NGC2366	4.16×10^6	2.07×10^6	16	5
NGC6789	1.09×10^6	5.88×10^5	19	13
UGC4459	3.03×10^4	2.04×10^4	3	1
UGC4483	1.98×10^5	9.94×10^4	14	8
Mean	1.04×10^7	5.08×10^6	16	7.4

Table 5.3: Column 2: Initial mass from STARLIGHT results; Column 3: Current Mass from STARLIGHT results; Column 4: Number of SSP spectra used in the final spectrum model of the galaxy. Column 5: Number of SSP spectra that contributed less than 1% to the total light.

with 150 elements, covering 6 metallicities and 25 different ages. The age range covers SSP from 1 Myr to 18 Gyr and the metallicities values are 0.0001, 0.0004, 0.004, 0.008, 0.02, 0.05. All elements of BC45 are included in the BC150 base. The results are depicted in figure 5.14 to 5.18.

The fitted models once more prededicted $\sim 20\%$ less flux for wavelenght lower than 4200 Å. The χ^2 values are smaller than the ones from the 45 elements base, but the difference is small, with a difference of 1 to 10% of previous χ^2. Since the largest base has 3 times more elements than the smaller, it would be expectable to generate higher quality fits. This may indicate that even only 45 spectra are already too alike and increasing the number of spectra does not improve the description of the observed spectra.

All the extinction factors found were below the upper value estimated from the [Calzetti et al., 1994] work and in general were lower than the previously fitted values. For UGC4459 galaxy, that had a negative $A(V)$ in the previous fit, a positive though small value was found (0.03), that further points to a low dust content. For each galaxy, the velocity dispersion values reached with both bases differ by a few dozens of km/s with the exception of UGC4483, that has a difference of 250Km/s. The velocity dispersion reflects the spectral line broadening and since the input spectra were the same it was expected that results with both bases would agree, as it is the case.

The galaxy masses were calculated as following exactly the same process as before. Fits with this extended base used, approximately two times more base spectra (16 SSP spectra, in average) than the previous one (an average of 8.6 spectra) and only in the UGC4459 case were less SSP components found than before. However, more than half of the recovered components accounted for less than 1% of the total light (column 9 of table 5.3). There are few components common to both fits, even with two bases calculated with the same code, with once more suggests that there is a great level of degeneracy. In general, results with the 150 element base tend to have older populations. This might in part be explained by the fact that this is the only base that contains elements older than 13 Gyr and it can be one more indication of flux calibration problems. If the spectra are not correctly calibrated the algorithm will perhaps give a greater weight to more exotic (older than the Universe) elements.

For all galaxies, and independently of the used base, only one SSP with age between 1×10^7 and 1×10^8 was found (in UGC4459). Although it is not impossible that this is the true scenario, it seems

unlikely and most probably is a consequence of the method. It can indicate that the fitting technique favors only young (age $< 1 \times 10^7$ yr) and old populations (age $> 1 \times 10^8$ yr), possibly to describe the blue and the red part of the spectra respectively, and is not sensitive enough to recover intermediate ages, that in the observed wavelengths do not have strong and characteristic spectral lines that could constrain the fit.

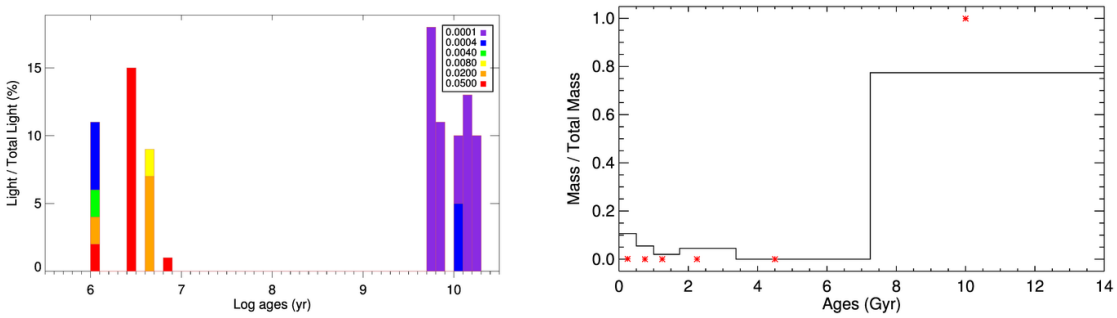


Figure 5.14: NGC1569: *Left* - Simple stellar populations that contributed to the model spectrum (in logarithmic time scale) *Right* - Comparision between the derived stellar mass fractions, from the two methods, in the age intervals 0-0.5, 0.5-1, 1-1.5, 1.5-3, 3-6, 6-14 Gyr. In red stars: FSF; In black: CMD from [McQuinn et al., 2010]

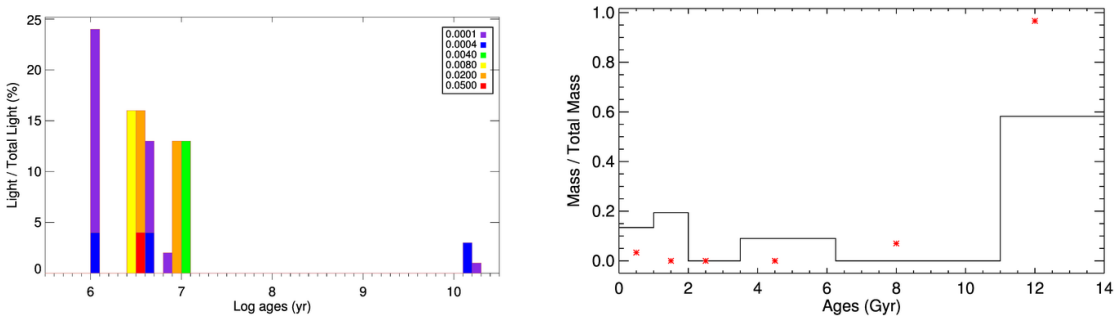


Figure 5.15: NGC2366: *Left* - Simple stellar populations that contributed to the model spectrum (in logarithmic time scale) *Right* - *Left* - Comparision between the derived stellar mass fractions, from the two methods, in the age intervals 0-1, 1-2, 2-3, 3-6, 6-10, 10-14 Gyr. In red stars: FSF; In black: CMD from [Weisz et al., 2011]

In what concerns the relative masses, there was a general increase in the mass created more than 6 Gyr ago, that matched the period where the higest mass contribution was expected from the CDM results. However, the accuracy for younger ages was much diminished, and only NGC6789 and UGC4483 show significant amounts of created mass in these periods. This confirms that increasing the number of base elements does not necessarily translates into better results. Moreover, these results show that full spectral fitting can retrieve old stellar populations, even when recent bursts have occured, which was one of the probable problems of the method, but with very low accuracy. This can be caused by an impossibility of finding all the SSP that are relevant to describe the stellar mass of the galaxy or because the light-to-mass ratios of the stellar populations are not calibrated for this method.

It is worth noticing that the accuracy of the estimated masses could not be predicted from the fit's quality, that actually had better χ^2 values than the fits with the 45 SSP spectra.

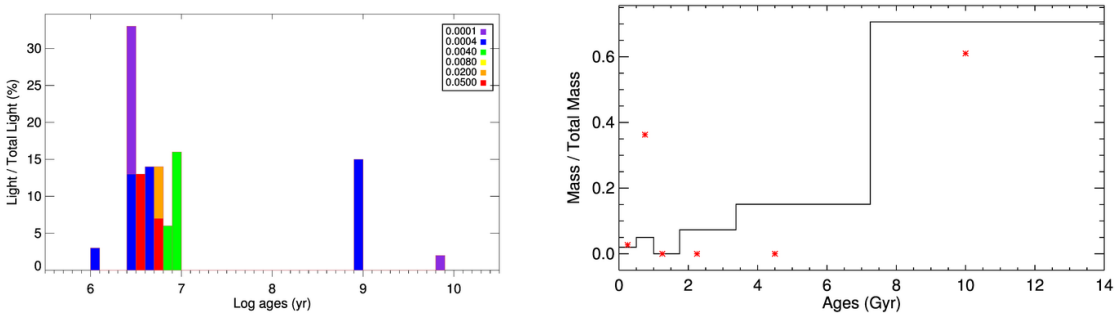


Figure 5.16: NGC6789: *Left* - Simple stellar populations that contributed to the model spectrum (in logarithmic time scale)*Right* - Comparision between the derived stellar mass fractions, from the two methods, in the age intervals 0-0.5, 0.5-1, 1-1.5, 1.5-3, 3-6, 6-14 Gyr. In red stars: FSF; In black: CMD from [McQuinn et al., 2010]

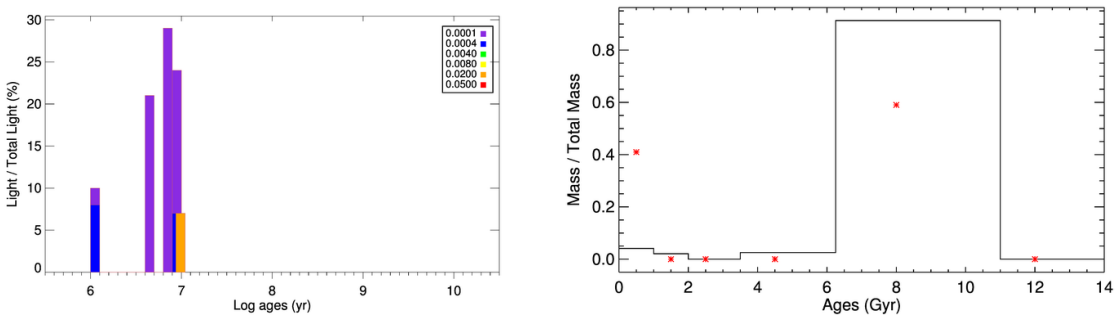


Figure 5.17: UGC4483: *Left* - Simple stellar populations that contributed to the model spectrum (in logarithmic time scale) *Right* - *Left* - Comparision between the derived stellar mass fractions, from the two methods, in the age intervals 0-1, 1-2, 2-3, 3-6, 6-10, 10-14 Gyr. In red stars: FSF; In black: CMD from [Weisz et al., 2011]

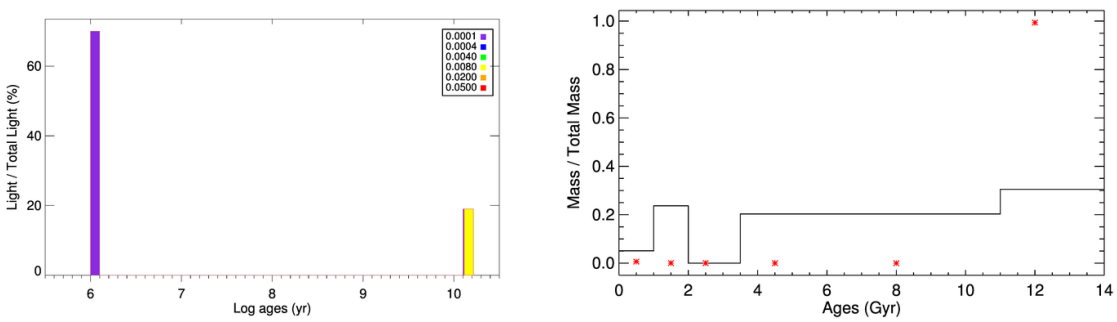


Figure 5.18: UGC4459: *Left* - Simple stellar populations that contributed to the model spectrum (in logarithmic time scale) *Right* - *Left* - Comparision between the derived stellar mass fractions, from the two methods, in the age intervals 0-1, 1-2, 2-3, 3-6, 6-10, 10-14 Gyr. In red stars: FSF; In black: CMD from [Weisz et al., 2011]

5.5 Attenuation law effect

To test the influence of the attenuation law in the fit outcome, NCG1569 (the spectrum with higher S/N) was fitted with other attenuation laws. The Cardelli, Clayton and Mathis [1989] (CCM) law, derived for galactic attenuation was tried. Two attenuation laws from Gordon et al.,[2003], one derived from for the Large Magellanic Cloud super shell (GD2) and other from the global Large Magellanic Cloud (GD3),

were also tested since this galaxy has some similarities with the sample. The results can be seen in figures 5.19, 5.20 and 5.21.

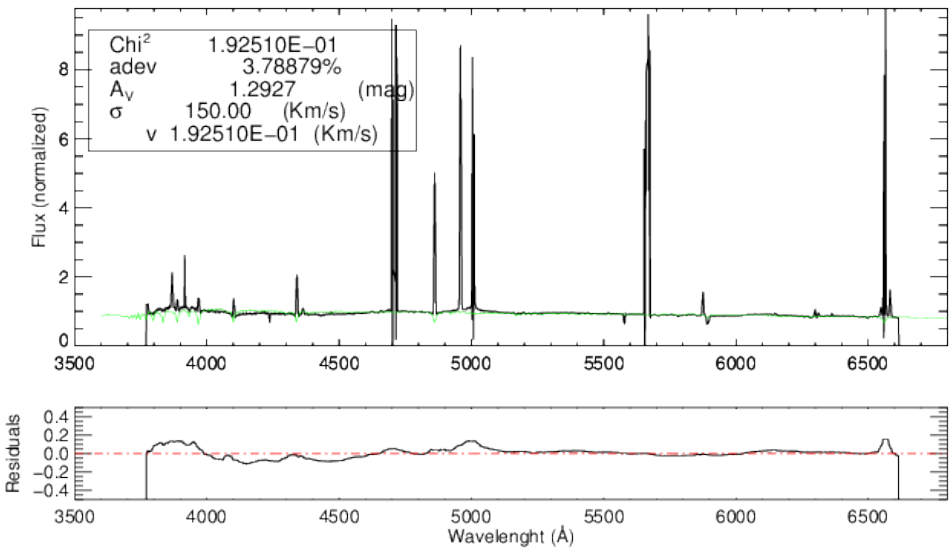


Figure 5.19: NGC1569 with CCM attenuation law. Input spectrum in black, model in green and residuals (in the lower panel).

With the CCM law, the extinction parameter $A(V)$ found was even higher than the one found with the Calzetti law. There were no significative differences in the continuum shape of the fitted model, and the flux discrepances found earlier persisted. The estimated SSP are similar, with the difference that an intermediate age stellar population was found in this fit. The estimated older population were a few Gyr younger.

The estimated $A(V)$ with the GD2 law is lower than with the Calzetti law, but remains quite higher then the one derived in the [Calzetti et al., 1994] work. Also in this case no significant influence in the continuum shape of the model spectra was found. A lower number of young stellar populations was found, although the age distribution has the same very young and extremely old dichotomy as the results with the Calzetti law. The results with GD3 were quite similar to the GD2 ones.

In summary, although the net $A(V)$ estimated values varied from law to law, the overall continuum shape of the model spectrum did not change significantly. The recovered SSP were different, although not completely unrelated with each other and the Calzetti results. The influence of these differences will become more evident in the star formation history analysis.

The recovered SSP show a similar temporal and mettallicity distribution with all attenuation laws - relativelly metal rich and young star populations and metal poor old populations, with more than 1 Gyr. The previous problem found with the Calzetti atteniation law of underestimation of mass created in older ages was present in all results. The calculated relative masses differ mostly for the period between 1 and 4 Gyr, but all attibute a much to high mass fraction to this period, and none shows a particular higher accuracy. This results were also checked for the remaining four galaxies, that also did not show a significant improvement in the predicted mass fractions. This results show that the chosen attenuation

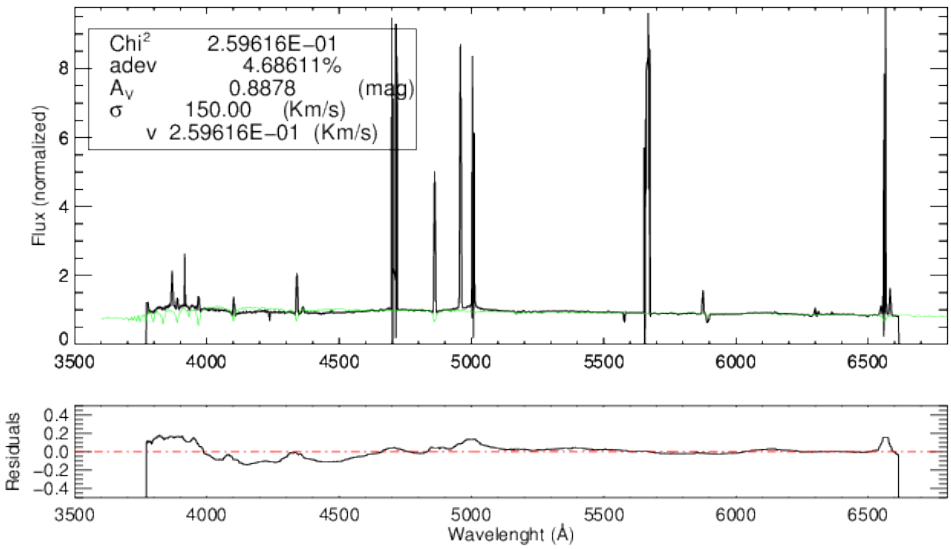


Figure 5.20: NGC1569 with GD2 attenuation law. Input spectrum in black, model in green and residuals (in the lower panel).

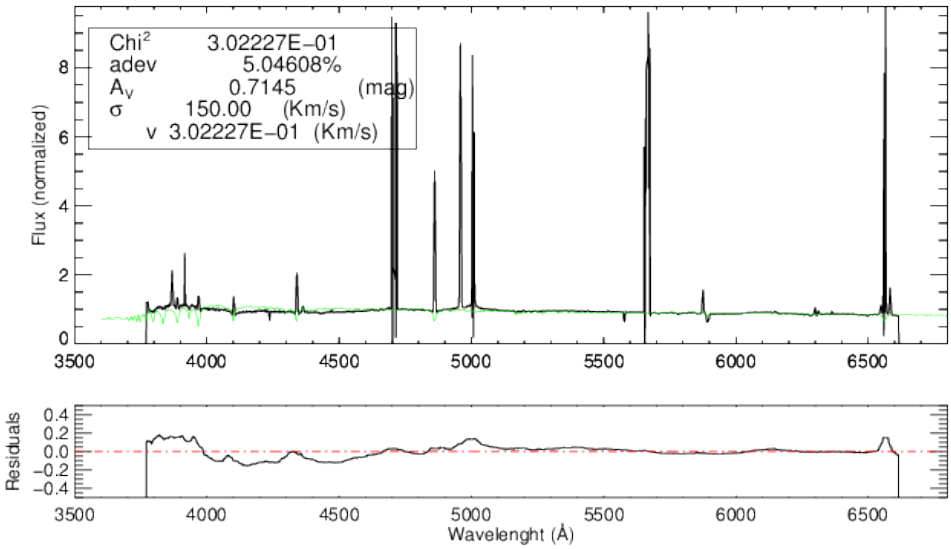


Figure 5.21: NGC1569 with GD3 attenuation law. Input spectrum in black, model in green and residuals (in the lower panel).

law plays an important role in the relative mass temporal distribution, although its effects are not evident in the quality of the fit or very clear in the age distribution of the found SSP.

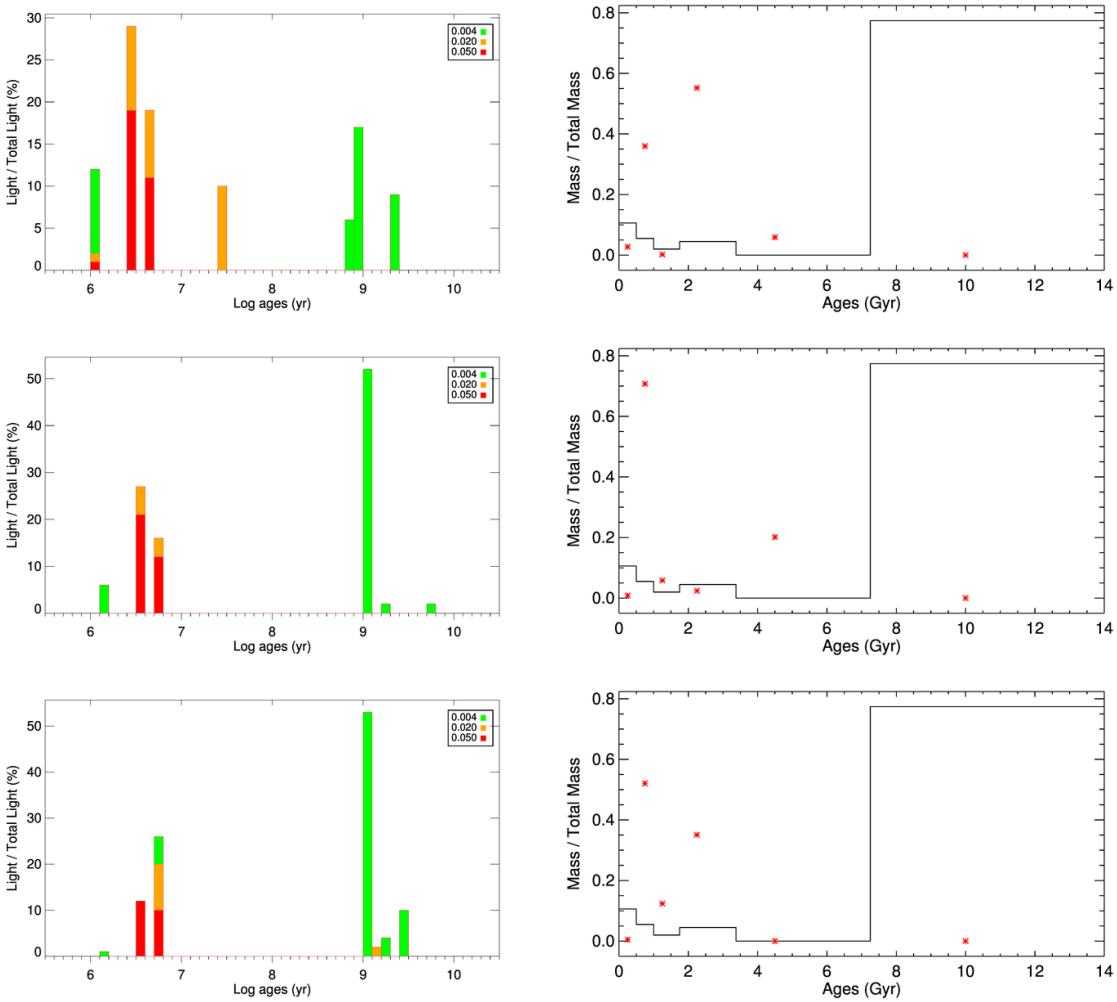


Figure 5.22: Recovered stellar population and relative masses with alternative attenuation laws. *Upper*: CCM law; *Middle*: GD2 law; *Lower:* GD3 law.

Chapter 6

Conclusions

The aim of this work was studying the accuracy of the simple stellar populations and their relative mass contribution to the total galaxy's mass derived from full spectral fitting. These are the first results of the Starfish project, led by Myriam Rodrigues, which has the goal of establishing the accuracy of the physical parameters (age, metallicity, star formation rates and masses) derived from full spectral fitting, validating and calibrating the method with real galaxy spectra for the first time.

From a first inspection on recovered star creation periods and comparing with the CMD results, for the three galaxies with higher S/N (a minimum of 22 in the blue and 20 in the red) the correct number of bursts was recovered, although with a age deviation of 0.5 Gyr in young ages and 3 Gyr in older populations. For the remaining two galaxies, older bursts with more than 1 Gyr were not found. The relative mass created within three broad age bins was also compared. Only one galaxy (NGC2366) shows a good agreement. The other galaxies results typically underestimated the mass contribution in the 1 - 10 Gyr interval and overestimated the created mass between 300 Myr - 1 Gyr.

The simple stellar populations found in each galaxy, match with reasonable accuracy recent known bursts (less than 10 Myr). However, the mass fraction attributed to these populations is lower than the expected, as it can be seen from the broad age bins comparison. Only in the case of UGC4459 was a population with age between 10 and 100 Myr found. This is not a completely unexpected result, since these stellar populations do not have characteristic absorption lines in the considered wavelengths, which implies that its fit depends only on the continuum shape and therefore is more difficult. For ages higher than 4 Gyr, the expected stellar populations were not found or their mass fraction underestimated. The only exception to this is the $\sim$ 5 Gyr old SSP found in NCG2366, that accounted for the majority of the galaxy's mass. Nevertheless, even in this case, older stellar formation expected around 10 Gyr could not be found. This confirms one of the problems of the method, already pointed by previous studies: older populations - specially after an intense recent burst - are difficult to recover. Due to the starburst nature of the presented sample is impossible to know in this study if full spectral fitting cannot find populations older than a few Gyr in any case, or if this problem is only caused by recent (few Myr) bursts. In summary, the method was able to find the very recent predicted SSP, and some older ones mostly around $\sim$1-2 Gyr old, but the estimated mass fractions do not match the predictions from CMD

analysis.

To access influence of the SSP spectral models used to perform the fits, another base was also used to perform the analysis. This base was also constituted by SSP Bruzual and Charlot [2003] models, but this time 150 elements were used, that spanned 6 metallicity values and 25 ages (from 1 Myr to 18 Gyr). In average, twice the number of SSP were found and, in general, more old components were recovered. Despite this, the quality of the SFH did not improve. This confirms that the SSPs models used as base elements are degenerate and an increase of the number of elements does not necessarily mean an improvement of the results. Besides the referred age-metallicity degeneracy, that already imposes limits to the method, the algebraic approach followed tries to decompose a spectrum in a combination of several SSP spectra that are not linearly independent. Because the SSP spectra used in the models are built with different amounts of stellar spectra, they will always be degenerate to same degree. A future approach to full spectral fitting will probably involve the determination of (more) independent base elements.

Several attenuation laws were also tried: the Cardelli, Clayton and Mathis [1989] galactic attenuation law, and two Gordon et al. [2003] attenuations derived from the Large Magellanic Cloud. For all attenuation laws, neither the model spectrum or the recovered SSP deviated much from the results obtained with the Calzetti law. The calculated mass fractions did differ however. Star formation was found in the same age periods but the relative strength of the bursts varying considerably. In conclusion, the chosen attenuation law has a important role in the estimated mass contribution of the SSPs, which strengthens the importance of the modelization of the dust in this problem.

It is important not to forget that these are still preliminary results, since the results derived from the integrated spectra were compared to the results for the whole galaxy, and not only to the observed area. Moreover, the masses derived from the CMD were calculated with a Kroupa like IMF, that predicts a much lower percentage of more than 10 $M\odot$ stars than the Chabrier model used in the fitting models. Moreover, all spectra models revealed a consistent $\sim$20% underestimation of the flux in the blue, that may introduce important errors in the derived mass fractions. Further investigation of the origin of this deviation and its eventual correction will be performed.

Future work on this analysis will include the verification of the flux calibration, through the comparison with the photometric fluxes that will be retrieved from the CMD. This will allow to understand if the blue flux problem arises from the reduction process or if it is a modelization problem. Furthermore, the mass fractions from color-magnitude diagrams will also be recomputed to match the area that was observed. Within the Starfish project, a wider sample of dwarf galaxies will be reduced and analyzed, following a similar methodology described where. Other models SSP models will be tested as base elements as well as different fitting algorithms. With these results the accuracy of the full spectral fitting method will be accessed and the light-to-mass ratios will be calibrated for this method which will permit more robust estimates of the star formation histories for unresolved galaxies.

Appendix A

Fiber efficiency determination for the VIMOS IFU

As refered in chapter 2, the sample used in the Starfish project was observed both with the GMOS and VIMOS instruments. One of the problems with VIMOS data for this project is to subtract the intrument signature of the spectra. In a student's project at the Paranal Observatory, under the supervision of the ESO-Chile fellow Fernando Selman, I investigated part of this problem, namely whether sky flats would allow a better correction of the data in comparision with the usual lamp flats.

A.1 Motivation

The current VIMOS Integral Field Unit (IFU) data reduction is done using screen flats, for which two lamps are available. As expected, the light of the lamps is not perfectly 'white', meaning that it doesn't have an uniform flux in the observed wavelengths, emitting more in the redder region. As a result, once that the spectrum trace in the CCD are reconstructed with the flat field, the extraction of the bluer wavelengths is not as accurate as in the other regions.

Some observers have claimed that twilight flats avoid this problem an would be a better alternative. The aim of the project is then to determine the efficiency of the fiber transmission obtained with twilight flats and to compare it with the usual efficiency calculated with screen flats.

A.2 Methodology

A.2.1 Observations

A set of 6 flats were taken during twilight in high (HR) and low resolution (LR) configuration, for which the *HR_blue* grism with no filter and the *LR_blue* grism with the OS-blue filter were respectively used. The time exposure was of 1 second with IFU shutter on on the high resolution configuration and no shutter of the low. The regular screen flats for the usual calibration were also obtained.

A.2.2 Data Reduction

The VIMOS pipeline in Gasgano was used to perform the reduction of the data. The description of the reductions steps is done in the following subsections. The following subsections describe the reduction steps. Frames signed with * are used in the next step of the pipeline.

Master Bias

Several bias frames are mean combined to form a Master Bias frame.

Function: vmbias

Inputs: 5 Bias frames

Outputs: Master Bias *

IFU Calibration

VIMOS pipeline version 2.9.3 (and at least all previews versions) does not recognize twilight flats as flats. To be able to use the vmifucalib recipe the user must change the Data Organizer (DO) category of the files from $UNCLASSIFIED$ to IFU_SCREEN_FLAT, which can be easily done in the VIMOS pipeline in the recipe menu, once the files are loaded into it, changing the 'Classification' item.

Function: vmifucalib

Inputs: 6 Twilight flats / 6 Screen flats (per quadrant); Arc lamp exposure; Master bias; Arc line catalog; IFU fiber identification

Outputs:Master screen flat; Extracted flat spectra; Extracted arc spectra; IFU trace*; IFU wavelength calibration*; IFU fiber transmission

Efficiency Map (Field of view) and Data Cube

The Field Of View (FOV) image is determined integrating the signal in the spectral dimension, although not on the entire wavelength range. Instead, a fixed $\Delta\lambda$ is used, and the central wavelength where to sum the signal is chosen to be one with the highest signal.

Since the calibration process does not produce the FOV of the flat fields, the science reduction recipe was applied to the raw flat fields. Once more, it's necessary to change the classification of the input flats from IFU_SCREEN_FLATS or $UNCLASSIFIED$ to $IFU_SCIENCE$ in order to the the pipeline to work nicely. In this reduction, the images were not corrected for transmission (unselect the option in the recipe menu) since in order to obtain the relative transmission.

Function: vmifuscience

Inputs: 6 Twilight flats / 6 Screen flats (per quadrant); Master bias; IFU Wavelength calibration; IFU Trace

Outputs: FOV*; Reduced flat spectra

Quadrant Combine

The FOV and reduced spectra of the 4 quadrants of each exposure were combined in the VIMOS pipeline.
Function: vmifucombine
Inputs: 4 Twilight flats / 4 Screen flats (x 6 exposures)
Outputs: FOV*
Recipe: vmifucombinecube
Inputs: 4 Reduced screen/twilight flat spectra (x 6 exposures)
Outputs: FOV*

Exposure Combine

The 6 combined FOV and data cubes for the twilight and screen flats were combined into a single frame. This was performed in iraf.
Function: imcombine
Parameters: combine: median
rejection: none

A.2.3 Data Analysis

For the data visualization and manipulation the QFitsView package was used, as well as IRAF and IDL developed for this analysis. Quadrants are identified as [x1:x2,y1:y2] as the range in FOV that they occupy.

FOV reconstructed by the VIMOS pipeline

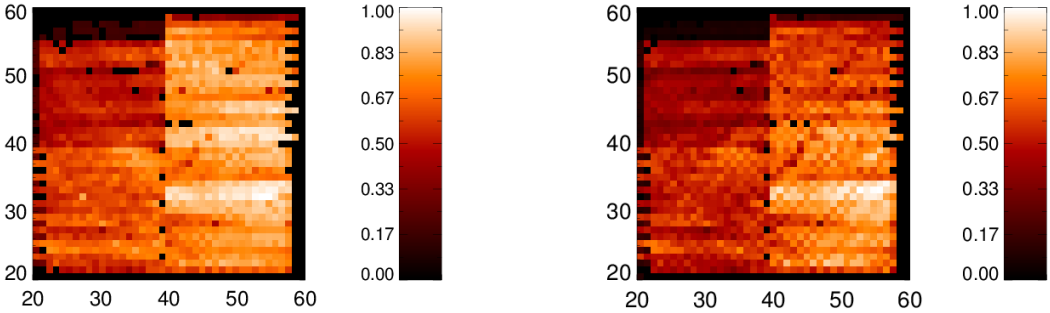
The maximum value of the FOV for each of the 4 or 16 blocks (for high or low resolution) were determined with IRAF imstat and the ratio between the maximum of each block and the absolute maximum was calculated. In the figures below the reconstructed FOV, normalized to the maximum, are followed by the respective ratio map.

Spectral response in broad wavelength bands

The data cubes were sliced in 4 parts, along the spectral dimension, and the resulting cubes were then collapsed in 2D images (an equivalent of the FOV for that given spectral range). Finally the images were normalized to the 5000 to 5350 Å FOV, resulting on a map of the relative efficiency of the fibers on a given wavelength range compared to the redder range.

A.2.4 Field of view quality

To access the quality of the FOV reconstruction obtained with the VIMOS pipeline, the M25 cluster was observed in High Resolution. The data reduction process followed was silimar to the previously described.



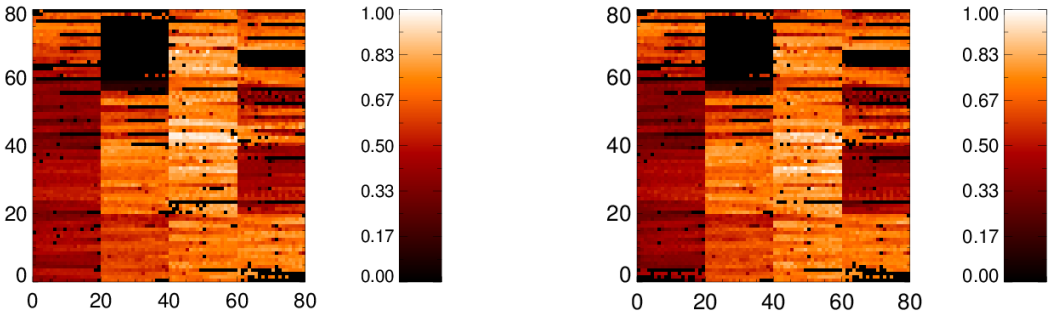
0.79	1.00
0.84	0.95

Absolute maximum: 17632

0.63	0.89
0.76	1.00

Absolute maximum: 108244

Figure A.1: VIMOS high resolution flats fluxes: *Left* - Scren flat; *Left* - Twilight Flat



0.76	0.71	1.00	0.84
0.48	0.79	1.00	0.85
0.58	0.84	0.95	0.59
0.56	0.74	0.82	0.83

Absolute maximum=17632

0.80	0.69	0.93	0.82
0.46	0.78	0.96	0.81
0.58	0.87	1.00	0.57
0.51	0.72	0.86	0.85

Absolute maximum=60076

Figure A.2: VIMOS low resolution flats fluxes: *Left* - Scren flat; *Left* - Twilight Flat

In high resolution mode *HR_orange* grims was used with the GG435 filter, resulting on a FOV of 27"x 27" and a magnification of 0.33"/pixel.

A.3 Discussion and conclusions

We concluded that in both situations, twilight and screen flat, quadrant [40:60,20:40] of the FOV has the best transmission efficiency.

In high resolution, there is a significant increase in the relative flux of the twilight flat field when compared with the screen flat flied for shorter wavelengths than 4000 Å, since the sun has a higher flux in this range than the lamp. This seams to indicate that the use of twilight flats could increase the accuracy of the trace and diminish the errors in flux in this range.

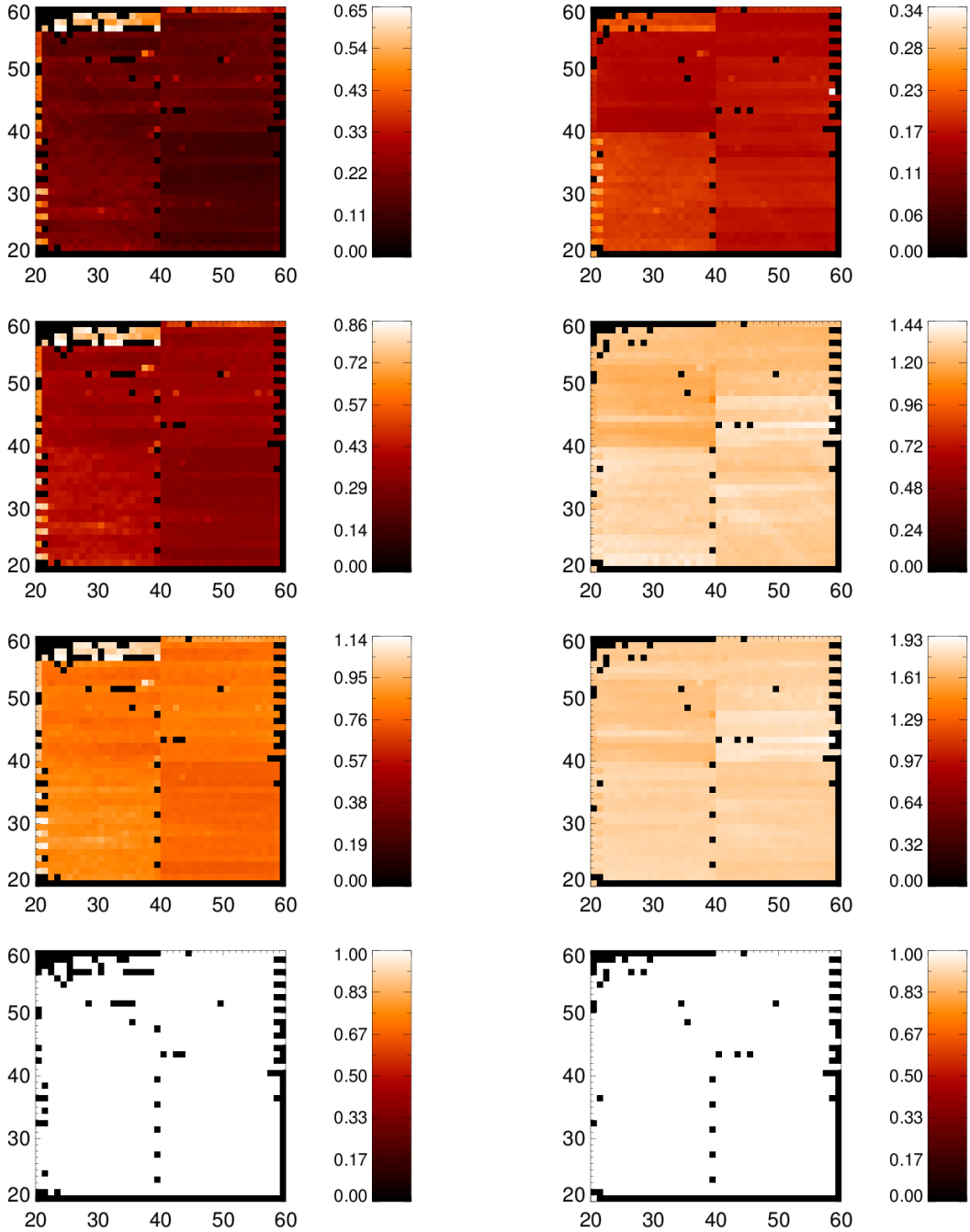


Figure A.3: VIMOS high resolution spectral response: *Left* - Screen flat; *Left* - Twilight Flat; First panel: 3450 to 4000 Å ; Second panel: 4000 to 4500 Å ; Third panel: 4500 to 5000 Å ; Fourth panel: 5000 to 5350 Å

The efficiency determined is the convolved effect of the fiber with the pixel sensitivity, which explains why the central part of the FOV and data cubes, common to the high and low resolution modes, does not have the same behavior in the two modes, since the combination fiber/pixel in each of them is different.

The reconstruction of the M25 cluster in high resolution matches the expected field.

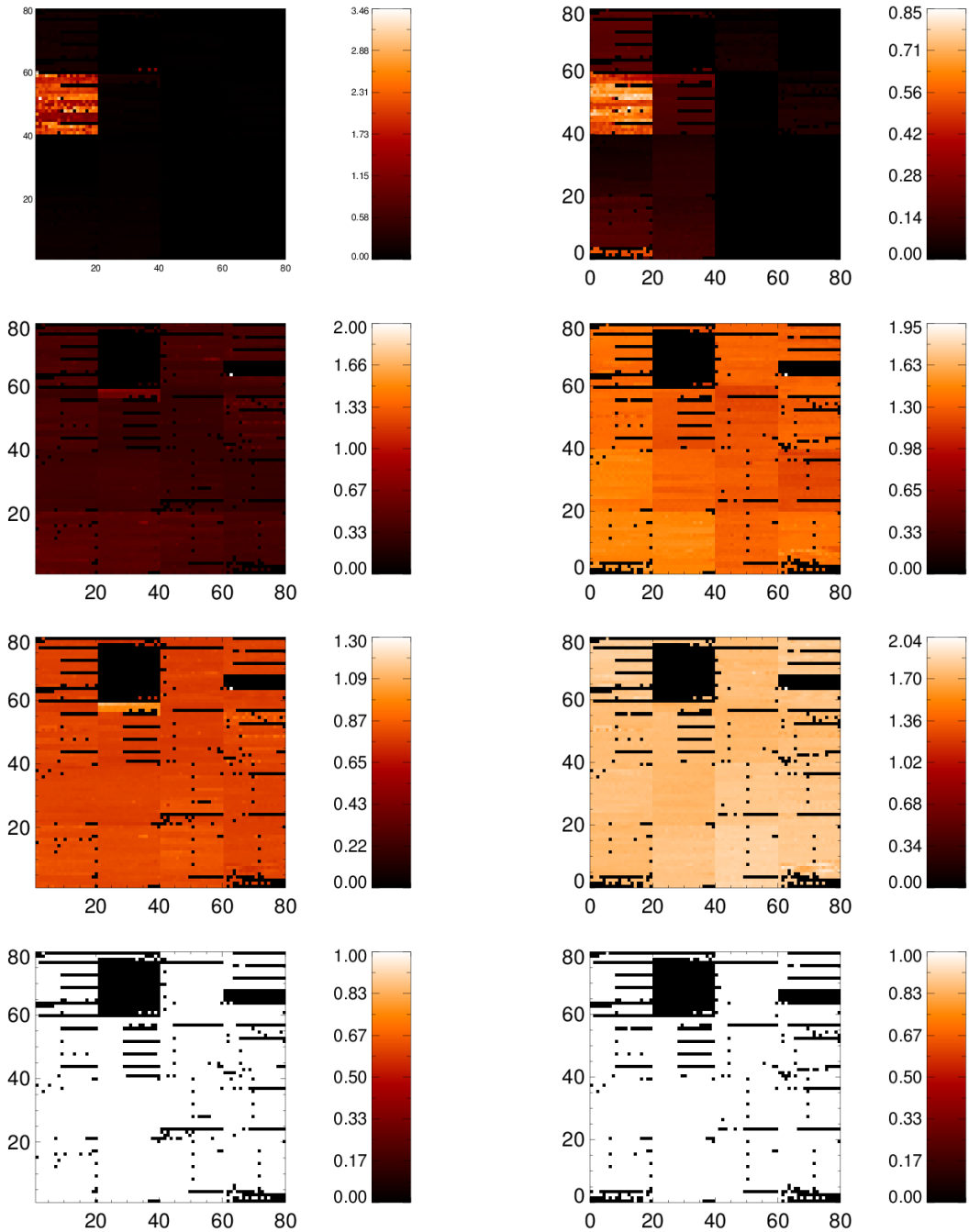


Figure A.4: VIMOS low resolution spectral response: *Left* - Screen flat; *Left* - Twilight Flat; First panel: 3450 to 4000 Å ; Second panel: 4000 to 4500 Å ; Third panel: 4500 to 5000 Å ; Fourth panel: 5000 to 5350 Å

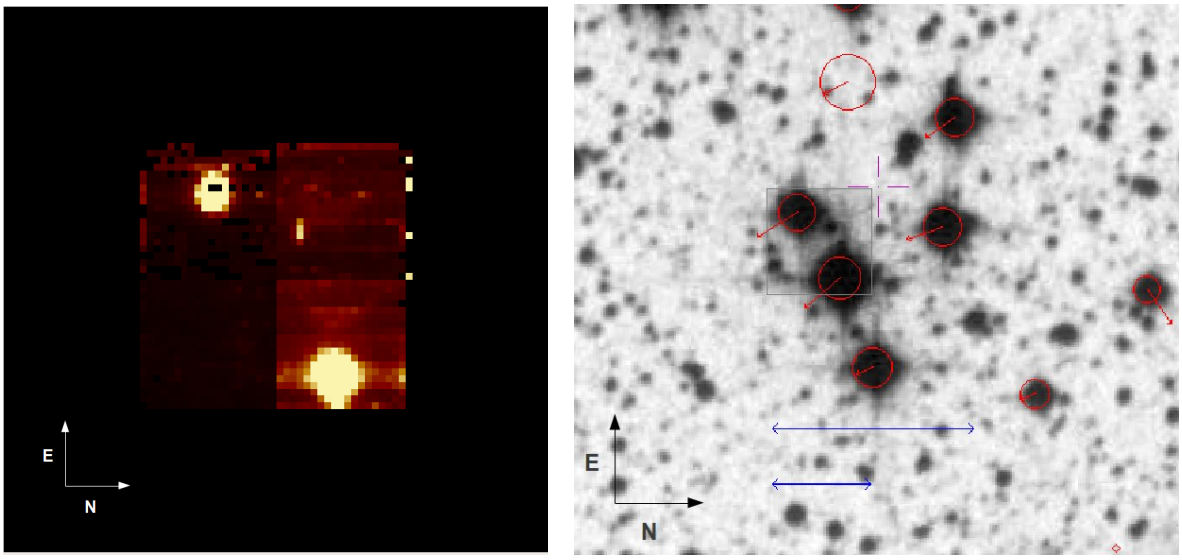


Figure A.5: *Left* - VIMOS reconstructed vield of fiew of M25 cluster in high resolution mode; *Left* - DSS image chart of the same field. Grey square corresponds to the instrument FOV in high resolution. (image produced with *Aladin* software)

Bibliography

D. Alloin, C. Gallart, E. Fleurence, E. Pompei, D. Raimann, U. Fritze-von Alvensleben, and S. Yi. Deriving Star Formation Histories: Evolutionary or Population Synthesis Codes versus Color-Magnitude Diagrams. *apss*, 281:109–113, July 2002. doi: 10.1023/A:1019592523194.

N.V. Asari, R. Cid Fernandes, G. Stasińska, J.P. Torres-Papaqui, A. Mateus, L. Sodré, W. Schoenell, and J.M. Gomes. The history of star-forming galaxies in the Sloan Digital Sky Survey. *mnras*, 381: 263–279, October 2007. doi: 10.1111/j.1365-2966.2007.12255.x.

M.S. Bessell. Measuring the Balmer Jump and the Effective Gravity in FGK Stars. *pasp*, 119:605–615, June 2007. doi: 10.1086/519981.

G. Bruzual and S. Charlot. Stellar population synthesis at the resolution of 2003. *mnras*, 344:1000–1028, October 2003. doi: 10.1046/j.1365-8711.2003.06897.x.

D. Calzetti, A.L. Kinney, and T. Storchi-Bergmann. Dust extinction of the stellar continua in starburst galaxies: The ultraviolet and optical extinction law. *apj*, 429:582–601, July 1994. doi: 10.1086/174346.

D. Calzetti, L. Armus, R.C. Bohlin, A.L. Kinney, J. Koornneef, and T. Storchi-Bergmann. The Dust Content and Opacity of Actively Star-forming Galaxies. *apj*, 533:682–695, April 2000. doi: 10.1086/308692.

G. Chabrier. Galactic Stellar and Substellar Initial Mass Function. *pasp*, 115:763–795, July 2003. doi: 10.1086/376392.

R. Cid Fernandes and R.M.G. Delgado. Testing evolutionary synthesis models: Empirical feedback to model makers. In G. R. Bruzual and S. Charlot, editors, *IAU Symposium*, volume 262 of *IAU Symposium*, pages 69–72, April 2010. doi: 10.1017/S1743921310002541.

R. Cid Fernandes, A. Mateus, L. Sodré, G. Stasińska, and J.M. Gomes. Semi-empirical analysis of Sloan Digital Sky Survey galaxies - I. Spectral synthesis method. *mnras*, 358:363–378, April 2005. doi: 10.1111/j.1365-2966.2005.08752.x.

P. Coelho, B. Barbuy, J. Meléndez, R. P. Schiavon, and B. V. Castilho. A library of high resolution synthetic stellar spectra from 300 nm to 1.8 μm with solar and α-enhanced composition. *aap*, 443: 735–746, November 2005. doi: 10.1051/0004-6361:20053511.

C. Conroy. Modeling the Panchromatic Spectral Energy Distributions of Galaxies. *araa*, 51:393–455, August 2013. doi: 10.1146/annurev-astro-082812-141017.

I. Dicaire, C. Carignan, P. Amram, O. Hernandez, L. Chemin, O. Daigle, M.-M. de Denus-Baillargeon, C. Balkowski, A. Boselli, K. Fathi, and R.C. Kennicutt. Hα kinematics of the Spitzer Infrared Nearby Galaxies Survey - II. *mnras*, 385:553–605, April 2008. doi: 10.1111/j.1365-2966.2008.12868.x.

L. Drissen, J.-R. Roy, C. Robert, D. Devost, and R. Doyon. The Star Formation History of the Starburst Region NGC 2363 and its Surroundings. *aj*, 119:688–704, February 2000. doi: 10.1086/301204.

I.O. Drozdovsky, R.E. Schulte-Ladbeck, U. Hopp, M.M. Crone, and L. Greggio. The Stellar Content of NGC 6789, A Blue Compact Dwarf Galaxy in the Local Void. *apjl*, 551:L135–L138, April 2001. doi: 10.1086/320021.

E.L. Fitzpatrick and D. Massa. An analysis of the shapes of ultraviolet extinction curves. III - an atlas of ultraviolet extinction curves. *apjs*, 72:163–189, January 1990. doi: 10.1086/191413.

R. M. González Delgado and R. Cid Fernandes. Testing spectral models for stellar populations with star clusters - II. Results. *mnras*, 403:797–816, April 2010. doi: 10.1111/j.1365-2966.2009.16152.x.

F. Hammer, H. Flores, D. Elbaz, X.Z. Zheng, Y.C. Liang, and C. Cesarsky. Did most present-day spirals form during the last 8 Gyr?. A formation history with violent episodes revealed by panchromatic observations. *aap*, 430:115–128, January 2005. doi: 10.1051/0004-6361:20041471.

M. Hamuy, A.R. Walker, N.B. Suntzeff, P. Gigoux, S.R. Heathcote, and M.M. Phillips. Southern spectrophotometric standards. *pasp*, 104:533–552, July 1992. doi: 10.1086/133028.

I.M. Hook, I. Jørgensen, J.R. Allington-Smith, R.L. Davies, N. Metcalfe, R.G. Murowinski, and D. Crampton. The Gemini-North Multi-Object Spectrograph: Performance in Imaging, Long-Slit, and Multi-Object Spectroscopic Modes. *pasp*, 116:425–440, May 2004a. doi: 10.1086/383624.

I.M. Hook, I. Jørgensen, J.R. Allington-Smith, R.L. Davies, N. Metcalfe, R.G. Murowinski, and D. Crampton. The Gemini-North Multi-Object Spectrograph: Performance in Imaging, Long-Slit, and Multi-Object Spectroscopic Modes. *pasp*, 116:425–440, May 2004b. doi: 10.1086/383624.

H. Krueger, U. Fritze-v.Alvensleben, and H.H. Loose. Optical and near infrared spectral energy distributions. of blue compact galaxies from evolutionary synthesis. *aap*, 303:41, November 1995.

T. Lilly and U. F.-V. Alvensleben. Analysing globular cluster observations. Models and analysis tools for Lick/IDS indices. *aap*, 457:467–476, October 2006. doi: 10.1051/0004-6361:20052986.

T. Lilly and U. Fritze-v. Alvensleben. Deriving star formation histories from integrated light: Colors and indices. *ArXiv Astrophysics e-prints*, July 2005.

L. Makarova, M. Koleva, D. Makarov, and P. Prugniel. Star Formation History of KDG 61 and KDG 64 From Spectroscopy and Colour-Magnitude Diagrams. In M. Koleva, P. Prugniel, and I. Vauglin,

editors, *EAS Publications Series*, volume 48 of *EAS Publications Series*, pages 65–66, July 2011. doi: 10.1051/eas/1148012.

P. Massey. A User's Guide to CCD reduction with IRAF. 1997.

P. Massey and M.M. Hanson. Astronomical Spectroscopy. *ArXiv e-prints*, October 2010.

P. Massey and M.M. Hanson. *Astronomical Spectroscopy*, page 35. 2013.

K.B.W. McQuinn, E.D. Skillman, J.M. Cannon, J. Dalcanton, A. Dolphin, S. Hidalgo-Rodríguez, J. Holtzman, D. Stark, D. Weisz, and B. Williams. The Nature of Starbursts. I. The Star Formation Histories of Eighteen Nearby Starburst Dwarf Galaxies. *apj*, 721:297–317, September 2010. doi: 10.1088/0004-637X/721/1/297.

H. Mo, F. Van den Bosh, and S. White. *Galaxy Formation and Evolution*. Cambridge University Press, 2010.

M. Rodrigues. *Evolution of intermediate mass galaxies up to z˜0.7 and studies of SNe Ia hosts*. PhD thesis, PhD Thesis, 2010, 2010.

E.F. Schlafly and D.P. Finkbeiner. Measuring Reddening with Sloan Digital Sky Survey Stellar Spectra and Recalibrating SFD. *apj*, 737:103, August 2011. doi: 10.1088/0004-637X/737/2/103.

Sparke and Gallagher. *Galaxies in the Universe, An Itroduction*. Cambridge University Press, 2nd edition, February 2007.

T. X. Thuan and Y. I. Izotov. A HST Study of the Stellar Populations in the Cometary Dwarf Irregular Galaxy NGC 2366. *apj*, 627:739–753, July 2005. doi: 10.1086/430338.

E. Tolstoy, V. Hill, and M. Tosi. Star-Formation Histories, Abundances, and Kinematics of Dwarf Galaxies in the Local Group. *araa*, 47:371–425, September 2009. doi: 10.1146/annurev-astro-082708-101650.

J. Walcher, B. Groves, T. Budavári, and D. Dale. Fitting the integrated spectral energy distributions of galaxies. *apss*, 331:1–52, January 2011. doi: 10.1007/s10509-010-0458-z.

W.H. Waller. Relics of an eruptive starburst in NGC 1569. *apj*, 370:144–157, March 1991. doi: 10.1086/169799.

D.R. Weisz, J.J. Dalcanton, B.F. Williams, K.M. Gilbert, E.D. Skillman, A.C. Seth, A.E. Dolphin, K.B.W. McQuinn, S.M. Gogarten, J. Holtzman, K. Rosema, A. Cole, I.D. Karachentsev, and D. Zaritsky. The ACS Nearby Galaxy Survey Treasury. VIII. The Global Star Formation Histories of 60 Dwarf Galaxies in the Local Volume. *apj*, 739:5, September 2011. doi: 10.1088/0004-637X/739/1/5.

M.J. Wolf, N. Drory, K. Gebhardt, and G.J. Hill. Ages and Metallicities of Extragalactic Globular Clusters from Spectral and Photometric Fits of Stellar Population Synthesis Models. *apj*, 655:179–211, January 2007. doi: 10.1086/509768.

G. Worthey. The Age-Metallicity Degeneracy. In I. Hubeny, S. Heap, and R. Cornett, editors, *Spectrophotometric Dating of Stars and Galaxies*, volume 192 of *Astronomical Society of the Pacific Conference Series*, page 283, 1999.

G. Worthey, S. M. Faber, J. J. Gonzalez, and D. Burstein. Old stellar populations. 5: Absorption feature indices for the complete LICK/IDS sample of stars. *apjs*, 94:687–722, October 1994. doi: 10.1086/192087.