Study of two radio gravitational lenses: insight into the high-redshift Universe and properties of mass distribution

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Filomena Volino

aus

Potenza, Italia

Bonn 2011

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

1. Referent: Dr. Olaf Wucknitz
2. Referent: Prof. Dr. Peter Schneider

Tag der Mündlichen Prüfung: 07 September 2011

Alla mia famiglia:
mamma, papà e
Piercarmine

Contents

List of Tables

List of Figures

Constants and quantities used throughout

constant/quantity	value
speed of light	$c = 2.99792458 \times 10^8$ m s^{-1}
gravitational constant	$G = 6.67 \times 10^{-11}$ m^3 kg^{-1} s^{-2}
solar mass	$M_\odot = 2 \times 10^{30}$ kg
solar luminosity	$L_\odot = 3.839 \times 10^{26}$ Watt
Jansky	Jy= 10^{-26} Watts m^{-2} Hz^{-1}
parsec	pc= $3.08568025 \times 10^{16}$ m

Overview

Similarly to a magnifying glass, a gravitational field deflects light rays that travel through it. This phenomenon, predicted by the theory of General Relativity by A. Einstein (1915), is referred to as gravitational lensing, and it is the core of the projects presented in this thesis. The study of gravitational lenses yields the mass distribution of the lensing object together with the properties of the background source. We have studied two of these systems using radio interferometers, achieving angular resolution within the range $0.001 - 4$ arcseconds. The first system is the 8 o'clock arc (J002240+143130). Main goal is the detection of radio emission from a high redshift source exploiting lensing as *natural telescope*; eventually the properties of the intervening mass could then be studied from the observed lensed configuration. The second system is the gravitational lens MG J0414+0534. The goal is to deduce the properties of the foreground mass distribution and the background source.

In Chapter 1 we describe the theoretical astrophysical framework on which is based the work presented in this thesis. We start describing our current understanding of cosmic expansion, structure formation and evolution in the Universe. Afterwards we describe the theory of gravitational lensing providing the basic concepts and definitions necessary to analyse a gravitational lens. Next we review some important recent results from the field, and highlight the astrophysical applications of gravitational lensing that are relevant for the projects discussed here.

In Chapter 2 we continue with lensing and describe first the analytical mass models used for mass modelling, then the techniques used to study the properties of the foreground mass distribution and the background source. Then we explain the procedure used to estimate the uncertainty on the lens model parameters and describe the set of constraints that a lensed configuration provides, this last aspect is important in order to interpret the goodness of fit of lens model. Afterwards we also discuss the main degeneracies to consider when interpreting the results from lens modelling.

In Chapter 3 we start describing the principles of radio-interferometry, calibration strategies and methods to reconstruct images of the radio sky. We discuss issues like smearing and sensitivity, that are relevant for lensing studies from radio observations, and in particular for the systems studied here. We conclude the chapter giving more details on the imaging technique which we have used to produce the maps of both systems presented in this work.

In Chapter 4 we present the 8 o'clock arc system. Our target is the high redshift galaxy ($z \sim 2.73$) J002240.78+143113.9, which is gravitationally lensed by a galaxy at $z \sim 0.38$. The system was first discovered at optical wavelengths (Allam et al., 2007), and later on studied also in the near-infrared (Finkelstein et al., 2009); both

studies have confirmed this system as an active star-forming region. In the NVSS catalogue (The National Radio Astronomy Observatory Very Large Array Sky Survey; Condon et al. 1998) the system is associated with a radio source of 5 mJy, which would predict a radio-derived star-formation rate more then one order of magnitude larger than what predicted by the existing studies. In order to investigate the source of the 5 mJy radio emission we have conducted radio-interferometric observations at 1.4 and 5 GHz with the Very Large Array (VLA) telescope. In this chapter we present results from the VLA observations. Detailed and accurate studies of these early episodes of star-formation are only possible when the emission is boosted by a gravitational lens. In particular, radio observations of these targets are becoming feasible now that the sensitivity of the radio telescopes has been improved by at least one order of magnitude. We find that most of the radio emission is associated with the lens galaxy. From our study we could place only an upper limit on the radio flux density and hence the radio-derived star-formation rate, which we find in good agreement with the predictions from optical and near-infrared observations.
In Chapter 5 we present the study of the gravitational lens MG J0414+0534. The lensed source is a QSO at redshift $z \sim 2.64$, while the elliptical lens galaxy is at redshift $z \sim 0.96$; the latter has a luminous companion, whose nature is probably a dwarf galaxy. We have conducted new global-VLBI observations at 1.7 GHz in order to map with high sensitivity the emission of the extended structures and hence obtain a rich set of constraints for lens-modelling purposes. In this chapter we present the new VLBI maps and lens-modelling results. The system is a very well know radio-loud lens. This type of lenses are the ideal laboratories where the presence of CDM structures can be directly revealed by independent bent of the radio jet or flux-ratios mismatches between the observed configuration and the prediction of a smooth mass model that does not account for such a mass structure. Indeed for this system, mid-infrared observations have reported the presence of a flux-anomaly which could indicate the presence of a dark-matter clump (Minezaki et al., 2009). Previous Very Long Baseline Interferometric (VLBI) observations have shown four resolved lensed images, separated by up to 2 arcseconds (Ros et al., 2000; Trotter et al., 2000). Deep optical observations, obtained with the Hubble Space Telescope, shows the four compact images of the lensed QSO, the lens galaxy and its luminous satellite, as well as an extended structure, probably connected to the QSO (Kochanek et al., 1999). We first use the new data to test the existing lens models, afterwards we constrain the model with a modelling technique that uses all the structures seen in the radio maps; this is a new modelling approach for this system and it also allows us to reconstruct the unlensed source structure. The source reconstruction obtained with our best mass model has shifts of ~ 5 milli-arcseconds that could be explained by a non-smooth mass distribution. This would also be consistent with the flux-anomaly problem reported by mid-infrared studies. At the end, we discuss if these data are sensitive to probe other mass models than the one we have considered.
To conclude, in Chapter 6 we summarize the research projects we have carried out and describe the possible future work.
In the appendix we describe the model fitting techniques used to describe the features seen in the radio maps.

Imagination is more important
than knowledge

A. Einstein

1 Introduction

In this chapter the theoretical astrophysical framework for this thesis is introduced. In Sections 1.1 and 1.2 we discuss the aspects of relativistic cosmology that are relevant for the science cases presented here. In Section 1.3 and 1.4 we describe the theory of gravitational lensing and some of its astrophysical applications.

1.1 The expanding Universe

While the first scientific results from the Planck mission have been recently delivered, new results are produced by the Wilkinson Microwave Anisotropy Probe (WMPA) satellite, which place tighter constraints on the currently most widely accepted standard model of cosmology, namely the *Cold Dark Matter and Cosmological Constant in a flat Universe*.
The cosmological principle, the Weyl postulate[1] and the framework provided by the theory of General Relativity gravity build the theoretical framework for this model that explains how the Universe has evolved since the primordial hot and dense phase of the Big Bang, which occurred ~ 13.7 billion years ago. Since then, as time proceeded, the small inhomogeneities that were set by fluctuations of the quantum field were amplified by gravitational instabilities, leading to the formation of the large-scale structures in the Universe that we observe today.
Hubble's experiment in 1929 first opened the picture of an *expanding Universe*. In 1959, observations of the spiral galaxy M33 showed its rotation curve not consistent with Keplerian dynamic (Volders, 1959), evidence found later on also for many other galaxies (Rubin et al., 1978). Observations were implying an amount of a non-luminous matter, named *dark matter* extending far from the centre of the galaxy. Already Zwicky in 1933 had evidence for it while studying the Coma galaxy cluster.

[1] An intuitive interpretation of this postulate would allow us to see as cosmological models only those ones given by hyperslices which are are everywhere orthogonal to the world lines of the cosmological fluid particles. In other terms, this postulate states how space is evolving over time, and guarantees the measurement of time.

Afterwards, the discovery of the Cosmic Microwave Background (CMB) by Penzias & Wilson in 1965 as relic radiation from the Big Bang, provided the first evidence of the early Universe. Finally, another important piece for the whole picture has come by observations of distant type Ia supernovae (Conley et al. 2011 and references therein) which have indicated that the Universe is undergoing an accelerated expansion at present times, explained by a *dark energy* component, acting as a negative pressure.
The WMAP satellite has studied in great detail the power spectrum of the primordial inhomogeneities, providing an enormous amount of information on all cosmological parameters showing no deviations from the currently accepted model that, although requiring its main ingredient being *dark*, seems to be the only one in agreement with the observational evidences achieved in the last 10−15 years. Among them, beside what is mentioned above, they need to be mentioned here the statistical properties of the large-scale structures that have been studied in large redshift surveys like the Sloan Digital Sky Survey (SDSS) and the Two Degree Field Galaxy redshift Survey (2dFGRS) (Padmanabhan et al., 2007; Hawkins et al., 2003), the spectra of high-redshift quasars (Tytler et al., 2000), and clusters studies (Eke et al., 1996) that allow one to constrain the primordial abundance of light elements and thus the baryon and total matter content of the Universe.
The latest results the WMAP mission are given in Larson et al. (2011) and Komatsu et al. (2011). Dark Energy constitutes 72.1% of the Universe energy density, Dark matter 23.3% and the remaining 4.6% is in form of baryonic matter. Other methods, like Baryonic Acoustic Oscillations (e.g. Zhai et al. 2010; Eisenstein et al. 2005), time delays measurements from gravitational lensing systems (Suyu et al., 2010) and cosmic shear measurements are all complementary to break the degeneracies between the cosmological parameters, and thus in our understanding of the components of the Universe in which we live.

1.1.1 Dynamics of the expanding Universe

As seen from the Earth, when smoothed on large scales, our Universe is isotropic. Furthermore, accepting the Copernican principle, the same property holds also for other observers. Thus we conclude that on large scales the Universe looks the same in all directions for an observer at any place. This statement is usually referred to as *cosmological principle* that in mathematical terms translate that the Universe is homogeneous and isotropic on large scales. From symmetry considerations it follows that the metric in such a Universe can be written as

$$ds^2 = c^2dt^2 - a^2(t)\left[d\chi^2 + f_k^2(\chi)\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right]. \tag{1.1}$$

Independently, Robertson (1935) and Walker (1936) demonstrated that this is the general form for the line element in a spatially homogeneous and isotropic space-time. It was used for the first time by Friedmann in 1922, and therefore is usually called the Friedmann-Robertson-Walker (FRW) metric. In Equation (1.1) c is the speed of light, $a(t)$ is a time-dependent scale factor; the coordinates θ and ϕ are angular coordinates, while χ is the coordinate in the radial direction; t is the time coordinate has measured by a set of comoving observers at coordinates (χ, θ, ϕ), which are called *comoving coordinates*. The factor $f_k(\chi)$ depends upon the geometry

of the Universe, which depends on the curvature parameter k. For a flat Universe (suggested by the observations of the CMB) $k = 0$, the factor $f_0(\chi) = \chi$ and the metric equation is

$$ds^2 = c^2dt^2 - a^2(t)\left[d\chi^2 + \chi^2\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right]. \tag{1.2}$$

At a given time t, a *comoving distance* is the distance between comoving observers. The *physical distance* between two points in space is $x = \chi a(t)$ [2].
In the theory of General Relativity, the metric is related to the matter distribution in the Universe by the Einstein field equations:

$$G_{\mu\nu} - \Lambda g_{\mu\nu} = \frac{8\pi G}{c^2}T_{\mu\nu}. \tag{1.3}$$

In the above equation, G is the gravitational constant, $G_{\mu\nu}$ is the Einstein tensor, representing the metric, while $T_{\mu\nu}$ is the energy-momentum tensor.The parameter Λ is the cosmological constant. The labels μ and ν run over the time index (0) and the three spatial indices (1,2,3).
The homogeneity and isotropy of the Universe implies that the matter contents must be a perfect fluid with density $\rho(t)$ and pressure $p(t)$. Inserting the metric equation (Eq. 1.2) in the field equations (Eq. 1.3), the components of Equation (1.3) reduce to two equations for the scale factor $a(t)$,

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda}{3} \tag{1.4}$$

and

$$\frac{\ddot{a}}{a} = -\frac{4}{3}\pi G\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda}{3}. \tag{1.5}$$

These are called *Friedmann's equations*. Equations (1.4) and (1.5) can be combined to give the *adiabatic equation* that expresses the first law of thermodynamics in the cosmological context,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[a^3(t)\rho(t)c^2\right] + p(t)\frac{da^3(t)}{dt} = 0. \tag{1.6}$$

Equation (1.6) says that the expansion of the Universe must be adiabatic, thus the total entropy does not change.
A third equation, the *equation of state* is now necessary to fully describe how the Universe expands, [3]

$$p = w\rho c^2, \tag{1.7}$$

which relate the pressure and energy density of the components of the Universe. The parameter w is dimensionless, and assumes different values for each of the components of the Universe, that is $w = 0$ for non-relativistic matter, $w = 1/3$ for radiation and relativistic matter, $w = -1$ for the vacuum energy component. The total energy density of the Universe is then the sum of the three components,

$$\rho_t = \rho_m + \rho_r + \rho_\Lambda, \tag{1.8}$$

[2] assuming for simplicity $\theta = 0$ and $\phi = 0$
[3] Note that of among the three equations (1.4), (1.5) and (1.6) only two of them are independent

with $\rho_\Lambda = \Lambda/8\pi G$ and ρ_m, and ρ_r refer to matter and radiation. With this decomposition, the Friedmann equations, (Eq. (1.4) and (1.5)), can be used without the Λ term, and its density and pressure are included in the variables $\rho(t)$ and $P(t)$. Combining Equations (1.7) and (1.6) gives the matter energy density falling as a^{-3}, consistent with the conservation of number of particles, and the radiation energy density as a^{-4}, which reflects the conservation of number of particles and the decrease of the energy density due to the expansion. Note that for $w = -1$, the energy density ρ stays constant with the expansion of the Universe. The ratio

$$\frac{\dot{a}}{a} \equiv H(t) \tag{1.9}$$

expresses the expansion rate of the Universe, and is called the *Hubble parameter*. For $\Lambda = 0$ and $k = 0$ (vanishing spatial curvature), there is a *critical density* whose value depends on the Hubble parameter,

$$\rho_{\rm crit}(t) = \frac{3H^2(t)}{8\pi G}. \tag{1.10}$$

This value is conveniently used to define the mean energy density of each of the components of the Universe as

$$\Omega_i(t) = \frac{\rho_i(t)}{\rho_{\rm crit}(t)} \tag{1.11}$$

Combining Equations (1.4) and (1.8), and using the definitions given in Equation (1.11), the expansion equation becomes

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0^2 \left[\frac{\Omega_r}{a^4} + \frac{\Omega_m}{a^3} + \frac{Kc^2}{a^2 H_0^2} + \Omega_\Lambda\right] = H^2(t). \tag{1.12}$$

At the current time, $H(t) = H_0$ and $a = 1$, Equation (1.12) yields an expression for the curvature of the Universe in terms of its total density today,

$$k = (\Omega_r + \Omega_m + \Omega_\Lambda - 1)\frac{H_0^2}{c^2}. \tag{1.13}$$

Defining $\Omega_0 = \Omega_r + \Omega_m + \Omega_\Lambda$, from the above equation it follows that the sign of $\Omega_0 - 1$ agrees with that of k, that is the total energy density determines the curvature of the Universe.

1.1.2 Light propagation in the expanding Universe

Photons of wavelength λ_1 and emitted at scale factor a_1 reach the observer today ($t = t_0$) with wavelength λ_0 related to the scale factor by

$$\frac{\lambda_0}{\lambda_1} = \frac{1}{a_1}. \tag{1.14}$$

In the expanding Universe, the observed wavelength (λ_0) is longer than λ_1; this is expressed in terms of the *redshift parameter* z, defined as

$$z = \frac{\lambda_0 - \lambda_1}{\lambda_1}. \tag{1.15}$$

It follows that the scale factor can be written as function of redshift as

$$a = \frac{1}{1+z}. \tag{1.16}$$

Cosmological redshift is an observational evidence, seen for the first time in 1912 by V. M. Slipher while taking spectra of distant galaxies. At that time astronomers did not agree on the nature of these distant galaxies, which were referred to as *spiral nebulae*. The issue was debated in the Great Debate in 1920 between H. Shapley and H. Curtis, and solved only in 1929 with the work of E. Hubble.

1.1.3 Distances in the expanding Universe

Distance measures in cosmology are given along the backward light cone, and as the Universe expands distances change. There is not a unique meaning of distance, but methods can be specified on how to measure distances, and define distances according to the methods used to measure them.

Angular diameter distance

The physical size of an object, ξ_1, at redshift z_1 is related to its angular size, θ_1, as measured today ($z = 0$) as

$$\xi_1 = D_{\rm ang}(0, z_1)\theta_1. \tag{1.17}$$

Consider an observer at redshift $z = 0$ situated at the centre of a circle, which has radius given by the comoving distance χ_1. In a flat space, the angle subtended by the circle is 2π and its physical size at z_1 is $2\pi\chi_1 a_1$. Consider now an element on the circle ξ_1 with angular size θ_1, then

$$\frac{\xi_1}{2\pi\chi_1 a_1} = \frac{\theta_1}{2\pi}. \tag{1.18}$$

Comparing Equations (1.17) and (1.18), the *angular distance* of the circle as seen from $z = 0$ is

$$D_{\rm ang}(0, z_1) = a_1\chi_1. \tag{1.19}$$

In gravitational lensing (see Sect. 1.3) physical distances at different redshifts have to be converted into angles seen at a given redshift. Hence distances are conveniently expressed as *angular diameter distances*.
The angular diameter distance of a source at redshift z_2 as measured by an observer at redshift $z_1 < z_2$ is

$$D_{\rm ang}(z_1, z_2) = a_2(\chi_2 - \chi_1). \tag{1.20}$$

Luminosity distance

Another way to measure distances is to relate the flux S, observed today at redshift $z = 0$, of a source at redshift z_1, to its luminosity, L,

$$S = \frac{L}{4\pi D_{\rm l}^2} \tag{1.21}$$

where D_l is the *luminosity distance*.
The above two distance measurements are related by the following expression (derived by Etherington in 1933),

$$D_{\rm l}(z) = (1+z)^2 D_{\rm ang}(z). \tag{1.22}$$

The k-correction

The quantities S and L in Equation (1.21) refer to the flux and the luminosity integrated over frequencies. Because of redshift, the observed flux at frequency ν, S_ν, is related to the luminosity of the source at frequency $\nu_{\rm e} = (1+z)\nu$. The observed flux is

$$S_\nu = \frac{(1+z)L_{\nu_{\rm e}}}{4\pi D_{\rm l}^2}, \tag{1.23}$$

which can also be written as

$$S_\nu = \frac{L_\nu}{4\pi D_{\rm l}^2}\left[\frac{L_{\nu_{\rm e}}}{L_\nu}(1+z)\right], \tag{1.24}$$

where the first factor is the relation between the bolometric quantities while the second one represents the spectral shift. This is referred to as *k-correction*, and it depends on the spectrum of the source.

1.2 Structure formation

Observations show that the Universe is homogeneous on large scales, while this is not the case at smaller scales. We observe anisotropies imprinted in the CMB, and the distribution of galaxies in the sky is highly anisotropic as well. We believe that today we observe the evolution over cosmic time of density inhomogeneities, that have evolved from smaller fluctuations in the early Universe, through *gravitational instability*. The mathematical framework within which we study their evolution is provided by linearised equations of gravity, for small perturbations, and numerical methods, when linear approximation breaks down. Within the theory of structure formation we are able to study the evolution over cosmic time of these early fluctuations, whose growth depends on the matter content of the Universe and the nature of dark matter. The gravitationally bound structures in which galaxies and clusters form are the dark matter halos.
Below we introduce two questions regarding the properties of these structures which can be both investigated using gravitational lensing.

1.2.1 Halo profiles

Numerical simulations show that the halos density profiles seem to have a universal functional form, described by

$$\rho(r) = \frac{\rho_{\rm s}}{(r/r_{\rm s})(1+r/r_{\rm s})^2}, \tag{1.25}$$

where ρ_s is the amplitude of the density profile, and r_s specifies a characteristic radius. This result was first reported by Navarro et al. (1997). The inner part of

the density profile follows r^{-1}, whereas the outer part follows r^{-3}. No analytical argument has been found for the existence of such a profile, and comparison of these profiles with observations is not simple because the dark matter is not directly observable. However the rotation curves of low surface brightness galaxies provide no evidence of a cusp in the central distribution.

1.2.2 Halos abundance

The comparison of theoretical predictions with observations is necessary as it allows to test our understanding of structures formations. For Milky Way type halos, numerical simulations based on a ΛCDM cosmology predict a sub-halos population of hundreds of dark-matter satellites, that is one order of magnitude higher than the observed population of satellite galaxies in the Local Group (Diemand et al., 2007; Moore et al., 1999). This discrepancy is referred to as the *missing satellite problem* (Fig. 1.1). Until recently, one possibility of solving the discrepancy has been the interpretation that they indeed exist but star formation is suppressed, thus they they are not observable (Klypin et al., 1999). In the last few years, the number of faint satellites has increased, most of them being found from the Sloan Digital Sky Survey (SDSS; York et al. 2000), and the discrepancy now is within a factor of few of the predicted number (e.g. Belokurov et al. 2007; Zucker et al. 2006a,b). However, recent works based on high-resolution simulations (Lactea I and II, and Acquarious simulations; Madau et al. 2008; Springel et al. 2008) show discrepant results, implying that out understanding of the problem is still far from being clear. Being a pure gravity effect, gravitational lensing (see next section), provides us a powerful tool to study the matter distribution in our Universe and address the important question regarding the abundance of dark matter halos (see Sect. 1.4.1).

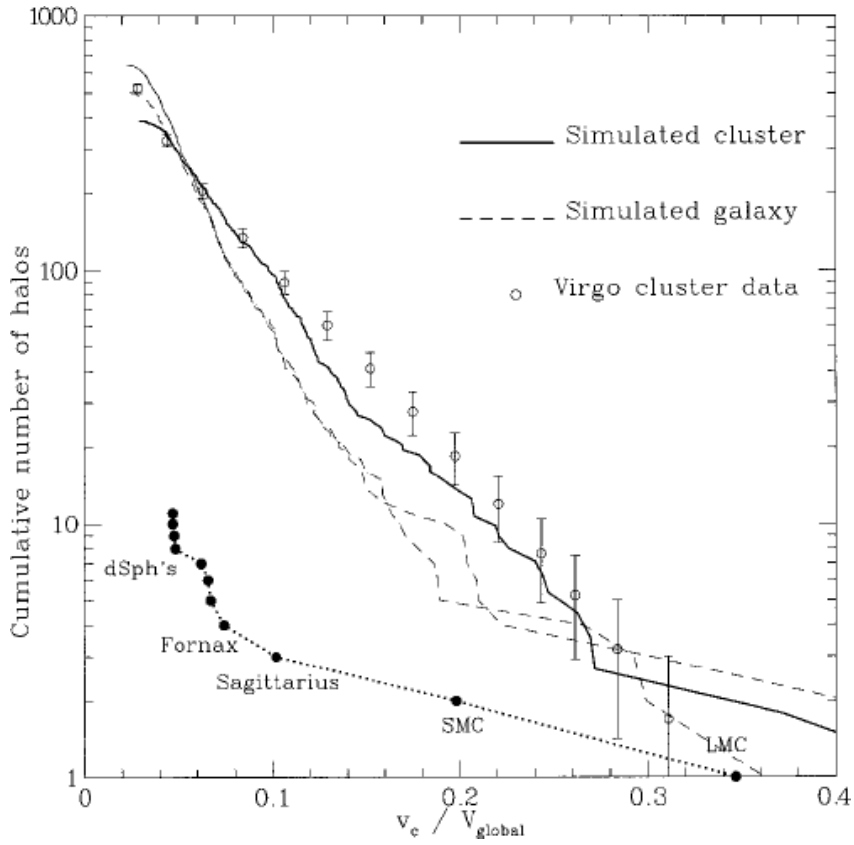


Figure 1.1: The plot shows the number density of sub-halos as function of their Keplerian velocity v_c for a cluster and galaxy mass. The velocity v_c is measured in units of the rotational velocity of the main halo. The open circles represent the number of sub-halos in the Virgo cluster, the filled ones the satellite galaxies in the Milky Way. The solid (dashed) curve shows the simulated cluster (galactic) mass halo (for the galactic halo there are two lines, at two different epochs). Adopted from Moore et al. (1999)

1.3 Theory of gravitational lensing

The theory of gravitational lensing describes how light rays from a background source propagate through a foreground mass distribution. Before reaching the observer, light rays are deflected by the gravitational potential along the line of sight, such that the source appears with a distorted shape. For a source whose distance in the sky from the line of sight of the potential is large, the effect of the lens potential is weak and one image is formed, whose shape is distorted. This is called the *weak lensing* regime that is observed for example from a background galaxy distribution emanating radiation through the large-scale structures of the Universe. On the other hand, when the distance in the sky between the source and the line of sight of the potential is small, the effect of the potential is higher and multiple images of the same background source may be formed. This is called *strong lensing* regime. The properties of the observed image configurations (which will be introduced in the next sections) carry information on the mass distribution that is causing the effect, thus this phenomenon provides an interesting tool to probe the mass distribution in the Universe; in particular in the cosmological framework it allows us to address questions related to the dark matter halos properties, small dark matter clumps and the large scale-structures as well.
For details on the different regimes and applications of the effect see Schneider et al. (1992). In this section the theory of the strong lensing regime is reviewed.

1.3.1 Basic equations of lens theory

The deflection angle

In Figure 1.2 is sketched the typical situation usually considered in gravitational lensing, that is the deflection of a light ray emitted from a (background) source when it travels through a foreground mass distribution. The *deflection angle* $\hat{\alpha}$ depends on the the mass distribution and on the impact parameter ξ. The angle $\hat{\alpha}$ can be written as

$$\hat{\alpha} = \frac{2}{c^2} \int \nabla_{\perp} \Phi \, \mathrm{d}l, \tag{1.26}$$

where the gradient of the Newtonian potential Φ is taken perpendicular to the light path and the integral is taken along the trajectory. Equation (1.26) relies on two assumptions: 1) the Newtonian potential is small (i.e. $\Phi \ll c^2$), 2) the lens has a small peculiar velocity (i.e. $v \ll c$). For a point-like mass M, the Newtonian potential is given by $\Phi(r_3, \xi) = -GM/\sqrt{\xi^2 + r_3^2}$, where r_3 is the distance component along the ray (see Fig. 1.2). Substituting the potential in Equation (1.26) gives the deflection angle of a point mass M,

$$\hat{\alpha} = \frac{4GM}{c^2 \xi}. \tag{1.27}$$

To calculate the deflection angle of a three-dimensional (extended) mass distribution the so-called *thin lens approximation* is made. The light ray is assumed to have a trajectory described by (ξ_1, ξ_2, r_3), where the coordinates are chosen such that the ray is propagating along r_3.We consider the deflected light ray, which is actually smoothly curved near the deflector, as a straight line. The distances of the

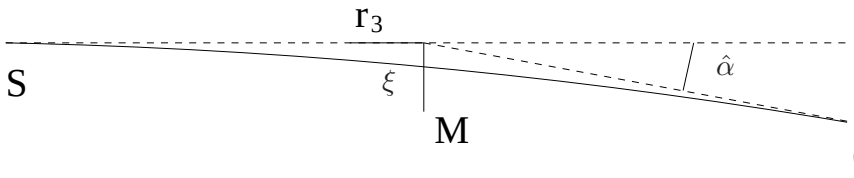


Figure 1.2: Light rays travelling from the source (S)to an observer (O) and being deflected by the mass M. The deflection corresponds to the angle $\hat{\alpha}$.

background source and the mass (i.e. the lens) from the observer are very large compared to extension of the lens. Under these assumptions[4] the deflection angle is given by integrating for each mass element and summing the deflection angles due to the mass elements at every ξ in the sky[5]:

$$\hat{\alpha}(\boldsymbol{\xi}) = \frac{4G}{c^2} \int \mathrm{d}^2\xi' \Sigma(\boldsymbol{\xi}') \frac{\boldsymbol{\xi} - \boldsymbol{\xi}'}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|^2} \tag{1.28}$$

where

$$\Sigma(\xi') = \int \rho(\xi_1, \xi_2, r_3) dr_3 \tag{1.29}$$

is the *surface mass density*, that is the mass density projected onto a plan perpendicular to the incoming light ray. Note that in Equation (1.29) $\xi' = (\xi_1, \xi_2)$ is a two-dimensional vector.

The lens equation

Consider a situation sketched in Figure 1.3. The two planes identify the source, the lens and the source plane connected by the optical axis. A light ray originating from the source at $\boldsymbol{\eta}$ intersects the lens plane at $\boldsymbol{\xi}$ and reaches the observer O. The two triangles $\widehat{oo's'}$ and $\widehat{oll'}$ are similar and the following relation holds:

$$\boldsymbol{\eta} = \frac{D_s}{D_d}\boldsymbol{\xi} - D_{ds}\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}). \tag{1.30}$$

Replacing the distance vectors $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ (in the source and the lens plane) with the angular quantities $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$,

$$\boldsymbol{\eta} = D_s\boldsymbol{\beta} \quad \text{and} \quad \boldsymbol{\xi} = \mathrm{D_d}\boldsymbol{\theta} \tag{1.31}$$

where D_s and D_d, and D_ds (see Fig. 1.3) are the angular diameter distances involved, the *lens equation* then follows as

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}), \tag{1.32}$$

where the quantity $\boldsymbol{\alpha}$ is the *scaled deflection angle* that is related to the deflection angle $\hat{\boldsymbol{\alpha}}$ by

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{D_{ds}}{D_s}\hat{\boldsymbol{\alpha}}(D_d\boldsymbol{\theta}); \tag{1.33}$$

[4] we are also implying here that the lens is the only mass distribution, which together with the two mentioned in the text, is a good assumption in the strong lensing regime

[5] the deflection angle of the deflector is given by the superposition of all deflection angles due to each mass element

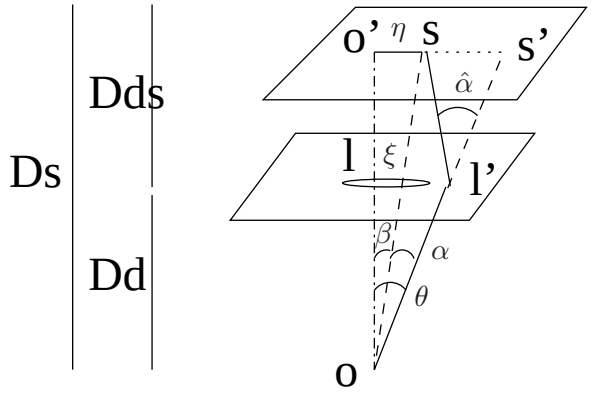


Figure 1.3: Geometry of strong gravitational lensing. The observer o sees the source s at a different angular position because of the deflection of the light. The distances between the observer and the source, the observer and the deflector, and the deflector and the source are indicated by D_s, D_d and D_{ds}.

the ratio $D_{\rm ds}/D_{\rm s}$ is defined as *lens strength efficiency*.

The scaled deflection angle can be expressed in terms of the surface mass density. Let us define the *critical surface mass density*

$$\Sigma_{crit} = \frac{c^2}{4\pi G}\frac{D_s}{D_{ds}} \tag{1.34}$$

and the *dimensionless surface mass density* or *convergence*

$$\kappa = \frac{\Sigma}{\Sigma_{\rm crit}}. \tag{1.35}$$

Now, re-writing Equation (1.28) using $\boldsymbol{\xi}$ as defined in Equation (1.31), and using the two definitions given in Equations (1.34) and (1.35), the scaled deflection angle becomes

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi}\int {\rm d}^2\theta' \kappa(\boldsymbol{\theta}')\frac{\boldsymbol{\theta}-\boldsymbol{\theta}'}{|\boldsymbol{\theta}-\boldsymbol{\theta}'|^2}. \tag{1.36}$$

The Fermat potential

From the identity $\nabla \ln|\boldsymbol{\theta}| = \boldsymbol{\theta}/|\boldsymbol{\theta}|^2$ it follows that the scaled deflection angle can be written as gradient of the function ψ

$$\boldsymbol{\alpha} = \nabla\psi \tag{1.37}$$

where ψ is the *deflection potential*,

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi}\int d^2\theta' \kappa(\boldsymbol{\theta}')\ln|\boldsymbol{\theta}-\boldsymbol{\theta}'|. \tag{1.38}$$

We recover the *Poisson equation* if we apply the Laplacian to Equation (1.38)

$$\nabla^2\psi = 2\kappa. \tag{1.39}$$

Using Equation (1.37) the lens equation can be written as

$$\boldsymbol{\beta} = \nabla \left(\frac{1}{2}\boldsymbol{\theta}^2 - \psi(\boldsymbol{\theta}) \right), \tag{1.40}$$

which becomes

$$\nabla \left(\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\beta})^2 - \psi(\boldsymbol{\theta}) \right) = 0. \tag{1.41}$$

Defining the scalar function ϕ (Schneider, 1985)

$$\phi(\boldsymbol{\theta}, \boldsymbol{\beta}) = \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\beta})^2 - \psi(\boldsymbol{\theta}), \tag{1.42}$$

Equation (1.41) can be written as

$$\nabla_\theta \phi(\boldsymbol{\theta}, \boldsymbol{\beta}) = 0 \tag{1.43}$$

which is equivalent to the lens equation (1.32).
The function ϕ is called the *Fermat potential* and is associated with the arrival time of the lensed images (see Sect.1.3.3) (Blandford & Narayan, 1986).

The Einstein radius

For an axi-symmetric mass distribution, the images are collinear with the centre of the lens; from the lens equation, $\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta})$, we see that in case of symmetry the source position is also aligned with the centre of the lens. It follows that the lens equations can then be written in one dimension. For an axisymmetric mass distribution, using Equation (1.27), the deflection angle is

$$\hat{\alpha} = \frac{4GM(\xi)}{c^2 \xi}, \tag{1.44}$$

where $M(\xi)$ is total mass within a circle of radius ξ. For this distribution the lens equation is given by

$$\beta = \theta - \frac{4GM(\theta)}{c^2 \theta} \frac{D_{\rm ds}}{D_{\rm d} D_{\rm s}}. \tag{1.45}$$

In the very special case when the source lies on the optical axis (see Fig. 1.3), $\beta = 0$. If in Equation (1.45) is $\beta = 0$, it follows

$$\theta_{\rm E} = \left(\frac{4GM}{c^2} \frac{D_{\rm ds}}{D_{\rm d} D_{\rm s}} \right)^{1/2}. \tag{1.46}$$

Therefore, due to the symmetry of the mass distribution, and the position alignment, the source is imaged into a ring of radius θ_E, called *Einstein ring*.
Such structures have indeed been observed (see Fig. 1.4), and provide an excellent observable to constrain the mass enclosed within the *Einstein radius*, θ_E. If the source is moved away from the optical axis, the ring will break up into multiple images, whose separation is approximately $\Delta\theta \approx 2\theta_E$.

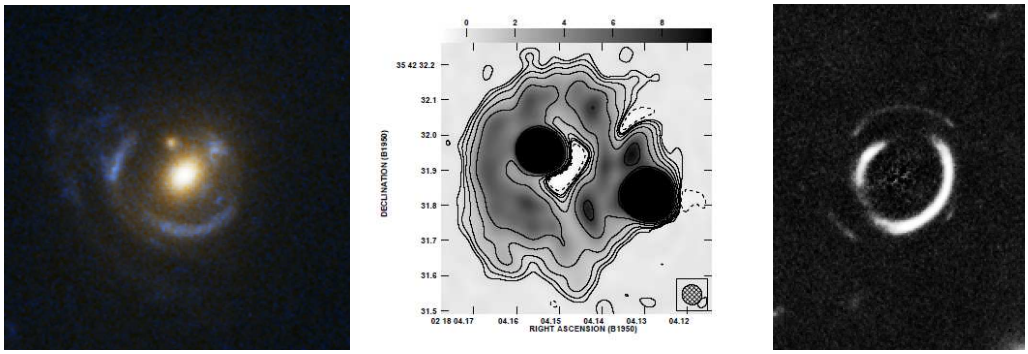


Figure 1.4: Three examples of Einstein's rings are shown. Left panel: system SDSS J095629.77+510006.6 (credit: A. Bolton). Right panel: Double Einstein ring, SDSSJ0946+1006. Two partial concentric ring-structures are visible (the foreground galaxy has been subtracted; credit: R. Gavazzi and T. Treu). Middle panel: System B0218+357 (Merlin/VLA 5 GHz map Biggs et al. 2001).

Time delay

When a lens produces multiple images the light travel time along the different light paths will be different. The difference in the arrival time of the images is called *time delay* which is the only dimensional observable of gravitationally lensed systems; if a source is variable the time delay can be measured by flux monitoring programs. There are two effects that contribute to the light-travel time (Cooke & Kantowski, 1975). First, light rays that are deflected are geometrically longer than undeflected ones, thus there is a *geometrical time delay*. Second, there is a *potential time delay*, which is due to the gravitational potential of the lens (this is known as *Shapiro delay*). For a light ray that crosses the lens plane at $\boldsymbol{\theta}$ the travels from the source to the observer and is given by the function

$$T(\boldsymbol{\theta},\boldsymbol{\beta}) = \frac{D_{\rm d}D_{\rm s}}{D_{\rm ds}c}(1+z_d)\left(\frac{(\boldsymbol{\theta}-\boldsymbol{\beta})^2}{2}-\psi(\boldsymbol{\theta})\right). \tag{1.47}$$

Recalling the scalar function defined in Equation (1.42), the time delay between two images at $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_j$ is

$$\Delta\tau_{ij} = \frac{D_{\rm d}D_{\rm s}}{D_{\rm ds}c}(1+z_d)[\phi(\boldsymbol{\theta}_i,\boldsymbol{\beta})-\phi(\boldsymbol{\theta}_j,\boldsymbol{\beta})]. \tag{1.48}$$

The distance factor in the above equation is proportional to $1/H_0$ (Eq. 1.9),

$$\Delta\tau \propto H_0^{-1}. \tag{1.49}$$

Thus, provided the redshifts of the lens and of the background source, and that the mass distribution is known, time delays measurement can constrain the value of the Hubble constant (Refsdal 1992, and references therein). This method for estimating the Hubble constant is independent of distance ladder methods that calibrate distances to high redshift galaxies with standard candles from low redshift, and that are affected by systematic uncertainties. A recent work by Suyu et al. (2010) shows that an accurate statistical study, based on Bayesian analysis, of a single gravitational lensing system, can provide precise measurements of H_0.

1.3.2 Magnification

The flux of a source density with intensity I is proportional to the solid angle $d\omega_0$ subtended by the source on the sky ($S_0 = I \times d\omega_0$). In the presence of a lens, due to the deflection of light rays, the solid angle is modified to the solid angle $d\omega$. Since the light deflection does not change the intensity I, the flux of the image is modified to $S = I \times d\omega$. The ratio of the two solid angles is called *magnification*, which is also the factor by which the observed source flux is changed,

$$\mu = \frac{S}{S_0} = \frac{d\omega}{d\omega_0}. \tag{1.50}$$

Let us define a Jacobian matrix that describes the lens transformation

$$A(\boldsymbol{\theta}) = \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\theta}} \ ; \tag{1.51}$$

using Equation (1.40), the elements of this matrix are

$$(A_{ij}) = (\delta_{ij} - \psi_{ij}) = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \tag{1.52}$$

The subscripts in ψ imply partial derivative of the potential with respect to θ_i and θ_j. A new quantity has been introduced, the *shear* $\gamma \equiv \gamma_1 + \mathrm{i}\gamma_2$, with

$$\begin{array}{ll} \gamma_1 = \frac{1}{2}(\psi_{,11} - \psi_{,22}) & \text{and} \\ \gamma_2 = \psi_{,12} \ . & \end{array} \tag{1.53}$$

The shear causes the shape distortion (due to the gravitational potential) of the images. The convergence κ is related to the potential through Equation (1.39), it contributes to isotropic magnification. The trace of matrix (1.52) is $\mathrm{tr}A = 2(1-\kappa)$, and its the eigenvalues are $a_{1,2} = 1 - \kappa \pm |\gamma|$, that give the factor of stretching along the direction given by the eigenvectors. The ratio of the solid angles subtended by an image and the unlensed source is the absolute value of the determinant of the matrix A

$$\mu = \frac{1}{\mathrm{det}A} = \frac{1}{(1-\kappa)^2 - |\gamma|^2} \ . \tag{1.54}$$

As defined in Eq. (1.54) the magnification can have positive or negative sign. The *parity* of an image is defined as the sign of μ. If both the eigenvalues have the same sign (+ or −), then the parity is positive; if the eigenvalues have different signs, the parity of the image is negative.

1.3.3 Image classification

Ordinary images

For a given source position $\boldsymbol{\beta}$, the Fermat potential defines a two-dimensional surface. *Ordinary images* form at points $\boldsymbol{\theta}$ where $\nabla\phi$ vanishes, thus they are located at the stationary points (local extrema and saddle points) of the two-dimensional arrival time surface defined by the Fermat potential (see the equivalence between Equations (1.42) and (1.32)).
The following three types of ordinary images can occur:

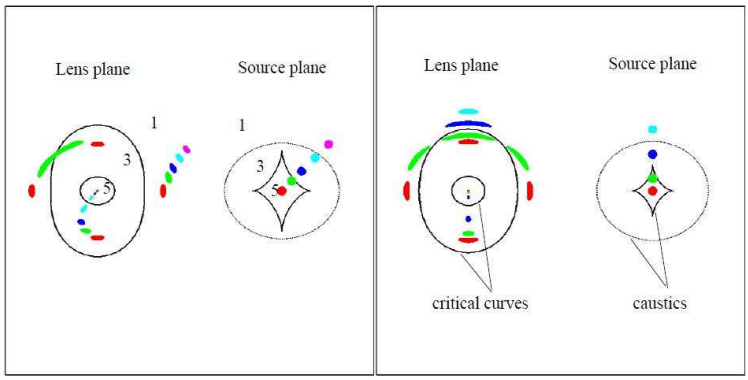


Figure 1.5: Critical curves and caustics for an elliptical mass distribution. The colours represent different source positions, and relative image positions. In the left panel the source approaches the centre through the *fold* caustics, in the right one through the *cusp*. The numbers 1,3,5 mark the regions in the source plane characterized by 1,3, or 5 number of images in the image plane. (Adapted from Narayan & Bartelmann 1996)

1. For $\det A > 0$, $\mathrm{tr} A > 0$, the image is formed at the minimum of the arrival time surface. This is called Type I image.

2. For $\det A < 0$, the image is formed at the saddle point and is called Type II image.

3. For $\det A > 0$, $\mathrm{tr} A < 0$, the image is formed at the maximum and is called Type III image. If this image exists it is located closer to the centre of the potential and hence is rarely observed because of low magnification and obscuration.

Critical images

For $\det A = 0$ the magnification becomes infinite. The set of points that satisfy this condition are called *critical curves* in the image plane, *caustics* in the source plane. In reality however, due to their finite sizes, when a source is close to a caustic, highly stretched images with high but finite magnification, are formed; these usually appear as arc like or rings. In Fig. 1.5, the image positions are shown given an elliptical mass distribution and different source positions. The outer smooth curve is the *radial* caustic, while the inner diamond is the *tangential* caustic.
Caustics and critical curves provide a useful qualitative understanding of a lens geometry. Critical curves divide the image plane into regions of different parities, that is images on either sides of a critical curve correspond to opposite parity. Caustics divide regions in the source plane with different multiplicity, and as a source moves across a caustic the number of images change by two (see Fig.1.5).

1.3.4 The odd number and magnification theorem

Consider a thin matter distribution whose surface mass density $\kappa(\boldsymbol{\theta})$ is smooth and decreases faster than $|\boldsymbol{\theta}|^{-2}$ for $|\boldsymbol{\theta}| \to \infty$. Such a lens has then finite total mass,

and continuous deflection angle. Under these assumptions, the *odd-number-theorem* (Burke, 1981) states that the total number of ordinary images is finite and odd. Under the same assumptions, the *magnification theorem* (Schneider, 1984) states that the image of a source that arrives first forms in the global minimum of the Fermat potential, with a magnification $\mu \geq 1$.

1.4 Gravitational lensing as tool

Gravitational lensing has a wide-spread set of astrophysical applications. Since it is sensitive to any kind of matter distribution, regardless if luminous or not, and it is an achromatic effect, it is used to probe different mass scales, stars, galaxies, clusters and large-scale structure as well (e.g. Suyu et al. 2010; McKean et al. 2010; Bauer et al. 2011). The combined analysis of strong lensing and stellar dynamics can be used to investigate galaxy evolution models (e.g. Ruff et al. 2011). Time delays measurements can be used to constrain cosmological parameters (Suyu et al., 2010; Coe & Moustakas, 2009). An intensive search for extra-solar planet objects has being carried out with the help of lensing within the OGLE (Optical Gravitational lensing experiment) program. When multiple images occur, the properties of the interstellar medium in lensing galaxies can be studied since the images are seen through different lines of sight (e.g. Falco et al. 1999), as well as differential Faraday rotation (Biggs et al., 2003). Within survey programs it is possible to carry out statistical studies in lensing, which yield cosmological constraints. If so far these studies have yielded only weak bounds (e.g Oguri et al. 2008), within the next decade, with the help of new instruments (e.g. JWST, ALMA, SKA) they will certainly benefit from the increase of the sample size.
In the following we describe the lensing applications on which the results of this thesis are based.

1.4.1 Mass substructures

Gravitational lenses described as smooth mass distributions fail in some cases to fit the image positions and in many cases to fit the flux-ratios. This problem is usually referred to as *astrometric anomalies and flux mismatches*. In this section we explain how this problem is related to the distribution of dark matter halos and how it is used to investigate the *missing satellite* problem which was introduced earlier.

Flux mismatches

Mao & Schneider (1998) argued that small-scale structures in the mass distribution may be a plausible explanation for the flux-ratio discrepancies. From equation (1.52), the gravitational magnification depends on second derivatives of the potential, whereas the deflected image positions depend on first derivatives. The presence of small-scale mass perturbations can thus change the flux ratios, nearly without changing the image positions. Studying lens systems which have a radio-

loud source[6], Dalal & Kochanek (2002) & Kochanek & Dalal (2004) concluded that flux anomalies can be explained by a substructure fraction of few percent $(0.6-7\%)$ of the smooth mass model, well in agreement with CDM predictions.
However the puzzle cannot be considered resolved as 1) high resolution numerical simulations have predicted a fraction of CDM substructures $\lesssim 0.5\%$ within a scale of typical image separations produced by lens galaxies (Mao et al., 2004), 2) when radio flux anomaly lenses are investigated it is often found that substructures are visible, thus luminous (e.g. MG 0414+0534 Schechter & Moore 1993; CLASS 2045+265 McKean et al. 2007; MG 2016+112 More et al. 2009).
There is however strong evidence that favours a lensing origin of the flux mismatches. In most cases, flux mismatches are such that the brightest saddle point is demagnified with respect to the magnification predicted by a smooth mass model. This effect is known as *parity-dependence* of flux ratios. Chen (2009) shows that this effect is unlikely to be produced by luminous satellite and that more substructures than predicted by simulation may be required to solve this problem. The use of larger sample of observed lenses, as well as simulations of dark matter and baryons, might be necessary for further studies of CDM substructures using strong gravitational lensing.

Astrometric anomalies

Besides the anomalies in the flux-ratios, discrepancies between observed and expected image positions can be used as well as evidence of dark matter clumps (substructures). Chen et al. (2007) shows that the image positions are perturbed on milli-arcsecond (mas) scales by substructures that project clumps near the Einstein radius of the main lens halo. Best candidate systems to detect those are lenses with extended sources, e.g. jets. The idea is that when a gravitational lens produces multiple images of an extended source, a non-smooth mass distribution may cause independent features in each image, which could reveal the presence of small structures along a line of sight. This is illustrated in Figure 1.6. Lensed AGN radio jets, imaged on mas scales using Very Long Baselines Technique (VLBI; see Chap. 3), where kink are detected are B1152+199 (Rusin et al., 2002; Metcalf, 2002), Q0957+561 (Haschick et al., 1981; Walsh et al., 1979; Garrett et al., 1996); MG0414+0534 (Hewitt et al., 1992). The last one is discussed in Chapter 5 of this thesis.

1.4.2 High-z Universe

Although the surface brightness of a lensed source is conserved, the gravitational magnification increases its observed flux density (see Eq. (1.50)), therefore they would appear brighter than they would without a lens. In some cases, this effect is essential to detect these sources in first place, provided that their lensed brightness is higher than the detection threshold of a survey or a current instrument sensitivity.

[6] flux anomalies could also be due to propagation effects, or to the gravitational lensing effect by stars within the lens galaxy; however, at radio wavelengths propagation effects do not occur or are frequency dependent, and hence can be corrected for, and the lensing by star can be excluded because stars have small Einstein radii compared to the size of a radio-loud source. Hence observations of lenses of a radio-loud source allows these studies

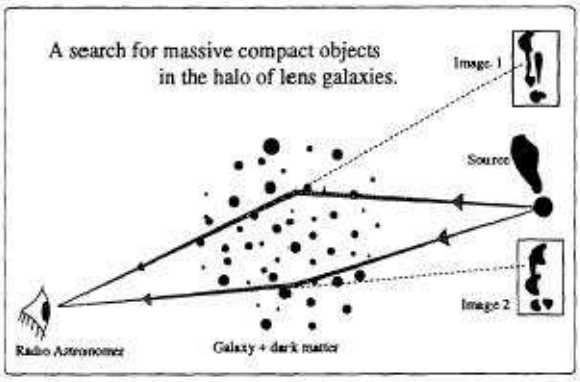


Figure 1.6: The picture has been adopted by (Garrett et al., 1996). The cartoon shows how lensed extended structures may reveal the presence of a non-smooth mass distribution, in this case independent bends in the lensed images could be produced.

Lensing capitalized as *natural telescope* has yielded to the discovery of high-redshift galaxies behind cluster lenses, as e.g. star-forming galaxies or sub-millimetre sources (Seitz et al., 1998; Garrett et al., 2005); in other cases it has allowed very detailed kinematic studies of distant galaxies (Nesvadba et al., 2006; Coppin et al., 2007) providing unique insight in the early Universe. A foreground lens as magnifier also provides higher angular resolution. Recent results from the Herschel telescope (Negrello et al., 2010) are very promising in the detection of strongly-lensed sub-millimetre galaxies, confirming lensing as powerful cosmological probe particularly at sub-millimetre wavelengths for the study of statistical and individual properties of high redshift star-forming galaxies. Together with sensitive and high resolution radio observations the study of high-z sub-sub-millimetre galaxies has implications on disentangling the emission of AGN from normal galaxies, which has implications on the cosmic history of star-formation and the growth of super-massive black holes.

Do not Bodies act upon Light at a distance, and by their action bend its Rays; and is not this action...strongest at the least distance?

I. Newton

2 Mass modelling of gravitational lenses

In the strong lensing regime multiple images of the same background source can be produced, whose observed configuration (relative positions, magnification and arrival times) is determined by the properties of the mass distribution. The goal of mass modelling is to find the best model for a lens that can explain the positions and flux ratios of the lensed images, as seen in the data. In reality, in lens modelling the intrinsic source structure needs to be modelled too, as the unlensed properties of the source are unknown as well.
Over the years several techniques have been developed that allow complex lens models to be constrained by high quality data, e.g. `LENSTOOL` (Kneib et al., 1993), `GRAVLENS` (Keeton, 2001b), `LENSCLEAN` (Kochanek & Narayan, 1992; Wucknitz, 2004) and `PIXELENS` (Saha & Williams, 2004). Combining different data sets results have been achieved concerning galaxy density profiles, their evolution with cosmic time (Ruff et al., 2011), the distribution of galaxies in the fundamental plane (Treu et al., 2006), the distribution of mass substructures along the line-of-sight and galaxies environments (Thanjavur et al., 2010).
For this thesis we have used the `GRAVLENS` software; below we will describe the lens-modelling techniques implemented in the code as well as the parametrization of mass models.
The chapter is organized as follow: in Sect. 2.1 we describe parametrizations for standard mass models. In Sect. 2.2 we describe the algorithm we used for mass modelling. We then explain, in Sect. 2.3, a technique suitable for modelling extended sources. Subsequently give details on error estimates and the number of constraints. Finally in Sect. 2.6 we describe typical degeneracies that occur in strong lensing.

2.1 Parametric mass models

In order to study a mass distribution of a gravitational lens, one approach is to parametrize it with a generic model, and apply model-fitting techniques that allow us to find the best set of parameters reproducing the observed properties of

a lens system. Alternatively, there are *non-parametric* methods, which use a grid parametrization of the surface mass density κ; the constraints are written as linear equations of κ and different mass distributions constrained by the data can be investigated (Saha & Williams, 2004).
We have followed the first approach and used analytical mass models which are good approximations to the real distribution of mass. In this section we describe the ones implemented in `GRAVLENS` which we have used in this thesis (Keeton, 2001a).

2.1.1 Power-law models

The density distribution in three dimensions is assumed to be a power-law given by

$$\rho \propto r^{-\gamma}. \tag{2.1}$$

The surface mass density is the projection of the three-dimensional density distribution along the line of sight, and hence it is proportional to $r^{1-\gamma}$.
Given a deflecting potential described by a power-law of the form

$$\psi(\theta) = \frac{b^2}{3-\gamma}\left(\frac{\theta}{b}\right)^{3-\gamma}, \tag{2.2}$$

the surface mass density κ, the deflection angle α and the shear γ_s are given by the following expressions:

$$\kappa(\theta) = \frac{3-\gamma}{2}\left(\frac{\theta}{b}\right)^{1-\gamma}, \tag{2.3}$$

$$\alpha(\theta) = b\left(\frac{\theta}{b}\right)^{2-\gamma} \tag{2.4}$$

and

$$\gamma_{\rm s}(\theta) = \frac{\gamma-1}{2}\left(\frac{\theta}{b}\right)^{1-\gamma} \tag{2.5}$$

where b is the Einstein radius.

2.1.2 Singular Isothermal Sphere (SIS) profile

This is a special case of the power law lens models that can account for the lensing properties of many galaxies and clusters. The radial density distribution is given by

$$\rho \sim r^{-2}. \tag{2.6}$$

It corresponds to a self-gravitating spherically-symmetric ideal gas whose temperature is constant at all radii, hence the name 'isothermal'. A spherical distribution is characterized by flat rotation curve, which is observed for spiral galaxies (Rubin et al., 1978). For elliptical galaxies the velocity dispersion of stars acts as kinetic temperature which is then constant with radius (Binney & Tremaine, 1987). This model is indeed found to apply well to the mass distribution as seen in galaxies (Koopmans et al., 2006).
This mass distribution has however two non-physical properties: the central density diverges as $\rho \propto r^{-2}$, and hence the name 'singular', and the total mass diverges as

$r \longrightarrow \infty$. We will address the former feature in the next section; the distribution for large r does not affect the lensing properties for the inner regions of the lens.
For a SIS the surface mass density at a projected radius ξ is given by

$$\Sigma(\xi) = \frac{\sigma_v^2}{2G\xi}, \tag{2.7}$$

where σ_v is the one-dimensional velocity dispersion.
Since $\kappa = \Sigma/\Sigma_{\rm crit}$, using the expression for $\Sigma_{\rm crit}$ (Equation. 1.34), the convergence for a SIS is

$$\kappa(\theta) = \frac{2\pi}{\theta}\frac{D_{\rm ds}}{D_{\rm s}}\left(\frac{\sigma_{\rm v}}{c}\right)^2, \tag{2.8}$$

where we have used $\xi = \theta D_{\rm d}$. Comparing Equation (2.8) with Equation (2.3) for $\gamma = 2$ gives

$$b = \frac{4\pi\sigma_{\rm v}^2 D_{\rm ds}}{c^2 D_{\rm s}}, \tag{2.9}$$

the deflection angle and the shear are given by

$$\alpha(\theta) = b \tag{2.10}$$

and

$$\gamma_{\rm s}(\theta) = \frac{1}{2}\frac{b}{\theta}. \tag{2.11}$$

For these models, in `GRAVLENS` the Einstein radius is parametrized as given in Equation (2.9).

2.1.3 Non-Singular Isothermal Sphere (NIS) profile

The central singularity for a SIS does not occur if the density is nearly constant within a core of radius r_c. The surface mass density at the projected radius ξ is given by

$$\Sigma(\xi) = \frac{\sigma_v^2}{2G\sqrt{\xi^2 + \xi_{\rm c}^2}}, \tag{2.12}$$

where ξ_c is the projected core radius. For radii much larger than $\xi_{\rm c}$ the surface mass density approaches the SIS model.
The lens properties are then,

$$\kappa(\theta) = \frac{b}{2\sqrt{\theta^2 + \theta_{\rm c}^2}}, \tag{2.13}$$

where $\xi_c = D_d\theta_{\rm c}$;

$$\alpha(\theta) = \frac{b(\sqrt{\theta^2 + \theta_{\rm c}^2} - \theta_{\rm c})}{\theta}, \tag{2.14}$$

and

$$|\gamma_{\rm s}(\theta)| = \frac{b(\sqrt{\theta^2 + \theta_{\rm c}^2} - \theta_{\rm c})}{2(\sqrt{\theta^2 + \theta_{\rm c}^2})(\sqrt{\theta^2 + \theta_{\rm c}^2} + \theta_{\rm c})}. \tag{2.15}$$

The value of ξ_c depends on the central density as ρ_c^{-2}. The fact that for lens galaxies an odd number of images is not observed is usually interpreted as the missing image is located very close to the centre of the lens galaxy and is highly demagnified. This yields a lower limit for the central density and thus an upper limit for the core radius, which for galaxy scales is expected to be small.

2.1.4 Truncated density distributions

A more general description for the three-dimension density distribution is power law with central cusp, which declines asymptotically and has a break radius a, given by

$$\rho \propto \frac{1}{r^{\gamma}} \frac{1}{(a^{\alpha} + r^{\alpha})^{(m-\gamma)/\alpha}}. \tag{2.16}$$

A density distribution of this family has a central cusp with $\rho \propto r^{\gamma}$, for large values of r, $r \gg a$ it declines as $\rho \propto r^{m}$. For $\gamma = 1$, $m = 3$ and $\alpha = 1$, it corresponds to the NFW (Navarro et al., 1997) and for $\gamma = 2$, $m = 4$ and $\alpha = 1$ to the Jaffe model (Jaffe, 1983).
The projected surface mass density is

$$\kappa(\theta) \propto \frac{1}{\theta^{\gamma-1}} \frac{1}{(\theta_a^{\alpha} + \theta^{\alpha})^{(m-\gamma)/\alpha}}. \tag{2.17}$$

Truncated isothermal (Pseudo-Jaffe) sphere (TIS)

For strong lensing modelling the Jaffe profile is modified such that the three-dimension density distribution is

$$\rho \propto (r^2 + s^2)^{-1}(r^2 + a^2)^{-1}, \tag{2.18}$$

where a is the break radius and s is a core radius, $s < a$ with central density $\rho_{\rm c}$. For radii $s \lesssim r \lesssim a$ the density ρ goes as r^{-2} as for the isothermal sphere, and in the outer regions, $r \gg a$, it falls as $\rho \sim r^{-4}$. In `GRAVLENS` the scaled surface mass density for a truncated isothermal sphere is written as:

$$\kappa(\theta) = \frac{b}{2}\left[\frac{1}{\sqrt{\theta_{\rm s}^2 + \theta^2}} - \frac{1}{\sqrt{\theta_{\rm a}^2 + \theta^2}}\right], \tag{2.19}$$

where, $\xi_s = D_{\rm d}\theta_{\rm s}$, $\xi_a = D_{\rm d}\theta_{\rm a}$.
The deflection angle is

$$\boldsymbol{\alpha}(\theta) = b f(\theta/\theta_{\rm s}, \theta/\theta_a) \tag{2.20}$$

where

$$f(\theta/\theta_s, \theta/\theta_a) \equiv \left(\frac{\theta/\theta_{\rm s}}{1 + \sqrt{1 + (\theta/\theta_{\rm s})^2}} - \frac{\theta/\theta_{\rm a}}{1 + \sqrt{1 + (\theta/\theta_{\rm a})^2}}\right). \tag{2.21}$$

The magnitude of the corresponding shear is

$$\gamma(\theta) = \frac{b}{2}\left[2\left(\frac{1}{\theta_{\rm s} + \sqrt{\theta_{\rm s}^2 + \theta^2}} - \frac{1}{\theta_{\rm a} + \sqrt{\theta_{\rm a}^2 + \theta^2}}\right) + \left(\frac{1}{\sqrt{\theta_{\rm s}^2 + \theta^2}} - \frac{1}{\sqrt{\theta_{\rm a}^2 + \theta^2}}\right)\right]. \tag{2.22}$$

In the above equations, b represents the Einstein radius, $\theta_{\rm E}$, which for a truncated isothermal mass distribution is related to the velocity dispersion $\sigma_{\rm v}$ by

$$\sigma_{\rm v} = c\left(\frac{\theta_{\rm E}}{6\pi}\frac{D_{\rm s}}{D_{\rm ds}}\right)^{1/2} \tag{2.23}$$

(Elíasdóttir et al., 2007).

2.1.5 Singular Isothermal Ellipsoid

Non-symmetric mass distributions are needed in order to explain quadruples images, which cannot be reproduced by spherical mass distributions. Let us now replace in Equation (2.7) the projected radius ξ by the quantity

$$\zeta = \sqrt{\xi_1^2 + q^2\xi_2^2}; \tag{2.24}$$

which is constant on ellipses with minor axis ζ, major axis ζ/q and axis-ratio q. We find that the surface mass density of a singular isothermal ellipsoid (SIE) is

$$\Sigma(\xi_1, \xi_2) = \frac{\sigma_\mathrm{v}^2}{2G}\frac{\sqrt{q}}{\sqrt{\xi_1^2 + q^2\xi_2^2}} = \frac{\sqrt{q}\sigma_\mathrm{v}^2}{2G}\frac{1}{\zeta}, \tag{2.25}$$

where the normalization is chosen such the mass inside an elliptical iso-density contour for a fixed Σ is independent of q (Kormann et al., 1994).
In `GRAVLENS` the scaled surface mass density for a SIE is given as

$$\kappa(\theta_1, \theta_2) = \frac{b}{2[(1-\epsilon)\theta_1^2 + (1+\varepsilon)\theta_2^2]^{1/2}}, \tag{2.26}$$

where the minor and major axis components are $\xi_1 = D_\mathrm{d}\theta_1$ and $\xi_2 = D_\mathrm{d}\theta_2$. In the above Equation ϵ is related to the axis ratio q by

$$q = \sqrt{\frac{1-\epsilon}{1+\epsilon}} \tag{2.27}$$

and b is related to the velocity dispersion σ_v by

$$b = q\sqrt{\frac{2}{1+q^2}}4\pi\left(\frac{\sigma_\mathrm{v}}{c}\right)^2\frac{D_\mathrm{ds}}{D_\mathrm{s}}. \tag{2.28}$$

The deflection angle cannot anymore be reduced to a one dimensional scalar, instead it has two components, α_1 and α_2, along the the two axis θ_1 and θ_2,

$$\alpha_1 = \frac{b}{\sqrt{1-q^2}}\tan^{-1}\left(\frac{\theta_1\sqrt{1-q^2}}{[q^2\theta_1^2 + \theta_2^2]^{1/2}}\right) \tag{2.29}$$

and

$$\alpha_2 = \frac{b}{\sqrt{1-q^2}}\tanh^{-1}\left(\frac{\theta_2\sqrt{1-q^2}}{[q^2\theta_1^2 + \theta_2^2]^{1/2}}\right). \tag{2.30}$$

The magnitude of the shear $|\gamma_s(\theta_1, \theta_2)|$ equals the surface mass density $\kappa(\theta_1, \theta_2)$.

2.2 Algorithms for mass modelling

The aim of mass modelling is to find the mass model that can explain the properties of the observed images (positions, flux ratios and arrival time delays). The problem has two unknowns, the intrinsic, unlensed, source properties, and the foreground mass distribution. Given a mass model and an observed configurations of multiple

lensed images, for each image the corresponding source position can be found independently from the other images, and if the mass model is correct, each of the lensed image should correspond to the same source position within the uncertainties of the image positions. Furthermore for the given source position all the observed images, but the central demagnified ones, should be predicted. This is the basis of the algorithm used to determine the lens model. The single steps are:

1. assume a simple parametrized mass model as starting model
2. given the mass model and the observed image positions, use the lens equation to find the corresponding source positions and error-weighted mean of the source position
3. use the mass model and the mean source position to determine the properties of the lensed images (number of images, positions, flux ratios, parities and arrival times)
4. based on the predictions of the model and the observed image properties, assign a measure of the goodness-of-fit, χ^2
5. adjust the parameters of the model to minimize the value of the χ^2

Steps 2, 3 and 4 are the core of the algorithm and are performed in every step of the procedure minimizing the χ^2.
This algorithm is implemented in the `LENSMODEL` application within the `GRAVLENS` software (Keeton, 2001b).

2.2.1 Solving the lens equation

The source position $\boldsymbol{\beta}$ and the position of a lensed image $\boldsymbol{\theta}$ are related by the lens equation, $\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta})$. For a given set of parameters describing the mass model, the deflection angle $\boldsymbol{\alpha}$ can be determined for every $\boldsymbol{\theta}$ using Equation (1.37). But for a given source position, $\boldsymbol{\beta}$, the equation is not linear and may have multiple solutions, which are the multiple image positions $\boldsymbol{\theta}_i$. To find all the positions of the images, a numerical root finder is needed which will find all the roots of the lens equation in the image plane.
In order for the root finder to work, the number of images and their approximate location must be specified. In this way, reading the lens equation from right to left, each image position can be taken and mapped to a unique source position. The location of the images can be found if the image plane is described by a grid divided in tiles. The vertices of every tile in the grid can be mapped to the source plane via the lens equation leading to a tiling of the source plane. Every point in the source plane is covered by at least one tile, and points covered by more than one tile are *multiply imaged*. Thus, given the source position, the image plane tiles, which map to the tiles that enclose the source, can be identified. These tiles are the regions which can be provided to the numerical root finder to solve the lens equation and to refine all the image positions for a particular source. In Fig. 2.1 we show the tiling in the image and in the source plane for the quadrupole lens system MG0414+0534.

In the image plane, the tiling has higher resolution near the critical curves and at the position of a secondary lens galaxy. Near the critical curves, the lens mapping folds on itself, its determinant changes sign and the number of images changes; in any tile containing a critical curve the code generates a sub-grid to ensure to resolve any close pair of images. The code generates also sub-grid in any tile containing any galaxy other than the main lens to resolve its critical curves.

2.2.2 Optimizing the model

Once a mass model is given, the first guess for the source is given by the error-weighted average source positions (see Eq. 2.38) obtained by mapping the multiple images back to the source plane. This position is then used to find the positions of all the corresponding images in the image plane. The deviations of the quantities predicted by the model from the observed ones constitute the χ^2 for the model.
The χ^2 term for the image positions, evaluated in the *image plane*, is then

$$\chi^2_{\rm img,pos} = \sum_i \delta\boldsymbol{\theta}_i^T \cdot S_i^{-1} \cdot \delta\boldsymbol{\theta}_i, \tag{2.31}$$

where

$$\delta\boldsymbol{\theta}_i = \boldsymbol{\theta}_{{\rm obs},i} - \boldsymbol{\theta}_{{\rm mod},i}. \tag{2.32}$$

The sum extends over all the images, $\boldsymbol{\theta}_{{\rm obs},i}$ and $\boldsymbol{\theta}_{{\rm mod},i}$ are the observed and modelled positions of image i. S_i represents the covariance matrix, describing the error ellipses as follows:

$$S_i = R_i^T \begin{pmatrix} \sigma^2_{1,i} & 0 \\ 0 & \sigma^2_{2,i} \end{pmatrix} R_i, \tag{2.33}$$

$$R_i = \begin{pmatrix} -\sin\theta_{\sigma,i} & \cos\theta_{\sigma,i} \\ -\cos\theta_{\sigma,i} & -\sin\theta_{\sigma,i} \end{pmatrix}, \tag{2.34}$$

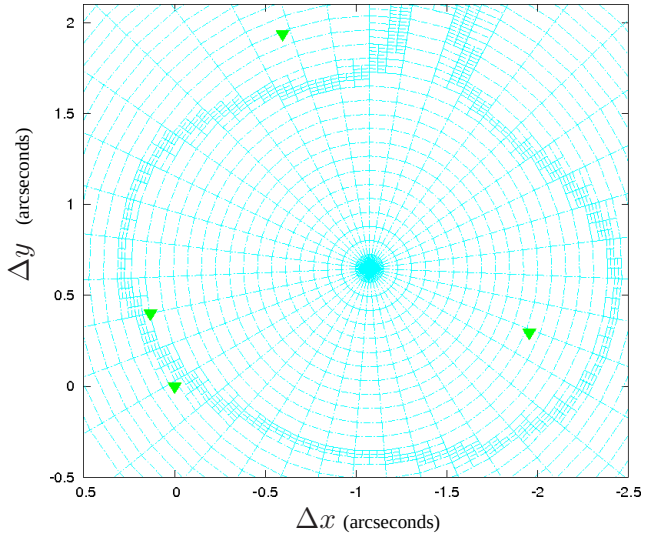
$\sigma^2_{1,i}$ and $\sigma^2_{2,i}$ are the semi-major and semi-minor axis of the error ellipse, and with position angle $\theta_{\sigma,i}$, measured East of North; R_i is the rotation matrix. The assumption is that position uncertainties for different point images are independent and Gaussian. When flux-ratios are included in the set of constraints, the χ^2 of the model has an additional term, $\chi^2_{\rm flux}$. If the observed flux of image i is $f_i \pm \sigma_{f,i}$, the χ^2 due to the deviations in the flux densities of the images is

$$\chi^2_{\rm flux} = \sum_i \frac{(f_i - \mu_i f_{\rm src})^2}{\sigma^2_{f,i}}, \tag{2.35}$$

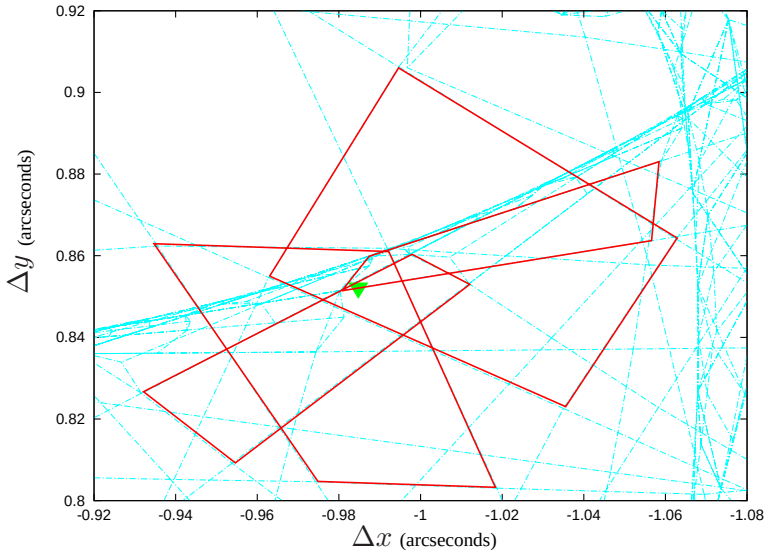
where μ_i is the modelled magnification of image i;

$$f_{\rm src} = \frac{\Sigma_i f_i \mu_i / \sigma^2_{f,i}}{\Sigma_i \mu_i^2 / \sigma^2_{f,i}} \tag{2.36}$$

is the best-fit source flux.
Other constraints can be included in the optimization: time delays or the structure of the source (see Sect. 2.3) or for example priors on the position angle and ellipticity of the host halo (which can be fitted from the surface brightness profile), or on its



(a)



(b)

Figure 2.1: Top panel shows the tiling in the image plane; the positions of the lensed images are marked by green triangles. At the bottom, panel b) shows a zoom of the tiling in the source plane; the position of the source is marked by the green triangle. Each tile in the image plane encompassing a lensed image has been mapped back onto the source plane using the lens equation. As expected in this plane the four tiles overlap, and the source is found in the regions where tiles overlap.

position from accurate optical measurements; each of them contributes to the χ^2 of the model, which is then the sum of all the terms,

$$\chi^2_{\rm tot} = \chi^2_{\rm pos} + \chi^2_{\rm flux} + \chi^2_{\rm oth}. \tag{2.37}$$

When the predicted number of images or the parity is incorrect, the code assigns an arbitrarily high value to the $\chi^2_{\rm tot}$.
To narrow down the search in the parameter space the χ^2 can be evaluated in the *source plane*. This is another minimization routine, which does not require to solve the lens equation and find all the images of a source; it provides a good starting model for the image plane minimization allowing a much faster convergence of the minimization routine in the image plane.
The χ^2 in the source plane is defined as,

$$\chi^2_{\rm src,pos} = \sum_i \delta\boldsymbol{\beta}_i^T \cdot M_i^T \cdot S_i^{-1} \cdot M_i \cdot \delta\boldsymbol{\beta}_i, \tag{2.38}$$

where

$$\delta\boldsymbol{\beta}_i = \boldsymbol{\beta}_{\rm obs,i} - \boldsymbol{\beta}_{\rm mod} \tag{2.39}$$

are the deviations between the modelled source position, $\boldsymbol{\beta}_{\rm mod}$, and the source position corresponding to image i, $\boldsymbol{\beta}_{\rm obs,i} = \boldsymbol{\theta}_{\rm obs,i} - \boldsymbol{\alpha}(\boldsymbol{\theta}_{\rm obs,i})$. For small $\delta\boldsymbol{\beta}$, $\delta\boldsymbol{\theta} \approx M_i \cdot \delta\boldsymbol{\beta}_i$ is a good approximation, and any deviations in the source plane multiplied by the magnifications matrix M_i gives an approximate image plane deviation; it follows that Equation (2.38) is a good approximation for the χ^2 evaluated in the image plane (Eq. 2.31). $\boldsymbol{\beta}_{\rm mod}$ is calculated as follows,

$$\boldsymbol{\beta}_{\rm mod} = A^{-1} \cdot \boldsymbol{b}, \tag{2.40}$$

where

$$A = \Sigma_i M_i^T \cdot S_i^{-1} \cdot M_i, \tag{2.41}$$

$$\boldsymbol{b} = \Sigma_i M_i^T \cdot S_i^{-1} \cdot M_i \cdot \boldsymbol{\beta}_{\rm obs,i}; \tag{2.42}$$

in the above equations, M_i is the magnification matrix for image i, S_i is the covariance matrix as in Equation (2.33); the modelled source position is then the error-weighted mean of $\boldsymbol{\beta}_{\rm obs,i}$.

2.3 Modelling extended sources

In the algorithm described above, the lensed configuration is treated as a collection of point-like components. However, in most cases the observations reveal extended lensed structures which could be used as well to constrain a lens model. Since the unlensed source properties are unknown, lens-modelling based on this approach requires to model in a self-consistent way the mass distribution and the source structure. An ideal lens model would then reproduce the observed properties of the images pixel by pixel, mapping the intrinsic surface brightness through the lensing potential and convolving it with the response of the telescope.
Given a true sky brightness distribution $\boldsymbol{I}_s$, and the observed one $\boldsymbol{I}_{obs}$, where the

vector notation expresses that we are considering the brightness for every pixel, the following lens equation needs to be inverted,

$$\boldsymbol{I}_{obs} = \mathcal{B}\mathcal{L}\boldsymbol{I}_s, \tag{2.43}$$

where $\mathcal{B}$ is is the matrix describing the telescope point-spread-function, and $\mathcal{L}$ is the matrix describing the lens mapping. Lens modelling based on this approach consists of two layers. An inner one, in which the lens model is fixed, and the source model is determined, and an outer one where the mass model parameters are varied to minimize the χ^2 function. Minimizing a χ^2 defined as

$$\chi^2 = \frac{|\boldsymbol{I}_{\rm obs} - \mathcal{B}\mathcal{L}\boldsymbol{I}_{\rm s}|^2}{\sigma^2}, \tag{2.44}$$

where σ^2 is the variance of the pixel errors, corresponds to invert Equation (2.43). Algorithms based on this idea are presented in Kochanek & Narayan (1992); Wucknitz (2004); Suyu et al. (2009). Lens systems constrained in this way, e.g. B1608+656 (Suyu et al., 2010), and B0218+357 (Wucknitz et al., 2004), are the basis to constrain the Hubble parameter using gravitational lensing.
The use of much larger sets of constraints requires however a higher computational effort, thus wider ranges of models are usually explored by faster techniques. The `LENSMODEL` application provides the *curve-fitting* technique for modelling extended images which uses the information contained in jets or arcs without the effort of building a full model for the intrinsic source brightness distribution (see below).

2.3.1 The *curve-fitting* technique

The technique can be applied when the lensed source has extended structure which is multiply imaged; it is based on the idea that surface brightness is conserved in gravitational lensing. For resolved features, contours showing the same level of surface brightness identify regions (sub-components) that must transform into each other under the lensing mapping. Taken a point on such a contour all the other images must lie on contours of the same level of surface brightness.

In reality, points on these curves do not always lie on contours of the same level, for no other reason than PSF smearing, however the approach is reasonable as it uses the information contained in the extended resolved images to constrain the model, without modelling the intrinsic source distribution.
Let us consider the case of two images of the same background source, each of them described by one curve, say curves A and B. Similarly to the case of a point-like source, given a lens model, for each of the curves, a curve describing the source structure can be found independently, and if the model is correct each curve should correspond within the uncertainties in the image plane to the curve describing the source. Furthermore, given the curve in the source plane, all the curves in the image plane must be predicted. Thus, taken curve A the intrinsic curve can be found, which can be mapped to the image plane and all the other images of the curve can be found. Let A_1 be an image of the curve A; within the uncertainty, A_1 must coincide with B. Thus a useful χ^2 is defined as the perpendicular distance between A_1 and B, integrated along the length of the curves. To handle the different lengths

of the curves, the "shorter" one is taken as "test" curve which is compared to the "longer" reference curves. To handle different sampling of the curves, points on the test curve are compared with segments on the reference curve. The geometry on which the algorithm is based is illustrated in Fig. 2.2a.
Let $\boldsymbol{y}$ be a test point, whose projection onto the segment connecting the reference points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is $\boldsymbol{y}'$. The *proper* projection is

$$\boldsymbol{y}'_{12}(\mathrm{y}) = [1 - \xi_{12}(\boldsymbol{y})]\boldsymbol{x}_1 + \xi_{12}(\boldsymbol{y})\boldsymbol{x}_2 \tag{2.45}$$

with ξ_{12} in the range $0 \leq \xi_{12} \leq 1$, in this case the projection lies on the segment between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, and

$$\xi_{12}(\boldsymbol{y}) = \frac{(\boldsymbol{y} - \boldsymbol{x}_1) \cdot (\boldsymbol{x}_2 - \boldsymbol{x}_1)}{|\boldsymbol{x}_2 - \boldsymbol{x}_1|^2}. \tag{2.46}$$

The perpendicular distance from $\boldsymbol{y}$ to the line segment is then

$$\mathrm{d}_{12}(\boldsymbol{y}) = |\boldsymbol{y} - \boldsymbol{y}'_{12}(\boldsymbol{y})|. \tag{2.47}$$

The reference curve may bent and the test point is outside the bend, in this case there may be no proper projections (Fig. 2.2b)[1], or if the test point is inside the bend there may be two proper projections (Fig. 2.2c). These cases are handled considering the distance of the test point to the nearest reference point. Thus the χ^2 contribution for the test point j is

$$\chi^2_{\mathrm{test,j}} = \min\left(\frac{|\boldsymbol{y}_\mathrm{j} - \boldsymbol{x}_\mathrm{i}|^2}{\sigma_\mathrm{i}^2}, \frac{\mathrm{d}_{\mathrm{i}-1,\mathrm{i}}(\boldsymbol{y}_\mathrm{j})^2}{\sigma_{\mathrm{i}-1}\sigma}\right) \tag{2.48}$$

where i runs over the reference points; for each segment the error-bars are taken to be the geometric mean of the error-bars at the endpoints. When there are multiple reference curves, the algorithm finds the smallest contribution among all of them. The total χ^2 for the curve constraints is conveniently defined as the sum over all the test points,

$$\chi^2_{\mathrm{crv}} = \Sigma_j \chi^2_{\mathrm{test,i}}. \tag{2.49}$$

The above expression reduces to Equation (2.31), if each curve consists of only a single point. In order to avoid that a good fit is found by reducing the number of points, the following statistics is specified $\bar{\chi}^2_{\mathrm{crv}} \equiv \chi^2_{\mathrm{crv}}/\mathrm{N_{crv}}$, with $\mathrm{N_{crv}}$ mean number of points. This prevents the code getting a lower value for the χ^2_{crv} with fewer points.

2.4 Errors on the parameters

The errors on the best-fitting model parameters can be expressed in terms of *confidence limit* from the likelihood function of the parameters. The integral of this function within such an interval gives the probability that the true value of the parameter lies within the specified limits.

[1] similarly, there is no proper projection if the test point lies beyond the end of the reference curve.

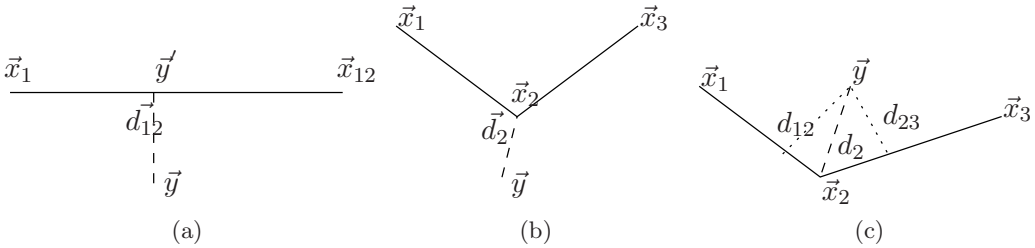


Figure 2.2: The figures show different geometries for curve fitting. A point $\boldsymbol{y}$ on the test curve is compared with segments on the reference curve. Panel (a) shows the most simple case, while panels (b) and (c) show more complicate cases when there is no proper projection (panel b) or there are two (panel c).

From the n-dimensional parameter space of the χ^2 distribution, the confidence levels on the probability of each parameter can be given individually from the distribution $\Delta\chi^2 = \chi^2 - \chi^2_{\rm min}$. For a Gaussian distribution, the interval within which is $\Delta\chi^2 = \chi^2 - \chi^2_{\rm min} \leq 1$ corresponds to the region in the parameter space such that a parameter will assume a value in this region with 68% (1σ) probability. Unless otherwise said, for each lens-model parameter, we have estimated the uncertainties in the region within which $\Delta\chi^2 \leq 1$ as each parameter is varied and all the others were fitted.

2.5 How many constraints?

The ratio $\chi^2_{\rm tot}/N_{\rm dof}$ is usually quoted as measure of the agreement between the data and the model, given that the mean value of the χ^2 distribution is equal to the number of degrees of freedom. In general if the model has p parameters, the number of degrees of freedoms is equal to $q - p$ where q is the number of constraints. Let us consider the case of N lensed images. For a source of n components, there are $2 \times n \times N$ position constraints, if each of the components is described as point-like; in addition when the flux ratios are included in the set of constraints, there are $n \times N$ more constraints. However, as explained in the previous sections, the intrinsic source position and flux are unknown, and have to be modelled too, thus the total number of constraints given by the positions is $2 \times n \times (N-1)$, and by the fluxes is $n \times (N-1)$. In presence of an extended structure, parametrized as described in Sect. 2.3.1, for each curve there is only one constraint given by the dimension of the curve, the model is indeed constrained by minimizing the perpendicular distance between reference and test curves. Thus for a source parametrized with m curves there are $m \times N$ constraints given by the curves. Similarly to what was said above, the total number of constraints is $m \times (N-1)$, in order to account for the unknown source structure. Note that the definition of χ^2 as given in Equation (2.49) is arbitrary and thus it has to be considered as an approximation of the goodness-of-fit. In Table 2.1 we give a short summary of the number of constraints given by the observed configuration for the lens systems discussed in this thesis. Clearly, the use of all the constraints, and when possible, the extended structure, provides more constraints for lens modelling.

Table 2.1: Constraints for modelling a lensed source consistently with the parametrization of Keeton (2001b). n is the number of sub-components a lensed source has, N is the number of lensed images, m is the number of curves when the lensed source is described as extended one. The two systems discussed in this thesis are over-viewed.

lens system	N n m	position	position and flux	curve
	N n m	$2n \times (N-1)$	$(2n+n) \times (N-1)$	$(3n+m) \times (N-1)$
J002240.91+1431110.4	4 1 1	6	9	12[a]
MG J0414+0534	4 3 3	18	27	36

[a] the system shows a bright arc counted here as one curve. However, the `LENSMODEL` application provides the *ring-fitting* algorithm which allows to parametrize the ring structure and constrain the lens model and the shape of the source (Kochanek et al., 2001).

2.6 Degeneracies in mass models

Although strong gravitational lensing provides very accurate mass measurements of the mass of distant galaxies, there are some degeneracies that affect lens mass models and that need to be considered for a correct interpretation of lens-modelling results.

2.6.1 Radial profile degeneracy

When the lensed images lie at a similar distance from the centre of the mass distribution, the observed configuration provides tight constraints only on the mass enclosed by the images but not on how it is distributed. Consider an axi-symmetric mass model, the two images form at θ_A and θ_B on opposite sides of the lens. The lens equation than reads

$$\beta = \theta_A - \alpha(\theta_A) = -\theta_B + \alpha(\theta_B). \tag{2.50}$$

Let us consider a power-law for the radial profile which leads to a power-law deflection angle $\alpha = b^{\gamma-1}/\theta^{\gamma-2}$. Inserting this into Equation (2.50), we find that the Einstein radius is given by

$$b = \left(\frac{\theta_A + \theta_B}{\theta_A^{2-\gamma} + \theta_B^{2-\gamma}} \right)^{1/(\gamma-1)}. \tag{2.51}$$

From the above equation, and recalling Equation (1.46), it nicely follows that the location of the ring b, and thus the mass scale, can be determined from the image configuration. However it is very difficult determining the radial profile of the mass distribution because of a strong degeneracy between b and the slope γ. If the images have similar angular distance from the lens, the dependence of b on the radial profile is very weak, of order of $(\delta\theta/\theta)$. This degeneracy can be broken if the data show lensed structures within the anulus $\delta\theta = \theta_A - \theta_B$ (Kochanek, 2006).

2.6.2 Shear-ellipticity degeneracy

This degeneracy relates to the angular structure of the lensing potential. Its quadrupole moment has contributions from the ellipticity of the lens galaxy as well as the external shear, and these two are indistinguishable; they are

$$\phi(r,\theta) = -\frac{\varepsilon r}{2}\cos 2(\theta - \theta_\varepsilon) \tag{2.52}$$

where θ_ε is the position angle measured East of North; and

$$\phi(r,\theta) = -\frac{\gamma_s r^2}{2}\cos 2(\theta - \theta_{\gamma_s}) \tag{2.53}$$

where γ_s represent the contribution of external shear. Models with only one of these two components are usually well constrained, but the fit is improved dramatically if both of them are included in the model.

He who loves practice without theory is like the sailor who boards ship without a rudder and compass and never knows where he may cast.

L. da Vinci

3

Radio interferometry and techniques for data reduction

For this thesis, we have made use of astronomical data taken with radio interferometers, namely the Very Large Array (VLA) and the global Very Long Baseline Interferometry (VLBI) array (consisting in the joint action of the European VLBI Network and the Very Long Baseline Array). In this chapter we give an overview on the radio-interferometry technique, and highlight the aspects that are most relevant for the work presented in the following chapters.
There are some advantages in studying gravitationally lensed systems at these frequencies: 1) Propagation effects on the incoming radiation are very small or frequency-dependent. 2) Microlensing is less likely to cause flux anomalies, thus flux ratio measurements are much more accurate than at other wavelengths, and can be used to constrain the lens model. 3) The high angular resolution allows us to investigate the presence of small-scale non-luminous matter ($M \leq 10^{8-9}$ $M_{\odot}$) that are expected to perturb the image positions on milliarcsecond scales. 4) Thanks to the latest technical developments, which have provided the upgrade of some of the existing radio arrays (Expanded-Very Large Array, EVLA, and Expanded-Multi Radio Linked Interferometer Network, e-Merlin), or the construction of new radio telescopes (Low Frequency Array, LOFAR), it is now possible to perform much more sensitive observations, or revealing the sky at new frequencies, opening the way to new lenses being found or high-redshift background source being detected. Therefore, the study of gravitationally lensed systems from radio observations is a highly efficient and promising tool for future discoveries in cosmology and galaxy evolution studies.

Table 3.1: Angular resolutions achieved nowadays by interferometric techniques.

telescope	aperture	$\lambda_{\rm min}$	baseline length	resolution
VLT	8.2 m	2 μm	200 m	$\sim$ 2.4 mas
ALMA	12 m[a]	0.4 mm	10 km	$\sim$ 10 mas
EVLA	25 m	6 mm	36.4 km	$\sim$ 44 mas
VLBA	25 m	3 mm	6000 km	$\sim$ 0.12 mas[b]
GMVA	25 m[c]	3 mm	15000 km	$\sim$ 50 μ arcsec

[a] upon completion ALMA will have 12 7-m dishes comprising the Atacama Compact Array

[b] longest VLBA baseline length at 3 mm

[c] the array comprises also the Effelsberg telescope (100 m), Pico Valeta (30 m), Plateau de Bure (6 $\times$ 15 m), Onsala (20 m) and Metsahovi (14 m)

3.1 Basic principles of radio-interferometry

3.1.1 What do we measure?

The image of a source observed through an aperture of a given size larger than the source (point-like source approximation) will result in alternating bright and dark bands called *fringes*; for an aperture whose size is comparable with that of the source, the fringes are no longer visible. The quantity that measures the relative amplitude of the fringes is the *visibility function*:

$$\mathcal{V} = \frac{I_{\rm max} - I_{\rm min}}{I_{\rm max} + I_{\rm min}},$$

where $I_{\rm max}$ and $I_{\rm min}$ are the maximum and minimum value of the intensity of the fringes. The larger is the value of $\mathcal{V}$, the better the source is observed. In the ideal case, for a point like source, $I_{\rm min}$ will reach zero, giving $\mathcal{V} = 1$. However, a source has finite size, and thus $\rm I_{min}$ will always be larger than zero, and the visibility smaller than one.

The first application of interferometric techniques in astronomy dates back to 1920, when for the first time the diameter of the super-giant star Betelgeuse could be measured with high angular resolution (Michelson, 1920; Michelson & Pease, 1921). Figure 3.1(a) shows the interferometer used by Michelson & Pease (1921) to measure the diameter of α-Orionis. Light rays from the star fall on two mirrors and are reflected on the same detector. By varying the distance between mirrors $\rm M_1$ and $\rm M_4$ fringe patterns were obtained. The corresponding visibility function depends on the total flux density on the aperture and on the distance between the two aperture (see Fig. 3.1 for more details). By varying the distance of the two mirrors, Michelson and Pease were able to determine the location of the first maximum and estimate the diameter of the source (Michelson & Pease, 1921).

The resolution of an instrument is proportional to the ratio between the observing wavelength and the size of the aperture ($\sim \lambda/D$). In order to achieve higher angular resolution, larger telescopes or higher frequencies are necessary. This very basic concept has led since 1920 to the development of many applications of interferometry in astronomy. Interferometry is nowadays applied very efficiently in those windows

of the electromagnetic spectrum that is possible to investigate from the ground allowing for angular resolution of order of milliarcsecond resolution (Table 3.1). Let us now consider an aperture of finite size x that is receiving radiation of wavelength λ. The mathematical Fourier Transform relation existing between the electric field in a generic point P, $E(\varphi)$, and the electric field distribution along the aperture with which we observe $E(x)$, can be proved:

$$E(\varphi) = \int_{-\infty}^{\infty} E(x) \mathrm{e}^{-2\pi \mathrm{i}(x/\lambda)\sin\varphi} dx, \tag{3.1}$$

E(φ) is also known as *radiation field pattern* or *far field*. It can be shown that for a single, constant illuminated aperture of size $x_\lambda = x/\lambda$, the far field will be $\propto \sin(\sin(\varphi)\pi x_\lambda)/\sin(\varphi)\pi x_\lambda$. The first null of this function occurs for $\sin(\varphi) = 1/x_\lambda$. The position of the first null corresponds to the angular resolution of the aperture as given by Rayleigh's criterion $\sim \lambda/x$, thus the larger is x_λ the smaller is the first maximum (main lobe of the antenna diagram) which is the smaller angular scale that can be resolved. An interferometer may be seen as a collection of such apertures, with $d_\lambda = d/\lambda$ (d_λ is the distance between the apertures in unit of λ).

3.1.2 Response of an interferometer

Consider two antennas at a distance $\mathbf{b}$ pointing towards the same source, in a direction given by the unit vector $\mathbf{s}$. The incoming wave-front will reach the two antennas with a time lag τ_g, called *geometrical time delay* and given by $\tau_g = \mathbf{b} \cdot \mathbf{s}/c$. In order to combine the signals coherently an instrumental delay is then introduced in the path of the second antenna. Figure 3.2 shows how the signals from both antennas are processed. The incoming signals at radio frequency ν_{RF} is down-converted to an intermediate frequency with a local oscillator at frequency ν_{LO}. This is done for practical reasons as it is technically more convenient to perform amplification, filtering, delaying and cross-correlation. The latter is performed in the *correlator*, which is a voltage multiplier followed by a time averaging (integrator) circuit. If the input voltages are $V_1(t)$ and $V_2(t)$, the output is the complex quantity called *visibility*

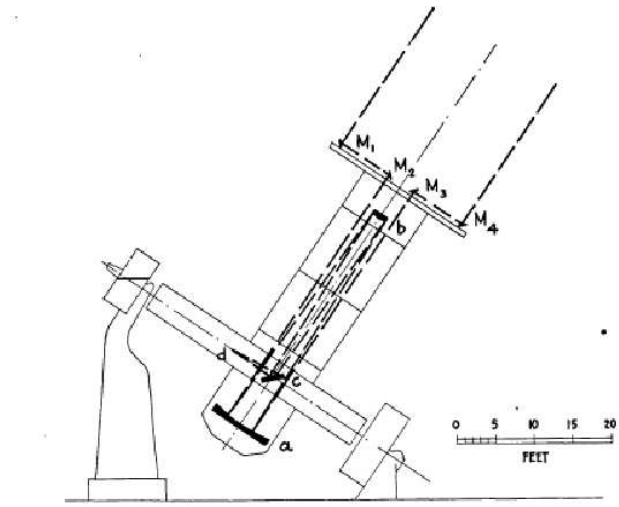
$$\mathcal{V} = \langle V_1 V_2^* \rangle.$$

Visibilities are measured as function of the antenna separations and can be used to reconstruct the source intensity distribution.
Let us consider the source of the electric field far away, such that all we measure is the surface brightness of the emitting source[1]; assuming that the source is spatially incoherent, the *spatial coherence function*, measured as function of the antennas separation (baseline) $\mathbf{b}$=$\mathbf{r}_1$−$\mathbf{r}_2$ is:

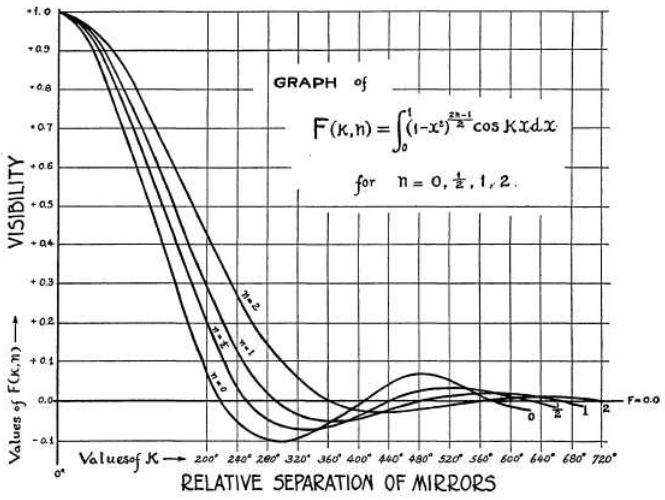
$$\mathcal{V}(\mathbf{r}_1, \mathbf{r}_2) \approx \int I_\nu(\mathbf{s}) \mathrm{e}^{-2\pi \mathrm{i}\nu \mathbf{s}(\mathbf{r}_1 - \mathbf{r}_2)/c} d\Omega. \tag{3.2}$$

In the above equation, I_ν is the source intensity distribution. The function $\mathcal{V}$ only depends on the separation vector $\mathbf{r}_1 - \mathbf{r}_2$ of the two points, not on their absolute

[1] We may conceive this as what we learn is the distribution of the source of electric field on a sphere of given radius $\mathbf{R}$, which indicates the position vector of the source.



(a)



(b)

Figure 3.1: Panel (a) shows the light path through the interferometer used by Michelson & Pease (1921). M_1, M_2, M_3, M_4, the mirrors; a, the 100 inches paraboloid; b, cover mirror; c, flat mirror; d, focus. In panel (b), the function $F(k, n)$ describes the flux distribution on the aperture as function of the distance from the centre. On the y-axis the plot shows the visibility function of the light of a source for different radial profiles, indicated by various values of n. On the x-axis, k is a parameter depending on the wavelength λ, b is the distance between the two mirrors and d is the size of a source, $k = \frac{2\pi b}{\lambda d}$. The curves show the fringe patterns for various light distributions as function of the parameter k. Given λ and b, the position of the first null of the visibility function corresponds to a value of k form which the diameter of the source, d, can be obtained. (Adopted from Michelson & Pease 1921)

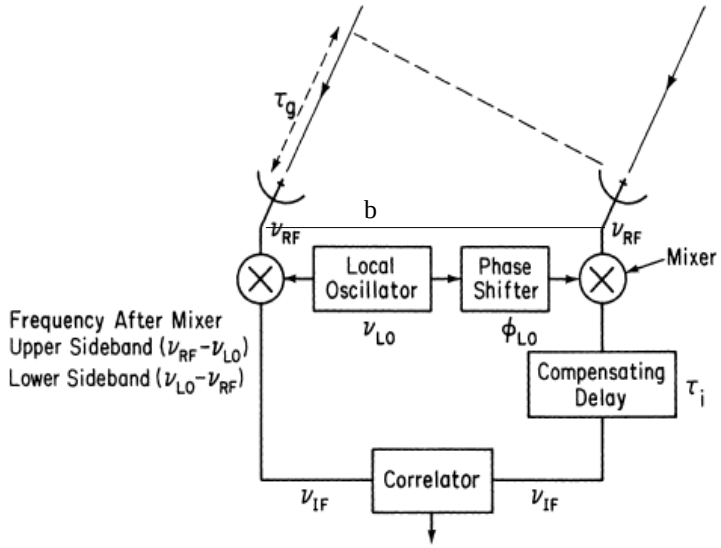


Figure 3.2: The incoming wave-front will reach the two antennas with a time lag τ_g that is proportional to the baseline length, **b**. The sky signal at radio frequency $\nu_{\rm RF}$ is combined with the signal given by the local oscillator and converted to an intermediate frequency $\nu_{\rm IF}$. Afterwards the signals are multiplied and averaged in the correlator whose output is the Fourier Transform of the sky brightness distribution. (Thompson 1999).

locations $\mathbf{r}_1$ and $\mathbf{r}_2$. Therefore, by holding one point fixed and moving the other around, the properties of I_ν can be reconstructed by $\mathcal{V}(\mathbf{r}_1, \mathbf{r}_2)$. This function is the complex visibility function measured by an interferometer.

3.1.3 Synthesis Imaging

Let us now define a coordinate system (u,v,w), where w points towards the centre of the field that is observed, u and v towards the East and the North in the local tangential plane (see Fig. 3.3). In this system, the vector spacing in the coherence function has components u, v and w measured in wavelengths $\lambda = \mathrm{c}/\nu$ at the centre frequency of the signal band;the baseline vector becomes $\mathbf{r}_1 - \mathbf{r}_2 = \lambda(u,v,w)$. In this coordinate system positions on the sky are defined in l, m, and n which are the direction cosines measured with respect to the u, v and w axis (Fig. 3.3). The response of the interferometer can be expressed as function of (u,v,w), since these coordinates represent the antenna separations with respect to the direction of the field. If we consider that the radiation comes from a small portion of the celestial sphere, $l \ll 1$ and $m \ll 1$ [2], the projected vector distance between two antennas has coordinates $x_u = \lambda u$ and $y_u = \lambda v$, and Equation (3.2) becomes

$$\mathcal{V}(u,v) = \int\int I(l,m) e^{-2\pi \mathrm{i}(ul+vm)} dl\ dm, \tag{3.3}$$

[2] the primary beam of parabolic antennas is usually $< 1^\circ$, which is a small area in the sky

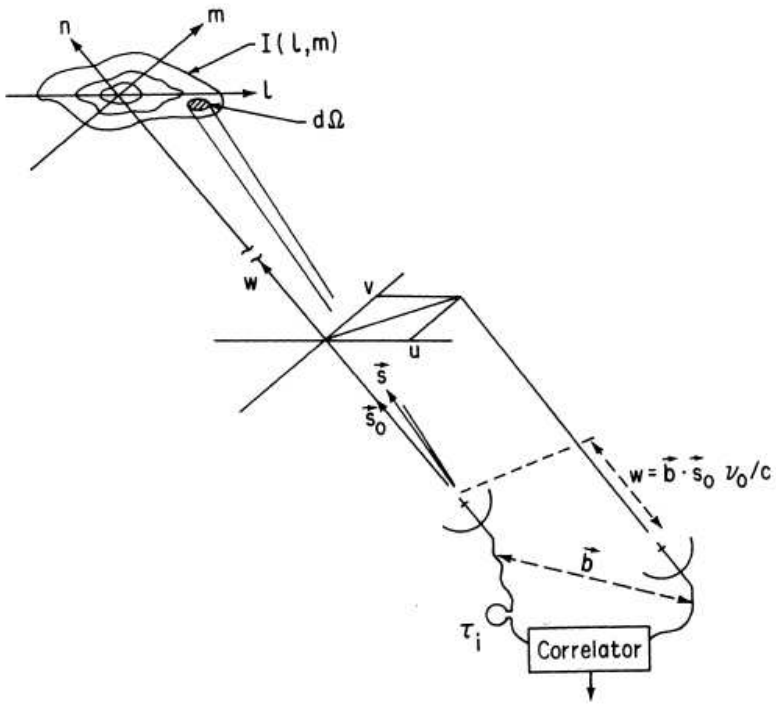


Figure 3.3: (u,v,w) and (l,m,n) coordinate systems used to express the antenna separations and the source brightness distribution. (Thompson 1999).

which is a 2-dimensional Fourier Transform relation between the spatial coherence function and the sky intensity written as $I(l,m)$. Equation (3.3) is the basis for radio interferometry. This is usually referred to as the van Cittert-Zernike theorem [3].

If the visibility function is measured in all the points of the (u,v) plane, the sky intensity $I(l,m)$ may be recovered by inverting Equation (3.3). However is not necessary to measure all the Fourier components simultaneously. Consider the very simple case of having only two antennas, one of which is mobile: by moving the second antenna from place to place all the required Fourier components may be measured. The method of gradually building up all the required Fourier components and using them to image a source in the sky is called *aperture synthesis* (for their work in developing this technique, M. Ryle & A. Hewish were awarded the Nobel Prize for physics in 1974). As seen from a source, the spatial separation between two antennas changes continuously as the Earth rotates, such that visibilities are measured at different points in the uv-plane. This technique is called *Earth Rotation Synthesis*.

The function that describes the sampling of the visibility function is known as *transfer function* or *sampling function*, indicated by $S(u,v)$; this is a function of the source declination as well as the antenna separations. This function gives the

[3]The theorem shows that the spatial coherence function only depends on the separation vector $\mathbf{r}_1$-$\mathbf{r}_2$ and that if all the measurements are in a plane, then there is a Fourier Transform relationship between the spatial coherence function and the intensity distribution.

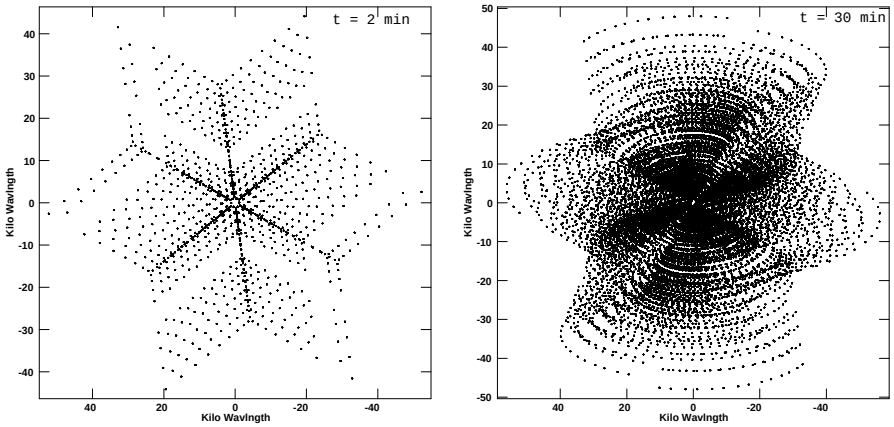


Figure 3.4: Distribution of visibilities using VLA. The uv-plane is filled with the rotation of the Earth over the time. The source declination is $\delta = 17°$.

values for u and v for which the visibility function is sampled, and hence it is zero everywhere no data have been taken; it is given by the following equation, which represents a sum of two-dimensional δ-Dirac function:

$$S(u,v) = \sum_k w_k \delta(u - u_k, v - v_k), \tag{3.4}$$

where w_k are the weighting factors; (u_k, v_k) are the points where the visibility measurements are recorded. The sampled visibility function is given by

$$\mathcal{V}^S(u,v) = S(u,v) \otimes \mathcal{V}(u,v). \tag{3.5}$$

The image that is reconstructed from the sampled visibility function is called *dirty image*, and can be obtained by Fourier transforming the sampled visibility function:

$$I^D = FT^{-1}(\mathcal{V}^S(u,v)) = FT^{-1}(S(u,v)) \otimes FT^{-1}(\mathcal{V}(u,v)) \tag{3.6}$$

where FT represents a Fourier Transform and the right-hand side equation follows from the convolution theorem and represents the convolution of the dirty beam B^D with the real image I:

$$I^D = B^D \otimes I. \tag{3.7}$$

B^D is the *point spread function* corresponding to the Fourier Transform of the sampling function, $\mathrm{B}^D = \mathrm{FT}^{-1}(S(u,v))$.
Equation (3.7) shows that the dirty beam B^D must be deconvolved from the dirty image I^D to recover the real image intensity distribution I. We will discuss in section 3.3 and 3.4 methods that are used to recover the true intensity distribution.

3.1.4 The effect of bandwidth in radio imaging

Let ν_0 be the centre frequency of the observing band $\Delta\nu$. The visibility data are processed as if they were all observed at the centre frequency ν_0. In particular, in the (u,v) plane components of the projected baseline vectors are calculated for ν_0.

Let these be (u_0,v_0), for any other frequency ν within $\Delta\nu$ they are (u,v)=(fu_0,fv_0), with $f = \nu/\nu_0$. Pratically this means that within $\Delta\nu$, at different frequencies the (u,v) plane is sampled in the radial direction. Due to the Fourier relation between the (u,v) and the image plane, the (l,m) coordinates of the brightness function are scaled too. The effect of averaging the data over the full bandwidth $\Delta\nu$ can be seen as averaging images of the brightness distribution each with a different scale factor in the direction aligned with the centre of the field. This effect is called *bandwidth smearing*. It is proportional to the fractional bandwidth, $\Delta\nu/\nu_0$, and the distance from the phase centre[4] scaled by the beam-width, $\theta_0/\theta_{\rm HPBW}$ (see Fig. 3.5 for an example). If the visibilities averaged over $\Delta\nu$ centred at ν_0 are changing significantly, features in a map are radially smeared and suffer loss of peak flux density far from the phase centre.
Strong gravitational lensing systems can be as wide as $4''$ (Browne et al., 2003), thus when mapping such fields, features far from the phase centre can only be maintained by not averaging the data over the whole observing bandwidth (see e.g. Garrett et al. 1999).

3.1.5 The effect of time averaging

The time averaging of the visibility data at the correlator results in another form of smearing in the image. Due to the rotation of the Earth, the visibility function rotates through an angle $\omega_e\tau$, resulting in different points corresponding to different times in the interval δt. Because of the Fourier Transform relation between the uv-plane and the image plane, in the latter the same rotation (of the image) occurs. At a point (l,m) the smearing is proportional to $\omega_e\tau\sqrt{l^2+m^2}$, with τ_a being the time interval between the visibilities are averaged.

3.1.6 Sensitivity

The measure of the weakest emission that can be detected is the *sensitivity*. For an image obtained from N antennas observing with a bandwidth $\Delta\nu$ and integrating over a time $t_{\rm int}$ it is given by:

$$\Delta I = \frac{1}{\eta_s}\frac{SEFD}{\sqrt{{\rm n_{pol}}N(N-1)\Delta\nu t_{\rm int}}}. \tag{3.8}$$

In Equation (3.8) $n_{\rm pol}$ is the number of polarizations included in the image, η_s represents the system efficiency and SEFD is the "system equivalent flux density", defined as the flux density of a radio source that doubles the system temperature:

$$\text{SEFD} = \frac{T_{\rm sys}}{K} = \frac{T_{\rm sys}2k_b}{\eta A}. \tag{3.9}$$

The SEFD is a measure of the overall antenna performance, it is expressed in Jansky (unit for spectral density, used especially in radioastronomy; 1Jy = 10^{-26}Wm^{-2}Hz^{-1}). K is the antenna gain, which consists of the efficiency (η) and collecting area of the

[4]This is defined as the centre of the field, chosen as the point for which the time delay τ_g is calculated

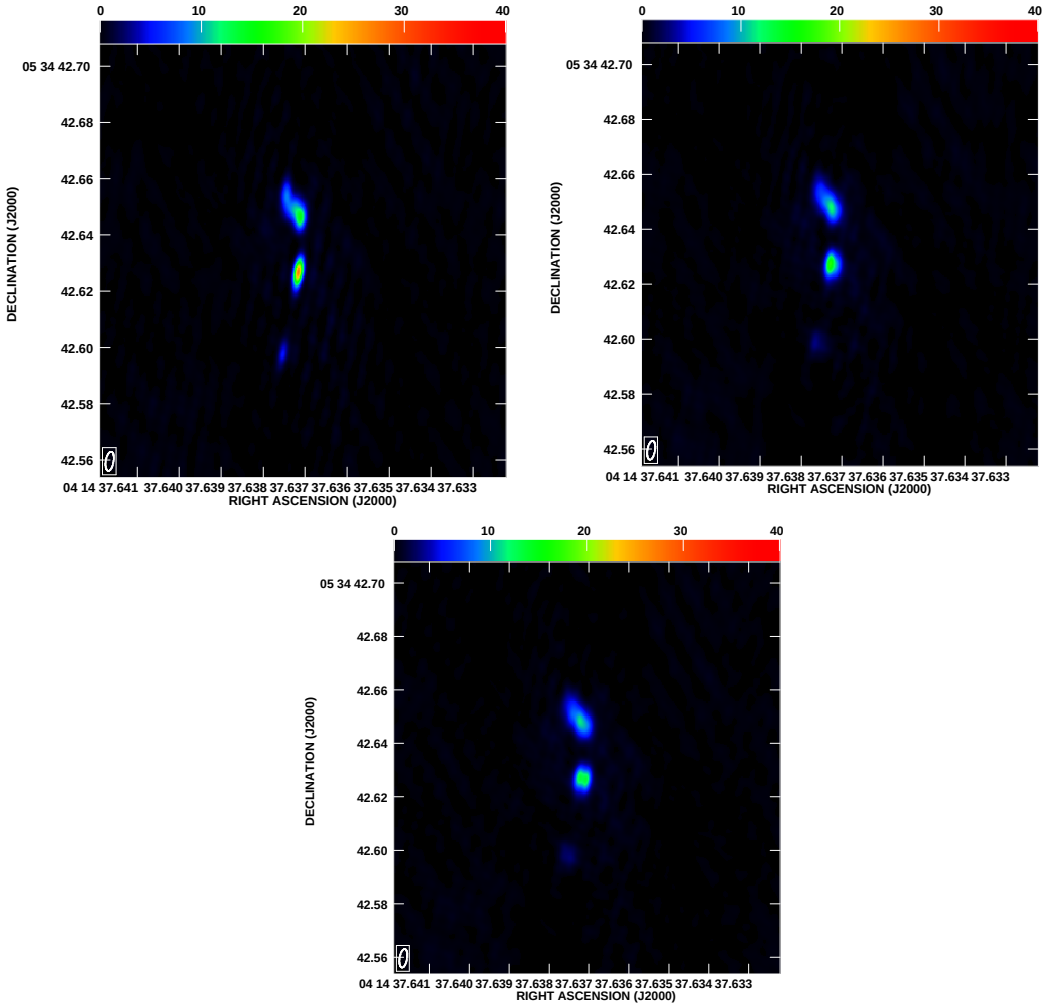


Figure 3.5: The effect of bandwidth smearing over 0.5, 2 and 8 MHz are shown. The figure shows the effect of radial smearing for a feature situated $\sim 2''$ away from the phase-centre. The map shows one of the 4 lensed images of the gravitational lensing system MG J0414+0534, which will be discussed in Chap. 5. From left to right, the panels show the effect of smearing occurring when averaging visibilities over 0.5 MHz, 2 MHz and 8 MHz. When visibilities are averaged over 0.5 MHz the shape of the feature is less affected and the percent peak loss is minimum; when averaging over a larger bandwidth, up to the complete observing bandwidth of 8 MHz, the effects of smearing are clearly visible: the peak flux density is reduced, and the shape of the three components appears is significantly changed.

antenna (A), and the Boltzmann's constant k_b. From Equation (3.8) is clear that the image sensitivity may be improved increasing the number of antennas, the observing bandwidth or the total integration time. The latest upgrade of previously existing radio facilities, such as EVLA and e-Merlin, will have larger observing bandwidth and improved receivers, that will improve the sensitivity by at least a factor of 10. Integrating the visibilities over the full bandwidth is referred to as *continuum observing mode*; when this is not possible because of the effects described in the previous section, and sensitivity has to be maintained many observing bandwidth (channels) are required. This is called *spectral line observing mode* (see Sect. 3.2.3).

3.2 Calibration and Editing

The spatial coherence function defined in Equation (3.3) is usually referred to as the *true* visibility function. An array of antennas samples this function at many different locations, after that the signals are collected, amplified, converted, correlated and averaged. A data point for each antenna pair is then recorded. These are called *observed* visibilities. *Calibration* is the process of determining and correcting for difference factors between observed and real visibilities. Standard calibration techniques are discussed in Sect. 3.2.1 and Sect. 3.2.2.

Editing

The process of identifying and discarding invalid data is called *editing*. This is an important step of the data reduction process since bad visibilities strongly degrade the imaging when Fourier transforming the data. It requires a careful inspection of the data in order to identify corrupted measurements, which may be antenna- or baseline-dependent. Depending on the observing wavelength, observations may be severely affected by radio-frequency interference, and when this happens these data need to be removed. Sometimes the antenna is off-source, or some scans are bad because of bad weather conditions; these scans need to be removed as well in order to avoid incorrect interpolation of the calibration.

3.2.1 Phase and amplitude calibration

The relation between the complex visibility observed at time t on the i-j baseline and the true one $\mathcal{V}_{ij}$(t) is:

$$\mathcal{V}_{ij}^{\rm obs}(t) = G_{ij}(t)\mathcal{V}_{ij}(t) + \epsilon_{ij}(t) + \eta_{ij}(t); \tag{3.10}$$

where the factor $G_{ij}(t)$ is the complex baseline based gain, $\epsilon_{ij}(t)$ is the baseline based complex offset, and $\eta_{ij}(t)$ is the baseline based complex noise, representing the thermal noise [5]. The factor $G_{ij}(t)$ is composed by the complex gains of the array elements i and j and a non-factorable part $g_{ij}(t)$ which for well designed arrays is within a one percent of unity,

$$G_{ij}(t) = g_i(t)g_j^*(t)g_{ij}(t) = a_i(t)a_j(t)e^{i[\phi_i(t)-\phi_j(t)]}, \tag{3.11}$$

[5]The use of complex quantities describes the combination of the two outputs of the correlator into a complex one

where $a_i(t)$ is an antenna based amplitude correction and $\phi_i(t)$ is the antenna based phase correction.
One way to perform amplitude and phase calibration is to observe a calibrator source, whose observations are used to determine the amplitude and phase corrections as function of time. The interpolation of these solutions then provides good values to correct the target visibility data. The image of the target made using this technique is referred to *phase-referenced image*.
In order to achieve good results, the phase calibrator should be unresolved at most frequencies and bright; its position should be known with high accuracy and within $< 2^\circ$ from the target source.
However, this method has a drawback when temporal and spatial variations in the atmosphere are present. Values for the $g_i(t)$ inferred from observing a calibrator may not apply to a source observed at another time and in another part of the sky. Thus, some residuals may remain. In case of VLBI observations imaging would be completely prevented by these effects and self-calibration is necessary (see Sect. 3.6).

Considering only the amplitude term in Equation (3.11), the raw correlation coefficient from the correlator is

$$\tilde{A}_{ij}(t) = a_i(t)a_j(t)S_{ij}, \tag{3.12}$$

where $a_{i/j}(t)$ are the amplitude terms of the antennas gains, and S_{ij} is the calibrated flux density which in terms of antenna performance and system temperature becomes

$$S_{ij} = \tilde{A}_{ij} b \sqrt{\frac{T_{\mathrm{sys}_i} T_{\mathrm{sys}_j}}{K_i K_j}}, \tag{3.13}$$

where the factor b accounts for signal-to-noise ratio losses due to effects like digitalization, the K_i are the antenna sensitivities measured in Kelvin/Jansky and T_{sys_i} are the system temperatures in Kelvin. During an observation the system temperature is measured frequently since it is time and elevation dependent. Using $\mathrm{T_{sys}}$ and K_i to determine the amplitude calibration is called *a-priori* amplitude calibration. This technique provide amplitude calibration consistent to better than 10%. An improvement to this can be made by observing strong non-varying sources. Models (or images) of these sources are used to perform self-calibration and determine constant offsets between the stations. However, this requires that the phases are very well calibrated.

Fringe-fitting

Very Long Baseline Interferometry (VLBI) is a technique where the telescopes are not physically connected. Their signal is recorded on the stations in a storage device (tapes in the past, disks at present) which are then brought together, multiplied, Fourier Transformed, and averaged in the correlator[6]. In this way, with the use of

[6] e-VLBI is a modern VLBI program conducted by the European VLBI network (EVN); data from the radio telescopes are transferred real time to the correlator via optical fibre cables; in e-VLBI high performing hardware and software are used that enable real time correlation of the data streams.

the longest baselines on ground (and space) is possible to achieve the highest angular resolution of the order of sub-milliarcseconds. *Fringe-fitting* is an important step when analysing VLBI data.
Residuals between the true and the observed visibility phase are due to variations of the phases with time and frequency; these originate from instrumental effects, geometrical delays and the propagation in the atmosphere and ionosphere. The error on the phase of the measured visibility is given by

$$\Delta\phi_{t,\nu} = \phi_0 + \frac{\partial\phi}{\partial\nu}\Delta\nu + \frac{\partial\phi}{\partial t}\Delta t \tag{3.14}$$

where the phase slope in frequency and time are respectively the residual delay and the residual rate. In order to correlate the data, a model has to be assumed that describes the geometrical delays and fringe rates. For connected interferometers residual delays and rates are small, and the model used by the correlator can properly account for these terms. This is not the case for VLBI arrays. Geometric errors are due to uncertainties in the terrestrial and celestial coordinates. Instrumental errors occur because of the signal propagating through different electronics and because of the use of different clocks at each station. The antennas can be up to $\sim$ 12000 kilometres apart, thus the signal travels through different geometric path in the atmosphere and ionosphere. The presence of electrons in the ionosphere introduces an additional phase delays, as well as different refractive indices in the atmosphere. Residuals phase delay and rates can then be significant, and need to be corrected. The *fringe-fitting* operation solves and removes for any residual delays and rates offsets that are not properly accounted by the model used by the correlator. *Fringe-fitting* estimates the phase delays and rates, and correct for them in the data. The frequency-time data (as delivered from the correlator) are Fourier transformed to the delay rate domain; here the peak of the function will appear as an isolated point, which gives the lowest residuals. Finding the signal is rather simple for strong sources, but for weak sources it can be tricky, since the signal-to-noise ratio is not enough to distinguish the true peak from noise peaks. Given a source model which replaces the true visibilities in Equation (3.10), it is possible to use all the data to determine delays and rates. This process of solving for antenna-dependent delays and fringe rates is called *global fringe-fitting*. Using a reference antenna, for which the phase, delay and rate are set to zero, solutions are found for each antenna; these solutions are then used to find a least-square solution where antenna phases, fringe rates and delays are determined from all the baselines [7].

3.2.2 Closure quantities and self-calibration

In order to image a target source is fundamental to align the phases from different antennas. In the early times of radio interferometry, Roger Jennison introduced a new concept that became crucial in the development of algorithms to produce images of the radio sky, the *closure quantities* (Jennison, 1958). He realized that

[7] Alternatively, there is the *baseline-based fringe fitting* that determines delays and rates for each baseline, and afterwards these are obtained for the single antennas. In order to use this method the source has to be detected on all the baselines, otherwise baselines with a weak or non-detection will not be calibrated.

the sum of visibility phase around a closed loop of baselines is free of antenna-based errors. Considering only the phase term of Equation (3.11), the relation between the measured and the true phases is

$$\tilde{\phi}_{ij}(t) = \phi_{ij}(t) + \theta_i(t) - \theta_j(t) \tag{3.15}$$

where $\tilde{\phi}_{ij}$ is the measured visibility phase and $\phi_{ij}(t)$ is the true one. Consider now the sum of phases around a closed triangle of antennas i, j and k,

$$\tilde{\phi_{ij}}(t) + \tilde{\phi_{jk}}(t) + \tilde{\phi_{ki}}(t) = \phi_{ij}(t) + \phi_{jk}(t) + \phi_{ki}(t) \tag{3.16}$$

where the terms θ_i, θ_j and θ_k cancel because the appear twice with opposite sign. The quantity

$$\tilde{C}_{ijk}(t) = \tilde{\phi_{ij}}(t) + \tilde{\phi_{jk}}(t) + \tilde{\phi_{ki}}(t). \tag{3.17}$$

is known as observed *closure phase* and apart from noise term is identical to the 'real' closure phase

$$C_{ijk}(t) = \phi_{ij}(t) + \phi_{jk}(t) + \phi_{ki}(t). \tag{3.18}$$

In the same way, for a loop of 4 elements a *closure amplitude* is defined as the ratio of the visibility amplitudes which cancels out the antenna based gains,

$$\Gamma_{ijkl}(t) = \frac{|\tilde{V}_{ij}(t)||\tilde{V}_{kl}(t)|}{|\tilde{V}_{ik}(t)||\tilde{V}_{jl}(t)|}. \tag{3.19}$$

These two quantities are then good observable. For an array of N elements there are $N(N-1)/2-(N-1)$ independent closure phases, and $N(N-1)/2-N$ closure amplitudes which can be used to constraint the true intensity sky distribution. However, they could not be used for imaging until the late 1970s, when the advent of fast computers made possible to produce images consistent with the informations contained in the closure quantities. The first iterative scheme was produced by Readhead and Wilkinson in 1978, and became known as *Hybrid mapping*. An initial model of the source is used to determine the phases on baselines with the help of closure phases. A new model is produced from the observed visibilities amplitudes and the just predicted visibility phases. These steps are then repeated until a satisfying model is obtained.

Self-calibration

Similarly to method just described, *self-calibration* is based on an iterative scheme, in which the antennas gains are allowed to be free parameters when determining the true sky intensity distribution improving the quality of the images. As mentioned in Sect. 3.2.1, VLBI phase calibration relies on this process. The iterative scheme consists of two parts, 1) determining the antenna based corrections, 2) determining the source brightness distribution. Basically the aim is to produce a model of the sky $\tilde{I}$, the Fourier transform $\hat{V}$ of which, when corrected by the complex gain factor, reproduces the observed visibilities within the noise $\tilde{V}$. This is done by minimizing the sum of the squares residuals

$$S = \sum_k \sum_{i,j;i\neq j} w_{ij}(t_k)|\tilde{V}_{ij}(t_k) - g_i(t_k)g_j^*(t_k)\hat{V}_{ij}(t_k)|^2 \tag{3.20}$$

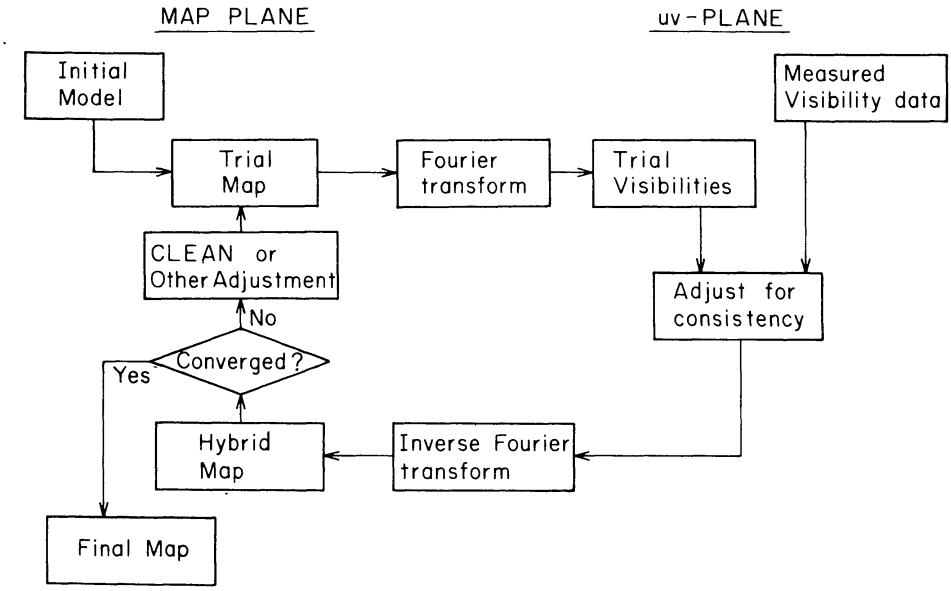


Figure 3.6: Iterative algorithm of self-calibration. Adapted from Pearson & Readhead (1984)

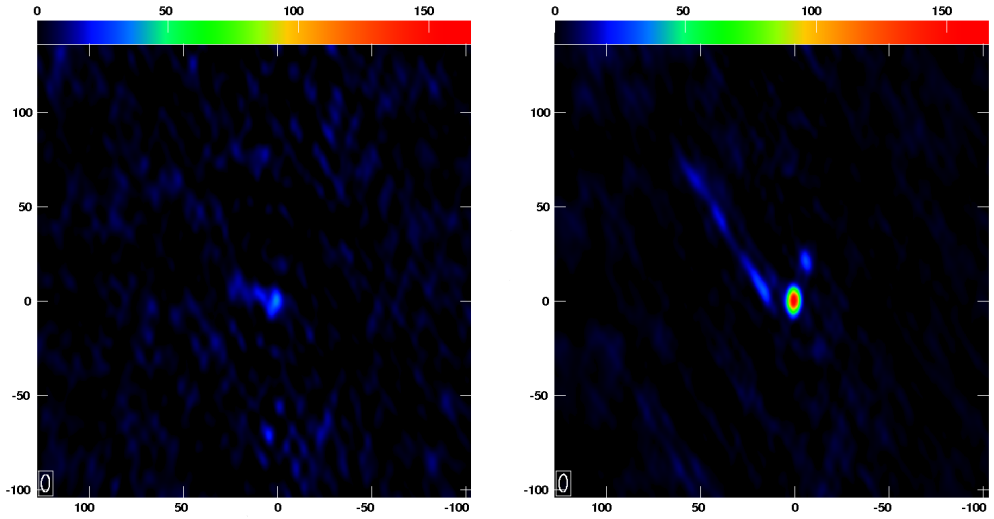
where the $w_{ij}(t_k)$ are the weights purely determined from the uncertainties in the observed visibilities (see Sect. 3.3). Figure 3.6 illustrates the process of self-calibration. A good starting model is required. The observed visibilities are then divided by this model in order to solve for the antenna gains which are then used to correct the visibilities,

$$V_{ij,\mathrm{corr}}(t) = \frac{\tilde{V}_{ij}(t)}{G_{ij}(t)G^*_{ij}(t)}. \tag{3.21}$$

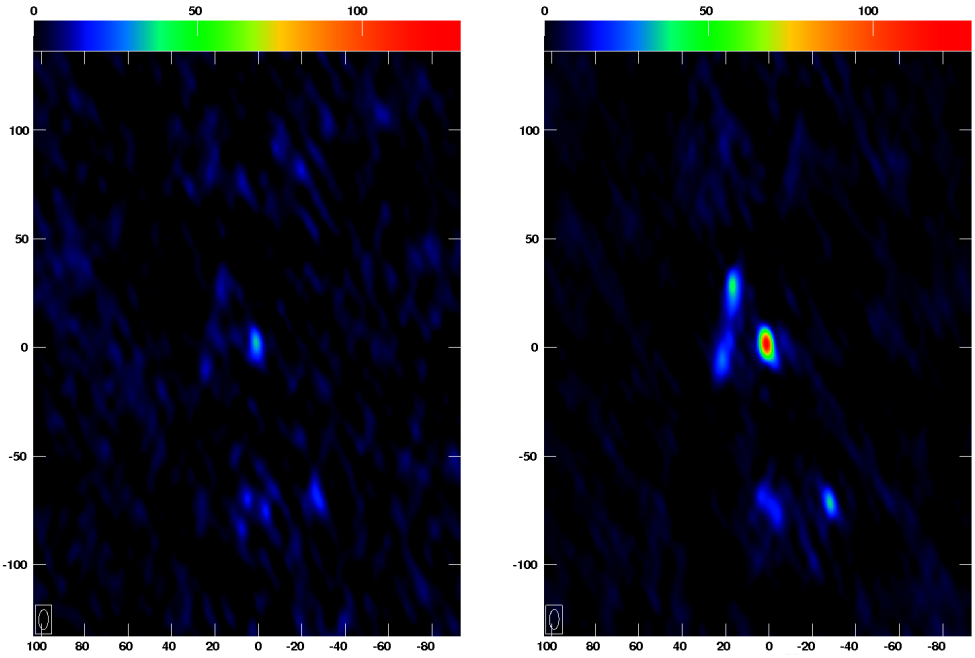
A new model is produced from the corrected visibilities which is then used for the next iteration. These steps are repeated until a good satisfactory sky intensity map is produced. Basic requirements for this process to work are that the total number of free parameters (free gains and parameters in the sky intensity distribution) should not be greater than the number of independent visibility measurements, the data has a good signal-to-noise ratio and the source structure is not too complex (relative to the model).
In Figures 3.7 and 3.8 are shown VLBA 1.7 GHz phase-referenced (left panel) and phase-self-calibrated (right panels) maps of the 4 lensed images of MG J0414+0534 (see Chap. 5). The brightest features in the left panels of Figure 3.7(a) and Figure 3.7(b) were used as starting model. The figure shows just the first iteration, but the improvement in dynamic range is clearly visible.

3.2.3 Bandpass calibration

When observations are made in spectral line mode (Sect. 3.8) a further step is needed in the calibration process in order to compensate for the change of gains with frequency, the so called *bandpass calibration*. The frequency response across

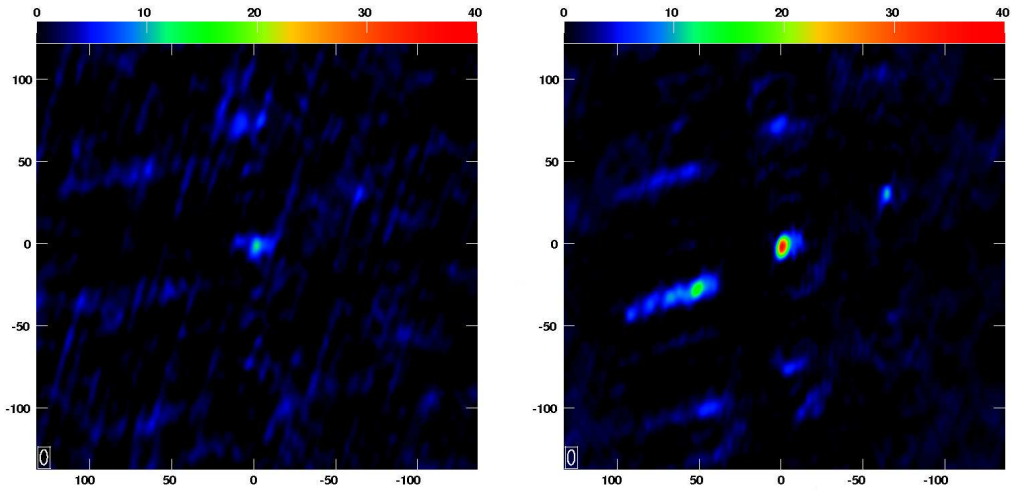


(a)

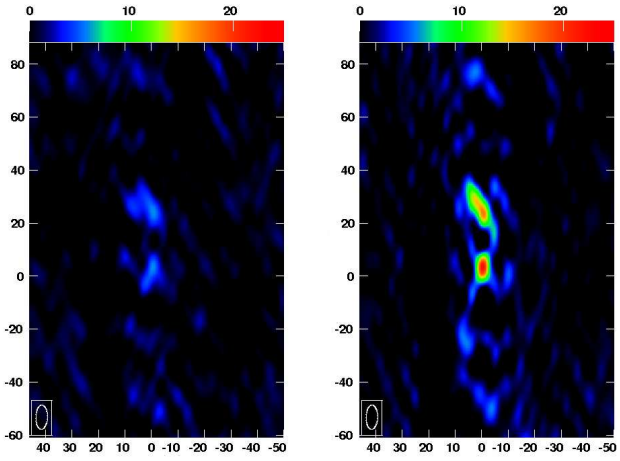


(b)

Figure 3.7: First iteration of self-calibration. Left panels: phase referenced maps; right panels: self-calibrated maps. The figures show the lensed images A1 (figure a)) and A2 (figure b)) of the lens system MG J0414+0534. The axes units are the relative right ascension (east of the phase centre) and declination (north of the phase centre) in milli-arcseconds.



(a)



(b)

Figure 3.8: Continue from Fig. 3.7. The figures show the lensed images B (figure a)) and C (figure b)) of the lens system MG J0414+0534. The axes units are the relative right ascension (east of the phase centre) and declination (north of the phase centre) in milli-arcseconds.

the frequency channels can be determined by observing a strong calibrator source whose spectrum is known over the frequency band.

3.3 Imaging

Equation (3.6) is a discrete approximation of

$$I(l,m)A(l,m) = \int\int \mathcal{V}(u,v)e^{2\pi i(ul+vm)}du\ dv \tag{3.22}$$

where we have taken into account the finite size of the interferometers elements, which have sensitivity dependent to the arrival direction of the radiation. This is expressed by the additional term $A(l,m)$ that describes the sensitivity as function of direction, that is called *primary beam*. From the recorded visibility measurements, at each $(u_k,\ v_k)$, $k = 1,...N$, Equation (3.22) is used to obtain an estimate of the modified sky brightness from the observed visibilities. This is done using the Direct Fourier Transform (DFT) or the Fast Fourier Transform (FFT) methods. FFT methods need the visibility plane to be gridded; because visibilities are not regularly sampled an interpolation needs to be done. Hence, a smoothing function is applied to the observed visibility function, which is then regularly sampled at regular spaced intervals (basically the convolution function is evaluated only on the grid points); afterwards the data are Fourier transformed, and divided by the FT of the convolution. The weighting factors w_k (introduced in Equation (3.4)) are $w_k = R_kT_kW_k$, which are non-zero where visibility points are measured. The coefficients R_k account for the reliability of each data point. This depends on the integration time, the system temperature and the bandwidth for that point. The T_k are given by a smooth tapering function, usually a Gaussian, used to down-weight the data at higher spatial frequencies. This can improve the quality of the image, if the higher spatial frequencies are sparsely sampled and noisy. The W_k describe the density weighting function which can be used to emphasize visibility points from different spatial frequencies, there are two choices for it. The *natural weighting* scheme, which sets the weights as the reciprocal of the variance of the measured visibilities; this is the optimal approach from a statistical point of view, it increases the sensitivity in a map down-weighting the data points which have poor or no measurements. The *uniform weighting* in which weights are set as the inverse of the density of the visibility points within a region of the uv-plane; this scheme down-weights points from smaller spatial frequencies, increasing the resolution. The *robust weighting* scheme is an intermediate approach, in which the weights are divided by the local density of visibilities plus a constant. In this way, one avoids to assigns high weights to noisy points only because the density is low.
Recalling equation(3.7), the dirty map is the convolution of the dirty beam B^d and the true map $I(l,m)$. Let Z be an intensity distribution which correspond to the unmeasured visibilities in the (u,v) plane, then $B_d * Z = 0$. Therefore any $I + \alpha Z$, where α is a number, is a solution of the convolution equation. This implies that there is no unique solution for the deconvolution problem and therefore non-linear methods are needed in order to find the optimal solution.
In the next Section we describe the two predominant algorithms used to find a solution to the convolution equation.

3.4 Deconvolution: the CLEAN algorithm

This algorithm provides one solution to the convolution equation by representing the radio sky as a collection of point sources. It was first implemented by Högbom (1974). It proceeds as follow:

1. The position and the strength of the brightest point in the dirty image are found.
2. The dirty beam multiplied by the peak strength and by a damping factor $\gamma \leq 1$ is subtracted at peak position in the dirty image. The factor γ is called *loop gain*, typically its value is $\sim 1\% - 5\%$.
3. The peak position and its strength are recorded as model component, the so called CLEAN component.
4. These steps are repeated for a fixed number of iterations or until no flux density peak higher than a threshold is found. Both conditions are user-specified. The model subtracted dirty map is now referred to as dirty image.
5. The collection of CLEAN components is convolved with a CLEAN beam, usually an elliptical Gaussian beam with the Full-Width at Half Maximum (FWHM) of the central lobe of the dirty beam. The image produced is the CLEAN image.
6. The remainder of the dirty map is added to the CLEAN image.

An improvement to this algorithm is the Clark algorithm (Clark, 1980) which works in the image and the (u,v) plane. It consists of two cycles, the minor and the major one. In the minor cycle, CLEAN components are selected in the dirty image if their intensity, above threshold fraction of the image peak, is greater than the highest sidelobe of the beam; a Högbom CLEAN is performed until there are no components to be selected in the dirty image. In the major cycle, the CLEAN components are Fourier transformed to the visibility plane, convolved with the sampling function, transformed back to the image plane and subtracted from the dirty map. Further improvement is given by the Cotton-Schwab algorithm in which the subtraction of the CLEAN components is performed on the *ungridded* visibility data. This has the advantage that any gridding errors are eliminated, provided that the Fourier transform of the CLEAN components to the u,v sample is done with high accuracy.
This algorithm is implemented in the Astronomical Image Processing Software (AIPS) used for the data reduction of the lens systems in this thesis.

The Maximum Entropy Method

The Maximum Entropy Method (MEM) reconstructs the brightness distribution by model fitting to the uv-data maximizing the *entropy* of the brightness distribution. One possible form for the *entropy* function is

$$\mathcal{H} = -\sum_k I_k \ln \frac{I_k}{M_k e}, \tag{3.23}$$

where M_k is a "default" image that allows an a-priori knowledge to be used. A low resolution image can be used for this purpose. A comparison between MEM and CLEAN is given in Narayan & Nityananda (1984). Since its invention CLEAN has been the most used method for deconvolution in radio astronomy, but it has not been applied to other fields, probably because of the decomposition into point sources which is not always feasible in other types of images. On the contrary MEM has spread to many different fields. Generally CLEAN is faster for small and simple sources, whereas for extended structures MEM is faster and gives better results.

3.5 Wide-field imaging

When imaging weak sources in a large field of view, where strong background sources are present far from the phase centre, observations must be taken with narrow bandwidths in order to preserve the wide field of view (see Fig. 3.9). In Sect. 3.1.4 we have seen that averaging the visibilities over the bandwidth corresponds to rescaling the image plane in the radial direction respect to the phase centre. When imaging weak sources over a large field of view, the effects of sidelobes from strong background sources would be hard to mitigate if the field of view is not preserved (namely if the data are averaged over the frequency band thus reducing the field of view).
When imaging wide field of view using the two-dimensional FT the $w = 0$ approximation (see Section 3.1.3) is no longer a valid approximation and serious errors may occur which increase as the square of the angular distance from the phase reference point. Following Cornwell & Perley (1992), images of a bigger area are made up of smaller facets. The algorithm is also implemented in AIPS. For each facet, the projected baselines coordinates and the observed phases are rotated to the phase reference position at the facet centre, so that the $w = 0$ condition is still valid for that facet; the sky brightness $I(l, m)$ can now be approximated with a two-dimensional FT and standard CLEAN can be performed (as described in the previous section); note however that for each facet being cleaned there is a local point-spread function which is used to deconvolve the brightness; when a facet is cleaned, the model in that facet is subtracted from the the residual data so the next facet to be cleaned is imaged using the latest residual visibility data; in this way sidelobes of a strong source in one facet are removed from the others before these are cleaned. Figure 3.9 represents the L-band primary beam of the Very Large Array (VLA). At the frequency of 1.4 GHz, it has a size of ~ 30 arcminutes. The target of this observations is a 5 mJy source (see Sect. 4.6). Note the flux scale of the image is in Jansky. It was therefore necessary image all the brighter sources, and remove their sidelobes. In the figure, the white rectangle shows the targeted region.

3.5.1 Multi-scale imaging

The CLEAN algorithm is based on the idea that the sky brightness distribution can be approximated as collection of point like sources (Sect. 3.4). In the *multi-scale* CLEAN, also implemented in AIPS, extended emission is modelled with extended components. This algorithm is well suited when imaging low signal-to-noise extended emission. However some steering by the user is required in order for this algorithm

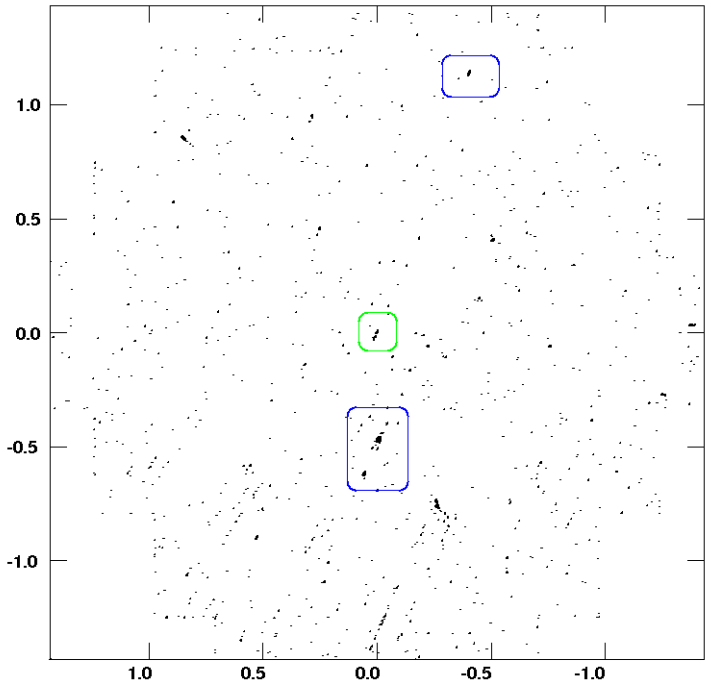


Figure 3.9: Field size ~ 30 arcminutes. The green box represents the target region. The blue ones correspond to regions with flux density at 1.4 GHz of $\simeq 40$ mJy (the southern) and $\simeq 60$ mJy (the northern), which are ≥ 8 times the flux density of our target. The axes units are the relative right ascension (east of the phase centre) and declination (north of the phase centre) in kilo-arcseconds.

to work well. Different flux cut-off scales need to be specified for different scales of emission, such that lower resolution images have higher flux cut-off.

The most precious things in life are not those you get for money

A. Einstein

4

The brightest Lyman Break Galaxy

The work presented in this chapter is based on the refereed publication Volino et al. (2010).
Goal of this project was to study the properties of a high-redshift star forming galaxy from radio observations and to compare them with results from observations at shorter wavelengths, which may suffer of systematic uncertainty due to propagation effects. So far this is the first case that galaxy strong lensing is exploited at radio wavelengths to study a high-redshift source. Moreover, if emission is detected from the multiply lensed star-forming region, the lens galaxy could be studied as well using lens-modelling techniques.
The chapter is organized as follows: in Section 4.1 we introduce the properties of this high-redshift galaxy population. In Section 4.2 we describe the astrophysical framework on which this study is based, namely the radio-FIR correlation. Afterwards, in Section 4.3 we describe how it is possible to study star-forming regions at other wavelengths but radio, and compare the systematic effects for different observing band. In Section 4.4 we introduce the system and the motivations for this project. In Section 4.5 we outline some aspects of lens-modelling relevant for the interpretation of the intrinsic source properties. We then present the radio observations and imaging results in Sections 4.6 and 4.7. In Sections 4.8 and 4.9 we present our results on the radio-derived star forming properties in the high-redshift source. Finally we give our conclusions on this project in Section 4.10.
Throughout the chapter we assume a Λ-CDM cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$.

4.1 Lyman Break Galaxies

Lyman Break Galaxies (LBGs) belong to a population of high redshift objects whose general properties [star-formation rates (SFRs), space density and mass] suggest that they are the progenitors of present-day luminous ellipticals and star-forming galaxies. Their spectra, characterized by a blue ultraviolet (UV) continuum, lines from massive stars, weak Lyα emission, strong interstellar absorption, and dust extinction, are extremely similar to those of nearby star-forming galaxies (Steidel et al., 1996b). Figure 4.1 shows spectra of two LGBs compared to a spectrum of a local starburst.

LBGs have been detected using the *Lyman Break technique*, a colour-selection criteria that exploit the Lyman discontinuity in the UV part of the rest-frame spectral energy distribution (Steidel et al., 1996a,b).

The Lyman discontinuity is an important feature of the UV spectrum of high-redshift galaxies whose photons emitted at $\lambda < 1216$ Å are absorbed by the intergalactic gas. Furthermore, radiation at $\lambda < 912$ Å is also heavily suppressed by the interstellar medium of the galaxy itself. A source of UV radiation at redshift z would not been detected if observed with a filter with central wavelengths $\lambda_1 \lesssim (1+z)912$ Å but only with redder ones. Hence, high-redshift galaxies become hidden when observations go towards shorter wavelengths (see Fig. 4.2). Large sample of LBGs have provided precise clustering measurements showing strong spatial correlation (r_0, correlation length, of order of $\sim 3-4$ Mpc h^{-1}) and thus implying that these galaxies are associated with the more massive dark matter halos [1], (Adelberger et al., 2005; Bielby et al., 2010). Understanding the properties of this high z galaxy population (their clustering properties, colour and star-formation rates) is indeed very important in order to test the paradigm of galaxy formation.

These galaxies are characterized by a wide variety of morphological properties. In general, they are not classified in terms of Hubble types because of the difficulty in identifying their structural components (Giavalisco, 2002).

Detailed investigations of these early episodes of star-formation require the additional magnification provided by gravitational lensing (see Sect. 1.4.2). Unfortunately, an advantageous strong lensing geometry is very rare (lensing probabilities are $\sim 10^{-3}$), and there are only a few cases where studies of an LBG boosted by gravitational lensing have been carried out; for example, MS1512−cB58 (Yee et al., 1996; Pettini et al., 2000; Siana et al., 2008), LBG J213512.73−010143 (Smail et al., 2007; Coppin et al., 2007), the 'arc-core' in the cluster 1E 0657−56 (Nesvadba et al., 2006).

4.1.1 Are radio observations of LBGs sensible?

Star-forming galaxies are dust embedded regions, and therefore unbiased studies are provided by observations in the far-infrared (FIR) to mm bands, where the total stellar and gas content of these dusty environments can be directly probed by measuring the thermal emission of the dust. On the other hand, at radio wavelengths, the non-thermal emission traces recent massive star-formation activity, which gives

[1] The clustering properties of the galaxies are determined from observations of their distribution in the sky.

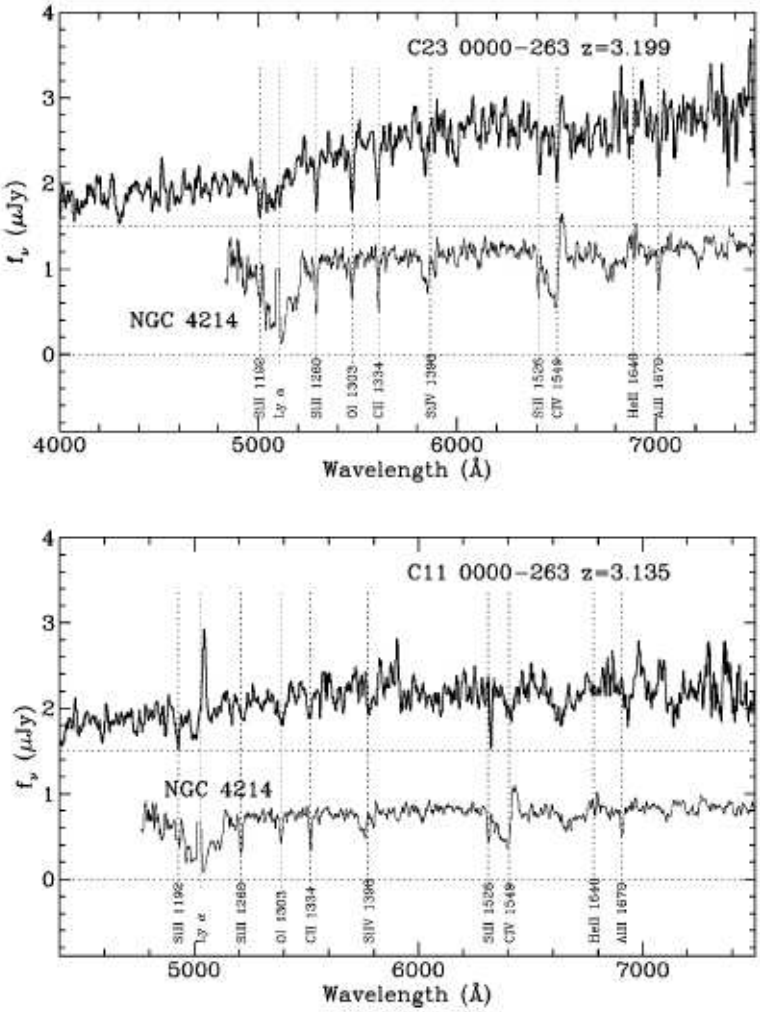


Figure 4.1: A comparison between spectra of Lyman Break Galaxies and the starburst galaxy NGC 4212 is shown. The spectra show strong interstellar absorption lines and blue continuum, denoting that formation of massive stars is taking place. Adopted from Giavalisco (2002).

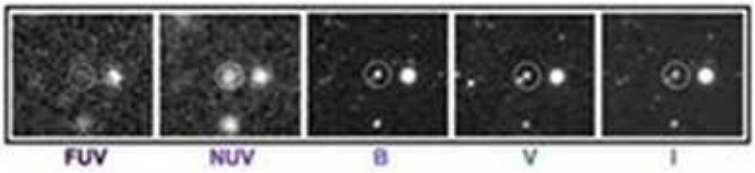


Figure 4.2: Illustration of the Lyman-Break technique: a galaxy observed at all wavelength seems to "disappear" from long (I, infrared; right most panel) to short wavelengths (FUV, far-UV; left most panel). Credit to Burgarella et al. (2006)

us another unobscured view of the dust embedded star-forming galaxies. Based on the FIR-radio correlation (see next Sect.) and a typical SFR of $\sim$ 30 $M_\odot$ yr^{-1} (Shapley et al., 2001), a non-thermal 1.4 GHz luminosity of $\sim 10^{22}$ WHz^{-1} is expected for these galaxies, which implies a sub-mJy flux density. Stacking techniques (Ivison et al., 2007; Carilli et al., 2008) or deep imaging are necessary to detect these objects. In addition to the technical improvements provided by new facilities at cm wavelengths, gravitational lensing is a powerful tool to investigate the radio emission of LBGs. Garrett et al. (2005) started exploiting this tool using clusters of galaxies as lenses. With this work, we want to extend the method to individual galaxies as lenses.

4.2 The FIR−radio correlation

At radio wavelengths thermal and non-thermal emission of galaxies trace recent massive star formation activity, and the galaxy luminosity can be modelled using the SFR of high mass stars as the only free parameter (Condon, 1992). Stars more massive than 8 $M_\odot$ end their life as type II and type Ib supernovae; their remnants are considered to be the source of acceleration of most of the relativistic electrons in galaxies; hence responsible for the non-thermal and thermal radio emission. Massive stars have lifetimes $\lesssim 10^7$ years, while relativistic electrons live probably $\lesssim 10^8$ years; in both cases, the lifetimes are much shorter than the Hubble time, thus the radio emission of a galaxy probes recent processes of star-formation.
The following equation describes the relation between the total radio luminosity of a galaxy and the massive SFR,

$$\mathrm{SFR} = \left\{ \frac{L_\nu}{\mathrm{WHz}^{-1}} \left[5.3 \times 10^{21} \left(\frac{\nu}{\mathrm{GHz}} \right)^{-0.8} + 5.5 \times 10^{20} \left(\frac{\nu}{\mathrm{GHz}} \right)^{-0.1} \right]^{-1} \right\} M_\odot \ \mathrm{yr}^{-1} \tag{4.1}$$

where the first term in the denominator represents the synchrotron radio emission and the second one the thermal emission from the HII regions. For a detailed description of the radio emission in normal galaxies the reader is referred to Condon (1992).
On the other hand, massive stars are thought to be born in dusty giant molecular clouds, such that almost all of the UV radiation is re-emitted in the FIR. This is the basis for the radio-FIR correlation, first discovered and reported by Van der Kruit (1971). Indeed, there is a very tight relation between radio and FIR luminosity at $z < 0.5$ (e.g. Jarvis et al. 2010), with a possible evolution up to $z \sim 4$ (e.g. Garrett

2002; Michałowski et al. 2010; Sargent et al. 2010).

4.3 Indicators of star formation

Information on the star-formation properties of a galaxy come from flux measurements at different wavelengths, such as ultraviolet continuum, optical lines, far-infrared continuum and lines, and radio continuum (non-thermal radiation). These are direct tracers of star formation activity, and are briefly discussed in this section.
In the UV, the continuum is dominated by the emission of young stars, so that the SFR scales linearly with the luminosity, under the assumption that the SFR remains constant over time scales longer than the lifetime of the dominant UV emitting stellar population ($< 10^8$ years). The main advantages of this technique are that it is directly connected to the photospheric emission of the young stellar population and it can be applied to star-forming galaxies over a wide range in redshift. The weak points are its sensitivity to dust extinction and to the assumed IMF. In order to obtain an estimate of the extinction, models are required, which take into account the geometrical dust distributions and the grain shape; alternatively, one needs to make use of the reddening information from the Balmer decrement or infrared (IR) lines (Calzetti et al., 1994; Calzetti, 1997). For high redshift star forming galaxies, the latter is the most accurate.
From early- to late-type galaxies, a comparison of the integrated spectra shows an increase in the strengths of the emission lines (Kennicutt, 1998). The integrated ionizing radiation that is necessary to explain these lines (mostly Hα and Hβ), is due to very massive stars ($M > 10\ M_{\odot}$) whose lifetime is of order of mega-years; it follows that emission lines provide a nearly instantaneous measure of the SFR. The big advantage of this method is its high sensitivity to the massive SFR, while the limitations are its sensitivity to the IMF, uncertainties in the extinction, and the assumption that all of the massive star formation is traced by the ionized gas. Systematic uncertainties due to dust extinction can be compensated by observations at longer wavelengths (continuum FIR), where the star formation is directly traced by the stellar light that is absorbed by dust and then re-emitted beyond a few μm. The efficacy of this method depends on the contribution of young stars to heat the dust, and on the optical depth of the dust itself. However, the weak point of using this method to trace the star formation activity, is that it is sensitive to the choice of the IMF (Kennicutt, 1998).
Finally, the tight correlation between the FIR and the radio (non-thermal) luminosities allows a direct measurement of the SFR from the radio luminosity. At these wavelengths the reddening by dust does not occur, and hence the observed flux density can be converted to a direct measurement of the star-formation activity. However, as discussed in the previous section, the radio emission traces the massive star-forming regions, and hence one needs to use conversion factors when comparing radio-derived to other wavelengths SFR.
Other methods to derive the SFR are based on synthesis modelling which provide relations between the SFR per unit mass or luminosity and the integrated colour of the stellar population. Also in these cases, systematic uncertainties from an incorrect initial mass function, reddening or star-formation history needs to be accounted

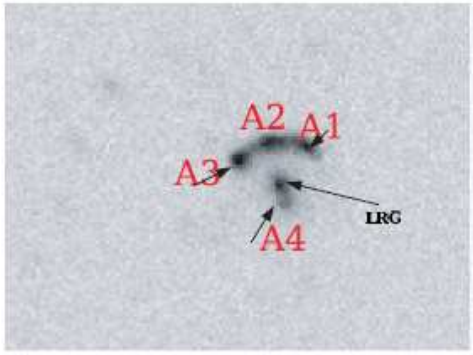


Figure 4.3: The 8 o'clock arc system. The 4 lensed images and the lens galaxy are shown. Credit to Allam et al. (2007)

for. These last methods are mainly used for faint galaxies.

4.4 The 8 o'clock arc

The 8 o'clock arc system was first identified by Allam et al. (2004) while imaging the data of the Sloan Digital Sky Survey Data Release 4, and follow up spectroscopic observations confirmed its lensing nature (Allam et al., 2007) [2]. The two components of the system are SDSS J002240.91+1431110.4, a luminous red galaxy (LRG), and SDSS J002240.78+143113.9 a very blue and elongated arc (hereafter referred to as the LBG). The LRG is at redshift $z = 0.38$ and acts as the lens. The lensed arc of the LBG subtends $9.6''$, and consists of three components: A1, A2 and A3 (with $i = 20.13$, 20.11 and 20.21, respectively[3]; Allam et al. 2007). The SPIcam[4] g–band image of Allam et al. (2007) also shows a faint fourth component, identified as the counterimage, 5 arcseconds away from the main arc and on the opposite side of the lens galaxy (Fig. 4.3). The redshift of the arc was measured to be $z = 2.73$ (Allam et al., 2007; Finkelstein et al., 2009).
Even taking into account a lensing magnification of ≈ 12.3 (from the lens model of Allam et al. 2007), the arc is 2.6 mag more luminous ($\approx$ a factor of 11 in luminosity) than L_* for LBGs (where L_* is the characteristic Schechter luminosity for LBGs; Steidel et al. 1999). This suggests that the system is going through a vigorous process of star formation. Allam et al. (2007) estimated a SFR ~ 230 $M_\odot$ yr^{-1} using the relation given in Pettini et al. (2000) scaled to MS1512−cB58. From their optical and near-infrared (NIR) studies, Finkelstein et al. (2009) found a more robust dust-corrected and de-lensed SFR of 266 ± 74 $M_\odot$ yr^{-1}. Their result confirms that this system is undergoing a very active process of star formation, and shows that the SFR is higher than $\sim 85\%$ of the high-redshift LBGs studied by Shapley et al. (2001, 2005).

[2]The authors have named the system 8 o'clock arc from the time of its discovery.
[3]Allam et al. 2007 obtained exposures using the SDSS *gri* filters.
[4]SPIcam CCD imager is mounted on the Astrophysical Research Consortium 3.5 m telescope at the Apache Point Observatory.

Table 4.1: SDSS photometry of the 8 o'clock arc system. The table has been adopted from Allam et al. (2007). The estimated error on the magnitudes is ±0.1

	g	r	i
LRG	20.14	18.62	18.16
A1	21.18	20.21	20.13
A2	20.99	20.40	20.11
A3	21.27	20.68	20.21
A4	~ 22	–	–
arc	19.95	19.22	22.40

4.4.1 NVSS identification of the 8 o'clock arc and motivations for our study

In the radio, the 8 o'clock arc lens system is coincident with an NVSS[5] (45 arcsecond resolution; Condon et al. 1998) radio source with a 1.4 GHz flux-density of ~ 5 mJy. If due to star formation the flux density at 1.4 GHz would imply a huge SFR $\sim$ 11 000 $M_\odot$ yr^{-1} (using Equations 4.2 and 4.3), which would contradict the estimates from the optical and NIR spectroscopy. Taking the optical and NIR derived SFR of ~ 270 $M_\odot$ yr^{-1} and a total magnification of $\mu \sim 12$, the gravitationally lensed 1.4 GHz flux density for the LBG is expected to be just ~ 0.12 mJy.

4.5 Lens modelling

Here we highlight some aspects of the lens models that are relevant for the interpretation of the de-lensed properties of the 8 o'clock arc system.
Allam et al. (2007) show that a singular isothermal ellipsoid (SIE) model for the lens galaxy provides a good fit to the optical data. Their model predicts the source position to be within the cusp (see Fig.4.4) of the tangential caustic curve, and the total magnification to be $\mu = 12.3^{+15}_{-3.6}$, where each of the three arc components has $\mu_{\rm i} \gtrsim 4$ (Allam et al., 2007). However lensing theory predicts that for this image configuration the magnifications of the A1 and A3 components (which lie outside the critical curve) should be equal, and their sum should be equal to the magnification of A2, which is expected to have the largest flux-density. This relation is known in gravitational lensing studies as the *cusp relation* (Schneider & Weiss, 1992). Using a magnification of $\mu_{\rm i} \sim 4$ (see below) for each component could lead to biased estimates of the de-lensed emission from the "8 o'clock arc" components and thus all of the parameters derived from it, e.g., the SFRs. This consideration is necessary here since we compare our results to those of Finkelstein et al. (2009) who derive the SFR from A3 and A2 components and correct for a total magnification $\mu_{\rm (A3+A2)} \sim 8$. Furthermore, the observed optical flux ratios of the lensed images indicate that the arc components have a similar magnification (see Table 4.1); this is pointing out a violation of the cusp relation which might be explained by small scale (milli-arcsecond scale) structures in the lens galaxy or extinction. Hence, the lens model would need further investigation. Nevertheless, in order to compare our

[5]The National Radio Astronomy Observatory Very Large Array Sky Survey.

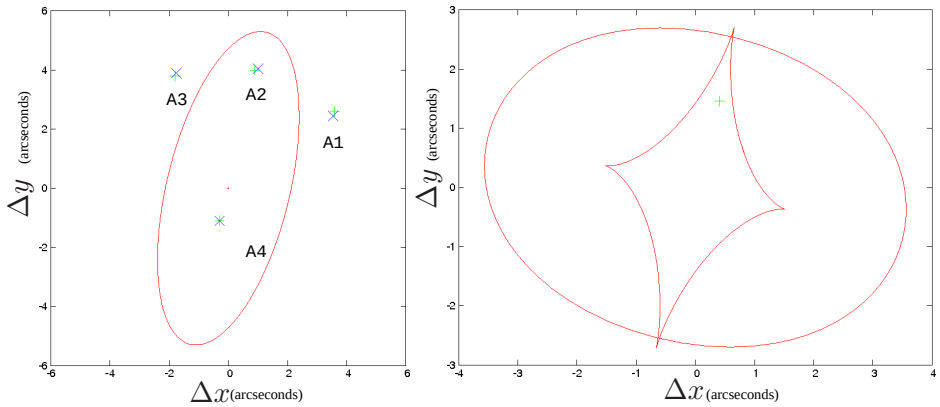


Figure 4.4: A simple Singular Isothermal Ellipsoid (SIE) has been fitted to the optical positions of the arc components (Table 4.1). The left (right) panel shows the image (source) plane. In the image plane, the green crosses mark the observed image positions and the blue the modelled ones. In the source plane, the cross indicates the modelled source position.

results with Finkelstein et al. (2009), we have corrected the flux density using $\mu_{\rm i} \sim 4$ per component.

Another aspect of this lensing system which needs further investigation, which was also pointed out by Allam et al. (2007), concerns the mass distribution outside the main lensing galaxy. The image separation between the arc and the faint counter arc (see Fig. 4.3) is $\simeq 5$ arcsec. Oguri et al. (2005) show that for lens systems with image separations of θ larger than 3 arcsec, the contribution of the lens environment may be important, with an external convergence larger than 10%. Moreover, Oguri (2006) has predicted that the contribution of dark matter halos surrounding the main lensing galaxy is significant in systems with $\theta \gtrsim 3$ arcsec and has to be taken into account in the lens modelling. Indeed, recent studies of group-scale gravitational lens systems have found that the lens environment is important, changing the convergence and producing density profiles that can be either shallower or steeper than isothermal (McKean et al., 2010; Thanjavur et al., 2010; Verdugo et al., 2010). A simple isothermal model for the lens galaxy of the 8 o'clock arc may not be the best parametrization for the mass model. However, given that the lens system is dominated by the massive LRG, we expect the model to be close to isothermal, and hence, for the magnification to be close to what was found by Allam et al. (2007).

4.6 Very Large Array observations of the system

We observed the 8 o'clock arc in November 2007 using the VLA in B-configuration at 1.4 GHz [synthesized half power beam width (HPBW) of ~ 3.9 arcsec] and at 5 GHz (synthesized HPBW ~ 1.2 arcsec), and again at 1.4 GHz one year later, when the telescope was in A-configuration (HPBW ~ 1.4 arcsec). The higher resolution and more sensitive 1.4 GHz data (second observing run) were necessary to disentangle the weak radio emission of the lensed LBG from the lens (see Fig. 4.5 and 4.6). For the first observing run, the total time on source was 6 hours, divided equally

between the two observing frequencies, while one year later we observed the system for 7 hours. In order to reduce the potential effect of bandwidth smearing, at 1.4 GHz we used the correlator in spectral line mode, using 25 MHz of bandwidth and 3.125 MHz wide channels. At this frequency a larger field of view is required for imaging (in order to remove the response of unwanted sources), and therefore the use of a narrower bandwidth is required to reduce chromatic aberrations which cannot be neglected. In addition to this, continuum observations in spectral line mode are usual preferred at this observing frequency as narrow interferences, which can be a problem at this or longer wavelength, can be identified. For both runs, 3C 48 and $0010 + 174$ were used as the flux-scale and gain (amplitude and phase) calibrators, respectively. The data were taken in 2 IFs, but for both 1.4 GHz observations, one of the IFs was corrupted, and therefore was not used in imaging. For the first run (November 2007), at 5 GHz we used a switching cycle of 1.5 and 5 min between the calibrator (0010+174) and the lens system, while at 1.4 GHz we used 1.5 and 20 min; in the second run (November 2008), the array configuration was larger (A-array), therefore we used a cycle-time of 1.5 and 10 min in order to compensate for the less stable phases.
The data reduction was performed using the AIPS (Astronomical Imaging Processing Software) package, provided by the National Radio Astronomy Observatory. The calibration strategy was the following: the flux-scale was set using 3C 48; afterwards amplitude and phase calibration was performed using 0010+174. In both cases, for the 1.4 GHz data, bandpass calibration was necessary in order to identify variations of amplitudes and phases across the band; 3C 48 was used for this. This step was not required for the 5 GHz data which were taken in continuum mode. At 1.4 GHz the primary beam has a half-width at half maximum of ~ 18 arcminutes while at 5 GHz it is ~ 5 arcminutes, therefore wide field imaging techniques and deep cleaning were necessary in order to map all of the sources in the field-of-view, and remove their interfering sidelobes from the region of interest around the lens system. At both frequencies the whole primary beam was mapped.

4.7 Imaging results

In Figure 4.5 and 4.6, we show the 1.4 and 5 GHz images of the 8 o'clock arc system that were taken with the VLA. The crosses indicate the positions of the lensed images and the lens galaxy. As already mentioned, in the NVSS catalogue the lens system is identified with a radio source that has a 1.4 GHz flux-density of ~ 5 mJy at 45 arcsec resolution (Condon et al., 1998).
Figure 4.5 shows the L-band map from the first observing run. At the optical position of the system we are interested in, this radio map shows extended emission associated with both components of the lens system, the lens and the arc. Towards the south-east direction, a second extended component is visible. The integrated flux density over region N is 2.44 ± 0.13 mJy, over region S it is 2.00 ± 0.17 mJy, well in agreement with the NVSS detection. However the resolution of ~ 4 arcsecond is not enough to disentangle the emission of the lens and the arc.
Figure 4.6(a) shows the C-band map from the same observing run (B array). The higher resolution of ~ 1.2 arcsecond shows emission extending towards the north-

west direction. The peak of the radio emission is coincident with the optical position of the lens. The south-east component seems now to be aligned with the northern extended structure, implying an AGN morphology. No radio emission above the 3σ level is detected at the optical position of the lensed arc components.
Imaging results from the L-band observations at higher resolution are shown in Fig. 4.6(b). These observations better resolve the radio emission of the system which we find dominated by an AGN associated with the lens galaxy. The peak of the radio emission is coincident with the optical position of the lens. In Figure 4.8 we show the spectral index map of the system between 1.4 and 5 GHz. We find that the peak in the radio emission has a spectral index of $\alpha \sim +0.25$, where $S_\nu \propto \nu^\alpha$, which is consistent with a synchrotron self-absorbed radio source. Therefore, we associate the brightest radio component, which is coincident with the optical position of the lensing galaxy, as the compact core of the AGN. At both frequencies, we see that there is radio emission that propagates from the core region towards the north-west. This radio emission has a steeper spectral index of $\alpha \sim -0.80$, consistent with an optically-thin radio jet. This would suggest that the AGN has a morphology that is consistent with an FR I type radio galaxy. Extended emission is also seen opposite to the jet direction with respect to the AGN core (towards the south-east), which could be from a radio lobe; the steep spectral index ($\alpha \sim -0.70$) and morphology agree with this classification. The integrated flux density of the system at 1.4 GHz is ~ 5 mJy, in agreement with the NVSS flux density.

4.7.1 Radio flux measurements

In Figure 4.7, we show an overlay of the 1.4 GHz radio contours and the SDSS *i*-band image. We find that the radio jet covers about two thirds of the arc, contaminating the radio emission from components A1 and A2 of the LBG. However, component A3 of the arc is relatively free from the emission of the radio jet and can be used to estimate the SFR of the LBG. The lensed 1.4 GHz flux-density at the position of A3 is 80 $\pm$ 36 μJy beam^{-1}, which is a 2.2 σ detection, and therefore, is not statistically significant. Based on the rms noise in our VLA map, we place a firm 3 σ upper limit of 108 μJy beam^{-1} for the 1.4 GHz flux-density of component A3. We have also measured the flux-density of component A2 and the total flux-density of the arc (A1 + A2 + A3), but these data are contaminated by the emission from the radio-jet. All our measurements are presented in Table 4.2.

4.8 Radio-derived star formation rate

Assuming the FIR-radio correlation the radio emission of a star-forming region can be converted in rate of massive stars ($M > 5$ M$_\odot$) expected to form. However in order to make a fair comparison between the results from the optical/NIR and the radio data for the 8 o'clock arc we must measure the SFR using the same initial mass function (Salpeter, 1955) and down to the same stellar mass limit (≥ 0.1 M$_\odot$). This needs the relation between the SFR and the radio emission to be modified. This was done by Yun et al. 2001 and found to be,

$$\frac{\mathrm{SFR}}{(\mathrm{M}_\odot\ \mathrm{yr}^{-1})} = (5.9 \pm 1.8) \times 10^{-22}\ \frac{L_{1.4\ \mathrm{GHz}}}{(\mathrm{W\ Hz}^{-1})}. \tag{4.2}$$

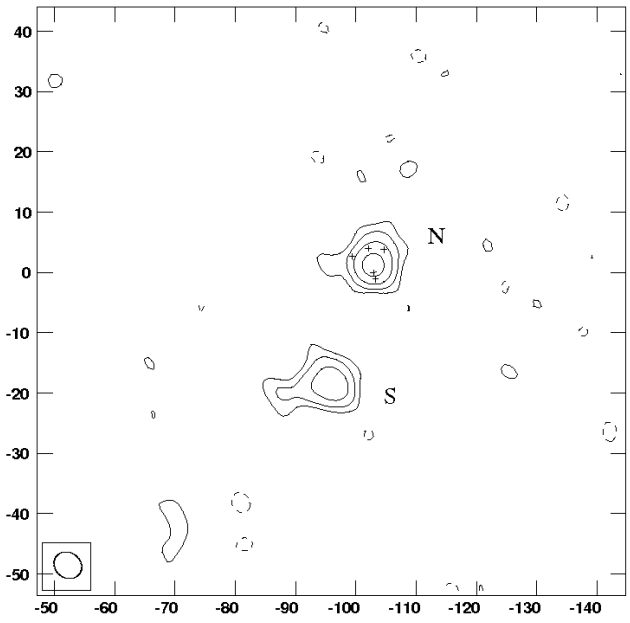
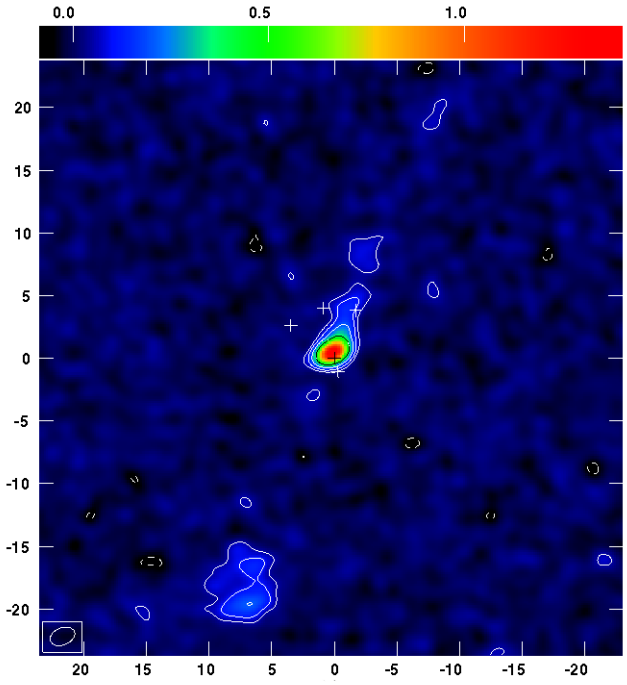
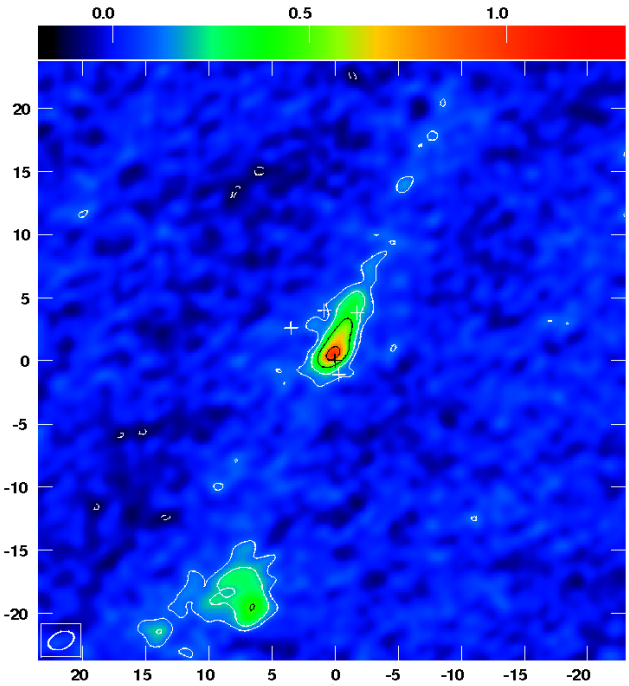


Figure 4.5: L-band, B array. The restoring beam has size of 4.8×4.2 arcsec at a position angle of 52. The contours are shown at $(-3,3,6,12,24)\times 42\mu$Jy, the rms map noise. The theoretical rms of the map is 35 μJy beam^{-1}. The axes units are the relative right ascension (east of the phase centre) and declination (north of the phase centre) in arcseconds. When these observations were carried out, there was an offset problem with the VLA correlator for 25 MHz bandwidth spectral-line observations, a fictitious source would appear after few hours of integration time at the phase centre. The problem was avoided shifting the phase centre by few arcsecond. In our case this was 100 arcsecond.



(a)



(b)

Figure 4.6: Panel **a)**:C-band, B-array. Panel **b)**: L-band, A-array. **a)**The 5 GHz map has a restoring beam of 2.1×1.4 arcsec at a position angle of −68.1. The contours are shown at(−3,3,6,12,24)×17μJy, the rms map noise. **b)**The 1.4 GHz image has been restored with 2.1 × 1.4 arcsec beam at a position angle of −67. The contours are shown at (−3,3,6,12,24)×36 μJy, the rms map noise. The theoretical rms of the 1.4 and 5 GHz maps are 25 and 15 μJy beam^{-1}, respectively. The colour-scale is in mJy beam^{-1}. The axes units are the relative right ascension (east of the phase centre) and declination (north of the phase centre) in arcseconds.

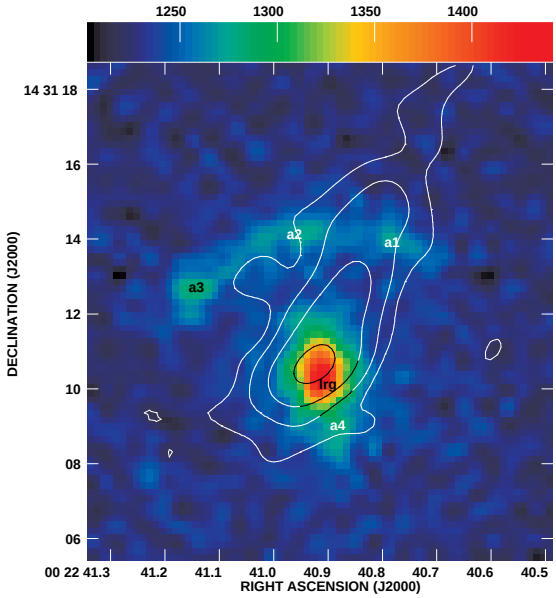


Figure 4.7: The 8 o'clock arc system: superposition of radio continuum 1.4 GHz contours on optical image. The contours are shown at the same levels as the top Fig. 4.6.

The rest-frame luminosity at 1.4 GHz is calculated using,

$$L_{1.4\ \mathrm{GHz}} = 4\pi D_L^2 S_{1.4\ \mathrm{GHz}} (1+z)^{-(1+\alpha)} \tag{4.3}$$

where D_L is the luminosity distance, α is the spectral index of the LBG (assumed to be -0.8) and $S_{1.4\ \mathrm{GHz}}$ is the observed monochromatic flux density. The factor $(1+z)^{-(1+\alpha)}$ accounts for the k-correction (Schmidt & Green, 1986). To calculate the SFR of the LBG we must also correct the observed flux-density for the lensing magnification. The lensing model constructed by Allam et al. (2007) gives a total magnification of 12.3 for the arc, with each component (A1, A2 and A3) having a magnification of about 4. It is likely that there could be differential magnification across the full extent of the arc given its large size. However, since we have at best an upper-limit for the radio emission from the arc and that the resolution of our radio data matches the size of the individual components, we have adopted the magnifications found by Allam et al. (2007) in our calculations. This gives also consistency when comparing the radio-derived SFR and the optical one. The estimated unlensed luminosity and SFR of the LBG is given in Table 4.2.

4.9 Radio emission from the 8 o'clock arc

We have observed the 8 o'clock arc gravitational lens system to investigate the source of the excess radio emission at 1.4 GHz. Previous observations from NVSS found a radio source with a flux-density of ~ 5 mJy. If this flux density is associated with the lensed LBG, the implied SFR of $\sim 11000\ \mathrm{M}_\odot\ \mathrm{yr}^{-1}$ is in conflict with the optical/NIR derived value of $\sim 270\ \mathrm{M}_\odot\ \mathrm{yr}^{-1}$. Three possible explanations could account for this;

Table 4.2: A summary of the radio emission observed from the 8 o'clock arc lens system. From our VLA observations we quote the observed (gravitationally lensed) flux-densities and the rest-frame (unlensed) luminosities at 1.4 GHz. We give the estimated SFRs, corrected for the lensing magnification. For comparison, we also show the data from NVSS and the expected radio emission based on the optical/NIR derived SFR.

Component	μ [a]	$S_{1.4\ \mathrm{GHz}}$ (μJy)	$L_{1.4\ \mathrm{GHz}}$ (W Hz^{-1})	SFR ($M_\odot$ yr^{-1})
A3 ($2.2\,\sigma$ detection)	4	80 ± 36 [b]	$(9 \pm 4) \times 10^{23}$	560 ± 300
A3 ($3\,\sigma$ upper limit)	4	≤ 108 [b]	$\leq 1.3 \times 10^{24}$	≤ 750
A2 (including jet)	4	95 ± 36 [b]	$(1.1 \pm 0.4) \times 10^{24}$	660 ± 310
A1+A2+A3 (including jet)	12	430 ± 70	$(1.7 \pm 0.3) \times 10^{24}$	1000 ± 350
All (NVSS)	12	4700 ± 500	$(1.9 \pm 0.2) \times 10^{25}$	11000 ± 3500

Prediction from optical/NIR-based SFR

SFR ($M_\odot$ yr^{-1})	$S_{1.4\ \mathrm{GHz}}$ (μJy)	$L_{1.4\ \mathrm{GHz}}$ (W Hz^{-1})
270±75	120 ± 50	$(4.6 \pm 1.9) \times 10^{23}$

[a] magnification correction
[b] flux density per beam

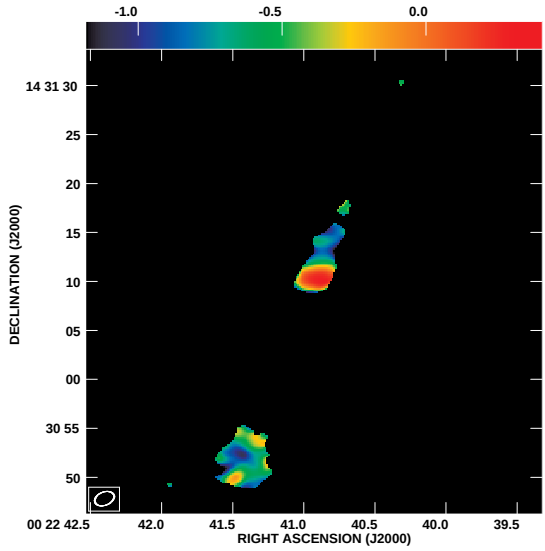


Figure 4.8: Spectral index map between 1.4 and 5 GHz, where $S(\nu) \propto \nu^{\alpha}$. The map has been made using only the emission detected at $\geq 3\sigma$ in the 1.4 and 5 GHz images. The colour-scale is the spectral index.

i) there is a radio-loud AGN within the lensing galaxy, ii) there is a radio-loud AGN associated with the lensed LBG and iii) there is a genuine discrepancy between the optical/NIR and radio methods for determining the SFR. Our observations with the VLA have shown that there is a radio-loud AGN within the lensing galaxy that dominates the radio emission from the 8 o'clock arc gravitational lens system. In principle, our observations should have been sensitive enough to detect the radio emission from the LBG as the predicted flux-density from the arc is 120 μJy. Due to losing half of the data from a corrupted IF and the radio jet from the AGN covering most of the arc, we could only place a limit on the SFR from the uncontaminated region of the lensed arc. Our limit of ≤ 750 $M_\odot$ yr^{-1} ($3\,\sigma$) is consistent with the SFR found from the optical/NIR data.
We were unfortunate that this system has a radio-loud lensing galaxy whose emission covered part of the LBG; from surveys of lens systems at radio wavelengths about 1 in 10 lensing galaxies, which are typically massive early-types, are radio-loud. For example, the Cosmic Lens All-Sky Survey (CLASS; Browne et al. 2003; Myers et al. 2003) found 22 gravitational lenses and two systems had radio-loud lensing galaxies (B2108+213; McKean et al. 2005; More et al. 2008, and B2045+265; Fassnacht et al. 1999; McKean et al. 2007).

4.10 Radio emission from other LBGs and future prospects

Carilli et al. (2008) showed the first robust statistical detection of sub-μJy radio emission for a sample of high redshift LBGs from the COSMOS field[6] at $z \sim 3$. Their results found an average flux density of 0.90 ± 0.21 μJy at 1.4 GHz, which implies a total SFR of $\simeq 31$ $M_\odot$ yr^{-1}, based on the radio–FIR correlation for low redshift star-forming regions. The comparison of this result with those obtained from UV data, for star-forming galaxies at the same redshift, showed a discrepancy between the SFR implied by the non-thermal radio luminosity and that derived from the UV spectra. In particular, the authors find that the ratio of radio- to UV-based SFRs is 1.8, indicating either a smaller dust attenuation factor (standard values for LBGs are ~ 5) or an attenuation of the radio luminosity to SFR conversion factor at $z \sim 3$. Discriminating between these two possibilities requires a deeper look into the interaction between CMB photons and relativistic electrons at such redshifts (increased electron cooling due to Inverse Compton scattering off the CMB; Carilli et al. 2008), and into the properties of LBGs. The radio-FIR correlation may as well require a deeper insight as the radio-derived SFRs rely on the assumption that this correlation is as tight as for local galaxies. A discrepancy between UV and radio-derived SFR is also found by Ivison et al. (2007), in their 1.4 GHz VLA survey of starburst galaxies up to $z \leq 1.3$.

For lensed systems, Spitzer IR spectroscopy and photometry of two LBGs, namely, MS 1512−cB58 (Siana et al., 2008) and the LBG J213512.73 − 010143 (also known as the Cosmic Eye; Siana et al. 2009) show that the UV spectral slope overpredicts the reddening by dust and thus the SFR measurements. In particular, for the system MS1512−cB58 ($z = 2.73$), Siana et al. (2008) find a SFR of $\sim 20 - 40$ $M_\odot$ yr^{-1},

[6]COSMOS is an Hubble Space Telescope project to survey a 2 degrees equatorial field to study formation and evolution of galaxies as function of redshift.

consistent with the SFR derived from the dust corrected Hα luminosity, but 5 times lower than the UV-derived SFR. For the Cosmic Eye ($z = 3.074$), Siana et al. (2009) find that the SFR inferred from the IR luminosity is 8 times lower than that predicted from the rest-frame UV properties. In addition, for the Cosmic Eye, CO studies indicate a SFR ~ 60 $M_\odot$ yr^{-1} (Coppin et al., 2007). These numbers imply that calculating a radio continuum dust unbiased SFR for a sample of star-forming galaxies could help resolve the conflicting SFRs that have been found from IR and UV studies.

The new generation of radio facilities, such as the Expanded Very Large Array (EVLA), will have significantly better surface brightness sensitivity due to improved receivers and larger instantaneous bandwidths. L-band observations with 1 GHz bandwidth using the EVLA will reach sensitivities of ~ 8 μJy in 1 hour, which corresponds to a SFR of ~ 220 $M_\odot$ yr^{-1} at redshift 2.7. To reach the average SFR for LBGs of ~ 30 $M_\odot$ yr^{-1} (Shapley et al., 2001, 2005) will require of order 60 hours of integration. Clearly, studies of lensed LBGs will also benefit from the increased sensitivity. The added advantage is that the lensing magnification will allow more systems to be observed over a shorter amount of time and give higher-resolution imaging of the structure of these galaxies. There has recently been a large increase in the numbers of lensed LBGs being found making the prospects of detailed studies of meaningful samples of these galaxies possible in the near future.

A theory is something nobody believes, except the person who made it. An experiment is something everybody believes, except the person who made it.

A. Einstein

5

The non-smooth mass distribution in the gravitational lens MG J0414+0534

Lensed QSO are ideal laboratories to study CDM substructures with strong gravitational lensing. If the foreground mass distribution is non-smooth on scales smaller than $10^{8\div9}M_{\odot}$ independent bends in the jet images on milli-arcsecond scales may be seen. The system discussed in this chapter, the gravitational lens MG0414+0534, is one example of multiply imaged QSO, that at high angular resolution may reveal the presence of mass substructures. Moreover, for this system, existing lens models (Ros et al., 2000; Minezaki et al., 2009) predict a parity-dependent flux-ratio anomaly, which can be used to address the study of cold dark matter substructures.
The system deserves a special investigation also because the lensed QSO was discovered to be the host of a powerful water maser (Impellizzeri et al., 2008), with a luminosity L$\sim 10^4 L_{\odot}$. Up to now, this is the only lensed water maser at a cosmological distance ($z > 1$) (McKean et al., 2011).
In this chapter we present the latest VLBI observations and results on lens modelling constrained by these data. Main goal of this project was the modelling of the foreground mass distribution using the information contained in the extended lensed structures revealed by the observations; for this target this is the first time that the modelling approach mentioned above is used. Note that the results of this study can be used to apply lens-modelling techniques with full source reconstruction (as explained in Chapter 2) which would provide a unique insight into the properties of the lensed source.
The chapter is organized as follows: in Section 5.1 we introduce the system; in Sections 5.2 and 5.3 we review the previous radio observations and existing lens models. In Section 5.4 we give details on the new VLBI observations, the calibration steps and present the imaging results. Afterwards, in Section 5.5 we explain the lens modelling strategy, and describe the set of parameters and constraints. We discuss details and results of the modelling in Sections 5.6 and 5.7. In Section 5.8 we discuss our results and compare with other studies. Then, in Section 5.9 we study the origin

of the emission lensed into the arc, and in Section 5.9 describe a new model we have tested for one of the mass components of this system.
Throughout the chapter we assume a Λ-CDM cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$.

5.1 The gravitational lens system MG0414+0534

The system consists of a main elliptical lens galaxy (Schechter & Moore, 1993) at redshift $z = 0.9584 \pm 0.0002$ (Tonry & Kochanek, 1999) and four images of a $z = 2.639 \pm 0.002$ (Lawrence et al., 1995) QSO separated by up to 2 arcseconds. It was discovered in the VLA-MIT survey (Bennett et al., 1986) and later on found by other lens surveys (CLASS, and JVAS; Myers et al. 2003). Hewitt et al. (1992) identified this system as gravitational lensing candidate from observations with the VLA telescope (Hewitt et al., 1992) with an angular resolution of $0.4''$. They found that at radio wavelengths the system is resolved into four components, two brighter ones separated by $0.4''$ and two fainter ones, the latter are separated about $2''$ from each other and from the pair. Based on the relative angular positions and the flux ratios the authors argued the lensing nature of this system. They did model the lens as a singular isothermal ellipsoid, but the final fit was rather poor; indeed, in their paper they quote it only as "lens candidate".
Later on, the system was observed at optical wavelengths by Schechter & Moore (1993). Goal of these observations was accurate photometry for the system components, to establish the presence of optical variability. The optical flux-ratios were found significantly different from the radio measurements, especially for the close pair A1−A2 in excess of a factor of two. They concluded that the difference could be explained by the extended structure of the lensed source (revealed by Hewitt et al. 1992). In their paper indeed, they *speculate* on the possibility that the pair *A1−A2 lies close to a caustic, and therefore different parts of the extended source might be magnified by different amounts.*
Figure 5.1 shows a combined optical and infrared Hubble-Space-Telescope HST images from the CfA-Arizona Space Telescope LEns Survey (CASTLES) (Kochanek et al., 1999). The lens galaxy is labelled as G, the four lensed images are A1, A2, B and C. The HST images show extended structures from a partial Einstein ring, connecting the pair A2-B. The arc has a significantly bluer colour then the image cores. The colour differences among the four images, and along the arc, indicate that the reddening is more likely intrinsic to the host galaxy of the QSO rather than the lens galaxy (Falco et al. 1997; see also discussion below). The sixth object is most likely not related to the lensed source, but rather associated with the lens or being a lens itself. It was shown by Falco et al. (1997) that lens models which only include the main lens cannot account for the properties of the lensed images (see Sect. 5.3). More recent studies (Ros et al., 2000; Minezaki et al., 2009) have shown that better fits are obtained if object-X is included in the lensing configuration, with the observed positions of the four images being well explained by a lens model consisting of both the main lens and object-X.

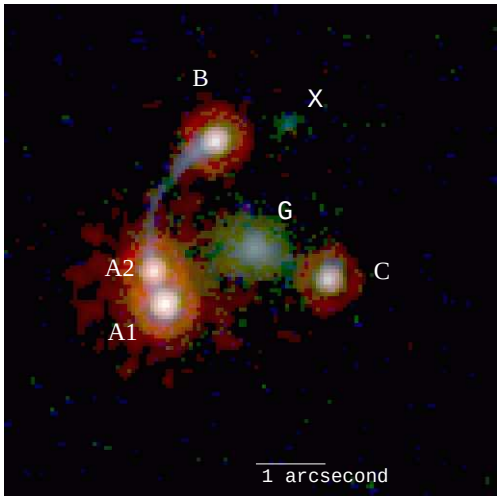


Figure 5.1: HST image of the gravitational lens MG0414+0534

Object-X

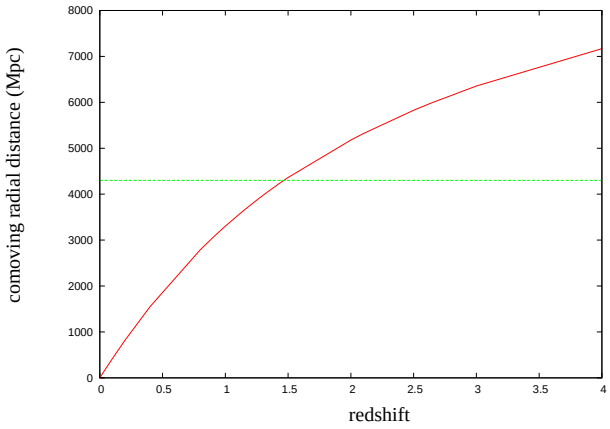
So far, observations have not constrained the redshift of this object, and for lens studies the assumption that it is at the same redshift of the lens galaxy has always been made. We consider three possibilities for its redshift. If object-X lies at the same redshift as the source it would be expected to have at least another counter-image for which there is no evidence in the HST map and it is unlikely that it has been missed. Object-X could also be at an intermediate redshift, $z_{\rm x}$ between the lens and the source. In this case the observed image configuration allows us to say that $z_{\rm x} \lesssim 1.5$. We explain this statement in the following: let us assume that object-X is the only one image of a source at redshift $z_{\rm x}$; for an axi-symmetric mass distribution, we can say that the radius within which we expect multiple images of the source is $\Delta\theta \simeq 2\theta_{\rm E}({\rm X})$, where $\theta_{\rm E}({\rm X})$ is the Einstein radius. Hence, given that we only see one image of object-X at $\theta_{\rm X} \simeq 1.2''$, we say that $\theta_{\rm E}({\rm X}) \lesssim \theta_{\rm X}/2$. From the observed image configuration the Einstein radius of the lens galaxy is $\theta_{\rm E} \simeq \theta_{\rm X}$. It follows that $\theta_{\rm X}({\rm X}) \lesssim \theta_{\rm E}/2$. Recalling Equation (1.46), the relation between the angular diameter distances and the size of the Einstein radius is $\theta_{\rm E} \propto D_{\rm ds}/D_{\rm s}D_{\rm d}$, and $\theta_{\rm E(X)} \propto D_{\rm dx}/D_{\rm x}D_{\rm d}$. Hence we can write,

$$\frac{D_{\rm dx}}{D_{\rm x}D_{\rm d}} \lesssim \frac{1}{2}\frac{D_{\rm ds}}{D_{\rm s}D_{\rm s}}. \tag{5.1}$$

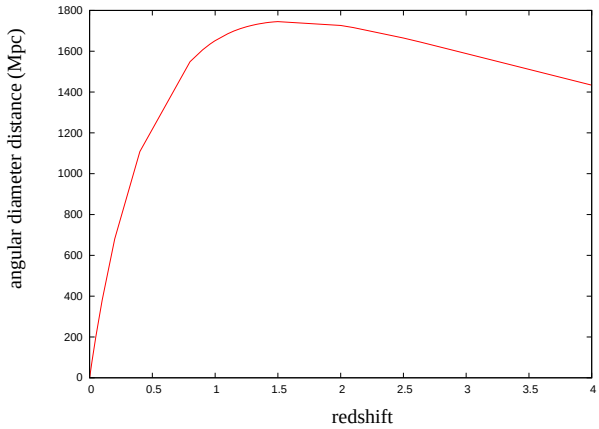
Since we know the redshift of the lens galaxy and of the QSO, we can re-arrange Equation (5.1), and using Equations (1.19) and (1.20), it follows

$$1 - \frac{\chi_{\rm d}}{\chi_{\rm x}} \lesssim 0.233. \tag{5.2}$$

In the above equations, the subscripts refer to the lens galaxy (d), object-X (x) and the source (s), and χ is the comoving distance. We can now solve Equation (5.2) with respect to $\chi_{\rm x}$, and obtain that $\chi_{\rm x} \lesssim 4300$ Mpc. In Figure 5.2 we read that this condition corresponds to $z_{\rm x} \lesssim 1.5$. The third possibility is that object-X is



(a)



(b)

Figure 5.2: The top (bottom) panel shows the comoving radial distance (angular diameter) distance as function of redshift for a flat Universe, with $\Omega_{\rm m} = 0.3$, $\Omega_{\Lambda} = 0.7$ and $H_0 = 70$ km s^{-1} Mpc^{-1}. In the top panel, the green line identifies the upper limit for the redshift of object-X so that Equation 5.1 is satisfied.

at the same redshift than the lens galaxy or in its foreground. Not knowing its redshift, the assumption $z_{\rm X} \simeq z_{\rm d}$ is a reasonable one if one considers that this is a faint object. The object appears being extended, thus it can be considered as dwarf galaxy near the redshift of G. Very recently, the work by Curran et al. (2011) finds absorption features at redshift $z \sim 0.38$ that may be associated with this object, in addition to the absorption occurring in the lens galaxy and in the host galaxy of the source. However, for the work and the results presented here, based on the above considerations, we have assumed that it lies at the same redshift of the lens galaxy.

Table 5.1: HST deep imaging: magnitudes and relative positions from the core centroid of image B. (*CASTLES* database).

	G	X	A1	A2	B	C
RA	−472±3	−857±9	600±3	732±3	0±3	−1342±3
Dec	−1.277±3	180±11	-1942±3	−1549±3	0±3	−1650±3
filter						
F160W	17.54±0.02	20.12±0.32	15.54±0.01	15.87±0.01	16.56±0.01	17.41±0.01
F110W	19.21±0.01		17.59±0.01	18.08±0.01	18.54±0.01	19.37±0.01
F205W	16.70±0.11		14.59±0.01	14.79±0.02	15.63±0.04	16.49±0.01
F555W	24.17±0.08		25,53±1.26	26.69±0.22	26.08±2.07	26.68±0.09
F675W	22.58±0.13	24.63±0.30	22.61±0.02	23.80±0.02	23.34±0.03	24.17±0.04
F814W	20.91±0.04	23.35±0.21	20.43±0.06	21.36±0.06	21.24±0.06	22.10±0.07

The blue optical arc

This structure was revealed by deep HST imaging (Falco et al., 1997), with a much bluer colour than the four compact images; no colour differences are found as function of the arc length, nor between among the lensed cores; hence the colour difference between the core of the images and the arc is likely to be explained by extinction occurring in the host galaxy of the lensed QSO rather than in the lens galaxy. In the latter case, because of the patchy extinction, the colour differences would appear on scales smaller than $0.4''$ (which is the smallest separation between the lensed images). It is however also possible that the arc is a lensed (distorted) region not physically connected with the core of the QSO. We go into more details on this in Section 5.9.

5.2 Existing Radio observations

Radio observations of MG0414+0534 with the VLA radio-telescope have provided flux measurements of the lensed components (Katz et al., 1997). The system was observed in several bands between 1.4 and 22 GHz. The spectral indices of the four components were found consistent with each other, with a best estimate for the averaged spectral index[1] of $\alpha = 0.80 \pm 0.02$ (Katz et al., 1997). The radio spectrum (Fig. 6 in Katz et al. 1997) shows a peak near 1 GHz suggesting that MG0414 belongs to the GHz-peaked spectrum radio sources. These are compact radio sources with sub kilo-parsec sizes. At higher angular resolution (Patnaik & Porcas, 1996; Porcas, 1998) the system shows compact radio-cores and extended structures in three of the lensed images. Figure 5.3 shows the 1.7 GHz (L-band) map of the system, made by Patnaik & Porcas (1996) using the European-VLBI network (EVN) radio-telescope.
Higher resolution and more sensitive observations at 8.4 GHz (X-band) (global-VLBI observations by Ros et al. 2000) and 5 GHz (C-band) (VLBA observations by Trotter et al. 2000) have later-on revealed more details on the extended structures in all the four images, namely a core region and jet-like structures (see Figures 5.4 and 5.5).

[1] the spectrum is modelled as $S \propto \nu^{-\alpha}$, where ν is the observing frequency

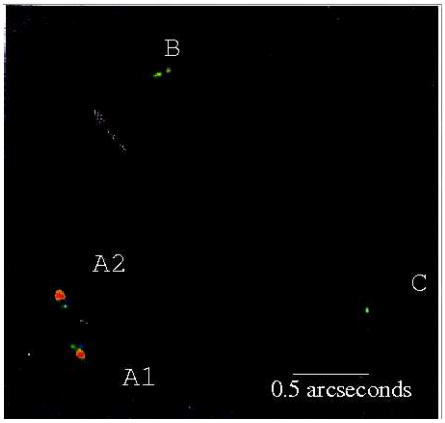


Figure 5.3: EVN maps at 1.7 GHz (Patnaik & Porcas, 1996; Porcas, 1998). Images A1 and A2 are brighter than the other two, image C is the faintest one. These maps already show the extended emission of the lensed QSO, but higher resolution is necessary to better understand the source properties and identify the sub-components among the images. This is important in order to use these features to constrain a lens model.

VLA data by Katz et al. (1997) have shown no emission at the position of object X or the lens galaxy.

5.3 Existing Lens models

The VLBI observations presented in this chapter were carried out with the goal of exploring lens models using a richer set of constraints, provided by the lensed jets of the QSO. Before we discuss the results we find, here we briefly overview the lens modelling constrained by existing VLBI observations.
VLBI observations in C and X band (Figures 5.4 and 5.5) have revealed a lot of details on the structure of the lensed source. The combination of the intrinsic source structure and the lensing transformation results in a quite complex geometry, that at first glance may mislead the lensing interpretation. The works by Ros et al. (2000) and Trotter et al. (2000) show that the image configuration revealed by the radio observations can be explained by a gravitational lensing scenario.
Ros et al. (2000) parametrize the mass distribution with a singular isothermal ellipsoid, describing the main lens, and a singular isothermal sphere, describing object-X; in addition, an external shear term was include as well. This model is found to account for the observed positions of the compact radio cores, with the total χ^2 dominated by the flux-ratio constraints.
Trotter et al. (2000) use a more general mathematical description for the lensing potential, namely a multipole expansion. They parametrize the source with four point-like components (corresponding to the features labelled p, q, r and s, Fig. 5.4), whose positions were used to constrain the model. In the best model the centroid position for the deflecting mass coincides with the measured optical position of the lens galaxy. However the authors find that the mass distribution is highly asymmetric or an unseen perturber has to be considered.

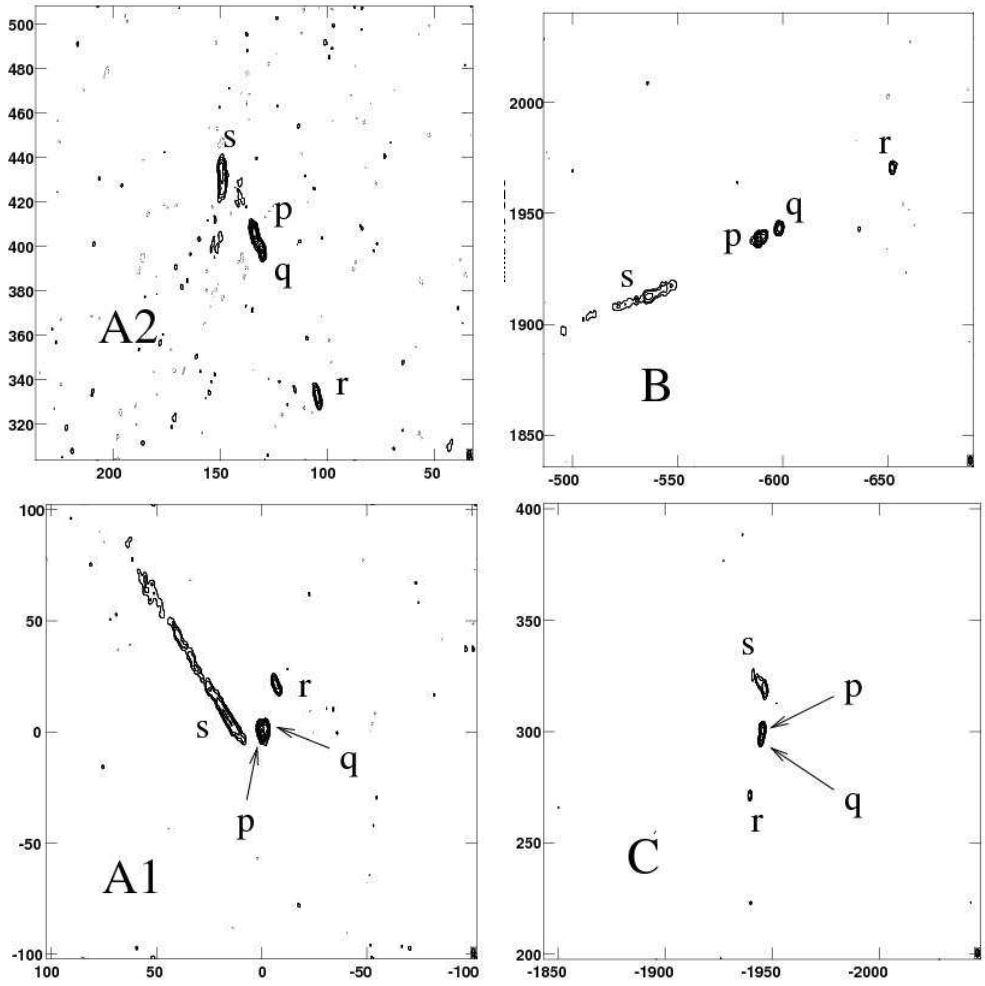


Figure 5.4: C-band (5 GHz) VLBI maps of MG J0414+0534 (Trotter et al., 2000). On the axes are plotted the relative right ascension (east of A1-p) and declination (north of A1-p) in milli-arcseconds. For this thesis we have adopted the same notation as these authors for the sub-components of the lensed images.

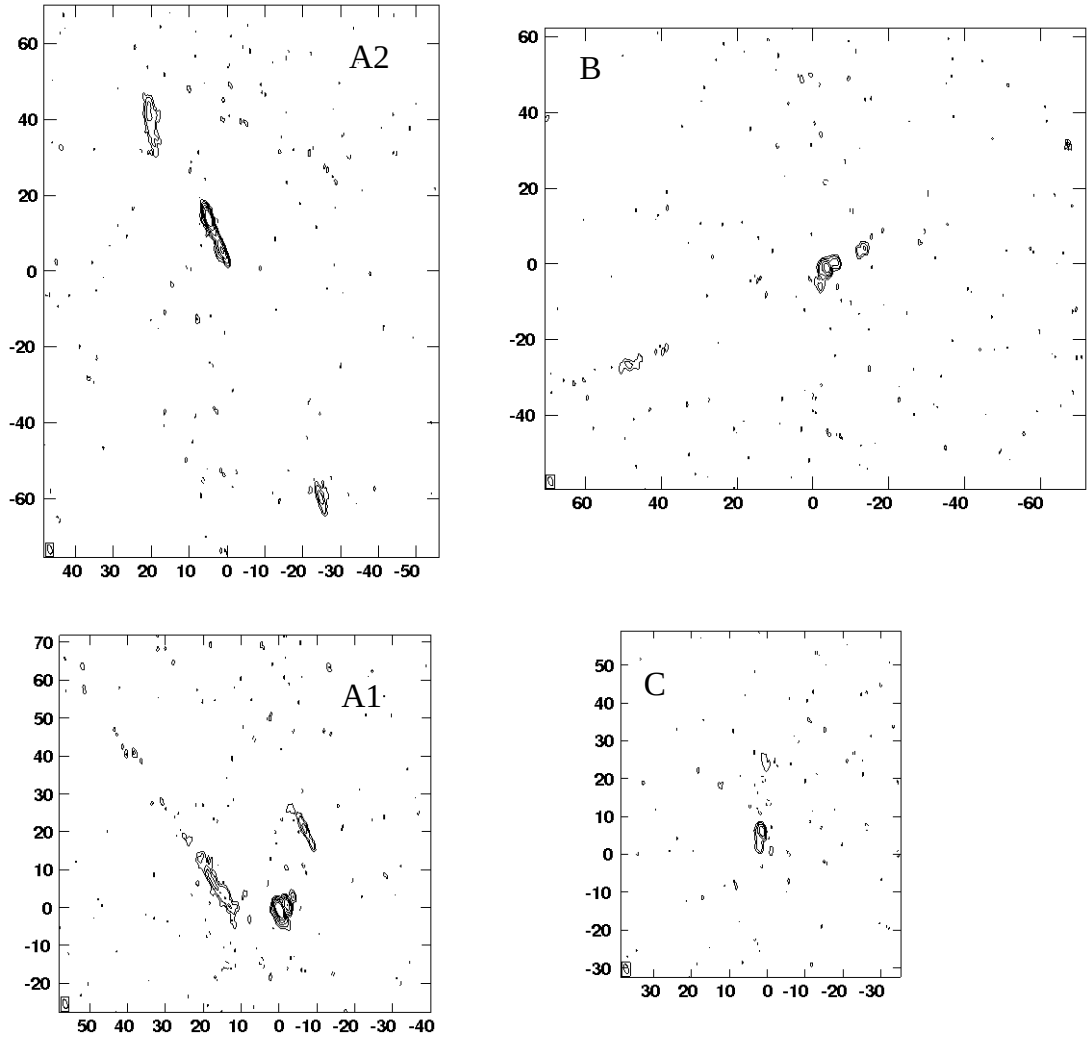


Figure 5.5: X-band (8 GHz) VLBI maps of MG J0414+0534 (Ros et al., 2000). On the axes are plotted the relative right ascension (east of the phase centre) and declination (north of the phase centre) in milli-arcseconds. Note that the four fields were imaged using the wide-field approach (see Sect. 3.5).

Note that neither group of authors fully exploit the wealth of structures seen at radio wavelengths, using a rather simplified description for the lensed source. The configuration contains a lot of information on the foreground mass distribution which could be studied in great detail.

5.3.1 The 'Flux-ratio' anomaly

As described above, since the discovery of the system, the observed flux ratios at different wavelengths have always been a puzzle.
Minezaki et al. (2009) report on mid-infrared observations of this system carried out with the Subaru telescope. Table 5.2 summarizes the flux-ratio measurements at different wavelengths from optical to radio. Within the uncertainties, measurements in the infrared and radio regimes are consistent with each other. Emission at these wavelengths is not affected by microlensing caused by stars in the foreground lens galaxy, given that the Einstein rings of the stars have sizes of few percents of the regions where the mid-infrared and the radio emission originate, and at radio wavelengths extinction by dust does not occur (and some propagation effects, like scattering, are wavelength-dependent, thus can be accounted for). This is not the case at smaller observing wavelengths, where dust extinction is largely affecting the radiation from the lensed source. The nature of this extinction is however not full understood. Elíasdóttir et al. (2006) found images A2 and A1 significantly more extinguished than images B and C, however they did not find just one extinction law, rather different ones giving similar results.
The pair A1−A2 is found to show the so-called 'flux-ratio' anomaly: for these two images, the radio and mid-infrared flux-ratios do not agree with the predictions of lens models. The *smooth* mass model, consisting of an isothermal ellipse, an isothermal sphere and an external shear, is found to account for the mid-infrared observed positions of the core images, but not the for the flux ratios; the pair A2/A1 is indeed predicted to have a value of $\simeq 1.06$ whereas we observe ~ 0.90 (see Table 5.2). Image A2 would then be the brighter negative parity image demagnified with respect to the model predictions. Gravitational lensing by a CDM substructure could be the explanation to this problem. Minezaki et al. (2009) infer a limit on the mass of a possible CDM substructure which is causing this anomaly, finding that a CDM sub-halo of mass $M > 360 M_{\odot}$ is likely to affect the radio and mid-infrared flux ratio of the A1-A2 pair.

5.4 New Global-VLBI observations at $\lambda = 18$ cm

The purpose of the new 1.7 GHz global-VLBI observations was to map the extended structures of the source combining the steep spectrum of the lensed source and the sensitivity of the data. This is a high priority in order to have a good data set which will be used to study the lensing properties of the foreground mass distribution.

Observations details

The system was observed at a wavelength of 18 cm on 7 June 2008 for 19 hours with a global VLBI array comprising the following 21 telescopes: Effelsberg (100 m),

Table 5.2: Measured flux-ratios from optical to radio wavelengths

	Flux ratios		
observing wavelength	A2/A1	B/A1	C/A1
0.8 μm[a]	0.425±0.033	0.474±0.040	0.215±0.018
1.1 μm[a]	0.637±0.024	0.417±0.016	0.194±0.007
1.6 μm[a]	0.738±0.025	0.391±0.013	0.179±0.005
2.1 μm[a]	0.832±0.017	0.384±0.015	0.174±0.002
11.7 μm[d]	0.90±0.04	0.36±0.02	0.12±0.03
22 GHz[b]	0.881±0.013	0.397±0.004	0.157±0.003
15 GHz[b]	0.8905±0.0049	0.3896±0.0015	0.1515±0.0013
8 GHz[b]	0.8974±0.0016	0.3887±0.0005	0.1492±0.0004
5 GHz[b]	0.898±0.010	0.386±0.003	0.144±0.002
1.7GHz[c]	0.93±0.09	0.37±0.03	0.14±0.01

[a] CASTLES
[b] Katz et al. (1997)
[c] work presented in this thesis (Sect. 5.4). The uncertainties here assume a 10% of uncertainty on the flux measurements, which may be higher than the accuracy on the flux-ratios, but accounts for calibration offsets.
[d] Minezaki et al. (2009)

Westerbork tied-array (equivalent to 93 m), Jodrell Bank (76 m), Onsala (25 m), Medicina (32 m), Noto (32 m), Torun (32 m), Urumqi (25 m), Shanghai (25 m), Hartebeesthoek (26 m), Arecibo [2](305 m), 9 antennas (25 m) of the VLBA and the VLA tied-array (equivalent to 130 m). The 19 hours of allocated time for this experiment were scheduled in this way: 5 hours of EVN, 3 hours of EVN and VLBA, 1.5 hours EVN - VLBA - VLA, 2.5 hours EVN - VLBA - VLA - ARECIBO, for the remaining 7 hours VLBA and VLA. Data were recorded in dual-polarization mode at 512 Mbps, distributed over 8 subbands per polarization, each of 8 MHz bandwidth. The data were correlated at the EVN MkIV Data Processor at the Joint Institute for VLBI in Europe (JIVE) in three different portions: an initial (for the EVN) and a final one (for VLBA and VLA), correlated in one single pass; a middle one having more than 16 stations which required 3 different passes in order to correlate all the baselines. There were 16 frequency points (channels) per subband. The integration time was 1 second.
The flux density of the target source is high enough (Katz et al., 1997) so that phase-referencing was only needed to create an initial model prior to self-calibration. The observing cycle consisted of one hour blocks repeated through the whole observing time. They were scheduled as follow: 15 minutes cycles of alternate scans between the target MG J0414 + 0534 and the phase referencing source J0422 + 0219, at 3.85 degrees distance; once per hour a fringe finder was observed for 4 minutes. For the 2.5 hours with Arecibo the 15 minutes scans alternated between the target and J0412 + 0438, 1.07 degrees away. J0409 + 1217, J0319 + 4130, J0555 + 3948, J0530 + 1331 and 3C286 were used as fringe finders and calibrators for polarization.

[2]our target, at declination 5°, can be tracked for ∼2 hours at Arecibo,close to the maximum of 2h:46m

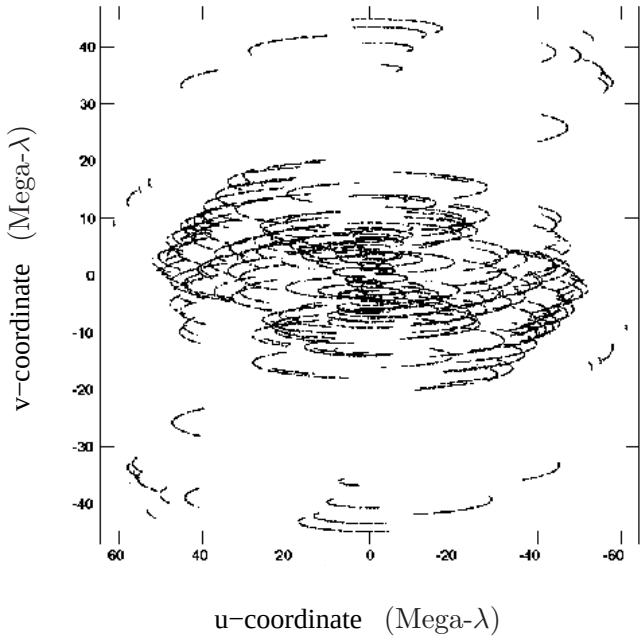


Figure 5.6: Global-VLBI observations: (uv)-plane coverage for the target source, MG0414+0534. The core of the plane is filled with uv-spacing of the EVN and VLBA antennas, the points at the edges in the north-south direction identify the baselines with the South-Africa antenna, and the most East-West points are given by the baselines to the Shanghai and Urumqi antennas. The target was tracked for almost 19 hours.

5.4.1 Data reduction

The data reduction and analysis was performed using the standard Astronomical Image Processing System (`AIPS`, version 31Dec10). The visibility phases were corrected for changes in parallatic angle to account for the apparent change in position angle as the source moves across the sky. The *a priori* amplitude calibration was determined from measured system temperatures and antenna gains.
Fringe fitting was performed using the two phase-referencing sources (J0422 + 0219 and J0412+0438). Changes of gains as function of frequency (across all the channels) were corrected via bandpass calibration. The spectral response of each telescope was analysed from the data of the fringe-finders, and the final bandpass correction for each telescope was constructed using only the scan with the flattest amplitude and phase response. Subsequently, the two phase referencing sources were mapped by performing self-calibration, initially phase-only but finally phase and amplitude. The phase and amplitude corrections from the best models for both sources were then applied to the target which was then imaged. The data on the lens system were then phase self-calibrated with solution intervals of 4, 2 and 1 minutes. During self-calibration, the telescopes were not given their full statistical weight. Note that the array is made out of several 25-meter diameter dishes and a few big telescopes. If the

telescopes are given their statistical weight, bigger telescopes will have a significantly larger weight, as well as all the baselines to these telescopes (see Equation (3.20)). Practically this means that the visibilities from these baselines are considered to have a better calibration than the rest, though it is not the case. Hence, imaging will be dominated by calibration errors. Therefore during these steps we did not use the full statistical weights. Through the whole process data were never averaged in frequency or time, in order to preserve a large field of view.

5.4.2 Imaging

The system is $2''$ wide, and therefore wide-field techniques are required for imaging (Garrett et al., 1999). Instead of mapping one single field ($2'' \times 2''$ wide) with milli-arcseconds (mas) resolution, which would require a large computing time, the `AIPS` task `IMAGR` provides a wide-field mapping procedure based on a multi-windowing approach. We mapped 3 sub-fields, centred on image B, C and the mid-point of image A1 and A2. Imaging the four fields separately was found giving poor results for fields A1 and A2.
The final images were produced using a weighting scheme between uniform and natural, setting the parameter `robust` to 0. This was found to give the best combination between sensitivity and resolution (see Sect. 3.3). The use of very tight clean boxes was found to be crucial in order to map the extended structures with high SNR. Furthermore a large number of clean iterations (10000; with a loop gain factor of 0.08 see Sect. 3.4) has helped in improving the SNR.
The rms noise on the final single image produced combing the three subfields is 0.11 mJy beam^{-1}; the rms noise on the three maps are 0.12, 0.10 and 0.08 mJy beam^{-1}. The peaks of emission in the 4 fields are 140, 140, 30 and 25 mJy/beam. Imaging results are shown in Figures 5.7 and 5.8. In both figures, a label in each panel identifies the lensed image and in Fig. 5.8 the sub-components are labelled adopting the same notation as Trotter et al. (2000): p and q for the two core components, r for the single jet component on one side of the core, s for the more extended jet regions on the other side and t for the outer part of this. The restoring beam has size of $\simeq 5 \times 3$ mas^2, and position angle P.A.$\simeq -7$ deg (measured East of North). In Fig. 5.7, the maps displays contours level at $(-1, 2, 4, 8, 16, 32) \times 0.5$ mJy/beam, corresponding to $\gtrsim 4$ times the rms noise in three subfields; in Fig. 5.7, the maps show contours level at $(-1, 2, 4, 8, 16, 32) \times 1.4$ mJy/beam, corresponding to the 10% of the peak of emission in the A1-A2 field. Previous high resolution VLBI observations of this system (Ros et al., 2000; Trotter et al., 2000) have showed the presence of a resolved core region and 2 jet-like structures in each of the 4 lensed images. Our new global VLBI data has higher sensitivity to the extended structures. A comparison of our maps with Fig. 5.4, shows that the structures are consistent with those shown in previous images; in images A1, A2 and B we confirm the feature seen in the outer part of the more elongated jet. Image B shows a resolved core region extending over ~ 20 mas and two jet like features (r and s/t) not exactly aligned with it. This image is highly stretched in the direction tangential to that of the lensing galaxy. Images A1 and A2 show the most complex structures, highly distorted in the tangential directions. Image A1 has the two jet-like features parallel to each other and not aligned with the central core. The brighter jet (s/t) is highly

Table 5.3: Radio flux densities for the fields shown in Fig. 5.7 and for the sub-components of each lensed image. Values reported in the first three columns are computed from the `clean` components (`cc`) in these regions. The forth column reports the total flux for each image, computed by the `cc`.

	core[a] (mJy)	jet1[b] (mJy)	jet2[c] (mJy)	tot[d] (mJy)
a1	238	322	60	620
a2	233	292	53	578
b	80	135	14	229
c	33	44	8	85

a total flux density for the core feature (labels *p,q*)
b total flux density for the jet labelled *s,t*
c total flux density for the jet labelled *r*
d this value is used to calculate the flux ratios reported in Table 5.2

elongated and extends for more than 80 mas in the NE direction. In image A2 the two extended jet-like features are also not aligned with the core. The brighter (*s/t*) is elongated in the NS direction over $\sim$ 60 mas, while *r* is $\sim$ 80 mas away in the SW direction. The core region extends over $\sim$ 20 mas, showing a two-component structure. Image C shows a NS structure, with the core in between *r* and *s/t*; the northern jet component is brighter than the southern, and is more elongated. This image is the least bright (and hence least resolved) amongst the four. Existing lens models predict that the 2 pairs of images, A1&A2 and B&C, should have opposite parities. The extended structures seen at radio wavelengths confirms this although the lensing transformations are highly non linear.
The total flux densities measured for each subfield agree with the flux ratios between the lensed images given by Katz et al. (1997). We present in Table 5.3 the flux density for the core-like and jet-like regions which we have computed from the `clean` components in each of these regions. Elliptical Gaussian model fitting of the 4 images was carried out via the `AIPS` task `JMFIT` and the results are given in Table 5.4. The errors were determined following Fomalont (1999). Following this formalism, the uncertainty on the position of the modelled component depends on the fitted size of the component, its peak flux density and the thermal-noise on the map. Most of the components are highly elongated, thus in Table 5.4 the errors on their positions are given as error ellipses. For the core features, the higher resolution data from Ros et al. (2000) (Figure 2) show that our description is consistent; higher resolution maps show indeed 3,2,4 and 2 sub-components in the core of images A1, A2, B and C. In Table 5.5 we summarize results from model fitting of the peaks of emission, carried out in `AIPS` using the procedure `MAXFIT`. This procedure is best suited to study the position of the peaks, but it does not account properly for the extended structures of the components. We refer the reader to Appendix A for a description on the two tasks and the procedure used to derive the errors on the fitted parameters.

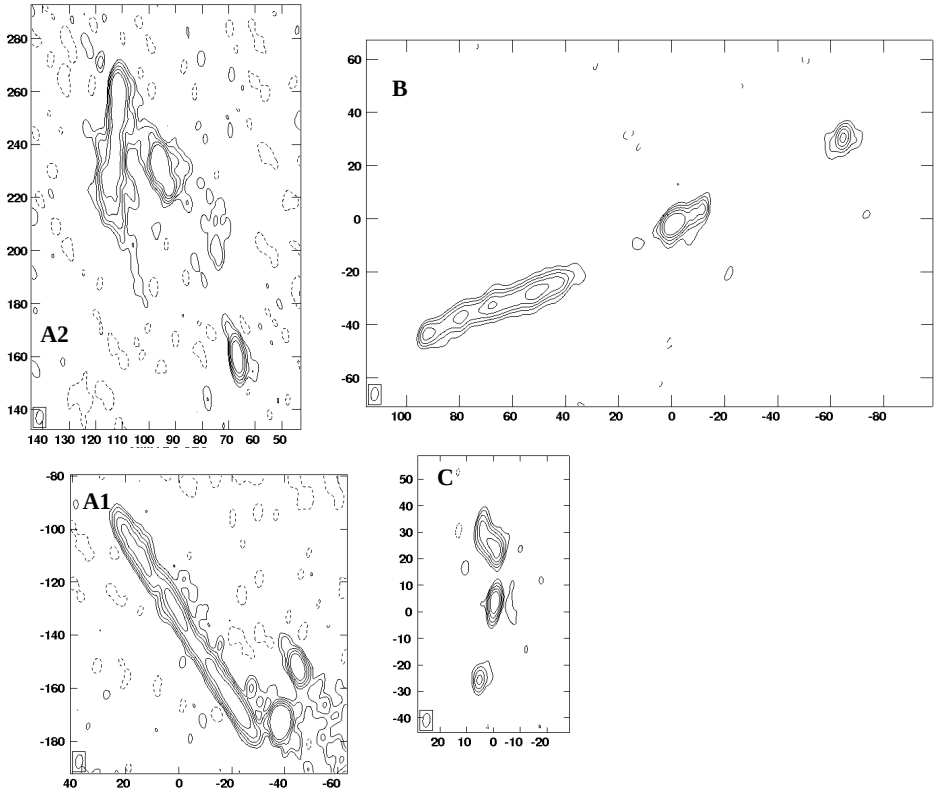


Figure 5.7: 1.7 GHz global VLBI observations: imaging results. The rms noise levels in the 4 images are 0.12, 0.12, 0.10 and 0.08 mJy/beam for A1, A2, B and C respectively; these were measured as the off-source rms noise levels. The maps display contours at $(-1, 2, 4, 8, 16, 32) \times 0.5$ mJy/beam. The restoring beam is plotted on the lower left corner of each panel. It has size of 5×3 mas^2 at position angle $-7°$. On the axes are plotted the relative right ascension (east of the phase centre) and declination (north of the phase centre), in milli-arcseconds. Note that the system was mapped using 3 sub-fields, centred on the mid-point of image A1&A2, image B and image C.

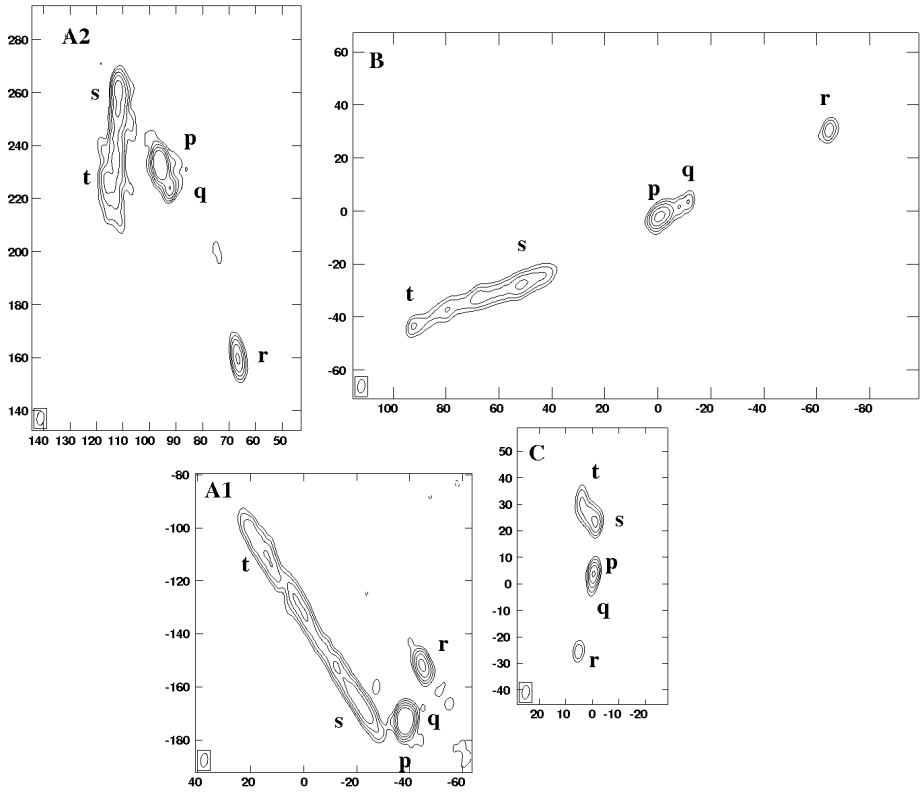
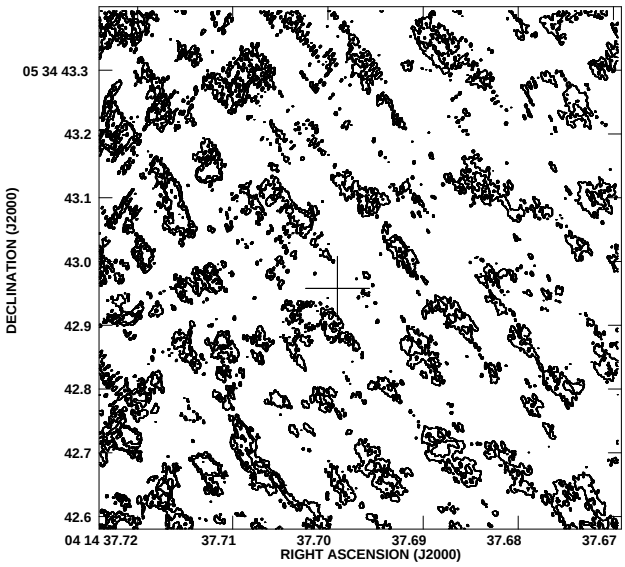


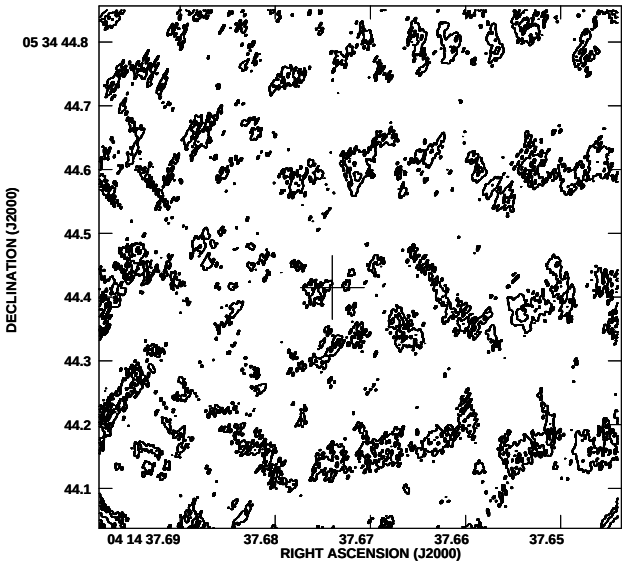
Figure 5.8: In each panel the labels identify the lensed image and different regions of the source using the same notation used by Trotter et al. 2000. The maps display contours at $(-1, 2, 4, 8, 16, 32) \times 1.4$ mJy/beam. On the axes are plotted the relative right ascension (east of the phase centre) and declination (north of the phase centre), in milli-arcseconds.

Emission from the lens and object X

Previous radio observations of this system did not report on the emission from the lens galaxy or object-X (Katz et al., 1997). Any highly demagnified lensed image located near the main lens G is not expected to be extended, but rather point-like. Hence a dirty-image was produced to check the presence of any point-like source in a field centred around the optical position of the lens galaxy. In the same way we have checked the presence of emission in a field centred on object-X. Figure 5.9 shows 3σ contours of the dirty maps of the two fields, each of sizes $0.8'' \times 0.8''$. The rms-noise in the maps is 0.4 mJy beam^{-1} for both fields; the cross marks the optical position of the lens and object-X. The 1σ level does not show the presence of a point-like source associated with the the lens galaxy or object-X.
Following Zhang et al. (2007) we have furthermore cleaned both fields using the same weighting scheme described above. Note that it is likely that sidelobes from the A1-A2 pair (image B) are present in the lens (object-X) field, and are removed only by further cleaning the field. Zhang et al. (2007) find that cleaning with a large number of iterations ($\sim 10^4$) one single field, centred at the position of the expected image, leads to lower rms noise in the map, and to the recovery of 70% of the flux of any point like image located near the lens galaxy centre. With this approach the final rms noise in the maps are 87 and 80 μJy beam^{-1}, corresponding to flux density limit (3σ levels) of ~ 0.26 (image located near the lens) and ~ 0.24 mJy^{-1} (any emission located near object-X), which we do not detect in our maps.



(a)



(b)

Figure 5.9: Dirty maps of two fields centred on the optical position (marked by the cross) of the lens galaxy (top panel) and object-X (bottom panel). The dirty maps do not show the typical beam pattern centred on the optical position marked by the cross which would indicate a peak of emission at these positions.

Table 5.4: Results from fitting of Gaussian components to the structures seen in the global-VLBI 1.7 GHz map. For each fitted components the position of the peak and the total flux densities are given. The uncertainties were estimated following the approach described in Appendix A. The first component of image A1 is taken as origin of the coordinates.

image	RA (mas)	Dec (mas)	error ellipse (mas, mas, deg)	S_{peak} (mJy beam^{-1})	S_{tot} (mJy)
	0.00	0.00	(0.04, 0.02, 5.5)	23.55±0.18	65.95±2.80
A1,core	−0.11	0.35	(0.01, 0.02, 172.9)	123.59±0.18	159.85±1.29
	−2.72	2.26	(0.06, 0.02, 170.7)	11.44±0.18	11.44±1.00
A1,r	−6.84	21.05	(0.02, 0.01, 7.1)	31.23±0.18	55.95±1.79
A1,jet[3]	14.15	2.29	(0.04, 0.02, 19.8)	20.55±0.18	50.42±2.45
	16.99	8.01	(0.06, 0.03, 177.1)	10.63±0.18	14.61±1.37
	19.51	11.46	(0.10, 0.03, 175.8)	6.80±0.18	7.42±1.09
	21.95	14.32	(0.07, 0.03, 27.1)	16.35±0.18	62.30±3.81
	25.55	20.30	(0.06, 0.03, 10.7)	14.85±0.18	34.31±2.31
	36.82	38.23	(0.10, 0.03, 6.1)	9.99±0.18	21.73±2.18
	41.18	45.69	(0.07, 0.03, 17.4)	14.64±0.18	44.62±3.05
	51.02	60.26	(0.11, 0.04, 12.8)	11.91±0.18	43.73±3.68
	57.67	70.46	(0.14, 0.04, 22.6)	11.00±0.18	49.24±4.45
	134.66	405.99	(0.01, 0.01, 0.185)	116.67±0.18	184.82±1.58
A2,core	131.12	397.56	(0.04, 0.02, 174.3)	20.91±0.18	38.02±1.82
A2,r	105.17	332.81	(0.03, 0.01, 3.6)	28.37±0.18	53.06±1.87
	150.75	435.54	(0.03, 0.01, 174.7)	22.71±0.18	20.02±0.88
A2,jet	150.29	429.12	(0.04, 0.01, 176.7)	36.12±0.18	128.00±3.54
	151.57	406.51	(0.09, 0.02, 179.6)	21.14±0.18	113.20±5.35
	154.96	398.20	(0.07, 0.02, 6.5)	15.77±0.18	38.66±2.45
	−590.47	1937.94	(0.02, 0.01, 166.1)	26.66±0.17	37.77±1.41
B,core	−593.23	1939.46	(0.05, 0.02, 171.7)	11.72±0.17	12.23±1.04
	−595.63	1940.34	(0.27, 0.13, 123.1)	5.11±0.17	32.91±6.44
	−602.06	1944.27	(0.15, 0.06, 170.6)	3.39±0.17	2.56±0.76
B,r	−655.43	1970.84	(0.07, 0.03, 167.7)	10.04±0.17	15.22±1.52
	−539.47	1913.00	(0.11, 0.05, 118.1)	12.76±0.17	73.41±5.75
B,jet	−523.21	1907.98	(0.12, 0.07, 125.7)	8.27±0.17	36.25±4.38
	−510.13	1903.32	(0.20, 0.12, 135.5)	5.13±0.17	23.23±4.52
	−500.33	1897.80	(0.23, 0.13, 142.8)	4.03±0.17	13.96±3.46
	−497.61	1897.22	(0.15, 0.06, 172.1)	3.87±0.17	3.80±0.98
	−1954.49	300.26	(0.01, 0.01, 170.9)	27.10±0.09	31.13±1.15
C,core	−1953.41	292.61	(0.14, 0.07, 170.5)	1.69±0.09	1.23±0.72
C,r	−1948.88	271.43	(0.08, 0.03, 169.7)	5.03±0.09	6.50±1.29
	−1955.26	318.83	(0.03, 0.01, 175.3)	10.95±0.09	12.87±1.17
C,jet	−1952.99	322.87	(0.04, 0.03, 8.6)	9.59±0.09	23.97±2.50
	−1950.12	328.72	(0.05, 0.02, 175.2)	6.89±0.09	8.60±1.25

Table 5.5: Position and strength of the emission peaks. The position of the core peak in A1 is taken as the origin of coordinates. This is offset from the brightest component of Table 5.4, not surprising given that two different procedure are used to determine the emission peaks and the parameters of the Gaussian components describing the features. See Appendix A for more details.

image	component	RA (mas)	Dec (mas)	S_{peak} (mJy beam^{-1})
	core	0.00±0.004	0.00±0.004	148.89±0.18
	r	−6.66±0.02	20.84±0.02	30.74±0.18
	jet	14.67±0.03	2.63±0.03	21.87±0.18
		19.57±0.03	10.53±0.03	20.99±0.18
		16.90±0.03	6.55±0.03	22.14±0.18
A1		25.83±0.04	20.28±0.04	14.99±0.18
		37.91±0.05	40.07±0.05	13.47±0.18
		41.02±0.04	45.03±0.04	15.48±0.18
		56.71±0.05	68.17±0.05	11.44±0.18
		52.05±0.04	61.42±0.04	13.83±0.18
	core	134.80±0.01	405.70±0.01	118.98±0.18
	r	105.32±0.02	332.49±0.02	28.33±0.18
	jet	150.67±0.01	433.95±0.01	46.37±0.18
A2		150.58±0.02	427.49±0.02	39.20±0.18
		151.97±0.03	407.74±0.03	21.98±0.18
		154.29±0.03	398.19±0.03	20.86±0.18
		−590.82±0.02	1937.89±0.02	32.25±0.17
	core	−601.75±0.09	1943.48±0.09	6.16±0.17
	r	−655.33±0.05	1970.62±0.05	10.81±0.17
B	jet	−538.46±0.04	1912.26±0.04	14.95±0.17
		−512.96±0.06	1907.35±0.06	9.59±0.17
		−510.48±0.09	1903.12±0.09	6.12±0.17
		−497.78±0.08	1896.68±0.08	6.85±0.17
	core	−1954.37±0.01	300.00±0.01	27.16±0.09
C	r	−1948.69±0.06	270.79±0.06	5.13±0.09
	jet	−1954.98±0.02	319.59±0.02	15.09±0.09

5.5 Lens modelling with the new radio observations

The main goal of this project was to improve the existing lens models including the set of constraints from the extended features seen at radio wavelengths. In order to do so, we started with a simple point-like description of the lensed source, and used only the core features as constraints. The best lens model is then used as starting model for the next step in which we use the position of all the fitted components to constrain the shape of the source using the *curve-fitting technique*. These results will be discussed in Sections 5.6 and 5.7. Here we give details on the set of constraints and parameters, and describe the modelling strategy.

5.5.1 Constraints from radio and optical measurements

We have parametrized the lens galaxy as Singular Isothermal Ellipsoid (SIE) and object-X as Singular Isothermal Sphere (SIS) (these mass parametrizations were described in detail in Chapter 2 (Section 2.1). The lens model includes also an external shear component. An independent shear can be produced by perturbations in the lens galaxy environment or by the internal angular structure of the lens itself. As shown by Keeton et al. (1997), in quadrupole lens systems this may lead to better fits. From optical measurements (Falco et al., 1997), the ellipticity and position angle (P.A.) of the lens galaxy are well constrained. However we have assumed no prior on the ellipticity nor on the P.A. in order to compare at the end our results with existing models and with optical results. The constraints on the positions of the lensed images were taken from the new VLBI observations (see Table 5.4). For the point-like model source we only used the core features. This region is resolved in all of the four images but, as shown also by the results of the Gaussian fitting, a different number of Gaussian components is needed to describe it; identifying these components among the lensed images is difficult thus we have constrained the position of each core-feature with the position of the centroid of the Gaussian components describing it. The smooth macro-model of Ros et al. (2000) can account relatively well for the positions of the lensed images within 0.5 milli-arcseconds. Hence, we have assumed the position of the VLBI cores to be known with an accuracy of 0.5 milli-arcseconds, even though the accuracy provided by the VLBI measurements is higher.
Constraints on the flux-ratios are taken from the flux of the `CLEAN` components for the core regions as given in Table 5.3 with an accuracy of 10% (on the relative fluxes).
Since the lens galaxy and object-X are detected in optical observations, their positions were constrained by the deep HST NICOMOS H-band image ($\lambda = 1.6\mu$m) from the CfA-Arizona Space Telescope Lens Survey (CASTLES; Kochanek et al. 1999). We assumed a position accuracy of 20 milli-arcseconds. Note that at optical wavelengths the position of the lens is known with an accuracy of ~ 3 milli-arcseconds, however we have to account for the fact that the radio and the optical emission may not originate from the same region in the source, which leads to an uncertainty in aligning the radio and the optical maps. We have estimated this uncertainty measuring per each image the scatter between its optical position and the position of the centroid of the radio VLBI core features; the largest value of 20 milli-arcseconds was taken as a measure of the uncertainty in aligning the two maps. Hence, the positions

Table 5.6: Lens model parameters

SIE+shear+SIS and point-like source	
Einstein radius of SIE	b_l
position of SIE	x_l,y_l
ellipticity of SIE	e, θ_e
shear	γ, θ_γ
Einstein radius of SIS	b_x
position of SIS	x_x,y_x
position of the source	x_s,y_s
flux of the source	f_s

of the lens galaxy G and object-X were constrained within ±20 milli-arcseconds from the optical measured value. The model has thirteen parameters, ten describing the lens and object X and three the unlensed source. Table 5.6 lists the parameters needed for the parametrization just mentioned.

5.5.2 Model fitting approach

The positions and flux-ratios of the core centroids provide in total twelve constraints (8 from the positions and 4 from the relative fluxes), and four more constraints are given by the positions of both lenses. Hence when fitting an SIE+shear+SIS there are 3 degrees-of-freedom when all the parameters of Table 5.6 are fitted. Several attempts were necessary to find out the best fitting procedure, especially regarding the fit of the lens position. At the beginning we fixed the positions of both lenses and only varied the Einstein radii while fitting for the ellipticity and shear. Once we found a good fit for these parameters, the position of the lens galaxy was allowed to vary in both directions (x and y). Afterwards all the parameters were optimized at the same time. As explained in Chapter 2 errors on the lens model parameters were determined from the range over which $\Delta\chi^2 \leq 1$ as each parameter was varied while the others were fitted.

5.6 Results from lens modelling using a point-like source

In this Section we summarize the lens models constrained by the positions and flux-ratios of the core images. In order to compare the various models we summarize the results all together in Table 5.7.

SIE+shear

Falco et al. (1997) showed that a lens model only including the main lens galaxy and an external shear term cannot account for the observed image configuration. Here we show a similar result obtained constraining the same model with the VLBI observations. This model has a reduced χ^2 of ~ 4.6, which has the largest contribution given by the position of the lens galaxy. Results are given in the top panel of Table 5.7.

5.6.1 SIE+shear+SIS

Results of this model are given in the middle panel of Table 5.7. The top row of the panel refers to the model constrained only by the positions of the centroid radio cores, the bottom row refers to the model constrained by positions and flux-ratios. When not using the flux-ratios, the problem has 0 degrees-of-freedom (DoF), which implies that the reduced χ^2, namely a measure of the goodness-of-fit, cannot be given even though this model has a total $\chi^2 \simeq 0$. Adding the flux ratios in the set of constraints gives 3 more DoF, with a reduced $\chi^2 \sim 0.5$ given by the flux ratios ($\chi^2_{\rm flux} \simeq 1.2$) and the optical constraints ($\chi^2_{\rm oth} \simeq 0.4$). The model predicts a flux ratio for A2/A1 $\simeq 1.04$. The fitted value of the ellipticity ($e = 0.20 \pm 0.07$) agrees with the one measured at optical wavelengths ($e_{\rm opt} = 0.20 \pm 0.02$ and $\theta_{e,{\rm opt}} = 71^\circ \pm 5^\circ$ Falco et al. 1997), with the position angles (P.A.) slightly misaligned ($\theta_e = -89^\circ \pm 13^\circ$). The χ^2 function along the P.A. axis (second raw, right panel of Figure 5.11) does not show a well shaped parabola around the $\chi_{\rm min}$, this might be due to degeneracies present in our model. A strong degeneracy is present between the shear and ellipticity, which is discussed in Figure 5.12. We find that the ratio between the Einstein rings is $b_{\rm x}/b_{\rm l} = 0.13 \pm 0.03$. Within the uncertainty in aligning the radio and the optical frame, the positions of the lens galaxy and object X are consistent. Ros et al. (2000) have constrained the same model using higher resolution VLBI data; in their model the lens galaxy has an ellipticity of $e = 0.21 \pm 0.05$, with a P.A $(-83 \pm 6)^\circ$, the ratio between the two Einstein rings is $b_{\rm x}/b_{\rm l} = 0.17 \pm 0.02$. Results from lens modelling using VLBI observations are consistent with each other. Both studies find the total χ^2 dominated by the flux-ratios term, and a comparable misalignment between θ_e and $\theta_{\rm opt}$.

Predicted image configurations and source reconstruction

In Figures 5.13 we show the critical curves in the source and in the image plane, with the observed and modelled positions of the lensed images, the lens galaxy and object-X. This model predicts image B in the absolute minimum of the time-delay surface, followed by image A1 ($\Delta t \sim 16.5$ days), A2 ($\Delta t \sim 16.9$ days) and C ($\Delta t \sim 70.5$ days). In Figure 5.14(a) we show the source reconstructions obtained projecting into the source plane the observed image configurations. In the source plane, the modelled core regions, predicted by each of the four images, agree rather well with each other. The modelled positions of the r features show offsets of ~ 5 milli-arcsecond in the x-direction. For the more elongated jet, the reconstruction obtained from image B shows the least agreement with the ones from the the other three images, predicting the outer components of the jet crossing the caustic; indeed, in the image plane, the source reconstruction by image B does not predict all the components for the A1-A2 pair (see Fig.5.14(b)).
As next step we will use a larger set of constraints, which includes the jet features (see Sect. 5.7); hence how well the model can account for these structures will be described later.

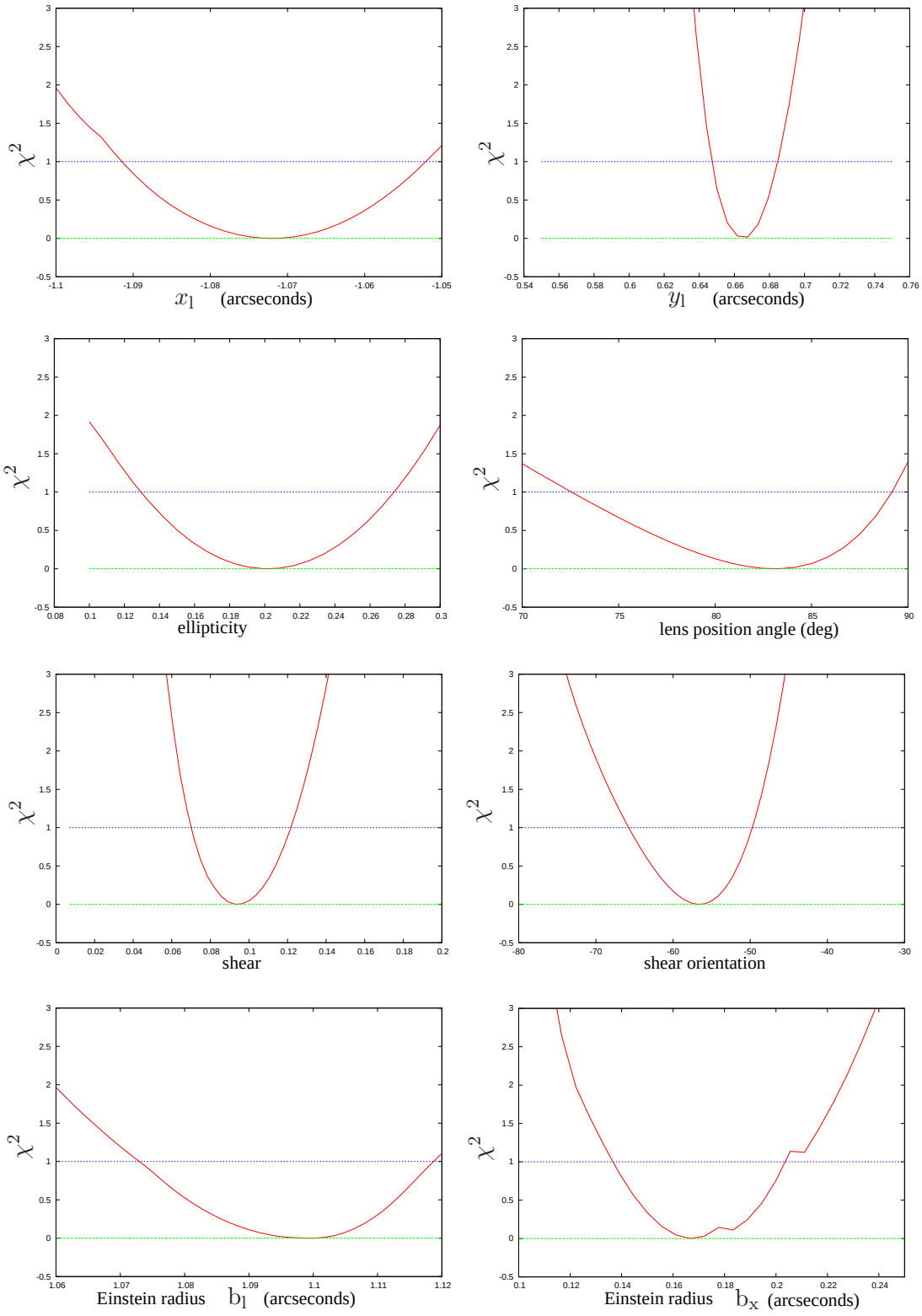


Figure 5.10: SIE+shear+SIS model constrained by the positions of the 1.7 GHz cores. These plots show the χ^2 function along the parameter-axis, varying each parameter while fitting all the others. Plotted in green is the value of $\chi^2_{\rm min}$ and in blue is the value $\chi^2_{\rm min} + 1$. Errors on the model parameters were determined in this range. For this model the total $\chi^2 \simeq 0$, however there are 0 DoF.

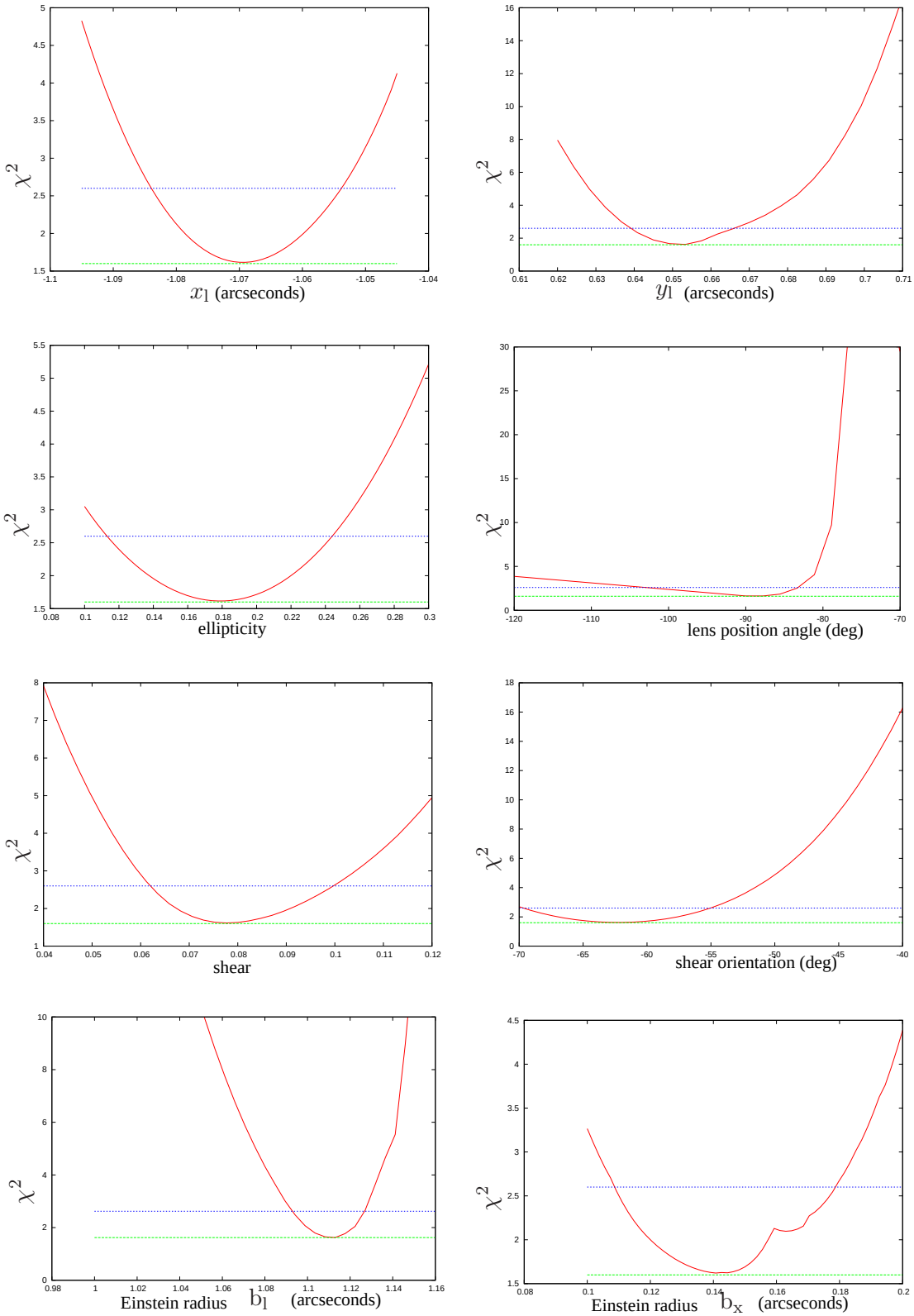


Figure 5.11: SIE+shear+SIS model constrained by the positions and relative flux-ratios of the 1.7 GHz core. The total $\chi^2 \simeq 1.6$, with 3 DoF. The χ^2 function along the P.A. axis does not show a parabola, as well as the χ^2 function of the shear orientation, which moves away from the absolute minimum for slight changes of this parameter. We discuss model parameters degeneracies in the next Figure.

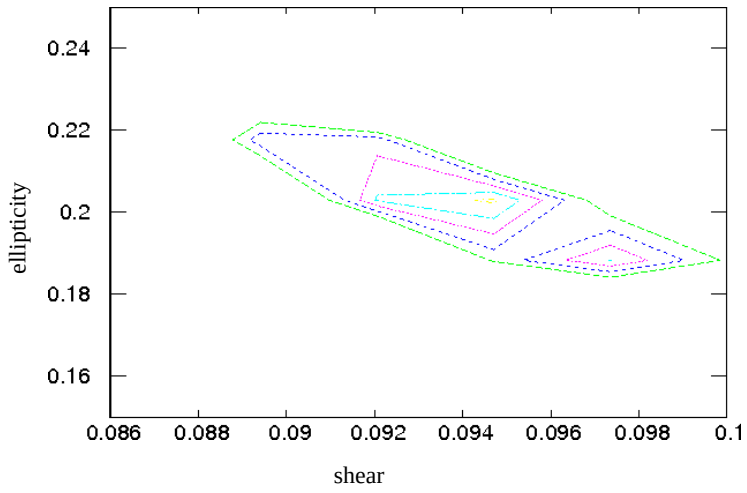


Figure 5.12: Degeneracy between shear and ellipticity. The plots show the χ^2 for the model SIE+shear+SIS while moving along the ellipticity and shear axis. Contours are shown for χ^2 values from 4 to 25.

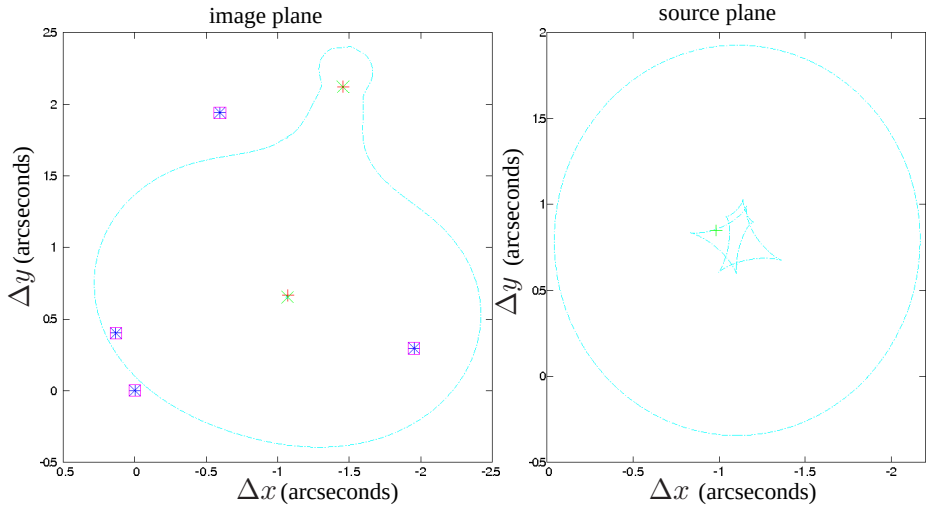
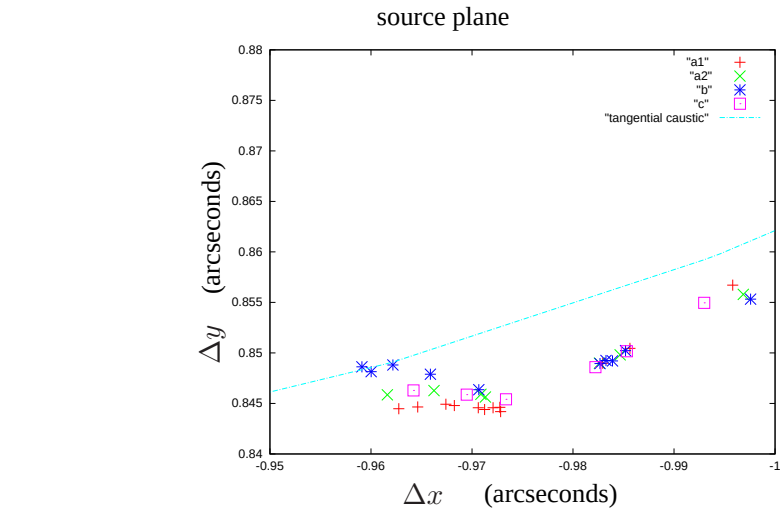
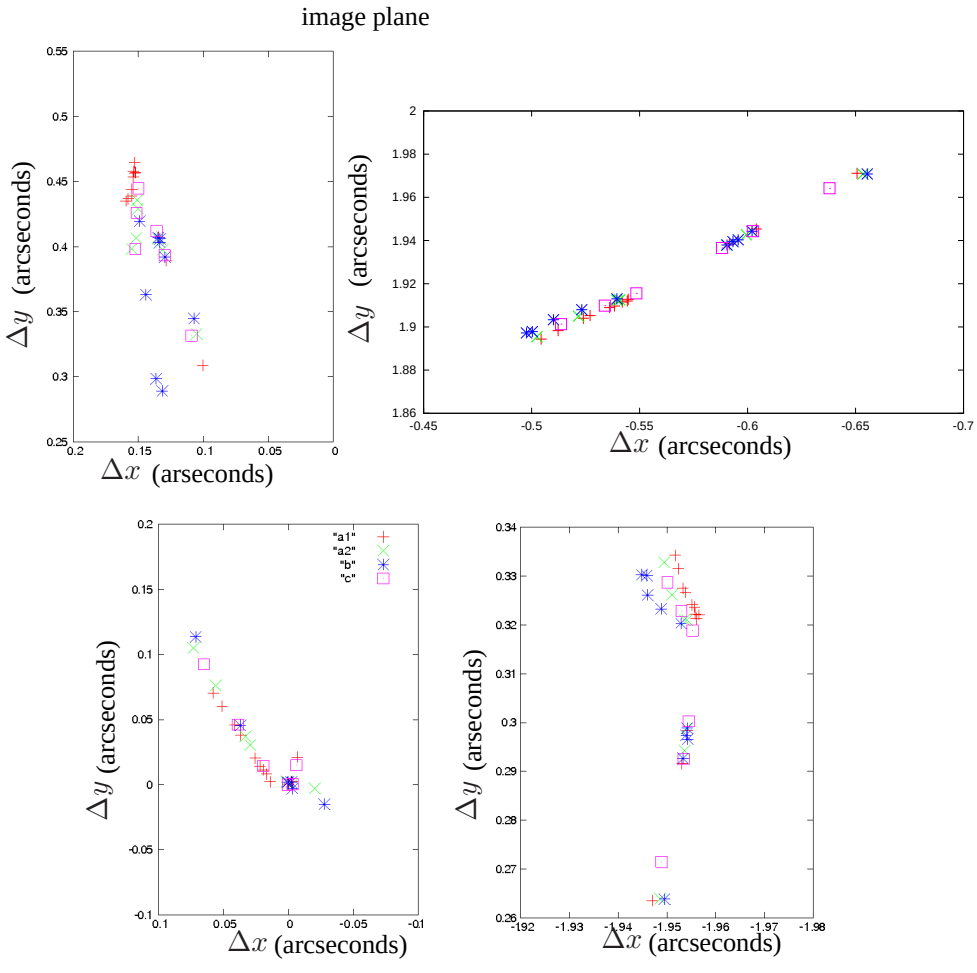


Figure 5.13: Critical curves for the model SIE+shear+SIS constrained by the positions and flux ratios of the 1.7 GHz core. The red (green) crosses indicate the observed (modelled) positions of the lens galaxy and object-X, while the blue asterisks (magenta squares) indicate the observed (modelled) positions of the core features. In the source plane, the source (the green cross) lies near the tangential caustic, as found by previous models.



(a)



(b)

Figure 5.14: Top panel: source reconstructions obtained per each lensed image. Bottom panel: predicted image configuration from model SIE+shear+SIS constrained by the 1.7 GHz core features. In the source plane, image B predicts the outer components of the elongated jet on the other side of the caustics, not in agreement in the image plane, with the observed configuration for the A1-A2 pair. The r component predicted by image C is offset by ~ 5 milli-arcsecond with respect to the other three.

The feature r as constraint

The same model (SIE+shear+SIS) was fitted using the positions and flux-ratios of feature r alone. Results are given in the lower panel of Table 5.7. The model just discussed is expected to account as well for observed configuration of the four images of this component.
This feature is not seen at optical wavelengths, nevertheless the positions of the lens and object-X can still be constrained by the optical measurements by using the relative astrometry between the r components and the core centroids, and assuming that within 20 milli-arcseconds these are aligned with their optical counterparts. Hence the position of the lens and object X is constrained within 20 mas of the optical measurement assuming the core centroid of image A1 as origin of coordinates. At the beginning we have constrained the model only using positional constraints, afterwards including the flux ratios. Interestingly, if the first model agrees with the one constrained by the core centroids, this is not the case for the second one. This model has a reduced $\chi^2 \sim 3.3$ given mainly by the optical constraints ($\chi^2_{\rm oth} \simeq 6.3$) with a contribution also by the flux-ratios ($\chi^2_{\rm flux} \simeq 3.7$). Similarly to what discussed above, when only using positional constraints the problem has 0 DoF, and hence even though the total $\chi^2 \simeq 0$, a measure of the goodness-of-fit cannot be given. When including the flux-ratios, the b parameter for object X is not constrained anymore, and a larger shear is required. The contribution to the χ^2 given by the flux-ratios is larger than the previous model, this may be due to mass perturbations that are not accounted by the model, but it may also indicate that the flux measurements for the r components are less accurate than the cores. The main contribution to the total χ^2 is given by the optical constraints.

Lens modelling using two point-like components

In the next row we show the results obtained constraining the same model with the observed positions and flux ratios of the core centroids and r components. In this case, the SIE+shear+SIS model is fitted very poorly, with a reduced $\chi^2 \sim 6.8$. In particular, the model seems to favour the main lens offset more offset 40 from the optical measurements, and a larger ellipticity. Figures 5.15 shows the predicted image configurations where we see that this model cannot account for the position of the VLBI features. The model predicts the lensed features offset within $\lesssim 5$ milli-arcseconds, with the largest offset for the A1-A2 pair. These results may be due a radial profile different than isothermal. However we will address this question in Section 5.7 when including the extended structures in the set of constraints.

Table 5.7: Model 1. SIE + shear + SIS and point-like source. We modelled the lensed source with only one component. In this Table we give the total χ^2, and in the following columns we give the contributions from the different constraints. The uncertainties on the parameters are given below each model. We adopt the same format for all the next tables.

model	feature	N_{dof}	b_l (mas)	(x_l,y_l) (mas)	e	θ_e (deg)	γ	θ_γ (deg)	(x_x,y_x) (mas)	bx (mas)	$\chi^2_{\rm tot}$	$\chi^2_{\rm pos}$	$\chi^2_{\rm flux}$	$\chi^2_{\rm oth}$ [a]
SIE+γ	p,q[b]	3	1166	(−1008,641)	0.24	−21.1	0.13	84.0	–	–	14	0	3	11
			±1	(±38,±25)	±0.02	±19.0	±0.01	±4.0	–					
SIE+SIS+γ	p,q[b]	0	1099	(−1072, 665)	0.20	83.1	0.094	−56.6	(−1457, 2122)	167	0	0	0	0
			±26	(±20, ±17)	±0.07	±10.5	±0.028	±9.1	(±20, ±20)	±36				
	p,q[b]	3	1113	(−1069,653)	0.17	−89.2	0.078	−62.6	(−1457, 2122)	143	1.6	0	1.2	0.4
			±20	(±24, ±14)	±0.07	−13,+5	±0.022	±7.5	(±20, ±20)	±36				
	r	0	1110	(−1072, 665)	0.17	83.6	0.091	−59.4	(−1457, 2122)	148	0	0	–	–
			±2	(±20, ±17)	±0.01	±0.9	±0.002	±5.0	(±20, ±20)	±1				
	r	3	1160	(−1060, 616)	0.23	−47	0.103	79	(−1457, 2122)	15	10.0	0.002	3.7	6.3
			±10	(±17, ±12)	±0.05	± 7	± 0.027	±6	(±20, ±20)	±15				
	p,q, r	12	1062	(−1137, 627)	0.360	−86	0.072	−46	189 (−1441, 2128)		82	49	25	8
			±8	(±20, ±8)	±0.003	±1	±0.011	±3	±11 (±24, ±38)					

[a] contribution from the optical constraints

[b] the position of their centroid was used

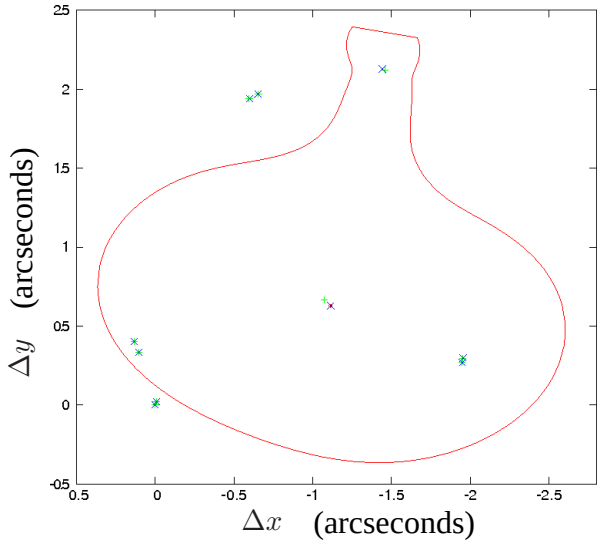


Figure 5.15: The plot shows the comparison between the observed (green crosses) and the modelled (blue crosses) configuration of lensed imaged; the observed and modelled position of the lens galaxy and object-X are also shown. Positions and relative flux-ratios of the core and r features were used to constrain the model.

5.7 Lens modelling with the extended-source structure

The lens model just constrained is used now as starting model for the next step: we use a more complex source parametrization, namely we include in the set of constraints the position of all the Gaussian components listed in Table 5.4, and use the *curve-fitting* technique to model the foreground mass distribution with an SIE+shear+SIS. As explained in Section 2.3, with this algorithm information on the shape of the source is used to constrain the model. Images A1, A2 and B were considered to be *reference-images* with respect to image C, which was the *test-image*. During the optimization routine, the family of curves describing image C was compared with the curves of A1, A2 and B. In this way the software accounts for the fact that image C is described by the least number of components.
We have assumed an accuracy of 1 milli-arcsecond on each of the jet components. Note however that the curve-fitting technique uses the uncertainties only for the references curves.
In Table 5.8 we report the best fitting model parameters and their uncertainties. These were estimated via a statistical approach: the jet components configuration was perturbed within ±10 mas the observed configuration, and each perturbed set of components was used to constrain the model using the curve fitting technique. The standard deviations of each lens parameter were then calculated and used to give the level of uncertainties. Comparing this table with the middle panel of Table 5.7 we see that within the uncertainties the two model agree, however the model constrained by the jet components has a slightly larger ellipticity. The model has a reduced $\chi^2 \sim 3.5$, mainly given by the curve term.

Predicted image configuration and source reconstruction

In Figure 5.16 we show the critical curves of this model, with the predicted positions of the main lens and object-X (image plane) and the source reconstruction (source plane). In Figure 5.17 we show the source reconstruction and the predicted image configuration given by the lens model discussed above. The comparison of Figure 5.17(a) and 5.14(a) shows that for images A1, A2 and C, the offsets between the jet components are reduced, and through the lens mapping one could match and identify the components among them, especially for the core and the more elongated jet. The modelled positions of feature r are still offset by $\simeq 5$ milli-arcsecond. The source reconstruction in Fig. 5.17(a) shows that, for this feature, the modelled position from image C is offset by the other three. Given that this component has a lower signal-to-noise ratio among all the features, it could be that its measured position is affected by a larger uncertainty, and that the offset in the source plane is just due to this. Note also that, though we describe this feature with one Gaussian component, in images A1 and A2 it shows a bit of extended structure. This could be another explanation for the offset in the source reconstruction, namely the uncertainty in identifying this feature among the four images. Image B appears to be not consistent with the other three, particularly in the outer part of the elongated jet, where we measure $\lesssim 5$ milli-arcseconds deviations between the reconstruction from image B and the other three images. In the image plane, the jet components in all the images, but image B, are consistent in all the four image predictions. Offsets in the parallel direction might also indicate uncertainties in the position of the components. For the core feature the model predicts for A2/A1 a flux ratio of $\simeq 1.05$. The source has an angular size of ~ 50 milli-arcseconds, corresponding to a linear size ~ 0.4 kilo-parsec; these numbers are consistent with a sub kilo-parsec scale expected for GHz peaked sources.

The radial profile of the lens galaxy

Recent studies find that the density profile of galaxies has become slightly steeper over cosmic time, with $\partial\gamma/\partial z_{\rm d} \simeq -0.25$ in the redshift range $0.05 - 0.1$, Ruff et al. (2011). In Section 2.6 we have shown that for an axi-symmetric mass distribution the slope of the density profile is degenerate with the size of the Einstein ring, and that in order to break this degeneracy information from lensed structures at different radii is needed. Here we investigate if the VLBI features provide us a useful set of constraints to a radial profile different than isothermal for the lens galaxy. Note that the stronger constraints come from the radial structures, e.g. the jet feature in image C or the core in image A2. We have checked the following profiles $\gamma = 2.01, 2.02, 2.03, 2.05, 2.1$. The total χ^2 for the best model are respectively $\chi^2_{\gamma=2.01} = 40.75$, $\chi^2_{\gamma=2.02} = 42.87$, $\chi^2_{\gamma=2.03} = 41.46$, $\chi^2_{\gamma=2.05} = 42.53$ and $\chi^2_{\gamma=2.1} = 41.05$. In the bottom panel of Table 5.8 we report results for $\gamma = 2.05$, which is the value for γ predicted by the evolution of Ruff et al. (2011). Within the uncertainties the two model are consistent with each other, though we find that for the power-law models the contribution to the total χ^2 from the VLBI core features is slightly larger. Hence we find no strong evidence for a density profile steeper than isothermal. Rather we conclude that the VLBI features are not sensitive to constrain the slope of the density profile.

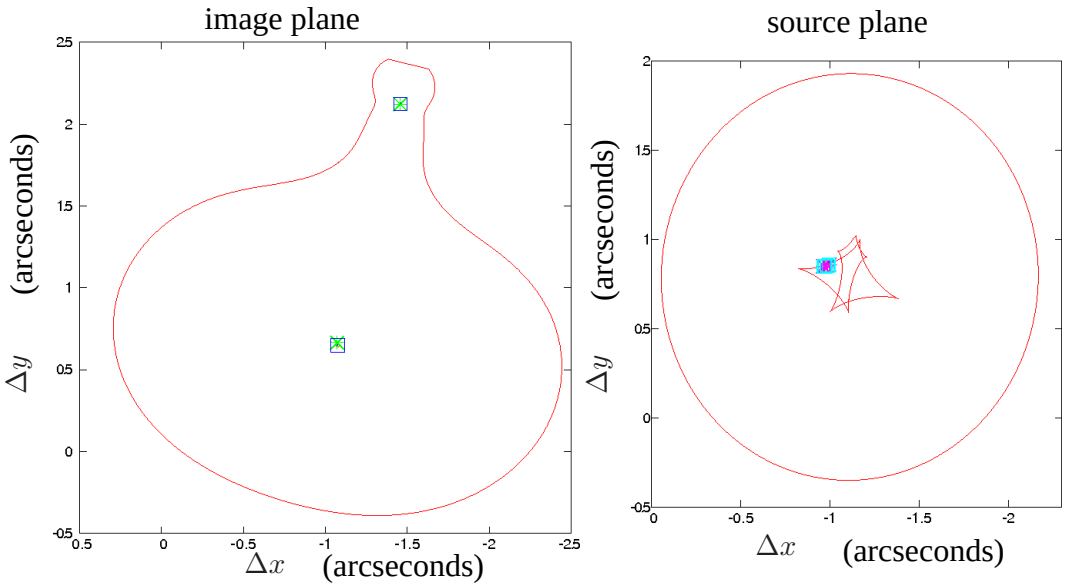


Figure 5.16: SIE+shear+SIS constrained by the extended structures. The plots show the critical curves. In the left panel, the blue squares (green crosses) are the modelled (observed) positions of the lens galaxy and object-X. The source lies near to the tangential curve, as already found for a point-like source description.

5.7.1 Velocity dispersion of the lens galaxy and object-X

For an axi-symmetric mass distribution, the Einstein radius b and the total mass within this radius are related by the following relation

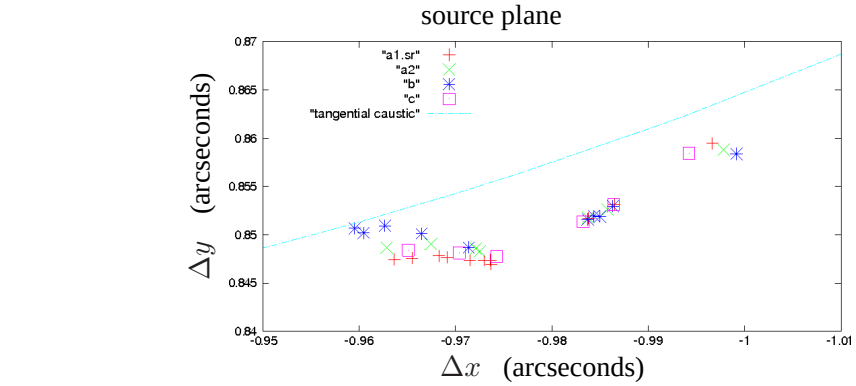
$$M = \frac{b^2 c^2}{4G} \frac{D_\mathrm{d} D_\mathrm{s}}{D_\mathrm{ds}}. \tag{5.3}$$

Recalling Equation (2.9), for a SIS, the Einstein radius can be related to the line of sight velocity dispersion σ_v by

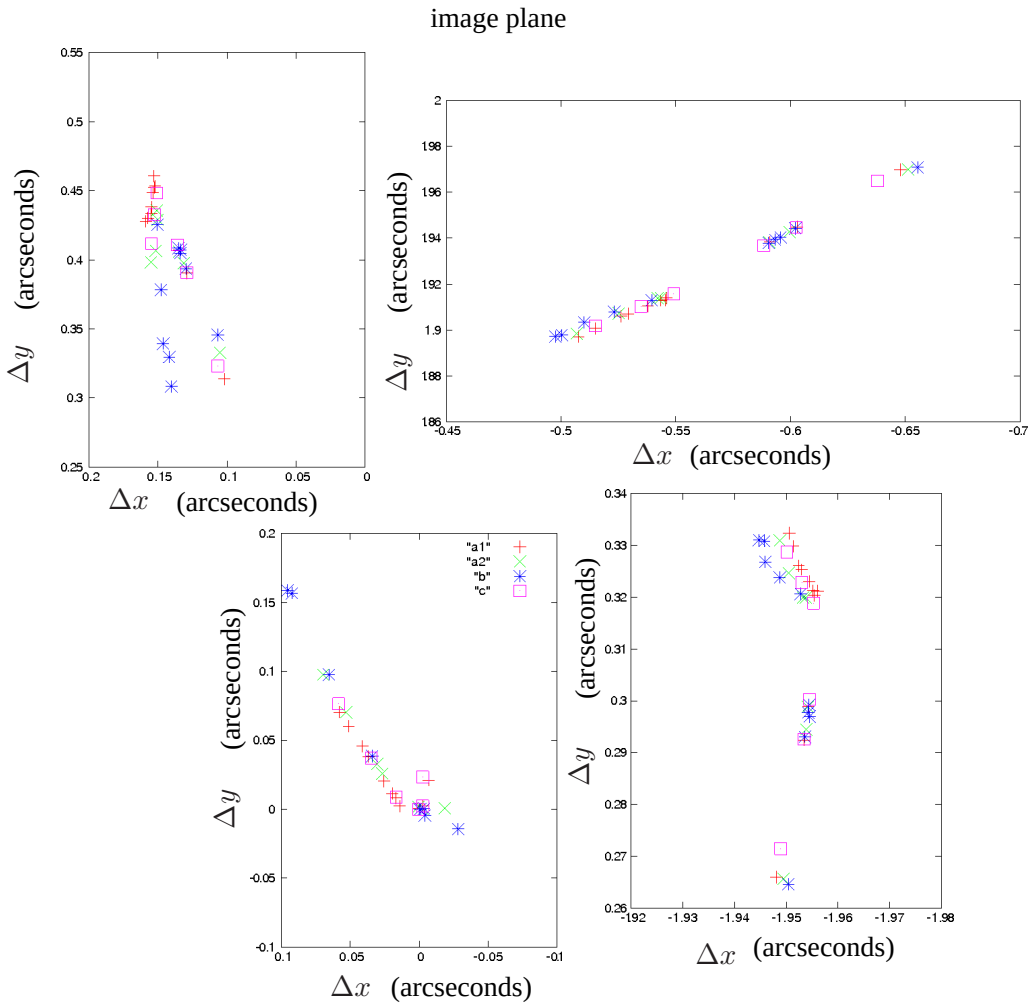
$$b = \frac{4\pi\sigma_\mathrm{v}^2}{c^2} \frac{D_\mathrm{ds}}{D_\mathrm{s}}. \tag{5.4}$$

From the best fitting Einstein radii we have predicted the line of sight velocity dispersion for the lens galaxy. Note that to the lens galaxy G we have fitted a singular isothermal ellipsoid for which there is no simple relation as Equation (5.3); therefore we have predicted its velocity dispersion using an SIS of similar mass. We find the velocity dispersion of the lens galaxy G to be $\sigma_\mathrm{v} = 275 \pm 30$ km s^{-1}. This value is found to be well in agreement with measured values of σ_v for lenses in the Strong Lenses Legacy Survey (SL2S) with median redshift $z_\mathrm{d} = 0.5$ (Ruff et al., 2011). Equations (5.3) and (5.4) imply that the ratio between the Einstein radii of the lens and object-X scales as

$$\frac{b_\mathrm{x}}{b_\mathrm{l}} \propto \left(\frac{\sigma_{\mathrm{v_x}}}{\sigma_{\mathrm{v_l}}} \right)^2. \tag{5.5}$$



(a)



(b)

Figure 5.17: SIE+shear+SIS constrained using the extended structures. Top panel: source reconstructions. Bottom panel: predicted image configurations. Image B does not seem to be consistent with the predictions of the other images. In the source plane, the modelled position of the r component from image C is offset ~ 5 milli-arcsecond with respect to the other images. The reconstruction for the elongated jet from A1, A2 and C images are well consistent with each other. The modelled source from image B is offset in the outer part of the jet by $\simeq 5$ milli-arcseconds with respect the other three images.

With the above equation and the predicted $\sigma_{\rm v}$ for the lens galaxy G the velocity dispersion for object-X is found to be $\sigma_{\rm v_x} \simeq 100$ km s^{-1}. This is consistent with the result given in Chen et al. (2007), who found a value $81 < \sigma_{\rm v_x} < 102$ km s^{-1} at a 95% confidence. The authors only present results for the Einstein radius (or velocity dispersion), they do not give details on their best model, and hence our data could not be used to compare their best model parameters with what we find.

Table 5.8: The model SIE+external shear+SIS constrained using the position of all the Gaussian components in the curve-fitting algorithm.

N_{dof} [c]	b_l (mas)	(x_l,y_l) (mas)	e	θ_e (deg)	γ	θ_γ (deg)	(x_x,y_x) (mas)	b_x (mas)	χ^2_{tot}	χ^2_{pos}	χ^2_{flux}	χ^2_{oth} [a]	χ^2_{crv} [b]
12	1108	(−1074, 648)	0.20	−87	0.073	−61	(−1459, 2124)	146	42.2	0.8	1	0.7	39.5
	±7	(±10, ±10)	±0.05	±3	±0.012	±2	(±10, ±28)	±5					
					Power-law+shear+SIS ($\gamma = 2.05$).								
12	1109	−1076,648	0.20	−87	0.075	−61	(−1456,2124)	151	42.53	1.75	1.22	0.77	38.79
	±8	±12, ±10	±0.04	±2	±0.008	±2	(±16, ±33)	±8					

[a] contribution from the optical constraints

[b] contribution from the curve term

[c] the model is constrained by the positions and flux ratios of the core features, and in addition per each image three curves were used to describe the core region and the two jets

5.8 Implications for a non-smooth mass distribution

In the previous sections we have discussed the lens models constrained by the new 1.7 GHz global-VLBI observations. Here we give an overview on what we find and discuss what the implications are for the mass distribution.
From previous studies, the lens system MG J0414+0534 is a candidate for the detection of CDM substructures (Minezaki et al., 2009). Mismatches between the modelled and observed flux-ratios, or astrometric-anomalies of the lensed sub-components, are the observational evidences that would prove such a mass distribution. Let us consider first the flux anomaly. We have discussed in Section 5.3.1 that mid-infrared flux-ratio measurements are not in agreement with the prediction of a smooth mass model for the pair A2/A1, which is found to be demagnified with respect to the prediction from such a model. The measured flux-ratios from radio observations (this work and all the previous ones, see Table 5.2) are found to be in agreement with mid-infrared observations. For the lens systems Q0957+561 (Garrett et al., 1994) and PKS1830-211 (Jin et al., 2003), by modelling the relative magnification matrix between images, the authors infer limits on the substructures in the mass distribution. The method is based on the fact that ideally, the surface brightness of lensed images is conserved, hence the observed flux density ratios of the images should be equal to the ratio of their solid angle. This is defined to be the determinant of the relative magnification matrix between the images, which gives the transformation of one image into another. For a given pair of images whose observed flux ratio is not affected by substructures the measured value should correspond within the uncertainties to the determinant of the modelled matrix. We could not apply successfully this method to our data, since as among the four images we can match only two sub-components, the core and r features. The components of the elongated jets are not easy to identify because of the high magnification gradients.
Let us now consider the astrometry of the system. Astrometric perturbations from substructures are of order of $10 \lesssim$ milli-arcseconds. However, when CDM substructures are present in the mass distribution, it is possible to detect astrometric anomalies if the position of the lens galaxy is fixed within $\lesssim 10$ milli-arcseconds; when the centre of the potential is not fixed, instead is let to float, any non-smooth mass model, and hence the astrometric anomalies, is degenerate with the fitted smooth mass model (Chen et al., 2007). These considerations are important to understand our results (the χ^2 of the model and the source reconstruction). Our best mass model does not account well for the extended structure of the source, with offsets of $\simeq 5$ milli-arcseconds (in the source plane). Also, we have assumed an uncertainty of 20 milli-arcseconds on the position of the lens galaxy (to account for the uncertainty on the registration of the radio and optical data), and that we have let this parameter free within this range; hence if substructures are present in the mass distribution, our smooth mass model may be degenerate with the presence of small mass clumps, responsible for the 5 milli-arcseconds shift that we measure from the source reconstruction.
The χ^2 of our best mass model has always a small contribution from the flux-ratios, and the model prediction for the A2/A1 flux ratio is not consistent with what observed, namely image A2 appears to be demagnified with respect to model prediction. Thus the conclusion that the mass distribution is characterized by multi

components seems to be realistic. However, it is not intuitive but rather arbitrary to determine the parameters of the mass clumps.
Trotter et al. (2000), who also aimed at modelling the mass distribution including all the structures revealed by the radio VLBI observations, were able to fit the centre of the gravitational potential consistently with optical measurements of the lens galaxy positions, but their fit suggests that the mass distribution is highly asymmetric. Though their description of the mass distribution is very difficult to correlate with a physical parameters, we interpret this result consistent with a non-smooth mass distribution.

5.9 The optical arc

Deep HST-imaging by Falco et al. (1997) revealed the presence of an extended arc connecting image A2 and B, bluer than the core images. It appears stretched in the tangential direction relative the centre of the lens galaxy, hence most likely it is due to an extended structure laying close to the tangential critical curve, in the vicinity of the lensed QSO. Here we address this question.
In Figure 5.18 we show the i-band image from the CASTLES project (Kochanek et al., 1999) where the signal of the arc is stronger. Our aim is to find the corresponding position in the source plane. In order to achieve this goal, we have placed circular regions along the arc (these are the black circles in Figure 5.18) in a way that the collection of all the regions describes the whole arc from image B to A2. Note that the arc extends over $\sim 1''$. The positions of the centres of the regions, referred to image A1, were projected into the source plane using the model of Table 5.8, and for each component in the source plane the corresponding image positions were predicted. Figures 5.20 and 5.19 show the results of this exercise. In the source plane the arc components correspond to positions close to the tangential caustic, in the proximity of the lensed QSO. In the image plane they are mapped into four regions, which we label D,E,F and G. The arc is produced by the merging pair D-E which is the brightest one. The other two lensed regions, F and G, are produced very close to the images A1 and C of the QSO, making it very hard to disentangle their emission from the lensed emission of A1 and C. The left panel in Figure 5.20 shows a zoom on the merging pair D-E, illustrating how the pair moves for different source positions. In the next section we give an interpretation of this result.

5.9.1 Interpretation on the nature of this region

Up to now there is no information on the redshift of the arc, which we assume to be equal to $z_{\rm QSO}$. The emission from the arc, as seen in the optical image, seems to be naturally extending from the bright peaks of images B and A2, thus the assumption that the arc is a lensed structure connected to the QSO is reasonable. In the source plane (Figure 5.19), the source lensed into the arc extends over ~ 20 milli-arcseconds in the x-axis, and ~ 40 milli-arcseconds in the y-axis, at a project distance from the QSO of ~ 70 milli-arcseconds. From the redshifts of the lens and the source galaxies, $1''$ in the source (lens) plane corresponds to 7.97 (7.92), hence the projected separation between the QSO and the extended source is ~ 0.56 kilo-parsec. These numbers are consistent with the interpretation that the emission of the arc is lensed

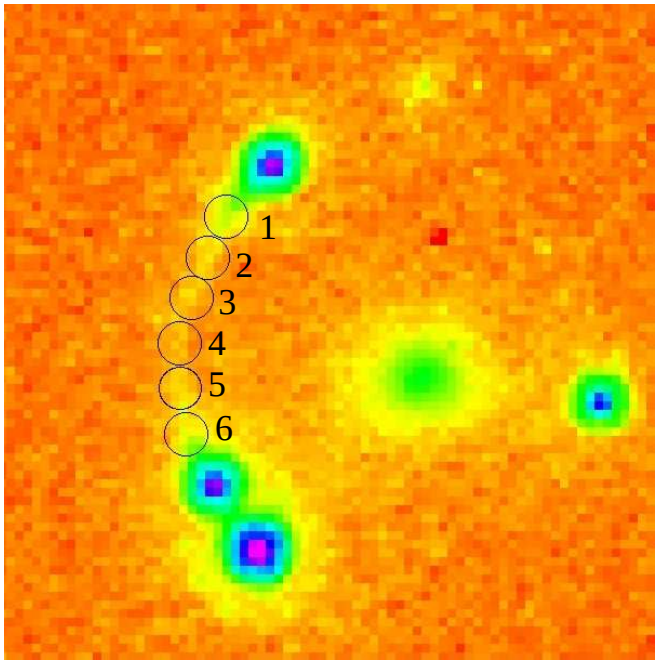


Figure 5.18: The figure shows the R-band image of the system from the CASTLES project (Kochanek et al., 1999) project. The black circles (radius of 0.113″) trace the optical arc. To identify the circles in the image and source plane (see also next figures) we have used numbers from 1 to 6 to label them.

emission from the QSO host galaxy. The R−I colour of the arc is 1.3±0.3 mag (Falco et al., 1997); similar colours are found for galaxy-type objects at the same redshift of the QSO. Hence, it is likely that this is lensed emission of a star-forming region in the host galaxy of the QSO, probably induced by a merging event in the Early Universe. Note that the sizes for this region would imply a system of star-forming regions rather than just a single one. Further investigation on the colour of the host galaxy is currently being done, by which we aim at identifying the galaxy template.

5.10 Truncation radius for object-X

Strong gravitational lensing has been successfully used to study the properties of galaxy dark matter halos, in particular to constrain sizes of halos truncated by tidal stripping in the gravitational potential of clusters or group of galaxies. In particular, the works by Suyu et al. (2010) and Vegetti et al. (2010) show that when lensed structures are seen at distances as large as the size of the *truncation radius* itself strong gravitational lensing is able to constrain directly the size of this radius without any assumption on scaling relations (e.g Halkola et al. 2007; Natarajan et al. 2009) nor on the light profile of the satellite (e.g. Donnarumma et al. 2011). Within this radius the density profile of the mass distribution is assumed to be isothermal, namely the density ρ falls as r^{-2}, while for larger radii the density declines as r^{-4}. Note that from a physical point of view the truncation radius of the satellite has

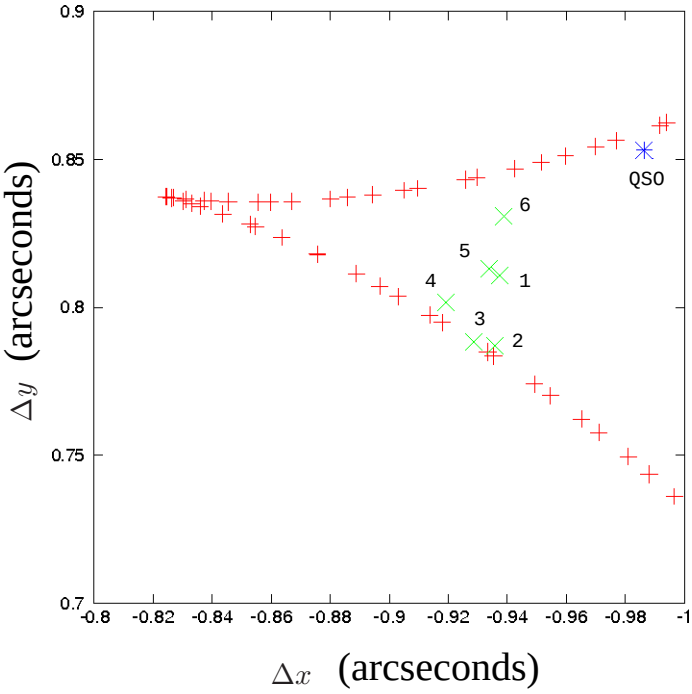
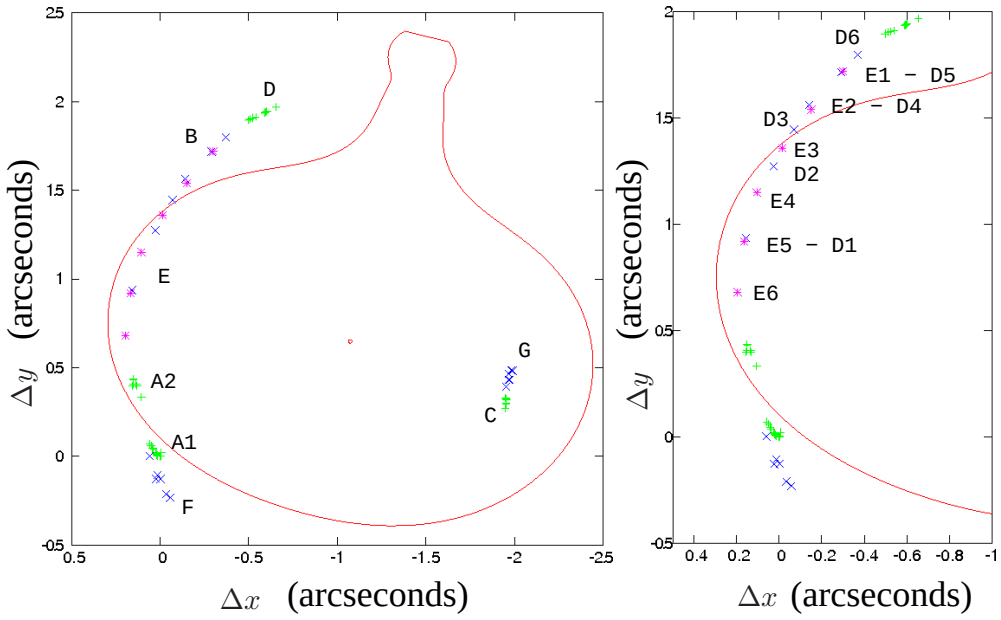


Figure 5.19: The figure shows the position of the lensed QSO and the extended region in the source plane. The source of the arc is found at approximately 70 (projected) milli-arcseconds of the QSO, close to the tangential caustic curve.

Figure 5.20: The purple asterisks correspond to the centre of the regions describing the arc. These were projected into the source plane and the corresponding images in the image plane were predicted. These are shown here with the blue crosses. The green ones represent the observed image configuration of the lensed QSO. The right panel shows a zoom in the same plane, on the merging pair E and D. See text for more details on the model.



to be consistent with its tidal radius, which defines the three-dimensional distance between the main galaxy and the satellite where the gravitational tidal forces due to the galaxy are equal to the gravitational force of the satellite. Mass particles beyond the truncation radius of the satellite fall into the gravitational potential of the main galaxy. For cluster member galaxies, or in general not isolated objects, this radius is found to be within $\lesssim$ 15 kilo-parsec (see Fig.10 in Suyu et al. 2010). In the lens system discussed here, object-X lays within $0.9''$ (7.1 kilo-parsec) from image B, $\sim 1.5''$(11.8 kilo-parsec) from the lens galaxy, and $\sim 2.5''$ (19.8 kilo-parsec) from image A1. Given the angular distances discussed above, we can only say that the truncation radius for object X could be constrained by the VLBI data if within $0.9'' - 2.5''$. We have fitted a lens model consisting of a singular ellipsoid (SIE), external shear and a truncated isothermal sphere (TIS). Results are summarized in Table 5.9. The reduced χ^2 is 0.8, with the main contribution given by the flux ratios. Figure 5.21 shows the χ^2 function of this parameter, over the range within which t_r is varied while fitting all the other parameters. The two lines identify the region corresponding to $\Delta\chi^2 \leq 1$. For any value in this range the $\Delta\chi^2$ function is not a parabola centred on $\chi_{\rm min}$; within this range the function is rather constant, implying that the set of constraints is not sensitive to this model. Note that for radii smaller than $0.6''$(4.7 kilo-parsec) we find $\Delta\chi^2 > 1$, hence we constrain the truncation radius to be at least 4.7 kilo-parsec. The comparison of this model with the SIE+γ+SIS model (middle panel of Table 5.7 shows that the reduced χ^2 is slightly larger for the SIE+γ+TIS model. There is no evidence for a finite truncation radius for object-X.

Table 5.9: Results for the modelSIE+shear+TIS.

N_{dof}	b_l (mas)	(x_l,y_l) (mas)	e	θ_e (deg)	γ	θ_γ (deg)	(x_x,y_x) (mas)	b_x (mas)	t_r (mas)	χ^2_{tot}	χ^2_{pos}	χ^2_{flux}	χ^2_{oth}
2	1140	(−1078,654)	0.22	−88	0.075	−59	(−1455, 2122)	168	1615	1.6	0	1.2	0.4
	±13	(±15,±15)	±0.08	±2	±0.016	±7	(±9, ±8)	±33	–				

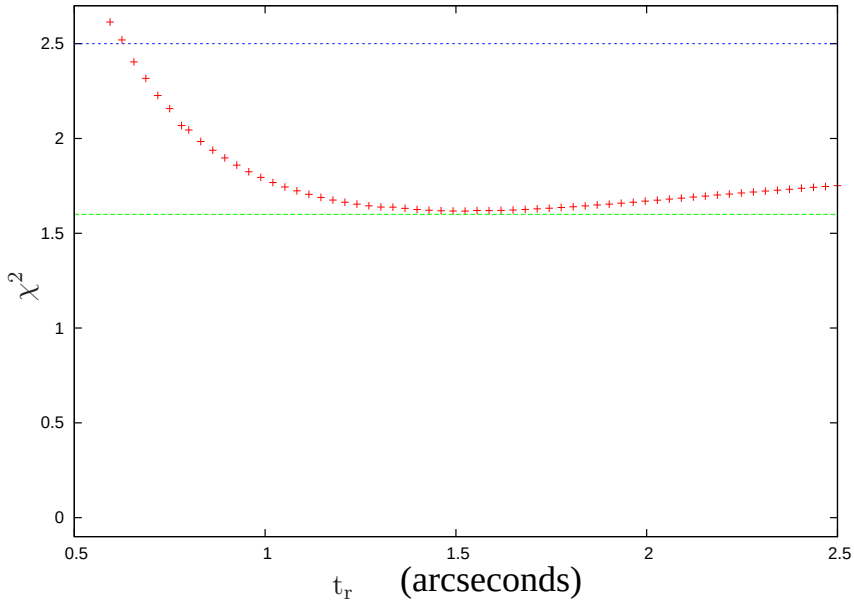


Figure 5.21: The curve shows a cut of the χ^2 surface along the truncation radius. The green line corresponds to $\chi^2_{\rm min}$, and the blue one to the $\chi^2_{\rm min} + 1$. The value of t_r that gives $\Delta\chi^2$ larger than 1 sets the lower limit for this parameter at $0.6''$. (see text for more details)

5.11 Conclusions

New sensitive and high-angular resolution observations for the gravitational lens MG J0414+0534 were carried out to investigate lens mass models that account for the complex extended lensed structures. We have presented new radio maps obtained from observations taken with a global-VLBI array at 1.7 GHz. The new maps confirm the extended structures revealed by previous observations at higher frequencies (Trotter et al., 2000; Ros et al., 2000). We have placed a 3σ upper limit on the flux density of an odd image near the centre of the lens galaxy, or emission from the lens, of 0.4 mJ beam^{-1}.

We have modelled the mass distribution in the foreground lensing plane using two different source models, the standard point-like one and subsequently a more complex one describing the lensed extended structures. The best lens model (reduced $\chi^2 \simeq 3.5$), consisting of a SIE+shear+SIS, with both lenses at the same redshift ($z \simeq 0.958$) accounts for the complex structure revealed by the radio observations. In the source plane the source is found to have linear sizes of ~ 0.4 kilo-parsec, well in agreement with the GHz peaked structure implied by the spectral analysis (Katz et al., 1997). However we find that a non-smooth mass distribution is a realistic description for the lensing potential. This result is consistent with studies done at mid-infrared wavelengths that investigate the flux-anomaly problem in the A1/A2 pair, though it is not intuitive what would be a realistic distribution of CDM structures that would account for the observed astrometry and flux-densities. Moreover fitting the same mass model using a different set of constraint, leads to a poorer fit (reduced $\chi^2 \simeq 6.8$) probably due a radial profile for the lens galaxy which is not isothermal. However we find that VLBI constraints are not sensitive enough to constrain the slope of the density profile. Our best model accounts for a lensed arc revealed by optical observations, which we find to be lensed emission of a source connected with the AGN. The predictions for the velocity dispersions of both lenses and a lower limit for the tidal radius of object-X are found to be in agreement with other independent studies.

It is the theory that decides what can be observed.

A. Einstein

6 Conclusions and future work

In this thesis two gravitationally lensed systems were studied using radio interferometric observations, J002240+143130 and MG J0414+0534. We were interested in exploiting gravitational lensing to probe the mass distribution and the properties of the lensed source, though this last point was addressed differently in the two projects.

The 8 o'clock arc (J002240+143130)

From optical and near-infrared observations, the system is known to be a Lyman Break galaxy at $z \simeq 2.73$, forming stars at a rate of $\sim 270\ M_{\odot}\mathrm{yr}^{-1}$ (Allam et al., 2007; Finkelstein et al., 2009). It is lensed into an arc by a lens galaxy at $z \simeq 0.38$. The rate at which the galaxy is forming stars is implying a radio flux density at 1.4 GHz of $\sim 120\mu$Jy. However, in the NVSS catalogue (Condon et al., 1998), the lens system is associated with a 5 mJy source which would predict a much larger star-formation rate. Main goal of this project was to study and understand the origin of the excess radio emission in the lensed galaxy.
We have observed the system J002240+143130 at 1.4 and 5 GHz with the VLA telescope in the B-array configuration, and again at 1.4 GHz in the A-array to obtain a higher angular resolution. We find that most of the 5 mJy radio emission is associated with an AGN radio galaxy whose peak coincides with the optical position of the lens galaxy. From the spectral analysis, carried out between 1.4 and 5 GHz, we find that the AGN is a FRI type radio galaxy, characterized by a compact core component, a steep spectrum jet extending from the core of the AGN, and an extended radio lobe, aligned with a jet propagating from the core in the opposite direction with respect to the lobe. We find no evidence for radio emission above the 3σ level associated with the lensed arc. Our observations should have been sensitive enough to detect any radio emission from the lensed arc. However we could not use half of the data, since part of the observing frequency band was lost, and the radio jet is covering most of the arc. Hence we could only infer an upper limit on the radio

emission from the arc of 108μ Jy beam^{-1}, which corresponds to a radio-derived star formation rate $\leq 750 M_{\odot} yr^{-1}$. This limit agrees with what found from optical and near-infrared data.

Future work

In order to study the properties of the star-forming activity in the lensed LBG we have started a new project whose goal is to quantify its content of gas mass. In February 2011, we have observed the J $1 \longrightarrow 0$ transition in the lensed LBG with the Effelsberg telescope. The data analysis has not yet been performed. After H_2, CO is the most abundant molecule whose rotational levels are excited by collision with H_2. Hence CO is a good tracer of molecular hydrogen, fuel for star formation. From these observations we aim at obtaining the total gas mass from the luminosity of the line, and from the comparison with the existing estimates of the SFR, we will obtain constraints on the efficiency of the star-formation process in the lensed galaxy. Furthermore, results from these observations may lead to new investigations (through higher resolution spectral lines observations and observations of higher CO transitions) to understand the dynamical state of the gas content, its temperature and density.

MG J0414+0534

The system is an example of image configuration produced by an elliptical mass distribution. At optical wavelengths we observe the four images of the QSO, the lens galaxy and a dwarf galaxy, most probably associated with the main lens, and an arc connecting two of the images. Previous observations at radio wavelengths have shown multi-components for each of the lensed images, and no radio emission associated with the lens galaxy, the dwarf galaxy or the arc. We have analysed new VLBI observations at 1.7 GHz. We resolve the four lensed images and their sub-components, namely a compact core region and two extended jet structures. The 1.7 GHz maps confirm a wealth of structures consistently with previous observations at higher frequencies (Ros et al., 2000; Trotter et al., 2000). In particular we map with a higher signal-to-noise the outer regions of the radio jets. We have performed Gaussian model fitting on the VLBI maps to describe the position, strength of emission and size of the sub-components in each of the four images, and built the set of constraints which was used to study various lens models. Main goal of this project was indeed to improve existing lens models, which where so far constrained with simpler source parametrization. Previous lens models constrained by VLBI observations (Ros et al., 2000) find that the observed configurations of the radio core features can be accounted by a lens system consisting of an isothermal ellipsoid (SIE; describing the main lens galaxy), an isothermal sphere (SIS; describing the dwarf galaxy) and an external shear contribution (γ). We have tested this mass model using the 1.7 GHz VLBI cores. We find that the best mass model has a reduced $\chi^2 \sim 0.5$ and, within the uncertainties on the lens model parameters, it is consistent with results from Ros et al. (2000). However, this model accounts for the observed configuration of the core features, not for the jet structures. Furthermore, model-fitting with other sets of constraints, (e.g., using point-like source parametrization

for another radio feature, but the core, or using a two point-like source parametrization) suggests that the radial density profile may be different than isothermal. Next, we have tested mass models using a source parametrization that accounts for the extended structure of the source, namely the core and both jets. We then use the new constraints and study the SIE+γ+SIS model. We find that the best mass model has a reduced $\chi^2 \sim 3.5$, and a slightly larger ellipticity than model discussed above. The source reconstructions obtained projecting into the source plane the four lensed images shows deviations $\lesssim 5$ milli-arcseconds for one of the images. This could be due to substructures mass not accounted with the SIE+γ+SIS mass distribution. Note that this system is known to be a candidate for mass substructures: existing lens mass models (e.g., Minezaki et al. 2009), and also the one just discussed, predict the so-called parity dependent flux-ratio anomaly which is considered to be a signature of mass substructures. However, it is not intuitive to predict what would be the properties of a realistic multi-components mass distribution.
Aiming at constraining a radial profile different than isothermal for the lens galaxy, we have tested power-law models for the lens galaxy (the model also includes the SIS+γ for the dwarf galaxy and the shear component). From a recent study (Ruff et al., 2011), the radial density profile for the lens galaxy at redshift $z \simeq 0.958$ is expected to be slightly steeper than isothermal. However, we find no evidence for a similar radial profile and interpret this result with the VLBI features not being sensitive enough to constrain the slope of the radial density profile. An attempt was made to constrain the size of the truncation radius of the dwarf galaxy, fitting a a truncated isothermal sphere for it (in addition, the model also includes the SIE+γ for the main lens and shear, as above). We find no evidence for a finite truncation radius for this galaxy, instead we infer a lower limit for it of ~ 4.7 kilo-parsec.
Finally, we find that the optical arc is most likely lensed emission of a star-forming region in the host galaxy of the QSO, and obtain a source reconstruction for it that explains the arc structure which is observed.

Future work

We are currently investigating the colour of the lensed source (from available HST data) in order to characterize the template of the host galaxy of the QSO (work in collaboration with J.P. McKean).
Another aspect that certainly needs to be investigated concerns the properties of the non-smooth mass distribution which could account for the observed flux-ratios and astrometry of this system. Hence, multi-components mass distributions or a lensing potential with higher order multipoles will be tested as well. Moreover, data reduction and analysis of the latest 8 GHz VLBI observations (O. Wucknitz et al. 2008) should be performed; results of this analysis compared to the 8 GHz VLBI data of Ros et al. (2000) will reveal morphological changes of the lensed source. An accurate astrometry (as the one provided by VLBI measurements) of quad gravitational lenses has the potential of detecting the proper motion in the lens galaxy, which is otherwise very difficult to perform in optical astronomy (see Kochanek et al. 1996).

Outlook

In the near future, thanks to the improvements made for observations from FIR to radio bands, e.g., the new *Herschel* instrument, ALMA, and the upgraded EVLA telescope, it will possible to obtain a direct view into the properties of star-forming galaxies. Detailed studies of statistically significant samples of LBGs will be possible, which will allow us to address important questions on galaxy formation. On the other hand, beside benefit of the gravitational magnification that enables the detection of high-redshift galaxies, studies of lensed star-forming galaxies will give us constraints for lens modelling purposes. With the advent of new facilities (LOFAR and SKA), the radio sky will be seen with new resolution and better sensitivity, opening the way to more gravitational radio-lens systems being found.
Practically this means that studies of the properties of the lens galaxies, their global mass distribution and radial profiles, will be carried out on larger samples yielding new informations on the structure and evolution of galaxies.

A

Model-fitting the VLBI maps

Features seen in astronomical images are usually described by a set of parameters, so they can be compared at different observing wavelengths, or with observations taken at different epochs. In most cases it is convenient to specify a model for the intensity distribution, and determine the parameters that describe a component within the specified model.
The `AIPS` software provides two tasks that allow us to fit physical parameters that describe the components seen in the radio map, MAXFIT and JMFIT.
For a given map array, MAXFIT determines the position and strength of its peak fitting a quadratic function using the least square method. This task does not perform the goodness of fit nor error estimates. It is conveniently used to find the location and the intensity of the peak, which can be then used as starting model for a more complex fit. JMFIT fits a two-dimensional elliptical Gaussian to a region of the image around the component. It uses the least square method, since it assumes that the pixels error distribution is Gaussian with constant rms over the component. However this is not the case for features in radio interferometric images, where over the map the pixel errors are correlated with each other by the beam pattern. To estimate the errors on the fitted parameters we have used the analytical approach, described in Fomalont (1999), that accounts for the fact that Gaussian model fitting in radio interferometric maps has uncertainties which depend on the signal-to-noise of the detection. This is described below.
Let σ_{rms} be the rms noise of the residual map, then the uncertainty on the modelled peak-flux S_p is

$$\sigma_{S_{peak}} = \sigma_{rms}; \tag{A.1}$$

the uncertainty on the modelled total flux density S_{tot} is approximated by

$$\sigma_{S_{tot}} = \sqrt{\sigma_{rms}^2 + (S/P)^2}. \tag{A.2}$$

Let d be the size of the component, then the uncertainty on the modelled position[1] is

$$\sigma_{pos} = \frac{\sigma_{rms} d}{2 S_{tot}}. \tag{A.3}$$

These equations give rough approximations of the true errors, which may be larger. When fitting several components, not well distinguished, their uncertainties may be larger than what given above.

[1] For components smaller than the beam size, the beam size is taken as d, and for components larger than the beam size, the size of the component is taken as d

Acknowledgements

And here I am, almost at the very end of this work, to acknowledge all the people without whom I would probably not finish like this. My supervisor, Dr. Olaf Wucknitz who gave me the chance to work on this project. For being all the times ready and willing to carefully answer my questions. His highly critical comments were always source of inspiration to me. Thanks to his *silent*, discreet and supporting presence without which I would probably be less conscious of my work and my understanding of physics and astronomy. A deep 'thank you' goes to Dr. Richard W. Porcas for all the quizzes and *tortures* he had for me, his high curiosity and interest for nature. For always sharing his ideas and being honest, and for chairing all Friday-night events. Thanks to Prof. Peter Schneider for the interest in this work; he has never hesitated to check the status of my project nor to express his opinion. Thanks to Dr. Andreas Brunthaler for being, together with all the people mentioned so far, a member of my thesis committee. A special thanks goes to Prof. Eduardo Ros for his very open mind to science and to mankind. Never without an answer or advice to give. I am more than glad to have you as a friend and collaborator. And you know, I think the world would need more than just one 'Eduardo'. Simply, thanks. Thanks to Dr. John McKean for his great enthusiasm for astronomy that makes him a perfect collaborator. Thanks to the International Max Planck Research School (IMPRS) coordinated by Prof. Eduardo Ros and Dr. Manolis Angelakis, with the collaboration of Frau Gabi Brauer and Simone Pott. The strength they had in carefully organizing lectures, seminars, retreats, talks and parties was of great motivation for me to participate and enjoy all the activities. Thanks to Dr. Jose A. Muñoz for the valuable collaboration during the second part of my project. Thanks to the Gravitational-Lensing&Cosmology Group lead by Prof. Peter Schneider and Prof. Cristiano Porciani, and to all the lensing-seminar speakers, for the intriguing discussions on the physics of the Universe. Among all the lensers and cosmologists, thanks to Sherry Suyu, Leonardo Castaneda, Xinghong Er, Xun Shi, Holger Israel, Ismael Tereno, Jacqueline Chen, Zahra Sheikhbahaee, Stephan Hilbert, Jan Hartlab, Jacqueline Chen, Emilio Pastor Mira and Anna Elia, in one way or another you were all inspiring to me. Thanks to Frau Christina Stein-Schmitz, Frau Kate Schrüfer, Frau Kristina Sörgel, Frau Elisabeth Kramer and Frau Elisabeth Danne for all the help with documents and bureaucracy, most of the times more than what you are asked to do in your job; your presence at the Argelander-Institut is very precious. Thanks to my officemates Yu-Ying Zhang and Hanhieh Mahmoudian. I can say I shared most of my time in Bonn with you, thanks for your every-day smiles.
And now here I come to acknowledge all the people whose presence has been crucial and of fundamental help in my life outside the Institute. My family, mamma papà e Piercarmine. Mamma&papà because I have never felt alone and far from

you. For always being interested and curious in what I do, in what life abroad is like, in learning new technology so that we could overcome the distance, and the unconditional Love and support that I receive from you. Piercarmine for his big trust in me and in my interests. 'Doctor' Luciano Cerrigone, perfect neighbour and 'Italian abroad', without whom I cannot imagine what my life in Bonn could be like. For all the meals we shared, deep discussions about science, humanity, politics, democracy and religion, and for all the *stupid questions* discussed together. It has been a huge pleasure knowing you, and I wish wherever you go a lot of happiness. Thanks to Marios, Luis, Gabriele&Alessia, Yoon&Jan, Jorge, Laura, Felipe, Lisa, Brenda, Amir, John Morgan, Frank, Chin-Shin, Kirill, Christian, Sybille, Mar, Carlos, Alberto, Michelle for all the nice time outside the Institute, all the parties with good food, the fun and adventures we had in and around Bonn. I truly wish that you all find what you are looking for in life. To the whole 'Friday night crowd' and the ethnic place for just being there, both the former and the latter. To my Italian friends, and among all Maria&Cristiano, Federica&Andrea, Carmen&Giorgio, Concetta, Alberto, Ale, Mariella, Giuseppe, Sandro, Rossella, Loredana and Antonio, for just being a phone-call or email away. Because thanks to you all, I can say that time and distance do not matter in life and friendship. And finally, and certainly not the least, Gian Franco without whom I would not have dreams. For all the deep Love and emotions that you give me, your constant presence and support. For all this, and because I am here right now, I owe you more than a thank you. Grazie.

Bibliography

Adelberger, K. L., Steidel, C. C., Pettini, M., et al. 2005, ApJ, 619, 697

Allam, S. S., Tucker, D. L., Lin, H., et al. 2007, ApJ, 662, L51

Allam, S. S., Tucker, D. L., Smith, J. A., et al. 2004, AJ, 127, 1883

Bauer, A. H., Seitz, S., Jerke, J., et al. 2011, ApJ, 732, 64

Belokurov, V., Zucker, D. B., Evans, N. W., et al. 2007, ApJ, 654, 897

Bennett, C. L., Lawrence, C. R., Burke, B. F., Hewitt, J. N., & Mahoney, J. 1986, ApJS, 61, 1

Bielby, R., Shanks, T., Weilbacher, P. M., et al. 2010, ArXiv e-prints

Biggs, A. D., Browne, I. W. A., Muxlow, T. W. B., & Wilkinson, P. N. 2001, MNRAS, 322, 821

Biggs, A. D., Wucknitz, O., Porcas, R. W., et al. 2003, MNRAS, 338, 599

Binney, J. & Tremaine, S. 1987, Galactic dynamics, ed. Binney, J. & Tremaine, S.

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

Browne, I. W. A., Wilkinson, P. N., Jackson, N. J. F., et al. 2003, MNRAS, 341, 13

Burgarella, D., Pérez-González, P. G., Tyler, K. D., et al. 2006, A&A, 450, 69

Burke, W. L. 1981, ApJ, 244, L1+

Calzetti, D. 1997, AJ, 113, 162

Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, ApJ, 429, 582

Carilli, C. L., Lee, N., Capak, P., et al. 2008, ApJ, 689, 883

Chen, J. 2009, A&A, 498, 49

Chen, J., Rozo, E., Dalal, N., & Taylor, J. E. 2007, ApJ, 659, 52

Clark, B. G. 1980, A&A, 89, 377

Coe, D. & Moustakas, L. A. 2009, ApJ, 706, 45

Condon, J. J. 1992, ARA&A, 30, 575

Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, AJ, 115, 1693

Conley, A., Guy, J., Sullivan, M., et al. 2011, ApJS, 192, 1

Cooke, J. H. & Kantowski, R. 1975, ApJ, 195, L11+

Coppin, K. E. K., Swinbank, A. M., Neri, R., et al. 2007, ApJ, 665, 936

Cornwell, T. J. & Perley, R. A. 1992, A&A, 261, 353

Curran, S. J., Whiting, M. T., Tanna, A., Bignell, C., & Webb, J. K. 2011, MNRAS, L231+

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

Diemand, J., Kuhlen, M., & Madau, P. 2007, ApJ, 667, 859

Donnarumma, A., Ettori, S., Meneghetti, M., et al. 2011, A&A, 528, A73+

Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, ApJ, 633, 560

Eke, V. R., Cole, S., & Frenk, C. S. 1996, MNRAS, 282, 263

Elíasdóttir, Á., Hjorth, J., Toft, S., Burud, I., & Paraficz, D. 2006, ApJS, 166, 443

Elíasdóttir, Á., Limousin, M., Richard, J., et al. 2007, ArXiv e-prints

Falco, E. E., Impey, C. D., Kochanek, C. S., et al. 1999, ApJ, 523, 617

Falco, E. E., Lehar, J., & Shapiro, I. I. 1997, AJ, 113, 540

Fassnacht, C. D., Blandford, R. D., Cohen, J. G., et al. 1999, AJ, 117, 658

Finkelstein, S. L., Papovich, C., Rudnick, G., et al. 2009, ApJ, 700, 376

Fomalont, E. B. 1999, in Astronomical Society of the Pacific Conference Series, Vol. 180, Synthesis Imaging in Radio Astronomy II, ed. G. B. Taylor, C. L. Carilli, & R. A. Perley, 301–+

Garrett, M. A. 2002, A&A, 384, L19

Garrett, M. A., Calder, R. J., Porcas, R. W., et al. 1994, MNRAS, 270, 457

Garrett, M. A., Knudsen, K. K., & van der Werf, P. P. 2005, A&A, 431, L21

Garrett, M. A., Nair, S., Porcas, R. W., & Patnaik, A. R. 1996, in IAU Symposium, Vol. 173, Astrophysical Applications of Gravitational Lensing, ed. C. S. Kochanek & J. N. Hewitt, 189–+

Garrett, M. A., Porcas, R. W., Pedlar, A., Muxlow, T. W. B., & Garrington, S. T. 1999, New Astronomy Rev., 43, 519

Giavalisco, M. 2002, ARA&A, 40, 579

Halkola, A., Seitz, S., & Pannella, M. 2007, ApJ, 656, 739

Haschick, A. D., Moran, J. M., Reid, M. J., Davis, M., & Lilley, A. E. 1981, ApJ, 243, L57

Hawkins, E., Maddox, S., Cole, S., et al. 2003, MNRAS, 346, 78

Hewitt, J. N., Turner, E. L., Lawrence, C. R., Schneider, D. P., & Brody, J. P. 1992, AJ, 104, 968

Högbom, J. A. 1974, A&AS, 15, 417

Impellizzeri, C. M. V., McKean, J. P., Castangia, P., et al. 2008, Nature, 456, 927

Ivison, R. J., Chapman, S. C., Faber, S. M., et al. 2007, ApJ, 660, L77

Jaffe, W. 1983, MNRAS, 202, 995

Jarvis, M. J., Smith, D. J. B., Bonfield, D. G., et al. 2010, MNRAS, 409, 92

Jennison, R. C. 1958, MNRAS, 118, 276

Jin, C., Garrett, M. A., Nair, S., et al. 2003, MNRAS, 340, 1309

Katz, C. A., Moore, C. B., & Hewitt, J. N. 1997, ApJ, 475, 512

Keeton, C. R. 2001a, ArXiv Astrophysics e-prints

Keeton, C. R. 2001b, ArXiv Astrophysics e-prints

Keeton, C. R., Kochanek, C. S., & Seljak, U. 1997, ApJ, 482, 604

Kennicutt, Jr., R. C. 1998, ARA&A, 36, 189

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

Kneib, J. P., Mellier, Y., Fort, B., & Mathez, G. 1993, A&A, 273, 367

Kochanek, C. S. 2006, in Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro, ed. G. Meylan, P. Jetzer, P. North, P. Schneider, C. S. Kochanek, & J. Wambsganss, 91–268

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

Kochanek, C. S., Falco, E. E., Impey, C. D., et al. 1999, in American Institute of Physics Conference Series, Vol. 470, After the Dark Ages: When Galaxies were Young (the Universe at 2 ¡ Z ¡ 5), ed. S. Holt & E. Smith, 163–175

Kochanek, C. S., Keeton, C. R., & McLeod, B. A. 2001, ApJ, 547, 50

Kochanek, C. S., Kolatt, T. S., & Bartelmann, M. 1996, ApJ, 473, 610

Kochanek, C. S. & Narayan, R. 1992, ApJ, 401, 461

Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18

Koopmans, L. V. E., Treu, T., Bolton, A. S., Burles, S., & Moustakas, L. A. 2006, ApJ, 649, 599

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

Larson, D., Dunkley, J., Hinshaw, G., et al. 2011, ApJS, 192, 16

Lawrence, C. R., Elston, R., Januzzi, B. T., & Turner, E. L. 1995, AJ, 110, 2570

Madau, P., Diemand, J., & Kuhlen, M. 2008, ApJ, 679, 1260

Mao, S., Jing, Y., Ostriker, J. P., & Weller, J. 2004, ApJ, 604, L5

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

McKean, J. P., Auger, M. W., Koopmans, L. V. E., et al. 2010, MNRAS, 404, 749

McKean, J. P., Browne, I. W. A., Jackson, N. J., et al. 2005, MNRAS, 356, 1009

McKean, J. P., Impellizzeri, C. M. V., Roy, A. L., et al. 2011, MNRAS, 410, 2506

McKean, J. P., Koopmans, L. V. E., Flack, C. E., et al. 2007, MNRAS, 378, 109

Metcalf, R. B. 2002, ApJ, 580, 696

Michałowski, M., Hjorth, J., & Watson, D. 2010, A&A, 514, A67+

Michelson, A. A. 1920, ApJ, 51, 257

Michelson, A. A. & Pease, F. G. 1921, ApJ, 53, 249

Minezaki, T., Chiba, M., Kashikawa, N., Inoue, K. T., & Kataza, H. 2009, ApJ, 697, 610

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

More, A., McKean, J. P., More, S., et al. 2009, MNRAS, 394, 174

More, A., McKean, J. P., Muxlow, T. W. B., et al. 2008, MNRAS, 384, 1701

Myers, S. T., Jackson, N. J., Browne, I. W. A., et al. 2003, MNRAS, 341, 1

Narayan, R. & Bartelmann, M. 1996, ArXiv Astrophysics e-prints

Narayan, R. & Nityananda, R. 1984, in Indirect Imaging. Measurement and Processing for Indirect Imaging, ed. J. A. Roberts, 281–+

Natarajan, P., Kneib, J.-P., Smail, I., et al. 2009, ApJ, 693, 970

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

Negrello, M., Hopwood, R., De Zotti, G., et al. 2010, Science, 330, 800

Nesvadba, N. P. H., Lehnert, M. D., Eisenhauer, F., et al. 2006, ApJ, 650, 661

Oguri, M. 2006, MNRAS, 367, 1241

Oguri, M., Inada, N., Strauss, M. A., et al. 2008, AJ, 135, 512

Oguri, M., Keeton, C. R., & Dalal, N. 2005, MNRAS, 364, 1451

Padmanabhan, N., Schlegel, D. J., Seljak, U., et al. 2007, MNRAS, 378, 852

Patnaik, A. R. & Porcas, R. W. 1996, in IAU Symposium, Vol. 173, Astrophysical Applications of Gravitational Lensing, ed. C. S. Kochanek & J. N. Hewitt, 305–+

Pearson, T. J. & Readhead, A. C. S. 1984, ARA&A, 22, 97

Pettini, M., Steidel, C. C., Adelberger, K. L., Dickinson, M., & Giavalisco, M. 2000, ApJ, 528, 96

Porcas, R. W. 1998, in Astronomical Society of the Pacific Conference Series, Vol. 144, IAU Colloq. 164: Radio Emission from Galactic and Extragalactic Compact Sources, ed. J. A. Zensus, G. B. Taylor, & J. M. Wrobel, 303–+

Refsdal, S. 1992, in Lecture Notes in Physics, Berlin Springer Verlag, Vol. 406, Gravitational Lenses, ed. R. Kayser, T. Schramm, & L. Nieser, 61–+

Ros, E., Guirado, J. C., Marcaide, J. M., et al. 2000, A&A, 362, 845

Rubin, V. C., Thonnard, N., & Ford, Jr., W. K. 1978, ApJ, 225, L107

Ruff, A. J., Gavazzi, R., Marshall, P. J., et al. 2011, ApJ, 727, 96

Rusin, D., Norbury, M., Biggs, A. D., et al. 2002, MNRAS, 330, 205

Saha, P. & Williams, L. L. R. 2004, AJ, 127, 2604

Salpeter, E. E. 1955, ApJ, 121, 161

Sargent, M. T., Schinnerer, E., Murphy, E., et al. 2010, ApJS, 186, 341

Schechter, P. L. & Moore, C. B. 1993, AJ, 105, 1

Schmidt, M. & Green, R. F. 1986, ApJ, 305, 68

Schneider, P. 1984, A&A, 140, 119

Schneider, P. 1985, A&A, 143, 413

Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses, ed. Schneider, P., Ehlers, J., & Falco, E. E.

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

Seitz, S., Saglia, R. P., Bender, R., et al. 1998, MNRAS, 298, 945

Shapley, A. E., Steidel, C. C., Adelberger, K. L., et al. 2001, ApJ, 562, 95

Shapley, A. E., Steidel, C. C., Erb, D. K., et al. 2005, ApJ, 626, 698

Siana, B., Smail, I., Swinbank, A. M., et al. 2009, ApJ, 698, 1273

Siana, B., Teplitz, H. I., Chary, R.-R., Colbert, J., & Frayer, D. T. 2008, ApJ, 689, 59

Smail, I., Swinbank, A. M., Richard, J., et al. 2007, ApJ, 654, L33

Springel, V., Wang, J., Vogelsberger, M., et al. 2008, MNRAS, 391, 1685

Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, ApJ, 519, 1

Steidel, C. C., Giavalisco, M., Dickinson, M., & Adelberger, K. L. 1996a, AJ, 112, 352

Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996b, ApJ, 462, L17+

Suyu, S. H., Marshall, P. J., Auger, M. W., et al. 2010, ApJ, 711, 201

Suyu, S. H., Marshall, P. J., Blandford, R. D., et al. 2009, ApJ, 691, 277

Thanjavur, K., Crampton, D., & Willis, J. 2010, ApJ, 714, 1355

Tonry, J. L. & Kochanek, C. S. 1999, AJ, 117, 2034

Treu, T., Koopmans, L. V., Bolton, A. S., Burles, S., & Moustakas, L. A. 2006, ApJ, 640, 662

Trotter, C. S., Winn, J. N., & Hewitt, J. N. 2000, ApJ, 535, 671

Tytler, D., O'Meara, J. M., Suzuki, N., & Lubin, D. 2000, Physica Scripta Volume T, 85, 12

Van der Kruit, P. C. 1971, A&A, 15, 110

Vegetti, S., Koopmans, L. V. E., Bolton, A., Treu, T., & Gavazzi, R. 2010, MNRAS, 408, 1969

Verdugo, T., Motta, V., Muñoz, R. P., et al. 2010, ArXiv e-prints

Volders, L. M. J. S. 1959, Bull. Astron. Inst. Netherlands, 14, 323

Volino, F., Wucknitz, O., McKean, J. P., & Garrett, M. A. 2010, A&A, 524, A79+

Walsh, D., Carswell, R. F., & Weymann, R. J. 1979, Nature, 279, 381

Wucknitz, O. 2004, MNRAS, 349, 1

Wucknitz, O., Biggs, A. D., & Browne, I. W. A. 2004, MNRAS, 349, 14

Yee, H. K. C., Ellingson, E., Bechtold, J., Carlberg, R. G., & Cuillandre, J. 1996, AJ, 111, 1783

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579

Yun, M. S., Reddy, N. A., & Condon, J. J. 2001, ApJ, 554, 803

Zhai, Z.-X., Wan, H.-Y., & Zhang, T.-J. 2010, Physics Letters B, 689, 8

Zhang, M., Jackson, N., Porcas, R. W., & Browne, I. W. A. 2007, MNRAS, 377, 1623

Zucker, D. B., Belokurov, V., Evans, N. W., et al. 2006a, ApJ, 650, L41

Zucker, D. B., Belokurov, V., Evans, N. W., et al. 2006b, ApJ, 643, L103