Exploring Dark Matter Properties from the Smallest to the Largest Scales

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms Universität Bonn

vorgelegt von
Maruša Bradač
aus
Maribor, Slowenien

Bonn, September 2004

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn.

Promotionskomission:

1. Mitglied: Prof. Dr. Peter Schneider (Institut für Astrophysik und Extraterrestrische Forschung, Universität Bonn)
2. Mitglied: Prof. Dr. Pavel Kroupa (Sternwarte, Universität Bonn)
3. Mitglied: Prof. Dr. Norbert Wermes (Physikalisches Institut, Universität Bonn)
4. Mitglied: Prof. Dr. Hermann Karcher (Mathematisches Institut, Universität Bonn)

Tag der Promotion: 3. November 2004

Contents

Summary

One of the main problems in cosmology is to understand the formation and evolution of galaxies, galaxy clusters, and large-scale structure. Whereas the basics of the current Cold Dark Matter (CDM) paradigm for structure formation are widely accepted, some controversial issues for CDM (e.g. cuspiness of dark-matter halos, substructure crisis) still remain. In order to test CDM predictions, one needs to investigate bound objects from the smallest (dwarf galaxies) to the largest (galaxy clusters) scales at different redshifts. In particular, their matter-content, luminous and dark, has to be studied in detail. We address these issues by using gravitational lensing. Specifically, we explore the properties of mass-substructure and the mass-profiles of galaxies, as well as the mass-profiles of galaxy clusters, using strong and weak gravitational lensing.

In the beginning we give a short preface to the subject, intended for a non-expert reader. Following the general introduction to gravitational lensing we address the lens properties of quadruply imaged systems, lensed by numerically simulated galaxies. We focus on the substructure in the lens galaxies. We compare signatures of substructure on gravitational lensing properties and compare these with the observed lensed systems. The second part of this work deals with clusters of galaxies. We describe a method that improves the mass estimates of clusters considerably and develop a new cluster mass reconstruction technique which combines strong and weak lensing. We test our method using simulations and find that it can very successfully reconstruct both the enclosed mass and the mass distribution of a cluster. Finally we apply this technique to an observed cluster RX J1347−1145, the most X-ray luminous cluster.

The signature of substructure on gravitational lensing in the ΛCDM cosmological model: We present a study of the lens properties of quadruply imaged systems, lensed by numerically simulated galaxies. We investigate a simulated elliptical and disc galaxy drawn from high-resolution simulations of galaxy formation in a concordance ΛCDM universe. The simulations include the effects of gas dynamics, star formation and feedback processes. Flux-ratio anomalies observed in strong gravitational lensing potentially provide an indicator for the presence of mass substructure in lens galaxies relative to that predicted from CDM simulations. We concentrate in particular on the prediction that, for an ideal cusp caustic, the sum of the signed magnifications of the three highly magnified images should vanish when the source approaches the cusp. Strong violation of this cusp relation indicates the

presence of substructure, regardless of the global, smooth mass model of the lens galaxy. We draw the following conclusions: (1) the level of substructure present in simulations produces violations of the cusp relation comparable to those observed, (2) higher-order catastrophes (e.g. swallowtails) can also cause relative changes of the order of 0.6 in the cusp relation as predicted by a smooth model, (3) the flux anomaly distribution depends on the image parity and flux, and both the brightest minimum and saddle-point images are more affected by substructure than the fainter images. In addition, the brightest saddle point is demagnified w.r.t. the brightest minimum. Our results are fully numerical and properly include all mass scales, without making semi-analytic assumptions. They are ultimately limited by the mass resolution of single particles in the simulation determined by current computational limits, however show that our results are not dominated by shot-noise due to the finite number of particles.

Mass-sheet degeneracy: Fundamental limit on cluster mass reconstruction from statistical (weak) lensing: Weak gravitational lensing is considered to be one of the most powerful tools to study the mass and the mass distribution of galaxy clusters. However, weak lensing mass reconstructions are plagued by the so-called mass-sheet degeneracy – the surface mass density κ of the cluster can be determined only up to a degeneracy transformation $\kappa \to \kappa' = \lambda\kappa + (1 - \lambda)$, where λ is an arbitrary constant. This transformation fundamentally limits the accuracy of cluster mass determinations if no further assumptions are made. We describe here a method to break the mass-sheet degeneracy in weak lensing mass maps using the distortion and redshift information of background galaxies and illustrate this by two simple toy models. Compared to other techniques proposed in the past, it does not rely on any assumptions about the cluster potential; it can be easily applied to non-parametric mass-reconstructions and no assumptions on boundary conditions have to be made. In addition it does not make use of weakly constrained information (such as the source number counts, used in the magnification effect). Our simulations show that *we are effectively able to break the mass-sheet degeneracy for supercritical lenses*, but that for undercritical lenses the mass-sheet degeneracy is very difficult to break, even under idealised conditions.

Strong and weak lensing united I: the combined strong and weak lensing cluster mass reconstruction method: Here we present a novel method for a cluster mass reconstruction which combines weak and strong lensing information on common scales. We extend the weak lensing formalism to the inner parts of the cluster, use redshift information of the background sources and combine these with the constraints from multiply-imaged systems. We demonstrate the quality of the method with simulations, finding an excellent agreement between the input and reconstructed mass also on scales that differ from the Einstein radius. Using a single multiple image system and photometric redshift information of the background sources we find that we are effectively able to break the mass-sheet degeneracy, therefore removing the main limitation on cluster mass estimates. We conclude that with high resolution (e.g. HST) imaging data the method can more accurately reconstruct cluster masses and their profiles than currently existing lensing techniques.

Strong and weak lensing united II: the cluster mass distribution of the

most X-ray luminous cluster RX J1347−1145: We apply the method described above to the ground-based multi-colour data of RX J1347−1145, the most X-ray luminous cluster known to date. The analysis of the data revealed an additional arc to those that were previously known for this cluster. The combined strong and weak lensing reconstruction confirms that the cluster is indeed very massive. If the redshift and identification of the multiple-image system as well as the redshift estimates of the source galaxies used for weak lensing are correct, we determine the enclosed cluster mass to $M(< 360\,\mathrm{h}^{-1}\mathrm{kpc}) = (1.2 \pm 0.3) \times 10^{15} M_{\odot}$. The reconstruction also shows a south-east mass extension that is compatible with X-ray measurements. With higher resolution (e.g. HST) imaging data, reliable multiple imaging information could be obtained and the reconstruction further improved.

1

Preface

Incidentally, disturbance from cosmic background radiation is something we have all experienced. Tune your television to any channel it doesn't receive, and about 1 percent of the dancing static you see is accounted for by this ancient remnant of the Big Bang. The next time you complain that there is nothing on, remember that you can always watch the birth of the universe.

Bill Bryson, A Short History of Nearly Everything

This short introduction is intended to explain the topics of this work for a non-expert reader. A person who is familiar with the field of cosmology and gravitational lensing might want to skip this chapter and start reading in the next one.

The Hot Big Bang Model assumes that the universe was born from a very dense and hot state. Shortly afterwards it consisted of a mixture of many interacting elementary particles. Under these conditions electromagnetic radiation was scattered very efficiently by matter and this kept the universe in thermal equilibrium. After 380,000 years the Universe had cooled enough (to 3000 K) for the electrons and nuclei to combine into atoms. When this process, called recombination, completed the mean-free-path of the photons became large and they began to propagate (nearly) freely through the Universe. Before recombination the photons were in thermal equilibrium, thus having a black-body spectral distribution corresponding to a temperature of 3000 K. Since the universe expands the temperature of the radiation changes and so we see today (13.6 Gyrs after the Big Bang) the photons having a black-body spectrum corresponding to a temperature of 2.728 K. These photons in fact give a snapshot of the universe in its infancy and are seen today in the form radiation in the microwave regime, called Cosmic Microwave Background radiation (CMB). Therefore the CMB has provided us with an excellent diagnostic to study the history of the universe.

While the existence of the background radiation itself was originally discovered by Penzias & Wilson in 1965, the first instrument to measure the CMB temperature on the whole sky was the COBE satellite. It confirmed the result of other CMB probes, that we observe a black body spectrum with a temperature of 2.728 K on the whole sky. In addition, however, each pointing COBE made showed slightly different temperature; these small fluctuations have been measured on the level of $\Delta T/T \sim 10^{-5}$. Such fluctuations in temperature across the sky are the precursors of structures we see today (galaxies, galaxy clusters, etc.); they have grown through gravitational instability from $\Delta\rho/\rho \sim 10^{-5}$ to the highly concentrated structures we observe around us. These results, however are in contradiction with the hypothesis that the universe consist only of the matter we see (i.e. luminous matter, also called baryonic). Namely we observe structures today with density contrast $\Delta\rho/\rho \gg 1$. For these structures the fluctuations should be at least 10^{-3} at the time of recombination. This is, however, not what we observe. The solution to the problem is given by the presence of dark-matter which dominates the matter content in the universe and has $\Delta\rho/\rho \sim 10^{-3}$ at the epoch of recombination (i.e. they started to grow at earlier times). Baryons on the other hand were coupled to the photons up to that time, and the pressure kept the baryons from collapsing under their own gravity. The idea of the existence of dark matter has become increasingly accepted in cosmology, with CMB measurements giving strong evidence for the presence of dark matter at the largest scales.

The first evidence for the existence of dark matter is however much older than CMB results, it came from dynamical studies of astronomical objects with Zwicky's estimation of the mass of Coma cluster in 1937. In addition the idea of the existence of dark matter is supported from other fields of research; X-ray studies and gravitational lensing among others. Unfortunately, we still do not know what the nature of dark matter is. Nowadays the most probable candidates are considered to be elementary particles that rarely interact with ordinary matter (e.g. weakly interacting massive particles or WIMPs); but such a particle has not been detected as of yet. The interface of particle physics and cosmology is therefore a fruitful field of research. Cosmologists need to study the properties of dark-matter on macroscopic scales, while a suitable candidate particle fitting these constraints needs to be detected in particle physics research.

One of the most recent successes in measuring the temperature over the whole sky with an exquisite sensitivity and resolution was achieved by the WMAP satellite. The results of WMAP together with studies of large scale structure strengthened the idea that the sum of luminous and dark matter does not represent the total matter and energy content in the universe. In addition to ordinary (luminous and dark) matter, there exist an additional form of matter or energy. In particular, the measurement that the universe is accelerating has come as a great surprise, since gravitational pull exerted by the matter in the universe slows down the expansion. The amazing conclusion was that the universe contains a combination of dark matter and dark energy, the latter being gravitationally repulsive, and the nature of both we do not understand. Cosmologists gave it the name dark energy (or also the cosmological constant Λ).

We argue that one of the best tools to study the dark matter distribution from the scales corresponding to stars to the large-scale-structure is gravitational lensing. In addition lensing can also be used to investigate the nature of dark energy, by studying the evolutionary effects of dark matter. We will focus on dark matter only, in particular we will study the mass and the mass distribution in galaxies and galaxy clusters. In this chapter we will give a very short introduction to both in the context of structure formation, mostly illuminating the problems of determining the mass distributions of galaxies and galaxy clusters and highlighting the questions we want to address in this work.

1.1 Structure formation in the universe

We have briefly mentioned before that the structures in the universe we see today grew from the primordial density fluctuations. However we did not explain their origin. The most probable (but not necessary correct) explanation we have to date is that they result from quantum vacuum fluctuations which were amplified by a rapid accelerated expansion (called inflation) in the very early universe (10^{-35} s after the big bang).

The model of structure formation, which describes how these fluctuations grew to the present state has three key ingredients. One needs to specify the cosmological model (the expansion rate, which follows from the cosmological density parameters) to define the background geometry of the universe. Further, a model for fluctuation generation and the amount and type of dark matter need to be known. For example light neutrinos might be considered as dark matter particles, however since they were relativistic when their thermal equilibrium was frozen as well as at the time the structures started to form, they can easier escape the gravitational potential than non-relativistic particles. This influences the growth of structures and the results show that neutrinos are not a suitable candidates for a dark matter particle.

Having specified these ingredients one is able to predict the evolution of dark matter. If the underlying assumptions are correct, the predicted structure formation scenario matches the observations, i.e. the CMB fluctuations and the matter structure we see in the universe today. The most successful model nowadays is the cold dark matter theory including dark energy (ΛCDM), with initial conditions from inflation. Although it might still be proven incorrect in detail, it at present the model that most comprehensively explains the observations. It assumes nearly scale-invariant initial fluctuations resulting from inflation and a type of dark matter that is cold, meaning that particles are non-relativistic at the time the structures started to form and can not free-stream out of the potential wells at the scales of cosmological interest. Since the density fluctuations can only start growing when the universe is matter dominated, the growth process depends upon the cosmological parameters.

In the framework of this model we can now investigate how the structures evolve to the present state (galaxies, clusters, large-scale-structure). In the ΛCDM model the numerical simulations predict that we have a so-called bottom-up or hierarchical scenario of structure formation. This means that the small-scale structures formed first and the more massive clumps were assembled in the process of merging. By comparing the predictions from N-body simulations to observed systems we can test the hypotheses of the structure formation model and make a step further in understanding the cosmological history and future of the universe.

1.2 Gravitational lensing

A consequence of an inhomogeneous universe is the effect of gravitational lensing. Light bundles travelling through an inhomogeneous universe will be deflected and therefore focused (or defocused). The inhomogeneities in fact act as a gravitational telescope (thus the term lensing). The gravitational lensing effects offer an excellent tool to study the matter in the universe, since it is independent of whether the matter is luminous or dark. Since other methods to determine the mass distribution depend on the information we get from objects through their luminous parts, it is usually difficult to convert this information to reveal the total matter distribution. Gravitational lensing, on the other hand, probes the distribution directly and therefore it has become one of the major tools in cosmology.

In this thesis we mostly concentrate on two applications of lensing; measuring the mass distribution of individual galaxies and galaxy clusters. In the case where source, observer and the lens are aligned, and the lens is strong enough, the light bundles can travel on different paths to the observer, and multiple images of the same source can be seen. Such systems are of particular interest in gravitational lensing. To date about 100 such systems with galaxies acting as lenses have been discovered. The properties of these multiple images (their positions w.r.t. the lens galaxy centre and fluxes) can be used to draw conclusions about the mass distribution of the lens galaxy. In particular, it turns out, that the flux ratios of the images are sensitive not only to the main lens potential but also to its small-scale structure (i.e. dwarf galaxies orbiting the main galaxy). Their properties can thus be investigated. An example of such a lens system is presented in Fig. 1.1. Four quasar images belong to the same source. The lens galaxy is seen as a green shade in the figure.

In clusters of galaxies the same effect can be observed as well. In addition to these strong distortions (also called strong lensing effect), many background sources are weakly distorted (weak lensing effect). The statistical analysis of such distortion can give us further information (up to larger radii than strong lensing) about the mass and mass distribution of clusters. The latter can not be used to detect galaxies as lenses individually, since only the cluster potential is strong enough for this effect to be observed in practice. An excellent example is given in Fig. 1.2. The latest camera, ACS, on board the Hubble Space Telescope offers an exquisite resolution and many strongly distorted multiply imaged sources can be seen. The weakly distorted sources can not be readily recognised by eye, only when their statistical properties are studied do they reveal information about the mass-distribution of the cluster.

1.2.1 Gravitational lensing by galaxies as cosmological probes

The main evidence in the past that large amounts of dark matter exist in galaxies is their flat rotation curves. The constant rotation velocities in the outer parts of the galaxy, instead of a Kepplerian decline (typical if only visible matter is taken into account), indicate that each galaxy is surrounded by significant amounts of dark matter. This has made galaxies particularly interesting in terms of cosmological research.

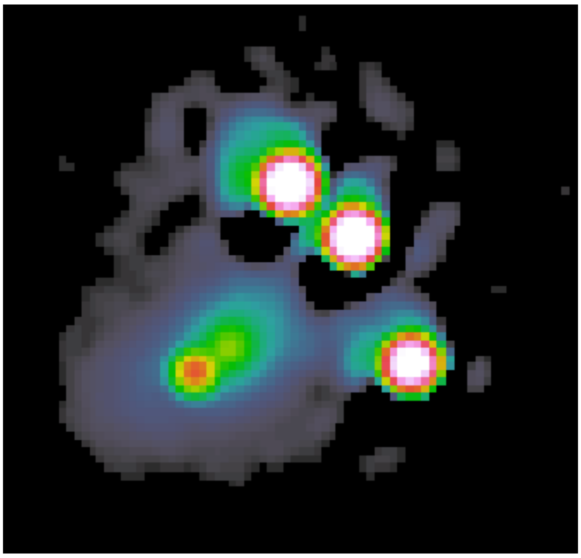


Figure 1.1: The footprint of a Yeti: a galaxy acting as a lens. The gravitationally lensed system B1422+231 with a quasar being lensed to four images. The green shade is the light from the galaxy acting as a lens (from CASTLES webpage).

We argued that the standard CDM paradigm is very successful in predicting the observed properties of dark matter on different mass scales. Therefore results that are in apparent contradiction with this theory are of extreme interest. In galaxies we seem to be facing two problems of this kind.

From the numerical simulations one can identify collapsed structures (also called halos) and study their properties. It turns out that the density profile of halos has a universal form, almost independent of the mass of the halo. This result was first found by Julio Navarro, Carlos Frenk & Simon White in the nineties, and the functional form of the profile is simply referred to as the NFW profile. It is a very remarkable result, since if the simulations are correct, galaxies and clusters of galaxies should have the same universal form (NFW) of the dark matter density distribution. In the very centre of the halos the density distribution is especially sensitive to the assumed nature of the dark matter. The CDM paradigm predicts a very steep dark matter profile in the centres of halos. This seems to be at odds with what we observe. However, on galaxy scales it is difficult to test this prediction accurately both from observations and simulations. Namely, these simulations were done with dark matter only and, especially on galaxy scales, the baryons are believed to play a very important role. Not only would they modify the density profile we measure (with gravitational lensing we measure together dark and luminous matter), they also influence the density profile of dark matter itself throughout the process of halo formation.

The second problem is the substructure problem. In simulations representing halos similar to our Galaxy, the number of substructures (halos with mass smaller than the main galaxy) does not match to what we observe in the Milky Way. The problem is especially striking for low mass halos (10^6 times less massive than the main galaxy). This problem was independently discovered by Ben Moore and Anatoly

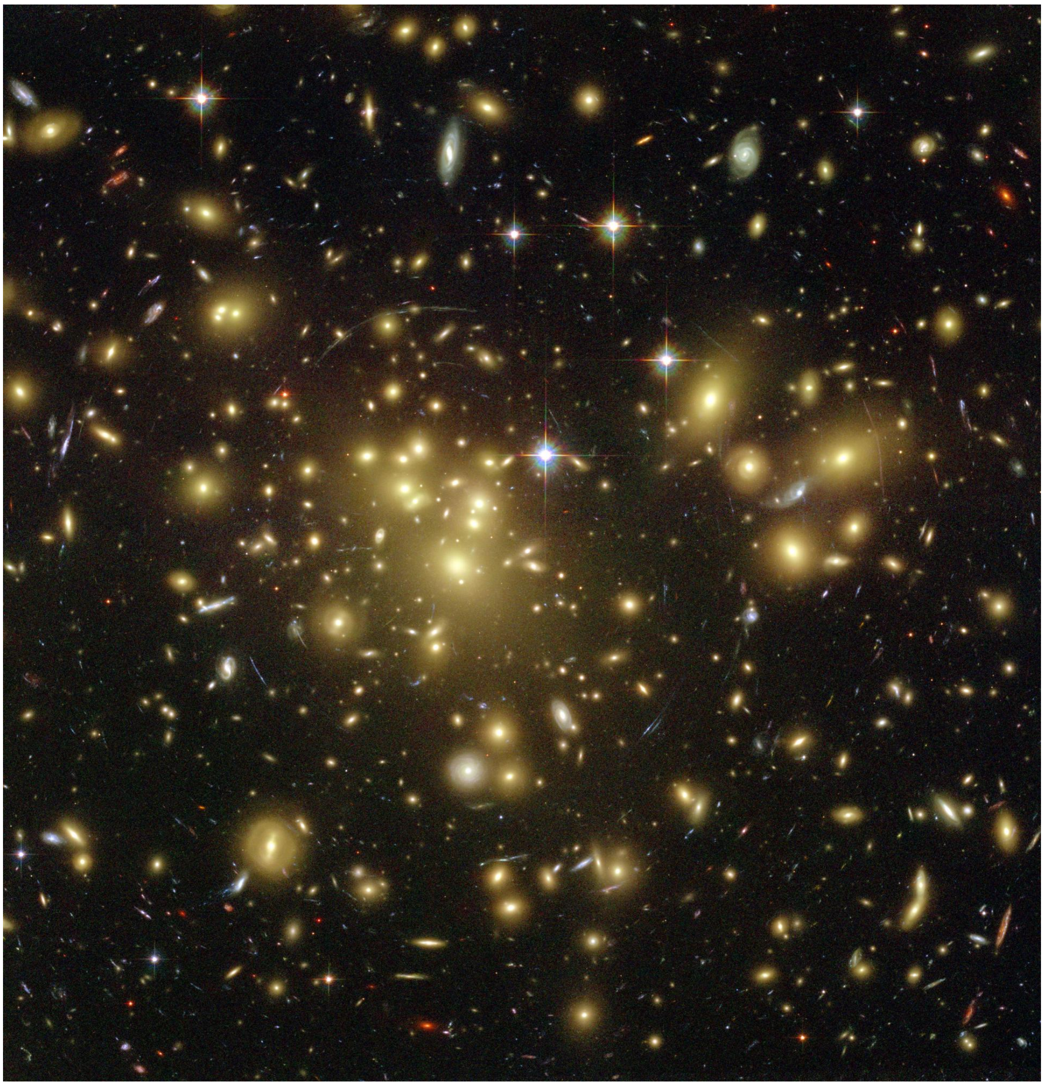


Figure 1.2: A galaxy cluster acting as a lens. Many strongly distorted multiply imaged sources (arcs) can be seen (from HST press release).

Klypin and colleagues in 1998 and has ever since posed a serious challenge to the CDM paradigm. The solution to this problem is that either the simulations and/or assumptions entering the structure formation model are wrong on these scales, or the sub-halos are indeed present in the Milky Way, however they have lost baryons in the past and are too dark to be observed. In Chapter 3 we will investigate this discrepancy for galaxies other than the Milky Way using the gravitational lensing technique.

1.2.2 Gravitational lensing by clusters of galaxies as cosmological probes

Galaxy clusters have been the focus of a very intense ongoing research in the past decades. They are the most massive bound structures in the Universe; moreover, their large dynamical time scale allows them to retain the information about their formation history. A cluster contains hundreds or even thousands of galaxies, which however do not contain a large fraction of cluster mass. The galaxy clusters are believed to be a "fair sample" of the matter content of the universe, therefore the ratio of their baryonic (i.e. luminous) mass to their total mass resembles the overall baryonic fraction in the universe. In addition the evolution in their number density as a function of mass and the time of their formation is a very important probe which can help us determine the matter content as well as the dark energy content in the universe.

It is therefore crucial to be able to make reliable estimates of cluster masses. This can be achieved by e.g. measuring the velocities of individual galaxies or from the properties of the X-ray emission of the intra-cluster gas. The latter technique has been successfully applied to many clusters. Unfortunately assumptions about hydrostatic equilibrium and spherical symmetry need to be made for to convert the luminosity of the X-ray emitting gas to the cluster mass. Gravitational lensing can also measure cluster mass to high accuracy and is free of the assumptions of baryonic physics. The problems of using gravitational lensing to determine the cluster mass are not in the underlying physics, which is well understood, but rather in the technical implementation. However, such problems are being resolved with advancing observational techniques and at present the measurements of cluster masses with gravitational lensing are a great success.

In this work we try to improve the existing methods and develop a new method to determine the enclosed mass and mass distribution of galaxy clusters using gravitational lensing. In Chapter 4 we discuss a method to overcome the largest technical difficulty of the reconstructions, the mass-sheet degeneracy. In Chapter 5 we develop a method which combines the two regimes of lensing, the regime where images are highly distorted and the one where lensing needs to be treated in the statistical sense. We show that accurate determination of the mass distribution can be performed. In Chapter 6 we apply this method to the data for the most X-ray luminous cluster at present and successfully measure its mass and mass distribution.

2

Introduction to gravitational lensing

Gravitational lensing theory rests on one of the main predictions of Einstein's theory of General Relativity, namely that light rays are bent when they propagate through an inhomogeneous medium. More precisely, light propagates along null-geodesics of the space-time metric, i.e. along the curves in the space-time with minimum 'length' between the two points.

Using the full theory, the exact calculations are hardly manageable. Fortunately, in all situations relevant to this work (as well as in most of the ones relevant for astrophysics) the corresponding equations can be linearised and the propagation of the light ray is well described by the thin lens and small angle approximations. For the description of lensing by galaxy clusters and individual galaxies these are sufficient since the deflecting mass is localised in a region much smaller compared to the distance between the observer and the source.

2.1 Gravitational lensing theory

The geometry of gravitational lensing is presented in Fig. 2.1. The mass distribution is located in the deflector plane $\mathcal{D}$ at the redshift $z_{\rm d}$, corresponding to the angular diameter distance $D_{\rm d}$ from the observer $\mathcal{O}$. The source is located in the source plane $\mathcal{S}$, at a redshift $z_{\rm s}$, with angular diameter distances $D_{\rm s}$ from the observer and $D_{\rm ds}$ from the lens. Following the thin lens approximation, the light ray between the observer and the lens plane and between the lens plane and the source plane is described by a straight line with a kink near the deflector. The effects of the space-time curvature due to a smooth homogeneous Universe are then all incorporated in the angular diameter distances. The kink is represented by the deflection angle vector $\hat{\boldsymbol{\alpha}}$, which depends on the mass distribution in the lens plane and the impact vector $\boldsymbol{\xi}$ of the light ray.

2.1.1 Lens properties

We define the optical axis to be the line connecting the observer and centre of the mass distribution (note that the precise choice of the optical axis is arbitrary). The source would be seen at an angular position $\boldsymbol{\beta}$ in the absence of the lens. Further we

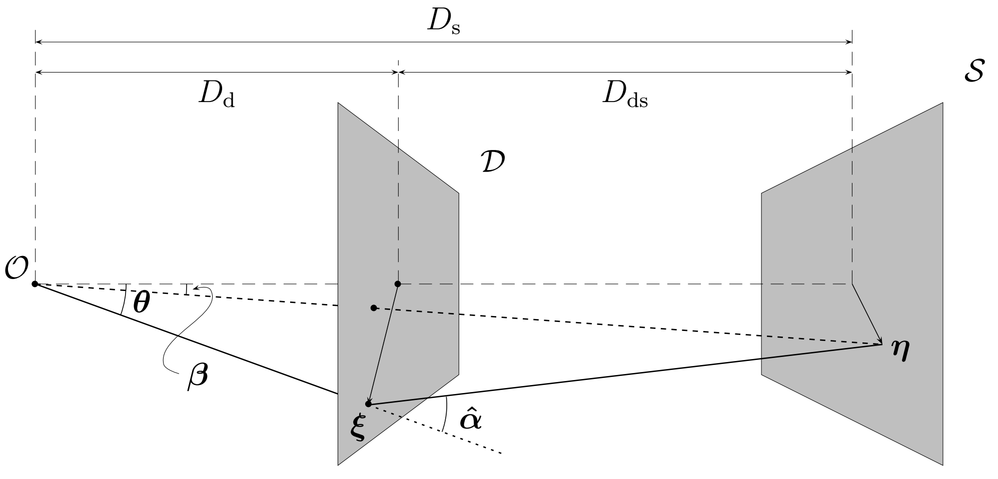


Figure 2.1: A sketch of a typical gravitational lens situation.

denote $\boldsymbol{\eta}$ to be the distance of the source from the optical axis in the source plane, and $\boldsymbol{\xi}$ the 'impact parameter' (more precisely the impact vector) in the lens plane. We can write by pure geometrical considerations, assuming a single lens plane

$$\boldsymbol{\eta} = \frac{D_{\rm s}}{D_{\rm d}}\boldsymbol{\xi} - D_{\rm ds}\,\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi})\;, \tag{2.1}$$

or equivalently

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{\rm ds}}{D_{\rm s}}\hat{\boldsymbol{\alpha}}(D_{\rm d}\boldsymbol{\theta}) = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta})\;. \tag{2.2}$$

$\boldsymbol{\alpha}(\boldsymbol{\theta})$ is referred to as the scaled deflection angle. The last equation is very important for lensing, thus its name *lens equation* or *ray-tracing equation.* It gives us the angular position $\boldsymbol{\theta}$ at which we see the lensed images for a given source that is located at $\boldsymbol{\beta}$. The scaled deflection angle $\boldsymbol{\alpha}(\boldsymbol{\theta})$ usually depends on $\boldsymbol{\theta}$ in a non-linear fashion, and therefore it is possible to obtain multiple solutions of $\boldsymbol{\theta}$ for a single source position $\boldsymbol{\beta}$; in such case we are talking about multiple imaging or strong lensing.

Consider now the deflection angle of a point mass M. If the impact parameter ξ (i.e. the absolute value of $\boldsymbol{\xi}$) of a light ray is much larger than the Schwarzschild radius of the lens, then from GR it follows that:

$$|\hat{\boldsymbol{\alpha}}| = \frac{4\,G\,M}{c^2\,\xi}. \tag{2.3}$$

As already mentioned before, the thin-lens approximation is justified here, and we can calculate the deflection angle of a three-dimensional mass distribution by adding up (integrating) the contributions of individual mass elements (each regarded as a point mass) resulting in a two-dimensional deflection angle. Defining the projected mass density $\Sigma(\boldsymbol{\xi})$:

$$\Sigma(\boldsymbol{\xi}) \equiv \int {\rm d}r_3\;\rho(\boldsymbol{\xi}, r_3)\;, \tag{2.4}$$

the deflection angle can be written in a compact form as

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4\,G}{c^2} \int_{\Re^2} \mathrm{d}^2\xi'\, \Sigma(\boldsymbol{\xi}') \frac{\boldsymbol{\xi} - \boldsymbol{\xi}'}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|^2} \;. \tag{2.5}$$

The properties of the mass distribution of the lens enter the lens equation only in the form of a scaled deflection angle $\boldsymbol{\alpha}$, which depends in turn on the projected mass density $\Sigma(D_\mathrm{d}\boldsymbol{\theta})$ and on the geometry of the system (i.e. on angular diameter distances between source, observer, and the lens). The 'strength' of the lens can be described in terms of a scaled surface mass density $\kappa(\boldsymbol{\theta})$ where

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(D_\mathrm{d}\boldsymbol{\theta})}{\Sigma_\mathrm{cr}} \;, \tag{2.6}$$

where Σ_cr is the critical surface mass density given by

$$\Sigma_\mathrm{cr} = \frac{c^2\, D_\mathrm{s}}{4\,\pi\, G\, D_\mathrm{d}\, D_\mathrm{ds}} \;. \tag{2.7}$$

For $\kappa(\boldsymbol{\theta}) \ll 1$ we are dealing with the *weak lensing* regime. Such a lens namely can not produce multiple images. A sufficient (but not necessary) condition for producing multiple images is $\kappa(\boldsymbol{\theta}) > 1$ for at least one θ, this lensing regime is called the *strong lensing* regime. The distinction between the two regimes is not always sharp and we will elaborate on this later.

Σ_cr depends upon the cosmological parameters. Rewriting $\boldsymbol{\alpha}$ in terms of κ gives

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi} \int_{\Re^2} \mathrm{d}^2\theta'\, \kappa(\boldsymbol{\theta}')\, \frac{\boldsymbol{\theta} - \boldsymbol{\theta}'}{|\boldsymbol{\theta} - \boldsymbol{\theta}'|^2} \;. \tag{2.8}$$

Now recall that $\boldsymbol{\theta}$ is a dimensionless quantity (it is a vector measured in radians). Since the lens equation (2.2) depends only upon $\kappa(\boldsymbol{\theta})$, one cannot measure the physical surface mass density Σ, without knowing Σ_cr, thus without knowing the redshifts of the lens and source. In addition, there is a (relatively weak) dependence on the cosmological parameters in Σ_cr.

The form of the equation (2.8) implies that deflection angle can be written in the gradient form

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \nabla\psi(\boldsymbol{\theta}) \;, \tag{2.9}$$

with $\psi(\boldsymbol{\theta})$ given as

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi} \int_{\Re^2} \mathrm{d}^2\theta'\, \kappa(\boldsymbol{\theta}')\, \ln|\boldsymbol{\theta} - \boldsymbol{\theta}'| \;. \tag{2.10}$$

Lens equation (2.2) thus represent the gradient mapping, analogous to what we are used to in optics. Further the gradient of the deflection vector and thus $\psi(\boldsymbol{\theta})$ satisfies the analogue to the Poisson equation from the Newtonian gravitational theory:

$$\nabla^2\psi(\boldsymbol{\theta}) = 2\kappa(\boldsymbol{\theta}) \;. \tag{2.11}$$

For a source with a size much smaller than the typical length on which the lensing properties change, the lens equation can be linearised. The lens mapping can then be described by Jacobian matrix $\mathcal{A}(\boldsymbol{\theta})$

$$\mathcal{A}(\boldsymbol{\theta}) = \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\theta}} = \left(\delta_{ij} - \frac{\partial^2 \psi(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}\right) = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} , \tag{2.12}$$

where the complex quantity $\gamma = \gamma_1 + \mathrm{i}\gamma_2$ is called *shear* and is defined using second-order derivatives of ψ

$$\gamma_1 = \frac{1}{2}(\psi_{,11} - \psi_{,22}) \ , \quad \gamma_2 = \psi_{,12} = \psi_{,21} \ . \tag{2.13}$$

The interpretation of shear and scaled surface mass density (also called *convergence*) now becomes clear. Consider a small circular source and a lens with $\gamma_1(\boldsymbol{\theta}) = \gamma_2(\boldsymbol{\theta}) = 0$. The source would only scale in size and remain circular. In the presence of shear, however, the source shape (as well as size) will changed (this is presented later also in Fig. 2.5). We define shear as a complex quantity, in addition we remind the reader that shear is not a vector. Since it transforms as $\gamma \rightarrow |\gamma|\, \mathrm{e}^{2\mathrm{i}\varphi}$ under the rotation of coordinate system, the shear can also be represented as a rank two symmetric tensor.

Note that the above approach is valid only in the case of a single lens-plane; in the case of multiple lens planes the potential ψ is in general not sufficient to describe the lens system, and the Jacobian matrix is no longer symmetric (Blandford & Narayan, 1986).

Lensing locally transforms circular sources into ellipses with semi axes being the eigenvalues of Jacobi matrix $\mathcal{A}(\boldsymbol{\theta})$. The ratio of the solid angles a source occupies in the source plane $\mathcal{S}$ and in the lens plane $\mathcal{D}$ is given by the determinant of the matrix $\mathcal{A}$. Therefore the determinant corresponds to the inverse of the magnification factor μ, i.e.

$$\mu(\boldsymbol{\theta}) = \frac{1}{\det \mathcal{A}} = \frac{1}{(1-\kappa)^2 - |\gamma|^2} \ . \tag{2.14}$$

In other words, if an isophote in the source plane encloses an area a (small enough, so we can apply local linearisation), the area enclosed by the same isophote in the lensed image will be μa. This comparison can be made, since lensing conserves the surface brightness.

Finally a semantic remark. Above we have shortly mentioned the question of which lenses are called weak, and which are called strong. In general, there is no strict definition of weak lensing and strong lensing. For our purposes however, we will call studies where we deal with multiple imaging ‘strong lensing’, and ‘weak lensing’ corresponds to analyses where lensing needs to be treated in statistical sense. In practice, if we can observe only single images of the sources (as in the case of weak lensing), we need to consider an ensemble of such sources to estimate the properties of the lens. For multiple imaging only a single source is sufficient. Now we see that even if lens is ‘strong’ enough to produce multiple images, one can treat the background population of sources by either using strong lensing (for sources that happen to form observable multiple images, thus using images which are not independent) or weak lensing principles (for all sources that are independent from each other). In the ideal case one combines both methods to obtain detailed results on the lensing properties. This is particularly useful for clusters where we observe many background sources that (projected onto the lens plane) lie within distances from the cluster centre where lensing effect is still noticeable, even though the sources are not multiply imaged.

2.2 Strong gravitational lensing

Now we concentrate on the lensing regime where multiple images are formed. Mathematically this means that multiple solutions of the lens equation (2.2) exist and the topic of this section will be to introduce the formalism of strong lensing.

For convenience we introduce dimensionless coordinates

$$\boldsymbol{x} = \frac{\boldsymbol{\xi}}{\xi_0} \quad ; \quad \boldsymbol{y} = \frac{\boldsymbol{\eta}}{\eta_0} \,, \tag{2.15}$$

where the length scale in the image plane ξ_0 can be chosen arbitrarily and the corresponding length scale in the source plane η_0 is given by $\eta_0 = \xi_0 D_\mathrm{s}/D_\mathrm{d}$. The lens equation in new coordinates reads

$$\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{\alpha}(\boldsymbol{x}) \tag{2.16}$$

and can be rewritten in terms of a gradient mapping $\boldsymbol{x} \mapsto \boldsymbol{y}$. Introducing a scalar function ϕ (also called Fermat potential for lensing)

$$\phi(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2} \left(\boldsymbol{x} - \boldsymbol{y}\right)^2 - \psi(\boldsymbol{x}) \tag{2.17}$$

we see that images of a source located at $\boldsymbol{y}$ are formed at those points $\boldsymbol{x}$ where $\nabla\phi = 0$ and $\det\mathcal{A}$ does not vanish. In a typical non-singular elliptical lens (having monotonically declining density profile) with a source, lens centre and the observer aligned, the equation $\nabla\phi = 0$ would have five distinctive solutions $\boldsymbol{x}_i^{(0)}$, each corresponding to one of the images. At critical curves, i.e. where $\det\mathcal{A} = 0$, the lens mapping properties change and we will study them in the next section.

2.2.1 Lens mapping close to the critical curves

In general, the Jacobian matrix $\mathcal{A}(\boldsymbol{\theta})$ is not necessary invertible, i.e. there may exist points in the lens plane $\boldsymbol{\theta}$ where the determinant of the matrix vanishes. They form closed curves and are called *critical curves*. If we map them back to the source plane we obtain *caustic curves*. The image of a point source that lies on the caustic curve is formally infinitely magnified. In practice, any deviation from a source being point like causes a finite (still very high) magnification. Physically, however, even point sources would have finite magnification, since in all derivations we use a description of lensing based on geometrical optics, which fails very close to critical curves. Namely two images located close to the critical curve have a very small differential time delay and one expects interferences to occur if this time delay is smaller than the coherence time of the light (Schneider et al., 1992). Nevertheless, the corrections in the framework of wave optics are very small and the geometric description is more than adequate for the purposes we are dealing with here.

We illustrate the image formation for a source crossing the caustic curve in Fig. 2.2 through an example of a typical non-singular isothermal elliptical (NIE) lens. An isothermal lens model throughout this work refers to a lens whose 3-D density follows $\rho(r) \propto r^{-2}$ and the surface mass density scales as $\kappa(\boldsymbol{\theta}) \propto |\boldsymbol{\theta}|^{-1}$ (in the spherically symmetric case). An NIE model is an approximate description for the projected mass density of a typical elliptical galaxy. In the left panel the lens plane is plotted and the critical curves are presented as solid lines. Through the lens mapping, the inner critical curve (i.e. inner ellipse) transforms onto the outer caustic curve (see right panel of the same figure, presenting the source plane) and the outer critical curve transforms into the inner, asteroid caustic. The asteroid caustic has four fold caustics (curved lines) and where two fold caustics join we get a so-called cusp.

The sketch of the formation of such asteroid caustic is given in Fig. 2.3. It shows a region of the lens plane around the critical curve which is pleated and projected onto the source plane, forming the caustic. If one now considers a source inside such a caustic (denoted as a small ellipse in the plot), it will correspond to two extra images in the lens plane. Therefore at each caustic crossing the image multiplicity changes by two. The 'inside' of a caustic is defined in such a way that the number of images is bigger than the number of images 'outside'. A typical *circular symmetric, singular* lens (with an isothermal or steeper profile) has two critical curves where the determinant vanishes. The caustic of the inner critical curve is a circle and the caustic of the outer critical curve degenerates into a point due to the circular symmetry. In Fig. 2.2 we present the case of a non-singular elliptical lens, whose critical curves are ellipses. The caustic of the inner critical curve is an ellipse and the caustic of the outer critical curve has the before mentioned asteroid shape. Since images forming close to the inner critical curve are distorted preferentially in the radial and those forming close to the outer critical curve preferentially in the tangential direction, the inner critical curve is called 'radial' and the outer one 'tangential'.

A source located in the centre of the asteroid caustic (c.f. green or dark blue image in Fig. 2.2) lies inside the inner as well as the outer caustic and therefore we observe 2 times 2 extra images, in total thus five. They are presented in the left panel and the fifth image (close to the centre) is highly demagnified. For the source close to a cusp (dark blue image, top panel) three of the images get highly magnified (which will eventually merge into a single remaining image when the source crosses the cusp). In the case of the source moving close to the fold (dark-blue image, bottom panel) two of the images are highly magnified and will eventually merge and dissappear on the fold caustic crossing. Finally we concentrate on sources located close to and inside a fold (dark blue source in the lower panel of Fig. 2.2) or a cusp (dark blue source in the upper panel of Fig. 2.2). For the non-singular elliptical lens both source positions will have five solutions of the lens equation (i.e. five images) correspond to two minima, two saddle points and a maximum of the Fermat potential. The Fermat potential ϕ for two different source positions is represented in Fig. 2.4. For both

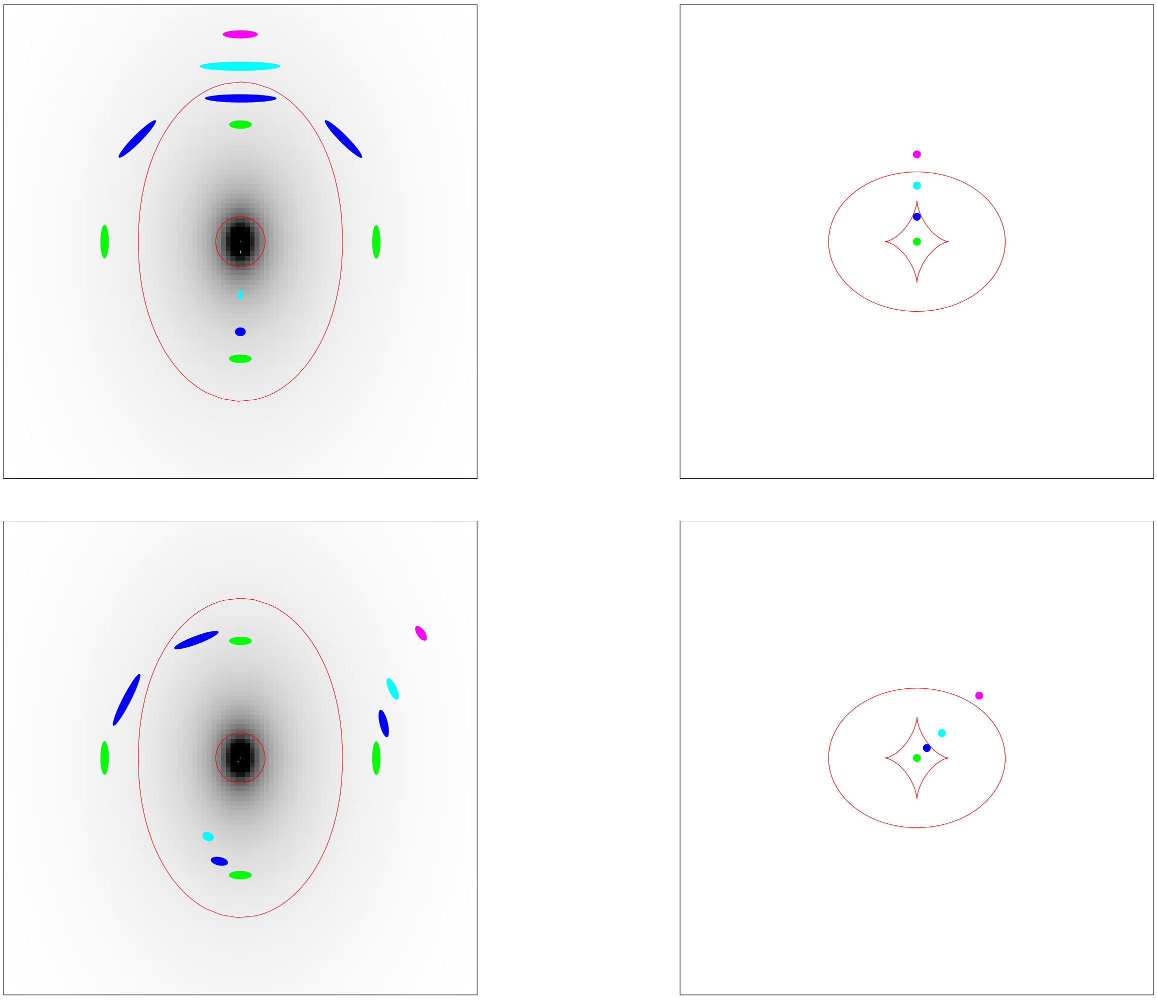


Figure 2.2: Compact source crossing a cusp caustic (top) and a fold caustic (bottom) of a non-singular elliptical lens. The diagram on the left shows critical curves and image positions and the diagram on the right shows caustics and source positions.

configurations (fold in panel **(a)** and cusp in **(b)**) one can clearly see how images form at the minima (red) and saddle points in both cases. The maximum is in both cases highly demagnified and presented by a single dot (barely visible) close to the centre. In summary:

Saddle point images are images corresponding to the saddle points of ϕ and have negative parity, i.e. $\det \mathcal{A} < 0$. As a consequence their magnification factor is negative (which can sometimes be a source of confusion, as we will speak then of images with negative magnifications - though their flux is proportional to the absolute value of magnification). They are also sometimes referred to as Type II images.

Minima or Type I images have a positive parity and trace, i.e. $\det \mathcal{A} > 0$ and $\mathrm{tr}\mathcal{A} > 0$. They are never demagnified; i.e. $|\mu| = \mu \geq 1$ (for an infinitesimal source).

Maxima are Type III images and also have a positive parity, i.e. $\det \mathcal{A} > 0$, however $\mathrm{tr}\mathcal{A} < 0$. For example, in the case of a non-singular isothermal lens the central

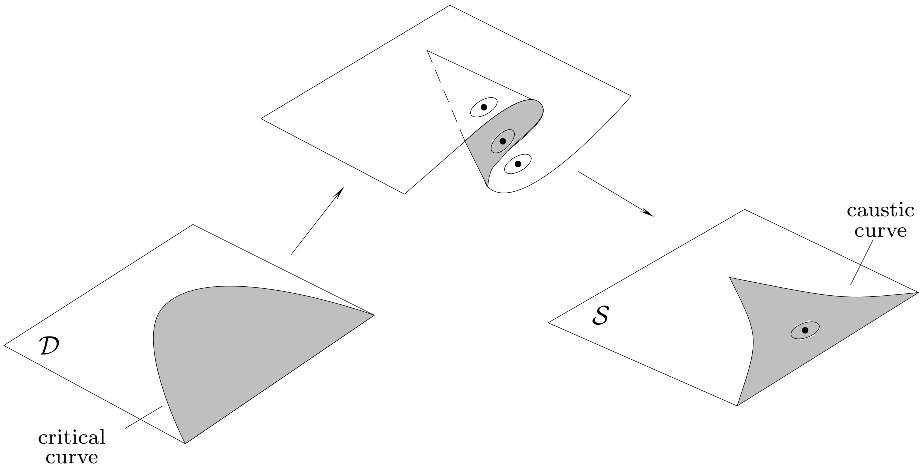


Figure 2.3: A sketch of the process of an asteroid caustic formation. A local region of the lens plane gets pleated and projected onto the source plane. A single source in the source plane (ellipse) has two extra corresponding images in the image plane (adapted from Petters et al., 2001; Gilmore, 1981).

image, corresponding to a maxima, will be highly demagnified and will in general escape observations.

Cusp and fold relations

These two specific configurations are interesting, since the images form in a configuration satisfying model-independent predictions for the magnifications of highly magnified images (Blandford & Narayan, 1986; Blandford, 1990; Schneider & Weiss, 1992; Mao, 1992).

These types of images form at positions $\boldsymbol{x}$ satisfying $\det\mathcal{A}(\boldsymbol{x}) \simeq 0$. For such images the general mapping can be approximated by Taylor expanding the Fermat potential ϕ around $\boldsymbol{x}_i^{(0)}$ (exactly on the critical curve) including the third order in $\boldsymbol{x}$ (for details on the rules for this truncation and for full derivation of what follows see Chapter 6 in Schneider et al., 1992).

Fold images: The approximation of the general mapping around the fold is thus given by a second order polynomial mapping. By studying the inversion of this transformation the properties of the two images that form close to the fold can then be studied. If the source approaches the fold from the 'inner' side, two of the lensed images form closer and closer together and finally merge and disappear when the source crosses the caustic. In the limit of the source approaching the fold, the magnifications of the two merging images (denoted by A and B) have opposite parity, however the same absolute value. Therefore asymptotically the following relation holds,

$$R_{\text{fold}} = \frac{|\mu_A + \mu_B|}{|\mu_A| + |\mu_B|} \to 0 \, . \tag{2.18}$$

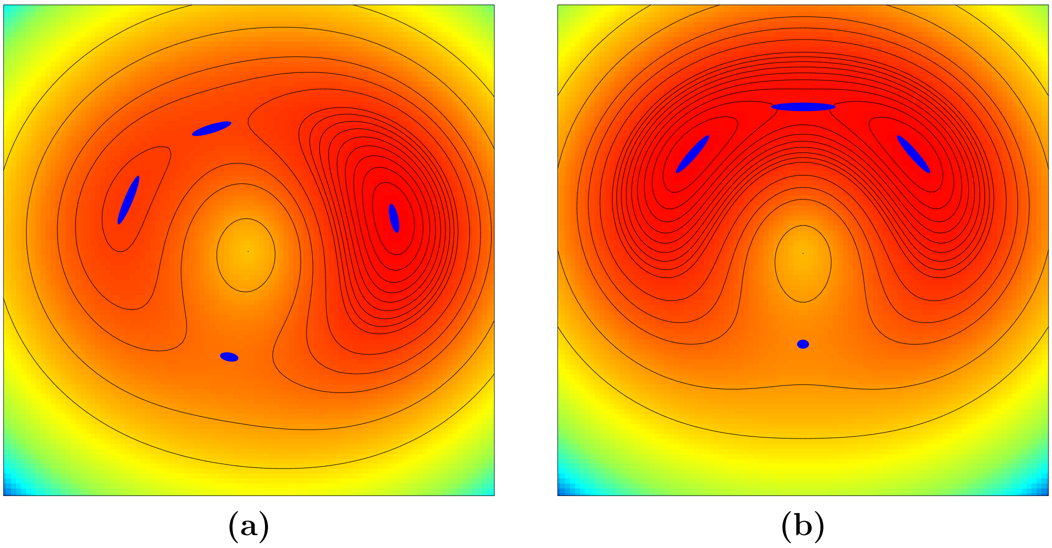


Figure 2.4: The Fermat potential $\phi(\boldsymbol{x}, \boldsymbol{y})$ for two different source positions $\boldsymbol{y}$ plotted in colour (red = most negative) and contours. Panel **(a)** represents a source close to a fold, and **(b)** close to the cusp.

We scaled the relation for convenience, since we only observe fluxes and not magnifications. The absolute value of magnification close to the fold varies inversely as the square root of the distance between the source and the (closest) point on the caustic.

Cusp images: Similarly the properties of the mapping near a cusp can be considered. The resulting mapping is a third order polynomial mapping. For a given source position (close to the cusp) three solutions represent the three merging images. Two of them will have the same parity (we denote these two by A and C), and the third one (B) the opposite parity. If the source approaches the cusp from inside, image B will move on a line perpendicular to the critical curve, and A and C approach the critical curve tangentially. The sum of magnifications of all three images vanish when the source approaches the cusp,

$$R_{\mathrm{cusp}} = \frac{|\mu_A + \mu_B + \mu_C|}{|\mu_A| + |\mu_B| + |\mu_C|} \rightarrow 0\ . \tag{2.19}$$

Both, the fold and cusp relations are a very powerful, model independent predictions. It would be useful to know more accurately the asymptotic behaviour of R_{cusp}, unfortunately the mathematics is too cumbersome for this to be done in practice. In the next chapter we study the behaviour of R_{cusp} numerically. In particular we study to what extent this relation is broken when moving the source further away from the cusp and how much it can be influenced by the substructure on scales smaller than the image separation in the main lens potential.

2.3 Statistical (weak) gravitational lensing

In the strong lensing regime we were dealing with the multiple images of a single source. Although we do not know the properties of the source itself, this unknown can be eliminated through constraints from multiple images. In weak lensing this is no longer the case. The weak lensing sources are background galaxies and they are intrinsically elliptical. An elliptical source transforms into an elliptical image under the influence of lensing and therefore we have almost no extra information to constrain the underlying mass distribution using individual images. For this reason lensing effects need to be treated in a statistical sense (performing ensemble averaging over the images) and is often also called statistical (weak) lensing.

Throughout this thesis we use the complex ellipticity $\epsilon = \epsilon_1 + \mathrm{i}\epsilon_2$ defined in terms of the second brightness moments Q_{ij} as

$$\epsilon \equiv \frac{Q_{11} - Q_{22} + 2\mathrm{i}Q_{12}}{Q_{11} + Q_{22} + 2(Q_{11}Q_{22} - Q_{12}^2)^{1/2}} \tag{2.20}$$

(see e.g. Blandford et al., 1991), where

$$Q_{ij} = \frac{\int \mathrm{d}^2\theta \, q_I[I(\boldsymbol{\theta})] \, (\theta_i - \bar{\theta}_i) \, (\theta_j - \bar{\theta}_j)}{\int \mathrm{d}^2\theta \, q_I[I(\boldsymbol{\theta})]} \ , \quad i, j \in \{1, 2\} \ . \tag{2.21}$$

$I(\boldsymbol{\theta})$ is the surface brightness of a galaxy image at the position $\boldsymbol{\theta}$ and $q_I(I)$ is a suitably chosen weight function (for practical reasons this definition is usually modified such that the weight function only depends on the position $\boldsymbol{\theta}$ and not on the surface brightness). For an image with circular isophotes we thus have $Q_{11} = Q_{22}$ and $Q_{12} = 0$. The trace of Q_{ij} corresponds to the size of the image, if $q_I(I) = H(I - I_0)$ is a Heaviside step function and the size is given by the solid angle enclosed by the limiting isophote I_0.

In terms of semi-major axis a and semi-minor axis b of an ellipse, the absolute value of the ellipticity is given by $|\epsilon| = \frac{a-b}{a+b}$ and φ is the angle between the x_1 axis of the coordinate frame and the semi major axis (see Fig. 2.5). Note that an ellipse turns into itself by a rotation of 180°, therefore we have similar transformation properties as in the case of the shear, thus $\epsilon = |\epsilon| \, \mathrm{e}^{2\mathrm{i}\varphi}$. For visualisation of ϵ see the left panel of Fig. 2.6.

We have already noted that under the influence of convergence and shear, images are anisotropically stretched. Using the fact that the surface brightness of a source is conserved when lensed, the second brightness moments of the source $Q^{(s)}$ and its image Q at position $\boldsymbol{\theta}$ are related through the transformation

$$Q^{(s)} = \mathcal{A} \, Q \, \mathcal{A}^T = \mathcal{A} \, Q \, \mathcal{A} \ , \tag{2.22}$$

where $\mathcal{A} \equiv \mathcal{A}(\boldsymbol{\theta}, z)$ is the Jacobi matrix of the lens equation. We include the redshift dependence here to remind the reader that γ and κ in general depend upon the redshift of the source and the lens.

For convenience we rewrite the Jacobi matrix in the following form

$$\mathcal{A} = [1 - \kappa(\boldsymbol{\theta}, z)] \begin{pmatrix} 1 - g_1(\boldsymbol{\theta}, z) & -g_2(\boldsymbol{\theta}, z) \\ -g_2(\boldsymbol{\theta}, z) & 1 + g_1(\boldsymbol{\theta}, z) \end{pmatrix} \ , \tag{2.23}$$

where we introduced the *reduced shear* $g(\boldsymbol{\theta}, z)$ as

$$g(\boldsymbol{\theta}, z) = \frac{\gamma(\boldsymbol{\theta}, z)}{1 - \kappa(\boldsymbol{\theta}, z)} \tag{2.24}$$

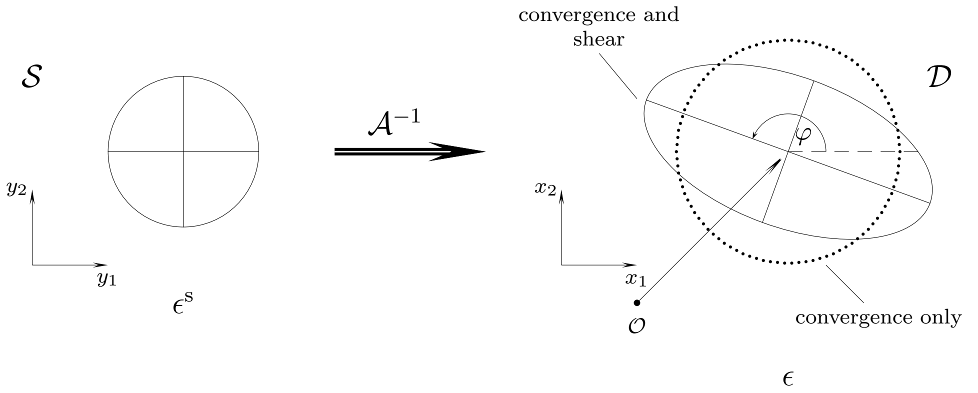


Figure 2.5: Circular source ϵ^{s} transforms into an ellipse under the influence of convergence and shear (adapted from Narayan & Bartelmann, 1997).

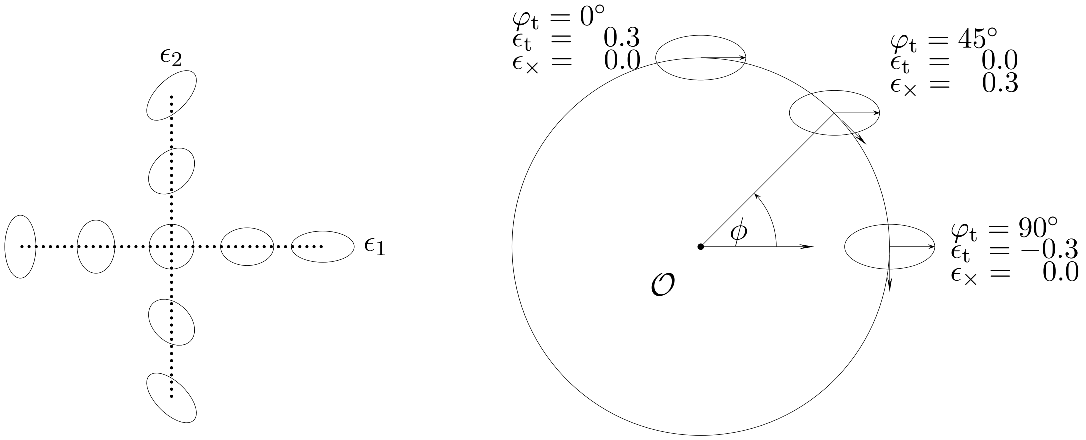


Figure 2.6: (left) Visualisation of components of the ellipticity $\epsilon = \epsilon_1 + \mathrm{i}\epsilon_2 = \frac{a-b}{a+b}\mathrm{e}^{2\mathrm{i}\varphi}$. (right) Visualisation of the tangential ϵ_{t} and cross $\epsilon\times$ component of the ellipticity. They are defined as $\epsilon_{\mathrm{t}} = -\Re\left[\epsilon\,\mathrm{e}^{-2\mathrm{i}\phi}\right]$ and $\epsilon_\times = -\Im\left[\epsilon\,\mathrm{e}^{-2\mathrm{i}\phi}\right]$. φ_{t} is the angle from the tangent to the major axis of the source (adapted from Courbin & Minniti, 2002).

The transformation between the source ellipticity $\epsilon^{(s)}$ and image ellipticity ϵ is given as a function of reduced shear $g(\boldsymbol{\theta}, z)$ (see Seitz & Schneider, 1997)

$$\epsilon^{(s)} = \begin{cases} \dfrac{\epsilon - g(\boldsymbol{\theta}, z)}{1 - g^*(\boldsymbol{\theta}, z)\epsilon} & \text{for } \left|g(\boldsymbol{\theta}, z)\right| \leq 1 \ , \\ \\ \dfrac{1 - g(\boldsymbol{\theta}, z)\epsilon^*}{\epsilon^* - g^*(\boldsymbol{\theta}, z)} & \text{for } \left|g(\boldsymbol{\theta}, z)\right| > 1 \ . \end{cases} \tag{2.25}$$

The inverse transformation is simply given by swapping ϵ^{s} with ϵ and by substituting $-g$ with g.

Since we will work with sources at different redshifts, we factorize the redshift dependence of the lens convergence κ and of the shear γ as

$$\kappa(\boldsymbol{\theta}, z) = Z(z)\kappa(\boldsymbol{\theta}) \ , \qquad \gamma(\boldsymbol{\theta}, z) = Z(z)\gamma(\boldsymbol{\theta}) \ , \tag{2.26}$$

where $Z(z)$ is the so-called 'cosmological weight' function:

$$Z(z) \equiv \frac{\lim_{z\to\infty} \Sigma_{\mathrm{cr}}(z_{\mathrm{d}}, z)}{\Sigma_{\mathrm{cr}}(z_{\mathrm{d}}, z)} \, \mathrm{H}(z - z_{\mathrm{d}}) \ . \tag{2.27}$$

The Heaviside step function $\mathrm{H}(z - z_{\mathrm{d}})$ accounts for the fact that the sources that are in front of the lens plane are not lensed. Note that, as suggested by its name, $Z(z)$ is cosmology dependent. For an Einstein-de Sitter cosmology it becomes

$$Z(z) = \frac{\sqrt{1+z} - \sqrt{1+z_{\mathrm{d}}}}{\sqrt{1+z} - 1} \, \mathrm{H}(z - z_{\mathrm{d}}) \ . \tag{2.28}$$

In Lombardi & Bertin (1999) cosmological weights were calculated for different cosmologies. The authors have shown that the differences between Einstein-de Sitter and the nowadays assumed standard cosmology (i.e. $\Omega = 1$, $\Omega_{\mathrm{m}} = 0.3$, $\Omega_\Lambda = 0.7$) are not significant for the purpose of cluster-mass reconstructions.

Following (2.24) the redshift-dependent reduced shear $g(\boldsymbol{\theta}, z)$ is given by

$$g(\boldsymbol{\theta}, z) = \frac{Z(z)\gamma(\boldsymbol{\theta})}{1 - Z(z)\kappa(\boldsymbol{\theta})} \ . \tag{2.29}$$

For an intrinsically circular source $\epsilon^{\mathrm{s}} = 0$, the image ellipticity for $|g(\boldsymbol{\theta}, z)| < 1$ case is simply $\epsilon = g$. Reduced shear is thus a measure for the image ellipticity of an intrinsically circular source. As mentioned before, background sources in the case of statistical (weak) lensing are galaxies and therefore thir intrinsic ellipticities are far from circular. However, the orientations of galaxies is very close to random, there is no preferred direction in the sky to point at (an exception being possibly the galaxies that belong e.g. to the same cluster, the effect of this intrinsic alignment is however negligible in the case of cluster lensing). Therefore by performing ensemble averaging of the image ellipticities the properties of the lens can be studied. More precisely, we can Taylor expand the expression (2.25) (for the case of $|g(\boldsymbol{\theta}, z)| < 1$ and $\left|\epsilon^{(s)}\right| < 1$ by definition)

$$\epsilon(z) = \frac{\epsilon^{(s)} + g(\boldsymbol{\theta}, z)}{1 + g^*(\boldsymbol{\theta}, z)\epsilon^{(s)}} = \left(\epsilon^{(s)} + g(\boldsymbol{\theta}, z)\right) \sum_{k=0}^{\infty} (-1)^k (g^*(\boldsymbol{\theta}, z))^k (\epsilon^{(s)})^k \,. \tag{2.30}$$

The consequence of isotropic distribution of intrinsic galaxy ellipticities can be reworded by stating that the source ellipticity distribution $p_{\epsilon^{\mathrm{s}}}(\epsilon^{\mathrm{s}})$ depends only on the modulus of ϵ^{s}. Equivalently this implies $\left\langle {\epsilon^{(s)}}^k \right\rangle = 0$ for each $k > 0$, where $\langle\rangle$ denotes the ensemble averaging. By performing ensemble averaging of (2.30) we note that all terms except $k = 0$ vanish. The expectation value for the image ellipticity at redshift z becomes

$$\left\langle \epsilon(z) \right\rangle = \begin{cases} g(\boldsymbol{\theta}, z) & \text{if } \left|g(\boldsymbol{\theta}, z)\right| < 1 \\ \dfrac{1}{g^*(\boldsymbol{\theta}, z)} & \text{otherwise .} \end{cases} \tag{2.31}$$

This relation is particularly simple due to the convenient definition of ϵ in terms of Q.

Finally a note on the tangential and cross component of the ellipticity. Suppose that in a given direction ϕ (see Fig. 2.6) we observe an image with ellipticity ϵ. We then define the tangential ϵ_{t} and cross component $\epsilon_\times$ of the ellipticity as $\epsilon_{\mathrm{t}} = -\Re\left[\epsilon\, \mathrm{e}^{-2\mathrm{i}\phi}\right]$ and $\epsilon_\times = -\Im\left[\epsilon\, \mathrm{e}^{-2\mathrm{i}\phi}\right]$. In some cases it is more convenient to work with these components, since for a circularly-symmetric mass distribution the resulting expectation value of lensed ellipticity will be always oriented tangent to the direction of the lens centre (i.e. centre of symmetry) and therefore $\left\langle \epsilon_\times \right\rangle = 0$. Even without assuming circular symmetry of the lens, the images will still be preferentially tangentially aligned to the direction of the mass concentration and so $\langle \epsilon_{\mathrm{t}} \rangle > 0$ and $\langle \epsilon_\times \rangle \simeq 0$. In practice, the signal in the $\langle \epsilon_\times \rangle$ component indicates the presence of systematic (noise) effects in the data.

3

Strong Gravitational Lensing & CDM substructure

Whereas the current Cold Dark Matter (CDM) paradigm for structure formation is widely accepted, two major problems for CDM still remain. While simulations predict cuspy dark matter halos (Moore, 1994), observed rotation curves of low surface brightness galaxies indicate that their dark matter halos have cores (Kravtsov et al., 1998; Swaters et al., 2000; van den Bosch & Swaters, 2001; de Blok & Bosma, 2002). The other is the apparent over-prediction of the small-scale power in CDM simulations. As was shown by Moore et al. (1999) and Klypin et al. (1999), the number of satellite halos seen in N-body simulations appears to far exceed the number of dwarf galaxies observed around the Milky Way. Particular discrepancies have been found for satellite masses $\lesssim 10^9 M_\odot$.

If one takes the observed satellites of the Milky Way as the total amount of substructure present, one needs to modify the CDM paradigm and include e.g. self-interacting dark matter (Spergel & Steinhardt, 2000) or a cut-off in the power spectrum (which might occur in a warm dark matter-dominated universe, Bode et al. 2001; Colín et al. 2000). Another possibility pointed out by Stoehr et al. (2002) and Hayashi et al. (2003) is that the problem might be the misidentification of the observed satellites with the substructure seen in N-body simulations. Many mechanisms have been proposed which might prevent star formation in halos of mass $\lesssim 10^9 M_\odot$, making them too dim for observations (Bullock et al., 2000; Benson et al., 2002; Somerville, 2002; Springel & Hernquist, 2003). Zentner & Bullock (2003) investigate a semi-analytic model that describes the properties of galaxy-sized halo substructure. The effects of the "tilt" and overall normalisation of the primordial power spectrum are discussed. All these predictions need to be tested within our own Galaxy as well as at cosmological distances.

Gravitational lensing is at present the only tool for investigating CDM substructure in galaxies outside the local group. As first noted by Mao & Schneider (1998), mass-substructure other than stars on scales less than the image separation can substantially affect the observed flux ratios in strong gravitational lens systems. Chiba (2002), Dalal & Kochanek (2002), Metcalf & Madau (2001), Metcalf & Zhao (2002), Keeton (2001), Mao & Schneider (1998) and Bradač et al. (2002) have argued that substructure can provide the explanation for the flux anomalies in various systems. Dalal & Kochanek (2002) further conclude that the amount of substructure needed to explain the flux ratios of quadruply-imaged systems broadly agrees with the CDM predictions. At least for some systems the flux mismatches are probably not just an artifact of oversimplified macromodels of the main lens galaxy (see e.g. Evans & Witt 2003; Metcalf & Zhao 2002). As discussed by Keeton (2003) and Chen et al. (2003), fluxes can be further affected by clumps of matter at a redshift different from that of the lens, lying along the line of sight between the observer and the source; however, this effect is not dominant. It is also possible that the small scale structure does not consist of compact CDM clumps, also tidal streams or offset disc components can affect the flux ratios (see Möller et al. 2003; Quadri et al. 2003).

Keeton (2001) and Gaudi & Petters (2002) recently focused on the magnification relations that should be satisfied by particular four-image geometries (so called "fold" and "cusp" configurations, described in detail in Sect. 2.2.1). Cusp and fold relations are model-independent predictions for the magnifications of highly magnified images (Blandford & Narayan, 1986; Blandford, 1990; Schneider & Weiss, 1992; Mao, 1992). Strictly speaking, however, they hold only for ideal "fold" or "cusp" configurations and it is therefore difficult in some cases to disentangle the effects of the source being further away from the cusp from that of substructure, purely by employing these relations.

The influence of substructure can not only be seen on image flux ratios, but also in the structure of multiple-imaged jets. The lens system B1152+199 is a case of doubly-imaged jets, one of which appears bent, whereas the other is not (Metcalf, 2002). Another explanation is that an intrinsic bend in the jet is simply magnified in one image, and produces only a small effect in the other.

Flux ratio anomalies can also be introduced by propagation effects in the interstellar medium (ISM) in the lens galaxy. Microlensing can change the flux ratios not only in the optical (e.g. Woźniak et al. 2000), but also at radio wavelengths (Koopmans & de Bruyn, 2000). Flux ratios can further be affected by galactic scintillation and scatter broadening (Koopmans et al., 2003). Fortunately, these effects are frequency dependent and one can recognise them using multi-frequency observations.

These electromagnetic phenomena are similar for images of different parities. For substructure, however, Schechter & Wambsganss (2002) found that magnification perturbations should show a dependence on image parity. Microlensing simulations showed that the probability distributions for magnifications of individual images are no longer symmetric around the unperturbed magnification. The distribution can, depending on image parity, become highly skewed in the direction of demagnification. This skewness causes a substantial probability for the brightest saddle point image to be demagnified. The two saddle point images in the quadruply-imaged system are the ones with negative parity, the other two have positive parities and are called minima (they correspond to the local minima and saddle points in the arrival time surface - see Sect. 2.2.1 for details). Observed lens systems seem to show this image parity dependence (Kochanek & Dalal, 2004), and this indicates that the flux ratio anomalies arise from gravitational lensing, rather than propagation effects.

All these possible effects on flux ratios have placed some doubt as to whether the existence of substructure can be rigorously tested with strong lensing and on what the expected signal is. Several groups have been testing these predictions using a semi-analytic prescription for substructure (Metcalf & Madau, 2001; Keeton et al., 2003; Dalal & Kochanek, 2002; Kochanek & Dalal, 2004).

In this thesis, however, we use the direct output of different halos and projections of a galaxy obtained in N-body+gasdynamics simulations as a lens galaxy. By using a semi-analytic prescription one overcomes the biggest problem we face, namely the problem of shot-noise in N-body simulations. Further, the problem of modelling becomes simpler, because one has an analytic model for the underlying smooth component of the mass distribution of the lens. However, by using the direct output of an N-body simulated galaxy, one does not make any assumption about the mass profiles of the macro model, or the substructure. Down to the resolution scales of the simulation we therefore believe we have a better comparison with a realistic galaxy.

This chapter of the thesis is in large extent published in Bradač et al. (2004d). In what follows we will first describe in more detail the theory of the influences of substructure on strong gravitational lensing. Further we will present the N-body simulations and the methods we used to predict its strong lensing properties. Results, comparison to the observational data, and conclusions are presented at the end of the first part.

3.1 The influence of substructure on strong gravitational lensing

It was first noted by Mao & Schneider (1998) that gravitational lensing is a unique tool for detecting substructure at cosmological distances. The underlying theory for this is simple. Consider first a lens system with multiple images, where the lens can be described by a smooth macromodel. With smooth we refer to a model which does not show structure in the potential on scales smaller than typical image separation. An image forming at position $\boldsymbol{\theta}$ thus has magnification $\mu^{(0)}$ given by

$$\mu^{(0)} = \frac{1}{\left(1-\kappa^{(0)}\right)^2 - |\gamma^{(0)}|^2} \ , \tag{3.1}$$

where $\kappa^{(0)}$ and $\gamma^{(0)}$ are calculated at the position of the image. Without loss of generality we chose the coordinate system such, that only the γ_1 component contributes to the total shear. Consider now a perturber, located close to where the image is formed and for simplicity assume it is causing perturbations $\delta\gamma_1 = \delta\kappa$ and $\delta\gamma_2 = 0$ to the unperturbed convergence and shear at the image position. The perturbed magnification is then given by

$$\mu = \frac{1}{\left(1-(\kappa^{(0)}+\delta\kappa)\right)^2 - \left(\gamma^{(0)}+\delta\gamma_1\right)^2} \ . \tag{3.2}$$

The ratio $\mu/\mu^{(0)}$ is, up to the first order in $\delta\kappa$, given by

$$\frac{\mu}{\mu^{(0)}} = 1 - 2\delta\kappa \frac{1}{1-(\kappa^{(0)}+\gamma^{(0)})} \ . \tag{3.3}$$

For a typical lens (such as the first example used to detect substructure B1422+231) we have at the position of the image showing the largest ($\sim 20\%$) deviation in flux from what the smooth model predicts $\gamma^{(0)}+\kappa^{(0)} \sim 0.9$. Thus for a 20% change in magnification, we typically need very small perturbations, $\delta\kappa \sim 0.01$.

The magnification of multiple images is therefore very sensitive to the small-scale perturbation in the lens potential. This makes lensing a golden tool to detect substructure. By measuring flux ratios of images (together with image positions to constrain the macro-model) we can in principle directly constrain the presence of the substructure. Unfortunately, the flux measurements are not as reliable as the position measurements. Therefore one would think it would be much easier to detect substructure by simply measuring image positions. However they are a lot less sensitive to small scale perturbations as compared to the fluxes.

To explain this we look at the gravitational potential in Fourier space. Equation (2.9) can be rewritten in the Fourier space as

$$\tilde{\boldsymbol{\alpha}}(\boldsymbol{k}) = -\mathrm{i}\boldsymbol{k}\tilde{\psi}(\boldsymbol{k}) \ , \tag{3.4}$$

where $\tilde{\psi}(\boldsymbol{k})$ and $\tilde{\boldsymbol{\alpha}}(\boldsymbol{k})$ are the Fourier transforms of the potential and the deflection angle, respectively. Similarly, the Fourier transforms of the convergence $\tilde{\kappa}$ and shear $\tilde{\gamma}$ can be calculated by considering Eqs.(2.11) and (2.13) resulting in

$$\tilde{\kappa}(\boldsymbol{k}) = -\frac{1}{2}\left(k_1^2+k_2^2\right)\tilde{\psi}(\boldsymbol{k}) \ , \tag{3.5}$$

$$\tilde{\gamma_1}(\boldsymbol{k}) = \frac{1}{2}\left(-k_1^2+k_2^2\right)\tilde{\psi}(\boldsymbol{k}) \ , \tag{3.6}$$

$$\tilde{\gamma_1}(\boldsymbol{k}) = -k_1k_2\tilde{\psi}(\boldsymbol{k}) \ , \tag{3.7}$$

where $\boldsymbol{k}=(k_1,k_2)$ is the wave vector. The norm of $\boldsymbol{k}$ is inversely proportional to the length scale; thus smaller scale fluctuations correspond to large $\boldsymbol{k}$.

This excursion to Fourier space now helps us understand why small-scale variations in the potential affect the fluxes of the images more than their positions. Namely, the absolute value of deflection angle $|\tilde{\boldsymbol{\alpha}}(\boldsymbol{k})|$ is proportional to $|\boldsymbol{k}|$, while the inverse of magnification factor contains terms proportional to $|\boldsymbol{k}|^4$. Therefore, although the measurements of the image positions can be done with higher accuracy than the measurements of the fluxes, measuring the flux ratios is up to now a more promising tool to directly detect the substructure.

To express this last statement in numbers, a typical $10^6M_\odot(10^9M_\odot)$ clump causes a shift in source position of typically ~ 1 mas(~ 30 mas), whereas it can in principle change the fluxes by an arbitrary factor. Unfortunately, measuring the fluxes can often be affected by microlensing (especially optical flux ratios) (Woźniak et al., 2000) and galactic scintillation (Koopmans et al., 2003). With high-resolution VLBI observations,however, we will be able to measure the influence of substructure on image position in the near future as well and this will offer a more robust method for direct detection on substructure.

Finally a word on measuring the Hubble constant with gravitational lensing. Strong lensing can be used to measure the Hubble constant H_0 (see Refsdal 1964). From positions and fluxes of individual images, one can constrain the mass model of a galaxy and from the measured time delays between the light curves of images, the Hubble constant can be determined. Since the time delay is proportional to the gravitational potential ψ, the small-scale perturbations due to the substructure in a lens galaxy change it at only a per cent level. The difficulty, however, lies in constraining the smooth mass model of a galaxy. One needs to constrain the average surface density $\langle\kappa\rangle$ near the images for which the time delays are measured. These predictions can vary substantially depending on the assumptions entering the lens model. At present therefore the way forward is to fix the Hubble constant to the value from e.g. HST key project and use the time delays to put tight constraints on e.g. the profile of the galaxy mass distribution.

3.2 Strong lensing by a simulated galaxy

N-body simulations can provide a powerful benchmark for testing the effects of substructure on strong lensing. One can simulate ideal conditions in which propagation effects of the ISM can be ignored and search for the signature of substructure. The drawback of this method at present lies in the resolution available for simulations that include dark matter, gas and star particles. This limits our analysis to mass clumps of masses $\gtrsim 10^7 M_\odot$. However, since the mass resolution is advancing rapidly, this will soon no longer be a problem.

We use the cosmological N-body simulation data for several realisations of galaxies including gas-dynamics and star formation (Steinmetz & Navarro, 2003). We investigate two different halos, each of them in three different projections. The simulations were performed using GRAPESPH, a code that combines the hardware N-body integrator GRAPE with the Smooth Particle Hydrodynamics technique (Steinmetz, 1996).

In Table 3.1 the properties of the halos are listed. In both cases the original simulated field contains approximately 300 000 particles. The simulation is contained within a sphere of diameter 32 Mpc which is split into a high-resolution sphere of diameter 2.5 Mpc centred around the galaxy and an outer low-resolution shell. Gas-dynamics and star formation are restricted to the high-resolution sphere, while the dark matter particles of the low-resolution sphere sample the large scale matter distribution in order to appropriately reproduce the large scale tidal fields (see Navarro & Steinmetz 1997 and Steinmetz & Navarro 2000 for details on this simulation technique). From the original simulated field we take a cut-out cube of size $\sim 200^3$ kpc^3 centred on a single galaxy. This area lies well within the high-resolution sphere and is void of any massive intruder particles from the low-resolution shell.

All simulations were performed in a ΛCDM cosmology ($\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$, $\Omega_\mathrm{b} = 0.019/h^2$, $\sigma_8 = 0.9$). They have a mass resolution of $1.26 \times 10^7 M_\odot$ and a spatial resolution of 0.5 kpc. A realistic resolution scale for an identified substructure is typically assumed to be ~ 40 particles which corresponds to $5 \times 10^8 M_\odot$. The quoted mass resolution holds for gas/stars. The high-resolution dark matter particles are about a factor of 7 ($= \Omega_0/\Omega_\mathrm{b}$) more massive. A detailed analysis of the photometric and dynamical properties of the simulated halos we used here was carried out in Meza et al. (2003) for the elliptical and Abadi et al. (2003a,b) for the disc galaxy.

3.2.1 Delaunay tesselation smoothing technique

From the particle distribution, irregularly sampled in the simulation box, we need to construct the underlying convergence κ field. In general, determining a continuous field from discretely sampled measurement points is a non-trivial process and is often discussed in various astrophysical applications.

In Bradač et al. (2002) we used a very simple approach, namely smoothing the field using the Gaussian kernel of a fixed size. Albeit simple, this method did not perform too well. In the high particle density regions (e.g. close to the central cusp) it smooths out the overall profile making it more core-like and thus changes the overall strong-lensing properties of the galaxy. This can be cured by choosing a smaller smoothing scale; however in this case the shot noise of individual particles is not effectively smoothed out and can influence the magnifications of a lens substantially (as discussed in Sect. 3.1). Third, when dealing with disc galaxies we need to be careful not to smooth out the disc (especially when seen edge on) as the disc contributes to the small-scale structure in our potential and thus again effects the lensing properties.

Therefore a method is needed that adapts the kernel size and shape in order to increase the signal to noise of the reconstructed field, while we still represent the strong lensing properties of the simulation correctly. For this purpose, we make use of the Delaunay tesselation technique from Schaap & van de Weygaert (2000), (see

Table 3.1: Properties of the two simulated halos we used. z_l denotes the redshift of the halo, z_s is the redshift of the source. N_bar, N_DM, and N_str are the numbers of baryonic, dark matter particles and "stars" respectively, present in the cut-out of the simulation we used (note that even within one family particles have different masses). M_tot is the total mass of the particles we used.

Halo	Elliptical	Spiral
z_l	0.81	0.33
z_s	3.0	3.0
N_bar	12 000	20 000
N_DM	17 000	26 000
N_str	70 000	110 000
M_tot	$1.5\ 10^{12} M_\odot$	$0.5\ 10^{12} M_\odot$

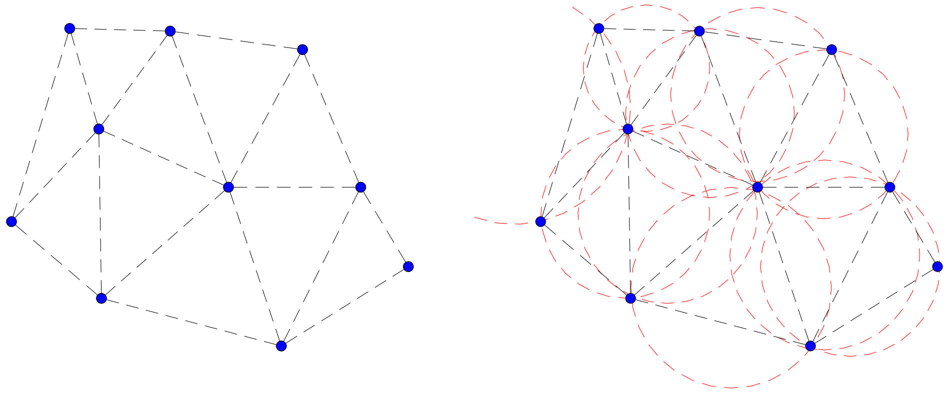


Figure 3.1: Sketch of dividing up space into Delaunay polygons. The space is divided into polygons with $M+1$ vertices, where M is the dimensionality of the problem. In this simple case we have $M=2$, the space is divided up into triangles (left) such, that in each circumcircle (right) there is no extra point.

also Lombardi & Schneider (2002) for more detailed explanations).

We decided to use the Delaunay tesselation technique over other adaptive smoothing techniques for the following reasons:

- Neither the shape nor the size of the smoothing kernel is a free parameter. The method automatically probes the high density regions at the maximum possible resolution; while it effectively smooths the low density regions, thus minimising the effects of the shot noise. It is also very successful in recovering anisotropic features, because the kernel shape is adaptive (in our case this is extremely important for the simulated disc galaxy we use).

- It is superior over the adaptive smoothing techniques where smoothing length depends on the distance to the nearest neighbours. In Delaunay tesselation, each point is connected to its nearest neighbour by a side in the triangulation. Delaunay triangulation is a planar graph with at most $3N-6$ edges and at most $2N-5$ triangles in 2-D. Therefore if you want to find the closest neighbour to each of the vertices, you only have to look at $3N-6$ pairs, instead of all $N(N-1)/2$ possible pairs, where N is the number of all vertices. Since Delaunay Tesselation is itself $\mathcal{O}(N)$ routine (Schaap & van de Weygaert, 2000), the full process performs much faster than conventional adaptive smoothing techniques.

The density estimation then proceeds as follows. We have N simulation particles (vertices) in M-dimensional space ($M=3$). The space is then tiled with polygons consisting of $M+1$ vertices, in our case we are thus dealing with tetrahedra. This is done in such a way that the circumsphere defined by the vertices does not contain any extra vertex (see Fig. 3.1 for explanation). For this process we make use of the QHULL algorithm (Barber et al., 1996).

Having done the tiling, we write down the density estimator $\rho(\boldsymbol{x}_i)$ at each vertex

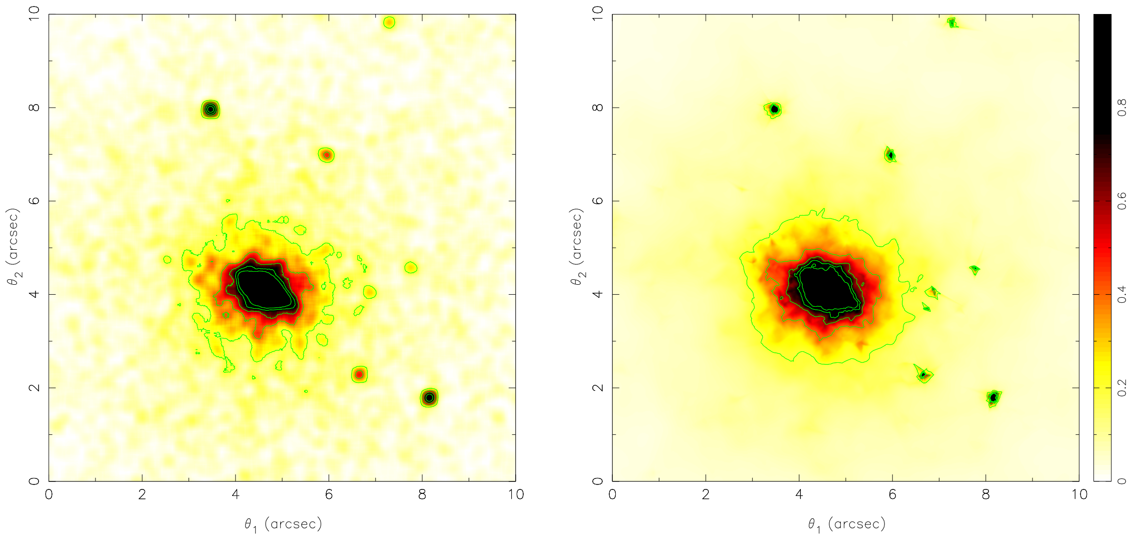


Figure 3.2: Comparison of the κ maps smoothed using a Gaussian smoothing kernel characterised by $\sigma_\mathrm{G} \sim 1$ kpc (left), and smoothed using the Delaunay smoothing technique (right). Colour coding and contour levels are equal for both panels. One arcsecond in the *lens* plane corresponds to approximately 6 kpc.

$\boldsymbol{x}_i$ as

$$\rho(\boldsymbol{x}_i) = \frac{m\,(1+M)}{W_i}\,, \tag{3.8}$$

where m is the mass of the i-th vertex and W_i is the sum of volumes of all K tetrahedra sharing this vertex, i.e. $W_i = \sum_{j=1}^{K} V_{\mathrm{Del,j}}$. Using the volume W_i in the density estimator guarantees the mass conservation. We need to properly normalise the estimate by recalling that each Delaunay tetrahedron is used for the density estimates at $1+M$ points, we get the pre-factor in the above expression.

In order to obtain the κ-map on a grid we first interpolate values of the density at each three-dimensional grid point. Finally, we project the resulting values onto a two-dimensional grid and scale Σ_{ij} with Σ_{crit} for a source at a redshift $z_\mathrm{s} = 3$ to obtain the convergence κ_{ij}.[1] Since the N-body simulations contain three independent classes of particles (gas, stars, and dark matter particles, each having different masses), we perform the procedure described above for each group separately and obtain the final κ map by adding up the contributions from all three classes. The Delaunay tesselation method performs very well in comparison to the standard Gaussian smoothing technique used in Bradač et al. (2002), or an adaptive Gaussian smoothing technique (see also Schaap & van de Weygaert 2000). In Fig. 3.2 we show a qualitative result of both methods. While in the case of Gaussian smoothing we still see the effects of the shot noise in the outer parts, the parts with high particle density are over-smoothed, the steep profile became shallower (compare the contour levels between left and right panel). This is not the case for the smoothing using Delaunay Tesselation.

---

[1] For distance calculations throughout this chapter we assume for simplicity an Einstein-de-Sitter Universe and the Hubble constant $H_0 = 65\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

3.2.2 Estimating the noise properties

A drawback of using the Delaunay tesselation method is that the noise evaluation for the final surface mass density map is non-trivial. For example, with a Gaussian kernel one can get an estimate for the noise by simply looking at the number of particles in a smoothing area (for a more detailed estimate, see Lombardi & Schneider 2002). When using the tesselation technique, such an approach is not viable.

One possible approach for estimating the error is to use the bootstrapping technique (see e.g. Heyl et al. 1994). Ordinarily we calculate physical properties (e.g. κ-map) by using all N particles from the simulation. In order to create a bootstrap image one has to randomly select N particles out of this simulation with replacement; i.e. some of the particles from the original simulation would be included more than once, some not at all. In other words, we randomly generate N numbers from 1 to N and they represent the bootstrapped set of particles. E.g. if a particle is included x times in a bootstrapped map, we put a particle at the same position with x times its original mass. After creating an ensemble of such images and calculating the desired physical quantity for each of them, one can then make an error estimation.

A problematic issue in our case is the creation of an ensemble of κ-maps from the bootstrapped images. Whereas the tesselation itself is done very quickly, the interpolation of density on a grid is a slow process. For each tetrahedron we need to find all grid points (in 3-D) that belong to the tetrahedron and estimate the contribution to its density by linear interpolation. Since the number of grid points needs to be large to retain the resolution, the process is computationally expensive. We therefore limit ourselves to 10 bootstrapped maps, as described above, and compare them with the original maps. We perform this analysis for the elliptical halo only, because it is computational expensive and we do not expect systematic differences.

For each pixel i we calculate the associated error using

$$\sigma^2\left(\kappa_i\right) = \frac{1}{Q-1}\sum_{q=1}^{Q}\left(\kappa_i^q - \langle\kappa_i\rangle\right)^2 \tag{3.9}$$

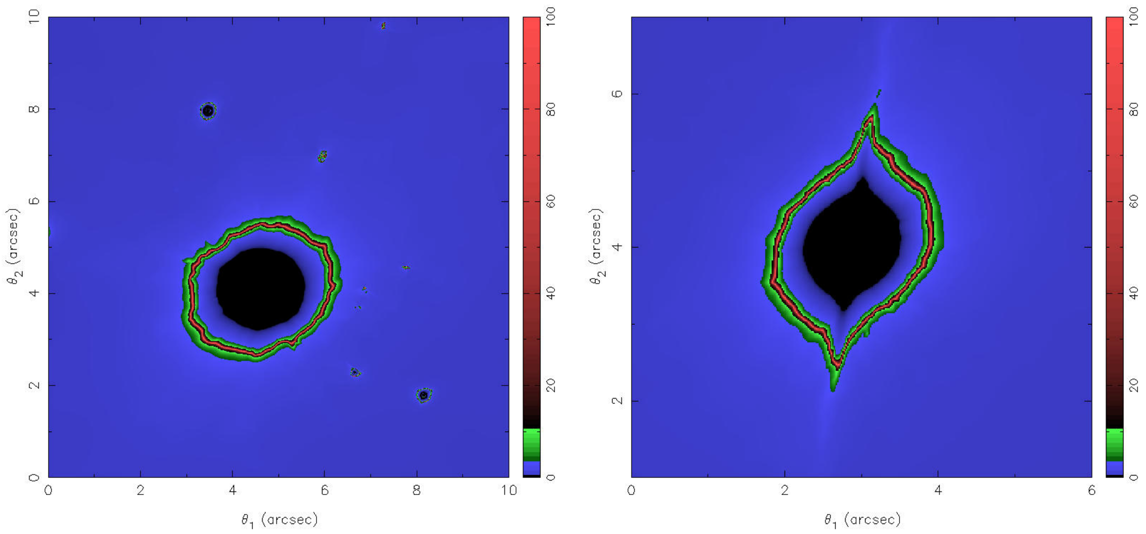
where $\langle\kappa_i\rangle$ is the average value of κ at pixel i averaged over all $q = Q$ bootstrapped maps; in our case $Q = 10$.

This procedure gives us an estimate for the error on κ. Within the critical curves, i.e. $\kappa \gtrsim 0.5$ the average noise $\sigma(\kappa_i)/\langle\kappa\rangle$ is of the order of $\lesssim 5\%$. We will return to this bootstrapping technique later on in this chapter.

3.2.3 Strong lensing properties

Having obtained the κ-map, we then calculate the lens properties on the grid (2048 × 2048 pixels). The lensing potential $\boldsymbol{\psi}$ (and therefrom all the lensing properties) can be calculated directly by performing the integration in (2.10). This is however time consuming.

A far better approach is suggested by the fact that $\boldsymbol{\alpha}$ and ψ are convolution-type integrals over κ. In Fourier space, the differentiation by x_i is replaced by multiplication by $-\mathrm{i}k_i$ and therefore Eqs.(3.5),(3.6), (3.7) are easy to deal with. In practice



(a) **(b)**

Figure 3.3: The magnification map of the simulated elliptical **(a)** and edge-on spiral **(b)** galaxy. External shear is added in the evaluation of the magnification map to account for neighbouring galaxies (see text). Lighter (red) regions represent high magnifications. The units on the axes are arcseconds, one arcsecond in the *lens* plane corresponds to approximately 7 kpc in **(a)** and 5 kpc in **(b)**.

one thus performs the Fourier transformation on κ and therefrom obtain the Fourier transforms of the potential $\tilde{\psi}$ by simple multiplication for $\boldsymbol{k} \neq \boldsymbol{0}$. For $\boldsymbol{k} = \boldsymbol{0}$ the $\tilde{\psi}$ is undetermined, however since this corresponds to a constant in ψ it is not important (a constant in ψ does not change γ, κ or $\boldsymbol{\alpha}$) and can be set to zero.

From this point one can proceed in two ways. One can calculate the two components of the shear $\gamma_{1,2}$ and the deflection angle $\boldsymbol{\alpha}$ by multiplying the potential in Fourier space by an appropriate kernel. However, since we calculate the Fourier transform of the potential on a grid with cells of finite size, we effectively filter out high spatial frequency modes. By multiplying the transform with different kernels for $\gamma_{1,2}$ and $\boldsymbol{\alpha}$, these final maps do not necessarily correspond to the same κ map. The effect is in most cases very small, but as we approach the critical curves it influences the calculations of image positions and fluxes substantially. Therefore, it is better to only calculate the lens potential ψ in the Fourier space and obtain the lens properties by finite differencing of ψ in the real space. The latter method is also less CPU-time and memory consuming.

We perform the transformation by the DFT (Discrete Fourier Transformation) method. In the definition of Fourier transform the integrals extend over the whole $\mathbb{R}^2$ space; therefore we can use the DFT provided κ decreases "fast enough" toward the boundaries of our field. Further, the discrete Fourier transform $\tilde{\kappa}$ is in fact the periodic continuation of the mass distribution in the real space. As a consequence the calculated lensing properties are the properties of the considered mass distribution plus all its periodic continuations. In order to reduce these periodic effect, padding was introduced. More precisely, the DFT was performed on a 4096 × 4096 grid, where one fourth of the grid contains the original κ map and the rest is set to zero.

Finally, if we calculate $\psi(\boldsymbol{x})$ from $\tilde{\psi}(\boldsymbol{k})$ via the inverse Fourier transform, we get a potential that does not correspond to κ but rather to $\kappa - \bar{\kappa}$, where $\bar{\kappa}$ is the average value of κ on the grid. This is due to the fact that when we calculate $\tilde{\kappa}$ from κ on a grid using the discrete Fourier transform we in fact calculate the transform of κ plus all its periodic continuations in $\boldsymbol{x}$-space. A periodic mass distribution with all elements having a positive total mass will have an infinite mass and as a consequence a diverging deflection angle. Therefore at the end, one has to add a term $\bar{\kappa}\,|\boldsymbol{x}|^2/2$ to the potential $\psi(\boldsymbol{x})$.

The biggest advantage of performing the calculation of ψ in Fourier space is that the calculations can be performed using Fast Fourier Transform (FFT) routines. This also allows us to choose large field sizes and large number of grid points. For the practical calculations, we use the publicly available C library FFTW ("Fastest Fourier Transform in the West") written by Frigo & Johnson (1998).

The simulated galaxy is a field galaxy. However, most of the lenses in quadruple image systems are members of groups. To make our simulated galaxies as closely related to realistic systems as possible, we add external shear to the Jacobi matrix (evaluated at each grid point). The shear components were the same for each projection and all halos, and chosen to be

$$\gamma_1^{\rm ext} = -0.04\ , \quad \gamma_2^{\rm ext} = -0.16\ ;$$

they correspond to the shear of the best-fit singular-isothermal ellipse model with external shear for the lens B1422+231 in Bradač et al. (2002). Figure 3.3 show the magnification maps of the elliptical and edge-on spiral galaxy. The corresponding caustic curves (for a source at $z_{\rm s} = 3$) were obtained by projecting the points of high magnification from the image plane to the source plane using deflection angle information. The critical curves are plotted in red in Fig. 3.3a for the elliptical and Fig. 3.3b for the edge-on spiral galaxy. The corresponding caustic curves are plotted in Fig. 3.6a and 3.9c respectively.

Four-image systems of N-body simulated galaxies

Finally we generate different four-image systems using individual simulated galaxies. For each halo and projection direction, regions in the source plane where five images form were determined. At first glance this seems to be straightforward, all you need to do is look for *all* solutions $\boldsymbol{\theta}_i$ of the lens equation (2.2) belonging to the source position $\boldsymbol{\beta}_{\rm s}$, thus in other words simply solve the equation

$$|\boldsymbol{\beta}(\boldsymbol{\theta}_i) - \boldsymbol{\beta}_{\rm s}|^2 = |\boldsymbol{\theta}_i - \boldsymbol{\alpha}(\boldsymbol{\theta}_i) - \boldsymbol{\beta}_{\rm s}|^2 = 0\ , \tag{3.10}$$

Since $\boldsymbol{\theta}_i$ enters the equation in a non-linear manner, more than one solution is possible, giving rise to the multiple images as we mentioned before. No robust algorithms are available to deal with problems of this sort in general, and therefore there is no guarantee of finding *all* solutions (and the number of them).

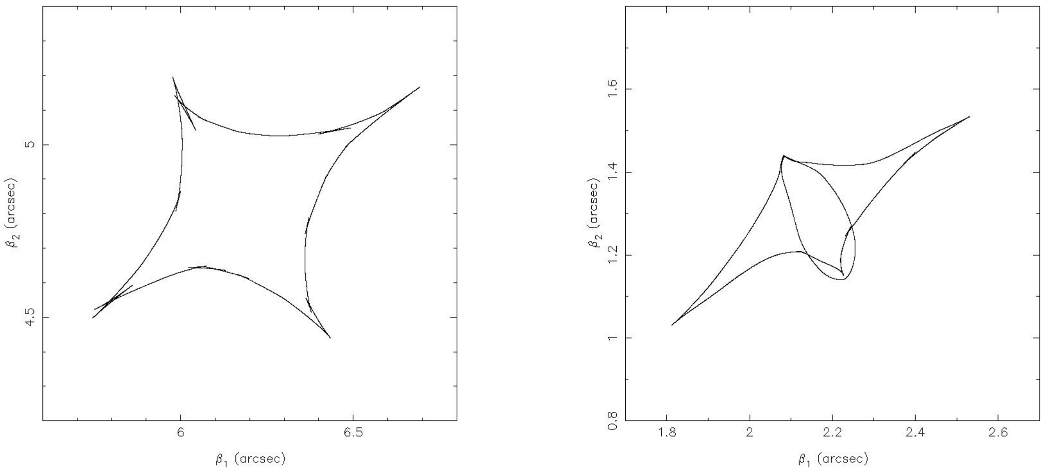
The best guess one can make is to scan the grid of solutions. In the context of lensing the grid search method has been described by Blandford & Kochanek (1987). For this method one defines a grid in the image plane (m, n) with positions $\boldsymbol{\theta}_{m,n}$, and each grid point is projected back onto the source plane using the lens equation (2.2), yielding $\boldsymbol{\beta}(\boldsymbol{\theta}_{m,n})$. For each grid point we now define two triangles with vertices (m, n) $(m+1, n)$ $(m+1, n+1)$ and $(m+1, n+1)$ $(m, n+1)$ (m, n) respectively. These two triangles, when projected back, correspond to two triangles in the source plane. Each grid point in the source plane that lies within the projected triangle gets assigned the label (m, n). After the whole lens-plane grid is transformed, points in the source plane get assigned one or more labels (provided the physical size of the lens plane grid has been chosen large enough), their number corresponding to the image multiplicity for this source position. Labels (m, n) also correspond to the approximate image positions $\boldsymbol{\theta}_{m,n}$.

Since we are only interested in the four image systems (i.e. those that have fold and cusp configurations), we choose all the source positions with image multiplicity of five. Namely, what are called quadruply imaged systems are actually five imaged systems. However the fifth image is in most cases highly demagnified and escapes the observations. The approximate image positions provide the initial conditions for the MNEWT routine from Press et al. (1992), which is used to search for the roots of equation (3.10). Because the deflection angle needs to be defined for arbitrary $\boldsymbol{\theta}$, we need to make use of interpolation. We use bilinear interpolation and bicubic spline, and both methods give comparable results. Once we have the image positions, the magnification factors are calculated and the four brightest images are selected; these then represent the "observable" images.

3.2.4 The importance of baryons

As mentioned above, the effect of baryons is very important in lens galaxies. Since we are interested in the effects of substructure, it is desirable to use N-body simulations that have the highest possible dynamical range in density. At present this is achieved in high-resolution N-body simulations that only include dark matter. However, if baryons are not included, the central regime of the lens potential is more shallow than what is typically observed in lens galaxies. All quadruply-imaged systems for which the inner slope of the mass distribution has been measured, are well described by a total mass density profile $\sim r^{-\gamma}$ with $\gamma \simeq 2$ (Kochanek, 1995; Cohn et al., 2001; Treu & Koopmans, 2002a,b; Koopmans & Treu, 2003; Treu & Koopmans, 2004), consistent with the combined mass distribution of dark matter and baryons seen in the simulations. Hence, dark matter-only simulations do not accurately represent the overall properties of lens galaxies, and instead we need to use hydro-dynamical simulations.

To investigate the importance of baryons we simulated an elliptical halo using only dark matter particles and performed the same lensing analysis as described above. We project the density distribution approximately along the long axis of the halo, thus maximising the central density. In Fig. 3.4 we show the corresponding caustic curves for the halo simulations; in (a) the simulation includes baryons (DM+B), in panel (b)



(a) **(b)**

Figure 3.4: The caustic curves of two simulated galaxies. For panel **(a)** we use the simulation which includes baryonic and dark-matter particles, whereas for panel **(b)** the simulation includes dark-matter particles only. The radial caustic for the dark-matter only simulation is almost entirely enclosed within the asteroid caustic, prohibiting formation of quadruply-imaged systems with cusp images. For simulation which include baryons the radial critical curve is located close to the galaxy centre, thus the corresponding caustic would lie well outside the asteroid caustic.

we use the DM-only simulation (DM). The radial critical curve in the DM+B halo is located close to the galaxy centre and is not well resolved (but also not important for our purposes) - the corresponding caustic is therefore not plotted.

The two caustic curves are very different, indicating very different overall strong lensing properties of these two simulations. Whereas the DM+B simulation has a steep inner mass profile (very close to singular isothermal), the DM simulation has a caustic configuration typical for a lens with a shallow density profile (see e.g. Wallington & Narayan, 1993). The radial critical curve is smaller compared to the DM+B halo and the corresponding caustic curve is almost entirely enclosed by the asteroid caustic. The prominent naked cusp region is a three-image region. This configuration is extremely rare, if not absent among the observed lensed systems. For a source located in such a region one observes three highly magnified images. There is only one possible example of a triply imaged quasar out of ~ 50 doublets and triplets (Evans & Hunter, 2002), APM 08275+5255 (Ibata et al., 1999). The vast majority of systems similar to B1422+231 can not be generated by such a potential. We conclude that baryons have to be included in the simulations, and we use in the rest of this chapter DM+B simulations only, discussing their limits where necessary.

3.3 Cusp relation in N-body simulated galaxy

In the previous section we have discussed a method to generate synthetic four-image systems from N-body simulations. There are three basic configurations of four image systems as presented in Fig. 2.2: the fold, cusp, and cross configurations. They correspond to a source located inside the asteroid caustic, close to a fold (dark blue source in the lower panel of Fig. 2.2), a cusp (dark blue source in the upper panel) or near the centre (green source) respectively. Analogous to classical geometric optics, images form at the local extrema of the Fermat potential ϕ (see Eq. 2.17 and Fig. 2.4). In a four-image system (with a generic elliptical lens) two images form at the minima and two at the saddle points of ϕ. A minimum thus has positive parity, i.e. $\mu > 0$ and a saddle point has negative parity $\mu < 0$. If the lens has a non-singular mass distribution or mass-distribution more shallow than isothermal at the centre, the fifth, highly demagnified image would correspond to a maximum. All configurations have been observed, and even though one would naively think that fold and cusp images are rare among observed lenses, they are in fact frequently observed due to the large magnification bias in flux-limited samples. In this section we will mainly concentrate on cusp image configurations.

The behaviour of gravitational lens mapping near a cusp was first studied by Blandford & Narayan (1986), Blandford (1990), Schneider & Weiss (1992), and Mao (1992), who investigated the magnification properties of the cusp images and concluded that the sum of their signed magnification factors approaches zero as the source approaches the cusp. For completeness we repeat here the equation (2.19), namely

$$R_{\rm cusp} = \frac{|\mu_{\rm A} + \mu_{\rm B} + \mu_{\rm C}|}{|\mu_{\rm A}| + |\mu_{\rm B}| + |\mu_{\rm C}|} \rightarrow 0, \ \text{for } \mu_{\rm tot} \rightarrow \infty \ ,$$

in which $\mu_{\rm tot}$ is the unsigned sum of magnifications of all four images, and A, B and C is the triplet of images forming the smallest opening angle $\theta_{\rm open}$ - see Fig. 3.5. The opening angle is measured from the galaxy centre and is the smallest angle that is spanned by the two images of equal parity. The third image lies inside this opening angle.

3.3.1 The cusp relation of an N-body simulated elliptical galaxy

The cusp relation (3.3) is an asymptotic relation and holds when the source approaches the cusp from inside the asteroid. One can derive the properties of lens mapping close to critical curves using a Taylor expansion of the Fermat potential around a critical point (see e.g. Schneider & Weiss 1992). Such calculations are very cumbersome and therefore it is difficult (if not impossible) to analytically explore the influence of arbitrary substructure. In practice, we can numerically calculate $R_{\rm cusp}$ for the N-body simulated systems. Smoothing the original κ-map on different scales then gives an indication of the influence of substructure of different length or mass scales on $R_{\rm cusp}$.

The three cusp images [designated as A, B and C in (2.19)] are chosen according to the image geometry. Since we know the lens position, this procedure is straight-

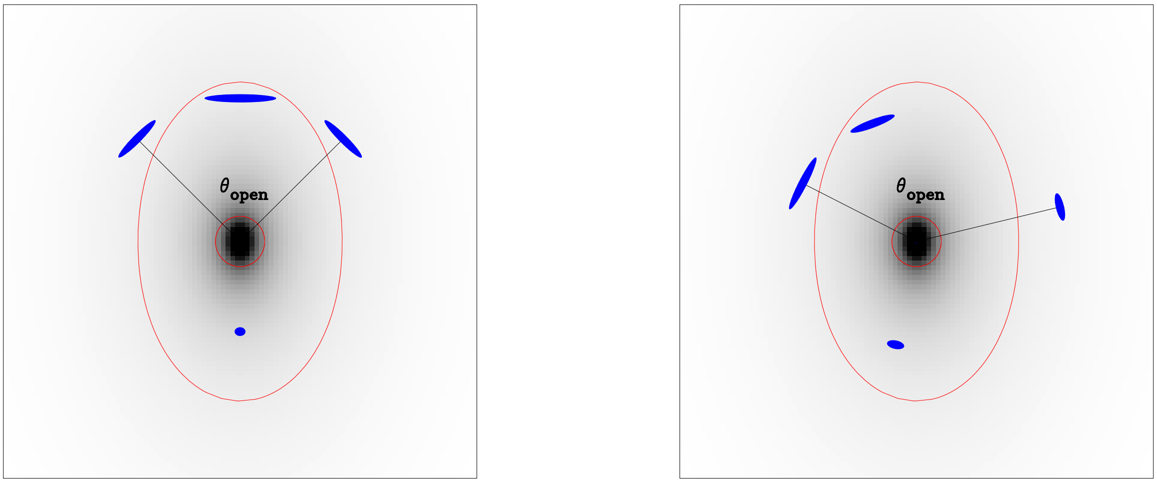


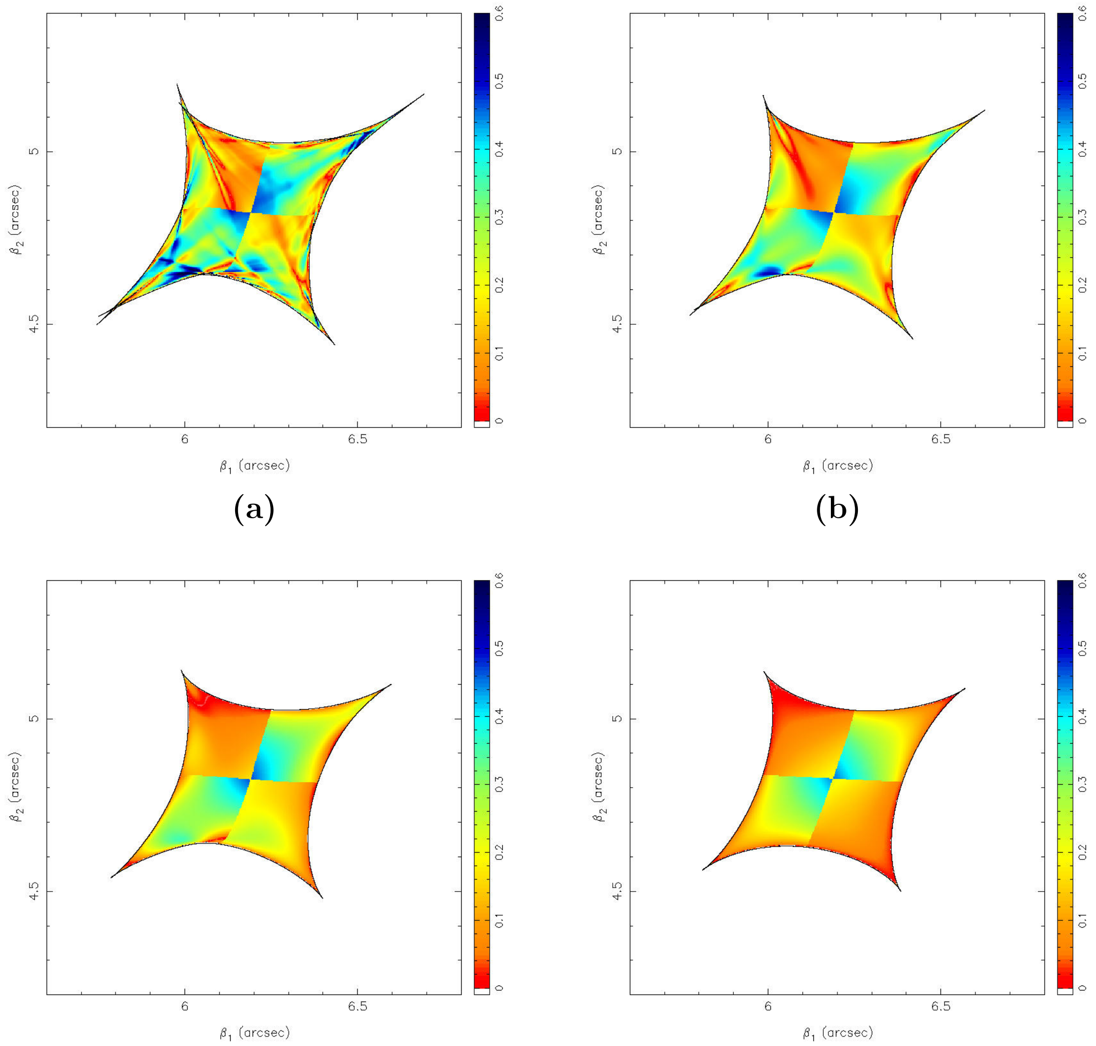
Figure 3.5: A lensed image configuration of a compact source close to a cusp caustic (left) and a fold caustic (right) of a typical non-singular elliptical lens. The panels also show critical curves (red) and the opening angle θ_{open}. In gray scale the convergence κ is plotted.

forward and foolproof. We identify the triplet of images belonging to the smallest opening angle (described above). Since we know the image parities and magnifications, one is tempted to identify the three brightest images as the cusp images and assign different parity to the brightest one than to the other two (e.g. as in Möller et al. 2003). However, due to the presence of shear and substructure this could lead to misidentifications.

Figure 3.6a shows the caustic curve in the source plane for the simulated elliptical at a redshift of $z_{\text{l}} = 0.81$. The source is at a redshift of $z_{\text{s}} = 3$. Approximately 30 000 independent lens systems are generated with source positions inside the asteroid caustic. R_{cusp} is plotted in colour. The apparent discontinuities originate from different image identification. In the very centre of the caustic the meaning of “cusp image” is ill defined. As the source moves in the direction of the minor or major axes we chose different triplet of images to calculate R_{cusp} and therefore the discontinuity arises.

The remaining panels of Fig. 3.6 show the effect of smoothing the small-scale structure in the surface mass density κ map with Gaussian kernels characterised by standard deviation σ_{G}. The values for σ_{G} were chosen to be $\sigma_{\text{G}} \sim 1, 2, 5$ kpc for panels (b), (c) and (d) respectively. Note that we do not smooth the κ-map directly. First we obtain the smooth model κ_{smooth} for the κ-map by fitting elliptical contours to the original map using the `IRAF.STSDAS` package `ellipse`. We subtract κ_{smooth} from κ and smooth the difference using different Gaussian kernels. We then add the resulting map back to the κ_{smooth}. In this way the overall radial profile of the mass distribution is not affected.

The effect of smoothing on the cusp relation is clearly visible. In Fig. 3.6d one sees that the substructure is completely washed out when smoothing on scales of



(a) **(b)**

(c) **(d)**

Figure 3.6: The cusp relation $R_{\rm cusp}$ for the N-body simulated elliptical galaxy at a redshift of $z_{\rm l} = 0.81$. The sources were put at a redshift of $z_{\rm s} = 3$ and approx. 30 000 source positions that lie inside the asteroid caustic were used. $R_{\rm cusp}$ is plotted in colour, for sources close to the cusp the smooth models would predict $R_{\rm cusp} \sim 0$ (i.e. red colour). The deviations are due to the substructure. Due to magnification bias most of the observed lenses correspond to the fold and cusp configurations. Discontinuities in the maps arise when the source moves in the direction of the minor or major axes, since a different triplet of images needs to be considered to calculate $R_{\rm cusp}$. On top we plot the caustic curve. Panel **(a)** shows $R_{\rm cusp}$ for the original image, whereas panels **(b)**–**(d)** show the cusp relation for the models where we additionally smoothed the substructure (see text) with a Gaussian kernel characterised by standard deviation $\sigma_{\rm G} \sim 1\,{\rm kpc}$ **(b)**, $\sigma_{\rm G} \sim 2\,{\rm kpc}$ **(c)** and $\sigma_{\rm G} \sim 5\,{\rm kpc}$ **(d)**.

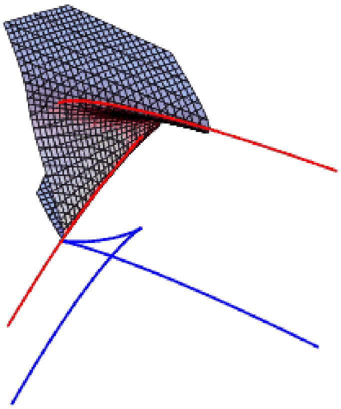


Figure 3.7: A sketch of a formation of a swallowtail. A local region of the lens plane gets pleated and projected onto the source plane forming a swallowtail (blue) - from `http://mathworld.wolfram.com`.

$\sigma_{\rm G} \sim 5\,{\rm kpc}$ is applied. As we go to smaller smoothing scales, the effects of substructure become clearly visible. The caustic structure in Fig. 3.6a shows swallowtails, a singularity of the lens mapping, presented in Fig. 3.7 (note that they are "thinner" in Fig. 3.6a and thus seen mostly as lines close to a fold). In their extension there is a region where the cusp relation is strongly violated (with $R_{\rm cusp} \sim 0.6$, where the smooth model predicts $R_{\rm cusp} \lesssim 0.1$). However, further out, a swallowtail can cause the cusp relation to change the trend and go to zero (due to high-magnification systems being formed in such region).

Finally, the cusp relation behaves differently for the source on the major or the minor axis (see especially Fig. 3.6d). This is a generic feature for smooth elliptical models and can easily be calculated for e.g. an elliptical isopotential model (see e.g. Schneider et al. 1992). We use this model since it is analytically tractable for source positions along the major and minor axis. In Fig. 3.8 we plot the cusp relation for the source moving along the major (minor) axis as a thick (thin) solid line for the elliptical isopotential model with $\epsilon = 0.15$. As the source approaches the cusp, $R_{\rm cusp} \to 0$ for both source positions, however the slope is different. We also plot the total magnification factor of the three cusp images, i.e. $\mu_{\rm A+B+C} = |\mu_{\rm A}| + |\mu_{\rm B}| + |\mu_{\rm C}|$ as a thick(thin) dashed line.

3.3.2 The cusp relation in an N-body simulated disk galaxy

The procedure described above was applied to three different projections of the elliptical halo and very similar conclusions can be drawn. Another question is how much these conclusions depend on the specific morphological type of the galaxy. To investigate this, we follow a similar procedure for a simulated disk galaxy.

In this case, however, we do not look at the effects of additional smoothing. It is difficult to subtract a smooth mass model, since the galaxy consists of a bulge, warped disk and extended halo, which can not simply be fitted by ellipses. If we were to smooth the edge-on projection with a Gaussian kernel, we would also wash out the

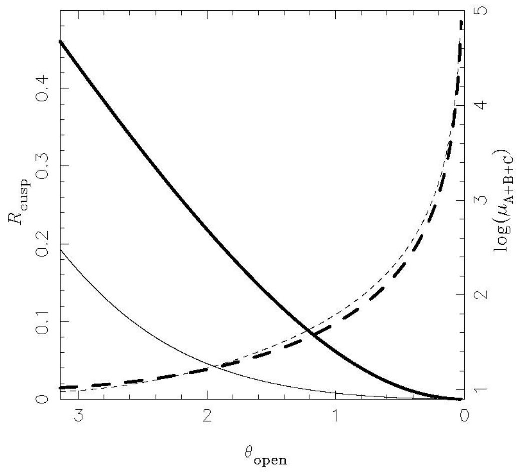


Figure 3.8: $R_{\rm cusp}$ for a simple elliptical isopotential lens model with $\epsilon = 0.15$. $R_{\rm cusp}$ is plotted as a thick (thin) solid line for sources along the major (minor) axis. $\theta_{\rm open}$ represents the angle measured from the position of the galaxy, spanned by two "outer" cusp images (A and C). The opening angle π means that the source is located at the centre (images A, C and the galaxy lie on the same line). When the source approaches the cusp, $\theta_{\rm open} \to 0$. The total magnification for the three cusp images $\mu_{\rm A+B+C}$ is plotted as a thick (thin) dashed line for sources along the major (minor) axis.

(in our case warped) disc. We only include the analysis of the cusp relation for this halo only to show the effect of the disc on $R_{\rm cusp}$.

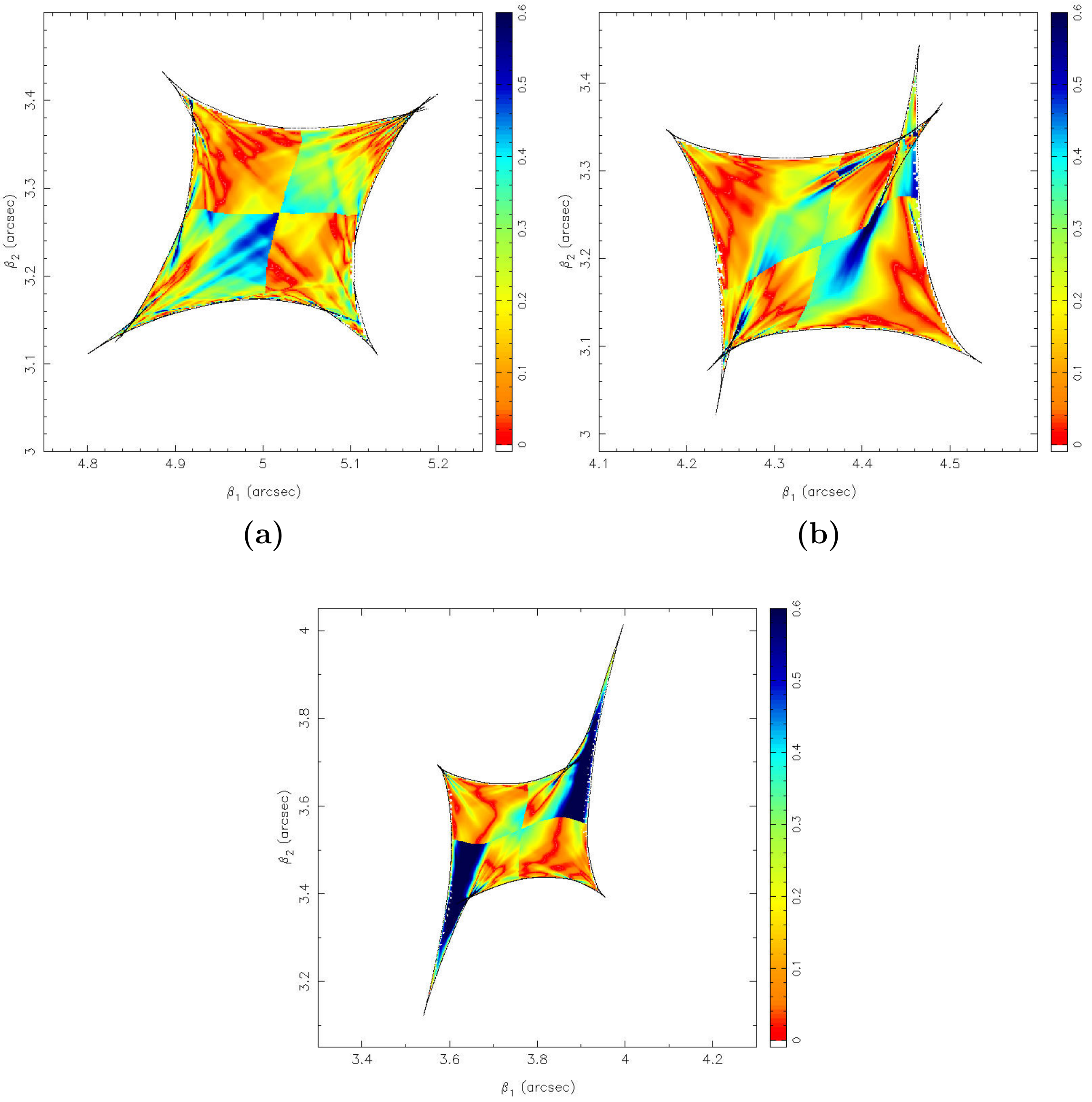
Figure 3.9 shows the cusp relation of an N-body simulated disk galaxy in three projections. The disc clearly plays a role for the edge-on projection (see Fig. 3.9b,c), whereas the face-on projection is similar to the elliptical galaxy. Especially in Fig. 3.9c, where the disc extent in projection is smaller than the typical image separation, the cusp relation in the upper-right and lower-left cusp is strongly violated. This direction corresponds to the orientation of the disc.

3.3.3 Observed cusp relation

Unfortunately the number of observed systems is not large; seventeen four-image systems have been published (see Keeton et al. 2003 for a summary), and four of them are believed to show a typical cusp geometry (see Table 3.2). Even so, a comparison between these observations and our simulations can be made.

Note that the relation (3.3) is model independent and can only be broken in the presence of substructure on scales smaller than the image separation. Hence, if a smooth model provides an adequate description of the lens galaxy, one expects $R_{\rm cusp} \sim 0$ for these lenses. This is clearly not the case and it is therefore difficult to explain their flux ratios with simple, smooth mass models.

If we make a comparison with simulations, one can see that the pattern in Fig. 3.6d clearly does not explain the observed $R_{\rm cusp}$ of these four lenses. On the other hand,



(a) **(b)**

(c)

Figure 3.9: The cusp relation for the N-body simulated spiral galaxy at a redshift of $z_l = 0.33$. The source was put at a redshift of $z_s = 3$ and approx. 10 000 0 source positions that lie inside the asteroid caustic were used. $R_{\rm cusp}$ is plotted in colour. On top we plot the caustic curve. Panels **(a)**, **(b)**, and **(c)** show three different orthogonal projections of the halo, **(a)** corresponds to the face-on, **(b)** and **(c)** to the edge-on projection. Panel**(c)** corresponds to the magnification map in Fig. 3.3b.

substructure on few kpc scales and below provides enough perturbations to explain the observed values of R_{cusp}.

The question arises, however, whether we can conclude from the value of R_{cusp} that we indeed see the effects of CDM substructure. Keeton et al. (2003) argue that the cusp relation alone does not reveal anomalous flux ratios in B1422+231. Still, detailed modelling for this system shows that the flux ratios are anomalous. The difficulties in modelling B1422+231 are not only a consequence of a violation in cusp relation, but also that image D is a fainter than predicted from smooth models. On the other hand, the simulated disk galaxy shows violation of R_{cusp} even though there are no clear mass clumps present in the region where images are formed. Hence one has to be cautious when making conclusions about the presence of CDM substructure based on the value of R_{cusp} alone.

3.3.4 The influence of noise in simulations on the cusp relation

The simulated cusp-relation can be reliably compared with observations only if we know the noise properties (i.e. fluctuations in the potential due to the finite number of particles) of our simulations. The effects of noise and physical substructures need to be disentangled.

From the bootstrapping procedure (Sect. 3.2.2) we also get an estimate of the rms fluctuations in R_{cusp}. We estimated the rms on R_{cusp} using a similar technique as for κ in (3.9). We do not perform the analysis directly in the source plane by subtracting the maps pixel-by-pixel. The problem is that bootstrapping somewhat changes the shape of the caustic curve (see also Fig. 3.11). Since we never observe the source plane directly, we can not distinguish between two shifted, but identical caustics. We therefore have to match different bootstrap maps in the image plane. We compare the image positions generated by each source in the original frame with those generated by bootstrapped lenses. Thus for each source position in the original map (see Fig. 3.6a), we search for the source position in the bootstrapped map such that the four image positions from both maps differ as little as possible.

In principle, one can redo the ray-tracing and calculate image fluxes for the position in the original map using the bootstrapped-lens properties. However this is not

Table 3.2: The values for R_{cusp} taken from Keeton et al. (2003) for four lens systems showing a typical cusp geometry.

Lens	R_{cusp}
B2045+265	0.52 ± 0.04
B0712+472	0.26 ± 0.06
RX J0911+0551	0.22 ± 0.06
B1422+231	0.18 ± 0.06

necessary, since the pixel size in the source plane is small enough and we only need an approximate error estimate. Figure 3.10 shows the estimated absolute error $\sigma(R_{\mathrm{cusp},i})$ for the elliptical halo from Fig. 3.6a. As described above, each value of $\sigma(R_{\mathrm{cusp},i})$, plotted in colour, does not correspond to the error for a source position, but refers to the error of the system with image positions similar to the original ones [2].

The absolute error becomes slightly larger in the regions of the swallowtails. This is, however, not the effect of substructure vanishing in individual bootstrapped realisations. It is rather the effect of slight position changes of individual clumps. If one looks at the individual caustic curves of bootstrapped halos (see Fig. 3.11) the swallow-tails are present in all realisations, although they change their positions slightly. Since this hardly affects the image positions we cannot perfectly match the source positions i' with the original source position i in these regions; thus we are overestimating the true error.

We conclude that the values of R_{cusp} in the close proximity of the cusp can be as high as $R_{\mathrm{cusp}} = 0.6$, with the error of ± 0.1 as determined from the bootstrapping. Smoothing the substructure on scales as large as $\sim 1\,\mathrm{kpc}$ does not remove this effect, but reduces it. This is expected, since smoothing changes the profile of substructures.

Finally, we investigate how well we can sample the smooth mass distribution given the number of particles in the original simulation. We take an ellipsoidal power-law density profile following $\rho(r) \propto r^{-2.9}$. The power-law index was chosen to closely reproduce the surface mass density κ of the simulation which follows $\kappa(\theta) \propto \theta^{-1.9}$ in the vicinity of the critical curve (see Sect. 3.4.1 for more details). The number of particles we use is the same as in the original N-body simulation. To each particle we assign the average mass of the original sample, leaving the total mass of the lens galaxy unchanged. We sample the density profile using a rejection method (e.g. Press et al., 1992). The resulting particle distribution was again adaptively smoothed using Delaunay tesselation, and we perform the lensing analysis as in all previous cases.

The resulting cusp relation is given in Fig. 3.12a. We have chosen the profile such that the particle densities around the outer critical curves are similar in both cases. Only in that case can we reliably compare the noise properties of R_{cusp}. The fact that the caustic is larger than in the case of the N-body simulated elliptical halo (compare to Fig. 3.6a) is here of lesser importance; it arises due the difference in the central profile, far from the critical curve.

The absence of strong violations of R_{cusp} close to the caustic in Fig. 3.12 as compared to Fig. 3.6a also confirms that deviation of R_{cusp} from zero is not dominated by the shot noise of the particles, but is due to genuine substructure in the N-body simulation. For a more quantitative comparison we show the probability distribution of R_{cusp} for systems with $\mu_{\mathrm{tot}} > 20$ in Fig. 3.12b as a solid line for the sampled smooth halo and as a dashed line for the original. The much tighter distribution for the sampled halo confirms the absence of strong violations of R_{cusp} in the sampled smooth halo compared to the original simulation.

Finally we note that by randomly sampling the density profile using rejection

---

[2] Note that this is precisely the procedure that is followed in lens modelling when fitting smooth models to lens galaxies with non-smooth potentials.

method the noise is Poissonian, i.e. behaving as $\propto 1/\sqrt{N}$. In reality, however, studies of errors of N-body simulations using smoothed particle hydrodynamics (Monaghan, 1992) show that the errors are significantly smaller, and behave as $\propto \log N/N$. Therefore our calculations give an upper limit on the expected noise.

This analysis also shows the advantage in smoothing with Delaunay tesselation. If we only use a Gaussian kernel (of the same size as in Fig. 3.6b), the deviations in the cusp-relation for the sampled particle distribution are much larger.

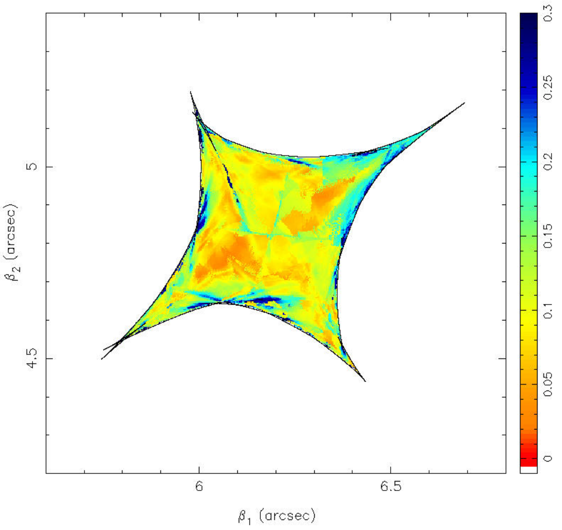


Figure 3.10: The estimated absolute uncertainty $\sigma(R_{\mathrm{cusp},i})$ of the cusp relation, calculated using the bootstrap analysis described in Sect. 3.2.2. Note that the colour coding has changed as compared with the other figures in this chapter. The errors plotted for each source position were calculated using the systems from bootstrapped maps having similar image positions as the system generated by the original lens; i.e. from Fig.3.6a (see text).

3.4 Saddle point demagnification in N-body simulated galaxy

As we have seen in the previous sections the presence of substructure in the smooth macro-potential can cause additional magnification or demagnification of images. In addition, however, these perturbations depend upon the image parity.

It was first noted by Witt et al. (1995) that the expected flux changes due to stellar-mass perturbers differ between saddle points and minima images. This was further investigated by Schechter & Wambsganss (2002) who conclude that for fold images a stellar-mass component added to a smooth mass distribution can cause a substantial probability for the saddle point image to be demagnified w.r.t. what smooth model would predict. Whereas the saddle point images (negative parity) can

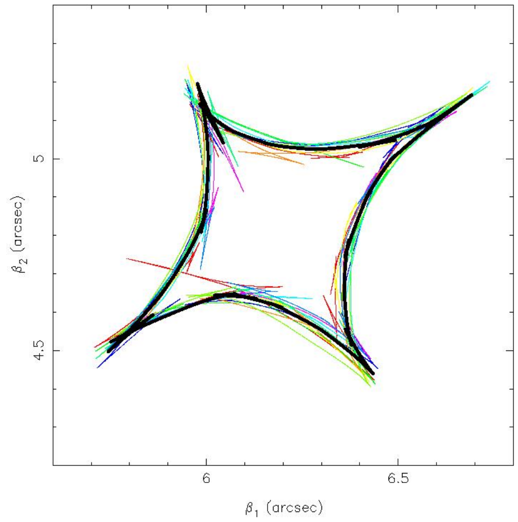


Figure 3.11: The original caustic curve (thick line) of the halo from Fig. 3.6a and the corresponding caustic curves from the ten bootstrapped maps plotted on top (thin lines).

also become brighter, the probability that they are demagnified is substantially larger (Keeton, 2003). This effect is referred to as saddle point demagnification.

The saddle point demagnification is observed in many lens systems and is especially noticeable in the systems showing a fold configuration. For a fold configuration one expects the two highly magnified images to have the same unsigned magnifications (in the absence of substructure). In the presence of substructure, this is not only violated, but in almost all cases the brightest saddle point image is fainter than the brightest minimum (see e.g. Kochanek & Dalal, 2004).

For these studies it is preferable to use lenses observed in radio frequencies; a typical quasar is believed to have a much larger radio emission region than its optical continuum counterpart. For this reason the radio-fluxes are less susceptible to stellar-mass microlensing which can mimic variations due to the CDM substructure. At present, there are $\sim$ 20 published quadruply imaged systems, and only a subsample of them is detected at radio frequencies. Kochanek & Dalal (2004) used 8 radio lenses, potentially introducing problems of “small number statistics”. However, if the saddle point demagnification shows up in a large sample of observed lenses (available in the future), this would be a major step forward in identifying flux anomalies either with substructure, or with propagation effects. If the observed flux anomalies depend on the image parity and its magnification we can set limits on the influence of the ISM; which influence fluxes of e.g. radio observations independent from the image parity.

To further study the effect of the saddle point demagnification, we fit all synthetic image systems using a singular isothermal ellipse model with external shear (SIE+SH) (Kormann et al., 1994). We do not include the flux ratios in the fit, since they are

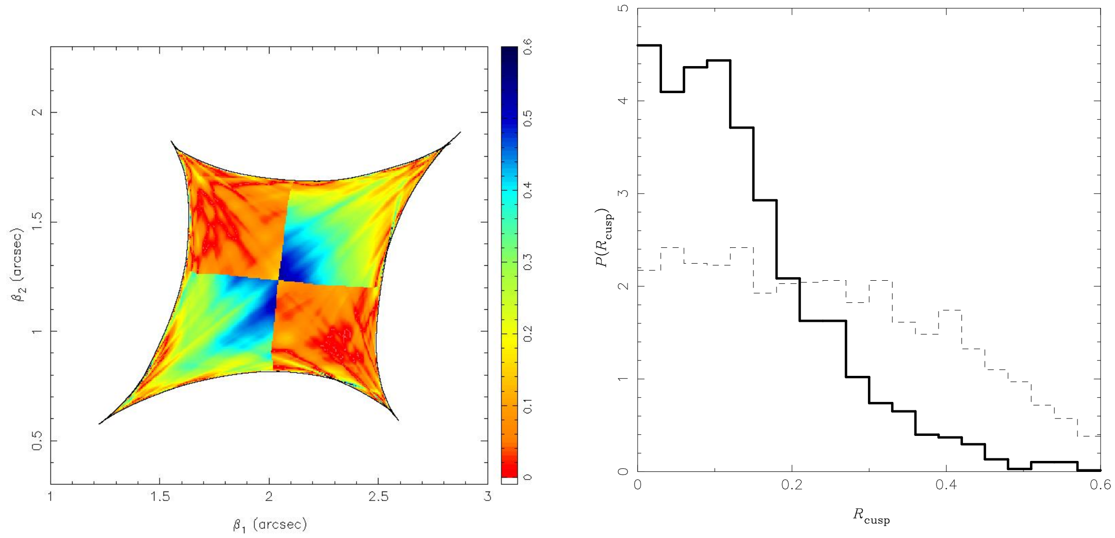


Figure 3.12: **(a)** The cusp relation for the smooth ellipsoidal model sampled with the same number of particles as present in the original N-body simulated elliptical galaxy. The redshifts of the source and the lens were kept at $z_{\rm l} = 0.81$ and $z_{\rm s} = 3$. $R_{\rm cusp}$ is plotted in colour. On top we plot the corresponding caustic curve. The absence of large fluctuation in $R_{\rm cusp}$ as compared to Fig. 3.6a shows that the signal is not dominated by the shot noise of the particles. **(b)** The probability distribution of $R_{\rm cusp}$ for systems with $\mu_{\rm tot} > 20$. The solid line gives the probability distribution for the sampled smooth ellipsoidal model, while the dashed line corresponds to the original N-body simulation (cf. Fig. 3.6a).

affected by the substructure. We also keep the lens position fixed. Using only image positions we have 7 free fitting parameters and 8 constraints. The parameters that we used are lens strength, two components of the ellipticity of the lens, two external shear components and the source position. The 4 image positions provide 8 constraints.

3.4.1 Surface mass density profile

We find that the average unsigned magnifications predicted by the SIE+SH model are higher than those from N-body simulations. This is true for all systems generated with the same halo; a consequence of the mass profile being steeper than isothermal around the typical position where the images are formed.

We have calculated the κ profile for the N-body simulated elliptical by fitting it with elliptical contours. For this purpose we use the `IRAF.STSDAS` package `ellipse`. In the inner region $r \lesssim 500$ pc the profile is very close to isothermal with the slope -1.0 ± 0.2. In the outer parts the profile becomes steeper than isothermal with slope -1.9 ± 0.2. The break radius, where the profile becomes steeper than isothermal, is smaller than the radius where images form.

We therefore over-predict the magnifications using an isothermal profile. In order to deal with this problem, one would preferably use a power-law profile with the index mentioned above for the lens modelling, instead of SIE. However, these profiles require the evaluations of hypergeometric functions and/or numerical integrations (Grogin & Narayan, 1996; Barkana, 1998). Modelling is therefore computationally too expensive to apply to several 10 000 sources. However, this is not critical, since we are not searching for the best fit macro-model, but rather pretending that we observe the systems and try to fit them with the generic model used for many observed lenses. Since we only observe the flux ratios and not directly the magnifications, it is impossible to compare magnification factors in practice. Thus one cannot determine the difference in profiles using only this consideration when dealing with real lens systems.

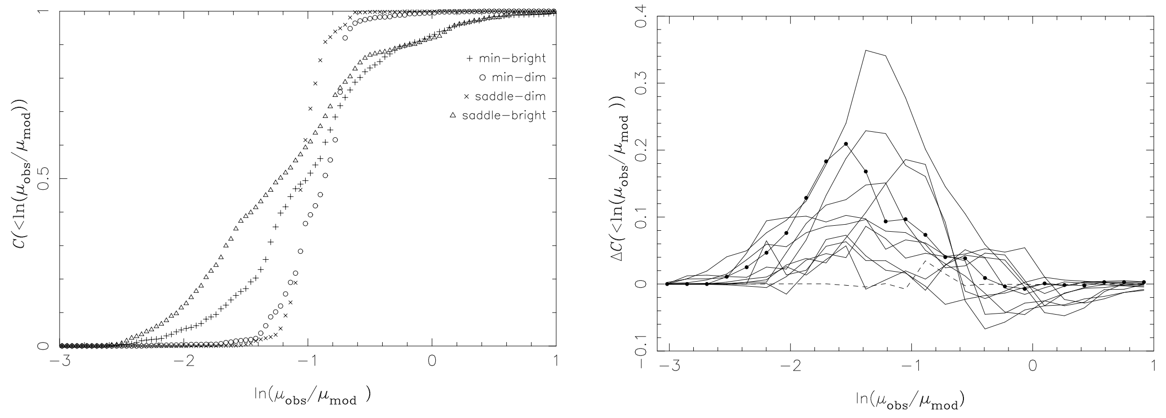
3.4.2 Suppressed saddle points

Having obtained the best-fit SIE+SH model parameters for the simulated synthetic systems we proceed as follows. As in Kochanek & Dalal (2004) we define the cumulative probability distribution of flux residuals $C\,(<\ln(\mu_{\rm obs}/\mu_{\rm mod}))$, where $\mu_{\rm obs}$ is the "observed" magnification of a simulated image and $\mu_{\rm mod}$ is the magnification predicted by the best-fitting smooth SIE+SH model (as described above). In Fig 3.13a we plot $C\,(<\ln(\mu_{\rm obs}/\mu_{\rm mod}))$ for the systems with "observed" total unsigned magnifications $\mu_{\rm tot} > 20$ for the simulated elliptical halo (see also Fig. 3.6a). We choose systems that have highly magnified images, because they are affected by substructure the most. We repeat the whole procedure with the same halo, but smoothed on a scale of $\sigma_{\rm G} \sim 5$ kpc; in this case essentially no difference is seen between images of different parity, as expected.

For the original halo the cumulative distribution is much broader for the brighter minimum and saddle point. This is in accordance with the conclusions from Mao & Schneider (1998); the higher the magnification, the more is the image flux affected by substructure. Among the two most magnified images, the saddle point is on average more demagnified compared to the brighter minimum. We have examined the two other orthogonal projections of the mass density of the same halo and we find that qualitatively the results do not change with projection.

The effect, however, is not as pronounced as in Kochanek & Dalal (2004). The reason is two-fold. First, our simulations have a resolution of $\sim 10^8\,M_\odot$ and structure at this scale and below is suppressed when using Delaunay tesselation. Kochanek & Dalal (2004), however, used SIS clumps with masses of $10^6\,M_\odot$. Since these are more numerous, they can enhance the effect. Further, fitting SIE to the global mass profile is not fully justified. The mass profile is known only for a handful of observed lens galaxies. Whereas the lens galaxies in MG1654+134, MG2016+112, 0047−281, and B1933+503 have a nearly isothermal profile (see Kochanek 1995; Cohn et al. 2001; Treu & Koopmans 2002a,b; Koopmans & Treu 2003), the one of PG1115+080 seems to be steeper (Treu & Koopmans, 2002b). Besides the absence of substructure on scales $< 10^8\,M_\odot$, our synthetic systems and their modelled quantities closely resemble the properties of realistic lenses and the way they are modelled.

In the analysis of Kochanek & Dalal (2004) the synthetic images were generated using an SIE macromodel with SIS substructure. This simplifies the model fitting and explains why they get a transition of cumulative distribution exactly at $\ln(\mu_{\rm obs}/\mu_{\rm mod}) \sim 0$.



(a) **(b)**

Figure 3.13: **(a)** The cumulative flux residuals for each type of image. Synthetic image systems were taken from the elliptical halo (see also Fig. 3.6a). Only image positions from the systems with $\mu_{\rm tot} > 20$ generated from the elliptical halo were fitted using SIE+SH. $\mu_{\rm obs}$ are the magnification factors taken directly from N-body simulations and $\mu_{\rm mod}$ are the ones from best fit SIE+SH models. **(b)** The difference of the cumulative distribution of flux residuals $\Delta C\,(<\ln(\mu_{\rm obs}/\mu_{\rm mod}))$ between the brightest saddle point and the brightest minimum images calculated using the original halo (solid line with dots, which also indicate the grid points used to evaluate ΔC), bootstrapped images (solid lines), and additionally smoothed halo on a scale of $\sigma_{\rm G} \sim 5\,{\rm kpc}$ (dashed line).

In addition, we have looked at the cumulative flux mismatch distribution in the bootstrapped maps, to investigate the significance of our results. In the bootstrapped images we confirm the broader distribution for bright minimum and saddle point images. In Fig 3.13b we plot the difference of the cumulative distribution of flux residuals $\Delta C\,(<\ln(\mu_{\rm obs}/\mu_{\rm mod}))$ between the brightest saddle point and the brightest minimum images for the original halo (solid line with dots), bootstrapped images (solid lines) and halo when additionally smoothed on a scale of $\sigma_{\rm G} \sim 5\,{\rm kpc}$ (dashed line). Positive values of ΔC thus denote saddle point demagnification. In all bootstrapped images ΔC is positive, except for few points (corresponding to $\Delta C\,(<\ln(\mu_{\rm obs}/\mu_{\rm mod})) \sim 1$). This confirms that the effect of the saddle point demagnification is present and comparable to the original halo, whereas in the smoothed halo this effect is not seen. We conclude that substructure on mass scales $\gtrsim 10^8\,M_\odot$ significantly contributes to the saddle point demagnification.

3.5 Conclusion

In this part of my thesis we have studied strong gravitational lensing properties of N-body simulated galaxies. In particular, we concentrated on the influence of substructure on flux ratios in lens systems with high (unsigned) magnifications. This analysis is crucial to fully understand the lensing signal that we observe in real lenses and to disentangle the influence on lensed-image fluxes due to propagation effects in the ionised ISM and mass substructure.

We have examined two strong lensing signatures of substructure, i.e. the violation of cusp relation observed in images that show a typical cusp configuration and the saddle point suppression. The saddle point suppression has been previously studied using semi-analytic prescriptions of substructure (e.g. Schechter & Wambsganss, 2002; Kochanek & Dalal, 2004). The effect of substructure on the cusp relation, however, has up to now not been studied in detail.

To determine the magnitude of both effects we use N-body simulated galaxies. The difference compared to the works of Schechter & Wambsganss (2002) and Kochanek & Dalal (2004) is that we are using a representation of substructure that is as realistic as possible and does not make any assumptions on the mass function and abundance of sub-halos. The drawback, however, is the resolution of the simulations. We are therefore not able to extrapolate the analysis to masses $< 10^8 M_\odot$. Still, the signatures of both effects are clearly present, and in the future we plan to use higher-resolution N-body plus gasdynamical simulations to explore their effects in even greater detail.

The main question when dealing with N-body simulations is how much are the magnification factors, that we use for synthetic image systems, affected by noise which can mimic substructure of $< 10^8\, M_\odot$. We show that the average relative noise in the surface mass density $\sigma(\kappa)/\kappa$ lies below the $\sim 5\%$ level for $\kappa \gtrsim 0.25$. Second, in Sect. 3.3.4 we show that the results for $R_{\rm cusp}$ are not significantly affected by the noise, and are dominated by physical substructure. The signal is dominated by several resolved mass clumps, which in projection lie close to the Einstein radius. Similarly, in the case of the phenomenon of suppressed saddle points, the bootstrap analysis shows that the signal is not dominated by noise.

The behaviour of $R_{\rm cusp}$ for sources close to a cusp is a very promising tool to detect substructure. Its main advantage is that it makes definite, model-independent predictions for the image magnifications. These predictions can only be broken in the presence of structure in the potential on scales smaller than the image separation or propagation effects. In Fig. 3.6, where we used a simulated elliptical galaxy to calculate $R_{\rm cusp}$, we clearly see these effects. When smoothing the substructure on larger scales we witness the transition to what is common for generic smooth SIE lens model.

However, the disk in the disk galaxy can also help to destroy the cusp relation for sources in the vicinity of a cusp. We have calculated the cusp relation pattern for the simulated disk galaxy, and even in the absence of obvious substructure in the form of clumps we can see strong violations of the cusp relation. Since the edge-on disk gives κ-variations on scales smaller than the image separation, violations of $R_{\rm cusp}$ are expected. However, one cannot conclude from a broken cusp relation alone that we observe the signatures of mass substructure in the form of clumps. Since the observations show that most observed lenses are ellipticals we have concentrated on this morphology. Despite the apparent simplicity of such lens galaxies, detailed modelling is required in most cases (e.g. B1422+231) to clearly see the effect.

The phenomenon of suppressed saddle points is a very strong prediction that rules out a significant influence of the ISM on flux anomalies, if confirmed by other observations of more lens systems. If flux anomalies depend on parity and magnification, they must clearly be caused by lensing of galaxies with significant substructure present. Observations so far show a clear parity dependence, which is more obvious for highly magnified images.

Finally, our analysis shows that the two brighter images are more affected by substructure than the two fainter ones. We confirm that the brightest saddle point image in N-body simulated systems has a higher probability to be demagnified, in accordance with predictions from microlensing and from semi-analytic work by Kochanek & Dalal (2004). It is therefore necessary that all mass scales are properly accounted for in these calculations, in order to correctly compare observations with theoretical predictions in detail.

For future work, we plan to study jet curvature using N-body simulations. At present, there is only one case of a curved jet observed that might be caused by gravitational lensing (Metcalf, 2002). It will be interesting, however, to investigate the probability of more of these occurrences. We plan to investigate the signal one expects on average for multiple-imaged jets; this signal is also less affected by noise in the simulation and low-mass substructure, since it depends mostly on first order derivatives of the potential.

In summary, gravitational lensing is a very powerful tool for testing the existence of CDM substructure. N-body simulated galaxies seem to produce similar effects as seen in observed lens systems. In addition, systematics on flux anomalies (scatter broadening, scintillation, microlensing) can be efficiently ruled out or confirmed by multi-frequency observations of lenses. The statistical analysis of large samples of lenses can directly probe the properties of CDM substructure in galaxies to a redshift of $z \sim 1$. This provides a unique tool to measure the evolution of these structures with cosmic time, as predicted in the hierarchical structure formation scenario.

4

Weighing galaxy clusters using statistical (weak) lensing

Galaxy clusters have been the focus of a very intense ongoing research in the past decades. As we note in the introduction, the mass distribution of galaxy clusters is particularly important for cosmological studies because they provide a critical test of the Cold Dark Matter (CDM) paradigm.

Statistical (weak) lensing offers an extremely useful tool to measure the cluster mass distribution. Clusters are massive enough to be detected individually by the statistical weak lensing effect, in contrast to individual galaxies where the combined signal from an ensemble of galaxies acting as a lens (for an ensemble of sources) is investigated and only the distribution function of galaxy properties can be inferred. This technique is also known as galaxy-galaxy lensing and is *not* a subject of this thesis.

I will concentrate on the cluster-mass reconstruction method that uses image distortion information. In the introduction, the theory on how an elliptical source is transformed under the influence of lensing is explained; in this chapter we discuss the practical approaches.

The effect of statistical weak lensing has long been known and observed (see Tyson et al., 1990). However it was only after the pioneering work of Kaiser & Squires (1993; see e.g. Clowe & Schneider, 2001, 2002 for further applications) that the field began to flourish. Since then several other mass reconstruction techniques and applications of weak lensing have been developed (see e.g. Kaiser et al., 1995; Lombardi & Bertin, 1999; Seitz & Schneider, 1997; Bridle et al., 1998; Marshall et al., 2002; Hoekstra et al., 1998).

The dominant problem one needs to overcome in the case of cluster mass reconstruction is to separate the distortion effects due to lensing from secondary effects caused by e.g. light propagation through the atmosphere and through the telescope optics. This separation can be done by measuring the shapes of the stars in the field, for which we can assume that they are intrinsically circular. The correction applied is highly technical and was developed especially for the weak lensing purposes by Kaiser et al. (1995).

An alternative to using the image distortion is to use the magnification properties of lensing (see Broadhurst et al., 1995). The local number count of background sources is directly related to the magnification of the lens, which to first order is given by $1 + 2\kappa$ (for $\kappa \ll 1$). As a result, a careful estimate of the local density of background sources can lead to a direct measurement of the projected density κ (see Fort et al., 1997; Taylor et al., 1998 for reported detections of this effect). In addition, other methods to measure cluster masses via magnification effects have been proposed (e.g. Bartelmann & Narayan, 1995 suggested to use the sizes of galaxies at a given surface brightness as a measure of the local lens magnification). All these methods, however, require a precise calibration of external parameters, for example the number counts of unlensed sources, which is very difficult to obtain with desired accuracy (Schneider et al., 2000). Hence, we concentrate only on methods using image distortions and discussing the possible use of magnification methods where necessary.

This chapter of the thesis is in large extent published in Bradač et al. (2004a).

4.1 Cluster mass reconstruction methods

The basic idea of using image distortions for cluster mass reconstructions was described in Sect. 2.3. From the measured image ellipticities one can obtain an unbiased estimate of the reduced shear by performing ensemble averaging –c.f.(2.31). The reduced shear depends on the shear γ and the surface mass density κ; both can be iteratively determined. From κ one can in principle determine the mass distribution if there were not for the mass-sheet degeneracy. This very important issue of weak lensing is described in the next section.

4.1.1 The problem of the mass-sheet degeneracy

Unfortunately, all statistical weak lensing can determine the projected surface mass density κ only up to a degeneracy transformation $\kappa \to \kappa' = \lambda\kappa + (1 - \lambda)$, where λ is an arbitrary constant. This invariance (first recognised by Falco et al. 1985) fundamentally limits the accuracy of cluster mass determinations if no further assumptions are made. In particular, this transformation leaves the relation between the intrinsic and observed ellipticity unchanged and therefore it can not be broken by using only measurements of the distortion of the background sources, if these all lie at the same distance from the observer.

A naive solution to this problem is to constrain λ by making simple assumptions about κ. For example, one can assume that the surface mass density is decreasing with distance from the centre, implying $\lambda > 0$. In addition, κ is likely to be non-negative, and so one can obtain an upper limit on λ (for $\kappa < 1$). More quantitatively, with the use of wide field cameras one might try to assume that $\kappa \simeq 0$ at the boundary of the field, far away from the cluster center. However, if we consider for example a $M_{\rm vir} = 10^{15} M_\odot$ cluster at redshift $z = 0.2$, we expect from N-body simulations to have a projected dimensionless density of about $\kappa \simeq 0.005$ at 15 arcmin from the cluster center (Douglas Clowe, private communication). Hence, even with the use of a 30×30 arcmin camera we expect to underestimate the virial mass of such a cluster by $\sim 20\%$.

The mass-sheet degeneracy can not be lifted by using the shapes of the background sources alone (Seitz & Schneider, 1997), and so one needs to make use of additional information. One promising solution to this problem is provided by the so-called magnification effect (see Broadhurst et al., 1995). Indeed, as mentioned before the local number counts of background sources is directly related to the magnification of the lens, which to first order is given by $1 + 2\kappa$. As a result, a careful estimate of the local density of background sources can lead to a direct measurement of the projected density κ and therefore breaks the mass-sheet degeneracy. However, all these methods require a good knowledge of external parameters, and the accuracy needed to break the mass-sheet degeneracy is extremely high and therefore it is more favourable to use distortion information (Schneider et al., 2000).

Here I focus on the possible methods to break the mass-sheet degeneracy using both distortion and redshift information of background sources. Seitz & Schneider (1997) already considered this problem and showed that the mass-sheet degeneracy can be *weakly* broken, provided that the probability distribution of the redshifts for the observed sources is known with good accuracy. Photometric redshifts based on multi-band photometry can be extremely accurate (see e.g. Benítez, 2000). Hence, it is sensible to assume that the *individual* redshifts of all background sources can be known in weak lensing studies. We make use of this information in order to investigate to what extent the mass-sheet degeneracy can be broken in practice.

4.1.2 The mass-sheet degeneracy in theory

In the simple case of background sources all having the *same redshift*, the mass-sheet degeneracy can be understood just using equations given in Sect. 2.3. Indeed, consider for a moment the transformation of the potential ψ

$$\psi(\boldsymbol{\theta}, z) \to \psi'(\boldsymbol{\theta}, z) = \frac{1-\lambda}{2}\boldsymbol{\theta}^2 + \lambda\psi(\boldsymbol{\theta}, z) \; , \tag{4.1}$$

where λ is an arbitrary constant. κ and γ are related to the potential ψ through its second partial derivatives (denoted by subscript), namely

$$\kappa = \frac{1}{2}(\psi_{,11} + \psi_{,22}) \; , \; \gamma_1 = \frac{1}{2}(\psi_{,11} - \psi_{,22}) \; , \; \gamma_2 = \psi_{,12} \; . \tag{4.2}$$

From (4.2) it follows that κ transforms as

$$\kappa(\boldsymbol{\theta}, z) \to \kappa'(\boldsymbol{\theta}, z) = \lambda\kappa(\boldsymbol{\theta}, z) + (1 - \lambda) \; , \tag{4.3}$$

and similarly the shear changes as

$$\gamma(\boldsymbol{\theta}, z) \to \lambda\gamma(\boldsymbol{\theta}, z) \; . \tag{4.4}$$

Therefore the reduced shear $g(\boldsymbol{\theta}, z)$ remains invariant.

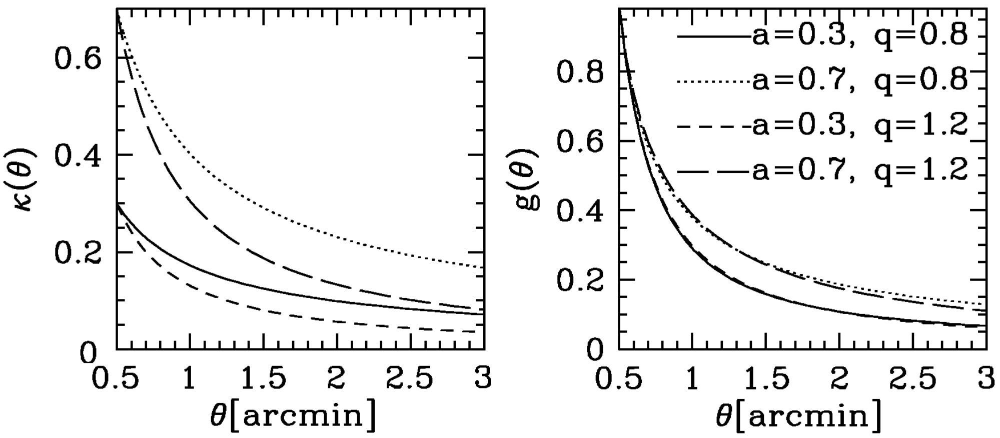


Figure 4.1: The surface mass density $\kappa(\theta)$ and reduced shear $g(\theta)$ for a spherically symmetric power law model with amplitude a and slope q. The models are pairwise nearly degenerate, i.e. different mass profiles have very similar reduced shear (Schneider et al., 2000).

This is in addition presented in Fig. 4.1 on an example of a simple spherically-symmetric power-law model with

$$\kappa(\theta) = a \left(\frac{\theta}{\theta_{\rm E}} \right)^{-q} \quad \text{for } \theta > \theta_{\rm E} \,, \tag{4.5}$$

and having mean surface mass density inside the Einstein radius $\langle\kappa\rangle = 1$. In Fig. 4.1 (calculated for $\theta_{\rm E} = 0.5$ arcmin) the models with very different slopes q would have very similar reduced shear and thus we would not be able to distinguish these models using weak lensing (if all sources would lie at the same redshift, thus reduced shear g would depend on θ only).

Seitz & Schneider (1997) have shown that in the case of a known *redshift distribution*, a similar form of the mass-sheet degeneracy holds to a very good approximation for non-critical clusters; i.e. for clusters with $|g(\boldsymbol{\theta}, z)| \leq 1$ for all source redshifts z. In such a case the standard weak-lensing mass reconstruction is affected by the degeneracy

$$\kappa \rightarrow \kappa' \simeq \lambda\kappa + \frac{(1-\lambda)\,\langle Z(z)\rangle}{\langle Z^2(z)\rangle} \,, \tag{4.6}$$

where $\big\langle Z^n(z)\big\rangle$ denotes the n-th order moment of the distribution of cosmological weights. As a result, *standard* weak-lensing reconstructions are still affected by the mass-sheet degeneracy even for sources at different redshifts; moreover, simulations show that the degeneracy is hardly broken even for the lenses close to critical.

In this work we therefore use the information of *individual redshifts* of background sources to break this degeneracy. Suppose for simplicity that half of the background sources are located at a known redshift $z^{(1)}$, and the other half at another known redshift $z^{(2)}$. The weak lensing reconstructions based of the two populations provides two different mass maps, $\kappa'(\boldsymbol{\theta}, z^{(1)})$ and $\kappa'(\boldsymbol{\theta}, z^{(2)})$, leading to two different forms of the mass-sheet degeneracy. In other words, the two mass reconstructions ($i = 1, 2$) are given by

$$\kappa'(\boldsymbol{\theta}, z^{(i)}) = \lambda^{(i)} \kappa_{\rm t}(\boldsymbol{\theta}, z^{(i)}) + \left(1 - \lambda^{(i)}\right) \tag{4.7}$$

where we have denoted with $\kappa_{\rm t}(\boldsymbol{\theta}, z^{(i)})$ the true projected κ of the lens at the angular position $\boldsymbol{\theta}$ for sources at redshift $z^{(i)}$. Since the transformation (4.7) holds for any $\boldsymbol{\theta}$, we have a system of equations to be solved for $\lambda^{(1)}$ and $\lambda^{(2)}$. The relation between $\kappa_{\rm t}(\boldsymbol{\theta}, z^{(1)})$ and $\kappa_{\rm t}(\boldsymbol{\theta}, z^{(2)})$ is known, namely from Eq. (2.26) follows

$$\kappa_{\rm t}(\boldsymbol{\theta}, z^{(1)}) Z(z^{(2)}) = \kappa_{\rm t}(\boldsymbol{\theta}, z^{(2)}) Z(z^{(1)}). \tag{4.8}$$

Suppose one measures both $\kappa'(\boldsymbol{\theta}, z^{(i)})$ at N different positions $\boldsymbol{\theta}_j$, this gives us a system of $2N$ equations to be solved for $\lambda^{(i)}$ and $\kappa_{\rm t}(\boldsymbol{\theta}_j)$. This theoretically allows us to break the mass-sheet degeneracy.

It is interesting to observe that this argument only applies to relatively "strong" lenses. For "weak" lenses, i.e. lenses for which we can use a first order approximation in κ and γ, the expectation value of measured image ellipticities is $\left\langle \epsilon(z) \right\rangle = \gamma(\boldsymbol{\theta}, z)$. In such case the degeneracy of the form

$$\psi(\boldsymbol{\theta}, z) \to \psi'(\boldsymbol{\theta}, z) = \frac{1 - \lambda}{2} \boldsymbol{\theta}^2 + \psi(\boldsymbol{\theta}, z) \tag{4.9}$$

leaves the observable $\gamma(\boldsymbol{\theta}, z)$ unchanged. As a result, the method described above cannot be used to break the mass-sheet degeneracy for these lenses. Only when the $(1 - Z(z)\kappa)$ term in the reduced shear becomes important and $g(\boldsymbol{\theta}, z)$ can be distinguished from $\gamma(\boldsymbol{\theta}, z)$ in the (noisy) data, will we be able to make unbiased cluster mass reconstructions.

The aperture mass measures are very convenient for measuring the weak lensing signal. A simple form of such measures is the so-called ζ statistics (Kaiser, 1995),

$$\zeta(\boldsymbol{\theta}; \vartheta_1, \vartheta_2) \equiv \bar{\kappa}(\boldsymbol{\theta}; \vartheta_1) - \bar{\kappa}(\boldsymbol{\theta}; \vartheta_1, \vartheta_2) \ . \tag{4.10}$$

It measures the difference between the mean surface mass densities in a circle of radius ϑ_1 at position $\boldsymbol{\theta}$ and in an annulus between the radii ϑ_1 and ϑ_2. As a consequence it measures the weighted lensing signal (in the weak lensing limit) and can not be used to determine the true κ. Therefore despite the fact that it is not affected by the mass-sheet degeneracy transformation of the form (4.9), it can not be used for determining the cluster mass.

The use of individual redshifts of background sources therefore provides the most direct method to break the mass-sheet degeneracy. The simple argument sketched above for two redshift planes is therefore considered in more detail in this thesis and generalised to background galaxies at different redshifts. In this chapter we use mock catalogues of image ellipticities, generated under optimistic assumptions. Such an approach allows us to investigate the minimum lensing strength needed in order to lift the degeneracy in forthcoming observations. As we will see below, this straightforward method can be successfully used to break the mass-sheet degeneracy in the strong lensing regime, and is particularly effective for critical lenses. In the next chapter we describe a new non-parametric cluster mass reconstruction method and apply the approach of breaking the mass sheet degeneracy to real data.

4.1.3 χ^2 approach

For the purpose of making the mass reconstruction it is very useful to adopt a Bayesian approach and to write the likelihood function of a given observed configuration of galaxy ellipticities. Here we will start with the simplest approach. Consider a set of N_g observed (complex) ellipticities ϵ_i and let us assume for a moment that the errors on their distribution only originate from the intrinsic ellipticity distribution. Each component of the intrinsic ellipticity distribution is, to a very good approximation, distributed according to a Gaussian with dispersion $\sigma = \sigma_{\epsilon^\mathrm{s}_1} = \sigma_{\epsilon^\mathrm{s}_2}$. Under these assumptions one can define a χ^2 function for the $|g| \ll 1$ case through the difference between the observed ellipticity (in our case ϵ_i) and the expectation value $g(\boldsymbol{\theta}, z)$

$$\chi^2_\mathrm{wl}(\pi_k) = \sum_{i=1}^{N_\mathrm{g}} \frac{|\epsilon_i - g(\boldsymbol{\theta}, z)|^2}{\sigma^2_{\epsilon^\mathrm{s}}} \ . \tag{4.11}$$

The total intrinsic ellipticity dispersion is $\sigma^2_{\epsilon^\mathrm{s}} = \sigma^2_{\epsilon^\mathrm{s}_1} + \sigma^2_{\epsilon^\mathrm{s}_2} = 2\sigma^2_{\epsilon^\mathrm{s}_1}$. If the cluster mass distribution is parametrised by a set of model parameters π_k, minimising χ^2_wl with respect to these will give a best-fit model to the data. As long as χ^2_wl is of the order of the number of degrees-of-freedom ($N_\mathrm{dof} = N_\mathrm{g}$), the model fits the data well.

However, in the regime where the $|g| \ll 1$ approximation does not hold, χ^2_wl needs to be rewritten. First, the expectation value of image ellipticities needs to be inserted from (2.31) and the observed ellipticity dispersion needs to be properly accounted for. χ^2 is therefore given by

$$\chi^2(\pi_k) = \sum_{i=1}^{N_\mathrm{g}} \frac{|\epsilon_i - \langle\epsilon\rangle|^2}{\sigma_i^2} \ , \tag{4.12}$$

where

$$\sigma_i^2 = \left(1 - \left|\langle\epsilon\rangle\right|^2\right)^2 \sigma^2_{\epsilon^\mathrm{s}} + \sigma^2_\mathrm{err} \ . \tag{4.13}$$

However, simulations show that such approach does not give an unbiased result for critical lenses. This has two main reasons. First, the expression (4.13) is only an approximation for the true variance, that contains higher (even) order moments of the distribution of $|\epsilon^\mathrm{s}|$ (Lombardi, 2000). Second and more important, ellipticity errors are not normally distributed for lenses with large reduced shears, $|g| \simeq 1$ (see Geiger & Schneider, 1998, Fig. 5). The real distribution of errors are cumbersome to evaluate, but need to be properly accounted for.

4.1.4 Maximum-likelihood approach

The maximum-likelihood approach is simply a generalisation of the χ^2-fitting. Indeed they are the same in the case when ellipticity errors are normally distributed. As I mentioned above, this is the case only if we work in the regime where the weak lensing approximation is valid. In the general case, however, the probability distribution of observed ellipticities needs to be properly accounted for.

Let us call $p_{\epsilon^{\mathrm{s}}}(\epsilon^{\mathrm{s}})$ the probability distribution of *intrinsic* source ellipticities. For simplicity and better understanding we will assume here the following truncated Gaussian distribution

$$p_{\epsilon^{\mathrm{s}}}(\epsilon^{\mathrm{s}}) = \frac{1}{2\pi\sigma^2\bigl[1-\exp\bigl(-1/2\sigma^2\bigr)\bigr]}\exp\bigl(-\bigl|\epsilon^{\mathrm{s}}\bigr|^2/2\sigma^2\bigr) \ . \tag{4.14}$$

Note, however, that the results in this section are independent of the particular choice for $p_{\epsilon^{\mathrm{s}}}$. The *observed* distribution $p_\epsilon(\epsilon \mid g)$ is related to the intrinsic one by (see Geiger & Schneider, 1998)

$$p_\epsilon(\epsilon \mid g) = p_{\epsilon^{\mathrm{s}}}\bigl(\epsilon^{\mathrm{s}}(\epsilon \mid g)\bigr)\left|\frac{\mathrm{d}^2\epsilon^{\mathrm{s}}}{\mathrm{d}^2\epsilon}\right|(\epsilon \mid g) \ . \tag{4.15}$$

The Jacobian determinant from the previous equation is given by

$$\left|\frac{\mathrm{d}^2\epsilon^{\mathrm{s}}}{\mathrm{d}^2\epsilon}\right|(\epsilon \mid g) = \begin{cases} \dfrac{(|g|^2-1)^2}{|\epsilon\, g^\star - 1|^4} & \text{for } |g| \le 1 \ , \\[2ex] \dfrac{(|g|^2-1)^2}{|\epsilon - g|^4} & \text{for } |g| > 1 \ . \end{cases} \tag{4.16}$$

In the weak lensing limit, $|g| \ll 1$, the Jacobian determinant is simply unity.

In general, the *measured* ellipticity ϵ^{m} will differ from ϵ because of measurement errors:

$$\epsilon^{\mathrm{m}} = \epsilon + \epsilon^{\mathrm{err}} \ . \tag{4.17}$$

The error ϵ^{err} is a random variable whose distribution depends on the details of the ellipticity measurement algorithm. In the following, for simplicity we will assume that ϵ^{err} is distributed according to a Gaussian distribution (4.14) with dispersion σ_{err}. Note that the actual shape measurement which corrects for PSF smearing (see e.g. Kaiser et al., 1995) can yield measured ellipticities with $|\epsilon^{\mathrm{m}}| > 1$. In our case, we write the probability distribution for the measured ellipticities as

$$p_{\epsilon^{\mathrm{m}}}(\epsilon^{\mathrm{m}} \mid g) = \int p_\epsilon(\epsilon \mid g)\, p_{\epsilon^{\mathrm{err}}}(\epsilon - \epsilon^{\mathrm{m}})\mathrm{d}\epsilon \ , \tag{4.18}$$

where $p_{\epsilon^{\mathrm{err}}}(\epsilon^{\mathrm{err}})$ is the probability distribution of measurement errors. This convolution takes into account the aforementioned measurement errors and therefore we do not need to discard galaxies with $|\epsilon^{\mathrm{m}}| > 1$.

The likelihood function $\mathcal{L}$ is the product of probability densities for observed image ellipticities ϵ_i^{m}, and depends on the model parameters $\pi = \{\pi_k\}$ through the reduced shear at the image positions $\{g_i\}$:

$$\mathcal{L}(\pi) = \prod_{i=1}^{N_{\mathrm{g}}} p_{\epsilon^{\mathrm{m}}}(\epsilon_i^{\mathrm{m}} \mid g_i) \ . \tag{4.19}$$

Here N_{g} is the number of observed galaxies with measured ellipticities ϵ_i^{m}. It is more convenient to deal with the log-likelihood function, defined as

$$l(\pi) \equiv -\ln \mathcal{L}(\pi) = -\sum_{i=1}^{N_{\mathrm{g}}} \ln p_\epsilon(\epsilon_i \mid g) \equiv \sum_{i=1}^{N_{\mathrm{g}}} l_i \ . \tag{4.20}$$

By minimising $l(\pi)$ we obtain the most likely parameters $\pi_{\max}$ given the observations.

For each of the $N_{\rm g}$ galaxies we need to calculate the probability distribution $p_{\epsilon^{\rm m}}(\epsilon^{\rm m} \mid g)$ [cf. Eq. (4.18)] of measured image ellipticity $\epsilon_i^{\rm m}$ given the reduced shear g. The evaluation of the full likelihood function is therefore non-trivial, since $N_{\rm g}$ integrals need to be calculated. On one hand, we are dealing with two Gaussian probability distributions for source ellipticities $\epsilon^{\rm s}$ and the measurement error $\epsilon^{\rm err}$. However, the final measured errors for observed image ellipticities are not distributed according to a Gaussian in general, since the source ellipticities are first being lensed and only then the measurement errors are added (i.e. the Jacobian determinant in (4.15) is not unity).

If the reduced shear is small (i.e. for *undercritical* lenses), this problem can be solved by including the measurement errors in the *source* probability distribution $p_{\epsilon^{\rm s}}$: in other words, if $|g| \ll 1$, the Jacobian determinant is unity and $p_\epsilon(\epsilon \mid g) = p_{\epsilon^{\rm s}}\big(\epsilon^{\rm s}(\epsilon \mid g)\big)$. We can therefore interchange the convolution with $p_{\epsilon^{\rm err}}$ appearing in Eq. (4.18) with the lensing transformation of Eq. (4.15). The calculations are then trivial, since $p_{\epsilon^{\rm s}}$ and $p_{\epsilon^{\rm err}}$ can be taken to be Gaussians with dispersions $\sigma^2_{\epsilon^{\rm s}}$ and $\sigma^2_{\rm err}$ respectively; hence, we can just use for $p_{\epsilon^{\rm s}}$ a Gaussian with dispersion $\sigma^2_{\epsilon^{\rm s}} + \sigma^2_{\rm err}$.

This is, however, not the case for critical lenses. As we mentioned, the approximate form for the measured ellipticities [cf. Eq. (4.15)] gives biased results for lenses with a large fraction of background sources having $|g| \simeq 1$. For the same reason, we should not discard galaxies with $|\epsilon^{\rm m}| > 1$ without properly accounting for their removal in Eq. (4.18). Therefore, unless otherwise stated, we will use the probability distribution for measured ellipticity given in Eq. (4.18). It still does not account for the measurement errors on the redshifts of background sources, however these are less important.

4.1.5 Parametric vs. non-parametric reconstruction

In Sect. 4.1.3 and 4.1.4 we have assumed that the lens is described by a family of models with model parameters π_k. The resulting mass-model will therefore depend upon the special parametrisation we chose in each case, the reconstruction is thus *parametric*. This is a drawback with respect to the inversion techniques (e.g. Kaiser & Squires, 1993) where κ is calculated directly by performing a linear inversion of the γ field which is readily obtained from image ellipticities (for the weak lensing case; it can, however, also be extended to the non-weak case). These inversion methods for mass reconstructions are in theory non-parametric; however since a smoothing scale is needed to obtain e.g. continuous shear maps the methods are not non-parametric in practice.

In addition, the methods presented in Sect. 4.1.3 and 4.1.4 can easily be improved by using ‘generalised’ lens models; we will refer to them as *non-parametric*. The simplest way is to put the potential ψ on a regular grid, each pixel value $\pi_k = \psi_k$ is a free parameter. κ and γ can then be obtained by finite differencing methods. This method will be described in more detail in Chapter 5. An alternative of putting a potential on a grid is to use directly κ on a grid. However, using the potential is preferable, since κ and γ depend locally on the potential ψ and therefore both can be readily obtained. γ, on the other hand, is a non-local quantity of κ and can be calculated from κ only if we have an infinite field. Needles to say that this is in practice not the case.

4.2 Generating mock catalogues for weak lensing

To test whether we can break the mass-sheet degeneracy by using redshift information on the background sources, we first need to generate a mock catalogue on which the parametrised model fitting will later be applied.

The general idea of generating the weak lensing catalogues is to assume a background galaxy population, transform their source ellipticities into image ellipticities using the lensing transformation (assuming a certain lens potential) and then add measurement errors on the ellipticities.

In more detail, we proceeded as follows:

1. 2000 galaxies are distributed randomly across the field of 6×6 arcmin^2, giving a density of 55 galaxies arcmin^{-2}.

2. The intrinsic ellipticities $\epsilon^{\rm s}$ are drawn from a Gaussian distribution (4.14) characterised by $\sigma = \sigma_{\epsilon_1^{\rm s}} = 0.15$.

3. We draw the redshifts of the background sources from a gamma distribution (Brainerd et al., 1996)

$$p_{\rm z}(z) = \frac{z^2}{2\, z_0^3} \exp\left(-z/z_0\right) \ , \tag{4.21}$$

   with $z_0 = 2/3$; the mean redshift is $\langle z \rangle = 3 z_0 = 2$, and the mode is $z_{\rm mode} = 2 z_0 = 4/3$. For simplicity the corresponding cosmological weights are evaluated assuming an Einstein-de Sitter cosmology. We put the lens at redshift $z_{\rm d} = 0.2$. To break the mass-sheet degeneracy one might think it is advantageous to have a higher redshift lens, e.g. $z_{\rm d} = 0.4$, in which case most of the galaxies lie on the steep part of the function $Z(z)$ and have a higher scatter of Z_i values. However, this effect is compensated for by the fact that the average reduced shear is lower for a lens at higher redshift. What further favours the low redshift cluster is the fact that more galaxies we can observe lie in the background of the cluster, although this is only a marginal effect.

4. For each galaxy, we evaluate the local shear according to the lens model, its position and redshift; then we lens the galaxy ellipticities. Note that the positions of the source galaxies are not transformed, i.e. we neglected here the magnification effect of the lens.

5. The measurement error $\epsilon^{\rm err}$ on the observed ellipticities is drawn from the distribution (4.14) with $\sigma = \sigma_{\rm err} = 0.1$ and added to the lensed ellipticities.

6. In most cases we consider measurement errors on the redshifts of the galaxies to simulate the use of photometric redshifts. These errors are drawn from a Gaussian distribution with $\sigma_{\mathrm{zerr}} = 0.06\,(1+z_i)$ (see Benítez, 2000); when adding the errors we ensure that the resulting redshifts are always positive.

In generating the mock catalogues we have tried to simulate an ideal case, since we are trying to answer the question of whether and in which situations the mass-sheet degeneracy can be broken in statistical lensing mass reconstructions. From a comparison between independent observations of the same galaxies in the COMBO-17 survey, $\sigma_{\mathrm{err}} \simeq 0.1 - 0.15$ (for each component of the ellipticity) can be estimated (Kleinheinrich, 2003). The typical values of $\sigma_{\epsilon_1^{\mathrm{s}}}$ are optimistically $\simeq 0.2$. However even in a less optimistic case ($\sigma_{\epsilon^{\mathrm{s}}} = 0.3$, $\sigma_{\mathrm{err}} = 0.15$) our conclusions are not changed.

4.2.1 Lens models

For transforming source ellipticities to image ellipticities one needs to assume an underlying mass distribution. As illustration of the method, we assume two simple lens models; a lens with a constant κ and a spherically symmetric lens with isothermal profile at a large distance. The same models are also used for the maximum-likelihood minimisation. The model parameters that can vary are the ones most closely related to the mass-sheet degeneracy transformation.

We consider first the simplest model possible; a constant sheet of mass with external shear (Family I). More precisely, we set $\kappa(\boldsymbol{\theta}) = C_1$, $\gamma_1(\boldsymbol{\theta}) = C_2$, and $\gamma_2(\boldsymbol{\theta}) = C_3$, with C_i being constants. In this model the convergence and shear for each galaxy depend only on the redshift of the source and not on its position. Although this model is unrealistic, it is very useful as a test of our method. It is reasonable to assume that if we are not able to break the mass-sheet degeneracy in this simple and favourable situation, it is unlikely that we can break it in more realistic and complicated cases. This simple model can also be used as an indicator of how strong the lens should be in order to obtain a reliable estimate for the average surface mass density κ across the field.

The second family we used is a non-singular model that approximates an isothermal sphere at large distances (see Schneider et al., 1992) plus a constant sheet of surface mass density (hereafter Family II). The dimensionless surface mass density is given by

$$\kappa(\theta/\theta_{\mathrm{c}}) = \kappa_0 \frac{1+\theta^2/\left(2\theta_{\mathrm{c}}^2\right)}{\left(1+\theta^2/\theta_{\mathrm{c}}^2\right)^{3/2}} + \kappa_1 \;, \tag{4.22}$$

where κ_0, κ_1 are dimensionless constants and θ_{c} is the core radius.

When fitting these two models to the simulated data, we vary the parameters that most closely resemble the mass-sheet degeneracy transformation for the case when the redshift distribution is not known; i.e. $\kappa(\boldsymbol{\theta}) \to \kappa'(\boldsymbol{\theta}) = \lambda\kappa(\boldsymbol{\theta}, z) + (1-\lambda)$ and $\gamma(\boldsymbol{\theta}) \to \gamma'(\boldsymbol{\theta}) = \lambda\gamma(\boldsymbol{\theta})$. These are the scaling of κ and γ and the constant sheet in κ. In other words, we allow all three constants for model Family I and κ_0 and κ_1 in Family II to vary; the remaining parameters are kept fixed at the values that were used to generate the data.

If the mass-sheet degeneracy is still present, this would show up in the distribution of the best-fit parameters. Namely, for a *single redshift plane* we expect a linear dependence between e.g. κ and $|\gamma|$ for Family I. More precisely, eliminating λ from (4.3) and (4.4) we get the correlation:

$$\gamma'(\boldsymbol{\theta}) = \kappa'(\boldsymbol{\theta}) \left(\frac{\gamma_{\rm t}}{\kappa_t - 1} \right) - \left(\frac{\gamma_{\rm t}}{\kappa_t - 1} \right) , \tag{4.23}$$

where quantities with subscript "t" denote the input (true) values. A similar relation can be derived for the Family II, in the case of κ_0 vs. κ_1.

In the case of a known *redshift distribution* for non-critical lenses Eq. (4.6) needs to be used. The two transformation to consider in this case are

$$\begin{aligned} \gamma' &\simeq \lambda \gamma_{\rm t} \\ \kappa' &\simeq \lambda \kappa_{\rm t} + \frac{(1-\lambda)\ \langle Z(z) \rangle}{\langle Z^2(z) \rangle} \end{aligned} \tag{4.24}$$

The correlation is expected to be of the form

$$\gamma'(\boldsymbol{\theta}) \simeq \kappa'(\boldsymbol{\theta}) \left(\frac{\gamma_{\rm t}}{\kappa_t - \langle Z(z) \rangle / \langle Z^2(z) \rangle} \right) - \left(\frac{\gamma_{\rm t}}{\kappa_t \langle Z(z) \rangle / \langle Z^2(z) \rangle - 1} \right) . \tag{4.25}$$

Note that the mass-sheet degeneracy transformation in (4.24) is only an approximate relation and has been derived by Seitz & Schneider (1997) for non-critical lenses only. In the case of a *single* source redshift the mass-sheet degeneracy transformation keeps the reduced shear, and therefore the probability distribution of the observed ellipticities invariant. For the sources *distributed* in redshift this is however no longer the case. Thus in principle the mass sheet degeneracy is broken when more than a single source redshift is used. In practice, however, the effect on the probability distribution of observed ellipticities is very small. Using simulations we will therefore test whether the correlation described in (4.25) is still present if we use information on *individual* redshifts.

From the discussion in Sect. 4.1.3, it is apparent that galaxies with $|g| \simeq 1$ contribute significantly to the removal of the mass-sheet degeneracy for two reasons:

- The dispersion of the observed ellipticities is proportional to $\left(1 - \left|\langle\epsilon\rangle^2\right|\right)^2$ [cf. Eq. (4.13)], and so galaxies with $|g| \simeq 1$ provide more accurate estimates of the local reduced shear;
- Recalling the argument discussed after Eq. (4.7), galaxies with $|g| \ll 1$ present a simple mass-sheet degeneracy [cf. Eq. (4.9)] which cannot be broken using the method described here. Therefore one needs the information from galaxies with $|g| \simeq 1$ to break the degeneracy.

It is important to realize that the likelihood function calculated using (4.18) already takes into account the first point mentioned above. As a result, galaxies with $|g| \simeq 1$ contribute most to the breaking of the mass-sheet degeneracy, while those with $|g| \ll 1$

do not contribute at all (as sketched in Sect. 4.1.1). It is sensible, therefore, to take into account this point when using the approximate form of the mass-sheet degeneracy presented in Eq. (4.6). When calculating moments of distribution of cosmological factor $\langle Z^n(z) \rangle$, we will use the weighted averaging;

$$\langle Z^n(z) \rangle \rightarrow \langle Z^n(z) \rangle' = \frac{\sum_{i=1}^{N_\mathrm{g}} w_i Z_i^n}{\sum_{i=1}^{N_\mathrm{g}} w_i} \ , \tag{4.26}$$

where w_i are the weight factors. E.g. if $w_i = \mathrm{const.}$ is chosen, we recover the results of Seitz & Schneider (1997).

As mentioned above, however, galaxies with $|g| \simeq 1$ are expected to contribute more to the signal. Therefore they are expected to play a dominant role in breaking the mass-sheet degeneracy. For this reason we also use a different weighting scheme with $w_i = 1/\sigma_i^2$ (where σ_i is the the observed ellipticity dispersion defined in (4.13)). σ_i is only an approximation of the true dispersion, however, as we show later it is adequate enough to describe the approximate behaviour of the mass-sheet degeneracy. For galaxies with $|g| \simeq 1$ the w_i will therefore be larger than for the rest and this will reflect into a different linear relation between κ and γ as given in (4.25). In the following we will investigate both schemes with $w_i = \mathrm{const.}$ and $w_i = 1/\sigma_i^2$ weighting.

4.3 Results of the model fitting to the simulated data

For each set of true model parameters π_t we generate 100 mock catalogues. Using these, we search for the most likely parameters π_max by minimising the log-likelihood function (4.20). We perform the minimisation with `C-minuit` (James & Roos, 1975), a routine which is part of the CERN Program Library. `C-minuit` is designed to minimise a multi-parameter function and analyse its shape around the minimum. It simultaneously makes use of the gradient as well as downhill simplex method.

Despite the robustness of the routine, one has to be very careful when performing the minimisation, because the function (4.20) has logarithmic singularities for $|g| = 1$ if $\sigma_\mathrm{err} = 0$ is assumed. If, for a particular parameter set, a background galaxy happens to have $|g| \simeq 1$, the log-likelihood function (4.20) diverges and the minimisation procedure has difficulties to "climb" over such regions, possibly leading to convergence to secondary minima.

In the case of $\sigma_\mathrm{err} \neq 0$ this is less likely to happen, as the convolution with $p_{\epsilon^\mathrm{err}}$ in (4.18) is effectively smoothing the probability distribution. Unfortunately the integration in (4.18) has to be performed numerically. The noise resulting from numerical calculations is therefore present in the log-likelihood function and the minimisation routine is not able to search for a minimum, since the function is not smooth. In addition, the integrations are very CPU time consuming; one needs to perform them for each galaxy separately. To avoid these problems, we note that function p_{ϵ^m} (4.18) depends on three parameters (without a loss of generality we can assume g to be real and accordingly transform ϵ^m). We therefore evaluate p_{ϵ^m} on a three-dimensional grid (with GNU Scientific Library `qags` routine, `http://www.gnu.org/software/gsl`) and use tri-linear interpolation to evaluate the log-likelihood.

Finally, we stress that the model fitted to the data belonged, for all simulations, to the same family as the original model used to generate the galaxy catalogue. Although unrealistic, this assumption allows us to directly compare the deduced parameters with the original ones, and thus simplifies the evaluation and interpretation of the simulation results. The free parameters of both models are $\pi = \{\kappa, \gamma_1, \gamma_2\}$ for Family I, $\pi = \{\kappa_0, \kappa_1\}$ for Family II. As mentioned before, these correspond to the mass sheet degeneracy transformation for the case where redshift information is not known. By investigating the distribution of best fit parameters π for both models, we can answer the question of how strong the lens needs to be to lift the abovementioned degeneracy under idealised conditions.

Family I: Lens model with uniform κ and γ

We first generate mock catalogues using a lens model with constant κ and γ. In particular we use $\gamma_1 = \gamma_2 = 0.1$ (see Fig. 4.2, panel a) and $\gamma_1 = \gamma_2 = 0.2$ (panel b). Ten different values of κ ranging from 0.5 to 1.5 were chosen for both sets, they are given as squares in both plots. For each model we fit 100 simulated mock catalogues (generated as described above) and we mark with crosses the resulting best fit parameters (the original values of the parameters are represented as squares). In bottom panels we plot the relative bias and variance for the recovered values of κ defined as

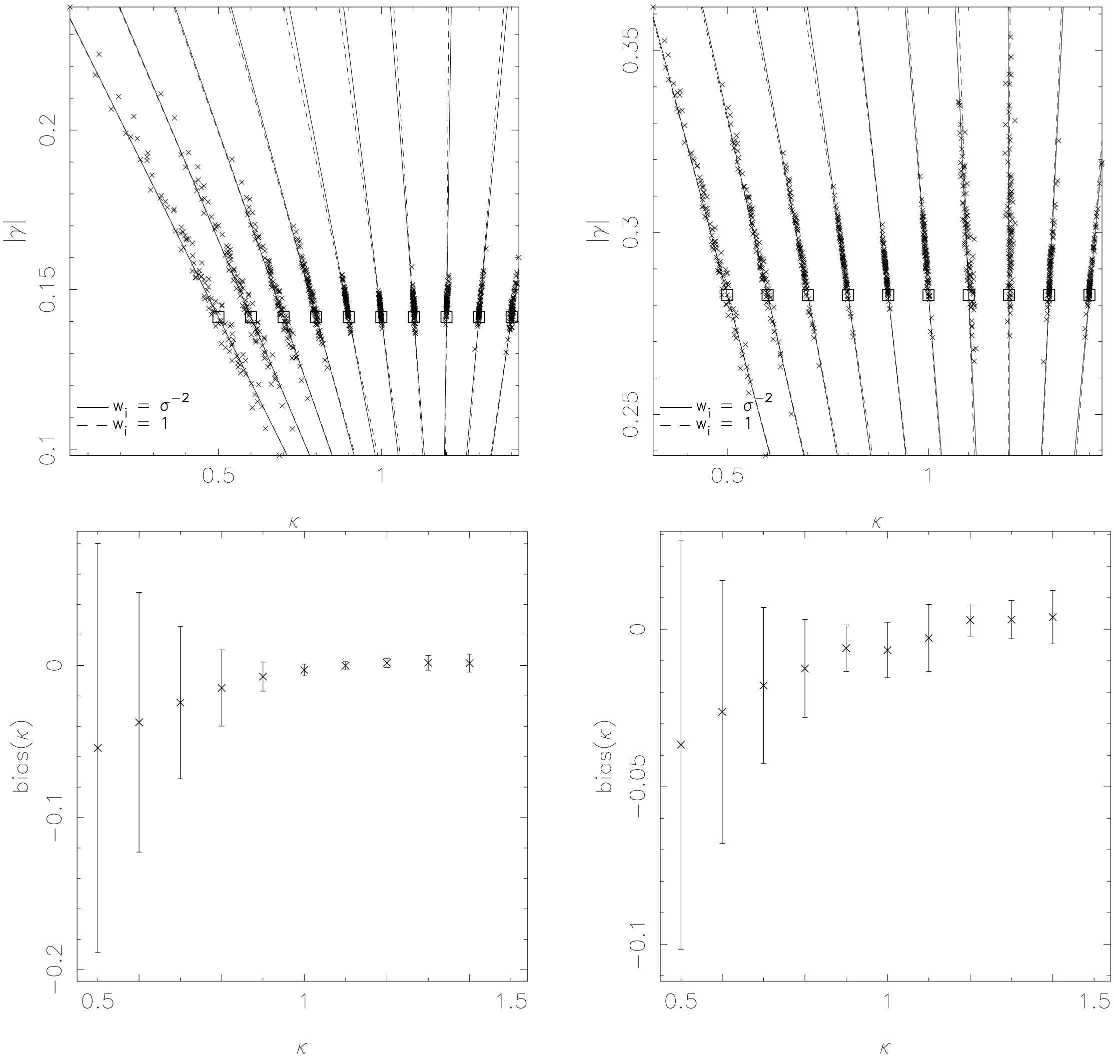
$$b(\kappa) = \sum_{i=1}^{N_\mathrm{m}} \left(\kappa_i - \langle\kappa\rangle\right) \; ; \qquad V(\kappa) = \frac{1}{N_\mathrm{m}} \sum_{i=1}^{N_\mathrm{m}} \left(\kappa_i - \langle\kappa\rangle\right)^2 \; , \tag{4.27}$$

where we sum over N_m mock catalogues and $\langle\kappa\rangle$ is the mean of best fit values κ_i. In order to explore the importance of redshift measurement errors we generate catalogues with and without redshift errors as described above (in Fig. 4.2 we show only the latter).

Surprisingly, even for a relatively strong lens with $\kappa = 0.5$ we are *not* able to break the mass-sheet degeneracy with good accuracy. For this model, the relative error on κ is of the order of 20%–40%, which should be considered very large, given the idealised conditions used in our simulations. Hence, the results of our simulations suggest that it is *very difficult* to break the mass-sheet degeneracy for non-critical lenses using shape information only. As expected, for data without redshift errors the fit slightly improves, although not significantly.

On the other hand, for critical and close to critical lenses, we are *effectively able to break the mass-sheet degeneracy*. The surface mass density is well constrained for $\kappa \gtrsim 1$. The resulting best fit parameters lie along the line calculated using (4.26) and $w_i = 1/\sigma_i^2$ rather than $w_i = \text{const}$. For one, this tells us that we still see a (rather weak) correlation left over from the mass-sheet degeneracy. More importantly, the results show that galaxies with $|g| \simeq 1$ contribute substantially to the removal of the degeneracy. If the redshift uncertainty is not added the fit improves and the constraints are tighter.

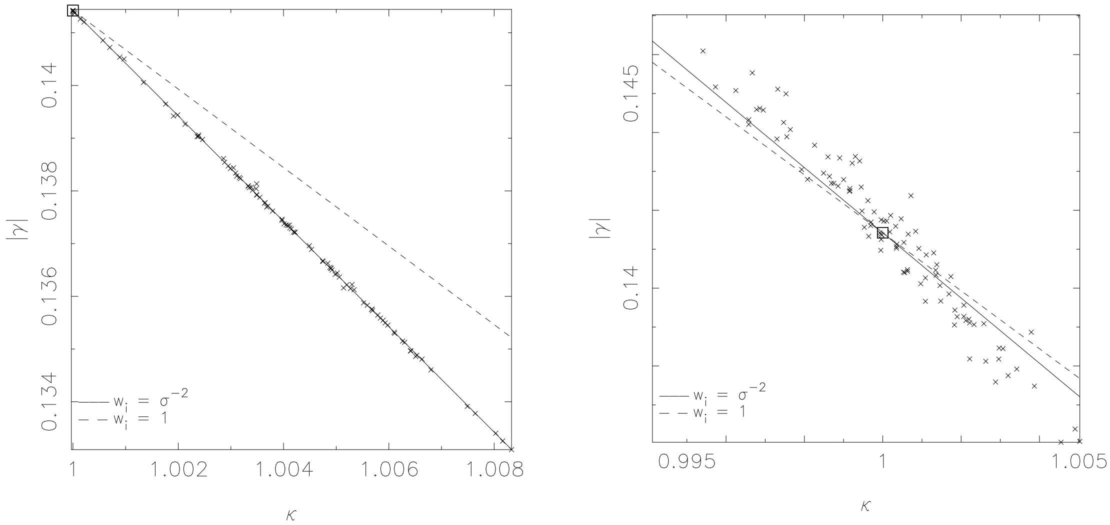
Finally, it is interesting to observe that in some cases (e.g. $\kappa = 1$ and $|\gamma| = 0.14$) we obtain biased results if we use p_ϵ instead of $p_{\epsilon^\mathrm{err}}$ for calculating log-likelihood (see



(a) **(b)**

Figure 4.2: Top: Recovered parameter values (crosses) as a result of minimising the log-likelihood function (4.20). For each of the ten sets of parameters (panel **(a)** $\gamma_1 = \gamma_2 = 0.1$, panel **(b)** $\gamma_1 = \gamma_2 = 0.2$) – denoted by squares – 100 mock catalogues were created using model Family I and the same family was used to fit the data. We use 3 free parameters (κ, γ_1, and γ_2) for fitting; we plot κ and $|\gamma| = \sqrt{\gamma_1^2 + \gamma_2^2}$. Lines correspond to the expected mass-sheet degeneracy calculated using the weighting scheme given by $w_i = 1/\sigma_i^2$ (solid lines) and $w_i = \text{const.}$ (dashed lines) in (4.26). Bottom: Bias and variance for the recovered values of κ for the data described above.

Fig. 4.3). For example, in Fig. 4.3a we use the likelihood function that first convolves the data with $p_{\epsilon^{\rm err}}$ and only then performs the lensing transformation [see discussion after Eq. (4.20)]. This can severely bias results for lenses where we have many galaxies for which $|g| \simeq 1$. Note that the expected mass-sheet degeneracy, calculated using $w_i = 1/\sigma_i^2$ in Eq. (4.26), needs to be calculated with $\sigma_{\rm err} = 0$, because we are not



(a) **(b)**

Figure 4.3: Recovered parameter values (crosses) as a result of minimising the log-likelihood function Eq. (4.20). Panel **(a)** shows results obtained using $p_\epsilon(\epsilon \mid g)$ – Eq. (4.15) – for the observed probability distribution for ellipticities, while for the panel **(b)** we use (as in the rest of this Chapter) $p_{\epsilon^{\rm m}}(\epsilon^{\rm m} \mid g)$ – Eq. (4.18). For the parameter set $\kappa = 1.0$, $\gamma_1 = \gamma_2 = 0.1$ – denoted by a square – 100 mock catalogues were created using model Family I and the same family was used to fit the data. Redshift errors were *not* added. We use 3 free parameters (κ, γ_1, and γ_2) for fitting, here we plot κ and $|\gamma|$. Lines correspond to the expected mass-sheet degeneracy calculated using Eq. (4.26). We used the weighting scheme given by $w_i =$ const. for the dashed lines. For the solid lines $w_i = 1/\sigma_i^2$ was used, however $\sigma_{\rm err} = 0$ was employed for panel **(a)** only. This is due to the fact that we are using $p_\epsilon(\epsilon \mid g)$ in **(a)** for the observed probability distribution for ellipticities, rather than $p_{\epsilon^{\rm m}}(\epsilon^{\rm m} \mid g)$.

correctly accounting for the measurement errors in this case.

In addition to the global, mass sheet invariance, there is a local point invariance. Looking at a single image ellipticity we can not distinguish between g and $1/g^*$ (Schneider & Seitz, 1995). This invariance is in practice easily broken, either by assuming a profile or simply by stating that κ increases towards the cluster centre (as is expected from the simulations). Now consider all sources being at the same redshift and we use the model Family I, i.e. $\kappa =$ const.. The invariance becomes a global invariance (we can not distinguish between g and $1/g^*$ over the whole field). In this case the log-likelihood function has two equivalent minima, one corresponding to g and one to $1/g^*$. However, since we do have the redshift information, this invariance is broken in most cases.

For our special case with $\kappa = 1$ and $|\gamma| = 0.14$ this is not so simple. The average absolute value of reduced shear for the galaxies is $\langle |g| \rangle \simeq 1$ and therefore for most galaxies the two minima lie close together. If the probability distribution of lensed ellipticities is properly accounted for, the resulting solutions give an unbiased estimate of the input parameters (see Fig. 4.3b). If this is not the case, as in Fig. 4.3a, the results can be biased.

This special case was presented as an extreme example. In reality such lenses do not exist. However, it clearly shows that critical lenses need to be treated with caution when using statistical lensing.

Family II: "isothermal" lens model

Mock catalogues for Family II were generated using $\kappa_0 = 0.9$, $\theta_\mathrm{c} = 1.5$ arcmin, and 4 different values of κ_1 ranging from 0.2 to 0.5. We fit these data with the same family of models and use κ_0 and κ_1 as free parameters; note that we fixed the core radius to the same value as the original profile. Figure 4.4a shows the results of log-likelihood minimisation for the data with added redshift errors. Solid and dashed lines in the figure correspond to the expected mass-sheet degeneracy calculated using the weighting of $w_i = 1/\sigma_i^2$ and $w_i = \mathrm{const.}$ in Eq. (4.26) respectively; dot-dashed lines give $\kappa_0 + \kappa_1 = \mathrm{const.}$

Finally we show the bias and the variance of the recovered values for κ_1 in Fig. 4.4b. When the lens becomes critical, the errors are small enough and we are able to constrain κ_1 with high accuracy. For these lenses we are therefore *effectively able to break the mass-sheet degeneracy*. This result is in accordance with the conclusions from the constant-lens model.

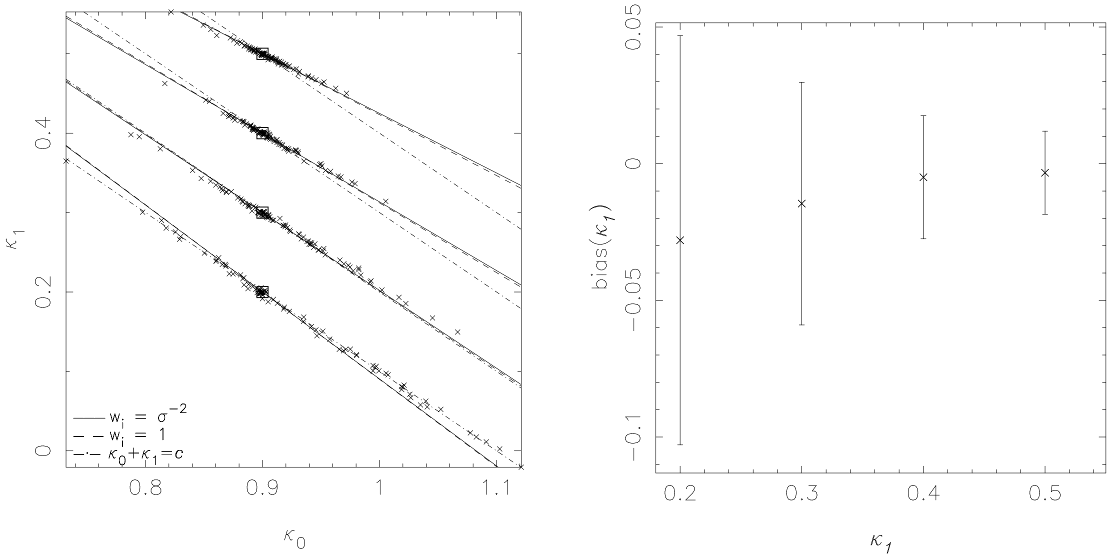
Surprisingly, for the case $\kappa_0 = 0.9$ and $\kappa_1 = 0.2$ we do not see the expected mass-sheet degeneracy, rather, the best-fit parameters lie along the line $\kappa_0 + \kappa_1 = \mathrm{const.}$ Likely, this is due to the fact that for this model we have galaxies with $|g| \simeq 1$ close to the centre, where the surface mass density is $\kappa(0) = \kappa_0 + \kappa_1$. Since the galaxies with $|g| \simeq 1$ are the ones that contribute most in breaking the mass-sheet degeneracy, it is not surprising for this model to see a degeneracy along the line $\kappa_0 + \kappa_1 = \mathrm{const.}$

If we take θ_c to be a free parameter in the fitting, we obtain the degeneracy $\kappa_0 + \kappa_1 = \mathrm{const}$ for all four sets of true model (input) parameters π_t. Since θ_c is allowed to vary, it can adjust so that the resulting best-fit model has, for lower values of κ_0, lower values of θ_c. The region of $|g| \simeq 1$ is then approximately unchanged as long as $\kappa_0 + \kappa_1 = \kappa_{0\mathrm{t}} + \kappa_{1\mathrm{t}}$.

4.3.1 Ensemble-averaged log-likelihood

It is common to summarise the properties of a distribution in terms of confidence limits. Knowing the confidence limits of the distribution one can than make statements such as "there is 99% chance that the true parameter value falls within this given region". An example of a two-dimensional distribution and the confidence limits is presented in Fig. 4.5.

Using the log-likelihood function (4.20) we can therefore in principle obtain the best fitting parameters π_max given the observations and the confidence regions on these parameters. To obtain the expected errors of parameters obtained from a single realization of data, one can calculate the ensemble-averaged log-likelihood (to simulate e.g. multiple repetition of the measurement of model parameters). Ensemble-averaging also provides a useful test for the behaviour of the log-likelihood function



(a) **(b)**

Figure 4.4: **(a)** Recovered parameter values (crosses) as a result of minimising the log-likelihood function Eq. (4.20). For each of the four sets of parameters (denoted by squares) 100 mock catalogues were created . We use model Family II and the same family was used to fit the data. The model has 2 free parameters (κ_0 and κ_1) – see Eq. (4.22). Solid and dashed lines correspond to the expected mass-sheet degeneracy calculated using $w_i = 1/\sigma_i^2$ and $w_i = \text{const.}$ in Eq. (4.26) for the weighting scheme respectively (both almost overlap). Dot-dashed lines are given by $\kappa_0 + \kappa_1 = \text{const.}$ **(b)** Relative bias and variance for the recovered values of κ_1.

in the asymptotic limit.

The log-likelihood function depends upon the source positions, ellipticities, and redshifts, as well as the lens properties. Using the true model values (denoted by the subscript "t", used to calculate the probability distributions) the ensemble average quantity $\langle l \rangle$ is then given by (Schneider et al., 2000)

$$\langle l \rangle(\pi) = \left[\prod_{i=1}^{N_{\rm gal}} \int \mathrm{d}z\, p(z) \int \mathrm{d}^2\theta_i\, p_{\rm t}(\boldsymbol{\theta}_i) \int \mathrm{d}^2\epsilon_i\, p_{\rm t}(\epsilon_i) \right] l(\pi) \,, \tag{4.28}$$

where $p(z)$ is the redshift distribution of the background sources (which is model independent), $p_{\rm t}(\boldsymbol{\theta}_i)$ is the probability that galaxy i lies at $\boldsymbol{\theta}_i$ (in our case a constant) and $p_{\rm t}(\epsilon_i)$ is the probability for galaxy i to have a measured ellipticity ϵ_i, given the true model. The operator in the brackets is understood as

$$\left[\prod_{i=1}^{N} \int \mathrm{d}x_i p(x_i) \right] X(x_i) := \int \mathrm{d}x_1 p(x_1) \int \mathrm{d}x_2 p(x_2) \cdots \int \mathrm{d}x_N p(x_N) X(x_1, \ldots, x_N) \,. \tag{4.29}$$

This means that in order to evaluate (4.28) one would need to evaluate an $5 \times N_{\rm gal}$–dimensional integral. However, since the log-likelihood function is a sum over all contributions l_i – c.f. (4.20), the integration operators work term by term, and we in fact get a sum of $N_{\rm gal}$ equal contributions. The ensemble-average is thus given by

$$\langle l \rangle(\pi) = n_{\rm g} \int \mathrm{d}^2\theta \int \mathrm{d}z\, p_z(z) \int \mathrm{d}^2\epsilon\, p_\epsilon(\epsilon|g) l_i(\pi) \;, \tag{4.30}$$

where $n_{\rm g}$ gives the number of galaxies per unit area, and $p_\epsilon(\epsilon|g)$ is the probability distribution of lensed ellipticities calculated using lens parameters $\pi_{\rm t}$.

Asymptotically (i.e. when the number of source galaxies $N_{\rm g}$ is very large), the quantity $2\Delta l$, where

$$\Delta l = \langle l \rangle(\pi) - \langle l \rangle(\pi_{\rm t}) \;, \tag{4.31}$$

behaves as a random variable following a χ^2_M distribution, where M is the number of free parameters. In particular, for Family I, where $M = 3$, we expect 68.3%, 90%, 95.4%, and 99% points within the levels $2\,\Delta l = \{3.53, 6.25, 8.02, 11.2\}$, while for Family II ($M = 2$) the corresponding levels are $2\,\Delta l = \{2.30, 4.61, 6.17, 9.21\}$. It is interesting to test the behavior of this quantity and the accuracy of the asymptotic limit in our case. Note that we also do not expect any bias if we are indeed in asymptotic regime; namely the log-likelihood is by definition asymptotically unbiased.

As shown by Geiger & Schneider (1998), the ellipticity distribution is generally skewed even for relatively small reduced shears ($|g| \simeq 0.6$); hence, we need to evaluate all integrations of Eq. (4.30), including the one on ϵ, numerically. The integration has been carried out using the GNU Scientific library. In Fig. 4.7 we show an example of the ensemble-average log-likelihood for one set of simulated parameters for model Family II; the other sets give similar results. We simulated the data using parameters $\kappa_0 = 0.9$ and $\kappa_1 = 0.2$ without adding the redshift errors. This is the extreme case mentioned above, where the distribution of recovered parameters lies in a very narrow valley along the line where $\kappa_0 + \kappa_1 = \mathrm{const}$ (dash-dotted line) rather than along the expected mass-sheet degeneracy line (solid line and dashed line for different weighting schemes). The ensemble-average log-likelihood calculations confirm the anomalous behavior of the degeneracy.

The $2\,\Delta l = 9.21$ contour in Fig. 4.7 should enclose 99% of all points. This is not satisfied in our case; this is mainly due to the fact that we are not in the regime of the asymptotic limit (seen by the fact that the points in the plot are not distributed according to a two dimensional Gaussian distribution). On the other hand, we observe that our results are unbiased and that the data are well described by the probability distribution $p_{\epsilon^{\rm m}}(\epsilon^{\rm m} \mid g)$ given in (4.18) that we use.

4.4 Conclusions

In this chapter we consider a new method to break the mass-sheet degeneracy in weak lensing mass reconstructions using shape measurements only. A detailed analysis of this method has shown that breaking the mass-sheet degeneracy is *very difficult* even in optimal conditions; arguably, it is extremely difficult in normal observational conditions.

Our main conclusions are:

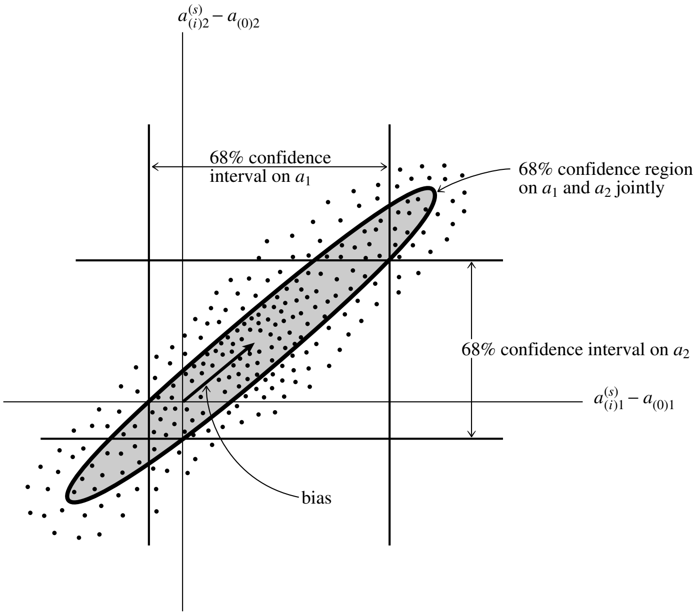


Figure 4.5: Describing a distribution in terms of confidence limits for a two dimensional space of parameters $\boldsymbol{a}$. If the points in the plot are distributed according to a Gaussian, 68% of all points should lie in the shaded confidence interval. The shape of the confidence region is ellipsoidal, which is due to the fact that we are dealing with Gaussian distribution here. (taken from Press et al., 1992)

1. The mass-sheet degeneracy can be broken by using the redshift information of the individual sources. However, this is effective for critical clusters only, i.e. for clusters that have sizable regions where multiple imaging is possible (and thus perhaps observed). The statistical lensing analysis has to be extended close to and inside the critical curves of the cluster. In the regions far outside the critical curve, where weak lensing mass reconstructions are normally performed, the lens is too weak for the mass-sheet degeneracy to be broken by using redshift and distortion information only, even when idealised conditions are employed.

2. Using simulations we find that the correlation remaining from the mass-sheet degeneracy transformation for critical lenses is well described by Eq. (4.6), $\kappa \rightarrow \kappa' \simeq \lambda\kappa + (1-\lambda)\left\langle Z(z)\right\rangle / \left\langle Z^2(z)\right\rangle$, where the moments of the cosmological weights are calculated using $w_i = 1/\sigma_i^2$ in (4.26).

3. In order to break the mass-sheet degeneracy with current data it is necessary to extend the statistical lensing analysis closer to the cluster centre and to simultaneously perform a weak and strong lensing analysis of the cluster. This will be subject of the next chapter.

In the next chapter we will describe a method which combines weak and strong lensing mass reconstruction techniques simultaneously in a non-parametric fashion. Weak and strong lensing data has been previously combined by e.g. Kneib et al.

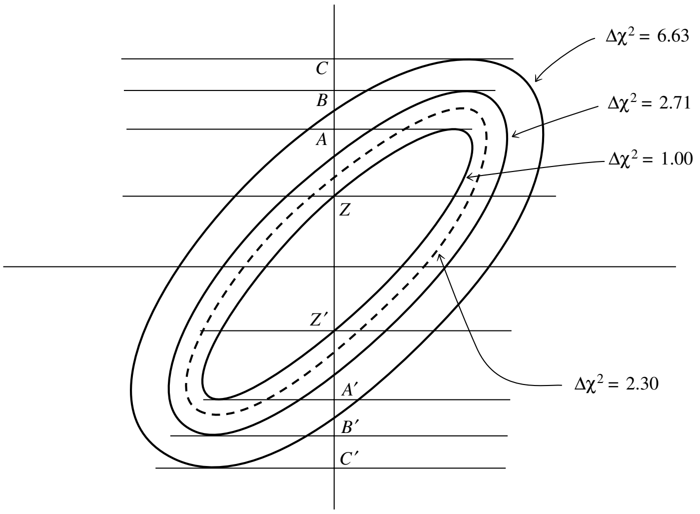


Figure 4.6: Asymptotically, the ensemble-averaged log-likelihood $2\Delta l$ behaves as a random variable following a χ^2_M distribution, where M is the number of free parameters. If we reached the asymptotic limit (having large enough number of galaxies) we expect 68.3%, 90%, and 99% points within the levels $2\,\Delta l = \{2.30, 4.61, 9.21\}$ for Family II ($M = 2$). The ellipse that contains 68.3% of normally distributed data is shown here dashed. The solid curves with $2\,\Delta l = \{1.00, 2.71, 6.63\}$ project onto one dimensional intervals AA', BB', and CC'. These intervals (not ellipses) also contain 68.3%, 90%, and 99% of distributed data and give 1-dimensional confidence intervals on y-axis parameter (taken from Press et al., 1992).

(2003); however, only the combination of weak lensing signal on scales > 500 kpc with the strong lensing signal at ~ 100 kpc was taken into account. The non-parametric statistical lensing reconstruction technique needs to be extended to within the critical regions, only then the breaking the mass-sheet degeneracy is possible in practice. Such a method does not rely on any assumption of the parametric form of the potential.

In summary, although breaking the mass-sheet degeneracy has proven to be surprisingly difficult, we have shown that it is in principle possible if one combines constraints on different scales (note that we assumed the knowledge of the cluster-mass profile; in practice one obtains the profile by using standard weak-lensing mass reconstructions). The mass-sheet degeneracy is probably the most severe limit of current weak lensing mass reconstructions, and is generally responsible for a significant fraction of the final error on the "total" mass of the cluster. Hence, breaking the mass-sheet degeneracy in practice is one of the most important challenges of weak lensing studies of clusters in the near future and is the subject of the next chapter.

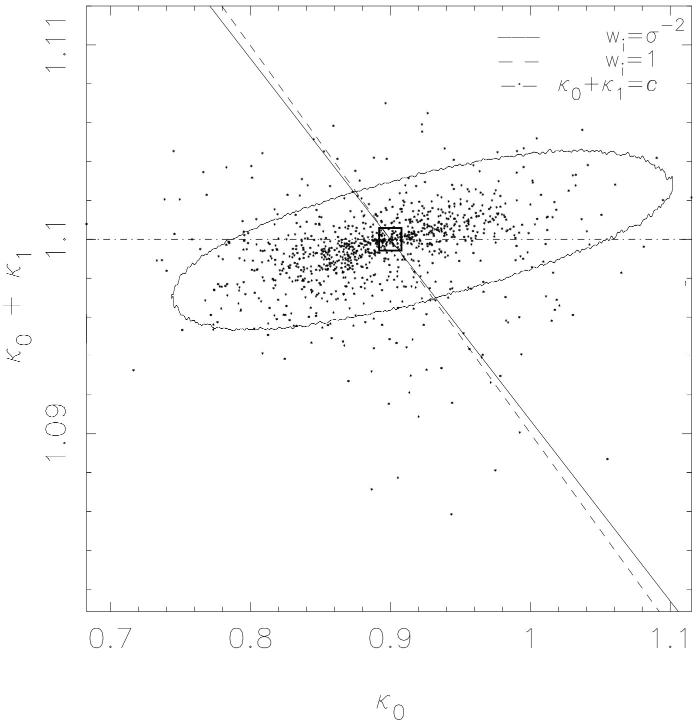


Figure 4.7: Ensemble-averaged log-likelihood for model family II with $\kappa_0 = 0.9$ and $\kappa_1 = 0.2$ (square). The elliptical contour gives the difference of ensemble-averaged log-likelihoods with a level of $2\,\Delta l = 9.21$. In contrast to previous figures, 1000 recovered model parameters are plotted as dots and $\kappa_0{+}\kappa_1$ vs. κ_1 is plotted (note the small range on y axis). The solid line gives the expected mass-sheet degeneracy line calculated with $w_i = 1/\sigma_i^2$ in Eq. (4.26) and the dashed line is calculated with $w_i =$ const. (both almost overlapping). The dash-dotted line is given by $\kappa_0 + \kappa_1 =$ const.

5

Strong and weak lensing united I: Cluster mass reconstruction using combined strong and weak lensing

In this chapter we develop a new cluster mass reconstruction technique using combined strong and weak lensing. Such a method is applicable to many known clusters that show highly distorted images, i.e. arcs. With upcoming higher resolution observations their number will increase substantially (see Fig. 1.2 for the first example of cluster being imaged with the new ACS camera).

We have shown in the previous chapter that the main limitation for a precise weak lensing mass-estimate is the mass-sheet degeneracy. The transformation of the surface mass-density $\kappa \rightarrow \kappa' = \lambda\kappa + (1-\lambda)$, where $\lambda \neq 0$ is an arbitrary constant, leaves the observable g unchanged in the absence of redshift information. We show that this degeneracy can be lifted using information on *individual* source redshifts and by extending the weak lensing reconstruction to the critical parts of the cluster. However such a reconstruction has proven to be surprisingly difficult.

In order to successfully perform the cluster mass reconstruction in practice we need to include additional information to the weak lensing data. The most natural is to use strong lensing (i.e. multiple image system) information. Since we have shown that we need clusters which have critical regions to break the mass-sheet degeneracy, it is very likely that these clusters will also show observable multiple image systems.

Only a combination of strong and weak lensing might be the key to success. Note that a single multiple-image system does not by itself break the degeneracy, since the transformation does not affect the image positions (and the same is true for the magnification ratios). However with (even a single) multiple image system with known redshift of the lens and the source we can determine the enclosed mass within the Einstein radius, giving less freedom for the weak lensing constraints. We perform a reconstruction of the gravitational potential ψ, since it locally determines both the lensing distortion (for weak lensing) as well as the deflection (for strong lensing). The method follows the idea from Bartelmann et al. (1996); Seitz et al. (1998). Its novel feature is that we directly include strong lensing information and the weak lensing reconstruction is extended to the critical parts of the cluster. We also include individual redshift information of background sources as well as of the source(s) being multiply imaged.

The results of this part of the thesis are currently in preparation to be submitted to A&A (Bradač et al., 2004b).

5.1 The cluster mass reconstruction methods

The idea of combining strong and weak lensing constraints is not new, it has been previously discussed by Abdelsalam et al. (1998); Kneib et al. (2003); Smith et al. (2004) and others. The method presented here has however some important advantages. E.g., in Abdelsalam et al. (1998) the authors reconstruct the pixelized version of the surface mass density κ. We argue, however, that using the potential ψ is favourable, since κ, γ, and $\boldsymbol{\alpha}$ locally depend upon the potential ψ – c.f. (2.11), (2.13), (2.9) – and all can be readily obtained. γ and $\boldsymbol{\alpha}$ on the other hand are a non-local quantities of κ and can be calculated from κ only if we have an infinite field. If a finite field is used, one usually employs Fourier analysis and in this case γ in fact corresponds to original κ and all its periodic continuations.

Further, even though not easy to implement, we decided to keep the parametrisation of the mass-distribution as general as possible. In Kneib et al. (2003) and Smith et al. (2004), on the other hand, the strong and weak lensing constraints were compared in Bayesian approach in the form of parametric modelling (c.f. Sect. 4.1.5). In addition, the weak lensing constraints were not used to the very centre of the cluster and individual redshift information was not included.

5.1.1 The outline of the method

Following the idea of Bartelmann et al. (1996) we parametrise the cluster mass-distribution by the values of a potential ψ_k on a regular grid. We determine κ, γ, and $\boldsymbol{\alpha}$ on a grid by replacing derivatives in (2.11), (2.13), and (2.9) with finite differencing method. Finally we evaluate $\kappa(\boldsymbol{\theta})$, $\gamma(\boldsymbol{\theta})$, and $\boldsymbol{\alpha}(\boldsymbol{\theta})$ at any position in the image plane $\boldsymbol{\theta}$ using bilinear interpolation. The reduced shear $g(\boldsymbol{\theta}_i, z_i)$ at each galaxy position $\boldsymbol{\theta}_i$ and redshift z_i, as well as the deflection angle $\boldsymbol{\alpha}(\boldsymbol{\theta}_m, z_\mathrm{s})$ at the positions $\boldsymbol{\theta}_m$ and redshift z_s corresponding to the multiple-image system can thus be readily evaluated.

We define a χ^2-function as follows

$$\chi^2(\psi_k) = \chi^2_\epsilon(\psi_k) + \chi^2_\mathrm{M}(\psi_k) + \eta R(\psi_k)\ . \tag{5.1}$$

$\chi^2_\epsilon(\psi_k)$ contains information from statistical weak lensing, whereas in $\chi^2_\mathrm{M}(\psi_k)$ we include the multiple imaging properties. $R(\psi_k)$ is a regularisation term multiplied by the regularisation parameter η. The regularisation is a function of the potential and disfavours any small-scale fluctuations in the potential.

The method described in Bartelmann et al. (1996) minimises the $\chi^2(\psi_k)$ w.r.t. the parameters ψ_k using standard numerical routines. However, if the number of grid points is large such an approach is not practical. We intend to solve the following system of equations

$$\frac{\partial \chi^2(\psi_k)}{\partial \psi_k} = 0 \, . \tag{5.2}$$

This is in general a non-linear system of equations (as will be shown in Sect. 5.1.3). We linearise the equations in terms of ψ_k and keep the terms not linear in ψ_k fixed at each iteration step. Before each iteration these terms are "updated" with the solution from the previous iteration (or from initial conditions in the first step) and the process is repeated until convergence. On the first sight it seems that we turned the problem of $N_{\rm dim}$-dimensional minimisation to a more difficult problem in general, namely solving a linear system of equations with $N_{\rm dim}$ unknowns. Inverting the resulting matrix of $\sim N_{\rm grid}^2$ elements is even for grids with small number of grid cells difficult in general. However, as it turns out, the resulting matrix is sparse and therefore is solving the resulting system computationally inexpensive.

The reconstruction is performed in a two-level iteration process, outlined in Fig. 5.1. We will refer to the iteration process mentioned above for solving the linear system of equations as first-level, where steps n_1 are repeated until convergence of κ. The second-level iteration is performed for the purpose of regularisation (as described in detail in Sect. 5.1.3). In order to penalise small-scale fluctuations in the potential, we start the reconstruction with a coarse grid (large cell size). Then for each n_2 step we increase the number of grid points in the field and compare the new reconstructed $\kappa^{(n_2)}$ with the one from the previous iteration $\kappa^{(n_2-1)}$ (or with initial input value $\kappa^{(0)}$ for $n_2 = 0$). The second-level iterations are performed until the final grid size is reached and the convergence is achieved.

5.1.2 The technical aspects

In this section we will briefly describe some technical aspects of how we calculate the lensing quantities κ, γ, and $\boldsymbol{\alpha}$ at an arbitrary position within the field from the potential ψ_k on a grid.

We consider the potential ψ_k on an equidistant grid of $(N_{\rm x}+2) \times (N_{\rm y}+2)$ points. On the inner $N_{\rm x} \times N_{\rm y}$ grid one can calculate κ, γ and $\boldsymbol{\alpha}$ using differencing methods; the extention by one row and column at each side is needed to be able to perform the finite differencing at each inner grid point. We use the finite differencing method with 9 grid points to calculate κ, 5 points for γ, and 4 points for $\boldsymbol{\alpha}$ (see Abramowitz & Stegun, 1972). The coefficients used for κ and γ are given in Fig. 5.2, the case of $\boldsymbol{\alpha}$ is discussed in the Sect. 5.2.2. To evaluate $\kappa(\boldsymbol{\theta})$, $\gamma(\boldsymbol{\theta})$, and $\boldsymbol{\alpha}(\boldsymbol{\theta})$ at a position $\boldsymbol{\theta}$, bilinear interpolation is used. E.g. the value of $\kappa(\boldsymbol{\theta})$ is expressed in terms of values of κ on four nearest-neighbouring grid points using bilinear interpolation. Since for each of these four values we need 9 (partly common) points with ψ_k, we express $\kappa(\boldsymbol{\theta})$ in terms of ψ_k at 16 grid points. The coefficients in this expression also depend upon the relative distance of $\boldsymbol{\theta}$ to the nearest grid point (needed to do bilinear interpolation).

Note, however, that the dimensionality of the problem is not $(N_{\rm x}+2) \times (N_{\rm y}+2)$. Namely, the transformation $\psi(\boldsymbol{\theta}) \rightarrow \psi(\boldsymbol{\theta}) + \psi_0 + \boldsymbol{a} \cdot \boldsymbol{\theta}$ leaves κ and γ invariant,

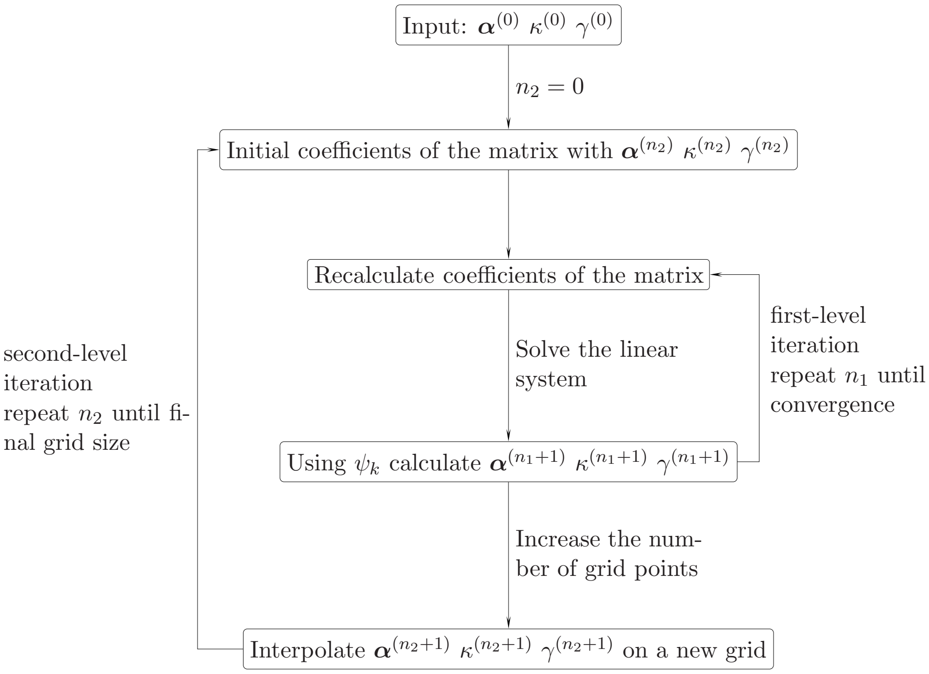


Figure 5.1: The outline of the two-level iteration process.

and therefore the potential needs to be fixed at three points (see Seitz et al., 1998; Bartelmann et al., 1996). In addition, even though the transformation $\psi(\boldsymbol{\theta}) \rightarrow \psi(\boldsymbol{\theta}) + \boldsymbol{a} \cdot \boldsymbol{\theta}$ affects the deflection angle $\boldsymbol{\alpha}$, it only causes a translation of the source plane, which is not an observable. Therefore, even in the presence of strong lensing, three points of the potential need to be held fixed.

Our ultimate aim is to make a reliable estimate for the cluster mass, therefore the mass-sheet degeneracy needs to be lifted. The degeneracy transformation of the potential is given by (4.1), thus $\psi \rightarrow \psi' = 0.5(1-\lambda)\boldsymbol{\theta}^2 + \lambda\psi$. The degeneracy can, however, be broken if the reconstruction is extended to the regime where the lens becomes critical and by using *individual redshifts* of background sources (see Chapter 4). Hence in contrast to Seitz et al. (1998) the potential ψ_k, is *not* held fixed at an additional, fourth point. The dimensionality of the problem is thus $N_{\text{dim}} = (N_{\text{x}}+2)(N_{\text{y}}+2) - 3$.

5.1.3 The χ^2-function

In this section we will describe each contribution to the χ^2-function in turn, starting with the statistical weak lensing.

For N_{g} galaxies with measured ellipticities ϵ_i the χ^2_ϵ was defined in equation (4.12), for completeness we simply rewrite the definitions here.

$$\chi^2_\epsilon(\psi_k) = \sum_{i=1}^{N_{\text{g}}} \frac{|\epsilon_i - \langle\epsilon\rangle|^2}{\sigma_i^2} \, , \tag{5.3}$$

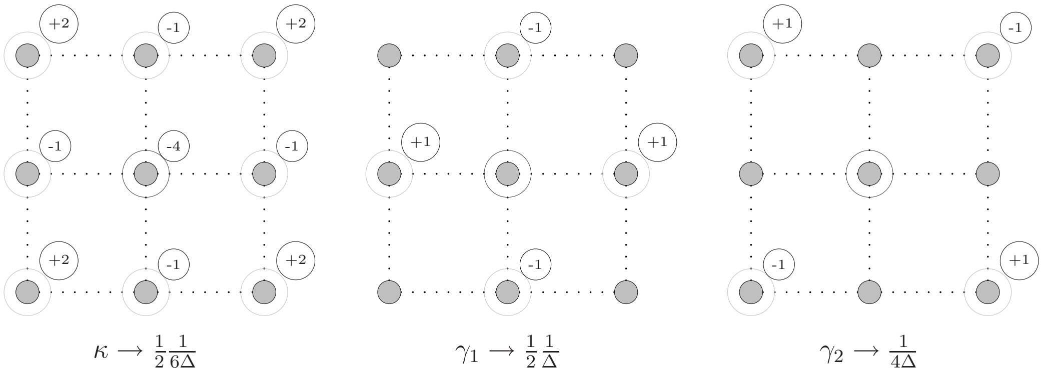


Figure 5.2: The finite differencing coefficients for κ (left), γ_1 (middle) and γ_2 (right). E.g. for κ we use a formula including 9 points, the multiplicative factor is given on the bottom, the individual coefficients in the circle. Thus for the middle point $(0,0)$ we get $\kappa(0,0) = \frac{1}{2}\frac{1}{6\Delta}\big(2\,[\psi(-1,1)+\psi(1,1)+\psi(-1,-1)+\psi(1,-1)] - [\psi(0,1)+\psi(-1,0)+\psi(1,0)+\psi(0,-1)] - 4\psi(0,0)\big)$.

where

$$\sigma_i^2 = \left(1 - \left|\langle\epsilon\rangle\right|^2\right)^2 \sigma_{\epsilon^{\rm s}}^2 + \sigma_{\rm err}^2 \ . \tag{5.4}$$

is (approximately) the variance of the measured ellipticities. The expectation value of measured ellipticities $\langle\epsilon\rangle$ is given in (2.31). In Chapter 4 we also argue that χ^2_ϵ can give biased results for lenses for which many galaxies have $|g| \simeq 1$. It would be better to work with a log-likelihood function with probability distribution that properly describes the distribution of observed ellipticities. Unfortunately such an approach is not viable here (as will become obvious later on). However, in general clusters do not have a large fraction of galaxies with $|g| \simeq 1$ and we showed that for these lenses the χ^2-minimisation is sufficient.

One of the major strengths of this statistical weak lensing reconstruction technique is the possibility to simultaneously include constraints from multiple images to the weak lensing data in a relatively straightforward manner. The simplest approach to strong lensing is to perform the so-called "source plane" modelling; i.e. to minimise the projected source position difference. Consider a multiple image system with the source at a redshift $z_{\rm s}$ and with $N_{\rm M}$ images located at $\boldsymbol{\theta}_m$. The corresponding χ^2-function is given by

$$\chi^2_{\rm M} = \sum_{m=1}^{N_{\rm M}} \frac{\left|\boldsymbol{\theta}_{\rm m} - Z(z_{\rm s})\boldsymbol{\alpha}(\boldsymbol{\theta}_m) - \boldsymbol{\beta}_{\rm s}\right|^2}{\sigma^2_{{\rm s},m}} \ , \tag{5.5}$$

where $\sigma_{{\rm s},m}$ are the errors on image positions, projected to the source plane. $\boldsymbol{\beta}_{\rm s}$ is the average source position and we calculate it using the deflection angle information from previous iteration $n_1 - 1$. For simplicity, however, we do not perform a projection of the error ellipse from the image plane onto the source plane. Instead, we keep $\sigma_{{\rm s},m}$ constant throughout the reconstruction. Therewith we avoided the numerical problem of having the $\chi^2_{\rm M}$ function diverging when one of the images would lie close to the critical curve. We are aware of the fact that the approach we use is not optimal (see e.g. Kochanek, 2004). If only multiple imaging is used, the resulting best-fit model would be biased towards high magnification factors, since errors on the source plane are magnified when projected back to the image plane (this information we do not use). In our case, however, the model also needs to take into account the constrains from statistical (weak) lensing and therefore the high magnification models will in fact be discarded. In addition, if e.g. for an elongated image error on its positions is considered isotropic, this would in general mean a highly anisotropic error in the source plane. We therefore argue that considering errors in the source plane to be isotropic is in fact a better approximation, since sources are on average more circular than their multiple images.

Regularisation needs to be employed, since the minimisation of χ^2_ϵ would otherwise lead to a potential which closely follows the noise pattern of the data. Even without the measurement errors, the intrinsic ellipticities would still produce pronounced small-scale noise peaks in the final reconstruction. In addition, the method presented here has an intrinsic invariance if no multiple imaging information is used and the weak lensing approximation $g \simeq \gamma$ applies. Namely, we can alternately add/subtract a constant a along diagonals of the potential (as sketched in Fig. 5.3). This transformation would on the one hand not affect γ, but on the other it would cause a similar change (with a constant $2a/3$) in κ – compare with Fig. 5.2. Thus in the $|g| \ll 1$ regime, where $\langle\epsilon\rangle = g \simeq \gamma$ these stripes will show up in the resulting κ map. This problem can, however, be very efficiently cured with regularisation.

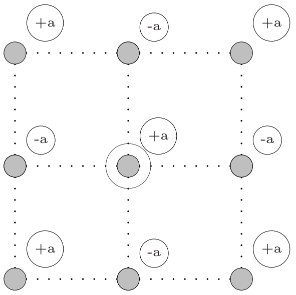


Figure 5.3: The intrinsic invariance of the method. If we alternately add/subtract a constant a along the diagonals the shear γ does not change (cf. Fig. 5.2), but κ changes in the similar way with a constant now being $2a/3$.

Since we want to measure the cluster mass, the regularisation should not influence breaking of the mass-sheet degeneracy. For example, one of the possibilities proposed in Seitz et al. (1998) for regularisation function was $R = \sum_{\rm grid} |\nabla\kappa|^2$. However, as the authors mentioned such regularisation would tend to flatten the profile and therefore affect the mass-sheet degeneracy breaking. Their maximum entropy (ME) regularisation is more suitable for this purpose, however it is very cumbersome to express its derivative in linear terms of ψ_k. Motivated by the success of moving prior in ME regularisation, we choose a very simple prescription for the regularisation function. We start off by a relatively coarse grid. Namely if the number of grid points $N_{\rm dim}$ is much smaller than the number of galaxies, the resulting reconstruction will not be able to follow the noise pattern. In each second-level iteration step we increase the grid size and compare the resulting κ map $\kappa^{(n_2)}$ with that from the previous iteration $\kappa^{(n_2-1)}$ interpolated on a finer grid, thus

$$R = \sum_{i,j=1}^{N_{\rm x},N_{\rm y}} \left(\kappa_{ij}^{(n_2)} - \kappa_{ij}^{(n_2-1)}\right)^2 . \tag{5.6}$$

For the case of $n_2 = 0$ we use an initial guess for κ which can in practice be obtained from strong lensing, direct finite-field reconstruction, parametrised (non-general) model fitting to weak lensing data, or also simply set to a constant. This method is very easy to implement and is in addition very efficient to remove the stripes (mentioned above) in the final reconstruction. If enough n_2 iteration steps are used it also does not substantially affect the mass-sheet degeneracy, since the information of initial $\kappa^{(0)}$ is lost.

Finally a word on the regularisation constant η. This parameter should in theory ensure $\chi^2/N_{\rm dof} \sim 1$, however it is in practice difficult to determine (in the critical lensing regime). As outlined in Geiger & Schneider (1998) the errors on measured ellipticities are not normally distributed and therefore the actual value of χ^2 at its minimum has no particular meaning. In practice, however, setting η such that the resulting $\chi^2/N_{\rm dof} \sim 1$ is valid is a good guess for this parameter. In addition, one adjusts η low enough for the method to have enough freedom to adapt to the information in the data and large enough for not allowing the solutions that follow the noise pattern. As a rule of thumb it is usually better to set η high and increase the number of iteration, allowing κ to change only slowly. Since the reconstruction is done in a two-level iteration and in addition multiple-image information is included, the method can successfully adapt to the data and the results are not very sensitive to the precise value η.

5.1.4 Initial conditions

The initial conditions are needed for the regularisation method as given by (5.6). As we will show later, this point is not very important if multiple imaging is included directly in χ^2. In that case one can simply use $\kappa^{(0)} =$ const.

However, the initial conditions can be very helpful to break the internal degeneracy, i.e. to distinguish galaxies that have $|g| \leq 1$ from those with $|g| > 1$ in the first step, and thus allow for a faster convergence. Breaking this degeneracy is desirable, since otherwise the method has difficulties in "climbing" over the $|g| = 1$ region (especially if *no* multiple-image systems are included). Since we use individual redshifts of background galaxies we do not have well defined critical curves (i.e. positions in the source plane where $|g| = 1$), since their position depends on the source redshift. In spite of this fact, the transition still poses a difficulty.

The possible presence of multiple image systems in a cluster can be used to provide the initial conditions. In particular if the redshift and identification of some of the systems is not completely secure and one would not like to include it in the χ^2_{M} directly, the information can still be added for to calculate $\kappa^{(0)}$. Throughout this work we will use multiple imaging directly in the χ^2-function and in certain cases also for the initial values of κ, γ, and $\boldsymbol{\alpha}$ (used to calculate the initial coefficients of the linear system).

5.2 The linear problem for ψ_k

In this section we present details of the method outlined in Sect. 5.1. They are highly technical and the fast reader might skip this part and go directly to the Sect. 5.3.

We aim to solve the equation

$$\frac{\partial \chi^2_\epsilon(\psi_k)}{\partial \psi_k} + \eta \frac{\partial R(\psi_k)}{\partial \psi_k} + \frac{\partial \chi^2_{\mathrm{M}}(\psi_k)}{\partial \psi_k} = 0 \; . \tag{5.7}$$

As outlined in Sect. 5.1 this is in general a non-linear system of equations. We try to solve it in an iterative way by linearising the equation in terms of ψ_k and keeping the non-linear terms fixed at each iteration step. The resulting system will be written in the form

$$\mathcal{B}_{jk}\psi_k = V_j \; , \tag{5.8}$$

where the matrix elements $\mathcal{B}_{jk}$ and vector V_j will contain the contributions from non-linear part. In the following sections we will describe contributions to (5.7) in turn.

5.2.1 The weak lensing analysis

The χ^2_ϵ for the weak lensing case is given in (5.3). From now on we will consider in detail only the $|g| \leq 1$ case; for $|g| > 1$ the calculations are done in the same fashion. First we plug into (5.3) the expectation value of observed ellipticities (i.e. the reduced shear g) giving

$$\chi^2_\epsilon(\psi_k) = \sum_{i=1}^{N_{\mathrm{g}}} \frac{\left|\epsilon - \frac{Z\gamma}{1-Z\kappa}\right|^2}{\sigma^2} = \sum_{i=1}^{N_{\mathrm{g}}} \frac{\left|\epsilon - Z\epsilon\kappa - Z\gamma\right|^2}{\left(1 - Z\kappa\right)^2 \sigma^2} \; , \tag{5.9}$$

where κ, γ, and σ depend on $\boldsymbol{\theta}_i$ only and Z depends on the redshift of the i-th source. Note that for simplicity we omit writing the index i to ϵ, κ, γ, σ and Z, although they are different for every galaxy and thus depend on i. As we described in Sect. 5.1.2 using finite differencing and bilinear interpolation we can write κ and γ at each galaxy position as a linear combination of ψ_k. This is expressed in the following matrix notation

$$\gamma_1(\boldsymbol{\theta}_i) = \mathcal{G}^{(1)}_{ik}\psi_k \; , \quad \gamma_2(\boldsymbol{\theta}_i) = \mathcal{G}^{(2)}_{ik}\psi_k \; , \quad \kappa(\boldsymbol{\theta}_i) = \mathcal{K}_{ik}\psi_k \; , \tag{5.10}$$

where the matrices $\mathcal{G}^{(1)}_{ik}$, $\mathcal{G}^{(2)}_{ik}$, and $\mathcal{K}_{ik}$ are composed of numerical factors described in Sect. 5.1.2. Now we consider the denominator of (5.9) $\hat{\sigma}^2_{\leq} = \left(1 - Z\kappa\right)^2 \sigma^2$ fixed at each iteration step (the subscript $\leq$ denotes the $|g| \leq 1$ case) and differentiate the following term of (5.9)

$$\begin{aligned}
&\frac{1}{2}\frac{\partial}{\partial\psi_k}\frac{|\epsilon - Z\epsilon\kappa - Z\gamma|^2}{\hat{\sigma}^2_{\leq}} = \\
-&\frac{Z}{\hat{\sigma}^2_{\leq}}\left[(\epsilon_1 - Z\epsilon_1\kappa - Z\gamma_1)\left(\epsilon_1\frac{\partial\kappa}{\partial\psi_k} + \frac{\partial\gamma_1}{\partial\psi_k}\right)\right. \\
+&\left.(\epsilon_2 - Z\epsilon_2\kappa - Z\gamma_2)\left(\epsilon_2\frac{\partial\kappa}{\partial\psi_k} + \frac{\partial\gamma_2}{\partial\psi_k}\right)\right] = \\
&\frac{Z^2}{\hat{\sigma}^2_{\leq}}\left[\mathcal{G}^{(1)}_{ij}\mathcal{G}^{(1)}_{ik} + \mathcal{G}^{(2)}_{ij}\mathcal{G}^{(2)}_{ik} + \epsilon_1\left(\mathcal{G}^{(1)}_{ij}\mathcal{K}_{ik} + \mathcal{K}_{ij}\mathcal{G}^{(1)}_{ik}\right)\right. \\
+&\left.\epsilon_2\left(\mathcal{G}^{(2)}_{ij}\mathcal{K}_{ik} + \mathcal{K}_{ij}\mathcal{G}^{(2)}_{ik}\right) + \left(\epsilon_1^2 + \epsilon_2^2\right)\mathcal{K}_{ij}\mathcal{K}_{ik}\right]\psi_k \\
-&\frac{Z^2}{\hat{\sigma}^2_{\leq}}\left[\epsilon_1\mathcal{G}^{(1)}_{ij} + \epsilon_2\mathcal{G}^{(2)}_{ij} + \left(\epsilon_1^2 + \epsilon_2^2\right)\mathcal{K}_{ij}\right]
\end{aligned} \tag{5.11}$$

where ϵ_1 and ϵ_2 are the two components of measured ellipticity of galaxy i (again omitting the index and dividing (5.9) for simplicity). We sum over all galaxies used for the weak lensing analysis and obtain a linear problem for ψ_k at each iteration step. The same approach can be made for $|g| > 1$ case, where $\hat{\sigma}^2_{>}$ kept constant is given by $\hat{\sigma}^2_{>} = Z_i^2\,|\gamma|^2\,\sigma_i^2$.

5.2.2 The strong-lensing term

Following the prescription from the previous section we now write the deflection angle in a matrix form

$$\alpha_1(\boldsymbol{\theta}_m) = \mathcal{D}^{(1)}_{ik}\psi_k\ , \quad \alpha_2(\boldsymbol{\theta}_m) = \mathcal{D}^{(2)}_{ik}\psi_k\ . \tag{5.12}$$

Both matrices give the finite differencing form for the gradient of the potential, in particular we use the central differencing formula, i.e. $\alpha_1(0,0) = \frac{1}{2\Delta}\left(\psi(1,0) - \psi(-1,0)\right)$ and $\alpha_2(0,0) = \frac{1}{2\Delta}\left(\psi(0,1) - \psi(0,-1)\right)$.

The χ^2 contribution to strong lensing is given in (5.5). The source position $\boldsymbol{\beta}_{\rm s}$ is kept constant at every iteration step, and is evaluated using the deflection angle information $\boldsymbol{\alpha}^{(n_1-1)}$ from previous iteration

$$\boldsymbol{\beta}_{\rm s} = \frac{1}{N_{\rm M}}\sum_{m=1}^{N_{\rm M}}\left(\boldsymbol{\theta}_{\rm m} - Z(z_{\rm s})\boldsymbol{\alpha}^{(n_1-1)}(\boldsymbol{\theta}_m)\right)\ . \tag{5.13}$$

We differentiate the following term in $\chi^2_{\rm M}$ for the x_1-coordinate

$$\frac{1}{2}\frac{\partial}{\partial\psi_k}\frac{\left(\theta_{\rm m,1} - Z(z_{\rm s})\alpha_1(\boldsymbol{\theta}_m) - \beta_{\rm s,1}\right)^2}{\sigma^2_{{\rm s},m(1)}} = -\frac{\left(\theta_{m,1} - \beta_{\rm s,1}\right)\mathcal{D}^{(1)}_{ij} - Z(z_{\rm s})\mathcal{D}^{(1)}_{ij}\mathcal{D}^{(1)}_{ik}\psi_k}{\sigma^2_{{\rm s},m(1)}}\ .$$

The expression for x_2-coordinate is obtained by exchanging $1 \leftrightarrow 2$. After summation of both terms over all images m we get a set of equations which is linear in ψ_k and can be readily included in (5.8).

5.2.3 The final result

Describing a way to linearise the contributions of weak and strong lensing, we can write the coefficients in the equation (5.8). Note that the contribution of the regularisation term (with χ^2-contribution given in (5.6)) is already linear in ψ_k and therefore the full matrix $\mathcal{B}_{jk}$ from is given in the form

$$
\begin{aligned}
\mathcal{B}_{jk} \;=\; & \sum_{i=1}^{N_{\mathrm{gal}}} \Big[a_{11}(i)\mathcal{G}^{(1)}_{ij}\mathcal{G}^{(1)}_{ik} + a_{22}(i)\mathcal{G}^{(2)}_{ij}\mathcal{G}^{(2)}_{ik} + a_{13}(i)\left(\mathcal{G}^{(1)}_{ij}\mathcal{K}_{ik} + \mathcal{G}^{(1)}_{ik}\mathcal{K}_{ij}\right) \\
& \qquad + a_{23}(i)\left(\mathcal{G}^{(2)}_{ij}\mathcal{K}_{ik} + \mathcal{G}^{(2)}_{ik}\mathcal{K}_{ij}\right) + a_{33}(i)\left(\mathcal{K}_{ij}\mathcal{K}_{ik}\right)\Big] \\
+ \; & \eta \sum_{g} \mathcal{K}_{gj}\mathcal{K}_{gk} + \sum_{m=1}^{N_{\mathrm{M}}} b_{11}(m)\mathcal{D}^{(1)}_{mj}\mathcal{D}^{(1)}_{mk} + b_{22}(m)\mathcal{D}^{(2)}_{mj}\mathcal{D}^{(2)}_{mk} \;,
\end{aligned}
$$

where indices i, g and m denote summation over all galaxies with ellipticity measurement, all grid points, and all images in the multiple imaged system, respectively. The V_j vector carries the information of all constant terms in (5.7)

$$
\begin{aligned}
V_j \;=\; & \sum_{i=1}^{N_{\mathrm{gal}}} a_1(i)\mathcal{G}^{(1)}_{ij} + a_2(i)\,\mathcal{G}^{(2)}_{ij} + a_3(i)\,\mathcal{K}_{ij} + \\
& \eta \;\sum_{g} \kappa^{(n_2-1)}\mathcal{K}_{gj} + \sum_{m=1}^{N_{\mathrm{M}}} b_1(m)\mathcal{D}^{(1)}_{mj} + b_2(m)\,\mathcal{D}^{(2)}_{mj} \;.
\end{aligned}
$$

The coefficients a now differ depending whether we are in the $|g| \leq 1$ or $|g| > 1$ regime. For $|g| \leq 1$ the coefficients are given by

$$
\begin{aligned}
a_{11}(i) &= a_{22}(i) = \frac{Z^2}{\hat{\sigma}^2_{\leq}} \;; \quad a_{13}(i) = \frac{Z^2}{\hat{\sigma}^2_{\leq}}\epsilon_1 \;; \\
a_{23}(i) &= \frac{Z^2}{\hat{\sigma}^2_{\leq}}\epsilon_2 \;; \quad a_{33}(i) = \frac{Z^2}{\hat{\sigma}^2_{\leq}}\,|\epsilon|^2 \;; \\
a_1(i) &= \frac{Z}{\hat{\sigma}^2_{\leq}}\epsilon_1 \;; \; a_2(i) = \frac{Z}{\hat{\sigma}^2_{\leq}}\epsilon_2 \;; \; a_3(i) = \frac{Z}{\hat{\sigma}^2_{\leq}}\,|\epsilon|^2
\end{aligned}
\tag{5.14}
$$

Similarly for $|g| > 1$

$$
\begin{aligned}
a_{11}(i) &= a_{22}(i) = \frac{Z^2}{\hat{\sigma}^2_{>}}\,|\epsilon|^2 \;; \quad a_{13}(i) = \frac{Z^2}{\hat{\sigma}^2_{>}}\epsilon_1 \;; \\
a_{23}(i) &= \frac{Z^2}{\hat{\sigma}^2_{>}}\epsilon_2 \;; \quad a_{33}(i) = \frac{Z^2}{\hat{\sigma}^2_{>}} \;; \\
a_1(i) &= \frac{Z}{\hat{\sigma}^2_{>}}\epsilon_1 \;; \; a_2(i) = \frac{Z}{\hat{\sigma}^2_{>}}\epsilon_2 \;; \; a_3(i) = \frac{Z}{\hat{\sigma}^2_{>}} \;.
\end{aligned}
\tag{5.15}
$$

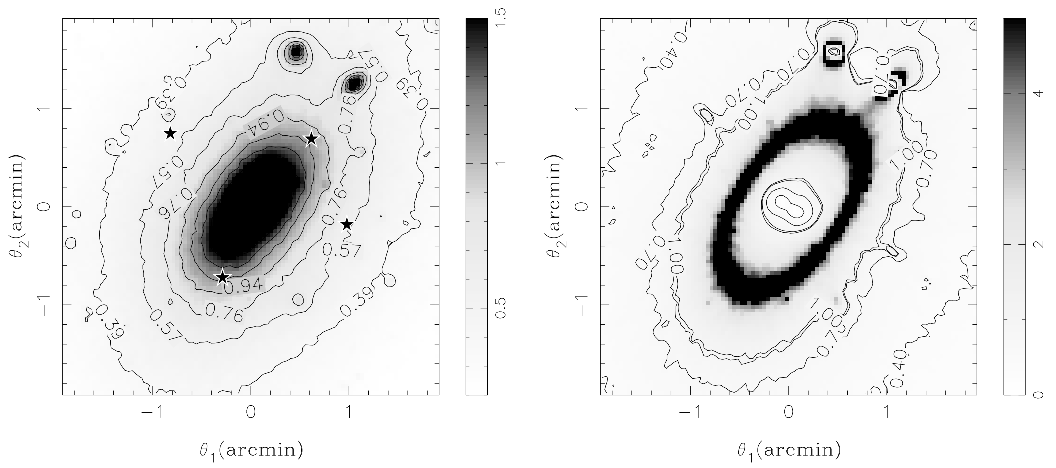
The coefficients b are the coefficients carrying the information about the multiple imaged system:

$$\begin{aligned} b_{11}(m) &= \frac{Z^2(z_\mathrm{s})}{\sigma^2_{\mathrm{s},m(1)}} \ ; \quad b_{22}(m) = \frac{Z^2(z_\mathrm{s})}{\sigma^2_{\mathrm{s},m(2)}} \ ; \\ b_1(m) &= \frac{Z(z_\mathrm{s})\,(\theta_{m,1}-\beta_{\mathrm{s},1})}{\sigma^2_{\mathrm{s},m(1)}} \\ b_2(m) &= \frac{Z(z_\mathrm{s})\,(\theta_{m,2}-\beta_{\mathrm{s},2})}{\sigma^2_{\mathrm{s},m(2)}} \ ; \end{aligned} \tag{5.16}$$

where $\theta_{m,1}$ and $\theta_{m,2}$ are booth coordinates of image m and $\sigma_{\mathrm{s},m(1)}$ and $\sigma_{\mathrm{s},m(2)}$ are the corresponding measurement errors projected to the source plane (the latter are in our case set equal). z_s is the redshift of the source being multiply imaged.

5.3 Cluster mass reconstruction from simulated data

5.3.1 Mock catalogues



(a) (b)

Figure 5.4: The gravitational lensing properties of a simulated cluster used for generating mock catalogues for statistical weak lensing and for the multiple image system. **a)** The surface mass density κ, **b)** the absolute value of the reduced shear $|g|$ for a source at $z_\mathrm{s} \rightarrow \infty$ are given in gray-scale and contours. The stars in **a)** denote the image positions of a four-image system at $z_\mathrm{s} = 1.76$ we use for the reconstruction.

To generate the mock catalogues we use a cluster from the high-resolution N-body simulation by Springel et al. (2001). The cluster we use is taken from the S4 simulation (for details see the aforementioned paper) and was simulated in the framework of the ΛCDM cosmology with density parameters $\Omega_\mathrm{m} = 0.3$ and $\Omega_\Lambda = 0.7$, shape parameter $\Gamma = 0.21$ and the normalisation of the power spectrum $\sigma_8 = 0.9$ and Hubble constant

$H_0 = 70\text{km s}^{-1}\,\text{Mpc}^{-1}$. The cluster simulation consists of almost 20 million particles, each with a mass of $4.68 \times 10^7\ M_\odot$ and a gravitational softening length of $0.7\,h^{-1}$ kpc. Due to the high mass resolution, the surface mass density κ-map can be obtained by directly projecting the particles (in our case along the z-axis) onto a 1024^2 grid (of a side length 4 Mpc) using the NGP (nearest gridpoint) assignment.

In what follows we have tried to generate the weak and strong lensing data such as to resemble as close as possible the data on cluster RX J1347−1145 we will use in Chapter 6. The surface mass density of the cluster was therefore scaled such as to have a sizeable region where multiple imaging is possible within 3.8×3.8 arcmin^2 field for sources at redshifts $z \gtrsim 1$ and a cluster at $z_\text{d} = 0.4$. The Einstein radius for a fiducial source at $z \to \infty$ is roughly $\theta_\text{E} \sim 1'$, giving a mass within this radius of $\sim 5 \times 10^{14} M_\odot$. The resulting κ map we use can be seen in Fig. 5.4a.

The lensing properties are calculated as described in detail in Chapter 4. The Poisson equation for the lens potential ψ – c.f. Eq. (2.11) – is solved on the grid in Fourier space with a DFT (Discrete Fourier Transformation) method using the FFTW library written by Frigo & Johnson (1998). The two components of the shear $\gamma_{1,2}$ and the deflection angle $\boldsymbol{\alpha}$ are obtained by finite differencing methods applied to the potential ψ. These data are then used to generate the weak lensing catalogues as well as the multiple image systems. The absolute value of the reduced shear (for a source with $z \to \infty$, thus $Z(z) = 1$) is given in Fig. 5.4b.

The weak lensing data are obtained by placing N_g galaxies on a 3.8×3.8 arcmin^2 field. We have simulated two different catalogues, one with $N_\text{g} = 145$ galaxies with positions corresponding to those from R-band weak lensing data of the cluster RX J1347−1145 and one with $N_\text{g} = 210$ galaxies corresponding to the I-band data used in the following Chapter. In this way we simulate the effects of “holes”, resulting from cluster obscuration and bright stars in the field.

The intrinsic ellipticities ϵ_s are drawn from a Gaussian distribution, each component is characterised by $\sigma = \sigma_{\epsilon^\text{s}} = 0.2$. We use the same redshifts as those measured in the R and I-band data respectively, both catalogues have average redshifts for background sources of $\langle z \rangle = 1.3$. The corresponding cosmological weights were evaluated assuming the ΛCDM cosmology (the same parameters were used as for the cluster simulations).

The lensed ellipticities are obtained using (2.25) and interpolating the quantities κ, and γ at the galaxy position using bilinear interpolation. The measurement errors ϵ^err on the observed ellipticities were drawn from a Gaussian distribution with $\sigma = \sigma_\text{err} = 0.1$ (each component) and added to the lensed ellipticities. We considered also the measurement errors on the redshifts of the galaxies to simulate the use of photometric redshifts. These have $\sigma_\text{zerr} = 0.06\,(1 + z_i)$ (see Benítez, 2000); in adding the errors we ensured that the resulting redshifts are always positive.

5.3.2 Multiple imaging

To obtain a four-image system from the simulation we used the method described in detail in Chapter. 3. We use the MNEWT routine from Press et al. (1992) to solve the lens equation for a given source position inside the asteroid caustic. The source was assumed at the redshift of $z_\mathrm{s} = 1.76$. We interpolate the deflection angle between the grid points using bilinear interpolation. Once we have the image positions, their magnifications are calculated and the four brightest images are chosen.

The errors on image positions can be conservatively estimated (for the data we use latter on) to $\sim 0\farcs3$. Since we need errors in the source plane, we set them by a factor of five smaller, $\sigma_{\mathrm{s},m} = 0\farcs06$ for each coordinate. The errors are assumed to be the same in both coordinates (see discussion in Sect. 5.1.3).

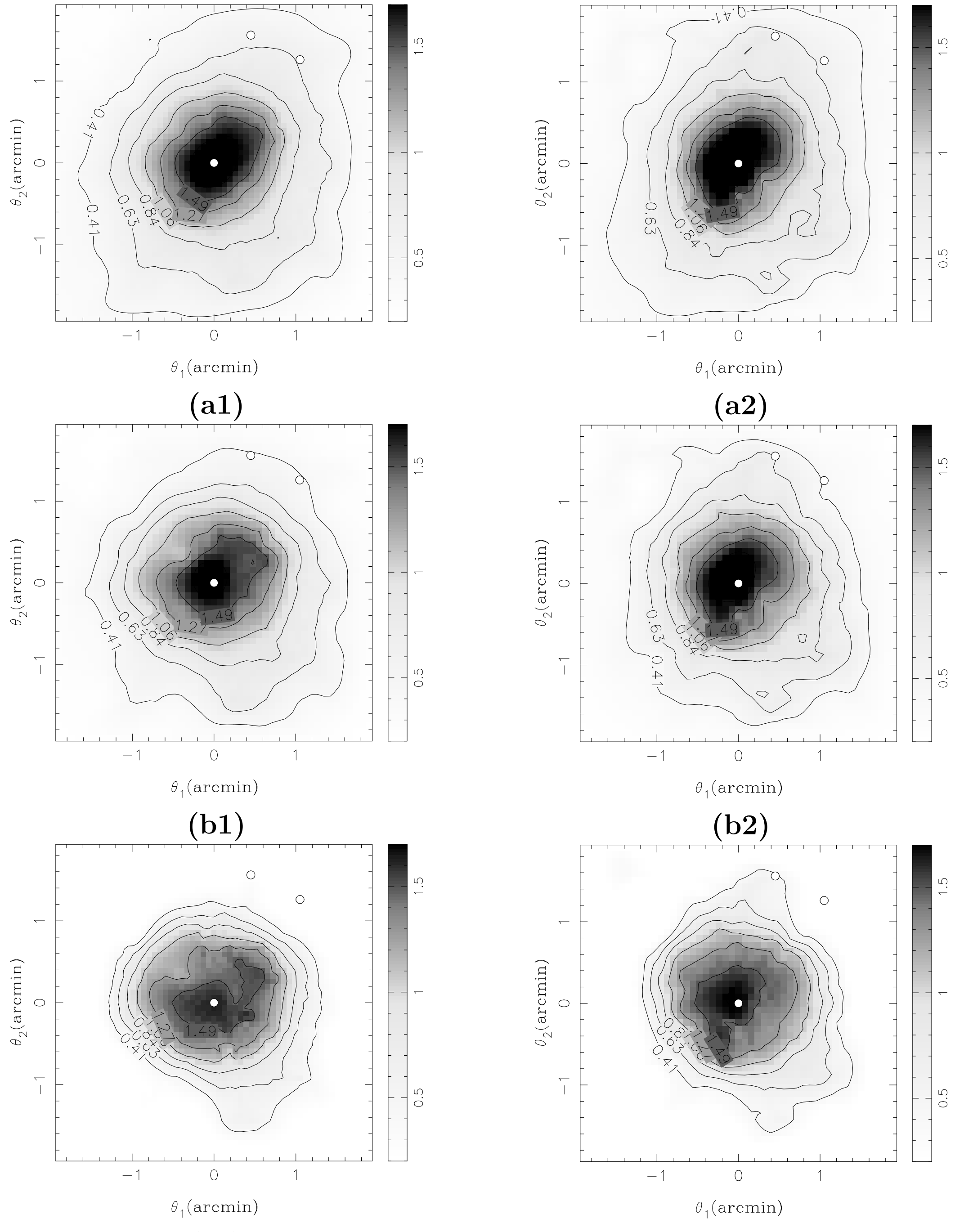
We also use this system to obtain one of the $\kappa^{(0)}$ models. We fit (in the image plane) a non-singular isothermal ellipsoid model (Kormann et al., 1994) where we allow the scaling b_0, ellipticity $|\epsilon_\mathrm{g}|$ and position angle ϕ_g to vary. The best fit model for this system has values of $\{b_0, |\epsilon_\mathrm{g}|, \phi_g\} = \{0\farcm97, 0.30, 1.01\}$. We fix the core radius r_c to $0\farcm1$. For model fitting throughout this Chapter we use `C-minuit` (James & Roos, 1975), a routine which is part of the CERN Program Library.

5.3.3 Weak lensing mass-reconstruction using simulated data

We used the mock catalogues to test the performance of the reconstruction method. For the initial model $\kappa^{(0)}$ we use three different scenarios: $\kappa^{(0)} = 0$ (and $\gamma^{(0)} = 0$) across the whole field (here after *I0*), $\kappa^{(0)}$ taken from the best fit model for the multiple image system (here after *IM*), and a non-singular isothermal sphere model NIS with scaling and core radius being the same as in *IM* (here after *IC*). The same models were used also to obtain the initial coefficients of the linear system (c.f. Sect. 5.2) at each galaxy position (for *I0* we use $\gamma_{1,2} = 0$). These initial models will help us explore the effects of regularisation and the capability of the reconstruction method to adapt to the data.

Finally, the linear system is solved using the `UMFPACK` routine for solving asymmetric sparse linear systems (Davis & Duff, 1999). We perform 30 second-level iterations, each time increasing N_x and N_y by one, starting with an initial 20×20 grid. The results of the reconstructions are shown in Fig. 5.5 for $N_\mathrm{g} = 210$ galaxies distributed in the same manner as the I-band data and $N_\mathrm{g} = 145$ galaxies distributed as the R-band data of the cluster. The regularisation parameter was set to $\eta = 400$ for I-band $\eta = 200$ for R-band data. The different regularisation parameters are used since the numbers of galaxies are different. In such a case the model has enough freedom to change, the final χ^2 had a value of the order of N_dof. It is very comforting to observe that the reconstructed maps do not depend crucially on the initial κ-model we use. This also indicates that we have broken the mass-sheet degeneracy.

From the reconstructed maps we have also estimated the mass within the radius of $1\farcm5$ around the cluster centre (for a redshift $z_\mathrm{d} = 0.4$ this corresponds to $340h^{-1}$ kpc). For this purpose we generated 10 mock catalogues for each band and did the reconstruction again with three different initial conditions. We list the resulting mass estimates in Table 5.1 for both the I and R-band mock catalogues. All the mass estimates are similar, note, however, that the two galaxy catalogues have galaxies partly in common and the errors are therefore correlated. We determine the enclosed mass of the simulated cluster to be $(1.0 \pm 0.1) \times 10^{15} M_\odot$, which is very close to the input value of $M_\mathrm{s}(< 340h^{-1}\ \mathrm{kpc}) = 0.99 \times 10^{15} M_\odot$. The error has been estimated



(a1) (a2)

(b1) (b2)

(c1) (c2)

Figure 5.5: κ-maps obtained from statistical lensing reconstruction using simulated data. **Left** panels show the reconstructions using $N_{\mathrm{g}} = 210$ galaxies distributed in the same manner as the I-band, while for the **right** panels we use $N_{\mathrm{g}} = 145$ galaxies distributed in the same manner as the R-band weak lensing data for RX J1347−1145 (see Chapter 6). The galaxies have been lensed with an N-body simulated cluster. Different initial conditions were used for the reconstruction. In **a1-a2)** we use best fit model from the multiple image system *IM* (see Sect. 5.3.2) in **b1-b2)** we use the *IC* model, an NIS model with the same scaling and core radius as *IM* and in **c1-c2)** we use *I0*, i.e. $\kappa^{(0)} = 0$ on all grid points. The regularisation parameter was set to $\eta = 400$ for the left panel and $\eta = 200$ for the right panel. The positions of the cluster centre and two major subclumps are plotted as white circles.

from the variance of mass determinations form different mock catalogues.

The results shows that our method is *effectively* able to break the mass-sheet degeneracy and is, as a consequence, very efficient in reproducing the cluster mass. Note that a single multiple-image system does not by itself break this degeneracy, since the transformation does not affect the image positions (and the same is true for the magnification ratios). We would need at least two multiple image systems with sources at different redshifts to be able to break the mass-sheet degeneracy using strong lensing data alone.

Unfortunately we can not resolve both clumps present in the simulations. This is due to the fact that the number density of background sources is low and the internal smoothing scale (i.e. the average distance between two source galaxies) is large. If one increases the number density to $\sim 100/\mathrm{arcmin}^2$ the clumps can be resolved.

We have also performed additional reconstructions in which we increased/decreased the original κ of the simulated cluster by 25%. This enables us to confirm that the agreement between input mass and reconstructed mass is not just accidental. We have generated new multiple image systems and new mock catalogues as before. We however do not perform a new strong-lensing reconstruction, for $\kappa^{(0)}$ we intentionally use the same (i.e. in this case "wrong") initial conditions as before. The old *IM* model would not fit the image positions any longer, since they have changed with the scaling of κ. The reconstructed masses of the increased κ simulation are in excellent agreement with the input values. The differences between different models are comparable (slightly smaller) to the ones shown in Table. 5.1. For the lower κ simulation, the reconstructed values are on average the same as the input value, however the scatter is higher. This is expected, since the lens is in this case weaker and the region where we are able to break the mass-sheet degeneracy is smaller than before.

For an additional test we modified the redshift distribution of the source galaxies used for weak lensing by placing some galaxies at higher redshift (with the resulting average distribution of $\langle z \rangle = 1.6$) and regenerated the mock catalogues. While this does not change substantially the accuracy of the determination of the enclosed mass, it does help to better reconstruct the shape of the mass distribution. This is expected, since high-redshift galaxies (when their shape is measured reliably) contribute most to the signal and improve the accuracy of the reconstruction.

5.4 Conclusions

In this Chapter we have developed a new method based on Bartelmann et al. (1996) to perform combined weak and strong lensing cluster mass reconstruction. The particular strength of our method is that we extent the weak lensing analysis to the critical parts of the cluster. In turn, this enabled us to directly include multiple imaging information to the reconstruction. Both are performed on the same scales, in contrast to similar methods proposed in the past where weak lensing on radii larger than the Einstein radius θ_{E} was combined with strong lensing information (see e.g. Kneib et al., 2003).

We have tested the performance of the method on simulated data and conclude that when a quadruply imaged system combined with weak lensing data and individual photometric redshifts is used, the method can very successfully reconstruct the cluster mass distribution. With a relatively low number density of background galaxies, 15/ arcmin^2 we are effectively able to reproduce the main properties of the simulated cluster. In addition, with larger number densities ~ 100/ arcmin^2 of background sources accessible by HST, the substructure in the cluster can be resolved.

We determine the enclosed mass of the simulated cluster to be $(1.0 \pm 0.1) \times 10^{15} M_\odot$, which is very close to the input value of $M_\mathrm{s}(< 1\farcm5) = 0.99 \times 10^{15} M_\odot$. Again, the accuracy of the mass-distribution determination can be further improved by using higher number densities of the sources already available to date. However, even with the data quality we use we are effectively able to break the mass-sheet degeneracy and therefore obtain the mass and mass-distribution estimates without prior assumptions on the lensing potential.

This method can be improved in many ways. First, we have used for the multiply imaged system only the information of the image positions. The reconstruction method can, however, be modified to include the morphological information of each extended source. Instead of using a regular grid, one would have to use adaptive grids and decrease the cell sizes around each of the images. This will be a subject of future work. Second, the photometric redshift determination does not only give the most likely redshift given the magnitudes in different filters, but also the probability distribution for the redshift. This information can be included in the reconstruction. In addition, source galaxies without redshift information can be used and different regularisation schemes can be considered.

Table 5.1: Reconstructed cluster mass within $340h^{-1}$ kpc radius around the cluster centre from simulations using mock catalogues resembling I-band (left) and R-band (right) weak lensing data and one 4-image system. Three different initial conditions have been used. We use best fit model from the multiple image system *IM* (see Sect. 5.3.2), the *IC* model (NIS with same scaling and core radius as *IM*) and *I0* with $\kappa^{(0)} = 0$ on all grid points. In the last line the input mass from the simulation M_s is given. In brackets we give for comparison the velocity dispersion of an SIS having the same enclosed mass within $340h^{-1}$ kpc.

	$M_{I,\mathrm{s}}(< 340h^{-1}$ kpc) $[10^{15} M_\odot]$	$[\sigma_\mathrm{SIS}]$ $[\mathrm{kms}^{-1}]$	$M_{R,\mathrm{s}}(< 340h^{-1}$ kpc) $[10^{15} M_\odot]$	$[\sigma_\mathrm{SIS}]$ $[\mathrm{kms}^{-1}]$
IM	1.04	[1720]	1.03	[1710]
IC	1.02	[1700]	0.99	[1670]
I0	0.85	[1560]	0.83	[1530]
	$M_\mathrm{s}(< 340h^{-1}$ kpc)		0.99	[1670]

6

Strong and weak lensing united II: Cluster mass reconstruction of the cluster RX J1347−1145

Encouraged by the success of our combined cluster mass reconstruction present in Chapter 5, we here apply the method to the weak and strong lensing data for the redshift $z_{\rm d} = 0.451$ cluster RX J1347−1145 (Schindler et al., 1995). Due to its record holding, this cluster has been a subject of many studies in X-ray (Schindler et al., 1995, 1997; Ettori et al., 2001; Allen et al., 2002) and optical (Fischer & Tyson, 1997; Sahu et al., 1998; Cohen & Kneib, 2002; Ravindranath & Ho, 2002). It has also been detected through the Sunyaev-Zel'dovich effect (Pointecouteau et al., 2001; Komatsu et al., 2001; Kitayama et al., 2004). Unfortunately, studies based on X-ray properties, SZE effect, velocity dispersion measures, strong and weak lensing have all yielded discrepant results for the mass estimate (see Cohen & Kneib, 2002 for a summary of the results).

For the purpose of mass reconstruction we use VLT/FORS data on the field of $3.8 \times 3.8\ {\rm arcmin}^2$ in UBVRI bands. We also use J and H-band data from NNT/SOFI and K-band data from VLT/ISAAC to make more reliable photometric redshift estimates. The shape measurements for the weak lensing reconstruction is performed on two FORS bands, R and I. The strong lensing properties of this cluster are analysed. From previous data sets five arc candidates were reported (Schindler et al., 1995; Sahu et al., 1998); using multi-colour data we conclude that only two possibly belong to the same multiple imaged system. We search for additional images possibly belonging to this system and present what we think is the most probable candidate. In addition we report a discovery of a new arc with two components, located at a distance of 1 arcmin from the brightest cluster galaxy.

This Chapter is organised as follows. In Sect. 6.1 we describe the observations and give a brief outline of the data reduction process. In Sect. 6.2 we describe how we search for multiply imaged systems. In Sect. 6.3 we give the results of combined strong and weak lensing reconstruction and we conclude in Sect. 6.4. The results of this part of the thesis are currently in preparation to be submitted to A&A (Bradač et al., 2004c).

6.1 Observations and data reduction process

The observations were obtained in UBVRI bands with FORS1@VLT/UT1 instrument in high resolution mode in the semester between April and September 2001. The field-of-view (FOV) of the FORS instrument is 3.4×3.4 arcmin^2. In addition we took from the ESO archive the J and H band images from SOFI@NTT (with the FOV 5.5×5.5 arcmin^2) and K band from ISAAC@VLT/UT1 (with the FOV 2.5×2.5 arcmin^2).

The data reduction was performed with the data reduction pipeline also used for the Garching-Bonn Deep Survey (Schirmer et al., 2003). We describe here briefly the essential steps of this reduction:

1. Flatfielding: CCD pixels have different sensitivity, and to account for that the flatfielding needs to be performed. In our particular case a combination of skyflats and smoothed superflats were used to calibrate the pixel sensitivity.

2. Masking: Bad pixels, cosmic rays, satellite tracks were identified and masked before co-addition.

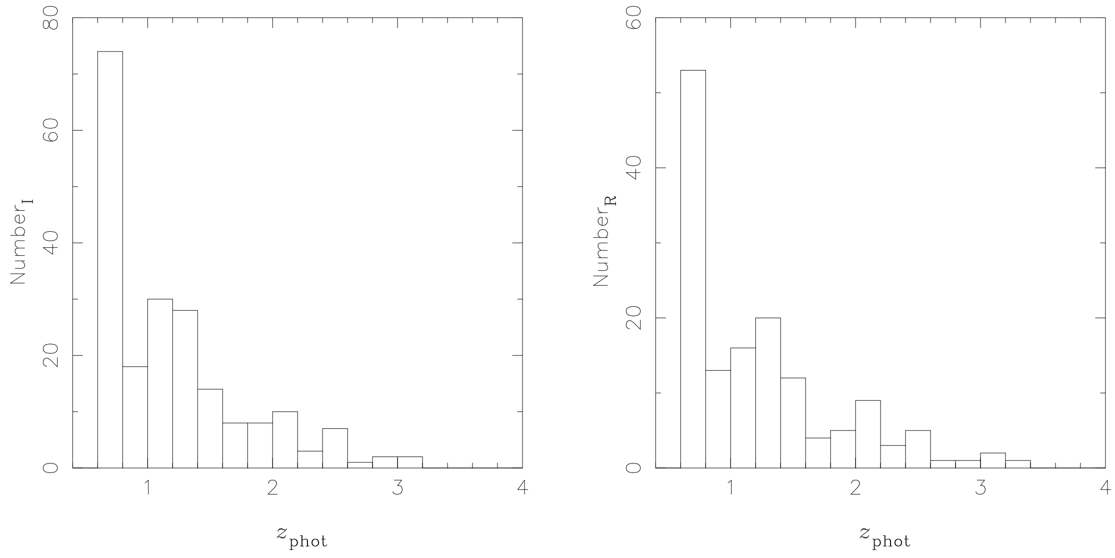
3. Defringing: Fringe pattern can form due to reflections within the CCD (especially problematic for instruments observing in longer wavelengths and having thinner CCDs). For the FORS1@VLT this is however not an issue (see e.g. Maoli et al., 2001).

4. Co-addition: after pre-processing individual exposures are co-added into a so-called sum-frame. An astrometric solution was obtained by comparing simultaneously object positions with standard sources from USNO-A2 catalogue (Monet et al., 1998) and with overlap objects in the individual frames. The IRAF package DRIZZLE (Fruchter & Hook, 2002) was used to align different pointings in sub-pixel accuracy. Since dithering was used (individual frames were slightly offset w.r.t. each other to avoid having gaps in the final image) the resulting image (for UBVRI bands) has a size of 3.8×3.8 arcmin^2, which is larger than the individual frames.

6.1.1 Redshift distribution of background sources

For the photometric redshift determination we estimate magnitudes of each object using SExtractor (Bertin & Arnouts, 1996). This program is designed to build a catalogue of objects from an astronomical image. We use it in double-image mode, meaning that the I-band image serves as a reference to build the object catalogue, ant the other images are used to measure properties of these selected ones. The magnitudes of the objects are measured (in different filters) by placing apertures around each of them in different images. Using the resulting magnitudes we determine the photometric redshifts of the objects using the HyperZ package (Bolzonella et al., 2000).

Special care needs to be taken when doing photometry for the purpose of the photometric redshift determination. Especially important is the choice of the aperture size SExtractor is using. For cluster members, due to their large apparent size, we argue that it is best to use SExtractor in Magiso mode, meaning that the magnitude will be measured within an isophote level that depends on the photometric error information. To measure the magnitudes of the faint sources, we use the Magaper mode, meaning that we give a fixed size for the aperture. The aperture was set to twice the FWHM of the seeing disc. In our case the seeing was $0'.7$ in BVRIK, $1'.0$ in U and J, and $1'.1$ in H. The reason for doing photometry in this way is that the use of the Magiso algorithm is not appropriate for measuring the magnitudes of faint sources. The noise for these objects is usually underestimated and the resulting aperture would be too large. On the other hand the small aperture used for background sources is not applicable for the large cluster members and their analysis was therefore performed separately.



(a) **(b)**

Figure 6.1: The redshift distributions of background sources used for weak lensing analysis (only sources with $z_{\mathrm{phot}} > 0.6$ were considered) for I-band **a)** and R-band catalogue **b)**. The mean photometric redshift of both samples is $\langle z \rangle = 1.3$.

6.1.2 Final weak lensing catalogues

For the weak lensing analysis the R and I band exposures were used. The I-band exposure is the deepest and has the best seeing of all and as a consequence has the highest number-density of sources that can be used for weak lensing. Special care needs to be taken when measuring the galaxy shapes and doing PSF correction for weak lensing analysis. Although SExtractor does measure ellipticities of objects as well, it does not take into account the PSF variation across the images. These measurements are thus not optimal for weak lensing. Instead special software packages have been developed in last 10 years that measure the shapes of the objects and simultaneously correct for the PSF anisotropy. We use the IMCAT software from Nick Kaiser (`http://www.ifa.hawaii.edu/~kaiser`), which is based on the KSB method described in Kaiser et al. (1995). The catalogues were obtained as described in Erben et al. (2001).

For the final weak lensing catalogue only the sources having photometric redshift estimate $z_{\mathrm{phot}} > 0.6$ were considered. This conservative cut was done in order to avoid having foreground sources in the catalogue. For the R-band data we thus have $N_{\mathrm{g}} = 145$ background sources (giving 10 galaxies per arcmin^2, and for the I-band $N_{\mathrm{g}} = 210$ (15/arcmin^2). The resulting distribution for the I and R band catalogues are given in Fig. 6.1.

6.2 Searching for multiply imaged candidates of RX J1347−1145

Up to date, five arcs have been reported in the cluster. The first two were discovered by Schindler et al. (1995), and the HST images revealed three additional ones (Sahu et al., 1998). Unfortunately these five arcs (A1-A5 as labelled by Sahu et al., 1998) do not belong to the same multiple image system. However two of them (A4 and A5) do have the same colours and we use them for the purpose of multiple imaging. We note that the arc A3 considered by Allen et al. (2002) to belong to this system as well has different colours.

The advantage of having J, H, and K band images allows the detection of new arcs. In particular, we report here on the discovery of a red double-component arc to the south-west of A4, we designate it with labels B1 and B2. The arc formed in the middle of a concentration of cluster members. In addition we detect in the vicinity a long thin arc, which was also presented in Lenzen et al. (2004) as number 3 (see Fig. 6.3). We use SExtractor (note that for the purpose of strong lensing this is sufficient) to measure the ellipticities of the arcs from the I-band (system A and C) and K-band images (system B).

We search for additional images in an automated fashion. The magnitudes of an image in either $N_{\mathrm{f}} = 8$ of $N_{\mathrm{f}} = 5$ filters $m_{\mathrm{i},f}$ are compared with the magnitudes $m_{j,f}$ of all other images in the field. We use again the χ^2 approach

$$\chi^2_{i,j} = \sum_{f=1}^{N_{\mathrm{f}}} \frac{\left(m_{\mathrm{i},f} - (m_{j,f} + \mu_{i,j})\right)^2}{\sigma^2_{i,f} + \sigma^2_{j,f}} \,, \tag{6.1}$$

where $\mu_{i,j}$ is the relative magnification between the images i and j, and $\sigma_{i,f}$ and $\sigma_{j,f}$ are the magnitude measurement errors. Since lensing is achromatic we can evaluate $\mu_{i,j}$ by forcing $\partial\chi^2_{i,j}/\partial\mu_{i,j} = 0$ to hold. The resulting $\chi^2_{i,j}$ function follows a χ^2-distribution with $N_{\mathrm{f}} - 1$ degrees of freedom. The best fitting images are then further visually analysed and tested for the conservation of surface brightness. For the case of arcs A4 and A5 we find a possible candidate counter image. All three all encircled dashed-yellow in Fig 6.2.

Using 8 colours for the redshift determination of A4-A5 and 5 colours for the counter image (it is located at the edge of the J, H, and K-band images and therefore the NIR photometry is not reliable) we find all three images are consistent at being at the redshift of $z_{\rm s} = 1.76$. This is in agreement with Ravindranath & Ho (2002) who, based on the absence of $O[II]$ line in their spectroscopy predict the redshift of A4 to be > 1.04. Unfortunately our redshift estimate is not very secure and can substantially influence the combined cluster mass reconstruction (the position of the critical curve changes with redshift). This is not important for the determination of the initial lens model $\kappa^{(0)}$ (we saw in the previous section that the choice of vastly different $\kappa^{(0)}$-models does not substantially influence the result), it is however important when including strong lensing information directly in the reconstruction. There are more candidate multiple image systems in this field; they will be a subject of a future study.

Table 6.1: The properties of the arcs A4,A5 and the candidate counter-image used in strong lensing analysis. We also present additional arc candidates (systems B and C). The properties of systems A and C are measured from the I-band, while B is measured from the K-band image. The positions are given w.r.t. the brightest cluster member.

Arc	θ_1 [arcmin]	θ_2 [arcmin]	$\|\epsilon\|$	PA[deg]
A4	0.3314	−0.4980	0.713	43.2
A5	−0.2891	−0.7009	0.333	167.9
	−1.0241	0.6509	0.327	23.4
B1	0.7440	−0.6873	0.282	67.5
B2	0.6971	−0.7653	0.593	33.6
C	0.3222	−0.7532		

Table 6.2: The photometric properties of the arcs A4 and A5. Given are three colours ($m_{\rm B} - m_{\rm I}$, $m_{\rm V} - m_{\rm I}$, and $m_{\rm R} - m_{\rm I}$) in magnitudes, and VRI peak surface brightnesses $S_{\rm V,R,I}$ (in rel. units). If objects belong to the same source the colours and surface brightness need to be conserved. We estimate a typical magnitude measurement error to be 0.05 mag in VRI and 0.1 mag in B-band.

	$m_{\rm B} - m_{\rm I}$	$m_{\rm V} - m_{\rm I}$	$m_{\rm R} - m_{\rm I}$	$S_{\rm V}$	$S_{\rm R}$	$S_{\rm I}$
A4	0.99	0.81	0.53	0.32	0.41	0.37
A5	1.09	0.88	0.57	0.33	0.42	0.41

6.3 Cluster mass reconstruction of RX J1347−1145

In this section we present the mass modelling of the data described above for the cluster RX J1347−1145. The full account of the method can be found in Chapter 5.

6.3.1 Initial conditions for the method

For the purpose of obtaining the initial values for $\kappa^{(0)}$, $\gamma^{(0)}$, and $\boldsymbol{\alpha}^{(0)}$ we investigate the signal from the averaged tangential ellipticities (hereafter called *IS* scenario)

$$\epsilon_{\rm t} = -\Re\left[\epsilon\,{\rm e}^{-2{\rm i}\phi}\right] \tag{6.2}$$

where ϕ specifies the direction to the source galaxy w.r.t. the the brightest cluster galaxy (BCG) (in the case of RX J1347−1145 this is the western of the two central cluster galaxies). We use the ellipticity measurements in I and R band and in Fig. 6.4 we plot the tangential ellipticity vs. projected radius $\langle\theta\rangle$ in radial bins centred on the BCG, containing 50 (40) galaxies each. Both I and R-band data give comparable results. We note that the tangential ellipticity signal is high at the edge of the field $\langle\epsilon_{\rm t}\rangle \sim 0.1$, thus making this data inaccessible for standard weak lensing techniques aiming to determine the mass, since on this relatively small field one cannot break the mass-sheet degeneracy by simply assuming $\kappa \sim 0$ at the field edges.

We fit the singular isothermal sphere (SIS) profile to the binned tangential ellipticities. The individual redshifts of the sources were therefore not used here, instead we use the average source redshift of $\langle z\rangle = 1.3$ (see Sect. 6.1.1). The resulting line-of-sight velocity dispersion is $\sigma_{\rm I,SIS} = 1010 \pm 60$ kms^{-1} for the I-band data and $\sigma_{\rm R,SIS} = 1060 \pm 90$ kms^{-1} for the R-band (both 1σ error bars). In addition, the central mass concentration is detected with more than 10σ significance in both bands.

The estimates for the line-of-sight velocity dispersion are higher than the measured velocity dispersion from Cohen & Kneib (2002), and lower than previous weak, strong and X-ray measurements. However, in the optical it is evident that the cluster has a lot of structure and therefore the SIS profile does not describe the cluster adequately. It has at least two main components; in addition there is X-ray emission off-centred from the BCG. Furthermore, at the scales $\lesssim 400h^{-1}$ kpc where we measure the profile of the cluster is probably not isothermal (see e.g. Navarro et al., 2004). Therefore, the values of σ obtained in this manner should not be trusted, we only use them for one of the initial models for $\kappa^{(0)}$

Another possibility to obtain initial conditions is to use the multiple image information for the cluster. We perform a very rough analysis by using the data for the arc system A presented in Sect. 6.2. In addition to the image positions we also use image ellipticities as constraints. The model consists of a non-singular isothermal ellipse (NIE) model in which we allow the centre of the cluster potential $\boldsymbol{\theta}_{\rm cl}$, the scaling b_0, ellipticity $|\epsilon_{\rm g}|$, and the position angle $\phi_{\rm g}$ to vary, and we fix the core radius to $0\overset{\prime\prime}{.}3$. Following the prescription of Kneib et al. (1996) we also include 10 brightest cluster members (selecting them according to their I-band magnitude and having photometric redshift estimate between 0.4 and 0.5) to the model. They have been modelled as non-singular isothermal spheres with line-of-sight velocity dispersion $\sigma_{\rm nis}$ and core radius $r_{\rm c,nis}$ following

$$\sigma_{\rm nis} \propto L^{1/4}\ , \qquad\qquad r_{\rm c,nis} \propto L^{1/2}\ , \tag{6.3}$$

The proportionality constants were chosen such that the I-band magnitude $m_I = 17.5$ galaxy would have $\sigma_{\rm nis} = 300\ {\rm km\,s^{-1}}$ and $r_{\rm c,nis} = 0\overset{\prime}{.}1$ (the BCG has $m_I = 17.8$). We also fix the core radius of the cluster to $r_{\rm c} = 0\overset{\prime}{.}3$. These constants were *not* allowed to vary. The particular, best fit model for this system has values of $\{\theta_{\rm cl,1}, \theta_{\rm cl,2}, b_0, |\epsilon_{\rm g}|, \phi_g\} = \{-0\overset{\prime}{.}21, -0\overset{\prime}{.}10, 0\overset{\prime}{.}97, 0.3, 0.8\}$.

We stress here that it was not our aim to obtain a detailed strong lensing cluster-mass model, since it will only be used for the initial values of reconstruction. The multiple imaged system used here is independently included to the non-parametric reconstruction. We have shown in Chapter 5 (and will also confirm this in Sect. 6.3) that the reconstruction depends little upon the details of the initial model we use. For this reason a detailed modelling is not needed. For the same reason we also do not include additional multiply imaged candidates to the analysis.

6.3.2 Combined weak and strong lensing mass reconstruction of RX J1347−1145

We apply the mass-reconstruction method to the I and R-band data of RX J1347−1145. We again use three different initial models for $\kappa^{(0)}$, *IM* comes from the best fit model from the strong lensing analysis of the cluster presented in Sect. 6.3.1, the *IS* model is the best fit SIS model to binned tangential ellipticities – see Sect. 6.3.1 (centred on the brightest cluster member) and *I0* has $\kappa^{(0)} = 0$. The regularisation parameter is set to $\eta = 200$ for the R-band and $\eta = 400$ for the I-band. The resulting κ-maps are given in Fig. 6.6. We also plot the contours from Fig. 6.6 a1 to the colour composite image in Fig. 6.2. We estimate the mass within the radius of $1\overset{\prime}{.}5$ (for a redshift $z_{\rm d} = 0.45$ this corresponds to $360h^{-1}$ kpc). These are given in Table 6.3. Results for both bands are consistent and mass estimates for different models give similar results. We estimate the mass of the cluster to be $M(<360\,h^{-1}{\rm kpc}) = (1.2 \pm 0.3) \times 10^{15} M_{\odot}$. The error is larger than in the case of simulated images, this is mostly attributed to the fact that we only use a three-image system. In addition, when computing the redshifts for the mock catalogues we did not include the outliers (i.e. galaxies with redshift estimates which can be wrong by more than factor 2). We conclude that the enclosed mass of the cluster is significantly higher than what is obtained under the assumption of circular symmetry in Sect. 6.3.1.

In Fig. 6.5 we plot all objects having estimated photometric redshifts $2.05 < z_{\rm phot} < 2.15$. The figure suggests that they form an arc-like structure $\theta_a \sim 1'$ to the south of the BCG. If these objects trace the critical curve at the redshift of $z_{\rm a} = 2.1$, as is suggested by this figure, then to a first approximation the corresponding mass enclosed within a circle of radius $\theta_{\rm a}$ is given by

$$M(<\theta_{\rm a}) \approx \pi\,(D_{\rm d}\theta_{\rm a})^2\,\Sigma_{\rm cr}(z = z_{\rm a})\ , \tag{6.4}$$

where we assumed the spherical symmetry of the cluster. In our case this yields an estimate of $M(< 240\ \mathrm{kpc}h^{-1}) \approx 8 \times 10^{14} M_{\odot}$. Note that this is a very simple approach to determine the enclosed mass, its reliability is limited since arcs form preferentially along the major axis of clusters. The presence of substructure can further bias the result. In our case, due to the orientation of the two brightest cluster galaxies (east-west) one might think the galaxies to the south trace the minor rather than the major axis. However, our detailed mass reconstruction (and X-ray measurements) show a south-east extention. In addition there are bright cluster members located to the south-west of the BCG (c.f. Fig. 6.6 and 6.2). Therefore the result should not be trusted in detail, however it is suggestive of a high mass for RX J1347−1145.

The resulting mass is larger than obtained by velocity dispersion measurement of Cohen & Kneib (2002). It is, however, consistent with X-ray data by Allen et al. (2002) and Ettori et al. (2001) and the previous weak-lensing results by Fischer & Tyson (1997). A possible explanation for the discrepant mass estimates was presented in Cohen & Kneib (2002). We agree that the cluster is most likely in a pre-merging process (with clumps merging preferentially in the plane of the sky). In such a scenario, until the merging is complete and the cluster is virialised, the dynamical cluster mass will be largely underestimated. The X-ray temperature can be increased in such merging processes and the high X-ray mass estimate for this cluster further supports the merger hypothesis. If the hypothesis is correct, gravitational lensing does provide the most accurate estimate for the cluster mass.

We note however, that our results are dependent upon the correct redshift determination and identification of the members of the multiple image system we use. If we put the multiple image system to a redshift of ~ 3 (~ 1.3) the estimated mass decreases(increases) by $\sim 10\%$. If the images do not belong to the same system, the changes might be even more drastic. However, at least for the two arcs A4 and A5 we consider this possibility less likely. In Table 6.2 we list the photometric properties of the two arcs. Within the errors they have the same colours as well as the same surface brightnesses. In addition, the photometric redshift estimation (using 8 filters) is the same for both images. Further, our results depend upon the correct determination of the photometric redshifts for the weak lensing sources. The random error of the determination is not crucial, the problem are the systematic uncertainties. It is namely not excluded that some foreground sources get assigned a high redshift and therefore high weight factor $Z(z)$. If the fraction of such sources compared to the total number of the sources used is large, the signal can be diluted (assuming these foreground sources are randomly oriented).

6.3.3 Rest-frame I-band brightness distribution and M/L-ratio of RX J1347−1145

In Fig. 6.3.2 we plot the $m_{\mathrm{V}} - m_{\mathrm{I}}$ vs. $m_{\mathrm{B}} - m_{\mathrm{V}}$ colours for the galaxies in our field. To select cluster members we cut the catalogue at all galaxies having I-band magnitude up to three magnitudes fainter than the BCG and with distance to the BCG smaller than $1'$. These are preferentially the cluster members and form a group around $m_\mathrm{V} - m_\mathrm{I} \sim 2$ and $m_\mathrm{B} - m_\mathrm{V} \sim 1$. The BCG (plotted blue in Fig. 6.3.2) is slightly bluer than the rest. The green polygon indicates the region from which we selected the cluster members. It is given by:

$$1 < m_\mathrm{B} - m_\mathrm{V} < 2\ ; \quad 0.7\,(m_\mathrm{B} - m_\mathrm{V}) + 0.7 < m_\mathrm{V} - m_\mathrm{I} < 0.7\,(m_\mathrm{B} - m_\mathrm{V}) + 1.3\ . \tag{6.5}$$

To obtain absolute rest-frame I and R-band magnitudes for the cluster members we apply the appropriate K-correction $K_\mathrm{I,R}(z_\mathrm{d})$ for elliptical galaxies from Poggianti (1997) and galactic extinction A to the measured isophotal I-band magnitudes. They are given by

$$M_\mathrm{I,R} = m_\mathrm{I,R} - 5\log_{10}\left(\frac{D_\mathrm{d}(1+z_\mathrm{d})^2}{10\ \mathrm{pc}}\right) - K_\mathrm{I,R}(z_\mathrm{d}) - A_\mathrm{I,R}\ , \tag{6.6}$$

where $K_\mathrm{I}(z_\mathrm{d} = 0.45) = 4.13$, $K_\mathrm{R}(z_\mathrm{d} = 0.45) = 4.28$, $A_\mathrm{I} = 0.121$, and $A_\mathrm{R} = 0.166$. We assume zero evolutionary correction. The resulting luminosities of the cluster members were smoothed using a Gaussian kernel characterised by $\sigma = 9''$, resulting in the brightness distribution shown (for I-band only) in Fig. 6.7.

The first concern with luminosity estimates is completeness. In Fig. 6.9 we plot the luminosity functions as a function of the absolute magnitudes M_I and M_R, respectively. We estimate the errors by taking into account the Poissonian fluctuations of the galaxy counts. From the plot we see that the cluster member counts are complete to at least the magnitude $M^*_\mathrm{I,R} - 3$ (where M^* is the magnitude at the "break" where the luminosity function changes the slope, see Fig. 6.9) and therefore the contribution from incompleteness will be negligible. A far larger concern is the contamination by non-cluster members and rejection of the actual members. In order to check against this one needs to investigate the galaxy population "outside" the cluster region (on images taken with the same photometric conditions and depth as the images we use). Unfortunately our observed fields span only ~ 450 kpc h^{-1}

Table 6.3: Reconstructed mass of RX J1347−1145 within $360h^{-1}$ kpc radius around the BCG from I-band (left) and R-band (right) weak lensing data and one candidate 3-image system. Three different $\kappa^{(0)}$ scenarios have been used. We use the best fit model from the multiple image system *IM* presented in Sect. 6.2, *IS* is the best fit SIS model from the process of parametrised fitting of weak lensing data and *I0* has $\kappa^{(0)} = 0$ on all grid points (see Sect. 6.3.1). In brackets we give for comparison the velocity dispersion of an SIS having the same enclosed mass within $360h^{-1}$ kpc.

	$M_I(< 360h^{-1}$ kpc) $[10^{15}M_\odot]$	$[\sigma_\mathrm{SIS}]$ $[\mathrm{kms}^{-1}]$	$M_R(< 360h^{-1}$ kpc) $[10^{15}M_\odot]$	$[\sigma_\mathrm{SIS}]$ $[\mathrm{kms}^{-1}]$
IM	1.46	[1950]	1.36	[1880]
IS	1.19	[1770]	1.19	[1770]
I0	1.10	[1700]	1.03	[1650]

around the brightest cluster galaxy and therefore this approach can not be used in our case. We conclude that the error budget on luminosity will be dominated by the systematics of the cluster member selection and contamination and is difficult to estimate. However in order to estimate the mass-to-light ratio the mass determination will be a dominant source of error.

To estimate the mass-to-light ratio we calculate the aperture luminosity L_a by adding the luminosities of the cluster members within $360h^{-1}$ kpc radius around the BCG. The resulting R and I band luminosities are $L_\mathrm{a,I}(< 360h^{-1}\mathrm{kpc}) = 2.6 \times 10^{12} L_\odot$ and $L_\mathrm{a,R}(< 360h^{-1}\mathrm{kpc}) = 1.8 \times 10^{12} L_\odot$, respectively. The mass-to-light ratios are $M/L_\mathrm{I} = 450 \pm 150 M_\odot/L_{\mathrm{I},\odot}$ and $M/L_\mathrm{R} = 650 \pm 150 M_\odot/L_{\mathrm{R},\odot}$. The high value results from the fact that the mass of the cluster is high. The cluster has only 200 members across this field, therefore it is under-luminous in optical bands. This is however not the case for the X-ray luminosity. In addition we are measuring the M/L ratio in the inner part of the cluster, which might not reflect the M/L ratios measured out to ~ 2 Mpc distances from cluster centres usually quoted in the literature.

6.4 Conclusions

The case of RX J1347−1145 has been a cause of many puzzles in the past. Very discrepant mass estimates are given in the literature, and unfortunately ours is not the only case where the mass measurements have proven to be difficult. We have applied the reconstruction method to deep optical data using a multiple-image system with three images selected based on their colours and redshifts.

We make the following conclusions.

1. The combined strong and weak lensing mass reconstruction confirms that the most X-ray luminous cluster is indeed very massive. If the redshift and identification of the multiple-image system, as well as redshifts used in weak lensing data, are correct we estimate the enclosed cluster mass within $360\,h^{-1}$kpc to $M(< 360\,h^{-1}\mathrm{kpc}) = (1.2 \pm 0.3) \times 10^{15} M_\odot$.

2. The reconstruction shows a south-east mass extension, compatible with the X-ray measurements.

3. A single component weak lensing model is not applicable for this cluster, a detailed modelling needs to be performed.

4. We have demonstrated the feasibility of breaking the mass sheet degeneracy in practice by using shape measurements and adding the information on individual redshifts, without any assumptions regarding the cluster potential.

In addition we measured the corresponding mass-to-light ratio of the cluster within $360\ h^{-1}$kpc. We find that the cluster is more luminous in the rest-frame I-band, which is expected due to the presence of many old (red) elliptical galaxies in clusters. The resulting mass-to-light ratios are high both in rest-frame I and R-band, giving $M/L_\mathrm{I} = 450 \pm 150 M_\odot/L_{\mathrm{I},\odot}$ and $M/L_\mathrm{R} = 650 \pm 150 M_\odot/L_{\mathrm{R},\odot}$. These values are higher than typical values for clusters claimed in the literature (~ 200 for R-band). However it is difficult to compare our results with existing measurements of the mass-to-light ratios, since they are usually performed at larger radii not accessible with our data.

In the course of this research we discovered one new extremely red arc (system B) at ~ 1 arcmin distance from the BCG. In addition we found an overdensity of a redshift 2.1 sources forming an arc-like configuration. If they indeed trace the $z = 2.1$ critical curve, the crude mass estimate confirms the detailed mass estimate from combined strong and weak lensing analysis. In addition, the high mass we obtain using the combined reconstruction also fits reasonably well the standard mass vs. X-ray luminosity relation (see Ettori et al., 2004).

One can test the reconstructed potential by using additional galaxies in the field that have confirmed high redshifts and search for their counter-images. Their image position (with redshift of the source known) is projected to the source plane and back to the image plane. If further image positions are possible given our model, the data can be further inspected there, and if the lens model is correct objects having the same colours and redshift as the original image should be found. In addition, in the vicinity of the critical curves background sources can be highly magnified. Since the critical curves corresponding to high redshift sources can be calculated from the model, one can search for objects at redshift and luminosity that would not be observed otherwise. This detailed strong lensing analysis of the multiply-imaged candidates in the field will be a subject of future work.

The mass-reconstruction of RX J1347−1145 can be significantly improved. Deep HST imaging would greatly help in identifying and confirming new multiple-image systems that we can use, thus allowing even more detailed modelling. In addition, not only the centre of the light for each of the arcs can be used as constraints, but also their morphology. As mentioned in Chapter 5, the reconstruction technique with an adaptive grid at the image positions can be used for these purposes. Further, spectroscopic redshifts need to be obtained for the multiple-image system candidates. Deep, wide-field imaging data of this cluster will help us improve the weak lensing constraints and the reconstruction can be performed at larger radii than presented here.

In conclusion, even without the best data quality currently accessible, we were able to perform a detailed cluster-mass reconstruction of the most X-ray luminous cluster RX J1347−1145. The method has also shown a high potential for the future. If the highest quality data is used, a combination of strong and weak lensing has proven to offer a unique tool to pin down the masses of galaxy-clusters as well as their profiles and accurately test predictions within the CDM framework.

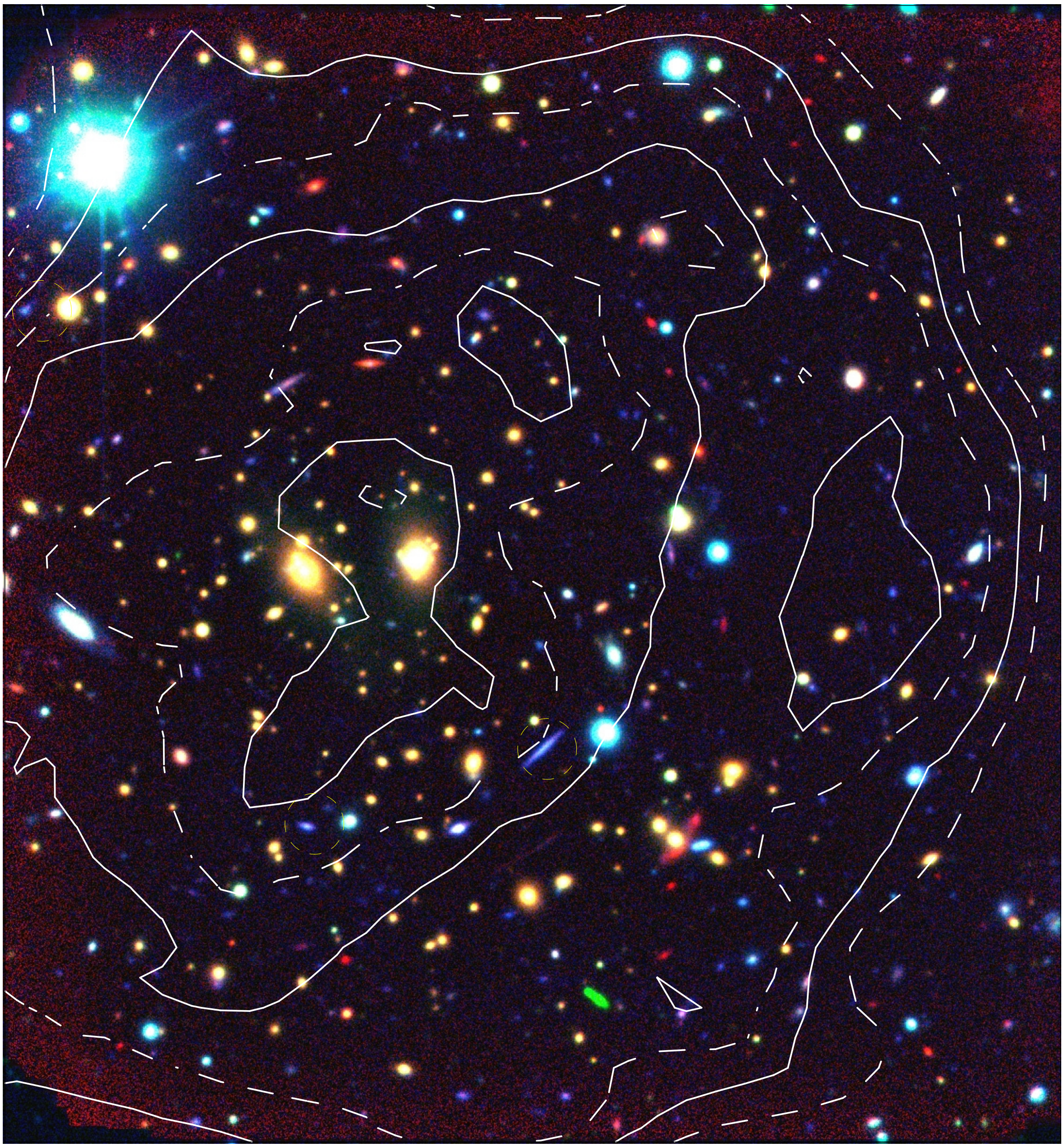
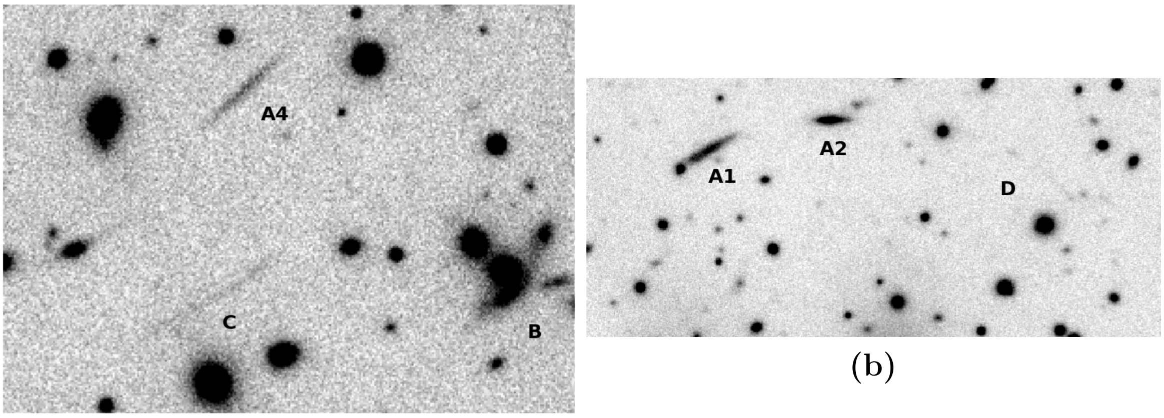
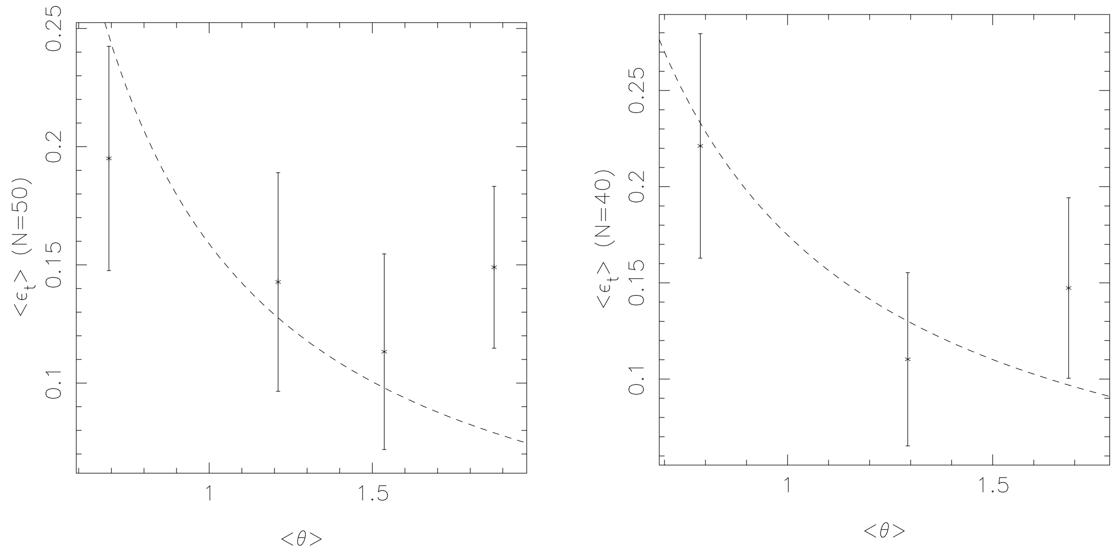


Figure 6.2: The BRK colour composite of the $\sim 3 \times 3.2\,\mathrm{amin}^2$ field of the RX J1347−1145 (from Mischa Schirmer). Overlaid is in white contours the combined weak and strong lensing mass reconstruction from Fig 6.6a. The contour levels are the same (the field here is smaller), for convenience we plot them with interchanging dashed and solid line. Yellow circles show the multiple image system we use. North is up and East is left.



(a) (b)

Figure 6.3: **a)** A ∼ $1'\!.0 \times 0'\!.6$ cutout of the K-band image of the RX J1347−1145 showing one of the arcs used for strong lensing (top), a newly discovered arc system B1-B2 (bottom right, c.f. Fig. 6.2) and the long thin arc C (bottom left) **b)** A ∼ $1'\!.0 \times 0'\!.2$ cutout of the I-band image of the RX J1347−1145 showing the arc system D, possibly consisting of two subcomponents(see Table 6.1).



(a) (b)

Figure 6.4: Tangential ellipticity ϵ_t vs. projected radius $\langle\theta\rangle$ in radial bins centred on the brightest cluster member containing 50 galaxies each for the I-band data **a)** and 40 galaxies for the R-band data **b)**. The errors are obtained by randomising the phases of the measured ellipticities, while preserving their absolute values. 100 randomisations were performed. Dashed line gives the best fit SIS profile to the data, for the I-band we obtain $\sigma_\mathrm{I,SIS} = 1010 \pm 60$ kms^{-1} and for the R-band $\sigma_\mathrm{R,SIS} = 1060 \pm 90$ kms^{-1} (both 1σ error bars).

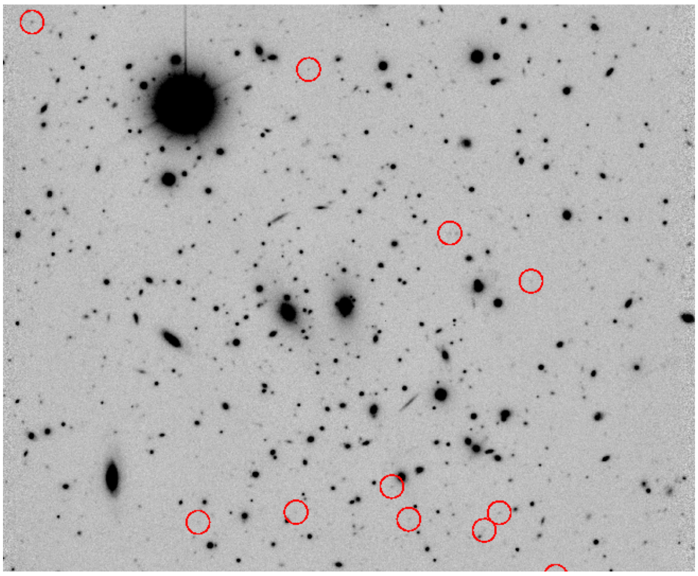
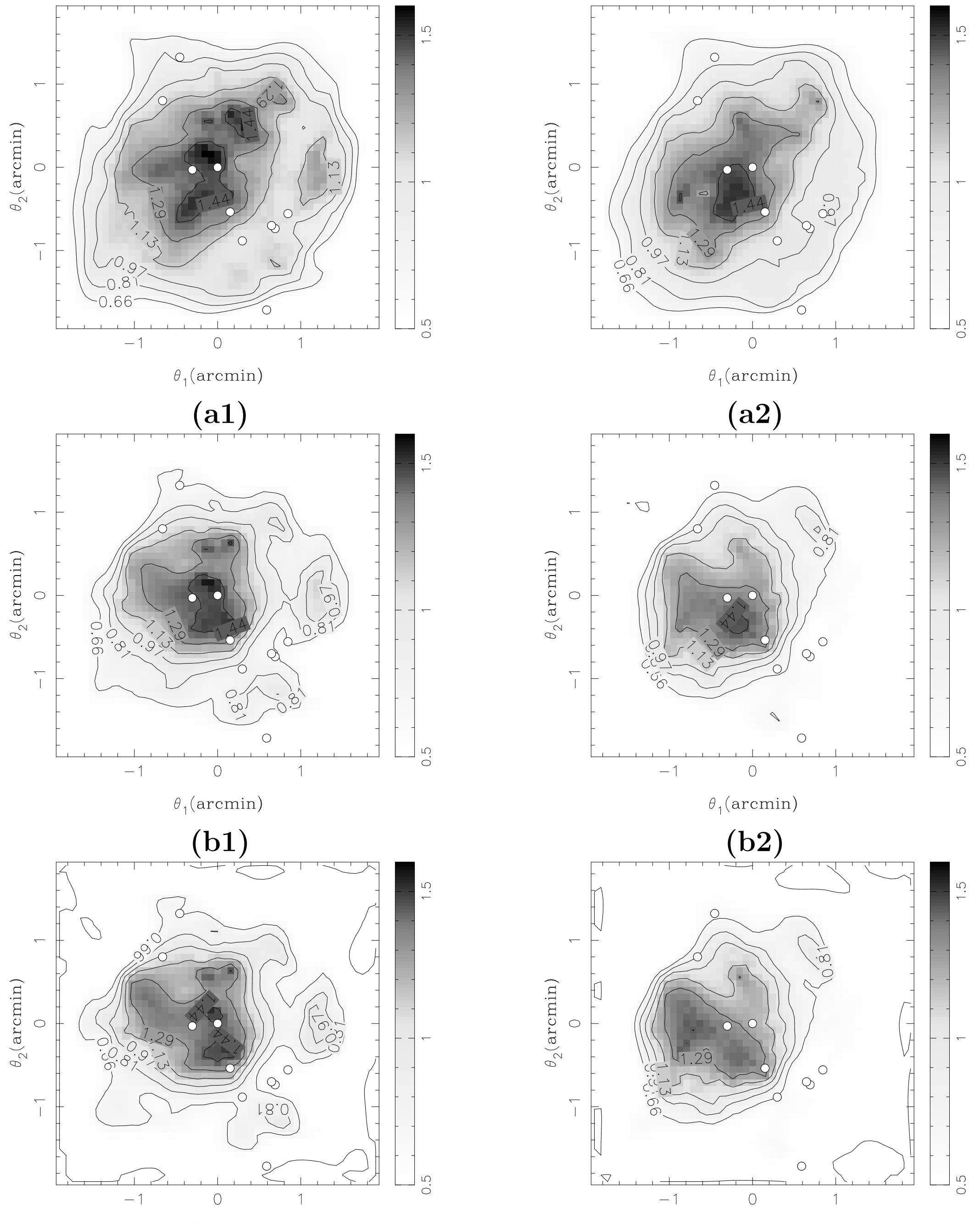


Figure 6.5: A $\sim 3\rlap{.}'5 \times 3\rlap{.}'5$ cutout of the I-band image of the RX J1347−1145. In circles we mark all the objects having photometric redshift estimation $2.05 < z_{\mathrm{phot}} < 2.15$. Objects to the south of the BCG are suggestive of a critical curve location for the sources at $z \sim 2.1$.



(a1) **(a2)**

(b1) **(b2)**

(c1) **(c2)**

Figure 6.6: κ-maps obtained from statistical lensing reconstruction using weak lensing data of the cluster RX J1347−1145. **Left** panels show the reconstructions using I-band data and for the ones on the **right** we use the R-band data. For R-band data we have $N_\mathrm{g} = 145$ background source galaxies, and for I-band $N_\mathrm{g} = 210$ all with known photometric redshifts. In **a1-a2)** we use best fit model from the strong lensing analysis of the cluster *IM* presented in Sect. 6.3.1, in **b1-b2)** we use the *IS* model (SIS model fitted to binned tangential ellipticities – see Sect. 6.3.1, centred on the brightest cluster member) and in **c1-c2)** we use *I0*, i.e. $\kappa^{(0)} = 0$ on all grid points. The regularisation parameter was set to $\eta = 200$ for R and $\eta = 400$ for the I-band. The positions of 10 brightest cluster members are plotted as white circles.

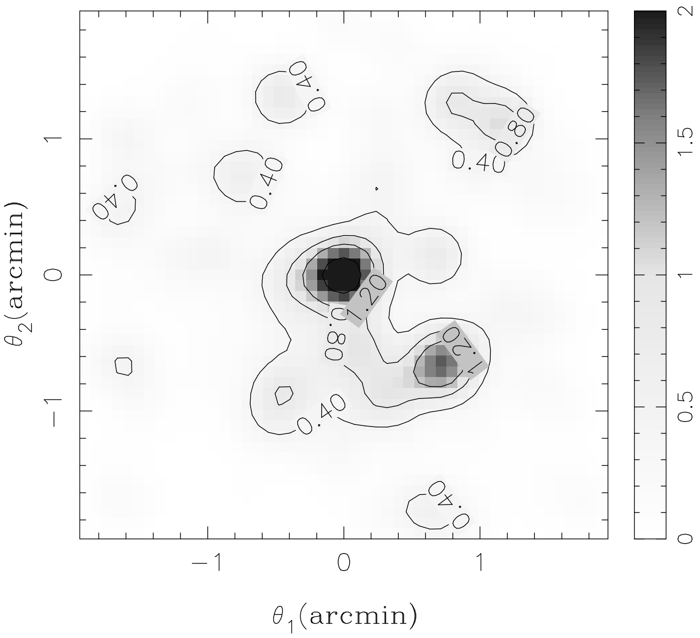


Figure 6.7: The I-band brightness distribution of the RX J1347−1145 in $10^{12} L_\odot/$ arcmin^2. The cluster members were selected using the colour cuts described in the text. Their luminosities have been smoothed using a Gaussian kernel characterised by $\sigma = 9''$.

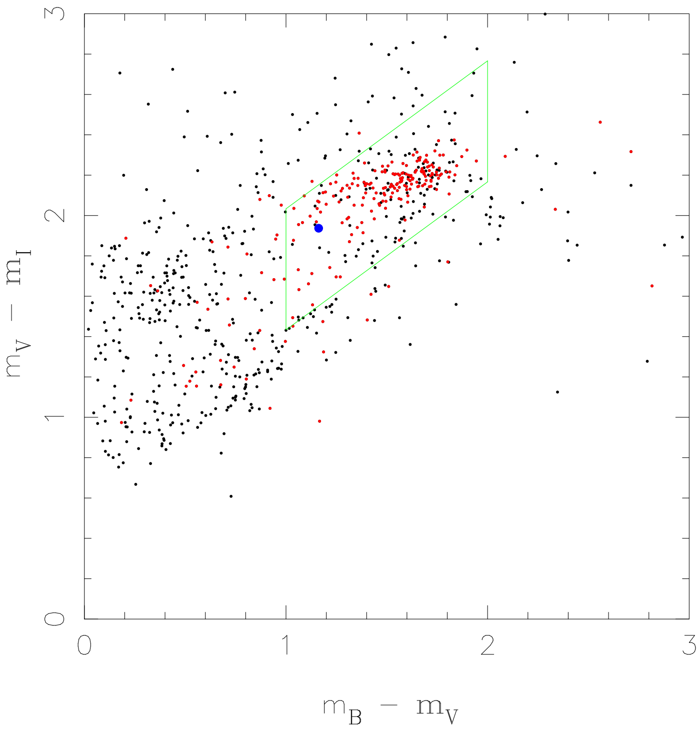


Figure 6.8: The $m_\mathrm{V} - m_\mathrm{I}$ vs. $m_\mathrm{B} - m_\mathrm{V}$ colours for the galaxies in our field. Cluster members were selected to lie inside the green polygon. BCG colours are given as a blue dot. In addition we plot in red all the galaxies which have photometric redshift estimate $0.35 < z_\mathrm{phot} < 0.55$.

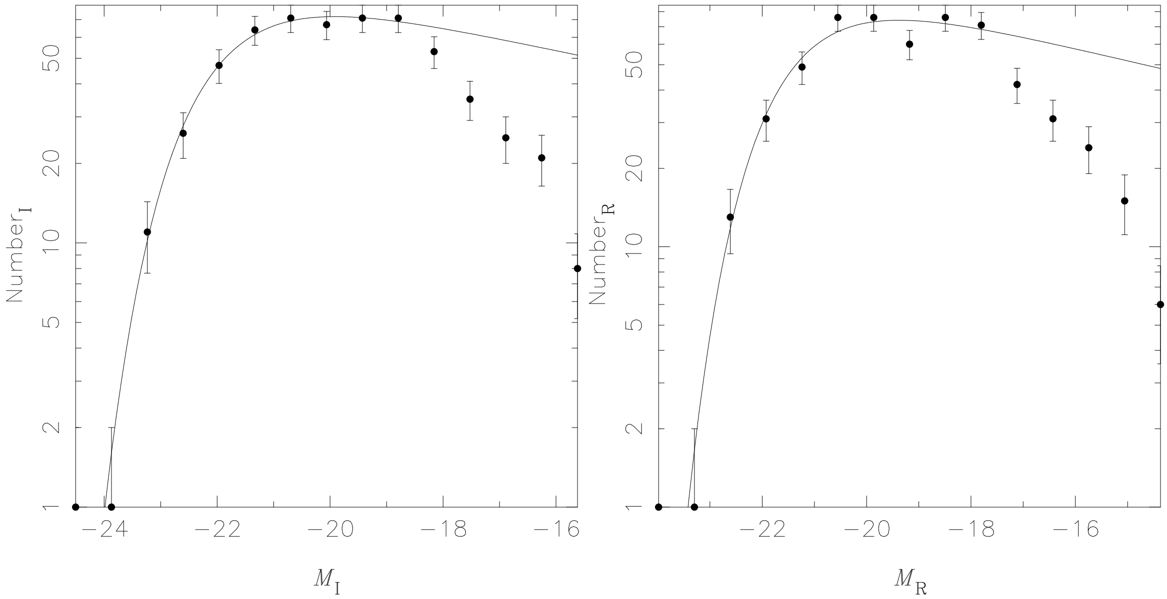


Figure 6.9: The galaxy luminosity function for the rest-frame I-band **(left)** and R-band **(right)** absolute magnitude for the cluster members selected using the colour cuts described in the text. We fitted (using `Mathematica`) the Schechter Luminosity function (Schechter, 1976) to the data, thereby excluding counts fainter than $M = -18$ in both bands and the BCG. The resulting characteristic magnitudes are $M^*_\mathrm{I} = -22.2$ and $M^*_\mathrm{R} = -21.7$ and the faint end slopes are given by $\alpha_\mathrm{I} = -0.9$, and $\alpha_\mathrm{R} = -0.9$.

7

Outlook

In this thesis we have studied how gravitational lensing can be used to determine the total (i.e. luminous and dark) mass distribution on different scales. We have shown how the small-scale structure in galaxies influences the observables in multiply imaged systems and how the combination of weak and strong lensing can be used to determine the mass and mass distribution in clusters of galaxies. Both these probes can be used to directly test the predictions from the ΛCDM paradigm. Our main findings have been summarised at the very beginning of this work, here we give an outlook and describe future plans.

The prediction of flux ratio anomalies and cusp relation violation has been observed in several systems and subsequently tested in great detail using simulations and semi-analytic models. There are other signatures of substructures, such as bending of jets and saddle point demagnification, for which observational evidence is present. We argue that these signatures of mass-substructure can best be tested using N-body simulations. However, a detailed numerical analysis has not yet been done. We plan to use higher-resolution N-body simulations which include baryons, to further test these predictions and therewith test the ΛCDM paradigm on these small scales.

Currently strong gravitational lensing is the only tool to directly detect possible presence of CDM substructure at cosmological distances. Unfortunately, only a handful of lenses are known which have the correct geometry and data quality (in radio frequencies) for the substructure to be studied. Therefore, the statistical studies are currently limited and comparisons with simulations difficult. The planned facilities, such as SKA, will change the situation dramatically by discovering multiply imaged systems in a daily pace. Not only large samples, but also the high resolution and monitoring capabilities of SKA will greatly improve the current status of strong-lensing research.

Turning to the second part of this work, in the two chapters describing combined strong and weak lensing mass reconstruction we have listed many points where improvements can be made. Using the HST ACS excellent imaging properties one can obtain deep high-resolution images of individual arcs (c.f. Fig. 1.2). We plan to further improve the method to include this information. In addition, wide-field imaging data is required for the weak lensing analysis of the cluster RX J1347$-$1145 to measure its mass at larger scales. Ultimately, the signals from different clusters need to be combined and e.g. the cluster mass function as a function of redshift can be obtained, which is a probe of the dark energy.

Future wide-field missions will therefore open a new window in strong lensing, weak lensing and cosmology. Two planed missions, namely LSST and SNAP, will produce complementary high-quality wide-field data especially suitable for weak and strong lensing analyses. The advantage of the ground-based 8-meter class telescope LSST over SNAP is that it will have larger field-of-view and faster photometry. On the other hand SNAP main advantage is high imaging quality, which is not limited by atmospheric seeing and therefore promises reduced and controlled systematics. In the future we plan to work on simulating the strong and weak lensing data and developing techniques particularly suitable for these extremely promising future missions.

Bibliography

Abadi, M. G., Navarro, J. F., Steinmetz, M., & Eke, V. R. 2003a, ApJ, 591, 499

Abadi, M. G., Navarro, J. F., Steinmetz, M., & Eke, V. R. 2003b, ApJ, 597, 21

Abdelsalam, H. M., Saha, P., & Williams, L. L. R. 1998, AJ, 116, 1541

Abramowitz, M. & Stegun, I. A. 1972, Handbook of Mathematical Functions (Handbook of Mathematical Functions, New York: Dover, 1972)

Allen, S. W., Schmidt, R. W., & Fabian, A. C. 2002, MNRAS, 335, 256

Barber, B. C., Dobkin, D. P., & Huhdanpaa, H. 1996, ACM Transactions on Mathematical Software, 22, 469

Barkana, R. 1998, ApJ, 502, 531

Bartelmann, M. & Narayan, R. 1995, ApJ, 451, 60

Bartelmann, M., Narayan, R., Seitz, S., & Schneider, P. 1996, ApJ, 464, L115

Benítez, N. 2000, ApJ, 536, 571

Benson, A. J., Frenk, C. S., Lacey, C. G., Baugh, C. M., & Cole, S. 2002, MNRAS, 333, 177

Bertin, E. & Arnouts, S. 1996, A&ASS, 117, 393

Blandford, R. & Narayan, R. 1986, ApJ, 310, 568

Blandford, R. D. 1990, QJRAS, 31, 305

Blandford, R. D. & Kochanek, C. S. 1987, ApJ, 321, 658

Blandford, R. D., Saust, A. B., Brainerd, T. G., & Villumsen, J. V. 1991, MNRAS, 251, 600

Bode, P., Ostriker, J. P., & Turok, N. 2001, ApJ, 556, 93

Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, A&A, 363, 476

Bradač, M., Lombardi, M., & Schneider, P. 2004a, A&A, 424, 13

Bradač, M., Schneider, P., Erben, T., & Lombardi. 2004b, in prep.

Bradač, M., Schneider, P., Erben, T., et al. 2004c, in prep.

Bradač, M., Schneider, P., Lombardi, M., et al. 2004d, A&A, 423, 797

Bradač, M., Schneider, P., Steinmetz, M., et al. 2002, A&A, 388, 373

Brainerd, T. G., Blandford, R. D., & Smail, I. 1996, ApJ, 466, 623

Bridle, S. L., Hobson, M. P., Lasenby, A. N., & Saunders, R. 1998, MNRAS, 299, 895

Broadhurst, T. J., Taylor, A. N., & Peacock, J. A. 1995, ApJ, 438, 49

Bullock, J. S., Kravtsov, A. V., & Weinberg, D. H. 2000, ApJ, 539, 517

Chen, J., Kravtsov, A. V., & Keeton, C. R. 2003, ApJ, 592, 24

Chiba, M. 2002, ApJ, 565, 17

Clowe, D. & Schneider, P. 2001, A&A, 379, 384

Clowe, D. & Schneider, P. 2002, A&A, 395, 385

Cohen, J. G. & Kneib, J. 2002, ApJ, 573, 524

Cohn, J. D., Kochanek, C. S., McLeod, B. A., & Keeton, C. R. 2001, ApJ, 554, 1216

Colín, P., Avila-Reese, V., & Valenzuela, O. 2000, ApJ, 542, 622

Courbin, F. & Minniti, D., eds. 2002, Gravitational Lensing: An Astrophysical Tool

Dalal, N. & Kochanek, C. S. 2002, ApJ, 572, 25

Davis, T. A. & Duff, I. S. 1999, ACM Trans. Math. Softw., 25, 1

de Blok, W. J. G. & Bosma, A. 2002, A&A, 385, 816

Erben, T., Van Waerbeke, L., Bertin, E., Mellier, Y., & Schneider, P. 2001, A&A, 366, 717

Ettori, S., Allen, S. W., & Fabian, A. C. 2001, MNRAS, 322, 187

Ettori, S., Tozzi, P., Borgani, S., & Rosati, P. 2004, A&A, 417, 13

Evans, N. W. & Hunter, C. 2002, ApJ, 575, 68

Evans, N. W. & Witt, H. J. 2003, MNRAS, 345, 1351

Falco, E. E., Gorenstein, M. V., & Shapiro, I. I. 1985, ApJ, 289, L1

Fischer, P. & Tyson, J. A. 1997, AJ, 114, 14

Fort, B., Mellier, Y., & Dantel-Fort, M. 1997, A&A, 321, 353

Frigo, M. & Johnson, S. 1998, in 1998 ICASSP conference proceedings, Vol. 3, 1381

Fruchter, A. S. & Hook, R. N. 2002, PASP, 114, 144

Gaudi, B. S. & Petters, A. O. 2002, ApJ, 574, 970

Geiger, B. & Schneider, P. 1998, MNRAS, 295, 497

Gilmore, R. 1981, Catastrophe theory for scientists and engineers (Dover, New York)

Grogin, N. A. & Narayan, R. 1996, ApJ, 464, 92

Hayashi, E., Navarro, J. F., Taylor, J. E., Stadel, J., & Quinn, T. 2003, ApJ, 584, 541

Heyl, J. S., Hernquist, L., & Spergel, D. N. 1994, ApJ, 427, 165

Hoekstra, H., Franx, M., Kuijken, K., & Squires, G. 1998, ApJ, 504, 636

Ibata, R. A., Lewis, G. F., Irwin, M. J., Lehár, J., & Totten, E. J. 1999, AJ, 118, 1922

James, F. & Roos, M. 1975, Computer Physics Communications, 10, 343

Kaiser, N. 1995, ApJ, 439, L1

Kaiser, N. & Squires, G. 1993, ApJ, 404, 441

Kaiser, N., Squires, G., & Broadhurst, T. 1995, ApJ, 449, 460

Keeton, C. 2001, astro-ph/0112350

Keeton, C. R. 2003, ApJ, 584, 664

Keeton, C. R., Gaudi, B. S., & Petters, A. O. 2003, ApJ, 598, 138

Kitayama, T., Komatsu, E., Ota, N., et al. 2004, PASJ, 56, 17

Kleinheinrich, M. 2003, Dark Matter halos of galaxies studied with weak gravitational lensing, PhD Thesis, University of Bonn

Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999, ApJ, 522, 82

Kneib, J., Hudelot, P., Ellis, R. S., et al. 2003, ApJ, 598, 804

Kneib, J.-P., Ellis, R. S., Smail, I., Couch, W. J., & Sharples, R. M. 1996, ApJ, 471, 643

Kochanek, C. S. 1995, ApJ, 445, 559

Kochanek, C. S. 2004, astro-ph/0407232

Kochanek, C. S. & Dalal, N. 2004, ApJ, 610, 69

Komatsu, E., Matsuo, H., Kitayama, T., et al. 2001, PASJ, 53, 57

Koopmans, L. & de Bruyn, A. 2000, A&A, 358, 793

Koopmans, L. V. E., Biggs, A., Blandford, R. D., et al. 2003, ApJ, 595, 712

Koopmans, L. V. E. & Treu, T. 2003, ApJ, 583, 606

Kormann, R., Schneider, P., & Bartelmann, M. 1994, A&A, 284, 285

Kravtsov, A. V., Klypin, A. A., Bullock, J. S., & Primack, J. R. 1998, ApJ, 502, 48

Lenzen, F., Schindler, S., & Scherzer, O. 2004, A&A, 416, 391

Lombardi, M. 2000, Ph.D. Thesis

Lombardi, M. & Bertin, G. 1999, A&A, 342, 337

Lombardi, M. & Schneider, P. 2002, A&A, 392, 1153

Möller, O., Hewett, P., & Blain, A. W. 2003, MNRAS, 345, 1

Mao, S. 1992, ApJ, 389, 63

Mao, S. & Schneider, P. 1998, MNRAS, 295, 587

Maoli, R., Van Waerbeke, L., Mellier, Y., et al. 2001, A&A, 368, 766

Marshall, P. J., Hobson, M. P., Gull, S. F., & Bridle, S. L. 2002, MNRAS, 335, 1037

Metcalf, R. 2002, ApJ, 580, 696

Metcalf, R. B. & Madau, P. 2001, ApJ, 563, 9

Metcalf, R. B. & Zhao, H. 2002, ApJ, 567, L5

Meza, A., Navarro, J. F., Steinmetz, M., & Eke, V. R. 2003, ApJ, 590, 619

Monaghan, J. 1992, ARA&A, 30, 543

Monet, D. B. A., Canzian, B., Dahn, C., et al. 1998, VizieR Online Data Catalog, 1252, 0

Moore, B. 1994, Nature, 370, 629

Moore, B., Ghigna, S., Governato, F., et al. 1999, ApJ, 524, L19

Narayan, M. & Bartelmann, M. 1997, in Proceedings of the 1995 Jerusalem Winter

School, eds. A. Dekel and J.P. Ostriker.

Navarro, J. & Steinmetz, M. 1997, ApJ, 478, 13

Navarro, J. F., Hayashi, E., Power, C., et al. 2004, MNRAS, 349, 1039

Petters, A. O., Levine, H., & Wambsganss, J. 2001, Singularity theory and gravitational lensing (Singularity theory and gravitational lensing / Arlie O. Petters, Harold Levine, Joachim Wambsganss. Boston : Birkhäuser, c2001. (Progress in mathematical physics ; v. 21))

Poggianti, B. M. 1997, A&ASS, 122, 399

Pointecouteau, E., Giard, M., Benoit, A., et al. 2001, ApJ, 552, 42

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical recipes in C. The art of scientific computing (Cambridge: University Press, 1992, 2nd ed.)

Quadri, R., Möller, O., & Natarajan, P. 2003, ApJ, 597, 659

Ravindranath, S. & Ho, L. C. 2002, ApJ, 577, 133

Refsdal, S. 1964, MNRAS, 128, 307

Sahu, K. C., Shaw, R. A., Kaiser, M. E., et al. 1998, ApJ, 492, L125

Schaap, W. E. & van de Weygaert, R. 2000, A&A, 363, L29

Schechter, P. 1976, ApJ, 203, 297

Schechter, P. L. & Wambsganss, J. 2002, ApJ, 580, 685

Schindler, S., Guzzo, L., Ebeling, H., et al. 1995, A&A, 299, L9

Schindler, S., Hattori, M., Neumann, D. M., & Boehringer, H. 1997, A&A, 317, 646

Schirmer, M., Erben, T., Schneider, P., et al. 2003, A&A, 407, 869

Schneider, P., Ehlers, J., & Falco, E. 1992, Gravitational Lenses (Gravitational Lenses, Springer-Verlag Berlin Heidelberg New York.)

Schneider, P., King, L., & Erben, T. 2000, A&A, 353, 41

Schneider, P. & Seitz, C. 1995, A&A, 294, 411

Schneider, P. & Weiss, A. 1992, A&A, 260, 1

Seitz, C. & Schneider, P. 1997, A&A, 318, 687

Seitz, S., Schneider, P., & Bartelmann, M. 1998, A&A, 337, 325

Smith, G. P., Kneib, J., Smail, I., et al. 2004, astro-ph/0403588

Somerville, R. S. 2002, ApJ, 572, L23

Spergel, D. N. & Steinhardt, P. J. 2000, Physical Review Letters, 84, 3760

Springel, V. & Hernquist, L. 2003, MNRAS, 339, 289

Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, MNRAS, 328, 726

Steinmetz, M. 1996, MNRAS, 278, 1005

Steinmetz, M. & Navarro, J. 2000, in ASP Conf. Ser. 197: Dynamics of Galaxies: from the Early Universe to the Present, eds. F. Combes, G. A. Mamon, and V. Charmandaris., 165

Steinmetz, M. & Navarro, J. F. 2003, New Astronomy, 8, 557

Stoehr, F., White, S. D. M., Tormen, G., & Springel, V. 2002, MNRAS, 335, L84

Swaters, R. A., Madore, B. F., & Trewhella, M. 2000, ApJ, 531, L107

Taylor, A. N., Dye, S., Broadhurst, T. J., Benitez, N., & van Kampen, E. 1998, ApJ, 501, 539

Treu, T. & Koopmans, L. V. E. 2002a, ApJ, 575, 87

Treu, T. & Koopmans, L. V. E. 2002b, MNRAS, 337, L6

Treu, T. & Koopmans, L. V. E. 2004, astro-ph/0401373

Tyson, J. A., Wenk, R. A., & Valdes, F. 1990, ApJ, 349, L1

van den Bosch, F. C. & Swaters, R. A. 2001, MNRAS, 325, 1017

Wallington, S. & Narayan, R. 1993, ApJ, 403, 517

Witt, H. J., Mao, S., & Schechter, P. L. 1995, ApJ, 443, 18

Woźniak, P. R., Udalski, A., Szymański, M., et al. 2000, ApJ, 540, L65

Zentner, A. R. & Bullock, J. S. 2003, ApJ, 598, 49

Acknowledgements

I am deeply indebted to the following people for various expressions of kindness, patience, support, help, care, cookies and fun:

Peter Schneider, for letting me go the way I wanted to take and being there at any cross-section to make sure I would not get lost. His ideas were always there when I most needed them. I would also like to thank him for carefully reading my manuscripts and thesis, and this is certainly not intended to sound like any other “we would like to thank the referee...” statement. The work behind correcting my writings is enormous and I am truly thankful to him for that.

The rest of my thesis committee, starting with Pavel Kroupa for agreeing to it in the first place, despite his busy schedule. I enjoyed many little discussions we had in the corridor. Many thanks also to Norbert Wermes for introducing me to the subject of astroparticle physics in the first seminar I took in Bonn, and Hermann Karcher for our cheerful discussion about dark matter and related problems.

There are so many great things about our institute I don’t even know where to start; the lack of hierarchy, the open doors, the never-ending conversations, the lack of a canteen, but most of all the people. Two of them have shared a very special place on the third floor with me. I therefore thank Marco Lombardi for keeping me in a good mood and for helping me with the first steps in the world of real science. I enjoyed it a lot, even with the side effects of rotten bananas. I learned a lot from him (Thank you Sir!). And Oliver Czoske for patiently sharing an office with me and all my sport belongings and helping me on many occasions.

Thomas Erben, Mischa Schirmer, Hendrik Hildebrandt, Douglas Clowe, and Jörg Dietrich have helped me with data and answered all my impossible questions. Most of all they contributed to the respect I gained for those little pixels. Anja von der Linden and Tim Schrabback have convinced me that Bonn is a nice place to be; I came for a semester and stayed for eleven of them. The number-of-cookies-I-stole times thanks for that (yours truly, cookie monster). I would like to thank Marco Hetterscheidt for showing me the ways through Siebengebierge (although not always the correct ones) and organising great parties at his place (Fried und Kriegen

rules). Peter “Footy” Watts for being a cheerful chap, always willing to help, and preparing the best Sunday lunch I ever had in Bonn. Thanks for many great ultimate-frisbee games, I enjoyed them a lot. Richard Porcas for being the official “shoulder to cry on”. The whole Lensing Group et. al. (including Lensing Group emeritus) for numerous interesting discussions during the lensing seminars. They all often showed me the way out of a big mess. Last but not least, thank-yous go to our lovely secretary Kathy (Houston, we have a problem), and our walking computer manuals Oliver Cordes and Ole Marggraf whose patience is truly unlimited.

It’s amazing what goes into a thesis. There’s Phil and Anže’s invitation for a weekend in Cambridge (needles to say none of the science was done there), there are friends back in Slovenia sharing my skiing enthusiasm. Saška, who reminded me again how much I like basketball. And Tommaso, who always managed to answer questions like what to eat at Maha’s and what to do with my life with the same ease. With him around, the world all of the sudden becomes a lot nicer place.

And then there is someone special. Léon Koopmans definitely deserves an award for his patience, understanding, and love. I think me finishing my thesis must have been tougher for him than for me. His love and caring made this entire trouble worthwhile. And finally my parents and my brother who never failed in their support. And it is their “fault” that the best place on Earth will for me always be the one where I grew up.