Carlos Roberto de Melo Carneiro

Testing General Relativity at Galactic Scales

Brasil

Agosto, 2021

Carlos Roberto de Melo Carneiro

Testing General Relativity at Galactic Scales

Dissertação submetida ao Programa de Pós Graduação em Física do Instituto de Física da UFRGS, como requesito parcial para obtenção do título de Mestre em Física, com ênfase em Astrofísica.

Universidade Federal do Rio Grande do Sul – UFRGS

Instituto de Física

Programa de Pós-Graduação

Supervisor: Cristina Furlanetto

Co-supervisor: Ana Chies Santos

Brasil

Agosto, 2021

Carlos Roberto de Melo Carneiro
Testing General Relativity
at Galactic Scales / Carlos Roberto de Melo Carneiro. – Brasil, Agosto, 2021-
169 p. : il. (algumas color.) ; 30 cm.

Supervisor: Cristina Furlanetto
Co-supervisor: Ana Chies Santos

Master Thesis – Universidade Federal do Rio Grande do Sul – UFRGS
Instituto de Física
Programa de Pós-Graduação, Agosto, 2021.

1. Gravitational lensing. 2. Galactic dynamics. 3. General Relativity I. Cristina Furlanetto. II. Universidade Federal do Rio Grande do Sul. III. Instituto de Física. IV. Testing General Relativity at Galactic Scales

CDU 02:141:005.7

Carlos Roberto de Melo Carneiro

Testing General Relativity at Galactic Scales

Dissertação submetida ao Programa de Pós Graduação em Física do Instituto de Física da UFRGS, como requesito parcial para obtenção do título de Mestre em Física, com ênfase em Astrofísica.

Trabalho aprovado. Brasil, 10 de Agosto de 2021:

Cristina Furlanetto
Orientadora

Ana Chies Santos
Co-Orientadora

Brasil
Agosto, 2021

Para Cintia e Henrique, em memória e com amor.

Acknowledgements

Se eu me permitisse aqui agradecer a todos e todas de maneira adequada, muito provavelmente essa dissertação teria o dobro de páginas (e partindo de mim, isso provavelmente não é um exagero completo). Portanto, deixo os agradecimentos calorosos e enfáticos para acontecerem pessoalmente, de preferência após todos e todas estarem devidamente vacinados. Contudo, isso não é motivo para deixar de registar aqui alguns agradecimentos mais que importantes. Dito isso, todo amor, carinho e respeito aos meus pais, sem os quais a trajetória que me trouxe até aqui teria sido impossível. O mesmo vale para meu irmão, por todo apoio e imensa paciência.

Um carinho especial eu dedico aos amigos e amigas que me acompanharam (mesmo que a distância) durante todo o mestrado e os meses (anos?) de isolamento, que me aturaram mesmo sendo monotemático nos assuntos ou reclamando constantemente da vida. Obrigado pelas receitas de bolo, risadas infinitas, partilhas importantes e apoio sem dimensões. Ayume, Brenno, Clarinha, Francisco, Maria, Luan, Octávio, Victória muito obrigado.

Um obrigado também para aqueles com quem um dia pude dividir um lar, e mesmo hoje, anos depois, ainda encontro conforto e a certeza de poder voltar. Emerson, Igor, Ivo, Hugo e Marcelo, um abraço apertado em todos.

À todos aqueles que me receberam tão bem em Porto Alegre, deixo aqui meus mais sinceros agradecimentos. Apesar do pouco tempo presencial que passamos juntos, a receptividade e carinho foram essenciais. Em particular ao Prof. Rogério por sua receptividade e atenção, e aos colegas de laboratório Emílio, Jamile, Mônica, Rodrigo e Theylor, que talvez mesmo sem perceber, foram importantes.

Um agradecimento mais que especial eu deixo as minhas orientadoras, Professoras Ana Chies e Cristina Furlanetto, não apenas por aceitarem participar de um projeto desafiador (seria essa a melhor palavra? who knows), mas também por todo apoio e humanidade que possuem. Obrigado pela paciência, dedicação, conhecimento e acolhimento. Citar Newton é clichê, mas se cheguei onde cheguei, foi por me apoiar em ombro de gigantas.

Agradeço ainda a todo o apoio fornecido pelos membros do grupo CosmObs, em particular ao Prof. Martin Makler, por todas as contribuições fornecidas e pelo acesso ao cluster CHE do CBPF, sem o qual esse trabalho não seria possível (literalmente). Aproveito para agradecer ainda a todas as demais pessoas que contribuíram direta ou indiretamente para esse trabalho, vocês são muitas e eu sou muito grato por toda a contribuição. Gostaria ainda de agradecer ao Departamento de Astronomia da UFRGS, e à própria UFRGS, por sua excelência e compromisso com a ciência e a educação. Em tempos sombrios a educação

pública, universal e de qualidade sempre representará um farol guia.

Não menos importante, gostaria de agradecer ao CNPq, órgão que financiou essa pesquisa e a tornou possível. E ainda em tempo, gostaria de agradecer ao LNCC e ao cluster SDumont por fornecer sua infraestrutura de supercomputadores.

Nós começamos confusos, e terminamos confusos em um nível mais elevado.
Alan Chalmers

Abstract

General Relativity (GR) has been successfully tested in Solar system scales. However, for galactic scales, this theory has been poorly tested. Moreover, several galaxy data analysis depend on the value of the Hubble constant (H_0), which is currently under discussion. In addition, over the decades, some tensions arose, and GR has not been able to solve them. In this work we implement another test to GR in galactic scales, combining mass measurements for an elliptical lens galaxy. These measurements can be obtained by stellar kinematics and gravitational lensing, which are connected by a $\gamma_{\rm PPN}$ parameter, in the Parametrized Post-Newtonian formalism. This parameter can be derived, in our context, from the ratio of two different potentials: the Newtonian potential Φ, which acts in massive and non-relativistic particles; and the curvature potential Ψ, related to the curvature of space-time and more important to the motion of relativistic and massless particles. If GR is assumed, $\gamma_{\rm PPN} = 1$. This approach has been implemented by Collet et al. (2018) using a system in which the lens galaxy is at $z_l = 0.035$. In this work they find $\gamma_{\rm PPN} = 0.97 \pm 0.09$. We apply the same methodology for the SDP.81 system in which the lens galaxy is at $z_l = 0.299$, and therefore is much more distant than the object studied by Collett et al. We use data from three state-of-the-art observatories nowadays: ALMA, HST and VLT/MUSE. Besides the test of GR, we are able to study the mass distribution of the lens galaxy, and we found that about 45% of its mass within $1R_e = 1.156''$ is due to the presence of dark matter, which can be described by a Navarro-Frenk-White profile, in accordance with previous studies using galaxies at similar redshift. However, the main result is the inference of $\gamma_{\rm PPN}$. We notice that our most likely model was deviating from GR within 1σ, but with a very small statistical error, which shows that our inference is dominated by systematic errors. Previous studies that tried to constrain $\gamma_{\rm PPN}$ had the same problem. Using the expected systematics derived by these previous studies, we are able to draw two possible scenarios. In the worst-case scenario, the systematic effects are of the order of 25%, leaves us with $\gamma_{\rm PPN} = 1.0092 \pm 0.0001(\text{stat}) \pm 0.2523(\text{sys})$, which agrees with GR within 1σ. In the best-case scenario, the systematics are of the order of 9%, and the final inference $\gamma_{\rm PPN} = 1.0092 \pm 0.0001(\text{stat}) \pm 0.0908(\text{sys})$, still in accordance with the GR within 1σ.

Key words: Gravitational lensing. Galactic dynamics. General Relativity.

Resumo

A Relatividade Geral (RG) tem sido testada com sucesso em escalas do Sistema Solar. No entanto, para escalas galácticas, esta teoria foi pouco testada. Além disso, várias análises de dados de galáxias dependem do valor da constante de Hubble (H_0), que está atualmente em discussão. Contudo, ao longo das décadas, muitas questões não resolvidas surgiram, e até o presente momento a RG não foi capaz de resolvê-las. Neste trabalho implementamos outro teste para a RG em escalas galácticas, combinando medidas de massa para uma galáxia elíptica que atua como lente gravitacional. Essas medidas podem ser obtidas por cinemática estelar e lentes gravitacionais, e são conectadas por um parâmetro $\gamma_{\rm PPN}$, no formalismo Pós-Newtoniano Parametrizado. Este parâmetro, em nosso contexto, pode ser obtido a partir da razão de dois potenciais diferentes: o potencial Newtoniano Φ, que atua em partículas massivas e não relativísticas; e o potencial de curvatura Ψ, relacionado à curvatura do espaço-tempo, e mais importante para o movimento de partículas relativísticas e sem massa. Assumindo a RG, $\gamma_{\rm PPN} = 1$. Essa abordagem foi implementada por Collett et al. (2018) usando um sistema cuja galáxia lente está em $z_l = 0.035$, no qual eles encontram $\gamma_{\rm PPN} = 0.97 \pm 0.09$. Nós aplicamos essa mesma metodologia para o sistema SDP.81, no qual a galáxia lente está em $z_l = 0.299$ e, portanto, muito mais distante do que o objeto estudado por Collett et al. Para isso, utilizamos dados provenientes de três instrumentos dos melhores observatórios da atualidade: ALMA, HST e VLT/MUSE. Além do teste de RG, fomos capazes de estudar a distribuição de massa da galáxia lente, e descobrimos que cerca de 45% de sua massa dentro de $1R_e = 1.156''$ é devido à presença de matéria escura descrita por um perfil Navarro-Frenk-White, estando nossa estimativa de acordo com estudos anteriores que utilizaram galáxias em redshift semelhante. No entanto, o resultado principal é a inferência de $\gamma_{\rm PPN}$. Percebemos que nosso modelo mais provável indica um desvio da RG dentro de 1σ, mas com um erro estatístico muito pequeno, o que mostra que nossa inferência é dominada por erros sistemáticos. Estudos anteriores que restringiam $\gamma_{\rm PPN}$ com uma abordagem similar a nossa apresentam o mesmo problema. Usando a sistemática derivada desses estudos anteriores, fomos capazes de desenhar dois cenários possíveis. No pior cenário, os efeitos sistemáticos são da ordem de 25%, o que nos deixa com $\gamma_{\rm PPN} = 1.0092 \pm 0.0001({\rm stat}) \pm 0.2523({\rm sys})$, que concorda com a RG dentro de 1σ. Já no melhor dos cenários, a sistemática é da ordem de 9% e a inferência final $\gamma_{\rm PPN} = 1.0092 \pm 0.0001({\rm stat}) \pm 0.0908({\rm sys})$ ainda permanece em acordo com a RG, dentro de 1σ.

Palavras-chave: Lentes gravitacionais. Dinâmica galáctica. Relavitividade Geral.

List of Figures

Figure 1 – Tension present in different measurements of the Hubble constant. . . . 25
Figure 2 – Representation of MUSE integral field spectroscopy. 31
Figure 3 – Schematic representation of the gravitational lensing effect. 32
Figure 4 – Einstein ring in the lens galaxy LRG 3-757. 32
Figure 5 – Gravitational lensing phenomenon at galaxy cluster MACS J0416. . . . 47
Figure 6 – Illustration of lens configuration scheme. 48
Figure 7 – Arbitrary mass distribution for the lens object. 51
Figure 8 – Example of an *exaggerated* external shear on the lensed source. 57
Figure 9 – SDP.81 HST/WFC3-F160W image. 60
Figure 10 – SDP.81 gravitational arcs in the HST/WFC3-F160W. 61
Figure 11 – SDP.81 MUSE collapsed image. 62
Figure 12 – Schematic representation of an array of antennas. 63
Figure 13 – SDP.81 ALMA observations. 65
Figure 14 – SDP.81 data contours. 67
Figure 15 – pPXF fitting of NGC4636 galaxy spectrum from SDSS. 71
Figure 16 – MGE power-law fitting example. 74
Figure 17 – MGE example of the parametrization of the one-dimensional NFW mass profile. 91
Figure 18 – Comparison between deflection angles: MGE parametrization and analytical. 92
Figure 19 – Representation of the "slicing" in the total prior volume. 96
Figure 20 – Nested Sampling for five steps with a collection of three points. 98
Figure 21 – Simulated kinematical map. 104
Figure 22 – Simulated source galaxy. 104
Figure 23 – Simulated deflected source. 105
Figure 24 – JAM modelling result of simulated data. 107
Figure 25 – Lens modelling result of simulated data. 108
Figure 26 – Source reconstruction of simulated data. 109
Figure 27 – SDP.81 MUSE SNR map. 111
Figure 28 – SDP.81 Voronoi binning map. 112
Figure 29 – pPXF model of SDP.81 central bin. 113
Figure 30 – Bootstrapping realization for the central spectrum of SDP.81 lens. . . . 114
Figure 31 – SDP.81 $V_{\rm rms}$ map - all spaxels. 114
Figure 32 – SDP.81 kinematical maps - selected spaxels. 115
Figure 33 – SDP.81 summary of MGE fitting procedure. 117
Figure 34 – MGE model of lens galaxy SDP.81. 117

Figure 35 – SDP.81 HST F160W isophotes. 118
Figure 36 – Rotate SDP.81 data. 120
Figure 37 – SDP.81 most probable lens model. 124
Figure 38 – SDP.81 most probable kinematical model. 125
Figure 39 – SDP.81 source reconstruction. 126
Figure 40 – Two dimensional marginalized posterior for some of the parameters of the combined analysis. 127
Figure 41 – Two dimensional posterior distributions for the parameters of most probable model of SDP.81. 166

List of Tables

Table 1 – Priors of the simulated data. $\mathcal{U}$(lower, upper) represents a non-informative prior, with lower and maximum values accepted. $\mathcal{G}$(mean, sigma) is a Gaussian prior, with mean and dispersion values. All the values are in appropriated physical units. 106
Table 2 – MGE parametrization of HST ePSF. For PSFs, the method assumes circular Gaussians with the same center. 116
Table 3 – MGE parametrization of HST F160W image of SDP.81. First column is the total counts for each Gaussian component, second column is the dispersion in pixel units, and the third column is the observed axial ratio of each Gaussian. 118
Table 4 – Gaussians components used as input for the JAM modelling of SDP.81 lens galaxy. 119
Table 5 – MGE parametrization of MUSE ePSF. 119
Table 6 – Priors of the model. $\mathcal{U}$(lower, upper) represents a non-informative prior, with lower and maximum values accepted. $\mathcal{G}$(mean, sigma) is a Gaussian prior, with mean and dispersion values. Υ_0 is in solar mass per solar luminosity, δ is in arcsec^{-1}, $\log M_{\rm SMBH}$ is in logarithmic solar mass, r_s in arcsec, and shear$_\phi$ is in degrees. All other parameters are dimensionless. 122
Table 7 – Newer Results: The inferred median and 68% confidence intervals (1σ) for the most probable model, assuming a baryonic and dark matter content. 123

List of abbreviations and acronyms

SR	Special Relativity
GR	General Relativity
ΛCDM	Λ-cold dark matter
PPN	Parametrized Post-Newtonian
CT18	Collett et al. (2018)
IFS	Integral field spectroscopy
IFU	Integral field unit
BT08	Binney; Tremaine (2008)
DF	Distribution function
LOSVD	Line-of-sight velocity distribution
MG16	Meneghetti (2016)
H-ATLAS	Herschel Astrophysical Terahertz Large Area Survey
HST	Hubble Space Telescope
MUSE	Multi Unit Spectroscopic Explorer
ALMA	Atacama Large Millimeter/Submillimeter Array
WFC3	Wide Field Camera 3
MAST	Barbara A. Mikulski Archive for Space Telescopes
PSF	Point spread function
ePSF	Effective Point spread function
FWHM	Full width at half maximum
VLT	Very Large Telescope
ESO	European Southern Observatory
CASA	Common Astronomy Software Application

SDSS	Sloan Digital Sky Survey
MGE	Multi-Gaussian Expansion
SMBH	Supermassive black hole
JAM	Jeans Anisotropic Modelling
SLI	Semilinear Inversion Method
NFW	Navarro-Frenk-White
MCMC	Markov Chain Monte Carlo
HPC	High-Performance Computing
SNR	Signal-to-Noise ratio

Contents

1 INTRODUCTION . . . 23
1.1 General Relativity and Alternatives . . . 24
1.2 Gravitational Slip and Post-Newtonian Parameter . . . 26
1.3 Galaxy Dynamics . . . 29
1.4 Gravitational Lensing . . . 31
1.5 When dynamics meets lensing . . . 33

2 THEORY OF GALACTIC DYNAMICS . . . 35
2.1 Collisionless Boltzmann Equation . . . 35
2.1.1 From Theory to Observation . . . 38
2.2 Jeans Equations . . . 41
2.2.1 Jeans Equations in an axisymmetric system . . . 43

3 GRAVITATIONAL LENS THEORY . . . 47
3.1 Lens Equation . . . 48
3.2 Lens Potential . . . 50
3.3 Mapping . . . 53
3.3.1 Jacobian Matrix and transformations between planes . . . 53
3.3.2 Distortion and Magnification . . . 55

4 DATA . . . 59
4.1 HST data . . . 60
4.2 MUSE data . . . 61
4.3 ALMA data . . . 62

5 NUMERICAL METHODS . . . 69
5.1 Integral Field Kinematics . . . 69
5.2 Multi-Gaussian Expansion Formalism . . . 72
5.3 Jeans Anisotropic Modelling . . . 76
5.4 Lens Modelling . . . 79
5.4.1 Semilinear Inversion Method . . . 80
5.4.2 Semilinear Inversion Method with Regularization . . . 82
5.4.3 Semilinear Inversion Method and a Bayes contribution . . . 83
5.4.4 AutoLens . . . 84
5.4.5 AutoLens and mass models . . . 85
5.4.6 Deflection angle and the connection with PPN approach . . . 89

5.5 Self-consistent mass modelling . . . 89
5.6 Mixing everything with dynesty . . . 93
5.6.1 An introduction to Bayesian inference faster than light . . . 93
5.6.2 Nested Sampling Method . . . 95
5.6.2.1 Sampling from the prior volume . . . 97
5.6.2.2 Nested Sampling procedure . . . 97
5.6.2.3 Terminator Criterion . . . 98
5.6.2.4 Nested Sampling Errors . . . 99
5.6.3 DYNESTY . . . 99
5.6.3.1 Bounding Distribution . . . 100
5.6.3.2 Sampling Proposal . . . 100
5.6.3.3 Parallelization . . . 101
5.6.4 The Combined Statistical Model . . . 102
5.7 Proof of Concept . . . 103

6 RESULTS . . . 111
6.1 Kinematical analysis . . . 111
6.2 MGE Models . . . 115
6.3 Preparing the data and setting the simultaneous modelling . . . 119
6.4 Most probable model and Results . . . 122
6.5 Some Final Considerations about the Modelling and Results . . . 128

7 SUMMARY & PERSPECTIVES . . . 131

BIBLIOGRAPHY . . . 135

APPENDIX 149

APPENDIX A – POINT MASS DEFLECTION ANGLE IN PPN APPROACH . . . 151

APPENDIX B – LIST OF SOFTWARES AND PACKAGES . . . 155

APPENDIX C – MGE AND LENSING FORMALISM . . . 157
C.1 A note about units . . . 162

APPENDIX D – TWO DIMENSIONAL POSTERIOR DISTRIBUTION . . . 165

APPENDIX E – PRESS RELEASE: TESTANDO A RELATIVIDADE GERAL EM ESCALA DE GALÁXIAS . . . 167

1 Introduction

From the fundamental interactions of the standard model of particle Physics known to date (gravitational, electromagnetic, weak and strong interaction), gravitation is the most common in our daily lives, and perhaps for this reason, the first to receive a formal mathematical treatment still in the 17th century by Isaac Newton (1643-1797). Probably, the idea of a gravitational behavior of the bodies could be assigned to Aristoteles (385–323 a.C.), even in a primordial way, as shown in Fabris (2020). However, in terms of modern science, the first approach to understand the gravity is Newton's Law of Universal Gravitation.

Roughly speaking, the Universal Gravitational Law could be stated as: *Every massive body attracts every other massive body in the Universe, with a force which is directly proportional to product of its masses, and inversely proportional to the square distance between them.* Or in the mathematical language,

$$\boldsymbol{F} = -G\frac{Mm}{r^2}\hat{\boldsymbol{r}}, \tag{1.1}$$

where G is the Universal Gravitational constant, M and m the masses of the massive bodies, r the distance between them and $\hat{\boldsymbol{r}}$ a versor parallel to the line connecting the bodies.

The Newtonian Mechanics proved to be successful for almost three centuries, however with the advent of the Electromagnetic theory proposed by James Clerk Maxwell (1831-1879) in the 19th century, the Newtonian theory started to show some conflicts with the Electromagnetic theory, mainly related to the nature of light and to the Physics described by different inertial observers.

For the discrepancy between Newtonian Mechanics and the arising Electromagnetic theory, Albert Einstein (1879-1955) developed in 1905 his Theory of Special Relativity (SR; Einstein 1905; Piattella 2020[1]). The main goal of SR was to make the Physics completely covariant in face of a change between inertial coordinate systems, i.e., different observers in different inertial coordinate systems should describe the same physical phenomenon.

The SR brought solutions for many problems involving electromagnetism and Newtonian mechanics, especially the ones which involved switching between inertial coordinate systems. In the light of this new theory, new phenomena have emerged, such as the time contraction, and the unification between space and time, now understood as a unified entity, called space-time. However, SR left aside the phenomenon of gravity.

[1] Einstein's original paper, and Portuguese translate direct from the original in German, by Professor Dr. Oliver F. Piattella.

The next step given by Einstein was to extend SR to non-inertial coordinate systems, so in 1915 Einstein published the Theory of General Relativity (GR; Einstein, 1915). One important discovery made by Einstein during this time was to realize the equivalence between non-inertial references and local gravitational fields, which later was known as the Weak Equivalence Principle.

From the perspective of GR, gravitational phenomenon is seen as a deformation of the space-time where all the particles are immersed. From this point of view, the gravitational field is interpreted as a deformation of the geometry of this space-time, caused by the presence of mass or energy (represented by the energy-momentum tensor, in the general case). For this reason, GR is called a geometric theory of gravity. Because of this geometric description of gravity, John Wheeler (1911-2008) wrote: *"Space-time tells matter how to move; matter tells space-time how to curve"*.

1.1 General Relativity and Alternatives

General Relativity has been tested in many ways (Baker; Psaltis; Skordis, 2015), since it was published in 1915. Precise tests in the Solar System (Bertotti; Iess; Tortora, 2003), Milky Way (Hees et al., 2017), and even strong regimes like supermassive black holes (Akiyama et al., 2019), show that GR is currently the most successful theory of gravity so far (Will, 2014; Ferreira, 2019). All these tests (and many others) place GR in a prominent and confident position. However, over the decades, some tensions arose, and GR has not been able to solve them to date.

One of the tensions that remains open is the so-called Hubble tension (e.g., Riess, 2019; Freedman, 2020). There are no general consensus about the measurement of the Hubble constant[2] (H_0) obtained from different observables. The discrepancy between different measurements of H_0, using measurements of the early Universe (e.g., Abbott et al., 2018; Aghanim et al., 2020) and the more direct measurements from the late Universe (e.g. Wong et al., 2019; Freedman et al., 2020; Riess et al., 2021) can be seen in Fig. 1. Although this discrepancy is noticeable, it is necessary to say that it is non necessarily an issue with GR, but instead could indicate a problem with our knowledge about the current cosmological model.

Another point of discussion is the requirement to include a dark energy component (Alcaniz, 2006; Mortonson; Weinberg; White, 2013) to explain the current expansion of the Universe. Besides requiring the existence of such component, the current cosmological model Λ-cold dark matter (ΛCDM), which is based in GR, also requires the existence of dark matter, that was firstly inferred by Fritz Zwicky (1898-1974; Zwicky 1933; Andernach;

[2] The present rate of the expansion of our Universe.

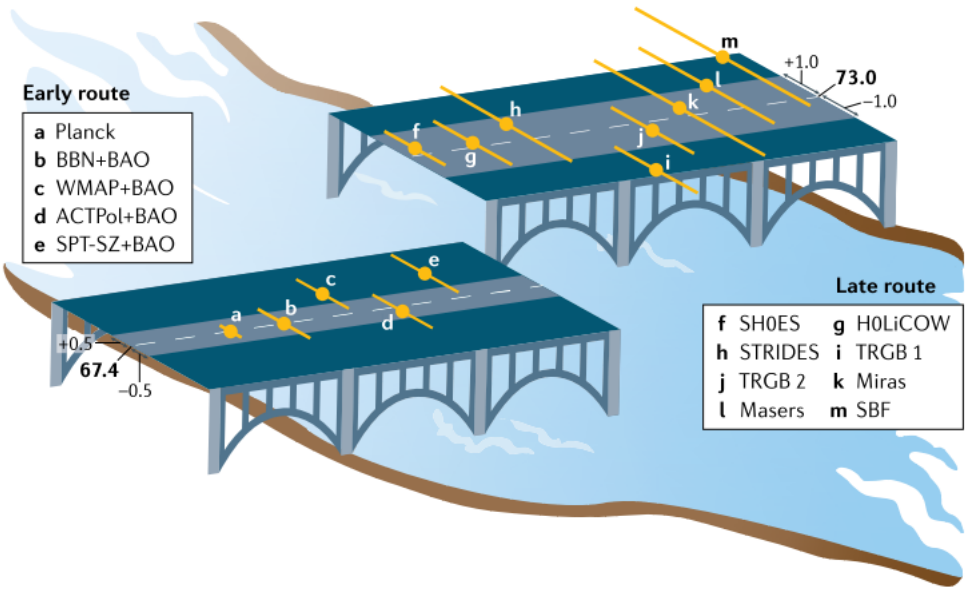


Figure 1 – Tension present in different measurements of the Hubble constant. Late route indicates the observables of the late (present) Universe, while early route are related to observables of the young Universe. Each of the measurements represents a different tracer or a different experiment. This figure was taken from Riess (2019).

Zwicky 2017[3]) from velocity dispersion of galaxies in Coma cluster, but not measured directly.

Finally, GR is not compatible with the principles of Quantum Mechanics. An example of this incompatibility is that from the point of view of Quantum theory, interactions act locally through the exchange of well-defined "quanta", while the gravitational field described by GR is a continuous quantity in nature. Another interesting question is the role of time in each theory, since for Quantum Mechanics time is treated as an independent parameter through which states evolve, while for GR time is a dynamical coordinate in space-time.[4] It is important to highlight that to date, although several alternatives have been proposed, none of them were successful to solve these inconveniences. For all the reasons above mentioned, modified theories of gravity have been proposed to address these tensions (Ishak, 2019).

The modified theories of gravity encompass a large range of models (Clifton et al., 2012), with proposals ranging from modifying the gravity itself to models that include quantum Physics or a different dark sector (other alternatives to dark matter and dark energy) in the Universe. Nevertheless, some of these theories, like DGP braneworld

[3] Original paper, and English Translation by H. Andernach.

[4] A good overview of such inconveniences can be found at <https://en.wikipedia.org/wiki/Quantum_gravity#Points_of_tension>. See also an interesting discussion at <https://physics.stackexchange.com/questions/387/a-list-of-inconveniences-between-quantum-mechanics-and-general-relativity>.

(Dvali; Gabadadze; Porrati, 2000) and $f(R)$ (Multamäki; Vilja, 2006), may have the same expansion history in large scale, which means that, for large distances, only measurements of the scale factor, $a(t)$, are not enough to distinguish GR from other theories of gravitation (Allemandi et al., 2006; Multamäki; Vilja, 2006).

Fortunately, other tests have been proposed using different parameters and ranges of applicability (Bertschinger, 2011; Baker; Psaltis; Skordis, 2014). Considering the Newtonian limit and small scales, i.e., sub-horizon length scales where the Hubble Flow[5] can be considered near constant, a common approach is to use the Parametrized Post-Newtonian (PPN) formalism (Will, 2014; Will, 2018), a well-known framework for testing weak-field regime of gravity. In particular, the γ parameter of this parametrization has received a lot of attention over the last decades (e.g., Bertotti; Iess; Tortora, 2003; Schwab; Bolton; Rappaport, 2010; Cao et al., 2017).

On the other hand, when considering cosmological scales, i.e., horizon length scales where the Hubble Flow is no more constant over the entire domain considered, tests involving the linearized gravity (Carroll, 2004; Mo; van den Bosch; White, 2010) are more common. In this work, we focus our attention on the gravitational slip parameter, which is defined here by the ratio of two scalar potentials that appears in the linear perturbed metric. In the next section we clarify some aspects of this choice, and present the connection between the slip parameter and the γ parameter in the scope of this dissertation.

1.2 Gravitational Slip and Post-Newtonian Parameter

The PPN approach is based on an expansion of the Newtonian metric around small quantities ($\epsilon \sim v/c$)[6], where the terms of the expansion describe corrections to Newtonian gravity. Besides that, the PPN parametrization also requires that the space-time for the system should be asymptotically flat, and the matter described by a perfect fluid. Once we assume a weak-field regime and an expansion around the Newtonian metric, we also expect a Newtonian limit well defined.

As showed by Will (2018), ten terms arise from the PPN expansion, each of them connected to a different correction in gravity, such that *"they measure or indicate general properties of metric theories of gravity"* (see Table 2, pg.31 of Will, 2014). Under reasonable considerations, many of these parameters can be left aside, however one of these parameters has received a special attention from the scientific community, the γ parameter, hereafter γ_{PPN}, which is associated to how responsive the curvature of space-time is to mass.

This particular parameter is interesting due to two main reasons. First, it appears in the first order approximation of the PPN formalism. Second, and maybe more important,

[5] The expansion rate of the Universe.
[6] v is the velocity and c the speed of light.

this parameter directly impacts the electromagnetic waves propagation, such that it is possible to use electromagnetic radiation to impose constraints on it. Furthermore, when assuming GR, $\gamma_{\rm PPN} = 1$, although for alternative theories of gravity its value can be different. Therefore, measurements of $\gamma_{\rm PPN}$ could indicate deviations from GR in scales where we probe it, if its value is statistically different from the unit.

Indeed, the $\gamma_{\rm PPN}$ was successfully tested in different regimes already. Bertotti; Iess; Tortora (2003), using the Cassini spacecraft mission and the Shapiro time delay, found a value for $\gamma_{\rm PPN} = 1 + (2.1 \pm 2.3) \times 10^{-5}$, in agreement with GR. Using measurements of velocity dispersion and gravitational lensing Schwab; Bolton; Rappaport (2010), found $\gamma_{\rm PPN} = 1.01 \pm 0.05$ for a sample of 53 galaxies with redshifts in the range $z \sim 0.1 - 0.3$. Following the same methodology, Cao et al. (2017) extends the sample size of Schwab; Bolton; Rappaport (2010) to 80 lens galaxies in the redshift range $z \sim 0.08 - 0.94$, thereby determining a value of $\gamma_{\rm PPN} = 0.995^{+0.037}_{-0.047}$. Both measurements are in agreement with GR, although systematic uncertainties[7] in lens modelling and velocity dispersion measurement are dominant in this kind of study.

In the cosmological context, many tests of gravity are based on the linearized gravity and growth of small perturbations (Mukhanov et al., 1992; Ferreira, 2019). Considering only scalar perturbations, the conformal Newtonian gauge[8] (Mukhanov et al., 1992; Ma; Bertschinger, 1995) is a simple gauge choice for the linear metric perturbation. In this case, the perturbation is characterized by two scalar potentials, Φ and Ψ, such that the line element is given by

$$dS^2 = -\left(1 + 2\frac{\Phi}{c^2}\right) c^2 dt^2 + \left(1 - 2\frac{\Psi}{c^2}\right) h_{ij} dx^i dx^j, \tag{1.2}$$

where c is the speed of light, t is the time, h_{ij} is the three-metric tensor of constant curvature space, and dx^i are space coordinates.

The first potential Φ is nothing more than the classical Newtonian potential. It is important to note that, for non-relativistic particles ($v^2/c^2 << 1$), dx^i/dt is very small and the spatial term of the metric is negligible, making the Newtonian potential more important to the non-relativistic particles. On the other hand, the second potential Ψ is more relevant to the motion of relativistic particles ($v^2/c^2 \sim 1$), and because it is associated with the spatial curvature of the metric, it is called curvature potential.

According to Mukhanov et al. (1992), the physical interpretation of the potentials Φ and Ψ is quite simple: they measure the amplitude of the scalar perturbations in the metric, given our gauge choice, i.e., the Newtonian gauge. A more phenomenological interpretation is given by Simpson et al. (2012), which bases their interpretation of the

[7] Around 25% for the second sample.
[8] Sometimes called longitudinal gauge.

potentials in the question: "*is the strength of gravity the same on cosmological scales as it is here on Earth?*". If the answer is no then, maybe, the motion of non-relativistic and relativistic particles can be differentially modified due the presence of these two potentials.

In this scenario, we define the gravitational slip parameter as the ratio of the potentials[9]

$$\eta = \frac{\Psi}{\Phi}, \tag{1.3}$$

which can be understood as an effective gravitational coupling between light and matter, since the potentials, in principle, can act differently in both components. Although not explicit in the expression (1.3), the gravitational slip parameter can be a function of time and scale (Ma; Bertschinger, 1995; Bertschinger; Zukin, 2008), and also can assume different values for different gravitational theories. For the standard GR, with a vanish anisotropic stress tensor[10], $\Phi = \Psi$, such that the gravitational slip $\eta = 1$. So, a deviation of this parameter from the unit could indicate a violation of the standard gravitational model based on GR.

Due to the nature of the slip parameter, complementary measurements that are sensible to the different potentials Ψ and Φ can be used to probe it. A straightforward approach is using measurements involving only the classical dynamics (e.g. cluster dynamics, stellar dynamics), which are sensible only to the Newtonian potential, associated with measurements of massless particles (e.g. photons), which is sensible to the sum of the potentials (Bertschinger, 2011). Based on this approach and using data from the galaxy cluster MACS J1206.2-0847 at $z = 0.44$, Pizzuti et al. (2016) probed the slip parameter using galaxy cluster dynamics and gravitational lensing phenomena, finding a value of $\eta(r_{200}) = 1.01^{+0.31}_{-0.28}$, where r_{200} indicates that their inference was made for a fixed radius of a sphere with mass 200 times the critical density of the Universe at that redshift. A more recent test of the slip parameter was performed Collett et al. (2018; CT18), in galaxy scale combining high-resolution lensing data with dynamical data from integral field spectroscopy of the lens galaxy. With this test they imposed strong constraints on η in galactic scales and low redshift ($z = 0.035$). The value of $\eta = 0.97 \pm 0.09$ found in this study is one of the most precise in the literature nowadays, and still in agreement with GR.

Now, it is critical to say that, although the slip parameter and the PPN parameter carry some similarity, they are not necessarily equal or even have the same constraints.

[9] The definition of slip parameter can be a little bit confusing in the literature, but the idea behind all of them is the same: the slip parameter measures a possible variation on the gravitational interaction perceive by relativist and non-relativistic particles.

[10] Here it is important to highlight that, even assuming GR, the unit of the slip parameter is reached only if the anisotropic stress is negligible, which means that for the radiation era (where the anisotropic stress is important) η could assume different values even within GR (Caldwell; Cooray; Melchiorri, 2007; Jain; Zhang, 2008).

Note that the hypotheses behind the two approaches are different, such that the results (or constraints) obtained for one of the parameters cannot be directly generalized for both. In particular, the PPN parametrization requires: i) a constant $\gamma_{\rm PPN}$; ii) a well-defined Newtonian limit. However, neither of these conditions are needed for the η parameter, defined through the line element (1.2), especially when we consider gravitational fields on large scales, i.e., cosmological scales. A recent discussion which stress this difference is done by Toniato; Rodrigues (2021), where the authors emphasize the differences between the parameters and present a relevant generalization for this problem. Another interesting approach, which aim at linking the standard PPN with the cosmological context, given by the line element (1.2), is developed by Sanghai; Clifton (2017). In this work, the authors claimed a new approach, called "parameterized post-Newtonian cosmology" (PPNC), which is based on a set of parameters that are able to consistently model weak-fields and cosmology.

Although not generally equal, under some assumptions, it is possible to compare $\gamma_{\rm PPN}$ and η. To do so, in this work we fix the following assumptions:

i. The space-time metric is given by the line element (1.2), which is in the Newtonian gauge and consider only scalar perturbations;

ii. There is a well-defined Newtonian limit, where the potentials Φ and Ψ still follow the Poisson equation;

iii. The gravitational slip parameter is constant on the relevant scales being studied.

Under these considerations, the metric described by (1.2) becomes equivalent to the PPN metric, and $\gamma_{\rm PPN} \equiv \eta$, such that it is possible to find bounds for both parameters using the same observable constraints. In fact, these assumptions are assumed in CT18, although not all of them are explicitly mentioned.

Finally, since our assumption are more strongly correlated to the hypothesis of PPN parametrization, in this work we will use the notation $\gamma_{\rm PPN}$ and the name Post-Newtonian parameter to refer to parameter of interest. However, keep in mind that it is only possible due to the considerations made.

1.3 Galaxy Dynamics

A way of quantifying the information about the gravitational potential of a given galaxy is via dynamics of its stellar content. This dynamics, in turn, could be accessed through modelling of the kinematic measurements.

For spiral galaxies, the kinematics of stars (or gas) in the disk can be considered (most of the cases) as coherent, thus making it possible to fit a rotation curve for the stars

(or gas) around the galactic center. These rotation curves allow us to measure the mass of the galaxies within a certain radius (Sofue; Rubin, 2001) and, maybe more important, the particular "flat shape" of such curves (Rubin; Ford, 1970) composes one of the first evidences for the existence of dark matter (see Chapter 3 of Courteau et al., 2014).

On the other hand, in the majority of the elliptical galaxies the stellar motion is dominated by dispersion. In this case, it is not possible to build a rotation curve model. However, this disordered motion causes a broadening of the spectral lines observed in the galaxy spectrum, allowing us to measure the velocity dispersion by only fitting the broadening of these lines. Generally speaking, the motion of the stars shifts the spectral line more to the red part of the spectrum if the stars are moving away from the observer, and more to the blue part if the stars are moving towards the observer (the Doppler effect). However, when the stellar motion is disordered (e.g. dispersion dominated), the final spectrum observed is a combination of both effects, making the spectral line broad as the velocity dispersion of the emitters increases. Lastly, assuming a Virial regime, it is possible, from the velocity dispersion, to measure the dynamical mass of an elliptical galaxy (Cappellari et al., 2006; Cappellari et al., 2013).

Many other properties of galaxies could be derived looking at their kinematics, such as its dark matter content, and morphology. But since it is not our goal in this dissertation, we refer the reader to the following reviews: Courteau et al. (2014), and Cappellari (2016b).

To make possible the dynamical modelling of galaxies, spectroscopy observations are needed, in order to obtain their spectral signatures. For a long time, only integrated spectra were available, usually measured within a certain aperture radius, limiting the possibility of decomposing different kinematic characteristics, especially for those galaxies at higher redshift. The next step for the spectroscopical observation in astronomy was the introduction of the long slit spectroscopy, which made possible to collect spectroscopic data along a given axis, previously decided upon in the experimental setup. The long slit spectroscopy was rapidly improved by the introduction of the multi-slit or multi-fiber spectroscopy, which allows the observation of many targets at the same time. However, the last breakthrough in instrumentation comes from the integral field spectroscopy (IFS)[11], which allowed to resolve spectrally and spatially thousands of galaxies in different redshift and wavelength ranges.

The idea behind the IFS is to collect a spectrum in each pixel of an image giving rise to what we call spaxel. Since each spaxel has its own spectrum and position in the image, it is possible to build a data cube with two spatial dimensions and one spectral dimension, which splits the light into different wavelengths (see Fig. 2).

[11] Also known as integral field unit (IFU).

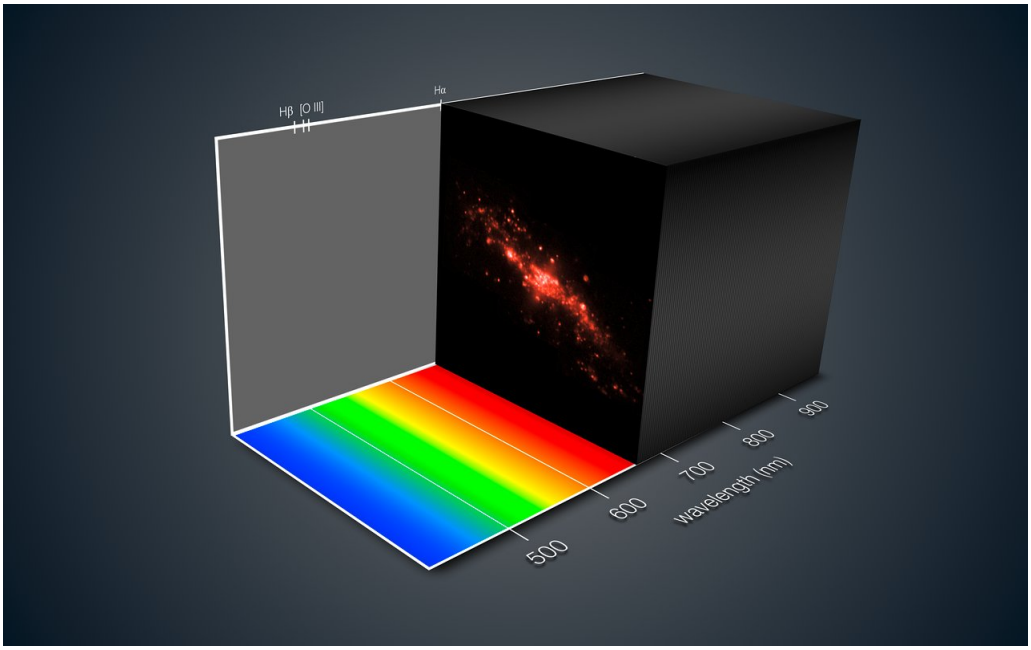


Figure 2 – Representation of MUSE integral field spectroscopy. We can see an image of galaxy NGC 4650A along the two dimensional axes, while in the third axis we see the variation in the wavelength. Image taken from: ESO/MUSE consortium/R. Bacon/L. Calçada

As a complement, it is important to highlight that the stellar motion is described only by the Newtonian potential Φ, and we left the details of the theory for Chapter 2.

1.4 Gravitational Lensing

In GR, the gravitational field is interpreted as a deformation of space-time. This deformation is due to the presence of energy and mass. If two objects are aligned in the line-of-sight, and the one in the foreground is very massive, the light coming from the object behind it (in the background) can be deflected as the result of the deformation of space-time. It means that, since space-time is curved by the presence of energy and matter, the light rays approaching this region of influence will follow curved paths, as shown in Fig. 3. This phenomenon is called gravitational lensing, where the foreground object is called the lens and the background object is the source.

Due to this effect, the light coming from the source can appear in a different position than the actual source. Moreover, in some cases, multiple and distorted images can form (Suyu et al., 2013; Wong et al., 2016). If the lens is spherical, and it is perfectly aligned with the source, the image of the source is distorted into an Einstein ring (see Fig. 4). The radius of the Einstein ring can be expressed as a function of the mass of the lens, the spatial curvature and the ratio of three angular diameter distances, to know: the distances between the observer, lens and source.

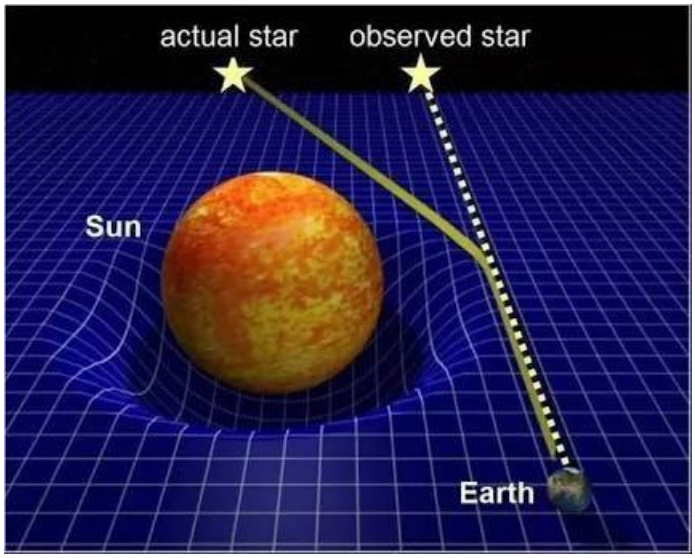


Figure 3 – Schematic representation of the gravitational lensing effect. Assuming an observer on Earth, the Sun acts like a lens in the foreground, and the star is the source in the background. In the presence of the Sun the space-time, represented by the blue grid, becomes curved and the light follows the curved path shown in continuum yellow. The dashed line shows the apparent position of the source as seen by the observer. The deflection of light is strongly exaggerated as a means to emphasize the effect. Image taken from Quora - What is gravitational lensing? Is it different from normal lensing?

Finally, due to the relativistic nature of the photons that compose the light, the gravitational lensing effect is sensible to the two different potentials, Φ and Ψ, described in Section 1.2.

Figure 4 – Einstein ring in the lens galaxy LRG 3-757, also known as Horseshoe. Near the center, more in red, we can see the lens galaxy itself, surrounded by a beautiful gravitational arc, in blue. Image from Hubble Space Telescope, WFC3. Credit: ESA/Hubble & NASA

Following Meneghetti (2016), and making use of the line element (eq. 1.2), it is

relatively easy to find the deflection angle for a point mass object. After some considerations (see Appendix A for demonstration), the deflection angle for a point mass is

$$\tilde{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = (1+\gamma_{\rm PPN})\frac{2GM}{c^2}\frac{\hat{\boldsymbol{\xi}}}{\xi}. \tag{1.4}$$

In the GR framework, when $\gamma_{\rm PPN} = 1$, this expression is equal to that found in the literature

$$\tilde{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4GM}{c^2}\frac{\hat{\boldsymbol{\xi}}}{\xi}. \tag{1.5}$$

The formalism that describes the lensing phenomena will be present in Chapter 3.

1.5 When dynamics meets lensing

Combining lensing information with kinematical measurements is a good way to break degeneracies in the parameter space of models describing the lens potential (Koopmans, 2006; Chirivì et al., 2020), especially when the parameters are complementary[12]. Besides that, several other parameters of the lens model can be tested, such as mass-to-light ratio of the lens galaxy, dark matter density slope and kinematic anisotropy of the stellar motion (Koopmans; Treu, 2003; Barnabè et al., 2012). Among them, the most important to us, is the Post-Newtonian parameter, $\gamma_{\rm PPN}$.

As previously stated, Post-Newtonian approach was already probed using gravitational lensing (in particular the strong regime) and galaxy dynamics. However, these probes contain some limitations. Using a sample of lenses, Cao et al. (2017) were able to probe the $\gamma_{\rm PPN}$ in a redshift range covering $0.08 < z_l < 0.94$, and found a value in agreement with GR. At the same time, the approach used is limited by a simple mass model and a single kinematic measure (the integrated velocity dispersion along the line-of-sight). On the other hand, CT18 using a nearby lens system ($z_l = 0.035$) were able to constrain the Post-Newtonian parameter with much more precision, due to the quality of the data (in particular the spatially-resolved kinematic map available from IFS observations), and the more sophisticated and self-consistent mass model used in their analysis.

In this work we propose to test GR using a gravitational lens system H-ATLAS J090311.6+003906 (SDP.81; Negrello et al., 2014) at intermediate redshift ($z_l = 0.299$), using the same methodology found in CT18, although with some additional data coming from interferometry for the lensing part of the analysis (better explained in Chapter 4). We want to probe the $\gamma_{\rm PPN}$ parameter by modelling lensing and dynamical masses (the two potentials described above) simultaneously in a self-consistent approach. In addition

[12] One good example is when the lens mass profile is assumed to be a singular isothermal sphere, characterized by a velocity dispersion, and we have the velocity dispersion information coming from kinematical constraints.

to that, we expect to be able to reconstruct the source image, and study other physical properties of the lens galaxy, such as dark matter content and stellar velocity dispersion distribution.

The dissertation structure is as follows: Chapter 2 presents galaxy dynamics theory, whose central focus will be Jeans equations. Chapter 3 is an overall of the lensing theory, focusing in the strong regime. In Chapter 4 we describe the system used in this work and the data available, while Chapter 5 describes the numerical implementation of the model. Finally, in Chapters 6 and 7 we present our results and summary, respectively. Some appendices are attached to the end of this dissertation as a complement to the text. Also, all the codes, scripts, softwares, and data are available (with some documentation still under development) at <https://github.com/carlosRmelo/Master-Degree.git>.

Throughout this dissertation, the cosmological parameters assumed are: $H_0 = 67.7$ Km s^{-1}Mpc^{-1}, $\Omega_\Lambda = 0.6911$, $\Omega_m = 0.3089$ (Ade et al., 2016).

2 Theory of Galactic Dynamics

In this Chapter we describe the Theory of Galactic Dynamics, and most of the following results are based on Chapter 4 of Binney; Tremaine (2008) (hereafter BT08), including the notation, for easy identification with the original. However, here we focus in the relevant aspects to this work, but trying not to completely lose sight of formalism. Some shortcuts can be found in Cappellari (2008, 2020), especially the connection between the theory, observation, and modelling.

2.1 Collisionless Boltzmann Equation

For collisionless systems, like elliptical galaxies without interactions[1], it is not practical (or even possible) to describe the orbits of all the stars bounded to the system. In this case, it is more appropriate to treat the question in a statistical way, i.e., finding the probability to spot a star in a six-dimensional volume $d^3\boldsymbol{r}d^3\boldsymbol{v}$ in the parameter space[2], also called phase space in this context, with a given position $\boldsymbol{r}$ and velocity (momentum) $\boldsymbol{v}$ ($\boldsymbol{p}$). In order to do that, we need a distribution function (DF), $f(\boldsymbol{r}, \boldsymbol{v}, t)$, such that $f(\boldsymbol{r}, \boldsymbol{v}, t)\, d^3\boldsymbol{r}d^3\boldsymbol{v}$ is the probability to find a star in the range $[\boldsymbol{r}, \boldsymbol{v}]$ and $[\boldsymbol{r} + d\boldsymbol{r}, \boldsymbol{v} + d\boldsymbol{v}]$ in the parameter space, given a time t. Assuming that every star has the same probability to be found, and by the definition of the DF, we have

$$\int d^3\boldsymbol{r}d^3\boldsymbol{v} f(\boldsymbol{r}, \boldsymbol{v}, t) = 1, \tag{2.1}$$

where the integral is evaluated over all phase space.

Now, let $\boldsymbol{\omega} = (\boldsymbol{r}, \boldsymbol{v})$ be the usual coordinate system in the parameter space, and μ a small volume in this space[3]. The probability to find a star in this volume is just

$$P = \int_\mu d^6\boldsymbol{\omega} f(\boldsymbol{\omega}). \tag{2.2}$$

Now, consider a new set of arbitrary parameters $\boldsymbol{W}$ and its own DF, $F(\boldsymbol{W})$. If μ is small enough, the DFs could be considered constant inside μ, so

$$P = f(\boldsymbol{\omega}) \int_\mu d^6\boldsymbol{\omega} = F(\boldsymbol{W}) \int_\mu d^6\boldsymbol{W}, \tag{2.3}$$

[1] Later we will discuss the validity limits for a system that can be considered collisionless.
[2] $\boldsymbol{r}$ is a three-dimensional position vector and $\boldsymbol{v}$ is a three velocity vector.
[3] In BT08 this small volume is denoted by ν, but here we choose to replace it by μ to avoid a possible confusion with other quantity that we will define further.

and if $\boldsymbol{W}$ are the canonical coordinates, $\int_\mu d^6\boldsymbol{\omega} = \int_\mu d^6\boldsymbol{W}$, letting us with $f(\boldsymbol{\omega}) = F(\boldsymbol{W})$[4].

Therefore, we conclude that the DF has the same value in an arbitrary volume in the parameter space for any canonical coordinate system. For convenience, we choose $\boldsymbol{\omega} = (\boldsymbol{q}, \boldsymbol{p})$ to be our canonical coordinates, where $\boldsymbol{q}$ are the generalized coordinates, and $\boldsymbol{p}$ are the generalized momenta associated to $\boldsymbol{q}$.

For those already familiar with Quantum Mechanics, the fact that the probability was a conservative quantity is nothing new. For those who are not familiar, think about a fluid that can change its position or state in time, but its total mass is still the same. We cannot create or destroy the mass, only change it. Similarly, the probability of finding a star in a given position in the parameter space can change with time, but the probability density described by DF is conserved. In this case, the continuity equation for the probability density is

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial \boldsymbol{\omega}} \cdot (f\dot{\boldsymbol{\omega}}) = 0, \tag{2.4}$$

where $\frac{\partial}{\partial \boldsymbol{\omega}} \equiv \boldsymbol{\nabla}_\omega$, and the above dotted notation represents the derivative with respect to time. In this expression, and most of the cases hereafter, we also relax the notation for $f = f(\boldsymbol{r}, \boldsymbol{v}, t)$.

Using Hamilton equations (Appendix D of BT08), where H is the Hamiltonian of the system,

$$\dot{\boldsymbol{q}} = \frac{\partial H}{\partial \boldsymbol{p}} \qquad \dot{\boldsymbol{p}} = -\frac{\partial H}{\partial \boldsymbol{q}}, \tag{2.5}$$

we can replace $\dot{\boldsymbol{\omega}} = (\dot{\boldsymbol{p}}, \dot{\boldsymbol{q}})$ in (2.4).

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{\omega}} \cdot (f\dot{\boldsymbol{\omega}}) &= \frac{\partial}{\partial \boldsymbol{q}} \cdot (f\dot{\boldsymbol{q}}) + \frac{\partial}{\partial \boldsymbol{p}} \cdot (f\dot{\boldsymbol{p}}) \\ &= \frac{\partial}{\partial \boldsymbol{q}} \cdot \left(f\frac{\partial H}{\partial \boldsymbol{p}}\right) - \frac{\partial}{\partial \boldsymbol{p}} \cdot \left(f\frac{\partial H}{\partial \boldsymbol{q}}\right) \\ &= \frac{\partial f}{\partial \boldsymbol{q}} \cdot \frac{\partial H}{\partial \boldsymbol{p}} + f\frac{\partial^2 H}{\partial \boldsymbol{q}\partial \boldsymbol{p}} - \frac{\partial f}{\partial \boldsymbol{p}} \cdot \frac{\partial H}{\partial \boldsymbol{q}} - f\frac{\partial^2 H}{\partial \boldsymbol{p}\partial \boldsymbol{q}}. \end{aligned} \tag{2.6}$$

However, since H should be differentiable and smooth, the second-order partial derivatives commute. Consequently,

$$\frac{\partial}{\partial \boldsymbol{\omega}} \cdot (f\dot{\boldsymbol{\omega}}) = \dot{\boldsymbol{q}} \cdot \frac{\partial f}{\partial \boldsymbol{q}} + \dot{\boldsymbol{p}}\frac{\partial f}{\partial \boldsymbol{p}}, \tag{2.7}$$

[4] A more careful reader will notice that we have hidden the explicit temporal dependence in the last passages. This was done not only to save notation, but also because such results must be valid for any time t.

where we use the Hamilton equations again. Replacing this equation in (2.4), we finally get:

$$\frac{\partial f(\boldsymbol{r},\boldsymbol{v},t)}{\partial t} + \dot{\boldsymbol{q}} \cdot \frac{\partial f(\boldsymbol{r},\boldsymbol{v},t)}{\partial \boldsymbol{q}} + \dot{\boldsymbol{p}} \cdot \frac{\partial f(\boldsymbol{r},\boldsymbol{v},t)}{\partial \boldsymbol{p}} = 0, \tag{2.8}$$

which is a partial differential equation for the distribution function f in the phase space. This equation is called **collisionless Boltzmann equation**.

Assuming an inertial Cartesian coordinate system, and a Hamiltonian in the form $H = v^2/2 + \Phi(\boldsymbol{r},t)$[5], where $\Phi \equiv \Phi(\boldsymbol{r},t)$ is the Newtonian potential per unit of mass (see footnote 5), the collisionless Boltzmann equation can be written as

$$\frac{\partial f(\boldsymbol{r},\boldsymbol{v},t)}{\partial t} + \boldsymbol{v} \cdot \frac{\partial f(\boldsymbol{r},\boldsymbol{v},t)}{\partial \boldsymbol{r}} - \frac{\partial \Phi}{\partial \boldsymbol{r}} \cdot \frac{\partial f(\boldsymbol{r},\boldsymbol{v},t)}{\partial \boldsymbol{v}} = 0. \tag{2.9}$$

Once we have the collisionless Boltzmann equation, we need to highlight its limitations. The first one comes from the fact that we assume that the objects described by the DF are conserved with time, although is not true in general. Stars are born and die all the time in a galaxy, then for a more precise description of the system, equation (2.8) should be corrected to

$$\frac{\partial f}{\partial t} + \boldsymbol{v} \cdot \frac{\partial f}{\partial \boldsymbol{r}} - \frac{\partial \Phi}{\partial \boldsymbol{r}} \cdot \frac{\partial f}{\partial \boldsymbol{v}} = \mathcal{B}_{rate} - \mathcal{D}_{rate}, \tag{2.10}$$

where $\mathcal{B}_{rate} = \mathcal{B}(\boldsymbol{r},\boldsymbol{v},t)$ is the birth rate of stars and $\mathcal{D}_{rate} = \mathcal{D}(\boldsymbol{r},\boldsymbol{v},t)$ is the death rate of stars. In equation (2.8), the right side is set equal to zero, because we assume that the birth and death rates compensate each other, or both are identical null. This approximation is valid for a large range of applications, especially for elliptical galaxies, and the main requirement is that the right side of equation (2.10) should be much smaller than the left side. As shown in BT08, equation (4.16), this approximation is valid if

$$\epsilon = \left| \frac{\mathcal{B}_{rate} - \mathbf{D}_{rate}}{f/t_{cross}} \right| << 1, \tag{2.11}$$

where t_{cross} is the crossing time, the time taken for a particular star to complete an entire orbit. For those galaxies dominated by old stellar populations, the birth rate is negligible, while the death rate is quite small, due to the relative long life of these stars[6]. Therefore, for a long time (Gyr), many crossing times can elapse, and the approximation is valid.

[5] This strange notation without mass is motivated by the fact that the gravitational energy of a body is proportional to its mass. Then, is "common" to write the gravitational potential energy as $V(\boldsymbol{r},t) = m\Phi(\boldsymbol{r},t)$, and define the total energy per unit mass as $v^2/2 + \Phi(\boldsymbol{r},t)$. In this notation, the kinetic energy per unit mass is just $v^2/2$, and sometimes referred as kinetic energy only, while $\Phi(\boldsymbol{r})$ is the gravitational potential per unit mass, also called simply by gravitational energy.

[6] See the discussion in BT08 Pgs. 278-279 and Chapter 5 of Binney; Merrifield (1998).

The second limitation comes from the fact that our DF assumes no correlation between the stars, i.e, our DF is separable, and the probability to find a star in a position $(\boldsymbol{r_1}, \boldsymbol{v_1})$ does not change the probability to find another star in a different position $(\boldsymbol{r_2}, \boldsymbol{v_2})$. However, since gravity is always acting in the stars, the position and velocity of a given star can be affected by its neighbors. Fortunately, for Newtonian Mechanics, the gravity force should decrease with the square distance from the source, see equation (1.1), turning the contribution of individual stars much smaller than the contribution of the entire galaxy, since the distances between stars are very large and its masses are considerably smaller than the total mass of the galaxy. Because of that, the motion of the stars in the galaxy is practically smooth, and we can consider it not affected by the neighbouring companions[7].

2.1.1 From Theory to Observation

For our Galaxy, the Milky Way, sometimes it is possible to measure the position of a given star and its velocity (Katz et al., 2018), but for the majority of the galaxies it is not possible due to the distance and instrumental resolution, and again we need to use an alternative way to obtain this information. Now we will introduce one alternative based on the DF presented previously.

Since our DF is defined by solving[8] equation (2.8), we can easily use it to obtain the expected value for different physical parameters, like the total luminosity or total mass, just multiplying $f(\boldsymbol{r}, \boldsymbol{v}, t)\, d^3\boldsymbol{r} d^3\boldsymbol{v}$ by the correspondent physical quantity. For example, using the definition of expected value[9], the expected total luminosity can be obtained by

$$\boldsymbol{E}[L] = \int L f(\boldsymbol{r}, \boldsymbol{v}, t) d^3\boldsymbol{r} d^3\boldsymbol{v}, \tag{2.12}$$

where the integral is over all the volume considered for the analysis, and L[10] is the luminosity of the stellar population in this volume.

Another important quantity to define is the probability per unit of volume, i.e., the probability to find a particular star at position $\boldsymbol{r}$ independent of its velocity. This is expressed as follows,

$$\nu(\boldsymbol{r}, t) = \int f(\boldsymbol{r}, \boldsymbol{v}, t) d^3\boldsymbol{v}. \tag{2.13}$$

Multiplying $\nu(\boldsymbol{r}, t)$ by the total number of stars of a given population, we have the number density of stars in the real space, $n(\boldsymbol{r}, t) = N\nu(\boldsymbol{r}, t)$. In the same way, multiplying

[7] More information can be found at <http://astro.utoronto.ca/~bovy/AST1420/notes/notebooks/05.-Equilibria-Spherical-Collisionless-Systems.html>.

[8] We will get back to this topic in Chapter 5.

[9] <https://en.wikipedia.org/wiki/Expected_value>

[10] We left aside the functional dependence of L because its depend on the model adopted.

(2.13) by the luminosity L gives us the three-dimensional luminosity density of a given stellar population,

$$j(\boldsymbol{r}, t) = L\nu(\boldsymbol{r}, t). \tag{2.14}$$

Note that this definition is in agreement with the previous definition of expected value, since it is enough to integrate the last equation in the volume $d^3\boldsymbol{r}$ to obtain the expected value of luminosity.

Therefore, assuming some model for the three-dimensional luminosity density, it is possible to derive $\nu(\boldsymbol{r})$. This is a key factor that facilitates the connection between theory and observation, and that will be addressed in Chapter 5, when we will talk about the adopted models.

Using the DF we can also obtain information on the velocity of stars. Dividing $f(\boldsymbol{r}, \boldsymbol{v}, t)$ by $\nu(\boldsymbol{r}, t)$, we get the velocity probability distribution for the stars in a given position $\boldsymbol{r}$,

$$P_{\boldsymbol{r}}(\boldsymbol{v}) = \frac{f(\boldsymbol{r}, \boldsymbol{v}, t)}{\nu(\boldsymbol{r}, t)}. \tag{2.15}$$

For distant galaxies, this distribution can only be inferred from the line-of-sight velocity distribution (LOSVD), which describes the velocity along the line-of-sight of stars with velocity in the range $[\boldsymbol{v}, \boldsymbol{v} + d\boldsymbol{v}]$. From the observational perspective, the LOSVD can be constrained by observations using spectroscopy, situation in which the velocity along the line-of-sight can be measured by the width of spectral lines.

We can describe the fraction of the stars with velocity ranging from $\boldsymbol{v}$ to $\boldsymbol{v} + d\boldsymbol{v}$ using equations (2.13) and (2.15). So, let $\hat{\boldsymbol{s}}$ be a versor parallel to the line-of-sight fixed in the galaxy center. For most of galaxies, the distance between the observer and any point $\boldsymbol{r}$ in the observed galaxy is large enough, making possible to assume that all vectors starting from the observer, to any point in the galaxy, is parallel to $\hat{\boldsymbol{s}}$. Assuming that, we can define,

$$\boldsymbol{r}_{\parallel} \equiv \hat{\boldsymbol{s}} \cdot \boldsymbol{r}, \quad \boldsymbol{v}_{\parallel} \equiv \hat{\boldsymbol{s}} \cdot \boldsymbol{v}, \tag{2.16}$$

as the components of $\boldsymbol{r}$ and $\boldsymbol{v}$ parallels to the line-of-sight, respectively. As a consequence, the components in the sky plane are

$$\boldsymbol{r}_{\perp} \equiv \boldsymbol{r} - \hat{\boldsymbol{s}} \cdot \boldsymbol{r}, \quad \boldsymbol{v}_{\perp} \equiv \boldsymbol{v} - \hat{\boldsymbol{s}} \cdot \boldsymbol{v}. \tag{2.17}$$

With these definitions, we can write the density probability fraction of finding a star in a position $\boldsymbol{r}_{\perp}$ in the sky plane and velocity $\boldsymbol{v}_{\parallel}$ along the line-of-sight as (BT08

equation 4.23):

$$F(\boldsymbol{r}_\perp, \boldsymbol{v}_\parallel) = \frac{\int d\boldsymbol{r}_\parallel d^2\boldsymbol{v}_\perp f(\boldsymbol{r}, \boldsymbol{v}, t)}{\int d\boldsymbol{r}_\parallel d^3\boldsymbol{v} f(\boldsymbol{r}, \boldsymbol{v}, t)}, \tag{2.18}$$

where we use (2.13) and (2.15). To make the interpretation clearer, consider separately the terms in this expression. First, the numerator gives us the probability density distribution of finding a star located at a given point in the sky plane (note that we integrate over all the possible values parallel to the line-of-sight) and with some velocity along the line-of-sight (we integrate over all velocities in the sky plane). While, the denominator gives the probability density distribution of finding a star at some point in the plane of the sky, regardless of its velocity. Then, $F(\boldsymbol{r}_\perp, \boldsymbol{v}_\parallel)$ is a fraction of the probability distribution, which gives us the chance of finding a star located in the plane of the sky with some velocity along the line-of-sight.

The LOSVD often is characterized by two main physical quantities: the mean velocity along line-of-sight $\overline{v}_\parallel(\boldsymbol{r}_\perp)$ and its velocity dispersion $\sigma_\parallel(\boldsymbol{r}_\perp)$[11]. For the mean velocity we have

$$\begin{aligned} \overline{v}_\parallel(\boldsymbol{r}_\perp) &= \int d\boldsymbol{v}_\parallel v_\parallel F(\boldsymbol{r}_\perp, \boldsymbol{v}_\parallel) = \frac{\int d\boldsymbol{r}_\parallel d^3\boldsymbol{v}\, v_\parallel\, f(\boldsymbol{r}, \boldsymbol{v}, t)}{\int d\boldsymbol{r}_\parallel d^3\boldsymbol{v}\, f(\boldsymbol{r}, \boldsymbol{v}, t)} \\ &= \frac{\int d\boldsymbol{r}_\parallel\, \nu(\boldsymbol{r}, t)\, \hat{\boldsymbol{s}} \cdot \overline{\boldsymbol{v}}}{\int d\boldsymbol{r}_\parallel\, \nu(\boldsymbol{r}, t)}, \end{aligned} \tag{2.19}$$

where we define the mean velocity at the point $\boldsymbol{r}$ as

$$\overline{\boldsymbol{v}}(\boldsymbol{r}) = \int d^3\boldsymbol{v}\, \boldsymbol{v}\, P_{\boldsymbol{r}}(\boldsymbol{v}) = \frac{1}{\nu(\boldsymbol{r}, t)} \int d^3\boldsymbol{v}\, \boldsymbol{v}\, f(\boldsymbol{r}, \boldsymbol{v}, t). \tag{2.20}$$

And finally, for velocity distribution $\sigma_\parallel(\boldsymbol{r}_\perp)$:

$$\sigma^2_\parallel(\boldsymbol{r}_\perp) = \int d\boldsymbol{v}_\parallel\, \left(v_\parallel - \overline{v}_\parallel\right)^2\, F(\boldsymbol{r}_\perp, \boldsymbol{v}_\parallel) = \frac{\int d\boldsymbol{r}_\parallel d^3\boldsymbol{v}\, \left(\hat{\boldsymbol{s}} \cdot \boldsymbol{v} - \overline{v}_\parallel\right)^2\, f(\boldsymbol{r}, \boldsymbol{v}, t)}{\int d\boldsymbol{r}_\parallel\, \nu(\boldsymbol{r}, t)}. \tag{2.21}$$

We take the opportunity to define the velocity dispersion tensor $\sigma^2_{ij}(\boldsymbol{r})$, which gives us the scatter (or spread) of the velocity around the mean:

[11] Here both quantities are scalars, since they are projection onto the sky plane.

$$\begin{aligned}\sigma_{ij}^2(\boldsymbol{r}) &= \frac{1}{\nu(\boldsymbol{r},t)}\int d^3\boldsymbol{v}\ (v_i-\overline{v}_i)\,(v_j-\overline{v}_j)\ f(\boldsymbol{r},\boldsymbol{v},t)\\ \sigma_{ij}^2(\boldsymbol{r}) &= \overline{v_i\,v_j}-\overline{v_i}\ \overline{v_j},\end{aligned} \tag{2.22}$$

where the indices (i,j) can assume values $(1,2,3)$, representing the coordinated system. Also note that this tensor is symmetric in its indices, $\sigma_{ij}^2(\boldsymbol{r}) = \sigma_{ji}^2(\boldsymbol{r})$, therefore only six components are independent.

2.2 Jeans Equations

In the last Section, we derive the Boltzmann equation in its colisioness form, and also its relation with some observables, such as mean velocity dispersion and luminosity. However, these connections, in certain ways, assume that we know the functional form for DF. Since equation (2.8) is quite general, and $f(\boldsymbol{r},\boldsymbol{v},t)$ is a function of seven variables, a large range of solutions can be obtained. These limitations require us to adopt certain assumptions and, in this Section, we will present an approach that tries to avoid such complications, focusing in deriving the velocity moments without the need of solving the Boltzmann equation for f.

During this Section, we will often use the Einstein summation convention for repeated indices. Briefly, a repeated index in the same term represents a summation over that index. Moreover, unless otherwise stated, all indices vary from $(1,2,3) = (x,y,z)$ in Cartesian coordinates.

We star integrating equation (2.9) over all the velocities,

$$\int d^3\boldsymbol{v}\,\frac{\partial f}{\partial t} + \int d^3\boldsymbol{v}\,v_i\frac{\partial f}{\partial r_i} - \frac{\partial\Phi}{\partial r_i}\int d^3\boldsymbol{v}\,\frac{\partial f}{\partial v_i} = 0. \tag{2.23}$$

Assuming that the velocity range over which we integrate does not depend on time, and the positions r_i an velocities v_i are independent, we can re-write the above expression as

$$\begin{aligned}\frac{\partial}{\partial t}\int d^3\boldsymbol{v}\,f + \frac{\partial}{\partial r_i}\left(\int d^3\boldsymbol{v}\,v_i\,f\right) - \frac{\partial\Phi}{\partial r_i}\int d^3\boldsymbol{v}\,\frac{\partial f}{\partial v_i} &= 0\\ \frac{\partial\nu(\boldsymbol{r},t)}{\partial t} + \frac{\partial}{\partial r_i}\left(\frac{\nu(\boldsymbol{r},t)}{\nu(\boldsymbol{r},t)}\int d^3\boldsymbol{v}\,v_i\,f\right) - \frac{\partial\Phi}{\partial r_i}\int d^3\boldsymbol{v}\,\frac{\partial f}{\partial v_i} &= 0\\ \frac{\partial\nu(\boldsymbol{r},t)}{\partial t} + \frac{\partial}{\partial r_i}\left(\nu(\boldsymbol{r},t)\,\underbrace{\frac{1}{\nu(\boldsymbol{r},t)}\int d^3\boldsymbol{v}\,v_i\,f}_{\overline{v_i}}\right) - \frac{\partial\Phi}{\partial r_i}\int d^3\boldsymbol{v}\,\frac{\partial f}{\partial v_i} &= 0\\ \frac{\partial\nu(\boldsymbol{r},t)}{\partial t} + \frac{\partial}{\partial r_i}\left(\nu(\boldsymbol{r},t)\,\overline{v_i}\right) - \frac{\partial\Phi}{\partial r_i}\int d^3\boldsymbol{v}\,\frac{\partial f}{\partial v_i} &= 0.\end{aligned} \tag{2.24}$$

However, once there is no stars with velocities arbitrarily large, we expect that $f(\boldsymbol{r}, \boldsymbol{v}, t) \to 0$ when $|\boldsymbol{v}| \to \infty$. Then, applying the divergence theorem in the last term in equation (2.24) it vanishes, resulting in

$$\frac{\partial \nu(\boldsymbol{r}, t)}{\partial t} + \frac{\partial \left(\nu(\boldsymbol{r}, t)\, \overline{v_i} \right)}{\partial r_i} = 0. \tag{2.25}$$

Now, we multiply (2.9) by v_j, and integrate again over all the velocities

$$\frac{\partial}{\partial t} \int d^3\boldsymbol{v}\, v_j\, f + \frac{\partial}{\partial r_i} \left(\int d^3\boldsymbol{v}\, v_i\, v_j\, f \right) - \frac{\partial \Phi}{\partial r_i} \int d^3\boldsymbol{v}\, v_j\, \frac{\partial f}{\partial v_i} = 0. \tag{2.26}$$

The last term can be replaced by,

$$\begin{aligned} \int d^3\boldsymbol{v}\, v_j\, \frac{\partial f}{\partial v_i} &= \cancelto{0}{v_j\, f \Big|_{-\infty}^{+\infty}} - \int d^3\boldsymbol{v}\, f\, \frac{\partial v_j}{\partial v_i} \\ &= -\int d^3\boldsymbol{v}\, f\, \delta_{ij} \\ &= -\delta_{ij}\, \nu(\boldsymbol{r}, t). \end{aligned} \tag{2.27}$$

Where we have used the fact that the velocities are independents, $\partial v_j / \partial v_i = \delta_{ji} = \delta_{ij}$, and $f \to 0$ when $|\boldsymbol{v}| \to \infty$. Coming back to equation (2.26),

$$\begin{aligned} \frac{\partial}{\partial t} \underbrace{\int d^3\boldsymbol{v}\, v_j\, f}_{\nu\, \overline{v_j}} + \frac{\partial}{\partial r_i} \left(\underbrace{\int d^3\boldsymbol{v}\, v_i\, v_j\, f}_{\overline{v_i v_j}\, \nu(\boldsymbol{r}, t)} \right) - \frac{\partial \Phi}{\partial r_i}\, (-\delta_{ij} \nu) &= 0 \\ \frac{\partial \left(\nu(\boldsymbol{r}, t)\, \overline{v_j} \right)}{\partial t} + \frac{\partial \left(\nu(\boldsymbol{r}, t)\, \overline{v_i v_j} \right)}{\partial r_i} + \nu(\boldsymbol{r}, t) \frac{\partial \Phi}{\partial r_j} &= 0. \end{aligned} \tag{2.28}$$

This expression can be written in a more useful form. For that, we multiply (2.25) by $\overline{v_j}$, and we subtract the result from equation (2.28),

$$
\begin{aligned}
\frac{\partial\left(\nu\,\overline{v_j}\right)}{\partial t}+\frac{\partial\left(\nu\,\overline{v_iv_j}\right)}{\partial r_i}+\nu\frac{\partial\Phi}{\partial r_j}-\overline{v_j}\left[\frac{\partial\nu}{\partial t}+\frac{\partial\left(\nu\,\overline{v_i}\right)}{\partial r_i}\right] &= 0\\
\cancel{\overline{v_j}\frac{\partial\left(\nu\right)}{\partial t}}+\nu\frac{\partial\left(\overline{v_j}\right)}{\partial t}+\frac{\partial\left(\nu\,\overline{v_iv_j}\right)}{\partial r_i}+\nu\frac{\partial\Phi}{\partial r_j}-\cancel{\overline{v_j}\frac{\partial\left(\nu\right)}{\partial t}}-\overline{v_j}\frac{\partial\left(\nu\,\overline{v_i}\right)}{\partial r_i} &= 0\\
\nu\frac{\partial\left(\overline{v_j}\right)}{\partial t}+\frac{\partial\left(\nu\,\sigma^2_{ij}(\boldsymbol{r})\right)}{\partial r_i}+\frac{\partial\left(\nu\,\overline{v_i}\,\overline{v_j}\right)}{\partial r_i}+\nu\frac{\partial\Phi}{\partial r_j}-\overline{v_j}\frac{\partial\left(\nu\,\overline{v_i}\right)}{\partial r_i} &= 0\\
\nu\frac{\partial\left(\overline{v_j}\right)}{\partial t}+\frac{\partial\left(\nu\,\sigma^2_{ij}(\boldsymbol{r})\right)}{\partial r_i}+\left(\nu\overline{v_i}\right)\frac{\partial\left(\overline{v_j}\right)}{\partial r_i}+\cancel{\overline{v_j}\frac{\partial\left(\nu\,\overline{v_i}\right)}{\partial r_i}}+\\
+\nu\frac{\partial\Phi}{\partial r_j}-\cancel{\overline{v_j}\frac{\partial\left(\nu\,\overline{v_i}\right)}{\partial r_i}} &= 0
\end{aligned}
$$

$$
\nu(\boldsymbol{r},t)\frac{\partial\overline{v_j}}{\partial t}+\left(\nu(\boldsymbol{r},t)\,\overline{v_i}\right)\frac{\partial\overline{v_j}}{\partial r_i}=-\,\nu(\boldsymbol{r},t)\frac{\partial\Phi}{\partial r_j}-\frac{\partial\left(\nu(\boldsymbol{r},t)\,\sigma^2_{ij}(\boldsymbol{r})\right)}{\partial r_i}. \tag{2.29}
$$

Above we use the definition of the velocity dispersion tensor, equation (2.22).

The equations (2.25) and (2.29) are powerful equations, since they make the connection between the DF and straightforward observables, like velocity and velocity dispersion. These equations are the known **Jeans Equations**.

Nevertheless, even knowing the potential $\Phi(\boldsymbol{r})$ and the distribution $\nu(\boldsymbol{r},t)$, only the Jeans equations are not enough, since they form a set of four equation[12], but we have nine unknown variables - three velocities components and six independent components of the velocity dispersion tensor. Thus, we need to impose more assumptions about the form of DF. For a detailed description about the possible choices, we recommend Section 4.8, pg. 347 of BT08. Here we will present only the axisymmetric assumption, which will be the framework used for our modelling.

2.2.1 Jeans Equations in an axisymmetric system

First of all, consider the Halmitonian H in cylindrical coordinates (R,ϕ,z), with $R=\sqrt{x^2+y^2}$ and $\phi=\arctan\left(\frac{y}{x}\right)$,

$$
H=\frac{1}{2}\left(p_R^2+\frac{p_\phi^2}{R^2}+p_z^2\right)+\Phi. \tag{2.30}
$$

Now, we can re-write the Boltzmann equation (2.8) in these coordinates (BT08, eq. 4.12)

$$
\frac{\partial f}{\partial t}+p_R\frac{\partial f}{\partial R}+\frac{p_\phi}{R^2}\frac{\partial f}{\partial\phi}+p_z\frac{\partial f}{\partial z}-\left[\frac{\partial\Phi}{\partial R}-\frac{p_\phi^2}{R^3}\right]\frac{\partial f}{\partial p_R}-\frac{\partial\Phi}{\partial\phi}\frac{\partial f}{\partial p_\phi}-\frac{\partial\Phi}{\partial z}\frac{\partial f}{\partial p_z}=0. \tag{2.31}
$$

[12] One from equation (2.25) and three for each velocity component in equation (2.29).

And then we are ready to make our assumptions:

i. The system is steady state, i.e, there are no time dependency.

ii. The system has axial symmetry with respect to z axis. It implies that, $\partial\Phi/\partial\phi = 0$ and $\partial f/\partial\phi = 0$.

iii. DF is a function of the Hamiltonian and angular momentum along the symmetry axis alone, $f = f(H, L_z)$. This constraint implies that crossed terms of velocity dispersion tensor are identically null (BT08, eqs. 4.38).

Under these assumptions, we will show that is possible to obtain a complete set of equations for all the velocities and velocity dispersion of interest.

Starting from the equation (2.31), following the assumptions, it is reduced to

$$p_R\frac{\partial f}{\partial R} + p_z\frac{\partial f}{\partial z} - \left[\frac{\partial\Phi}{\partial R} - \frac{p_\phi^2}{R^3}\right]\frac{\partial f}{\partial p_R} - \frac{\partial\Phi}{\partial z}\frac{\partial f}{\partial p_z} = 0. \tag{2.32}$$

Now, we multiply this equation by p_R, and then integrate over all the momenta

$$\begin{aligned}\int d^3\boldsymbol{p}\, p_R^2\frac{\partial f}{\partial R} + \int d^3\boldsymbol{p}\, p_R p_z\frac{\partial f}{\partial z} \;&-\; \frac{\partial\Phi}{\partial R}\int d^3\boldsymbol{p}\, p_R\frac{\partial f}{\partial p_R} + \\ &+\; \int d^3\boldsymbol{p}\, p_R\frac{p_\phi^2}{R^3}\frac{\partial f}{\partial p_R} - \frac{\partial\Phi}{\partial z}\int d^3\boldsymbol{p}\, p_R\frac{\partial f}{\partial p_z} = 0. \end{aligned}\tag{2.33}$$

But, we can express the momenta in terms of the velocities[13]: $p_R = v_R$, $p_\phi = Rv_\phi$, $p_z = v_z$. With this is mind, now we check each of the terms above.

For the first term:

$$\begin{aligned}\int d^3\boldsymbol{p}\, p_R^2\frac{\partial f}{\partial R} &= \frac{\partial}{\partial R}\int dv_R dv_\phi dv_z\, R\, v_R^2\, f \\ &= \frac{\partial}{\partial R}\left(R\underbrace{\int d^3\boldsymbol{v}\; v_R^2\, f}_{\nu\,\overline{v_R^2}}\right) \\ &= \frac{\partial}{\partial R}\left(R\nu\,\overline{v_R^2}\right).\end{aligned}\tag{2.34}$$

For the second term:

$$\begin{aligned}\int d^3\boldsymbol{p}\, p_R p_z\frac{\partial f}{\partial z} &= \frac{\partial}{\partial z}\int d^3\boldsymbol{v}\, v_R\, v_z\, R\, f \\ &= \frac{\partial}{\partial z}\left(R\nu\overline{v_R\, v_z}\right).\end{aligned}\tag{2.35}$$

[13] The lack of mass in this expressions is due to our choice of use energy per unit of mass instead of energy.

The third term:

$$\int d^3\boldsymbol{p}\, p_R \frac{\partial f}{\partial p_R} \underset{\text{by parts}}{=} \cancelto{0}{p_R f\Big|_{-\infty}^{+\infty}} - \int d^3\boldsymbol{p}\, f \frac{\partial p_R}{\partial p_R} R$$
$$= -\nu R. \tag{2.36}$$

The fourth term:

$$\int d^3\boldsymbol{p}\, p_R \frac{p_\phi^2}{R^3} \frac{\partial f}{\partial p_R} \underset{\text{by parts}}{=} \cancelto{0}{\frac{p_\phi^2\, p_R\, f}{R^3}\Big|_{-\infty}^{+\infty}} - \int d^3\boldsymbol{p}\, \frac{f}{R^3} \frac{\partial (p_\phi^2\, p_R)}{\partial p_R} \tag{2.37}$$

But

$$\frac{\partial (p_\phi^2\, p_R)}{\partial p_R} = \cancelto{0}{\frac{\partial p_\phi^2}{\partial p_R}} p_R + \frac{\partial p_R}{\partial p_R} p_\phi^2 = p_\phi^2, \tag{2.38}$$

so that,

$$\int d^3\boldsymbol{p}\, p_R \frac{p_\phi^2}{R^3} \frac{\partial f}{\partial p_R} = -\int \frac{d^3\boldsymbol{v}\, R^3 v_\phi^2 f}{R^3}$$
$$= -\nu\, \overline{v_\phi^2}. \tag{2.39}$$

The fifth term:

$$\int d^3\boldsymbol{p}\, p_R \frac{\partial f}{\partial p_z} \underset{\text{by parts}}{=} \cancelto{0}{p_R f\Big|_{-\infty}^{+\infty}} - \int d^3\boldsymbol{p}\, f \cancelto{0}{\frac{\partial p_R}{\partial p_z}}$$
$$= 0. \tag{2.40}$$

Merging all the above results in equation (2.33), then dividing it by R, we have:

$$\frac{\partial (\nu \overline{v_R^2})}{\partial R} + \frac{\partial (\nu \overline{v_R\, v_z})}{\partial z} + \nu \left[\frac{\overline{v_R^2} - \overline{v_\phi^2}}{R} + \frac{\partial \Phi}{\partial R} \right] = 0. \tag{2.41}$$

Following the same procedures above, but multiplying (2.31) by p_z or p_ϕ we can find:

$$\frac{1}{R} \frac{\partial (\nu \overline{v_R\, v_z})}{\partial R} + \frac{\partial (\nu \overline{v_z^2})}{\partial z} + \nu \frac{\partial \Phi}{\partial z} = 0. \tag{2.42}$$

$$\frac{1}{R^2} \frac{\partial (R^2\, \nu \overline{v_R\, v_\phi})}{\partial R} + \frac{\partial (\nu \overline{v_z\, v_\phi})}{\partial z} = 0. \tag{2.43}$$

The last three equations (2.41), (2.42), and (2.43) are quite general, since we do not apply our condition (iii). Applying it, equation (2.43) turns trivially null, while equation (2.42) results in a simple first order differential equation for $\overline{v_z^2}$,

$$\overline{v_z^2} = \frac{1}{\nu(R,z)} \int_z^\infty dz'\, \nu(R,z') \frac{\partial \Phi(R,z')}{dz'}. \tag{2.44}$$

And then, if we know $\overline{v_R^2}$, perhaps under certain assumptions about the anisotropy of the orbits, as we will do in Chapter 5, we can compute $\overline{v_\phi^2}$ using equation (2.41),

$$\overline{v_\phi^2} = \overline{v_R^2} + \frac{R}{\nu} \frac{\partial(\nu(R,z)\,\overline{v_R^2})}{\partial R} + R \frac{\partial \Phi(R,z)}{dR}. \tag{2.45}$$

To summarize this Section, let us point out the most important results:

- Equation (2.8) is the collisionless Boltzmann equation, that describes the probability density function (DF) to find a star in a given position $\boldsymbol{r}$ and some velocity $\boldsymbol{v}$, at time t, in the observed galaxy. From Boltzmann equation, we derive some connections between the DF and observables, like the luminosity density (2.14), the mean velocity (2.20), and the velocity dispersion (2.21).

- Finally, to be able to solve the Boltzmann equation without necessarily knowing the DF, we derive the four Jeans equations (2.25) and (2.29), and under some assumptions of axisymmetry of the system, we find more two equations, (2.44) and (2.45), that define a complete set of equations (with the Jeans ones) that make possible to find all the velocities $(\overline{v_R}, \overline{v_\phi}, \overline{v_z})$, and its dispersions $(\sigma_{RR}^2, \sigma_{\phi\phi}^2, \sigma_{zz}^2)$, through equations (2.22).

3 Gravitational Lens Theory

Throughout this Chapter we will focus our attention to the gravitational lens theory, i.e., the change of the light paths in the presence of massive bodies. This phenomenon can be addressed in two different ways, related to its intensity. One of them is called strong regime (or strong lensing), which is characterized by large distortions of the source object, large (de-)magnifications, and the possible formation of multiple images. This effect usually is seen as arcs around galaxies and galaxy clusters, that act as deflectors. The other one is called weak regime (or weak lensing), which also causes distortions and magnifications, but with significantly less intensity. In this case, statistical analysis over a large sample of deflected images are needed, once its effects are weaker. Fig. 5 shows both effects caused by the galaxy cluster MACS J0416. We can see both the gravitational arcs around the center of the cluster, and small magnifications and distortions of the background galaxies across the entire field.

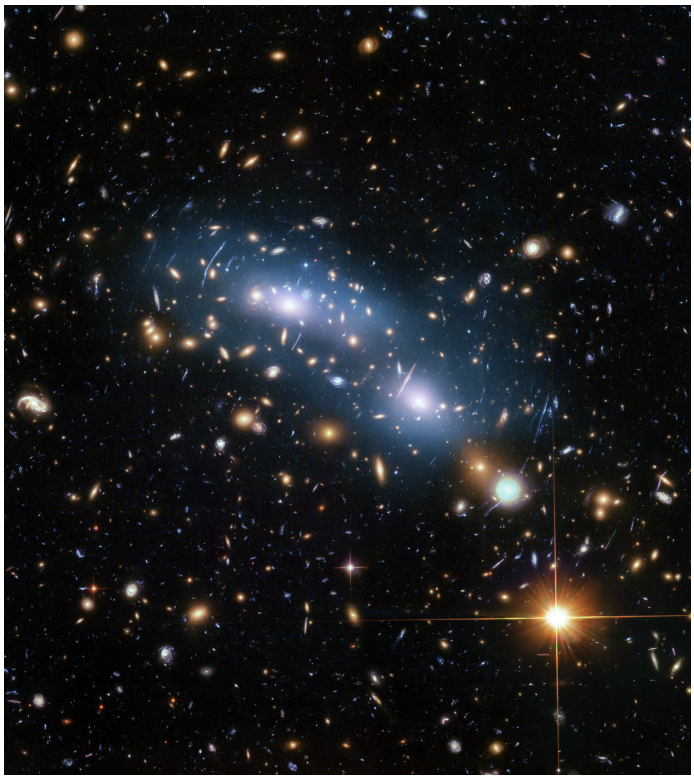


Figure 5 – Galaxy cluster MACS J0416. Image from Hubble Frontier Fields program. Image taken from: NASA, ESA and M. Montes (University of New South Wales)

For the purposes of this dissertation, here we briefly present the formalism for the strong lens regime. More complete derivations can be found in Meneghetti (2016), Schneider; Ehlers; Falco (1992), and FERREIRA (2008).

3.1 Lens Equation

According to the description of gravity by GR, when a light ray passes near a massive body, its path can be deflected by the curvature of space-time. When this curvature is strong enough, this deflection could be in such way that the emitter will be seen in a different position in the sky. This effect is known as the gravitational lensing.

Looking at Fig. 6, and labeling: source as the emitter body, lens the massive object responsible for the strong gravitational field that curves the space-time and the observer the one who receives the bent light. In this figure, D_{LS}, D_L and D_S are the angular diameter distances between the lens and source, lens and observer, and source and observer, respectively. Moreover, it is important to note that these distances depend on the cosmology, so, in general, $D_{LS} \neq D_L + D_S$.

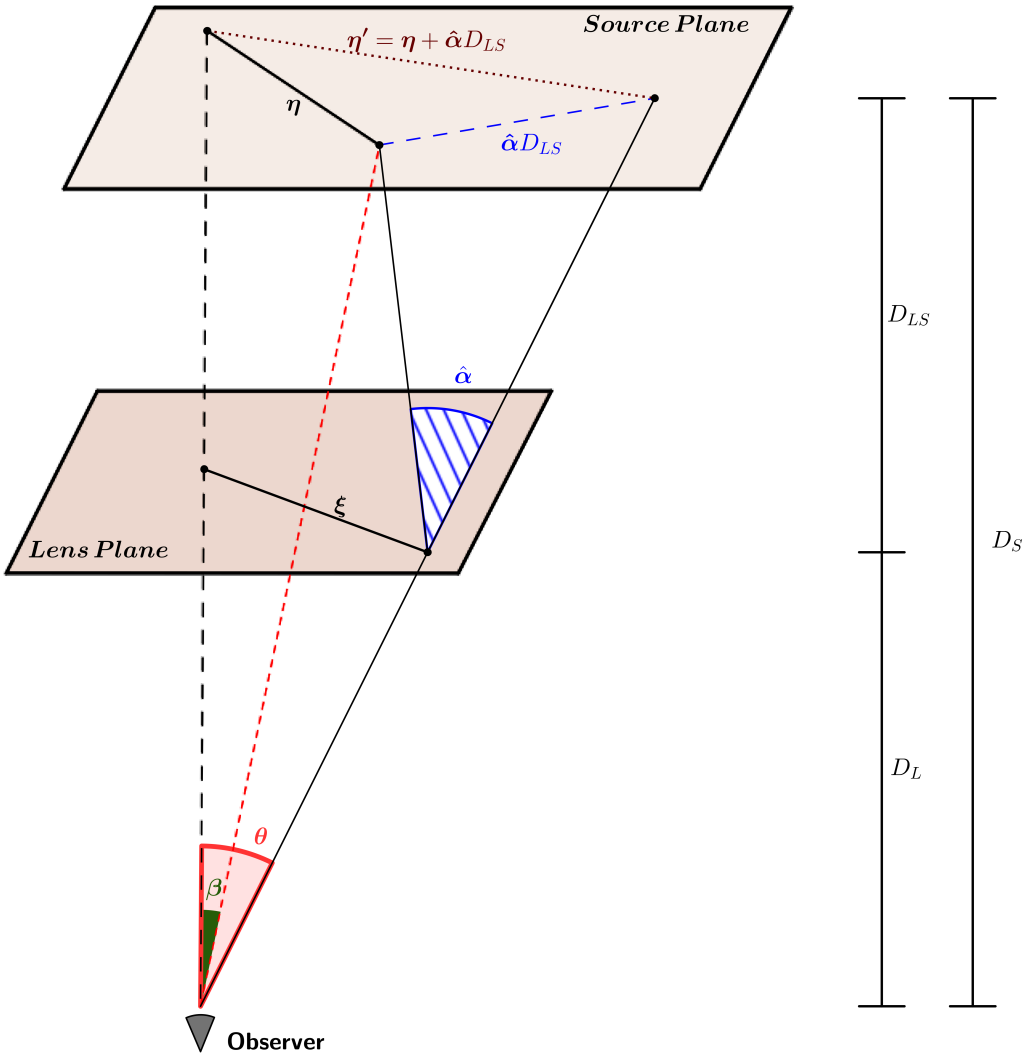


Figure 6 – Illustration of lens configuration scheme. The dimensions in this cartoon are extremely exaggerated to help visualisation.

Furthermore, in the majority of cases (except for the large scale structure), the lens dimensions are much smaller than distances involved in the lensing phenomenon. When

this is true, we can make use of the **thin lens approximation**. In this approximation, we assume that the light ray bends in only one point in its path, i.e, only in the plane that contains the lens object, the lens plane.

The lens plane is defined as a plane that crosses the lens center, and it is perpendicular to the line-of-sight of the observer, the optical axis (bolded dotted line in Fig. 6). In the same way, the source plane is the plane perpendicular to the line-of-sight that crosses the center of the source object and contains it.

From Fig. 6, let the true two-dimensional position of the source be labeled by $\boldsymbol{\eta}$ in the source plane, while the true two-dimensional angular position of the source be $\boldsymbol{\beta}$. Moreover, $\boldsymbol{\theta}$ is the two-dimensional angular apparent position of light, and $\tilde{\boldsymbol{\alpha}}$ is the deflected angle of the light ray coming from the source on the lens plane. For small angles ($\omega << 1$ rad) the Taylor approximation

$$\tan(\omega) = \omega - \frac{\omega^3}{3!} + \mathcal{O}^5, \tag{3.1}$$

is valid. Therefore, in first approximation, we can rewrite the angular position of the source as $\boldsymbol{\eta} = \boldsymbol{\beta} D_S$. Also, we define the impact parameter $\boldsymbol{\xi}$ as the position of the deflected light in the lens plane, and we can write it as $\boldsymbol{\xi} = \boldsymbol{\theta} D_L$.

Now, we can write the apparent source position (in the source plane) as the real position in the source plane plus the additional position seen by the observer due to the deflection caused by the massive body. Formally, this can be written as

$$\boldsymbol{\eta'} = \boldsymbol{\eta} + \tilde{\boldsymbol{\alpha}}(\boldsymbol{\xi}) D_{LS}, \tag{3.2}$$

where $\boldsymbol{\eta'} = \boldsymbol{\theta} D_S$ is the apparent observed position of the source. Note that, the deflection angle $\tilde{\boldsymbol{\alpha}}$ is a function of the impact parameter, once it depends on the position of the deflected light ray in the image plane, also called image position.

Finally, replacing $\boldsymbol{\eta'}$ and $\boldsymbol{\eta}$ in (3.2), and rearranging the terms, we obtain

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{LS}}{D_S} \tilde{\boldsymbol{\alpha}}(D_L \boldsymbol{\theta}). \tag{3.3}$$

If we define the reduced deflection angle as

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) \equiv \frac{D_{LS}}{D_S} \tilde{\boldsymbol{\alpha}}(D_L \boldsymbol{\theta}), \tag{3.4}$$

we can rewrite the equation (3.3) in a simplified way:

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}). \tag{3.5}$$

This equation is called the **lens equation**, and its interpretation is quiet simple: the real source position is equal to the observed image position minus the reduced deflection angle

seen by the observer. If there is more than one solution for the lens equation above, the source will be observed in different positions $\boldsymbol{\theta}$, that is, multiple images will be detected.

3.2 Lens Potential

From the thin lens approximation, we consider that light rays are curved only in the lens plane. Therefore, it is reasonable to consider instead of the three-dimensional mass density along the line-of-sight $\rho(\boldsymbol{r})$, the projected mass density along the line-of-sight, denoted by $\Sigma(\boldsymbol{\xi})$, where $\boldsymbol{\xi}$ is a two-dimensional vector in the plane of sky or the impact parameter in our case.

If we define our Cartesian coordinate system (x, y, z) with z-axis along the line-of-sight, the projected mass density is easily calculated by

$$\Sigma(\boldsymbol{\xi}) = \int_0^{D_S} \rho(\boldsymbol{r}) dz. \tag{3.6}$$

In this definition, we are considering only the mass between the observer and the source. However, as the density quickly drops to zero in the interval $(0, D_S)$, is common to replace the integration limits to $(-\infty, +\infty)$.

Now, consider the equation (1.5) for the deflection angle of a point mass,

$$\tilde{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4GM}{c^2} \frac{\hat{\boldsymbol{\xi}}}{\xi},$$

where $\hat{\boldsymbol{\xi}}$ is a versor in the lens plane which gives the direction of the impact parameter, and ξ the distance of the bent light ray to the optical axis.

In the following, consider Fig. 7, where $\boldsymbol{\xi'}$ is a two-dimensional vector in the lens plane which gives the position of a mass element dM with respect to the lens center O, and $\boldsymbol{\xi}$ is the impact parameter.

In this case the deflection, $d\tilde{\boldsymbol{\alpha}}(\boldsymbol{\xi})$, caused by an infinitesimal mass element dM can be expressed as:

$$d\tilde{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2} \frac{(\boldsymbol{\xi} - \boldsymbol{\xi'})}{|\boldsymbol{\xi} - \boldsymbol{\xi'}|^2} dM, \tag{3.7}$$

or in terms of the projected density (3.6),

$$d\tilde{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2} \frac{(\boldsymbol{\xi} - \boldsymbol{\xi'})}{|\boldsymbol{\xi} - \boldsymbol{\xi'}|^2} \Sigma(\boldsymbol{\xi'}) d^2\boldsymbol{\xi'}. \tag{3.8}$$

As a result, the total deflection angle is just the integral over (3.8),

$$\tilde{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2} \int \frac{(\boldsymbol{\xi} - \boldsymbol{\xi'})}{|\boldsymbol{\xi} - \boldsymbol{\xi'}|^2} \Sigma(\boldsymbol{\xi'}) d^2\boldsymbol{\xi'}, \tag{3.9}$$

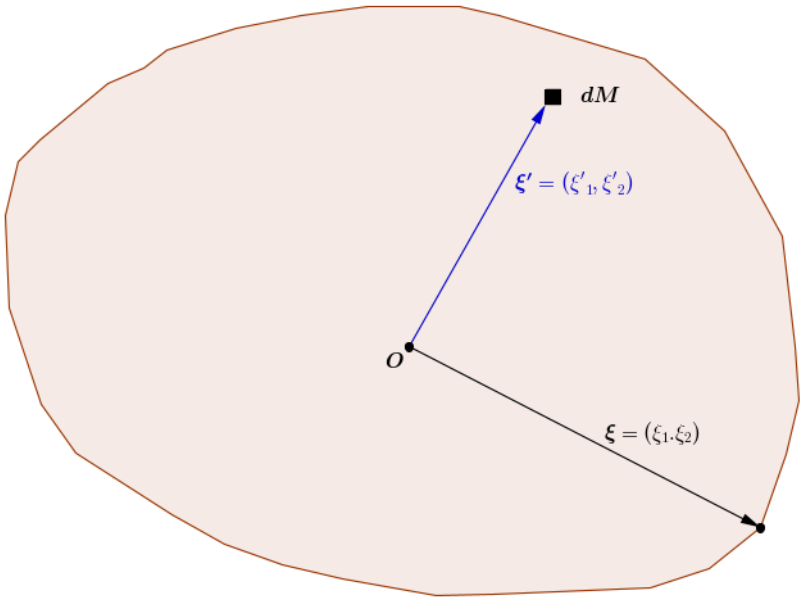


Figure 7 – Arbitrary mass distribution for the lens object.

valid for any mass distribution. Sometimes it is useful to rewrite the last equation in terms of $\boldsymbol{\theta}$. Using the relation $\boldsymbol{\xi} = \boldsymbol{\theta} D_L$, we have

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{D_L D_{LS}}{D_S} \frac{4G}{c^2} \int \frac{(\boldsymbol{\theta} - \boldsymbol{\theta'})}{|\boldsymbol{\theta} - \boldsymbol{\theta'}|^2} \Sigma(D_L \boldsymbol{\theta'}) d^2\boldsymbol{\theta'}, \tag{3.10}$$

where we use the expression (3.4) for the reduced deflection angle on the left side. Defining

$$\Sigma_c \equiv \frac{c^2}{4\pi G} \frac{D_S}{D_L D_{LS}}, \tag{3.11}$$

as a **critical density**, we rewrite $\boldsymbol{\alpha}(\boldsymbol{\theta})$ as

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi \Sigma_c} \int \frac{(\boldsymbol{\theta} - \boldsymbol{\theta'})}{|\boldsymbol{\theta} - \boldsymbol{\theta'}|^2} \Sigma(D_L \boldsymbol{\theta'}) d^2\boldsymbol{\theta'}. \tag{3.12}$$

However, equation (3.12) can be rewritten in a more elegant way if we define a new dimensionless quantity, the **convergence**

$$\kappa(\boldsymbol{\theta'}) \equiv \frac{\Sigma(D_L \boldsymbol{\theta'})}{\Sigma_c}, \tag{3.13}$$

which is the projected mass per unit of critical density.

The convergence is the important quantity to distinguish between the gravitational lensing regimes. A mass distribution with $\kappa \geq 1$ ($\Sigma \geq \Sigma_c$) will act as a strong lens, producing multiple images. On the other hand, when $\kappa \ll 1$ ($\Sigma \ll \Sigma_c$), we enter in the regime of weak lensing.

Now, assuming that the coordinates $\boldsymbol{\theta}$ and $\boldsymbol{\theta'}$ are independent, and denoting the gradient operator with relation to $\boldsymbol{\theta}$ by $\boldsymbol{\nabla}_{\boldsymbol{\theta}}$, we can write the following identity[1]

$$\boldsymbol{\nabla}_{\boldsymbol{\theta}} \ln(|\boldsymbol{\theta} - \boldsymbol{\theta'}|) = \frac{(\boldsymbol{\theta} - \boldsymbol{\theta'})}{|\boldsymbol{\theta} - \boldsymbol{\theta'}|^2}. \tag{3.14}$$

So the reduced deflection angle (3.12) can be written as

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \boldsymbol{\nabla}_{\boldsymbol{\theta}} \left[\frac{1}{\pi} \int \kappa(\boldsymbol{\theta'}) \ln(|\boldsymbol{\theta} - \boldsymbol{\theta'}|) d^2\boldsymbol{\theta'} \right], \tag{3.15}$$

where the exchange between the integral and the gradient is only possible because $\boldsymbol{\theta}$ and $\boldsymbol{\theta'}$ are independent. Finally, defining the **lens potential** function as

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi} \int \kappa(\boldsymbol{\theta'}) \ln(|\boldsymbol{\theta} - \boldsymbol{\theta'}|) d^2\boldsymbol{\theta'}, \tag{3.16}$$

we obtain a more compact and elegant form to the reduced deflection angle

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \boldsymbol{\nabla}_{\boldsymbol{\theta}} \psi(\boldsymbol{\theta}), \tag{3.17}$$

as the gradient of the lens potential $\psi(\boldsymbol{\theta})$.

However, the two-dimensional Dirac distribution function can be written as

$$\delta(\boldsymbol{\theta} - \boldsymbol{\theta'}) = \frac{1}{2\pi} \boldsymbol{\nabla}_{\boldsymbol{\theta}} \left[\frac{\boldsymbol{\theta} - \boldsymbol{\theta'}}{|\boldsymbol{\theta} - \boldsymbol{\theta'}|^2} \right], \tag{3.18}$$

then, applying the Laplacian operator[2] to the lens potential (3.16) we have

$$\begin{aligned}
\nabla^2_{\boldsymbol{\theta}} \psi(\boldsymbol{\theta}) &= \boldsymbol{\nabla}_{\boldsymbol{\theta}} \cdot \boldsymbol{\nabla}_{\boldsymbol{\theta}} \left[\frac{1}{\pi} \int \kappa(\boldsymbol{\theta'}) \ln(|\boldsymbol{\theta} - \boldsymbol{\theta'}|) d^2\boldsymbol{\theta'} \right] \\
&= \boldsymbol{\nabla}_{\boldsymbol{\theta}} \left[\frac{1}{\pi} \int \kappa(\boldsymbol{\theta'}) \frac{\boldsymbol{\theta} - \boldsymbol{\theta'}}{|\boldsymbol{\theta} - \boldsymbol{\theta'}|^2} d^2\boldsymbol{\theta'} \right] \\
&= \int \kappa(\boldsymbol{\theta'}) d^2\boldsymbol{\theta'} 2 \frac{1}{2\pi} \boldsymbol{\nabla}_{\boldsymbol{\theta}} \left[\frac{\boldsymbol{\theta} - \boldsymbol{\theta'}}{|\boldsymbol{\theta} - \boldsymbol{\theta'}|^2} \right] \\
&= 2 \int \kappa(\boldsymbol{\theta'}) \delta(\boldsymbol{\theta} - \boldsymbol{\theta'}) d^2\boldsymbol{\theta'} \\
&= 2\kappa(\boldsymbol{\theta}),
\end{aligned} \tag{3.19}$$

where we use the filtering property of the Dirac distribution. Also note that the expression (3.19) is just a Poisson equation for the lens potential.

[1] Actually this identity is valid for any non-null vector $\bar{\boldsymbol{r}} = \boldsymbol{r} - \boldsymbol{r'}$ in a Euclidean space.
[2] $\nabla^2_{\boldsymbol{\theta}} \equiv \boldsymbol{\nabla}_{\boldsymbol{\theta}} \cdot \boldsymbol{\nabla}_{\boldsymbol{\theta}}$

3.3 Mapping

Once we know how the light rays from the source are deflected by a lens object, (3.12) or (3.17), we need to know how the source coordinates $\boldsymbol{\beta} = (\beta_1, \beta_2)$ are mapped into the lens plane $\boldsymbol{\theta} = (\theta_1, \theta_2)$. This mapping is the responsible for generating the distorted and multiple images of the source in the lens plane, the most important effect of the gravitational lensing phenomenon.

The position and the shape of the images formed after the deflection can be determined solving the lens equation (3.5). Moreover, strictly speaking, the mapping is a smooth, local and continuum injector application. However, in the lens theory this mapping can be multi-valued, i.e, a point in the source plane can be mapped in more than one position in the lens plane.

3.3.1 Jacobian Matrix and transformations between planes

To better understand the mapping of source into the lens plane (or opposite mapping), consider two coordinate systems (x, y) and (u, v) and a set of smooth and continuous functions ϕ_i at any point, such that:

$$\begin{cases} x = \phi_1(u, v) \\ y = \phi_2(u, v) \end{cases} \tag{3.20}$$

Since the functions ϕ_i are smooth, they must be at least up to the first order differentiable. So, the differentials in respect to (x, y) coordinates are

$$\begin{cases} dx = \left(\frac{\partial \phi_1}{\partial u}\right) du + \left(\frac{\partial \phi_1}{\partial v}\right) dv \\ dy = \left(\frac{\partial \phi_2}{\partial u}\right) du + \left(\frac{\partial \phi_2}{\partial v}\right) dv \end{cases} \tag{3.21}$$

Or, matricially

$$\begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} \frac{\partial \phi_1}{\partial u} & \frac{\partial \phi_1}{\partial v} \\ \frac{\partial \phi_2}{\partial u} & \frac{\partial \phi_2}{\partial v} \end{bmatrix} \begin{bmatrix} du \\ dv \end{bmatrix}. \tag{3.22}$$

Written in this way, it is clear that the equation (3.22) represents a change of coordinate systems[3]. More precisely, it represents the mapping between coordinates (u, v) and (x, y). The matrix that makes this transformation between the coordinate systems is called Jacobian matrix, and it is denoted by $\boldsymbol{J}$. Formally, the determinant of $\boldsymbol{J}$ represents the deformation of a coordinate system into another (Malajovich, 2010, Theorem 10.11 or Chapter 2.10 of Carroll, 2003),

$$dxdy = \det(\boldsymbol{J}) dudv, \tag{3.23}$$

[3] A linear transformation.

where we define

$$\boldsymbol{J} \equiv \frac{\partial(x,y)}{\partial(u,v)} = \begin{bmatrix} \frac{\partial\phi_1}{\partial u} & \frac{\partial\phi_1}{\partial v} \\ \frac{\partial\phi_2}{\partial u} & \frac{\partial\phi_2}{\partial v} \end{bmatrix}, \tag{3.24}$$

and $\det(\boldsymbol{J})$ is the determinant of $\boldsymbol{J}$. For a more complete and rigorous discussion, we strongly recommend Nakahara (2003) and O'Neill (1983).

Finally, before continuing to the transformations between the source plane and image plane, we would like to highlight some aspects of these Jacobian transformations. First, we require that the connection functions ϕ_i (the coordinated functions), should be smooth and differentiable at any point, because we want to avoid discontinuities. Second, such transformations are local, and not global, transformations, this is why we desire coordinated functions differentiable at first order, at least. Third, as the functions ϕ_i are smooth, they have a well-defined inverse ϕ_i^{-1}. So, we can change from the (x,y) coordinate system to (u,v) through $\phi_i^{-1}(x,y)$, and the correspondent matrix is just the inverse Jacobian matrix $\boldsymbol{J^{-1}}$. Lastly, this can be extended to higher dimensions without loss of generality (check references cited above).

Now we will determine the matrix that performs the transformation between the coordinates in the lens plane to the coordinates in the source plane, i.e., the matrix transformation or the Jacobian transformation. In the lensing formalism, this matrix is denoted by $\boldsymbol{\mathcal{A}}$. So, looking at the equation (3.5), we can write

$$\begin{cases} \beta_1 = \theta_1 - \alpha_1 \\ \beta_2 = \theta_2 - \alpha_2 \end{cases} \quad \text{and} \quad \boldsymbol{\mathcal{A}} = \frac{\boldsymbol{\partial\beta}}{\boldsymbol{\partial\theta}} = \frac{\partial(\beta_1,\beta_2)}{\partial(\theta_1,\theta_2)}. \tag{3.25}$$

Using the definition (3.24), we find

$$\boldsymbol{\mathcal{A}} = \begin{bmatrix} 1 - \frac{\partial\alpha_1}{\partial\theta_1} & -\frac{\partial\alpha_1}{\partial\theta_2} \\ -\frac{\partial\alpha_2}{\partial\theta_1} & 1 - \frac{\partial\alpha_2}{\partial\theta_2} \end{bmatrix}, \tag{3.26}$$

or,

$$\boldsymbol{\mathcal{A}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} \frac{\partial\alpha_1}{\partial\theta_1} & \frac{\partial\alpha_1}{\partial\theta_2} \\ \frac{\partial\alpha_2}{\partial\theta_1} & \frac{\partial\alpha_2}{\partial\theta_2} \end{bmatrix}. \tag{3.27}$$

In terms of matrix elements, the last expression can be placed in the following form

$$\mathcal{A}_{ij} = \delta_{ij} - \frac{\partial\alpha_i}{\partial\theta_j}. \tag{3.28}$$

However, by (3.17) $\alpha_i = \partial\psi/\partial\theta_i$, therefore

$$\mathcal{A}_{ij} = \delta_{ij} - \frac{\partial^2\psi}{\partial\theta_i\partial\theta_j}, \tag{3.29}$$

where the partial derivatives can be commuted, if we assume θ_1 and θ_2 independent, and $\psi(\boldsymbol{\theta})$ continuous and differentiable up to second order.

Thus, if we know the deflection angle equation (3.12) or the lens potential (3.16), we can easily find how each point of the image plane is mapped into the source plane using equations (3.28) or (3.29), respectively. This mapping, of course, is strongly dependent on the lens mass distribution, as we have discussed. If the Jacobian matrix (3.24) has a non-zero determinant, it can be inverted, allowing us to write, in first approximation (Bartelmann; Maturi, 2017),

$$\delta\boldsymbol{\theta} = \boldsymbol{\mathcal{A}}^{-1}\delta\boldsymbol{\beta} = \boldsymbol{\mathcal{M}}\delta\boldsymbol{\beta}, \tag{3.30}$$

where we define the **magnification tensor** or **magnification matrix**, $\boldsymbol{\mathcal{M}} \equiv \boldsymbol{\mathcal{A}}^{-1}$. In other words, the above expression tells us how an infinitesimal vector on the source plane ($\delta\boldsymbol{\beta}$) is mapped into the lens plane ($\delta\boldsymbol{\theta}$). The name of $\boldsymbol{\mathcal{M}}$ will be clear soon.

3.3.2 Distortion and Magnification

As we saw above, due to the lensing phenomenon the image of a source can appear distorted or even multiplied in different positions to the observer. All this information is related to the Jacobian matrix $\boldsymbol{\mathcal{A}}$, and its inverse $\boldsymbol{\mathcal{M}}$ (Schneider; Ehlers; Falco, 1992).

However, the **magnification** effect, another important feature of gravitational lensing is noteworthy. Defined as the modulus of the ratio of the flux image (in some frequency) per unit of flux of source (in same frequency),

$$\mu \equiv \left| \frac{F_\nu^{(image)}}{F_\nu^{(source)}} \right|, \tag{3.31}$$

the magnification can be better understood as follows: since the deflection of light changes the path of the light beam without changing the number of photons emitted by the source or their frequency, i.e., it is an achromatic and conservative effect, the specific intensity[4] I_ν, also called spectral brightness, remains the same for the source and the image. Once the flux in a given frequency is the integral of specific intensity over the solid angle[5],

$$F_\nu = \int I_\nu d\Omega, \tag{3.32}$$

and considering I_ν approximately constant for a small source in the sky, we can write the flux as

$$F_\nu = I_\nu \Delta\Omega. \tag{3.33}$$

[4] Energy per unity of solid angle per unity of time per unity of area per unity of frequency.
[5] Here we are using Ω to denote the solid angle in the sky plane.

Nevertheless, as $I_\nu^{(image)} = I_\nu^{(source)}$, we have

$$\mu = \left| \frac{\Delta\Omega^{(image)}}{\Delta\Omega^{(source)}} \right|, \tag{3.34}$$

where $\Delta\Omega$ is the solid angle subtended by the object. It means that, if the solid angle which the source is observed changes, the measured flux for this object is magnified (or demagnified) by a factor μ. The magnification can also be interpreted as how much the area (or brightness) of the image is greater/smaller than that of the source. Keeping this in mind, and recalling the definition of solid angle, i.e,

$$d\Omega = \frac{dA}{r^2}, \tag{3.35}$$

where r is the angular diameter distance to the object, and A the area subtended by it in the sky, the equation (3.34) can be placed, for an infinitesimal solid angle, in the form

$$\mu = \left| \frac{dA^{(image)}}{dA^{(source)}} \frac{D_S^2}{D_L^2} \right| = \left| \frac{d\xi_1 d\xi_2}{d\eta_1 d\eta_2} \frac{D_S^2}{D_L^2} \right|. \tag{3.36}$$

Here $\boldsymbol{\xi} = (\xi_1, \xi_2)$ are coordinates in the image plane, and $\boldsymbol{\eta} = (\eta_1, \eta_2)$ coordinates in the source plane. Now, using the definitions from the beginning of the this Chapter, $\boldsymbol{\xi} = \boldsymbol{\theta} D_L$ and $\boldsymbol{\eta} = \boldsymbol{\beta} D_S^{-1}$, the last equation assumes the form

$$\mu = \left| \frac{d\theta_1 d\theta_2}{d\beta_1 d\beta_2} \right|, \tag{3.37}$$

which is, as seen before, the determinant of the magnification tensor $\boldsymbol{\mathcal{M}}$, thus explaining its name too. In addition, the physical interpretation of the Jacobian matrix determinant is clear, since it represents precisely the deformation of one coordinate system into another. Finally, using the inverse matrix determinant rule, the magnification is expressed as

$$\mu = |\det(\boldsymbol{\mathcal{M}})| = \left|\det(\boldsymbol{\mathcal{A}}^{-1})\right| = |\det(\boldsymbol{\mathcal{A}})|^{-1}. \tag{3.38}$$

Mathematically speaking, we avoid the possibility of the determinant of Jacobian matrix vanish, because we prefer continuous transformations. But, if we relax this requirement, there may be points where $\det(\boldsymbol{\mathcal{A}}) = 0$. When this occurs, the magnification diverge and formally tends to infinity[6]. These positions define a curve in the source plane called **caustics**, which are not necessarily smooth and can present cusps. On the other hand, **critical curves** are obtained by mapping the caustics to the lens plane through the lens equation.

As a last comment in this Chapter, we need to notice that, sometimes we need to include an additional contribution to the deflection angle to take in account some

[6] Although physically this will never happen, since infinite magnifications do not occur in nature.

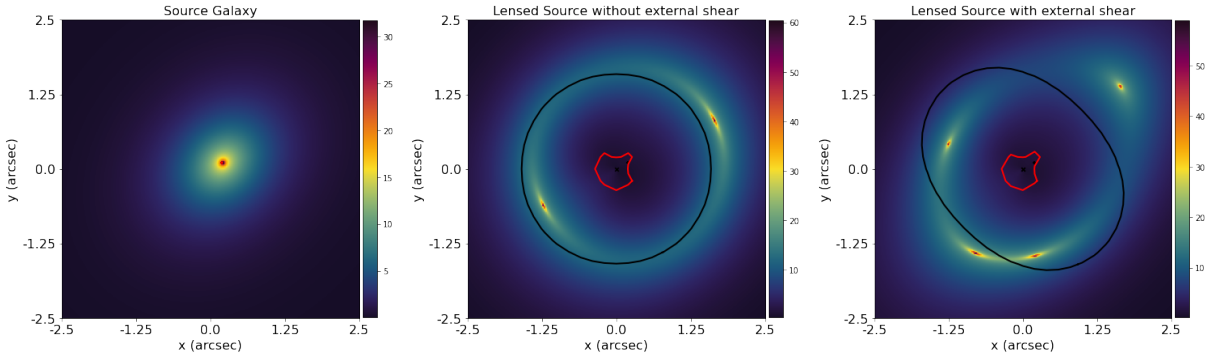


Figure 8 – Example of an *exaggerated* external shear on the lensed source. Left panel shows the source galaxy, the central panel the lensed source by a spherical isothermal lens model and no shear, and the right panel the source galaxy lensed by the same spherical isothermal lens model with an external shear with magnitude equal to 0.15 and angle equal to 45° counterclockwise from $x-$axis. The black and red lines are, respectively, the critical curve and caustic. The cross marks the lens center, while the black dot is the source center in the source plane. The scale bars are in arbitrary units.

contribution of neighboring masses (such as other galaxies or clusters near the lens galaxy), or even any sub-structures along the line-of-sight (Keeton; Kochanek; Seljak, 1997). This contribution is quantified by an external shear component, described by its magnitude and angle. The angle represents the preferred direction of stretching of the image due to external shear, while the magnitude is the strength of this stretching. Fig. 8 shows an example comparing the distortion of a lensing source in the presence of an *exaggerated* external shear and without an external shear. The black and red lines are, respectively, the critical curve and caustic.

For more rigorous details about the lensing theory, see also Bartelmann (2010).

4 Data

In this Chapter we present the data used for our combined analysis of lensing and dynamics in order to constraint $\gamma_{\rm PPN}$, that will be described in Chapters 5 and 6.

For the combined analysis intended, we need at least two kinds of data: integral field spectroscopy, for dynamical modelling, and photometric data, for lens and dynamical modelling, since we use the luminosity distribution as a tracer of the stellar content (see Chapter 5). The H-ATLAS J090311.6+003906 (SDP.81) system has both, with an additional: this system also has high-resolution interferometric data obtained using a very extended configuration, which leads to high-resolution angular imaging (in our case), perfect for gravitational lens modelling.

SDP.81 system was first detected by Negrello et al. (2010), as a part of the Herschel Astrophysical Terahertz Large Area Survey (H-ATLAS), as a candidate of submillimeter source galaxy being lensed by an elliptical foreground galaxy in redshift $z_l = 0.299$. In this work, Negrello et al. also confirm the nature of the source galaxy in redshift $z = 3.042$ using measurements of CO emission lines obtained with the Green Bank Telescope.

This system has already been studied by several groups (Negrello et al., 2014; Dye et al., 2014; Dye et al., 2015; Tamura et al., 2015; Rybak et al., 2015; Wong; Suyu; Matsushita, 2015), however never with the goal proposed here in our project. Furthermore, most of the cited works focus on the the reconstruction of the source galaxy and on its properties[1], not on the lens. Our methodology, on the other hand, allows us to infer physical properties of the lens galaxy, as well as testing GR.

Our dataset is composed of: photometric data from Hubble Space Telescope (HST)[2], integral field spectroscopy data from the Multi Unit Spectroscopic Explorer (MUSE)[3], and interferometric data from Atacama Large Millimeter/Submillimeter Array (ALMA)[4]. In the following we describe these datasets.

[1] This is due to the unprecedented tens-of-parsec resolution of the submillimeter data available for this system.

[2] <https://www.nasa.gov/mission_pages/hubble/main/index.html>

[3] <https://www.eso.org/sci/facilities/develop/instruments/muse.html>

[4] <https://www.almaobservatory.org/en/home/>

4.1 HST data

The HST data were obtained with Wide Field Camera 3 (WFC3)[5] in 2011 in two different bands: F160W[6] with 4417.62s of total integration time, and F110W[7] with 711.74s of total integration time. Both images are publicly available in the Hubble Legacy Archive[8] (PropID: 12194, PI: Negrello). The data were reduced using the WFC3 standard pipeline by the Barbara A. Mikulski Archive for Space Telescopes (MAST) team, and drizzled following Gonzaga et al. (2012).

We choose to work with the deepest image F160W, which has a pixel scale of 0.09″. The point spread function (PSF) was obtained fitting a collection of unsaturated stars in the field near the galaxy using the IRAF task DAOPHOT (Stetson, 1987). These stars are combined to generate a effective PSF (ePSF; Anderson; King, 2000), i.e., an empirical model describing how the light of the stars is spread in a particular pixel. The ePSF for the HST data has an FWHM[9] approximately equal to 0.108″, which corresponds to $\sim$ 495pc in the redshift of the lens galaxy. This ePSF will be later parametrized by a sum of 2D-Gaussians (see Section 5.2), in order to be included in the modelling.

A cutout around the lens galaxy can be seen in Fig. 9. In this figure the gravitational arcs are not clearly seen due to the emission of the foreground galaxy. In the top left side we see the emission of a star.

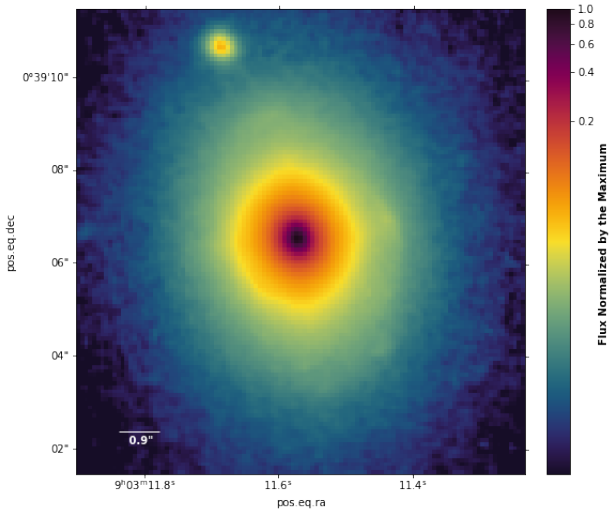


Figure 9 – SDP.81 lens galaxy. Image from HST/WCF3-F160W. Axis are RA and DEC. The gravitational arcs are not clearly seen in this image, because they are embedded in the lens galaxy light. North is up and East is left.

In order reveal the emission of the background source in the Fig. 9, we need to

[5] <https://www.stsci.edu/hst/instrumentation/wfc3>
[6] <http://svo2.cab.inta-csic.es/svo/theory/fps3/index.php?id=HST/WFC3_IR.F160W>
[7] <http://svo2.cab.inta-csic.es/svo/theory/fps3/index.php?id=HST/WFC3_IR.F110W>
[8] <https://hla.stsci.edu/>
[9] Full Width at Half Maximum.

disentangle their emission from the emission of the lens galaxy, and a detailed description of this procedure is done in Chapters 5 and 6. After modelling the lens galaxy light profile, and subtracting it from the original image, we reveal the arcs of the SDP.81 system, as seen in Fig. 10. In this figure we see the rest-frame optical emission of the background source.

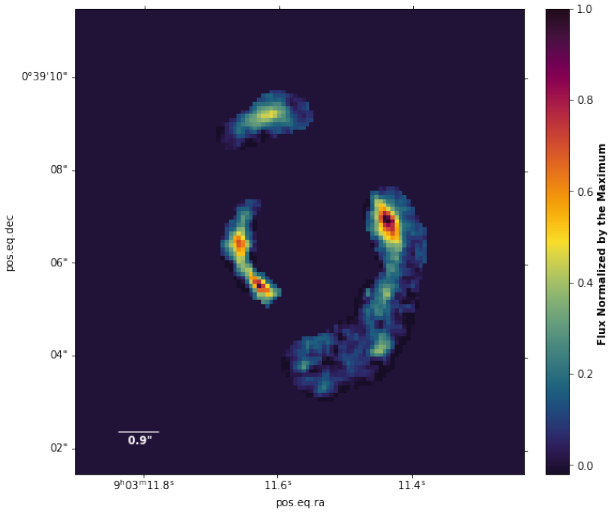


Figure 10 – SDP.81 gravitational arcs present in the HST/WFC3-F160W, after lens light subtraction. North is up and East is left.

4.2 MUSE data

MUSE is a panoramic integral-field spectrograph on-board of the Very Large Telescope (VLT) of the European Southern Observatory (ESO). As pointed out previously (Section 1.3), IFUs are powerful instruments for astronomy, since they are able to extract spectra in each pixel (named spaxel in IFU context) of an image, producing a final data cube with two spatial and one spectral coordinates. Between the most prominents IFUs today, MUSE stands out, due to the high quality of the data produced.

SDP.81 MUSE data were obtained from the ESO Science Archive Facility[10] (ProgID: 294.B-5042, PI: Gavazzi, Raphael), and the observations occurred in 2015 as a part of the program *"MUSE spectroscopy of SDP.81, the highest resolution Sub-mm ALMA gravitational lens: accurate mass profile & line of sight structure"*, with 13600s of total exposure time. The data cover the spectral range of $460 - 935$nm, with mean spectral Resolution (R) equal to 2989[11] and spectral scale of 1.25Å.

The field of view is $1.64' \times 1.64'$ and the pixel scale is $0.2''$. A collapsed image covering the spectral range of $480 - 800$nm can be seen in Fig. 11. The PSF was built following the same procedures of those used for the HST image. We use DAOPHOT task

[10] <http://archive.eso.org/cms.html>
[11] Which is roughly equivalent to a FWHM of 2.71Å in the galaxy frame.

to construct an ePSF (FWHM approximately equal to 0.421″, corresponding to ∼ 1930pc in the redshift of the lens galaxy) based on unsaturated stars near the galaxy, and once again, this ePSF will be parametrized by the MGE method (see Section 5.2).

The MUSE data were reduced as a product of ESO Phase 3 archive[12], following the standard pipeline described in Weilbacher; Streicher; Palsa (2016). The final data cube was submitted to the Zurich Atmosphere Purge code (ZAP; Soto et al., 2016), that performs an enhancement of the MUSE data cube, removing telluric lines, that may have been left in the previous reduction step.

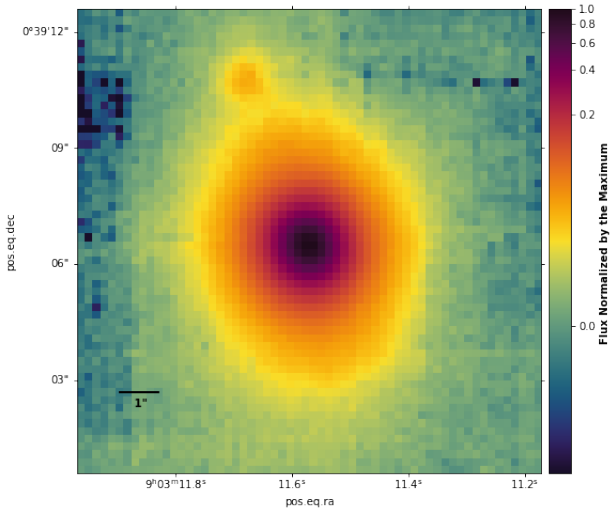


Figure 11 – SDP.81 MUSE collapsed image in the spectral range of 480 − 800nm. North is up and East is left.

4.3 ALMA data

ALMA is one of most powerful interferometers operating in submillimeter (submm) wavelengths, consisting of 64 antennas, of which 54 have 12m diameter, and 12 have 7m diameter[13], placed in Chajnantor plateau, Chile.

Working together, this array of antennas is able to measure the interference pattern caused by the delay of the measurement of the signal received from an object between two or more antennas separated by a baseline B. In other words, emission coming from different positions of the same source in the sky arrives in different times at each antenna of the array (Fig. 12). This delay in the signal received between each pair of antennas causes an interference pattern, which is related to the source brightness and position in the sky.

The interference pattern measured by the interferometer is located in the Fourier space (u, v), also called $uv-$plane, where each pair of antenna represents a point in the

[12] <http://archive.eso.org/cms/eso-data/phase3/ESO_reduced_data_products_description.pdf>
[13] <https://www.eso.org/public/brazil/teles-instr/alma/antennas/>

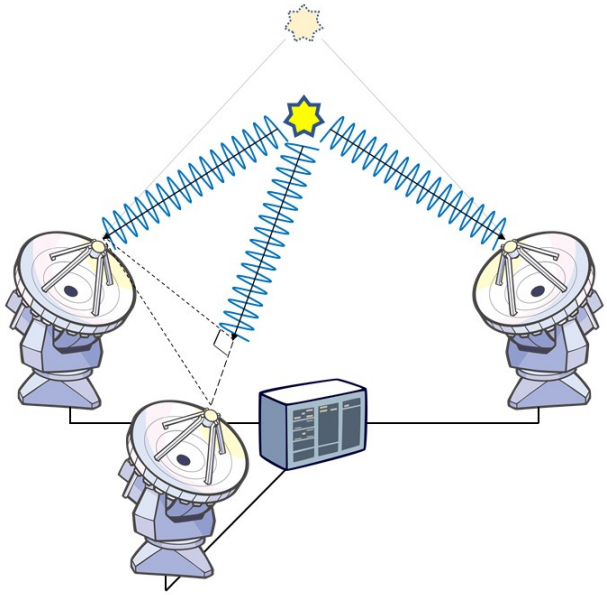


Figure 12 – Representation of signal coming from the same source, but arriving in the antennas at different position and time. After received by the antennas, the signal is processed by a correlator. Image taken from: ALMA (ESO/NAOJ/NRAO).

$uv-$plane (called visibility). This sample of discrete and finite number in the Fourier domain (each antenna pair) gives rise to the visibility function $V(u,v)$ in the $uv-$plane, which is the Fourier transform of the source surface brightness $T(x,y)$ (see Van Cittert–Zernike theorem[14]),

$$V(u,v) = \int T(x,y)\ \exp^{-2\pi i(ux+vy)}\, dxdy. \tag{4.1}$$

By the inverse Fourier transform we obtain the surface brightness

$$T(x,y) = \int V(u,v)\ \exp^{2\pi i(ux+vy)}\, dudv. \tag{4.2}$$

However, each pair of antennas only samples one point in the Fourier space, such that an array cannot sampling all the Fourier domain, resulting in a imperfect image. Fortunately, the Earth rotation makes possible different configurations of the array with respect to the source position, sampling better the $uv-$plane, and improving the quality of the final image.

Since not all Fourier domain is sampled, the image obtained by the inverse Fourier transform (equation 4.2) is not the "true" image of the sky, but an image of the true sky brightness distribution convolved with a dirty beam (also called *synthesized beam*). This synthesized beam means that there are no information about the true sky brightness

[14] <https://en.wikipedia.org/wiki/Van_Cittert%E2%80%93Zernike_theorem>

distribution on those specific angular scales. This image is also called *dirty image*. Nevertheless, dirty images can be improved applying deconvolution methods, whose intention is to mitigate the incompleteness of sampling in the $uv-$plane. One of such techniques is called "CLEAN"[15], which results in an image of the true sky brightness convolved with an *clean beam*, which is described by a two-dimensional Gaussian[16], described by

$$f(x,y) = A\exp\left\{-\left(\frac{4\ln(2)}{d_1^2}(\cos(\theta)x+\sin(\theta)y)^2+\frac{4\ln(2)}{d_2^2}(-\sin(\theta)x+\cos(\theta)y)^2\right)\right\}. \tag{4.3}$$

As stated in the documentation of CASA: *"[...] A is the amplitude (usually set to unity) and θ is the anti-clockwise angle from the x axis to the semi-major axis of the ellipse described by $f(x,y)$ in the $xy-$plane (x-axis and the semi-major axis should be coplanar). The d_1 and d_2 are respectively the semi-major and semi-minor axis of this ellipse, which is formed by the cross-section that lies parallel to the $xy-$plane, at a height so that d_1 is equal to the FWHM distance of the one dimensional Gaussian which lies on the plane formed by the z axis and d_1. Note that $d_1 \geq d_2 > 0$, since d_1 is the semi-major axis."* The *cleaned image* is the proper image for scientific usage.

More information about interferometry data can be found through the courses offers by Italian ALMA Regional Centre[17]. And for lens modelling using submm data, and interferometry see the website of Dr. Mattia Negrello[18]. A good overview, with data reduction, is given by Remijan, A. et al. (2015), as well.

SDP.81 data were obtained as part of Science Verification for the 2014 ALMA Long Baseline Campaign, using the most extended configuration of ALMA to date. These observations were taken in bands 4 ($\sim$ 2mm), 6 ($\sim$ 1.3mm) and 7 ($\sim$ 1.0mm) using between 23 and 36 antennas, and they are publicly available at the ALMA Science Portal[19]. These bands comprise both continuum emission and some emission lines of CO and H_2O. The data reduction and calibration were done following the pipeline provided by the ALMA Science Portal[20]. These procedures were implemented using the Common Astronomy Software Applications (CASA[21]; McMullin et al., 2007). All the detailed information about observation, data reduction and imaging of SDP.81 is presented in Vlahakis et al. (2015).

In this work we perform the modelling directly in the image plane, instead of the uv-plane. We choose to work with the image generated by the continuum band 7, which includes the $\sim 250\mu$m rest-frame emission, and has original pixel scale of 0.005″ (Amazing,

[15] <https://www.cv.nrao.edu/~abridle/deconvol/node7.html>
[16] For more details see **Definition Synthesized Beam** at CASA documentation in <https://casadocs.readthedocs.io/en/stable/index.html>
[17] <http://www.alma.inaf.it/index.php/Courses>
[18] <http://www.mattianegrello.com/?page_id=2488>
[19] <https://almascience.eso.org/>
[20] <https://almascience.nrao.edu/processing/science-pipeline>
[21] <https://casadocs.readthedocs.io/en/stable/index.html>

isn't it?). Fig. 13 shows the emission of the background source, its rest-frame submm emission. Here, it is important to highlight that due to the wavelength observed by ALMA, the lens light is not visible, and only the gravitational arcs from the background source are captured. Such feature is perfect for lens modelling, since the contamination due to the lens galaxy light is negligible.

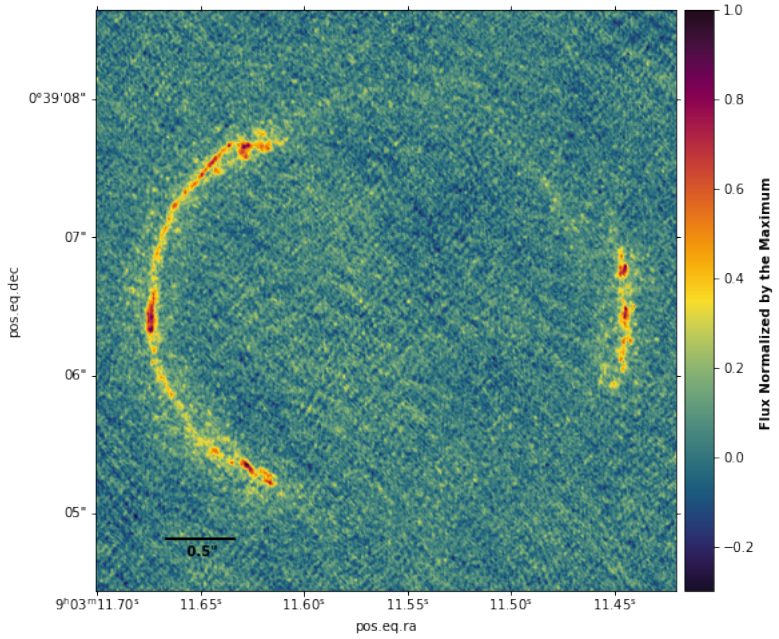


Figure 13 – ALMA observations of SDP.81 system, where only the background emission are visible, due to the spectral window of the observations (band 7 continuum). The beam size is not visible because it is on the order of the pixel size. Furthermore, the typical flux emission in the order of mJy. North is up and East is left.

The choice for performing the lens modelling in the image plane was motivated by two main reasons: first, modelling in the image plane allow us to mask and fit only the regions where the emission of the lensed source is is detected. Second, this choice turns the modelling more computationally efficient (we model thousands of pixels instead of millions of visibilities) and improves the goodness of fit, since extended areas of background sky are left aside. The main disadvantage of this choice is that we are modelling the Fourier transform of the data, instead of the interferometric visibilities. As a consequence of the procedure of obtaining an image from the visibilities (i.e. the Fourier transform), the image pixels becomes correlated by the beam, which could biases the image plane modelling, if the uncertainties do not take into account the covariance between the pixels. However, due to the high coverage of the $uv-$plane in these ALMA data, the error associated to the image plane is significantly reduced. As shown by Dye et al. (2015), the beam size for the observations of SDP.81 is of the order of the pixel scale of the cleaned image, such that the errors associated to the transformation of visibilities to the image are negligible.

In addition, Dye et al. (2018) performs a comparison between modelling the cleaned

image data and the visibility data directly. In this work they find minor differences in the reconstructed source and lens model, when a semilinear inversion (see Section 5.4.1) is applied. It is worth mentioning that in this work they model 6 galaxy-scale gravitational lens systems using ALMA data with much lower angular resolution (due to ALMA configuration and integration time, which resulted in a more sparse uv coverage) than the data obtained for the SDP.81 system. According to this work, the authors conclude that the differences in the modelling of the image data and the visibility data will become more notable when the coverage of the $uv-$plane becomes more sparse.

For lens modelling purposes, we have binned the band 7 continuum image from a pixel scale of $0.005''$ to a pixel scale of $0.02''$. This increases the modelling efficiency and also reduces the covariance between the pixels. For the modelling (see Chapter 6) we use the synthesized beam sizes as obtained by the ALMA team, with $0.0308'' \times 0.0235''$ ($\sim$ 141pc $\times$ 108pc at the lens redshift) for the beam semi-major and semi-minor axes respectively, and the position angle of 15°. This synthesized beam also has been binned in to the same pixel scale. We apply a small astrometric shift of $0.24''$ to align HST data to the ALMA data. We determine this shift using the SDSS Photometric Catalogue (DR12) Alam; et al. (2016) to match stars in the HST field. We notice that the shift applied in HST data results in a good agreement with the lens center determined by the H-ATLAS candidate lensed galaxies catalogue[22] (Negrello et al., 2017). Then, to align the MUSE data, we use the HST image[23] as a reference, and perform an astrometric shift equal to $0.18''$ in the MUSE data.

In Fig. 14 we show the near-infrared emission of the lensed source at $z = 3.04$ observed with HST F160W and ALMA band 7 continuum contours overlaid. As it is immediately seen in the figure, there is a distinct offset between the submm emission (white contours, tracing the dust) and the rest-frame optical emission (mainly tracing the stars). An additional structure is observed to the north of the HST image, as well as a much larger extent of the F160W emission of the western gravitational arc towards the south. Such offset and extended near-infrared emission were also reported in Dye et al. (2015), where the authors conclude, after the lens modelling reconstruction, that the most plausible scenario to explain these observations is that the background source consists of two objects which are merging. The authors also argue that this galaxy merger scenario is supported by the reconstruction of the source kinematics (from the CO emission), which reveals a rotating disc of gas and dust in a state of collapse.

For completeness, in Fig. 14 we also show the MUSE data as dashed yellow contours overlaid.

Now, in the next Chapter we will present the numerical methods used in this

[22] <https://vizier.u-strasbg.fr/viz-bin/VizieR?-source=J/MNRAS/465/3558>
[23] Already aligned with ALMA data.

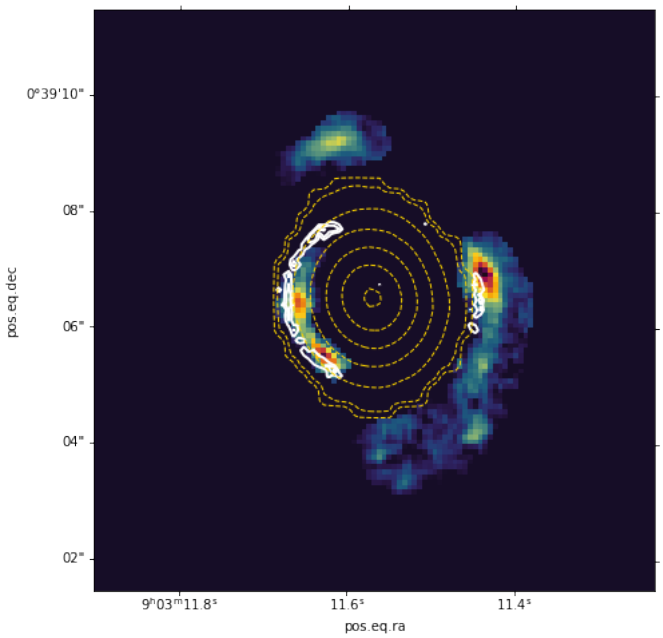


Figure 14 – Overlaid contours of MUSE (dashed yellow), and ALMA (white) in HST F160W imaging. North is up and East is left.

dissertation. The objective is to present the models and strategies used to solve both, the Jeans equations that describes the kinematics (Chapter 2), and the lens equation (Chapter 3) that describes the light deflection. Chapter 5 also describes the statistical method we use to combine gravitational lensing and galaxy dynamics in order to infer the most probable value for $\gamma_{\rm PPN}$ given the data presented in this Chapter.

5 Numerical Methods

In this Chapter we describe the numerical methods and implementations developed to constrain $\gamma_{\rm PPN}$ using a galaxy that acts as a gravitational lens at $z_l = 0.299$. All softwares and packages used are open source, as well as the scripts we developed, that can be accessed at <https://github.com/carlosRmelo/Master-Degree.git>. We belive that a free and open science is the best way to build a more robust and transparent Science, and also a way to promote equity and social justice, providing equal access to information, research and data. So, always consider making your research public.

As best practice, we organize all the software and packages used in Appendix B, including the installation method, version, website and any other considerations that we deem necessary.

5.1 Integral Field Kinematics

In this Section we describe the method we use to recover the LOSVD through the galaxy observed spectra. For this, we use the publicly available Penalized Pixel-Fitting (pPXF[1]; Cappellari; Emsellem, 2004) method, implemented in Python by Cappellari (2016a). For a complete discussion about the method, its implementation, limitations, and tests, we refer to the original aforementioned papers.

The idea behind pPXF is to recover the LOSVD parametrized by a Gauss-Hermite polynomial in the pixel space. There are a lot of benefits in choosing the pixel space instead of Fourier space, among them we highlight the capability to mask bad pixels more easily, and the possibility to simultaneously fit the stellar/gas kinematics and stellar population.

Assuming that a galaxy can be formed only by its stellar populations[2], the observed spectrum at a given position in the sky would be the sum (luminosity-weighted) of individual stellar spectra convoluted with their LOSVD. This makes the recovery of LOSVD a deconvolution problem of the galaxy spectrum. pPXF tries to do that by making a model spectrum $G_{mod}(x)$ of the galaxy, through the convolution of standard template spectra $T(x)$ and the parametrized LOSVD. Here, both galaxy and template spectra are rebinned in wavelength to a linear scale $x = \ln(\lambda)$, where λ is the wavelength. The model can also take into account multiplicative or additive polynomials. The model

[1] <https://www-astro.physics.ox.ac.uk/~mxc/software/#ppxf>

[2] More accurate results can be accessed including the contributions of the gas and the emission of the sky in the model spectrum, as we will see next.

spectrum can be described as follows (eq. 11 in Cappellari, 2016a)

$$G_{mod}(x) = \sum_{n=1}^{N} w_n \left\{ [T_n(x) * \mathcal{L}_n(cx)] \sum_{k=1}^{K} a_k \mathcal{P}_k(x) \right\} + \sum_{l=1}^{L} b_l \mathcal{P}_l(x) + \sum_{j=1}^{J} c_j \mathcal{S}_j(x), \quad (5.1)$$

where $\mathcal{L}_n(cx)$ are the LOSVDs, which may vary depending on the N template adopted[3], $T_n(x)$ are a set templates[4], $\mathcal{P}_k$ and $\mathcal{P}_l$ are multiplicative or additive orthogonal polynomials of degree k and l respectively (with a_k and b_l its coefficients), $\mathcal{S}_j(x)$ are a set of sky templates, $*$ denotes convolution, and finally w_n and c_j are the luminosity-weighted coefficients. Additive polynomials can be included to avoid mismatch between the galaxy spectrum and the templates adopted, while multiplicative polynomials can correct spectral calibration, and make the fit insensitive to reddening by dust. In this expression, K, L and J depend on the number of templates used, and the degree of polynomials (additive or multiplicative) adopted, which are defined by the user.

The LOSVD in pPXF, for both the stellar and gas components, is parametrized by a Gauss-Hermite[5] polynomial in the form

$$\mathcal{L}(v) = \frac{\exp\left(-0.5\, y^2\right)}{\sigma\sqrt{2\pi}} \left[1 + \sum_{m=3}^{M} h_m H_m(y) \right], \quad (5.2)$$

where $y = (v - V)/\sigma$, v is the line-of-sight velocity, V the mean velocity, and σ the velocity dispersion. Also, H_m are the Hermite polynomials[6], with M being the order of the momenta.

The best fit for the galaxy spectrum model $G_{mod}(x)$, eq. (5.1), can be found minimizing the χ^2, which measures the agreement between the model and the observed galaxy spectrum, over the set of I spectral good pixels

$$\chi^2 = \sum_{i=0}^{I} \left[\frac{G_{mod}(x_i) - G(x_i)}{\Delta G(x_i)} \right]^2, \quad (5.3)$$

where $G(x)$ is the observed galaxy spectrum, and $\Delta G(x_i)$ the 1σ uncertainty in each pixel where the galaxy spectrum was measured. This minimization is a non-linear least-squares optimization problem for the parameters $[V, \sigma, h_3, ..., h_M]$[7], and this minimization is what pPXF does efficiently!

[3] For example, we can choose to use different LOSVDs for stars and gas components.

[4] These templates can be formed by stellar templates, galaxy templates, single stellar population models or combinations of them.

[5] This parametrization was introduced by van der Marel; Franx (1993) and Gerhard (1993) as a way to "naturally" measure asymmetric (h_3) and symmetric (h_4) deviations from a Gaussian profile.

[6] <https://mathworld.wolfram.com/HermitePolynomial.html>

[7] And possibly other parameters of the equation 5.1, when included during the fit.

Another important feature in pPXF is that the algorithm *penalizes* non-Gaussian solution. This is motivated by the fact that the LOSVD of galaxies is generally well reproduced by a Gaussian (Bender; Saglia; Gerhard, 1994; Dekel et al., 2005). To allow that, the χ^2 in eq. (5.3) should be modified to

$$\chi^2_p = \chi^2 + \alpha\mathcal{P}, \tag{5.4}$$

where $\mathcal{P}$ is a penalty function, which penalizes non-Gaussian solutions based on the deviation of the LOSVD from a Gaussian shape, while α is an adjustable penalty, which depends on the data quality. The discussion about the form of $\mathcal{P}$ can be found in Cappellari; Emsellem (2004), Section 3.3, and the discussion about the non-linear fitting procedure in Cappellari (2016a), Section 3.4. An important note is that the procedure also takes into account the line spread function[8] of the instrument, i.e., the ability of the instrument to resolve two neighboring spectral lines.

An example of fit using pPXF can be seen in Fig. 15. To illustrate this method we used the NGC4636 galaxy spectrum obtained from the Sloan Digital Sky Survey (SDSS) during data release 12. The spectrum is available at <http://dr12.sdss3.org/>. For this fitting we use a subset[9] of the Single Stellar Population library by Vazdekis et al. (2010), also called MILES models, with 4 moments $[V, \sigma, h_3, h_4]$, and an additive polynomial of degree 12. This example is available at pPXF documentation under the name ppxf_example_kinematics_sdss.py. The best fit parameters and the reduced χ^2_{DOF} (the χ^2 per degree of freedom) are shown in the top left of the Figure, where the systemic velocity was not subtracted.

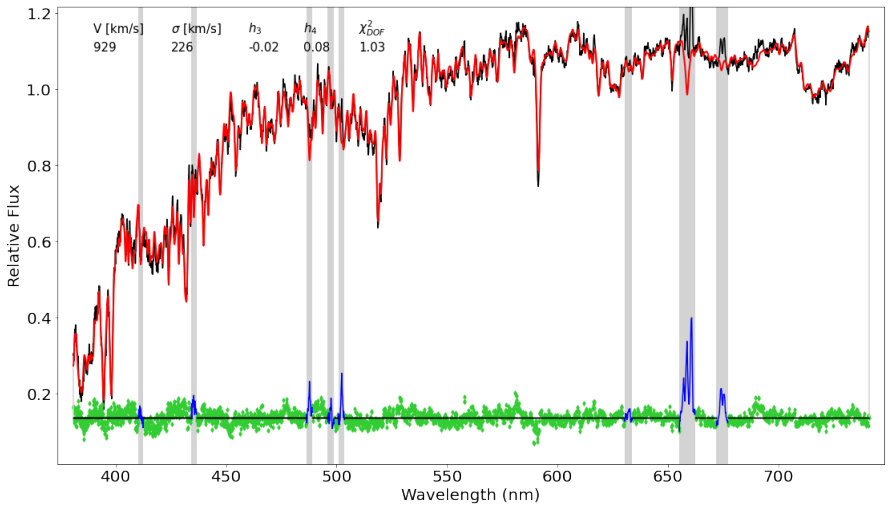


Figure 15 – pPXF fitting of NGC4636 galaxy spectrum from SDSS.

[8] A version of the point spread function, but in the spectroscopic context.
[9] This subset is provided jointly with the pPXF code.

5.2 Multi-Gaussian Expansion Formalism

Now we describe the Multi-Gaussian Expansion (MGE; Emsellem; Monnet; Bacon, 1994; Cappellari, 2002) approach, which will be used for describing the mass model used in both dynamical and lensing modelling. A more detailed description can be found in the above cited papers.

This approach consists in a feasible and robust form of parametrizing the galaxy surface brightness (assuming that it is monotonically decreasing), in such a way that we can use it as the tracer of mass and luminosity. Another advantage of this method is its capability to capture deviations of the isophotes from ellipses, making possible the modeling of multi-component objects such as lenticulars and spirals. With this parametrization, we are able to describe the surface mass density used to compute the lens deflection angle (eq. 3.6), and the luminosity density used as a tracer in dynamical modelling (eq. 2.14). More important, this parametrization ensures a self-consistent mass profile for both models.

This method was used in a vast range of applications, including dynamical modelling (Cappellari et al., 2007), gravitational lensing (van de Ven et al., 2010), and measurement of the mass of supermassive black holes (Cappellari et al., 2010).

Assuming that $I(x', y')$ is the projected surface brightness, we can parametrize it as a sum of 2D-Gaussians as follows:

$$I(x', y') = \sum_{j=1}^{N} \frac{L_j}{2\pi\sigma_j^2 q'_j} \exp\left[-\frac{1}{2\sigma_j^2}\left(x'^2_j + \frac{y'^2_j}{q'^2_j}\right)\right], \tag{5.5}$$

where (x',y',z') is a projected coordinate system centered in the galaxy center with z' along the line-of-sight, and N is the total number of Gaussians adopted[10]. Each Gaussian component has a total luminosity L_j, an observed projected axial ratio $0 \leq q'_j \leq 1$, and dispersion σ_j along the semi-major axis, which is assumed to be aligned with x'-coordinate.

To obtain the intrinsic three-dimensional luminosity density $\nu(\boldsymbol{r})$[11] it is necessary to deproject the luminosity surface density above, eq. (5.5). However, even if we knew the real inclination angle i (which is not true in general), the deprojection is not unique, except in the cases where the galaxy is edge-on ($i = 90°$). Fortunately, the deprojection becomes unique once a model is adopted. Following Cappellari (2002), and assuming an

[10] The number of Gaussians adopted is defined *on-the-fly* during the fit procedure, see the discussion in Cappellari, (2002), Sections 3.3 and 3.4.

[11] In Chapter 2, equation (2.14), we had used the notation $j(\boldsymbol{r})$ for the three-dimensional luminosity density, however it is common in the literature to merge the term L with $\nu(r)$, and treat $\nu(r)$ as the three-dimensional luminosity density. Even considering this a bad practice, we will adopt this notation hereafter as a way to maintain consistency with the available literature. We took the opportunity to remind that the system is considered a steady state, so there is no time dependency, and we left the t notation aside.

oblate axisymmetric model, the intrinsic luminosity density can be written, in cylindrical coordinates (R, ϕ, z) as

$$\nu(R,z) = \sum_{j=1}^{N} \frac{L_j}{(\sqrt{2\pi}\,\sigma_j)^3\, q_j} \exp\left[-\frac{1}{2\pi\sigma_j^2}\left(R^2 + \frac{z^2}{q_j^2}\right)\right], \tag{5.6}$$

where L_j and σ_j are the same as in (5.5), and q_j is the deprojected three-dimensional intrinsic axial ratio, related with the projected axial ratio by

$$q_j^2 = \frac{q_j'^2 - \cos^2 i}{\sin^2 i}. \tag{5.7}$$

The luminosity density can be easily converted to mass density, just assuming a mass-to-light ratio (M/L), and multiplying eq. (5.6) by it. In general, the M/L may change across the galaxy, such that to model such variation within the MGE framework one can consider that each Gaussian component will have its own mass-to-light ratio. For this reason, we can assume one M/L per Gaussian component as in the form

$$\Upsilon_j = M_j/L_j, \tag{5.8}$$

$$M_j = \Upsilon_j\, L_j, \tag{5.9}$$

where M_j is the mass of the j^{th} Gaussian component with L_j luminosity[12]. Nevertheless, many works (e.g. Cappellari et al., 2013; Cappellari et al., 2009; Williams; Bureau; Cappellari, 2009) show a good agreement with an assumption of an constant M/L, i.e. $\Upsilon_j = \Upsilon$. Then, assuming a "multicomponent" M/L, leads us to a mass density profile given by

$$\rho(R,z) = \sum_{j=1}^{N} \frac{M_j}{(2\pi)^{3/2}\sigma_j^3 q_j} \exp\left[-\frac{1}{2\pi\sigma_j^2}\left(R^2 + \frac{z^2}{q_j^2}\right)\right]. \tag{5.10}$$

Once we have obtained the mass density profile (5.10), the gravitational potential can be derived from the Homoeoid Theorem for densities stratified on similar concentric ellipsoids (see Binney; Tremaine (2008) equation 2.140)[13]. In the MGE case, the gravitational potential can be written as

$$\Phi(R,z) = -G\sqrt{\frac{2}{\pi}} \sum_{j}^{N} \frac{M_j}{\sigma_j} \tilde{\Phi}_j(R,z), \tag{5.11}$$

where G is the gravitational constant and $\tilde{\Phi}_j(R,z)$ is given by

$$\tilde{\Phi}_j(R,z) = \int_0^1 \frac{d\tau}{\sqrt{1-\zeta_j^2\tau^2}} \exp\left[-\frac{\tau^2}{2\sigma_j^2}\left(R^2 + \frac{z^2}{1-\zeta_j^2\tau^2}\right)\right], \tag{5.12}$$

[12] In the end, for a more robust approach, such values can be constrained by stellar population models and will depend on the color gradients across the considered galaxy.

[13] See also Chandrasekhar (1969) for the proof.

with $\zeta_j^2 = 1 - q_j^2$.

Besides the parametrization of surface brightness profile, the flexibility of MGE method also makes possible the parametrization of a vast range of decreasing one-dimensional functions (power-laws are the best candidates). For example, a function with a prominent peak can be approximated by a sum of Gaussians with higher amplitude and small dispersion, while smooth functions can be approximated by a sum of Gaussians with larger dispersion. An example is shown in Fig. 16, where we fit the function $y = \frac{100}{(1+x)^4}$ using the MGE technique. For this fitting, a set of 16 Gaussians was used, each of them represented by a different color. The sum of all Gaussians results in the blue bold line that fits the original function, in orange, with high accuracy as shown in the residuals. This example can be found at MgeFit documentation as mge_fit_example.py.

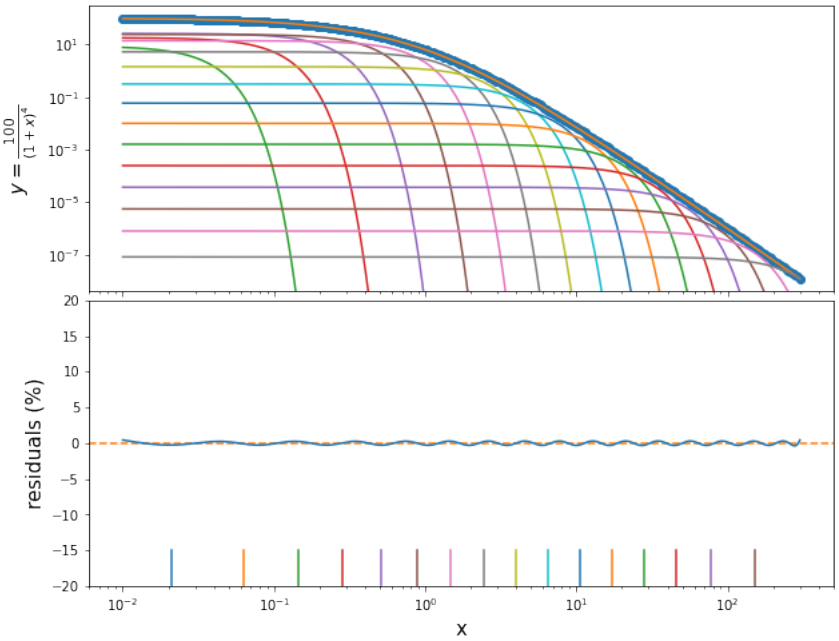


Figure 16 – Fitting of a power-law function using the MGE parametrization. For this parametrization, a set of 16 Gaussians are used, represented by different colors. The sum of all Gaussians (bold blue) fits very well the original function (in orange), as shows the residuals in the bottom panel. The x coordinate are logarithmically spaced, therefore the function are logarithmically sampled, as a requirement of MgeFit routine.

This capability of parametrizing function allows us to include other contributions in the mass density profile beyond those tracked by luminosity, like a dark matter content or an supermassive black hole (SMBH). We are able to do that because, in principle, we can parametrize a one-dimensional mass profile and then, assuming some axial ratio for each Gaussian in the parametrization, build a projected mass density, such as given by eq. (5.5) when it is multiplied by a mass-to-light ratio.

The MGE parametrization is one of the most crucial parts of what comes next. So, in

order to perform this parametrization, we use the Python version of MgeFit[14] software implemented by Cappellari (2002), to decompose the projected luminosity density of galaxies. We follow closely the procedure described in the paper, therefore, we recommend the reading of the related paper for more information on how the algorithm works. Here we just emphasize that the algorithm starts performing the fitting along one-dimensional sectors from the major-axis to the minor-axis (that covers the entire galaxy) logarithmically spaced in elliptical radii $m'^2 = x'^2 + y'^2/q'^2$, where q' is a representative axial ratio of the galaxy isophotes. Then a non-linear least-squares minimization is performed to ensure that the best parameters of eq. (5.5) are found. The algorithm also takes into account the PSF of the galaxy image.

To perform the conversion from the Gaussian parameters fitted by the MgeFit procedure to physical quantities, we follow the standard definitions presented in the documentation. Briefly, the MgeFit method returns three vectors of values $\mathrm{Tc}_j, \sigma_j^{\mathrm{pix}}$, and $\mathrm{q}_j^{\mathrm{Obs}}$ with length equal to the number of Gaussians used in the parametrization. Tc_j is the total counts of the j^{th} Gaussian component, σ_j^{pix} is the associated dispersion of the Gaussian in pixel units, and $\mathrm{q}_j^{\mathrm{Obs}}$ is the projected axial ratio, q'_j.

The total central counts Tc_j of each Gaussian can be converted to the corresponding central peak surface brightness C_j^0 (in counts per pixels^{-1}) using

$$C_j^0 = \frac{\mathrm{Tc}_j}{2\pi(\sigma_j^{\mathrm{pix}})^2 q_j^{\mathrm{Obs}}}. \tag{5.13}$$

To convert the central surface brightness C_j^0 in counts per pixels^{-1} into to a central surface brightness $\mu_j(X)$ in mag per arcsec^{-2}, we apply the standard photometry equation,

$$\mu_j(X) = Z_X + 5\log(\mathrm{SCALE}) + 2.5\log(\mathrm{EXPTIME}) - 2.5\log C_j^0 - A_X, \tag{5.14}$$

where X is the notation for the photometric band, Z_X is the photometric zeropoint in band X, SCALE is the scale plate, EXPTIME is the exposure time of the image, and A_X is the extinction in the X band. Thus, $\mu_j(X)$ is the central surface brightness in band X, measured in magnitudes per arcsec^{-2}, for the j^{th} Gaussian.

Finally, the central surface brightness $\mu_j(X)$, can be converted to the central surface density I_j in $\mathrm{L}_\odot\mathrm{pc}^2$ through

$$I_j = \left(\frac{64800}{\pi}\right)^2 10^{0.4\left(\mathrm{M}_{\odot,X} - \mu_j(X)\right)}, \tag{5.15}$$

where $\mathrm{M}_{\odot,X}$ is the absolute solar magnitude in the band X, and $\mathrm{L}_\odot$ is the solar luminosity. We have already anticipated the fact that these are the proper units used by the JAM

[14] <https://pypi.org/project/mgefit/>

modelling that will be presented in the next Section. The conversion of $\sigma_j^{\rm pix}$ from pixels to arcsec is much more simple, just multiply it by the SCALE value.

An useful physical quantity to define here is the total luminosity per Gaussian

$$L_j = 2\pi I_j \sigma_j^2 q'_j, \tag{5.16}$$

where the galaxy distance is needed to convert σ_j, originally given in arcsec, to pc.

5.3 Jeans Anisotropic Modelling

Here we describe the Jeans Anisotropic Modelling (JAM; Cappellari, 2008; Cappellari, 2020), method used for describing the kinematics of real galaxies using the Jeans equations. We start recalling the two equations, (2.44) and (2.45), for the second velocity moment derived in Section 2.2.1,

$$\overline{v_z^2} = \frac{1}{\nu(R,z)} \int_z^\infty dz'\, \nu(R,z') \frac{\partial \Phi(R,z')}{dz'}, \tag{5.17}$$

$$\overline{v_\phi^2} = \overline{v_R^2} + \frac{R}{\nu} \frac{\partial(\nu(R,z)\,\overline{v_R^2})}{\partial R} + R\frac{\partial \Phi(R,z)}{dR}. \tag{5.18}$$

Before applying the MGE formalism in this context, we need to make another assumption about the anisotropy of the velocity dispersion components, i.e., about the alignment of the velocity ellipsoid (the orbits followed by the stars when we assume an axisymmetric system). We already assumed an axisymmetric system, which led us to conclude that $\overline{v_R\, v_z} = 0$. Now, if we assume a constant flattening of the orbits in the meridional plane, we can write $\overline{v_R^2} = b\overline{v_z^2}$, where b is the anisotropy. Under this assumption, equation (5.18) can be solved uniquely, once we know $\overline{v_z^2}$ and $\Phi(R,z)$

$$\overline{v_\phi^2} = b\left[\frac{R}{\nu} \frac{\partial(\nu(R,z)\,\overline{v_z^2})}{\partial R} + \overline{v_z^2}\right] + R\frac{\partial \Phi(R,z)}{dR}. \tag{5.19}$$

Lastly, to make the comparison between the theory and observations, these intrinsic quantities, such as velocity dispersion defined above, should be integrated along the line-of-sight. A caveat in this procedure is that, in general, we do not know the galaxy inclination i, therefore the coordinate system (x, y, z) intrinsic to the galaxy does not necessarily have the z-axis along the line-of-sight.

Hence, we define a sky coordinate system (x', y', z') with z' along the line-of-sight, and x' parallel to the projected semi-major axis of the galaxy. If the inclination i is the angle between the axis z and z', which implies $i = 90°$ when the galaxy is edge-on, the

transformations between the velocities written in these two coordinate systems is given by (Malajovich, 2010; Cappellari, 2020)

$$\begin{bmatrix} v_{x'} \\ v'_y \\ v_{z'} \end{bmatrix} = \boldsymbol{S} \cdot \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} \quad \text{with} \quad \boldsymbol{S} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos i & -\sin i \\ 0 & \sin i & \cos i \end{bmatrix}. \tag{5.20}$$

However, we need to change the Cartesian coordinates (x, y, z) to cylindrical coordinates before, to maintain consistency with the coordinate system where the dispersions are written. Again, this is done by applying another transformation from (x, y, z) to (R, ϕ, z). coordinate system. The velocities are transformed according to

$$\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \boldsymbol{T_{cyl}} \cdot \begin{bmatrix} v_R \\ v_\phi \\ v_z \end{bmatrix} \quad \text{with} \quad \boldsymbol{T_{cyl}} = \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{5.21}$$

where $R^2 = x^2 + y^2$ and $\phi = \arctan(y/x)$ is the angle[15] in the $xy-$plane measured counter-clockwise from x-axis. Since both transformations, $\boldsymbol{S}$ and $\boldsymbol{T_{cly}}$, are linear transformations, we can define a *super-transformation* $\boldsymbol{ST} \equiv \boldsymbol{S} \cdot \boldsymbol{T_{cyl}}$, that maps (R, ϕ, z) into (x', y', z'). Then, the velocity components are

$$\begin{bmatrix} v_{x'} \\ v_{y'} \\ v_{z'} \end{bmatrix} = \boldsymbol{ST} \cdot \begin{bmatrix} v_R \\ v_\phi \\ v_z \end{bmatrix} \quad \text{with} \quad \boldsymbol{ST} = \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi\cos i & \cos\phi\cos i & -\sin i \\ \sin\phi\sin i & \cos\phi\sin i & \cos i \end{bmatrix}. \tag{5.22}$$

Therefore, the velocity along the line-of-sight $v_{z'}$ is

$$v_{z'} = (v_R \sin\phi + v_\phi \cos\phi)\sin i + v_z \cos i, \tag{5.23}$$

in the general case. But, what we observe is the projected second velocity moment luminosity-weighted, i.e, $I(x', y')\overline{v^2_{los}} \equiv I(x', y')\overline{v^2_{z'}}$, where $I(x', y')$ is the projected surface brightness. According to (2.20), we can compute this quantity as follows,

$$I(x', y')\overline{v^2_{los}} = \int_{-\infty}^{+\infty} \left(\int d^3\boldsymbol{v}\, f\, v^2_{z'} \right) dz', \tag{5.24}$$

where f is the DF and the integral over z' is the projection along the line-of-sight. We can simplify this expression writing

[15] Once again, is a bad practice use ϕ as coordinate and angle, but following the literature we will do that, but keep in mind this differentiation.

$$\int d^3\boldsymbol{v} f\, v_{z'}^2 = \sin^2\phi \sin^2 i \int d^3\boldsymbol{v} f\, v_R^2 + 2\sin\phi \sin^2 i \cos\phi \cancelto{0}{\int d^3\boldsymbol{v} f\, v_R v_\phi} +$$
$$+ \sin^2 i \cos^2\phi \int d^3\boldsymbol{v} f\, v_\phi^2 + \cos^2 i \int d^3\boldsymbol{v} f\, v_z^2 +$$
$$+2\cos i \sin i \sin\phi \cancelto{0}{\int d^3\boldsymbol{v} f\, v_R v_z} + 2\cos i \sin i \cos\phi \cancelto{0}{\int d^3\boldsymbol{v} f\, v_R v_\phi}$$
$$= \sin^2\phi \sin^2 i\, (\nu\overline{v_R^2}) + \cos^2\phi \sin^2 i\, (\nu\overline{v_\phi^2}) + \cos^2 i (\nu\overline{v_z^2}), \tag{5.25}$$

where the crossed term are null (see condition (iii) in Section 2.2.1). Combining this result with the previous one, we have:

$$\overline{v_{los}^2} = \frac{1}{I(x',y')} \int_{-\infty}^{+\infty} \left\{ \nu \left[\sin^2 i \left(\sin^2\phi\, \overline{v_R^2} + \cos^2\phi\, \overline{v_\phi^2} \right) + \cos^2 i\, \overline{v_z^2} \right] \right\} dz'. \tag{5.26}$$

The second moment $\overline{v_{los}^2}$ given by eq. (5.26) is a good approximation for the observed quantity $V_{rms}^2 = V^2 + \sigma^2$, also known the as root mean square velocity, where V is the observed mean stellar velocity and σ is the velocity dispersion. Finally, under the considerations made, we are able to model the velocity distribution along the line-of-sight, given a potential Φ, an inclination i and an anisotropy b.

The anisotropy and the inclination are two parameters where previous knowledge is sparse, and more robust techniques such as a non-linear search, are needed. We will return to this topic in later, when we will discuss the statistical method used for sampling all the unknown parameters. On the other hand, the gravitational potential can be tracked using the stellar luminosity, as we shown using the MGE method, equation (5.11). To do that, all we need is to assume a mass-to-light ratio, constant or even one per Gaussian component in the MGE parametrization.

Introducing the MGE formalism to solve the Jeans equations, results in a projected velocity dispersion along the line-of-sight given by (Cappellari (2008), eq. 28)

$$\overline{v_{los}^2}(x',y') = \frac{4\pi^{3/2} G}{I(x',y')} \int_0^1 \left\{ \sum_{k=1}^{N} \sum_{j=1}^{M} \nu_{0k} q_j \rho_{0j} u^2 \right.$$
$$\times \frac{\sigma_k^2 q_k^2 (\cos^2 i + b_k \sin^2 i) + \mathcal{D}_{jk} x'^2 \sin^2 i}{(1-\mathcal{C}_{jk})\sqrt{(\mathcal{A}_{jk} + \mathcal{B}_{jk}\cos^2 i)[1-(1-q_j^2)u^2]}}$$
$$\left. \times \exp\left(-\mathcal{A}_{jk} \left[x'^2 + \frac{(\mathcal{A}_{jk} + \mathcal{B}_{jk}) y'^2}{\mathcal{A}_{jk} + \mathcal{B}_{jk}\cos^2 i} \right] \right) \right\} du, \tag{5.27}$$

where $\nu_{0k} = \nu(0,0)$, $\rho_{0j} = \rho(0,0)$ and

$$\mathcal{A}_{jk} = \frac{1}{2}\left(\frac{u^2}{\sigma_j^2} + \frac{1}{\sigma_k^2}\right) \tag{5.28}$$

$$\mathcal{B}_{jk} = \frac{1}{2}\left(\frac{1-q_k^2}{\sigma_k^2 q_k^2} + \frac{(1-q_j^2)u^4}{\sigma_j^2[1-(1-q_j^2)u^2]}\right) \tag{5.29}$$

$$\mathcal{C}_{jk} = 1 - q_j^2 - \frac{\sigma_k^2 q_k^2}{\sigma_j^2} \tag{5.30}$$

$$\mathcal{D}_{jk} = 1 - b_k q_k^2 - \left[(1-b_k)\mathcal{C}_{jk} + (1-q_j^2)b_k\right] u^2. \tag{5.31}$$

Once there is no reason[16] to treat the gravitational potential as formed only by the stellar component, the Gaussians describing the potential and the luminosity are not necessarily equal. Thus, in the above expressions, the index k are related to the stellar content (i.e. luminosity) described by N Gaussians, and the index j is related to components of potential, described by M Gaussians (that could assigned to the stellar content or something else, like dark matter). Besides that, we let the constant anisotropy b_k be different for each luminous component k. In time, this luminosity-weighted anisotropy, is commonly parametrized as

$$\beta_z = 1 - \frac{\overline{v_z^2}}{\overline{v_R^2}} = 1 - \frac{\sum_k [\nu\overline{v_z^2}]_k}{\sum_k b_k [\nu\overline{v_z^2}]_k}, \tag{5.32}$$

where $[\nu\overline{v_z^2}]_k$ is a Gaussian component of the luminosity-weighted velocity dispersion, and the summation is over all the luminous Gaussians.

To compute the $\overline{v_{los}^2}$ in a robust and efficient way, we use the publicly available Jeans Anisotropic Modelling (JAM[17]; Cappellari 2008), that uses the MGE parametrization for the stellar mass profile (from the galaxy surface brightness distribution) and for the dark matter mass profile (see Section 5.4.5) and computes the predicted velocity second moment[18], under the assumptions discussed here. For an overview of the method, including the spherical solution of Jeans equations, see Cappellari (2020).

5.4 Lens Modelling

In this Section we will discuss the lens modelling, which involves the source reconstruction and the modelling of the lens mass. To solve both problems at the same time, we introduce the semilinear inversion method (SLI; Warren; Dye, 2003). The idea behind the method (and the reason for its name "semilinear") is that: once the mass profile

[16] Since the standard paradigm allows dark matter.
[17] Sometimes refer as JAMPY, which is the Python implementation of the code.
[18] Once more, convolving the model with the appropriate PSF associated with the IFU used.

is known, the source reconstruction is nothing more than an inversion matrix problem that can be solved very efficiently. Unfortunately, searching the mass profile parameters is still non-linear.

5.4.1 Semilinear Inversion Method

The SLI method assumes a pixelated source-plane and a parametrized lens mass model. Then, for a fixed mass model, the method performs a ray tracing using the lens equation (3.5), such that the lensed counts of the source plane pixels provides the best fit to the observed image. This procedure accounts for the PSF. The goodness of this fit is quantified by a merit function, G, which should be minimized in order to find the counts in the source plane pixels that better reproduce the counts in the observed image.

Let the source pixels be labeled by $i = 1, 2, ..., \mathcal{I}$, without any restriction on how the source plane is tessellated, and let the corresponding ray traced image pixel be labeled by $j = 1, 2, ..., \mathcal{J}$. The method assumes that each observed image pixel has a flux d_j and statistical uncertainty σ_j, this means that the method considers each pixel as an independent pixel, which is appropriated for CCD data.

Given a lens model, each j^{th} pixel of the lensed image can be mapped on each source pixel i, producing the source pixel surface brightness vector $\boldsymbol{s}$. This mapping is done by a matrix f_{ij}, such that the observed j^{th} pixel image can be computed as $\sum_i s_i f_{ij}$. The matrix f_{ij} accounts both the lens equation (3.5) and PSF. Finally, the ray traced j^{th} pixel can be compared with j observed image pixel that has a flux value d_j and uncertainty σ_j. Thus, the problem becomes to minimize the merit function G, in such a way that the vector $\boldsymbol{s}$ is the one that best reproduces the observed data. Assuming a χ^2 statistic, this can be place as

$$G = \chi^2 = \sum_{j=1}^{\mathcal{J}} \left(\frac{\sum_{i=1}^{\mathcal{I}} s_i f_{ij} - d_j}{\sigma_j} \right)^2 . \tag{5.33}$$

Minimizing this expression with respect to each of the s_i terms, yield a set of $\mathcal{I}$ equations in the following form

$$\frac{1}{2}\frac{\partial G}{\partial s_i} = 0 = \sum_{j=1}^{\mathcal{J}} \left(\frac{f_{ij} \sum_{k=1}^{\mathcal{I}} s_k f_{kj} - f_{ij} d_j}{\sigma_j} \right) . \tag{5.34}$$

The term $1/2$ is a mathematical trick, that will become clear soon.

We can rewrite this expression in matrix notation. To that, note that the summation over j can be placed inside the summation over k, then

$$\sum_{j=1}^{\mathcal{J}}\left(\frac{f_{ij}\sum_{i=k}^{\mathcal{I}} s_k f_{kj} - f_{ij}d_j}{\sigma_j}\right) = \sum_{k=1}^{\mathcal{I}} s_k \underbrace{\frac{\sum_{j=1}^{\mathcal{J}} f_{ij}f_{kj}}{\sigma_j}}_{F_{ik}} - \underbrace{\sum_{j=1}^{\mathcal{J}}\frac{f_{ij}d_j}{\sigma_j}}_{D_i}$$

$$= \sum_{k=1}^{\mathcal{I}} s_k F_{ik} - D_i, \tag{5.35}$$

and we write

$$\boldsymbol{sF} = \boldsymbol{D}, \tag{5.36}$$

where $\boldsymbol{F}$ is a symmetric $\mathcal{I}\times\mathcal{I}$ matrix, and $\boldsymbol{D}$ and $\boldsymbol{s}$ are column matrices of length $\mathcal{I}$. So, the solution $\boldsymbol{s}_{best}$ that minimizes the merit function is easily obtained just inverting the last equation, i.e.,

$$\boldsymbol{s}_{best} = \boldsymbol{F}^{-1}\boldsymbol{D}. \tag{5.37}$$

This procedure eliminates the requirement of searching for source parameters during the lens modelling, at the same time ensures that the global minimum is determined, given a lens model.

The errors in the source reconstruction can be accessed in a simple form (ignoring errors introduced by lens parameters), if we are able to compute the covariance matrix. But first, note that the Hessian matrix of our problem is given by

$$\frac{1}{2}\frac{\partial^2 G}{\partial s_i\,\partial s_k} = \sum_{j=1}^{\mathcal{J}}\frac{\partial}{\partial s_k}\left(\frac{f_{ij}\sum_{m=1}^{\mathcal{I}} s_m f_{mj} - f_{ij}d_j}{\sigma_j}\right)$$

$$= \sum_{j=1}^{\mathcal{J}}\frac{\sum_{m=1}^{\mathcal{I}} f_{ij}f_{mj}}{\sigma_j}\cancel{\frac{\partial s_k}{\partial s_m}}^{\delta_{mk}} - \sum_{j=1}^{\mathcal{J}}\cancel{\frac{\partial}{\partial s_k}\left(\frac{f_{ij}d_j}{\sigma_j}\right)}^{0}$$

$$= \sum_{j=1}^{\mathcal{J}}\frac{f_{ij}f_{kj}}{\sigma_j} = F_{ik}, \tag{5.38}$$

so the matrix $\boldsymbol{F}$ is one-half times[19] the Hessian matrix of our merit function G.

Now, if the variables of the problem can be considered Gaussian random variables (which is our case), the Hessian matrix, evaluated in its maximum, is equal to the inverse of the covariance matrix $\boldsymbol{C}$ (Yuen, 2010), which lead us with

$$\boldsymbol{C} = \boldsymbol{F}^{-1}, \tag{5.39}$$

and the errors in the source reconstruction are the diagonal terms of $\boldsymbol{C}$.

[19] Here is the reason of the trick used before.

5.4.2 Semilinear Inversion Method with Regularization

To mitigate noisy and possible nonphysical solutions of the source reconstruction, a different merit function can be designed, adding to the χ^2 a regularization term G_L,

$$G = \chi^2 + \lambda G_L \tag{5.40}$$

where λ is a scalar constant, referred to as the regularization weight, which quantifies the level of the regularization.

This regularization term G_L acts like a prior, heavily penalizing reconstructed sources which are less smooth. In order to preserve the linearity of the original merit function, the most straightforward form for the regularization term is a linear combination of terms $s_i s_k$,

$$G_L = \sum_{i,k} a_{ik} s_i s_k, \tag{5.41}$$

where a_{ij} are scalars.

Once the penalty terms are linear, the same analysis applied for the unregularized solution (Section 5.4.1) can be applied here. So, the regularized solution for the source pixels that minimizes the merit function (5.40) is

$$\boldsymbol{s}_{best} = [\boldsymbol{F} + \lambda \boldsymbol{H}]^{-1} \boldsymbol{D}, \tag{5.42}$$

where $\boldsymbol{H}$ is called regularization matrix, with elements given by

$$H_{ik} = \frac{1}{2} \frac{\partial^2 G_L}{\partial s_i \partial s_k}. \tag{5.43}$$

Again, the standard errors of the reconstructed source pixels are given by the diagonal terms of the covariance matrix $\boldsymbol{C}$, which is just

$$\boldsymbol{C} = [\boldsymbol{F} + \lambda \boldsymbol{H}]^{-1}, \tag{5.44}$$

as we see before.

There are three main forms for the regularization G_L in the literature (Warren; Dye, 2003), named: zeroth-order, gradient and curvature (also known as zero, constant and planar, respectively). The zeroth order is given by

$$G_L = \sum_{i}^{\mathcal{I}} s_i^2. \tag{5.45}$$

The simplest form for the gradient is

$$G_L = \sum_{x,y} \left(s_{x,y} - s_{x+1,y}\right)^2 + \sum_{x,y} \left(s_{x,y} - s_{x,y+1}\right)^2 , \tag{5.46}$$

that takes into account the difference between the neighboring pixels. This form of regularization ensures an additional step, the translation for the index i to the indices x, y.

In the same way, the simplest curvature regularization is

$$G_L = \sum_{x,y} \left[s_{x,y} - 0.5(s_{x+1,y} + s_{x-1,y})\right]^2 + \sum_{x,y} \left[s_{x,y} - 0.5(s_{x,y+1} + s_{x,y-1})\right]^2 . \tag{5.47}$$

5.4.3 Semilinear Inversion Method and a Bayes contribution

Suyu et al. (2006) introduced a Bayesian approach in the SLI methodology, which allows comparison between different models (give the lens parameters, source pixelization and regularization) in a confident way[20]. This approach aimed to find the most probable solution using the information coming from the data and the regularization scheme adopted, at the same time. For the lens problem, this approach tries to maximize the ϵ, defined by

$$\begin{aligned} -2\ln\epsilon = \sum_{j=1}^{\mathcal{J}} \left(\frac{\sum_{i=1}^{\mathcal{I}} s_i f_{ij} - d_j}{\sigma_j} \right)^2 &+ \ln\left[\det(\boldsymbol{F} + \lambda \boldsymbol{H})\right] - \ln\left[\det(\lambda \boldsymbol{H})\right] \\ &+ \lambda \boldsymbol{s}^T \boldsymbol{H} \boldsymbol{s} + \sum_{j=1}^{\mathcal{J}} \ln\left(2\pi\sigma_j^2\right). \end{aligned} \tag{5.48}$$

Equation (5.48) is able to quantify the goodness of the fit in many levels (Dye et al., 2008; Nightingale; Dye, 2015). One of them is quantifying the quality of the image reconstruction, which is related to the first term (i.e. the χ^2). Once the source reconstruction is a linear problem (a linear inversion), it is, in principle, possible to perfectly reconstruct the observed image, regardless of the image noise or the accuracy of the lens model. The regularization is necessary just to avoid these possible problems. Then, a "good" solution is achieved by equation (5.48), which assumes that the image data contain independent Gaussian noise in each image pixel (CCD data), therefore the best solution is the one whose the $\chi^2_{DOF} \sim 1$[21]. Solutions which $\chi^2_{DOF} < 1$ are penalized for being overly complex, whereas those with a $\chi^2_{DOF} > 1$ are penalized for not invoking a more complex source model, when the data supports it. This Bayesian approach quantifies the principle of Occam's razor[22]. Furthermore, for both circumstances, these penalties lead to a reduction

[20] This comparison is a good metric to be used if we want to rank different models in terms of goodness of the fit.

[21] It means that, the residuals between model and data are consistent with a Gaussian distribution as well.

[22] <https://en.wikipedia.org/wiki/Occam%27s_razor>

in $\ln \epsilon$, improving the fit. Another level of inference present in equation (5.48) is related to the terms two, three and four, which contains the regularization matrix $\boldsymbol{H}$. These terms are responsible for smooth source reconstruction, favouring simpler and less complex sources.

5.4.4 AutoLens

To perform the lens modelling using all the SLI methodology (including the Bayesian approach) presented in the last subsections, we use the open-source software AutoLens: Automated Modelling of a Strong Lens's Light, Mass and Source, by Nightingale; Dye; Massey (2018).

AutoLens is an entirely automated code for modelling and analysis of galaxy-scale strong gravitational lenses. With AutoLens, we are able to simultaneously model the lens galaxy light, lens galaxy mass profile, and reconstruct the source galaxy assuming a pixelated source or even a parametric model. AutoLens goes beyond the methodology present here, and allows the reconstruction of the source using an adaptive pixel-grid (Nightingale; Dye, 2015), and others regularization schemes.

As stated above, the SLI methodology finds the best source reconstruction given a lens mass profile, but the best mass profile is still unknown, and we need to find it. The most common approach is to assume a parametric mass model for the lens, and then do a non-linear search in the mass parameter space to find the best parameters (or the most probable ones). This non-linear search can be done using different techniques, such as sampling methods using Markov Chain Monte Carlo (MCMC), or even through optimizers which seek to find the best model regardless of the errors. Once again, AutoLens shows its completeness by providing a large set of lens mass models, and many different non-linear searchers[23].

An additional issue found in galaxy-galaxy strong lensing context, and solved by AutoLens, is the emission of the lens galaxy light. Normally, when using photometric data the lensed source are embedded in the lens galaxy light, so in order to reveal these arcs, we need to subtract the lens light. Again, the most common approach is to assume a parametric lens light model, find the best (or more probable) parameters, and subtract the lens light model to the original data. To do that, AutoLens offers different lens light profiles, allowing the lens light modelling and subtraction before modelling lens mass.

All these facilities are described and very well documented through lectures provided as Jupyter notebooks, at PyAutoLens[24] readthedocs. We strongly recommend these

[23] A tutorial about the different techniques implemented in AutoLens can be find at <https://pyautolens.readthedocs.io/en/latest/howtolens/chapter_optional.html>.

[24] The Python implementation of AutoLens. See <https://pyautolens.readthedocs.io/en/latest/index.html>

lectures, as they cover from theory to practical modelling, using simulated and real data.

Unfortunately, AutoLens does not provide the MGE parametrization as an alternative to light or mass profile. For this reason, it is necessary that we implement the MGE mass profile and the deflection angle associated in the context of AutoLens framework.

In technical terms, we create a new Python class describing the deflection angle produced by a MGE mass distribution. This new MGE mass class is then passed to AutoLens, that performs the semilinear inversion, i.e., given the mass model (MGE and its parameters), the pixelated grid source, and the regularization constant, AutoLens does the source reconstruction, computes the errors associated to it, reconstructs the observed image (the model image), and computes our goodness function $\ln \epsilon$, which quantifies the quality of the fit.

To our goals, the non-linear step of sampling the lens parameters is made externally to AutoLens, because we want to perform a combined analysis, searching for the most probable parameters that describe the lens mass profile and the dynamical model at the same time. In the following we describe the mass model used, and in the Section 5.6 we describe the sampling method.

5.4.5 AutoLens and mass models

As commented above, for AutoLens to perform the lens modelling, we need to setup the mass model to be used, and an important feature of the lens equation (eq. 3.12) is that it is linear in the mass term. It means that, it is possible to use different mass models, which generates different deflection angles, and the final deflection angle, at a given point, associated with the sum of the masses is just the sum of the individual deflection angles at that point. This is important because we can assume that the total mass responsible for bending the light path has two distinct origins: the stellar content and the dark matter content. And, AutoLens allows us to include as many mass components we want.

To describe the stellar mass content, we implement and use the mass profile obtained by the parametrization of the surface brightness distribution using the MGE approach, equation (5.10)[25]. The reduced deflection angle associated with this mass profile is derived

[25] The projected mass profile can be obtained integrating this equation along the line-of-sight (z'-axis), and you will found that the projected mass density is just the projected surface brightness (5.5) multiplied by the mass-to-light ratio (5.8). Just remember that we need to rotate the coordinate system (x, y, z) to the (x', y', z'), with x' along the semi-major axis, before the projection.

in Appendix C, and given by

$$\alpha_{x'}(x',y') = \frac{1}{\pi\, D_L^2 \Sigma_c} \int_0^1 udu \sum_j \frac{M_j}{\sigma_j^2} \frac{x'}{(1-\eta^2u^2)^{1/2}} \exp\left[-\frac{-u^2}{2\sigma_j^2}\left(x'^2 + \frac{y'^2}{1-\eta^2u^2}\right)\right] \tag{5.49}$$

$$\alpha_{y'}(x',y') = \frac{1}{\pi\, D_L^2 \Sigma_c} \int_0^1 udu \sum_j \frac{M_j}{\sigma_j^2} \frac{y'}{(1-\eta^2u^2)^{3/2}} \exp\left[-\frac{-u^2}{2\sigma_j^2}\left(x'^2 + \frac{y'^2}{1-\eta^2u^2}\right)\right], \tag{5.50}$$

where $\alpha_{x'}$ and $\alpha_{y'}$ are the components of the deflection parallel to x' and y', respectively, D_L is the angular diameter distance to the lens in Mpc, $\eta^2 = 1 - q_j'^2$, M_j is the total mass of the j^{th} Gaussian, and Σ_c is the critical surface density. Other constants are the same as defined in (5.5). In both expressions, x', y', and σ_j are given in radians, as well as the deflection angles.

To obtain the above deflection angle the same assumptions for the MGE formalism were made, i.e., we assume that the projected mass density is parametrized by a sum of elliptical concentric Gaussians, centered at the galaxy center (0,0), and semi-major axis parallel to the x'-axis. We recall that these assumptions are important, because we need to guarantee that the image we want to model accomplish these requirements.

For the dark matter content, we choose to parametrize it using the well-known Navarro-Frenk-White (NFW) mass density profile (Navarro; Frenk; White, 1997),

$$\rho(r) = \frac{\rho_s}{(r/r_s)(1+r/r_s)^2} = \frac{\kappa_s \Sigma_c}{r(1+r/r_s)^2}, \tag{5.51}$$

where ρ_s is the characteristic density, r_s is the scale radius, and $\kappa_s = \frac{\rho_s r_s}{\Sigma_c}$.

This profile was primarily obtained from dark matter N-body simulations in the 90s from Navarro, Frenk and White and since then has been used, with its generalization (ZHAO, 1996), for a large range of applications (e.g., Williams; Bureau; Cappellari, 2009; Wyithe; Turner; Spergel, 2001;Nitschai; Cappellari; Neumayer, 2020)

Usually, this profile is characterized by the two constants ρ_s and r_s, often parametrized by (Dúmet-Montoya, 2012)

$$\rho_s = \frac{\Delta\rho_{\rm ref}}{3} \frac{c_\Delta^3}{\left[\ln(1+c_\Delta) - \frac{c_\Delta}{1+c_\Delta},\right]} \tag{5.52}$$

and

$$r_s = \frac{r_\Delta}{c_\Delta} = \frac{1}{c_\Delta}\left(\frac{3M_\Delta}{4\pi\Delta\rho_{ref}}\right)^{1/3} \tag{5.53}$$

where $\rho_{\rm ref}$ is a reference density, r_Δ is the radius that encircles a region with density Δ times greater the reference density, M_Δ is the mass encompassed in r_Δ, and c_Δ is the halo concentration parameter, defined as

$$c_\Delta = \frac{r_\Delta}{r_s}. \tag{5.54}$$

This choice of re-parametrization gets clearer when we look at the mass enclosed within radius r

$$M(r) = \int_0^r 4\pi r' \rho(r') dr' = \frac{4\pi \rho_s r^3}{c_\Delta^3} \left[\ln(1 + c_\Delta) - \frac{c_\Delta}{1 + c_\Delta} \right], \tag{5.55}$$

where we note that the mass M diverges when $r \to \infty$. For that reason, a common approach is to define a reference density $\rho_{\rm ref}$[26], and a "cluster radius" r_Δ as stated above.

The main problem with this approach is that halo mass can evolve with time, which turns these parameters dependent on the mass range, cosmology and redshift (Diemer; More; Kravtsov, 2013; Wyithe; Turner; Spergel, 2001). To avoid such dependencies, we choose to let free the parameters r_s and κ_s in the analysis and modelling that we do in Chapter 6.

Now, coming back to the density profile (5.51), the projected mass density can be obtained integrating this equation along the line-of-slight (Wright; Brainerd, 2000), resulting in

$$\Sigma_{\rm NFW}(R') = 2 r_s \rho_s F(R'), \tag{5.56}$$

where

$$F(R') = \begin{cases} \frac{1}{R'^2 - 1} \left(1 - \frac{1}{\sqrt{R'^2 - 1}} \arctan\left(\sqrt{R'^2 - 1}\right) \right), & \text{if } R' > 1, \\ 1/3, & \text{if } R' = 1, \\ \frac{1}{R'^2 - 1} \left(1 - \frac{1}{\sqrt{1 - R'^2}} \operatorname{arctanh}\left(\sqrt{1 - R'^2}\right) \right), & \text{if } R' < 1. \end{cases} \tag{5.57}$$

and $R' = \xi / r_s = \sqrt{x'^2 + y'^2}/r_s$ is a dimensionless radius in the sky plane.

This implies a deflection angle of (Golse; Kneib, 2002; Bartelmann, 1996)

$$\alpha(R') = 4\kappa_s \frac{g(R')}{R'}, \tag{5.58}$$

[26] Often associated to the critical density of the Universe (Wright; Brainerd, 2000)

where

$$g(R') = \begin{cases} \ln(\frac{R'}{2}) + \frac{1}{\sqrt{R'^2-1}}\arctan\left(\sqrt{R'^2-1}\right) & \text{if } R' > 1, \\ 1 + \ln(\frac{1}{2}) & \text{if } R' = 1, \\ \ln(\frac{R'}{2}) + \frac{1}{\sqrt{1-R'^2}}\operatorname{arctanh}\left(\sqrt{1-R'^2}\right) & \text{if } R' < 1. \end{cases} \tag{5.59}$$

A more realistic model should take into account the possible ellipticity and the orientation of the mass profile, adding two more parameters to the model, the ellipticity itself and a position angle. However, the construction of the elliptical lens models is quite complicated, normally resulting in non-analytical and computationally expensive deflection angles.

Two main approaches in the literature were developed to treat this problem related to elliptical lens models. The more robust one (and more computationally expensive) consists in just change the radius $\xi = \sqrt{x'^2 + y'^2}$ to an elliptical version $\Xi = \sqrt{x'^2 + y'^2/q'^2}$ in the projected mass density, and then calculate the deflection angle using this new elliptical mass profile[27]. This approach gives rise to the so-called elliptical NFW (eNFW) profile. Besides that, the model can then be arbitrarily rotated by the desired position angle.

AutoLens contains an implementation of the eNFW mass model, which takes into account the position angle of the mass profile, the ellipticity, and the other parameters of the NFW profile.

The second approach is called pseudo-elliptical lens models, because the introduction of the ellipticity occurs not in the mass profile, but instead in the lens potential ψ. The great advantage of this class of models is that, if the spherical version of the deflection angle is analytical, then the elliptical one will be too!

However, the pseudo-elliptical models have some domains of validity, and should be used with care. Since this discussion is beyond the scope of this dissertation, and we opt to use the "true" elliptical approach described above (the eNFW profile), we refer the readers to Dúmet-Montoya (2012) and Dúmet-Montoya; Caminha; Makler (2012) for a better explanation on the validity of the models.

After all this, to build a model that takes into account both, stellar and dark matter content, is straightforward. We just need to sum the contributions coming from the equations (5.49), (5.50) and (5.59)[28] for each (x', y') position where we want to compute the model.

[27] This is exactly the same case of MGE lens mass model (see Appendix C), where the deflection angles are not analytical.

[28] Or its elliptical version.

5.4.6 Deflection angle and the connection with PPN approach

Until now, in this Section, we do not consider how the PPN approach changes things. We already have discussed how the lensing phenomena is sensible to the potential Ψ (see Sections 1.2 and 1.4), and how the $\gamma_{\rm PPN}$ appears in the equations (Appendix A). But it is not clear how we actually probe this parameter. We will make this clear now.

Suppose that the deflection angle for our model, which can take into account different mass contributions, is labelled by $\boldsymbol{\alpha}_{\rm GR}$. The subscript GR is to remember that this deflection angle is computed assuming GR, i.e, $\gamma_{\rm PPN} = 1$. Now, assuming that a possible deviation to GR can be accessed by the PPN approach, the deflection angle will be $\boldsymbol{\alpha}_{\rm PPN}$, and it is related with the GR prediction by

$$\boldsymbol{\alpha}_{\rm PPN} = \frac{(1+\gamma_{PPN})}{2}\boldsymbol{\alpha}_{\rm GR}. \tag{5.60}$$

Thus, if GR is correct[29], $\gamma_{\rm PPN} = 1$, and $\boldsymbol{\alpha}_{\rm PPN} = \boldsymbol{\alpha}_{\rm GR}$.

In terms of modelling, we will sample the $\gamma_{\rm PPN}$ simultaneously with the others parameters of the combined model, in order to find the most likely model that reproduces our gravitational lensing and kinematics data.

5.5 Self-consistent mass modelling

In the last Section, we discuss the lens modelling and, in particular, the mass content adopted, including contribution of stellar mass and dark matter. As discussed, taking into account these two different contributions is "easy" for the lens modelling. Once the deflection angle is linear with respect to mass, the contribution coming from different mass profiles can be added directly.

The problem that arises now is how to include the same mass profiles for both models (dynamical and lensing), in a self-consistent way, in order to constrain it by fitting simultaneously the lensing and kinematical data. In other words, we have to include the contribution of dark matter described by the NFW profile, which was already discussed in the lensing context, to the framework of the dynamical modelling presented in Section 5.3. The answer was given in Section 5.2: Parametrizing the NFW mass density profile using the MGE approach!

We can transform the one-dimensional fit of the NFW profile performed by the MGE algorithm into a projected mass density assuming a given ellipticity (or axial ratio). Then, the gravitational potential Φ, used during the JAM modelling (see Section 5.3), will consist of a sum of 2D-Gaussians representing the contributions from the stellar and dark matter contents.

[29] In the sense of best theory today.

This solution, in principle, can also be used for the lens modelling. Instead of use the contribution coming from two different deflection angles (one coming from the stellar content, parametrized by the MGE method, and another coming from the NFW mass model directly), we can use only the MGE parametrization of both mass profiles, and compute the deflection angle only through equations (5.49) and (5.50), with the summation up to $N + M$, where N is the number of Gaussians used to constrain the stellar content, and M the number of Gaussian used to parametrize the dark matter profile.

The latter approach ensures the self-consistency of the mass model, and is the heart of the methodology used in this dissertation, because in this way we are probing the two potentials, Ψ and Φ, at the same time and using the same mass model to describe them.

Sadly, using parametrized versions of both mass profiles results in a lot of Gaussians, increasing drastically the computational time needed to calculate the deflection angle provided by the MGE method. Even optimizing the numerical computation of the integrals of equations (5.49) and (5.50), we need to evaluate them at thousands of points[30], and for a thousands (or even more!) models, since we will sample the parameters in a statistical way. Therefore, we need to improve this calculation, and at the same time, guarantee that the same mass models are being used for lens and dynamical modelling.

Our solution was to use the eNFW profile implemented in AutoLens for the lensing modelling, and use its MGE parametrized version only for the dynamical modelling. This strategy ensures that we will be using the same mass model for both modelling techniques, and improves significantly the processing time required to compute the deflection angle, since the integral involved in the eNFW profile is less expensive to compute numerically than that provided by the MGE method, and the AutoLens is optimized for this kind of calculation.

We take care to look at the error associated with the MGE parametrization of the eNFW profile and how it impacts the deflection angle. After different tests we conclude that the error associated to this parametrization in the deflection angle is of the order of 0.04%. Such small error in the deflection angle shows that using the eNFW profile in detriment of its parametrized version does not represent a major impact on the final result, such that we feel safe to use the MGE parametrized version of the dark matter profile in the dynamic model and its analytical version[31] in the lens model.

As an example, Fig. 17 shows the MGE parametrization of the one dimensional NFW profile. For this example, we considered $\kappa_s = 1.00$, and $r_s = 52708.168$pc.

In Fig. 18 we show a comparison between the deflection angle generated by the

[30] Especially if we consider the ALMA resolution for the lensing data.

[31] Here we are using analytical in the sense of the original profile, not its MGE parametrization. However, we recall that the deflection angle of an eNFW profile is not analytical and demands numerical calculations.

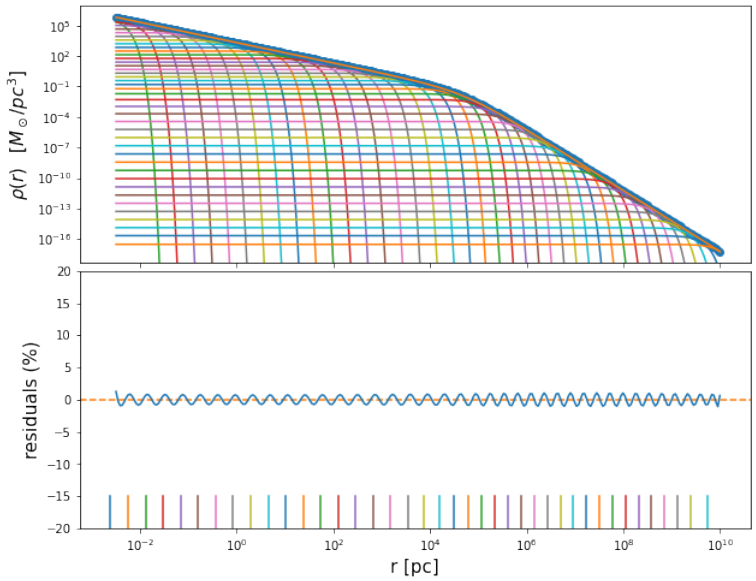
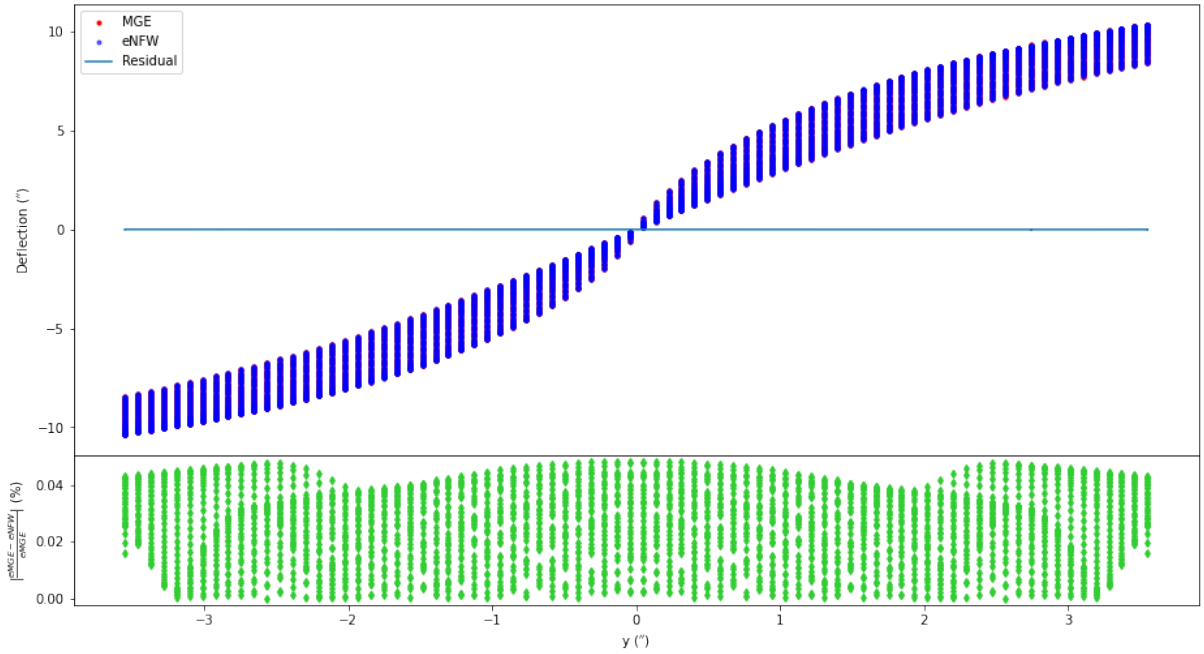
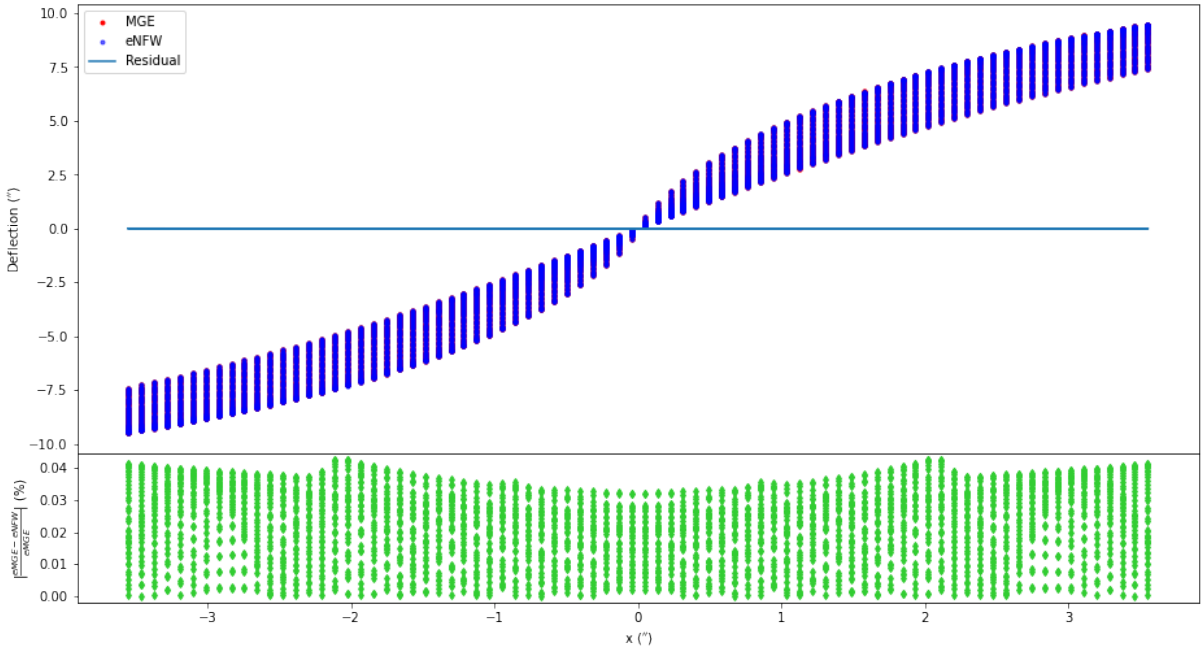


Figure 17 – MGE parametrization of the one-dimensional NFW mass profile for $\kappa_s = 1.00$, and $r_s = 52708.168$pc. A set of 40 Gaussians has been used. Upper panel shows in different colors the Gaussians, in bolded blue the reconstructed NFW profile from MGE parametrization, and in orange (under the bolded blue) the original profile. Lower panel shows the percentage residuals.

MGE parametrization (given by equations 5.49 and 5.50) and the deflection angle generated by the original eNFW profile. To compare, we generate a uniform grid with 80x80 pixels, and pixel scale equal to $0.09''$. Then, we calculate for each point (x', y'), the deflection angle using the eNFW lens mass model (as implemented in AutoLens), and the parametrization above. We set the axial ratio equal to $q' = 0.85$, and the position angle of the lens mass $PA_M = 0°$. The plot shows the deflection angle (α'_x or α'_y) as a function of the angular position, x' or y'. Since the same x' (or y') is associated to more than one y' (or x'), we can see many vertical lines in the plots of the deflection angle in one of the coordinates, formed by many deflections associated to the same point, but this effect is only due to the fact that each panel of Fig. 18 are showing "projections" along x'- or y'-axis.



(a) Deflection angle in y' direction as a function of y' position, for every possible x'.



(b) Deflection angle in x' direction as a function of x' position, for every possible y'.

Figure 18 – Comparison of the deflection angle calculated assuming the MGE parametrization of an eNFW profile (red dots), and the deflection calculated directly from the eNFW profile (blue dots). Due to the small difference, the dots are superposed. The *almost* horizontal line in each plot is the absolute difference, while the bottom panel of each plot shows the relative errors, in percentage (green dots). For more details, see the text.

5.6 Mixing everything with DYNESTY

Finally, we are able to present the statistical method to sample the unknown parameters of our combined model. Discussing all the nuances of statistical modelling is beyond the scope of this work, so we will only present an overview of the method adopted and how we use it.

The reader interested in knowing more about statistical methods applied to Astronomy can learn more through Feigelson; Babu (2012). And an excellent explanation of Markov Chain Monte Carlo (MCMC) methods are found in Hogg; Foreman-Mackey (2017).

5.6.1 An introduction to Bayesian inference faster than light

Bayesian inference is a well-known approach to estimate the distribution of a set of possible parameters, given a model, that consistently reproduces the data. Much of modern Astronomy is based on this approach (Aghanim et al., 2020; Nelson; Ford; Payne, 2014; Kipper et al., 2016; Di Valentino, 2021).

In this context, the Bayesian inference is a way to estimate the posterior probability $P(\boldsymbol{\Theta}|\boldsymbol{D},\boldsymbol{M})$ of a set of parameters $\boldsymbol{\Theta}$ for a given model $\boldsymbol{M}$ conditioned on some data $\boldsymbol{D}$. This information can be accessed through the Bayes Rule

$$P(\boldsymbol{\Theta}|\boldsymbol{D},\boldsymbol{M}) = \frac{P(\boldsymbol{D}|\boldsymbol{\Theta},\boldsymbol{M})P(\boldsymbol{\Theta}|\boldsymbol{M})}{P(\boldsymbol{D}|\boldsymbol{M})}, \tag{5.61}$$

where $P(\boldsymbol{D}|\boldsymbol{\Theta},\boldsymbol{M})$ is the likelihood of the model given the parameters, $P(\boldsymbol{\Theta}|\boldsymbol{M})$ is the prior over the parameters, which quantifies our initial knowledge about them, and

$$P(\boldsymbol{D}|\boldsymbol{M}) = \int_{\boldsymbol{V}_{\boldsymbol{\Theta}}} P(\boldsymbol{D}|\boldsymbol{\Theta},\boldsymbol{M})P(\boldsymbol{\Theta}|\boldsymbol{M})d\boldsymbol{\Theta} \tag{5.62}$$

is the evidence (or marginal likelihood) for the data, given the model. The integral is over the entire parameter volume space $\boldsymbol{V}_{\boldsymbol{\Theta}}$, i.e, over all possible parameter combinations. The evidence allows different model assumptions to be compared through the ratios of evidence values known as Bayes factor[32].

To save notation and keep consistency with what comes next, we will use the same notation as Speagle (2020). So, we rewrite the equation (5.61) as

$$\mathcal{P}(\boldsymbol{\Theta}_{\boldsymbol{M}}) = \frac{\boldsymbol{\mathcal{L}}(\boldsymbol{\Theta}_{\boldsymbol{M}})\boldsymbol{\pi}(\boldsymbol{\Theta}_{\boldsymbol{M}})}{\boldsymbol{\mathcal{Z}}_{\boldsymbol{M}}}, \tag{5.63}$$

where $\mathcal{P}(\boldsymbol{\Theta}_{\boldsymbol{M}}) \equiv P(\boldsymbol{D}|\boldsymbol{\Theta},\boldsymbol{M})$ is the posterior, $\boldsymbol{\mathcal{L}}(\boldsymbol{\Theta}_{\boldsymbol{M}}) \equiv P(\boldsymbol{D}|\boldsymbol{\Theta},\boldsymbol{M})$ is the likelihood, $\boldsymbol{\pi}(\boldsymbol{\Theta}_{\boldsymbol{M}}) \equiv P(\boldsymbol{\Theta}|\boldsymbol{M})$ is the prior, and $\boldsymbol{\mathcal{Z}}_{\boldsymbol{M}} \equiv P(\boldsymbol{D}|\boldsymbol{M})$ is the evidence. When the sub-index

[32] <https://en.wikipedia.org/wiki/Bayes_factor>

$\boldsymbol{M}$ is omitted, it means that we are only talking about a certain model, and not the collection of models used to build the posterior.

For a large range of applicability, the posterior $\mathcal{P}(\boldsymbol{\Theta}_{\boldsymbol{M}})$ often does not have an analytic form, and the only way to treat it is using numerical methods. One of the common approaches to do that consist is trying to estimate the posterior distribution constructing an algorithm that generate a set of samples $\{\boldsymbol{\Theta}_1, \boldsymbol{\Theta}_2, \boldsymbol{\Theta}_3, ..., \boldsymbol{\Theta}_N\}$, which we can use to approximate the posterior as a weighted set of discrete points

$$\mathcal{P}(\boldsymbol{\Theta}) \approx \hat{\mathcal{P}}(\boldsymbol{\Theta}) = \frac{\sum_i^N p(\boldsymbol{\Theta}_{\boldsymbol{i}})\delta(\boldsymbol{\Theta}_{\boldsymbol{i}})}{\sum_i^N p(\boldsymbol{\Theta}_{\boldsymbol{i}})}, \tag{5.64}$$

where $p(\boldsymbol{\Theta}_{\boldsymbol{i}})$ is the importance weight associated with each $\boldsymbol{\Theta}_{\boldsymbol{i}}$ and $\delta(\boldsymbol{\Theta}_{\boldsymbol{i}})$ is the Dirac delta function located at $\boldsymbol{\Theta}_{\boldsymbol{i}}$.

In the literature there are a lot of algorithms that perform this kind of sampling. Among them the most common are based on the MCMC methods, with well-known implementations, such as EMCEE by Foreman-Mackey et al. (2013), and PyMC3 by Salvatier; Wiecki; Fonnesbeck (2016). The MCMC-based method usually tries to generate samples "proportional" to the posterior, such that $p_i = 1$, and for that reason computing the evidence (and therefore performing model comparison) is a hard task for them. In addition to that, MCMC-based methods often are very sensitive[33] to the initial position where we start the exploration of the parameter space, tending to struggle when the posterior is composed of widely-separated modes.

One alternative to those samplers, is the Nested Sampling algorithm, introduced by Skilling (2006). Nested Sampling methods belong to a class of samplers that tries to estimate the posterior and the evidence jointly, solving some of the issues present in MCMC based methods.

According to Josh Speagle[34], Nested sampler also *"has a variety of appealing statistical properties, which include:*

- *well-defined stopping criteria for terminating sampling,*
- *generating a sequence of independent samples,*
- *flexibility to sample from complex, multi-modal distributions,*
- *the ability to derive how statistical and sampling uncertainties impact results from a single run, and*
- *being trivially parallelizable."*

[33] In the sense of spending time to reach the convergence, since the all ideal MCMC-based method should find the same solution if enough samples are generated.

[34] <https://dynesty.readthedocs.io/en/latest/overview.html>

For all these reasons we choose to use it for sampling the parameters of the posterior of our combined model. The next Section is dedicated to present the main ideas behind the Nested Sampling algorithm.

However, before we continue, we need to point two more comments. First, sampling algorithms are designed to sample things, not to find the best solution. The best solution is spotted by optimizers, not samplers. Sampling algorithms tell us the most probable set of parameters that reproduces the data, given a model and a prior.

Second, different sampling methods should be used for different strategies and problems. *"There is no One True Method to Rule Them All [...]"*[35]. So as advice, we suggest that, whenever possible, do tests with different methods, to ensure that the best strategy is being followed.

5.6.2 Nested Sampling Method

The Nested Sampling method tries to sample not the posterior itself, but the evidence $\mathcal{Z}$, instead. This can be done generating samples in nested "shells" of increasing likelihood, turning possible to estimate the evidence for complicated distributions. Then the final set of samples can be combined with their associated importance weights p_i to generate the associated posterior.

For the Nested Sampling, the problem can be solved applying the following procedure described in Speagle (2020):

1. *"slicing the posterior into many simpler distributions,*
2. *sampling from each of those in turn, and*
3. *re-combining the results afterwards."*

The problem then is how to efficiently estimate the complex integral involving the evidence $\mathcal{Z}$

$$\mathcal{Z} = \int_{\boldsymbol{V}_{\boldsymbol{\Theta}}} \boldsymbol{\mathcal{L}}(\boldsymbol{\Theta}_{\boldsymbol{M}})\boldsymbol{\pi}(\boldsymbol{\Theta}_{\boldsymbol{M}})d\boldsymbol{\Theta}. \tag{5.65}$$

This problem is solved by the Nested Sampling transforming the integral over all the possible values of $\boldsymbol{\Theta}_{\boldsymbol{M}}$ into a one-dimensional integral over the prior volume X of the enclosed parameter space

$$\mathcal{Z} = \int_{\boldsymbol{V}_{\boldsymbol{\Theta}}} \boldsymbol{\mathcal{L}}(\boldsymbol{\Theta})\boldsymbol{\pi}(\boldsymbol{\Theta})d\boldsymbol{\Theta} = \int_0^1 \mathcal{L}(X)dX. \tag{5.66}$$

[35] Speagle (2020)

In this equation, $\mathcal{L}$ defines an iso-likelihood contour that corresponds to the boundary of the volume X. The prior volume X, defined by

$$X(\lambda) = \int_{\boldsymbol{V}_{\boldsymbol{\Theta}}:\mathcal{L}\geq\lambda} \boldsymbol{\pi}(\boldsymbol{\Theta}_M) d\boldsymbol{\Theta}, \tag{5.67}$$

represents the fraction of the prior where the likelihood $\boldsymbol{\mathcal{L}}(\boldsymbol{\Theta})$ is greater than some threshold λ. The integration limits are due to the fact that the prior is normalized, so $X(\lambda = 0) = 1$ and $X(\lambda \to \infty) = 0$.

A schematic representation of this "slicing" is shown in Fig. 19. The left plot shows the iso-likelihood contours of a two-dimensional multi-modal (modes 4 and 5) likelihood function $\boldsymbol{\mathcal{L}}(\boldsymbol{\Theta})$. The contours are represented by different colors, and enclose a different fraction of the total prior volume. The right panel shows the likelihood as a function of the prior volume X, that encloses some fraction of the total prior volume. In this situation, the evidence is just the area under the curve, which is straightforward to evaluate.

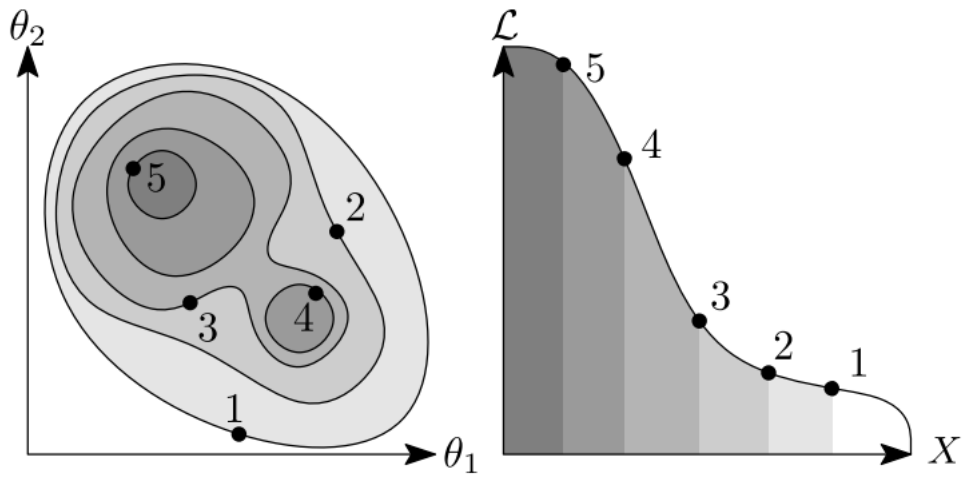


Figure 19 – Representation of the "slicing" in the total prior volume. The left panel shows the iso-likelihood contours represented by different colors, and enclosing a different fraction of the total prior volume. The right panel shows the likelihood as a function of the prior volume X. The evidence is the area under the curve. For more details, see the text. Image taken from Handley; Hobson; Lasenby (2015).

Once a final set of samples $\{\boldsymbol{\Theta}_1, \boldsymbol{\Theta}_2, \boldsymbol{\Theta}_3, ..., \boldsymbol{\Theta}_N\}$ are drawn, the one-dimensional evidence integral can be evaluated by standard numerical techniques, e.g., a Riemann sum

$$\hat{\mathcal{Z}} = \sum_i^N \mathcal{L}(\boldsymbol{\Theta}_i) \times (X_i - X_{i-1}) \equiv \sum_i^N \hat{p}(\boldsymbol{\Theta}_i), \tag{5.68}$$

where the hat notation is used to reminder that this is a noisy estimation, and $\hat{p}(\boldsymbol{\Theta}_i)$ is the estimated importance weight. These values can also be used to approximate the posterior (5.64):

$$\hat{\mathcal{P}}(\boldsymbol{\Theta}) = \frac{\sum_i^N \hat{p}(\boldsymbol{\Theta_i})\delta(\boldsymbol{\Theta_i})}{\sum_i^N \hat{p}(\boldsymbol{\Theta_i})}. \tag{5.69}$$

5.6.2.1 Sampling from the prior volume

As we see above, once we have the final set of samples, evaluating the evidence (and through it the posterior) is a straightforward task. The problem then becomes how to generate samples from the prior $\boldsymbol{\pi(\Theta)}$ subject to the likelihood constraint λ. One alternative is to just perform a simple sampling rejection algorithm, i.e., at a given iteration i generate samples $\boldsymbol{\Theta}_{i+1}$ from the prior $\boldsymbol{\pi(\Theta)}$ until $\mathcal{L}(\boldsymbol{\Theta}_{i+1}) \geq \mathcal{L}(\boldsymbol{\Theta}_i)$. However, this procedure becomes very inefficient as the sampling goes on, and the prior volume X compresses. Therefore, to improve the overall efficiency, Nested Sampling algorithm needs a way to generate samples directly from the constrained prior

$$\boldsymbol{\pi}_\lambda(\boldsymbol{\Theta}) \equiv \begin{cases} \boldsymbol{\pi(\Theta)}/X(\lambda), & \mathcal{L}(\boldsymbol{\Theta}) \geq \lambda \\ 0, & \mathcal{L}(\boldsymbol{\Theta}) < \lambda \end{cases} \tag{5.70}$$

Since the parameters we want to sample can assume a wide range of values, the prior $\boldsymbol{\pi}_\lambda(\boldsymbol{\Theta})$ can vary drastically from one position to another, making it hard to sample the parameters from the constrained prior. However, if the prior is a standard uniform (flat prior) for all the dimensions, the density constrained and interior to λ is constant, then the volume X is well behaved. Then, using an appropriate "prior transform" function τ, we are able to map a set of parameters $\boldsymbol{\mathcal{U}}$ with a uniform prior over the D-dimensional unit cube to the physical parameters $\boldsymbol{\Theta}$ of interest. Therefore, applying the transformation, the initial problem is reduced to the much more simple problem of sampling uniformly the transformed constrained prior

$$\boldsymbol{\pi}'_\lambda(\boldsymbol{\mathcal{U}}) \equiv \begin{cases} 1/X(\lambda), & \mathcal{L}(\boldsymbol{\Theta}) = \tau(\boldsymbol{\mathcal{U}}) \geq \lambda \\ 0, & \mathcal{L}(\boldsymbol{\Theta}) = \tau(\boldsymbol{\mathcal{U}}) < \lambda \end{cases} \tag{5.71}$$

To save notation, and avoid possible confusions, throughout the rest of the text we will assume $\boldsymbol{\pi}(\boldsymbol{\Theta})$ is a unit cube prior, unless otherwise specified.

5.6.2.2 Nested Sampling procedure

Now we have all the ground prepared to discuss the Nested Sampling procedure more heuristically.

Suppose that initially we have a prior volume X without any constraints. A set of N samples $\{\boldsymbol{\Theta}_1, \boldsymbol{\Theta}_2, \boldsymbol{\Theta}_3, ..., \boldsymbol{\Theta}_N\}$ is drawn from the initial unconstrained unit cube. Each of these samples corresponds to a likelihood $\{\boldsymbol{\mathcal{L}}(\boldsymbol{\Theta}_1), \boldsymbol{\mathcal{L}}(\boldsymbol{\Theta}_2), \boldsymbol{\mathcal{L}}(\boldsymbol{\Theta}_3), ..., \boldsymbol{\mathcal{L}}(\boldsymbol{\Theta}_N)\}$.

At iteration i the algorithm selects the worst sample, i.e., the sample with the lower likelihood $\boldsymbol{\mathcal{L}}_i^{\min}$ value and highest prior volume, and store it. This sample is now called dead point, and the remaining points are called live points. Now, the dead point is replaced by a new point $\boldsymbol{\Theta}'$, drawn by the constrained prior $\boldsymbol{\pi}_\lambda$, such that $\boldsymbol{\mathcal{L}}_{i+1}(\boldsymbol{\Theta}') \geq \lambda$.

By construction, Nested Sampling algorithm ensures that $\boldsymbol{\mathcal{L}}_{i+1}(\boldsymbol{\Theta}') \geq \boldsymbol{\mathcal{L}}_i^{\min}$, and $X_{i+1} \leq X_i$. Finally, repeatedly sampling in this way, we are able to find the most probable regions where the parameters reproduce the data more confidently.

Figure 20 illustrates this procedure for a one dimensional case. Initially three samples are drawn for the unconstrained prior volume $X = [0, 1]$ (the bottom of the figure). Then, the sample with the lowest likelihood is removed from the set of live points, and added to the list of dead points. Then a new point is randomly drawn from the new prior volume $[0, X_1]$ (step 1), and with the two survivors forms a new set of live points. Now we repeat the procedure, removing the lowest sample of the current set of live points, and replacing it by a new point drawn from the reduced range $[0, X_2]$ (step 2). At the end of the algorithm, the procedure generates a set of eight samples, where three of them are live points, and five are dead points. These points are used to evaluate the sum (5.68), and estimate the evidence $\hat{\mathcal{Z}}$.

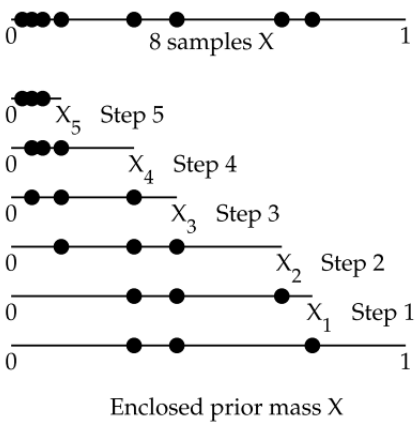


Figure 20 – Nested Sampling for five steps with a collection of three points. At each interaction, the lowest sample is replaced by a new one, drawn for the constrained prior volume (here called prior mass). Image taken from Skilling (2006).

5.6.2.3 Terminator Criterion

One of the advantages of Nested Sampling is the well-defined stop criterion, usually not available for MCMC-based methods. We can choose to stop the sampling when the remaining live and the dead points give us an integral that encompasses most of the

posterior. At this point, adding more live points no longer contributes significantly to the evidence, and we can stop the algorithm. Mathematically, this can be put as follows

$$\Delta \ln \hat{\mathcal{Z}} \equiv \ln (\hat{\mathcal{Z}}_i + \Delta \hat{\mathcal{Z}}_i) - \ln (\hat{\mathcal{Z}}_i) < \epsilon, \tag{5.72}$$

where $\Delta \hat{\mathcal{Z}}_i$ is the estimated remaining evidence, and ϵ determines the tolerance. The remaining evidence at iteration i is usually estimated constructing a strict upper limit, such as (see Speagle, 2020)

$$\hat{\mathcal{Z}}_i \leq \mathcal{L}_i^{\max} \hat{X}_i, \tag{5.73}$$

where $\mathcal{L}_i^{\max}$ is the maximum value of the likelihood at iteration i among the remain live points inside the estimated prior volume $\hat{X}_i$.

5.6.2.4 Nested Sampling Errors

All the sampling methods are subject to an error associated with the fact that we are approximating a continuous distribution by a set of discrete and finite number of samples. This error usually is called "sampling noise". However, we expect that this error becomes negligible as the number of samples[36] and interactions increase.

Unfortunately, Nested Sampling contains an additional source of error, associated with the fact that the algorithm uses noisy estimators for the prior volume, $X \to \hat{X}$. This error is translated to the evidence through the importance weight (eq. 5.68), raising an "statistical noise".

This discussion is far beyond the scope of this work, and we will not cover it here. But it is important to highlight this issue, and make clear that the final inference of Nested Sampling does take into account these two sources of errors. A fully discussion can be found at Higson et al. (2018b), and for a quick overview, refer to Speagle (2020).

5.6.3 DYNESTY

To explore the parameter space of our model and sampling the posterior, we use the public, open-source package DYNESTY (Speagle, 2020), a Python implementation of the Nested Sampling algorithm. DYNESTY present a lot of facilities, including a vast documentation[37], and an adaptive version of the Nested Sampling, known as Dynamic Nested Sampling (Higson et al., 2018a). In this dissertation we choose to work with the classical version presented earlier, also called Static Nested Sampling.

DYNESTY offers a large range of options for how the prior volume should be constrained, and how the new points should be proposed. The idea is to improve the

[36] Live points in Nested Samples methods, and walkers in MCMC methods.
[37] <https://dynesty.readthedocs.io/en/latest/#>

efficiency using different strategies depending on the problem we are dealing with. The former case is related to how the bounding distributions are constructed, and the second is related to the sampling method used to propose the new live points inside that constrained prior volume.

We will not discuss all the available options here and we limit ourselves only to present the combination used in this work. For more details, see Sections 4.1 and 4.2 of Speagle (2020), and the references therein.

5.6.3.1 Bounding Distribution

Exploring the entire unit cube can be very inefficient, especially in the cases where the posterior is multi-modal and the algorithm can be stuck in a local minimum. One alternative for such distributions is using a set of ellipsoids that can possibly be overlapped. These ellipsoids are constructed from the remaining live points using an interactive clustering scheme (Feroz et al., 2019).

Briefly, at the first iteration the entire collection of points are proposed uniformly over the entire unit cube, then the lowest likelihood point is storage as dead point. Next, the algorithm builds a bounding ellipsoid over the entire collection of the remaining live points. DYNESTY uses a scaled version of the empirical variance matrix of the current live points, centered at the empirical mean, to determine the size and shape of the ellipsoid (Mukherjee; Parkinson; Liddle, 2006). The scale is such that the ellipsoid encompasses all the available live points.

After constructing this ellipsoid, a K-means clustering with K= 2 is initialized at the endpoints of the major axes. The live points then are assigned to one of the clusters, and a new pair of bounding ellipsoids is constructed, one for each of the clusters. The resulting two ellipsoids are accepted if the combined volume of them is substantially smaller. This process continues recursively, until no decomposition is accepted. DYNESTY tries to control the decomposition of the live points, being more conservative in the clustering step, in order to avoid the *"shredding the posterior into many tiny islands of isolated live point clusters"* (Speagle, 2020).

This particular kind of Bounding Distribution allows the exploration of the multi-modal regions, without the issue of being stuck in a local minimum.

5.6.3.2 Sampling Proposal

Once we have determined the new constrained prior volume, we need to replace the dead point (or dead points) by new ones, restrict to $\boldsymbol{\mathcal{L}}_{\boldsymbol{\Theta}'} \geq \boldsymbol{\mathcal{L}}_i^{\min}$, where $\boldsymbol{\Theta}'$ is the proposal sample.

One useful approach for sampling the new points restricted to the bounding

distribution is to propose a new point by evolving a given live point $\boldsymbol{\Theta}_k \rightarrow \boldsymbol{\Theta}'$. The most straightforward approach is sampling the new point from the constrained prior using the Metropolis-Hastings MCMC algorithm. The algorithm is as follows:

1. Drawn the new sample $\boldsymbol{\Theta}'$ from the proposal distribution $Q(\boldsymbol{\Theta}'|\boldsymbol{\Theta}_k)$, centered in $\boldsymbol{\Theta}_k$;
2. Move to $\boldsymbol{\Theta}'$ with probability $A = \min\left[1, \frac{\pi_\lambda(\boldsymbol{\Theta}')}{\pi_\lambda(\boldsymbol{\Theta}_k)} \frac{Q(\boldsymbol{\Theta}_k|\boldsymbol{\Theta}')}{Q(\boldsymbol{\Theta}'|\boldsymbol{\Theta}_k)}\right]$. Otherwise, stay at $\boldsymbol{\Theta}_k$;
3. Repeat previous steps for N_{walks} interactions.

This procedure is called random walk.

However, since the constrained priors are uniform by construction, the ratio of the priors is always equal to one. And, if the proposal distribution is symmetric, i.e., $Q(\boldsymbol{\Theta}'|\boldsymbol{\Theta}_k) = Q(\boldsymbol{\Theta}_k|\boldsymbol{\Theta}')$, their ratio is equal to one either. Then, a new point is accepted if it is within the current constrained prior that we are sampling from[38], and rejected otherwise.

For DYNESTY, the proposal distribution $Q(\boldsymbol{\Theta}'|\boldsymbol{\Theta}_k)$ is a simple uniform distribution associated to the ellipsoid centered on $\boldsymbol{\Theta}_k$ and covariance C_b, where C_b is one of the bounding distributions that encompasses $\boldsymbol{\Theta}_k$, randomly selected (when $\boldsymbol{\Theta}_k$ belongs to more than one bounding distributions).

5.6.3.3 Parallelization

Given the computationally expensive likelihood of our model, we are forced to parallelize our script and adapt it to a high-performance computing (HPC) environment. DYNESTY is easily parallelizable, and also uses several built-in functions that make the adaptation friendly.

The principal advantage of the parallelization is to be able to compute individual likelihood (for the same interaction) in independent cores at same time. The computing time does not scale linearly with the number of cores, however using many cores in parallel processing improves a lot the overall computing time, making this project possible.

To parallelize our script we use the open-source SCHWIMMBAD (Price-Whelan; Foreman-Mackey, 2017) package, "that provides a uniform interface to parallel processing pools and deployment on a cluster or supercomputer (via, e.g., MPI or JobLib)." With SCHWIMMBAD, we are able not just to spread the calculation of the live points through multiple CPUs, but also over multiple nodes in a HPC environment.

[38] Remember that we have a noisy estimation of the prior volume, so we need to take care and accept only the samples with likelihood higher than the dead point discarded in the former iteration.

A bottleneck in the parallelization process is that each interaction is limited by the slowest-performing member of interaction. This occurs because we need to know all the values of the likelihood at each interaction before determining the dead point.

Finally, with our code parallelized, and the data properly reduced, we are able to perform the combined modelling! In the next Section, we provide a "Proof of Concept" of our code, to make sure that it is able to explore the real data of the SDP.81 system. Then, in the next Chapter we present the Results of our analysis.

Before we move on, we want to highlight that the majority of the tests performed to get the best version of the script and the final run used for the Results of this work was only possible thanks to the CHE[39], SDumont[40,41] and ADA-Lovelace[42] clusters.

5.6.4 The Combined Statistical Model

In the last Section, we discuss all the numerical methods that are applied to our data, and at the end we present the statistical framework we employ in order to infer the most probable value of the parameters of our model, especially $\gamma_{\rm PPN}$. However, to consider the work complete, we need to present one last detail related to how our posterior was defined. The posterior, in turns, is a way to summarize what we know about the problem after the data has been observed and a model was proposed.

So, in order to define our posterior, we need to fix a model, and a metric, that allows us to compare the model with the real observed data. This is what the likelihood does in a Bayesian framework.

Since the lensing data are independent of the dynamical data, we can consider the individual likelihoods of each model independently, such that we can produce a single likelihood for the combined modelling multiplying the individual likelihoods, it means

$$\mathcal{L}_{\rm Model} \equiv \mathcal{L}_{\rm Lens} \times \mathcal{L}_{\rm Dyn}, \tag{5.74}$$

where $\mathcal{L}_{\rm Lens}$ is the likelihood of the lens model, quantified by ϵ, equation (5.48), and $\mathcal{L}_{\rm Dyn}$ is the likelihood of the dynamical model, quantified by a χ^2 statistic between the observed $V_{\rm rms}$ map and the JAM model, described in Section 5.3.

With this likelihood, and some priors over the parameters, we finally[43] are ready to sample the posterior!

[39] CBPF
[40] LNCC
[41] The project was selected by a public call for the Standard allocation program, with 2073600 UA. More information at <https://sdumont.lncc.br/alloc.php?pg=alloc#>. See also our abstract at <https://sdumont.lncc.br/projects_view.php?pg=projects&status=ongoing>, under the name "Testando a Relatividade Geral em escala de galáxias".
[42] UFRGS
[43] We swear!

Just to summarize, here we highlight the possible parameters that our model can take into account:

- Υ_k: mass-to-light ratio[44]
- i: galaxy inclination
- β_z: luminosity-weighted anisotropy[45]
- $M_{\rm SMBH}$: mass of a central supermassive black hole
- κ_s: scale parameter of the NFW dark matter mass model
- r_s: the scale radius of the NFW dark matter mass model
- $q_{\rm DM}$: deprojected[46] axial ratio of the dark matter halo
- $\text{shear}_{\rm mag}$: lens external shear magnitude
- shear_{ϕ}: lens external shear angle, counterclockwise from the projected x-axis.
- $\gamma_{\rm PPN}$: Parametrized Post-Newtonian parameter.

5.7 Proof of Concept

To prove our implementation and ideas, in this Section we provide a self-consistent test using simulated kinematical and lensing data, for which we know the input parameters of the model used to create such simulation. The main goal is, given some simulated data (interpreted as the observed data), to sample the parameter space looking for the most probable model that reproduces our simulated data.

In order to create the simulated data, we adopt the following procedure:

1. We parametrize a eNFW profile using MGE method, assuming: $z_l = 0.299$, $\kappa_s = 0.075$, $q_{\rm DM} = 0.85$ and $r_s = 18'' \sim 82500$pc at the distance of the object at z_l;
2. Simulate a $V_{\rm rms}$ map, using the real SDP.81 surface brightness MGE decomposition[47], assuming a constant $\Upsilon = 7.20 {\rm M}_\odot/{\rm L}_\odot$, and the dark matter profile above. We assume a inclination of $i = 90°$, $\beta_z = 0.0$ constant, and the contribution of a supermassive black hole with total mass $M_{\rm SMBH} = 10^9 {\rm M}_\odot$, modeled as a Gaussian with dispersion equal to $0.01''$ at the galaxy center. We assume a pixel scale for the kinematic map equal to $0.2''$, same as the one of MUSE data used in this work;

[44] Can be a constant value, or one per Gaussian component k (free or correlated).
[45] Can be a constant value, or one per Gaussian component.
[46] It means, the real axial ratio, not the projected in the plane of the sky.
[47] More details see Chapter 6.

3. The kinematic map was convolve with the MUSE PSF(See Section 4.2), and we add a Gaussian noise (dispersion equal to 5% of the value) to each point of the map. Figure 21 shows the final map, used as real data.

4. Using a grid with 100×100 pixels, and pixel scale of $0.1''$, we simulate a source galaxy at $z_s = 4.100$ using an Elliptical Sérsic light profile[48] (see Figure 22).

5. We build a lens mass profile using the same mass model used for the kinematic map. In addition, we set an external shear with $\text{shear}_{\text{mag}} = 0.01$, and $\text{shear}_{\phi} = 88°$. We keep $\gamma_{\text{PPN}} = 1$.

6. Using the lens mass profile we ray-trace the emission of the source galaxy with the lens equation to obtain the gravitational arcs in the lens plane. This image is then convolved with a PSF with FWHM equal to $0.1''$, and added a constant noise of 0.1 counts in each pixel. The final image of the deflected source is shown in Figure 23.

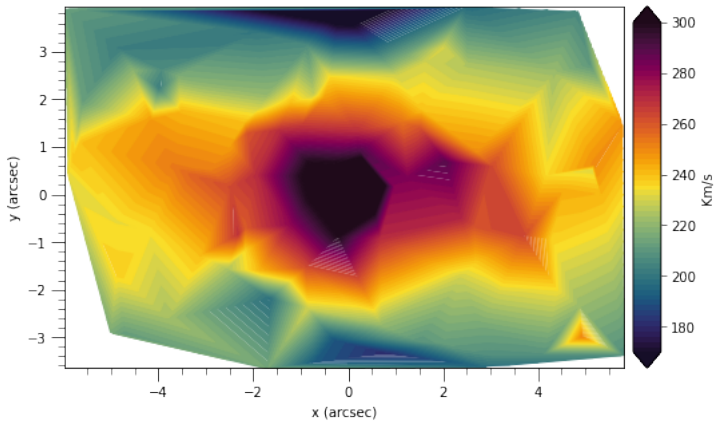


Figure 21 – Simulated V_{rms} map used as the real data in our proof-of-concept test.

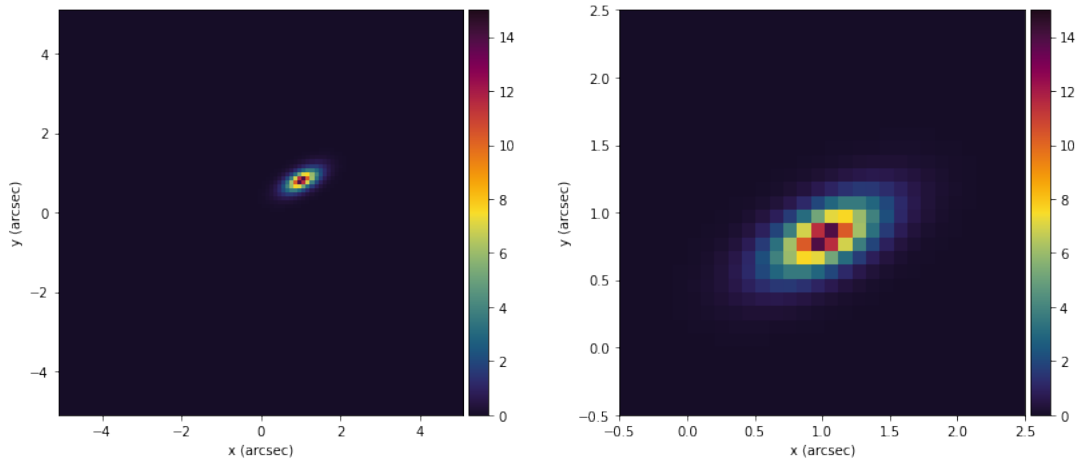


Figure 22 – Elliptical Sérsic source galaxy. The left panel shows the entire pixel grid used for the simulation, and the right panel shows a zoom in the galaxy.

[48] Since we are using the SLI method, the parameters of the source light profile are not of our main interest. But, you can check the parameters used for the simulation it in our GitHub.

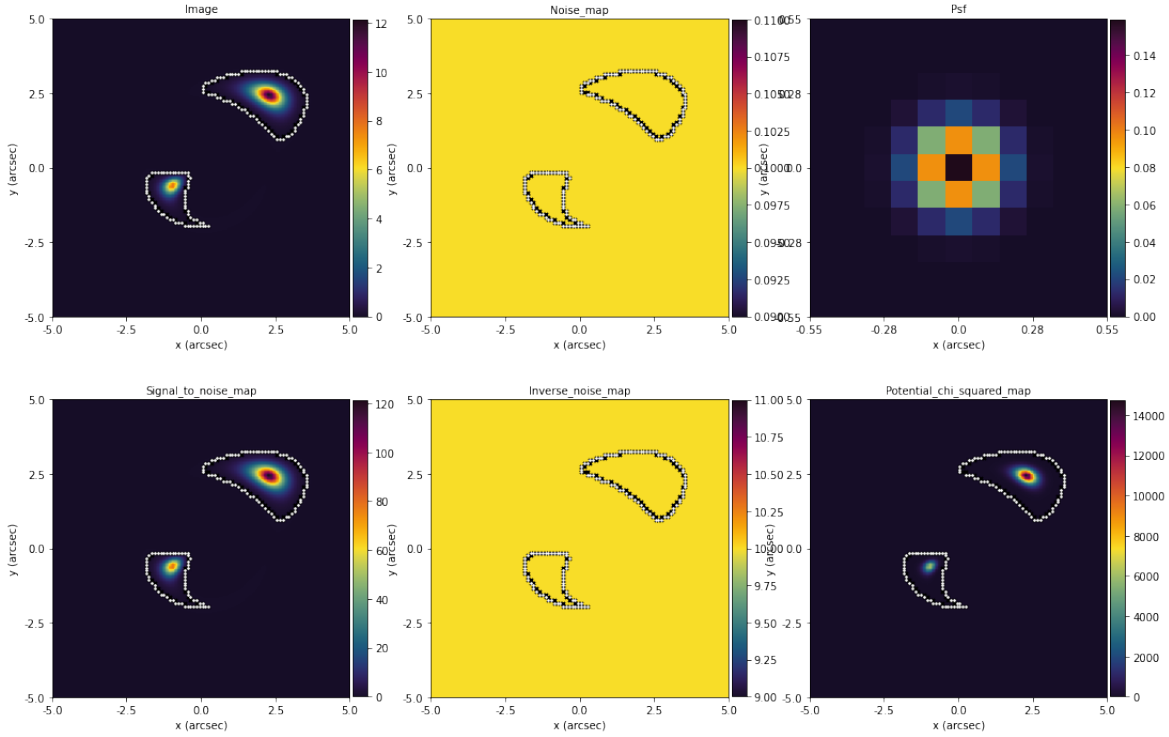


Figure 23 – Emission of the deflected background galaxy in the lens plane. From top left from bottom right we have: The deflected source, and in white a mask selecting only the arcs. The noise map. Image of the PSF used. Signal-to-noise map of the arcs. Inverse Noise Map. The potential χ^2 map of the imaging, which represents how much each pixel can contribute to the χ^2 map, assuming the model fails to fit it.

Now, using the methodology described in the last Sections, and the theoretical framework presented in this dissertation, we will try to find the most probable model for these simulated data. However, we need to make some disclaimers. First, our simulation has the sole purpose of validating the method, and should not be seen as a systematic test. Although, a systematic test using different simulations generated with different model parameters would be desirable, the processing time required to run such test was prohibitive, given the duration of this Masters project. Second, we introduce the error and the PSF convolution in the final images, not through small perturbations on the real parameters, so that we are not able to verify the real impact of this modification on the real parameters used. That said, our evaluation metric here will be to look at the final models and see how well they fit the input data.

In order to do that, we use DYNESTY to sample the parameter space, using the configuration presented in Section 5.6. We set a number of live points equal to $n_{\text{live}} = 100$, and the stop criteria equal to $\Delta \ln \hat{\mathcal{Z}} < 0.01$. The source-grid reconstruction is a rectangular grid with 40×40 pixels, and the regularization constant was fixed equal to 1.0.

To improve the computation time for the modelling, we fixed the dark matter

scale radius r_s during the sampling, so that our model has nine free parameters: (Υ, β_z, i, $M_{\rm SMBH}$, κ_s, $q_{\rm DM}$, shear$_{\rm mag}$, shear$_\phi$, $\gamma_{\rm PPN}$).

Instead of sampling the inclination directly, we choose to sample it from equation (5.7), in order to obtain a physical intrinsic axial ratio. This choice also avoids numerical problems during the deprojection step in the JAM modelling. The inclination can be inferred from the sampled axial ratio $q_{\rm inc}$, inverting (5.7)

$$i = \arctan\left(\sqrt{\frac{1 - q'^2_{\rm min}}{q'^2_{\rm min} - q^2_{\rm inc}}}\right), \tag{5.75}$$

where $q'^2_{\rm min}$ is the minimum value of the stellar projected axial ratio. In the same way, we sampled the logarithm of the mass of the supermassive black hole, due to the order of magnitude involved.

We also assume that the mass profile follows the light position angle, then our elliptical mass profile (stellar + dark matter) has the same position angle of the light profile, which is equal to zero, once we construct the profile using the MGE method, that assumes the semi-major axis coincident with $x'-$axis.

Table 1 shows the priors adopted for this modelling. We assume flat (non-informative) prior for all the parameters, considering the priors intervals as containing the typical values for the parameters found in the literature as well as the input values used in the simulation, except for the Post-Newtonian parameter, which we assume a Gaussian prior centered in 1.0 (GR) and dispersion equal to 0.05. This assumption is based on the CT18' result, and the belief of deviations of GR should be small.

Parameter	Prior
Υ	$\mathcal{U}$(0.5, 15)
β_z	$\mathcal{U}$(-3, 3)
$q_{\rm inc}$	$\mathcal{U}$(0.055, $q'_{\rm min}$)
$\log M_{\rm SMBH}$	$\mathcal{U}$(7, 11)
κ_s	$\mathcal{U}$(0, 1)
$q_{\rm DM}$	$\mathcal{U}$(0.1, 1)
shear$_{\rm mag}$	$\mathcal{U}$(0, 0.1)
shear$_\phi$	$\mathcal{U}$(0°, 179°)
$\gamma_{\rm PPN}$	$\mathcal{G}$(1.0, 0.05)

Table 1 – Priors of the simulated data. $\mathcal{U}$(lower, upper) represents a non-informative prior, with lower and maximum values accepted. $\mathcal{G}$(mean, sigma) is a Gaussian prior, with mean and dispersion values. All the values are in appropriated physical units.

The results of the kinematical modelling can be seen in Fig. 24. The top left panel shows the data, and the top right panel shows a symmetrized version of the data. This symmetrized map helps the visual inspection, once our assumed model (JAM) is

axisymmetric. The bottom left panel shows the most probable model inferred, and the bottom right shows the normalized error. Excepting some outliers in the edges, the model reproduces the data very well, with an error less than 10%.

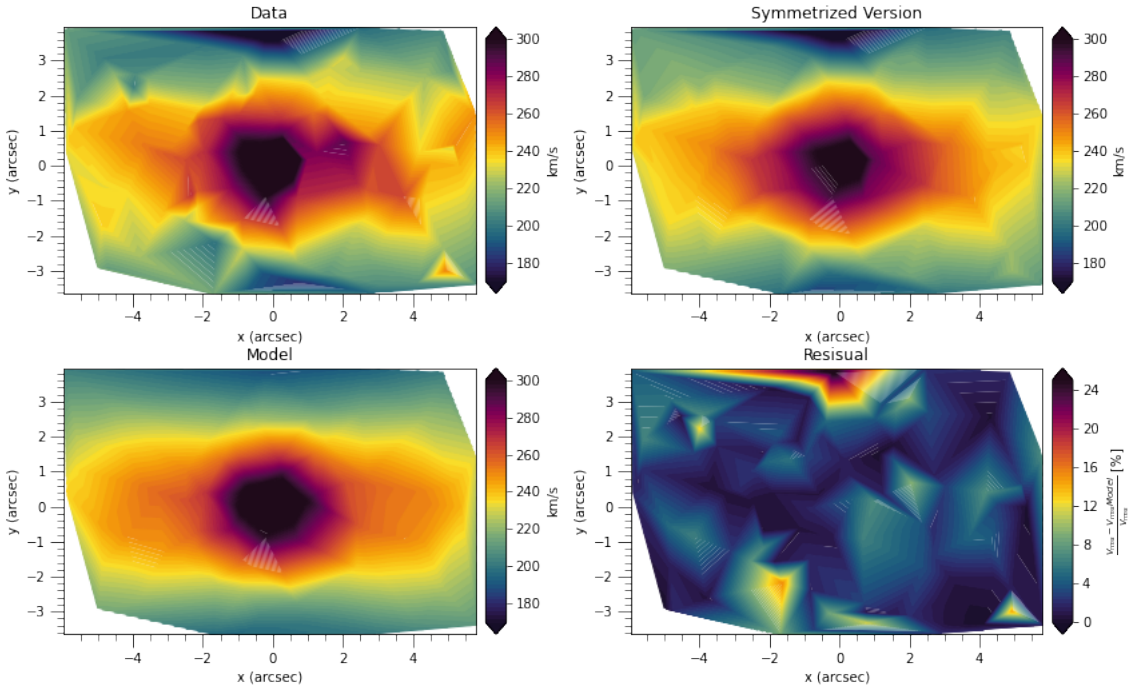


Figure 24 – JAM modelling result of simulated data. Top left is the data, and top right the symmetrized version of the data, to helps the visual inspection. Bottom left is the model, and bottom right the normalized error.

Figure 25 shows the model of the gravitational arcs. On the upper panel we have the data, the signal-to-noise map, and the most probable model. Note that these plots show only the region used for the modelling, i.e., the region inside the mask (black contour) where the data used for modelling the lens are located. The bottom panel shows, from left to right, the residual map, the normalized residual map by the noise, and the χ^2 map. Once again, we can see that the most probable model fits the data very well.

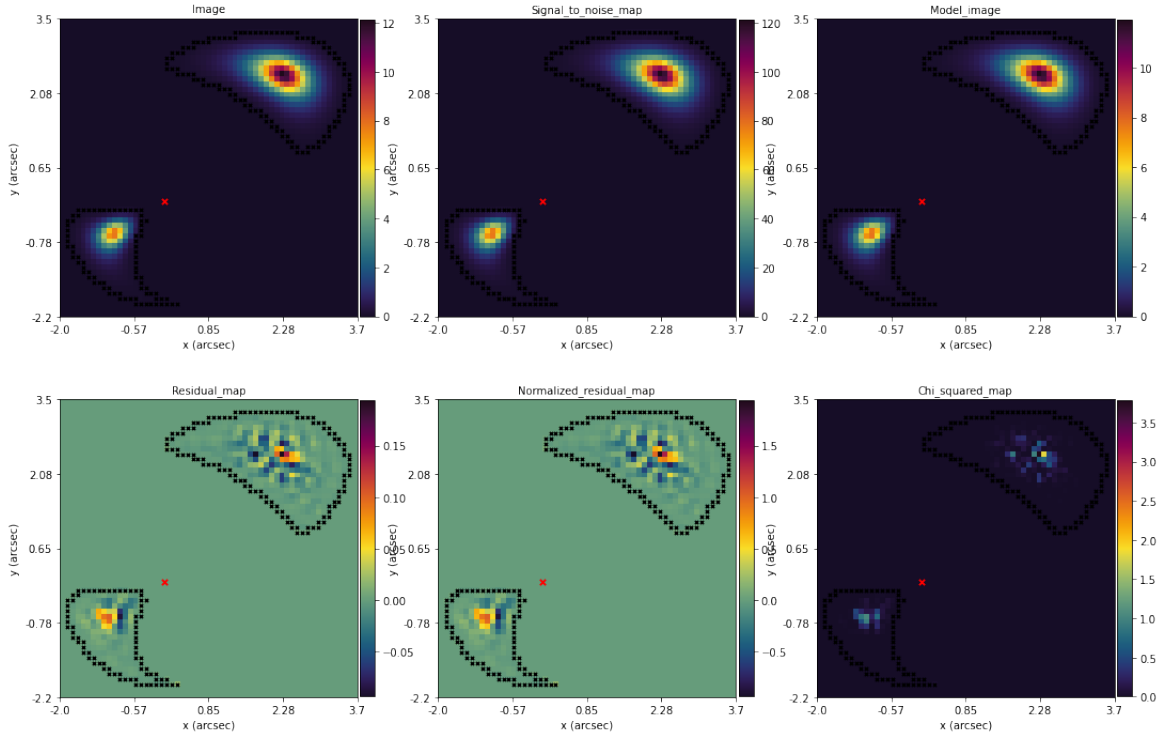


Figure 25 – Lens modelling results. Top panel from left to right we have: Data, signal-to-noise map, and most probable model. Bottom panel shows: residual map (left), normalized residual map (center), and χ^2 map (right).

Finally, Fig. 26 shows the lens model, and the source reconstruction. The central top panel shows the source reconstruction[49], and the white dots are the positions where the mask contours are traced in the source plane. The residual map is the difference between the data and the reconstructed source, when the data is mapped in the source plane using the most probable model. The regularization weight map shows that we use a constant value for all pixels (which could be different, if we use other regularization schemes or source pixelization methods, see Nightingale; Dye; Massey, 2018).

[49] Compare this wonderful reconstruction with the simulated source Fig. 22, and keep in mind that we do not make any assumptions about this surface brightness of morphology. Although this is a surprising reconstruction, this is not the general case, especially when dealing with real data that is subject to noise and systematics.

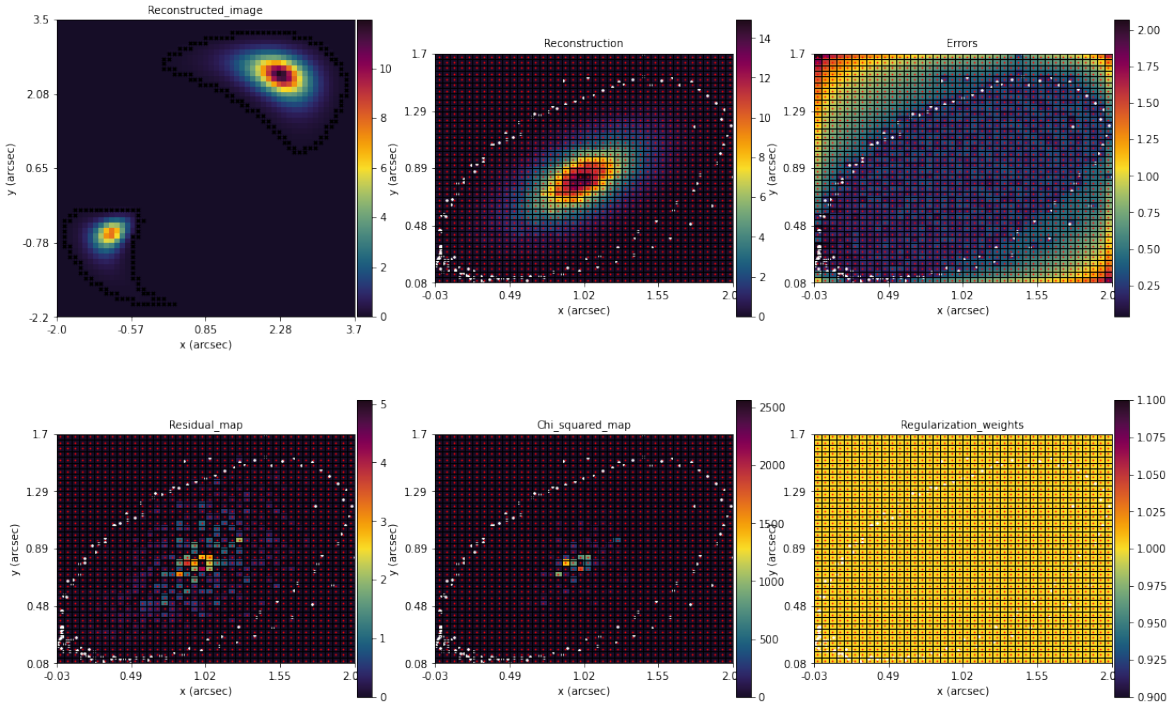


Figure 26 – The top central panel shows the source reconstruction using the SLI method, and the most probable model. More information, see the text.

After verifying the ability of our implementation to retrieve the most probable model, we consider our code to be able to probe the parameters through real data, and finally perform a real test of GR in a new, even particular, domain.

6 Results

In this Chapter we present the results of the SDP.81 modelling, which include the kinematical map (obtained fitting the IFU data with pPXF), the MGE lens light decomposition used to trace the stellar content of the galaxy, and the simultaneous reconstruction of the observed lensed arcs and resolved stellar kinematics using a single self-consistent mass model, whose most probable parameters, including $\gamma_{\rm PPN}$, are obtained using DYNESTY

6.1 Kinematical analysis

To obtain the kinematical information used as constraints during JAM modelling, we model the spectral data obtained with MUSE (Section 4.2) using the pPXF routine (Section 5.1).

In order to improve the spectral modelling, we first select only the spectra with a signal-to-noise (SNR) above 2.0. We measured this SNR as the ratio of the average signal and the average noise in the wavelength range that we fit ($5000 - 7000$Å in the galaxy frame). A total of 290 spaxels were selected, and Fig. 27 shows the SNR map.

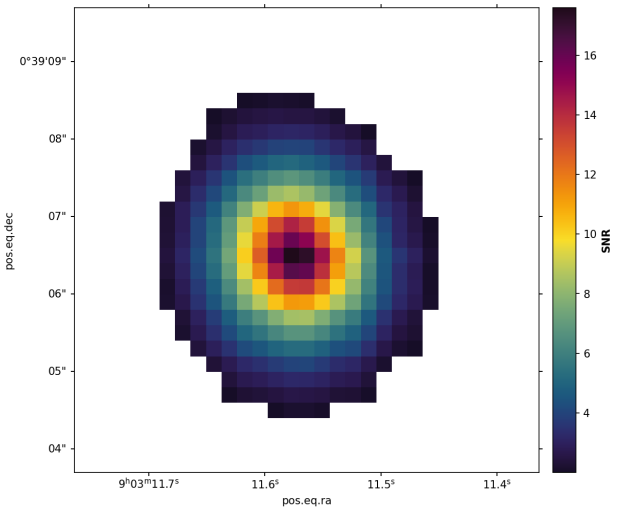


Figure 27 – MUSE SNR map of selected spaxels (SNR > 2.0). The SNR is the ratio of the average signal and the average noise in the wavelength range that we want to fit.

Then, to ensure a reliable stellar kinematics, we sum nearby spectra using the adaptive spatial two-dimensional binning scheme (VorBin; Cappellari; Copin 2003) to increase the SNR in each resulting Voronoi bin at a minimum of 25. The Voronoi binning

scheme results in 17 *goodbins*, which can be seen in Fig. 28. The small dots in the lower panel represent the older spaxels, and the red dots the new Voronoi bins.

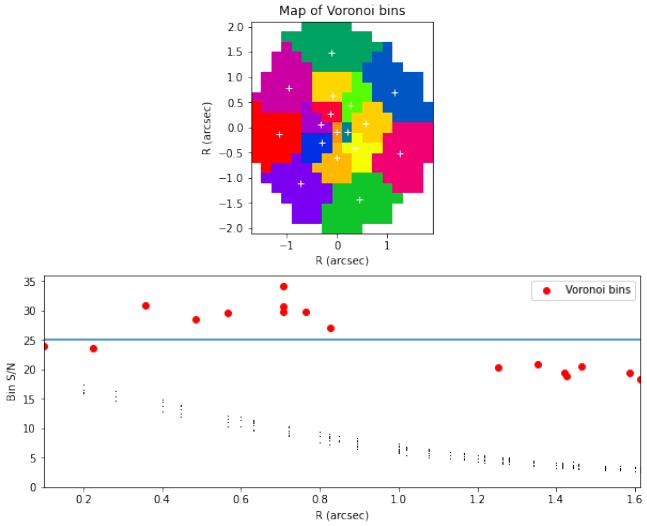


Figure 28 – Map of Voronoi bins. Each color in the upper panel represents one of the 17 Voronoi bins with SRN $\sim$ 25. In the lower panel we see a SNR radial profile of the spaxels/bins. Small dots are the older spaxels, and the red dots are the co-added Voronoi bins.

Each of the remaining Voronoi bins is modelled using pPXF as a combination of a select set[1] of templates from the Indo-US stellar library[2] (Valdes et al., 2004) in the optical region $3460 - 9464$Å. We choose to use the Indo-US templates due to its spectral resolution of FWHM$=$ 1.35Å, which is appropriate for the modelling of spectra at the redshift of our galaxy.

For each Voronoi bin we fit the two first moments of the Gauss-Hermite polynomials $[V, \sigma_V]$, with the inclusion of an additive polynomial of order 4 (Section 5.1). Besides that, we mask possible emissions lines, and two regions where the variance in the data was very prominent. We perform our modelling in the observer rest-frame, which means that we bring the galaxy spectra to the rest-frame before the fitting.

As an example of the spectral fitting procedure[3], in Fig. 29 we show the result for the central bin. The observed spectrum is shown in black, and the model in red. The residuals are green diamonds, and the gray vertical lines are the masked parts, while the bold blue lines are the masked lines. In light blue we see the variance in each pixel of the spectrum. The best-fit velocity and velocity dispersion are $V = 141$km/s, $\sigma_V = 246$km/s respectfully, with $\chi^2_{DOF} = 1.085$. The spectra are in the observer rest-frame.

To measure the uncertainty in the fitting parameters (velocity and velocity dispersion), we use a bootstrapping approach. In a few words, we perform a first fit, called best

[1] This set of templates are available at <https://github.com/remingtonsexton/BADASS3>.
[2] <https://www.noao.edu/cflib/>
[3] Other examples of model fitting can be accessed in our GitHub.

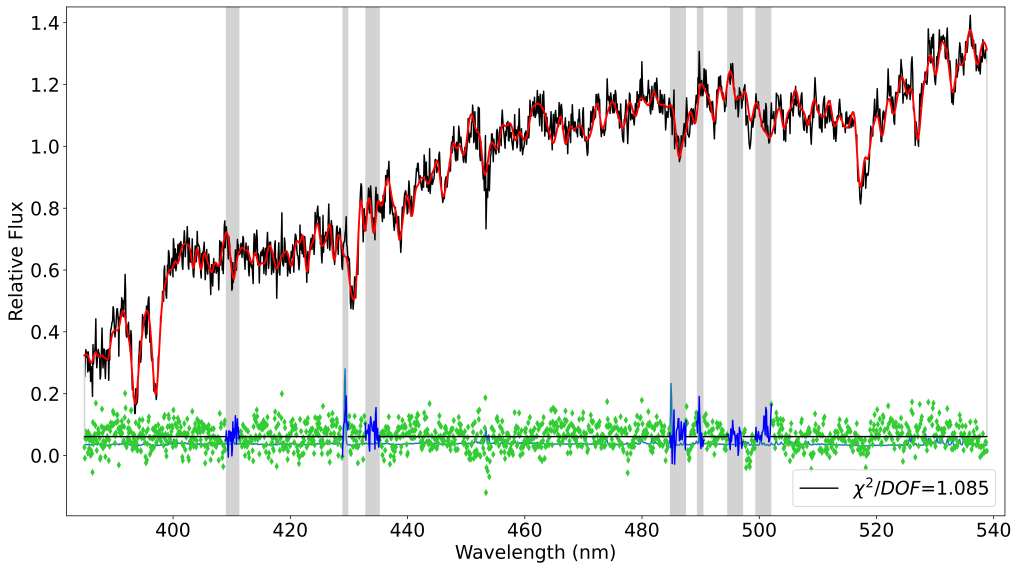


Figure 29 – Model of central spectrum obtained using pPXF, in the observer rest-frame. Observed data are in black, model in red, gray vertical lines are masked regions, and the residual are shown in green. The best fit parameters are $V = 141$km/s and $\sigma_V = 246$km/s. In the bottom right we see the reduced χ^2. For details of the spectral fitting, see the text.

fit, and compute the residuals for each pixel. Now, from each pixel in the best fit spectrum, use its value as a center of a Gaussian distribution, and the residual associated with this pixel as the dispersion of this Gaussian. This ensures a new value for each pixel in the spectrum, drawn from the Gaussian distribution. Following this procedure for all the pixels we generate a new full range spectrum. We fit this new spectrum with pPXF (setting the bias flag equal to zero) and save the outputs $[V, \sigma_V]$. We repeat this procedure generating 200 new spectra and measurements of V and σ_V. After that, we plot the distributions for velocity and velocity dispersion (see Fig. 30), and compute the 1σ of each as the uncertainty associated to the measurement[4].

After fitting all the 17 Voronoi bins, we construct the $V_{\rm rms}$ map, which is the root mean square velocity[5] calculated for each Voronoi bin. For that, we subtract the central velocity, i.e., the velocity of the central bin, from each of Voronoi bin velocities, in order to obtain the real velocity with respect to the galaxy center. We do not see any evidence of rotation. This $V_{\rm rms}$ map is shown in Fig. 31.

[4] This strategy is very similar to other bootstrapping approaches used by the community. In particular, our approach is based on two main codes. One by Jonathan Cohn (<https://github.com/jhcohn/ppxf/blob/master/ppxf_nifs_kinematics_witherr.py>), and the other by Remington Oliver Sexton (<https://github.com/remingtonsexton/BADASS3/blob/master/badass3_v7_7_6.py>). We thank all the community for making their research publicly available.

[5] $V_{\rm rms}^2 = V^2 + \sigma_V^2$

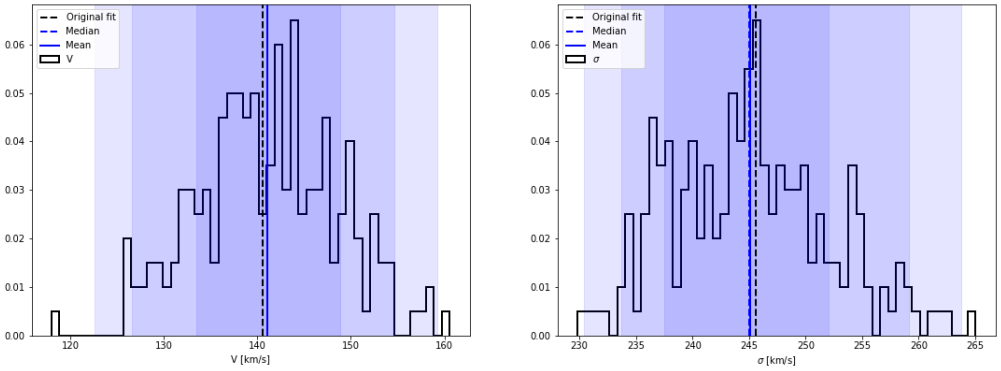


Figure 30 – Bootstrapping realization for the central spectrum of SDP.81 lens. The shaded areas shows the 1σ, 2σ and 3σ deviations. Left panel shows the results for the velocity, and the right panel the results for the velocity dispersion. Black dotted line shows the original best fit, i.e., the fit associated to the best fit (see text), the blue dotted line shows the median of the distribution, and the blue continuum line the mean. The vertical axis is the of occurrence of each bin normalized by the total of iterations.

We use the following error propagation to estimate the uncertainties in the $V_{\rm rms}$

$$\sigma_{\rm rms} = \frac{\sqrt{(V \times 1\sigma_{\rm vel})^2 + (\sigma_V \times 1\sigma_{\rm disp})^2}}{V_{\rm rms}}, \tag{6.1}$$

where $\sigma_{\rm rms}$ is the uncertainty in the root mean square velocity, $1\sigma_{\rm vel}$ is the 1σ uncertainty in the velocity, and $\sigma_{\rm disp}$ the 1σ uncertainty in the velocity dispersion.

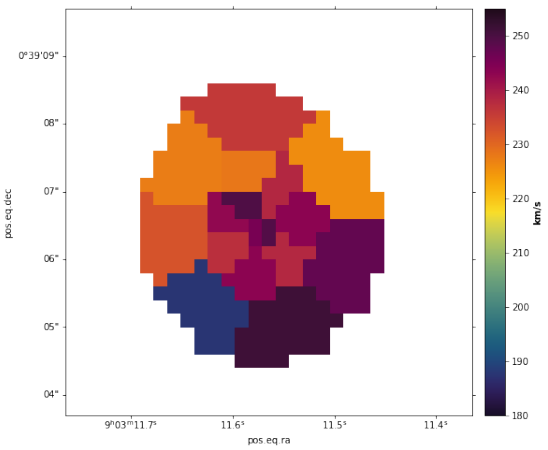


Figure 31 – Resulting $V_{\rm rms}$ map of SDP.81 lens galaxy. North is up and East is left.

We notice that two bins, which are close to each other near the border of the galaxy (the two pixels in the South of Fig. 31), present very discrepant values for $V_{\rm rms}$. We believe that these discrepant values (much higher and much lower) are being caused either by some contamination of the lensed source (whose emission is more intense in that region, as can be seen in Fig. 14) and/or by the low SNR at the edge of the galaxy.

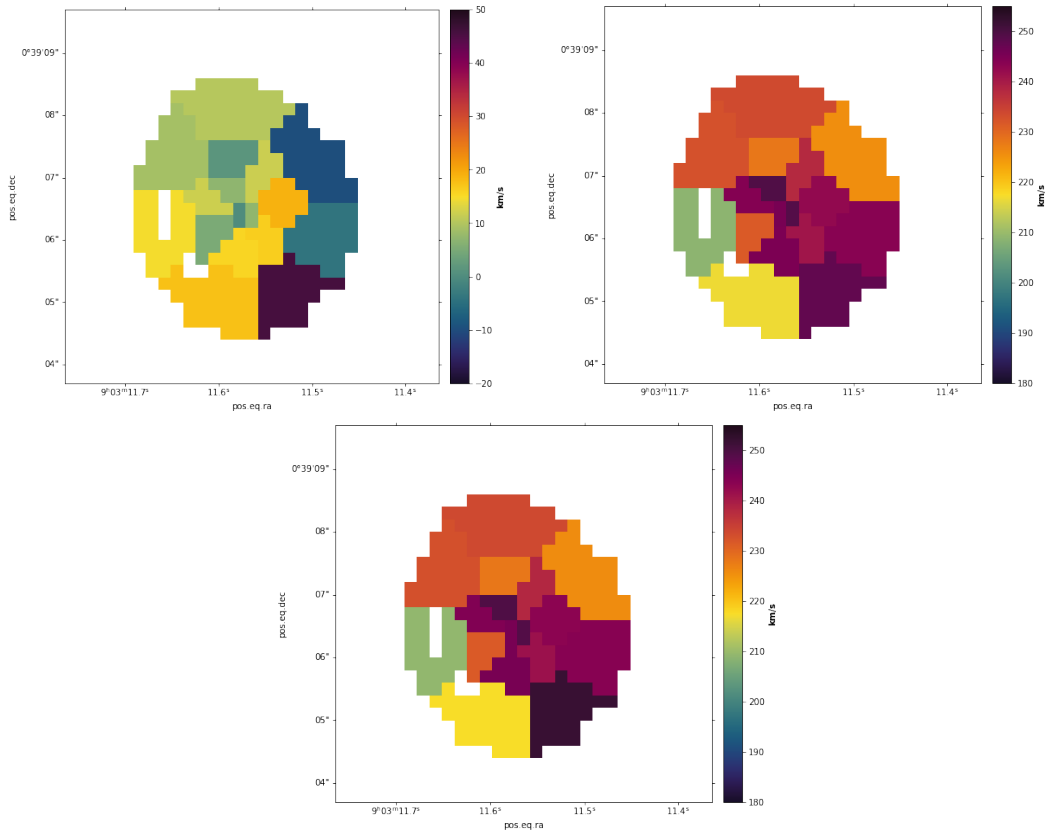


Figure 32 – Final kinematical maps of SDP.81 lens galaxy after remove possible contaminated spaxels. The blank areas are the removed spaxels. The scale bar as the same as Fig. 31. Top left shows the velocity map, top right the velocity dispersion map, and in bottom the $V_{\rm rms}$ map.

In order to minimize the possible contamination of the lensed source, we identify the emission peaks of the arcs in the HST image (Fig. 10), and after overlaying it to the MUSE emission, we remove the spaxels of the MUSE which correspond to those pixels. After removing the spaxels with possible contamination, we perform all the steps described above, however using only the remaining spaxels where we believe the emission of the source is negligible. The new kinematic maps can be seen in Fig. 32, where the blank areas indicate the removed spaxels. The scale bar is the same as Fig. 31, to make the comparison possible. We can see a small improvement in those pixels with discrepant values, and for this reason we choose to work with this final map towards the rest of the dissertation.

6.2 MGE Models

For modelling the stellar content in both dynamical and lensing analysis as well as for the modelling of the dark matter content in the dynamical analysis, we will use the MGE approach, as described in Section 5.2.

For the stellar content, we parametrize the surface brightness profile of the HST

F160W image. However, before parametrizing it, we parametrize the HST ePSF using a sum of 2D circular Gaussians, such that we can take it into account during the MGE parametrization of the galaxy surface brightness. A set of three Gaussians are needed to model the ePSF (see Section 4.1) and Table 5 shows its components.

Total Counts	σ [arcsec]
1.10	0.06
5.87	0.15
3.33	0.27

Table 2 – MGE parametrization of HST ePSF. For PSFs, the method assumes circular Gaussians with the same center.

Now, to perform the MGE parametrization of the lens galaxy light, we followed the next steps:

1. We use the FIND_GALAXY procedure available at MGEFIT package to determine the galaxy center, its position angle[6], and the ellipticity. As a result, we find a position angle equal to $PA = 11.9°$, and a ellipticity of $\epsilon = 0.150$.

2. Then we use the latter information as an input of the SECTORS_PHOTOMETRY procedure of MGEFIT, which is responsible for performing the photometry of the galaxy image along sectors equally spaced in angle. During this step we have masked the star in the North of the HST F160W image.

3. Finally, we apply the MGE_FIT_SECTORS_REGULARIZED procedure of MGEFIT package for each of the above photometric sectors. This is the actual MGE fitting procedure. In this step the parametrized PSF is taken into account.

These three steps are summarized in Fig. 33.

[6] Measured counterclockwise from North.

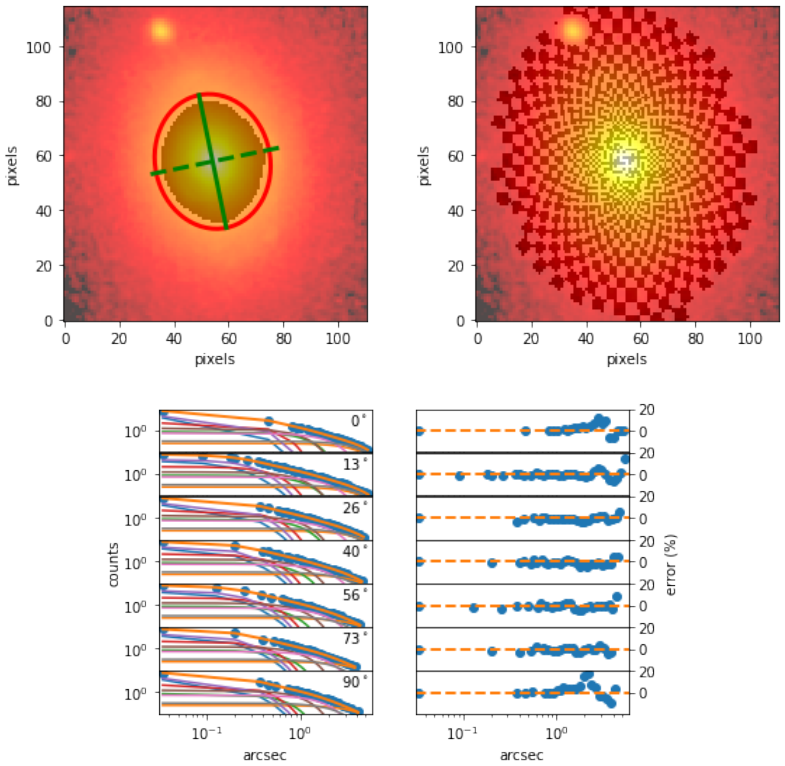


Figure 33 – Summary of MGE fitting procedure. Top left shows the FIND_MY_GALAXY step. Top right panel is the photometric sectors where the MGE fit occurs. The bottom panels show a few examples of the MGE fit along some of the sectors linearly spaced in angles between the major (0°) and the minor (90°) axis and their corresponding errors.

The best fit obtained with MGE, and the residuals are shown in Fig. 34.

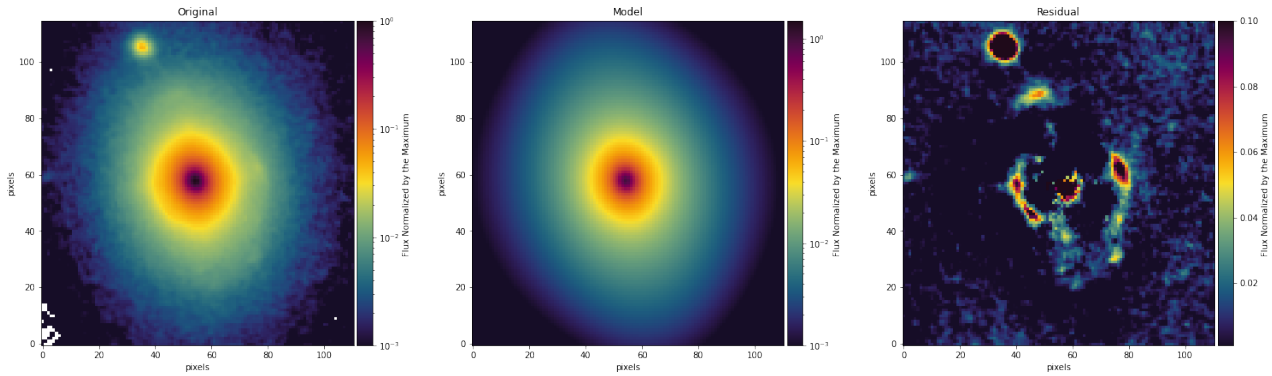


Figure 34 – MGE model of lens galaxy SDP.81. From left to right: image data, best MGE model, residual normalized by the maximum value. In the right panel we clearly see the emission of the background lensed source. North is up and East is left in all panels.

However, this is not our final MGE model used for the combined modelling. To mitigate possible contamination due to the emission of the lensed source in the lens surface brightness profile fitted, we exclude the lensed arcs from the original image to separate

the lens light from the background lensed galaxy emission. For that we interpolate[7] the image in the regions where the lensed arcs were excluded (right panel of Fig. 34), in order to obtain a lens image "without" the source emission.

Then, we perform the same steps described earlier in the interpolated image, to finally obtain the MGE model used in the rest of the analysis. This fit results in a set of 9 concentric Gaussians and co-aligned at a position angle of 11.9° (counterclockwise from North), and components given by Table 3. Figure 35 shows the isophotes as observed with HST overlaid by the MGE model, contours are in steps of 0.5 mag/arcsec^2. The yellow circle at North is the masked star. The right panel is a zoom-in in the galaxy center.

Total Counts	$\sigma^{\rm pix}$	$q^{\rm Obs}$
128.56	0.51	0.72
261.32	1.47	0.62
198.36	2.93	0.82
406.64	6.16	1.00
149.08	7.69	0.47
501.57	11.89	1.00
239.36	16.44	0.57
607.56	29.46	0.61
590.38	29.46	1.00

Table 3 – MGE parametrization of HST F160W image of SDP.81. First column is the total counts for each Gaussian component, second column is the dispersion in pixel units, and the third column is the observed axial ratio of each Gaussian.

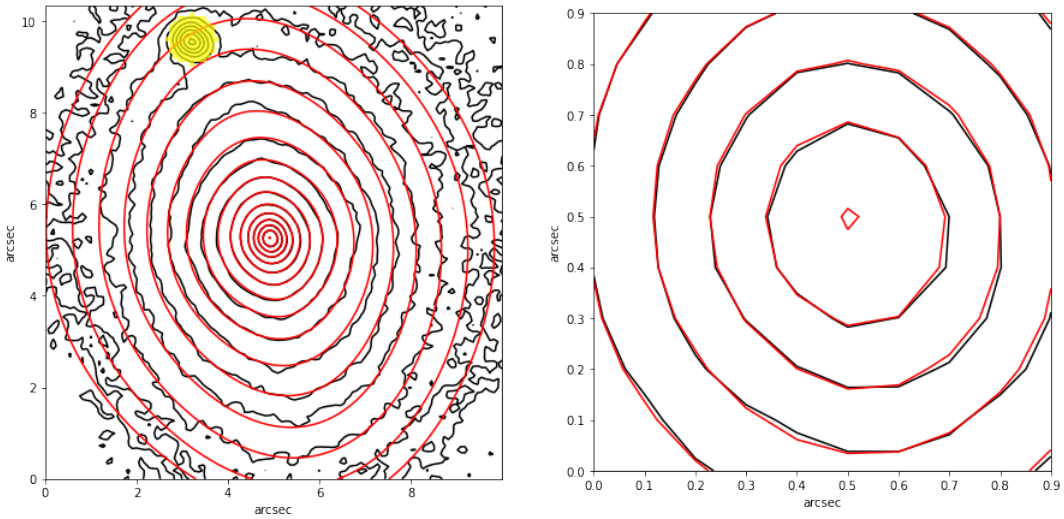


Figure 35 – HST F160W isophotes (black lines) overlaid by the final MGE model (red lines). The contours are in steps of 0.5 mag/arcsec^2. The right panel shows a zoom in the galaxy center. The yellow circle up to North is the masked star. North is up and East is left.

[7] We use the SCIPY.INTERPOLATE.GRIDDATA routine for a cubic interpolation.

Using equations (5.13), (5.14) and (5.15), we transform the total counts of each Gaussian (Table 3) into a surface intensity I in solar luminosity per parsec square. We assume a Vega zeropoint[8] in band F160W equal to $Z^{\rm Vega}_{\rm F160W} = 24.662$ mag, and an absolute solar magnitude equal to $\mathrm{M}^{\rm Vega}_{\odot,F160W} = 3.37$ (Willmer, 2018). Table 4 shows the Gaussian inputs used during JAM modelling.

I [$\mathrm{L}_\odot/\mathrm{pc}^2$]	σ [arcsec]	$q^{\rm Obs}$
17964.30	0.05	0.72
5062.47	0.13	0.62
728.03	0.26	0.82
277.24	0.55	1.00
139.13	0.69	0.47
91.66	1.07	1.00
40.06	1.48	0.57
29.55	2.65	0.61
17.58	2.65	1.00

Table 4 – Gaussians components used as input for the JAM modelling of SDP.81 lens galaxy.

We take the opportunity to present the MGE parametrization of the MUSE ePSF (see Section 4.2). We follow the same procedure used for the HST ePSF parametrization, assuming a set of circular and concentric Gaussians. Again, this is useful because we can introduce this parametrized ePSF in the JAM modelling. Table 5 shows the components of the parametrized MUSE ePSF.

Total Counts	σ [arcsec]
0.10	0.24
0.53	0.42
0.37	0.82

Table 5 – MGE parametrization of MUSE ePSF.

6.3 Preparing the data and setting the simultaneous modelling

Finally, after the long discussion and presentation made in the previous Chapters, we are ready to test the GR using the data from the SDP.81 system. Therefore, we need to make the final adjustments to the data and prepare the model that will be sampled.

The first adjustment of the data, and also the first assumption about the model, is that we assume that the lens mass center coincides with the lens luminous center determined in the last Section. So, we first find the coordinates (RA, DEC) of the center

[8] <https://www.stsci.edu/hst/instrumentation/wfc3/data-analysis/photometric-calibration/ir-photometric-calibration>

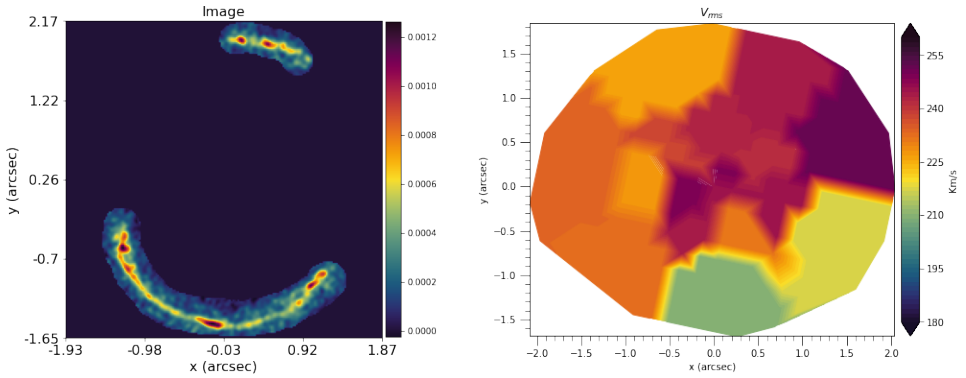


Figure 36 – Centered and rotate data used for modelling. Left panel shows the emission of the source galaxy seen by ALMA, while the right panel is the $V_{\rm rms}$ map linearly interpolated for visualization.

determined by the FIND_MY_GALAXY routine, and then we use this coordinates to center the image of the emission of the background galaxy observed by ALMA. This also assumes that the dark matter halo has the same center.

The second adjustment is a requirement of the JAMPY[9] package. It assumes that the x'-axis coincides with the galaxy projected major axis. So to fulfill this requirement, we rotate our kinematical and gravitational arcs data by the position angle determined earlier, such that the semi-major axis of the galaxy coincides with the $x'-$axis. This rotation was made keeping the center fixed[10].

An image with these transformation already applied can be seen in Fig. 36. We mask the ALMA data to model only the position where the emission from the background galaxy is clearly seen. In addition, the ALMA data have already been binned in 4×4 pixels, as explained in the Section 4.3.

With these two datasets we want to find the most probable model which is able to reproduce both data simultaneously. It means to use DYNESTY to sample the parameter space and find the most probable values of the parameters of the single self-consistent mass model that reproduces both lensing and kinematical data.

To make our model more realistic, and at the same time practicable, we use a mass-to-light ratio modulated by a Gaussian function[11], which allow us to set a different mass-to-light ratioΥ_j for each component of the MGE parametrization based on its dispersion. This Gaussian-modulate mass-to-light ratio has three parameters: the central stellar mass-to-light ratio Υ_0, a gradient parameter δ that describes the smoothness of the Gaussian, and a "lower" υ_0 parameter, which is the ratio between the central and the

[9] The Python implementation of JAM method.

[10] Since the MGE PSFs for MUSE and HST are circular, we do not need to rotate it, however the dirty beam of ALMA was rotated by the same angle.

[11] This approach was implemented by Hongyu Li at his GitHub page <https://github.com/HongyuLi2016/JAM>.

outermost mass-to-light ratio. Mathematically, we have

$$\Upsilon_j = \Upsilon_0 \left[\upsilon_0 + (1 - \upsilon_0) e^{-0.5(\sigma_j \, \delta)^2} \right], \tag{6.2}$$

where σ_j is the dispersion associated to the MGE component. There are two main advantages in this approach. One, we reduce the number of parameters related to the mass-to-light ratio from nine (one per MGE component) to three. Two, this ensures that the total baryonic mass-to-light profile is monotonically decreasing, consistent with observations of real galaxies.

Still in order to make our model more realistic and yet computationally possible, we allow the anisotropy of the system be described by a set of eight β_z^i, i.e, one anisotropy parameter for each Gaussian component in the MGE parametrization, excepting the two central Gaussians, which we fix to have the same $\beta_z^{1,2}$. We do that because the two central Gaussians have dispersions smaller than the MUSE pixel scale. Moreover, this choice reduces the number of parameters in our model. It is important to highlight that the individual anisotropies do not have any physical interpretation, and are not associated with different stellar populations. Assuming different anisotropies is just a way to parametrize the real anisotropy profile, such that the physical interpretation comes from the anisotropic profile $\beta(R)$, where R is the distance from the galactic center.

The dark matter component assumed is an eNFW profile with three free parameters: κ_s, r_s, and $q_{\rm DM}$ (see Section 5.4.5). It means that, at each iteration during the sampling, we fit the eNFW profile described by these parameters using the MGE method, and include it in the dynamical modelling. For the lens modelling, we include the analytical version of the eNFW profile, with the same parameters values. We are assuming that the dark matter halo has the same center of the lens galaxy, and has the same orientation, but we allow the axis ratio $q_{\rm DM}$ to vary. Our model also includes a central supermassive black hole at the center of the lens, modelled as a point mass.

For the lensed source reconstruction, we use a pixelated source grid with 40×40 pixels, and we keep the regularization constant (Section 5.4) fixed at $\lambda = 2.00$.

The model has other four parameters, two associated to a possible external shear (see Chapter 3) in the lens model ($\text{shear}_{\rm mag}$, shear_{ϕ}), the inclination i that we chose to sampling from the equation (5.75), and the $\gamma_{\rm PPN}$ parameter itself.

In summary, our model has nineteen parameters: three associated to the mass-to-light ratio, eight associated to the anisotropy profile, three describing the dark matter halo, one describing the mass of the supermassive black hole, one related to the inclination of the galaxy, two describing the external shear, and the Post-Newtonian parameter.

6.4 Most probable model and Results

We use DYNESTY (see Section 5.6) to sample the posterior probability of our model fit to the data, exploring a parameter space with nineteen dimensions: three describing the M/L, eight for the stellar anisotropy, one for the galaxy inclination, one for the mass of the supermassive black hole, three describing the dark matter profile, two for the external shear, and the γPPN. We use a multi-ellipsoid bounding distribution with a sampling proposal set to be a random walk and perform $N_{\rm walks} = 15$ walks. We use a set of 400 live points to explore the parameter space until the terminator criterion, set as $\Delta \ln \hat{\mathcal{Z}} < 0.1$, is reached.

We set our prior knowledge about the parameters to be quantified by non-informative priors (flat priors), excepting the $\gamma_{\rm PPN}$ which is quantified by a Gaussian distribution around the GR. For the flat priors, we use more conservative boundaries based on values for early-type galaxies found in the literature. On the other hand, for the $\gamma_{\rm PPN}$, we choose a Gaussian prior centered in GR ($\gamma_{\rm PPN} = 1.00$) and dispersion equal to 0.05. This narrow distribution for the Post-Newtonian parameter is based on the value found by CT18, which has this order of uncertainty. The model priors are listed in Table 6. A total of 274712 likelihoods were evaluated during the sampling.

The most probable parameters of our model are summarized in Table 7 with their statistical uncertainties (1σ confidence limit). The most probable lens model and the residuals are seen in Fig. 37, and the the most probable dynamical model in Fig. 38. In Fig. 39 we show the source reconstruction assuming the most probable model for the lensed arcs.

Parameter	Prior
Υ_0	$\mathcal{U}$(0.5, 15)
υ_0	$\mathcal{U}$(0, 1.0)
δ	$\mathcal{U}$(0.1, 2.0)
β_z^i	$\mathcal{U}$(-3.0, 3.0)
$q_{\rm inc}$	$\mathcal{U}$(0.0550, 0.4688)
$\log M_{\rm SMBH}$	$\mathcal{U}$(7.0, 11.0)
κ_s	$\mathcal{U}$(0, 2.0)
r_s	$\mathcal{U}$(2.0, 50.0)
$q_{\rm DM}$	$\mathcal{U}$(0.1, 1.0)
$\text{shear}_{\rm mag}$	$\mathcal{U}$(0, 0.1)
shear_ϕ	$\mathcal{U}$(0°, 179°)
$\gamma_{\rm PPN}$	$\mathcal{G}$(1.0, 0.05)

Table 6 – Priors of the model. $\mathcal{U}$(lower, upper) represents a non-informative prior, with lower and maximum values accepted. $\mathcal{G}$(mean, sigma) is a Gaussian prior, with mean and dispersion values. Υ_0 is in solar mass per solar luminosity, δ is in arcsec^{-1}, $\log M_{\rm SMBH}$ is in logarithmic solar mass, r_s in arcsec, and shear_ϕ is in degrees. All other parameters are dimensionless.

Parameter	Most Probable Value	Description
Υ_0	$3.4542^{0.0028}_{-0.0017}$	Central Mass-to-Light ratio [$L_\odot/M_\odot$]
υ_0	$0.0895^{0.0010}_{-0.0008}$	Lower value for the Mass-to-Light ratio
δ	$1.4550^{0.0025}_{-0.0025}$	Smoothness of the Mass-to-Light profile [arcsec^{-1}]
$\beta_z^{1,2}$	$0.2497^{0.0018}_{-0.0014}$	Anisotropy for MGE components 1 and 2
β_z^3	$1.6835^{0.0020}_{-0.0020}$	Anisotropy for MGE component 2
β_z^4	$0.5399^{0.0015}_{-0.0019}$	Anisotropy for MGE component 3
β_z^5	$-0.3448^{0.0037}_{-0.0049}$	Anisotropy for MGE component 4
β_z^6	$-1.8370^{0.0037}_{-0.0022}$	Anisotropy for MGE component 5
β_z^7	$0.0256^{0.0047}_{-0.0057}$	Anisotropy for MGE component 6
β_z^8	$0.0133^{0.0027}_{-0.0028}$	Anisotropy for MGE component 7
β_z^9	$0.5150^{0.0011}_{-0.0014}$	Anisotropy for MGE component 8
$q_{\rm inc}$	$0.3045^{0.0006}_{-0.0005}$	Deprojected axial ratio
$\log M_{\rm SMBH}$	$7.5392^{0.0027}_{-0.0034}$	Logarithmic mass of supermassive black hole [$M_\odot$]
κ_s	$0.2724^{0.0001}_{-0.0002}$	Scale factor of dark matter halo
r_s	$4.4937^{0.0041}_{-0.0053}$	Scale radius of dark matter halo [arcsec^{-1}]
$q_{\rm DM}$	$0.9002^{0.0003}_{-0.0002}$	Axial ratio of dark matter halo
$\text{shear}_{\rm mag}$	$0.0871^{0.0000}_{-0.0001}$	Shear magnitude
shear_ϕ	$0.8832^{0.0022}_{-0.0020}$	Shear angle counterclockwise from $x'-$axis [degree]
$\gamma_{\rm PPN}$	$1.0092^{0.0001}_{-0.0001}$	Post-Newtonian Parameter

Table 7 – Newer Results: The inferred median and 68% confidence intervals (1σ) for the most probable model, assuming a baryonic and dark matter content.

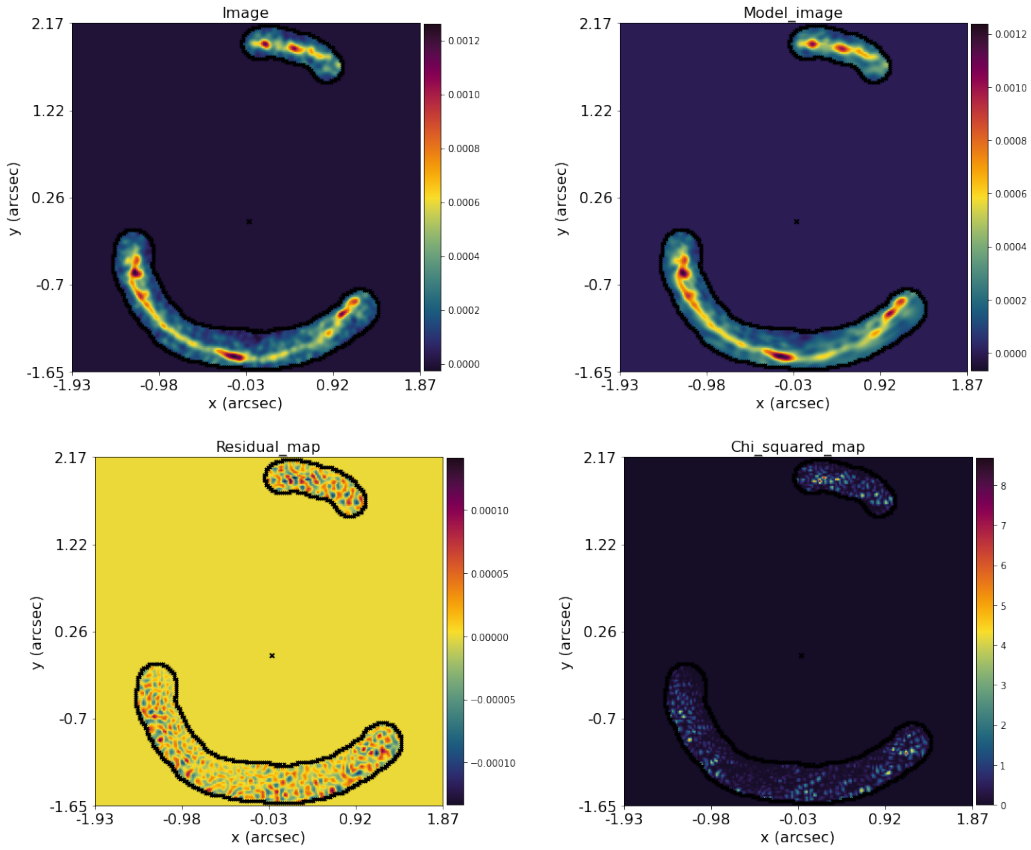


Figure 37 – Most probable lens model. Top row shows the image data (left) and the reconstructed model (right) assuming our most probable parameters. Bottom row shows the residual map (left) and the χ^2 map (right).

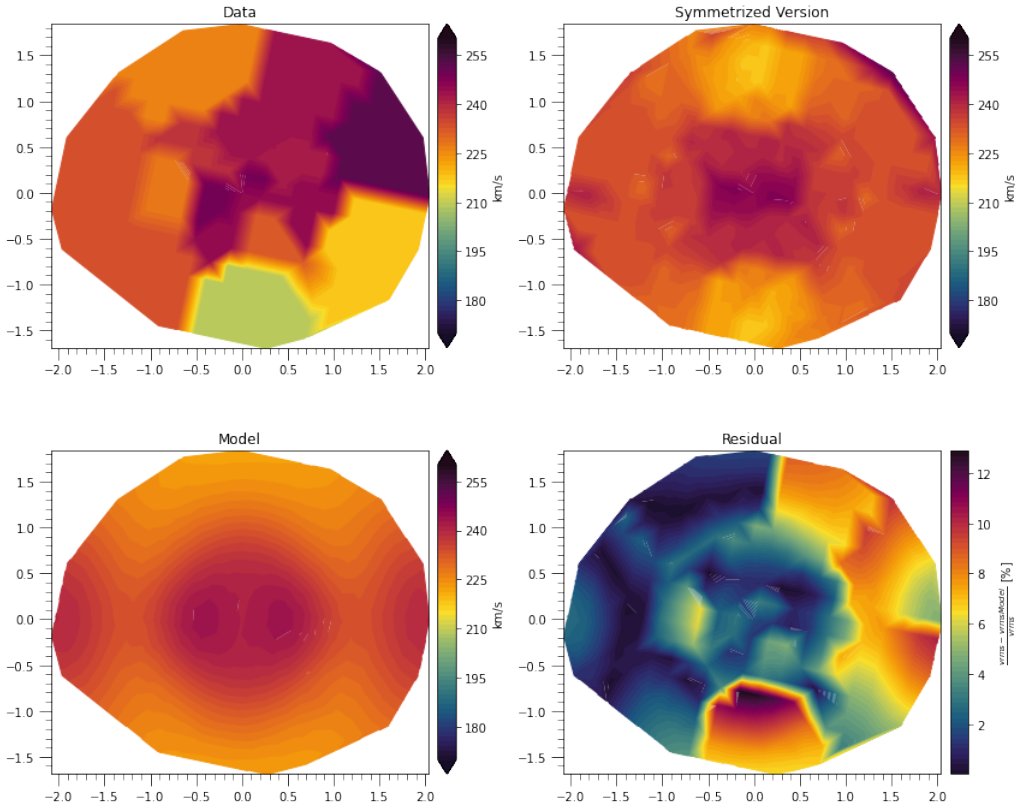


Figure 38 – Most probable kinematical map obtained from the dynamical modelling. In the top row are the $V_{\rm rms}$ map obtained from the observation (left) and its symmetrized version (right) to help with the visual inspection. In the bottom we see $V_{\rm rms}$ map obtained from the most probable dynamical model (left) and the residuals (right). All figures are linearly interpolated for better visualization.

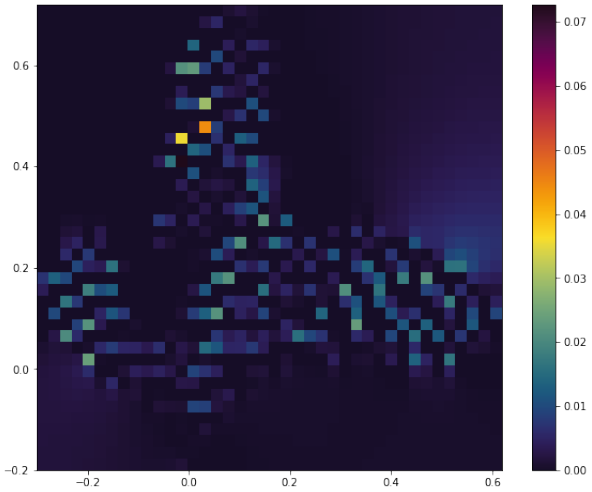


Figure 39 – Source reconstruction assuming a rectangular grid with 40×40 pixels, and the most probable model (Table 7).

The two-dimensional posterior distribution (the corner plots) for all the parameters can be seen in the Appendix D. Here, in Fig. 40, to save amount of space, we plot only the distribution for some selected parameters.

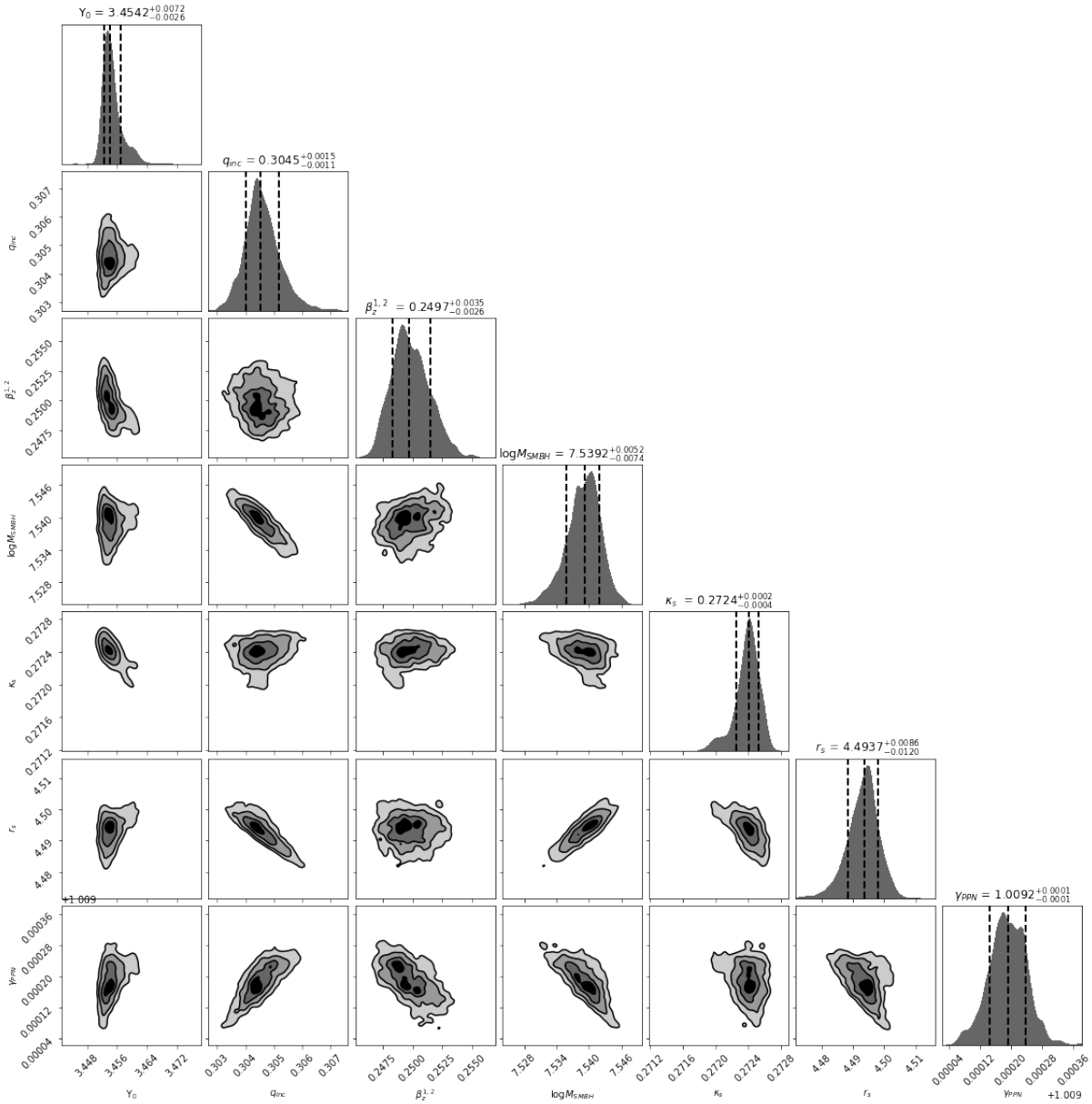


Figure 40 – Two dimensional marginalized posterior for some of the parameters of the combined analysis. The contours are 10%, 40%, 65%, and 85% quantiles, respectively. The dashed lines are 1σ credible interval.

6.5 Some Final Considerations about the Modelling and Results

The SDP.81 modelling performed contains a considerable amount of information, which requires a deeper analysis in order to derive any confident conclusion. Unfortunately, we do not have the time to carry out such analysis at this point, so we will limit ourselves to the discussion of the more direct implications of our model and will make some cautionary notes about the results obtained in the last Section.

The most straightforward comment is about the quality of the models. Looking at the residuals in Figures 37 and 38, we can see that our model simultaneously reproduces the lensing and kinematical data, except for a particular bin in the dynamical model. In Section 6.1 we have already pointed this particular issue in the kinematic map, and now we reinforce the necessity to make a deeper analysis into what could be causing such effect. On the other hand, the lens model seems to reproduce the data with high accuracy, and therefore can be considered an excellent result.

Unfortunately, the comparison between our mass model and other models previously presented in the literature (e.g., Dye et al., 2014; Rybak et al., 2015) using the same data for the lensing model, is difficult to do. This is because the majority of those works are interested in the source galaxy, not in the lens, and more critically, different authors have assumed different mass profiles for the lens (not all with a clear physical interpretation), making a fair comparison difficult.

In this sense, and considering the high quality data used in this work, we consider our lens mass profile the most complex and detailed for this system to date, allowing us to infer the distribution of stellar and dark matter, as well as the black hole mass. It is also worth mentioning that such level of complexity and detail was only possible due to the inclusion of the spatially-resolved stellar kinematic data to the analysis, which was not considered in previous studies of the SDP.81 system.

Less impressive is the source reconstruction, shown in Fig. 39. Even with the source reconstruction being a sub-product of our analysis, we expected a better source reconstruction due to the quality of the ALMA data and the excellent lens model. Comparing our source reconstruction with the analysis made by Tamura et al. (2015), we see a similar morphology distribution of the source galaxy (an extent morphology)[12], however we cannot see the distinct features as been shown by Dye et al. (2015) and Rybak et al. (2015). Once our lens modelling looks good enough, the main problem here could be addressed to the regularization constant or the source grid resolution chosen. We want to point that some works make use of an adaptive grid (Dye et al., 2015; Rybak et al., 2015), however due to the high computational cost to implement such methods in the context of this dissertation, we leave that for future work.

[12] Remember that our system are rotated by the PA angle of 11.9°.

For the most probable model parameters we can compute the total mass of the lens galaxy. For that, we use analytic expression derived by Mitzkus; Cappellari; Walcher (2017) to calculate the mass of an axisymmetric MGE decomposition within a sphere of a given radius[13].

Assuming an Einstein Ring of 1.56″ (Dye et al., 2014), the total mass (stellar + black hole + dark matter) inside a sphere with this radius is $2.265 \times 10^{11} M_{\odot}$, where the dark matter represents $\sim 54\%$ of this mass.

To make a fairer comparison with other works, we also compute the mass inside the half-light radius $R_e = 1.156''$, estimated following the procedure described[14] in Cappellari et al. (2013). We found a total mass of $1.650 \times 10^{11} M_{\odot}$, with a fraction of dark matter equal to 45%. This fraction of dark matter marginally agrees with lensing studies using lens galaxies at similar redshift and effective radius (see Figure 6 of Sonnenfeld et al.). We also notice that the external shear is small, such that the contribution of the SDP.81 lens galaxy is dominant in the lens modelling.

Finally, we can discuss the elephant in the room. Our model is excluding GR within 1σ!! However, the uncertainties provided by the Bayesian inference are statistical uncertainties, and a careful attention to them quickly reveals that our estimates are underestimated, and we need to estimate the systematics involved in the analysis. Indeed we already expect that, since the inference on $\gamma_{\rm PPN}$ is dominated by systematic effects (e.g., Cao et al., 2017; CT18).

Estimating the systematics involved in our analysis is not an easy task, once we need to take into account different mass models, stellar templates used to obtain the kinematical constraints for the dynamical modelling, the impact of the cosmological parameters, the assumptions of the orbits of the stars, and so on. This not only requires a lot of time to analyze, but also more computational power. Unfortunately, both are beyond what could be done during the two-year duration of the Masters, so we leave this analysis for a future work.

As a compromise, we expect our inference to have a systematic error between the one obtained by Cao et al. (2017) and the inference made by CT18. It means that, we expect an upper limit about of $\sim 25\%$ and a lower limit about of$\sim 9\%$ of systematic effects in our inference.

The worst-case scenario leaves us with $\gamma_{\rm PPN} = 1.0092 \pm 0.0001({\rm stat}) \pm 0.2523({\rm sys})$, which is in agreement with GR, inside 1σ. In the best-case scenario, the final inferred $\gamma_{\rm PPN} = 1.0092 \pm 0.0001({\rm stat}) \pm 0.0908({\rm sys})$, still in accordance with the GR, within 1σ confidence limit. It is important to remember that we assume that $\gamma_{\rm PPN}$ is constant across the relevant length scale of SDP.81 lens galaxy, and a well defined Newtonian limit exist.

[13] This can be done using the JAMPY routine MGE_RADIAL_MASS.
[14] And implemented in JAMPY package as MGE_HALF_LIGHT_ISOPHOTE.

This means that our result is valid for the galactic scale within the Einstein Ring (~ 7.150 kpc), where we perform our measurements.

We end this Section emphasizing the necessity for studies that seek to better understand what are the sources of systematic errors and how they impact the inference of the Post-Newtonian parameter, and also how the assumptions made could change our inference. Shedding light on this knowledge can help us better understand the limits of where we are testing Einstein' theory.

7 Summary & Perspectives

Since its foundation, General Relativity has been tested in different regimes and scales. And to this day, GR has proved to be the best theory of gravitation we have, even though some problems such as the Hubble Tension and the incompatibility with Quantum Mechanics are still unsolvable.

In this work we extend the limits of tests of GR to galaxy scales and intermediate redshift. For that, we use the SDP.81 system at redshift $z_l = 0.299$, and high-resolution data coming from the best observatories/instruments: ALMA, VLT/MUSE, HST/WFC3. This test was performed in a limit where we consider the Post-Newtonian approach valid under some assumptions, in such a way that we were able to simultaneously model the gravitational lens phenomenon and the lens galaxy dynamics. Such simultaneous modelling allow us to infer the Post-Newtonian parameter.

In order to perform this analysis, in Chapter 2 we presented the fundamentals of galaxy dynamics, starting from the distribution function that describes the motion of the stars in a steady-state system, until the Jeans Equations for an axisymmetric system. We emphasized that this phenomenon is described only by the Newtonian potential. We also make some connections between theory and observations.

In Chapter 3, we introduce the gravitational lens formalism, which describes the deflection of light due to gravitational fields. This phenomenon is a direct consequence of GR, and therefore it can be sensitive to both the classical Newtonian potential and a possible potential due to space-time curvature.

Chapter 4 is dedicated to the data used in this dissertation. We present some details about the instruments, and about the quality of the data we have. Although our deflected source (ALMA) and lens photometry (HST) data are great in terms of angular resolution, the IFU related data (MUSE) is not of such high quality, largely due to the redshift of the observed lens galaxy.

In Chapter 5 we spent a lot of time presenting all the numerical methods used in this work, which can be considered the central key of our modelling. We go through the methods for modelling the spectral information (pPXF) and obtain kinematic maps, the methods used for describe the lens mass model and the gravitational potential (MGE), the dynamical modelling (JAM), the semilinear inversion method used for lens modelling (AutoLens), and finally how we combine all the information in a self-consistent model, that can be study using a Bayesian framework (dynesty). We still provide a Proof of Concept as the way to validate our method.

This brings us to Chapter 6, where we finally present the results for our most probable model. Our combined analysis allowed us not only to test GR, but also to (briefly) study the mass profile of the lens galaxy. Our most likely model predicts a dark matter fraction of $\sim 45\%$ of the total mass within an effective radius, which is in agreement with previous studies of galaxies of the same mass and redshift. The reconstructed source presented similar morphology to that found in other studies of the same system using the same ALMA data, and the small differences in the reconstruction of the source can be related to differences in the source plane gridding and regularization parameter used for the lens reconstruction. The kinematical map derived from the most probable model showed a considerable agreement with the data, except for a small region in the lens galaxy, which coincides with the location of the more intense emission of the background source. As for the lens model, the agreement with the data is surprising.

The inference on the Post-Newtonian parameter showed to be tricky, once the most probable model $\gamma_{\rm PPN} = 1.0092 \pm 0.0001$ is deviating from GR within 1σ confidence limit. However, this inference is dominated by systematic errors, as has been reported in other works with a similar nature. Unfortunately, it was not possible to extend our analysis to the point of estimating the systematic uncertainties present in this work. So, as a compromise, we consider two possible scenarios based on previous studies.

The worst scenario considers that systematic errors are in the order of 25% (Cao et al., 2017), which would lead us to $\gamma_{\rm PPN} = 1.0092 \pm 0.0001({\rm stat}) \pm 0.2523({\rm sys})$, in agreement with GR, inside 1σ confidence limit. For the best scenario, the systematic error is on the order of 9% (Collett et al., 2018), such that $\gamma_{\rm PPN} = 1.0092 \pm 0.0001({\rm stat}) \pm 0.0908({\rm sys})$ at 1σ confidence limit, consistent with the prediction of one from GR. In other words, General Relativity has passed another test.

Modelling the SDP.81 system, we were able to study many features of the system, ranging from its dark matter content to its kinematics, ending perhaps in the most relevant information for this work, the inference of $\gamma_{\rm PPN}$ at intermediate redshift. To our knowledge, this is the more complex and detailed mass modelling performed for this lens system. However a lot of work still can be done.

The most important is to refine the inference in terms of the systematic uncertainties and exploration of the space parameter. This requires testing different mass models (e.g. the impact of different dark matter profiles, change the assumptions on mass-to-light ratio), testing different stellar templates in the kinematical analysis, a better understanding of the kinematic maps once we have two bins in which the spectra are possibly contaminated by the lensed galaxy emission, and quantifying the impact of the cosmological model assumed. In principle all of these can be done, but more time is needed.

A secondary goal is to refine the source reconstruction, searching for better choices of regularization coefficient and source grid resolution. It would also be interesting to study

how reconstruction can be better (or worse) using an adaptive grid.

Lastly, there is a lot of information present in this model, such that a deeper analysis of it can bring much more astrophysical information about the lens galaxy at $z_l = 0.299$. For example, we can use the information on the Υ (mass-to-light ratio) profile to investigate the Initial Mass Function (IMF) of the lens galaxy. We can also study the IMF from stellar population modelling of the MUSE data.

We end this work remembering that large data surveys are to come and with instruments with ever better resolutions. This paves the way for us to continue testing GR in even more extreme situations, and hopefully, in the future, we will better understand the tensions surrounding it.

Bibliography

Abbott, T. M. C. et al. Dark energy survey year 1 results: A precise h0 estimate from des y1, bao, and d/h data. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press (OUP), v. 480, n. 3, p. 3879–3888, Jul 2018. ISSN 1365-2966. Disponível em: <http://dx.doi.org/10.1093/mnras/sty1939>. Citado na página 24.

Ade, P. A. R. et al. Planck2015 results. *Astronomy Astrophysics*, EDP Sciences, v. 594, p. A13, Sep 2016. ISSN 1432-0746. Disponível em: <http://dx.doi.org/10.1051/0004-6361/201525830>. Citado na página 34.

Aghanim, N. et al. Planck 2018 results. *Astronomy Astrophysics*, EDP Sciences, v. 641, p. A6, Sep 2020. ISSN 1432-0746. Disponível em: <http://dx.doi.org/10.1051/0004-6361/201833910>. Citado 2 vezes nas páginas 24 and 93.

Akiyama, K. et al. First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole. *Astrophys. J. Lett.*, v. 875, p. L1, 2019. Citado na página 24.

Alam, S.; et al. VizieR Online Data Catalog: The SDSS Photometric Catalogue, Release 12 (Alam+, 2015). *VizieR Online Data Catalog*, p. V/147, mar. 2016. Citado na página 66.

Alcaniz, J. S. Dark energy and some alternatives: a brief overview. *Brazilian Journal of Physics*, scielo, v. 36, p. 1109–1117, 2006. ISSN 0103-9733. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-97332006000700002&nrm=iso>. Citado na página 24.

Allemandi, G. et al. Conformal aspects of the Palatini approach in Extended Theories of Gravity. *General Relativity and Gravitation*, v. 38, n. 1, p. 33–60, jan 2006. ISSN 0001-7701. Disponível em: <http://link.springer.com/10.1007/s10714-005-0208-7>. Citado na página 26.

Andernach, H.; Zwicky, F. English and Spanish Translation of Zwicky's (1933) The Redshift of Extragalactic Nebulae. *arXiv e-prints*, p. arXiv:1711.01693, nov. 2017. Citado na página 25.

Anderson, J.; King, I. R. Toward High-Precision Astrometry with WFPC2. I. Deriving an Accurate Point-Spread Function. , v. 112, n. 776, p. 1360–1382, out. 2000. Citado na página 60.

Baker, T.; Psaltis, D.; Skordis, C. Linking Tests of Gravity On All Scales: from the Strong-Field Regime to Cosmology. *Astrophysical Journal*, v. 802, n. 1, dec 2014. ISSN 15384357. Disponível em: <http://arxiv.org/abs/1412.3455http://dx.doi.org/10.1088/0004-637X/802/1/63>. Citado na página 26.

Baker, T.; Psaltis, D.; Skordis, C. LINKING TESTS OF GRAVITY ON ALL SCALES: FROM THE STRONG-FIELD REGIME TO COSMOLOGY. *The Astrophysical Journal*, American Astronomical Society, v. 802, n. 1, p. 63, mar 2015. Disponível em: <https://doi.org/10.1088/0004-637x/802/1/63>. Citado na página 24.

Barnabè, M. et al. The SWELLS survey - IV. Precision measurements of the stellar and dark matter distributions in a spiral lens galaxy. *Monthly Notices of the Royal Astronomical Society*, v. 423, n. 2, p. 1073–1088, jun 2012. ISSN 00358711. Disponível em: <https://academic.oup.com/mnras/article-lookup/doi/10.1111/j.1365-2966.2012.20934.x>. Citado na página 33.

Bartelmann, M. Arcs from a universal dark-matter halo profile. , v. 313, p. 697–702, set. 1996. Citado na página 87.

Bartelmann, M. Gravitational lensing. *Classical and Quantum Gravity*, IOP Publishing, v. 27, n. 23, p. 233001, Nov 2010. ISSN 1361-6382. Disponível em: <http://dx.doi.org/10.1088/0264-9381/27/23/233001>. Citado na página 57.

Bartelmann, M.; Maturi, M. Weak gravitational lensing. *Scholarpedia*, v. 12, n. 1, p. 32440, 2017. Revision #186553. Citado na página 55.

Bender, R.; Saglia, R. P.; Gerhard, O. E. Line-of-sight velocity distributions of elliptical galaxies. , v. 269, p. 785–813, ago. 1994. Citado na página 71.

Bertotti, B.; Iess, L.; Tortora, P. A test of general relativity using radio links with the Cassini spacecraft. *Nature*, v. 425, n. 6956, p. 374–376, 2003. ISSN 1476-4687. Disponível em: <https://doi.org/10.1038/nature01997>. Citado 3 vezes nas páginas 24, 26, and 27.

Bertschinger, E. One gravitational potential or two? Forecasts and tests. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 369, n. 1957, p. 4947–4961, 2011. ISSN 1364503X. Citado 2 vezes nas páginas 26 and 28.

Bertschinger, E.; Zukin, P. Distinguishing modified gravity from dark energy. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, v. 78, n. 2, p. 1–13, 2008. ISSN 15507998. Citado na página 28.

Binney, J.; Merrifield, M. *Galactic Astronomy*. [S.l.: s.n.], 1998. Citado na página 37.

Binney, J.; Tremaine, S. *Galactic Dynamics: Second Edition*. [S.l.: s.n.], 2008. Citado 4 vezes nas páginas 19, 35, 73, and 158.

Caldwell, R.; Cooray, A.; Melchiorri, A. Constraints on a new post-general relativity cosmological parameter. *Physical Review D*, American Physical Society (APS), v. 76, n. 2, Jul 2007. ISSN 1550-2368. Disponível em: <http://dx.doi.org/10.1103/PhysRevD.76.023507>. Citado na página 28.

Cao, S. et al. Test of parametrized post-Newtonian gravity with galaxy-scale strong lensing systems. *The Astrophysical Journal*, v. 835, n. 1, p. 92, jan 2017. ISSN 1538-4357. Disponível em: <http://arxiv.org/abs/1701.00357%0Ahttp://dx.doi.org/10.3847/1538-4357/835/1/92http://stacks.iop.org/0004-637X/835/i=1/a=92?key=crossref.d066e7b8013870245a1c339c4439180fhttp://arxiv.org/abs/1701.00357http://dx.doi.org/10.3847/1538-4357/835/1/92>. Citado 5 vezes nas páginas 26, 27, 33, 129, and 132.

Cappellari, M. Efficient multi-gaussian expansion of galaxies. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press (OUP), v. 333, n. 2, p. 400–410, Jun 2002. ISSN 1365-2966. Disponível em: <http://dx.doi.org/10.1046/j.1365-8711.2002.05412.x>. Citado 3 vezes nas páginas 72, 75, and 156.

Cappellari, M. Measuring the inclination and mass-to-light ratio of axisymmetric galaxies via anisotropic Jeans models of stellar kinematics. *Monthly Notices of the Royal Astronomical Society*, v. 390, n. 1, p. 71–86, oct 2008. ISSN 00358711. Disponível em: <https://academic.oup.com/mnras/article-lookup/doi/10.1111/j.1365-2966.2008.13754.x>. Citado 5 vezes nas páginas 35, 76, 78, 79, and 156.

Cappellari, M. Improving the full spectrum fitting method: accurate convolution with gauss–hermite functions. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press (OUP), v. 466, n. 1, p. 798–811, Nov 2016. ISSN 1365-2966. Disponível em: <http://dx.doi.org/10.1093/mnras/stw3020>. Citado 4 vezes nas páginas 69, 70, 71, and 156.

Cappellari, M. Structure and kinematics of early-type galaxies from integral field spectroscopy. *Annual Review of Astronomy and Astrophysics*, Annual Reviews, v. 54, n. 1, p. 597–665, Sep 2016. ISSN 1545-4282. Disponível em: <http://dx.doi.org/10.1146/annurev-astro-082214-122432>. Citado na página 30.

Cappellari, M. Efficient solution of the anisotropic spherically aligned axisymmetric Jeans equations of stellar hydrodynamics for galactic dynamics. *Monthly Notices of the Royal Astronomical Society*, v. 494, n. 4, p. 4819–4837, jun 2020. ISSN 0035-8711. Disponível em: <http://arxiv.org/abs/1907.09894https://academic.oup.com/mnras/article/494/4/4819/5818356>. Citado 4 vezes nas páginas 35, 76, 77, and 79.

Cappellari, M. et al. The SAURON project - IV. The mass-to-light ratio, the virial mass estimator and the Fundamental Plane of elliptical and lenticular galaxies. , v. 366, n. 4, p. 1126–1150, mar. 2006. Citado na página 30.

Cappellari, M.; Copin, Y. Adaptive spatial binning of integral-field spectroscopic data using Voronoi tessellations. *MNRAS*, v. 342, p. 345–354, 2003. Citado 2 vezes nas páginas 111 and 156.

Cappellari, M.; Emsellem, E. Parametric recovery of line-of-sight velocity distributions from absorption-line spectra of galaxies via penalized likelihood. *Publications of the Astronomical Society of the Pacific*, IOP Publishing, v. 116, n. 816, p. 138–147, Feb 2004. ISSN 1538-3873. Disponível em: <http://dx.doi.org/10.1086/381875>. Citado 2 vezes nas páginas 69 and 71.

Cappellari, M. et al. The SAURON project - X. The orbital anisotropy of elliptical and lenticular galaxies: revisiting the (V/ ,) diagram with integral-field stellar kinematics. *Monthly Notices of the Royal Astronomical Society*, v. 379, n. 2, p. 418–444, aug 2007. ISSN 0035-8711. Disponível em: <https://academic.oup.com/mnras/article-lookup/doi/10.1111/j.1365-2966.2007.11963.x>. Citado na página 72.

Cappellari, M. et al. The atlas3d project – xx. mass–size and mass– distributions of early-type galaxies: bulge fraction drives kinematics, mass-to-light ratio, molecular gas fraction and stellar initial mass function. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press (OUP), v. 432, n. 3, p. 1862–1893, May 2013. ISSN 0035-8711. Disponível em: <http://dx.doi.org/10.1093/mnras/stt644>. Citado na página 30.

Cappellari, M. et al. Testing Mass Determinations of Supermassive Black Holes via Stellar Kinematics. In: Debattista, V. P.; Popescu, C. C. (Ed.). *Hunting for the Dark: the Hidden Side of Galaxy Formation*. [S.l.: s.n.], 2010. (American Institute of Physics Conference Series, v. 1240), p. 211–214. Citado na página 72.

Cappellari, M. et al. The mass of the black hole in Centaurus A from SINFONI AO-assisted integral-field observations of stellar kinematics. , v. 394, n. 2, p. 660–674, abr. 2009. Citado na página 73.

Cappellari, M. et al. The ATLAS^{3D} project - XV. Benchmark for early-type galaxies scaling relations from 260 dynamical models: mass-to-light ratio, dark matter, Fundamental Plane and Mass Plane. , v. 432, n. 3, p. 1709–1741, jul. 2013. Citado 2 vezes nas páginas 73 and 129.

Carroll, S. *Spacetime and Geometry: An Introduction to General Relativity*. [S.l.]: Pearson, 2003. ISBN 0805387323,9780805387322. Citado na página 53.

Carroll, S. M. *Spacetime and geometry. An introduction to general relativity*. [S.l.: s.n.], 2004. Citado na página 26.

Chandrasekhar, S. *Ellipsoidal figures of equilibrium*. [S.l.: s.n.], 1969. Citado na página 73.

Chirivì, G. et al. GLaD: Gravitational Lensing and Dynamics, combined analysis to unveil properties of high-redshift galaxies. p. 1–17, mar 2020. Disponível em: <http://arxiv.org/abs/2003.08404>. Citado na página 33.

Clifton, T. et al. Modified gravity and cosmology. *Physics Reports*, Elsevier B.V., v. 513, n. 1-3, p. 1–189, 2012. ISSN 03701573. Disponível em: <http://dx.doi.org/10.1016/j.physrep.2012.01.001>. Citado na página 25.

Collett, T. E. et al. A precise extragalactic test of General Relativity. *Science*, v. 360, n. 6395, p. 1342–1346, jun 2018. ISSN 0036-8075. Disponível em: <http://www.sciencemag.org/lookup/doi/10.1126/science.aao2469https://www.sciencemag.org/lookup/doi/10.1126/science.aao2469>. Citado 3 vezes nas páginas 19, 28, and 132.

Courteau, S. et al. Galaxy masses. *Reviews of Modern Physics*, American Physical Society (APS), v. 86, n. 1, p. 47–119, Jan 2014. ISSN 1539-0756. Disponível em: <http://dx.doi.org/10.1103/RevModPhys.86.47>. Citado na página 30.

Dekel, A. et al. Lost and found dark matter in elliptical galaxies. , v. 437, n. 7059, p. 707–710, set. 2005. Citado na página 71.

Di Valentino, E. A combined analysis of the h0 late time direct measurements and the impact on the dark energy sector. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press (OUP), v. 502, n. 2, p. 2065–2073, Jan 2021. ISSN 1365-2966. Disponível em: <http://dx.doi.org/10.1093/mnras/stab187>. Citado na página 93.

Diemer, B.; More, S.; Kravtsov, A. V. The pseudo-evolution of halo mass. *The Astrophysical Journal*, American Astronomical Society, v. 766, n. 1, p. 25, Mar 2013. ISSN 1538-4357. Disponível em: <http://dx.doi.org/10.1088/0004-637X/766/1/25>. Citado na página 87.

Dúmet-Montoya, H. S. *Modelagens Semianalíticas para Arcos Gravitacionais: Seção de Choque e Método Perturbativo em Lentes Pseudoelípticas*. 2012. Tese (Doutorado) - CBPF - Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, 2012. Disponível em:<http://cbpfindex.cbpf.br/index.php?module=main&moduleFile=pubDetails&pubId=5823&typeId=9>. Acesso em: 02 jul. 2021. Citado 2 vezes nas páginas 86 and 88.

Dúmet-Montoya, H. S.; Caminha, G. B.; Makler, M. Domain of validity for pseudo-elliptical NFW lens models. Mass distribution, mapping to elliptical models, and arc cross section. , v. 544, p. A83, ago. 2012. Citado na página 88.

Dvali, G.; Gabadadze, G.; Porrati, M. 4d gravity on a brane in 5d minkowski space. *Physics Letters B*, Elsevier BV, v. 485, n. 1-3, p. 208–214, Jul 2000. ISSN 0370-2693. Disponível em: <http://dx.doi.org/10.1016/S0370-2693(00)00669-9>. Citado na página 26.

Dye, S. et al. Models of the Cosmic Horseshoe gravitational lens J1004+4112. , v. 388, n. 1, p. 384–392, jul. 2008. Citado na página 83.

Dye, S. et al. Modelling high-resolution alma observations of strongly lensed highly star-forming galaxies detected by herschel. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press (OUP), v. 476, n. 4, p. 4383–4394, Feb 2018. ISSN 1365-2966. Disponível em: <http://dx.doi.org/10.1093/mnras/sty513>. Citado na página 65.

Dye, S. et al. Revealing the complex nature of the strong gravitationally lensed system H-ATLAS J090311.6+003906 using ALMA. *Monthly Notices of the Royal Astronomical Society*, v. 452, n. 3, p. 2258–2268, 2015. ISSN 13652966. Citado 4 vezes nas páginas 59, 65, 66, and 128.

Dye, S. et al. Herschel*-ATLAS: Modelling the first strong gravitational lenses. *Monthly Notices of the Royal Astronomical Society*, v. 440, n. 3, p. 2013–2025, 2014. ISSN 13652966. Citado 3 vezes nas páginas 59, 128, and 129.

Dyson, F. W.; Eddington, A. S.; Davidson, C. A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations Made at the Total Eclipse of May 29, 1919. *Philosophical Transactions of the Royal Society of London Series A*, v. 220, p. 291–333, jan. 1920. Citado na página 153.

Einstein, A. Zur elektrodynamik bewegter körper. *Annalen der Physik*, v. 322, n. 10, p. 891–921, 1905. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19053221004>. Citado na página 23.

Einstein, A. Die Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin*, p. 844–847, jan. 1915. Citado na página 24.

Emsellem, E.; Monnet, G.; Bacon, R. The multi-gaussian expansion method: a tool for building realistic photometric and kinematical models of stellar systems I. The formalism. , v. 285, p. 723–738, maio 1994. Citado na página 72.

Fabris, J. C. Aristóteles e a gravitação. *Cadernos de Astronomia*, v. 1, n. 1 SE - Seção Temática, p. 6–16, jul 2020. Disponível em: <https://periodicos.ufes.br/astronomia/article/view/30464>. Citado na página 23.

Feigelson, E. D.; Babu, G. J. *Statistical Methods for Astronomy*. 2012. Citado na página 93.

Feroz, F. et al. Importance nested sampling and the multinest algorithm. *The Open Journal of Astrophysics*, The Open Journal, v. 2, n. 1, Nov 2019. ISSN 2565-6120. Disponível em: <http://dx.doi.org/10.21105/astro.1306.2144>. Citado na página 100.

FERREIRA, P. d. C. *Arcos Gravitacionais: relações entre parâmetros locais e não locais das imagens através de simulações.* 2008. TCC (Graduação) - Curso de Astronomia, UFRJ - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2008. Disponível em:<http://cbpfindex.cbpf.br/index.php?module=main&moduleFile=pubDetails&pubId=1905&typeId=11>. Acesso em: 11 jun. 2021. Citado na página 47.

Ferreira, P. G. Cosmological Tests of Gravity. *Annual Review of Astronomy and Astrophysics*, v. 57, n. 1, p. 335–374, aug 2019. ISSN 0066-4146. Disponível em: <https://iopscience.iop.org/article/10.1088/0034-4885/79/4/046902https://www.annualreviews.org/doi/10.1146/annurev-astro-091918-104423>. Citado 2 vezes nas páginas 24 and 27.

Foreman-Mackey, D. et al. emcee: The mcmc hammer. *Publications of the Astronomical Society of the Pacific*, IOP Publishing, v. 125, n. 925, p. 306–312, Mar 2013. ISSN 1538-3873. Disponível em: <http://dx.doi.org/10.1086/670067>. Citado na página 94.

Freedman, W. L. Measurements of the Hubble Constant: Tensions in Perspective. 6 2021. Citado na página 24.

Freedman, W. L. et al. Calibration of the tip of the red giant branch. *The Astrophysical Journal*, American Astronomical Society, v. 891, n. 1, p. 57, Mar 2020. ISSN 1538-4357. Disponível em: <http://dx.doi.org/10.3847/1538-4357/ab7339>. Citado na página 24.

Gerhard, O. E. Line-of-sight velocity profiles in spherical galaxies: breaking the degeneracy between anisotropy and mass. , v. 265, p. 213, nov. 1993. Citado na página 70.

Golse, G.; Kneib, J.-P. Pseudo elliptical lensing mass model: Application to the nfw mass distribution. *Astronomy Astrophysics*, EDP Sciences, v. 390, n. 3, p. 821–827, Aug 2002. ISSN 1432-0746. Disponível em: <http://dx.doi.org/10.1051/0004-6361:20020639>. Citado na página 87.

Gonzaga, S. et al. *The DrizzlePac Handbook*. [S.l.: s.n.], 2012. Citado na página 60.

Handley, W. J.; Hobson, M. P.; Lasenby, A. N. POLYCHORD: Next-generation nested sampling. *Monthly Notices of the Royal Astronomical Society*, v. 453, n. 4, p. 4384–4398, 2015. ISSN 13652966. Citado na página 96.

Harris, C. R. et al. Array programming with NumPy. *Nature*, Springer Science and Business Media LLC, v. 585, n. 7825, p. 357–362, set. 2020. Disponível em: <https://doi.org/10.1038/s41586-020-2649-2>. Citado na página 156.

Hees, A. et al. Testing general relativity with stellar orbits around the supermassive black hole in our galactic center. *Phys. Rev. Lett.*, American Physical Society, v. 118, p. 211101, May 2017. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.118.211101>. Citado na página 24.

Higson, E. et al. Dynamic nested sampling: an improved algorithm for parameter estimation and evidence calculation. *Statistics and Computing*, Springer Science and Business Media LLC, v. 29, n. 5, p. 891–913, Dec 2018. ISSN 1573-1375. Disponível em: <http://dx.doi.org/10.1007/s11222-018-9844-0>. Citado na página 99.

Higson, E. et al. Sampling Errors in Nested Sampling Parameter Estimation. *Bayesian Analysis*, v. 13, n. 3, p. 873–896, mar. 2018. Citado na página 99.

Hogg, D. W.; Foreman-Mackey, D. Data Analysis Recipes: Using Markov Chain Monte Carlo. *arXiv*, IOP Publishing, v. 236, n. 1, p. 11, 2017. ISSN 23318422. Disponível em: <http://dx.doi.org/10.3847/1538-4365/aab76e>. Citado na página 93.

Hunter, J. D. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, IEEE COMPUTER SOC, v. 9, n. 3, p. 90–95, 2007. Citado na página 156.

Ishak, M. Testing general relativity in cosmology. *Living Reviews in Relativity*, Springer International Publishing, v. 22, n. 1, p. 1, dec 2019. ISSN 2367-3613. Disponível em: <https://doi.org/10.1007/s41114-018-0017-4http://link.springer.com/10.1007/s41114-018-0017-4>. Citado na página 25.

Jain, B.; Zhang, P. Observational tests of modified gravity. *Physical Review D*, American Physical Society (APS), v. 78, n. 6, Sep 2008. ISSN 1550-2368. Disponível em: <http://dx.doi.org/10.1103/PhysRevD.78.063503>. Citado na página 28.

Katz, D. et al. Gaia data release 2. *Astronomy Astrophysics*, EDP Sciences, v. 616, p. A11, Aug 2018. ISSN 1432-0746. Disponível em: <http://dx.doi.org/10.1051/0004-6361/201832865>. Citado na página 38.

Keeton, C. R.; Kochanek, C. S.; Seljak, U. Shear and ellipticity in gravitational lenses. *The Astrophysical Journal*, American Astronomical Society, v. 482, n. 2, p. 604–620, Jun 1997. ISSN 1538-4357. Disponível em: <http://dx.doi.org/10.1086/304172>. Citado na página 57.

Kipper, R. et al. Stellar kinematics using a third integral of motion: method and application on the Andromeda galaxy. *Monthly Notices of the Royal Astronomical Society*, v. 460, n. 3, p. 2720–2730, aug 2016. ISSN 0035-8711. Disponível em: <https://academic.oup.com/mnras/article-lookup/doi/10.1093/mnras/stw1194>. Citado na página 93.

Koopmans, L. Gravitational lensing stellar dynamics. *EAS Publications Series*, EDP Sciences, v. 20, p. 161–166, 2006. ISSN 1638-1963. Disponível em: <http://dx.doi.org/10.1051/eas:2006064>. Citado na página 33.

Koopmans, L. V. E.; Treu, T. The Structure and Dynamics of Luminous and Dark Matter in the Early-Type Lens Galaxy of 0047281 at z = 0.485 . *The Astrophysical Journal*, v. 583, n. 2, p. 606–615, 2003. ISSN 0004-637X. Citado na página 33.

Lam, S. K.; Pitrou, A.; Seibert, S. Numba: A llvm-based python jit compiler. In: *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*. New York, NY, USA: Association for Computing Machinery, 2015. (LLVM '15). ISBN 9781450340052. Disponível em: <https://doi.org/10.1145/2833157.2833162>. Citado na página 156.

Ma, C.-P.; Bertschinger, E. Cosmological Perturbation Theory in the Synchronous and Conformal Newtonian Gauges. *The Astrophysical Journal*, v. 455, p. 7, 1995. ISSN 0004-637X. Citado 2 vezes nas páginas 27 and 28.

Malajovich, G. *Álgebra Linear*. 3. ed. [s.n.], 2010. Disponível em: <https://www.labma.ufrj.br/~gregorio/livro/al2.pdf>. Citado 2 vezes nas páginas 53 and 77.

McMullin, J. P. et al. CASA Architecture and Applications. In: Shaw, R. A.; Hill, F.; Bell, D. J. (Ed.). *Astronomical Data Analysis Software and Systems XVI*. [S.l.: s.n.], 2007. (Astronomical Society of the Pacific Conference Series, v. 376), p. 127. Citado na página 64.

Meneghetti, M. Introduction to gravitational lensing. *Lecture note*, n. 0, p. 103, 2016. Disponível em: <https://www.ita.uni-heidelberg.de/~jmerten/misc/meneghetti_lensing.pdf>. Citado 4 vezes nas páginas 19, 32, 47, and 151.

Mitzkus, M.; Cappellari, M.; Walcher, C. J. Dominant dark matter and a counter-rotating disc: MUSE view of the low-luminosity S0 galaxy NGC 5102. , v. 464, n. 4, p. 4789–4806, fev. 2017. Citado na página 129.

Mo, H.; van den Bosch, F. C.; White, S. *Galaxy Formation and Evolution*. [S.l.: s.n.], 2010. Citado na página 26.

Mortonson, M. J.; Weinberg, D. H.; White, M. *Dark Energy: A Short Review*. 2013. Citado na página 24.

Mukhanov et al., V. F. Theory of cosmological perturbations. , v. 215, n. 5-6, p. 203–333, jun. 1992. Citado na página 27.

Mukherjee, P.; Parkinson, D.; Liddle, A. R. A nested sampling algorithm for cosmological model selection. *The Astrophysical Journal*, American Astronomical Society, v. 638, n. 2, p. L51–L54, Jan 2006. ISSN 1538-4357. Disponível em: <http://dx.doi.org/10.1086/501068>. Citado na página 100.

Multamäki, T.; Vilja, I. Cosmological expansion and the uniqueness of the gravitational action. *Phys. Rev. D*, American Physical Society, v. 73, p. 024018, Jan 2006. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevD.73.024018>. Citado na página 26.

Multamäki, T.; Vilja, I. Spherically symmetric solutions of modified field equations in f(R) theories of gravity. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, v. 74, n. 6, p. 1–5, 2006. ISSN 15507998. Citado na página 26.

Nakahara, M. *Geometry, topology and physics*. [S.l.: s.n.], 2003. Citado na página 54.

Navarro, J. F.; Frenk, C. S.; White, S. D. M. A Universal Density Profile from Hierarchical Clustering. , v. 490, n. 2, p. 493–508, dez. 1997. Citado na página 86.

Negrello, M. et al. The Herschel-ATLAS: a sample of 500 μm-selected lensed galaxies over 600 deg^2. , v. 465, n. 3, p. 3558–3580, mar. 2017. Citado na página 66.

Negrello, M. et al. The Detection of a Population of Submillimeter-Bright, Strongly Lensed Galaxies. *Science*, v. 330, n. 6005, p. 800–804, nov 2010. ISSN 0036-8075. Disponível em: <https://www.sciencemag.org/lookup/doi/10.1126/science.1193420>. Citado na página 59.

Negrello, M. et al. Herschel *-ATLAS: Deep HST/WFC3 imaging of strongly lensed submillimetre galaxies. *Monthly Notices of the Royal Astronomical Society*, v. 440, n. 3, p. 1999–2012, 2014. ISSN 13652966. Citado 2 vezes nas páginas 33 and 59.

Nelson, B.; Ford, E. B.; Payne, M. J. Run DMC: An efficient, parallel code for analyzing radial velocity observations using N-body integrations and differential evolution Markov Chain Monte Carlo. *Astrophysical Journal, Supplement Series*, v. 210, n. 1, 2014. ISSN 00670049. Citado na página 93.

Nightingale, J. W.; Dye, S. Adaptive semi-linear inversion of strong gravitational lens imaging. , v. 452, n. 3, p. 2940–2959, set. 2015. Citado 2 vezes nas páginas 83 and 84.

Nightingale, J. W.; Dye, S.; Massey, R. J. AutoLens: automated modeling of a strong lens's light, mass, and source. , v. 478, n. 4, p. 4738–4784, ago. 2018. Citado 2 vezes nas páginas 84 and 108.

Nightingale, J. W. et al. 'pyautolens': Open-source strong gravitational lensing. *Journal of Open Source Software*, The Open Journal, v. 6, n. 58, p. 2825, 2021. Disponível em: <https://doi.org/10.21105/joss.02825>. Citado na página 156.

Nitschai, M. S.; Cappellari, M.; Neumayer, N. First Gaia dynamical model of the Milky Way disc with six phase space coordinates: a test for galaxy dynamics. , v. 494, n. 4, p. 6001–6011, jun. 2020. Citado na página 86.

O'Neill, B. *Semi-Riemannian Geometry With Applications to Relativity, 103, Volume 103 (Pure and Applied Mathematics)*. [S.l.]: Academic Press, 1983. ISBN 0125267401. Citado na página 54.

Piattella, O. F. O artigo fundador da teoria da relatividade restrita: Sobre a eletrodinâmica dos corpos em movimento. *Cadernos de Astronomia*, v. 1, n. 1, p. 157–176, 2020. Disponível em: <https://periodicos.ufes.br/astronomia/article/view/31681>. Citado na página 23.

Piqueras, L. et al. *MPDAF - A Python package for the analysis of VLT/MUSE data*. 2017. Citado na página 156.

Pizzuti, L. et al. CLASH-VLT: Testing the nature of gravity with galaxy cluster mass profiles. *Journal of Cosmology and Astroparticle Physics*, v. 2016, n. 4, 2016. ISSN 14757516. Citado na página 28.

Price-Whelan, A. M.; Foreman-Mackey, D. schwimmbad: A uniform interface to parallel processing pools in python. *The Journal of Open Source Software*, The Open Journal, v. 2, n. 17, sep 2017. Disponível em: <https://doi.org/10.21105/joss.00357>. Citado 2 vezes nas páginas 101 and 156.

Price-Whelan, A. M. et al. The astropy project: Building an open-science project and status of the v2.0 core package. *The Astronomical Journal*, American Astronomical Society, v. 156, n. 3, p. 123, Aug 2018. ISSN 1538-3881. Disponível em: <http://dx.doi.org/10.3847/1538-3881/aabc4f>. Citado na página 156.

Remijan, A. et al. *ALMA Cycle 3 Technical Handbook Version 1.0*. [S.l.: s.n.], 2015. 1–192 p. ISBN 9783923524662. Citado na página 64.

Riess, A. G. The expansion of the universe is faster than expected. *Nature Reviews Physics*, Springer Science and Business Media LLC, v. 2, n. 1, p. 10–12, Dec 2019. ISSN 2522-5820. Disponível em: <http://dx.doi.org/10.1038/s42254-019-0137-0>. Citado 2 vezes nas páginas 24 and 25.

Riess, A. G. et al. Cosmic distances calibrated to 1edr3 parallaxes and hubble space telescope photometry of 75 milky way cepheids confirm tension with cdm. *The Astrophysical Journal*, American Astronomical Society, v. 908, n. 1, p. L6, Feb 2021. ISSN 2041-8213. Disponível em: <http://dx.doi.org/10.3847/2041-8213/abdbaf>. Citado na página 24.

Rubin, V. C.; Ford, W. K. J. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. , v. 159, p. 379, fev. 1970. Citado na página 30.

Rybak, M. et al. ALMA imaging of SDP.81-I. A pixelated reconstruction of the far-infrared continuum emission. *Monthly Notices of the Royal Astronomical Society: Letters*, v. 451, n. 1, p. L40–L44, 2015. ISSN 17453933. Citado 2 vezes nas páginas 59 and 128.

Salvatier, J.; Wiecki, T. V.; Fonnesbeck, C. Probabilistic programming in python using PyMC3. *PeerJ Computer Science*, PeerJ, v. 2, p. e55, abr. 2016. Disponível em: <https://doi.org/10.7717/peerj-cs.55>. Citado na página 94.

Sanghai, V. A.; Clifton, T. Parameterized post-Newtonian cosmology. *Classical and Quantum Gravity*, IOP Publishing, v. 34, n. 6, 2017. ISSN 13616382. Citado na página 29.

Schneider, P.; Ehlers, J.; Falco, E. E. *Gravitational Lenses*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1992. 285–286 p. (Astronomy and Astrophysics Library). ISBN 978-3-540-66506-9. Disponível em: <http://link.springer.com/10.1007/978-3-662-03758-4>. Citado 3 vezes nas páginas 47, 55, and 151.

Schwab, J.; Bolton, A. S.; Rappaport, S. A. GALAXY-SCALE STRONG-LENSING TESTS OF GRAVITY AND GEOMETRIC COSMOLOGY: CONSTRAINTS AND SYSTEMATIC LIMITATIONS. *The Astrophysical Journal*, v. 708, n. 1, p. 750–757, jan 2010. ISSN 0004-637X. Disponível em: <http://stacks.iop.org/0004-637X/708/i=1/a=750?key=crossref.cdd62a8f9202a4043559b31c60389076>. Citado 2 vezes nas páginas 26 and 27.

Science Software Branch at STScI. *PyRAF: Python alternative for IRAF*. 2012. ascl:1207.011 p. Citado na página 156.

Simpson, F. et al. Cfhtlens: testing the laws of gravity with tomographic weak lensing and redshift-space distortions. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press (OUP), v. 429, n. 3, p. 2249–2263, Dec 2012. ISSN 0035-8711. Disponível em: <http://dx.doi.org/10.1093/mnras/sts493>. Citado na página 27.

Skilling, J. Nested sampling for general Bayesian computation. *Bayesian Analysis*, v. 1, n. 4, p. 833–860, 2006. ISSN 19360975. Citado 2 vezes nas páginas 94 and 98.

Sofue, Y.; Rubin, V. Rotation curves of spiral galaxies. *Annual Review of Astronomy and Astrophysics*, Annual Reviews, v. 39, n. 1, p. 137–174, Sep 2001. ISSN 1545-4282. Disponível em: <http://dx.doi.org/10.1146/annurev.astro.39.1.137>. Citado na página 30.

Sonnenfeld, A. et al. The sl2s galaxy-scale lens sample. v. dark matter halos and stellar imf of massive early-type galaxies out to redshift 0.8. *The Astrophysical Journal*, American Astronomical Society, v. 800, n. 2, p. 94, Feb 2015. ISSN 1538-4357. Disponível em: <http://dx.doi.org/10.1088/0004-637X/800/2/94>. Citado na página 129.

Soto, K. T. et al. ZAP - enhanced PCA sky subtraction for integral field spectroscopy. , v. 458, n. 3, p. 3210–3220, maio 2016. Citado na página 62.

Speagle, J. S. dynesty: a dynamic nested sampling package for estimating bayesian posteriors and evidences. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press (OUP), v. 493, n. 3, p. 3132–3158, Feb 2020. ISSN 1365-2966. Disponível em: <http://dx.doi.org/10.1093/mnras/staa278>. Citado 5 vezes nas páginas 93, 95, 99, 100, and 156.

Stetson, P. B. DAOPHOT: A Computer Program for Crowded-Field Stellar Photometry. , v. 99, p. 191, mar. 1987. Citado na página 60.

Suyu, S. H. et al. Two accurate time-delay distances from strong lensing: Implications for cosmology. *The Astrophysical Journal*, IOP Publishing, v. 766, n. 2, p. 70, Mar 2013. ISSN 1538-4357. Disponível em: <http://dx.doi.org/10.1088/0004-637X/766/2/70>. Citado na página 31.

Suyu, S. H. et al. A Bayesian analysis of regularized source inversions in gravitational lensing. , v. 371, n. 2, p. 983–998, set. 2006. Citado na página 83.

Tamura, Y. et al. High-resolution ALMA observations of SDP.81. I. The innermost mass profile of the lensing elliptical galaxy probed by 30 milli-arcsecond images. *Publications of the Astronomical Society of Japan*, v. 67, n. 4, p. 72, aug 2015. ISSN 0004-6264. Disponível em: <http://arxiv.org/abs/1702.02157http://dx.doi.org/10.1093/pasj/psx011https://academic.oup.com/pasj/article-lookup/doi/10.1093/pasj/psx011https://academic.oup.com/pasj/article-lookup/doi/10.1093/pasj/psv040>. Citado 2 vezes nas páginas 59 and 128.

Toniato, J. D.; Rodrigues, D. C. Post-newtonian γ-like parameters and the gravitational slip in scalar-tensor and f(r) theories. *Physical Review D*, American Physical Society (APS), v. 104, n. 4, Aug 2021. ISSN 2470-0029. Disponível em: <http://dx.doi.org/10.1103/PhysRevD.104.044020>. Citado na página 29.

Valdes, F. et al. The Indo-US Library of Coudé Feed Stellar Spectra. , v. 152, n. 2, p. 251–259, jun. 2004. Citado na página 112.

van de Ven, G. et al. The einstein cross: Constraint on dark matter from stellar dynamics and gravitational lensing. *The Astrophysical Journal*, American Astronomical Society, v. 719, n. 2, p. 1481–1496, Jul 2010. ISSN 1538-4357. Disponível em: <http://dx.doi.org/10.1088/0004-637X/719/2/1481>. Citado 2 vezes nas páginas 72 and 162.

van der Marel, R. P.; Franx, M. A New Method for the Identification of Non-Gaussian Line Profiles in Elliptical Galaxies. , v. 407, p. 525, abr. 1993. Citado na página 70.

van der Velden, E. CMasher: Scientific colormaps for making accessible, informative and 'cmashing' plots. *The Journal of Open Source Software*, v. 5, n. 46, p. 2004, fev. 2020. Citado na página 156.

Vazdekis, A. et al. Evolutionary stellar population synthesis with MILES - I. The base models and a new line index system. , v. 404, n. 4, p. 1639–1671, jun. 2010. Citado na página 71.

Virtanen, P. et al. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature Methods*, Springer Science and Business Media LLC, v. 17, n. 3, p. 261–272, Feb 2020. ISSN 1548-7105. Disponível em: <http://dx.doi.org/10.1038/s41592-019-0686-2>. Citado na página 156.

Vlahakis, C. et al. The 2014 alma long baseline campaign: Observations of the strongly lensed submillimeter galaxy hatlas j090311.6+003906 at z = 3.042. *The Astrophysical Journal*, American Astronomical Society, v. 808, n. 1, p. L4, Jul 2015. ISSN 2041-8213. Disponível em: <http://dx.doi.org/10.1088/2041-8205/808/1/L4>. Citado na página 64.

Warren, S. J.; Dye, S. Semilinear Gravitational Lens Inversion. , v. 590, n. 2, p. 673–682, jun. 2003. Citado 2 vezes nas páginas 79 and 82.

Weilbacher, P. M.; Streicher, O.; Palsa, R. *MUSE-DRP: MUSE Data Reduction Pipeline*. 2016. ascl:1610.004 p. Citado na página 62.

Will, C. M. The confrontation between general relativity and experiment. *Living Reviews in Relativity*, v. 17, 2014. ISSN 14338351. Citado 2 vezes nas páginas 24 and 26.

Will, C. M. *Theory and Experiment in Gravitational Physics*. Cambridge: Cambridge University Press, 2018. ISSN 0736-8046. ISBN 9781316338612. Disponível em: <https://www.cambridge.org/core/product/identifier/CBO9781107415324A009/type/book_parthttp://www.ncbi.nlm.nih.gov/pubmed/2359728https://www.cambridge.org/core/product/identifier/9781316338612/type/bookhttp://arxiv.org/abs/1011.1669http://dx.doi.org/10.> Citado na página 26.

Williams, M. J.; Bureau, M.; Cappellari, M. Kinematic constraints on the stellar and dark matter content of spiral and S0 galaxies. , v. 400, n. 4, p. 1665–1689, dez. 2009. Citado 2 vezes nas páginas 73 and 86.

Willmer, C. N. A. The absolute magnitude of the sun in several filters. *The Astrophysical Journal Supplement Series*, American Astronomical Society, v. 236, n. 2, p. 47, Jun 2018. ISSN 1538-4365. Disponível em: <http://dx.doi.org/10.3847/1538-4365/aabfdf>. Citado na página 119.

Wong, K. C. et al. H0licow – iv. lens mass model of he 04351223 and blind measurement of its time-delay distance for cosmology. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press (OUP), v. 465, n. 4, p. 4895–4913, Nov 2016. ISSN 1365-2966. Disponível em: <http://dx.doi.org/10.1093/mnras/stw3077>. Citado na página 31.

Wong, K. C. et al. H0licow – xiii. a 2.4 per cent measurement of h0 from lensed quasars: 5.3 tension between early- and late-universe probes. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press (OUP), v. 498, n. 1, p. 1420–1439, Sep 2019. ISSN 1365-2966. Disponível em: <http://dx.doi.org/10.1093/mnras/stz3094>. Citado na página 24.

Wong, K. C.; Suyu, S. H.; Matsushita, S. THE INNERMOST MASS DISTRIBUTION of the GRAVITATIONAL LENS SDP.81 from ALMA OBSERVATIONS. *Astrophysical*

Journal, IOP Publishing, v. 811, n. 2, p. 115, 2015. ISSN 15384357. Disponível em: <http://dx.doi.org/10.1088/0004-637X/811/2/115>. Citado na página 59.

Wright, C. O.; Brainerd, T. G. Gravitational Lensing by NFW Halos. , v. 534, n. 1, p. 34–40, maio 2000. Citado na página 87.

Wyithe, J. S. B.; Turner, E. L.; Spergel, D. N. Gravitational lens statistics for generalized nfw profiles: Parameter degeneracy and implications for self-interacting cold dark matter. *The Astrophysical Journal*, American Astronomical Society, v. 555, n. 1, p. 504–523, Jul 2001. ISSN 1538-4357. Disponível em: <http://dx.doi.org/10.1086/321437>. Citado 2 vezes nas páginas 86 and 87.

Yuen, K.-V. Appendix A: Relationship between the Hessian and Covariance Matrix for Gaussian Random Variables. In: _____. *Bayesian Methods for Structural Dynamics and Civil Engineering*. Chichester, UK: John Wiley Sons, Ltd, 2010. p. 257–262. ISBN 9780470824566. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470824566.app1http://doi.wiley.com/10.1002/9780470824566.app1>. Citado na página 81.

ZHAO, H. Analytical models for galactic nuclei. *Monthly Notices of the Royal Astronomical Society*, Oxford University Press (OUP), v. 278, n. 2, p. 488–496, Jan 1996. ISSN 1365-2966. Disponível em: <http://dx.doi.org/10.1093/mnras/278.2.488>. Citado na página 86.

Zwicky, F. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, v. 6, p. 110–127, jan. 1933. Citado na página 24.

Appendix

APPENDIX A – Point mass deflection angle in PPN approach

Following Meneghetti (2016) (hereafter, MG16) we want to show some insights for the calculation of the point mass deflection angle using the line element (1.2) under the assumption made in the Section 1.2, just to remember: i) The space-time metric is given by the line element (1.2), which is in the Newtonian gauge and consider only scalar perturbations; ii) There is a well-defined Newtonian limit, where the potentials Φ and Ψ still follow the Poisson equation; iii) The gravitational slip parameter is constant on the relevant scales being studied. For a more rigorous derivation of the deflection angle, see also Schneider; Ehlers; Falco (1992).

Starting by the line element (1.2), consider a null space-time interval, i.e, $dS^2 = 0$. In Cartesian coordinates and flat space (remember, we are considering the weak field approximation) we have:

$$(1 + 2\frac{\Phi}{c^2})c^2dt^2 = (1 - 2\frac{\Psi}{c^2})dr^2, \tag{A.1}$$

were we denote $dr = dxdydz$. So, the speed of light in the gravitational field is

$$c' \equiv \frac{dr}{dt} = c\sqrt{\frac{(1 + 2\Phi/c^2)}{(1 - 2\Psi/c^2)}}. \tag{A.2}$$

By construction, we have assumed that Φ/c^2 and Ψ/c^2 are much smaller than one, so we can expand (A.2) in terms of Taylor series,

$$\begin{aligned}
\sqrt{1 + 2\Phi/c^2} &= 1 + \frac{1}{2}\frac{2\Phi}{c^2} + \mathcal{O}^2(\Phi/c^2), \\
\frac{1}{\sqrt{1 - 2\Psi/c^2}} &= 1 + \frac{1}{2}\frac{2\Psi}{c^2} + \mathcal{O}^2(\Psi/c^2), \\
\sqrt{\frac{(1 + 2\Phi/c^2)}{(1 - 2\Psi/c^2)}} &= \left[1 + \Phi/c^2 + \mathcal{O}^2(\Phi/c^2)\right]\left[1 + \Psi/c^2 + \mathcal{O}^2(\Psi/c^2)\right] \\
&= 1 + \frac{(\Phi + \Psi)}{c^2} + \mathcal{O}^2 \approx 1 + \frac{(\Phi + \Psi)}{c^2}.
\end{aligned} \tag{A.3}$$

Thus, the refraction index eq. (1.20) of MG16 is just,

$$n = c/c' = \frac{1}{1 + (\Phi + \Psi)/c^2} \approx 1 - \frac{\Phi + \Psi}{c^2}. \tag{A.4}$$

Comparing this expression with the one found in MG16 (eq. 1.20), we can see that the refraction index in PPN is modified only by the Ψ potential, and when $\Psi = \Phi$ (GR case), we recovery the standard one. Now, assuming the parametrization $\gamma_{PPN} = \Psi/\Phi$, we rewrite the refraction index above in terms of $\gamma_{\rm PPN}$,

$$n = 1 - (1 + \gamma_{\rm PPN})\frac{\Phi}{c^2}. \tag{A.5}$$

And, if $\gamma_{\rm PPN}$ is constant, we can just replace $2\Phi \to (1+\gamma_{\rm PPN})\Phi$ in MG16 equations, where Φ still the Newtonian potential. Therefore, the deflection of a light ray coming from the $+\hat{z}$ direction and passing through the lens in $|z| = 0$, with impact parameter $\boldsymbol{\xi}$, is given by

$$\tilde{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{(1+\gamma_{\rm PPN})}{c^2}\int_{-\infty}^{+\infty} \boldsymbol{\nabla}_\perp \Phi dz, \tag{A.6}$$

where $\boldsymbol{\nabla}_\perp$ is the gradient perpendicular to the light path. This equation is the same as (1.36) in MG16, when GR is assumed and $\gamma_{\rm PPN} = 1$.

For completeness, consider a lens formed by a point mass. Its gravitational potential is

$$\Phi(r) = -\frac{GM}{r}, \tag{A.7}$$

with $r = \sqrt{x^2+y^2+z^2}$ and M the total mass. If the light is traveling parallel to z-axis, the modulus of the impact parameter is $|\boldsymbol{\xi}| = \sqrt{x^2+y^2} \equiv \xi$, and the potential can be written as

$$\Phi(r) = -\frac{GM}{\sqrt{\xi^2+z^2}}.$$

On the other hand,

$$\boldsymbol{\nabla}_\perp \Phi = \begin{bmatrix} \nabla_x \Phi(x,y,z) \\ \nabla_y \Phi(x,y,z) \end{bmatrix} = \frac{GM}{r^3}\begin{bmatrix} x \\ y \end{bmatrix},$$

$$\tilde{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = (1+\gamma_{\rm PPN})\frac{GM}{c^2}\begin{bmatrix} x \\ y \end{bmatrix}\int_{-\infty}^{+\infty}\frac{1}{(\xi^2+z^2)^{3/2}}dz. \tag{A.8}$$

This integral can be solved applying a trigonometric transformation ($z = \xi\tan\theta$)[1]. The final result is just

$$\tilde{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = (1+\gamma_{\rm PPN})\frac{2GM}{\xi^2 c^2}\begin{bmatrix} x \\ y \end{bmatrix} = (1+\gamma_{\rm PPN})\frac{2GM}{c^2}\frac{\hat{\boldsymbol{\xi}}}{\xi}, \tag{A.9}$$

where we define the versor $\hat{\boldsymbol{\xi}} = \boldsymbol{\xi}/\xi$, perpendicular to the light ray.

[1] Or using the computer algebra software you like more.

Once again, if we assume GR, the last equation is reduced to the deflection angle derived by Einstein in 1915,

$$\tilde{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4GM}{c^2}\frac{\hat{\boldsymbol{\xi}}}{\xi}, \tag{A.10}$$

and confirmed in 1920 by Dyson; Eddington; Davidson (1920).

APPENDIX B – List of softwares and packages

All the scripts and data used in this work can be accessed through the link <https://github.com/carlosRmelo/Master-Degree.git>. Some documentation is still under developmente, but most of the scripts contains comments and Jupyter Notebooks[1] (which come with a '.ipynb' suffix) guiding the reader. Comments, suggestions and doubts are always welcome, so feel free to send an email to carlos.melo@ufrgs.br.

The scripts were written mostly for the Python Programming Language[2], using the Python version 3.7.6. They are also tested with Python version 3.6.3, and no bugs are found. The Operational system where the scripts were developed is Ubuntu 20.04.2 LTS (Focal Fossa)[3], with architecture x86_64.

We also emphasize that whenever possible we try to make our plots more accessible, using the CMasher color library, that tries to make colormaps more accessible, and informative to color-vision deficiency (CVD; color blindness) people.

The table bellow shows the principal Python packages used in this work. The first column is the package name, second column is the version of package used in our implementation, third column the installation method, and finally the fourth column is the related paper. Some of these packages have their own dependencies, that may not be listed in this table. We strongly recommend the reader check the documentation of the packages for more information, and even for tutorials and *"how to use"* protocols.

[1] <https://jupyter.org/>
[2] <https://www.python.org/>
[3] <https://releases.ubuntu.com/20.04/>

Python Packages List			
Package	Version	Installation method	Related paper
Astropy	4.1	pip	Price-Whelan et al., 2018
CMasher	1.6.2	pip	van der Velden, 2020
dynesty	1.1	pip	Speagle, 2020
JamPy	6.1.2	pip	Cappellari, 2008
Matplotlib	3.3.3	pip	Hunter, 2007
MgeFit	5.0.13	pip	Cappellari, 2002
MPDAF	3.4	pip	Piqueras et al., 2017
Numba	0.50.1	pip	Lam; Pitrou; Seibert, 2015
NumPy	1.18.5	pip	Harris et al., 2020
PlotBin	3.1.3	pip	——
pPXF	7.2.1	pip	Cappellari, 2016a
PyAutoLens	1.1.4	pip	Nightingale et al., 2021
PyRAF	2.1.15	Astroconda Channel[4]	Science Software Branch at STScI, 2012
schwimmbad	0.3.1	pip	Price-Whelan; Foreman-Mackey, 2017
SciPy	1.5.2	pip	Virtanen et al., 2020
VorBin	3.1.4	pip	Cappellari; Copin, 2003

[4] The instalation follows this tutorial: <https://youtu.be/3EXJRdH2mSM>

APPENDIX C – MGE and Lensing Formalism

According to MGE formalism (Section 5.2), the projected surface brightness for a galaxy can be written as

$$I(x',y') = \sum_{j=1}^{N} \frac{L_j}{2\pi\sigma_j^2 q'_j} \exp\left[-\frac{1}{2\pi\sigma_j^2}\left(x_j'^2 + \frac{y_j'^2}{q_j'^2}\right)\right]. \tag{C.1}$$

This parametrization assumes that all Gaussians are concentric and centered in the galaxy center, whose position coincides with the origin of the coordinate system. We can easily convert this superficial brightness distribution into a superficial mass distribution multiplying it by a mass-to-light ratio

$$\Upsilon_j = M_j/L_j. \tag{C.2}$$

Thus, we define the surface mass density as

$$\Sigma(x',y') = \sum_{j=1}^{N} \frac{M_j}{2\pi\sigma_j^2 q'_j} \exp\left[-\frac{1}{2\pi\sigma_j^2}\left(x_j'^2 + \frac{y_j'^2}{q_j'^2}\right)\right] \tag{C.3}$$

Once the gravitational lensing phenomenon is sensible to the projected mass along the line-of-sight, the expression (C.3) is appropriate to compute the deflection angle (3.12). However, instead of computing the deflection angle directly from the projected mass, we prefer (for the MGE case) to do that through the lens potential $\psi(x',y')$. From the gravitational lensing theory, the deflection angle is related to the lens potential by (eq. 3.17),

$$\boldsymbol{\alpha} = \boldsymbol{\nabla}\psi(x',y'), \tag{C.4}$$

where $\boldsymbol{\alpha}$ is the reduced deflection angle, $\boldsymbol{\nabla}$ is the gradient operator in two dimensions written in coordinates of the sky plane, and $\psi(x',y')$ is the lens potential. The lens potential can be obtained from the Poisson equation (3.19),

$$\nabla^2\psi(x',y') = 2\kappa(x',y'), \tag{C.5}$$

where $\kappa(x',y')$ is the convergence, or the projected mass density per unit of critical density, eq. (3.13).

To calculate the deflection angle, we consider that the projected mass density is given by (C.3), and from it we derive the lens potential using (C.5).

We start pointing out that the distribution (C.3) is given by a set of concentric ellipses, in a such a way that we can use the Homoeoid Theorem (Binney; Tremaine 2008, equation 2.140). The theorem states that, if a mass distribution is in the form

$$\rho = \rho(m^2), \quad m^2 = a_1 \sum_i \frac{x_i}{a_i^2}, \tag{C.6}$$

then the gravitational potential associated is

$$\Phi(\boldsymbol{r}) = -\pi G \left[\frac{a_2 a_3}{a_1}\right] \int_0^\infty \frac{[\Gamma(\infty) - \Gamma(m(\tau))]}{[(\tau + a_1^2)(\tau + a_2^2)(\tau + a_3^2)]^{1/2}} d\tau, \tag{C.7}$$

where

$$\Gamma(m(\tau)) = \int_0^{m^2(\tau)} \rho(m^2) dm^2, \quad m^2(\tau) = a_1 \sum_i \frac{x_i}{a_i^2 + \tau}. \tag{C.8}$$

In these expressions, m^2 defines an elliptical concentric coordinate, the index i varies from $(1, 2, 3) = (x, y, z)$[1], and the coefficients a_i are the semi-axis of the ellipsoid formed by the coordinate m^2. Beside that, the parameter $\tau \geq 0$ labels the concentric surfaces of the ellipsoid.

This theorem assumes that the Poisson equation is in the classical form $\nabla^2\Phi(\boldsymbol{r}) = 4\pi G\rho(m^2)$. For this reason, we re-write the equation (C.5) as

$$\begin{aligned} \nabla^2\psi &= 2\frac{\Sigma(x', y')}{\Sigma_c} = 2\frac{4\pi G}{c^2}\frac{D_L D_{LS}}{D_S}\Sigma(x', y') \\ &= 4\pi G\left[\frac{2}{c^2}\frac{D_L D_{LS}}{D_S}\Sigma(x', y')\right] \\ &= 4\pi G\hat{\Sigma}(x', y'), \end{aligned} \tag{C.9}$$

where we define

$$\hat{\Sigma}(x', y') = \frac{2}{c^2}\frac{D_L D_{LS}}{D_S}\Sigma(x', y'). \tag{C.10}$$

Therefore, we will use the equations (C.7) and (C.8) with Φ replaced by ψ, and ρ replaced by $\hat{\Sigma}$.

Finally we can use the equation (C.7) to compute the lens potential $\psi(x', y')$. We define the coordinate

$$m_j^2 = x'^2 + \frac{y'^2}{q_j'^2}, \tag{C.11}$$

so, the projected mass profile is

$$\hat{\Sigma}(x', y') = \hat{\Sigma}(m^2) = \frac{2}{c^2}\frac{D_L D_{LS}}{D_S}\sum_j \frac{M_j}{2\pi\sigma_j^2 q_j'} \exp\left[-\frac{1}{2\pi\sigma_j^2} m_j^2\right]. \tag{C.12}$$

[1] It is important to note that the theorem still valid for different dimensions, without losing the generality.

In addition, each Gaussian component is independent, which turns possible to treat only one j component and in the end to sum over all j. So, we define

$$\hat{\Sigma}_j(x', y') = A_j \exp\left[-\frac{1}{2\pi\sigma_j^2} m_j^2\right], \tag{C.13}$$

with

$$A_j \equiv \frac{2}{c^2}\frac{D_L D_{LS}}{D_S}\frac{M_j}{2\pi\sigma_j^2 q_j'}. \tag{C.14}$$

We start by calculating the term $\Gamma(\infty)$. Using equation (C.8), we get

$$\Gamma_j(\infty) = \int_0^\infty \hat{\Sigma}_j(x', y') dm_j^2 = 2A_j \int_0^\infty m_j \exp\left[-\frac{1}{2\sigma_j^2} m_j^2\right] dm_j, \tag{C.15}$$

which is a simple Gaussian integration, whose result is

$$\begin{aligned} \Gamma_j(\infty) &= -2A_j\sigma_j^2 \exp\left[-\frac{1}{2\sigma_j^2} m_j^2\right]\Bigg|_0^\infty \\ \Gamma_j(\infty) &= 2A_j\sigma_j^2. \end{aligned} \tag{C.16}$$

Similarly,

$$\begin{aligned} \Gamma_j(m_j(\tau)) &= -2A_j\sigma_j^2 \exp\left[-\frac{1}{2\sigma_j^2} m_j'^2\right]\Bigg|_0^{m_j(\tau)} \\ \Gamma_j(m_j(\tau)) &= -2A_j\sigma_j^2 \left(\exp\left[-\frac{1}{2\sigma_j^2} m_j^2(\tau)\right] - 1\right). \end{aligned} \tag{C.17}$$

Thus,

$$\Gamma_j(\infty) - \Gamma_j(m_j(\tau)) = 2A_j\sigma_j^2 \exp\left[-\frac{1}{2\sigma_j^2} m_j^2(\tau)\right]. \tag{C.18}$$

Identifying $a_1 = 1$ and $a_2 = q_j'$ in equation (C.11), the potential (C.7), here the two-dimensional lens potential, can be placed as

$$\begin{aligned} \psi_j(x', y') &= -\pi G q_j' \int_0^\infty \frac{2A_j\sigma_j^2 \exp\left[-\frac{1}{2\sigma_j^2}\left(\frac{x'^2}{\tau+1} + \frac{y'^2}{\tau+q_j'}\right)\right]}{\left[(\tau+1)(\tau+q_j'^2)\right]^{1/2}} d\tau. \\ \psi_j(x', y') &= -2\pi G q_j'\sigma_j^2 A_j \int_0^\infty \frac{\exp\left[-m_j^2(\tau)/(2\sigma_j^2)\right]}{\left[(\tau+1)(\tau+q_j'^2)\right]^{1/2}} d\tau. \end{aligned} \tag{C.19}$$

Now, we do the substitution

$$u^2 = \frac{1}{1+\tau} \;\rightarrow\; 2u\,du = -\frac{1}{(1+\tau)^2} d\tau, \tag{C.20}$$

which when $\tau \to 0, u \to 1$, and $\tau \to \infty, u \to 0$. This substitution also allows to write

$$du = \frac{-1}{2u(1+\tau)^2} d\tau \;\; \rightarrow \;\; d\tau = \frac{-2}{u^3} du, \tag{C.21}$$

as well as,

$$\tau = \frac{1-u^2}{u^2}, \quad \text{and} \quad q_j'^2 + \tau = \frac{1-(1-q_j'^2)u^2}{u^2}. \tag{C.22}$$

Therefore,

$$\begin{aligned} (1+\tau)(q_j'^2+\tau) &= \frac{1}{u^2}\frac{1-(1-q_j'^2)u^2}{u^2} \\ \frac{1}{\left[(1+\tau)(q_j'^2+\tau)\right]^{1/2}} &= \frac{u^2}{\left[1-(1-q_j'^2)u^2\right]^{1/2}}. \end{aligned} \tag{C.23}$$

On the other hand, the coordinate $m_j^2(\tau)$ is

$$\begin{aligned} m_j^2(\tau) &= \frac{x'^2}{1+\tau} + \frac{y'^2}{q_j'^2+\tau} \\ m_j^2(\tau) &= u^2\left[x'^2 + \frac{y'^2}{1-(1-q_j'^2)u^2}\right], \end{aligned} \tag{C.24}$$

where we use equations (C.20) and (C.22). Using the above expression and (C.23), we can rewrite (C.19) as

$$\begin{aligned} \psi_j(x',y') &= -2\pi G q_j' \sigma_j^2 A_j \int_1^0 \frac{\exp\left[-\frac{-u^2}{2\sigma_j^2}\left(x'^2 + \frac{y'^2}{1-(1-q_j'^2)u^2}\right)\right]}{\left[1-(1-q_j'^2)u^2\right]^{1/2} u^{-2}} \left(-\frac{2}{u^3}\right) du. \\ \psi_j(x',y') &= -4\pi G q_j' \sigma_j^2 A_j \int_0^1 \frac{f_j'(u)}{u} du, \end{aligned} \tag{C.25}$$

where we define the function $f_j'(u)$ as being

$$f_j'(u) \equiv \exp\left[-\frac{-u^2}{2\sigma_j^2}\left(x'^2 + \frac{y'^2}{1-\eta^2u^2}\right)\right]\frac{1}{(1-\eta^2u^2)^{1/2}}, \tag{C.26}$$

with $\eta^2 = 1 - q_j'^2$.

Now, let us take a look ate the constant factor in (C.25).

$$\begin{aligned} 4\pi G q'_j \sigma_j^2 A_j &= 4\cancel{\pi} G \cancel{q'_j} \cancel{\sigma_j^2} \frac{\cancel{2}}{c^2} \frac{D_L D_{LS}}{D_S} \frac{M_j}{\cancel{2\pi} \cancel{\sigma_j^2} \cancel{q'_j}} \\ &= \frac{4\,G\,M_j}{c^2} \frac{D_L D_{LS}}{D_S} \times \underbrace{\left(\frac{\Sigma_c}{\Sigma_c}\right)}_{\text{multiply by one}} \\ &= \cancel{\left(\frac{4\,G}{c^2}\right)\left(\frac{c^2}{4\pi G}\right)} \overset{\frac{1}{\pi}}{\cancel{D_L D_{LS}}} \frac{M_j}{\Sigma_c} \frac{\cancel{D_S}}{\cancel{D_L D_{LS}}} \\ &= \frac{M_j}{\pi \Sigma_c}. \end{aligned} \tag{C.27}$$

And finally, the lens potential, component j is given by

$$\psi_j(x', y') = -\frac{M_j}{\pi \Sigma_c} \int_0^1 \frac{f'_j(u)}{u} du. \tag{C.28}$$

The reduced deflection angle (C.4), now is given by

$$\alpha_{x'}(x', y') = \sum_j \frac{\partial \psi_j(x', y')}{\partial x'}, \tag{C.29}$$

$$\alpha_{y'}(x', y') = \sum_j \frac{\partial \psi_j(x', y')}{\partial y'}, \tag{C.30}$$

where the summation is over all the Gaussians that describe the projected mass density profile, which generates the lens potential, and $\boldsymbol{\alpha}(x', y') = (\alpha_{x'}, \alpha_{y'})$.

Also, as the variable u is independent of (x', y'), we can change the order of the integral and derivative in (C.29), such that

$$\frac{\partial \psi_j}{\partial x'} = \frac{\partial}{\partial x'} \left[-\frac{M_j}{\pi \Sigma_c} \int_0^1 \frac{f'_j(u)}{u} du \right] = -\frac{M_j}{\pi \Sigma_c} \left[\int_0^1 \frac{\partial f'_j(u)}{\partial x'} \frac{du}{u} \right], \tag{C.31}$$

But,

$$\begin{aligned} \frac{\partial f'_j(u)}{\partial x'} &= \frac{1}{(1 - \eta^2 u^2)^{1/2}} \frac{-u^2}{2\sigma_j^2} 2x' \exp\left[-\frac{-u^2}{2\sigma_j^2} \left(x'^2 + \frac{y'^2}{1 - \eta^2 u^2} \right) \right] \\ \frac{\partial f'_j(u)}{\partial x'} &= -\frac{u^2\, x'}{\sigma_j^2} f'_j(u). \end{aligned} \tag{C.32}$$

Therefore,

$$\frac{\partial \psi_j}{\partial x'} = -\frac{M_j\, x'}{\pi \sigma_j^2 \Sigma_c} \int_0^1 u\, f'_j(u) du. \tag{C.33}$$

Proceeding in a similar way for the derivative with respect to y', we get

$$\frac{\partial \psi_j}{\partial y'} = -\frac{M_j\, y'}{\pi \sigma_j^2 \Sigma_c} \int_0^1 \frac{u\, f_j'(u)}{(1-\eta^2 u^2)} du. \tag{C.34}$$

Thus, the reduced deflection angle is written in its final form as

$$\alpha_{x'}(x', y') = \frac{1}{\pi \Sigma_c} \int_0^1 u du \sum_j \frac{M_j}{\sigma_j^2} \frac{x'}{(1-\eta^2 u^2)^{1/2}} \exp\left[-\frac{-u^2}{2\sigma_j^2}\left(x'^2 + \frac{y'^2}{1-\eta^2 u^2}\right)\right], \tag{C.35}$$

$$\alpha_{y'}(x', y') = \frac{1}{\pi \Sigma_c} \int_0^1 u du \sum_j \frac{M_j}{\sigma_j^2} \frac{y'}{(1-\eta^2 u^2)^{3/2}} \exp\left[-\frac{-u^2}{2\sigma_j^2}\left(x'^2 + \frac{y'^2}{1-\eta^2 u^2}\right)\right]. \tag{C.36}$$

These two equations are the same as derived by van de Ven et al. (2010). Also these two integral can be solved by well-known numerical Gaussian quadrature.

C.1 A note about units

Units of measure are very important, so let us give them some attention now.

Assuming that the angular diameter distance is given in megaparsec (Mpc), and labelling it as D, we can convert an angular measurement in arcsec ($L['']$) to parsec ($L[\mathrm{pc}]$), multiplying it by a constant factor $\mathrm{pc} = \frac{D\pi}{0.648}$[2], such that

$$L[\mathrm{pc}] = \mathrm{pc} \times L[''] = \frac{D\pi}{0.648} L['']. \tag{C.37}$$

Then, a measure in parsec can be transformed in a measure in arcsec just inverting the above equation

$$L[''] = \frac{0.648}{D\pi} L[\mathrm{pc}]. \tag{C.38}$$

On the other hand, a measure in arcsec ($L['']$) can be transformed in a measure in radians ($L[\mathrm{rad}]$), applying the conversion

$$L[\mathrm{rad}] = \frac{\pi}{180 \times 3600} L[''] \qquad \rightarrow \qquad L[''] = \frac{180 \times 3600}{\pi} L[\mathrm{rad}]. \tag{C.39}$$

This lead us with,

$$L[\mathrm{pc}] = \frac{D\pi}{0.648} \frac{180 \times 3600}{\pi} L[\mathrm{rad}] = 10^6 D\, L[\mathrm{rad}], \tag{C.40}$$

[2] The factor 0.648 comes from the conversion of degress to arcsec, and takes in account the order of magnitude M = 10^6.

where we use (C.37) and (C.39). Therefore, to obtain a measure in units of radians, we need to multiply the corresponding value, in parsecs, by $\frac{1}{10^6 D}$.

Now suppose that the critical surface density is given in units of solar masses per megaparsec square, i.e., $[\Sigma_c] = \mathrm{M}_\odot/(\mathrm{Mpc}^2)$. Applying conversion (C.40) to its length dimension, we have

$$\mathrm{Mpc}^2 = 10^{12}\mathrm{pc}^2 \quad \times \quad \underbrace{\frac{1}{(10^6\, D)^2}}_{conversion\, factor} \quad = \mathrm{rad}^2 \frac{1}{D^2}. \tag{C.41}$$

So,

$$\frac{M_\odot}{\mathrm{Mpc}^2} = \frac{M_\odot}{(\mathrm{rad}/D)^2} = \frac{M_\odot}{\mathrm{rad}^2} D^2. \tag{C.42}$$

And the critical surface density can be expressed in units of solar masses per radian square, just multiplying the correspond value in solar masses per megaparsec square by the square of luminosity distance in parsec

$$\Sigma_c[\frac{M_\odot}{\mathrm{rad}^2}] = \Sigma_c[\frac{M_\odot}{\mathrm{Mpc}^2}]\, D^2[\mathrm{pc}]. \tag{C.43}$$

Using this units for the critical surface density, we are able to use x', y' and σ_j in radians too, and the correspondent reduced deflection angle also is given in radian units.

APPENDIX D – Two dimensional posterior distribution

Figure 41 shows the two-dimensional posterior distribution for the most probable model. The shaded areas are, respectively, 10%, 40%, 65%, and 85% quantiles. The dashed lines are 1σ credible interval.

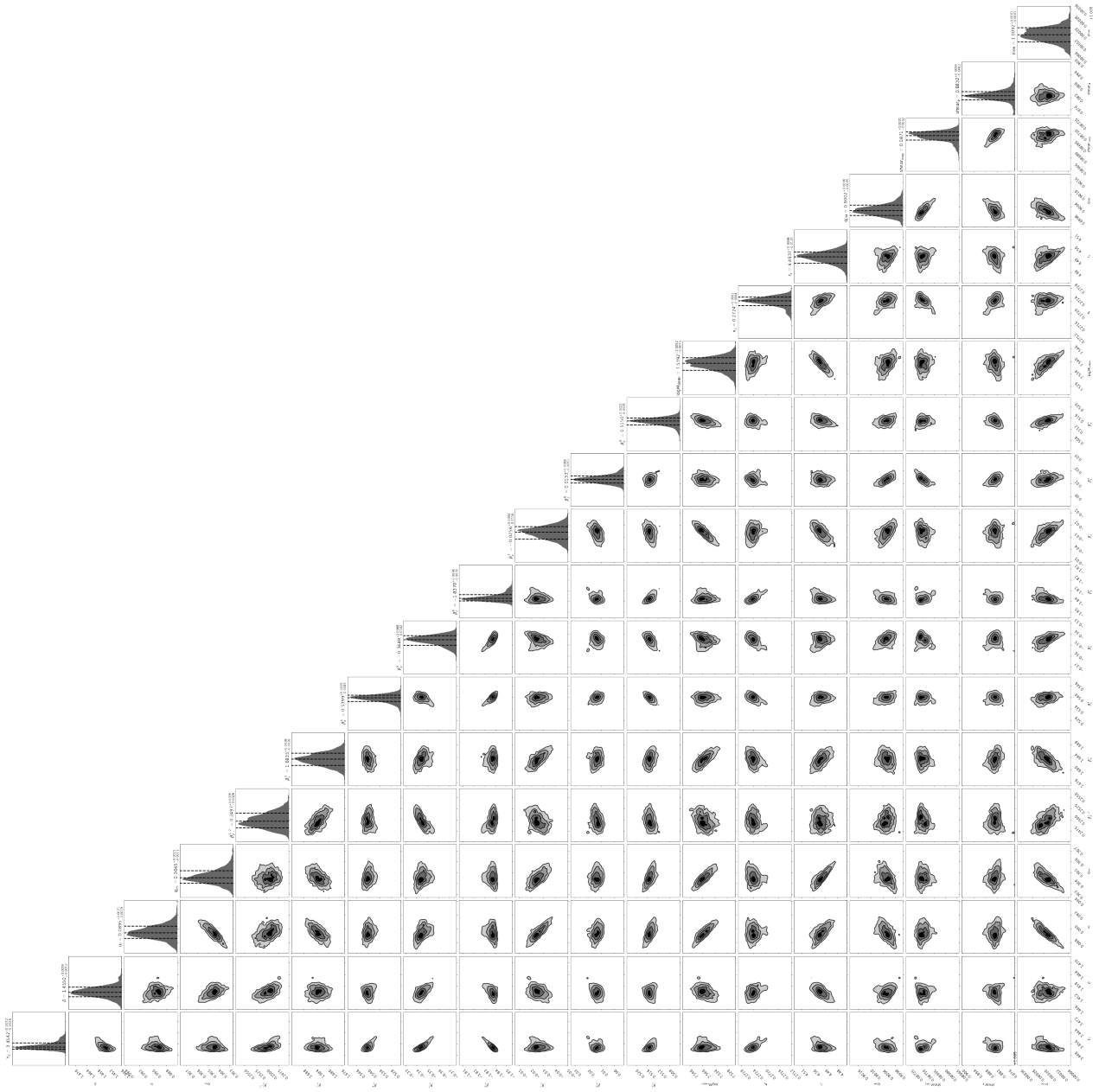


Figure 41 – Two dimensional posterior distributions for the parameters of most probable model of SDP.81.

APPENDIX E – Press Release: Testando a Relatividade Geral em Escala de Galáxias

Se você é um terráqueo e já foi criança algum dia, então com certeza já sentiu os efeitos da gravidade na própria pele. O interessante talvez seja notar que é essa mesma gravidade a responsável por você cair de uma árvore em direção ao chão e por manter o planeta Terra em órbita do Sol.

Atualmente, a melhor teoria capaz de explicar o fenômeno da gravidade é a Relatividade Geral, teoria fundada por Einstein no começo do século XX. Essa teoria passa a ver a gravidade não mais como uma força que atraía todos os corpos com massa, mas sim como uma deformação do próprio espaço-tempo (leia-se deformação do próprio Universo) causada por corpos que possuem massa e energia. Essa deformação do próprio Universo é o que sentimos por gravidade. E mais, qualquer partícula (fosse uma pessoa, um planeta ou a própria luz) que se movesse nesse espaço-tempo curvado deveria se mover seguindo essas curvas. O exemplo mais famoso deste efeito são as lentes gravitacionais, que ocorrem quando dois corpos estão alinhados na linha de visão de um observador, e o corpo mais próximo é muito massivo. Devido a grande massa do corpo próximo ao observador, o espaço-tempo será curvado de tal maneira que a luz emitida pelo corpo mais distante poderá ser vista pelo observador, embora ela apareça de maneira distorcida. Nesse contexto, o objetivo massivo causador da distorção é chamado de lente gravitacional, e o objeto mais distante é chamado de fonte.

A Figura 1 pode ajudar a entender esse fenômeno. Se entre você e esse simpático cachorrinho à esquerda houvesse um corpo muito massivo, digamos uma galáxia inteira, pela teoria do Newton você não seria capaz de observar o cachorrinho ao fundo, pois sua visão seria bloqueada pela galáxia. Porém, pela Teoria da Relatividade Geral, essa galáxia muito massiva curvaria o caminho que a luz deve percorrer, de tal maneira que você conseguiria ver o cachorrinho, embora de uma maneira bem estranha, como mostra o painel da direita.

Contudo, mesmo sendo a melhor explicação para o fenômeno de gravitação que conhecemos na atualidade, a Relatividade Geral possui alguns problemas com os quais não consegue lidar. Em particular, ela não é capaz de explicar algumas observações feitas sem precisar incluir novos elementos (como matéria e energia escura) nunca observados diretamente. Por essa, e outras razões, é necessário que continuemos testando a Relatividade Geral em novas escalas e ambientes.

E é justamente isso o que foi feito no projeto do mestrando em Astrofísica realizado

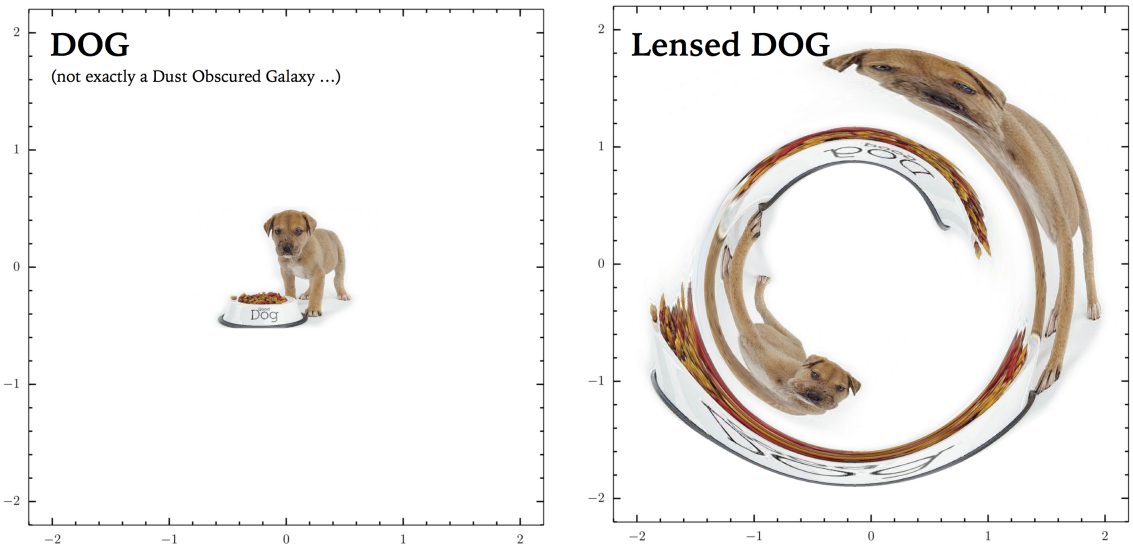


Figura 1: Efeito de lentes gravitacionais em um cachorrinho. Fonte: Dr. Mattia Negrello website

no Instituto de Física da UFRGS, Carlos Roberto. Em seu trabalho, ele estende um teste da teoria desenvolvida por Einstein para novas distâncias, procurando entender se a Relatividade Geral continua sendo válida, mesmo quando nos distanciamos (e muito) da nossa vizinhança galáctica.

Nesse teste, ele usa uma abordagem já estabelecida por outros autores, que consiste em usar o fenômeno de lentes gravitacionais em combinação com a descrição do movimento das estrelas da galáxia que está agindo como lente gravitacional. Isto é, observando o movimento das estrelas e a distorção causada por sua massa, somos capazes de verificar a validade da teoria einsteiniana.

No projeto, o mestrando Carlos utilizou o sistema conhecido como SDP.81 (Figura 2), cujo galáxia lente está a aproximadamente 5 bilhões de anos luz de distância, e que possuía todos os dados necessários para esse tipo de análise, incluindo dados adicionais de alta resolução. A análise preliminar dos resultados obtidos mostraram que mais uma vez a Relatividade Geral continua em sua posição de destaque como a teoria que melhor descreve o fenômeno da gravitação, pois mais uma vez ele se mostrou bem sucedida no teste que foi realizado.

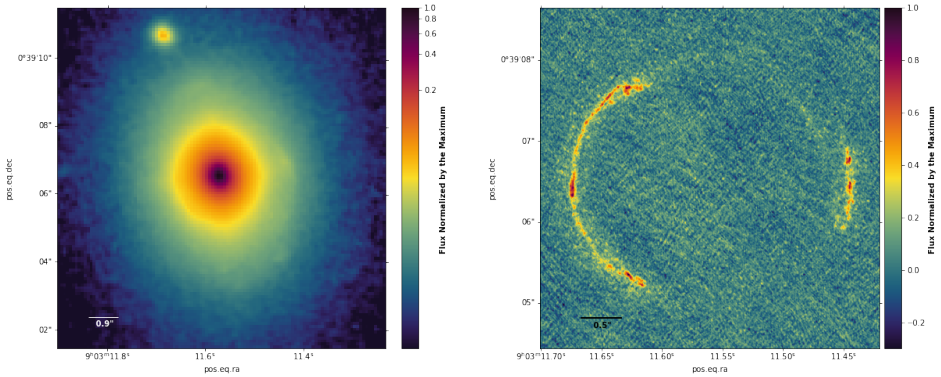


Figura 2: Imagem do sistema SDP.81. À esquerda está uma imagem capturada pelo telescópio espacial Hubble da galáxia que age como lente gravitacional, deformando o espaço-tempo. Já à direita, vemos a imagem distorcida da galáxia fonte, capturada pelo interferômetro ALMA. Devido aos diferentes comprimento de onda em que operam os instrumentos, a galáxia que age como lente não é vista na imagem da direita.