DOCTORAATSPROEFSCHRIFT

2010 | Faculteit Wetenschappen

Genetic algorithms for the non-parametric inversion of gravitational lenses

Proefschrift voorgelegd tot het behalen van de graad van
Doctor in de Wetenschappen, richting fysica, te verdedigen door:

Jori LIESENBORGS

Promotor: prof. dr. Philippe Bekaert
Copromotor: prof. dr. Sven De Rijcke

D/2010/2451/18

universiteit hasselt

Preface

For some reason I always find it difficult to start writing a preface. The thought of skipping the preface altogether and subsequently claiming it must have gotten lost in the print process brought a smile to my face, yes perhaps even a chuckle. Nevertheless, there are some things that need to be said, and people to be thanked, thereby ending this small fantasy of mine.

The dissertation that lays before you concludes a number of years of research, that actually started when working on my master's thesis. If I hadn't chosen the topic about gravitational lensing at the time, or if that thesis subject simply wasn't available then, you would be reading a completely different text, if any. I would therefore like to thank Sven De Rijcke, not only for introducing me to this topic years ago, but also for his continuing guidance in the following years. Of course, my gratitude also goes to my promoter, Philippe Bekaert, without whom pursuing this doctoral degree would not have been possible.

When asked about my research topic, I like to start with saying that it belongs to the category "computational astrophysics". Apart from enjoying the fact that this immediately impresses people, I actually believe that it covers the content fairly well. Furthermore it stresses that this is an interdisciplinary work, tackling an astrophysics problem with computational methods. It was therefore very convenient to have my desk at the EDM, a computer science research institute of the Hasselt University, as this gave me access to its high-quality computational infrastructure. I would like to thank Eddy Flerackers, Frank Van Reeth and Peter Vandoren for making this possible.

The fact that I have enjoyed performing this research during these years is closely related to the nice work environment that everyone of the EDM and of the NVE group that I'm part of, contributes to. Some friends and colleagues that I would like to thank in particular are Wim Lamotte, Peter Quax, Panagiotis Issaris, Kris Luyten and Tom Van Laerhoven.

Unfortunately it hasn't all been fun and games, especially the last year of my PhD. Saying that this was a difficult period would be an understatement at the least, and I wouldn't be finishing this dissertation right now without the support of colleagues, friends and family. I would therefore like to thank my parents, whom I can always count on to find stability and support in difficult times. Furthermore, they have made it possible to start studying physics after finishing computer science studies, certainly a privilege that not everyone can enjoy. I would also like to thank Stijn Neuteleers, a friend for nearly twenty years now. Although we don't meet that often, I always enjoy our conversations as they help me put things into perspective. Finally, I would like to thank my wife for her support and care. So Mieke, thank you ... for everything.

Jori Liesenborgs
April 12, 2010

Samenvatting

De titel van deze thesis, *"Genetische algoritmes voor de niet-parametrische inversie van gravitationele lenzen"*, verdient enige verduidelijking. Laat ons van achter naar voren werken, en beginnen met het sterrenkundige begrip *gravitationele lens*, een soort kosmische luchtspiegeling. Figuur S.1 maakt duidelijk hoe dit in zijn werk gaat: een zeer ver weg gelegen bron zendt lichtstralen uit die in de buurt komen van een uitermate massief object, zoals een galaxie of een cluster van galaxieën. Dankzij de algemene relativiteitstheorie van Einstein weten we dat de zwaartekracht van zo'n object in staat is om deze lichtstralen af te buigen. Het object dat de afbuiging veroorzaakt wordt een gravitationele lens genoemd. Na hun afbuiging zetten de lichtstralen dan hun koers richting Aarde verder, om deze uiteindelijk vanuit verschillende richtingen te bereiken. Doordat de lichtstralen van eenzelfde bron een waarnemer op Aarde vanuit verschillende richtingen bereiken, zal deze waarnemer kopieën van eenzelfde bron zien op verschillende plaatsen aan de hemel. In het voorbeeld op de figuur zal de waarnemer dus niet één heldere bron recht achter de cluster van galaxieën zien, maar zal hij drie zogenaamde *beelden* van die bron waarnemen. Rechts-boven in de figuur is een opname te zien die gemaakt is met de Hubble ruimtetelescoop. De vier heldere objecten die duidelijk te zien zijn, zijn eigenlijk vier beelden van eenzelfde bron, veroorzaakt door het gravitationele-lenseffect van een cluster van galaxieën, de gele objecten.

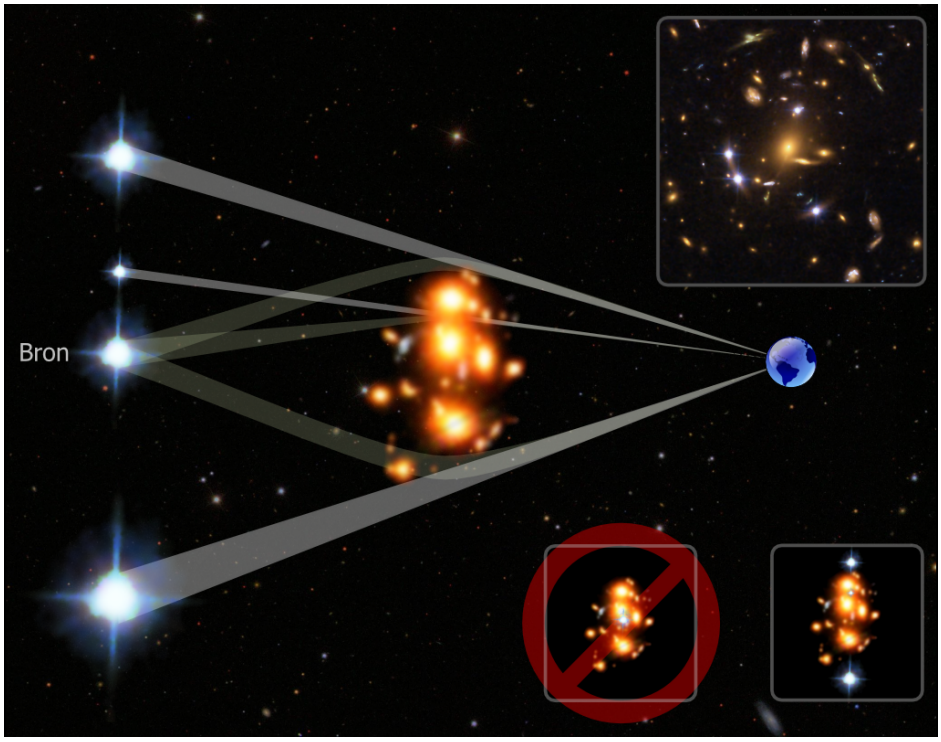
De exacte plaatsen waar de verschillende beelden van eenzelfde bron te zien zijn, hangen af van de zwaartekrachtwerking van de gravitationele lens, en dus van diens massaverdeling. De beelden encoderen daardoor informatie over deze massaverdeling en men kan proberen op basis van enkel de beelden die informatie weer te decoderen en zo een idee te krijgen van de massaverdeling van de lens. Men spreekt dan van *inversie* van de gravitationele lens[1]. Hiervoor zou men een bepaalde vorm van de gravitationele lens kunnen vooropstellen, bijvoorbeeld een zekere elliptische massaverdeling. Inversie van de lens betekent dan het zoeken naar juiste parameters voor zo'n model, zoals bijvoorbeeld de ellipticiteit of de oriëntatie van de ellipsen. Typisch gaat het over een relatief klein aantal parameters, en men noemt zulke technieken *parametrische methodes*. Hiertegenover staan de zogenaamde *niet-parametrische methodes*, waarbij men tracht geen aannames te maken over de vorm van de massaverdeling. De ironie wil dat men dit typisch doet door de massaverdeling op te bouwen uit een zeer groot aantal basisfuncties, waarvan men vervolgens de gewichten nog moet bepalen. In werkelijkheid gebruiken niet-parametrische methodes dus net een zeer groot aantal parameters.

Er bestaan verschillende technieken om goede waarden te vinden voor parameters voor een model, al dan niet parametrisch. In deze thesis wordt gebruik gemaakt van zogenaamde *genetische algoritmes*, een optimalisatiemethode die geïnspireerd is

[1] Als men alles weet over de bron en de gravitationele lens, kan men vrij makkelijk berekenen hoe de beelden er uitzien, en men noemt dit ook wel het *voorwaartse probleem*. Als men de situatie omkeert en op basis van de beelden probeert iets te achterhalen over de lens en de bron, spreekt men van het *inverse probleem*.

door de evolutietheorie van Darwin, door het principe van *"survival of the fittest"*. Concreet zal men met een groot aantal willekeurige modellen beginnen, en men zal deze steeds combineren en muteren, naar analogie met voortplanting in biologische organismen. Hierbij wordt ervoor gezorgd dat betere modellen een groter nageslacht krijgen, zodat generatie na generatie gunstige eigenschappen van bepaalde modellen benadrukt worden, en ongunstige uitsterven. Op deze manier tracht men dus goede oplossingen voor een probleem te *kweken*.

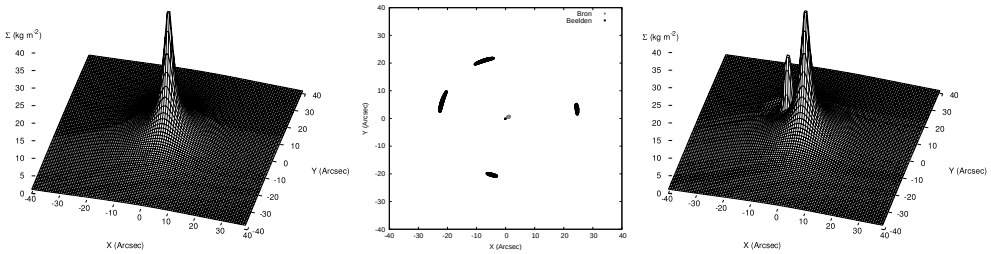
Alle ingrediënten van de titel van dit werk zijn hiermee beschreven. Deze thesis gaat over de ontwikkeling van methodes – gebaseerd op genetische algoritmes – om de massaverdeling van een gravitationele lens te achterhalen. Hierbij worden enkel de beelden die door de lens veroorzaakt worden, gebruikt als invoer; er zullen geen aannames gemaakt worden over de vorm van de massaverdeling. Het nut van mas-



Figuur S.1: De afbuiging van lichtstralen van een verre bron door een tussenliggend erg massief object, zoals bijvoorbeeld een cluster van galaxieën, ligt aan de basis van het gravitationele-lenseffect. Doordat de lichtstralen een waarnemer op aarde zo via verschillende wegen kunnen bereiken, zal de waarnemer verschillende beelden van dezelfde bron te zien krijgen. Rechts-boven wordt een foto getoond, gemaakt met de Hubble ruimtetelescoop. Daarin zijn duidelijk vier zeer heldere beelden van een quasar te zien, veroorzaakt door de cluster van galaxieën die te zien is in het geel.

sabepaling op basis van het lenseffect, is dat ze onmiddellijk vat kan geven op *alle* materie in de gravitationele lens: niet alleen op dat deel dat we zowiezo al met bijvoorbeeld een telescoop kunnen waarnemen, maar ook op de meer mysterieuze *donkere materie*. Vergelijking van zulke massareconstructies met theoretische modellen, kan ons informatie verschaffen over de eigenschappen van donkere materie.

Om te achterhalen hoe betrouwbaar een massareconstructie is, is het belangrijk de zogenaamde *ontaardingen* in dit vraagstuk te begrijpen. Het blijkt namelijk dat wanneer men één oplossing van het inversieprobleem gevonden heeft, men ook andere, gelijkwaardige oplossingen kan reconstrueren, die even compatibel zijn met de waargenomen beelden. Men spreekt dan van ontaarde oplossingen. Een voorbeeld hiervan is te zien in figuur S.2. Een eenvoudige berekening toont dat de massaverdeling aan de linkerzijde de cirkelvormige bron in het middenpaneel transformeert tot vijf beelden. In deze thesis wordt ondermeer aangetoond dat de massaverdeling aan de rechterzijde precies hetzelfde effect heeft. Bij het maken van deze alternatieve oplossing, werd in een cirkelvormig gebied wat massa *'geleend'* om een extra piek in de massaverdeling te creëren. Het bestuderen van deze en andere soorten ontaardingen, brengt meer duidelijkheid wanneer men een gevonden massareconstructie wil interpreteren. Het voorbeeld in de figuur maakt al duidelijk dat niet alle eigenschappen van zo'n reconstructie even betrouwbaar zullen zijn.



Figuur S.2: De massaverdeling aan de linkerzijde transformeert de cirkelvormige bron in het middelste paneel tot vijf beelden. De massaverdeling aan de rechterzijde, die duidelijk verschilt van de eerste massaverdeling, heeft echter precies hetzelfde effect. Op basis van enkel de beelden, zou men dus geen enkel onderscheid tussen beide modellen kunnen maken; ze zijn beide even goed mogelijk. Dit is een voorbeeld van zogenaamde ontaardingen in het vraagstuk van gravitationele-lensinversie.

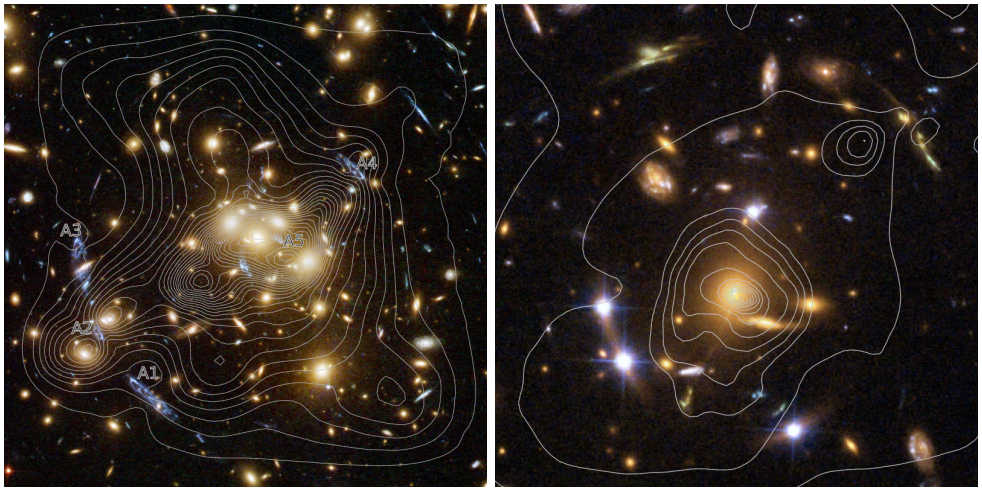
Zoals al eerder gezegd werd, maakt een gravitationele lens meerdere kopieën – die men "beelden" noemt – van eenzelfde bron zichtbaar en voor de eigenlijke inversiemethode is het nodig om deze beelden te identificeren. In een eerste fase wordt dan op basis hiervan een vrij ruwe schatting van de massaverdeling gemaakt, wat neerkomt op het gebruik van een relatief laag aantal basisfuncties. Op basis van deze eerste schatting wordt vervolgens een nieuwe massaverdeling gemaakt, met een hogere resolutie in gebieden die – volgens de eerdere schatting – meer massa bevatten. In die gebieden worden de basisfuncties daarom wat dichter bij elkaar geplaatst. Deze verfijningsstap kan uiteraard een aantal keer herhaald worden, tot men een aanvaardbaar model voor

de massaverdeling gevonden heeft.

Telkens er in een van zulke stappen een massaverdeling bepaald moet worden, wordt een genetisch algoritme gebruikt om de gewichten van de basisfuncties vast te leggen. Om het principe van "survival of the fittest" na te bootsen, is er een criterium nodig om te beslissen welke modellen beter zijn dan andere, een zogenaamd *fitness criterium*. Hiervoor wordt nagegaan in welke mate een model compatibel is met de waargenomen beelden. Terugkijkend naar figuur S.1 kan men de lichtstralen ook achterwaarts volgen, en voor een goed model moet men zo voor elk beeld weer bij eenzelfde bron uitkomen. In het begin van de procedure, waarbij de modellen nog zeer willekeurig zijn, zal dit echter niet het geval zijn. Wanneer men de lichtstralen dan volgt, zal men voor elk beeld een andere bronpositie bekomen. Als fitness criterium wordt dan gebruikt hoe ver zulke voorspelde bronposities van elkaar verwijderd liggen. Door modellen die ze dichter bij elkaar brengen te bevoordelen, worden er generatie na generatie oplossingen gekweekt die hier steeds beter in zijn. Uiteindelijk bekomt men zo een model waarbij de beelden telkens door één enkele bronpositie verklaard worden. Naast dit belangrijke criterium kan een genetisch algoritme ook verschillende criteria tegelijk optimaliseren, zodat de methode kan uitgebreid worden met alle bruikbare informatie. Een belangrijk voorbeeld is dat een goed model niet alleen de waargenomen beelden moet voorspellen, maar ook moet vermijden dat extra beelden voorspeld worden.

Om na te gaan of deze aanpak werkt, wordt er gebruik gemaakt van gesimuleerde lenssituaties. Concreet worden een massaverdeling en een aantal bronposities vooropgesteld, en op basis daarvan worden de posities van de verschillende beelden berekend. Deze beelden dienen dan als invoer voor de procedure, en de reconstructie van de massaverdeling kan dan vergeleken worden met de oorspronkelijke. De voorgestelde inversiemethode blijkt succesvol te zijn in het reconstrueren van de massaverdeling; de nauwkeurigheid waarmee dit gebeurt hangt echter af van de hoeveelheid en de locaties van de verschillende beelden. Enkele toepassingen van de methode op waargenomen gravitationele-lenssystemen zijn te zien in figuur S.3, waarin contouren van de massareconstructie getoond worden bovenop beelden gemaakt met de Hubble ruimtetelescoop. In deze situaties waren er telkens een beperkt aantal beelden beschikbaar, maar zelfs op basis van enkel deze informatie wordt een massaverdeling voorspeld die in grote lijnen de zichtbare massa volgt.

Er bestaan uiteraard ook andere methodes voor lensinversie. Net als andere niet-parametrische methodes heeft de voorgestelde techniek als voordeel dat er geen aannames over de vorm van de massaverdeling gemaakt worden. Het grootste voordeel van deze procedure is echter dat het gebruik van een genetisch algoritme toelaat om zeer flexibel allerlei soorten beschikbare gegevens te gebruiken; men moet enkel kunnen aangeven of een bepaalde voorgestelde oplossing beter is dan een andere. De prijs van deze flexibiliteit is een wat langere berekening, die wel verlicht kan worden door het rekenwerk op te splitsen over meerdere processoren.



(a) Cl 0024+1654 (b) SDSS J1004+4112

Figuur S.3: De ontwikkelde inversiemethode werd toegepast op twee bekende gravitationele lenzen, namelijk de clusters Cl 0024+1654 en SDSS J1004+4112. Op de achtergrond is telkens een opname te zien, gemaakt met de Hubble ruimtetelescoop. De reconstructie van de massaverdeling op basis van het gravitationele-lenseffect is te zien als contouren bovenop deze afbeeldingen. Hoewel er in beide gevallen slechts weinig beelden gebruikt werden, vindt men al snel een massaverdeling terug die gelijkenissen vertoont met de verdeling van het licht. De kleurenbeelden komen uit het Space Telescope Science Institute (STScI) archief[a,b].

[a] `http://hubblesite.org/newscenter/archive/releases/2007/17/image/b/`
[b] `http://hubblesite.org/newscenter/archive/releases/2006/23/image/a/`

Contents

—1—

Introduction

The title of this dissertation, *Genetic algorithms for the non-parametric inversion of gravitational lenses*, is certainly a mouthful and demands some clarification. The purpose of this first chapter is therefore to introduce two important parts: gravitational lenses and genetic algorithms. This should give the reader a firm grasp on what this work tries to accomplish, before delving into the more technical aspects in the remaining chapters.

1.1 The gravitational lens effect

The gravitational lens effect is a spectacular astrophysical phenomenon, with a myriad of applications. This section shows the effect itself, what causes it and provides a brief historical overview of research into gravitational lenses.

1.1.1 Example and working

The easiest way to explain what the gravitational lens effect is, is by showing it. Figure 1.1 shows the cluster of galaxies Cl 0024+1654 and is a beautiful example of the lens effect. The cluster galaxies themselves have a yellowish color, and together form an object that is so massive that it causes a significant deflection of the light rays that pass nearby. Because of this, a kind of cosmic mirage occurs: several distorted images of the same blue galaxy – which is located much farther away than the cluster itself – can be seen.

Figure 1.2 illustrates how the deflection of light can cause an observer to see several copies of the same object. A light ray starts at the distant astronomical source and

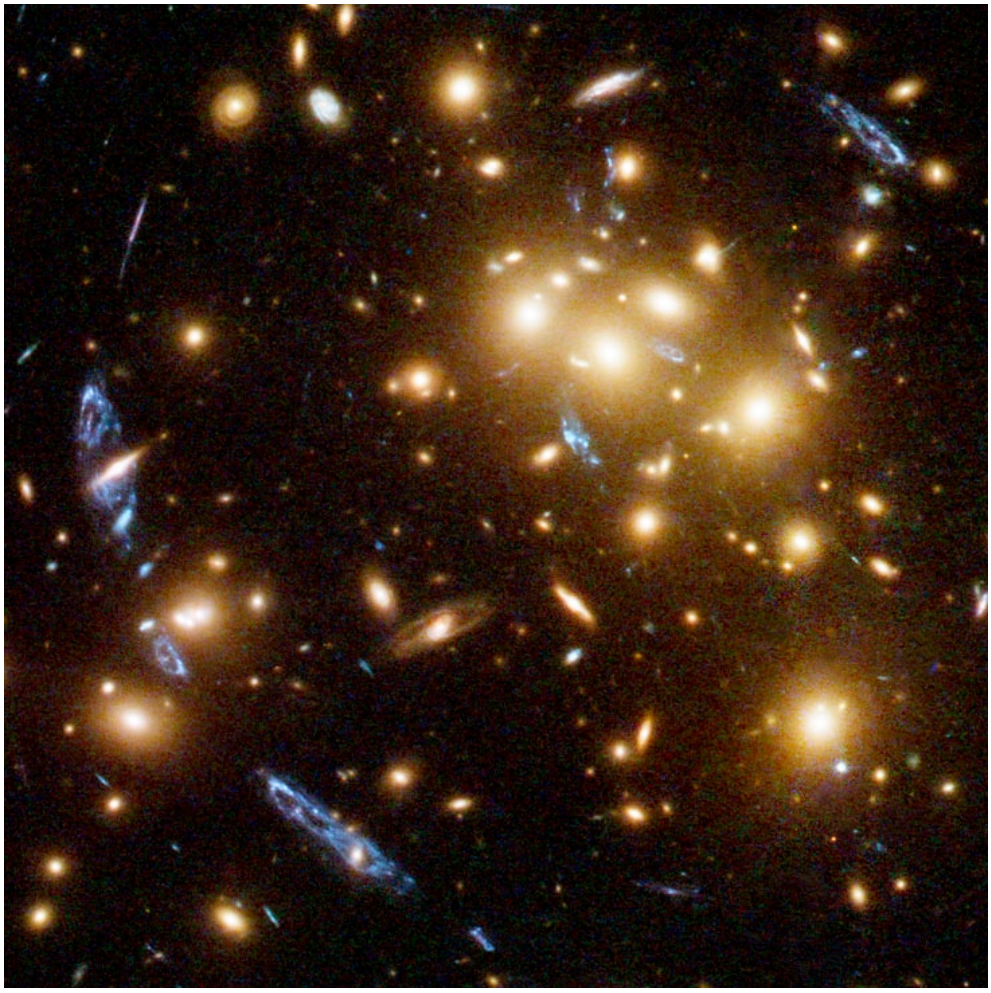


Figure 1.1: The distant cluster of galaxies Cl 0024+1654 (yellow objects) has such a strong gravitational field that it causes a significant deflection of light rays. Because of this, multiple – somewhat distorted – copies of the same blue background galaxy can be seen at several locations. This is one of the larger gravitational lens systems, in which the largest image separation is approximately 1 arcmin. The color image was taken from the Space Telescope Science Institute (STScI) archive[a].

[a] http://hubblesite.org/newscenter/archive/releases/2007/17/image/b/

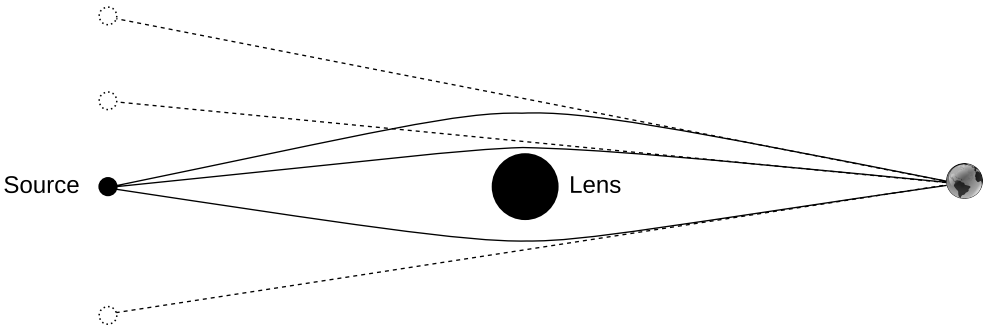


Figure 1.2: In a gravitational lensing scenario, light rays from a distant source are deflected by the gravitational field of an intermediate object – the gravitational lens – before continuing towards the observer on earth. This can cause the observer to see the source in a different direction, or can even cause multiple copies of the same source to be visible.

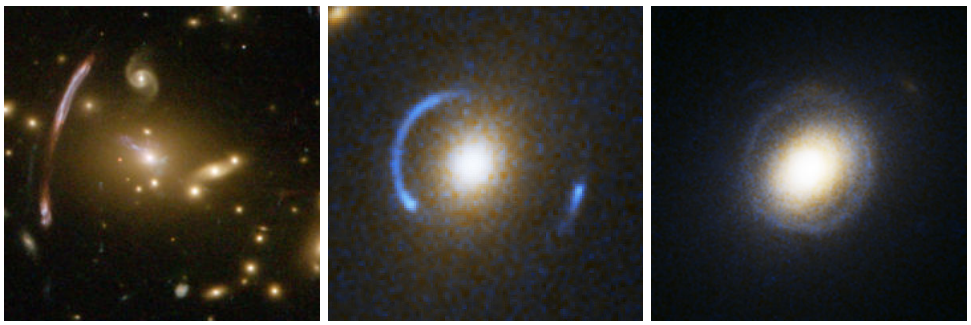
travels in the general direction of the observer. Due to the strong gravitational field of an intermediate object, e.g. the cluster of galaxies in the previous example, the light ray gets deflected however. This will cause the observer on earth to see the source in a different direction than would be the case if the intermediate object were not present. As is illustrated in the figure, in case of a very strong gravitational field it is possible that light from the source can reach the observer by means of several different paths. This will then cause the observer to see multiple copies of the same object. The intermediate object responsible for the deflection of light rays is called the *gravitational lens*; the copies of the source that an observer sees due to the lens effect are usually called *images* of that source. The term *source* itself is usually reserved for the object that one would see if the deflection of light rays could be turned off in some way. Depending on the strength of the deflection, the shape of the source and the complexity of the gravitational field of the lens, the images can show significant distortion.

1.1.2 Regimes and applications

Depending on the scale of the lens effect and the alignment between source, lens and observer, the lens effect is usually subdivided into a number of regimes. Below we shall describe the *strong lensing*, *weak lensing* and *microlensing* regimes, as well as some of their possible applications.

Strong lensing

In the strong lensing regime, the gravitational field of the lens is sufficiently strong and the alignment between source, lens and observer is well enough so that multiple images can be formed. The cluster lens Cl 0024+1654 is a very nice example of a



(a) Abell 2667 (b) SDSS J1205+4910 (c) SDSS J2321+0939

Figure 1.3: If the alignment between source, gravitational lens and observer is well enough, the images of a source can show a large amount of deformation. This can cause arc-like images, partial and even full rings. In case a full ring is visible, one often speaks of an Einstein ring. The color images were taken from the STScI archive [a,b,c].

[a] http://hubblesite.org/newscenter/archive/releases/2007/12/image/a/
[b] http://hubblesite.org/newscenter/archive/releases/2005/32/image/d/
[c] http://hubblesite.org/newscenter/archive/releases/2005/32/image/i/

strong lensing system. Often one will encounter highly deformed images, possibly even arc-like or ring-like features, of which some examples are given in figure 1.3.

The precise location of the images depends on the source position and the gravitational field of the lens, i.e. on its mass distribution. The images therefore encode some information about said mass distribution, and one can try to decode this information and reconstruct the mass density of the gravitational lens. Often, this procedure is called *gravitational lens inversion* and several methods exist. In *parametric* methods, a specific shape of the lens is assumed, e.g. an elliptical lens with a particular radial profile. Inverting the lens then corresponds to finding a relatively small number of appropriate parameters that can explain the observed properties of the source. In *non-parametric* methods, one tries to avoid making assumptions about the shape of the mass distribution. This can be done by using a large amount of basis functions for example[1]. As lens inversion is the focus of this dissertation, a more detailed overview of lens inversion methods will be given in chapter 4. The lens effect depends on the distribution of all the matter in the gravitational lens, both luminous and dark. Having such an independent probe of the mass distribution is very useful of course, as comparing the result with the visible mass provides some information about the distribution of dark matter.

The strong lens effect not only causes multiple images of the same source to appear, but most often these images are also deformed. This makes it possible for images to become much larger than the actual source, which in turn makes them much easier to

[1] For this reason, the term "non-parametric" is somewhat misleading, as it usually involves finding a *large* number of appropriate parameters.

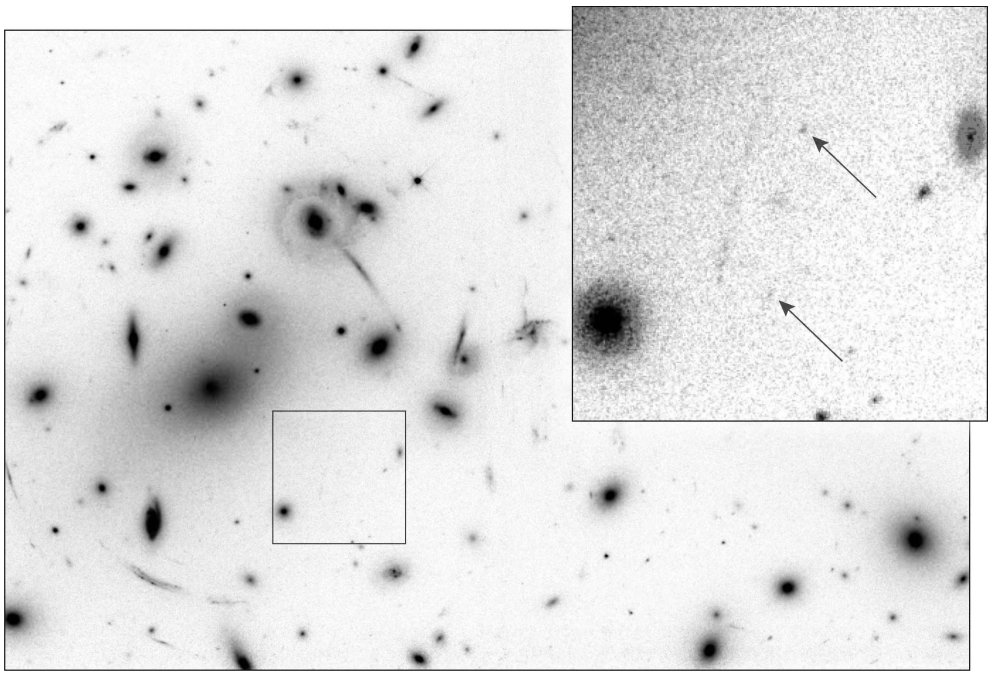


Figure 1.4: The cluster lens Abell 2218 creates multiple highly magnified images of the same distant background galaxy. The source itself is located at a redshift of $z \approx 5.6$ and would otherwise be impossible to detect [29].

see. Gravitational lenses are therefore a kind of *cosmic telescope*, allowing an observer to view objects which would otherwise be impossible to detect. Figure 1.4 shows a very nice example of this type of application. Here, two images of the same distant source are highly magnified by a gravitational lens, the cluster Abell 2218.

When light can reach the observer by means of different paths, the light travel time for each of these paths will differ. If the source itself has a variable brightness, each image will contain similar brightness fluctuations at a different time, allowing the time delays between different paths to be measured. It was already noted by S. Refsdal in 1964 that this time delay depends both on the mass distribution of the gravitational lens and on the *Hubble constant* [86]. This makes it possible to provide an independent measure of the Hubble constant by analyzing gravitational lens systems (e.g. [90]).

Statistical information about gravitational lenses, can, at least in principle, help constrain several other cosmological parameters, like Ω_{m} or Ω_{Λ}, or even the equation of state parameter w of dark energy (see e.g. [40] or [36]). Currently however, the best constraints still come from the analysis of the Cosmic Microwave Background radiation.

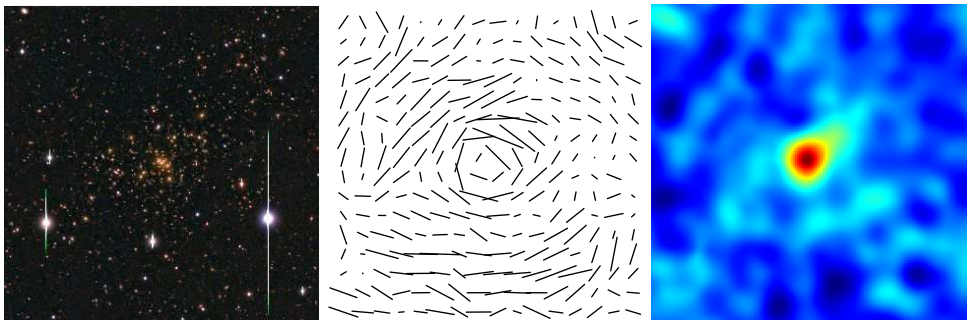


Figure 1.5: Left panel: central 8×8 arcmin^2 region of the cluster Cl 0024+1654. Center panel: shear map determined from the distortion of background galaxies in this cluster. Right panel: mass reconstruction based on the shear map in the center panel (images from [104]).

Weak lensing

When source, lens and observer are less well aligned, the observer will only see a single image of the source and one speaks of the weak lensing regime. The image will still show some deformation, but since the real shape of the source is not known, this deformation cannot be fixed based upon the information contained in the image alone. However, if a large group of images, located close to each other in the sky, is analyzed, it is possible to detect how these images were deformed.

To see how this is possible, first suppose that one still has only one source, but perfectly round in shape. When this source shape is affected by a gravitational lens, the observer will no longer see a circular object, but instead will see an elliptical one. By inspecting the ellipticity and orientation of the ellipse, one can then learn how the gravitational lens has affected the source shape. Now, in practice each individual source is not round, but when a group of sources is considered in absence of lensing, one expects them to be round on average, that is, they will not show a preferential alignment. When there is a gravitational lens that distorts the intrinsic shapes, the average shape that one obtains in this way will no longer be round, but will now show some preferential direction.

By considering a large number of background galaxies this way, one can construct a map of these distortions, usually called a *shear map*. In turn, this shear map can be used to obtain an estimate of the mass distribution of the gravitational lens (e.g. using [51]). The most important application of weak lensing is therefore related to mass reconstruction. An example of a shear map and the corresponding mass reconstruction can be seen in figure 1.5.

Since weak lensing does not have the requirement that multiple images are formed, it can provide information about the mass in regions where the gravitational field is less strong. Typically this is in the regions which lie further away from the central mass concentration. On the other hand, in the central region itself, the weak lensing

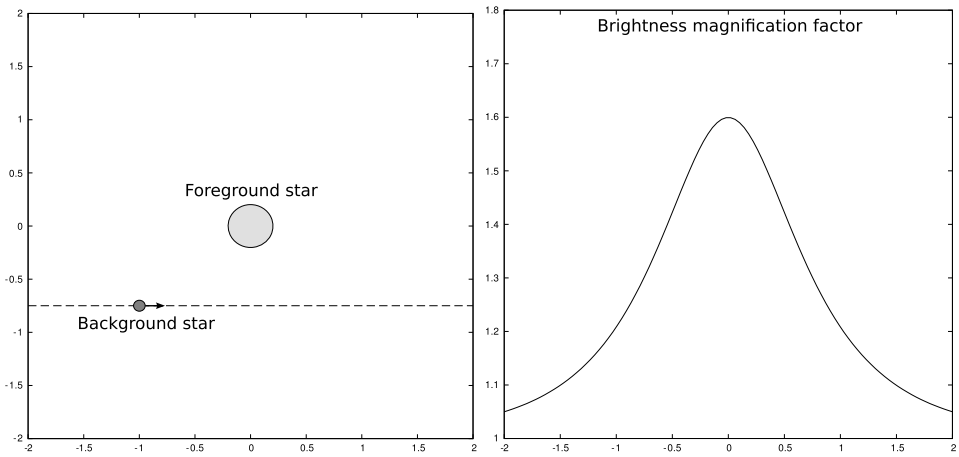


Figure 1.6: When a background star moves with respect to a foreground star, the received flux from this background star will change due to the gravitational lens effect. For the trajectory shown in the left panel, the expected brightness variation is shown in the right panel.

approach usually fails because the distortions of the images become too large. Trying to take the best of both worlds, some methods analyze both weak and strong lensing at the same time, to obtain a mass reconstruction for a very large region (e.g. [26]).

Microlensing

The term "microlensing" is used to describe lensing on a smaller scale, in which there may be multiple images, but the effect is too small to be resolved. A typical example is lensing of one star by another star. Of course, if the scale of the effect is too small to be resolved, the deformation in the image(s) cannot be seen directly. However, because the size of the images, although not resolved, does change, the lens effect has also an impact on the flux that we receive of the background star. More specifically, if source star and lens star move with respect to each other, the received flux of the background star will change in a very specific manner (see figure 1.6).

This brightness variation was used to search for Massive Compact Halo Objects, or MACHOs. Such objects, like brown dwarfs or planets, are a candidate for dark matter in our own Galaxy, and by statistically analyzing the number of lensing events, constraints can be placed on this form of dark matter. Results from the so-called MACHO Project indicate that observed lensing events are likely to explain the existence of roughly 20% of dark matter in this form [2].

This brightness variation is typical for lensing by a star, or a compact object in general, but will be different if the lens configuration is modified somewhat. For example, it is certainly possible that instead of a single star, a binary star is found instead, leading

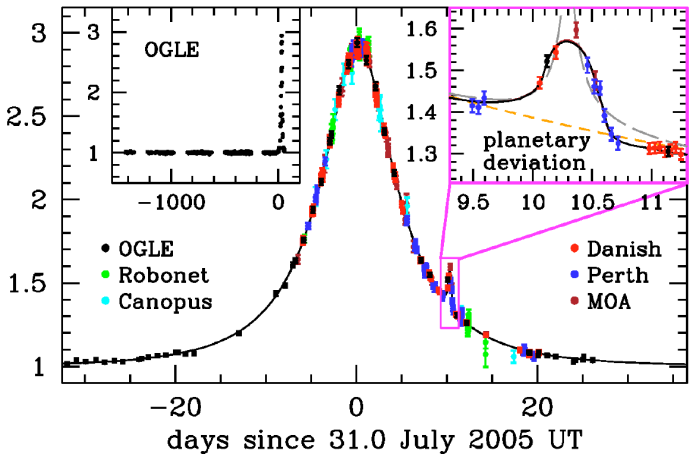


Figure 1.7: If the gravitational lens is a star with a planet orbiting around it, the typical brightness variation of the background source will differ from the expected pattern. This image shows such a deviation, and indicates the presence of a planet [3].

to a different brightness pattern. It is also possible that around the star a planet is orbiting, also leading to different brightness patterns. Currently, several planets have already been found by inspecting brightness fluctuations in this way. A nice example of such a brightness variation can be seen in figure 1.7.

The term microlensing is also often used in another context. When a background quasar is lensed by a galaxy for example, usually the image positions can be explained by a relatively simple model. The fluxes that are received from each image however, usually differ from the expected results. Apart from intrinsic quasar variability, it is also possible that such fluxes are influenced by individual stars in the lensing galaxy. If this is the case, one also speaks of microlensing, since the star itself will not have much influence on the position of an image, but can have a significant impact on the received flux. If the deviations are caused by larger substructures, often the term millilensing is used.

1.1.3 A brief history of gravitational lensing

Below a short account of the history of research into the gravitational lens effect will be given. The interested reader is referred to [96] for a thorough historical review.

The deflection of light rays by gravity can be traced back to Isaac Newton himself, who already suspected that light could indeed be influenced by a gravitational field. The first published result however, is attributed to Johann Georg von Soldner, who showed in 1801 that using Newtonian mechanics, the size of the deflection angle $\hat{\alpha}$

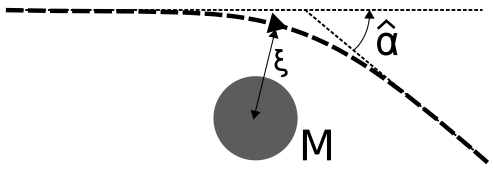


Figure 1.8: A light ray that passes close to a very massive object undergoes a noticeable deflection. For a point mass deflector and in the limit of a small deflection angle $\hat{\alpha}$, this angle only depends on the mass M of the deflecting object and the distance of closest approach ξ.

(see figure 1.8) in the limit of a small deflection is given by the formula

$$\hat{\alpha}_{\text{Newtonian}} = \frac{2GM}{c^2\xi}.$$

Here, M is the mass of the deflecting object, G is the gravitational constant, c is the speed of light and ξ is the distance of closest approach.

More than one hundred years later however, Albert Einstein, having developed his General Theory of Relativity, published a new result. He calculated that for the deflection of a light ray, the deflection angle is actually twice the Newtonian result:

$$\hat{\alpha} = 2\hat{\alpha}_{\text{Newtonian}} = \frac{4GM}{c^2\xi}.$$

If one calculates this deflection angle for light rays that pass close to the sun, i.e.

$$M = M_\odot = 1.99 \times 10^{30}\text{kg},$$

$$\xi = R_\odot = 6.96 \times 10^{8}\text{m},$$

one obtains a value of 1.74 arcsec.

By measuring the positions of stars very close to the sun, and comparing these measurements to the predicted positions of the stars without considering the lens effect, one can try to measure the value of this deflection angle. For obvious reasons, one has to wait for a solar eclipse to occur to attempt such a measurement, and this is precisely what Arthur Eddington did in 1919. The result he obtained was that the measurements of the deflection angle were compatible with the value predicted by Einstein, taking into account the 20% error margin. Currently, this error margin has been reduced to 0.02%[66].

In the early days after the confirmation of this gravitational lens effect, astronomers studied the problem of lensing of a background star by some foreground star. It was also Eddington who pointed out in 1920 that, if these stars are sufficiently well aligned, multiple images can occur. A few years later, in 1924, O. Chwolson remarked that if the alignment is perfect, the background star would be transformed into a ring-like

image. Ironically, today such a ring-like feature is called an Einstein ring instead of a Chwolson ring. The term "gravitational lens" was first mentioned by R. W. Mandl in the 1930's, although O. J. Lodge had already introduced the term "lens" in 1919.

As Einstein remarked in 1936, the optical telescopes then available would not be able to resolve multiple images from a background star. A year later however, Fritz Zwicky published an article in which he claimed that lensing by extragalactic nebulae – as galaxies were still called then – would have a much better chance at being observed. Zwicky also mentioned the importance of the gravitational lens effect as an independent mass probe and as a natural telescope, allowing very distant objects to be studied more easily.

The years that followed were rather silent with respect to gravitational lensing publications. About one quarter of a century later however, at around the same time quasars were discovered, the field of gravitational lensing received new input. In the 1960s, Yu. G. Klimov, S. Liebes and S. Refsdal independently studied the lens effect again, from lensing by stars to lensing by galaxies. Refsdal also considered that the difference in light travel time for multiply-imaged sources may be measurable if the source shows brightness fluctuations. Furthermore, this time delay was shown to depend on the Hubble constant, and therefore offered the possibility to measure its value.

Until 1971, the deflecting object was always modeled as a point mass. At that time however, N. Sanitt studied the influence of an extended mass distribution as a model for a galaxy-lens. For the departure from spherically symmetric lens models, one still had to wait two more years for the works of R. R. Bourassa, in which deflection by a more general spheroidal mass distribution was studied. Sanitt also noted that, in the case of quasars, the difference in size between continuum and line-emission regions should lead to a different magnification for these two features. A more general expression for the time delay between images was obtained by J. H. Cooke and R. Kantowski in 1975. They showed that it could be decomposed into a part which corresponds to the different lengths of the paths by which light rays reached the observer, and into a part that corresponds to a difference in the gravitational potential that each light ray has to cross. As shown by K. Chang and S. Refsdal in 1979 however, brightness variations can also be caused by small structures, e. g. a star in a galaxy-lens, which makes it more difficult to measure such time delays.

The status of theoretical curiosity that gravitational lenses still had, ended in 1979. D. Walsh, R. F. Carswell and R. J. Weymann announced the detection of the two quasars QSO 0957+561 A,B with remarkably similar spectra, located very close to each other, at about 6 arcsec. The gravitational lensing nature of this system was confirmed when the lensing galaxy was detected close to one of the quasar images. The detection of the first observed gravitational lensing system drew much interest to the lensing topic, both observational and theoretical, and the number of articles related to gravitational lensing has seen an enormous increase since then.

1.2 Genetic algorithms

Genetic algorithms are an optimization strategy inspired by the principle of natural selection, or survival of the fittest. The story therefore necessarily starts with Charles Darwin's famous work, *On the origin of species* [21], which he published in the year 1859. The central idea is that in a certain population, those individuals that are adapted better to the living circumstances, have a better chance at survival and consequently a better chance at passing on specific traits to their offspring. Darwin himself was not aware of how these characteristics were passed on from one generation to the next. It was Gregor Mendel who offered the first insights into these matters a while later, although it took a few decennia for his work to become rediscovered.

In a genetic algorithm, one effectively tries to *breed solutions* to a problem. One starts with an initial *population* of trial solutions, encoded in a user defined way. In analogy to biological systems, an encoded representation of a trial solution can be called a *chromosome* or a *genome* for example. The first population is usually initialized in some random way, and using this population, a new one will be created. To do so, one chooses genomes for *recombination* (also called crossover) or *cloning*, mimicking sexual and asexual reproduction respectively. When selecting genomes for this step, it is important to apply *selection pressure*: one would like better trial solutions to produce more offspring. One therefore needs a so-called *fitness measure*, which allows the algorithm to detect which genomes in the current population encode better trial solutions. As a final step, *mutations* are introduced into the new population to ensure genetic variety, to avoid premature convergence to a local optimum. Using the new population these steps are simply repeated, and so, generation after generation, improving trial solutions are generated.

Genetic algorithms belong to a wider class of *evolutionary computation* techniques or *evolutionary algorithms*, optimization procedures applying methods inspired by evolution in biological systems. The history of evolutionary computation is quite complex, with different researchers using slightly different techniques and early work not getting the attention of a wide audience. One of the first published results are the works of Barricelli, starting around 1953. Barricelli was using John von Neumann's high speed computer in Princeton, where he performed some trials in the area of artificial life, evolving generation after generation. Around 1957, Fraser was working on the modeling of genetic systems. In his work, a population of P individuals produced P' offspring, of which a number were eliminated until a new population of size P was obtained again. The offspring themselves were created by recombination of solutions from the parent population. Around the same time, in 1958, Bremermann was studying models in which a binary string represented a trial solution, and in which reproduction, selection and mutation were used for creating a new population. He also provided some theoretical perspectives on the working of his models. Also in 1958, Friedberg was trying to evolve machine language programs to perform simple tasks.

In the 1960s, Ingo Rechenberg and Hans-Paul Schwefel worked on what they called

evolution strategies. Using an evolutionary model, Rechenberg studied evolution towards an optimal airfoil shape. Schwefel used a similar technique to determine the optimal shape of a nozzle in a specific problem. While both in principle did use the concept of a population, the population was only two individuals in size: a parent and its offspring, a mutated version of the parent. In the early 60s, Lawrence Fogel studied simulations of evolution for creating artificial intelligence, called evolutionary programming. In his work only mutation was used to create a new population.

At the University of Michigan, John Holland started his research in the 60s as well. His group studied adaptation in nature and looked for ways to apply similar techniques to artificial systems. The emphasis was not on any specific problem though, but more on the theoretical aspects. This led to his well known work *"Adaptation in natural and artificial systems"*, published in 1975 [42]. Holland was the first to introduce the name "genetic algorithm", and used chromosomes consisting of strings of ones and zeros. To advance to the next generation, recombination, selection and mutation were used. It is interesting to note that a few years earlier, in 1970, Fraser and Burnell also published a book, called *"Computer models in genetics"*, about a model which resembled the genetic algorithm of Holland. For some reason however, this work did not receive as much attention at the time. The work of Ken De Jong, a graduate student of Holland, helped to raise the interest in genetic algorithms. Also in 1975, he wrote his doctoral thesis in which he studied several test optimization problems.

In the years that followed, genetic algorithms became ever more popular as an optimization method. The increasing computing power and inherent parallelism of genetic algorithms undoubtedly played no small role in this trend. The brief historical information in this section is based on [35], [74], [85] and [32]; the interested reader is referred to these works for more detailed information.

To sum up, genetic algorithms are a non-local heuristic optimization method, inspired by the principle of natural selection. Thanks to the use of recombination, the search space is sampled in a non-local way, contrary to a gradient based search for example. It is a heuristic search method, since it is difficult to analyze theoretically. Fortunately however, the many successful applications of genetic algorithms indicate that it does indeed provide a good optimization method. The line which separates genetic algorithms from other evolutionary algorithms is rather vague. In this work, contrary to what was called a genetic algorithm by Holland himself, the notion of a genetic algorithm will not be that strict. Anything that uses a population of trial solutions and which advances to the next generation by means of recombination, selection and mutation will be termed a genetic algorithm.

A few of the many interesting applications of genetic algorithms are:

- a method for fitting quasar spectra [101],
- analysis of supernova data [5],
- search for gravitational wave signatures of inspiraling black hole binaries (in simulated data) [83],

- design of antennae elements for a radio frequency telescope [55],
- allocation of railway platforms [18],
- the analysis of seismic wave data [95],
- method for finding an optimal molecular geometry [22], and
- protein folding studies [105].

1.3 Motivation and thesis outline

Hopefully, this introductory chapter has made it clear to the reader what this dissertation will try to accomplish. Using genetic algorithms as the optimization scheme, the images in strong gravitational lens systems will be analyzed and used to attempt to reconstruct the mass distribution of the gravitational lens. Furthermore, this will be done in such a way that assumptions about the shape of the mass distribution are avoided.

The gravitational lens effect depends on all the matter which is present, independent of whether it emits light or is dark. Without altering the gravitational force, the presence of dark matter is needed to explain for example the velocities of galaxies in clusters (e.g. [115]), or to explain the rotational velocities of disc galaxies (e.g. [88]). Having an independent probe of the total mass distribution, comparing this to the visible mass allows one to obtain an estimate of the dark matter distribution.

Simulations of the evolution of cold dark matter, predict a universal profile for the mass distributions, ranging from dwarf galaxies to clusters of galaxies. In this so-called Navarro-Frenk-White (NFW) profile, for the outer regions of a mass distribution one finds $\rho \sim r^{-3}$, while in the central regions one finds a cuspy $\rho \sim r^{-1}$ dependence [76]. Furthermore, an NFW halo is predicted to have many subhalos.

However, observations do not always seem to agree with these simulations. One discrepancy is called the *cusp-core problem*: certain observations seem to suggest that dark matter at least in some cases does not have a cuspy profile, but instead a flat central core is found (e.g. [94]). Another problem is the *missing satellite problem*, i.e. the Milky Way has a lot less small satellite galaxies than predicted.

Clearly, theory and observations do not match as well as desired, and having a good method of probing the mass distribution directly is therefore of great value. Before describing the inversion method proposed in this dissertation, some background in gravitational lensing theory is provided in chapter 2. To be able to interpret the results of a gravitational lens inversion, one has to know which degeneracies may affect the reconstruction. These are described in chapter 3, after which the inversion procedure is explained and tested using simulations in chapter 4. The method is applied to observed lensing systems in chapter 5 and this work is concluded in chapter 6.

—2—

Lensing theory

To be able to invert a gravitational lens, i.e. to determine its mass distribution based on the observed images, one must first obtain a deeper understanding of how exactly the lens effect is created. It is precisely this that this chapter tries to accomplish: the bending of light rays is studied more rigorously and in doing so, the central equation aptly called the *lens equation*, will be derived.

The theory in this section is largely based upon the information in [15], [75] and [96]. This last reference is recommended for an in-depth treatment of the theory.

2.1 Deflection of a light ray

We shall be interested in deflection caused by weak sources of gravity, and for our purposes the space-time curvature around such an object is well described by the so-called *static weak field metric* or *first order post-Newtonian metric* (e.g. [39]) :

$$\mathrm{d}s^2 = \left(1+\frac{2\Phi}{c^2}\right)c^2\mathrm{d}t^2 - \left(1-\frac{2\Phi}{c^2}\right)(\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2) = \left(1+\frac{2\Phi}{c^2}\right)c^2\mathrm{d}t^2 - \left(1-\frac{2\Phi}{c^2}\right)\mathrm{d}\sigma^2,$$

where Φ is the Newtonian gravitational potential of the object. A necessary condition for this metric to be valid is that $\Phi \ll c^2$, which is certainly the case for galaxies and clusters of galaxies. In a cosmological setting, as usual it is assumed that this metric is a local perturbation of the global metric of the expanding universe. This means that this metric is assumed to be valid in the neighborhood of the deflecting object.

Light rays travel along null-geodesics, i.e. $\mathrm{d}s = 0$, so that the equation becomes:

$$\left(1+\frac{2\Phi}{c^2}\right)c^2\mathrm{d}t^2 = \left(1-\frac{2\Phi}{c^2}\right)\mathrm{d}\sigma^2.$$

Taking into account that $\Phi \ll c^2$, this can easily be rewritten as

$$c\frac{\mathrm{d}t}{\mathrm{d}\sigma} \approx 1 - \frac{2\Phi}{c^2}.$$

This equation makes it clear that the situation is formally equivalent with a light ray traveling through Euclidean space, but containing a location-dependent refractive index:

$$n(\boldsymbol{r}) = \frac{c}{v(\boldsymbol{r})} = 1 - \frac{2\Phi(\boldsymbol{r})}{c^2} \qquad \text{in which } v = \frac{\mathrm{d}\sigma}{\mathrm{d}t}.$$

Note that no wavelength dependence is present in this equation: a gravitational lens is achromatic.

For a specific path, the light travel time is given by:

$$\begin{aligned} t &= \frac{1}{c}\int_S^O n(\boldsymbol{r})\mathrm{d}\sigma \\ &= \frac{1}{c}\int_S^O \mathrm{d}\sigma - \frac{1}{c}\int_S^O \frac{2\Phi(\boldsymbol{r})}{c^2}\mathrm{d}\sigma \qquad (2.1) \\ &= t_{\mathrm{geom}} + t_{\mathrm{grav}}, \qquad (2.2) \end{aligned}$$

in which the integral is taken along a path from source (S) to observer (O). As is indicated in these equations, the light travel time can be split into a geometrical part, which is clearly just the length of the path divided by the speed of light, and a gravitational part which depends on the potential of the deflector. To determine the physical paths that light rays can follow, the principle of Fermat is used, which says that the actual paths are those along which the light travel time is extremal.

It is show in appendix B that for such a physical path, the following relation is valid:

$$\boldsymbol{T}_S - \boldsymbol{T}_O = \frac{2}{c^2}\int_S^O \boldsymbol{\nabla}_\perp \Phi \, \mathrm{d}\sigma. \qquad (2.3)$$

Here, the vectors $\boldsymbol{T}_S$ and $\boldsymbol{T}_O$ are tangent vectors to the path of the light at the source and observer respectively. The quantity $\boldsymbol{\nabla}_\perp\Phi$ is the projection of $\boldsymbol{\nabla}\Phi$ on the plane perpendicular to the path, in a specific point.

That the difference between these two unit vectors is related to the deflection angle can be seen in figure 2.1. If one takes the z-axis orientation from source to observer, the projection of this difference on the xz-plane is the following:

$$(\boldsymbol{T}_S - \boldsymbol{T}_O)_{xz} = (\cos\alpha_1 - \cos\alpha_2)\boldsymbol{e}_z + (\sin\alpha_1 + \sin\alpha_2)\boldsymbol{e}_x$$

$$\Rightarrow (\boldsymbol{T}_S - \boldsymbol{T}_O)_{xz} \approx (\alpha_1 + \alpha_2)\boldsymbol{e}_x = \alpha_x \boldsymbol{e}_x,$$

if we consider only small deflection angles.

Of course, the right side of equation (2.3) still depends on the actual path followed by the light, since one has to consider a plane perpendicular to it in a specific point to

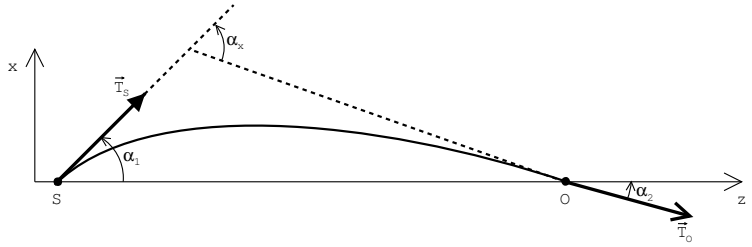


Figure 2.1: This image shows the tangent vectors $\boldsymbol{T}_S$ and $\boldsymbol{T}_O$. Projected onto the xz-plane, these unit vectors form an angle $\hat{\alpha}_x$.

be able to calculate $\boldsymbol{\nabla}_\perp \Phi$. Instead of using the precise path to calculate this integral, the small deflection angle approximation is used again, and the path is taken to be a straight line, with the impact parameter of the original path. For the deflection angle this then gives the following result:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \hat{\boldsymbol{\alpha}}(\xi_x, \xi_y) = \frac{2}{c^2} \int \left(\boldsymbol{e}_x \frac{\partial}{\partial \xi_x} + \boldsymbol{e}_y \frac{\partial}{\partial \xi_y} \right) \Phi(\xi_x, \xi_y, z) \mathrm{d}z \tag{2.4}$$

$$= \frac{2}{c^2} \boldsymbol{\nabla} \left(\int \Phi(\xi_x, \xi_y, z) \mathrm{d}z \right). \tag{2.5}$$

In this last equation, the gradient is a two-dimensional one, since the z-coordinate has been eliminated by integration. The two-dimensional vector $\boldsymbol{\xi}$ describes the impact parameter of the path.

2.2 Lens equation

In a typical gravitational lensing situation, the distances from source to lens and from lens to observer are extremely large compared to the dimensions of the deflector itself. One therefore usually employs the thin lens approximation: the mass distribution of the deflector is said to lie in in the *lens plane* and the deflection happens instantaneously in this plane. Similarly, the source is said to reside in the *source plane*. Figure 2.2 illustrates this situation.

Figure 2.3 shows the projection of this situation on the xz or yz plane, and illustrates how one comes to the fundamental equation of gravitational lensing, the *lens equation*. A light ray starts from the source S, travels in a straight line to the lens plane where its deflection is described by $\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi})$, and travels in a straight line again towards the observer O. Instead of seeing the source S in the direction described by $\boldsymbol{\beta}$, the light ray will now reach the observer from a direction $\boldsymbol{\theta}$ which consequently is where the image I of the source will be seen.

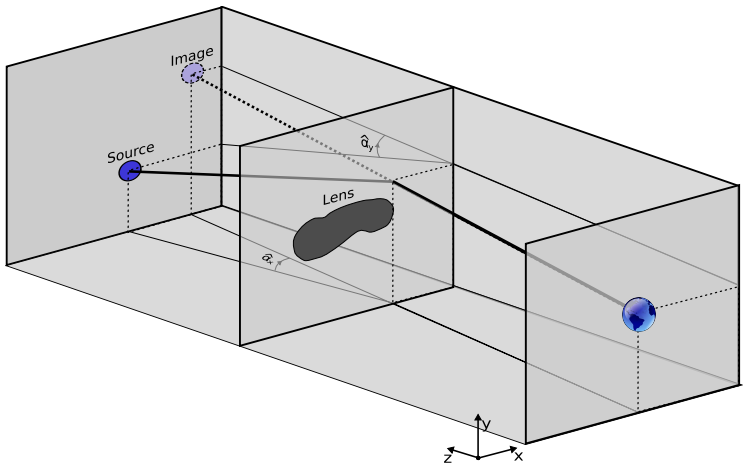


Figure 2.2: In the thin lens approximation, the source which is located in the source plane emits light rays that travel in a straight line towards the lens plane. There, a ray gets deflected instantaneously after which it continues on a straight line towards the observer again. The distances involved are very large compared to the dimensions on which the deflection takes place. This is why the lensing mass is said to be located in the lens plane.

From the figure it is clear that the following relation is valid if one considers the fact that all angles involved are small:

$$\boldsymbol{\beta}\, D_s + \hat{\boldsymbol{\alpha}}\, D_{\rm ds} = \boldsymbol{\theta}\, D_{\rm s}.$$

This can easily be rewritten to obtain the lens equation:

$$\boldsymbol{\beta}(\boldsymbol{\theta}) = \boldsymbol{\theta} - \frac{D_{\rm ds}}{D_{\rm s}}\, \hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}). \tag{2.6}$$

In this equation, the identity $D_{\rm d}\boldsymbol{\theta} = \boldsymbol{\xi}$ has been used to write the deflection angle in terms of θ rather than ξ.

The lens equation is a mapping from θ-space to β-space. The β-space clearly corresponds to the source plane, and describes what the observer would see if the gravitational lens effect could be turned off somehow. The θ-space is often called the *image plane* and describes what the observer sees because of the gravitational deflection of light. Because the mapping is from image plane to source plane, it is in general not possible to find a direct expression which gives the image positions for a specific source position. Instead, one can use a so-called ray-tracing procedure to calculate what the images look like if one knows the source shape and position and the deflection field $\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta})$. For each position in the image plane, one calculates the corresponding position in the source plane using the lens equation. If the point lies within the source, that point will be seen at the corresponding location in the image plane. For this reason the lens equation is also called the *ray-trace equation* and states that the observer will

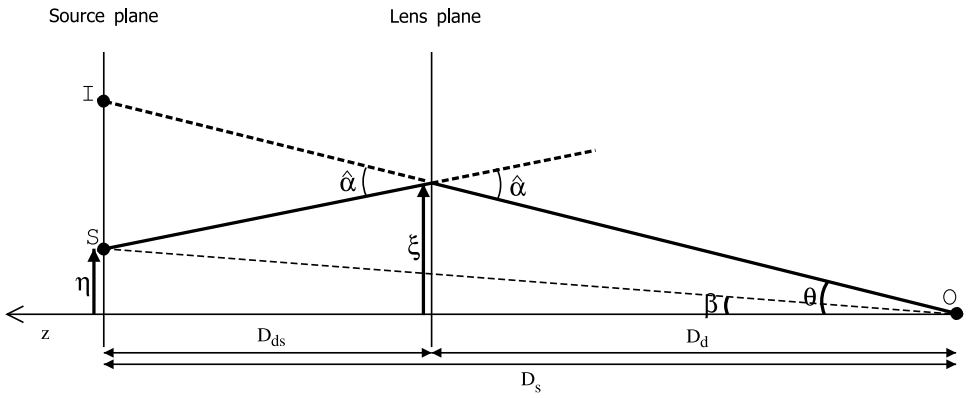


Figure 2.3: This figure illustrates the various quantities involved in deriving the lens equation. In essence, it is the same as figure 2.2, but showing only the projection on the xz or yz-plane.

see in a direction $\boldsymbol{\theta}$ what would be seen in the corresponding direction $\boldsymbol{\beta}$ if the lens effect could be disabled.

The distance measures D_d, D_ds and D_s are in principle *angular diameter distances*, as will become clear in section 2.3. For these kinds of distances, in general $D_\mathrm{s} \neq D_\mathrm{ds}+D_\mathrm{d}$.

Using the relationship between $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$, equation (2.5) can be rewritten as

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \frac{2}{c^2}\boldsymbol{\nabla}_\xi \left(\int \Phi(D_\mathrm{d}\theta_x, D_\mathrm{d}\theta_y, z)\mathrm{d}z \right) = \boldsymbol{\nabla}_\theta \hat{\psi}(\boldsymbol{\theta}),$$

in which $\hat{\psi}$ was defined as follows:

$$\hat{\psi}(\boldsymbol{\theta}) = \frac{1}{D_\mathrm{d}}\frac{2}{c^2}\int \Phi(D_\mathrm{d}\theta_x, D_\mathrm{d}\theta_y, z)\mathrm{d}z.$$

For convenience, when only a single source plane is being used, one often uses a scaled version of the deflection angle

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) := \frac{D_\mathrm{ds}}{D_\mathrm{s}}\,\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}),$$

for which:

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \boldsymbol{\nabla}_\theta \psi(\boldsymbol{\theta}). \tag{2.7}$$

The quantity $\psi(\boldsymbol{\theta})$ is usually called the *projected potential*, *effective lensing potential* or *lens potential*, and is defined as follows:

$$\psi(\boldsymbol{\theta}) := \frac{D_\mathrm{ds}}{D_\mathrm{s}}\hat{\psi}(\boldsymbol{\theta}) = \frac{2}{c^2}\frac{D_\mathrm{ds}}{D_\mathrm{s}D_\mathrm{d}}\int \Phi(D_\mathrm{d}\theta_x, D_\mathrm{d}\theta_y, z)\mathrm{d}z. \tag{2.8}$$

In this text, $\hat{\psi}(\boldsymbol{\theta})$ shall also be referred to as the lens potential and shall be used when multiple source redshifts, and hence multiple $D_\mathrm{ds}/D_\mathrm{s}$ fractions, can be present.

2.3 Lensing on a cosmological scale – time delays

To introduce the lens effect on a cosmological scale, we will start from equation (2.2). The derivation which follows is based on the information in [56] en [96].

2.3.1 Geometrical part

First, let us focus on the geometrical part of the time delay of a light ray and suppose that the large-scale geometry of the universe is well described by the Friedman-Robertson-Walker metric

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - R(t)^2\mathrm{d}\sigma^2$$

in which:

$$\mathrm{d}\sigma^2 = \frac{\mathrm{d}u^2 + u^2\mathrm{d}\theta^2 + u^2\sin^2\theta\mathrm{d}\varphi^2}{\left(1+\frac{k}{4}u^2\right)^2}.$$

Light rays travel along null geodesics, giving the equation

$$c\mathrm{d}t = R(t)\mathrm{d}\sigma. \tag{2.9}$$

Again, the thin lens approximation will be used: the light ray will travel from source S along a null geodesic until it reaches the lensing object. In the deflection point I, the light ray is deflected by the angle $\hat{\boldsymbol{\alpha}}$ after which it travels along a null geodesic towards the observer O. The spacetime coordinates (u, θ, φ) of these points are fixed. The physical distance between them varies only because of the overall expansion of the universe, encoded in $R(t)$.

Ultimately, physical paths between a fixed source and observer will require looking for paths which extremize the time delay. Because source and observer are fixed and adding a constant will not matter when looking for extrema, we shall evaluate the following expression:

$$t_{\mathrm{geom}} = \Delta t_{SI} + \Delta t_{IO} - \Delta t_{SO}.$$

From equation (2.9) one obtains:

$$\Delta\sigma = c\int\frac{1}{R(t)}\mathrm{d}t \approx c\frac{1}{R_0}\Delta t.$$

Here, we used the fact that the time differences involved are small with respect to the time over which the scale factor $R(t)$ changes, and have called R_0 the value of the scale factor at the current time. The geometrical part of the time delay then becomes (see also figure 2.4):

$$c\,t_{\mathrm{geom}} = R_0(\sigma_{ds} + \sigma_d - \sigma_s). \tag{2.10}$$

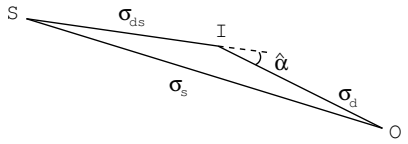


Figure 2.4: The coordinate distances and deflection angle in the cosmological setting of the gravitational lens effect. This is used to calculate the geometrical part of the time delay function.

In the following derivation it is assumed that $k = 0$, but an analogous derivation can be made for $k = \pm 1$. In the case under study, space is flat and again assuming small deflection angles, the coordinate distance becomes:

$$\begin{aligned} \sigma_s^2 &= \sigma_d^2 + \sigma_{ds}^2 - 2\sigma_d\sigma_{ds}\cos(\pi - \hat{\alpha}) = \sigma_d^2 + \sigma_{ds}^2 + 2\sigma_d\sigma_{ds}\cos\hat{\alpha} \\ &\approx \sigma_d^2 + \sigma_{ds}^2 + 2\sigma_d\sigma_{ds}(1 - \frac{1}{2}\hat{\alpha}^2) = (\sigma_d + \sigma_{ds})^2 - \sigma_d\sigma_{ds}\hat{\alpha}^2. \end{aligned}$$

Taking the square root of this expression, one obtains:

$$\begin{aligned} \sigma_s &\approx (\sigma_d + \sigma_{ds})\sqrt{1 - \frac{\sigma_d\sigma_{ds}\hat{\alpha}^2}{(\sigma_d + \sigma_{ds})^2}} \approx (\sigma_d + \sigma_{ds})\left(1 - \frac{\sigma_d\sigma_{ds}\hat{\alpha}^2}{2(\sigma_d + \sigma_{ds})^2}\right) \\ &\approx \sigma_d + \sigma_{ds} - \frac{\sigma_d\sigma_{ds}}{2\sigma_s}\hat{\alpha}^2, \end{aligned}$$

since in zeroth order in $\hat{\alpha}$ one finds that $\sigma_s \approx \sigma_d + \sigma_{ds}$.
Finally the expression (2.10) then reduces to:

$$c\, t_{\text{geom}} = R_0 \frac{\sigma_d\sigma_{ds}}{2\sigma_s}\hat{\alpha}^2. \tag{2.11}$$

To proceed it is necessary to introduce the concept of *angular diameter distance*, as illustrated in figure 2.5. Suppose that two light rays A and B are emitted by a source S, starting from a physical distance d from each other. These light rays travel towards the observer O and will arrive there with an angle θ between them. The angular diameter distance D_{s} to the source is then defined as:

$$D_{\text{s}} := \frac{d}{\theta}.$$

Call δ the coordinate distance which corresponds to d, and in contrast to d this coordinate distance is fixed in time. If R_s is the scale factor at the time the light rays are emitted, obviously

$$d = R_s\delta.$$

Since the coordinate distance δ is supposed to be extremely small, from the sine rule one finds:

$$\frac{\theta}{\delta} = \frac{1}{\sigma_s}.$$

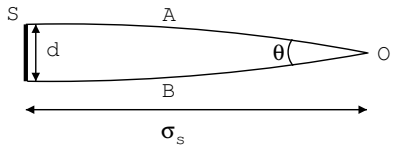


Figure 2.5: Definition of the angular diameter distance. If an observer sees an object of intrinsic size d under an angle θ, the angular diameter distance D to the object is such that $D\theta = d$.

The expression for the angular diameter distance then becomes:

$$D_{\mathrm{s}} = \frac{d}{\theta} = \frac{R_s \delta}{\theta} = \frac{R_s \theta \sigma_s}{\theta} = R_s \sigma_s.$$

Using the substitutions

$$\sigma_s = \frac{D_{\mathrm{s}}}{R_s} \qquad \sigma_{ds} = \frac{D_{\mathrm{ds}}}{R_s} \qquad \sigma_d = \frac{D_{\mathrm{d}}}{R_d}$$

equation (2.11) becomes:

$$c\, t_{\mathrm{geom}} = \frac{R_0}{R_d} \frac{D_{\mathrm{d}} D_{\mathrm{ds}}}{2 D_{\mathrm{s}}} \hat{\alpha}^2.$$

It is well known that the factor $\frac{R_0}{R_d}$ is simply $1 + z_d$, in which z_d is the redshift of the deflecting object. Again calling θ and β the directions in which the observer sees the image and the source respectively, and applying the sine rule to the triangle in figure 2.4, one finds:

$$\frac{\sin(\pi - \hat{\alpha})}{\sigma_s} = \frac{\sin(\theta - \beta)}{\sigma_{ds}},$$

which, for small angles, reduces to:

$$\hat{\alpha} \approx \frac{\sigma_s}{\sigma_{ds}} (\theta - \beta).$$

Finally, the geometrical part of the time delay becomes:

$$c\, t_{\mathrm{geom}} = (1 + z_d) \frac{D_{\mathrm{s}} D_{\mathrm{d}}}{2 D_{\mathrm{ds}}} (\boldsymbol{\theta} - \boldsymbol{\beta})^2.$$

2.3.2 Gravitational part

As was stated earlier, the gravitational interaction with the deflecting object is locally still described by the static weak field metric. From equation (2.1) the gravitational part of the time delay, when the light ray reaches the lens itself then becomes:

$$-\frac{1}{c} \int_S^O \frac{2\Phi(\boldsymbol{r})}{c^2} \mathrm{d}\sigma \approx -\frac{1}{c} D_{\mathrm{d}} \hat{\psi}(\boldsymbol{\theta}).$$

Because of the cosmological expansion, this time delay will be modified by a factor $1 + z_d$ when the observer is reached. The gravitational part of the time delay then becomes:

$$c\, t_{\text{grav}} = -(1 + z_d) D_{\text{d}} \hat{\psi}(\boldsymbol{\theta}) = -(1 + z_d) \frac{D_{\text{d}} D_{\text{s}}}{D_{\text{ds}}} \psi(\boldsymbol{\theta}).$$

2.3.3 Time delay function

Adding the expressions for the geometrical and gravitational parts of the time delay, one obtains the following *time delay function*:

$$t(\boldsymbol{\theta}) = \frac{1 + z_d}{c} \frac{D_{\text{d}} D_{\text{s}}}{D_{\text{ds}}} \left(\frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\beta})^2 - \psi(\boldsymbol{\theta}) \right). \tag{2.12}$$

For a given source position $\boldsymbol{\beta}$, the time delay (up to an arbitrary constant) for a light ray reaching the observer from direction $\boldsymbol{\theta}$ is given by $t(\boldsymbol{\theta})$. The surface defined in this way, is often referred to as the *arrival time surface*. Using the principle of Fermat, the stationary solutions give the true paths of the light rays, that is:

$$\boldsymbol{\nabla} t(\boldsymbol{\theta}) = 0 \Leftrightarrow \boldsymbol{\theta} - \boldsymbol{\beta} - \boldsymbol{\nabla} \psi = 0,$$

which is again the lens equation if one takes into account equation (2.7).

Appendix C shows how angular diameter distances can be calculated if one knows the cosmological parameters H_0, Ω_m, Ω_r and Ω_Λ, and one knows the redshift of the object.

2.4 Deflection by a mass distribution

Instead of working with the gravitational potential of the object, using which one can obtain the deflection angle from equation (2.4) or (2.5), one can also use the mass distribution of the lens.

2.4.1 Point mass deflector

The gravitational potential of a point mass of mass M is given by

$$\Phi(x, y, z) = -\frac{GM}{\sqrt{x^2 + y^2 + z^2}}.$$

Equation (2.4) then becomes:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{2}{c^2} \int \frac{GM}{(\xi_x^2 + \xi_y^2 + z^2)^{\frac{3}{2}}} (\xi_x \boldsymbol{e}_x + \xi_y \boldsymbol{e}_y) \mathrm{d}z$$

It is an easy exercise to show that this expression reduces to the value first reported by Einstein:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4GM\boldsymbol{\xi}}{c^2\xi^2}, \tag{2.13}$$

or in terms of $\boldsymbol{\theta}$:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \frac{4GM}{c^2 D_\mathrm{d}}\frac{\boldsymbol{\theta}}{\theta^2}. \tag{2.14}$$

The lens equation for a point mass lens becomes

$$\boldsymbol{\beta}(\boldsymbol{\theta}) = \boldsymbol{\theta} - \theta_E^2\frac{\boldsymbol{\theta}}{\theta^2} \quad \text{in which } \theta_E = \sqrt{\frac{D_\mathrm{ds}}{D_\mathrm{s}}\frac{4GM}{c^2 D_\mathrm{d}}}.$$

It is clear that in general, lens, source and images will lie on a straight line. Rewriting this as

$$\boldsymbol{\beta}(\boldsymbol{\theta}) = \boldsymbol{\theta}\left(1 - \frac{\theta_E^2}{\theta^2}\right),$$

it is clear that for $\theta \gg \theta_E$, $\boldsymbol{\beta} \approx \boldsymbol{\theta}$ and source and image positions nearly coincide. This is of course what one would expect: if the source and lens are badly aligned, the lens effect is negligible.

2.4.2 General mass distribution

For a general mass distribution, one can start from equation (2.4) and fill in the gravitational potential for a mass distribution $\rho(\boldsymbol{r})$:

$$\Phi(\boldsymbol{r}) = -G\int\frac{\rho(\boldsymbol{r'})}{|\boldsymbol{r}-\boldsymbol{r'}|}\mathrm{d}\boldsymbol{r'}.$$

Some straightforward algebra then gives the expression:

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

in which

$$\Sigma(\boldsymbol{\xi}) = \int\rho(\xi_x,\xi_y,z)\mathrm{d}z$$

is the *two-dimensional* or *projected mass density*. Note that this is the expression one would expect for a two-dimensional mass density built from mass elements $\Sigma(\boldsymbol{\xi})\ \mathrm{d}\boldsymbol{\xi}$, if the deflection angle for a point mass in the origin is described by equation (2.13).

In terms of $\boldsymbol{\theta}$ the expression for the deflection angle becomes:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \frac{4GD_\mathrm{d}}{c^2}\int\frac{\Sigma(\boldsymbol{\theta'})(\boldsymbol{\theta}-\boldsymbol{\theta'})}{|\boldsymbol{\theta}-\boldsymbol{\theta'}|^2}\mathrm{d}\boldsymbol{\theta'}. \tag{2.15}$$

2.4.3 Circularly symmetric mass distribution

Suppose that the projected mass density Σ only depends on the length of $\boldsymbol{\xi}$, i.e. it possesses circular symmetry. In that case, as can be expected, the deflection angle at a distance ξ is equal to the deflection angle of a point mass with mass equal to the total mass within this radius. Explicitly:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4GM(\xi)}{c^2}\frac{\boldsymbol{\xi}}{\xi^2} \quad \text{in which } M(\xi) = 2\pi\int_0^{\xi}\Sigma(\xi')\xi'\mathrm{d}\xi',$$

or again in terms of θ:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \frac{4GM(\theta)}{c^2 D_\mathrm{d}}\frac{\boldsymbol{\theta}}{\theta^2} \quad \text{in which } M(\theta) = 2\pi D_\mathrm{d}^2\int_0^{\theta}\Sigma(\theta')\theta'\mathrm{d}\theta'. \tag{2.16}$$

A derivation of this result using complex integration can be found in appendix D. The fact that for a circularly symmetric mass distribution the deflection angle depends only on the enclosed mass will be very useful when discussing degeneracies in chapter 3.

2.4.4 Einstein radius and critical mass density

If one has a circularly symmetric mass distribution centered on the origin and a source which lies directly behind it, the symmetry of the situation demands that if a strong lens effect is present, the image of the source will be a ring, usually called an *Einstein ring*. Substituting $\boldsymbol{\beta} = 0$ and the deflection for a circularly symmetric mass distribution (2.16) in the lens equation, one finds the condition

$$1 = \frac{D_\mathrm{ds}}{D_\mathrm{s}}\frac{4GM(\theta)}{c^2 D_\mathrm{d}}\frac{1}{\theta^2} \tag{2.17}$$

for the non-trivial solution. If, for a specific mass distribution with total projected mass profile $M(\theta)$, this condition has a solution for a real value of θ, a ring will be visible and the corresponding radius will be called the *Einstein radius* θ_E.

To gain some further insight into this condition, it is instructive to introduce $\bar{\Sigma}(\theta)$, the mean mass density within a radius θ:

$$\bar{\Sigma}(\theta) = \frac{M(\theta)}{\pi D_\mathrm{d}^2\theta^2}$$

The previous condition can then be rewritten as

$$\frac{\bar{\Sigma}(\theta)}{\Sigma_{cr}} = 1,$$

in which

$$\Sigma_{cr} = \frac{c^2}{4\pi G}\frac{D_\mathrm{s}}{D_\mathrm{ds}D_\mathrm{d}} \tag{2.18}$$

is called the *critical density*. This means that a source in the origin will produce an Einstein ring at a radius θ if the average density within that radius is equal to the critical density.

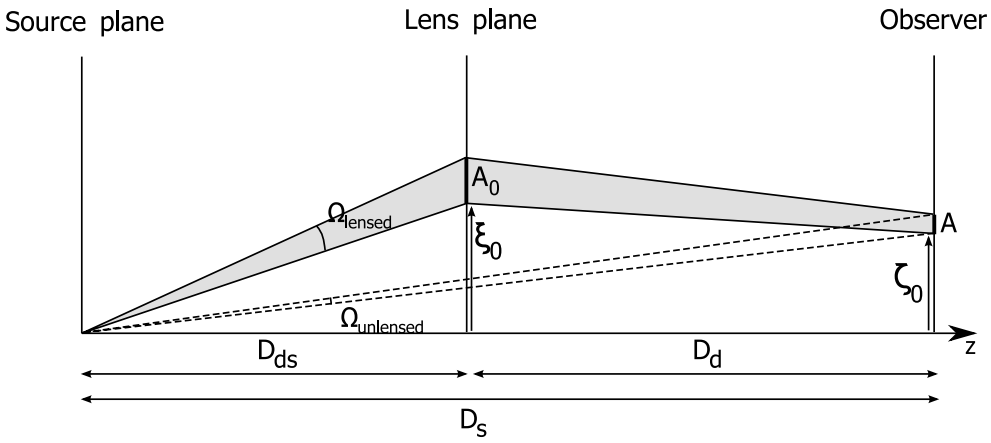


Figure 2.6: Illustration of the influence of the deflection of light on the flux received from an ideal point source. In the unlensed case, a detector of area A receives the photons emitted in the solid angle Ω_{unlensed}. Because of the lens effect however, a different amount of photons will reach the same detector, namely the ones emitted in solid angle Ω_{lensed}.

2.5 Magnification

Because the lens effect changes the size which an observer measures for an object, one can easily imagine that it can have an effect on the flux received from an astronomical source. In this section we shall investigate this, first for an ideal point source and next, more realistically, for an extended source.

2.5.1 Point source

To calculate the influence of the lens effect on the flux received from an isotropic point source, we shall first turn around the lensing situation somewhat, as is illustrated in figure 2.6. This time, the source is at the origin, and because of the gravitational deflection of light, a bundle of light rays hits an area A, for example the surface of a detector.

For an isotropic source, the flux received from the source on the detector will be proportional to the solid angle subtended by the detector, as seen from the source:

$$F_{\text{unlensed}} \propto \Omega_{\text{unlensed}} = \frac{A}{D_{\text{s}}^2}.$$

If the influence of the lens is taken into account, the photons in the solid angle Ω_{lensed} are received instead, so that the flux received from the source is:

$$F_{\text{lensed}} \propto \Omega_{\text{lensed}} = \frac{A_0}{D_{\text{ds}}^2}.$$

This means that the ratio of these fluxes, i.e. the *magnification* of the source is given by:

$$\mu = \frac{F_{\text{lensed}}}{F_{\text{unlensed}}} = \frac{A_0}{A}\frac{D_s^2}{D_{\text{ds}}^2}.$$

So to calculate the magnification, on needs to find the ratios of the areas A_0 and A. Suppose, for A_0, we take triangular area bounded by the vectors

$$\boldsymbol{\xi}_0,\ \boldsymbol{\xi}_1 = \boldsymbol{\xi}_0 + \mathrm{d}\boldsymbol{\xi}_1 = \boldsymbol{\xi}_0 + \mathrm{d}\xi_1 \boldsymbol{e}_x \text{ and } \boldsymbol{\xi}_2 = \boldsymbol{\xi}_0 + \mathrm{d}\boldsymbol{\xi}_2 = \boldsymbol{\xi}_0 + \mathrm{d}\xi_2 \boldsymbol{e}_y.$$

By inspecting figure 2.6, one easily sees that these vectors are transformed to the vectors

$$\boldsymbol{\zeta}_i = \frac{D_{\text{s}}}{D_{\text{ds}}}\boldsymbol{\xi}_i - \hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}_i) D_{\text{d}},$$

in which i is 0,1 or 2. These vectors again define a triangular area of which the area can be calculated by[1]

$$A = \frac{1}{2}(\mathrm{d}\boldsymbol{\zeta}_{1,x}\mathrm{d}\boldsymbol{\zeta}_{2,y} - \mathrm{d}\boldsymbol{\zeta}_{2,x}\mathrm{d}\boldsymbol{\zeta}_{1,y})$$

in which

$$\mathrm{d}\boldsymbol{\zeta}_i = \boldsymbol{\zeta}_i - \boldsymbol{\zeta}_0.$$

Some straightforward algebra then yields the result:

$$A = \frac{1}{2}\left[\left(\frac{D_{\text{s}}}{D_{\text{ds}}} - D_{\text{d}}\frac{\partial\hat{\alpha}_x}{\partial\xi_x}\right)\left(\frac{D_{\text{s}}}{D_{\text{ds}}} - D_{\text{d}}\frac{\partial\hat{\alpha}_y}{\partial\xi_y}\right)\mathrm{d}\xi_1\mathrm{d}\xi_2 - D_{\text{d}}^2\frac{\partial\hat{\alpha}_y}{\partial\xi_x}\frac{\partial\hat{\alpha}_x}{\partial\xi_y}\mathrm{d}\xi_1\mathrm{d}\xi_2\right]$$

Noting that $D_{\text{d}}\boldsymbol{\theta} = \boldsymbol{\xi}$ as before one finds

$$\mu^{-1} = \frac{A}{A_0}\frac{D_{\text{ds}}^2}{D_s^2} = \left(1 - \frac{D_{\text{ds}}}{D_{\text{s}}}\frac{\partial\hat{\alpha}_x}{\partial\theta_x}\right)\left(1 - \frac{D_{\text{ds}}}{D_{\text{s}}}\frac{\partial\hat{\alpha}_y}{\partial\theta_y}\right) - \frac{D_{\text{ds}}^2}{D_{\text{s}}^2}\frac{\partial\hat{\alpha}_y}{\partial\theta_x}\frac{\partial\hat{\alpha}_x}{\partial\theta_y} = \left|\frac{\partial\beta_i}{\partial\theta_j}\right|. \quad (2.19)$$

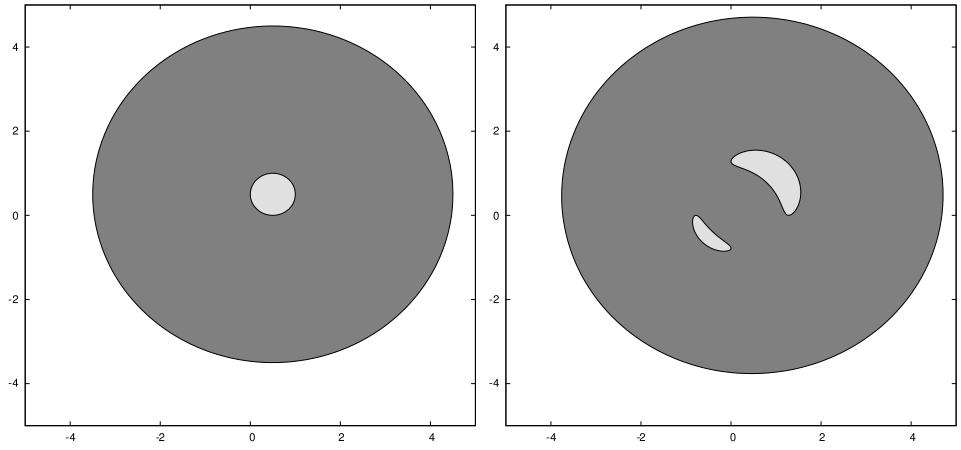
The inverse magnification of a point source is thus given by the Jacobian determinant of the lens equation.

2.5.2 Extended source

To simulate a small source of constant surface brightness, suppose that this source consists of a large number of point sources, all with the same intrinsic luminosity and distributed uniformly. If the source is small enough so that μ does not change over its area, each of the point sources will cause a flux that is larger by this factor.

To the observer however, the point sources are not perceived as lying at their original spacing since the observer sees an image of the source, not the source itself. The surface area of the image that the observer sees is larger by a factor that is precisely the inverse of the Jacobian determinant of the lens equation. This means that the

[1] Compare this to the z-component of a vector product.



(a) Source plane (b) Image plane

Figure 2.7: Lens effect of a point mass lens in the origin of the coordinate system, on a source which has different dimensions in different wavebands. The axes on this figure are in units of the Einstein radius of the point mass lens. Because one region is large compared to the Einstein radius, its magnification will be quite small. For a small region on the other hand, the magnification factor can be considerably larger. Combining these effects with limited resolving power, this means that the color of the source can still be different from the color of the image.

point sources in the image are brighter by a factor μ, but the image itself is also larger than the source by a factor μ. Therefore, the surface brightness of the image is the same as the surface brightness of the source. This is true in general: *lensing conserves surface brightness*. The fact that the flux of an object changes is purely geometrical as this can be attributed to the change in size.

Since it is just the image size that is changed, not the surface brightness, it is actually the resolution of the detector that will determine if the magnification of an image is well described by the magnification of a point source or not. If an image is much smaller than a CCD pixel for example, the pixel itself will receive the modified total flux of the image which is well described by μ. If the image is much larger than the typical size of a pixel, i.e. it is well resolved, each pixel will estimate the surface brightness at that point. Of course the flux of the image as a whole will still be changed by a factor that is approximately μ, but again this is a purely geometrical effect.

It is also interesting to see how color changes are still possible because of this. Figure 2.7 illustrates the situation, using the effect of a point mass lens with a certain Einstein radius θ_E. Suppose, in one waveband, that the source is quite large, for instance a disc of uniform surface brightness with a radius that is much larger than θ_E. The lensed version of this source will look very much like the source itself: the lensed radius will be almost equal to the real radius of the source, as it is much larger

that θ_E, and the surface brightness inside the lensed disc is unchanged. Suppose that in another waveband, this same source is small compared to θ_E and located close to the center of the lens. The point mass lens will generate two distinct images, each with a modified flux. The total flux in this waveband will also differ from the original one. If the resolution of our detector is not good enough to resolve the image, in one waveband the flux will not be magnified, but in another waveband it will. This of course leads to a modified color.

2.6 Convergence and shear

In equation (2.8) the projected potential was already introduced:

$$\psi(\boldsymbol{\theta}) = \frac{2}{c^2}\frac{D_{\rm ds}}{D_{\rm s}D_{\rm d}}\int \Phi(D_{\rm d}\theta_x, D_{\rm d}\theta_y, z){\rm d}z,$$

for which one finds that:

$$\boldsymbol{\nabla}\psi(\boldsymbol{\theta}) = \boldsymbol{\alpha}(\boldsymbol{\theta})$$

Calculating the two dimensional Laplacian of the projected potential one finds:

$$\begin{aligned}\nabla^2_\theta\psi &= \left(\frac{\partial^2}{\partial\theta_x^2}+\frac{\partial^2}{\partial\theta_y^2}\right)\psi = \frac{D_{\rm ds}}{D_{\rm s}D_{\rm d}}\frac{2}{c^2}\int D_{\rm d}^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\Phi{\rm d}z \\ &= \frac{D_{\rm ds}}{D_{\rm s}}\frac{2}{c}\int\left(\nabla^2\Phi - \frac{\partial^2\Phi}{\partial z^2}\right){\rm d}z = \frac{D_{\rm ds}}{D_{\rm s}}\frac{2}{c}\int 4\pi G\rho{\rm d}z - 0,\end{aligned}$$

where in the last step the fact that $\Phi \to 0$ at infinity was used. It is then immediately clear that

$$\frac{1}{2}\nabla^2\psi(\boldsymbol{\theta}) = \frac{\Sigma(\boldsymbol{\theta})}{\Sigma_{cr}} := \kappa(\boldsymbol{\theta}).$$

The quantity κ is called the *convergence* and is a dimensionless measure of the mass density. Since $\boldsymbol{\alpha}$ is the gradient of a scalar function, it is also clear that:

$$\frac{\partial\alpha_x}{\partial\theta_y} = \frac{\partial^2\psi}{\partial\theta_x\partial\theta_y} = \frac{\partial\alpha_y}{\partial\theta_x}.$$

One also defines the *shear* components γ_1 and γ_2 as:

$$\gamma_1(\boldsymbol{\theta}) = \frac{1}{2}\left(\frac{\partial^2\psi}{\partial\theta_x^2} - \frac{\partial^2\psi}{\partial\theta_y^2}\right) = \frac{1}{2}\left(\frac{\partial\alpha_x}{\partial\theta_x} - \frac{\partial\alpha_y}{\partial\theta_y}\right)$$

$$\gamma_2(\boldsymbol{\theta}) = \frac{\partial^2\psi}{\partial\theta_x\partial\theta_y} = \frac{\partial\alpha_x}{\partial\theta_y} = \frac{\partial\alpha_y}{\partial\theta_x}.$$

Consider the Jacobian matrix A of the lens equation, and call its elements A_{ij}:

$$A = \left(\frac{\partial\beta_i}{\partial\theta_j}\right) = \begin{pmatrix} 1-\frac{\partial\alpha_x}{\partial\theta_x} & -\frac{\partial\alpha_x}{\partial\theta_y} \\ -\frac{\partial\alpha_y}{\partial\theta_x} & 1-\frac{\partial\alpha_y}{\partial\theta_y}\end{pmatrix} = \begin{pmatrix} 1-\frac{\partial^2\psi}{\partial\theta_x^2} & -\frac{\partial^2\psi}{\partial\theta_x\partial\theta_y} \\ -\frac{\partial^2\psi}{\partial\theta_x\partial\theta_y} & 1-\frac{\partial^2\psi}{\partial\theta_y^2}\end{pmatrix},$$

so that, for the components of this matrix, one can write:

$$A_{ij} = \delta_{ij} - \frac{\partial^2 \psi}{\partial\theta_i \partial\theta_j}. \tag{2.20}$$

With the definitions of shear and convergence above, this can be rewritten as:

$$A = (1-\kappa)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \gamma_1 & \gamma_2 \\ \gamma_2 & -\gamma_1 \end{pmatrix}.$$

The convergence can therefore be considered responsible for a uniform scaling of an image, the shear will cause it to be deformed along a specific axis.

The eigenvectors of the traceless matrix are also eigenvectors of the identity matrix, and calculating these eigenvectors (not normed) one finds

$$\boldsymbol{V}_1 = (\gamma - \gamma_1, \gamma_2)$$

$$\boldsymbol{V}_2 = (\gamma_2, \gamma_1 - \gamma),$$

where

$$\gamma = \sqrt{\gamma_1^2 + \gamma_2^2}.$$

These vectors are clearly orthogonal, as could be expected since they are the eigenvectors of a symmetric matrix. The corresponding eigenvalues are

$$\lambda_1 = 1 - \kappa - \gamma$$

$$\lambda_2 = 1 - \kappa + \gamma$$

Calling ϕ the angle that the axis of $\boldsymbol{V}_1$ makes with the x-axis, the matrix A can be written as:

$$A = (1-\kappa)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \gamma\begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix}.$$

2.7 Critical lines and caustics

The inverse magnification at a specific point in the image plane is given by the Jacobian determinant of the lens equation. Setting $\mu^{-1} = 0$ defines lines in the image plane, called the *critical lines*. By using the lens equation to map these lines to the source plane, one obtains a different set of lines called the *caustics*.

If a point source is located on a critical line, in principle it will have an infinite brightness. Of course, no real point sources exist, and the magnification will be caused by the the change in geometrical size of the observed image. In any case, critical lines mark the regions of high magnification, and therefore in general, of large deformations.

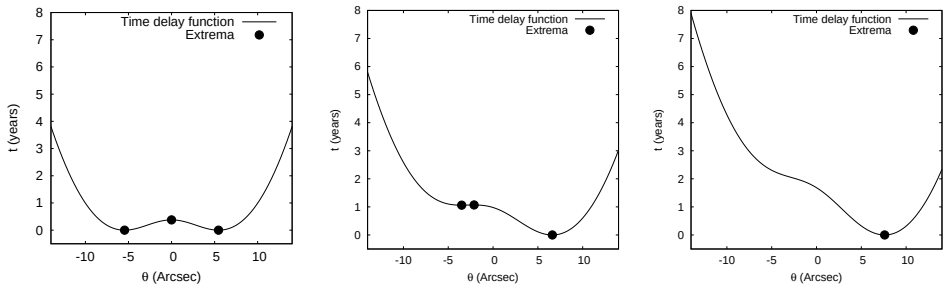


Figure 2.8: These figure show the time delay function for several positions of the source, for a circularly symmetric mass distribution. Left panel: the source is placed behind the symmetry center of the lens. Center panel: there is a distance of 1 arcsec between the source and the center of the lens. Right panel: source and lens are offset by 2 arcsec.

Since μ^{-1} becomes zero on a critical line, these lines also mark regions where the sign of the magnification changes. A negative sign simply indicates a mirrored image, and therefore the critical lines mark the regions of positive and negative *parity*.

To gain some further understanding it is instructive to take a closer look at the time delay function. If one calculates the Hessian matrix of the function (2.12), one finds:

$$\frac{\partial^2 t}{\partial\theta_i \partial\theta_j} = \frac{1+z_d}{c}\frac{D_\mathrm{d} D_\mathrm{s}}{D_\mathrm{ds}}\left(\delta_{ij} - \frac{\partial^2\psi}{\partial\theta_i\partial\theta_j}\right) = \frac{1+z_d}{c}\frac{D_\mathrm{d} D_\mathrm{s}}{D_\mathrm{ds}} A_{ij},$$

in which A_{ij} are again the elements of the Jacobian matrix of the lens equation. The curvature of the time delay surface is therefore related to the magnification: if the Hessian matrix has an eigenvalue which is zero, so will A_{ij} and the inverse magnification, being the product of the eigenvalues of A, will be zero. So, in general, a smaller curvature indicates a smaller inverse magnification, or correspondingly a higher magnification.

Now, for the time delay function which corresponds to a specific source position $\boldsymbol{\beta}$, the images will lie at the positions where the time delay function is extremal. The brightness of the images of a point source can then be derived by evaluating the Hessian matrix of the time delay function at these positions, a smaller curvature indicating a larger magnification. Figure 2.8 illustrates this for a source at different positions. For a circularly symmetric mass distribution[2], as was used in creating the panels, the images will lie along the line which connects the center of the mass distribution with the source, so we can limit ourselves to a one-dimensional cut of the time delay function.

The left panel shows the situation in which the source is directly behind the mass distribution. As shown in the image, there are three extrema and therefore three images. In reality, by symmetry the outer two extrema will form a ring, the Einstein

[2]The mass distribution used corresponds to a projected Plummer sphere. See section 2.8.2 for details.

ring. The center panel shows what happens if the source moves outwards slightly. The offset between the center of the lens and the source is 1 arcsec. It can be clearly seen that the extremum on the right moves outward as well, the other two extrema move towards each other. The right panel shows what happens if the source moves outwards even further: the offset is now 2 arcsec. In this case, the time delay function will only have one extremum anymore. This is to be expected: as the alignment between observer, lens and source deteriorates, the gravitational lens effect will diminish in strength and only one image of the source will be visible.

In between the situation in the center and right panels, it is clear that the two images will merge and disappear. At that point, the maximum of the center image and the minimum of the left image will merge as well and the curvature of the time delay function will become zero. Of course this means that the inverse magnification becomes zero at that point as well, or equivalently that *image pairs get created or destroyed on critical lines*. When images are merging or nearly merging, the magnification will be very large.

These one-dimensional versions of the time delay surface also suggest that in general there will be an odd number of images: one that converges to the source position as the source and lens become less well aligned, and possibly other image pairs. This is in fact a general result and does not only apply to circularly symmetric lenses [16]. Also, each time an image pair gets created on a critical line, the source crosses a caustic. This means that the position of the source with respect to the caustics, determines the number of images.

For a circularly symmetric lens, equation (2.16) shows that a vector $\boldsymbol{\theta}$ in the image plane corresponds to a vector $\boldsymbol{\beta}$ in the source plane which is oriented along the same direction. This means that angle under which an image is seen from the lens center, is the same angle as the one under which the source is seen. Figure 2.9 then shows that the lens equation for such a lens transforms a surface element $\beta\Delta\phi\Delta\beta$ to a surface element $\theta\Delta\phi\Delta\theta$. Or, stated differently, this means that the inverse magnification, being the ratio of these quantities, simplifies to:

$$\mu^{-1} = \frac{\beta}{\theta}\frac{\partial\beta}{\partial\theta},$$

in which

$$\beta(\theta) = \theta - \frac{4GM(\theta)}{c^2 D_{\mathrm{d}}\theta}$$

is a one-dimensional version of the lens equation for a circularly symmetric lens.

The equation of the inverse magnification above shows that in general there will be two critical lines: one for which $\frac{\beta}{\theta} = 0$ and one for which $\frac{\partial\beta}{\partial\theta} = 0$. The first one obviously corresponds to the caustic point $\beta = 0$. When the source is close to this point, the images will be stretched tangentially along the corresponding critical line. One therefore speaks of the tangential caustic, and tangential critical line. When the source is close to the other caustic, it will have images close to another critical line. Since these images are then elongated in the radial direction, one speaks of the radial caustic and radial critical lines. Figure 2.10 illustrates this.

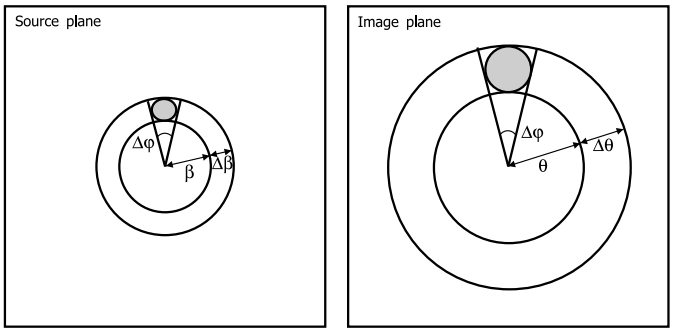


Figure 2.9: For a circularly symmetric lens, the angle under which the source is seen, is the same as the angle under which the images are seen. This causes a surface element $\beta\Delta\phi\Delta\beta$ to be transformed into a surface element $\theta\Delta\phi\Delta\theta$, leading to a particularly simple expression for the inverse magnification (see text).

2.8 Specific lenses

The concepts derived in the previous sections will be illustrated for a number of specific mass distributions. First, the lens effect for a family of mass distributions derived from the so-called singular isothermal sphere is shown. Next, because it is used frequently in the inversion procedure, the lens effect of a Plummer mass distribution is shown.

2.8.1 Isothermal ellipsoids

As a model for a galaxy for example, consider it as consisting of a large number of identical components of mass m. Furthermore, assume that such particles behave as an ideal gas, trapped in the combined spherically symmetric potential. Assume that the temperature is T everywhere in this ideal gas, so that the velocity dispersion σ_v of the particles is given by

$$m\sigma_v^2 = kT.$$

One particularly easy solution to the equation of hydrostatic equilibrium is given by the density of the so-called *singular isothermal sphere (SIS)*:

$$\rho(r) = \frac{\sigma_v^2}{2\pi G}\frac{1}{r^2}.$$

An interesting feature of such a mass distribution is that the speed v of a test particle on a circular orbit in the gravitational potential of such a mass distribution, is independent of the radius of the orbit:

$$v^2 = 2\sigma_v^2.$$

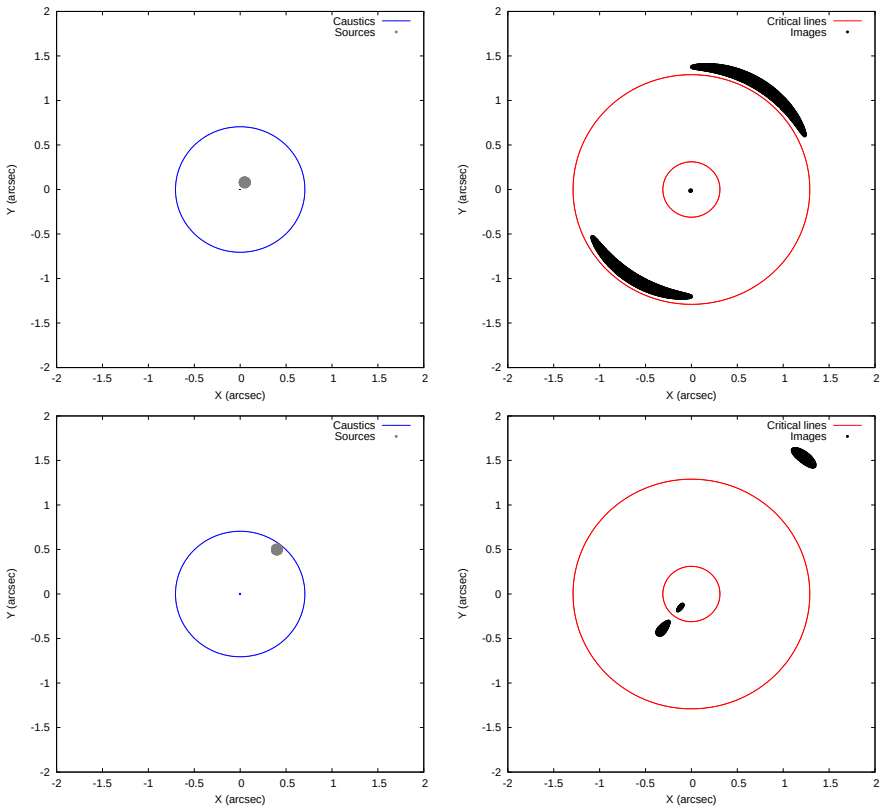


Figure 2.10: Left panel: a circular source close to the tangential caustic (top) and radial caustic (bottom). Right panel: corresponding images and critical lines, showing how images close to the tangential critical line are stretched tangentially and images close to the radial critical line are stretched radially.

This model can therefore predict flat rotation curves.

Starting from this mass density, one can easily calculate $\Sigma(\theta)$ and $M(\theta)$:

$$\Sigma(\theta) = \frac{\sigma_v^2}{2GD_\mathrm{d}\theta}, \quad M(\theta) = \frac{\pi D_\mathrm{d}\sigma_v^2\theta}{G}.$$

Equation (2.17) has the solution

$$\theta_E = \frac{D_\mathrm{ds}}{D_\mathrm{s}}\frac{4\pi\sigma_v^2}{c^2},$$

using which the expression of the deflection angle (2.16) becomes:

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \theta_E \boldsymbol{e}_\theta.$$

Here, $\boldsymbol{e}_\theta$ is a unit vector parallel to $\boldsymbol{\theta}$.

Some modifications of this mass distribution exist. The first one is called the *non-singular isothermal sphere (NSIS)*, in which the central singularity has been removed through the introduction of a core radius θ_c. The projected mass density and total mass profile then become:

$$\Sigma(\theta) = \frac{\sigma_v^2}{2GD_\mathrm{d}\sqrt{\theta^2+\theta_c^2}}, \quad M(\theta) = \frac{\pi\sigma_v^2 D_\mathrm{d}}{G}\left(\sqrt{\theta^2+\theta_c^2} - \theta_c\right).$$

One can also introduce an ellipticity f into the projected mass distribution. A first version, in which the central singularity is still present is called the *singular isothermal ellipse (SIE)*, which has the following mass distribution:

$$\Sigma(\boldsymbol{\theta}) = \frac{\sigma_v^2\sqrt{f}}{2GD_\mathrm{d}\sqrt{\theta_x^2 + f^2\theta_y^2}}.$$

It can be shown that this lens causes the following deflection angle[63]:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \frac{4\pi\sigma_v^2}{c^2}\frac{\sqrt{f}}{\sqrt{1-f^2}}\left[\mathrm{asinh}\left(\frac{\sqrt{1-f^2}}{f}\frac{\theta_x}{|\boldsymbol{\theta}|}\right)\boldsymbol{e}_x + \mathrm{asin}\left(\sqrt{1-f^2}\frac{\theta_y}{|\boldsymbol{\theta}|}\right)\boldsymbol{e}_y\right].$$

A final variation is the one in which both a core radius and ellipticity are introduced. The projected mass density of such a *non-singular isothermal ellipse (NSIE)* is then

$$\Sigma(\boldsymbol{\theta}) = \frac{\sigma_v^2\sqrt{f}}{2GD_\mathrm{d}\sqrt{\theta_x^2 + f^2\theta_y^2 + \theta_c^2}}.$$

The components of the deflection angle in this case are (see e.g. [73]):

$$\hat{\alpha}_x(\boldsymbol{\theta}) = \frac{4\pi\sigma_v^2}{c^2}\frac{\sqrt{f}}{\sqrt{1-f^2}}\mathrm{atanh}\left(\frac{\theta_x\sqrt{1-f^2}}{\sqrt{\theta_x^2 + f^2\theta_y^2 + \theta_c^2} + f\theta_c}\right)$$

$$\hat{\alpha}_y(\boldsymbol{\theta}) = \frac{4\pi\sigma_v^2}{c^2}\frac{\sqrt{f}}{\sqrt{1-f^2}}\text{atan}\left(\frac{f\theta_y\sqrt{1-f^2}}{f\sqrt{\theta_x^2+f^2\theta_y^2+\theta_c^2}+\theta_c}\right).$$

Figure 2.11 shows the gravitational lens effect of these lenses. In each case, the image plane and source plane are drawn on top of each other, so that both the source and images can be seen. Of course, in a real lensing situation only the images would be visible. The mass distributions of these lenses are not shown in the figure, only their effect is visible. These pictures were made using the ray tracing procedure explained before: for each point in the (gridded) image plane, the corresponding point in the source plane is calculated using the lens equation. If that point lies within the circular source, the pixel in the image plane is painted black, otherwise it is left blank. Note that this also means that if the images are projected back onto the source plane, they should form a consistent source shape since all of these images are derived from a single source.

2.8.2 Plummer mass distribution

The gravitational potential of the Plummer sphere [84] is:

$$\Phi(r) = -\frac{GM}{\sqrt{r^2+a_P^2}},$$

in which a_P is a characteristic width of the mass distribution. The Poisson equation $\nabla^2\Phi = 4\pi\rho$ can be used to easily show that this potential corresponds to the mass distribution

$$\rho(r) = \frac{3M}{4\pi}\frac{a_P^2}{(r^2+a_P^2)^{\frac{5}{2}}}.$$

Integration along the z-axis then yields the circularly symmetric projected mass density Σ:

$$\Sigma(\theta) = \frac{M}{\pi D_\mathrm{d}^2}\frac{\theta_P^2}{(\theta^2+\theta_P^2)^2},$$

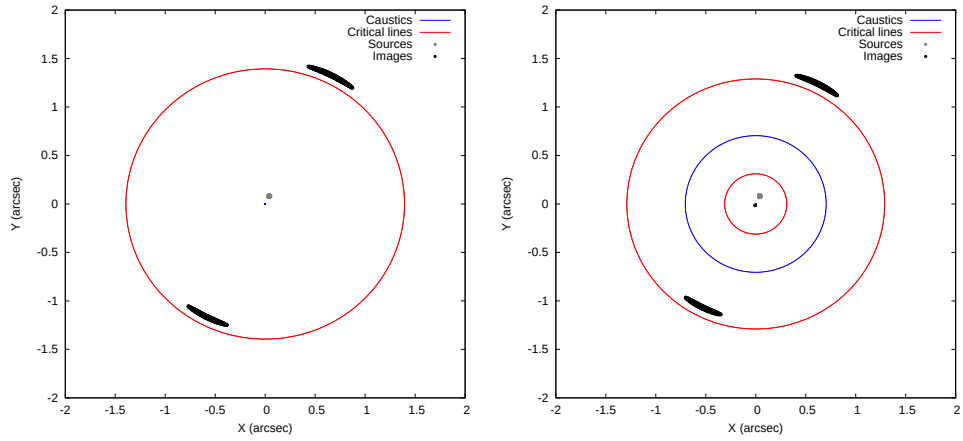
in which θ_P is a characteristic angular width, related to a_P through $D_\mathrm{d}\theta_P = a_P$. The total mass profile $M(\theta)$ then becomes

$$M(\theta) = M\frac{\theta^2}{\theta^2+\theta_P^2},$$

from which one sees that the parameter M is the total mass contained in the mass distribution.

By equation (2.16), the scaled deflection angle then becomes

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{D_\mathrm{ds}}{D_\mathrm{s}}\frac{4GM}{c^2 D_\mathrm{d}}\frac{\boldsymbol{\theta}}{\theta^2+\theta_P^2},$$



(a) SIS (b) NSIS

(c) SIE (d) NSIE

Figure 2.11: Images of a circular source in the case of a SIS, NSIS, SIE and NSIE lens. The critical lines and caustics are also shown in each case. For these simulated situations, the lens is at redshift $z_d = 0.25$, the source is at redshift $z_s = 1.5$ and the angular diameter distances were calculated in a flat cosmological model with $H_0 = 71$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.27$ and $\Omega_\Lambda = 0.73$. The lens in each case has a velocity dispersion of $\sigma_v = 250$ km s^{-1} and where appropriate a core radius $\theta_c = 0.1$ arcsec and ellipticity $f = 0.8$ were used.

which can be seen to originate from the projected potential

$$\psi(\theta) = \frac{D_{\rm ds}}{D_{\rm s}} \frac{2GM}{c^2 D_{\rm d}} \ln(\theta^2 + \theta_P^2).$$

—3—

Lensing degeneracies

Before talking about lens inversion, it is important to gain some further insight into what actually can be constrained by the images in a lensing system. This brings us to the important topic of *lensing degeneracies*. This chapter will also serve to illustrate further the theory of the lens effect.

3.1 Introduction

Since, in a strong lensing situation, the precise positions and deformations of the images of a source depend on the exact shape of the mass distribution, gravitational lensing holds the promise of constraining said mass distribution. In principle, this strong lens inversion sounds fairly straightforward: use a particular model and optimize its parameters so that it reproduces the observations as good as possible. In practice however, one is hindered by gravitational lensing degeneracies which allow a wide range of mass distributions to produce the exact same image configuration. On the level of parametric lens inversion results, this can manifest itself as parameter degeneracies (e.g. [109]), but at their core the degeneracies are of course present at the level of the mass distribution. An obvious example is that one cannot constrain the three dimensional mass distribution, as lensing is described by the two dimensional, projected density. But even when only considering this projected density, many degeneracies are present.

Several classes of degeneracies were first identified in [37] and were later reinterpreted in [89] in terms of changes in the arrival-time surface. Although we shall be focusing on strong lensing, degeneracies have also been studied in weak lensing systems (e.g. [97]) and even in the context of microlensing [81].

As with many inverse problems, lens inversion is often said to be *ill-posed*, indicating that there is no real unique solution to the problem. Of course, this is what lensing degeneracies are all about: given that at least one solution can be found for the inverse problem, is it possible to modify the corresponding mass distribution and still be compatible with the observational constraints?

3.2 The monopole degeneracy

If the constraints provided by the images in a strong lensing system seem to suggest a particular feature in the mass distribution, one may wonder if this is in fact a real feature of the mass map. To verify this, a question one can ask is the following: can this feature be removed from the reconstruction while still obtaining a good inversion, given the available constraints? Below, we shall describe how the monopole degeneracy can help to answer this question. The content of this section is based on [70].

For a circularly symmetric mass distribution $\Sigma(\theta)$, the expression for the deflection angle only depends on the total enclosed projected mass, as can be seen in equation (2.16), repeated here for convenience:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \frac{4GM(\theta)}{c^2 D_{\mathrm{d}}} \frac{\boldsymbol{\theta}}{\theta^2}.$$

From the same equation, it is clear that a circularly symmetric mass distribution of which the total mass is zero beyond a specific radius, does not produce a gravitational lens effect outside said radius. If such a mass distribution is added to an existing one, the original lens equation will be modified only inside the circular region in which it has non-zero mass.

Consider a lens mass map $M(x)$ specified by $M_A(x)$ in $[0, m]$, by $M_B(x)$ in $[m, 1]$ and which is zero beyond the unit radius:

$$\begin{aligned} M_A(x) &= -\frac{1}{4m^2}x^4 + \frac{1}{2}x^2 \\ M_B(x) &= \frac{m^2}{4(m-1)^3}\left[-2x^3 + 3(m+1)x^2 - 6mx + 3m - 1\right]. \end{aligned}$$

The shape of such a function is shown in the left panel of figure 3.1 for two different values of m, which specifies the position of the maximum. The right panel of the same figure shows the associated density profiles, which are composed of two parts as well:

$$\begin{aligned} \Sigma_A(x) &= \frac{1}{x}\frac{dM_A}{dx} = -\frac{1}{m^2}x^2 + 1 \\ \Sigma_B(x) &= \frac{1}{x}\frac{dM_B}{dx} = \frac{m^2}{4(m-1)^3}\left[-6x + 6(m+1) - \frac{6m}{x}\right]. \end{aligned}$$

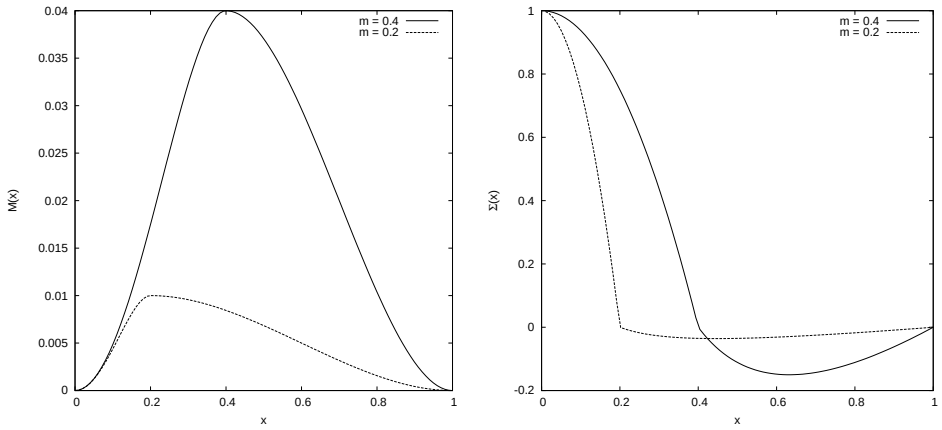


Figure 3.1: Left panel: shape of the total mass map of the circularly symmetric basis functions used to construct degenerate solutions (see text). The value of m determines the position of the maximum. Right panel: the total mass profiles shown in the left panel give rise to these density profiles. As the value of m becomes smaller, the amplitude of the negative density part decreases.

Clearly, the smaller the value of m, the flatter the density profile becomes after this point. Using such a profile, it is possible to introduce or erase a peak in an existing mass map without changing much to the rest of the distribution and while conserving the total mass.

A particular example of such a mass distribution, which we shall call a *monopole*[1], can be seen in figure 3.2. From a circular region, some mass is "borrowed" to build a central density peak. Now consider the mass distribution of the left panel of figure 3.3, which causes five images of the same circular source. If one adds the monopole from figure 3.2 to this mass distribution, placing it in the circular region indicated in the figure, it will only affect the lens equation within that circular region. The right panel of figure 3.3 shows the resulting mass distribution, which clearly differs much from the original one. One can also see in the bottom half of this panel that the circular source still produces the exact same images as before. Note that in principle it is possible that extra images will be predicted inside the circular region, since the lens equation was modified there. In this particular example, this was not the case however.

While this example may seem somewhat artificial, this type of degeneracy in fact allows one to redistribute the mass in between the images in a wide variety of ways. One only needs to use a large number of these monopoles, making sure that none of them overlap with the input images. For any set of weights of these basis functions, the images will be projected onto the same sources, since the lens equation was not modified at the location of the images. One only needs to make sure that no extra

[1]In a multipole expansion, only the monopole term is needed to describe this mass distribution.

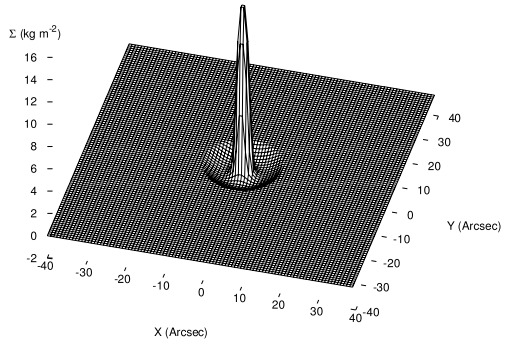


Figure 3.2: Example of a "monopole" mass distribution: a circularly symmetric mass distribution with a total mass of zero beyond a specific radius. Such a mass map will not cause any lens effect outside the outer radius.

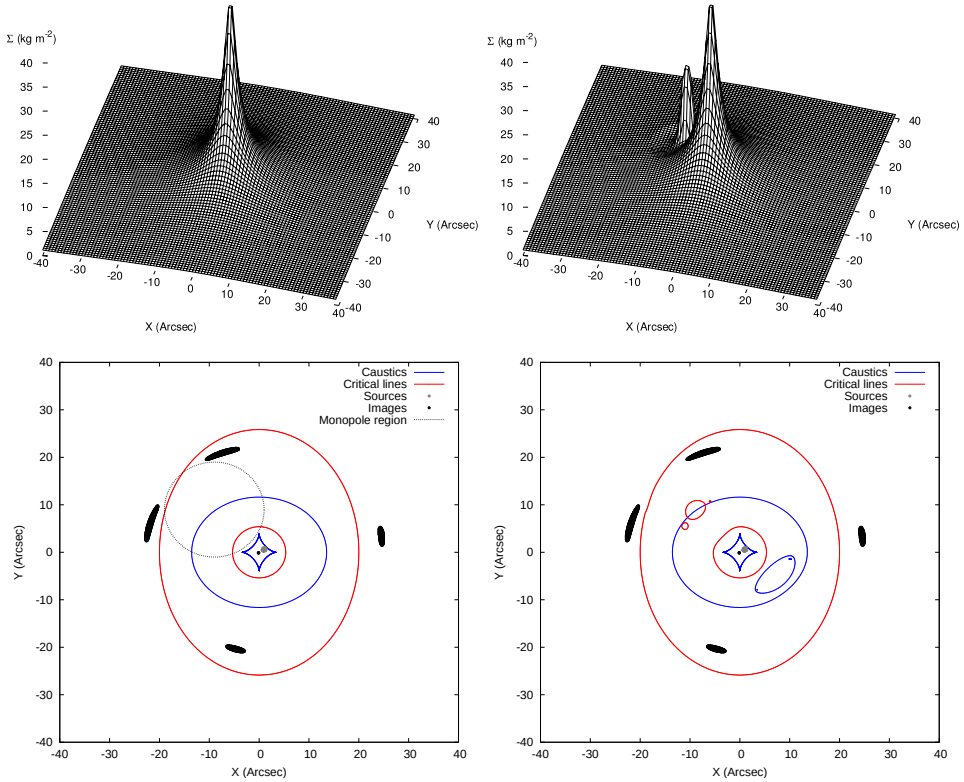


Figure 3.3: Left panel: elliptical mass distribution (top) transforms a circular source into five images (bottom). Right panel: after adding the "monopole" (see text), the resulting mass distribution (top) still transforms the source into the same five images (bottom).

images are predicted, but otherwise it is clear that a large number of mass maps will produce the same image configuration. We shall come back to this in section 5.1.4 where this type of degeneracy is applied to a particular lens inversion.

Since the deflection field is not changed at the location of the images, it is clear that the magnification of the images is not changed. Another way to see this is that the image sizes and source sizes are not affected by this degeneracy, so that the magnification is not modified either. Each individual monopole is circularly symmetric and does not produce a deflection field outside a specific radius. From $\boldsymbol{\nabla}\psi = \boldsymbol{\alpha} = 0$ it is then clear that the lens potential outside that radius is a constant. Adding a monopole to the mass distribution then has the effect of adding a constant to the time delay function at the location of all the images, but since only differences in time delays can be observed, this does not matter and the predicted time delays between the images stay the same.

3.3 The mass sheet degeneracy

The so-called *mass sheet or steepness degeneracy* (e.g. [31] or [92]) is undoubtedly one of the most famous degeneracies in lensing. In this section, the basic mechanism will be discussed; some generalizations of this degeneracy will be discussed in the next section.

Let us consider a strong lensing system with images coming from a single source, for example the strong lensing system in the left panel of figure 3.3. A uniform sheet of mass with density Σ_s produces a deflection described by

$$\hat{\boldsymbol{\alpha}}_\mathrm{s}(\boldsymbol{\theta}) = \frac{D_\mathrm{s}}{D_\mathrm{ds}}\frac{\Sigma_\mathrm{s}}{\Sigma_\mathrm{cr}}\boldsymbol{\theta},$$

in which the critical mass density Σ_cr for the current geometry is defined by equation (2.18). Note that Σ_cr depends on the redshift of the source via the angular diameter distances D_s and D_ds. Let $\Sigma_0(\boldsymbol{\theta})$ be a mass distribution that is compatible with the observed images. This means that the corresponding lens equation

$$\boldsymbol{\beta}_0(\boldsymbol{\theta}) = \boldsymbol{\theta} - \frac{D_\mathrm{ds}}{D_\mathrm{s}}\hat{\boldsymbol{\alpha}}_0(\boldsymbol{\theta})$$

projects the images onto the source plane in such a way that they overlap exactly. Without further constraints, this immediately yields an infinite number of alternative solutions. Indeed, if the mass distribution is replaced by

$$\Sigma_1(\boldsymbol{\theta}) = \lambda\Sigma_0(\boldsymbol{\theta}) + (1-\lambda)\Sigma_\mathrm{cr}, \tag{3.1}$$

the new lens equation becomes

$$\boldsymbol{\beta}_1(\boldsymbol{\theta}) = \boldsymbol{\theta} - \lambda\frac{D_\mathrm{ds}}{D_\mathrm{s}}\hat{\boldsymbol{\alpha}}_0(\boldsymbol{\theta}) - (1-\lambda)\frac{D_\mathrm{ds}}{D_\mathrm{s}}\hat{\boldsymbol{\alpha}}_\mathrm{s}(\boldsymbol{\theta}) = \lambda\boldsymbol{\beta}_0(\boldsymbol{\theta}). \tag{3.2}$$

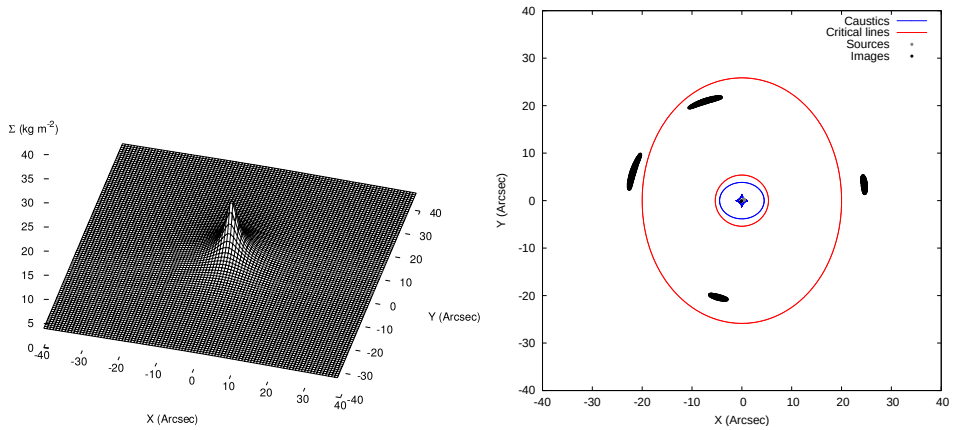


Figure 3.4: Left panel: the mass distribution in the left panel of figure 3.3 is rescaled by a factor λ, and a mass sheet of density $(1-\lambda)\Sigma_{\rm cr}$ is added. Right panel: as shown in equation (3.2), this causes a scaled version (scaled by factor λ) of the source plane to correspond to the same image plane.

The transformation (3.1) describes the so-called *mass sheet degeneracy* and simply rescales the source plane by the factor λ, as shown by equation (3.2), producing an equally acceptable source reconstruction. The dimensions of the source are different, but since these are not observable the new mass distribution is equally valid. Note that merely adding a mass sheet is not sufficient; one also needs to rescale the original mass distribution by the same factor λ, which justifies the alternative name of *steepness degeneracy*. The density of the mass sheet has to be precisely the critical mass density for this to work. For this reason, a mass sheet cannot be used when there are sources at different redshifts, since these would require different critical densities. Figure 3.4 illustrates the rescaling of the source plane.

If the source involved is variable, it may be possible to measure the time delays between the images:

$$\Delta t_{ij} = t(\boldsymbol{\theta}_i) - t(\boldsymbol{\theta}_j),$$

where $t(\boldsymbol{\theta})$ is the time delay function from equation (2.12), and $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_j$ are the positions of two of the images. For a sheet of mass consisting of precisely the critical density, the projected potential is

$$\psi_s(\theta) = \frac{1}{2}\theta^2,$$

so that the mass sheet degeneracy transforms an initial lens potential ψ_0 into

$$\psi_1(\boldsymbol{\theta}) = \lambda\psi_0(\boldsymbol{\theta}) + (1-\lambda)\psi_s(\boldsymbol{\theta}) = \lambda\psi_0(\boldsymbol{\theta}) + \frac{1-\lambda}{2}\theta^2$$

Noting that the mass sheet degeneracy changes a source position $\boldsymbol{\beta}_0$ to $\boldsymbol{\beta}_1 = \lambda\boldsymbol{\beta}_0$, a quick calculation shows that the relationship between the original time delay $\Delta t_{ij,0}$

and the time delay of the degenerate version $\Delta t_{ij,1}$ simply becomes:

$$\Delta t_{ij,1} = \lambda \Delta t_{ij,0}. \tag{3.3}$$

This means that time delay measurements break the mass sheet degeneracy, since a particular version of the degeneracy corresponds to a particular time delay.

The mass sheet degeneracy also has a particularly simple effect on the magnification factor μ. Since each dimension is scaled by a factor λ, the area of a small source is scaled by a factor λ^2. Keeping the sizes of the images constant, this means that the new magnifications of the source are given by $\mu_1 = \lambda^{-2}\mu_0$. Since absolute magnification values are in practice very difficult to obtain as the source is not observable and since all images are magnified by the same factor, brightness information of the images is in general not useful for breaking the mass sheet degeneracy.

3.4 Generalizations of the mass sheet degeneracy

An infinite sheet of mass cannot be used to create degenerate solutions when multiple sources at different distances, or different redshifts are present. For this reason, it is often claimed that the presence of two sources at different redshifts suffices to break the mass sheet degeneracy (e.g. [1]). Although this clearly breaks the degeneracy in its original form, below we shall see that even when multiple sources are present, it is possible to create rescaled versions of them. First we shall consider the case in which the sources are rescaled with the same factor (based on [69]), afterward the case in which each source is rescaled by a different factor shall be considered.

3.4.1 Identical scale factors

Let us first start by considering a single source again. As was shown in the previous section, it is fairly straightforward to produce equally compatible solutions using a mass sheet. An infinite sheet of mass, however, is not the only mass distribution which can be used to produce such degenerate solutions. If the mass density is circularly symmetric and equal to Σ_{cr} in an area large enough to encompass all the images, the same source scaling effect will arise, thanks to equation (2.16). The center of symmetry of such a distribution determines the center of the scaling (which is the origin of the coordinate system in case of an infinite sheet). This way, the mass sheet degeneracy is easily transformed into a mass disk degeneracy. In fact, the added mass density need not be constant inside such a disk to produce the same effect. As long as the *total* mass inside each image point is the same as for the mass disk, equation (2.16) ensures that the distribution can be used to construct a degenerate solution as well. This constraint automatically implies a density equal to Σ_{cr} inside the annuli in which the images reside, but otherwise allows a lot of freedom.

This freedom allows us to construct a mass distribution which, when added to a scaled version of an existing solution for the lens mass density, is equally compatible with

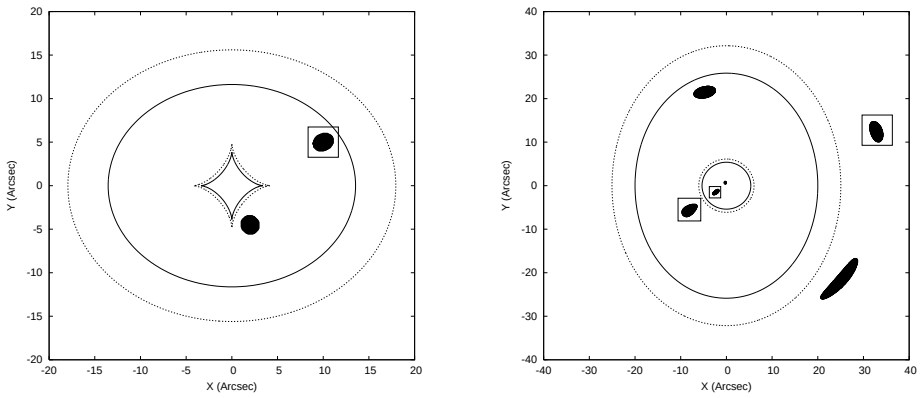


Figure 3.5: Left panel: original sources used to illustrate the construction of degenerate solutions. The source surrounded by a box is placed at redshift $z_1 = 1.2$, the second source is at $z_2 = 1.8$. The caustics created by the non-singular isothermal ellipse placed at $z = 0.5$ are also visible. The solid line corresponds to $z_1 = 1.2$, the dotted line to $z_2 = 1.8$. Right panel: images of the two sources used to illustrate the construction of degenerate solutions. The critical lines are also shown.

the observed images, but which will rescale the sources. The effect is therefore very similar to that of the mass sheet degeneracy, but this degeneracy is not necessarily broken by the presence of additional images of sources at different redshifts.

To illustrate the procedure, consider the two sources and their respective images in figure 3.5. The two sources are placed at redshifts $z_1 = 1.2$ and $z_2 = 1.8$ and the images are created by a non-singular isothermal ellipse at $z = 0.5$. This non-singular isothermal ellipse then provides us with the initial mass density $\Sigma_0(\boldsymbol{\theta})$. A flat cosmological model with $\Omega_\mathrm{m} = 0.27$, $\Omega_\Lambda = 0.73$ and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ was used to calculate the necessary angular diameter distances.

The circularly symmetric mass density $\Sigma_\mathrm{gen}(\theta)$ and corresponding $M_\mathrm{gen}(\theta)$ that we shall construct, must have the same effect as a mass sheet for both sources. This mass density will serve as the generator of the transformation which creates a degenerate solution $\Sigma_1(\boldsymbol{\theta})$ from an existing solution $\Sigma_0(\boldsymbol{\theta})$. The procedure is very similar to the mass sheet case:

$$\Sigma_1(\boldsymbol{\theta}) = \lambda\Sigma_0(\boldsymbol{\theta}) + (1-\lambda)\Sigma_\mathrm{gen}(|\boldsymbol{\theta} - \boldsymbol{\theta}_\mathrm{c}|), \tag{3.4}$$

in which $\boldsymbol{\theta}_\mathrm{c}$ is the center of symmetry of the generator. The mass distribution of the generator must satisfy constraints provided by the images: the mass enclosed by each image point must equal the mass of the corresponding constant-density mass sheet. Therefore, if a specific image of a source at redshift z lies in an annulus with inner radius θ_in and outer radius θ_out, the constraint provided by said image is the following:

$$\forall\theta \in [\theta_\mathrm{in}, \theta_\mathrm{out}] : M_\mathrm{gen}(\theta) = \pi D_\mathrm{d}^2\theta^2\Sigma_\mathrm{cr}(z), \tag{3.5}$$

in which the radii are measured with respect to the chosen center of symmetry $\boldsymbol{\theta}_\mathrm{c}$.

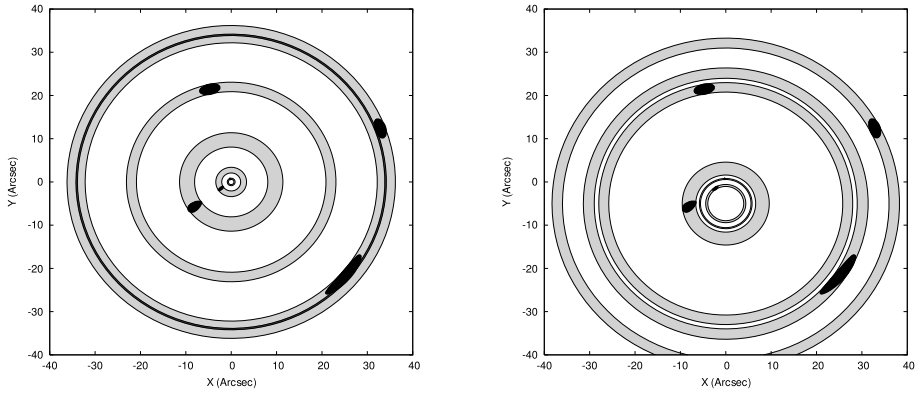


Figure 3.6: Left panel: the annuli in which the images reside, as seen from the origin of the coordinate system, are displayed as gray rings. The darker ring indicates the region in which the annuli of the outer images overlap. Because of this overlap, no suitable mass density can be constructed (see text). Right panel: similar to the left panel, but now the annuli are centered on $(0, -5)$. This center can be used in the construction of a degenerate solution since there are no longer overlapping annuli.

Consequently, within such an annulus the mass density must equal the critical density for an image at redshift z and in the region enclosed by the annulus, the mean density must equal the critical density. In the left panel of figure 3.6, we plot the annuli of the images, as seen from the center of the non-singular isothermal ellipse. Looking at the furthest image of each source, it is clear that no $\Sigma_{\rm gen}$ can be constructed. The mass density would have to be equal to $\Sigma_{\rm cr}(z_1)$ inside the annulus of one image and $\Sigma_{\rm cr}(z_2)$ inside the annulus of the other image. Since these regions overlap, as is indicated by the darker ring, this is impossible. However, if we take $(0, -5)$ as the center, there are no overlapping annuli as can be seen in the right panel of figure 3.6.

Once an appropriate center has been identified, the positions of the images of each source can be used to calculate parts of the total mass map $M_{\rm gen}(\theta)$, as specified by (3.5). In our example, these constraints are illustrated by thick black lines in figure 3.7 (left panel), when using $(0, -5)$ as the center of the distribution. The rest of the mass map can easily be interpolated, after which the full density profile of $\Sigma_{\rm gen}(\theta)$ can be derived. In the left panel of figure 3.7, a third degree polynomial was used to interpolate between the constrained regions. The resulting density profile is plotted in the right panel of the same figure and the critical densities for the two sources are indicated with dotted lines. Note that although this particular example does not require negative densities, in general it is possible that this is indeed necessary. This need not be a problem, since the resulting mass distribution will be combined with the existing distribution $\Sigma_0(\boldsymbol{\theta})$ (the non-singular isothermal ellipse in this example) and may still yield an overall positive density profile. Still, placing a positivity constraint on the overall density profile may help to alleviate this degeneracy.

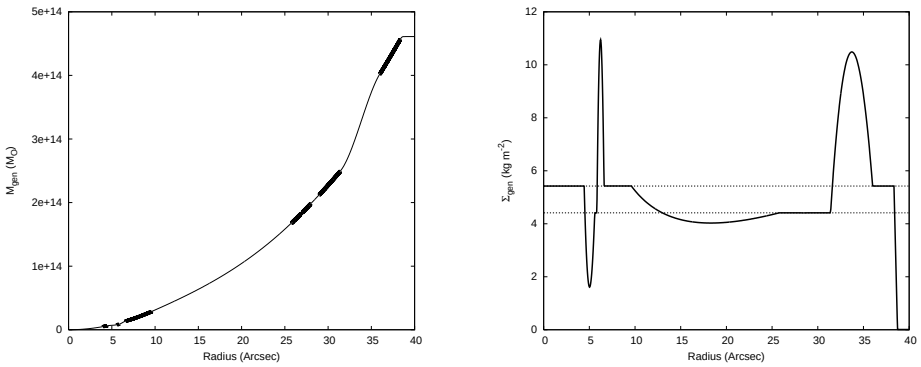


Figure 3.7: Left panel: the positions of the images of each source place constraints on the enclosed mass $M_{\rm gen}$ (thick lines on the mass profile). The regions in between can easily be interpolated. Right panel: the total mass profile in the left panel gives rise to the density profile $\Sigma_{\rm gen}$ shown here. The dotted lines indicate the critical mass densities for the two sources.

By construction, the procedure (3.4) has the same effect as the mass sheet degeneracy: the observed images are identical but the reconstructed sources are scaled versions of the original ones while the density profile of the lens has become less steep. The resulting density profile for $\lambda = 0.75$ can be seen in figure 3.8, in which the original profile is shown as well. Clearly, the central peak has become weaker while at larger radii a ring of excess density has been introduced. This figure illustrates nicely that the term steepness degeneracy still applies to this kind of degenerate solution. When the images of figure 3.5 are projected back onto their source planes using the new mass distribution Σ_1, the sources in figure 3.9 (solid lines) are retrieved. The fact that the images of a specific source overlap perfectly when projected onto the source plane proves that the constructed mass distribution is still compatible with the observed images and can therefore correctly be identified as a degenerate solution. Since each dimension is scaled by $\lambda = 0.75$, the reconstructed sources are smaller than the original ones (dotted lines). The image also clearly shows that the direction of the scaling is towards $(0, -5)$, the center of the circularly symmetric $\Sigma_{\rm gen}$ which was constructed.

Of course, since the positions of the images are not affected, one is free to repeat the entire procedure using the newly acquired Σ_1 as the "original" solution. In general, if it is possible to create N different circularly symmetric density distributions $\Sigma_{{\rm gen},i}$, each with another center of symmetry $\boldsymbol{\theta}_{{\rm c},i}$, it is easily derived that for any λ, the following mass distribution will still project the images back onto consistent sources:

$$\Sigma_N(\boldsymbol{\theta}) = \lambda^N \Sigma_0(\boldsymbol{\theta}) + (1-\lambda) \sum_{i=1}^{N} \lambda^{N-i} \Sigma_{{\rm gen},i}(|\boldsymbol{\theta} - \boldsymbol{\theta}_{{\rm c},i}|). \tag{3.6}$$

This way, the mass distribution which is added to Σ_0 need not possess circular symmetry anymore and its density profile can become much more complex. Equation

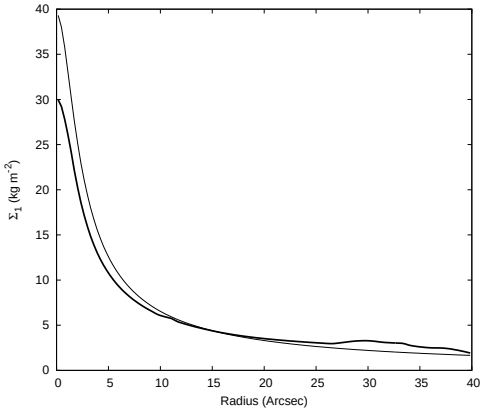


Figure 3.8: The profile of the degenerate solution (thick line) is compared to the profile of the original mass distribution, a non-singular isothermal ellipse. In this example, $\lambda = 0.75$ was used.

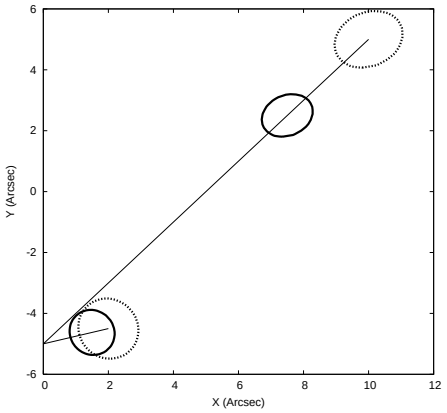


Figure 3.9: Sources recreated by the degenerate solution (thick solid lines); the original sources are indicated by dotted lines. The direction of the scaling is clearly towards $(0, -5)$.

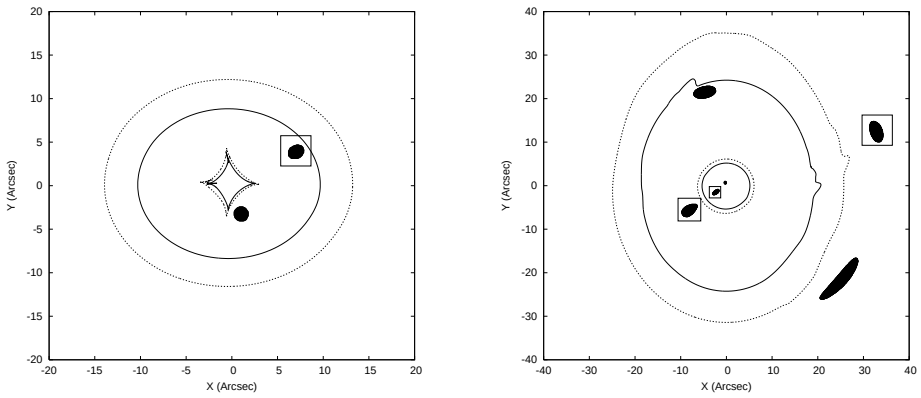


Figure 3.10: Left panel: sources and caustics predicted by a degenerate solution. Comparing with the left panel of figure 3.5 one sees that both sources and caustics are scaled versions of their original counterparts. Right panel: the reconstructed sources and caustics shown in the left panel predict the images and critical lines shown here. The same images as in the right panel of figure 3.5 can be seen and the critical lines still resemble the original ones.

(3.6) is important from a practical point of view: the target scale λ^N can easily be reached by using N generators, each producing only a very small effect. If a large number of suitable center positions can be found, this can severely reduce the amount of substructure introduced by the procedure.

An example of degenerate source and image planes obtained by using $N = 100$ different $\Sigma_{\mathrm{gen},i}$ can be seen in figure 3.10. Each source is scaled by a factor $\lambda^N = 0.75$ in each dimension; the caustics are scaled as well. The critical lines still show the same general structure. The mass distribution of the degenerate solution can be seen in figure 3.11 (left panel) and contains several ring-like structures, the most prominent one being centered on $(-8, -8)$. This can also be clearly seen when the density profile is calculated using $(-8, -8)$ as the center (right panel of figure 3.11). The first peak in this plot is due to the non-singular isothermal ellipse, the second one is caused by the ring-like substructure of the degenerate solution. Note that the ring is not caused by one particular generator, but is a combined effect.

As an aside, even when it is impossible to rescale the sources, it may still be possible to introduce ring-shaped features. One only needs a ring-shaped region without data points. A circularly symmetric distribution which is zero everywhere but which fluctuates in the ring-shaped region in such a way that the total mass inside the region is zero as well, can simply be added to the original, again thanks to equation (2.16). Doing so can obviously introduce a ring-like feature in the mass distribution of the lens. While such a construction only affects the deflection angles in the ring-shaped area, we shall see further on that such a modification may have an effect on the time delays in the system.

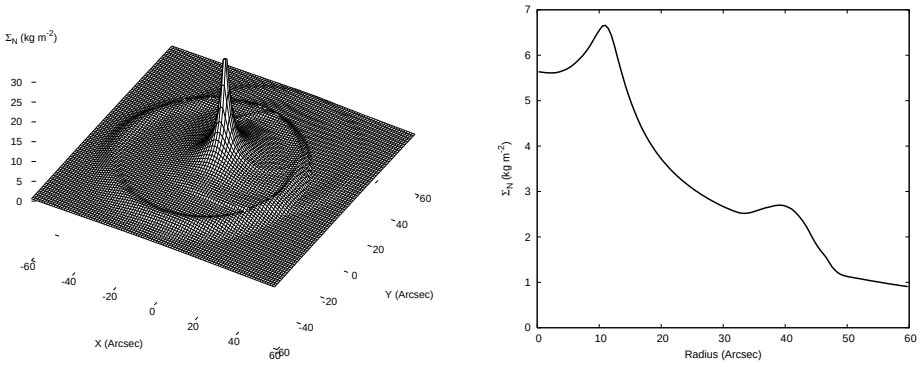


Figure 3.11: Left panel: the degenerate solution which gives rise to the source and image planes shown in figure 3.10. Several ring-like features can be seen, the most prominent one being centered on $(-8, -8)$. Right panel: density profile as seen from $(-8, -8)$. Apart from the peak of the non-singular isothermal ellipse, one can clearly see a ring-like feature.

3.4.2 Different scale factors

Using the generalization of the mass sheet degeneracy that was just shown, the sources are still rescaled with the same factor. But it is also possible to construct a degenerate solution which allows one to scale each source with a different factor. To illustrate this, we shall be using the situation shown in the left half of figure 3.12: an elliptical mass distribution causes two sources to produce three images each. In what follows, only source 1 will be rescaled with a certain factor, in a subsequent step the other source can be rescaled in a similar way, using a different factor. Focusing on only one source and temporarily forgetting about the other, the source can be rescaled using the classic mass sheet degeneracy:

$$\Sigma_1(\boldsymbol{\theta}) = \lambda\Sigma_0(\boldsymbol{\theta}) + (1-\lambda)\Sigma_{\mathrm{cr}}(z_1),$$

in which an initial mass distribution Σ_0 is replaced by a new one, Σ_1, and the critical mass density for the source at redshift z_1 is used. For simplicity, an infinite mass sheet was used here, but in principle this can be replaced by a mass disk or even a more complex mass distribution which has the same effect as a mass sheet for source 1.

This means that to the initial Σ_0, a mass distribution Σ_{add} was added to obtain Σ_1:

$$\Sigma_{\mathrm{add}}(\boldsymbol{\theta}) = \Sigma_1(\boldsymbol{\theta}) - \Sigma_0(\boldsymbol{\theta}) = (1-\lambda)\left[\Sigma_{\mathrm{cr}}(z_1) - \Sigma_0(\boldsymbol{\theta})\right].$$

Taking the second source back into account, it is clear that adding Σ_{add} to the original mass distribution will cause the images of one source to correspond to a scaled version of that source, but the effect on the second source is less clear.

However, suppose it is possible to modify Σ_{add} so that the deflection it causes for source 1 is unaltered, and no deflection is caused at the location of the images of

source 2. In that case, after adding $\Sigma_{\rm add}$, the images of source 1 would correspond to a scaled version of that source, and the images of source 2 would still be projected onto the same source area. Furthermore, since

$$\hat{\boldsymbol{\alpha}}_{\rm add}(\text{images of source 2}) = 0$$

and since

$$\Sigma_{\rm add} \propto \frac{\partial \hat{\alpha}_{x,add}}{\partial \theta_x} + \frac{\partial \hat{\alpha}_{y,add}}{\partial \theta_y}$$

this implies that

$$\Sigma_{\rm add}(\text{images of source 2}) = 0.$$

If one then constructs $\Sigma_1 = \Sigma_0 + \Sigma_{\rm add}$, in general one finds that

$$\begin{aligned} \Sigma_1(\text{images of source 1}) &= \lambda\Sigma_0(\text{images of source 1}) + (1-\lambda)\Sigma_{\rm cr}(z_1) \\ \Sigma_1(\text{images of source 2}) &= \Sigma_0(\text{images of source 2}) \end{aligned}$$

There still remains the question of how to modify $\Sigma_{\rm add}$ so that it does not produce a deflection at the location of the images of source 2. This can be done by using a large number of the monopole basis functions we met earlier. If none of these overlap with the images of source 1, the deflection there will be unaltered. To make sure that the deflection angle at the location of the images of source 2 vanishes, appropriate weights of these basis function must be sought. An example of this approach can be seen in the right half of figure 3.12. In this example, the weights were determined by a genetic algorithm.

Another approach can be used as well. From the discussion above, it is clear that the final version of the mass distribution $\Sigma_{\rm add}$ should produce a specific value of the deflection angle $\hat{\boldsymbol{\alpha}}_{\rm add}$ at the location of the images of source 1. On the other hand, at the location of the images of source 2, no deflection should be produced. We therefore know the deflection angle at specific locations, and at other locations the deflection angle can take a large number of values, the only real constraint being that no additional images should be produced. Interpolating a deflection field in such a way that $\boldsymbol{\nabla} \times \hat{\boldsymbol{\alpha}}_{\rm add} = 0$, or equivalently that the deflection originates from a scalar lensing potential $\hat{\psi}$, is precisely what is done in the LensPerfect lens inversion procedure [19]. Using a similar interpolation procedure to determine $\hat{\boldsymbol{\alpha}}_{\rm add}$ and therefore $\Sigma_{\rm add}$, yields the result shown in figure 3.13.

3.4.3 General considerations

It is clear from the discussions above that these generalizations of the mass sheet degeneracy necessarily introduce substructure into the mass distribution. If, in the most general case, source 1 is rescaled by a factor λ_1 and source 2 by λ_2, for the

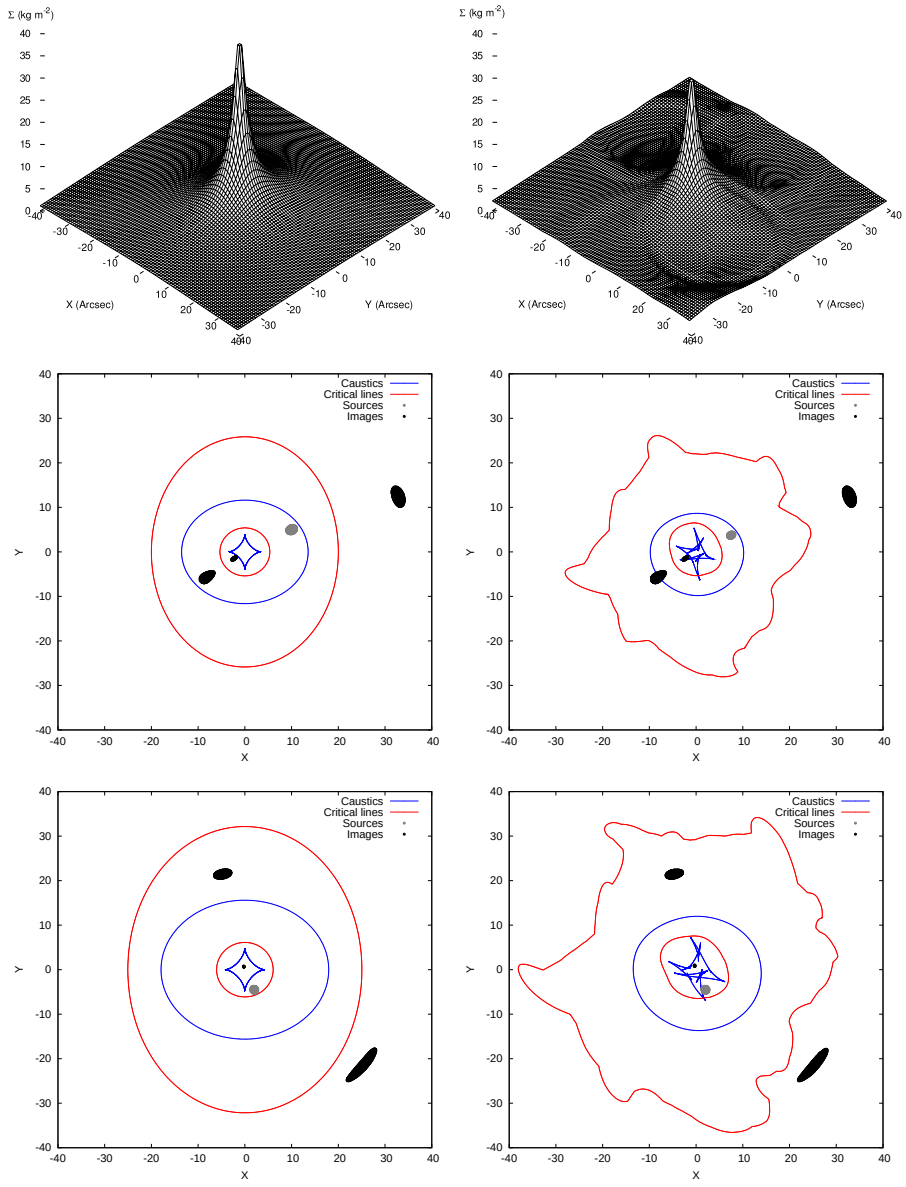


Figure 3.12: Left half: an elliptical mass distribution causes two sources at different redshifts to each produce three images. Right half: a modified version of the mass distribution, using the monopole functions defined earlier, causes the first set of images to correspond to a rescaled version of the source. The images of the second source still correspond to a source of the same size.

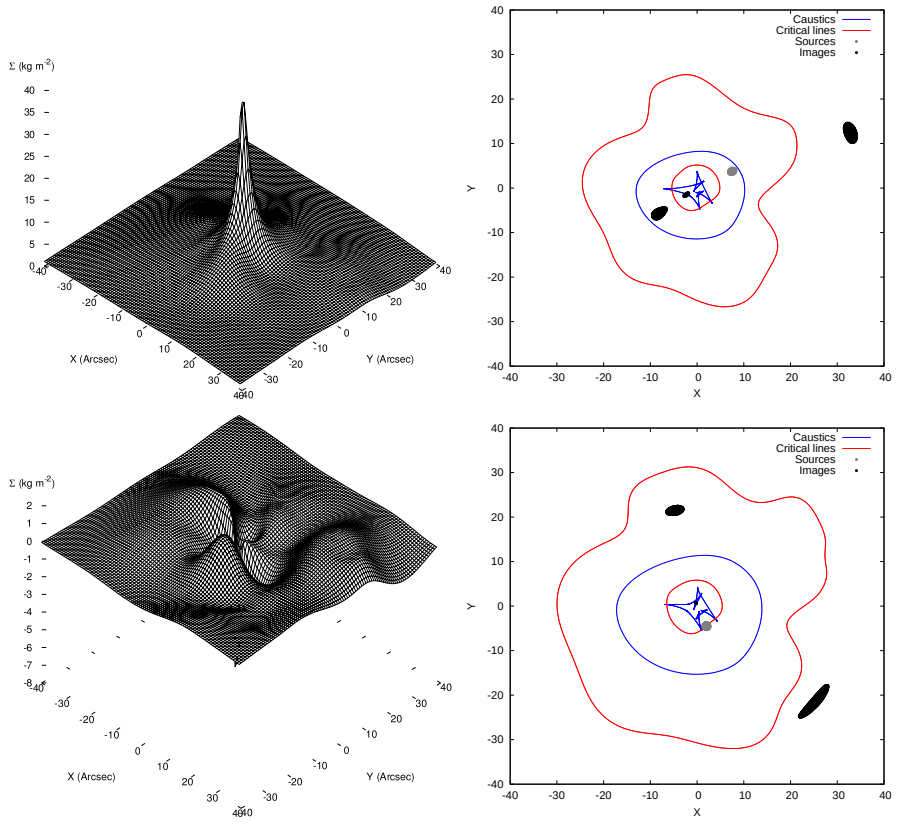


Figure 3.13: Left panel: another mass distribution which causes a scaling effect of source 1 (top). Because the difference with the original mass distribution is hard to see, the bottom half shows the difference $\Sigma_{\rm add}$ between the new and original mass maps. Right panel: this modified mass map also causes source 1 to be rescaled, while not changing the size of source 2.

projected density this implies:

$$\begin{aligned}\Sigma_1(\text{images of source 1}) &= \lambda_1\Sigma_0(\text{images of source 1}) + (1-\lambda_1)\Sigma_{\rm cr}(z_1)\\ \Sigma_1(\text{images of source 2}) &= \lambda_2\Sigma_0(\text{images of source 2}) + (1-\lambda_2)\Sigma_{\rm cr}(z_2).\end{aligned}$$

Note that this is precisely what happens by construction in the case of identical scale factors, as shown by equation (3.5) and in the right panel of figure 3.7. With respect to this degeneracy, the density at the location of the images is directly related to the source scale. Substructure that is introduced at other locations can always be manipulated using the monopole degeneracy, but this cannot change the density at the location of the images themselves.

The effect on the magnification is obvious, and the same as in the classic mass sheet degeneracy. The source area is again scaled with a factor λ^2, possibly different for each source in the most general version, so the magnification of each image of that source changes from μ_0 to $\mu_1 = \lambda^{-2}\mu_0$. The effect on the time delay is less clear than in the case of the simple mass sheet degeneracy, but there clearly is an influence since the overall lens potential changes. Precisely how this changes is more complex.

3.5 Magnification and time delay information

In the previous sections it was shown how the monopole degeneracy can be used to redistribute the mass in between the images, and how the mass sheet degeneracy – or its generalizations – is responsible for changing the size of a source. Now suppose that a particular model for an observed gravitational lens has been found, but that new measurements have shown that the predicted time delays or image brightnesses are wrong. In this section it will be shown how the existing mass map can be modified to yield the correct results, further illustrating lens degeneracies.

For a general lens, the deflection angle is produced by a scalar function $\hat{\psi}$:

$$\boldsymbol{\nabla}\hat{\psi}(\boldsymbol{\theta}) = \hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}),$$

which for a circularly symmetric lens, by using equation (2.16) for the deflection angle, can be seen to give the relation

$$\frac{\partial\hat{\psi}}{\partial\theta} = \frac{4GM(\theta)}{c^2 D_d \theta}.$$

If $\hat{\psi}(\theta)$ or $M(\theta)$ is known, it is straightforward to obtain the other quantity using the relations:

$$\hat{\psi}(\theta) = \frac{4G}{c^2 D_d}\int_0^\theta \frac{M(\theta')}{\theta'}\mathrm{d}\theta' \text{ and } M(\theta) = \frac{c^2 D_d}{4G}\theta\frac{\partial\hat{\psi}}{\partial\theta}. \tag{3.7}$$

Suppose one would like to modify the time delay at one of the images. Recalling the discussion about the time delay function (2.12), it is clear that the images of a source

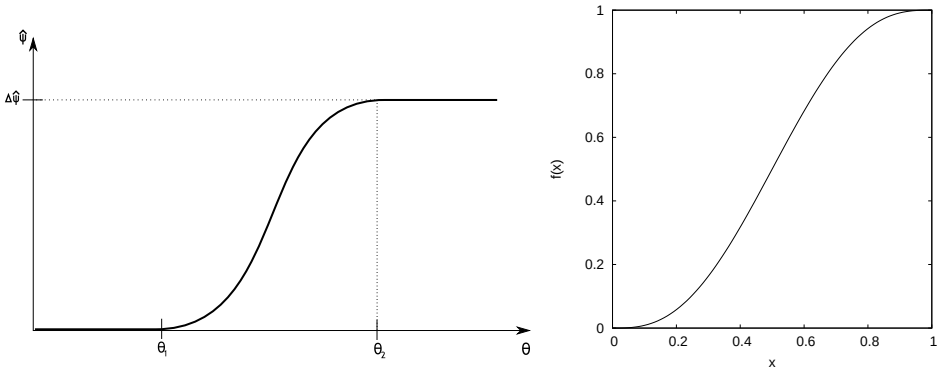


Figure 3.14: Left panel: a lens potential based on this profile can produce a different time delay at one of the image locations. One only needs to make sure that the non-constant part does not overlap with any of the images. Right panel: the non-constant part of the potential shown in the left panel can be modeled using this function $f(x)$.

lie at the positions where $\boldsymbol{\nabla} t = 0$, and at these positions the time delay function can be used to predict time delays. If it is possible to modify the lens potential in such a way that the image locations are not affected, but the time delay is, the goal of modifying the time delay at one of the images will be reached.

If a circularly symmetric potential $\hat{\psi}$ shown in the left panel of figure 3.14 is added to the original potential, centered on one of the images, and in such a way that the non-constant part does not overlap with any of the images, the time delay at the first image will be changed. Since $\hat{\boldsymbol{\alpha}} = \boldsymbol{\nabla}\hat{\psi}$ and $\hat{\psi}$ is constant at the image locations, the deflection field will not be changed there, which means that the same images will still be predicted. Furthermore, because of the extra potential difference $\Delta\hat{\psi}$ at one of the images, the time delay there will be modified. Specifically, if the time delay there should be altered by a value Δt, equation (2.12) shows that

$$\Delta\hat{\psi} = \frac{c\Delta t}{(1+z_d)D_d}.$$

The non-constant part of this $\hat{\psi}(\theta)$ profile can be modeled by the dimensionless function $f(x)$ where

$$f(x) = 6x^5 - 15x^4 + 10x^3,$$

which is shown in the right panel of figure 3.14. The lens potential can then be described by:

$$\hat{\psi}(\theta) = \begin{cases} 0 & \theta \in [0, \theta_1] \\ \Delta\hat{\psi}\ f\left(\frac{\theta-\theta_1}{\theta_2-\theta_1}\right) & \theta \in [\theta_1, \theta_2] \\ \Delta\hat{\psi} & \theta \in [\theta_2, \infty] \end{cases}.$$

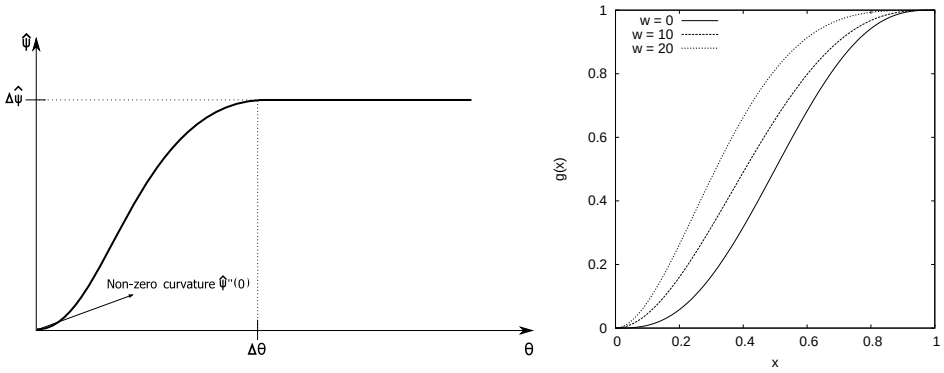


Figure 3.15: Left panel: a circularly symmetric lens potential based on this profile will, when centered on one image, change the predicted magnification there as it has non-zero curvature at the center. Right panel: the non-constant part of the profile in the left panel can be modeled by the function $g(x)$ (see text). The curvature in the origin can be controlled by the parameter w.

From equation (3.7), the total projected mass inside θ_1 and outside θ_2 is zero; in the interval $[\theta_1, \theta_2]$ the total mass profile needed to modify the time delay is given by:

$$M(\theta) = \frac{c^2 D_d}{4G} \frac{\theta}{\theta_2 - \theta_1} \Delta\hat{\psi} \frac{\mathrm{d}f}{\mathrm{d}x} \left(\frac{\theta - \theta_1}{\theta_2 - \theta_1} \right).$$

Something similar can be done to change the predicted magnification at the location of a specific image. Remember from the discussion in section 2.6 that it is the curvature of the lensing potential at the location of an image which determines the magnification. If it is possible to change the curvature at one of the image sites, without affecting the deflection angle there, the same image location will be predicted but the magnification will be changed. By adding the circularly symmetric potential from the left panel of figure 3.15 to the existing potential, centered on the image of which the magnification should be changed, this can be accomplished. The non-zero curvature in the origin will modify the magnification; the zero first derivative makes sure that the deflection angle is not changed at the image location. One then still has to make sure that the area of width $\Delta\theta$ does not overlap with any of the other images, otherwise the deflection field will be changed there and different image locations will be predicted.

The non-constant part of such a potential can be modeled by the function $g(x)$, in which

$$g(x) = \left(6 - \frac{w}{2}\right) x^5 - \left(15 - \frac{3w}{2}\right) x^4 + \left(10 - \frac{3w}{2}\right) x^3 + \frac{w}{2} x^2$$

and which has a dimensionless curvature w in the origin. A plot of this function for a few values of w can be seen in the right panel of figure 3.15. For $w = 0$ this obviously

reduces to the function $f(x)$ from before. The lens potential $\hat{\psi}$ is then given by

$$\hat{\psi}(\theta) = \begin{cases} \Delta\hat{\psi}\, g\left(\frac{\theta}{\Delta\theta}\right) & \theta \in [0, \Delta\theta] \\ \Delta\hat{\psi} & \theta \in [\Delta\theta, \infty] \end{cases} .$$

The curvature in the origin is given by

$$\hat{\psi}''(0) = \frac{\partial^2\psi}{\partial\theta^2}(0) = \frac{\Delta\hat{\psi}}{\Delta\theta^2} w,$$

so if the desired magnification and curvature of the existing lens potential are known, the necessary $\hat{\psi}''(0)$ can be calculated and $\Delta\hat{\psi}$, $\Delta\theta$ and w can easily be chosen to provide the necessary curvature.

Note that even for a fixed value of w, there still exists a further degeneracy between $\Delta\hat{\psi}$ and $\Delta\theta$. To obtain a certain value of the curvature $\hat{\psi}''(0)$, a smaller value of $\Delta\hat{\psi}$ then implies a smaller value of $\Delta\theta$. Only a small change to the existing lensing potential is therefore necessary to influence the curvature and hence the magnification at a specific location.

3.6 Conclusions

This chapter should have made it clear that degeneracies *plague* lens inversions, as one can easily construct another mass map compatible with the observations, once one solution has been found. This is of course to be expected, as lens inversion is an ill-posed problem.

The monopole degeneracy as described earlier, allows one to create degenerate solutions by shifting mass in between the images. Using the information in the images alone, this degeneracy is impossible to break, as each degenerate version predicts the exact same observable properties of the images. In a later chapter we shall see a very nice example of this: in an inversion of the cluster lens Cl 0024+1654 an interesting feature was present, but since this was located in between the images, it was readily removed using this degeneracy and cannot be considered a true feature of the mass map (or at least, not with any certainty). Because of this degeneracy the best one can hope is to obtain constraints on the projected density at the location of the images themselves. This automatically illustrates the need for lensing systems containing many images from many sources.

The mass sheet degeneracy on the other hand does change the density at the location of the images themselves, and therefore introduces some substructure. Since the monopole degeneracy cannot change the mass density at the images, it is actually the degree of smoothness one wishes the solution to have that breaks the degeneracy – at least when multiple sources are present. The degree of smoothness can depend implicitly on the method used to perform the inversion, e.g. in parametric techniques or procedures using overlapping, smooth basis functions, or explicitly as prior information or a regularization scheme, e.g. in the PixeLens method [1]. It is interesting

that neither terms mass sheet degeneracy or steepness degeneracy are very correct in the most general version, as is shown in figure 3.13. Perhaps the name of source scale degeneracy would be more appropriate.

The discussion in this chapter was from the viewpoint of strong lensing data, but the mass sheet degeneracy is present in weak lens inversions as well. This is because in practice one does not measure the shear, but the so called *reduced shear*, a combination of the shear and convergence. It is precisely this quantity that is invariant under the mass sheet transformation. However, as was shown by [6], using weak lensing data it is in principle possible to break the mass sheet degeneracy if individual source redshifts are available and if sources with a rather high distortion are included. Another way to break the degeneracy is to add information about the magnification, for example by using source number statistics [12] or Type Ia supernovae observations [43]. Additional information about stellar dynamics in the gravitational lens can also help to break the degeneracy [62]. Since the mass sheet degeneracy rescales the time delay surface, time delay measurements can be used to break it as well.

It was also shown that it is relatively straightforward to modify the predicted time delays. This does not even change the density at the location of the images themselves. The procedure discussed here introduces a ring-shaped feature in between the images, but this can of course be redistributed using the monopole basis functions.

The predicted magnification at an image location can also be altered quite easily. In principle one only needs to make very small changes to the mass distribution to change the predicted magnification, without modifying the predicted positions of the images. This is similar to the microlensing problem: in strongly lensed quasars, the magnification of the images is often different from the expected magnification based on a smooth lens model. But it is often less clear what kind of modification needs to be made to ensure compatibility with the observations. A possible explanation is microlensing by a single star, as this would not modify the image position by a measurable amount, but could change the magnification factor much. The discussion above shows that a wide variety of changes to the mass distribution can explain measurements, so that the scale of the object that causes the deviation from a smooth model is very hard to establish.

The degeneracies described here may introduce substructure, and therefore degenerate solutions may predict additional images. This means that the area where no images can be seen provides some information as well. In the next chapter we shall see how this so-called *null space* is used in the proposed inversion method.

—4—

Gravitational lens inversion

Having discussed lensing theory and the effect of possible degeneracies, the core topic of this dissertation is now opened: gravitational lens inversion. In this chapter, the inversion procedure based on genetic algorithms is explained, and is tested using simulated gravitational lens systems.

4.1 Parametric and non-parametric methods

As explained in the introductory chapter, lens inversion procedures are usually subdivided into two categories: parametric and non-parametric methods. In parametric methods, a model is chosen based upon the observed image configurations or the visible mass distribution. Because of this, much of the structure in the mass distribution of the lens is fixed before the actual inversion is applied. The inversion step then corresponds to finding appropriate parameters for the model, like the ellipticity or orientation of an elliptical lens model. In non-parametric methods on the other hand, one tries to avoid making assumptions about the shape of the mass distribution. This can be done using a large number of basis functions, allowing a wide range of shapes of the mass distribution.

One can also make the distinction between source plane and image plane based methods. In source plane based methods, the images of a source are projected back onto the source plane, and depending on the positions of the images there, a goodness-of-fit measure is assigned. Note that the uncertainties are measured in the image plane and must be transformed to be taken into account correctly. As the lens equation works in this direction, from image plane to source plane, this is a fast method. In image plane based methods on the other hand, the lens equation is solved: for a suggested

source position the corresponding image locations are calculated. Solving the lens equation for a particular source position is in general hard and time consuming, but is in principle more correct since the number of predicted images can be taken into account and the uncertainties on the image positions can be used entirely correctly.

Many parametric methods exist, some of the most popular being the GRAVLENS [53] and LENSTOOL [57][50] packages. Both can use a combination of source plane and image plane optimization, where a source plane based method is used for speed, and an image based method can be used for accuracy. The technique used in [114] uses a cluster scale component to describe the dark matter influence and adds lens models for the galaxies in a cluster as well. Their technique uses an image plane based goodness-of-fit, and is used to iteratively identify additional multiple image systems. One of the earliest accounts of parametric fitting can be found in [28], where models were fit to the observations of the system Q0957+561 A,B. There are also methods that use a parametric lens model, but allow the source surface brightness to be more complex, such as the RingCycle method [58], LensCLEAN [59], LensMEM [108] and methods derived from them, such as Visibility LensCLEAN [30] and LenTil [112]. In the work of [9] and subsequent articles a Bayesian technique was used to obtain a parametric model of the lens, as well as a gridded source shape. The same authors also investigated using a genetic algorithm to learn about the source shape [8]. A search for appropriate coefficients in a multipole expansion of the lens mass distribution was performed in [102].

In the category of non-parametric inversion procedures, the PixeLens method [91] certainly deserves mentioning. This method uses a lens mass distribution composed of a large number of square cells, of which the mass densities are determined in such a way that the image configuration is best explained. The SLAP method [24] uses basis functions on a dynamic grid, providing better resolution in regions containing more mass. This method was used to estimate the mass distribution of the cluster Abell 1689 [25] and was later modified to include weak lensing measurements in the WSLAP version [26]. The combination of weak and strong lensing data was also done in [7] and applied to the cluster RX J1347.5-1145. In the works [61] and [106], a non-parametric correction to a previously established lens potential is generated, to learn about substructure present in gravitational lenses.

In the realm of non-parametric techniques, one has to choose between optimizing for the lens potential or for the mass distribution. The advantage of working directly with the lens potential is that all lens properties can be derived from local properties of the lens potential. This means that no assumptions have to be made about the mass outside the area in which the procedure works. Since all mass contributes to the lens effect, working with the mass density always involves some assumptions about the region outside of which the inversion method works. On the other hand, making sure that the mass density does not become negative is straightforward if the mass density is modeled directly.

4.2 Inversion method

The inversion method described below belongs to the non-parametric class of methods. By using a large number of basis functions for the mass distribution, no assumptions are made about the mass density of the lens. As will become clear, only minimal constraints will be used to obtain an acceptable lens model. The method was first described in [67] and extensions were discussed in [68], [70] and [71].

4.2.1 Outline of the procedure

The reconstruction procedure requires the user to specify a square shaped area in which the algorithm shall try to recover the mass distribution. The user also needs to specify the images – these can be extended images or point images – as well as which images originate from the same source. The angular diameter distances to lens and sources, or equivalently the redshifts to these objects, are also required input.

In this square shaped area, the mass distribution will be modeled by a *large number of basis functions*. The basis functions that were chosen, are *projected Plummer spheres*, of which the lens effect was discussed in section 2.8.2. Such a mass distribution was chosen as a basis function since it is well behaved at all radii and has a finite total mass. For a mass distribution that consists of a number of Plummer mass distributions of total mass M_i, with width parameter $\theta_{P,i}$ and with symmetry center on $\boldsymbol{\theta}_{s,i}$, the projected mass distribution is easily seen to be:

$$\Sigma(\boldsymbol{\theta}) = \frac{1}{\pi D_{\mathrm{d}}^2} \sum_{i=1}^{N} M_i \frac{\theta_{P,i}^2}{\left(|\boldsymbol{\theta} - \boldsymbol{\theta}_{s,i}|^2 + \theta_{P,i}^2\right)^2}.$$

Since the deflection angle for a single Plummer sphere is known, equation (2.15) becomes:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \frac{4G}{c^2 D_{\mathrm{d}}} \sum_{i=1}^{N} M_i \frac{\boldsymbol{\theta} - \boldsymbol{\theta}_{s,i}}{|\boldsymbol{\theta} - \boldsymbol{\theta}_{s,i}|^2 + \theta_{P,i}^2},$$

which, knowing the projected potential of a single circularly symmetric mass distribution, can easily be seen to originate from:

$$\hat{\psi}(\boldsymbol{\theta}) = \frac{2G}{c^2 D_{\mathrm{d}}} \sum_{i=1}^{N} M_i \ln\left(|\boldsymbol{\theta} - \boldsymbol{\theta}_{s,i}|^2 + \theta_{P,i}^2\right).$$

For the magnification or shear components, one needs the derivatives of the deflection angle with respect to θ, which yields:

$$\frac{\partial \hat{\alpha}^i}{\partial \theta^j} = \frac{4G}{c^2 D_{\mathrm{d}}} \sum_{k=1}^{N} M_k \left(\frac{\delta^{ij}}{|\boldsymbol{\theta} - \boldsymbol{\theta}_{s,k}|^2 + \theta_{P,k}^2} - \frac{2(\theta^i - \theta_{s,k}^i)(\theta^j - \theta_{s,k}^j)}{\left(|\boldsymbol{\theta} - \boldsymbol{\theta}_{s,k}|^2 + \theta_{P,k}^2\right)^2} \right).$$

For clarity, the indices referring to vector components have been written in superscript in this last equation.

The $\boldsymbol{\theta}$ positions for which such quantities will be calculated – the image positions for example – are fixed before the actual inversion starts. This means that the deflection angle, its derivatives and the lens potential are all linear functions in the masses M_i of the basis functions, so that the relevant equations can be written in matrix form. For the lens equation, as was already noted in [24], this goes as follows. A given set of R points in the lens plane, $\boldsymbol{\theta}_k$, $k = 1 \ldots R$, is related to a corresponding set of R points in the source plane. Let Θ be a vector of length $2R$, containing the coordinates of the points in the image plane, in which x and y components alternate. Similarly, B is a vector of length $2R$ which will contain the coordinates of the corresponding points in the source plane. The masses M_i of the Plummer distributions that make up the mass distribution of the lens are stored in an N dimensional column vector M. The lens equation can then be rewritten as

$$B = \Theta - \gamma M,$$

with γ a $2R \times N$ matrix whose components are given by:

$$\begin{aligned} \gamma_{2k-1,l} &= \frac{D_{ds}}{D_d D_s} \frac{4G}{c^2} \frac{(\boldsymbol{\theta}_k - \boldsymbol{\theta}_{s,l})_x}{|\boldsymbol{\theta}_k - \boldsymbol{\theta}_{s,l}|^2 + \theta_{P,l}^2} \\ \gamma_{2k,l} &= \frac{D_{ds}}{D_d D_s} \frac{4G}{c^2} \frac{(\boldsymbol{\theta}_k - \boldsymbol{\theta}_{s,l})_y}{|\boldsymbol{\theta}_k - \boldsymbol{\theta}_{s,l}|^2 + \theta_{P,l}^2}. \end{aligned}$$

The problem of inverting a gravitational lens system is thus transformed into the problem of finding the vector M, given the matrices Θ and γ. By requiring the M_i to be positive, one automatically obtains a mass distribution which is non-negative in the entire region under consideration.

Apart from the Plummer basis functions, an infinite sheet of mass can also be present as a basis function, depending on the choice of the user starting the inversion procedure. This can be useful if enough data are present to break the mass sheet degeneracy. If one only considers the center of a cluster, a non-negligible density offset may be present, which otherwise would need to be modeled by a large amount of basis functions. Section 4.6.3 contains an illustration of this.

Inverting the gravitational lens observations then corresponds to finding appropriate weights M_i of the Plummer basis functions, and possibly a suitable weight of the mass sheet basis function. Searching for these weights is done using a genetic algorithm, possibly even a so-called *multi-objective genetic algorithm*. Before describing this in section 4.2.3, we shall first discuss how the Plummer basis functions are arranged in the square shaped area defined by the user starting the inversion procedure.

4.2.2 Adaptive grid

As was already described in the previous section, the procedure starts with a square shaped area, which the user deems large enough to encompass the projected mass

density of the lens. At first, this area in the lens plane is uniformly subdivided in square grid cells. At the center of each grid cell, a Plummer mass distribution is positioned. The width of each Plummer distribution is set proportional to the side of its grid cell. We tested which proportionality factor allows to best reproduce a wide range of mass densities and found that a value of 1.7 yields a good trade-off between smoothness and dynamic range. The same scale factor was subsequently used in our lens inversion simulations.

The genetic algorithm then breeds, for this given grid and corresponding weights of the basis functions, the best solution M (and possibly a suitable weight of the mass sheet basis function). Given this first approximation of the total mass density, a new grid is constructed by further subdividing grid cells that contain a large fraction of the total mass. This way, the new grid will allow a better approximation of the mass density, without wasting resources on areas which contain little mass. Note that in this step, the mass in the mass sheet is not considered as it does not contribute to the structure in the gravitational lens.

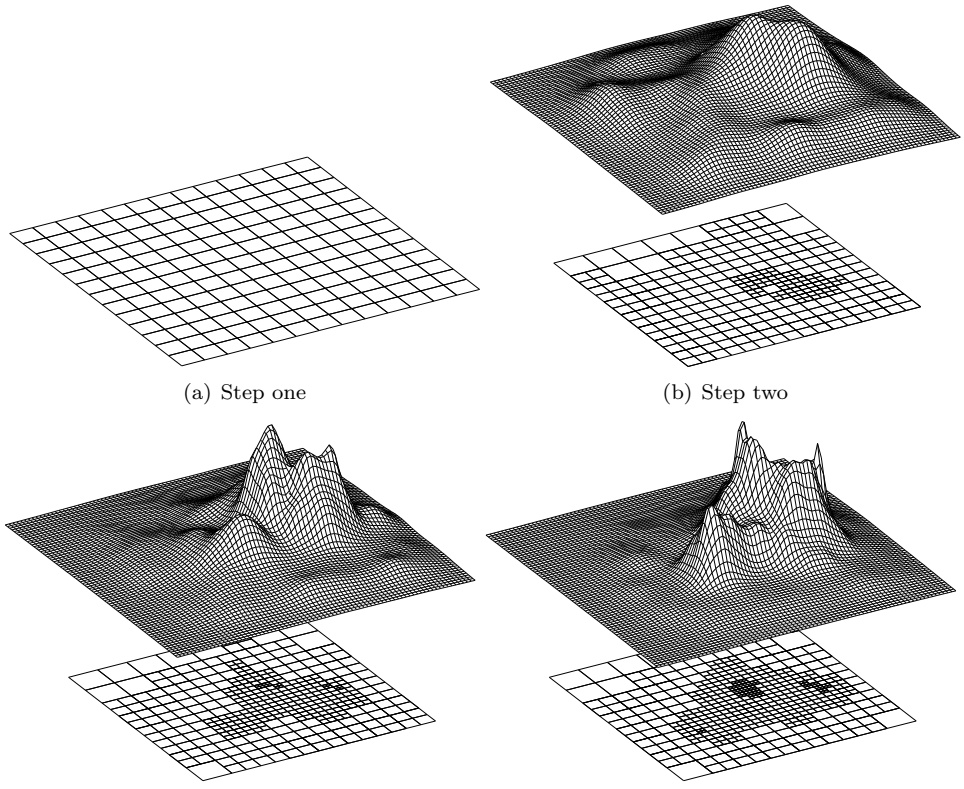
With each cell of this new grid a Plummer distribution is associated and the individual masses are determined by the genetic algorithm, as before. In our implementation, this procedure of refining the grid is repeated unless the number of grid cells exceeds a user defined number. In our tests using simulations, increasing the number of basis functions over one thousand rarely yielded an improved reconstruction. Figure 4.1 illustrates the procedure. At first, a uniform grid is used. With this grid, a first estimate of the distribution is found and this is used to create a new grid. The figure shows a few additional mass density estimates on which new grids are based.

4.2.3 Multi-objective genetic algorithms

As was already described in the introductory chapter, with genetic algorithms, one tries to breed good solutions to a given problem. A central concept is the genome, which is an encoded representation of a possible solution. Usually, the genome will encode the parameters of a specific model, in this case the weights of the basis functions.

For a particular genome, there has to be some kind of measure of how adequate it fits the data. This value is usually called the fitness of the genome. The algorithm starts with a random set of genomes which together form the initial population. From this population, a new one will be created using the following procedure:

- For each genome, the fitness is calculated.
- A new set of genomes is created by combining and copying genomes of the current population. Selection of genomes in this reproduction step should favor genomes with a better fitness.
- Finally, mutations are introduced in the new population to ensure genetic variety.



(a) Step one (b) Step two

(c) Step three (d) Step four

Figure 4.1: The use of a dynamic grid. The grid spacing is refined in those regions where individual grid cells contain a large fraction of the total mass of the lens.

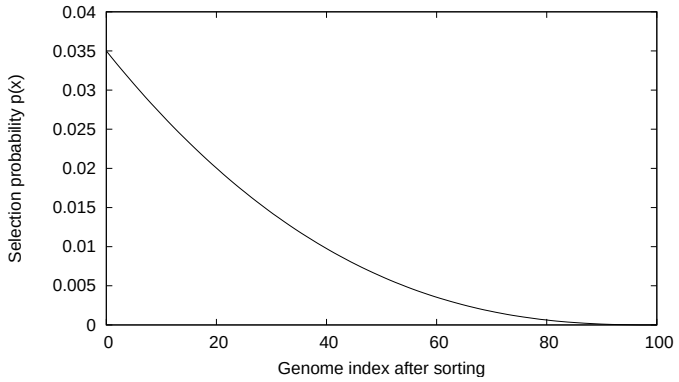


Figure 4.2: Probability distribution when selecting genomes in a population that is 100 genomes large, using a selection pressure parameter $\beta = 2.5$.

By repeating this procedure generation after generation, one tries to breed increasingly better solutions to a problem.

When selecting genomes in the process of creating a new population, a key ingredient of the genetic algorithm is to apply some form of selection pressure: genomes which are deemed more fit, should have a higher probability of creating offspring. This can be achieved by ranking the genomes according to their fitness measures and by assigning a higher selection probability to the genomes with a more favorable rank. To be specific, let s be the population size and suppose the genomes have been sorted according to their fitness measures. The probability of selecting a genome x is then

$$p(x) = \frac{1+\beta}{s}\left(1 - \frac{x}{s}\right)^{\beta},$$

where the parameter β controls the selection pressure. The algorithm described here typically uses a setting of $\beta = 2.5$, but somewhat lower values may be helpful if the genetic algorithm seems to easily get stuck in a non-optimal solution. Figure 4.2 shows this probability distribution for $\beta = 2.5$ and population size $s = 100$.

Other selection schemes exist, but this one has the advantage that the precise fitness value does not matter; one only needs to be able to tell if a genome is better than another one. After introducing some random mutations to ensure genetic diversity, the newly created population replaces the old one and the procedure is repeated until a stopping criterion is fulfilled. For an excellent introduction to the use of genetic algorithms in an astrophysical context, the interested reader is referred to [17].

In tests using lens inversion it was often found that it would be advantageous to use more than one fitness criterion. For example, one would not only like the reconstruction to be compatible with the observed images, one would also like the reconstruction not to predict the existence of additional images. Stated differently, in some cases the area in which no images are observed, i.e. the *null space*, can effectively be used

as well. To be able to use additional constraints, one has to be able to quantify how well a solution performs with respect to these constraints, which yields a set of fitness values. The genetic algorithm as described above only uses a single fitness measure, so one solution would be to combine all these fitness values into one number. In doing so, weight factors can be used to combine them, effectively stating how important each fitness measure is. While this is a possible approach, it would require the user to specify the weights to be used, and may result in a trial-and-error method to find appropriate weight factors. Furthermore, using a specific set of weight values will automatically bias the path followed in the search space as the optimization routine progresses.

Fortunately, there is no need to specify a weight factor. Genetic algorithms are excellent solvers of *multi-objective* or multi-criterion optimization problems. Below, some key concepts are introduced. The interested reader is referred to [23] for an in-depth treatment of this subject.

In a multi-objective genetic algorithm a genome has several fitness measures, each one related to a specific aspect of the optimization problem. A genome is said to *dominate* another genome if two criteria are met:

1. it is not worse in all fitness measures, and

2. it is strictly better in at least one fitness measure.

Using this concept of dominance, one can identify in a population the members which are not dominated by any other genome, resulting in the so-called *non-dominated set*. The concept of a non-dominated set is used to devise a new ranking scheme, as there is no longer a single fitness criterion to base the ranking procedure on. First, the non-dominated set of the entire population is calculated. The genomes in this set will receive the highest rank. After removing this set from the population, a new non-dominated set is identified and these genomes receive the next-to-highest rank. The procedure is repeated until all genomes of the population are processed. Afterwards, the genetic algorithm can proceed as before. This procedure is often referred to as *non-dominated sorting*.

When there are conflicting objectives, there is a whole range of optimal solutions: the Pareto-optimal front. There exist procedures to preserve a certain amount of diversity in the population, allowing the Pareto-optimal front to be well sampled by the genetic algorithm. In our specific case however, the objectives are not conflicting and no further modifications are required.

The genetic algorithm may be instructed to use *elitism*. In this case, in a single-objective genetic algorithm, the best trial solution in one generation is automatically copied to the next. In a multi-objective setting, one or a few genomes of the current set of best solutions can be introduced into each generation. This may help to obtain a better and faster convergence, but on the other hand it may also evolve the algorithm towards a sub-optimal solution identified early on in the search.

4.2.4 Finalizing and averaging

The dynamic grid method has one disadvantage: regions containing only a relatively small portion of the total mass will not be subdivided into smaller grid cells. As a result, the basis functions in such regions may lack the resolution needed for an accurate reconstruction. To overcome this problem, a finalizing step can be added to the procedure. A uniform grid of 64 by 64 grid cells for example, is then created and the associated basis functions are used as small corrections to the current estimate of the mass distribution. Because they are corrections, the weights of the basis functions are allowed to be negative. The genetic algorithm again determines appropriate values for these weights.

To create a single candidate solution, first the dynamic grid method is used to create a first good estimate of the mass distribution and afterward, small-scale corrections are introduced in the finalizing step. This entire procedure is then repeated a number of times, each time yielding a somewhat different mass distribution. One can then calculate the average mass distribution to inspect the features which are common in all reconstructions, and one can calculate the standard deviation to check in which regions the individual solutions disagree. Averaging the solutions will also increase the smoothness of the retrieved mass density.

Being able to create an averaged solution is an attractive feature of our approach, but it would be of little use if the resulting mass density would not be a good solution of the inversion problem (or a worse solution than the individual solutions). Using simulations (see section 4.6), we found that the averaged solution is indeed also a good solution, with a very high fitness, and in many cases even does a better job than many of the individual solutions. This is because the random mutations that occur during the reproduction process of the genomes, cause the best solution to oscillate around the "true" solution. Since averaging a set of solutions suppresses these random fluctuations, the averaged solution can be a more faithful realization of the true solution than any of the individual solutions.

4.3 Genetic algorithm details

The actual inversion of a gravitational lens system comes down to determining the weights of the basis functions, so it is clear that the genetic algorithm lies at the core of the inversion procedure.

4.3.1 Genome and basic fitness measure

The goal of the genetic algorithm is to determine good values for the masses of the individual Plummer distributions which are laid out according to a specific grid. Therefore, the genome in our genetic algorithm will encode the masses of these Plummer distributions.

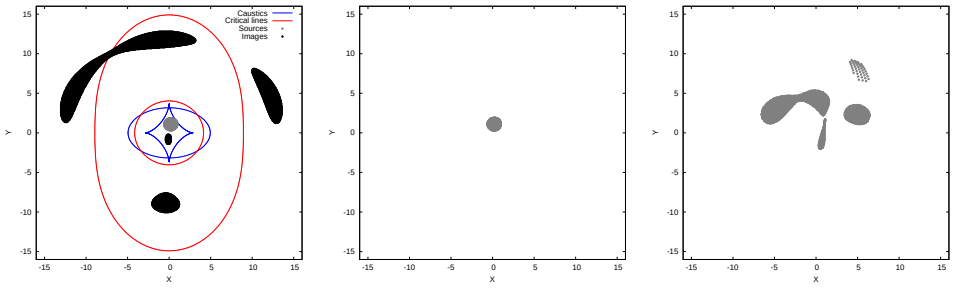


Figure 4.3: Left panel: the ray-trace procedure has been used to calculate the images of a circular source, being lensed by an elliptical mass distribution. Center panel: if the images of the left panel are projected back onto the source plane using the correct lens equation, a single, consistent source shape is recovered. This is only natural, since this source was used to create the different images in the first place. Right panel: if on the other hand, an incorrect lens equation, or correspondingly an incorrect mass distribution, is used to project the images back onto the source plane, each image will occupy a different region in the source plane. As the situation in the center panel is clearly superior, this suggests that the overlap between the back-projected images can be used to base a fitness measure on.

For a specific set of basis function weights, we need to define a way to evaluate how good the corresponding solution is. Since we are working in the strong lensing regime, it is assumed that the gravitational lens system produces multiple images of one or more sources. Figure 4.3 illustrates the basic idea to judge how good a trial solution is. Suppose that a specific lens causes the images of a source which can be seen in the left panel of this figure. If one would project these images back to the source plane using the correct lens equation, one would find that back-projected images of the same source will overlap perfectly, as can be seen in the center panel. However, if one were to use an incorrect gravitational lens to project the images back onto the source plane, a situation like the one shown in the right panel of this figure could arise. There, the back-projected images clearly do not correspond to a single, consistent source shape. For this reason, the degree in which back-projected images of the same source overlap will be used to determine how good the suggested solution actually is.

The basic way this is implemented is as follows, more details can be found in section 4.4. For a given solution of the mass distribution of the lens, the images of a single source are projected back to the source plane. The areas occupied by each image are surrounded by rectangles: two examples are shown in Figure 4.4. Corresponding corners of the rectangles are connected with imaginary springs. Consider two rectangles, each enclosing a back-projected image. For corresponding corners, the distance is calculated in absolute units, for example in units of arcminutes or arcseconds. In a previous step, a length scale was calculated as the average of the lengths of the sides of all the rectangles belonging to a specific source. The distance between two corresponding corners is then divided by this length, yielding a dimensionless distance d. The "potential energy" for this pair of corners is then simply d^2. Repeating this

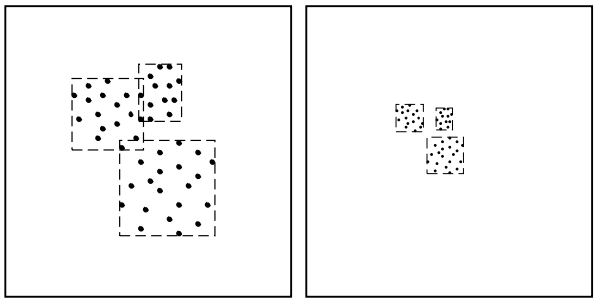


Figure 4.4: Evaluating the fitness of back-projected images. In the left panel, on an absolute scale, the value of the potential energy of the imaginary springs connecting the rectangles that enclose the back-projected images of a single source is higher than in the situation depicted in the right panel. However, the back-projected images in the left panel overlap, unlike those in the right panel. We therefore scale the rectangles enclosing the back-projected images of a single source with the mean size of the rectangles when calculating the fitness value.

for the other three corners and adding together the energies then gives the potential energy of these two rectangles. For a specific source, this procedure is then done for all pairs of back-projected images and the sum of these potential energies is then the potential energy contribution of this source. The fitness value of a given lens solution is the sum of the potential energies of all sources. It is important to take into account the scaling of the rectangles when calculating the potential energy values. Comparing the left and right parts of figure 4.4, it is clear that the left situation definitely corresponds to a better overlap, while on an absolute scale the potential energy of the right situation will be the lower one. For this reason, we express distances between corners of rectangles relative to the size of the rectangles, or, in other words, relative to the size of the source.

As was mentioned above, the genome represents the masses of the individual Plummer distributions. To be more precise, the genome only represents the relative contribution of each Plummer distribution: each Plummer mass is represented by a dimensionless number between 0 and 1000. These numbers are stored in the vector M and the matrix product

$$\Theta' = \gamma M \tag{4.1}$$

is calculated. For the dimensionless masses to be converted into real masses, the vector M needs to be multiplied with a factor λ, bringing the lens equation in the form

$$B = \Theta - \lambda\Theta'. \tag{4.2}$$

Since Θ and Θ' are constant column matrices, it is an easy and computationally inexpensive task to find, for a given M, the factor λ that maximizes the fitness, or, in other words, for which the back-projected images of the sources coincide best. The value of the fitness of that particular situation is then considered to be the fitness of

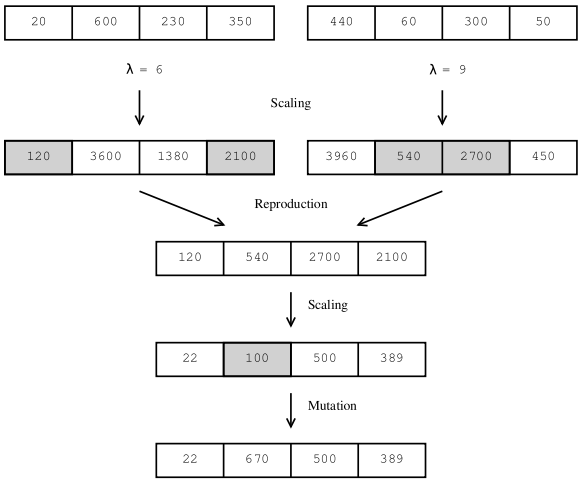


Figure 4.5: Reproduction and mutation of genomes. See text (section 4.3.2) for explanation.

the genome.

This search for an appropriate scale factor λ for the so-called *positional fitness* is always done, even when multiple fitness measures are being used. For example, when both a positional fitness measure and null-space fitness measure are being used (to avoid the prediction of extra images), the positional fitness alone is first used to determine the scale factor λ and using the situation described by this scale factor, the null-space fitness measure is calculated.

4.3.2 Cross-over and mutation

In our implementation, to obtain a new population, some genomes are copied from the original population while others are obtained by merging two genomes. These processes are called *cloning* and *cross-over* or *recombination*, and are intended to mimic the effects of asexual and sexual reproduction respectively.

The procedure of merging two genomes consists of a few steps which are illustrated in figure 4.5. At first, the values between 0 and 1000 of each genome are multiplied with their best λ value to obtain the true Plummer masses they represent. Then, for each Plummer distribution, the procedure selects at random the mass from one of the two genomes. In genetic algorithm literature this is often referred to as *uniform cross-over*. Finally, these values are rescaled in such a way that the largest number is 500.

When the new population is complete, mutations are introduced in some genomes. In early generations, some values are simply changed to a random number between 0 and 1000. It is for this reason that the previous step rescaled the Plummer masses

to a maximum value of 500. This way, a random change of the value will also allow a considerable increase in mass for that Plummer distribution.

When the best fitness values of successive generations start to converge, a new mutation rule is adopted. In this case, random numbers in the interval $[-200, 200]$ are generated and added to some of the genomes' values. Resulting values which are negative or larger than 1000, are set to zero or 1000 respectively. The first mutation rule makes sure that a large range of mass densities can be inspected. When the algorithm starts to converge near a good solution, the second mutation rule assures that the algorithm can more closely approach the best solution. If desired, extra steps can be introduced where even smaller mutations are introduced, but this tends to change very little to the overall mass distribution.

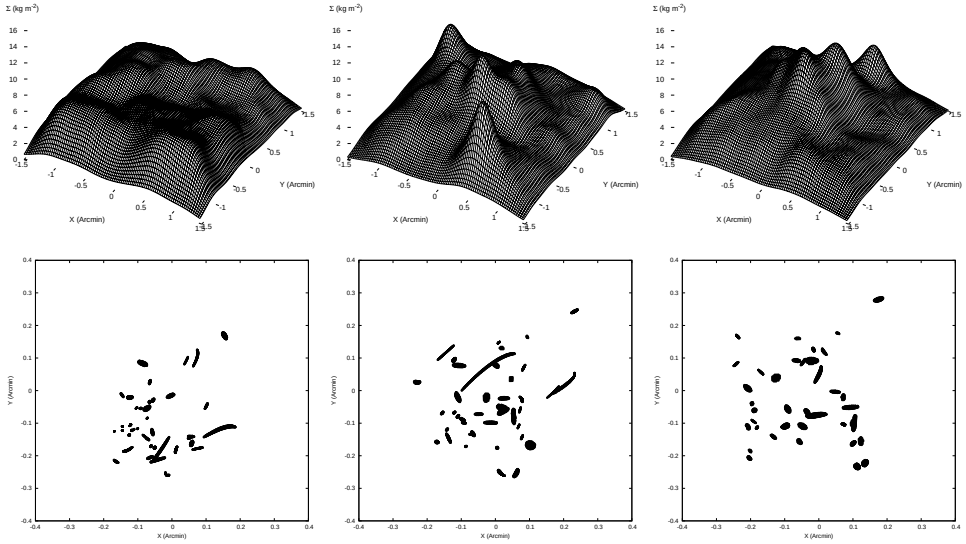
4.3.3 Evolution

It is instructive to see how the genetic algorithm evolves towards a solution. Figures 4.6 and 4.7 show the mass distribution corresponding to the best genome in a population at various stages, or generations, in the genetic algorithm. Also shown are the input images projected back onto their source planes. In this example, the only requirement was that the back-projected images should overlap, so the situation in the source plane is used to judge how well a trial solution is. In essence, the algorithm tries to pull the back-projected images of a source towards each other, in which it is clearly successful: as the genetic algorithm advances, images of a each source start to overlap increasingly well. At the same time, the corresponding mass distribution starts to show similarity to the true mass distribution, which is shown in the left panel of figure 4.12.

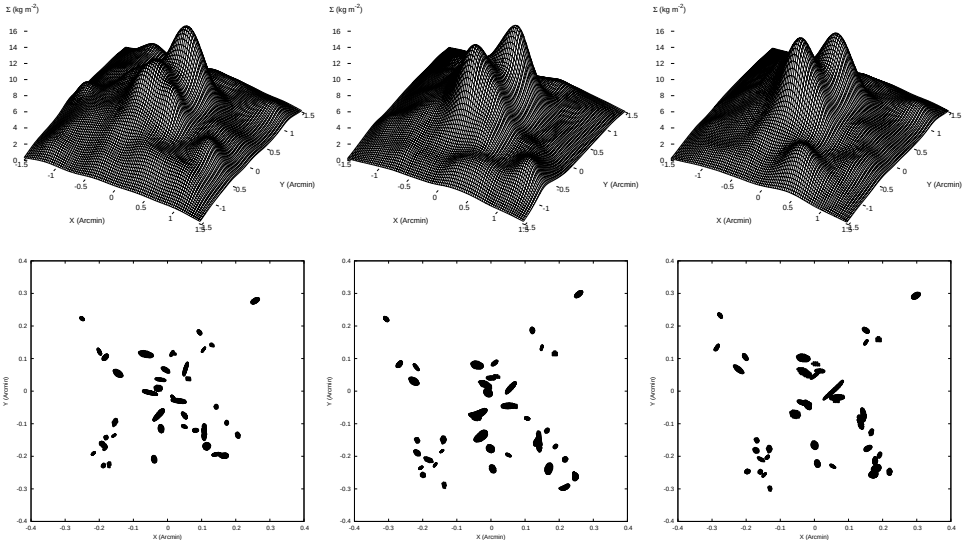
It is also clear from these figures that most of the structure of the resulting mass map is determined relatively early on in the process. The final stages will definitely cause a better overlap in the back-projected images, as can be seen by comparing generation 100 to generation 2000, but the corresponding changes in the mass distribution are minimal.

4.4 Fitness measures for lens inversion

The weights stored in a genome will determine what the mass distribution looks like, and will determine the deflection angle, its derivatives and the lens potential. This information can be used to calculate the back-projected positions of the images in the source plane, the magnification factors etc., and it is based on this situation that fitness measures are assigned. Below, different fitness measures that can be used, depending on the information at hand, are described.

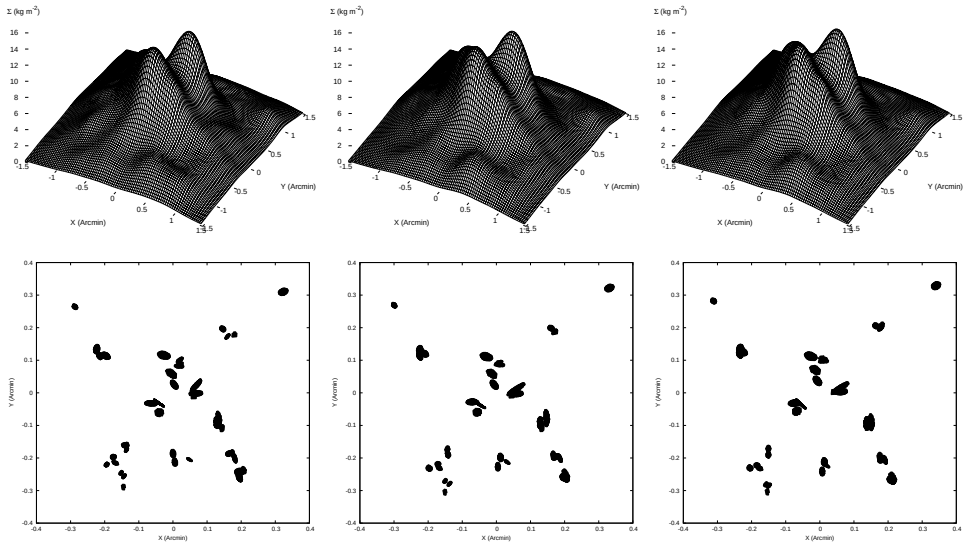


(a) Generation 1 (b) Generation 10 (c) Generation 20

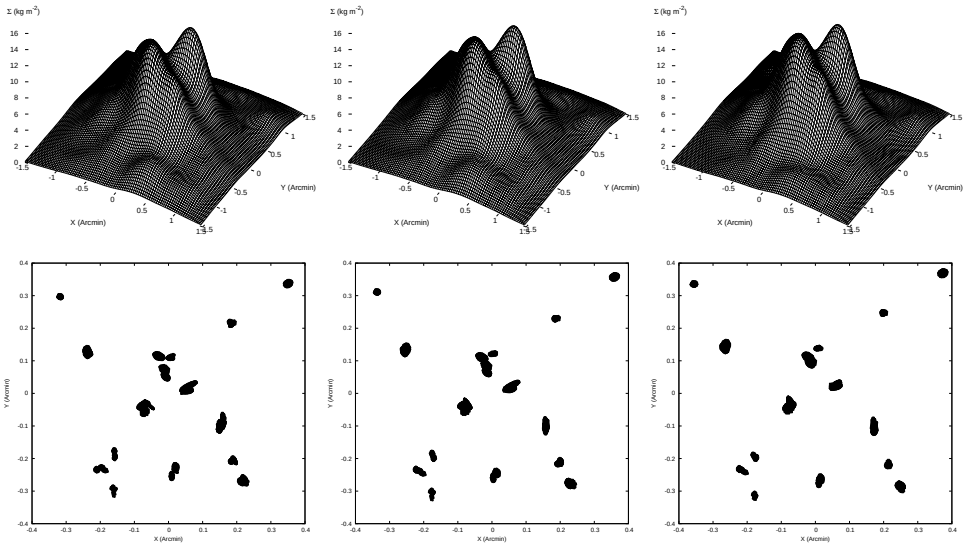


(d) Generation 30 (e) Generation 40 (f) Generation 50

Figure 4.6: This figure shows the mass distribution and back-projected images corresponding to the best genome in a population at various stages in the genetic algorithm. This allows one to look at the evolution of the genomes that is taking place. Shown here is the best genome in generation 1, 10, 20, 30, 40 and 50.



(a) Generation 100 (b) Generation 200 (c) Generation 400



(d) Generation 700 (e) Generation 1000 (f) Generation 2000

Figure 4.7: Continuation of figure 4.6. Shown here is the best genome in generation 100, 200, 400, 700, 1000 and 2000.

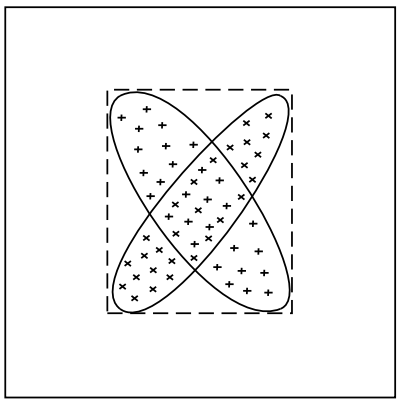


Figure 4.8: When different images of an elliptical source are projected back onto the source plane, the situation shown has a good fitness value, even though the images do not overlap well. This problem is easily resolved by allowing the enclosing rectangles to be rotated.

4.4.1 Positional fitness

The basic idea behind the so-called *positional fitness measure* was already encountered earlier. Since the images originate from a single source, when back-projected onto the source plane, the images of each source should overlap to form a consistent source shape.

Extended images

For extended images, the images of each source are projected onto the source plane. There, rectangles of which the sides are parallel to the coordinate axes are used to surround the images, and the distance between corresponding corners of these rectangles are used to calculate the overlap between the images. In doing so, the average size of the rectangles is used as the length scale to determine the distance. This makes sure that the genetic algorithm does not evolve towards a solution which over-projects the images. This can appear to be a better solution on an absolute scale, while on the size of the source itself the overlap is actually worse (see figure 4.4).

If few sources are present, it is possible that one encounters the following problem if this method is used, as is illustrated in figure 4.8. Because the rectangles used to determine the degree of overlap in the back-projected images are aligned with the coordinate axes, crossing elliptical shapes will be interpreted as overlapping images. If there are a large number of multiply imaged sources and hence a large number of independent constraints, this problem does not manifest itself. The solution to the problem is clear: the rectangles which are placed around the back-projected images should be rotated until they are aligned with their respective images. This can be done in an efficient way using the “rotating calipers” algorithm [98].

If inside each image several features can be identified which originate from the same source features, these image points can be identified as extra constraints. The principle is the same as before: connect the corresponding features with virtual springs, and the potential energy is used as as fitness measure. Here also, the estimated source size is used as the distance scale.

Point images

The original algorithm was designed to work with extended images. As is clear from the discussion above, this yields a natural scale size which can be used in determining how well the back-projected images overlap. However, in many studies (e.g. [10], [38] or [72]) point-like information is provided instead. For this reason, it would certainly be practical if this information could be used directly.

Unfortunately, when using point like images, the natural scale size which can be used to calculate the overlap of the images is no longer present. As shall be illustrated in section 4.6, the following approach seems to work very well. First, all the point-like images are projected back onto their source planes. Then, the size of the rectangle (again aligned with the coordinate axes) which encloses all these back-projected points is sought, and this is used as the length scale in measuring the distances between points originating from the same source. Using such a length scale not only makes the measure independent of an absolute scale size, it also prevents a general over-projection of the images. This would scale the source planes as a whole and therefore would automatically reduce the fitness value if an absolute scale is used.

4.4.2 Magnification information

It is possible that in an extended source there are features that effectively behave as images from a point source. Although the magnification is well described by the relation between source size and image size, the precise magnification may differ at the location of these point-like features. A typical example is a quasar-image as the point-like image, embedded in an image of its host galaxy.

Although the discussion in the chapter about degeneracies should have made it clear that such magnification information does not tell much about the overall mass distribution, it may be that for some reason one would like the model to correctly reproduce the brightnesses of the point-like images. In any case, adding extra information may help to find a valid reconstruction, since it will lead to a different exploration of the search space.

Because a consistent source shape has to be found when the images are projected back onto the source plane, the back-projected brightnesses of the point-like images should also be the same. Of course, to project these brightness values back onto the source plane, the magnification value at the position of the image has to be used. One can then compare the brightnesses of these back-projected images as well, and use them as a height value. Instead of using two-dimensional springs to base the fitness

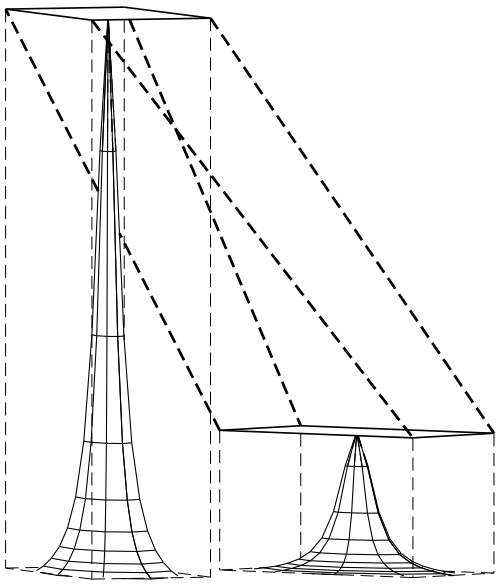


Figure 4.9: Back-projected images are surrounded by rotated rectangles and placed at a height corresponding to the maximal brightness values of the images. Corresponding corners are connected with virtual springs (indicated by thick dotted lines) and the potential energy of the situation is used as the fitness measure. This way, reconstructed sources will also overlap in the brightness domain.

measure on, in this case three-dimensional ones are used. Figure 4.9 illustrates this.

4.4.3 Null-space

The positional fitness measure described above makes the genetic algorithm evolve towards a solution which causes the image regions to be back-projected onto the same area in the source plane. However, no measures are taken to prevent other regions in the image plane from being projected onto the same area in the source plane as well. If that were the case, running the ray-trace procedure from before on the reconstructed lens and sources would clearly predict additional images, as is illustrated in figure 4.10.

One possible solution to avoid these spurious images, is to impose some kind of prior on the lensing mass. For example, one could modify the fitness criterion to favor smoother mass distributions. Undoubtedly, this would reduce the number of extra images since smoother mass distributions have less complex caustic structures. However, this goes against the spirit of our endeavor, which is to find out how much information can be retrieved non-parametrically from lensing systems, using only the information at hand.

In order to avoid introducing this kind of bias in the generated solutions, a different

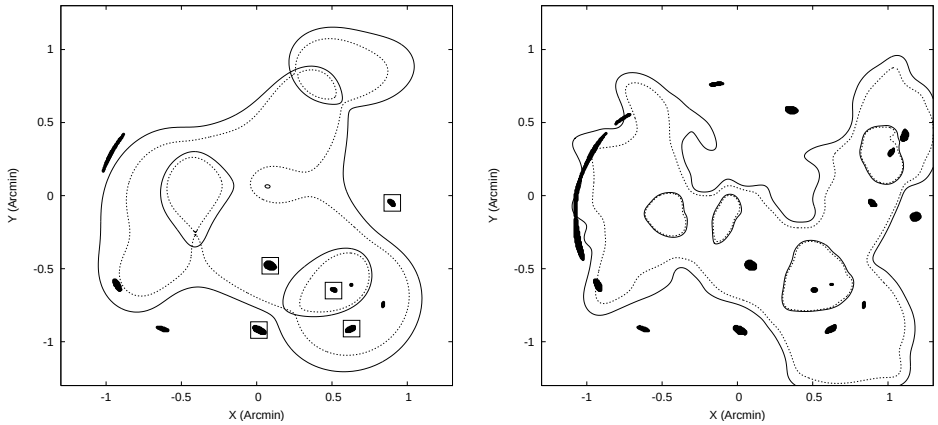


Figure 4.10: Left panel: images generated by two sources. The images of one of the sources are surrounded by rectangles. The critical lines corresponding to the distances to each source are visible as well. Right panel: the algorithm which takes only the positional fitness into account can easily produce solutions which cause the back-projected images to overlap well. However, the reconstructed sources and lens in general can cause undesired images to be produced, as is illustrated in this figure.

approach is used. Note that there is still much information available that has not yet been used. Points in the image plane which are not part of an image of the source, i.e. the null space, provide additional constraints: if they are projected back onto the source plane using the correct lens equation, none of these points will lie inside the source area. Otherwise, an image would be visible at that specific location. Using the null space was already suggested in [24], but there, only null space points adjacent to the images were used. This can avoid the acceptance of solutions that produce images larger than the observed ones, but it obviously fails to avoid the extra images.

To incorporate the information in the null space, two schemes have been explored. In the first one, a regular grid of null space points is generated and the points which fall inside an input image are removed. To avoid generating images which lie relatively far from the other images, and thereby generate more mass than is necessary, the area in which the null space points are generated should be chosen large enough, i.e. somewhat larger than the area of the images themselves. After projecting the images back onto the source plane, their envelope is calculated. This is the smallest convex polygon enclosing all the back-projected points, and is used as the current estimate of the source shape. Next, each point in the null space is projected back onto the source plane and the number of points which lie inside the reconstructed source are counted. Clearly, one would like this count to be as low as possible, since our objective is to remove the extra images. Note that by only considering a discrete number of points in the null space, it is possible that small images are generated which lie in between null space points.

Instead of simply using points, one can also divide the null space into a number of small triangles, similar to the approach in [4]. When these are projected back onto the source plane, for each triangle the amount of overlap with the reconstructed source is calculated. The corresponding area of the triangle in the image plane is then used as this triangle's contribution to the null space fitness measure. Again, this number should be as low as possible to avoid generating extra images. The advantage of this method is that it becomes easier to avoid small undesired images, at the cost of increased computational complexity.

A small variation of this scheme was implemented to be able to include the null space fitness measure in the case that one only has point-like images. While in the previous case, the area in which images could be seen, or in principle could be present (e.g. behind a bright galaxy), is typically excluded from the input grid, the approach here is not to do this. The point images of a source are projected back onto the source plane, and again the envelope of these points is calculated. The number of back-projected triangles of the null space grid that overlap with this shape is then used as an estimate of the amount of images that the source produces. Again, this value should be as low as possible.

Using the null-space is a natural way to avoid the introduction of unnecessary sub-structure in the lens mass distribution, as this typically produces extra images. It therefore has somewhat the same effect as a prior on the smoothness of the mass distribution, but is based on a simple observational argument.

4.4.4 Critical lines

In many cases it is obvious that images are not intersected by a critical line, i.e. that all points of an image have the same parity. This means that when the images are not merging, one would like to avoid critical lines intersecting the images in the reconstruction. It is easy to detect if a critical line intersects an image: one simply needs to calculate the sign of the magnification at each image point. Only if this is the same for all image points, no critical line intersects the image. In practice, the magnification signs of neighboring points are compared and the fitness measure simply counts the number of pairs for which the signs change. Using a fitness measure which disfavors such solutions can help avoid the genetic algorithm being trapped in a sub-optimal region of the solution space. While this method worked well in the case of Cl 0024+1654, applying the same method to SDSS J1004+4112 was far less successful.

Figure 4.11 illustrates the problem. In the left panel, the black regions mark two images of a single galaxy, and the points in each image should all have the same parity. For the constructed solution, the critical lines are shown and they clearly do not intersect the input images, meaning that no parity changes will be present in an image and that the solution will not be penalized. When the proposed mass map is used to project the images back onto the source plane, the situation in the right panel arises. Clearly, when the envelope of the back-projected images (gray area) is

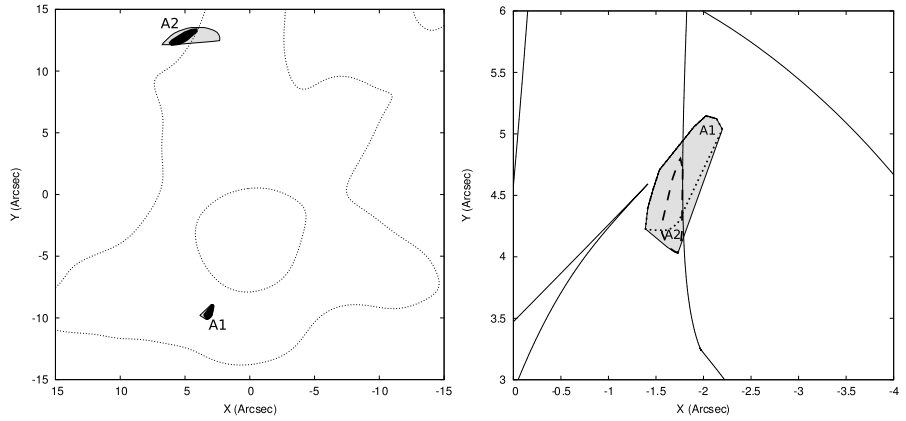


Figure 4.11: Illustration of the problem with the original fitness measure to penalize situations in which a critical line crosses an image. Suppose two input images (left panel, black) are known not to be intersected by a critical line. The critical lines of a certain trial solution indeed do not intersect the input images, so all the points in the input images will have the same parity. However, when the images are projected onto the source plane (right panel), the envelope of both images is in fact intersected by a caustic, causing a critical line to intersect the current prediction of the images (left panel, gray).

considered, a caustic does intersect this region and correspondingly when this shape is used to predict the images, a critical line will intersect an image as can be seen in the left panel. By not specifying precisely what type of solution one is interested in, the existing criterion can easily lead the genetic algorithm towards a sub-optimal reconstruction.

Instead of calculating the magnification information at the location of the images, the value of the magnification is now calculated on a relatively coarse grid covering the region of interest. This is used to create a rough estimate of the critical lines, which in turn are projected onto the source plane to provide an estimate of the caustics. The intersection of the caustics with the source shape is calculated and the total length is used as a fitness measure, a lower value indicating a better fitness.

4.4.5 Time-delay information

When time delay information is available for a number of images of a single source, one would like to use this information to constrain the allowed region in the solution space even further. By calculating the lensing potential at the image points for which time delay information is available, in principle equation (2.12) can be used to compare the predicted time delays with the observed ones. However, to do so, one needs to know the position $\boldsymbol{\beta}$ of the source. While the source position may be estimated once a good overlap of the images has been reached, this is in general not possible while

the genetic algorithm is still evolving, and certainly not near the start, when the trial mass maps are still quite random and the images are projected onto very different regions.

Having tested a number of possible fitness measures, we found that the following one works very well. Suppose that there are N images $\boldsymbol{\theta}_i$ with corresponding points in the source plane $\boldsymbol{\beta}_i$. It is possible that time delay information is not available for all images, so let us call T the set of image indices for which time delay information is at hand. The measured time delay between image i and j will be called $\Delta t_{\mathrm{obs},ij}$. Explicitly stating the $\boldsymbol{\beta}$ dependence in the time delay function:

$$t(\boldsymbol{\theta}, \boldsymbol{\beta}) = \frac{1 + z_d}{c} \frac{D_\mathrm{d} D_\mathrm{s}}{D_\mathrm{ds}} \left(\frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\beta})^2 - \psi(\boldsymbol{\theta}) \right),$$

the fitness measure is then given by:

$$\sum_{i \in T} \sum_{\substack{j \in T \\ j \neq i}} \sum_{k=1}^{N} \sum_{l=1}^{N} \left(\frac{[t(\boldsymbol{\theta}_i, \boldsymbol{\beta}_k) - t(\boldsymbol{\theta}_j, \boldsymbol{\beta}_l)] - \Delta t_{\mathrm{obs},ij}}{\Delta t_{\mathrm{obs},ij}} \right)^2. \tag{4.3}$$

Again, a lower value implies a better fitness of the trial solution.

4.5 Performance considerations

Calculating a reconstructed mass distribution of a lensing system, simulated or real, requires the evolution of a population of genomes, and for each genome one or more fitness values should be calculated. For this reason, it can be expected that the method is relatively slow. To still yield results in an acceptable amount of time (e.g. a few days), this section describes a number of measures that were taken. Appendix E provides a more general overview of what was implemented for this dissertation.

4.5.1 Matrix representation

As was described earlier, the equations to calculate the source positions, the deflection angle derivatives or lens potential values, are linear in the unknown masses M_i of the Plummer basis functions. These equations therefore lend themselves to a representation in matrix form. This makes it easy to use optimized functions to calculate the result of the multiplication of a row of one matrix, with another column vector. In the implementation of the algorithm this operation and others, like multiplying a vector with a constant, were implemented using the *Intel Integrated Performance Primitives* library[1].

When a grid is used to represent the null space or to calculate an approximation of the critical lines, several quantities may need to be calculated for the same points.

[1] `http://software.intel.com/en-us/intel-ipp/`

For example, if the same null space grid is used for two sources at a different redshift, the deflection angle at the points in the grid will be the same. When calculating the positions of these points in the source plane however, different values will be obtained because of different values of $D_{\mathrm{ds}}/D_{\mathrm{s}}$ for these two sources. Since this is a straightforward rescaling of the deflection angles and the deflection angles themselves require a more computationally intensive matrix multiplication, it is advantageous to only calculate the deflection angle once. In the case of the two null space grids, this would mean that only the points of one grid would be used to calculate the deflection angle since the second grid simply uses the same points. Care has been taken in the implementation to make sure that no such duplicate computations occur. This not only reduces the computational load, but also improves the memory requirements.

4.5.2 Floating point precision

Where real-valued numbers are necessary, the calculations in the genetic algorithm happen entirely using 32-bit floating point numbers, instead of using the common 64-bit, "double precision" representation. By doing so, not only does the calculation happen faster, it also requires less memory for storage of the data. Use of 32-bit floating point numbers is possible by scaling the relevant quantities appropriately. As none of the fitness measures depend on the exact scale used, for example in the positional fitness measure the size of the source is used as a length scale, this can easily be done.

4.5.3 Distributed computation

All the genomes in a population need to have their fitness measures evaluated before they can be ranked accordingly. The fitness measures of one genome do not depend on the other genomes, so this fitness calculation allows for parallelism. For this reason, the genetic algorithm was written in such a way that the fitness computation can be distributed over any number of CPUs, connected to each other by a computer network. A simple load-balancing procedure was implemented so that faster CPUs receive more genomes to evaluate.

4.5.4 CUDA

Graphical cards that are present in recent computers often contain a so-called Graphical Processing Unit, or GPU, that can be programmed. Furthermore, such GPUs allow a large amount of parallelism. Since the genetic algorithm itself benefits from a parallel implementation, one may wonder if a speedup of the evaluation of genomes can occur if this is implemented on the GPU instead. Unfortunately, trying to use the GPU for this would require some redesign of the genetic algorithm software and even so, implementing the fitness measures above on the GPU may prove to be difficult.

For another aspect of the genetic algorithm however, it was relatively straightforward to use the GPU. More precisely, the CUDA[2] programming language for NVidia GPUs was used for this implementation. In a multi-objective genetic algorithm, the non-dominated sort procedure requires one to evaluate which genomes dominate a specific genome, and which genomes are dominated by this genome. Each genome should therefore be compared to every other genome, with respect to each fitness measure. For a larger population size (e.g. ten thousand genomes), this step itself can become the bottleneck in the genetic algorithm, rather than calculating the fitness values. However, comparing each pair of genomes to each other is something that can easily occur in parallel. Although in principle this can be distributed over a number of CPUs, in this case a GPU-based implementation was done. For large population sizes, this reduced the time spent in this step of the algorithm by about a factor of four.

4.6 Simulations

Although many galaxy lenses are currently known and even more and more cluster lenses are detected, inverting such a lens does not provide one with objective information about how similar the reconstruction is to the true mass distribution, as the latter remains unknown. For this reason, simulations are used to find out how well the algorithm works. A projected mass distribution is generated and for a certain set of sources at different distances, the images are calculated using the ray tracing procedure. Then, using the images of each source as input, the algorithm tries to find a compatible mass map and at the same time, the source positions are predicted. By comparing the reconstructed mass density to the one used to calculate the images, one immediately obtains a grasp on which features of the lensing mass have been recovered.

4.6.1 Many extended images

In the first tests, many sources producing many images were used. The mass distributions of the lenses in these simulations were created by randomly adding a number of Plummer distributions. The number of sources, their positions and redshifts were also chosen at random. A wide variety of these gravitational lens systems were used to test the algorithm. In the example that we present below, a lens with mass of the order of $10^{15}\,M_{\odot}$ was positioned at $z = 0.45$ while the redshifts of the sources were sampled from a uniform distribution in the interval $[1.2, 4.0]$.

In this example, the relevant angular diameter distances were calculated using a standard CDM cosmology, with a matter density $\Omega = 1$ and a Hubble parameter $H_0 = 70$ km s^{-1} Mpc^{-1}. For these settings, the angular diameter distance between

[2] `http://www.nvidia.com/object/cuda_home.html`

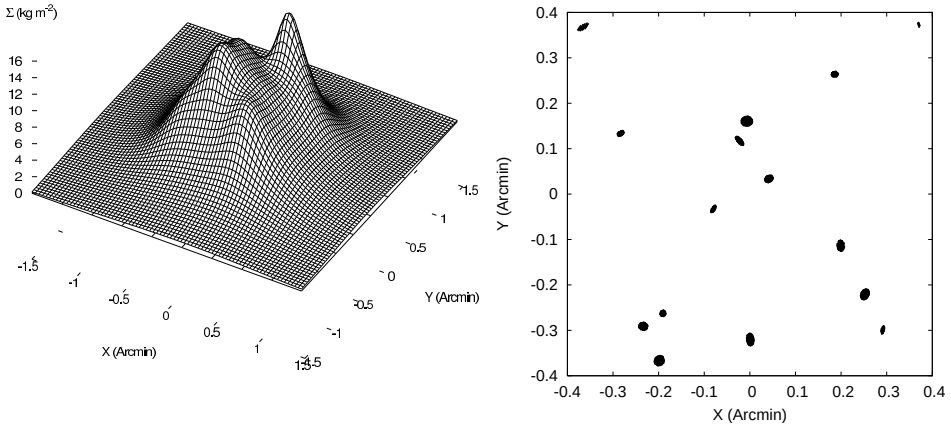


Figure 4.12: Left panel: the mass distribution of the input lens used in the simulation. The total lens mass within a radius of 1.5 arcmin, which is slightly further out than the position of the outermost image, is $0.95 \times 10^{15}\, M_\odot$. Right panel: the positions and shapes of the 15 sources used in the simulation within the source plane.

an observer at redshift z_1 and a source at redshift z_2 is given by

$$D(z_1, z_2) = \frac{2c}{H_0} \frac{1}{1+z_2} \left(\frac{1}{\sqrt{1+z_1}} - \frac{1}{\sqrt{1+z_2}} \right).$$

In figure 4.12, we show the mass distribution and the positions and shapes of the sources. The total mass of the lens within a radius of 1.5 arcmin, which is slightly further out than the position of the outermost image, is $0.95 \times 10^{15}\, M_\odot$ and the number of sources in this simulation is 15. This configuration was used to generate the images shown in the left panel of figure 4.13, which in turn serves as input for the inversion algorithm. The resolution of this image is 1024×1024 pixels. Critical lines and caustics for a source at redshift $z = 2.5$ are presented in the right panel of figure 4.13, in which the source positions are also indicated. The genetic algorithm then constructs a lens solution that projects images of a single source onto overlapping regions in the source plane, using only the overlap fitness measure as described above. For this particular simulation, the fitness converged for a grid containing about 400 Plummer mass distributions. As explained before, multiple applications of the inversion algorithm yield different solutions. Still, each solution manages to produce overlapping back-projected images and, while this is in no way enforced by the algorithm, the positions of the back-projected images are very close to the true source positions.

After applying the inversion routine 25 times and averaging the individual solutions, we obtained the final solution presented in the left panel of figure 4.14. This figure shows a striking resemblance to the left panel of figure 4.12. Clearly, the mass distribution of the lens is retrieved with very high accuracy. The fitness values of the

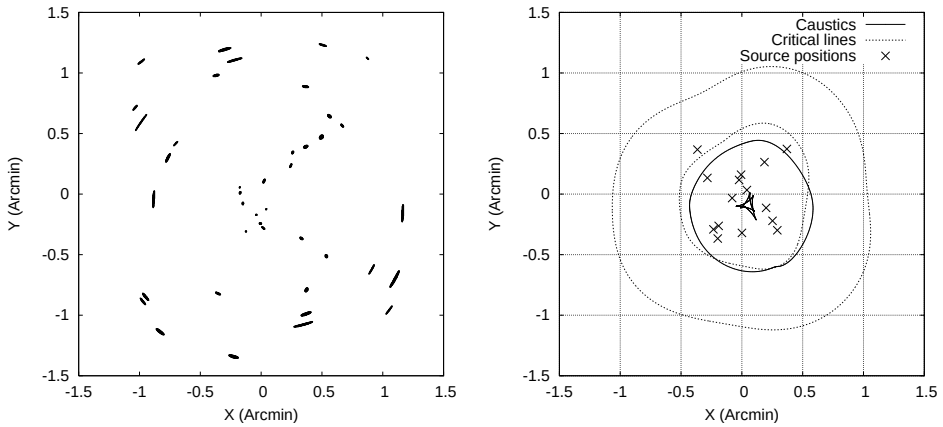


Figure 4.13: Left panel: the images used as input for the inversion routine in the simulation. The resolution of this image is 1024×1024 pixels. Right panel: the critical lines (dotted lines) and caustics (full lines) of the lens for a source at redshift $z = 2.5$. The crosses indicate the positions of the sources.

25 individual solutions and of the averaged solution are shown in the left panel of figure 4.15. Since averaging the individual solutions suppresses random generation-to-generation fluctuations, which can even prevent the solutions from further lowering the fitness value, and enhances their common traits, the averaged solution even outperforms each individual solution. When the images of figure 4.13 are projected back onto the source plane, we obtain the situation shown in the right panel of figure 4.14. The back-projected images overlap very well and are close to the true source positions. The critical lines and caustics of the averaged solution for a source at redshift $z = 2.5$ are shown in the right panel of figure 4.15, which can be compared with the right panel of figure 4.13. Again, the resemblance is striking. In the left panel of figure 4.16, we show the absolute value of the difference between the mass distributions of the input lens and of the averaged solution. In the right panel of figure 4.16, the standard deviation of the 25 individual solutions is presented. The first quantity is a measure for the quality of the fit, the second measures the disagreement between the individual solutions.

For a circularly symmetric lens, only the total mass enclosed within the radius of the outermost image can be determined. The lens employed in the simulation is not spherically symmetric but one can still surmise that we do not have a very good handle on the mass outside the outermost image. In figure 4.17, we show the circularly averaged density profiles of the input lens and of the averaged solution. As expected, both agree excellently with each other within the inner ~ 1.5 arcmin, which is about the position of the outermost image. Outside that radius, the density is no longer well constrained by the data and the profile of the best lens solution drops below that of the input lens. For the averaged solution, the mass enclosed within a radius of 1.5

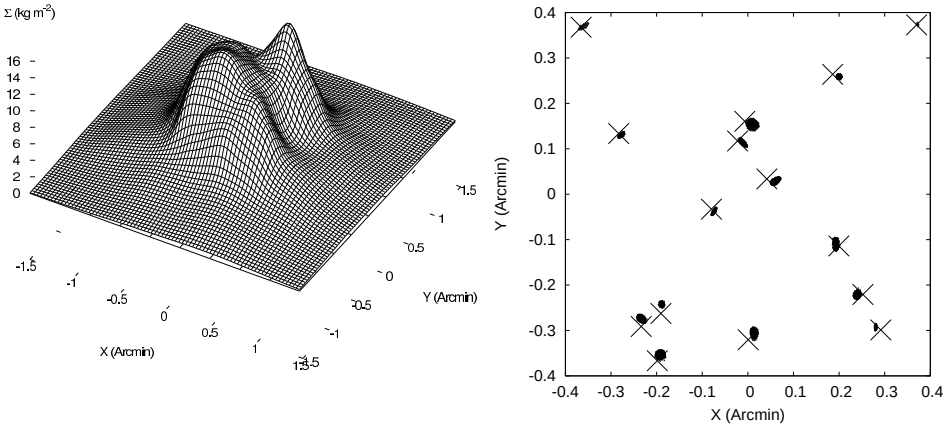


Figure 4.14: Left panel: the average of 25 individual solutions for the simulation. The total mass of this averaged lens solution within a radius of 1.5 arcmin is $0.96 \times 10^{15}\, M_{\odot}$. This figure can be compared with the left panel of figure 4.12. Right panel: the positions of the back-projected images within the source plane for the averaged solution of the simulation. The true source positions are marked with crosses. This figure can be compared with the right panel of figure 4.12.

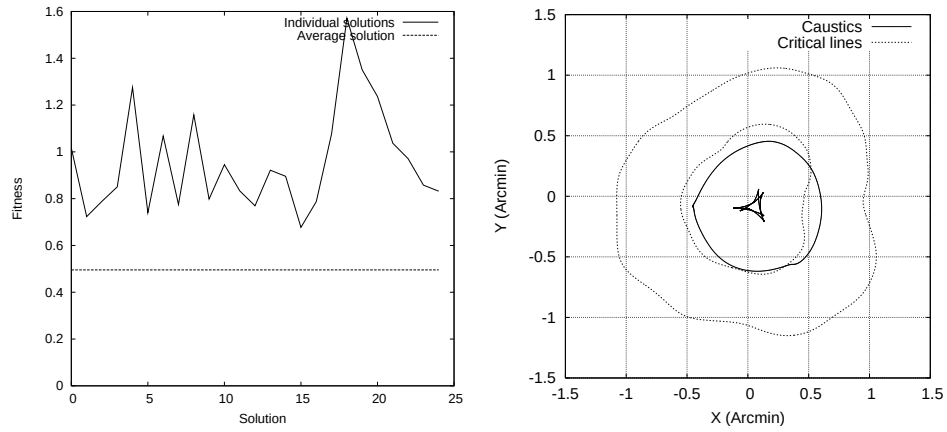


Figure 4.15: Left panel: the fitness values of the 25 individual solutions compared with the fitness value of the averaged solution for the simulation. The averaged solution is clearly superior to the individual reconstructions. Right panel: the critical lines and caustics for a source at redshift $z = 2.5$ of the averaged solution in the simulation. This figure can be compared with the right panel of figure 4.13.

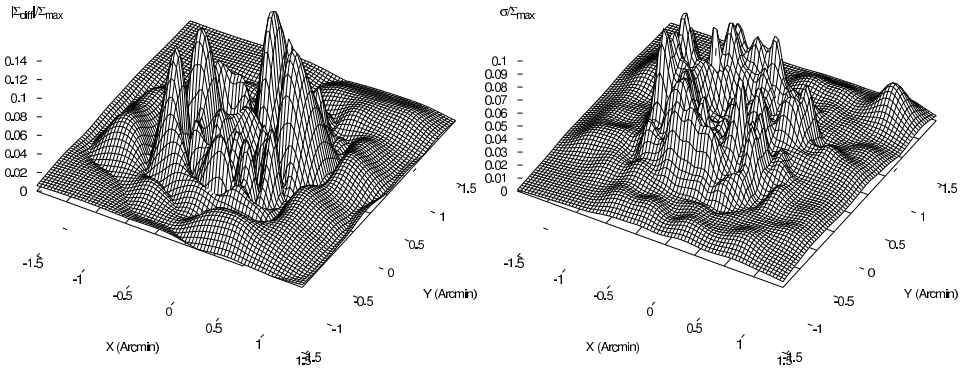


Figure 4.16: Left panel: the absolute value of the difference between the mass distributions of the input lens and of the averaged solution of the simulation, relative to the maximum mass density of the input lens. Right panel: the standard deviation of the 25 individual solutions of the simulation, relative to the maximum mass density of the input lens.

arcmin is $0.96 \times 10^{15}\, M_\odot$ which can be compared with the input lens, which comprises a mass of $0.95 \times 10^{15}\, M_\odot$ within the same radius.

The implementation used for this simulation employs a population of 250 genomes. The calculations were done in a distributed manner, using sixteen Intel® Xeon$^{\mathrm{TM}}$ 2.4 GHz processors of a computer cluster. Depending on the number of sources, creating a single solution may require several hours. To give a specific example, the 25 solutions used in the simulation were created in four days.

4.6.2 Few sources and null-space

If many sources are available, the previous example has shown that much of the lens mass distribution can be expected to be recoverable. While such data is currently available (e.g. [10]) and more of these systems are likely to be discovered in the future, they are currently far outnumbered by systems containing a handful of images of only one or a few sources.

As explained before, if the original procedure was used to determine the mass distribution of the lens, extra images were predicted, leading to the identification of the null space as an additional source of information. In the example below, we use the sources and the lens mass density shown in figure 4.18. The sources have an elliptical shape and are positioned at $z = 2.5$ and $z = 1.5$ respectively; the lensing mass is located at $z = 0.45$. Again, we use a standard CDM cosmology with a matter density $\Omega = 1$ and a Hubble parameter $H_0 = 70$ km s^{-1} Mpc^{-1} for simplicity. As can be derived from the caustic structures, each source produces five images, the positions of which are displayed in the left panel of figure 4.10. The critical lines corresponding to both redshifts are also indicated in this figure. The images of the source at $z = 1.5$ are surrounded by small squares.

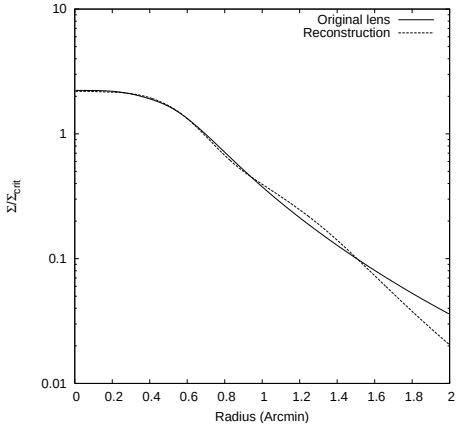


Figure 4.17: The circularly averaged density profiles of the input lens (full line) and of the averaged solution (dotted line), normalized to the critical density for a source at redshift $z = 2.5$. Within a radius of 1.5 arcmin, which coincides with the position of the outermost image, both profiles agree very well. Outside this radius, the density of the averaged solution is not constrained very well by the data and drops below that of the input lens.

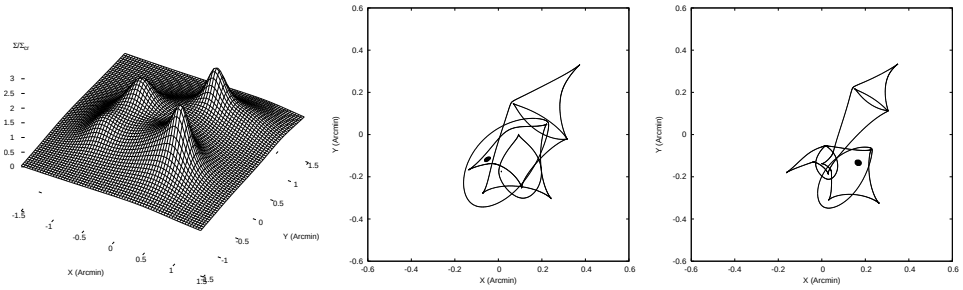


Figure 4.18: Left panel: mass distribution of the input lens (placed at redshift 0.45) used in the first simulation of the few sources situation. The value of Σ_{cr} was evaluated at redshift 2.5. Middle panel and right panel: the two sources, placed at redshifts 2.5 and 1.5 respectively, used in the first simulation, together with the caustic structures at these redshifts.

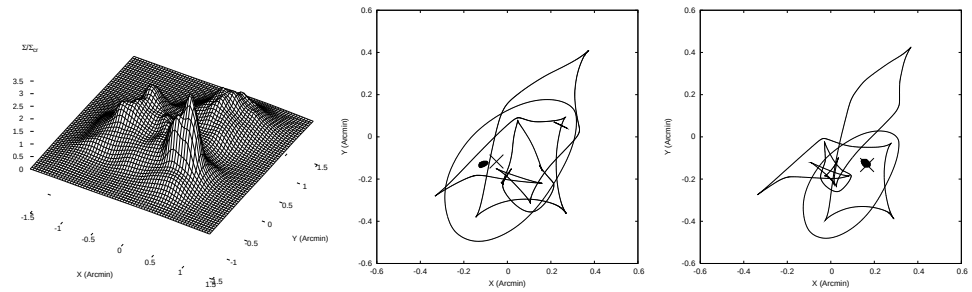


Figure 4.19: Left panel: when the null space is taken into account, this mass distribution was obtained after averaging twenty individual solutions. Comparing this to the left panel of figure 4.18 shows that much of the general appearance is retrieved. Middle panel and right panel: the reconstructed sources lie close to the true source positions (indicated by crosses). When these figures are compared to the middle and right panels of figure 4.18, it is clear that much of the same caustic structure is present in the reconstruction.

Using the positional and null-space fitness measures, it is now straightforward to generate solutions which indeed only produce the input images. Averaging twenty such solutions yields the mass density and reconstructed sources shown in figure 4.19. Comparing this figure with figure 4.18, one immediately notices the striking resemblance. This proves that our method is capable of reproducing, at least qualitatively, the mass density distribution of a gravitational lens based on the positions, redshifts, and shapes of very few images and on null space information. The peaks in the reconstructed mass density appear to be somewhat stronger than those of the input lens. The reconstructed sources are very similar in shape to the true sources and their positions lie very close to those of the input sources. The reconstructed sources are, however, more extended than the input sources. The caustic structure presented in the middle and right panels of figure 4.19 are strikingly similar to but also more extended than those of the input lens, presented in the middle and right panels of figure 4.18. In this example the method using null space triangles was used, based on a regular 64×64 grid, covering an area of 3.3×3.3 arcmin^2.

It is interesting to take a look at the difference in mass densities between original and reconstructed lens. The left panel of figure 4.20 shows that the mass density around the peak positions is not reconstructed very well. Inspecting the right panel of the same figure, which displays the standard deviation of the individual solutions, it is clear that precisely these regions differ strongly among the solutions. When the reconstructed source and lens are used to reproduce the images, the result shown in the left panel of figure 4.21 is obtained. The ten input images are indeed reproduced and the critical lines closely resemble those of figure 4.10 (left panel). One cannot ask more from any lens inversion algorithm. The right panel of figure 4.21 shows the circularly averaged mass density, centered on the mass density peak at $(0.5, -0.5)$. This plot visualizes once more the overestimated mass density in that region, although the same general features are clearly present in both original and reconstructed profile. When

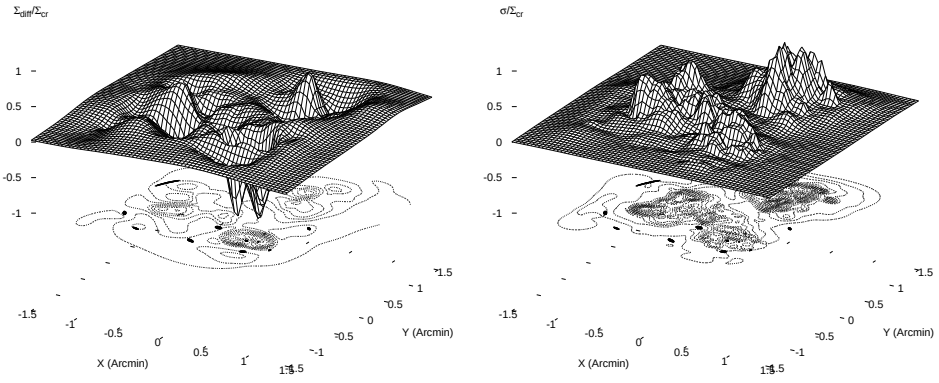


Figure 4.20: Left panel: the difference between the true mass density and the average solution, relative to the critical mass density for a source at redshift 2.5. This clearly shows that the true shape of the mass density peaks cannot be determined accurately. Right panel: standard deviation of the twenty individual reconstructed mass densities. Individual solutions clearly disagree about the exact shape of the mass density around the peak positions.

the total mass of the lens is calculated, the agreement is excellent within a radius of 1 arcmin, differing by only a few percent from the true mass, as is shown in figure 4.22. Beyond that radius, the difference starts to increase, again indicating that using only the strong lens effect, one cannot obtain a firm handle on the mass density outside the radius of the outer images.

One might argue that a system with two sources and ten images still provides a relatively large number of constraints. For this reason, let us now turn to a simple five-image system, created by an elliptical mass distribution. The source and its corresponding images are depicted in the left panel of figure 4.23. The redshifts of the source and the lens are 2.5 and 0.45 respectively.

In this case, the inversion algorithm seems to easily produce solutions causing a critical line to intersect the two rightmost images. To avoid this, the brightness overlap and positional overlap were used as two separate fitness measures. Since their combination can be seen as a weighting scheme, decoupling them allows a broader search in the model space. Afterwards, the genetic algorithm was indeed able to consistently find solutions which reproduce the input images. After averaging ten solutions obtained using a moderately subdivided grid, the results shown in the center panel of figure 4.23 were obtained. This clearly resembles the true situation, although the source is again larger. If the grid is subdivided further, this results in the situation depicted in the right panel of figure 4.23, showing more complex critical lines and corresponding caustic structure.

These examples clearly show the effect of the mass sheet degeneracy, as the reconstructed sources are in each case larger than the input sources. In fact, in these examples the degeneracy was broken artificially: the algorithm did not yet support

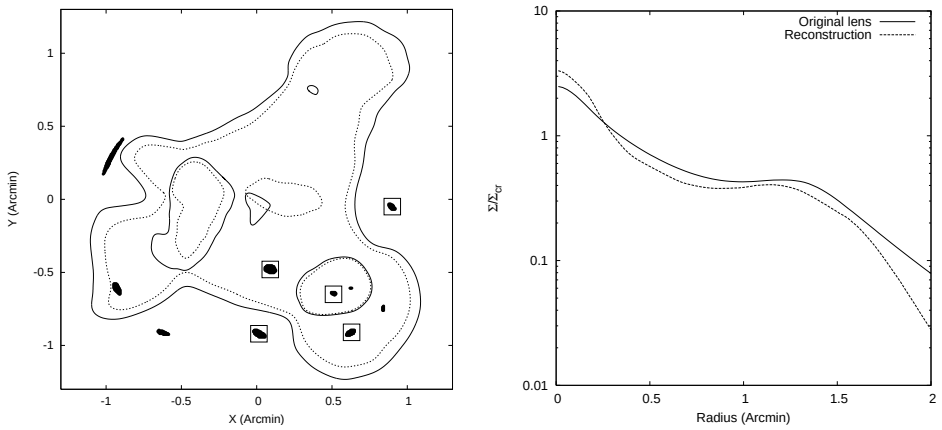


Figure 4.21: Left panel: the reconstructed source and lens of figure 4.19 generated these images and critical lines. The correspondence with the left panel of figure 4.10 is striking. Right panel: the circularly averaged mass densities of input lens and reconstruction, as seen from the mass density peak at $(0.5, -0.5)$.

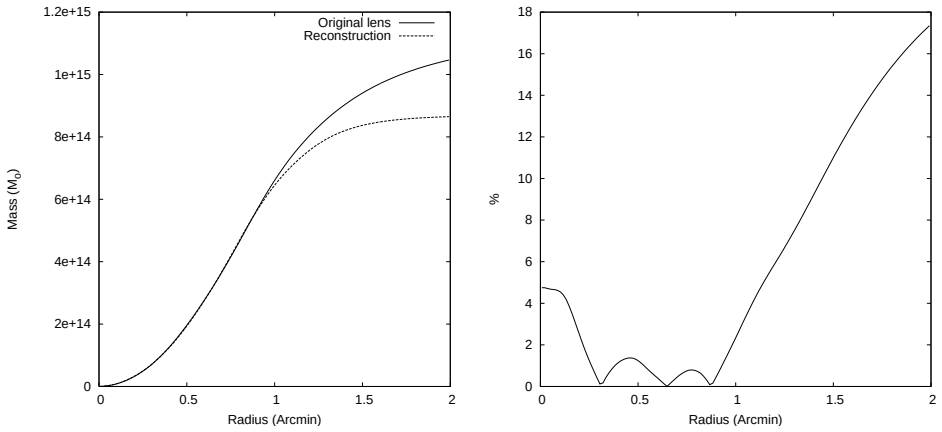


Figure 4.22: Left panel: total mass within a specific radius for both input lens and reconstruction, as seen from the origin. Clearly, the total mass is estimated well within a radius of 1 arcmin. Right panel: the amount of disagreement in total mass within a specific distance from the origin. Again, within a radius of 1 arcmin the agreement is very good.

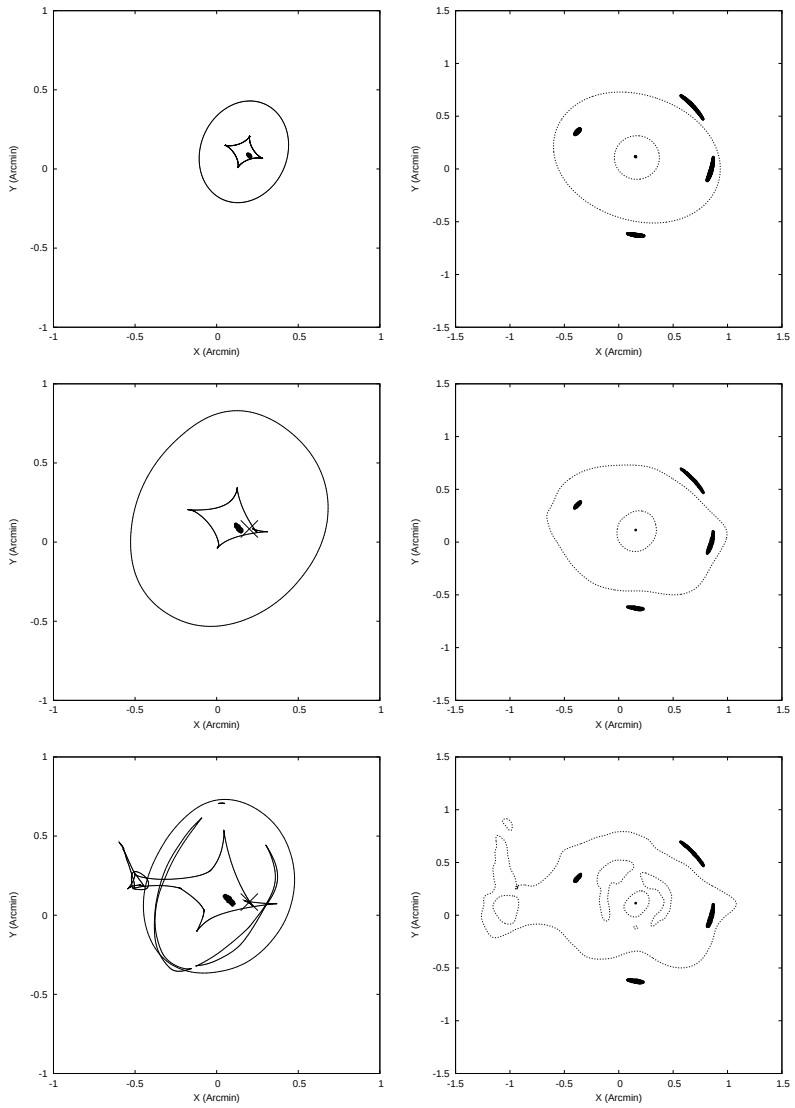


Figure 4.23: Top row: the source used in the simulation of a gravitational lens system with an elliptical mass distribution, relative to the caustic structure. This situation creates five images, which are used as the input of the inversion routine. Center row: the reconstructed source, relative to the predicted caustics when using a moderately subdivided grid and averaging ten individual solutions. Note that the source and the caustic structure are more extended than their original counterparts. Bottom row: similar to the center row, but using a finely subdivided grid. The fitness measures did not improve considerably compared to the situation in the center row.

a mass sheet basis function at the time, so the mass distribution was assumed to be zero outside the region in which the algorithm would search. However, the input mass maps in these examples, being composed of relatively wide Plummer functions and a NSIE distribution respectively, still contain some mass outside this region. The presence of the mass sheet degeneracy can also be seen in the right panel of figure 4.21.

4.6.3 Point-like images & mass sheet basis function

That the point-based fitness measure yields good results can be seen in figure 4.24. The input mass distribution can be seen in the left panel of this figure and consists of three NSIE mass distributions. The reconstruction can be seen in the right panel of this figure, and is the average of 30 individual solutions. In this inversion procedure, the algorithm was allowed to use a mass sheet basis function, but in the end result this is negligible. The left panel of figure 4.25 shows the true source positions as crosses, together with the reconstructed source positions. It should be noted that an offset of $(5.5, -3)$ arcsec was added to the reconstructed sources for easy comparison. Except for this offset, which is of course not observable, the reconstructed source positions are remarkably accurate. This indicates that the scale of the source plane is determined correctly, meaning that in this case the available data successfully break the mass sheet degeneracy. The importance of taking the overall scale of the source plane into account can be seen in the right panel of this figure. If an inversion is done using an absolute distance scale, an incorrect mass sheet is recovered. Especially in the beginning of the genetic algorithm, using an appropriate scale will be important as projecting the points onto a smaller general area by means of a considerable mass sheet will otherwise lead to an improved fitness measure. This can cause the genetic algorithm to get stuck in a local optimum.

A variation of this mass distribution shows how the presence of a mass sheet can be recovered, but that the mass sheet basis function is necessary to do so. Figure 4.26 shows the true mass distribution in the left panel, consisting of a scaled version of the mass distribution in figure 4.24, to which a sheet of constant density was added. When only Plummer basis functions are used, the result in the center panel is obtained. Clearly, this does not look like the input mass distribution at all. There are two causes of this undesirable behavior. First, the algorithm will have to try to mimic the effect of a mass sheet using the Plummer basis functions which is a rather difficult task, depending on the amount of constraints available. The second problem is that the subdivision scheme will be less effective. Since the mass sheet holds most of the mass, the subdivision procedure will not be successful at refining the grid in the central region. When the algorithm is instructed to not only use the Plummer basis functions, but a mass sheet basis function as well, the result is much improved as can be seen in the right panel of the figure.

In the previous examples, the input mass distribution and images were self-generated, so it is possible that some bias is still present in the reconstructions. It would certainly be interesting to see how well the algorithm performs on data generated by somebody

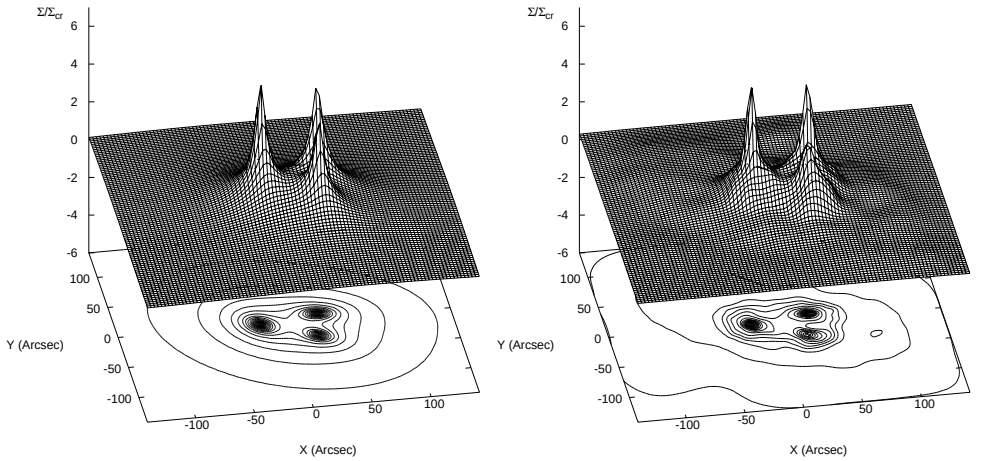


Figure 4.24: Left panel: mass distribution of the input lens, used to test the algorithm that works with point images. This lens is actually a superposition of three NSIE mass densities. Right panel: the reconstructed mass map, which is the average of 30 individual reconstructions, shows much resemblance to the input mass distribution of the left panel.

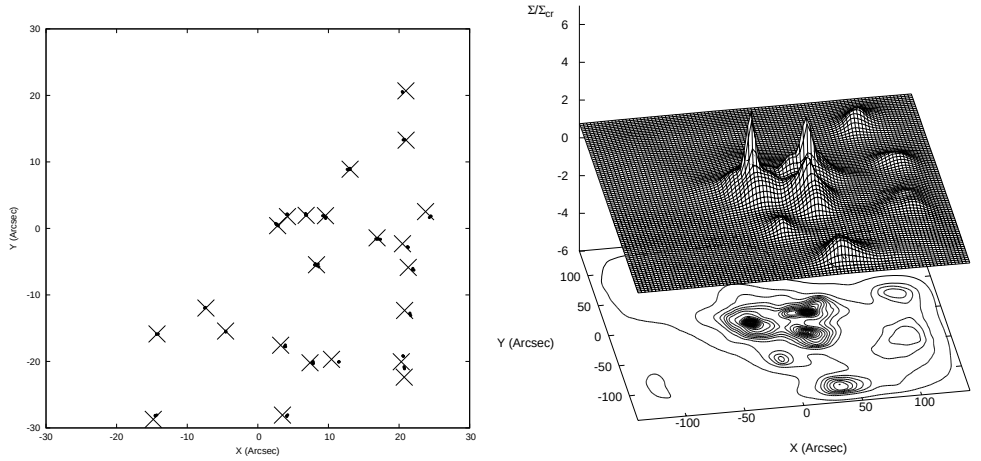


Figure 4.25: Left panel: comparison of the input source positions and the reconstructed source positions, i.e. the back-projected point images. Apart from the offset of $(5.5, -3)$ arcsec which was added to the reconstructed positions, the agreement is very good. Right panel: when the scale of the source plane is not used in calculating the overlap between the back-projected points, but an absolute distance scale is used instead, this is the result. Clearly, an incorrect mass sheet basis function has been recovered.

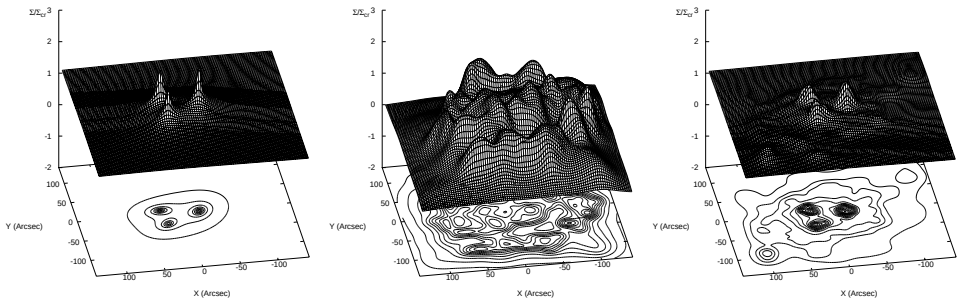


Figure 4.26: Left panel: true projected mass density of a lens used to test the inversion procedure. The mass distribution consists of a few relatively small perturbations on top of a sheet of mass. Center panel: when the original procedure is applied to the images produced by the input lens, it is not successful in creating an acceptable mass map (see text). Right panel: when a sheet of mass is added as a basis function, the algorithm again is able to create acceptable reconstructions of the projected density.

else. The author of the LensPerfect software, Dan Coe, made the input images and mass map of one of his own tests available. The input mass map is actually a model of the lensing cluster Abell 1689, and can be seen in the left panel of figure 4.27. The envelope of the outermost images used in the inversion can be seen as a thick black line. One can therefore only hope to constrain the mass distribution inside this region. The result that was obtained with the LensPerfect method can be seen in the right panel of this figure, and clearly shows much resemblance to the input mass map inside the region covered by the images.

The result of the point-based inversion procedure is shown in figure 4.28. For this test, the region in which mass should be recovered was chosen relatively large compared to the size of the images. To suppress unnecessary fluctuations, the null space fitness measure was included as well. The result shown in this figure arises from taking the average of 40 individual reconstructions. The left panel of the figure shows the obtained mass map as a contour map, in which the positions of the images are indicated as well. The right panel of the figure shows the same result, now using the same color scale as the LensPerfect images for easy comparison. Clearly, the method described here also performs well on these data, at least when only the region within the thick black line is considered. The fact that outside that region the density differs more and more as one looks further away from the center, is due to the use of the mass sheet basis function.

4.7 Automatic redshift determination?

One may wonder if it is also possible to let the genetic algorithm determine automatically the redshift of one or more sources, if this data is unknown. For this purpose,

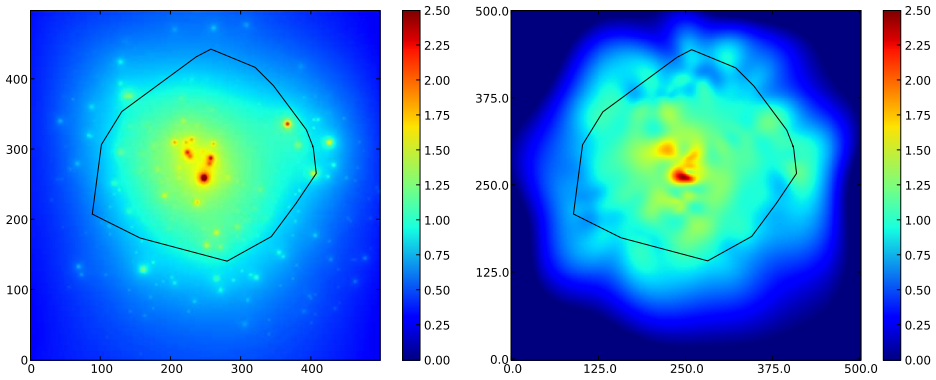


Figure 4.27: Left panel: input mass map used to test the LensPerfect lens inversion method. The thick black line indicates the envelope of the outermost images. Right panel: the mass distribution recovered by the LensPerfect method. Inside the region marked by the black line, there is clearly a good resemblance (images from [19]).

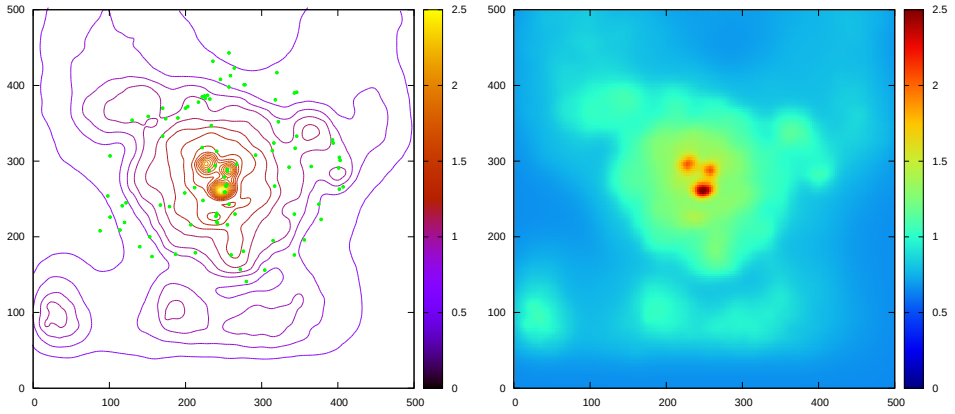


Figure 4.28: Left panel: after averaging 40 individual solutions, this is the obtained mass map if the point-based fitness measure and null space fitness measures are used. The locations of the input images are indicated as well. Right panel: the same result as in the left panel, now using the same color scheme as in figure 4.27. The method described here clearly also provides a good recovery off the mass distribution inside the region bounded by the outermost images.

the genetic algorithm was modified so that the $D_{\mathrm{ds}}/D_{\mathrm{s}}$ fraction of the sources with unknown redshift was also present in each genome, and correspondingly could be optimized.

The results were disappointing however. Even when the images of a source with unknown redshift were located close to images with a known redshift, it was difficult to predict if the determined distance fraction would be recovered accurately.

—5—

Applications

The previous chapter explained how a genetic algorithm is used to derive a mass map that is compatible with a given set of observations. In this chapter, the inversion procedure shall be applied to some existing strong lensing systems.

5.1 Cl 0024+1654

5.1.1 Introduction

One of the most spectacular examples of strong gravitational lensing can be seen in the cluster lens Cl 0024+1654, shown in figure 1.1. Using recent observations with the Advanced Camera for Surveys (ACS) of the Hubble Space Telescope (HST), one can easily see that five well resolved images depict a single source, but even before these five images were identified, it was clear that three arc segments were caused by a gravitational lens effect [60].

This strong lensing information was first used in [52]. The authors of this work noted that these arc segments do not obey the so-called length theorem [64], implying that no simple elliptical lens model can be used. They show that if perturbations by cluster members are added, the observed arc lengths can indeed be reconstructed. In [107], a more advanced reconstruction technique was used, consisting of a smooth lens model perturbed by some smaller galaxies and a non-parametric source model. Whereas previous work suggested that the main cluster potential was offset from the largest galaxy, these authors find that these positions, in fact, agree well.

After the first HST images clearly revealed the presence of five images, more lensing studies followed. The new, well resolved images were used in [20] to study the source

itself, a blue galaxy containing some interesting dark features and a bar-like structure. In [103] the images were used to find the parameters describing elaborate lens and source models. Their algorithm constructs the complete image plane based on a set of source and lens parameters and compares the result with the HST observations. They find that the mass distribution is dominated by a smooth dark matter component with a considerable core radius, centered at a position near the largest cluster member.

Much of the earlier mass uncertainties originated from the poorly established source redshift. In [11] a spectroscopic redshift of 1.675 was finally measured and was used in a new inversion. The authors found that the image positions can be accurately reproduced using a model which traces the locations of the brightest cluster members. In [48], a non-parametric method is used to invert the lens, using both strong and weak-lensing data. In the strong lensing region, the retrieved mass profile closely resembles the result of [11], but according to [99], the associated velocity dispersion is too high to correspond to the measured value of 1150 km s^{-1} [27].

Below, the procedure which makes use of the positional information of the images in the lensing system, the null space and which avoids critical lines to intersect the images shall be used to reconstruct the mass distribution of Cl 0024+1654. No information about the positions of cluster members is used. The content of this section is based on [70].

5.1.2 Input

The images of sources A and B as described in [48] are used, at redshifts of 1.675 and 1.3 respectively. The redshift of the lens itself is 0.395 and angular diameter distances were calculated in a flat cosmological model with $H_0 = 71$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.27$ and $\Omega_\Lambda = 0.73$. The inversion procedure constructs the lensing mass distribution in a square shaped area of 1.3 arcmin by 1.3 arcmin, centered on the brightest cluster galaxy. To avoid predicting images which are located relatively far away, the null space grid measured 3 arcmin by 3 arcmin, centered on the same galaxy. Initially, a uniform grid of 15×15 is used to place the Plummer basis functions on, and the grid is refined until approximately 800 basis functions are used (no mass sheet basis function was used). After this, the finalizing step is executed on a uniform 64×64 grid. Below we shall see that this leads to a very good source reconstruction.

The gravitational lens creates several large images of source A, a blue galaxy. A part of the source is mapped onto five easily identifiable sub-images, as can be seen in figure 5.1. The high resolution ACS images allow several corresponding features to be identified (up to twelve features in some images), which will be used to calculate how well the back-projected images overlap. It is assumed that no other images of the source are present, so that only the five images themselves are excluded from the null space for this source. Source B only has two images, the third one most likely being occluded by the central cluster members. The complete images are used to estimate the source size, but only a single point in each image is used to measure how well the images overlap when projected back onto the source plane (measured relative to

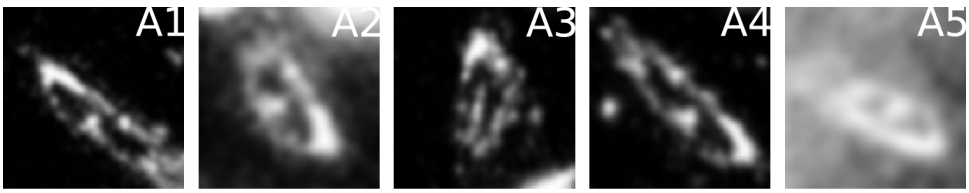


Figure 5.1: The image parts of the five images of source A which were used in the reconstruction, labeled in the same way as in the work [48]. Due to the extended nature of these images, several corresponding features are easily identified. The images shown here are not displayed on the same scale.

the estimated size of the source). In this case, not only the images themselves were excluded from the null space, but also the region in which the central cluster members reside. This allows the algorithm to predict an unobserved third image anywhere in that region. For both sources it is assumed that no critical lines intersect the images which are used. Here, the critical line fitness which penalizes parity changes inside an image is used.

5.1.3 Results

The mass map shown in the left panel of figure 5.2 was obtained after averaging 28 individual solutions. This number is dictated by the computer time it takes to generate the individual solutions and by the fact that after averaging together 15 solutions or more, the average solution does not change significantly. The largest fraction of the rather steep mass distribution coincides with the position of the central cluster members. The central image of source A is located between two density peaks, which resembles the situation shown in the ACS images. These facts can be clearly seen when the retrieved mass contours are drawn on top of the observed situation, as is shown in figure 5.3. This same figures also illustrates the remarkable accuracy with which the two cluster galaxies enclosing the image at $(0.5, -0.3)$ arcmin are retrieved. We would like to stress again that these were retrieved automatically; no prior information about the presence of these galaxies was used. It is these galaxies that cause the middle image of the three arc segments to be compressed, thereby causing the violation of the length theorem. The mass inside a circular region of radius 0.5 arcmin, centered on $(0.075, -0.075)$ arcmin is found to be $1.60 \times 10^{14} M_{\odot}$. This region is enclosed by a dotted line in figure 5.3.

When the input images of source A are projected back onto the source plane, a consistent source is produced, as can be clearly seen in figure 5.4. The size of the source is approximately 2.5 arcsec (corresponding to 21 kpc). This is larger than both the value of 1 arcsec mentioned in [20] and the value of 0.5 arcsec mentioned in [48], but the general appearance does agree very well with both works. We shall come back to this size difference later. The retrieved source positions and caustics

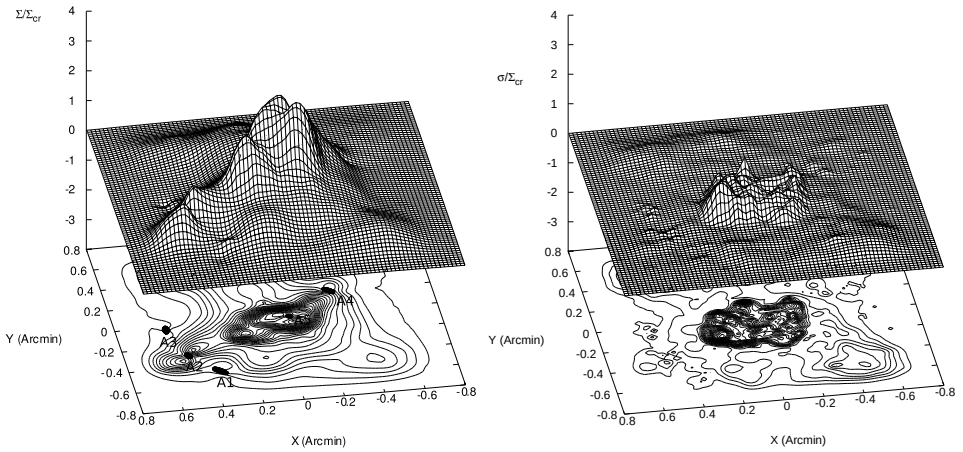


Figure 5.2: Left panel: after averaging 28 individual reconstructions, this is the resulting mass map for Cl 0024+1654 predicted by our procedure. The positions of the input images of source A are also indicated in this figure. The critical density used in this figure corresponds to a redshift $z = 3$. Right panel: the standard deviation of the individual reconstructions shows that the different solutions tend to disagree about the exact shape in the central part of the mass distribution. In particular, this figure suggests that the mass peak around $(0.2, -0.2)$ arcmin in the left panel should not be regarded as an actual feature.

at $z = 1.675$ are depicted in the left panel of figure 5.5. If these sources are used to calculate the image positions, the results shown in the right panel of figure 5.5 are retrieved. From this image it is clear that the multi-objective genetic algorithm succeeded in generating solutions which only predict one extra image (for source B) and which do not have critical lines intersecting the input images.

On closer inspection of the resulting mass map in figure 5.2, there seems to be an intriguing feature at $(0.2, -0.2)$ arcmin. At this location the mass map shows a clear peak, but in the ACS images no cluster member can be seen at this location (figure 5.3). Could this be evidence of dark matter in this cluster? Inspecting the standard deviation of the individual solutions helps to shed some light on this matter. As can be seen in the right panel of figure 5.2, the individual solutions do not agree well on the exact shape of the central part of the mass distribution. In fact, the largest uncertainty is located precisely around the position of this mysterious peak, which suggests that we should be very careful when trying to interpret this feature.

5.1.4 Monopole degeneracy

From the discussion about the monopole degeneracy, we already know that features in between the images can easily be manipulated. It was also explained that using basis functions of this type, one can build a more complex mass distribution that,

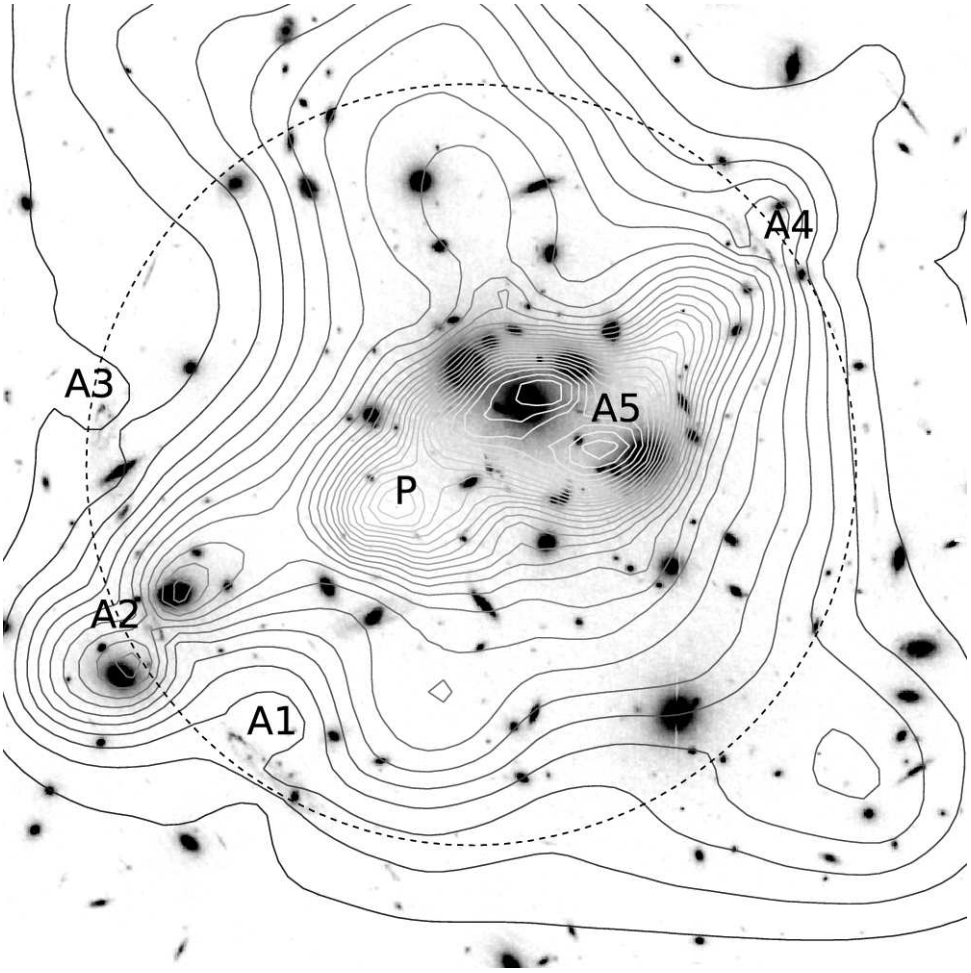


Figure 5.3: The contours of the retrieved mass map in the left panel of figure 5.2 are shown on top of the ACS image of the central cluster region (north is up, east is left). The lensing mass is found to be concentrated around the largest cluster galaxy and the central image A5 is found to be located between two mass peaks, which also resembles the observed situation. The positions of the two galaxies in the south-east region are retrieved very accurately as well. Note that this image also suggests that there is a density peak labeled P in a region where very few cluster light originates. The total mass in the region bounded by the dotted line is found to be $1.60 \times 10^{14} M_{\odot}$. The area displayed in this figure is approximately 1.3×1.3 arcmin^2.

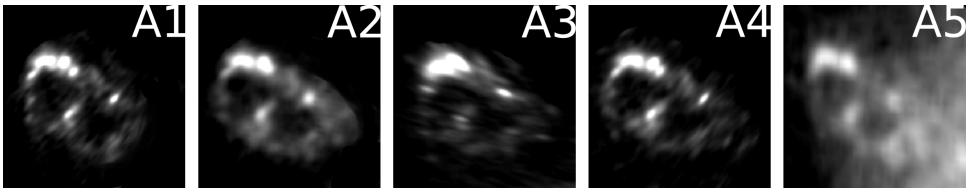


Figure 5.4: When the images shown in figure 5.1 are projected back onto the source plane, these source shapes are retrieved. Each figure shows the same region in the source plane, approximately 3 arcsec by 3 arcsec in size.

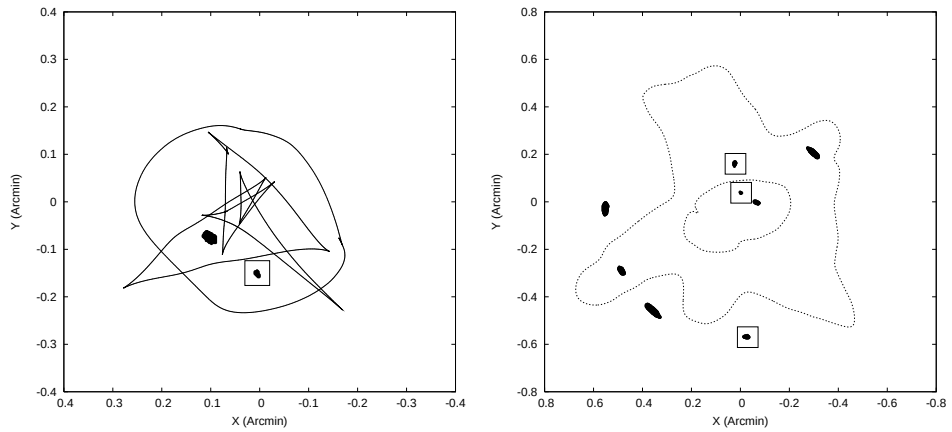


Figure 5.5: Left panel: this figure shows the predicted position of source A and B as well as the caustics corresponding to the redshift of source A. Source B is enclosed by a small square. Right panel: when the sources of the left panel are projected onto their image planes, these image positions arise. The image of source B which is closest to the origin was not part of the input; the model predicts an image at this location. Other than this image, no additional images were predicted.

when added to an existing gravitational lens reconstruction, will produce an equally acceptable solution. Using the recovered mass distribution of Cl 0024+1654 we shall illustrate this below.

The region of interest is subdivided into a number of square-shaped grid cells. For each grid cell, the distance from its center to the nearest image is calculated. If this distance is relatively large compared to the size of the grid cell – e.g. at least four times as large – a basis function is associated to this cell. The distance to the nearest image is used as the unit length; the width of the non-negative part Σ_A is set proportional to the size of the grid cell. This implies that for a specific basis function, all the images lie in the area within which the total mass of the basis function is zero. Since the lens equation for a circularly symmetric basis function only depends on the total mass within a specific radius, in this case the lens equation at the position of the images will be unaffected when such a basis function is added to the existing mass distribution. Similarly, when all the basis functions on the grid are considered, the lens equation at the location of the images will not be influenced, independent of the precise weight values of the basis functions. Everywhere else, the lens equation will indeed be modified, meaning that extra images may be predicted, depending on the precise weight values used.

In the case of Cl 0024+1654, it then becomes immediately clear that the peak at $(0.2, -0.2)$ arcmin can easily be removed by creating a degenerate solution. Even by adding a single basis function with an appropriate width and height to the existing solution, the feature can be eliminated. It can also automatically be removed using the grid-based procedure described above, as can be seen in the left panel of figure 5.6. In this example, a 32 by 32 grid was used, and the weights were determined by a genetic algorithm. The goal of the optimization was to keep the gradient of the resulting mass map as low as possible. To obtain a smooth result, the procedure was repeated for twenty of such grids, each with a small random offset. As can be seen in the figure, this does not only remove the peak at $(0.2, -0.2)$ arcmin, but also reduces the overall steepness. Also note that one of the peaks between which the central image of source A originally resided, has been erased almost entirely. The resulting mass map, consisting of one smooth component and two perturbing components, at least qualitatively resembles the models used by [52] and [107]. The right panel of figure 5.6 shows the critical lines at the redshift of source A, as well as the images predicted by the new solution. Because this newly created solution does not modify the lens equation in the regions of the images and because no extra images are created, the fitness values are exactly the same as those of the solution in figure 5.2. For this reason, both mass maps are equally acceptable solutions.

5.1.5 Discussion

One of the recent gravitational lensing studies of this lens, is that of [48], which used both strong and weak lensing data. The strong lensing mass of $1.60 \times 10^{14} M_{\odot}$ is less than the value of $(1.79 \pm 0.13) \times 10^{14} M_{\odot}$ found in their study, but it is still in

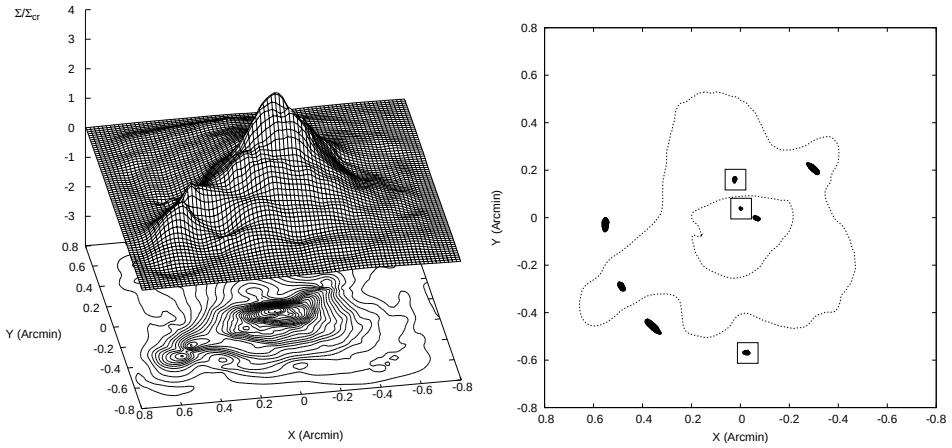


Figure 5.6: Left panel: if the grid based method to redistribute mass is applied to the mass map shown in the left panel of figure 5.2, this new distribution is obtained. The peak at $(0.2, -0.2)$ arcmin has automatically been removed and the overall distribution has become less steep. Right panel: the mass distribution in the left panel predicts the images shown in this figure, which are indistinguishable from the images in figure 5.5 (right panel). The critical lines on the other hand, do display some changes, reflecting the modifications to the lens equation.

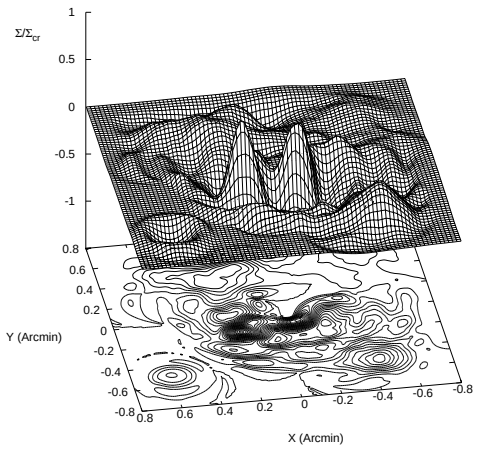


Figure 5.7: This plot shows the differences between the mass distributions in figure 5.2 and figure 5.6. Clearly, the structure has been altered in a way which does not display any particular symmetry.

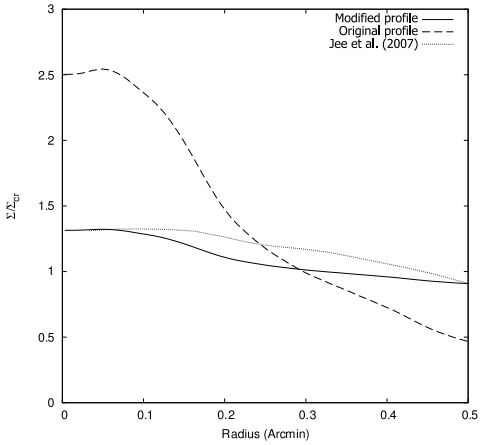


Figure 5.8: The density profile in the circular region indicated in figure 5.3 is described by the dotted curve. If this profile is scaled down by a factor of five and a mass sheet is added to keep the strong lensing mass constant, the the profile described by the thick black line is obtained. In the strong lensing region, it clearly resembles the profile shown in the work of [48], suggesting that the results shown here differ mainly by the mass-sheet degeneracy. This figure also clearly shows that the strong lensing mass estimate from this work differs from the one in [48].

good agreement. We mentioned earlier that our size estimate for source A is higher than found in other works. This is a well-known consequence of a generalized version of the mass sheet degeneracy. As was discussed earlier, this degeneracy is hard to break for lensing systems with only a handful of sources, even if these have different redshifts. Like the original mass sheet degeneracy, the generalized degeneracy leaves the observed images identical but the reconstructed sources are scaled versions of the original ones while the density profile of the lens becomes less steep.

The relation with the inversion of [48] can be revealed by comparing the predicted source sizes. The size of source A in our inversion is five times larger than in their work, thereby identifying the scale factor in the mass-sheet degeneracy. When we downscale our mass reconstruction by a factor of five and add a constant sheet of mass in such a way that the total strong lensing mass of $1.60 \times 10^{14} M_{\odot}$ is unaffected, the circularly averaged density profile in figure 5.8 is obtained (thick black line). This clearly shows much resemblance to the profile found in [48] in the strong lensing region. Note that since our reconstruction procedure only looks for mass in a region which is $1.3\times$ 1.3 arcmin^2 in size, the profile will quickly drop to zero beyond the range shown in the figure.

After this inversion was performed, new multiply imaged systems were identified in [114]. Not only does this work reveal the presence of the third image of source B close to the position predicted by the model shown here, the mass profile therein closely resembles the profile of figure 5.8.

When the monopole degeneracy was applied to the case of Cl 0024+1654, a simple optimization routine was used to remove substructure from the previously obtained mass distribution. However, there is no general rule as to how the mass map may be modified. For example, with some extra effort the existing mass map could have been transformed into one which followed the light more closely, or which corresponded better to the available X-ray data [80]. The only constraints which matter in this respect are the absence of unobserved images and possibly dynamic measurements. Image positions, fluxes and time delays are completely unaffected by this type of degeneracy, which allows you to redistribute matter in any number of ways. This freedom is illustrated in figure 5.7, which depicts the differences between the two mass distributions shown here.

5.2 SDSS J1004+4112

5.2.1 Introduction

The lensing cluster SDSS J1004+4112, located at a redshift of 0.68, can be seen in figure 5.9 and was revealed by the presence of a multiply imaged quasar as reported by [47]. The lensing system was first identified as a quadruply imaged quasar, but later a fifth central image of the quasar was detected [45] and spectroscopically confirmed [44]. Three multiply imaged galaxies were identified in HST/ACS images in [100] and time delay information for three of the quasar images was measured by [33], improving the earlier reported time delay between the two closest quasar images in [34]. This work did not only only invalidate earlier proposed models of the lensing system (e.g. [78], [110]), which predicted shorter time delays, it also confirmed that microlensing is the cause of the strange magnification patterns in the quasar images, present both in optical [87] and X-ray [65] measurements. With its separation of 14 arcsec, the multiply imaged quasar in SDSS J1004+4112 has held the record for being the widest lensed quasar for a number of years. The discovery of SDSS J1029+2623, a multiply imaged quasar with a separation of over 22 arcsec [46] broke this record recently. The statistics of multiply imaged quasars by clusters are studied in [41].

5.2.2 Input

Figure 5.10 shows the image systems that were used in the inversion of SDSS J1004+4112, using the same labeling as [100]. There are five spectroscopically confirmed images of a quasar at redshift 1.734, labeled Q1-Q5. Corresponding to the time delay measurements of [33], we used a time delay of 40.6 days between Q2 and Q1, and a time delay of 821.6 days between Q3 and Q1. No magnification information was used, as the quasar image magnifications are influenced by microlensing, introducing a large uncertainty. The positions of the quasar images were set to those reported in [45]. Four, possibly five images are present of a galaxy at redshift 3.332, labeled A1-A5, with image A5 being marked as uncertain by [100]. The third system used consists of

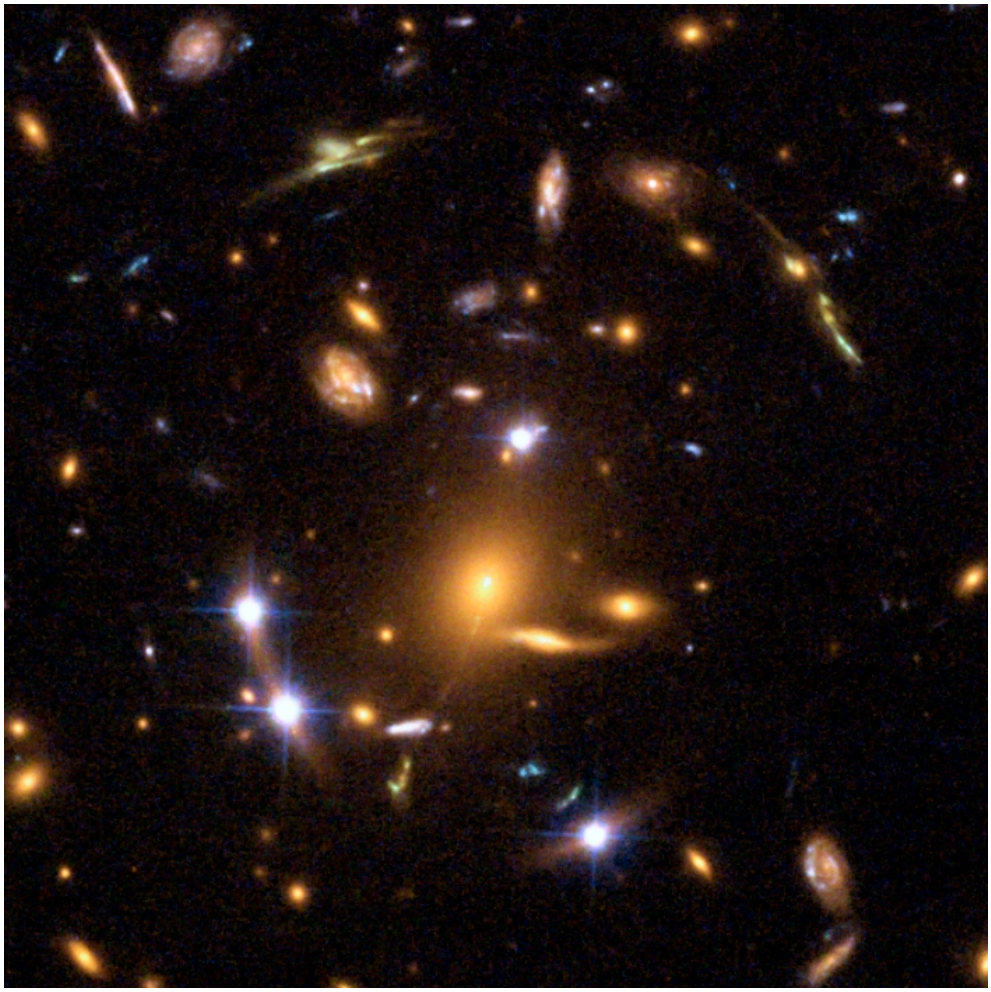


Figure 5.9: Image taken with the Hubble Space Telescope of the lensing cluster SDSS J1004+4112. Around the central elliptical galaxy, one can easily spot four bright images of the same quasar. As can be seen in figure 5.10, other multiply imaged systems exist as well. The color image was taken from the Space Telescope Science Institute (STScI) archive[a].

[a] http://hubblesite.org/newscenter/archive/releases/2006/23/image/a/

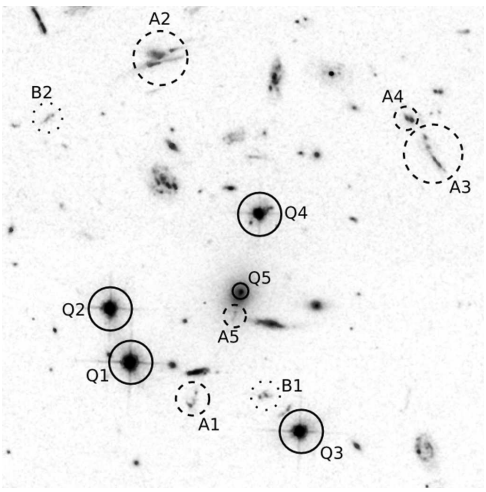


Figure 5.10: The multiple image systems which are used in the inversion of SDSS J1004+4112, using the same labeling as [100] (north is up, east is left). Five images of a quasar (Q1-Q5) are available, as well as four, possibly five images of a galaxy marked A1-A5, and two images of a second galaxy marked B1-B2. Between B1 and Q3 and to the left of B2 are two images of a third galaxy marked C1-C2 in [100], but this system was not used as no redshift is currently available.

two images of a galaxy at redshift 2.74, marked B1-B2. Note that another galaxy with two images was identified in the aforementioned work, but because of its unknown redshift, it was not used in the inversion. Angular diameter distances were calculated in a flat cosmological model with $H_0 = 71$ km s^{-1} Mpc^{-1}, $\Omega_\mathrm{m} = 0.27$ and $\Omega_\Lambda = 0.73$. Using the redshift information described above, this fixes the $D_\mathrm{ds}/D_\mathrm{s}$ ratios for the lensing systems, which is required input information in our method. The inversion results described below were first shown in [71].

5.2.3 Results

Unless noted otherwise, the uncertainties reported here specify a 68% confidence level.

First inversion

Since image A5 was marked as uncertain, the first inversion does not include it. The algorithm was instructed to look for mass in a square region, 35 arcsec wide, roughly centered on image Q4. The null space fitness measure was based on a square region, 60 arcsec wide, subdivided into a 64 by 64 grid. For each source, the image regions were excluded from the null space, and for systems A and B, the central cluster region was excluded as well, allowing the algorithm to predict the locations of the central

images of these systems. The null space is a relatively large region, but this avoids the introduction of unnecessary substructure at the edge of the mass reconstruction region, that would cause images to appear at larger distances. The critical line fitness was based on a square shaped region, 40 arcsec wide, subdivided into a 64 by 64 grid. After each inversion, a finalizing step was performed, as described earlier. This causes some minor modifications to be made to the mass map, to improve the positional and time delay fitness measures. Contrary to the case of Cl 0024+1654, no explicit mass redistribution step will be performed here.

The average solution of 28 individual inversions predicts the source positions and caustics shown in the left panel of figure 5.11. The source position of galaxy A is marked by a dashed rectangle, the position of galaxy B is marked by a dotted one. When these sources and the reconstructed lens are used to predict the image configurations, the result in the right panel of the same figure is obtained. The critical lines and caustics in these figures are calculated for the redshift of the quasar. The mass map itself is shown in the left panel of figure 5.12, with most of the mass in the same region as the brightest cluster galaxy (BCG). The standard deviation of the individual reconstruction can be seen in the right panel of the same figure, showing that the precise distribution of mass in the central region differs between the individual reconstructions. Figure 5.13 shows the average profile and its standard deviation. The large core clearly differs from the NFW-like behavior that one might expect.

When inspecting the right panel of figure 5.11, one sees that the average solution predicts central images of galaxies A and B. The predicted position of the central image of galaxy A coincides with the location of image A5, although the predicted shape is far less extended. Figure 5.14 shows the central region of the cluster, after subtracting the central cluster members using the GALFIT software [82]. In each of the filters, one can clearly see the central image of the quasar in the upper-left region. Image A5 can clearly be seen in the F555W and F814W images. Since the other constraints predict a central image of galaxy A at this location and since it indeed resembles a mirror image of A1, we feel confident that this is in fact the central image of galaxy A.

Second inversion

Including the central image of galaxy A will provide additional information that will lead to a different inversion since its true shape is different from the one predicted by the first inversion. For this reason, a second inversion was performed in which image A5 was added as an observational constraint. The rest of the constraints are the same as in the first inversion. Figure 5.16 shows the source and image configurations obtained in this case, using the average solution of 28 individual reconstructions. The central image of galaxy A is now clearly more extended than in the first inversion. When the images of galaxies A and B are projected back onto their source planes, the source shapes in figure 5.15 are reconstructed. The back-projected images of each

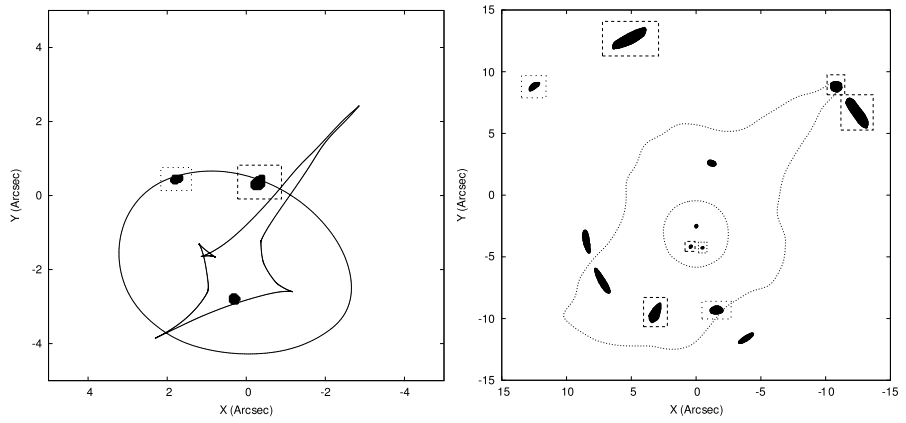


Figure 5.11: Left panel: when the input images are projected back on to the source plane using the average of 28 individual solutions, these source positions are obtained. Galaxy A is surrounded by a dashed rectangle, galaxy B by a dotted one. The caustics correspond to the redshift of the quasar. Right panel: when the sources and caustics of the left panel are used to predict the images and critical lines using the average solution, this configuration is obtained.

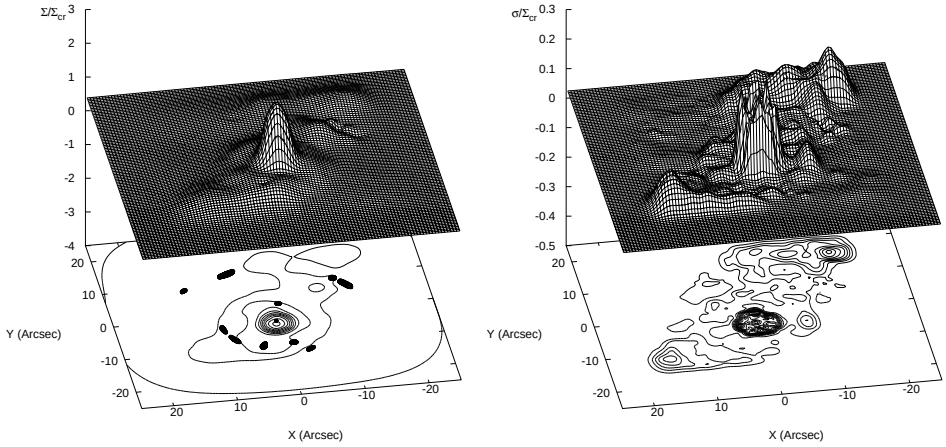


Figure 5.12: Left panel: average mass map of 28 individual solutions when image A5 is not taken into account. Most of the mass is found to coincide with the region of the BCG. The critical density was calculated at the redshift of the quasar. Right panel: standard deviation of the individual solutions, showing that the precise distribution near the center of the cluster is somewhat uncertain.

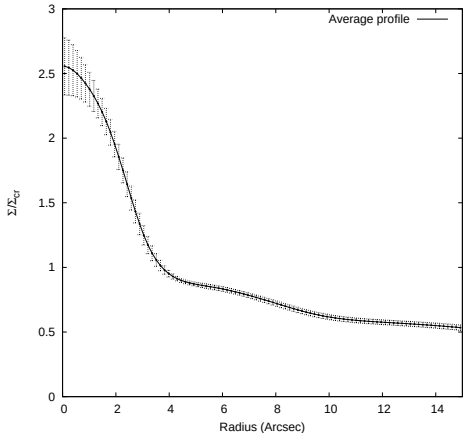


Figure 5.13: The circularly averaged profile of the inversions when image A5 is disregarded, together with the standard deviation.

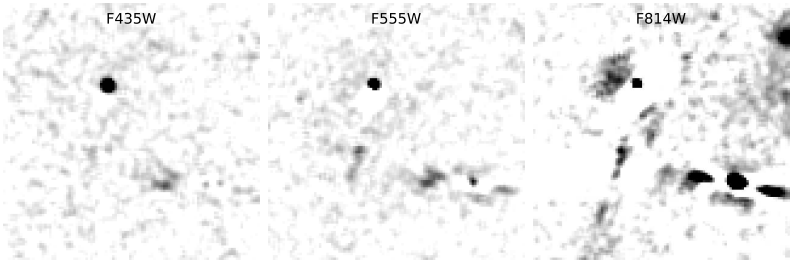


Figure 5.14: The central part of the cluster after removing the contribution of the central cluster members using GALFIT. The central quasar image can clearly be seen in each filter, in the upper left part of the image. Below and to the left of it, image A5 can be seen in the F555W and F814W images. More to the right, an extra object can be seen, where the inversion predicts the central images of galaxies B and C.

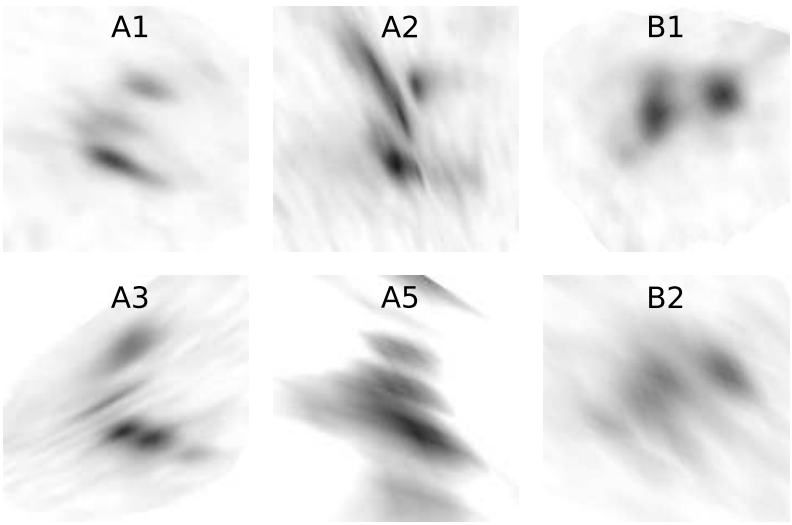


Figure 5.15: When using the model resulting from the second inversion to project the galaxy images back onto their source planes, these images are obtained. Note that image A4 is not shown here, as it is occluded by a cluster galaxy. The size of galaxy A is approximately 4 kpc, the size of galaxy B is approximately 2.5 kpc.

source clearly resemble each other, illustrating that a good positional fitness has been achieved. The estimated size of galaxy A is approximately 4 kpc, the size of galaxy B approximately 2.5 kpc.

The effect of the inclusion of image A5 can best be seen in the average mass map, as shown in the left panel of figure 5.17. Now, the mass distribution has clearly become much steeper in the central region, although some disagreement still remains between the individual solutions (right panel). A comparison with the visible matter can be seen in figure 5.18. The effect on the mass density can also be clearly seen in the circularly averaged profile, shown in the left panel of figure 5.19. It would definitely be interesting to see how much the resulting mass map resembles a NFW distribution; a fit to such a profile will be performed below.

When calculating the total mass within 60 kpc, corresponding to the region of the quasar images, and 110 kpc, the region bounded by the images of galaxy A, we find results of $2.5 \times 10^{13}\ M_\odot$ and $6.1 \times 10^{13}\ M_\odot$ respectively. These values can be compared to the findings of [110], who also find $2.5 \times 10^{13}\ M_\odot$, and of [100], who find $6 \times 10^{13}\ M_\odot$. This illustrates that the mass within the images is well constrained.

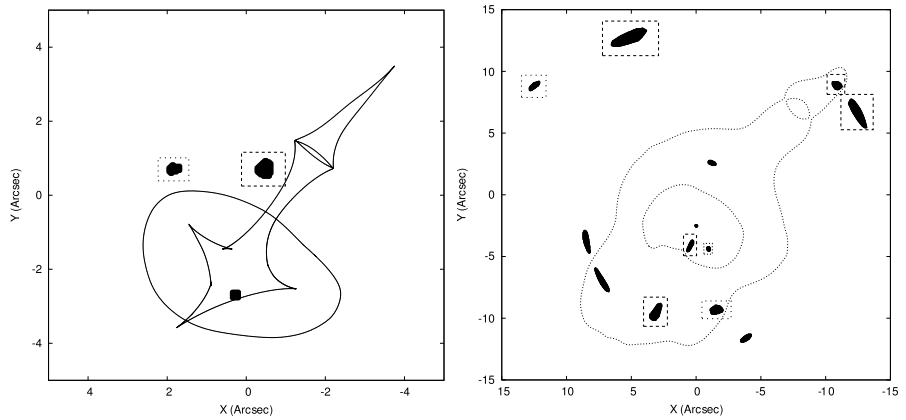


Figure 5.16: Left panel: when the average of 28 individual solutions is used to reconstruct the source plane when image A5 is included as a constraint, this result is obtained. The dashed box again indicates galaxy A, the dotted one galaxy B. Right panel: the sources and caustics in the left panel correspond to these images and critical lines. In this case, the central image of galaxy A is indeed more elongated. The critical lines and caustics again correspond to the redshift of the quasar.

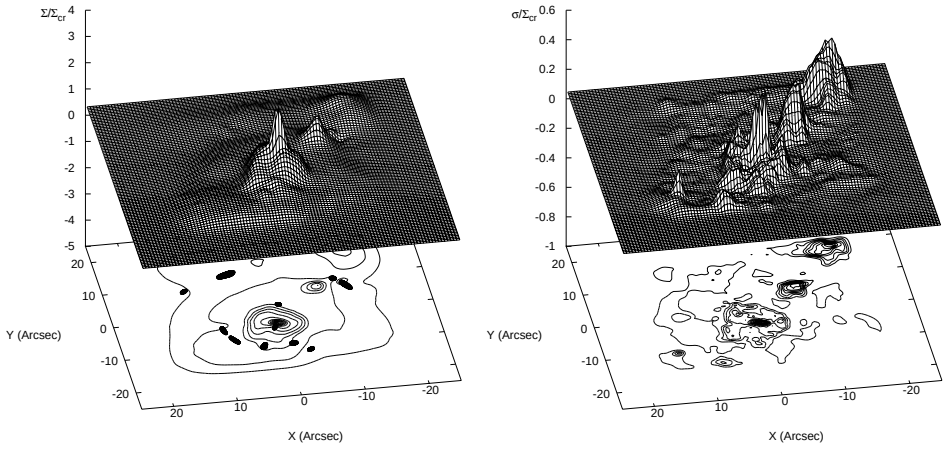


Figure 5.17: Left panel: average mass density of the 28 individual solutions. When image A5 is included, the central region clearly needs to be much steeper. Right panel: standard deviation of the individual solutions. The precise mass distribution in the central region differs somewhat among the reconstructions. The critical density again corresponds to the critical density at the redshift of the quasar.

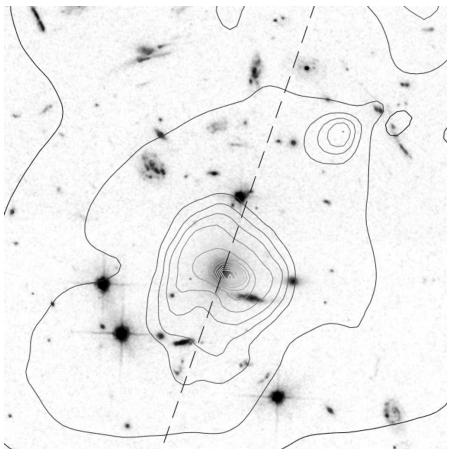


Figure 5.18: The average solution resulting from the second inversion is shown as a contour map on top of the ACS image. Most of the mass clearly lies in the same area as the central cluster members. The mass peak in the north-west part of the figure is not significant, as it can easily be redistributed. The dashed line indicates the orientation of the BCG.

5.2.4 Fitting a NFW profile

The NFW density profile [76] is described by:

$$\rho_{\rm NFW}(r) = \frac{\rho_{\rm s}}{(r/r_{\rm s})(1+r/r_{\rm s})^2},$$

in which $\rho_{\rm s}$ is a density scale factor and $r_{\rm s}$ is a characteristic radius. The density scale can be expressed in terms of $c_{\rm vir}$, which relates $r_{\rm s}$ to the virial radius $r_{\rm vir}$ through $r_{\rm vir} = c_{\rm vir} r_{\rm s}$. The virial radius itself is defined as the radius within which the mean density equals $\Delta_{\rm vir}$ times the mean matter density at the redshift of the halo. This virial overdensity $\Delta_{\rm vir}$ stems from the spherical collapse model, and for a flat cosmological model it can be approximated by (e.g. [13], [14])

$$\Delta_{\rm vir} \approx \frac{18\pi^2 + 82x - 39x^2}{\Omega(z)},$$

in which $x = \Omega(z) - 1$ and $\Omega(z)$ is defined as the ratio of the mean matter density to the critical density. Through lens inversion one recovers the projected density:

$$\Sigma_{\rm NFW}(R) = \int_{-\infty}^{\infty} \rho_{\rm NFW}(R,z)dz, \tag{5.1}$$

for which an analytical expression can be calculated (e.g. [111]):

$$\Sigma_{\rm NFW}(x) = \begin{cases} \frac{2r_{\rm s}\rho_{\rm s}}{(x^2-1)}\left(1 - \frac{2}{\sqrt{1-x^2}}\operatorname{arctanh}\sqrt{\frac{1-x}{1+x}}\right) & (x<1) \\ \frac{2r_{\rm s}\rho_{\rm s}}{3} & (x=1) \\ \frac{2r_{\rm s}\rho_{\rm s}}{(x^2-1)}\left(1 - \frac{2}{\sqrt{x^2-1}}\arctan\sqrt{\frac{x-1}{1+x}}\right]) & (x>1) \end{cases},$$

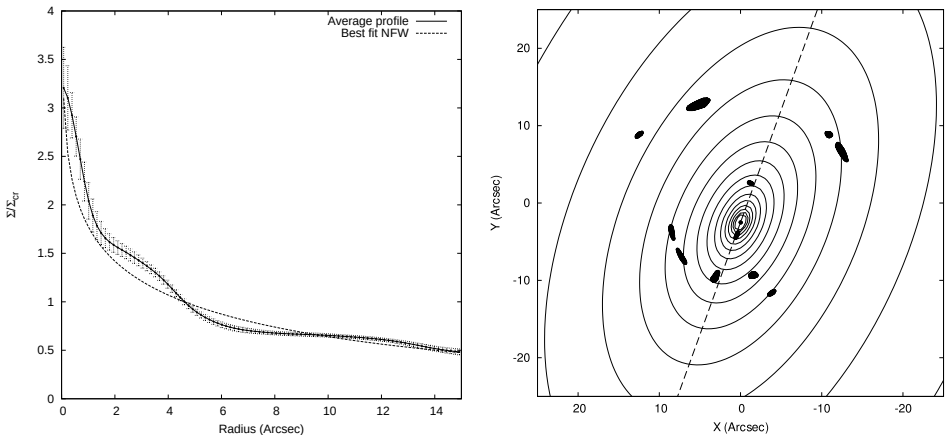


Figure 5.19: Left panel: average profile and standard deviation of the resulting mass distributions. The dashed line shows the best fit NFW profile. Right panel: when only the mass density at the location of the images is taken into account, this is the resulting best fit NFW. The center of the profile lies very close to Q5, as does the center of the BCG. The orientation is very similar to that of the BCG (dashed line), and corresponds to the general alignment of the cluster members [78].

in which $x = R/r_{\rm s}$.

Naively performing a fit of the profile in the left panel of figure 5.19 to a projected NFW profile, yields the best fit profile described by the dashed line in the same figure. One then finds $r_{\rm s} = 41.2^{+1.5}_{-1.3}$ arcsec, and $c_{\rm vir} = 5.37^{+0.14}_{-0.12}$. Although this seems to correspond well to the values found by [79], who reported $r_{\rm s} = 39^{+12}_{-9}$ arcsec and $c_{\rm vir} = 6.1^{+1.5}_{-1.2}$ (90% confidence) based on Chandra X-ray observations, the uncertainties found in this way are far too low. As explained previously, using the monopole degeneracy it is possible to redistribute the mass in between the images, without affecting any of the observable properties of the lensing system. This means that the uncertainty of the circularly averaged profile is actually much larger than obtained by simply calculating the standard deviation of the individual profiles. In turn, this translates to larger uncertainties on the parameters of the fit.

Since the mass distribution in between the images is not well constrained, it is interesting to see how much the density at the location of the images themselves constrains the NFW parameters. First, we calculated the average density and its standard deviation at the location of each image. Then, an elliptical generalization of $\Sigma_{\rm NFW}$ was fitted to these data points. An axis ratio f was introduced in the projected NFW profile by setting $R = (fx^2 + y^2/f)^{1/2}$ in equation (5.1). We prefer this substitution over $R = (x^2 + (y/q)^2)^{1/2}$ that would correspond to an axisymmetric NFW instead of a triaxial one, because the circularly averaged profile in the first case corresponds closely to the profile of a symmetric NFW with the same $r_{\rm s}$ and $c_{\rm vir}$ parameters. This

allows the obtained values to be compared directly to fits to the circularly averaged profile. After fitting the elliptical generalization of $\Sigma_{\rm NFW}$, the values $r_{\rm s} = 58^{+21}_{-13}$ arcsec and $c_{\rm vir} = 3.91 \pm 0.74$ are obtained. The best fit NFW is shown in the right panel of figure 5.19. Its orientation corresponds to that of the BCG and to the general configuration of the cluster members as reported in [78].

5.2.5 Predictions

In [100] a lens model was used to predict the redshift of galaxy C, of which the two images lie between B1 and Q3, and to the left of B2 respectively (see figure 5.10). Doing the same using the average model discussed above, we find that the back-projected images nearly overlap for a $D_{\rm ds}/D_{\rm s}$ ratio of 0.64, corresponding to a redshift of 3.35, slightly higher than the reported redshift of 2.94. After the inversions were completed, we have learned that the authors of the aforementioned work have now spectroscopically confirmed the redshift of galaxy C to be 3.288 (T. Broadhurst, private communication).

The right panel of figure 5.16 contains a prediction for the central image of galaxy B, lying to the right of image A5. Inspecting figure 5.14 again, there indeed seems to be an object at that location, which is especially clear in the F435W and F555W filters. It is important to note however that the model also predicts that the central image of galaxy C mentioned above, is located at almost the same location as the central image of galaxy B. For this reason, the object that can be seen in figure 5.14, is possibly a superposition of the central images of these two galaxies.

The predicted flux ratios for the quasar system – relative to the flux of Q1 – are shown in table 5.1 and are compared to the flux ratios from other works. Although no magnification information was used in the inversion, the general trend of the predictions matches the observations. Also note that the relatively large uncertainties show that the non-parametric technique can accommodate a wide number of flux ratios, without taking microlensing into account.

Finally, the model presented here predicts a time delay of slightly over 1300 days between images Q1 and Q4 of the quasar. This is still consistent with the constraint presented in [33] which specifies that this delay should be over 1250 days. The Q1-Q5 time delay is predicted to be of the order of 1900 days.

5.2.6 Discussion

The system under study only provides a few sources at different redshifts, which, in principle, still allows a generalized version of the mass sheet or steepness degeneracy [69]. It is for this reason that the available time delay information is of particular importance here, as it helps to break the degeneracy. The fact that the degeneracy is broken well can be seen in the low dispersion in the outer regions of the surface density (right panels of Figs. 5.12 and 5.17) which is of the order of $\Sigma/\Sigma_{cr} \approx 0.05$, indicating that in the genetic algorithm a similar mass sheet basis function is found

	Prediction	CASTLES[1]			I2005	F2008
Image		F160W	F555W	F814W		
Q1	1	1	1	1	1	1
Q2	1.03 ± 0.38	0.6486	1.0864	1.3428	0.732	0.724
Q3	0.54 ± 0.19	0.4487	0.4529	0.4656	0.346	0.592
Q4	0.29 ± 0.11	0.3191	0.6138	0.2489	0.207	
Q5	0.032 ± 0.029	0.0114	0.00024	0.0047	0.003	

Table 5.1: The predicted flux ratios of the quasar images, compared to data from the CASTLES project, [45] and [33] respectively. Note that only in this last work, the combined effect of the intrinsic variability of the source and the time delay has been taken into account. The general trend of the predicted values matches the observations, even though no magnification information was used in the inversion. The uncertainties show that this non-parametric inversion method can create a wide variety of flux ratios, even without having to consider microlensing.

in each individual reconstruction. It is interesting to compare the mass map of the second inversion to the mass map obtained by [93]. The outer contours of their reconstruction show a remarkably circular structure, causing a similar effect as the mass sheet basis function used in our work. The contour steps in that figure would correspond to $\Sigma/\Sigma_{\rm cr} = 0.22$, indicating that a similar mass density will be found near the edges of image system A as the inversion shown here.

Note that in the reconstruction of the projected mass density, relatively large structures seem to exist to the north and south of images A3 and A4. As already suggested by the large associated standard deviations, one should not place much confidence in the displayed shape of these features, as the mass in those regions can easily be redistributed without affecting any of the observable properties of the lensing system using the monopole degeneracy. For the same reason it is extremely difficult to make reliable statements about the nature of substructure that may be present near the cluster center. One can only hope to make reliable predictions about the projected density at the location of the images themselves, illustrating the need for lenses with many multiply-imaged systems. Furthermore, to probe the core regions of clusters, central images are of particular importance as is nicely illustrated by the difference in profiles between the two inversions shown in this article.

When studying the constraints provided by the density at the image locations, we find that the resulting best fit NFW bears great resemblance to the general cluster configuration. As is often the case (e.g. [54]) the fit has a very similar orientation as that of the central galaxy, which in this case also follows the general distribution of the cluster galaxies. In a recent study, the fact that lensing clusters are often over-concentrated was discussed [77]. Although the circularly averaged profile indeed suggests that this may be the case in this cluster as well, the more reliable two-dimensional fit yields an estimate of the concentration which is compatible with the

[1] http://www.cfa.harvard.edu/castles/

expected value $c_{\rm vir} \sim 4$ [79][14].

Acknowledgments

The Cl 0024+1654 and SDSS J1004+4112 image data presented here were obtained from the Multimission Archive at the Space Telescope Science Institute (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NAG5-7584 and by other grants and contracts.

—6—

Conclusion

This dissertation describes a method to extract information about the mass distribution in a gravitational lens system. The method is a so-called non-parametric lens inversion procedure, as it allows the mass distribution to have a wide variety of features, and no specific shape of the mass distribution is assumed. To do so, a large number of basis functions are used, the weights of which are determined by a genetic algorithm. The procedure uses Plummer basis functions, but is has been verified that Gaussian basis functions yield similar results; square shaped basis functions appear to be less suitable as they quickly cause complex caustic structures. With this in mind, it is interesting that in [49] the point was made that when basis functions overlap, the introduced correlation reduces the effective number of degrees of freedom.

Such a genetic algorithm allows one to specify the desired properties of a solution with great flexibility. One only has to be able to tell which trial solution is better than another, no assumptions about differentiability or even continuity of a goodness-of-fit measure are required. Using what is called a multi-objective genetic algorithm one can even specify different kinds of these fitness measures, which should be optimized simultaneously. In the context of gravitational lensing, this allows one to use several types of observations: the positions of the images, the region where no images are seen, i.e. the null space, and time delays between the images for example. In principle it should also be possible to extend this method to the weak lensing regime by adding a fitness measure based on the predicted and measured shear, but this requires further testing. The fitness criteria described here are based on physically motivated arguments only, apart from the assumption that there is no mass structure outside the square shaped region specified by the user, no assumptions are present.

The examples based on simulations shown in this thesis, illustrate that these types of information can be used to learn about the lensing mass distribution. What exactly

can be learned depends on the information at hand, as degeneracies play an important role in lens inversion. For example, in tests using the few-sources situation, in the beginning it was not clear why the size of the sources could not be retrieved. After all, the presence of two sources should break the mass sheet degeneracy. Of course, it is the generalized mass sheet degeneracy which is responsible for this.

The mass sheet degeneracy allows a scaled source to correspond to the same images, as long as one rescales the lens mass distribution and adds a specific sheet of mass, and is undoubtedly one of the most famous degeneracies in lensing. It was shown in this dissertation that this degeneracy can be generalized to multiple sources, even with different scale factors. An important fact is that rescaling a source is directly coupled to the mass density at the location of the corresponding images. This means that rescaling sources will necessarily alter the substructure of the lens. Therefore, it is actually the prior information about the smoothness of the mass distribution which breaks this degeneracy when there are multiple sources. In the method described here, overlapping Plummer basis functions are used, which implicitly determine a level of smoothness.

Another important degeneracy, here called the monopole degeneracy, allows one to easily redistribute mass in between the images without changing any of the observable properties of these images. This makes it clear that the amount of images and their coverage of the image plane are an important factor, as one can only hope to learn about the mass density at the location of these images. The importance of the local density of the images was also mentioned in [19], but was not linked to the monopole degeneracy. The monopole degeneracy seems to be under-appreciated: the only direct application that can be found is in [113], where circularly symmetric modifications of power-law models for PG 1115+080 were explored. Yet it is clear that the degeneracy is an important aspect of any gravitational lens inversion, as it can be used to introduce or remove many kinds of features, and can greatly enhance any uncertainties involved.

These two degeneracies alone already make the need for lensing systems with a great multitude of images clear. The method can also be applied to few-sources systems, as was done for the lensing clusters, Cl 0024+1654 and SDSS J1004+4112. In this case however, one has to be very careful when interpreting the results. For instance, it was explicitly shown using the recovered mass distribution of Cl 0024+1654 that it can easily be modified significantly and still yield the same predicted images.

In working with genetic algorithms, it has become clear that one has to be quite specific about what constitutes a good solution and what does not. Otherwise, the algorithm may evolve towards a solution which performs excellently according to its fitness measure, but which does not have the desired properties at all. A clear example is the prediction of extra images in systems with few images, if the null space is not used as an extra constraint.

The images in gravitational lens systems clearly encode information about the mass distribution of the lens, and genetic algorithms can help in extracting this information. The results can be expected to be most accurate when the image plane is sampled

well, but a general idea of how the mass is distributed can be obtained in few-sources systems as well, given that all the information present is used effectively.

—A—

Scientific contributions

A.1 Related publications

Below is a list of publications which contain material relevant to this dissertation:

- Jori Liesenborgs, Sven De Rijcke, Herwig Dejonghe, Philippe Bekaert, **'Non-parametric strong lens inversion of SDSS J1004+4112'**, 2009, MNRAS, 397, 341
- Jori Liesenborgs, Sven De Rijcke, Herwig Dejonghe, Philippe Bekaert, **'The search for dark matter via strong lens inversions of galaxy clusters using genetic algorithms'**, proceedings of the Identification of Dark Matter conference 2008 (IDM 2008), August 18-22, 2008, Stockholm, Sweden.
- Jori Liesenborgs, Sven De Rijcke, Herwig Dejonghe, Philippe Bekaert, **'Non-parametric strong lens inversion of Cl 0024+1654: illustrating the monopole degeneracy'**, 2008, MNRAS, 389, 415
- Jori Liesenborgs, Sven De Rijcke, Herwig Dejonghe, Philippe Bekaert, **'A generalization of the mass-sheet degeneracy producing ring-like artefacts in the lens mass distribution'**, 2008, MNRAS, 386, 307
- Jori Liesenborgs, Sven De Rijcke, Herwig Dejonghe, Philippe Bekaert, **'Non-parametric inversion of gravitational lensing systems with few images using a multi-objective genetic algorithm'**, 2007, MNRAS, 380, 1729
- Jori Liesenborgs, Sven De Rijcke, Herwig Dejonghe, **'A genetic algorithm for the non-parametric inversion of strong lensing systems'**, 2006, MNRAS,

367, 1209

A.2 Unrelated publications

This is a list of publications which are not related to this dissertation:

- Martijn van den Broek, Ralf Eichhorn, Christian Van den Broeck, Jori Liesenborgs, **'Theory and molecular dynamics simulation of a Brownian nanorefrigerator'**, in preparation.
- Tom Van Laerhoven, Jori Liesenborgs, Frank Van Reeth, **'Real-Time Watercolor Painting on a Distributed Paper Model'**, 2004, Proceedings of the Computer Graphics International Conference 2004, 640–643
- Jori Liesenborgs, Peter Quax, Wim Lamotte, Frank Van Reeth, **'Designing a Virtual Environment for Large Audiences'**, 2002, Lecture Notes in Computer Science, volume 2344, 585–595
- Jori Liesenborgs, Wim Lamotte, Frank Van Reeth, **'Voice over IP with JVOIPLIB and JRTPLIB'**, 2001, Proceedings of the 26th Annual IEEE Conference on Local Computer Networks

—B—

Deflection of a light ray

To determine the physical path of a light ray, we have to find the solutions which extremize the light travel time:

$$\begin{aligned} t &= \frac{1}{c}\int_S^O n(\boldsymbol{r})\mathrm{d}\sigma \\ &= \frac{1}{c}\int_S^O \mathrm{d}\sigma - \frac{1}{c}\int_S^O \frac{2\Phi(\boldsymbol{r})}{c^2}\mathrm{d}\sigma \end{aligned}$$

Suppose that the path is described by $\boldsymbol{r}(\lambda)$, in which $\lambda \in [\lambda_S, \lambda_O]$ and the positions of source and observer are given by:

$$\boldsymbol{r}_S = \boldsymbol{r}(\lambda_S)$$

$$\boldsymbol{r}_O = \boldsymbol{r}(\lambda_O).$$

The integral above then becomes

$$t = \frac{1}{c}\int_{\lambda_S}^{\lambda_O} n\left(x(\lambda), y(\lambda), z(\lambda)\right)\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\lambda}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}\lambda}\right)^2 + \left(\frac{\mathrm{d}z}{\mathrm{d}\lambda}\right)^2}\mathrm{d}\lambda,$$

in which $x(\lambda)$, $y(\lambda)$ and $z(\lambda)$ are the components of $\boldsymbol{r}(\lambda)$. We can rewrite this expression as

$$t = \frac{1}{c}\int_{\lambda_S}^{\lambda_O} L\left(x(\lambda), y(\lambda), z(\lambda), \frac{\mathrm{d}x}{\mathrm{d}\lambda}, \frac{\mathrm{d}y}{\mathrm{d}\lambda}, \frac{\mathrm{d}z}{\mathrm{d}\lambda}\right)\mathrm{d}\lambda$$

in which

$$L(x, y, z, \dot{x}, \dot{y}, \dot{z}) = n(x, y, z)\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}.$$

Note that in this expression $\dot{x}$ indicates the variable for which $\frac{\mathrm{d}x}{\mathrm{d}\lambda}$ shall be substituted in the integral.

Searching for a path for which $\delta t = 0$ is then equivalent to solving the differential equations which originate from the Lagrangian L. Working this out for the x-component yields:

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial x}$$

$$\Leftrightarrow \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(n(\boldsymbol{r}(\lambda))\frac{\mathrm{d}x}{\mathrm{d}\lambda}\frac{1}{\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\lambda}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}\lambda}\right)^2+\left(\frac{\mathrm{d}z}{\mathrm{d}\lambda}\right)^2}}\right) = \frac{\partial n}{\partial x}(\boldsymbol{r}(\lambda))\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\lambda}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}\lambda}\right)^2+\left(\frac{\mathrm{d}z}{\mathrm{d}\lambda}\right)^2}.$$

Note that:

$$\sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}\lambda}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}\lambda}\right)^2+\left(\frac{\mathrm{d}z}{\mathrm{d}\lambda}\right)^2} = \left|\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\right|.$$

Similar expressions can be obtained for the y- and z-components, so that in general one can write:

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(n(\boldsymbol{r}(\lambda))\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\frac{1}{\left|\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\right|}\right) = \boldsymbol{\nabla} n(\boldsymbol{r}(\lambda))\left|\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\right|$$

$$\Leftrightarrow \left(\boldsymbol{\nabla} n(\boldsymbol{r}(\lambda))\cdot\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\right)\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\frac{1}{\left|\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\right|} + n(\boldsymbol{r}(\lambda))\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\frac{1}{\left|\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\right|}\right) = \boldsymbol{\nabla} n(\boldsymbol{r}(\lambda))\left|\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\right|$$

$$\Leftrightarrow \left[\boldsymbol{\nabla} n(\boldsymbol{r}(\lambda)) - (\boldsymbol{\nabla} n(\boldsymbol{r}(\lambda))\cdot\boldsymbol{T}(\lambda))\,\boldsymbol{T}(\lambda)\right]\left|\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\right| = n(\boldsymbol{r}(\lambda))\frac{\mathrm{d}\boldsymbol{T}}{\mathrm{d}\lambda}.$$

In this expression we used

$$\boldsymbol{T}(\lambda) = \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\frac{1}{\left|\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\right|}.$$

It is clear that $\boldsymbol{T}(\lambda)$ is a vector tangent to the path, with unit length.

Since $\Phi \ll c^2$ one finds:

$$\begin{aligned}\frac{\mathrm{d}\boldsymbol{T}}{\mathrm{d}\lambda} &= -\frac{2}{c^2}\left[\boldsymbol{\nabla}\Phi(\boldsymbol{r}(\lambda)) - (\boldsymbol{\nabla}\Phi(\boldsymbol{r}(\lambda))\cdot\boldsymbol{T}(\lambda))\,\boldsymbol{T}(\lambda)\right]\left|\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\right| \\ &= -\frac{2}{c^2}\boldsymbol{\nabla}_\perp\Phi(\boldsymbol{r}(\lambda))\left|\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\right|.\end{aligned}$$

In this equation, $\boldsymbol{\nabla}_\perp\Phi$ is the projection of $\boldsymbol{\nabla}\Phi$ on the plane perpendicular to the path in a specific point.

Integrating the expression above then yields the result:

$$\boldsymbol{T}(\lambda_O) - \boldsymbol{T}(\lambda_S) = -\frac{2}{c^2}\int_{\lambda_S}^{\lambda_O}\boldsymbol{\nabla}_\perp\Phi(\boldsymbol{r}(\lambda))\left|\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\lambda}\right|\mathrm{d}\lambda$$

$$\Leftrightarrow \boldsymbol{T}_S - \boldsymbol{T}_O = \frac{2}{c^2} \int_S^O \boldsymbol{\nabla}_\perp \Phi \, \mathrm{d}\sigma.$$

C

Angular diameter distances

C.1 Introduction

In the research domain of gravitational lenses, one often has to supply angular diameter distances to perform specific calculations. Unfortunately, there is no way to measure such distances directly, but they have to be calculated from the observed redshifts of astronomical objects.

Below, an isotropic and homogeneous universe described by the Friedmann-Robertson-Walker (FRW) metric

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a(t)^2\left(\frac{\mathrm{d}r^2}{1-kr^2} + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\mathrm{d}\phi^2\right) \qquad (k = +1, 0 \text{ or } -1),$$

will be assumed to describe our universe well.

C.2 Angular diameter distances in a FRW metric

Figure C.1 shows an observer in Euclidean space, looking at an object at distance D, perpendicular to the line of sight and subtending an angle $\Delta\theta$. When the angle is small, the following relation holds:

$$d \approx \Delta\theta D.$$

This is exactly the way an angular diameter distance in a general metric is defined: the size of the object (at the time the light we receive now was emitted) must equal the corresponding angle time the angular diameter distance.

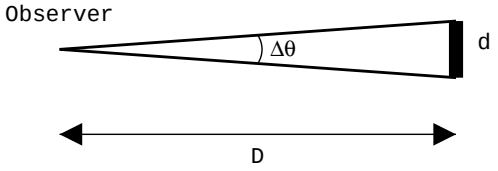


Figure C.1: Euclidean space

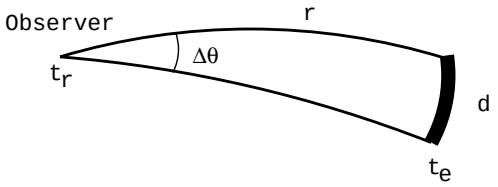


Figure C.2: FRW geometry

In a general FRW metric, the situation we are interested in is depicted in figure C.2. The coordinate system is chosen in such a way that the observer is at the origin, the object lies on a surface of constant ϕ and the radial coordinate of the endpoints is r. Note that the coordinates of the object being viewed are fixed.

Suppose that the light rays emitted at a time t_e are received by the observer at this instance, t_0. The proper size of the object at t_e is simply

$$d = a(t_e) r \Delta\theta.$$

On the other hand, the angular diameter distance D is *defined* in such a way that the relation

$$\Delta\theta D = d$$

holds, which yields the following expression for the angular diameter distance:

$$D = a(t_e) r.$$

The radial coordinate r can be calculated by noting that a light ray traces a null geodesic, so that the light rays emitted towards the observer obey the following equation:

$$c^2 \mathrm{d}t^2 = a(t)^2 \frac{\mathrm{d}r^2}{1 - kr^2}.$$

This leads to the following relations, depending on the specific geometry:

$$c \int_{t_e}^{t_0} \frac{\mathrm{d}t}{a(t)} = \int_0^r \frac{\mathrm{d}r'}{\sqrt{1 - kr^2}} = \begin{cases} \sin^{-1} r & (k = 1) \\ r & (k = 0) \\ \sinh^{-1} r & (k = -1) \end{cases}.$$

At this point, it should be clear that to actually calculate angular diameter distances, we have to know the evolution of the distance scale $a(t)$.

C.3 Dependence on cosmological parameters

The evolution of the distance scale in the FRW metric, is described by the following equation:

$$\dot{a}(t)^2 - \frac{8\pi G}{3}\rho(t)a(t)^2 + kc^2 = 0.$$

Assuming that the total energy density ρ can be split into three parts: the energy density of the vacuum ρ_v, the energy density of matter ρ_m and the energy density of radiation ρ_r. The previous equation can then be rewritten as follows:

$$\dot{a}(t)^2 - \frac{8\pi G}{3}\left[\rho_v(t) + \rho_m(t) + \rho_r(t)\right]a(t)^2 + kc^2 = 0. \tag{C.1}$$

The adiabatic expansion of the universe, provides additional information:

$$\mathrm{d}\left(\rho a^3\right) = -\frac{p}{c^2}\mathrm{d}a^3$$

$$\Leftrightarrow \left(\rho + \frac{p}{c^2}\mathrm{d}a^3\right) + a^3\mathrm{d}\rho = 0.$$

For matter, the energy density is dominated by the mass of the matter and the pressure can be set to zero:

$$p_m = 0$$

Assuming that matter, radiation and vacuum energy density evolve independently, the following relations hold:

$$\frac{\mathrm{d}a^3}{a^3} = -\frac{\mathrm{d}\rho_m}{\rho_m}$$

$$\Leftrightarrow \ln a^3 = -\ln \rho_m + \text{constant}$$

$$\Leftrightarrow a^3\rho_m = \text{constant}$$

$$\Leftrightarrow a(t)^3\rho_m(t) = a_0^3\rho_{0m},$$

In which $a_0 = a(t_0)$, the current scale factor of the universe. This way, we find an expression for $\rho_m(t)$ in terms of $a(t)$:

$$\rho_m(t) = \left(\frac{a_0}{a(t)}\right)^3 \rho_{0m}.$$

For radiation, the equation of state is

$$p_r = \frac{1}{3}\rho_r c^2,$$

leading to the following relation:

$$\rho_r(t) = \left(\frac{a_0}{a(t)}\right)^4 \rho_{0r}.$$

If the energy density of the vacuum is constant (corresponding to a true cosmological *constant* Λ), the relation

$$\rho_v(t) = \rho_{0v}$$

holds, corresponding to the following equation of state:

$$p_v = -\rho_v c^2.$$

Using these relations, equation (C.1) can be rewritten as follows:

$$\dot{a}(t)^2 - \frac{8\pi G}{3}\left[\rho_{0m}\left(\frac{a_0}{a(t)}\right)^3 + \rho_{0r}\left(\frac{a_0}{a(t)}\right)^4 + \rho_{0v}\right]a(t)^2 + kc^2 = 0.$$

Introducing the critical energy density

$$\rho_c = \frac{3}{8\pi G}H_0^2$$

and writing

$$\Omega = \frac{\rho}{\rho_c},$$

this yields:

$$\dot{a}(t)^2 - H_0^2\left[\Omega_{0m}\left(\frac{a_0}{a(t)}\right)^3 + \Omega_{0r}\left(\frac{a_0}{a(t)}\right)^4 + \Omega_{0v}\right]a(t)^2 + \frac{kc^2}{H_0^2}H_0^2 = 0$$

Dividing the equation by a_0^2 results in the expression

$$\left(\frac{\dot{a}(t)}{a_0}\right)^2 - H_0^2\left[\Omega_{0m}\left(\frac{a_0}{a(t)}\right)^3 + \Omega_{0r}\left(\frac{a_0}{a(t)}\right)^4 + \Omega_{0v}\right]\left(\frac{a(t)}{a_0}\right)^2 + \frac{kc^2}{H_0^2 a_0^2}H_0^2 = 0,$$

which can be simplified by introducing

$$R(t) = \frac{a(t)}{a_0}$$

and

$$\Omega_{0k} = -\frac{kc^2}{H_0^2 a_0^2}.$$

This way, we obtain

$$\dot{R}(t)^2 - H_0^2\left[\Omega_{0m}R(t)^{-3} + \Omega_{0r}R(t)^{-4} + \Omega_{0v}\right]R(t)^2 - \Omega_{0k}H_0^2 = 0$$

$$\Leftrightarrow \dot{R}(t)^2 - H_0^2\left[\frac{\Omega_{0m}}{R(t)} + \frac{\Omega_{0r}}{R(t)^2} + \Omega_{0v}R(t)^2 + \Omega_{0k}\right] = 0.$$

Evaluating this expression at t_0, one gets the following relation:

$$\Omega_{0m} + \Omega_{0r} + \Omega_{0v} + \Omega_{0k} = 1,$$

since

$$R(t_0) = \frac{a(t_0)}{a_0} = 1$$

and

$$\dot{R}(t_0) = \frac{\dot{a}(t_0)}{a_0} = H_0.$$

Parametrizing our ignorance of H_0 by h:

$$H_0 = \frac{h}{T_H} \qquad \text{where } T_H = (100\ km\ s^{-1}\ Mpc^{-1})^{-1}$$

and introducing

$$T = \frac{t}{T_H},$$

the evolution of an FRW universe is described by:

$$\left(\frac{\mathrm{d}R}{\mathrm{d}T}\right)^2 - h^2\left[\frac{\Omega_{0m}}{R(T)} + \frac{\Omega_{0r}}{R(T)^2} + \Omega_{0v}R(T)^2 + \Omega_{0k}\right] = 0.$$

Previously, the vacuum energy density was assumed to be constant, but more generally we can write:

$$p_v = w\rho_v c^2.$$

It is an easy exercise to obtain the following expression when using this modified equation of state:

$$\left(\frac{\mathrm{d}R}{\mathrm{d}T}\right)^2 - h^2\left[\frac{\Omega_{0m}}{R(T)} + \frac{\Omega_{0r}}{R(T)^2} + \frac{\Omega_{0v}}{R(T)^{1+3w}} + \Omega_{0k}\right] = 0 \tag{C.2}$$

From this expression one easily sees that

$$\frac{\mathrm{d}T}{\mathrm{d}R} = \frac{1}{h\sqrt{\frac{\Omega_{0m}}{R} + \frac{\Omega_{0r}}{R^2} + \frac{\Omega_{0v}}{R^{1+3w}} + \Omega_{0k}}},$$

which can be used to rewrite the integral at the end of the previous section:

$$c\int_{t_e}^{t_0}\frac{\mathrm{d}t}{a(t)} = cT_H\int_{T_e}^{T_0}\frac{\mathrm{d}T}{a_0R(T)} = c\frac{T_H}{a_0}\int_{R(T_e)}^{R(T_0)}\frac{\mathrm{d}R}{R}\frac{\mathrm{d}T}{\mathrm{d}R}$$

$$\Leftrightarrow c\int_{t_e}^{t_0}\frac{\mathrm{d}t}{a(t)} = c\frac{T_H}{a_0}\int_{R(T_e)}^{R(T_0)}\frac{\mathrm{d}R}{R}\frac{1}{h\sqrt{\frac{\Omega_{0m}}{R} + \frac{\Omega_{0r}}{R^2} + \frac{\Omega_{0v}}{R^{1+3w}} + \Omega_{0k}}}.$$

Noting that

$$R(T_e) = \frac{a(t_e)}{a_0} = \frac{1}{1+z}$$

where z is the observed redshift of the astronomical object, this can be written as

$$c\int_{t_e}^{t_0} \frac{\mathrm{d}t}{a(t)} = \frac{c}{h}\frac{T_H}{a_0}\int_{\frac{1}{1+z}}^{1} \frac{\mathrm{d}R}{R} \frac{1}{\sqrt{\frac{\Omega_{0m}}{R} + \frac{\Omega_{0r}}{R^2} + \frac{\Omega_{0v}}{R^{1+3w}} + \Omega_{0k}}}$$

$$\Leftrightarrow c\int_{t_e}^{t_0} \frac{\mathrm{d}t}{a(t)} = \frac{c}{h}\frac{T_H}{a_0}\int_{\frac{1}{1+z}}^{1} \frac{\mathrm{d}R}{\sqrt{\Omega_{0m}R + \Omega_{0r} + \Omega_{0v}R^{1-3w} + \Omega_{0k}R^2}}. \tag{C.3}$$

C.3.1 Flat space ($k = 0$)

In the $k = 0$ (and therefore $\Omega_0 k = 0$) case, the radial coordinate r is simply given by

$$r = c\int_{t_e}^{t_0} \frac{\mathrm{d}t}{a(t)},$$

from which one finds:

$$r = \frac{c}{h}\frac{T_H}{a_0}\int_{\frac{1}{1+z}}^{1} \frac{\mathrm{d}R}{\sqrt{\Omega_{0m}R + \Omega_{0r} + \Omega_{0v}R^{1-3w}}}.$$

Substituting this into the expression of the angular diameter distance, we get:

$$D = \frac{cT_H}{h}\frac{a(t_e)}{a_0}\int_{\frac{1}{1+z}}^{1} \frac{\mathrm{d}R}{\sqrt{\Omega_{0m}R + \Omega_{0r} + \Omega_{0v}R^{1-3w}}}$$

$$\Leftrightarrow D = \frac{1}{1+z}\frac{cT_H}{h}\int_{\frac{1}{1+z}}^{1} \frac{\mathrm{d}R}{\sqrt{\Omega_{0m}R + \Omega_{0r} + \Omega_{0v}R^{1-3w}}}$$

C.3.2 Curved space ($k \neq 0$)

Using the definition of Ω_{k0}

$$\Omega_{0k} = -\frac{kc^2}{H_0^2 a_0^2},$$

we can express a_0 as follows:

$$a_0 = \frac{cT_H}{h}|\Omega_{0k}|^{-\frac{1}{2}}.$$

In the $k = +1$ case, the radial coordinate can be calculated in the following way:

$$\sin^{-1} r = \frac{c}{h}\frac{T_H h}{cT_H}\sqrt{-\Omega_{0k}}\int_{1}^{\frac{1}{1+z}} \frac{\mathrm{d}R}{\sqrt{\Omega_{0m}R + \Omega_{0r} + \Omega_{0v}R^{1-3w} + \Omega_{0k}R^2}}$$

$$\Leftrightarrow \sin^{-1} r = \sqrt{-\Omega_{0k}} \int_1^{\frac{1}{1+z}} \frac{\mathrm{d}R}{\sqrt{\Omega_{0m}R + \Omega_{0r} + \Omega_{0v}R^{1-3w} + \Omega_{0k}R^2}}$$

$$\Leftrightarrow r = \sin\left[\sqrt{-\Omega_{0k}} \int_1^{\frac{1}{1+z}} \frac{\mathrm{d}R}{\sqrt{\Omega_{0m}R + \Omega_{0r} + \Omega_{0v}R^{1-3w} + \Omega_{0k}R^2}}\right].$$

In this case, the angular diameter distance is given by:

$$\begin{aligned} D &= a(t_e)r = \frac{a(t_e)a_0}{a_0} r \\ &= \frac{1}{1+z}\frac{cT_H}{h\sqrt{-\Omega_{0k}}} \sin\left[\sqrt{-\Omega_{0k}} \int_{\frac{1}{1+z}}^{1} \frac{\mathrm{d}R}{\sqrt{\Omega_{0m}R + \Omega_{0r} + \Omega_{0v}R^{1-3w} + \Omega_{0k}R^2}}\right]. \end{aligned}$$

A similar calculation for the $k = -1$ case yields:

$$D = \frac{1}{1+z}\frac{cT_H}{h\sqrt{\Omega_{0k}}} \sinh\left[\sqrt{\Omega_{0k}} \int_{\frac{1}{1+z}}^{1} \frac{\mathrm{d}R}{\sqrt{\Omega_{0m}R + \Omega_{0r} + \Omega_{0v}R^{1-3w} + \Omega_{0k}R^2}}\right].$$

C.4 Summary

It is easy to generalize these results to obtain an expression for the angular diameter distance between objects at redshifts z_1 and z_2 ($z_1 < z_2$):

$$D(z_1, z_2) = \frac{1}{1+z_2}\frac{cT_H}{h}$$

$$\times \begin{cases} \dfrac{1}{\sqrt{-\Omega_{0k}}} \sin\left[\sqrt{-\Omega_{0k}} \displaystyle\int_{\frac{1}{1+z_2}}^{\frac{1}{1+z_1}} \mathrm{d}R \left(\Omega_{0m}R + \Omega_{0r} + \Omega_{0v}R^{1-3w} + \Omega_{0k}R^2\right)^{-\frac{1}{2}}\right] & (k = +1) \\ \displaystyle\int_{\frac{1}{1+z_2}}^{\frac{1}{1+z_1}} \mathrm{d}R \left(\Omega_{0m}R + \Omega_{0r} + \Omega_{0v}R^{1-3w}\right)^{-\frac{1}{2}} & (k = 0) \\ \dfrac{1}{\sqrt{\Omega_{0k}}} \sinh\left[\sqrt{\Omega_{0k}} \displaystyle\int_{\frac{1}{1+z_2}}^{\frac{1}{1+z_1}} \mathrm{d}R \left(\Omega_{0m}R + \Omega_{0r} + \Omega_{0v}R^{1-3w} + \Omega_{0k}R^2\right)^{-\frac{1}{2}}\right] & (k = -1). \end{cases}$$

—D—

Circularly symmetric mass distribution

The expression for the deflection angle for a general projected mass distribution Σ is:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2}\int \frac{\Sigma(\boldsymbol{\xi'})(\boldsymbol{\xi}-\boldsymbol{\xi'})}{|\boldsymbol{\xi}-\boldsymbol{\xi'}|^2}\mathrm{d}\boldsymbol{\xi'}.$$

In case the mass distribution $\Sigma(\boldsymbol{\xi})$ is circularly symmetric, this equation can be simplified. Choosing the x-axis parallel to $\boldsymbol{\xi}$, one has:

$$\boldsymbol{\xi}-\boldsymbol{\xi'} = \xi\boldsymbol{e_x} - \xi'(\boldsymbol{e_x}\cos\varphi + \boldsymbol{e_y}\sin\varphi)$$

in which φ is the angle between $\boldsymbol{\xi'}$ and the x-axis and ξ' is the length of the same vector.

In taking the integral, the y-component will average out to zero. Supposing that the projected density Σ is circularly symmetric:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \boldsymbol{e_x}\frac{4G}{c^2}\int_0^\infty \mathrm{d}\xi' \int_0^{2\pi}\mathrm{d}\varphi \frac{(\xi-\xi'\cos\varphi)\xi'\Sigma(\xi')}{\xi^2+\xi'^2-2\xi\xi'\cos\varphi}.$$

Calling $\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \hat{\alpha}(\boldsymbol{\xi})\boldsymbol{e_x}$, the expression above becomes:

$$\begin{aligned}\hat{\alpha}(\boldsymbol{\xi}) &= \frac{4G}{c^2}\xi\int_0^\infty \mathrm{d}\xi'\Sigma(\xi')\xi'\int_0^{2\pi}\mathrm{d}\varphi\frac{1}{\xi^2+\xi'^2-2\xi\xi'\cos\varphi}\\ &- \frac{4G}{c^2}\int_0^\infty \mathrm{d}\xi'\Sigma(\xi')\xi'^2\int_0^{2\pi}\mathrm{d}\varphi\frac{\cos\varphi}{\xi^2+\xi'^2-2\xi\xi'\cos\varphi}.\end{aligned}$$

Appendix D. Circularly symmetric mass distribution

Considering first

$$\int_0^{2\pi} \mathrm{d}\varphi \frac{1}{\xi^2 + \xi'^2 - 2\xi\xi' \cos\varphi}$$

and calling $e^{i\varphi} = z$, one has:

$$\mathrm{d}\varphi = \frac{\mathrm{d}z}{iz} \qquad \cos\varphi = \frac{z + z^{-1}}{2}$$

$$\Rightarrow \int_0^{2\pi} \mathrm{d}\varphi \frac{1}{\xi^2 + \xi'^2 - 2\xi\xi' \cos\varphi} = -\frac{1}{i\xi\xi'} \oint \frac{\mathrm{d}z}{z^2 - z\left(\frac{\xi}{\xi'} + \frac{\xi'}{\xi}\right) + 1} = -\frac{1}{i\xi\xi'} \oint \frac{\mathrm{d}z}{\left(z - \frac{\xi}{\xi'}\right)\left(z - \frac{\xi'}{\xi}\right)}.$$

For the singularity $z = \frac{\xi}{\xi'}$, the residual is

$$Res\left(z = \frac{\xi}{\xi'}\right) = \frac{1}{\frac{\xi}{\xi'} - \frac{\xi'}{\xi}}. \tag{D.1}$$

The singularity $z = \frac{\xi'}{\xi}$ has residual

$$Res\left(z = \frac{\xi'}{\xi}\right) = \frac{1}{\frac{\xi'}{\xi} - \frac{\xi}{\xi'}}. \tag{D.2}$$

Since the integration path is a circle with unit radius in the complex plane, one has to use residual (D.1) when $\xi < \xi'$ and residual (D.2) when $\xi > \xi'$.

Now consider

$$\int_0^{2\pi} \mathrm{d}\varphi \frac{\cos\varphi}{\xi^2 + \xi'^2 - 2\xi\xi' \cos\varphi}$$

and again use the substitution $e^{i\varphi} = z$. One then finds:

$$\int_0^{2\pi} \mathrm{d}\varphi \frac{\cos\varphi}{\xi^2 + \xi'^2 - 2\xi\xi' \cos\varphi} = \frac{1}{2i} \oint \frac{(z + z^{-1})\mathrm{d}z}{z(\xi^2 + \xi'^2) - \xi\xi'(z^2 + 1)} = -\frac{1}{2i\xi\xi'} \oint \frac{(z^2 + 1)\mathrm{d}z}{z\left(z - \frac{\xi}{\xi'}\right)\left(z - \frac{\xi'}{\xi}\right)}.$$

The singularity at $z = 0$ then gives:

$$Res(z = 0) = 1. \tag{D.3}$$

For the singularity $z = \frac{\xi}{\xi'}$, the residual is:

$$Res\left(z = \frac{\xi}{\xi'}\right) = \frac{\frac{\xi^2}{\xi'^2} + 1}{\frac{\xi}{\xi'}\left(\frac{\xi}{\xi'} - \frac{\xi'}{\xi}\right)} = \frac{\xi^2 + \xi'^2}{\xi^2 - \xi'^2} \tag{D.4}$$

and for the $z = \frac{\xi'}{\xi}$ singularity, one finds the residual

$$Res\left(z = \frac{\xi'}{\xi}\right) = \frac{\xi'^2 + \xi^2}{\xi'^2 - \xi^2}. \tag{D.5}$$

In case $\xi < \xi'$ one has to use residuals (D.3) and (D.4), if $\xi > \xi'$ residuals (D.3) and (D.5) should be used.

One then finds the expression:

$$\begin{aligned}
\hat{\alpha}(\boldsymbol{\xi}) &= \frac{4G}{c^2}\xi\int_0^{\xi}\Sigma(\xi')\xi'\mathrm{d}\xi'\int_0^{2\pi}\frac{\mathrm{d}\varphi}{\xi^2+\xi'^2-2\xi\xi'\cos\varphi} \\
&+ \frac{4G}{c^2}\xi\int_{\xi}^{\infty}\Sigma(\xi')\xi'\mathrm{d}\xi'\int_0^{2\pi}\frac{\mathrm{d}\varphi}{\xi^2+\xi'^2-2\xi\xi'\cos\varphi} \\
&- \frac{4G}{c^2}\int_0^{\xi}\Sigma(\xi')\xi'^2\mathrm{d}\xi'\int_0^{2\pi}\frac{\cos\varphi\mathrm{d}\varphi}{\xi^2+\xi'^2-2\xi\xi'\cos\varphi} \\
&- \frac{4G}{c^2}\int_{\xi}^{\infty}\Sigma(\xi')\xi'^2\mathrm{d}\xi'\int_0^{2\pi}\frac{\cos\varphi\mathrm{d}\varphi}{\xi^2+\xi'^2-2\xi\xi'\cos\varphi} \\
&= -\frac{8\pi G}{c^2}\xi\int_0^{\xi}\Sigma(\xi')\xi'\frac{1}{\xi\xi'}\frac{1}{\frac{\xi'}{\xi}-\frac{\xi}{\xi'}}\mathrm{d}\xi' - \frac{8\pi G}{c^2}\xi\int_{\xi}^{\infty}\Sigma(\xi')\xi'\frac{1}{\xi\xi'}\frac{1}{\frac{\xi}{\xi'}-\frac{\xi'}{\xi}}\mathrm{d}\xi' \\
&+ \frac{8\pi G}{c^2}\int_0^{\xi}\Sigma(\xi')\xi'^2\frac{1}{2\xi\xi'}\left(1+\frac{\xi'^2+\xi^2}{\xi'^2-\xi^2}\right)\mathrm{d}\xi' + \frac{8\pi G}{c^2}\int_{\xi}^{\infty}\Sigma(\xi')\xi'^2\frac{1}{2\xi\xi'}\left(1+\frac{\xi^2+\xi'^2}{\xi^2-\xi'^2}\right)\mathrm{d}\xi' \\
&= -\frac{8\pi G}{c^2}\xi\int_0^{\xi}\frac{\Sigma(\xi')\xi'\mathrm{d}\xi'}{\xi'^2-\xi^2} - \frac{8\pi G}{c^2}\xi\int_{\xi}^{\infty}\frac{\Sigma(\xi')\xi'\mathrm{d}\xi'}{\xi^2-\xi'^2} \\
&+ \frac{8\pi G}{c^2}\int_0^{\xi}\Sigma(\xi')\xi'^2\frac{1}{\xi\xi'}\frac{\xi'^2}{\xi'^2-\xi^2}\mathrm{d}\xi' + \frac{8\pi G}{c^2}\int_{\xi}^{\infty}\Sigma(\xi')\xi'^2\frac{1}{\xi\xi'}\frac{\xi^2}{\xi^2-\xi'^2}\mathrm{d}\xi'.
\end{aligned}$$

The second and fourth terms cancel, and adding the two other terms gives:

$$\begin{aligned}
\hat{\alpha}(\boldsymbol{\xi}) &= \frac{8\pi G}{c^2}\int_0^{\xi}\Sigma(\xi')\frac{\frac{\xi'^3}{\xi}-\xi\xi'}{\xi'^2-\xi^2}\mathrm{d}\xi' \\
&= \frac{8\pi G}{c^2}\int_0^{\xi}\Sigma(\xi')\frac{\xi\xi'\left(\frac{\xi'^2}{\xi^2}-1\right)}{\xi^2\left(\frac{\xi'^2}{\xi^2}-1\right)}\mathrm{d}\xi' \\
&= \frac{8\pi G}{c^2\xi}\int_0^{\xi}\Sigma(\xi')\xi'\mathrm{d}\xi'.
\end{aligned}$$

Finally, one obtains the following expression:

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

in which

$$M(\xi) = 2\pi\int_0^{\xi}\Sigma(\xi')\xi'\mathrm{d}\xi'$$

is the projected mass within a radius ξ. Since, by construction, one has

$$\boldsymbol{e}_x = \frac{\boldsymbol{\xi}}{\xi}$$

one can write:

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

—E—

Implementation

The work described in this dissertation would not have been possible without writing supporting software. Below, a brief overview can be found of some of the major components and utilities. All software was programmed in the C++ programming language.

E.1 MOGAL

To make experimenting with genetic algorithms easier, a separate Multi-Objective Genetic Algorithm Library (MOGAL) was written. To write a genetic algorithm, the user only has to implement a representation of the genome, fitness measures, reproduction rules etc. In short, only the code that is specific to the problem under study needs to be implemented. The fitness calculation can be done either locally on a single CPU, or can be distributed over a number of CPUs, as desired by the user starting the genetic algorithm. As described in the text, the non-dominated sort routine necessary in a multi-objective algorithm, will be done on the GPU using the CUDA language if possible.

E.2 GRALE

The GRAvitational LEns (GRALE) library contains the essential components for simulating gravitational lens systems, and for inverting them, i.e. it contains the implementations of the genetic algorithms described in the dissertation. In case one has to write a very specific program, this library can be very useful. For most purposes

however, the tool GRALESHELL can be used, a program which allows one to easily access much of the functionality of the GRALE library without having to write a C++ program.

To visualize the result of a lens inversion, it may be useful to calculate the deflection field and its derivatives on a fine grid in the lens plane. The end result of an inversion can be quite complex, being the average of several individual inversion which consist of many hundreds of basis functions. For this reason, such a calculation can be done in a distributed way if the user desires so. Each CPU will then calculate these quantities for a specific part of the lens plane.

E.3 GRALESHELL

For most purposes, a text-mode shell around the GRALE library called GRALESHELL can be used. It contains commands to simulate gravitational lens systems. The user can specify several simple lens types, like the isothermal sphere models or the Plummer model, or can create a composite lens, i.e. a lens consisting of a superposition of several of these models, translated and rotated with respect to each other. Commands are available to visualize such a lens, using a 3D plot or a profile plot for example. Several source shapes can be used: a circular source, an elliptical one, a polygon shaped one, or even a pixelized source. The listing below shows a few example commands for simulating gravitational lenses.

```
lens/new/nsie 1272.46 1300 0.8 1.5
lens/save mylens.lensdata
lensplane/new/local -40 -40 40 40 1024 1024
lens/profile massprofile.txt 40 0 0 * * no

srcplane/new 1746.22 878.63
sources/add/ellipse 0 10 5 1 30 1.1 0.6 yes
imgplane/new/lensplane
imgplane/plot/gnuplot images1.gnuplot yes yes yes yes X Y yes
imgplane/timedelay 0 10 5
```

To invert gravitational lens data, the program contains commands to specify the input, to specify the grid on which the basis functions should be placed, and to perform the actual inversion. In the last step, the user can choose between an inversion which uses only the local CPU and one which is executed in a distributed way. The listing below shows a few commands which can be used to invert lensing data.

```
imgdata/new/points
  -24.72085695 -87.40427296
  27.49850916 14.17810948
  24.20549548 29.67064975
  end
```

```
imgdata/list/add 1785 1233.52

imgdata/new/points
  -7.309501299 -76.20927696
  23.88425492 12.48020577
  12.98410719 15.79977737
  end
imgdata/list/add 1775.42 1114.7

grid/new/uniform 0 0 250 15
invert/grid/local 1117.87 0 pointoverlap 256 * * * * * * * yes
lens/save inv1.lensdata
```

E.4 Point selection utility

In some cases it may be helpful to identify matching points in different images of the same source. For this reason, a point selection utility was created, of which a screenshot can be seen in figure E.1. The tool allows the user to select points in the input images, or points can be selected in the source plane if an initial reconstruction is available. This last feature can be helpful to identify additional features, since they may be more easily identified if the images are projected back onto the source plane, where they should originate from the same source.

E.5 GRALE Editor

To make it easier to select different images in real observed lenses as input for the inversion routine, the GRALE Editor was written. It allows one to load one or more FITS[1] images, using the World Coordinate System (WCS) information stored therein. As it is often easier to identify images in a color picture, the user can also load a PNG or JPG image and place it on top of a FITS image. If this picture has the same dimensions, i.e. it has the same pixel width and height, it is transformed to match the FITS image. If this is not the case, the user can select points in the FITS image which should match points in the PNG or JPG image and allow the program to compute the necessary transformation. Figure E.2 shown an overlay of a color image of the cluster Abell 1689 onto a FITS image of the same cluster which was made in this way.

The program allows the user to select points which should belong to different images in a system. It can also be specified which points should correspond to each other, like in the point selection utility, albeit somewhat less advanced. Time delay information can also be added to a specific point. Points can also be triangulated, allowing one

[1] Flexible Image Transport System. A file format used to store astronomical observations.

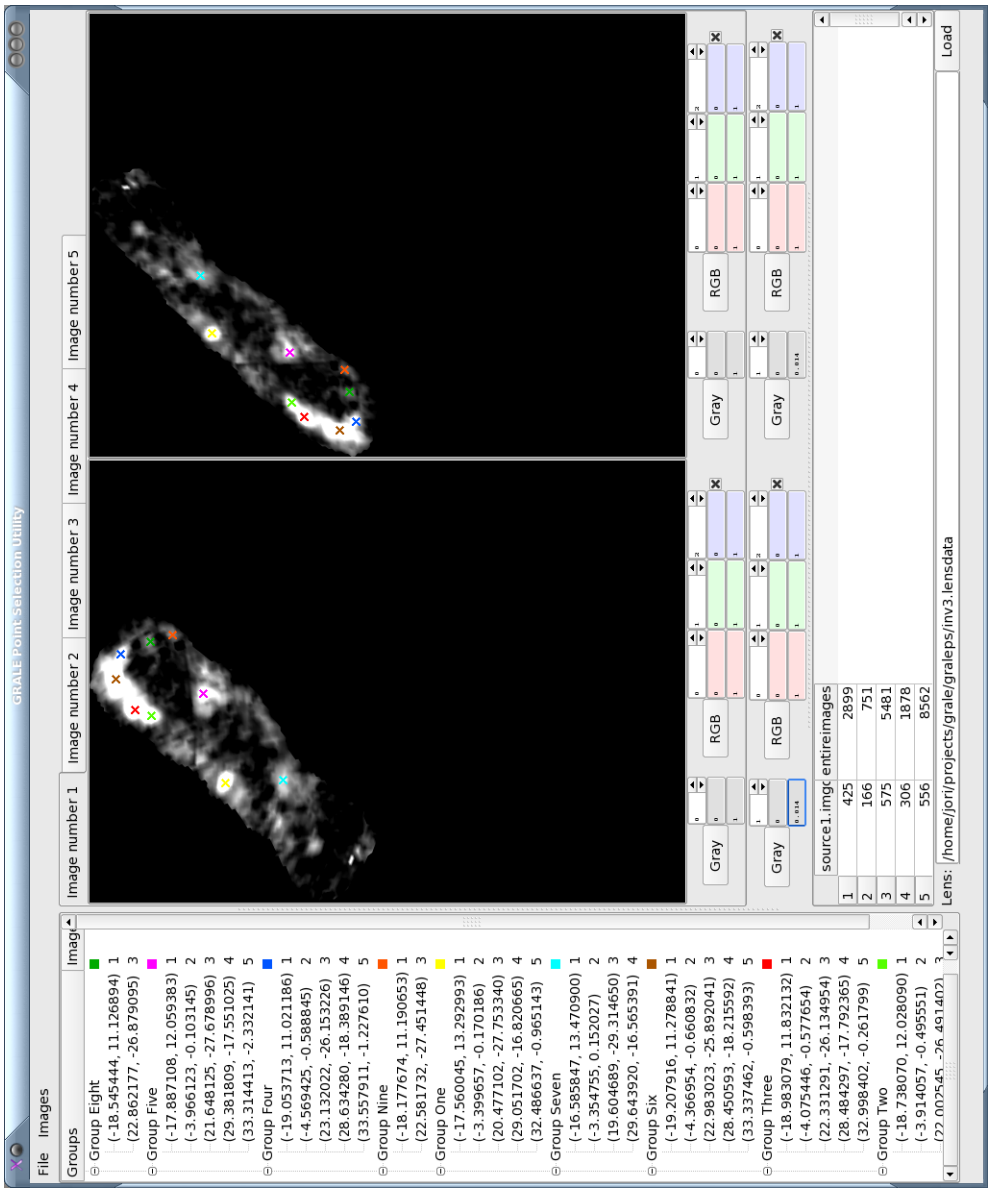


Figure E.1: Screenshot of the point selection utility. The right shape is the image as seen in the image plane. When an initial reconstruction is used to project the image back onto the source plane, this results in the shape on the left. Different images can be selected by means of the tabs at the top. Each color indicates a different point group, i.e. it identifies a point in several images that should originate from the same point in the source plane.

to manipulate grids which can be used for the null space fitness or critical line fitness for example. Figure E.3 shows points selected in two images of the same blue galaxy in the gravitational lensing cluster Cl 0024+1654.

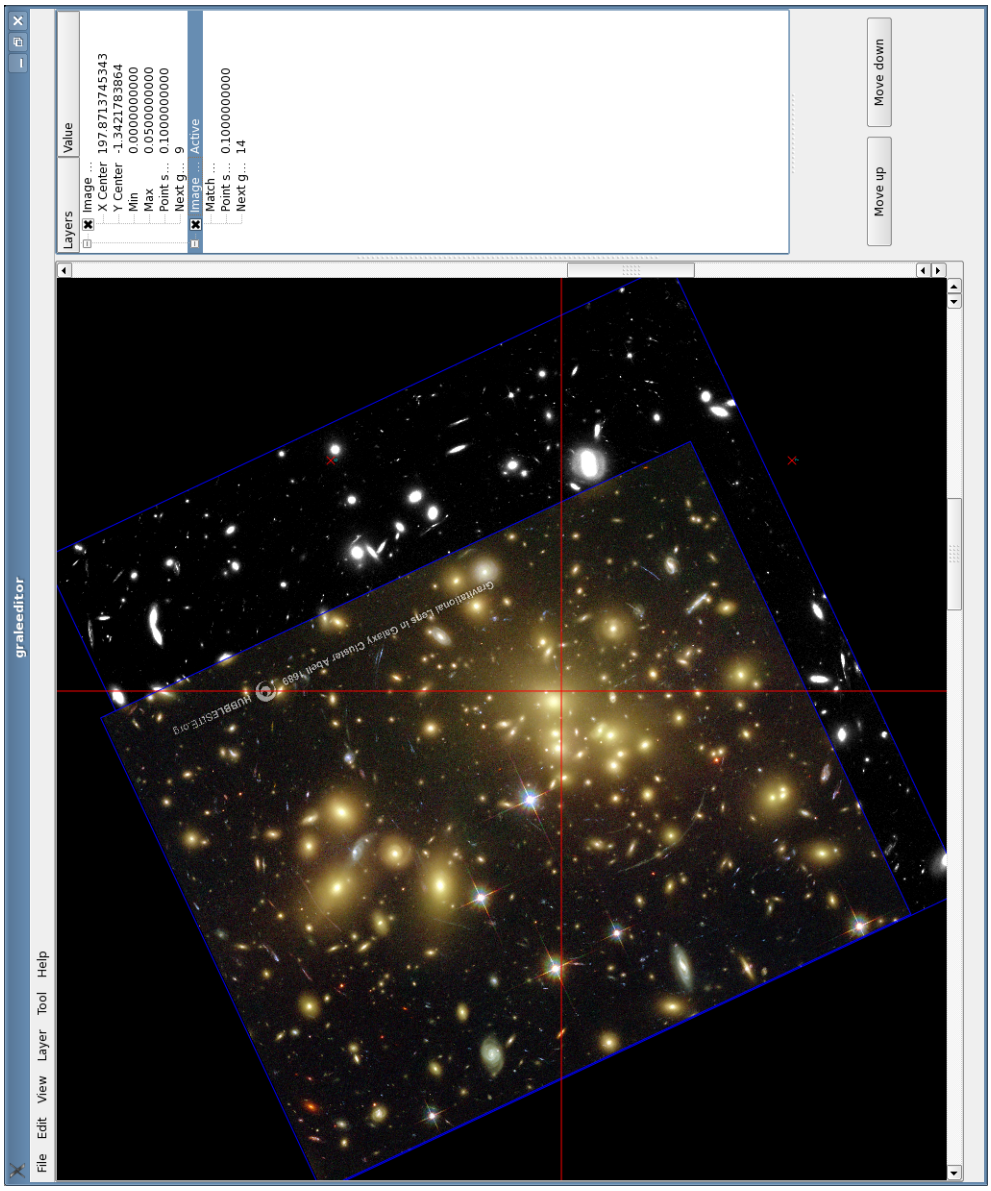


Figure E.2: The GRALE Editor was used to overlay a color image of the Abell 1689 cluster onto a FITS image of the same cluster.

Figure E.3: This figure shows two images of the same blue background galaxy in the lensing cluster Cl 0024+1654, in which points were selected to be used in the inversion routine.

Bibliography

[1] H. M. Abdelsalam, P. Saha, and L. L. R. Williams. Nonparametric Reconstruction of Abell 2218 from Combined Weak and Strong Lensing. *AJ*, 116:1541–1552, October 1998.

[2] C. Alcock, R. A. Allsman, D. R. Alves, T. S. Axelrod, A. C. Becker, D. P. Bennett, K. H. Cook, N. Dalal, A. J. Drake, K. C. Freeman, M. Geha, K. Griest, M. J. Lehner, S. L. Marshall, D. Minniti, C. A. Nelson, B. A. Peterson, P. Popowski, M. R. Pratt, P. J. Quinn, C. W. Stubbs, W. Sutherland, A. B. Tomaney, T. Vandehei, and D. Welch. The MACHO Project: Microlensing Results from 5.7 Years of Large Magellanic Cloud Observations. *ApJ*, 542:281–307, October 2000.

[3] J.-P. Beaulieu, D. P. Bennett, P. Fouqué, A. Williams, M. Dominik, U. G. Jørgensen, D. Kubas, A. Cassan, C. Coutures, J. Greenhill, K. Hill, J. Menzies, P. D. Sackett, M. Albrow, S. Brillant, J. A. R. Caldwell, J. J. Calitz, K. H. Cook, E. Corrales, M. Desort, S. Dieters, D. Dominis, J. Donatowicz, M. Hoffman, S. Kane, J.-B. Marquette, R. Martin, P. Meintjes, K. Pollard, K. Sahu, C. Vinter, J. Wambsganss, K. Woller, K. Horne, I. Steele, D. M. Bramich, M. Burgdorf, C. Snodgrass, M. Bode, A. Udalski, M. K. Szymański, M. Kubiak, T. Więckowski, G. Pietrzyński, I. Soszyński, O. Szewczyk, Ł. Wyrzykowski, B. Paczyński, F. Abe, I. A. Bond, T. R. Britton, A. C. Gilmore, J. B. Hearnshaw, Y. Itow, K. Kamiya, P. M. Kilmartin, A. V. Korpela, K. Masuda, Y. Matsubara, M. Motomura, Y. Muraki, S. Nakamura, C. Okada, K. Ohnishi, N. J. Rattenbury, T. Sako, S. Sato, M. Sasaki, T. Sekiguchi, D. J. Sullivan, P. J. Tristram, P. C. M. Yock, and T. Yoshioka. Discovery of a cool planet of 5.5 Earth masses through gravitational microlensing. *Nat*, 439:437–440, January 2006.

[4] R. D. Blandford and C. S. Kochanek. Gravitational imaging by isolated elliptical potential wells. I - Cross sections. II - Probability distributions. *ApJ*, 321:658–675, October 1987.

[5] C. Bogdanos and S. Nesseris. Genetic algorithms and supernovae type Ia analysis. *Journal of Cosmology and Astro-Particle Physics*, 5:6–+, May 2009.

[6] M. Bradač, M. Lombardi, and P. Schneider. Mass-sheet degeneracy: Fundamental limit on the cluster mass reconstruction from statistical (weak) lensing. *A&A*, 424:13–22, September 2004.

[7] M. Bradač, P. Schneider, M. Lombardi, and T. Erben. Strong and weak lensing united. *A&A*, 437:39–48, July 2005.

[8] B. J. Brewer and G. F. Lewis. When Darwin Met Einstein: Gravitational Lens Inversion with Genetic Algorithms. *Publications of the Astronomical Society of Australia*, 22:128–135, 2005.

[9] B. J. Brewer and G. F. Lewis. Strong Gravitational Lens Inversion: A Bayesian Approach. *ApJ*, 637:608–619, February 2006.

[10] T. Broadhurst, N. Benítez, D. Coe, K. Sharon, K. Zekser, R. White, H. Ford, R. Bouwens, J. Blakeslee, M. Clampin, N. Cross, M. Franx, B. Frye, G. Hartig, G. Illingworth, L. Infante, F. Menanteau, G. Meurer, M. Postman, D. R. Ardila, F. Bartko, R. A. Brown, C. J. Burrows, E. S. Cheng, P. D. Feldman, D. A. Golimowski, T. Goto, C. Gronwall, D. Herranz, B. Holden, N. Homeier, J. E. Krist, M. P. Lesser, A. R. Martel, G. K. Miley, P. Rosati, M. Sirianni, W. B. Sparks, S. Steindling, H. D. Tran, Z. I. Tsvetanov, and W. Zheng. Strong-Lensing Analysis of A1689 from Deep Advanced Camera Images. *ApJ*, 621:53–88, March 2005.

[11] T. Broadhurst, X. Huang, B. Frye, and R. Ellis. A Spectroscopic Redshift for the Cl 0024+16 Multiple Arc System: Implications for the Central Mass Distribution. *ApJ*, 534:L15–L18, May 2000.

[12] T. J. Broadhurst, A. N. Taylor, and J. A. Peacock. Mapping cluster mass distributions via gravitational lensing of background galaxies. *ApJ*, 438:49–61, January 1995.

[13] G. L. Bryan and M. L. Norman. Statistical Properties of X-Ray Clusters: Analytic and Numerical Comparisons. *ApJ*, 495:80–+, March 1998.

[14] J. S. Bullock, T. S. Kolatt, Y. Sigad, R. S. Somerville, A. V. Kravtsov, A. A. Klypin, J. R. Primack, and A. Dekel. Profiles of dark haloes: evolution, scatter and environment. *MNRAS*, 321:559–575, March 2001.

[15] Paul Bunn. Gravitational lensing: Center of light curves for a point mass lens system. Master's thesis, 2001. http://www.phy.duke.edu/ugrad/thesis/bunn/-thesis.ps.

[16] W. L. Burke. Multiple Gravitational Imaging by Distributed Masses. *ApJ*, 244:L1+, February 1981.

[17] P. Charbonneau. Genetic Algorithms in Astronomy and Astrophysics. *ApJS*, 101:309–+, December 1995.

[18] M. Clarke, C. J. Hinde, M. S. Withall, T. W. Jackson, I. W. Phillips, S. Brown, and R. Watson. Allocating Railway Platforms Using A Genetic Algorithm. In M. Bramer, R. Ellis, & M. Petridis, editor, *Research and Development in*

Intelligent Systems XXVI, , Volume . ISBN 978-1-84882-982-4. Springer-Verlag London, 2010, p. 421, pages 421–+, 2010.

[19] D. Coe, E. Fuselier, N. Benítez, T. Broadhurst, B. Frye, and H. Ford. LensPerfect: Gravitational Lens Mass Map Reconstructions Yielding Exact Reproduction of All Multiple Images. *ApJ*, 681:814–830, July 2008.

[20] W. N. Colley, J. A. Tyson, and E. L. Turner. Unlensing Multiple Arcs in 0024+1654: Reconstruction of the Source Image. *ApJ*, 461:L83+, April 1996.

[21] Charles Darwin. *On the origin of species by means of natural selection or the preservation of favoured races in the struggle for life.* John Murray, London, 1859.

[22] D. M. Deaven and K. M. Ho. Molecular geometry optimization with a genetic algorithm. In *eprint arXiv:mtrl-th/9506004*, pages 6004–+, June 1995.

[23] Kalyanmoy Deb. *Multi-Objective Optimization Using Evolutionary Algorithms.* John Wiley & Sons, Inc., New York, NY, USA, 2001.

[24] J. M. Diego, P. Protopapas, H. B. Sandvik, and M. Tegmark. Non-parametric inversion of strong lensing systems. *MNRAS*, 360:477–491, June 2005.

[25] J. M. Diego, H. B. Sandvik, P. Protopapas, M. Tegmark, N. Benítez, and T. Broadhurst. Non-parametric mass reconstruction of A1689 from strong lensing data with the Strong Lensing Analysis Package. *MNRAS*, 362:1247–1258, October 2005.

[26] J. M. Diego, M. Tegmark, P. Protopapas, and H. B. Sandvik. Combined reconstruction of weak and strong lensing data with WSLAP. *MNRAS*, 375:958–970, March 2007.

[27] A. Dressler, I. Smail, B. M. Poggianti, H. Butcher, W. J. Couch, R. S. Ellis, and A. J. Oemler. A Spectroscopic Catalog of 10 Distant Rich Clusters of Galaxies. *ApJS*, 122:51–80, May 1999.

[28] C. C. Dyer and R. C. Roeder. Possible multiple imaging by spherical galaxies. *ApJ*, 238:L67–L70, June 1980.

[29] R. Ellis, M. R. Santos, J.-P. Kneib, and K. Kuijken. A Faint Star-forming System Viewed through the Lensing Cluster Abell 2218: First Light at z $\approx$ 5.6? *ApJ*, 560:L119–L122, October 2001.

[30] J. D. Ellithorpe, C. S. Kochanek, and J. N. Hewitt. Visibility LensClean and the Reliability of Deconvolved Radio Images. *ApJ*, 464:556–+, June 1996.

[31] E. E. Falco, M. V. Gorenstein, and I. I. Shapiro. On model-dependent bounds on H(0) from gravitational images Application of Q0957 + 561A,B. *ApJ*, 289:L1–L4, February 1985.

[32] David B. Fogel. Foundations of evolutionary computation. In *Proceedings of the SPIE*, pages 1–13, 2006.

[33] J. Fohlmeister, C. S. Kochanek, E. E. Falco, C. W. Morgan, and J. Wambsganss. The Rewards of Patience: An 822 Day Time Delay in the Gravitational Lens SDSS J1004+4112. *ApJ*, 676:761–766, April 2008.

[34] J. Fohlmeister, C. S. Kochanek, E. E. Falco, J. Wambsganss, N. Morgan, C. W. Morgan, E. O. Ofek, D. Maoz, C. R. Keeton, J. C. Barentine, G. Dalton, J. Dembicky, W. Ketzeback, R. McMillan, and C. S. Peters. A Time Delay for the Cluster-Lensed Quasar SDSS J1004+4112. *ApJ*, 662:62–71, June 2007.

[35] L. Michael Freeman. *Industrial Applications of Genetic Algorithms*. CRC Press, Inc., Boca Raton, FL, USA, 1998.

[36] J. Gilmore and P. Natarajan. Cosmography with cluster strong lensing. *MNRAS*, 396:354–364, June 2009.

[37] M. V. Gorenstein, I. I. Shapiro, and E. E. Falco. Degeneracies in parameter estimates for models of gravitational lens systems. *ApJ*, 327:693–711, April 1988.

[38] A. Halkola, S. Seitz, and M. Pannella. Parametric strong gravitational lensing analysis of Abell 1689. *MNRAS*, 372:1425–1462, November 2006.

[39] J. B. Hartle. *Gravity: an introduction to Einstein's General Relativity*. Addison-Wesley, 2003.

[40] P. Helbig, D. Marlow, R. Quast, P. N. Wilkinson, I. W. A. Browne, and L. V. E. Koopmans. Gravitational lensing statistics with extragalactic surveys . II. Analysis of the Jodrell Bank-VLA Astrometric Survey. *A&A Supp.*, 136:297–305, April 1999.

[41] J. F. Hennawi, N. Dalal, and P. Bode. Statistics of Quasars Multiply Imaged by Galaxy Clusters. *ApJ*, 654:93–98, January 2007.

[42] John Holland. *Adaptation in natural and artificial systems*. University of Michigan Press, Ann Arbor, Michigan, 1975.

[43] D. E. Holz. Seeing Double: Strong Gravitational Lensing of High-Redshift Supernovae. *ApJ*, 556:L71–L74, August 2001.

[44] N. Inada, M. Oguri, E. E. Falco, T. J. Broadhurst, E. O. Ofek, C. S. Kochanek, K. Sharon, and G. P. Smith. Spectroscopic Confirmation of the Fifth Image of SDSS J1004+4112 and Implications for the M_BH-sigma* Relation at z = 0.68. *PASJ*, 60:L27+, October 2008.

[45] N. Inada, M. Oguri, C. R. Keeton, D. J. Eisenstein, F. J. Castander, K. Chiu, P. B. Hall, J. F. Hennawi, D. E. Johnston, B. Pindor, G. T. Richards, H.-W. R. Rix, D. P. Schneider, and W. Zheng. Discovery of a Fifth Image of the Large Separation Gravitationally Lensed Quasar SDSS J1004+4112. *PASJ*, 57:L7–L10, June 2005.

[46] N. Inada, M. Oguri, T. Morokuma, M. Doi, N. Yasuda, R. H. Becker, G. T. Richards, C. S. Kochanek, I. Kayo, K. Konishi, H. Utsunomiya, M.-S. Shin, M. A. Strauss, E. S. Sheldon, D. G. York, J. F. Hennawi, D. P. Schneider, X. Dai, and M. Fukugita. SDSS J1029+2623: A Gravitationally Lensed Quasar with an Image Separation of 22.5". *ApJ*, 653:L97–L100, December 2006.

[47] N. Inada, M. Oguri, B. Pindor, J. F. Hennawi, K. Chiu, W. Zheng, S.-I. Ichikawa, M. D. Gregg, R. H. Becker, Y. Suto, M. A. Strauss, E. L. Turner, C. R. Keeton, J. Annis, F. J. Castander, D. J. Eisenstein, J. A. Frieman, M. Fukugita, J. E. Gunn, D. E. Johnston, S. M. Kent, R. C. Nichol, G. T. Richards, H.-W. Rix, E. S. Sheldon, N. A. Bahcall, J. Brinkmann, Ž. Ivezić, D. Q. Lamb, T. A. McKay, D. P. Schneider, and D. G. York. A gravitationally lensed quasar with quadruple images separated by 14.62arcseconds. *Nat*, 426:810–812, December 2003.

[48] M. J. Jee, H. C. Ford, G. D. Illingworth, R. L. White, T. J. Broadhurst, D. A. Coe, G. R. Meurer, A. van der Wel, N. Benítez, J. P. Blakeslee, R. J. Bouwens, L. D. Bradley, R. Demarco, N. L. Homeier, A. R. Martel, and S. Mei. Discovery of a Ringlike Dark Matter Structure in the Core of the Galaxy Cluster Cl 0024+17. *ApJ*, 661:728–749, June 2007.

[49] E. Jullo and J.-P. Kneib. Multiscale cluster lens mass mapping - I. Strong lensing modelling. *MNRAS*, 395:1319–1332, May 2009.

[50] E. Jullo, J.-P. Kneib, M. Limousin, Á. Elíasdóttir, P. J. Marshall, and T. Verdugo. A Bayesian approach to strong lensing modelling of galaxy clusters. *New Journal of Physics*, 9:447–+, December 2007.

[51] N. Kaiser and G. Squires. Mapping the dark matter with weak gravitational lensing. *ApJ*, 404:441–450, February 1993.

[52] A. Kassiola, I. Kovner, and B. Fort. Perturbations of cluster cusps by galaxies - The triple arc in CL 0024+1654. *ApJ*, 400:41–57, November 1992.

[53] C. R. Keeton. Computational Methods for Gravitational Lensing. *ArXiv Astrophysics e-prints*, February 2001.

[54] C. R. Keeton, C. S. Kochanek, and E. E. Falco. The Optical Properties of Gravitational Lens Galaxies as a Probe of Galaxy Structure and Evolution. *ApJ*, 509:561–578, December 1998.

[55] A. Kerkhoff and H. Ling. Design of broadband antenna elements for a low-frequency radio telescope using Pareto genetic algorithm optimization. *Radio Science*, 44:6006–+, December 2009.

[56] T. P. Kling. *New approaches to gravitational lensing.* PhD thesis, 2000. http://artemis.phyast.pitt.edu/thesis/kling.ps.

[57] J. P. Kneib, Y. Mellier, B. Fort, and G. Mathez. The Distribution of Dark Matter in Distant Cluster Lenses - Modelling A:370. *A&A*, 273:367–+, June 1993.

[58] C. S. Kochanek, R. D. Blandford, C. R. Lawrence, and R. Narayan. The ring cycle - an iterative lens reconstruction technique applied to MG1131+0456. *MNRAS*, 238:43–56, May 1989.

[59] C. S. Kochanek and R. Narayan. LensClean - an algorithm for inverting extended, gravitationally lensed images with application to the radio ring lens PKS 1830-211. *ApJ*, 401:461–473, December 1992.

[60] D. C. Koo. *Large-Scale Motions in the Universe, ed. V. G. Rubin & G. V. Cayne*, pages 513–540. Princeton: Princeton Univ. Press, 1988.

[61] L. V. E. Koopmans. Gravitational imaging of cold dark matter substructures. *MNRAS*, 363:1136–1144, November 2005.

[62] L. V. E. Koopmans. Gravitational Lensing & Stellar Dynamics. In G. A. Mamon, F. Combes, C. Deffayet, and B. Fort, editors, *EAS Publications Series*, volume 20 of *EAS Publications Series*, pages 161–166, 2006.

[63] R. Kormann, P. Schneider, and M. Bartelmann. Isothermal elliptical gravitational lens models. *A&A*, 284:285–299, April 1994.

[64] I. Kovner. Cosmic gravitational diagnostics. In Y. Mellier, B. Fort, and G. Soucail, editors, *Gravitational Lensing*, volume 360 of *Lecture Notes in Physics, Berlin Springer Verlag*, pages 16–90, 1990.

[65] G. Lamer, A. Schwope, L. Wisotzki, and L. Christensen. XMM-Newton Observations and Simultaneous Optical Spectroscopy of the Gravitational Lens System SDSS J1004+4112. In A. Wilson, editor, *The X-ray Universe 2005*, volume 604 of *ESA Special Publication*, pages 641–+, January 2006.

[66] D. E. Lebach, B. E. Corey, I. I. Shapiro, M. I. Ratner, J. C. Webber, A. E. E. Rogers, J. L. Davis, and T. A. Herring. Measurement of the Solar Gravitational Deflection of Radio Waves Using Very-Long-Baseline Interferometry. *Physical Review Letters*, 75:1439–1442, August 1995.

[67] J. Liesenborgs, S. De Rijcke, and H. Dejonghe. A genetic algorithm for the non-parametric inversion of strong lensing systems. *MNRAS*, 367:1209–1216, April 2006.

[68] J. Liesenborgs, S. De Rijcke, H. Dejonghe, and P. Bekaert. Non-parametric inversion of gravitational lensing systems with few images using a multi-objective genetic algorithm. *MNRAS*, 380:1729–1736, October 2007.

[69] J. Liesenborgs, S. De Rijcke, H. Dejonghe, and P. Bekaert. A generalization of the mass-sheet degeneracy producing ring-like artefacts in the lens mass distribution. *MNRAS*, 386:307–312, May 2008.

[70] J. Liesenborgs, S. De Rijcke, H. Dejonghe, and P. Bekaert. Non-parametric strong lens inversion of Cl 0024+1654: illustrating the monopole degeneracy. *MNRAS*, 389:415–422, September 2008.

[71] J. Liesenborgs, S. de Rijcke, H. Dejonghe, and P. Bekaert. Non-parametric strong lens inversion of SDSS J1004+4112. *MNRAS*, 397:341–349, July 2009.

[72] M. Limousin, J. Richard, J.-P. Kneib, H. Brink, R. Pelló, E. Jullo, H. Tu, J. Sommer-Larsen, E. Egami, M. J. Michałowski, R. Cabanac, and D. P. Stark. Strong lensing in Abell 1703: constraints on the slope of the inner dark matter distribution. *A&A*, 489:23–35, October 2008.

[73] G. Meylan, P. Jetzer, P. North, P. Schneider, C. S. Kochanek, and J. Wambsganss, editors. *Gravitational Lensing: Strong, Weak and Micro*, 2006.

[74] Melanie Mitchell. *An Introduction to Genetic Algorithms*. MIT Press, Cambridge, MA, USA, 1998.

[75] Ramesh Narayan and Matthias Bartelmann. Lectures on gravitational lensing. 1996.

[76] J. F. Navarro, C. S. Frenk, and S. D. M. White. The Structure of Cold Dark Matter Halos. *ApJ*, 462:563–+, May 1996.

[77] M. Oguri, J. F. Hennawi, M. D. Gladders, H. Dahle, P. Natarajan, N. Dalal, B. P. Koester, K. Sharon, and M. Bayliss. Subaru Weak Lensing Measurements of Four Strong Lensing Clusters: Are Lensing Clusters Over-Concentrated? *ArXiv e-prints*, January 2009.

[78] M. Oguri, N. Inada, C. R. Keeton, B. Pindor, J. F. Hennawi, M. D. Gregg, R. H. Becker, K. Chiu, W. Zheng, S.-I. Ichikawa, Y. Suto, E. L. Turner, J. Annis, N. A. Bahcall, J. Brinkmann, F. J. Castander, D. J. Eisenstein, J. A. Frieman, T. Goto, J. E. Gunn, D. E. Johnston, S. M. Kent, R. C. Nichol, G. T. Richards, H.-W. Rix, D. P. Schneider, E. S. Sheldon, and A. S. Szalay. Observations and Theoretical Implications of the Large-Separation Lensed Quasar SDSS J1004+4112. *ApJ*, 605:78–97, April 2004.

[79] N. Ota, N. Inada, M. Oguri, K. Mitsuda, G. T. Richards, Y. Suto, W. N. Brandt, F. J. Castander, R. Fujimoto, P. B. Hall, C. R. Keeton, R. C. Nichol, D. P. Schneider, D. E. Eisenstein, J. A. Frieman, E. L. Turner, T. Minezaki,

and Y. Yoshii. Chandra Observations of SDSS J1004+4112: Constraints on the Lensing Cluster and Anomalous X-Ray Flux Ratios of the Quadruply Imaged Quasar. *ApJ*, 647:215–221, August 2006.

[80] N. Ota, E. Pointecouteau, M. Hattori, and K. Mitsuda. Chandra Analysis and Mass Estimation of the Lensing Cluster of Galaxies Cl 0024+17. *ApJ*, 601:120–132, January 2004.

[81] B. Paczynski. Gravitational microlensing at large optical depth. *ApJ*, 301:503–516, February 1986.

[82] C. Y. Peng, L. C. Ho, C. D. Impey, and H.-W. Rix. Detailed Structural Decomposition of Galaxy Images. *AJ*, 124:266–293, July 2002.

[83] A. Petiteau, Y. Shang, and S. Babak. The search for black hole binaries using a genetic algorithm. *Classical and Quantum Gravity*, 26(20):204011–+, October 2009.

[84] H. C. Plummer. On the problem of distribution in globular star clusters. *MNRAS*, 71:460–470, March 1911.

[85] Colin R. Reeves and Jonathan E. Rowe. *Genetic Algorithms – Principles and Perspectives*. Kluwer Academic Publishers, 2003.

[86] S. Refsdal. On the possibility of determining Hubble's parameter and the masses of galaxies from the gravitational lens effect. *MNRAS*, 128:307–+, 1964.

[87] G. T. Richards, C. R. Keeton, B. Pindor, J. F. Hennawi, P. B. Hall, E. L. Turner, N. Inada, M. Oguri, S.-I. Ichikawa, R. H. Becker, M. D. Gregg, R. L. White, J. S. B. Wyithe, D. P. Schneider, D. E. Johnston, J. A. Frieman, and J. Brinkmann. Microlensing of the Broad Emission Line Region in the Quadruple Lens SDSS J1004+4112. *ApJ*, 610:679–685, August 2004.

[88] V. C. Rubin, W. K. J. Ford, and N. . Thonnard. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/. *ApJ*, 238:471–487, June 1980.

[89] P. Saha. Lensing Degeneracies Revisited. *AJ*, 120:1654–1659, October 2000.

[90] P. Saha, J. Coles, A. V. Macciò, and L. L. R. Williams. The Hubble Time Inferred from 10 Time Delay Lenses. *ApJ*, 650:L17–L20, October 2006.

[91] P. Saha and L. L. R. Williams. A Portable Modeler of Lensed Quasars. *AJ*, 127:2604–2616, May 2004.

[92] P. Saha and L. L. R. Williams. Gravitational Lensing Model Degeneracies: Is Steepness All-Important? *ApJ*, 653:936–941, December 2006.

[93] P. Saha, L. L. R. Williams, and I. Ferreras. Meso-Structure in Three Strong-lensing Systems. *ApJ*, 663:29–37, July 2007.

[94] P. Salucci. The constant-density region of the dark haloes of spiral galaxies. *MNRAS*, 320:L1–L5, January 2001.

[95] M. Sambridge and G. Drijkoningen. Genetic algorithms in seismic waveform inversion. *Geophysical Journal International*, 109:323–342, May 1992.

[96] P. Schneider, J. Ehlers, and E. E. Falco. *Gravitational Lenses*. Springer-Verlag, 1992.

[97] P. Schneider and C. Seitz. Steps towards nonlinear cluster inversion through gravitational distortions. 1: Basic considerations and circular clusters. *A&A*, 294:411–431, February 1995.

[98] M. I. Shamos. *Computational Geometry*. PhD thesis, Yale University, 1979.

[99] P. R. Shapiro and I. T. Iliev. On the Mass Profile of Galaxy Cluster CL 0024+1654 Inferred from Strong Lensing. *ApJ*, 542:L1–L4, October 2000.

[100] K. Sharon, E. O. Ofek, G. P. Smith, T. Broadhurst, D. Maoz, C. S. Kochanek, M. Oguri, Y. Suto, N. Inada, and E. E. Falco. Discovery of Multiply Imaged Galaxies behind the Cluster and Lensed Quasar SDSS J1004+4112. *ApJ*, 629:L73–L76, August 2005.

[101] C. Stoughton, S. Jester, D. E. Vanden Berk, and R. Nichol. Quasar Spectral Fitting using a Genetic Algorithm. In *Bulletin of the American Astronomical Society*, volume 36 of *Bulletin of the American Astronomical Society*, pages 1586–+, December 2004.

[102] C. S. Trotter, J. N. Winn, and J. N. Hewitt. A Multipole-Taylor Expansion for the Potential of the Gravitational Lens MG J0414+0534. *ApJ*, 535:671–691, June 2000.

[103] J. A. Tyson, G. P. Kochanski, and I. P. dell'Antonio. Detailed Mass Map of CL 0024+1654 from Strong Lensing. *ApJ*, 498:L107+, May 1998.

[104] K. Umetsu, E. Medezinski, T. Broadhurst, A. Zitrin, N. Okabe, B.-C. Hsieh, and S. M. Molnar. The Mass Structure of the Galaxy Cluster Cl0024+1654 from a Full Lensing Analysis of Joint Subaru and ACS/NIC3 Observations. *ArXiv e-prints*, August 2009.

[105] Ron Unger and John Moult. genetic algorithms for protein folding simulations. *Journal of Molecular Biology*, pages 75–81, 1993.

[106] S. Vegetti and L. V. E. Koopmans. Statistics of mass substructure from strong gravitational lensing: quantifying the mass fraction and mass function. *MNRAS*, 400:1583–1592, December 2009.

[107] S. Wallington, C. S. Kochanek, and D. C. Koo. Gravitational lens inversions of CL 0024+1654. *ApJ*, 441:58–69, March 1995.

[108] S. Wallington, C. S. Kochanek, and R. Narayan. LensMEM: A Gravitational Lens Inversion Algorithm Using the Maximum Entropy Method. *ApJ*, 465:64–+, July 1996.

[109] J. Wambsganss and B. Paczynski. Parameter degeneracy in models of the quadruple lens system Q2237+0305. *AJ*, 108:1156–1162, October 1994.

[110] L. L. R. Williams and P. Saha. Models of the Giant Quadruple Quasar SDSS J1004+4112. *AJ*, 128:2631–2641, December 2004.

[111] C. O. Wright and T. G. Brainerd. Gravitational Lensing by NFW Halos. *ApJ*, 534:34–40, May 2000.

[112] O. Wucknitz. LENSCLEAN revisited. *MNRAS*, 349:1–13, March 2004.

[113] H. S. Zhao and B. Qin. Concordance of Kinematics and Lensing of Elliptical Galaxies with WMAP Cosmology. *Chinese Journal of Astronomy and Astrophysics*, 6:141–154, April 2006.

[114] A. Zitrin, T. Broadhurst, K. Umetsu, D. Coe, N. Benítez, B. Ascaso, L. Bradley, H. Ford, J. Jee, E. Medezinski, Y. Rephaeli, and W. Zheng. New multiply-lensed galaxies identified in ACS/NIC3 observations of Cl0024+1654 using an improved mass model. *MNRAS*, 396:1985–2002, July 2009.

[115] F. Zwicky. On the Masses of Nebulae and of Clusters of Nebulae. *ApJ*, 86:217–+, October 1937.