Alma Mater Studiorum - Università di Bologna

School of Science
Department of Physics and Astronomy
Master Degree Programme in Astrophysics and Cosmology

Your Title

Graduation Thesis

March 15, 2024

Presented by:
Your Name

Supervisor:
Chiar.mo Prof. Supervisor

Co-supervisor:
Dott. Co-supervisor

Academic year 2022-2023
Graduation date V

Nature uses only the longest threads to weave her patterns, so that each small piece of her fabric reveals the organization of the entire tapestry

---

Richard P. Feynman

Abstract

This is the abstract.

The primary objective of this thesis is to address these challenges by developing advanced Python algorithms based on differentiable programming paradigm, leveraging the capabilities of PyTorch and TensorFlow frameworks to enable precise modeling and analysis of gravitational lenses. By employing parametric models, these algorithms exploit automatic differentiation to backpropagate errors and compute gradients of a loss function, facilitating the optimization of high-dimensional parameter spaces. Through the training of these parametric models, relevant features can be extracted and key parameters of the lensing system can be estimated. The resulting models can then be applied to real observational data, improving the characterization and classification of strong lenses with enhanced accuracy and efficiency.

The structure of this thesis is the following: an introduction is presented in Chapters 1 and 2. Chapter 3 describes some examples. Finally, Chapter 4 specifically focus on the applications.

Sommario

Questo è il sommario.

L'obiettivo principale di questo lavoro di tesi è quello di affrontare tali questioni sviluppando algoritmi Python avanzati per la precisa modellizzazione e analisi di lenti gravitazionali forti, fondati su tecniche di programmazione differenziabile, implementate utilizzando i framework PyTorch e TensorFlow. Utilizzando modelli parametrici, questi algoritmi sfruttano la differenziazione automatica per retropropagare gli errori e calcolare i gradienti di una funzione di costo, facilitando l'ottimizzazione nello spazio dei parametri. Attraverso l'implementazione in tali modelli parametrici, è possibile estrapolare le caratteristiche rilevanti e stimare i parametri chiave del sistema in esame. I modelli risultanti possono essere applicati a dati osservativi reali, migliorando la caratterizzazione e la classificazione delle lenti con maggiore precisione ed efficienza.

La struttura di questa tesi è la seguente: nei Capitoli 1 e 2 viene presentata un'introduzione. Il Capitolo 3 descrive alcuni esempi. Il Capitolo 4 si concentra specificamente sull'applicazione.

Contents

List of Figures

List of Tables

1

Chapter 1

1.1 History

Throughout history, questions surrounding the origin, age, and size of the universe have fascinated humans. Plato believed in a Universe that remains constant, envisioning it as an entity that was created perfect, unaging, and free from decay (Plato, 2010). This view of a static universe prevailed for more than two millennia, and it was so entrenched in cosmological thinking that even Albert Einstein initially subscribed to it. To reconcile his field equations of general relativity with the notion of a static universe, Einstein introduced the cosmological constant, a term that provided a mathematical means to allow for static solutions to his equations (Einstein, 1917).

However, the early 20th century brought discoveries that challenged the long-standing paradigm of a static universe. Vesto Slipher made pivotal observations, noting that most galaxies recede from the Milky Way at high velocities. Building on this foundation, Edwin Hubble, in the 1920s, conducted an analysis of the escape velocities of distant galaxies, leading to a groundbreaking discovery. Hubble observed that the farther away galaxies are, the faster they appear to be moving away. He plotted the radial velocities of these galaxies against their distances and found that the data could be best described by a straight line, indicating a linear relationship (Fig. 1.1).

1.2 Friedmann Equations

1.2.1 Expansion rate

According to the *Cosmological Principle*, on sufficiently large scales, the universe is homogeneous and isotropic. Considering a sphere with homogeneous density and a test particle at location $\vec{x}$, and introducing a spherical coordinate system that is allowed to expand with time, due to the cosmological principle, the expansion and the time-dependent position $\vec{r}(t)$ can be expressed by

$$\vec{r}(t) = a(t)\vec{x}\,, \tag{1.1}$$

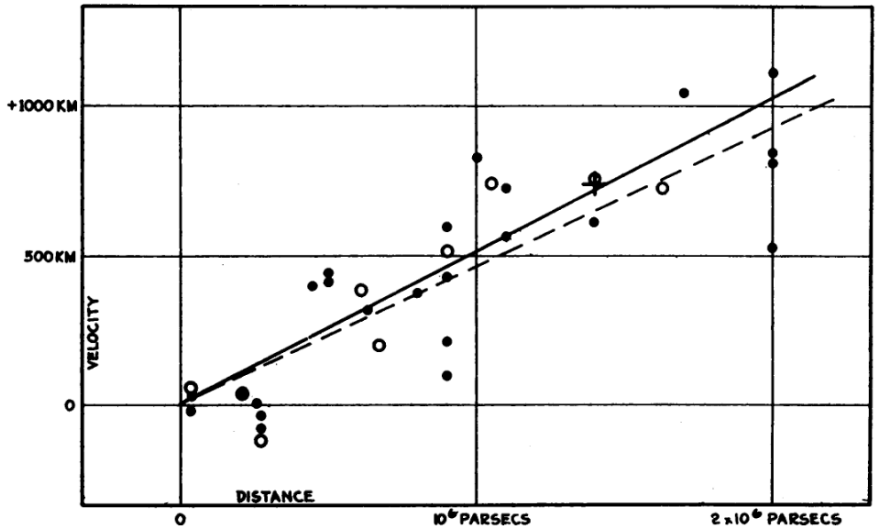

Figure 1.1. Radial velocities, corrected for solar motion, plotted against distances estimated from involved stars and mean luminosities of nebulae in a cluster.
Credits: Hubble (1929).

where $a(t)$ is the *cosmic scale factor*, which does not depend only on time. For $t_0 =$ today, the scale factor is conventionally set to $a(t_0) = 1$. Scalar quantities r, x can be used instead of $\vec{r}, \vec{x}$ due to isotropy.

1.2.2 Dynamics of the expansion

To derive the Friedmann Equations and to study the evolution of the scale factor $a(t)$, to better understand the development of the universe, one has to start from Einstein field's equation, that describes the geometry of space-time:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + g_{\mu\nu}\Lambda = \frac{8\pi G}{c^4}T_{\mu\nu}\,, \tag{1.2}$$

where $R_{\mu\nu}$ and R are the Ricci tensor and Ricci scalar, respectively, Λ is the *cosmological constant* and $T_{\mu\nu}$ is the *energy-momentum tensor*, which includes all contributions of energy and acts as the source of gravity.

The energy-momentum tensor of the universe is that of a homogeneous perfect fluid, characterized by its density $\rho(t)$ and pressure $p(t)$. Using the Robertson-Walker metric to describe homogeneity and isotropy, the Einstein's equations simplify to the

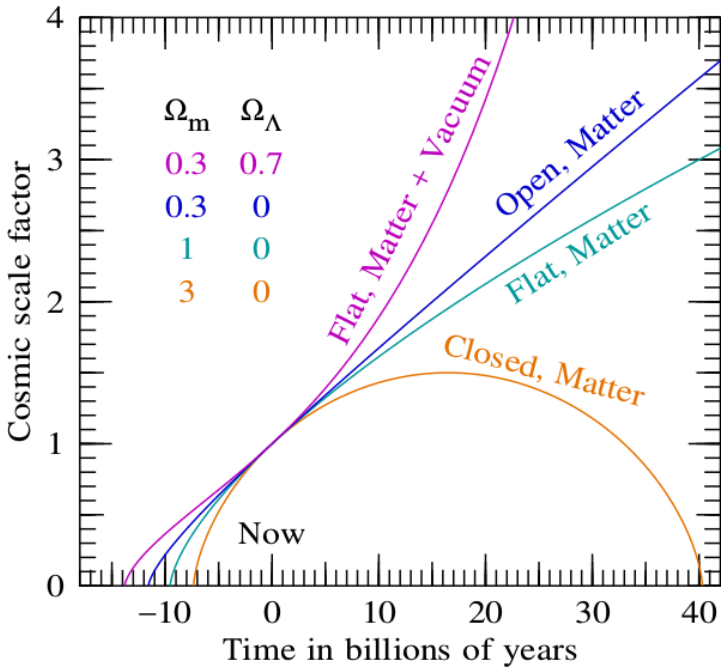

Figure 1.2. Numerical solutions for different cosmological models.
Credits: Hamilton (2019).

Friedmann equations (Friedman, 1922; Friedmann, 1924):

$$H(t)^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda c^2}{3} - \frac{Kc^2}{a^2}\,, \tag{1.3a}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda c^2}{3}\,, \tag{1.3b}$$

where K is a constant parameter that defines the curvature of spatial surfaces:

- $K = -1$ means open, hyperbolic space (i.e. infinite) with negative curvature;
- $K = 0$ means flat, Euclidean space;
- $K = 1$ means closed, spherical space with positive curvature.

I

First part

2

Chapter 2

In the realm of astrophysics and gravitational theory, the concept that gravity affects not only matter but also light has a long history, dating back to Newton's *Opticks* (Newton, 1704). The initial calculation of the Newtonian deflection of light rays passing near massive bodies was undertaken by Cavendish in 1784, and von Soldner subsequently published these findings in 1804 (Will, 1988, 2014). It was Albert Einstein himself who, in 1915, as he completed his General Theory of Relativity (Einstein, 1915), recognized that the Newtonian prediction of light-ray deflection near the Sun was only half the value foreseen by his revolutionary theory. The confirmation of this twice as large deflection during the solar eclipse of May 29, 1919, observed by Eddington (Dyson & Eddington, 1920; Will, 2015), marked a momentous confirmation of this nascent theory and captivated global attention. Not only did this observation represent the first in a series of triumphs for general relativity, but it also inaugurated the practical application of *gravitational lensing*, a method that would later become a cornerstone of observational astrophysics.

More than a century has passed since that pivotal moment, and gravitational lensing has evolved into a well-established and respected tool within the fields of astronomy and astrophysics. Presently, the study of lensing theory and its applications can be broadly categorized into three distinct components (Kochanek, 2004): *strong* lensing, characterized by non-linear deflection at the scale of galaxies and galaxy clusters, produces distinct phenomena such as multiple images, arcs, and rings, *weak* lensing, observed on both cluster and cosmological scales, is a subtle linear effect that gently aligns background galaxies with intervening matter. Statistical analysis of the observed distribution of light allows for the extraction of information regarding the distribution of intervening matter. Finally, *microlensing* involves the dynamic fluctuation of light when compact objects pass in front of background sources at scales too minute to be resolved.

2.1 Light deflection

According to Einstein's theory of relativity, objects with gravitational pull have the ability to alter the fabric of space-time, resulting in the bending of light rays (Narayan &

Bartelmann, 1997) . Gravitational lensing occurs when a substantial mass distribution can effectively curve and amplify the light emitted from a source positioned behind it.

The calculation of light deflection involves the examination of geodesic curves originating from the field equations of general relativity. Light deflection can also be understood through Fermat's principle[1], similar to how it is described in geometrical optics. The main approach is to consider light deflection within the framework of general relativity as a refraction problem, for which a refractive index n can be introduced.

To investigate the bending of light and to determine the refractive index, an initial approximation is made by assuming that the lens is "weak" and significantly smaller than the source-lens-observer optical system, an assumption true for nearly all astrophysical scenarios. This "weak field" approximation refers to a lens with a relatively small Newtonian gravitational potential, which means $\phi \ll c^2$, where c is the speed of light. Additionally, it is plausible to assume that light deflection occurs within a region small enough that the expansion of the universe can be disregarded. Leveraging the principle of equivalence, one can select a locally inertial frame in which space-time is flat and described by Minkowski's metric. In this context, the line element of the local metric tensor can be written as a small perturbation of the metric, such as

$$\mathrm{d}s^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu = \left(1+\frac{2\phi}{c^2}\right)c^2\,\mathrm{d}t^2 - \left(1-\frac{2\phi}{c^2}\right)\mathrm{d}\vec{x}^2\ . \tag{2.1}$$

Applying Fermat's principle, the total deflection angle of a photon is the integral over the gradient of the potential perpendicular to the light path along the proper light path. Thanks to the *Born approximation*[2], it can be shown (Schneider et al., 1992) that the deflection angle can be obtained integrating over the unperturbed light path:

$$\hat{\vec{\alpha}}(\vec{\xi}) = \frac{2}{c^2}\int_{-\infty}^{+\infty}\vec{\nabla}_\perp\phi(\vec{\xi},z)\,\mathrm{d}z\ , \tag{2.2}$$

where ξ is the impact parameter of the photon traveling along the $\vec{e}_z$ direction that passes through the lens at $z = 0$ (Fig. 2.1).

[1] Light travels between two points along the path that requires the least time.

[2] Simplification valid when the gravitational potential is small: deflection of light is treated like a linear process, neglecting higher-order corrections.

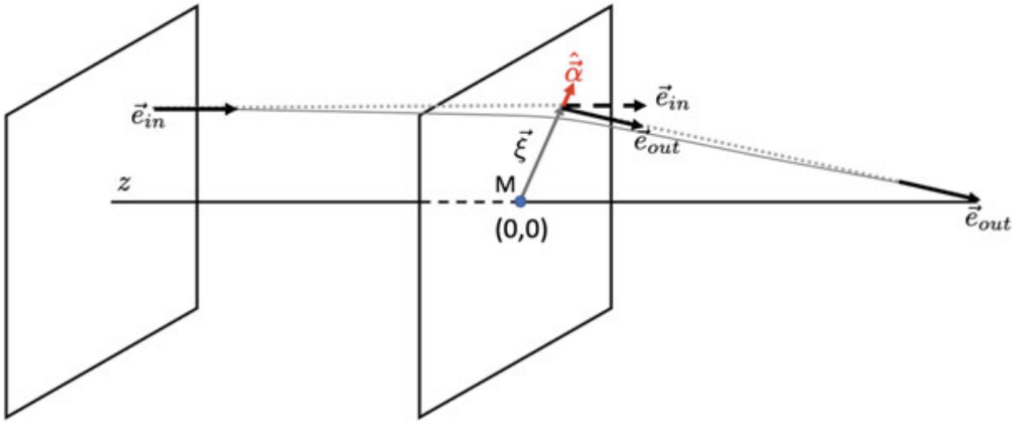

Figure 2.1. Schematics for the Born approximation.
Credits: Meneghetti (2021).

3

Chapter 3

This chapter delves into the intricate process of lens modeling, a pivotal technique for interpreting gravitational lensing observations. Gravitational lens models can be separated into two main categories: point-mass lenses (i.e. *microlenses*) and extended lenses, each possessing distinctive characteristics. Microlensing refers to the lensing effect caused by objects with relatively small masses, such as stars or planets, acting as lenses. Unlike their more massive counterparts, microlenses do not produce multiple discernible images of the background source. Instead, they induce magnification variations over time as the lens moves relative to the observer and the source. Transitioning to a grander scale, extended lenses involve massive structures like galaxies and galaxy clusters, capable of producing multiple, resolvable images of background sources.

3.1 Microlenses

This section is devoted to exploring the phenomenon of microlensing, which refers to the lensing effects caused by objects of relatively small mass in the universe, such as planets, stars, star clusters, and other compact objects located within the Milky Way or distant galaxies. Typically, these microlenses are considered, in a first-order approximation, to be point-masses or collections of point-masses.

3.1.1 Deflection angle and lensing potential

As already derived, by setting the lens position as the center of the reference frame and using the relation $\xi = D_L\theta$, the deflection angle for a point mass lens can be written as

$$\vec{\alpha}(\vec{\theta}) = \frac{D_{LS}}{D_S}\hat{\vec{\alpha}}(\vec{\theta}) = \frac{D_{LS}}{D_S}\frac{4GM}{c^2 D_L}\frac{\vec{\theta}}{|\vec{\theta}|^2}\,. \tag{3.1}$$

Given that $\vec{\alpha}(\vec{\theta}) = \vec{\nabla}\hat{\Psi}(\vec{\theta})$, the lensing potential of the point mass lens is

$$\hat{\Psi}(\vec{\theta}) = \frac{4GM}{c^2}\frac{D_{LS}}{D_L D_S}\ln|\vec{\theta}|\,. \tag{3.2}$$

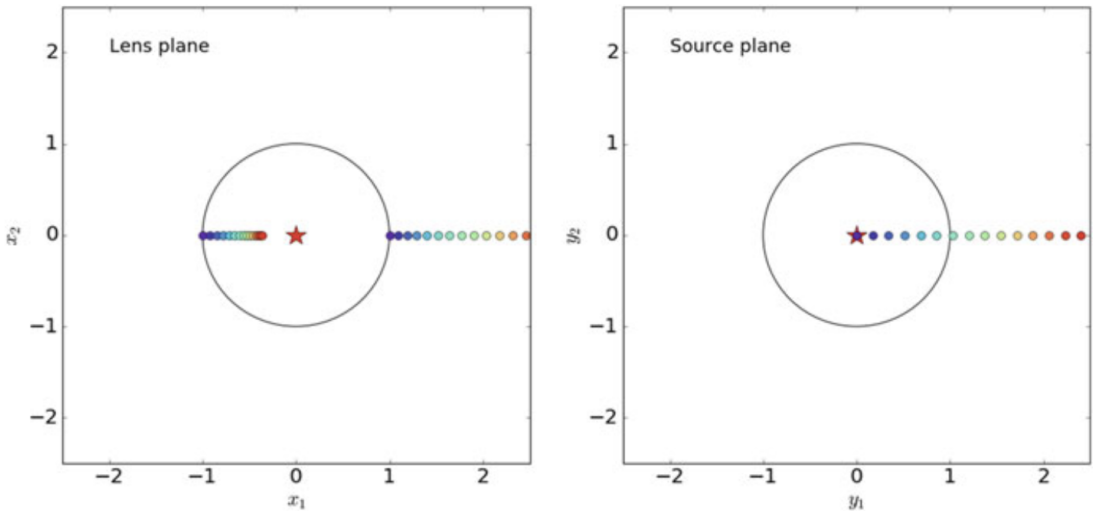

Figure 3.1. Solutions of the lens equation for a point-mass, with the lens represented by the star at the center. The Einstein ring is highlighted in black. In the right diagram, the locations of various sources are marked with colored circles. The images produced, as calculated using Eq. (3.5), are displayed in the left diagram.
Credits: Meneghetti (2021).

3.1.2 Lens equation and multiple images

Given the deflection angle of Eq. (3.1), the lens equation becomes

$$\beta = \theta - \frac{4GM}{c^2\theta}\frac{D_{LS}}{D_L D_S}\,, \tag{3.3}$$

where the vector signs can be omitted due to the fact that the vector $\hat{\vec{\alpha}}$ always points away from the lens.

As already anticipated in Section 2.1, the lens equation can be written in a more concise way by introducing a scale radius θ_E, i.e. the *Einstein radius* defined in Eq. (3.1), and setting $y = \beta/\theta_E$ and $x = \theta/\theta_E$, results:

$$\beta = \theta - \frac{\theta_E^2}{\theta} \quad \Rightarrow \quad y = x - \frac{1}{x}\,. \tag{3.4}$$

This equation is quadratic in θ (or x) and has two solutions:

$$x_\pm = \frac{y \pm \sqrt{y^2+4}}{2}\,, \tag{3.5}$$

which means that there always exist two images for a given source position.

In the right section of Fig. 3.1, some sources are arranged at varying angular distances

from the lens, which is marked by a red star. Each source is represented by a unique color to facilitate the identification of its corresponding images in the left section. Each source generates two images: one positioned at $x_+ > 0$ and the other in the range of $-1 < x_- < 0$. These images appear on either side of the lens, with the x_- image always lying within a circle of radius $x = 1$. This circle is equivalent to the image produced by a source directly behind the point lens at $y = 0$, resulting in a ring-shaped image with radius θ_E, the Einstein ring. The size of the Einstein radius is typically

$$\theta_E \approx 1'' \left(\frac{M}{10^{12} M_\odot}\right)^{1/2} \left(\frac{D}{\text{Gpc}}\right)^{-1/2} , \tag{3.6}$$

where

$$D \equiv \frac{D_L D_S}{D_{LS}} \tag{3.7}$$

is the *effective lensing distance*.

As the angular separation $y \to 0$, it is observed that $x_- \to 0$, whereas $x_+ \to y$. This indicates that when the angular distance between the lens and the source increases significantly, the source is unlensed. In theory, an image still exists at $x_- = 0$, but this central image has zero magnification.

3.1.3 Critical lines, caustics and magnification

The Jacobian determinant for a point-mass lens can be written as

$$\det A(x) = \frac{y}{x} \frac{\mathrm{d}y}{\mathrm{d}x} , \tag{3.8}$$

which means that the eigenvalues of the Jacobian matrix are

$$\lambda_t(x) = \frac{y}{x} = \left(1 - \frac{1}{x^2}\right) , \tag{3.9a}$$

$$\lambda_r(x) = \frac{\mathrm{d}y}{\mathrm{d}x} = \left(1 + \frac{1}{x^2}\right) . \tag{3.9b}$$

II

Second part

4

Chapter 4

Building on the foundational understanding of gravitational lensing and its computational challenges, a critical aspect of modern astrophysical research involves the optimization of parametric functions to model the complex phenomena of the Universe accurately. This optimization is essential in gravitational lensing studies, where precise models of the mass distribution within lensing objects are paramount for interpreting the observed lensing effects. The optimization process often requires navigating a high-dimensional parameter space to find the best-fit parameters that reconcile theoretical models with observational data, a task that can be computationally intensive and algorithmically complex.

In this context, the advent of frameworks such as PyTorch[1] (Paszke et al., 2019) and TensorFlow[2] (Abadi et al., 2016) represents a significant leap forward for astrophysical research. These open-source libraries, primarily developed for deep learning applications, offer powerful tools for automatic differentiation, a technique that facilitates the calculation of gradients automatically, a cornerstone for any optimization algorithm. The ability of PyTorch and TensorFlow to efficiently compute derivatives of highly complex, nested functions makes them exceptionally well suited for optimizing parametric models in gravitational lensing studies.

Moreover, PyTorch and TensorFlow are designed to exploit the capabilities of modern computing hardware, including GPUs and TPUs, enabling the parallel processing of large datasets and the acceleration of computational tasks. This feature is particularly beneficial for gravitational lensing research, where handling large amounts of observational data and running complex simulations is commonplace. Using the computational power offered by these frameworks, one can significantly reduce the time required for model optimization and data analysis, thus enhancing the efficiency of the investigations.

4.1 Differentiable programming

Differentiable programming is an advanced computational paradigm that unites traditional programming concepts with the principles of differentiation, a fundamental

[1] https://pytorch.org
[2] https://www.tensorflow.org

concept in calculus. This approach extends the idea of computing derivatives to entire programs, enabling the automatic calculation of gradients of program outputs with respect to inputs. Differentiable programming is particularly powerful in the context of optimization, machine learning, and artificial intelligence, where it facilitates efficient parameter tuning to minimize or maximize some objective function. Methods for computing derivatives in computer programs can be classified into four categories (see Fig. 4.1):

1. manually working out derivatives and coding them;

2. **numerical differentiation**, involves approximating the derivative of a function using values of the original function evaluated at some sample points (Burden et al., 2016). In its simplest form, it is based on the limit definition of a derivative. It is quite simple to implement and apply to a wide range of problems, especially when dealing with data-driven models or functions that lack a clear analytical representation. Its effectiveness diminishes in high-dimensional settings, where the complexity and computational demands increase exponentially, highlighting the method's limitations in handling complex, multi-variable functions efficiently;

3. **symbolic differentiation**, which refers to the automatic process of finding derivatives using the rules of differentiation to obtain an exact symbolic expression for the derivative of a given function (Grabmeier et al., 2003). This method works similarly to how humans perform differentiation "by hand", manipulating symbols according to mathematical laws;

4. **automatic differentiation** (AD), also called algorithmic differentiation,is a method that computes the derivative of a function efficiently and accurately by systematically applying the chain rule of calculus to the sequence of elementary operations (additions, multiplications, trigonometric functions, etc.) that constitute a computer program. All numerical computations are ultimately compositions of a finite set of elementary operations for which derivatives are known (Verma, 2000; Griewank & Walther, 2008), and combining the derivatives of the constituent operations through the chain rule of calculus[3] gives the derivative of the overall composition. AD is not an approximation like numerical differentiation, but rather computes derivatives to machine precision.

[3]The derivative of the composition of two (or more) differentiable functions f and g can be expressed in terms of the derivatives of the single functions f and g. If $h(x) = f(g(x))$, then $h'(x) = f'(g(x))g'(x)$.

Technique	Advantage(s)	Drawback(s)
Hand-coded analytical derivative	Exact and often fastest method.	Time consuming to code, error prone, and not applicable to problems with implicit solutions. Not automated.
Finite differentiation	Easy to code.	Subject to floating point precision errors and slow, especially in high dimensions, as the method requires at least D evaluations, where D is the number of partial derivatives required.
Symbolic differentiation	Exact, generates symbolic expressions.	Memory intensive and slow. Cannot handle statements such as unbounded loops.
Automatic differentation	Exact, speed is comparable to hand-coding derivatives, highly applicable.	Needs to be carefully implemented, alhough this is already done in several packages.

Figure 4.1. Summary of techniques to calculate derivatives.
Credits: Margossian (2019).

In particular, differentiable programming refers to using automatic differentiation in some way that allows a program to optimize its parameters to improve on some task. It only has three requirements:

1. a parameterized function/model to be optimized;

2. (automatic) differentiability of the object to be optimized;

3. a function suitable to measure the performance of the model.

The process of optimizing a model in the field of differentiable programming is often called *training*. In the subsequent portion of this section, a comprehensive overview of automatic differentiation and its role in the training process will be provided. This will be followed by an in-depth exposition of the process itself.

4.1.1 Automatic differentiation

Computational graph

As already stated, automatic differentiation operates on the fundamental principle that complex functions can be decomposed into a series of elementary arithmetic operations. Given a target composite function $f(x) = h \circ g(x) = h(g(x))$, with $x \in \mathbb{R}^n$, $g : \mathbb{R}^n \to \mathbb{R}^k$, and $h : \mathbb{R}^k \to \mathbb{R}^m$, applying the chain rule and elementary matrix multiplication, the

corresponding Jacobian matrix[4] J is thus:

$$J = J_{h \circ g} = J_h(g(x)) \cdot J_g(x)\,, \tag{4.1}$$

with $(i, j)^{th}$ element:

$$J_{ij} = \frac{\partial f_i}{\partial x_j} = \frac{\partial h_i}{\partial g_1}\frac{\partial g_1}{\partial x_j} + \frac{\partial h_i}{\partial g_2}\frac{\partial g_2}{\partial x_j} + \ldots + \frac{\partial h_i}{\partial g_k}\frac{\partial g_k}{\partial x_j}\,. \tag{4.2}$$

More generally, if f is the composite expression of L functions

$$f = f^L \circ f^{L-1} \circ \ldots \circ f^1\,, \tag{4.3}$$

the corresponding Jacobian matrix will be

$$J = J_L \cdot J_{L-1} \cdot \ldots \cdot J_1\,. \tag{4.4}$$

Hence, given a complex function f, it is possible to break down the action of the Jacobian matrix on a vector into simple components. So, following Griewank & Walther (2008) notation, a function $f : \mathbb{R}^n \to \mathbb{R}^m$ can be constructed using intermediate variables v_i such that

- variables $v_{i-n} = x_i$, $i = 1, \ldots, n$ are the input variables,
- variables v_i, $i = 1, \ldots, l$ are the intermediate variables,
- variables $y_{m-i} = v_{l-i}$, $i = m-1, \ldots, 0$ are the output variables.

The representation of all the elementary operations that take place to construct a certain function f is called the *evaluation trace*, which can also be pictured as a *computational graph* (Bauer, 1974), useful for visualizing the dependency relations between intermediate variables. Figure 4.2 shows the computation graph for an example function $f : \mathbb{R}^2 \to \mathbb{R}$:

$$f(x_1, x_2) = \ln(x_1) + x_1 x_2 - \sin(x_2)\,. \tag{4.5}$$

Given the computational graph of all elementary operations, AD can be implemented in two main modes: *forward* accumulation mode and *reverse* accumulation mode (backward mode).

[4]The Jacobian matrix of a vector-valued function of several variables is the matrix of all its first-order partial derivatives.

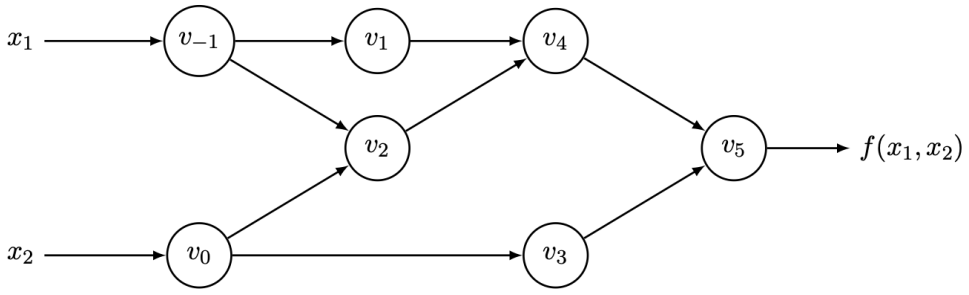

Figure 4.2. Computational graph of the example $f(x_1, x_2) = \ln(x_1) + x_1 x_2 - \sin(x_2)$. Credits: Baydin et al. (2018).

4.2 Light sources modeling

The prevalent approach to modeling light sources such as galaxies involves using a parametric profile (R), where R represents a measure of distance from the center of the object. These models often presuppose an excessively idealized level of symmetry, yet offer the advantage of being straightforward to define and apply. When building galaxy components, it is generally assumed that the profiles exhibit elliptical symmetry (Peng et al., 2002, 2010), and therefore R represents the elliptical radius

$$R(\vec{x}') = \sqrt{{x_1'}^2 + x_2'/q^2} \tag{4.6}$$

of a source with an axis ratio q. The coordinate system (x_1', x_2') of the source can be rotated by the *position angle* φ with respect to the coordinate system (x_1, x_2) of the observation. The resulting isophotes[5] of the surface brightness distribution of such an elliptical source are ellipses with semi-major axis R/q, semi-minor axis R and orientation φ with respect to the x_1 axis of observation.

4.2.1 Sérsic profile

The most common model to describe elliptical surface brightness distributions is the *Sérsic law* (Sérsic, 1963, 1968), which is given by the exponential

$$I(R) = I_e \exp\left\{-b_n \left[\left(\frac{R}{R_e}\right)^{\frac{1}{n}} - 1\right]\right\}, \tag{4.7}$$

[5]Curves of constant surface brightness.

where I_e is the surface brightness at the effective radius[6] R_e and $n > 0$ is called the *Sérsic index*, which characterizes the slope of the profile. The function b_n depends only on the Sérsic index and is defined by

$$\gamma(2n, b_n) = \frac{1}{2}\Gamma(2n)\,, \tag{4.8}$$

where Γ, γ are, respectively, the Gamma function and the incomplete Gamma function. It can be shown (Ciotti, 1991; Ciotti & Bertin, 1999) that, for a Sérsic index in the range $0.5 \leq n \lesssim 8$, b_n can be approximated by

$$b_n \approx 2n - \frac{1}{3} + \frac{4}{405n} \approx 1.9992n - 0.3271\,. \tag{4.9}$$

With this, the profile is now fully determined by the seven parameters for position x_1, x_2, effective radius R_e, Sérsic index n, intensity at the effective radius I_e, axis ratio q and position angle φ.

The Sérsic profile is a versatile model; by varying n it is possible to obtain many of the classical galaxy profiles as special cases, such as Gaussian profiles ($n = 0.5$), exponential profiles ($n = 1$) and de Vaucouleurs (de Vaucouleurs, 1948) profiles ($n = 4$).

[6] Also called the half-light radius, the radius within which half of the galaxy's luminosity is contained.

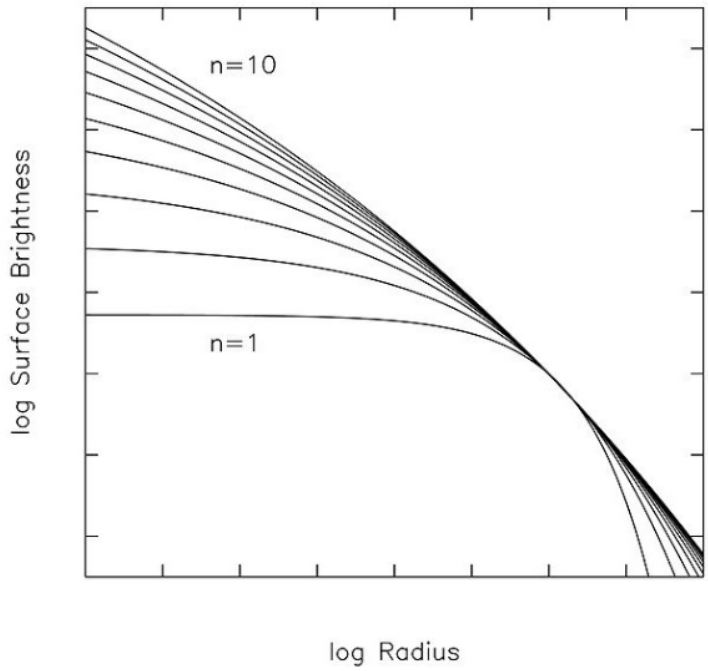

Figure 4.3. Sérsic profiles for different values of n. On average, $n \approx 2 - 10$ for bulges and elliptical galaxies, $n \approx 1$ for disk galaxies and $n \leq 0.5$ for bars and stellar clumps. Credits: Burke (2013).

III

Third part

5

Summary and conclusions

In this thesis, we have systematically explored the application of automatic differentiation (AD) and differentiable programming methods, utilizing the computational capabilities of PyTorch and TensorFlow, to address various aspects of gravitational lensing. This exploration was aimed at enhancing the optimization of parametric functions and models, which are critical for accurately modeling gravitational lensing phenomena. This concluding chapter aims to concisely summarize this work, emphasizing the integration of AD in gravitational lensing and detailing the specific topics addressed.

Looking ahead, AD promises to revolutionize gravitational lensing studies further by facilitating the development of sophisticated models for complex phenomena and enabling efficient analysis of vast astronomical datasets. Its integration into time-delay cosmography and non-parametric mass reconstruction could yield deeper insights into the universe's expansion and structure. Moreover, AD's role in processing data from upcoming large-scale surveys will be pivotal in uncovering new discoveries and enhancing our cosmological understanding.

In summary, AD's contribution to gravitational lensing fields is profound, offering a pathway to novel discoveries and a deeper understanding of the cosmos. Its continued development and application stand to unlock even greater potentials in astrophysical research.

IV

Appendix

A

Appendix name

A.1 Section header

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse

platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Bibliography

Abadi, M., Barham, P., Chen, J., et al. 2016, TensorFlow: A system for large-scale machine learning, arXiv. http://arxiv.org/abs/1605.08695

Bauer, F. L. 1974, SIAM Journal on Numerical Analysis, 11, 87, doi: 10.1137/0711010

Baydin, A. G., Pearlmutter, B. A., Radul, A. A., & Siskind, J. M. 2018, Automatic differentiation in machine learning: a survey, arXiv. http://arxiv.org/abs/1502.05767

Burden, R. L., Faires, J. D., & Burden, A. M. 2016, Numerical analysis, tenth edition edn. (Boston, MA: Cengage Learning)

Burke, C. 2013, PhD Thesis

Ciotti, L. 1991, Astronomy and Astrophysics, 249, 99. https://ui.adsabs.harvard.edu/abs/1991A&A...249...99C

Ciotti, L., & Bertin, G. 1999, Analytical properties of the R^(1/m) luminosity law, arXiv. http://arxiv.org/abs/astro-ph/9911078

de Vaucouleurs, G. 1948, Annales d'Astrophysique, 11, 247. https://ui.adsabs.harvard.edu/abs/1948AnAp...11..247D

Dyson, F. W., & Eddington, A. S. 1920, Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, 220, 291, doi: 10.1098/rsta.1920.0009

Einstein, A. 1915, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, 844. https://ui.adsabs.harvard.edu/abs/1915SPAW.....844E

—. 1917, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, 142. https://ui.adsabs.harvard.edu/abs/1917SPAW.....142E

Friedman, A. 1922, Zeitschrift für Physik, 10, 377, doi: 10.1007/BF01332580

Friedmann, A. 1924, Zeitschrift für Physik, 21, 326, doi: 10.1007/BF01328280

Grabmeier, J., Kaltofen, E., & Weispfenning, V., eds. 2003, Computer Algebra Handbook (Berlin, Heidelberg: Springer Berlin Heidelberg), doi: 10.1007/978-3-642-55826-9

Griewank, A., & Walther, A. 2008, Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation, Second Edition, 2nd edn. (Society for Industrial and Applied Mathematics), doi: 10.1137/1.9780898717761

Hamilton, A. J. S. 2019. https://jila.colorado.edu/~ajsh/courses/astr3740_20/evol.html

Hubble, E. 1929, Proceedings of the National Academy of Sciences, 15, 168, doi: 10.1073/pnas.15.3.168

Kochanek, C. S. 2004, The Saas Fee Lectures on Strong Gravitational Lensing, arXiv. http://arxiv.org/abs/astro-ph/0407232

Margossian, C. C. 2019, WIREs Data Mining and Knowledge Discovery, 9, e1305, doi: 10.1002/widm.1305

Meneghetti, M. 2021, Lecture Notes in Physics, Vol. 956, Introduction to Gravitational Lensing: With Python Examples (Cham: Springer International Publishing), doi: 10.1007/978-3-030-73582-1

Narayan, R., & Bartelmann, M. 1997, Lectures on Gravitational Lensing, arXiv. http://arxiv.org/abs/astro-ph/9606001

Newton, I. 1704, Opticks: or, A treatise of the reflections, refractions, inflexions and colours of light : also two treatises of the species and magnitude of curvilinear figures (London: Printed for Sam. Smith, and Benj. Walford), doi: 10.5479/sil.302475.39088000644674

Paszke, A., Gross, S., Massa, F., et al. 2019, PyTorch: An Imperative Style, High-Performance Deep Learning Library, arXiv. http://arxiv.org/abs/1912.01703

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, The Astronomical Journal, 124, 266, doi: 10.1086/340952

—. 2010, The Astronomical Journal, 139, 2097, doi: 10.1088/0004-6256/139/6/2097

Plato, ed. 2010, Plato's cosmology: the Timaeus of Plato, first issued in paperback edn., The international library of philosophy 4, Ancient philosophy : in 10 volumes No. 4 (Oxfordshire New York, NY: Routledge)

Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses, ed. I. Appenzeller, G. Börner, M. Harwit, R. Kippenhahn, J. Lequeux, P. A. Strittmatter, & V. Trimble, Astronomy and Astrophysics Library (Berlin, Heidelberg: Springer Berlin Heidelberg), doi: 10.1007/978-3-662-03758-4

Sérsic, J. L. 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41

—. 1968, Atlas de Galaxias Australes

Verma, A. 2000, Current Science, 78, 804. https://www.jstor.org/stable/24103956

Will, C. M. 1988, American Journal of Physics, 56, 413, doi: 10.1119/1.15622

—. 2014, Living Reviews in Relativity, 17, 4, doi: 10.12942/lrr-2014-4

—. 2015, Classical and Quantum Gravity, 32, 124001, doi: 10.1088/0264-9381/32/12/124001

Ringraziamenti

Nel momento in cui questo capitolo della mia vita si chiude e ne inizia uno nuovo, desidero esprimere la mia più profonda gratitudine a coloro che hanno reso possibile questo viaggio. La realizzazione di questa tesi magistrale non sarebbe stata possibile senza il sostegno, l'incoraggiamento e l'amore di molte persone straordinarie che mi hanno accompagnato lungo il percorso. È un privilegio avere l'opportunità di esprimere il mio sincero apprezzamento a queste persone speciali.

Grazie a tutti.