A Tale of Two Wide Separation Gravitational Lenses

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Anupreeta More

aus

Mumbai (India)

Bonn, 2008

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

1. Referent: Prof. Dr. Peter Schneider
2. Referent: Prof. Dr. Johannes Schmid-Burgk

Tag der Mündlichen Prüfung: 05 June 2008

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn, http://hss.ulb.uni-bonn.de/diss_online, elektronisch publiziert.

Contents

Constants and Quantities used throughout

Constants/Quantities	Values
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
Jansky:	Jy= 10^{-26} Watts m^{-2} Hz^{-1}
parsec:	pc= $3.08568025 \times 10^{16}$ m
Hubble constant:	$H_0 = 70$ km s^{-1} Mpc^{-1}
Energy density parameter of Matter:	$\Omega_m = 0.3$
Energy density parameter of Dark energy:	$\Omega_{\Lambda} = 0.7$

Overview

Two wide separation gravitational lens systems are investigated in this thesis. The wide separation refers to the angular separation of the gravitationally lensed images of a distant source. Since the mass of an intervening lensing object (a single or group of galaxies) which deflects light rays from the source is proportional to the square of the image separation, the wide image separation suggests that the lensing object is massive in both systems (See Chapter 1). The main aim of this thesis is to study these massive lensing objects in order to deduce the properties of the dark matter halos and possibly the background source. Beginning with a general introduction to the current understanding of cosmological structure formation and the distribution of mass in Chapter 1, an emphasis on Gravitational Lensing and its growing applications is given. Furthermore, a review of some important recent results from the fields is given which drives the motivation for the research presented in this thesis. Subsequent to that, a theoretical background of gravitational lensing, adequate for the main research carried out here, is formulated.

To continue further with the tools in lensing, in Chapter 2 attention is given to the parameterised mass modeling aspect of lensing in order to familiarize the reader with these concepts and techniques, which are used in the latter half of this thesis for both of the lens systems. An analytical framework of some of the standard models used for mass modeling is laid down. The main degeneracies in lensing that have to be confronted with, are also discussed.

The data that provide constraints to the mass models were obtained from radio interferometric observations of each lens system. In Chapter 3, the technique of Very Long Baseline Interferometry (VLBI) is described. Since the observations are performed in the *phase referencing* mode for both of the lens systems, the emphasis is on this aspect of the interferometric observations. Similarly, some issues like *smearing* of features in the radio images specific to these lens systems, are discussed.

In Chapter 4, new observations and mass models for the gravitational lens system MG 2016+112 are presented. This system has four lensed images in the radio with a maximum image separation of ~ 3.6 arcsec. An account of previous work including the results of up-to-date multi-frequency observations and several mass models is sketched. This system has the highest redshift lensing galaxy ($z = 1.01$) and a high redshift background quasar ($z = 3.27$). The lensing galaxy of MG 2016+112 is found to have luminous substructure in the form of a satellite galaxy. As a consequence of the satellite galaxy being in the proximity of the lensing galaxy, the astrometry of the lensed images of MG 2016+112 is found to be distorted.

Observations of satellite galaxies in the local Universe are found to be discordant with the Cold Dark Matter (CDM) predictions (e.g., Kauffmann et al. 1993). An estimate of the substructure mass fraction of the total mass of the main halo can be used to test predic-

tions of CDM cosmology (e.g., Dalal & Kochanek 2002). More importantly, unlike any other lens system, the satellite galaxy in MG 2016+112 is spectroscopically confirmed to be at the same redshift as the main lensing galaxy. Furthermore, the mass model gives an accurate mass estimate for the early-type lensing galaxy. This, combined with dynamical studies for several lensing galaxies can be studied statistically to learn about galaxy formation and evolution (e.g., Koopmans et al. 2006b). Moreover, the high magnification (a factor of $\gtrsim$100) of a pair of merging lensed images in MG 2016+112 allows a study of the properties of a high redshift quasar with unprecedented resolution and sensitivity.

The new radio observations carried out at three frequencies and different epochs using VLBI are presented. These observations have detected new lensed radio components. With the help of new and better observational constraints, the previous best working mass model by Koopmans et al. (2002b) and a number of new models, are tested. The new observational data are inconsistent with the model predictions of Koopmans et al. (2002b) and their model needs to be modified. The spectral analysis from the new observations is presented and the (non-) detection of substructure near the non-merging lensed images is discussed.

In Chapter 5, the focus is on CLASS B2108+213, the largest image separation lens system ($\sim$ 4.6 arcsec) discovered in the Cosmic Lens All Sky Survey (Browne et al. 2003). B2108+213 is doubly imaged by a massive early-type lensing galaxy (z = 0.365), which is found to belong to a group of galaxies. The mass distributions at galaxy-group scales are not well-studied as compared to those at galaxy or cluster scales, since lens systems with intermediate image separations, such as B2108+213, are few in number (Oguri et al. 2005; Oguri 2006). Furthermore, the nature of a third component in the radio situated close to the optical position of lensing galaxy is ambiguous which could either be an odd lensed image which are extremely rare to find due to the low magnification (Narasimha et al. 1986) or an Active Galactic Nucleus (AGN) of the lensing galaxy.

The previous observational and mass modeling results for B2108+213 are summarized in the beginning of Chapter 5. New VLBI radio observations at 1.7 and 5 GHz are made at different epochs and new MERLIN 5 GHz observations are presented. Subsequently, the fine structure found in the lensed images are used to test various parameterised mass models and to constrain the slope of the density profile of the dark matter halo. Furthermore, a spectral analysis of the radio components and the results of new mass models are discussed, which confirm the status of the third radio component as the AGN of the lensing galaxy in B2108+213. The background quasar in B2108+213 is a BL Lac type quasar with a featureless spectrum and the determination of its redshift has been elusive. Using the results of the new mass models presented here and from the new observations presented in McKean et al. (in prep), the redshift of the quasar was estimated. The conclusions from the new findings and mass modeling analysis are presented.

Lastly, the research presented here is summarized and a short description of the possible future work is given in Chapter 6. The relevant concepts and methods required in this thesis, are expounded in the Appendix.

1 Introduction

1.1 Background on Cosmology

A standard cosmological model has emerged over the last few decades based on the theoretical framework provided by the general theory of relativity. The Universe started in a very hot dense phase and cooled off as it expanded. Small inhomogeneities were set in the dense soup by the fluctuations of a quantum field. As time proceeded, these small fluctuations were amplified by *gravitational instabilities* and led to the formation of large scale structures in the Universe.

The above qualitative picture was given its present shape and a quantitative basis by a variety of observations. Hubble's discovery of the expansion of the Universe convinced cosmologists to abandon the idea of a static Universe. The expanding Universe naturally leads to the idea of a hot and dense initial phase. In addition, it predicts the existence of a relic background of photons that should be observable today. The observations of this relic background (e.g., Smoot et al. 1992) called the Cosmic Microwave Background (CMB), at the expected temperature (~ 3 K), gave a great deal of confidence in the belief of an expanding Universe.

A flat geometry of the Universe is possible only if the energy density of the Universe is equal to a critical value (see Appendix A). The energy density inferred from the light in stars alone was not enough to equal the critical density. The first observational evidence of this came from the study of cluster dynamics (Zwicky 1937). The observations of rotation curves of galaxies (e.g., Rubin et al. 1985) also indicated the presence of non-baryonic matter which was termed *dark matter*. But even after including the dark matter component, the energy density of the Universe was found not to be enough to match the critical density (Bahcall et al. 2000). Large redshift surveys like the Sloan Digital Sky Survey (SDSS) and the Two Degree Field Galaxy Redshift Survey (2dFGRS) enabled the statistical study of large scale structure (e.g., Norberg et al. 2002; Blanton et al. 2003). The galaxy clustering is particularly sensitive to the matter content in the Universe.

Large scale structure studies have consistently shown that the matter density is less than the critical density of the Universe. The missing part of the puzzle was provided by observations of distant supernovae (Perlmutter et al. 1999). These observations indicated that the Universe is undergoing an accelerated expansion at present times. This accelerated expansion can be explained by an energy density component which acts like a fluid with negative pressure. It is called *dark energy*[1] and compensates for the missing budget of energy density required for a flat Universe.

[1] In the theory of general relativity, the dark energy component is included in the form a constant called *cosmological constant* (Λ).

A flat Universe is not just a theorists dream but has observational support. The power spectrum of the inhomogeneities in the CMB provide constraints on the geometry of the Universe. The Wilkinson Microwave Anisotropy Probe (WMAP) experiment has studied the inhomogeneities in great detail and the results indicate that the geometry of the Universe to be flat (Spergel et al. 2003). The combined constraints established by precision observations of the CMB, large scale structure and distant supernovae have all contributed to the current concordance model called Lambda Cold Dark Matter (ΛCDM) (e.g., Dunkley et al. 2008). According to the ΛCDM model the current energy density of the Universe is dominated by dark energy (74%) and the remainder is dominated by matter. However, out of the total matter content, baryonic matter constitutes a small fraction of 4% whereas most of the matter is dark (22%). Although the ΛCDM model has enjoyed a tremendous success with observations, the main ingredients of this model, dark matter and dark energy, are as their names suggest, dark. They have been only inferred from astronomical observations and indirect evidences. The present studies in cosmology have revolved around trying to understand the nature of these two components.

1.1.1 Structure Formation

The small density perturbations present in the Universe at very early times can be seen as the inhomogeneities in the CMB. The evolution of these density perturbations can be studied using linear perturbation theory until the perturbation grows and the evolution becomes nonlinear. The growth of these perturbations depends upon the total matter density in the Universe and the nature of dark matter.

Structure formation in the universe can be studied by considering the abundance and the internal structure of dark matter haloes. A halo is a region where dark matter has collapsed and become gravitationally bound. The abundance of halos is quantified in terms of the halo mass function. The *halo mass function* is the number density of halos in a unit comoving volume[2] with masses in the range m and $m + dm$. The internal structure of each halo (i.e. the *halo profile*) is described as below.

Halo profile

Numerical simulations show that the dark matter density in halos of all masses follows a universal profile (see Fig. 1.1) which can be described by

$$\rho(r) \propto \left(\frac{r}{r_s}\right)^{-1} \left(1 + \frac{r}{r_s}\right)^{-2} \tag{1.1}$$

where the scale radius (r_s) is a function of the halo mass (Navarro, Frenk & White 1996, 1997). The inner part of the density distribution goes as r^{-1} and hence is cuspy. In the outermost part it falls of as r^{-3} and resembles an isothermal (r^{-2}) profile in the intermediate range. This profile is referred to as the Navarro-Frenk-White (NFW) profile. Although the profile of the halos is found to be cuspy in the inner regions from the ΛCDM simulations (e.g., Diemand et al. 2005), observations of the rotation curves in disk galaxies (e.g., Salucci & Burkert 2000) suggest a constant-density core in the inner regions.

[2] A volume defined in terms of the comoving co-ordinates (see Appendix A).

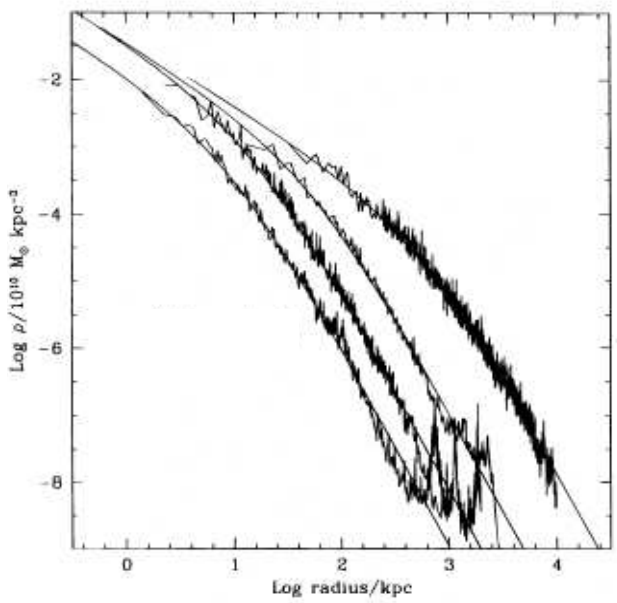

Figure 1.1: The density distribution of halos with masses ranging up to four orders of magnitude. The simulated density distributions of halos of different masses are found to fit well to a single profile given by NFW (Navarro, Frenk & White 1996).

Furthermore, halos are found to have not only ellipsoidal but triaxial mass distributions (e.g., Jing & Suto 2002). However, numerical simulations suggest that the baryons in the central regions of the halos (i.e. the stellar and gas components) lead to an isotropic and isothermal mass distribution thereby, making the halos nearly spherical in the centre (e.g., Kazantzidis et al. 2004).

Abundance of sub-halos

Numerical simulations of Milky Way-sized halos based on a ΛCDM cosmology suggest an abundant sub-halo population (see the right panel of Fig. 1.2), whereas the observed population of satellite galaxies in the Local Group is less by an order of magnitude (e.g., Moore et al. 1999). The left panel in the Fig. 1.2 is a cumulative plot of the satellite population for the Milky Way and the sub-halo population for a dark-matter-only simulation of a galaxy and a cluster scale halo. Although the level of substructure predicted in clusters is close to the observations, for galaxies there is a clear under-abundance observed. This discrepancy is referred to as the *missing satellite problem*. One of the proposed solutions to this problem is to prevent star-formation in the sub-halos of galaxies thereby making them dark (e.g., Klypin et al. 1999). On the other hand, recently, with the help of the SDSS star catalogue several new satellites have been discovered (e.g., Belokurov et al. 2007; Martin et al. 2007; Simon & Geha 2007). This has mitigated the discrepancy of substantially low observed population of the satellites. Furthermore efforts have been made to more accurately characterize the luminosity function of the Milky Way satellites by taking into account the selection effects(Koposov et al. 2007). The galaxy formation processes are suspected to prevent galaxy formation in the less massive sub-halos and is

under investigation (Koposov et al. in prep.).

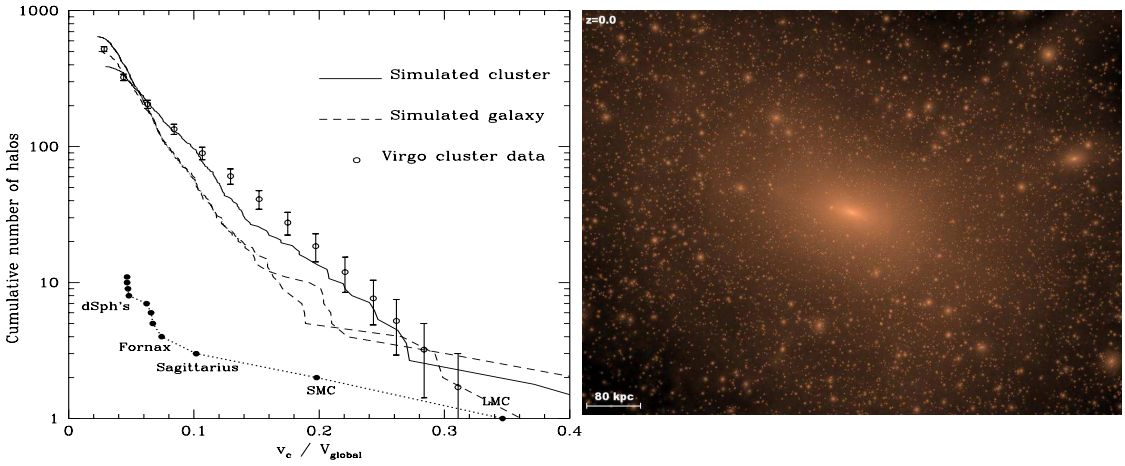

Figure 1.2: *Left*: Abundance of the sub-halo population as a function of their circular velocities normalized to the circular velocities of the host halo. The dotted line and the open circles are the observations of the Milky Way galaxy and Virgo cluster which are compared with the simulations of a typical galaxy and cluster, respectively (Moore et al. 1999). *Right*: A simulated population of *dark* matter sub-halos around a Milky Way-like galaxy (Diemand et al. 2007) most of which cannot be detected even for the most nearby galaxy−Milky Way.

1.2 Theory of Gravitational Lensing

Gravitational lensing is a phenomenon in which the light rays emanating from a distant background source are deflected, due to a gravitational potential along the line-of-sight, before reaching the observer. If the background source has a large projected separation from the line-of-sight to the lens potential, then the effect of the lens potential is weak and a single distorted image is formed. This regime is called *weak lensing* and in principle, it is observed for photons coming from sources in all directions. On the other hand, if the potential is strong and the projected position of the background source[3] is close to the centre of the lens potential, multiple images of the background source are formed. This regime is called *strong lensing*. Since gravitational lensing is sensitive to both luminous and dark matter, it is an excellent tool to study the properties of dark matter halos.

In this section, the theoretical framework of gravitational lensing is laid down to familiarize the reader with various definitions and terms that will be used throughout this thesis (for details, see e.g., Schneider et al. 1992; Narayan & Bartelmann 1996).

1.2.1 Deflection Angle and Lens Equation

Similar to the light rays being bent as they pass through a prism, light rays are bent in the presence of a gravitational field. The angle through which the rays emitted from a distant

[3]Note that the undeflected position of the background source is never known since it can not be observed.

source are deflected due to the presence of intervening mass in the line-of-sight is called the *deflection angle* and it can be written as

$$\hat{\boldsymbol{\alpha}} = -\int \bar{\nabla}_{\perp} n \, dl = \frac{2}{c^2} \int \bar{\nabla}_{\perp} \Phi \, dl \,, \tag{1.2}$$

where n is effectively the refractive index of the intervening mass. It is related to the gravitational potential by $n = 1 - 2|\Phi|/c^2$ (Schneider et al. 1992) and is based on the assumption that the Newtonian potential is small (i.e. $\Phi \ll c^2$) and also, the peculiar velocity of the mass is small (i.e. $v \ll c$).

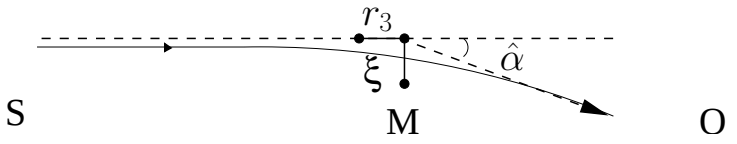

Figure 1.3: Light rays emitted from a source S deflected by a point mass M and observed by an observer O. The deflection takes place through an angle $\hat{\alpha}$.

For a point mass, the Newtonian potential is given by $\Phi = -GM/\sqrt{\xi^2 + r_3^2}$ where ξ is the physical separation perpendicular to the undeflected light ray and r_3 is a component of the physical separation of the order of ξ in the line-of-sight (see Fig. 1.3). ξ marks the point of closest approach and is called the *impact parameter*. Substituting the potential of a point mass in Eq. 1.2 gives

$$\hat{\alpha} = \frac{4GM}{c^2 \xi} \,. \tag{1.3}$$

This can be conveniently expressed in units of the Schwarzschild radius ($R_s = 2GM/c^2$) of a point mass as $\hat{\alpha} = 2R_s/\xi$.

To calculate the deflection angle of rays emitted by distant galaxies or quasars due to the intervening mass distributions of galaxies and/or clusters which are *extended* in reality, there are some assumptions involved. The deflection is assumed to occur due to the mass lying in a thin plane called the lens plane. The distances of the background source and the deflector (also referred to as lens) from the observer are very large compared to the spatial extent of the lens plane. Therefore, this thin lens approximation is reasonable.

Thus, the deflection angle can be quantified by integrating the line-of-sight density for each mass element and summing the deflections due to the mass elements at every ξ in the plane of the sky.

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \int \frac{4G\Sigma(\boldsymbol{\xi}')}{c^2} \frac{\boldsymbol{\xi} - \boldsymbol{\xi}'}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|^2} \, d^2\xi' \,. \tag{1.4}$$

where

$$\Sigma(\boldsymbol{\xi}') = \int \rho(\boldsymbol{\xi}', r_3) \, dr_3$$

is the surface mass density at ξ'.

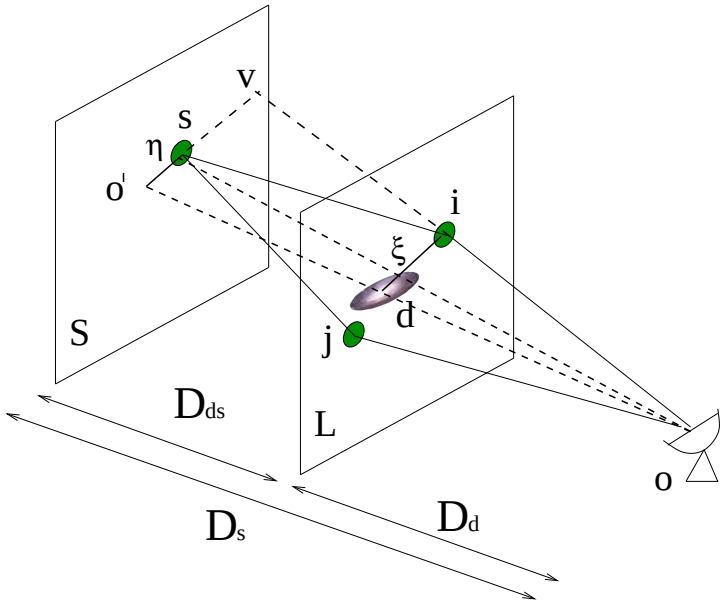

Figure 1.4: The ray-trace diagram showing the background source (in the source plane S) is deflected by a galaxy d (in the lens plane L) such that multiple images i and j are formed and seen by the observer at o.

Lens equation

Consider an observer at o, the lens d (e.g., a galaxy) in the lens plane L and a source s (e.g., a quasar) in the source plane S as shown in Fig. 1.4. Let odo′ be the optical axis connecting the three planes. Let ∠o′os = $\boldsymbol{\beta}$, ∠sov = $\boldsymbol{\alpha}$, ∠o′ov = $\boldsymbol{\theta}$ and ∠siv = $\hat{\boldsymbol{\alpha}}$ then it can be seen that

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}). \tag{1.5}$$

This ray-trace equation is called the Lens equation where

$$\boldsymbol{\beta} = \boldsymbol{\eta}/D_s \; ; \quad \boldsymbol{\theta} = \boldsymbol{\xi}/D_d, \tag{1.6}$$

and the scaled deflection angle α is simply related to the deflection angle ($\hat{\alpha}$) by

$$\boldsymbol{\alpha} = \frac{D_{ds}}{D_s}\hat{\boldsymbol{\alpha}}. \tag{1.7}$$

Here, D_s, D_d, and D_{ds} are the angular diameter distances (see Appendix A) between the observer and the source, the observer and the deflector, and the deflector and the source, respectively. Substituting Eq. 1.6 and 1.7 in Eq. 1.5 gives

$$\boldsymbol{\eta} = \frac{D_s}{D_d}\boldsymbol{\xi} - D_{ds}\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}), \tag{1.8}$$

where η is the projected source position. Now, let $\Sigma_{\rm crit}$ be a *critical surface mass density* defined such that

$$\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{D_s}{D_d D_{ds}}. \tag{1.9}$$

The scaled surface mass density is called the *convergence* and is therefore given by,

$$\kappa = \frac{\Sigma}{\Sigma_{crit}} . \tag{1.10}$$

Substituting Eq. 1.4 and 1.7 in the lens equation given by Eq. 1.8 and equating it to Eq. 1.5 gives the scaled deflection angle as

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi} \int \kappa(\boldsymbol{\theta}') \frac{\boldsymbol{\theta} - \boldsymbol{\theta}'}{|\boldsymbol{\theta} - \boldsymbol{\theta}'|^2} \, d^2\theta' . \tag{1.11}$$

Deflection and Fermat potential

The deflection angle in Eq. 1.11 can be written as a gradient of a function ψ using $\nabla \ln|\boldsymbol{\theta}| = \boldsymbol{\theta}/|\boldsymbol{\theta}^2|$. Hence,

$$\alpha = \nabla\psi . \tag{1.12}$$

Comparing Eq. 1.12 with Eq. 1.11 gives

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi} \int \kappa(\boldsymbol{\theta}') \ln|\boldsymbol{\theta} - \boldsymbol{\theta}'| \, d^2\,\theta' . \tag{1.13}$$

The Laplacian of Eq. 1.13 gives $\nabla^2\,\psi = 2\,\kappa$, which is a Poisson equation. Therefore, ψ is called the *deflection potential*.

The lens equation Eq. 1.5 can be expressed using Eq. 1.12 as

$$\boldsymbol{\beta} = \nabla \left(\frac{1}{2}\boldsymbol{\theta}^2 - \psi(\boldsymbol{\theta}) \right), \tag{1.14}$$

which can be rearranged to give

$$\nabla \left(\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\beta})^2 - \psi(\boldsymbol{\theta}) \right) = 0 . \tag{1.15}$$

The above equation can be written as $\nabla\,\phi(\boldsymbol{\theta}, \boldsymbol{\beta}) = 0$ where ϕ is a scalar function,

$$\phi\,(\boldsymbol{\theta}, \boldsymbol{\beta}) = \left(\frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\beta})^2 - \psi(\boldsymbol{\theta}) \right). \tag{1.16}$$

The potential ϕ is associated with the travel-time of light rays and is known as the *Fermat potential* (e.g., Schneider 1985; Blandford & Narayan 1986). The stationary points of the Fermat potential are satisfied by the solutions of the lens equation which is in accord with Fermat's principle (also, see section 1.2.3).

Einstein radius

The deflection angle for a circularly symmetric mass distribution, using Eq. 1.3, is

$$\hat{\alpha} = \frac{4GM(\xi)}{c^2\xi} \tag{1.17}$$

where $M(\xi)$ corresponds to the mass within the radius ξ. Therefore, the lens equation for such a distribution is given by

$$\beta = \theta - \frac{4GM(\theta)}{c^2\theta}\frac{D_{ds}}{D_d D_s}. \tag{1.18}$$

Consider the case when the source lies on the optical axis joining the observer and the lens. Then substituting $\beta = 0$ in Eq. 1.18 gives

$$\theta_E = \left(\frac{4GM(\theta_E)}{c^2}\frac{D_{ds}}{D_d D_s}\right)^{1/2}. \tag{1.19}$$

Therefore, owing to the symmetric mass distribution and aligned positions, the source is lensed as a ring with an angular radius θ_E which is termed the *Einstein radius* and the ring is called the *Einstein ring*. Generally, the image separation is approximately twice the Einstein radius. A sufficient condition to form multiple images is $\kappa \geq 1$ which is satisfied if the source position β is within the Einstein ring.

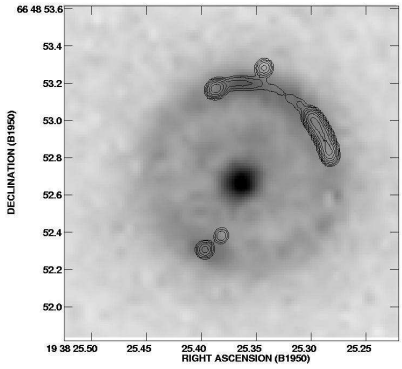

Figure 1.5: *Hubble Space Telescope (HST)* optical image of gravitational lens system B1938+666 with an Einstein ring and the lensing galaxy in the centre. An arc joining three images and a fourth image situated diametrically opposite is shown with the overlaid contours using Multi-Element radio linked Interferometry (MERLIN) in the radio (King et al. 1998).

Since the Einstein radius is an observable which is simply related to the mass enclosed within the Einstein ring, such rings provide excellent constraints to accurately measure the enclosed mass of galaxies or clusters. Thus, the wider the image separation, the more massive is the lens. From over a hundred known lens systems, several lens systems are found to have fairly complete rings or highly elongated arcs in the optical (e.g., Bolton et al. 2006a, also, see Fig 1.5). Thus, the mass of galaxies and clusters can be estimated independent of other studies like kinematics.

1.2.2 Magnification, Shape and Parity of Images

Let $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ be the positions of a source and a lensed image. Let the solid angle subtended by the source be ω_o in the absence of the lens. In the presence of a lens, due to the

deflection of light rays, the solid angle of the source in the sky is modified to the solid angle ω (which is seen as the image). This changes the flux of the image ($S = I \times \omega$) too, since the intensity I remains constant. The magnification is the ratio of the solid angle of the image to that of the source,

$$\mu = \frac{S}{S_o} = \frac{\delta\omega}{\delta\omega_o}\,. \tag{1.20}$$

Here, the shape of the image is determined by the Jacobian matrix $M(\boldsymbol{\theta}) = \frac{\partial\boldsymbol{\beta}}{\partial\boldsymbol{\theta}}$. The elements of the inverse magnification matrix M are given by using using Eq. 1.14 and 1.16,

$$M_{ij} = \phi_{ij} = (\delta_{ij} - \psi_{ij}) = \begin{pmatrix} 1-\kappa-\gamma_{s1} & -\gamma_{s2} \\ -\gamma_{s2} & 1-\kappa+\gamma_{s1} \end{pmatrix} \tag{1.21}$$

where the subscripts of ϕ and ψ imply partial derivatives of the quantities with respect to θ_i and θ_j. The convergence κ contributes to isotropic magnification whereas the shear $\gamma_s = \sqrt{\gamma_{s1}^2 + \gamma_{s2}^2}$ causes anisotropic magnification of the images. Thus, the magnification is written as

$$\mu = \frac{1}{|\det \boldsymbol{M}|} = \frac{1}{(1-\kappa)^2 - \gamma_s^2}\,. \tag{1.22}$$

The trace $\mathrm{tr}(M) = 2(1-\kappa)$ and the eigen values of the magnification matrix are $\lambda_{1,2} = 1-\kappa\mp\gamma_s$ which give the factor of stretching of the images along the direction of the eigen vectors.

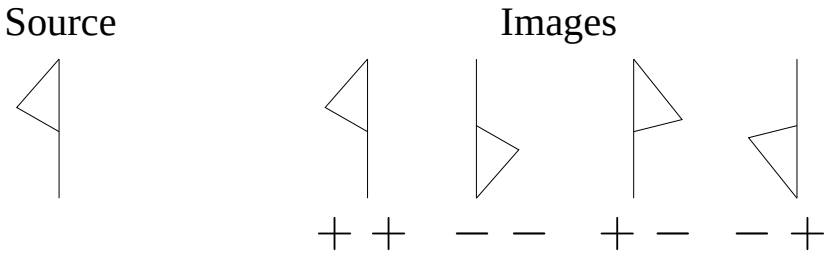

Figure 1.6: The source and the images with different parities. The '+' and '−' refer to the signs of the eigen values of the transformation matrix that maps the source to the images. The sign of the product of the eigen values determines the parities of the images.

The parity of an image is the handedness of a feature in the image with respect to that in the background source (e.g., jet components). This can be understood from the magnification matrix which is a mapping between the source and the images. If both the eigen values of the magnification matrix have the same sign (i.e. either both '+' or '-'), then the parity of the image is positive (see Fig. 1.6). If the eigen values have different signs, the parity of the image is the opposite of that of the source.

1.2.3 Time Delay of Images

The light rays emitted by the source before arriving at the position of the observer suffer two effects. The light rays take geometrically longer paths than the undeflected light rays

would take in the absence of a lens. Additionally, the light rays are delayed when traveling through the gravitational potential of the lens which is also known as the *Shapiro delay*. Therefore, the *total time delay* is a combination of these two effects and is given by

$$\tau(\boldsymbol{\theta}) = \frac{(1+z_d)}{c}\frac{D_d D_s}{D_{ds}}\left[\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\beta})^2 - \psi(\boldsymbol{\theta})\right], \tag{1.23}$$

The time delay of an image however, cannot be known since the travel-time of the rays in the absence of the lens cannot be measured to use as a reference. The only observable is the relative time delay of any two images of a source which can be measured in the event of variability in the source. For images at $\boldsymbol{\theta}_i$ and θ_j, the time delay between the two images is,

$$\Delta\tau_{ij} = \frac{(1+z_d)}{c}\frac{D_d D_s}{D_{ds}}\left[\phi\left(\boldsymbol{\theta}_i,\boldsymbol{\beta}\right) - \phi\left(\boldsymbol{\theta}_j,\boldsymbol{\beta}\right)\right]. \tag{1.24}$$

Since the factor of the angular diameter distances is,

$$\frac{D_d D_s}{D_{ds}} \propto \frac{1}{H_0},$$

the time delay $\Delta\tau \propto H_0^{-1}$ (See Appendix A). Thus, provided the redshifts of the lens and the background source, and the mass distribution are known, the time delay measurements can constrain the value of the Hubble constant. Such an estimate of the Hubble constant is independent of other methods which involve several uncertainties in the numerous steps, for example, distance ladder methods which use standard-candle stars from low-redshift to calibrate the distances to those in high-redshift galaxies. Although due to the uncertainties in the mass models and relative time delay measurements, the desired accuracy in the measurement of the Hubble constant using individual lens systems is not achieved yet (with the exception of a few lens systems, Koopmans et al. 2003; York et al. 2005). Nevertheless, a significant improvement in the statistical uncertainties is made by combining measurements from several lens systems (e.g., Saha et al. 2006; Oguri 2007). A recent estimate of the Hubble constant suggest a value of 71^{+6}_{-8} km s^{-1} Mpc^{-1} (Coles 2008).

1.2.4 Ordinary (non-critical) Image Properties

Given a source position ($\boldsymbol{\beta}$), a two-dimensional arrival-time surface is defined by the Fermat potential. The images that form at the extrema and the saddle points of the arrival-time surface (i.e. stationary points) are called *ordinary images*. Consider a single and thin lens plane consisting of a lens with a smooth density distribution. Let the density distribution (ρ) drop faster than r^{-3} for $r \to \infty$ to ensure finite enclosed mass. Based on these assumptions and for a source position other than that on a caustic (see section 1.2.5), the *odd number theorem* states that the total number of ordinary images is finite and odd. However, in case of any discontinuities or singularities in the mass distribution, this theorem does not hold. The *magnification theorem* states that provided $\kappa \geq 0$ there is always at least one image with positive parity and with a magnification $\mu \geq 1$. Three types of images can be formed.

1. For $det\ M > 0$, $\mathrm{tr}(M) > 0$, the image is formed at the minimum (L) of the arrival-time surface (see Fig. 1.7 and 1.8). This is called the Type I image. At least one

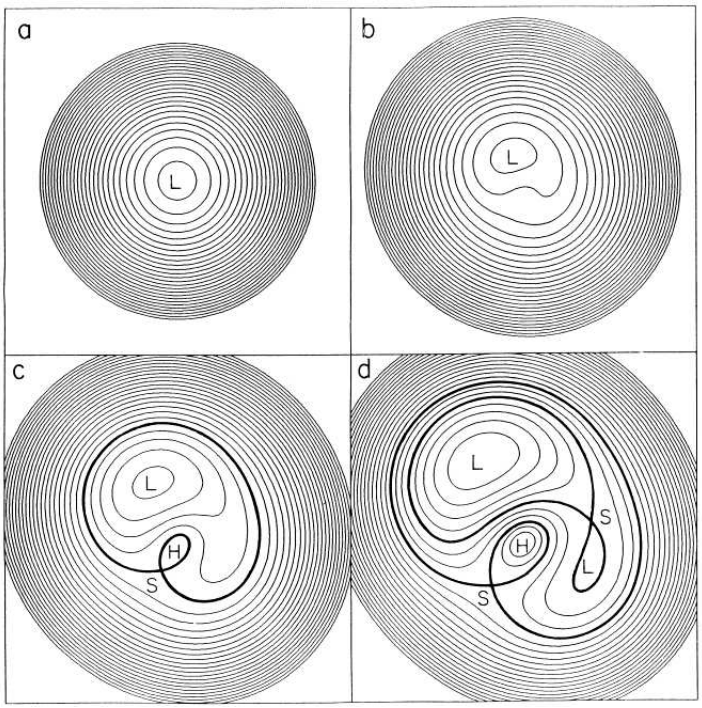

Figure 1.7: Types of lensed images formed at stationary points of the time-arrival surface for a fixed source position and an elliptical lens potential (Blandford & Narayan 1986). The panels show the effects of strengthening the lens potential, starting with no potential (panel a), a weak potential (panel b); and panels c and d show a strong enough potential to produce three and five images, respectively.

type I image exists irrespective of the presence of a lens from the magnification theorem.

2. For $det\ M < 0$, the image is formed at the saddle point (S) and is called the Type II image. It is known from the odd number theorem that when a type II image exists, it is imperative that such a lens system has multiple images.

3. For $det\ M > 0$, $\mathrm{tr}(M) < 0$, $\kappa > 1$, the image is formed at the maximum (H) and is called the Type III image. If the type III image exists, it is located closer to the centre of the potential and hence, rarely observed because of both obscuration and low magnification.

In the event of variability in the source, the Type I image is the first image to vary since the travel-time is minimum for this image, the next image to vary is the Type II image, and the Type III image is the last to vary. This forms a qualitative constraint in lens systems which a correct mass model must satisfy.

1.2.5 Critical Image Properties

For $det\ A = 0$, the magnification μ becomes infinite theoretically. The loci of the points in the image plane that satisfy this condition are known as *critical curves* and the loci of

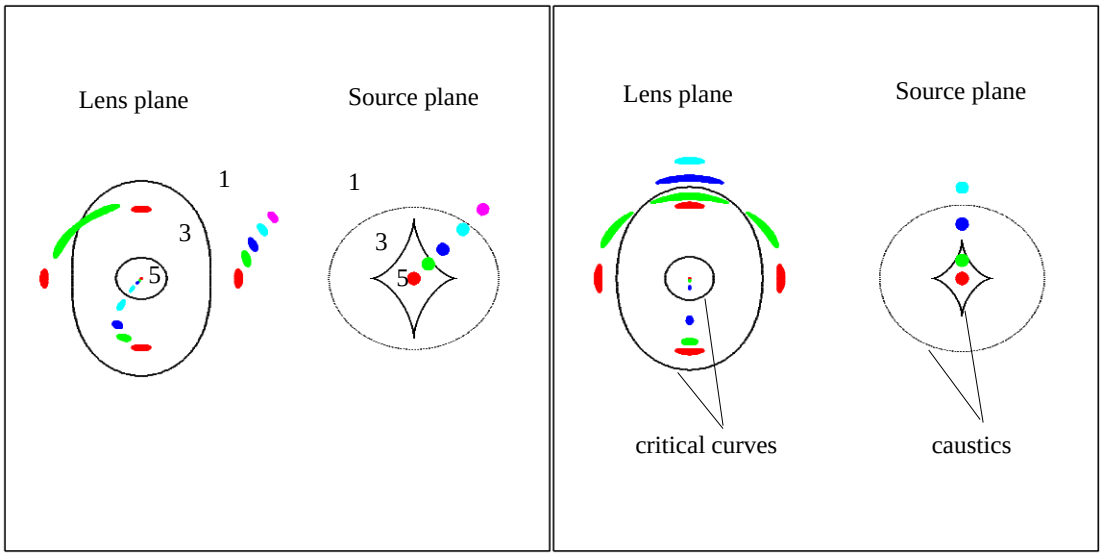

Figure 1.8: Critical curves in the lens (or image) plane and caustics in the source plane for an elliptical mass distribution. The colors represent different source positions and respective image positions. In the left panel, the source approaches the centre through the *fold* caustic and in the right panel, it approaches through the *cusp* (Narayan & Bartelmann 1996).

critical points when mapped to the source plane are called *caustics*. In reality, however, due to the finite sizes of galaxies highly stretched images like arcs and rings are formed with very high but finite magnification, since the total magnification is the flux-density weighted mean magnification over the source, that is,

$$\mu_t = \frac{\int d^2\beta\, I(\boldsymbol{\beta})\, \mu_p(\boldsymbol{\beta})}{\int d^2\beta\, I(\boldsymbol{\beta})} \tag{1.25}$$

where I($\boldsymbol{\beta}$) is the intensity and μ_p is the magnification of a point source at $\boldsymbol{\beta}$.

In Fig. 1.8, given an elliptical mass distribution and a source position (source plane), the corresponding image positions (image plane) are shown. The regions marked 1, 3 or 5 indicate a total of 1, 3 or 5 number of images formed in the image plane, respectively. In the case of an elliptical mass distribution, there are two caustics formed. The outer smooth caustic is called the *radial* caustic and the inner astroid is called the *tangential* caustic. Furthermore, the tangential caustic has two distinct features. A line singularity is called a *fold* caustic and the point at which a fold caustic changes direction such a singularity is called *cusp* caustic (e.g., Schneider et al. 1992).

Critical curves and caustics provide a useful qualitative understanding of a lens system (e.g., Blandford & Narayan 1986). The caustics demarcate regions with different multiplicity and as the source moves across the caustic the number of images change by two (see Fig. 1.8). Furthermore, 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. Fig. 1.9 shows images of real gravitational lens systems. The first two panels show both the ordinary images (a double or a quad) and the near-critical images (an Einstein

ring or an arc). The third panel has an unusual image configuration with multiple sources producing double or quadruple images.

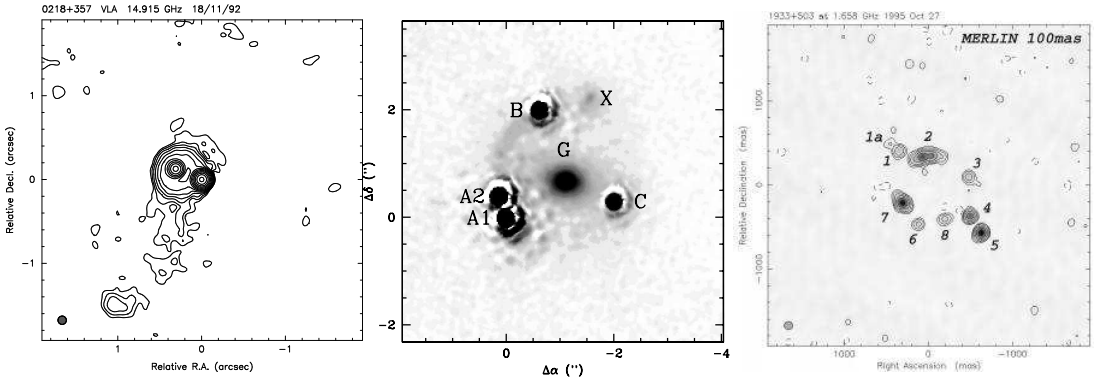

Figure 1.9: Left: Very Large Array (VLA) image of the smallest separation lens system B0218+357. It has two images and an Einstein ring (Biggs et al. 2001). Centre: The four images in the lens system MG 0414+054 with an arc and lensing galaxy in the centre as found in the optical (Ros et al. 2000). Right: A complicated lens system B1933+503 with three distinct sources lensed in an atypical configuration. One source is doubly imaged and the other two are quadruply imaged (Marlow et al. 1999).

1.3 Gravitational Lensing as a Tool

Gravitational lensing has branched into several fields with wide-spread applications. It can constrain cosmological parameters, for instance, the Hubble constant. An extensive search for extra-solar planets is possible with the help of lensing, for example, the Optical Gravitational Lensing Experiment (OGLE). Since multiple images of the same source are seen through different lines of sights in strong lensing, the properties like the interstellar medium of lensing galaxy (e.g., Falco et al. 1999; Winn et al. 2004; Mittal et al. 2007) and differential Faraday rotation in the polarization of images (e.g., Patnaik et al. 1993) can be studied.

Lensing is unbiased to the type of matter and is achromatic. Thus, it can be used to probe mass distributions at almost all observable spatial scales, for example, stars, galaxies, clusters and large-scale structures (Kochanek et al. 2004). With the help of a variety of surveys like SDSS, GEMS, GOODS and AEGIS[4], statistical analyses in lensing and other fields have been possible. For instance, the statistics of arcs, implications of image separations as a function of mass of the lensing halos and statistical significance of an image configuration over others (e.g., Bartelmann et al. 1998; Oguri 2006) have been carried out. Furthermore, the properties of galaxies like the formation and evolution have been explored (e.g., Kochanek et al. 2000; Treu et al. 2006; Koopmans et al. 2006b).

In the following, two applications of lensing are described in detail which form the subject of the research in this thesis.

[4]GEMS−Galaxy Evolution from Morphology and Spectral energy distribution Survey, GOODS−Great Observatories Origins Deep Survey and AEGIS−All-wavelength Extended Groth strip International Survey

1.3.1 Substructure

Gravitational lensing can be used as a probe to understand the problem of the missing satellites in the form of substructure. Numerous lens systems show lensed images with flux density ratios that are discrepant as compared to the expected flux density ratios from the theory of gravitational lensing. Such discrepancies in the observed flux densities are referred to as flux-ratio anomalies. Whereas the astrometric constraints of lens systems can be generally explained by smooth mass models, the anomalous flux-ratios found in a number of lens systems can be resolved by including substructure. The satellite mass fraction ($0.6 - 7\%$) inferred from the flux-ratio anomalies of a handful of lens systems are found to be consistent with the predictions of CDM cosmology (Dalal & Kochanek 2002). However, high resolution simulations of sub-halo population have suggested a much smaller substructure mass fraction of $\lesssim$0.5% within a scale of typical image separations produced by lens galaxies (Mao et al. 2004).

The anomalies in the flux-ratios of the lens systems can also be due to propagation effects like absorption and scattering (e.g., Mittal et al. 2007). However, any discrepancies found in the observed and expected positions of the lensed images can only be due to gravitation. This is called astrometric anomaly. Therefore, lens systems with an astrometric anomaly are interesting candidates, the mass modeling of which will provide a more direct evidence of substructure. Mass clumps within the Einstein ring are shown to produce measurable astrometric perturbations in artificial lens systems by populating them with realistic models of substructure distribution. Furthermore, limits on the mass of substructure clumps are placed from observed lens systems (e.g., Chen et al. 2007). Similarly, new methods are developed to study efficiently the contribution of substructure to the lens mass distribution provided enough constraints are available (Alard 2008). However, only a couple of lens systems are known to have signatures of an astrometric anomaly, for example, B0128+437 (Biggs et al. 2004) and MG 0414+054 (Trotter et al. 2000). The lens system MG 2016+112 is the clearest example of the astrometric anomaly and is dealt with in the first half of this thesis.

1.3.2 Galaxy Mass Distribution and Importance of Environment

The enclosed mass can be measured from the image separation as was shown in the Eq. 1.19 but the radial distribution of mass within, can only be probed if the lensed images are formed at different angular separations. Although this is possible (e.g., Cohn et al. 2001; Wucknitz 2004), not many lens systems can provide the required constraints.

Alternatively, combining the lensing analysis with the stellar dynamics in galaxies, the matter distribution at galaxy-scales in the host halos can be constrained since both of the methods provide mass estimates enclosed within different radii. Koopmans et al. (2006b) first implemented the technique of lensing and dynamics combined on a sample of lens systems selected from Sloan Lens Advanced Camera for Survey (SLACS) and Lenses Structure and Dynamics (LSD) surveys. The elliptical lens galaxies in their sample are found to follow an isothermal mass distribution ($\gamma = 2.01^{+0.02}_{-0.03}$ with confidence level of 68 per cent; see Fig. 1.10) with no significant evolution up to the investigated redshift $z = 1$. The average ellipticities and the position angles of the mass distribution are found to be consistent with those of the light distribution.

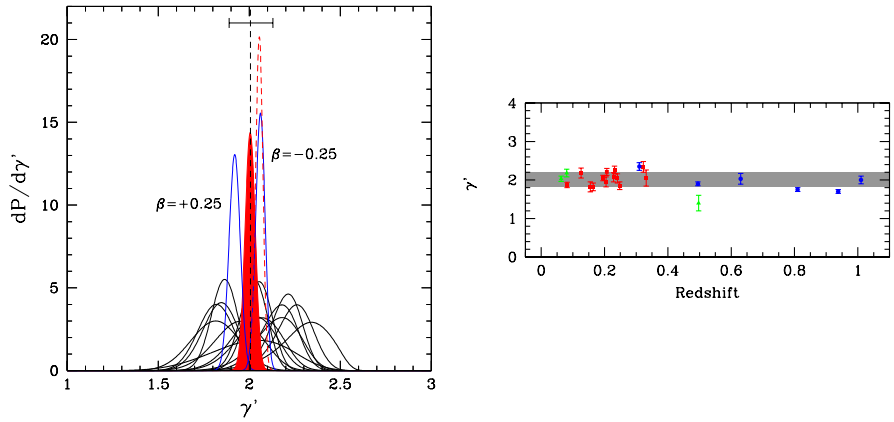

Figure 1.10: *Left*: Posterior probability distribution of the slope of the density profile. (Koopmans et al. 2006b). *Right*: Density profile slope as a function of spatial distribution of early-type galaxies of the lens systems from SLACS (red dots) and LSD (blue dots) survey.

A statistical analysis of lens galaxies versus non-lens galaxies showed that about 25 per cent of the lens galaxies lie in dense environments (Keeton et al. 2000b). Furthermore, lens systems with large image separations are predicted to have enhanced contribution from their environments (Oguri 2006). The consequences of not taking into account the effects of environments in the mass models could introduce undesired biases in the model parameters (Keeton & Zabludoff 2004). Moreover, the detection of lens systems has been used as tracers for identifying and studying galaxy groups at moderate redshifts (e.g., Williams et al. 2006).

Recently, interactions of galaxies in groups are predicted to steepen the profile from the typical isothermal profiles found for the early-type lens galaxies (Dobke et al. 2007). Subsequently, the SLACS lens systems which are best-modeled with steeper than isothermal profiles are also found to have companions as compared to those modeled with shallower than isothermal profiles (Auger 2008). B2108+213 is a wide separation lens system with a companion galaxy close to the massive lensing galaxy and is confirmed to lie in a galaxy group. In this thesis, the lens system B2108+213 is investigated with the aim of constraining the slope of the density profile and studying the contribution of the group in the image splitting by modeling the mass distribution.

2 Lens Mass Modeling

The intervening mass between a distant source and an observer acts as a lens and deflects the light rays emitted by the source. If the deflection potential of the intervening mass is strong enough, it can produce multiple images of the background source. The relative deflections and magnifications of the multiple images are governed by the properties of the lens. The principal aim of this chapter is to outline the method used to infer the properties of the deflection potential by modeling multiply-imaged gravitational lens systems.

The modeling of gravitational lenses has been advanced due to the increasing precision and sensitivity of the observed data. This has been largely possible due to high resolution imaging HST (e.g., Bolton et al. 2006a), in the optical/Near Infrared (NIR) and in the radio using VLBI (e.g., Biggs 2005). Dedicated monitoring of lens systems has made it possible to measure the flux ratios and relative time delays for an increasing number of lens systems (e.g., Saha et al. 2006). The flux densities, time delays and positions of lensed images constitute the primary constraints for modeling the mass distribution.

The advance in precision data has rendered simple models as inefficient to fit all the data quantitatively. A contribution from the environment or the line-of-sight substructure and various density profiles are some of the complications which have to be taken into account. Over the years, specific modeling algorithms have been developed like (Kneib et al. 1993), (Wucknitz 2004), (Saha & Williams 2004). These algorithms use different approaches to reconstruct the mass distribution or the potential which suit specific observational constraints. To study the mass distribution of the lensing galaxy of MG 2016+112 and B2108+213, the publicly available software package, (Keeton 2001) was used.

This chapter is divided into three sections. The algorithm used for mass modeling is described in section 2.1. Subsequently, analytical expressions for some of the standard mass models used in this thesis, are given in section 2.2. Some known degeneracies in lensing are discussed in the final section 2.3.

2.1 Algorithm for Parameterized Mass Modeling

The aim of mass modeling is to find a model for the lens that will explain the positions and flux ratios of the lensed images, as seen in the data. The true position of the source is unknown. For a given mass model, the source position can be determined from an image position and vice versa. In the case of strong lensing, multiple images of a single source are seen. Each image position can be mapped back to the source position independent of the other images. If the mass model is correct, then each of the multiple images should correspond to the same source position within the uncertainties of the measured image

positions. Furthermore, all of the images [1] predicted for this source position should be observable. This forms the basis of the algorithm used to determine a mass model. The steps adopted in the modeling algorithm are,

1. assume a simple parametric mass model
2. given this mass model and the observed image positions, use the lens equation to find the corresponding source positions and an error-weighted mean source position (see Eq. 2.8)
3. use this mean source position and the mass model to determine the properties of the lensed images (i.e. the positions, magnifications, parities and the total number of images)
4. assign a measure of goodness-of-fit (χ^2) based on the predictions of the model and the observed image properties
5. adjust the parameters of the mass model to minimize the χ^2

Steps 2 and 3 mentioned above involve solving the lens equation with the assumed model but with a subtle difference. In step 2 the source position needs to be calculated using the image position, while step 3 involves finding all of the image positions corresponding to a source position. These two steps are the workhorses of the algorithm and are performed in every step of the optimization (here, the χ^2 minimization) procedure. The optimization is carried out using the downhill simplex method (Press et al. 1992) implemented in

.

2.1.1 Solving the Lens Equation

Let $\boldsymbol{\beta}$ be the position of the source and $\boldsymbol{\theta}$ be the position of the corresponding image. These two are related by the lens equation, $\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta})$. If the parameters of the mass model are defined, then the deflection $\boldsymbol{\alpha}$ for every $\boldsymbol{\theta}$ can be determined (see section 2.2). Hence, given an image position it is easy to calculate the corresponding source position. But for a particular source position, $\boldsymbol{\beta}$, this non-linear equation has multiple solutions which gives multiple image positions, $\boldsymbol{\theta}_i$. To find the positions of all of these images, a numerical equation solver is needed which will find all the roots of the lens equation in a two dimensional plane.

Any numerical root finder can work if the expected number of images and disjoint regions that bound each of the images, are specified. The bounding regions can be found by dividing the surface of the image plane into a grid. The vertices of every tile in the grid can be mapped in a straightforward manner to the source plane via the lens equation leading to a tiling of the source plane. Regions in the source plane which are covered by more than one tile are *multiply imaged*. Thus, given the source position, the image plane tiles which map to the tiles encompassing the source, can be identified. These form the bounding regions which can be provided to a numerical root finder to solve the lens equation and to find all of the image positions for a particular source position.

[1] Barring the images which are demagnified below the flux limit of observations

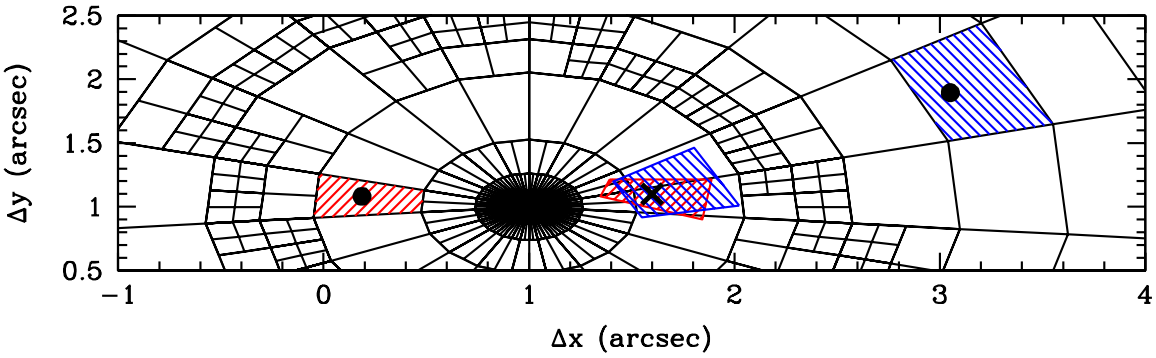

Figure 2.1: The image plane of a doubly imaged lens system gridded with tiles produced by (see text for details).

For instance, Fig. 2.1 shows the gridded image plane of a doubly imaged lens system. The cross marks the source position and the dots mark the image positions. Each tile shown in the image plane is mapped back to the source plane using the lens equation. The blue and the red hatched tile of the image plane are mapped to a blue and a red hatched tile of the source plane, respectively. As expected the source position is situated in the overlapped region of the two tiles of the source plane. Note that the gridded source plane is not shown except for the two tiles.

2.1.2 Optimization of the Model

Once a mass model is specified, the unknown source position is the error-weighted average of the source positions (see Eq. 2.8) obtained by mapping the multiple images back to the source plane. This position is used to find the positions of all of the images corresponding to this source in the image plane. The deviations of the model predicted quantities (positions and fluxes) from those of the observed images constitute the χ^2 for a model,

$$\chi^2_{pos} = \sum_i \delta\boldsymbol{\theta}_i^T \cdot S_i^{-1} \cdot \delta\boldsymbol{\theta}_i \ , \tag{2.1}$$

where

$$\delta\boldsymbol{\theta}_i = \boldsymbol{\theta}_{obs,i} - \boldsymbol{\theta}_{mod,i} \ . \tag{2.2}$$

The observed and modeled positions of the images i are given by $\boldsymbol{\theta}_{obs,i}$ and $\boldsymbol{\theta}_{mod,i}$. S_i is the error ellipse given by the covariance matrix,

$$S_i = R_i^T \begin{pmatrix} \sigma^2_{1,i} & 0 \\ 0 & \sigma^2_{2,i} \end{pmatrix} R_i \ , \tag{2.3}$$

where $\sigma^2_{1,i}$ and $\sigma^2_{2,i}$ are the semi-major and semi-minor axes of the error ellipse with a position angle $\theta_{\sigma,i}$ and

$$R_i = \begin{pmatrix} -\sin\theta_{\sigma,i} & \cos\theta_{\sigma,i} \\ -\cos\theta_{\sigma,i} & -\sin\theta_{\sigma,i} \end{pmatrix} , \tag{2.4}$$

is the rotation matrix. If the flux densities of images i are $f_i \pm \sigma^2_{f,i}$, then the χ^2 due to the deviations in the flux densities of the images is calculated as

$$\chi^2_{flux} = \sum_i \frac{(f_i - \mu_i f_{src})^2}{\sigma^2_{f,i}} \ , \tag{2.5}$$

where

$$f_{src} = \frac{\sum_i f_i \, \mu_i / \sigma^2_{f,i}}{\sum_i \mu_i^2 / \sigma^2_{f,i}} \tag{2.6}$$

is the flux density of the source and $\mu_i = \mid \det (M'_i) \mid = \mid \det (M_i) \mid^{-1}$ is the model predicted magnification of the images i relative to the source.

Other priors could be introduced in the optimization, for example, confining the position angle of the ellipticity of the host halo known from fitting the surface brightness profile of the lens galaxy or the mass scale known from the measured stellar velocity dispersion of the lens galaxy. If the number of images predicted for the source position is not correct, or the parity of the predicted images is incorrect, an arbitrarily high χ^2 is assigned to the model. All of the individual χ^2 are added to the χ^2 budget to give a total χ^2 for a model,

$$\chi^2_{tot} = \chi^2_{pos} + \chi^2_{flux} + \chi^2_{oth}. \tag{2.7}$$

Since the model optimization is carried out in the image plane it is called the *image plane χ^2 minimization*. Subtracting the total number of parameters that are optimized from the total number of constraints imposed gives the total number of *degrees of freedom (dof)* for a model. The goodness-of-fit of a model is then determined by the reduced χ^2 (i.e. χ^2_{tot}/dof) of the model. The model with a minimum reduced χ^2 ($\simeq$1, preferably) will be the best approximate mass model. Since all of the images have to be mapped back to the source plane (fast) and remapped to the image plane which involves numerical root finding (slow), this minimization is time consuming.

A minimization routine which is faster but also approximate, is the *source plane minimization*. The source plane χ^2 is defined by,

$$\chi^2_{src,pos} = \sum_i \delta\boldsymbol{\beta}_i^T \cdot M_i'^T \cdot S_i^{-1} \cdot M'_i \cdot \delta\boldsymbol{\beta}_i \ , \tag{2.8}$$

where,

$$\delta\boldsymbol{\beta}_i = \boldsymbol{\beta}_{obs,i} - \boldsymbol{\beta}_{mod} \ . \tag{2.9}$$

Here, the source positions are found from the observed positions of the images via $\boldsymbol{\beta}_{obs,i} = \boldsymbol{\theta}_{obs,i} - \boldsymbol{\alpha}(\boldsymbol{\theta}_{obs,i})$, the modeled source position $\boldsymbol{\beta}_{mod}$ is

$$\boldsymbol{\beta}_{mod} = \frac{\sum_i M_i'^T \cdot S_i^{-1} \cdot M'_i \cdot \boldsymbol{\beta}_{obs,i}}{\sum_i M_i'^T \cdot S_i^{-1} \cdot M'_i}, \tag{2.10}$$

and M'_i is the magnification matrix. The modeled source position is the error-weighted mean of $\boldsymbol{\beta}_{obs,i}$. Here, the χ^2 in the source plane (Eq. 2.8) is an approximation of the image plane χ^2 (Eq. 2.1). This approximation of $M'_i \cdot \delta\boldsymbol{\beta}_i \approx \delta\boldsymbol{\theta}_i$ holds good when $\delta\boldsymbol{\beta}$ is small. Any deviation in the source plane is multiplied by the magnification matrix M'_i to give an approximate image plane deviation. Using the best-fitting model from the approximate

source plane minimization as an initial guess model for the image plane minimization, allows the minimization in the image plane to converge faster. Since the lens equation is not solved to find all of the images of a source, this minimization can find best-fitting models which predict an incorrect number of images or do not conserve the parities of the images. Nevertheless, source plane minimization is very useful to narrow down the search in the parameter space which can be later efficiently tested with the image plane minimization.

2.1.3 Errors on the Model Parameters

When k parameters are varied, the distribution of $\Delta\chi^2 = \chi^2 - \chi^2_{\rm min}$ around the best-fitting parameters is given by the $\chi^2(k)$ distribution with k degrees of freedom. The errors on the best-fitting parameters of a model are expressed in terms of *confidence limits* using the posterior probability distribution for each of the parameters. The confidence limits indicate the probability that the true value of the parameter will lie within the specified limits with a certain per cent confidence. From the k-dimensional parameter space of the χ^2 distribution, the confidence levels on the probability of each parameter can be given individually (Press et al. 1992). For example, the boundary $\Delta\chi^2_{68\%}$ corresponds to an interval for each parameter and, statistically, a parameter will assume values within this interval with a 68 per cent (1σ) probability. The value of $\Delta\chi^2$ for a given confidence interval depends on the number of degrees of freedom k (see table in section 15.6 of Press et al. 1992).

2.2 Standard Mass Models

The solution of the lens equation is dependent on the specified mass model. In this section the analytical mass distributions of galaxy-scale halos used in this thesis are presented. These simple mass models are an approximation to the inferred mass distributions. However, by and large they represent the true distribution satisfactorily. Moreover, they can be combined to create arbitrarily complex mass distributions (i.e. multi-component mass distribution).

2.2.1 Power-law Density Profile of a Circular Lens

The radial light distribution in galaxies can be commonly fitted using a power-law (e.g., Caon et al. 1993). The density distribution in dark matter haloes can also be described with broken power-laws (i.e. NFW profile; Navarro, Frenk & White 1997). Since most galaxies and clusters of galaxies are well described by power-laws, the lensing properties of circularly symmetric density distributions are examined here. The density distribution in three dimensions is assumed to be a power-law given by

$$\rho \propto r^{-\gamma}. \tag{2.11}$$

The symmetry of the lens distribution implies that the images will be collinear with the centre of the lens and the source position. In this case, the vectorial lens equation can

be reduced to a scalar equation using $|\boldsymbol{\beta}|$ and $|\boldsymbol{\theta}|$. When the source lies directly behind the centre of the lens, an Einstein ring is formed due to symmetry.

The surface mass density κ is the projection of the line-of-sight density distribution on a plane in the sky and hence is proportional to $r^{1-\gamma}$. The surface mass density κ, the deflection angle α, and the shear γ_s are given by

$$\kappa(\theta) = \frac{3-\gamma}{2}\left(\frac{\theta}{b}\right)^{1-\gamma}, \tag{2.12}$$

$$\alpha(\theta) = b\left(\frac{\theta}{b}\right)^{2-\gamma}, \tag{2.13}$$

$$\gamma_s(\theta) = \frac{\gamma-1}{2}\left(\frac{\theta}{b}\right)^{1-\gamma}. \tag{2.14}$$

where b is the Einstein radius.

2.2.2 Singular Isothermal Sphere (SIS) Profile

A simple case of the power law lens models with $\gamma = 2$ applies fairly well to the mass distribution seen in galaxies (Koopmans et al. 2006b). The model corresponds to a self-gravitating spherically-symmetric ideal gas in steady state equilibrium with a constant temperature at all radii. Hence, it is called isothermal and the term singular reflects the fact that the density diverges at the centre. Flat rotation curves as observed in spiral galaxies are a characteristic of this distribution (e.g., Rubin et al. 1985). In the case of elliptical galaxies, the velocity dispersion of stars acts as a kinetic temperature which is constant with radius (e.g., Binney & Tremaine 1987).

The surface mass density at a projected radius ξ for a SIS is given by

$$\Sigma(\xi) = \frac{\sigma^2}{2G\xi} \tag{2.15}$$

where σ^2 is the one dimensional velocity dispersion. Since $\kappa = \Sigma/\Sigma_{\rm crit}$ and using Eq. 1.9 for $\Sigma_{\rm crit}$, the convergence for a SIS is

$$\kappa(\theta) = \frac{1}{2\theta}\frac{4\pi\sigma^2 D_{ds}}{c^2 D_s}\,. \tag{2.16}$$

Comparing Eq. 2.16 with Eq. 2.12 for $n = 2$ gives

$$b = \frac{4\pi\sigma^2 D_{ds}}{c^2 D_s}, \tag{2.17}$$

which is the same as the Einstein radius. Moreover, the deflection and shear are given by

$$\alpha(\theta) = b, \tag{2.18}$$

and

$$\gamma_s(\theta) = \frac{1}{2}\frac{b}{\theta} = \kappa(\theta)\,. \tag{2.19}$$

Thus, the deflection is a constant at all impact parameters.

2.2.3 Non-singular Isothermal Sphere (NIS) Profile

The singularity at the centre of a SIS can be avoided if the density distribution of the halo near the centre is replaced by a constant within a core radius (r_c). This is termed a non-singular (or softened) isothermal profile. Introducing a projected core radius ξ_c in Eq. 2.15 gives

$$\Sigma(\xi) = \frac{\sigma^2}{2G\sqrt{\xi^2 + \xi_c^2}}. \tag{2.20}$$

The lens properties are thus,

$$\alpha(\theta) = \frac{b\,(\sqrt{\theta^2 + \theta_c^2} - \theta_c)}{\theta}\,, \tag{2.21}$$

$$\kappa(\theta) = \frac{b}{2\sqrt{\theta^2 + \theta_c^2}}\,, \tag{2.22}$$

where $\xi_c = D_d\theta_c$ and,

$$|\,\gamma_s(\theta)| = \frac{b\,(\omega - \theta_c)}{2\,\omega\,(\omega + \theta_c)}\,. \tag{2.23}$$

where $\omega^2 = \theta^2 + \theta_c^2$.

2.2.4 Singular Isothermal Ellipsoid (SIE) Profile

Spherical mass distributions cannot produce quadruple images. Furthermore, numerical simulations predict ellipticity and triaxiality in the mass distribution of galaxies (e.g., Jing & Suto 2002) and observations show ellipticity in the surface brightness distribution. Therefore, any possible angular structure in the lens potential must be accounted for in the models. The spherically symmetric density profile can be modified using an axis ratio (q) to include ellipticity e (e.g., Kormann et al. 1994). The corresponding surface mass density is,

$$\Sigma(\xi_1, \xi_2) = \frac{\sigma^2}{2G}\frac{\sqrt{q}}{\sqrt{q^2\xi_1^2 + \xi_2^2}} = \frac{\sigma^2}{2G}\frac{\sqrt{q}}{\zeta}, \tag{2.24}$$

such that ξ_1 and ξ_2 are along the major axis and minor axis of the ellipsoid, respectively[2]. Here, $\zeta = (q^2\xi_1^2 + \xi_2^2)^{1/2}$ and $0 < q \leq 1$. Iso-density elliptical contours labeled by ζ correspond to ellipses with ζ as the minor axis and ζ/q as the major axis. The normalization used in this convention implies that the mass inside an iso-density contour for a fixed Σ is independent of q.

In , the major axis $\xi_1 = D_d\theta_1$, the minor axis $\xi_2 = D_d\theta_2$, the ellipticity $e = 1 - q$ and the scaled surface mass density κ is,

$$\kappa(\theta_1, \theta_2) = \frac{b}{2[(1-\epsilon)\,\theta_1^2 + (1+\epsilon)\,\theta_2^2]^{1/2}}. \tag{2.25}$$

[2]The axes labels ξ_1 and ξ_2 are flipped relative to those in Kormann et al. (1994) while maintaining consistency in their physical meaning.

Here ϵ is related to the axis ratio q by

$$q = \sqrt{\frac{1-\epsilon}{1+\epsilon}}, \tag{2.26}$$

and the mass scale b is related to the velocity dispersion σ by,

$$b = \sqrt{\frac{2q}{1+q^2}}\, 4\pi \frac{\sigma^2}{c^2} \frac{D_{ds}}{D_s}. \tag{2.27}$$

Since there is no circular symmetry, the deflection angle cannot be reduced to a one dimensional scalar and has two components, one along the θ_1 axis and the other along the θ_2 axis. These are denoted by α_1 and α_2 respectively and are given by,

$$\alpha_1 = \frac{b}{\sqrt{2\epsilon}} \tan^{-1}\left(\frac{\theta_1 \sqrt{1-q^2}}{[q^2\,\theta_1^2 + \theta_2^2]^{1/2}}\right) \tag{2.28}$$

and,

$$\alpha_2 = \frac{b}{\sqrt{2\epsilon}} \tanh^{-1}\left(\frac{\theta_2 \sqrt{1-q^2}}{[q^2\,\theta_1^2 + \theta_2^2]^{1/2}}\right). \tag{2.29}$$

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

2.3 Degeneracies in the Models

Although lensing is very useful to estimate the mass enclosed within the Einstein radius of the lensing halo and to constrain the Hubble constant independently, however, it has a set of drawbacks. Some parameters are degenerate and cannot be uniquely estimated from gravitational lensing alone. Two common degeneracies are discussed here.

Steepness degeneracy

The lens equation that relates the background source position to the position of the images is

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}) \tag{2.30}$$

where the deflection angle (α) depends on the lens mass distribution. Consider a circularly symmetric lens potential at the origin and two images formed at θ_A and θ_B on the opposite side of the lens. Using Eq. 2.30, the unknown source position can be eliminated to give

$$\theta_A - \alpha(\theta_A) = -\theta_B + \alpha(\theta_B). \tag{2.31}$$

where θ_B is in the negative quadrant. The deflection produced by a circularly symmetric power law profile ($\rho \propto r^{-\gamma}$) is given by $\alpha = b^{\gamma-1}\theta^{2-\gamma}$. If $\alpha(\theta)$ is substituted in Eq. 2.31, then the Einstein radius (b) or more generally, the critical radius can be given as

$$b = \left(\frac{\theta_A + \theta_B}{\theta_A^{2-\gamma} + \theta_B^{2-\gamma}}\right)^{1/(\gamma-1)}. \tag{2.32}$$

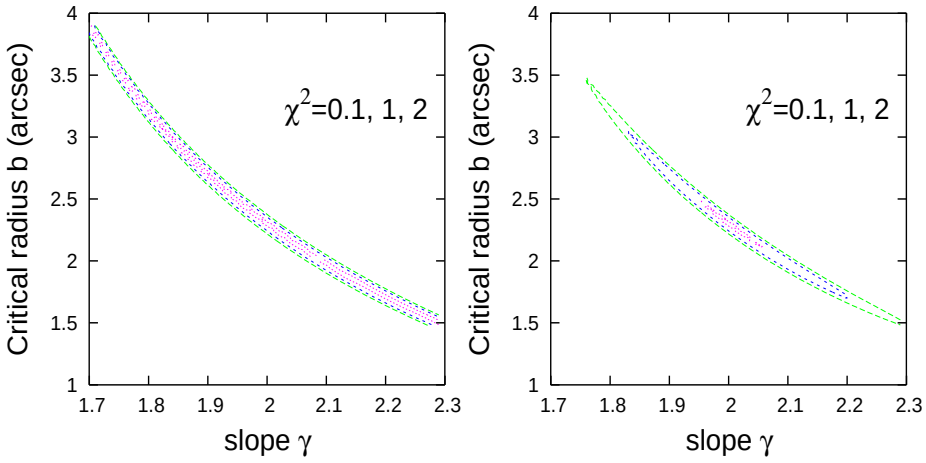

Figure 2.2: The degeneracy between the critical radius and the slope of the power law profile is broken by introducing the constraints from the structure in the two images.

Eq. 2.32 shows that for any doubly imaged lens system with an arbitrary circularly symmetric profile, and fixed values of the observables θ_A and θ_B, a degeneracy is introduced between the power law index γ and the Einstein radius b. In principle, this degeneracy is broken by using the flux density ratio. However, there are not enough constraints to take into account the effects of asymmetry in the lens potential. Using extra constraints from the structure in a doubly imaged lens system could break the degeneracy. This is illustrated below with a toy model using .

Consider a SIS+shear model with an Einstein radius of $b = 2.3$ arcsec at (0,0), a shear of 2 per cent with a position angle of 80 deg and a source at (1,1) arcsec. This toy model is used to generate mock data with two point-like images. These data, when fitted with a power law, show degeneracy between the slope of the density profile and critical radius b (and hence, the mass within). Fig. 2.2 shows the χ^2 surface plot for the critical radius b and slopes of the power law density profile. The panel on the left shows that several models are found to be equally good which have b and γ in different combinations when the source has two point-like images.

Now, consider a second source at (1.0008, 1.0011) arcsec introduced in the previous model. Addition of this source generates two more images close to the images generated by the first source. A model fitted to the data reproduces all the parameters very well as shown in the right panel of Fig. 2.2. The χ^2 surface is well-constrained giving more definite values for both b and γ. Thus, the structure in the images defined by an annulus $\delta\theta = \theta_A - \theta_B$ encompassing the images can be used to constrain the density profile within the annulus. (e.g., Kochanek et al. 2004).

Degeneracy between ellipticity and shear

The quadrupole moment of the lens potential introduces angular perturbations in the potential and the mass distribution which are observable. In the Taylor-expanded expression for the lens potential, further higher order terms also exist but are generally negligible. However, what is indistinguishable is the effect of individual factors, that is, the elliptic-

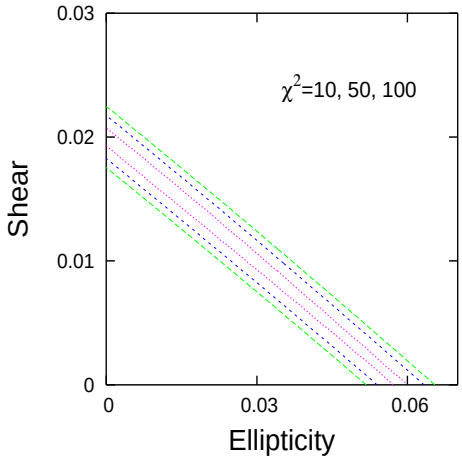

Figure 2.3: A contour plot of χ^2 as a function of shear and ellipticity showing the degeneracy between the two quantities.

ity, the internal shear (i.e. arising from the mass within the critical radius) and the external shear. This is demonstrated by their individual contributions to the potential as,

$$\psi(\theta,\chi) = -\frac{\epsilon\,\theta}{2}\cos 2(\chi - \chi_\epsilon) \ , \tag{2.33}$$

and,

$$\psi(\theta,\chi) = -\frac{\gamma_s\,\theta^2}{2}\cos 2(\chi - \chi_{\gamma_s}) \ , \tag{2.34}$$

where γ_s represents the internal and external shear together.

A toy model of an isothermal sphere and an external shear (2 per cent) with a position angle of 70 deg was used to generate multiple images of two artificial sources. The mock data thus generated was tweaked within the assigned uncertainties of 1 mas on the positions. A model with an isothermal ellipsoid and external shear was then used by fixing the position angle at 70 deg for both ellipticity and shear, and allowing the ellipticity and shear to fit the data. Fig. 2.3 shows that the values of the χ^2 for models with various combinations of ellipticity and shear are equally preferred. The ellipticity can be increased in the model and shear can be decreased to get an equally acceptable model. Similarly, the shear can be increased up to the initial value of 2 per cent by decreasing the ellipticity to 0. Thus, it is difficult to constrain the effects of ellipticity and shear individually.

3 Interferometry and VLBI techniques for data reduction

Gravitational lensing can produce multiple images of a distant background source that cannot be seen directly due to the intervening matter. The light coming from such distant sources is likely to interact with the line-of-sight particles on its way to the observer. Since lensing is an achromatic phenomenon, the light from lensed images can be observed at various wavelengths. The radio emission, although weak, can carry information from the distant Universe since it suffers little from propagation effects. Also, the emission is seen with radio from milli-arcsecond (mas) to arc-minutes scales. Moreover, the terrestrial atmosphere is transparent to a wider range in the radio regime making the less expensive ground based instruments a viable option as compared to other electromagnetic wave bands like x-rays where space based instruments are inevitable. Therefore, it is more efficient to study distant gravitationally lensed sources and their properties in the radio.

In an attempt to achieve higher resolutions, larger telescopes or higher frequencies are needed (resolution = wavelength/size of the aperture). For practical reasons, telescopes cannot be arbitrarily enlarged or the source of interest may not always emit at higher frequencies. Instead, signals from several telescopes separated over large distances can be combined. This is called *Interferometry*. The early radio interferometers (ca. 1950) were similar to Michelson interferometer where the principle is the same as in the optical. In a radio interferometer, the signals from two antennas are added and fed to a square-law detector where the output is proportional to the square of the sum of two antenna voltages. The next important development was in 1952 when Ryle introduced the phase switching interferometer. Here, the phase of one antenna is periodically reversed such that the output is proportional to the time average of the product of the voltages measured by each antenna, which is called *cross-correlation* of signals.

Advanced techniques have been developed and numerous radio arrays have been constructed since, for example, the Very Large Array (VLA) consisting of identical and mobile antennas spread over regular distance intervals or the Multi-Element Radio Linked Interferometer Network (MERLIN) with different sized and randomly distributed antennas. For both the VLA and MERLIN the data are correlated in real time. With increased separations between antennas, the technique used in combining the signals also becomes important, for example, it should be cost effective, the losses in the signal amplitude should be minimized, the signals should be synchronously combined etc. *Very Long Baseline Interferometry* (VLBI) is a technique wherein the antennas are not physically connected and the signal from each antenna is recorded on disks. These tapes are brought together to be multiplied and averaged in a *correlator*. Therefore, the phases of the signals

from all of the antennas need not be synchronized during the observation i.e in real time. This technique allows the use of the longest baselines possible on ground (and space) to give the highest angular resolution, typically of the order of sub-milliarcsecond. This allows celestial sources to be studied in detail, and astrometric and geodetic measurements to be made with great accuracy. However, accounting for measurement errors is much more difficult in VLBI compared to the errors in smaller and homogeneous arrays because the geographical/weather conditions at each antenna are different. Such differences have to be taken into account.

Throughout this thesis, high angular resolution imaging with interferometer arrays is used to study gravitational lens systems. The following sections are to familiarize the reader with the basic concepts involved in the process of making radio images using the technique of interferometry[1].

3.1 Introduction

This section compares the response of a single element aperture with a multi-element interferometer to show the importance the latter has in the observational aspect of this thesis project.

Resolution

Consider a single element antenna with an aperture of size D that receives a signal at wavelength (λ). The Fourier transform of the aperture illumination is the far-field pattern. In the one dimensional case,

$$E(\phi) = \int E(x)\, e^{-ikxsin\phi} dx \tag{3.1}$$

where $E(x)$ is the aperture field distribution and the power pattern is $P(\phi) = |E(\phi)|^2$. The power pattern is also known as the response of the antenna and has a maximum in the centre and weak side-lobes at the edges (assuming symmetry). A plot of a one-dimensional power pattern is shown in red in Fig. 3.1. The angular resolution of an antenna is given by Rayleigh's criterion as the separation between the maximum and the first null which occurs at $\sim \lambda/D$.

Suppose there are two antennas each with aperture D and separated by a distance B called the baseline. The combined response of the antennas is the interference pattern shown in green in Fig. 3.1 which is enveloped by the diffraction pattern of the single antenna. Here, the resolution is again given by the extent of the first null from the central maximum which occurs at $\sim \lambda/B$. Thus, as the baseline is increased higher, a higher resolution is achieved. In Fig. 3.1, baseline B is chosen four times the single antenna aperture D for illustrative purpose. Practically, a baseline is several times larger than the aperture of an antenna, for instance, the size of the VLA antennas is 25 m whereas the longest baseline is 36 km.

[1]The concepts discussed here are mainly relevant to the topic of the thesis. For detailed understanding, see Thompson et al. (1986), Taylor et al. (1999), Zensus et al. (1995) and Perley et al. (1989)

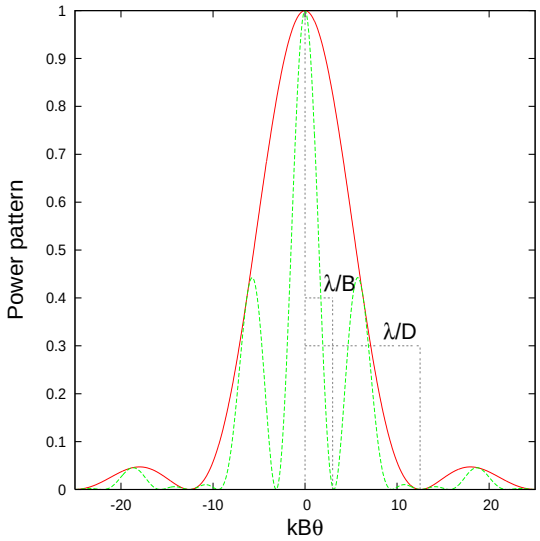

Figure 3.1: Interference fringes of the two antenna power pattern enveloped by the single antenna power pattern. Therefore, the resolution ($\sim \lambda/B$) achieved due to a two-antenna baseline B is much higher than the resolution ($\sim \lambda/D$) from a single antenna of size D.

What do we measure ?

The fringes obtained from the interference pattern of the two antennas have an intensity maximum (I_{max}) and an intensity minimum (I_{min}). Hence, the visibility is defined as

$$\mathcal{V} = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \, . \tag{3.2}$$

For a point source, the I_{min} in the fringes will reach zero thereby, giving $\mathcal{V} = 1$. This almost never happens in reality since a source has a finite size and the emission from different points in the extended source do not interfere destructively at the same point. Thus the $I_{min} > 0$ and so the visibility $\mathcal{V} < 1$.

It is convenient to measure the visibilities in the following way. Consider a distant source in an arbitrary direction of unit vector $\bar{s}$ making an angle θ with respect to a reference point called the phase centre (see Fig.3.2). The signals measured by two antennas at a distance B apart differ due to the difference in the paths traversed. The corresponding time delay in the signal received by the first antenna with respect to the second is called the geometric time delay ($\tau_g = \vec{B} \cdot \bar{s}/c$). In order to compensate for this and combine the signals coherently, an instrumental delay is introduced in the path of the second antenna. The signals received from each antenna are further processed digitally, as shown in the schematic diagram, and are filtered and down-converted to intermediate frequencies for practical convenience. The cross-correlation of the voltage signals is performed by multiplying and then averaging in an integrator to give a complex quantity called the visibility.

$$\mathcal{V} = \langle V_1 V_2^* \rangle \, . \tag{3.3}$$

The visibility from a pair of antennas at an instant gives a Fourier component of the brightness distribution of the source. If the visibilities are measured as a function of

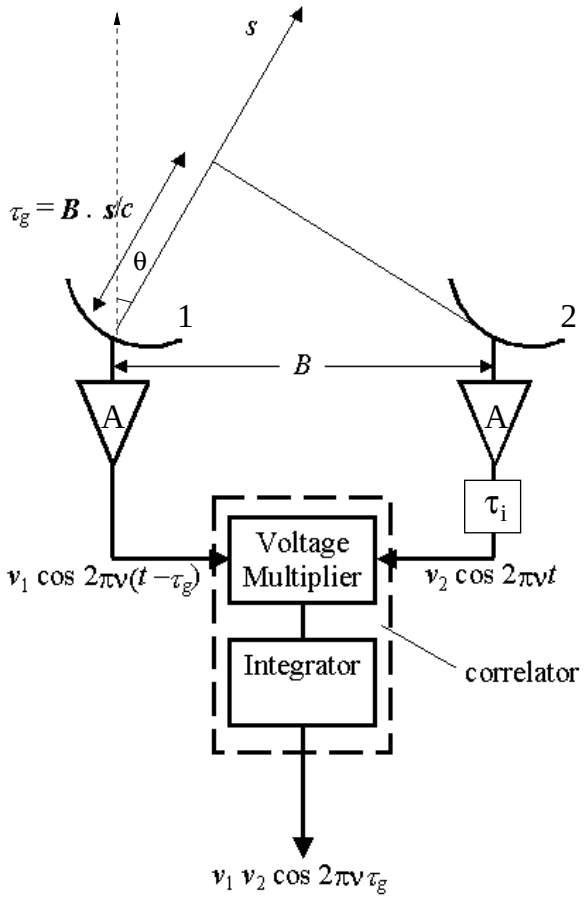

Figure 3.2: The radiation emanated from a source with a direction vector $\bar{s}$ is received by two antennas separated by baseline *B*. The measured signals are combined in the correlator to produce the coherence function which is the Fourier transform of the brightness distribution (adapted from Thompson 1999).

different antenna separations (baselines), a visibility function can be obtained and used to reconstruct the brightness distribution ($I(\bar{s_x})$). This visibility function is also called the spatial coherence function. Therefore, for a source $\bar{s_x}$ away from the reference direction, the two-dimensional visibility function is

$$\mathcal{V} = \int_{4\pi} A(\bar{s_x})\, I(\bar{s_x})\, e^{-i2\pi\nu\bar{B}\cdot\bar{s}/c} d\Omega = |\mathcal{V}|\, e^{i\phi} \tag{3.4}$$

where $A(\bar{s_x})$ is the power pattern of the antenna and ϕ is the phase of the cross-correlated signal.

3.1.1 Synthesis Imaging

Consider the projection of baselines in the *uv*-plane such that the *v*-axis points towards the North celestial pole and *u*-axis towards the east. Let the source brightness distribution (*I*) be defined in the *xy*-plane. Ideally, if the visibilities are measured at all points in

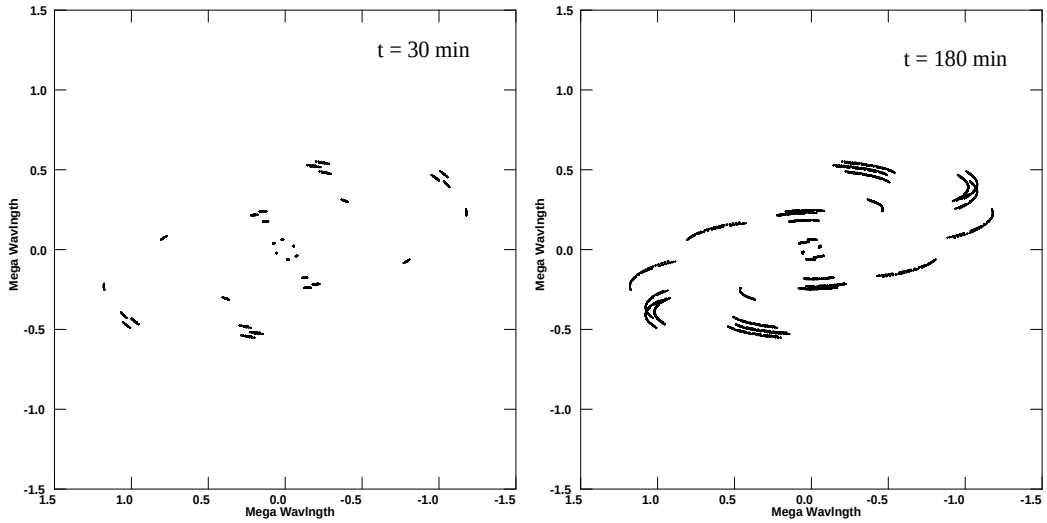

Figure 3.3: Snapshots of the distribution of the visibilities using MERLIN show increasingly filled *uv*-data points in the *u,v*-plane due to the motion of antennas with the rotation of the earth over time.

this plane, the brightness distribution (*I (x,y)*) can be calculated. Practically, this plane has large gaps since it is not economical to construct antennas to fill the visibility plane completely. Fortunately, as the Earth rotates, the projected baselines change as a function of time, thereby giving visibilities at different points in the *uv*-plane. This is called *Earth Rotation Synthesis*.

Fig. 3.3 shows the baselines of the MERLIN antennas which trace the visibility function along curves, filling the *uv*-plane at different instants of time. Each curve corresponds to a pair of antennas such that one of the antenna is at the origin and the other traces a curve with respect to the former. Fig. 3.4 is an illustration of the spatial extent of the *uv*-coverage of MERLIN and the Very Long Baseline Array (VLBA). The VLBA can map features at milli-arcsecond scales whereas MERLIN is best used for mapping sources extended over several sub-arcseconds. Note that low resolution arrays are equally as useful as high resolution ones since different sciences can be done with both of them.

From the van Cittert-Zernicke theorem (Born & Wolf 1959), the brightness distribution is given by the Fourier transform of the visibility ($\mathcal{V}$) which is a complex function measured by the antenna

$$I(x, y) \;=\; \mathscr{F}(\mathcal{V}(u, v)) \tag{3.5}$$

where $\mathcal{V}(u, v) = |\mathcal{V}|e^{i\phi}$ is the visibility in the *uv*-plane.

The visibility function is *sampled* at various points in the visibility plane (u_k, v_k). The sampling function *S(u,v)* is given by

$$S(u, v) = \sum_k w_k \, \delta(u - u_k) \, \delta(v - v_k) \tag{3.6}$$

where w_k is the weighting factor. The sampled visibility function is given by

$$\widetilde{\mathcal{V}}(u, v) = S(u, v) \times \mathcal{V}(u, v) \; . \tag{3.7}$$

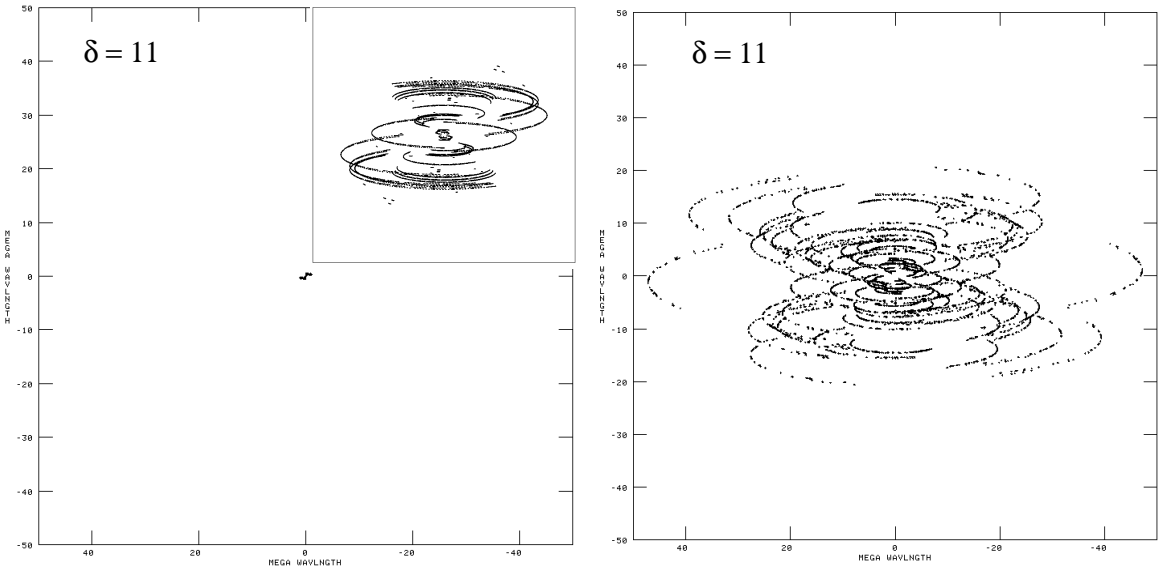

Figure 3.4: Distribution of the visibilities in the *uv*-plane showing MERLIN and VLBA baselines at the same scales. The inset shows a zoomed-in image of MERLIN baselines.

A dirty image is a reconstructed image of the source brightness distribution but with incomplete visibility function i.e. gaps hence, referred to as dirty. The dirty image $I^d(x,y)$ can be obtained by Fourier transforming the sampled visibility function,

$$I^d(x,y) = \mathscr{F}\Big(\widetilde{\mathcal{V}}(u,v)\Big) = \mathscr{F}\Big(S(u,v)\Big) * \mathscr{F}\Big(\mathcal{V}(u,v)\Big)\,. \tag{3.8}$$

The second equality comes from the convolution theorem, the right hand side is the convolution of the dirty beam B_d with the real image I i.e.

$$I^d(x,y) = B_d * I \tag{3.9}$$

where $B_d = \mathscr{F}\langle S(u,v)\rangle$ is the point spread function. Therefore, the dirty beam B_d must be deconvolved from the dirty image I^d to obtain the true image intensity distribution (I).

3.1.2 Basic Terms

Some basic terms which are used in this thesis are now defined,

Coherence time

The time t_c within which the phase of the fringes changes by less than a radian is called the *coherence time*. This quantity can depend on atmospheric effects, electronics, accuracy of system clocks etc. Sources which are weak need longer averaging time but this time is limited by the coherence of the phases within that time. If the data is averaged for $t > t_c$, it is difficult to find the peak of the fringes. This degrades the Signal-to-Noise Ratio (SNR).

Sensitivity

The measurement of the weakest feature of emission that can be detected is called the *sensitivity*. The sensitivity of an image obtained from N identical antennas observing with a bandwidth $\Delta\nu$ and integrated over time t_{int} for a single polarization is,

$$\Delta I_m = \frac{1}{\eta_s}\sqrt{\frac{T_{sys}}{N(N-1)\,t_{int}\,\Delta\nu\,K}}\,. \tag{3.10}$$

Factors like receiver noise, feed losses, spillover, atmospheric emission etc. are included in the system temperature (T_{sys}); η_s is the system efficiency which accounts for the effects of the electronics; $K = \eta A/2k$ is a measure of antenna performance consisting of antenna efficiency (η), area of the antenna (A) and Boltzmann's constant (k). Note: T_{sys} and K are assumed to be the same for all antennas which is generally not true for VLBI arrays since the antennas have different sizes and efficiencies. It is clear from Eq. 3.10 that antennas with low T_{sys} and/or large areas are most sensitive. Also, including more antennas in an array will improve the image sensitivity.

3.1.3 Smearing Effects

The frequency of a signal received by an antenna is generally converted to an intermediate frequency which is convenient for electronic processing of the signals. This is called the *intermediate frequency* (IF). Furthermore, the signals are not measured at a single frequency but at a range of frequencies which forms the *bandwidth* and each IF is divided into channels within that bandwidth. The consequences of averaging data over the bandwidth and time are discussed below.

Bandwidth smearing

The *uv*-plane is obtained by scaling the antenna separations with the wavelength of observation i.e. $u = x_u/\lambda$, $v = x_v/\lambda$ where x_u (and y_v) is the antenna separation projected along u (and v) axis. Given a baseline, the ratio u_λ/v_λ = constant. Therefore, the frequencies within a bandwidth ($\triangle\nu$) correspond to different *uv*-points in the radial direction. Due to the Fourier relation between the *uv*-plane and the image plane, the points in the image plane get rescaled too. If the visibilities averaged over a bandwidth ($\triangle\nu$) centred at ν_0 are changing significantly, then the features in a map are radially smeared and suffer loss of peak flux density far from the phase centre. This is called *bandwidth smearing*. This effect is given in terms of fractional loss in the peak intensity $\frac{I}{I_0} \propto \beta$. Here, $\beta = \frac{\triangle\nu\ \theta_0}{\nu_0\ \theta_{HPBW}}$ is proportional to the fractional bandwidth ($\triangle\nu/\nu_0$) and distance from the phase centre scaled by the beam-width (θ_0/θ_{HPBW}). In Fig. 3.5, a feature observed with ~8 mas resolution (of VLBA) at 1.7 GHz, situated at an angular separation of ~ 4.6 arcsec from the phase centre is shown. The effect of radial smearing can be clearly seen due to averaging over the complete bandwidth of 32 MHz in the left panel. When the visibilities are averaged over 8 MHz bandwidth, the peak intensity and the shape of the feature is less affected. In case of averaging over a bandwidth of 0.5 MHz, the smearing effects are minimum and the percent loss in the peak intensity is negligible.

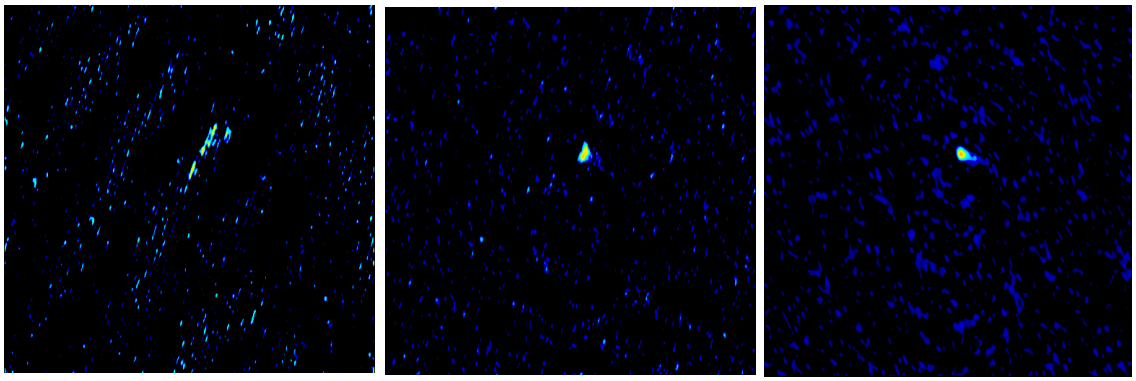

Figure 3.5: For frequencies averaged over 32, 8 and 0.5 MHz wide bands, the effects of bandwidth smearing of a feature in the map are shown.

Thus, when mapping a wide field with a single phase centre, the fidelity of the features far from the centre can be maintained by not averaging the data over the whole bandwidth.

Time average smearing

When the visibilities are averaged over a time interval δt, the averaged value corresponds to a time (τ_a). However, due to the rotation of the earth, the visibility function rotates through an angle $\omega_e \delta t$, resulting in different *uv*-data points corresponding to different times in the interval δt. This in turn causes smearing of the features in the image through the same angle and leads to a loss of peak flux density which is proportional to the time τ_a and the distance from the phase centre scaled by the beam-width (θ_0/θ_{HPBW}) i.e.

$$1 - \frac{I}{I_0} \propto \left(\frac{\theta_0}{\theta_{HPBW}}\right)^2 \tau_a^2 \,. \tag{3.11}$$

3.2 Phase Referencing

The phase measured by the correlator is the visibility phase along with the associated errors. These errors originate from the atmosphere, ionosphere, instruments and geometrical (astrometric) uncertainties. For VLBI, each of these terms can be significant and hence the phase needs to be corrected. Geometric errors can be due to uncertainties in the terrestrial or celestial co-ordinates. Instrumental errors come from the propagation of the signal through different antenna electronics, use of different clocks at different stations, inefficient instrumental response due to elevation dependent effects and axis-offsets due to rotation of the axes of antennas. The antennas are generally separated by several kilometers. Not only do they look through the atmosphere or ionosphere along different geometric path lengths but also the compounds with different refractive indices ($n \neq 1$) in the atmosphere. Also, the ionized particles in the ionosphere give rise to additional phase delays.

These phase errors and delays can be corrected (using fringe fitting / self calibration) provided the source has a high enough signal-to-noise ratio (SNR). Otherwise, a noise

spike can give a false signal detection when trying to align the phases from different antennas. This means only strong sources can be imaged. However, one could still image weak sources by taking the corrections from a strong source nearby (typically $< 2^{\circ}$). This is exactly what is done in *phase referencing*[2]. In this technique, the weak target source is observed at short intervals along with a nearby strong calibrator source. If the phases are measured in an interval short enough such that they can be added coherently, then the phase errors will not be significantly different for both of the sources. These phase errors determined for the calibrator can then be interpolated and used to correct for the target source.

The phase calibrator should satisfy the following criteria: a) it should be a point source (unresolved at most frequencies), b) it should be strong (loosely SNR $>$ 5 within the coherence time), and c) the calibrator source position should be known with high accuracy. Due to the high resolution of VLBI it is not always possible to find an unresolved strong source near to the target. Experience shows that phase referencing produces best results for a switching angle (θ_s, angle between target and calibrator) of $\sim$ 2 deg at 8.4 and 5 GHz, and $< 4 - 5$ deg at 1.7 GHz.

3.3 Editing

Editing is an important step in data reduction. Some careful editing is required to flag bad data which are baseline or antenna dependent. Sometimes the antenna is off source when it should be looking at the source or certain scans are bad (for example, high spikes) due to bad weather. Such scans should be completely flagged to avoid incorrect interpolation of the calibration. Even after the calibration has been applied to the data, some editing or flagging might be needed to take care of calibration problems. For instance, a scan on an antenna with an incorrect calibration factor shows low amplitude compared to the neighboring scans, hence it needs to be deleted. Fig. 3.6 shows examples of some real data from observations presented in this thesis. The top left panel shows the first few bad scans on baselines with Arecibo (AR) indicating it is not yet looking at the source. These scans should be flagged. The top right panel shows that for most of the scans on baselines with Fort Davis (FD) the antenna was on-source after a time delay of a few minutes. This means FD is slewing and needs flagging. Bottom left panel shows Torun (TR) with some amplitude calibration problems on all baselines whereas Effelsberg (EF) on the bottom right clearly shows amplitude of about 0.2 Jy for the same baselines. Thus, TR was completely rejected and a few bad data points were edited out for EF.

3.4 Calibration

The measured visibilities are not the same as the true visibilities owing to various factors. Correcting for these factors to get the value closest to the true visibilities is called *Calibration*. Calibration is required at several stages of data observing, recording, processing

[2]Note: Since the phases are measured by choosing the calibrator as the phase centre, only the relative position of the target with respect to the calibrator is known in phase referencing.

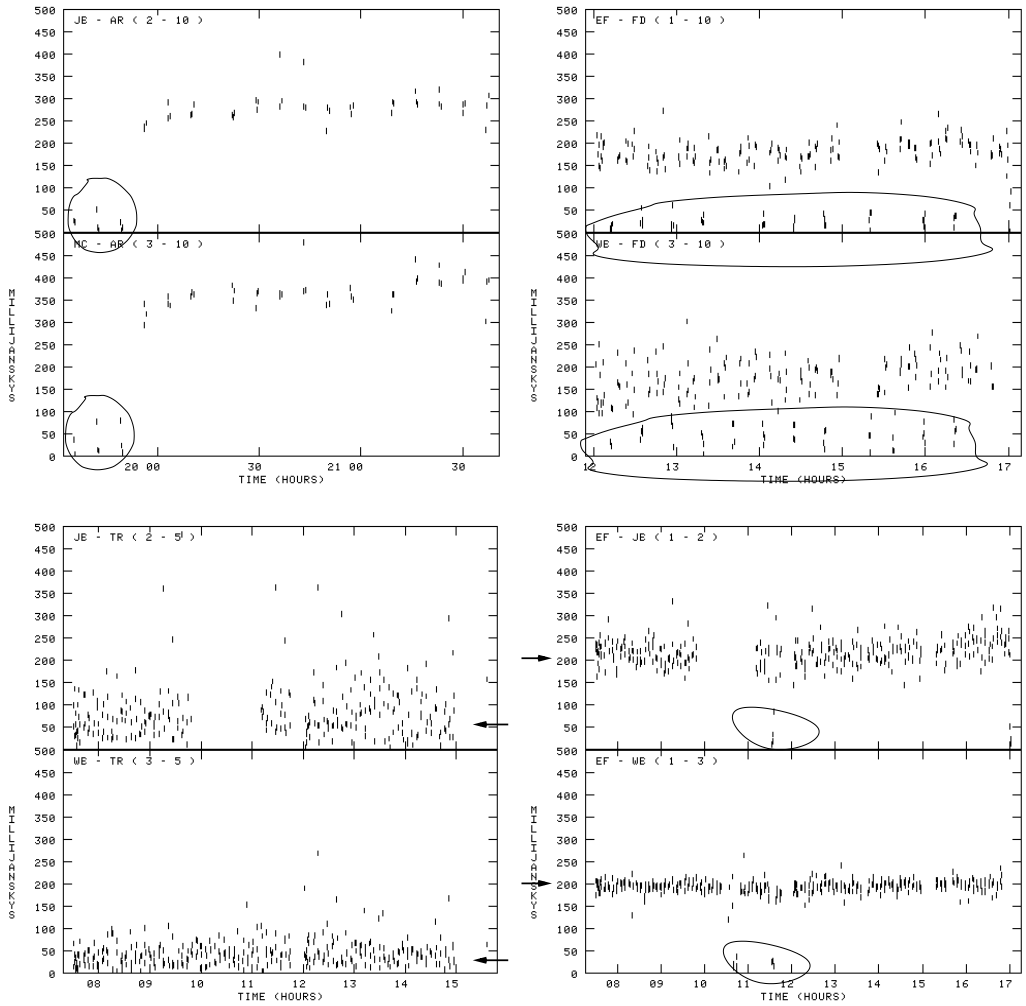

Figure 3.6: The panels in the figure highlight the bad data resulting from problems of different natures.

etc. Some initial calibration like accurate positions of the antennas, their pointing, an accurate source position and checking the gains of the antennas with the help of a calibrator are essential.

3.4.1 Amplitude Calibration

The measured visibility is related to the true visibility as follows,

$$\widetilde{\mathcal{V}}_{ij}(t) = \mathcal{G}_{ij}(t)\,\mathcal{V}_{ij}(t) + \epsilon_{ij}(t) + \eta_{ij}(t)\,, \tag{3.12}$$

where the subscripts i and j represent antennas, $\mathcal{G}_{ij}(t)$ is the baseline based complex gain, ϵ_{ij}(t) is the baseline based complex offset and η_{ij}(t) is the baseline based complex noise. The *baseline based complex gain* can be divided into complex gains of associated antennas and a baseline based residual gain, $g_{ij}(t)$ is called the *closure error*,

$$\mathcal{G}_{ij}(t) = g_i(t)\,g_j^*(t)\,g_{ij}(t) = a_i(t)\,a_j(t)\,e^{i(\phi_i(t)-\phi_j(t))}\,. \tag{3.13}$$

Generally, the closure error is a small factor and can be ignored. The antenna gain further consists of amplitude and phase corrections. Ignoring the noise terms in Eq. 3.12 and retaining only the amplitudes, the measured visibility amplitude (i.e. the raw correlation coefficient) is then

$$\Gamma_{ij} = a_i(t)\, a_j(t)\, S_{ij} \; . \tag{3.14}$$

Substituting for the amplitude terms of the antenna gains and rearranging the Eq. 3.14 gives,

$$S_{ij} = \Gamma_{ij}\, b \, \sqrt{\frac{T_{sys_i} T_{sys_j}}{K_i K_j}} \tag{3.15}$$

where b is the quantization correction factor, T_{sys} is the system temperature of an antenna in Kelvin and K_i is a measure of antenna performance in K Jy^{-1}. The system temperature is measured frequently in an observation run since it is time and elevation dependent. Using T_{sys} and K to determine the amplitude calibration is called the *a priori* amplitude calibration. An improvement to this calibration can be made by observing a strong non-varying source and performing amplitude self-calibration to get the corrections.

3.4.2 Fringe Fitting

The model used by the correlator to account for the geometric time delays and the fringe rates has errors, for instance, the antenna position, the clocks at different antennas and errors in the Earth and atmospheric model etc. These errors cause sloping phases and phase offsets across the frequency bands. The error in the interferometer phase expanded to a first order is given by,

$$\Delta\phi_{t,\nu} = \phi_0 + \frac{\partial\phi}{\partial\nu}\Delta\nu + \frac{\partial\phi}{\partial t}\Delta t \; . \tag{3.16}$$

The first, second and third terms are the phase error, the delay residual and the rate residual, respectively. *Fringe fitting* estimates these phase delays and phase rate residuals, and corrects for the same in the data. The data are divided into a two dimensional array of time and frequency. By taking the Fourier transform, the peak of the function is located in the delay and rate domain which gives the lowest residuals. Fringe fitting on weak sources is difficult since the signal-to-noise ratio (SNR) is not high enough for identifying the true peak. After fringe fitting, the data can be averaged over time and frequency to improve SNR and to reduce the size of the data at the expense of bandwidth/time smearing.

When the delays and rates are determined for each baseline and then obtained for the antennas, it is called *Baseline-based fringe fitting*. Here, the source has to be detected on all baselines or else the baselines with only a weak or non-detection will not be calibrated. Also, this method may not always determine phases, delays and rates that will satisfy the closure relationship (see section 3.4.3)

Global fringe fitting is used to generate a solution for antenna-based delay and fringe rate parameters simultaneously. Using a reference antenna, solutions for antennas are found directly as opposed to baseline based fringe fitting. The increased sensitivity of this method allows fringe fitting on weaker sources. Fig. 3.7 shows an example of global fringe fitting.

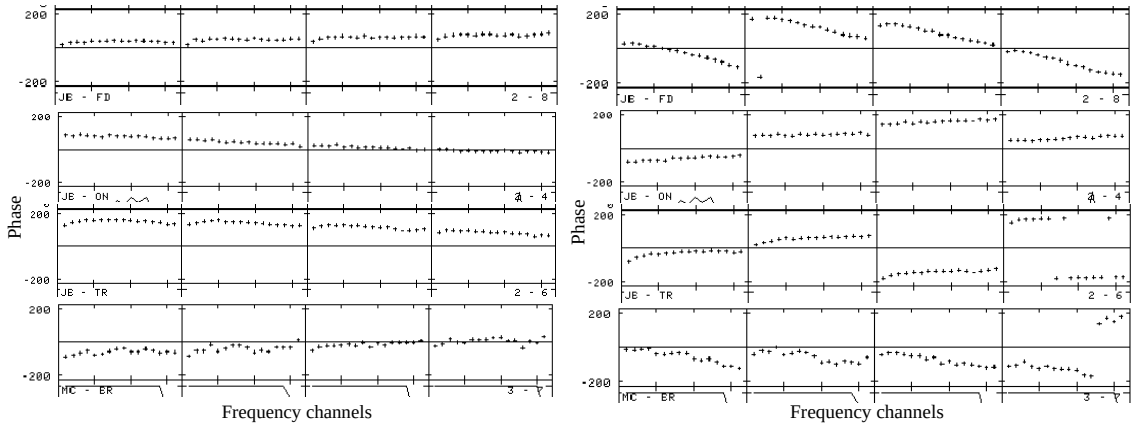

Figure 3.7: Left: Phase offsets across the frequency channels on various baselines **before** fringe fitting. Right: Aligned phases **after** fringe fitting for the same baselines. Note that the scatter in the phases has also reduced after fringe fitting.

3.4.3 Closure Phase and Amplitude

In the late 1960s, Roger Jennison realized that the sum of the visibility phases around a closed loop of baselines would be free of antenna based errors since they get canceled. Such a sum is called *closure phase*. The closure phase for a triangle of three baselines can be written as

$$\phi_{ijk} = \phi_{ij} + \phi_{jk} + \phi_{ki} \; . \tag{3.17}$$

Similarly, *closure amplitude* for a closed loop of four antennas is the ratio of the visibility amplitudes needed to cancel the antenna based gains,

$$A_{ijkl} = \frac{|\widetilde{\mathcal{V}}_{ij}|}{|\widetilde{\mathcal{V}}_{ik}|} \frac{|\widetilde{\mathcal{V}}_{kl}|}{|\widetilde{\mathcal{V}}_{jl}|} \; . \tag{3.18}$$

Both of these quantities are good observables, but were not enough to constrain the unknown visibilities and could not be used for imaging till the 1970s. Later, an iterative approach was developed by Readhead & Wilkinson (1978) (henceforth, RW) to use the closure phases and *a priori* information of the source for making maps which came to be known as *Hybrid mapping*. An initial model of the source is used to determine the phases on baselines with the help of the observed closure phase. A new model is made from the observed visibility amplitudes and the above determined phases. While mapping the source for the new model, a narrow window is chosen around the source in order to reject most of the surrounding regions which is mostly noise. The steps are repeated until a satisfactory map of the source is obtained.

3.4.4 Self-calibration

This method, also like hybrid mapping, consists of two parts. a) Determining the antenna based complex gain corrections i.e dealing with complex visibilities[3] and b) determin-

[3]The essential difference between Self-calibration and Hybrid mapping is that the former method explicitly solves for the antenna based corrections to remove them and the latter uses the closure quantities

ing the source brightness distribution which is done using the method, which is described in section 3.5.

Self-calibration is based on the principles of the RW method except that the antenna gains are allowed to vary in the process of determining the source brightness making the quality of the images better. Basically, the following equation is minimized by varying the antenna gains g_i and g_j, and the model visibilities $\widehat{\mathcal{V}}$,

$$\mathcal{S} = \sum_k \sum_{i,ji\neq j} w_{ij}(t_k)|\widetilde{\mathcal{V}}_{ij}(t_k) - g_{ij}(t)g^*_{ij}(t)\widehat{\mathcal{V}}_{ij}(t_k)|^2 \ . \tag{3.19}$$

Here, $w_{ij}(t_k)$ are the weights that can alter the way self-calibration proceeds. These weights are generally determined from the errors in the observed visibilities. A standard approach is to choose simply the reciprocal of the variance of the measured visibilities called *natural weighting*. This down weights the data points which have poor or no measurements and increases the sensitivity of the map.

Fig. 3.8 depicts the process of self-calibration. Here, a good starting model of the source is required. This model is subtracted from the observed visibilities in order to solve for the complex antenna gains. These solutions are used to determine corrected visibilities

$$\mathcal{V}_{ij,corr}(t) = \frac{\widetilde{\mathcal{V}}_{ij}(t)}{g_{ij}(t)g^*_{ij}(t)} \ . \tag{3.20}$$

Now, from the corrected visibilities a new model of the source is made which can be used for the next iteration. These steps are repeated until a satisfactory brightness distribution map is obtained. Since the model tries to fit the data exactly, self-calibration works well if the source is simple like a point source (or a source not much different from the model) and if the data has a good SNR.

3.5 Imaging

From the visibility function, the true brightness map can be obtained by taking a Fourier transform,

$$I(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{V}(u,v)e^{-2\pi i(ux+vy)}du\,dv \ . \tag{3.21}$$

This can be carried out using the Direct Fourier Transform (DFT) or the Fast Fourier Transform (FFT) methods. In either case, the visibility plane needs to be gridded to make the transform easier. Since the visibility plane is not regularly sampled, the visibilities may not always lie exactly on the grid points hence, some sort of interpolation is done. First, a smoothing function C is applied to the observed visibility function and then it is re-sampled only at regularly spaced intervals using the re-sampling operator R (see Zensus et al. (1995)). For a large number of grid points, FFT is generally faster than DFT since the number of operations it needs to perform is far less than that for DFT. Eq. 3.6 shows that the sampling function is weighted by $w_k = R_k T_k D_k$ and is non-zero only at points where the visibilities are measured. Here the coefficients R_k are associated with the reliability of each data point. It depends upon the integration time and bandwidth of the observation

which cancel these very antenna based complex gains

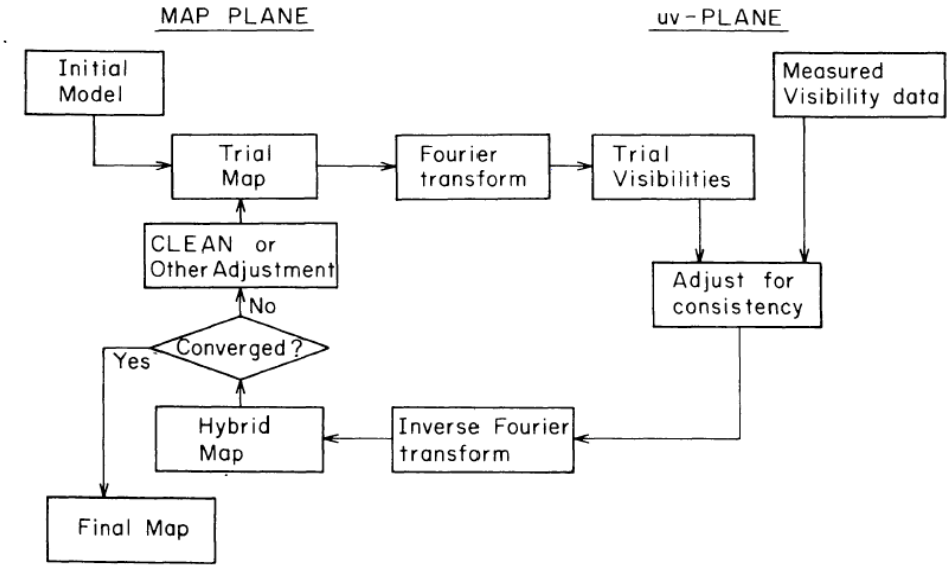

Figure 3.8: A flowchart summarizing self-calibration (Pearson & Readhead 1984).

and system temperature of the antenna. T_k is the tapering function (typically a Gaussian) and is generally used for downweighting the data at higher spatial frequencies. When the higher spatial frequencies are sparsely sampled and mostly consist of noise, tapering at such times can improve the quality of imaging. D_k is the density weighting function and as the name suggests is used to emphasize different regions in the visibility plane thereby optimizing the output in the desired direction. For instance to get high resolution, D_k can be set to the inverse of the density of the visibility points which down-weights the small spatial frequencies. Applying the convolution theorem to Eq. 3.8 gets Eq. 3.9,

$$I^d(x,y) = B_d(x,y) * I(x,y) \; . \tag{3.22}$$

It can be seen that $I^d(x,y)$, the dirty map is a convolution of the dirty beam $B_d(x,y)$ which is a Fourier transform of sampling function *S(u,v)* and true map *I(x,y)* which is a Fourier transform of the true visibilities. Here the additive noise term is ignored. The solution of convolution given by Eq. 3.22 is called the *principal solution* which is the dirty map $I^d(x,y)$. Now, suppose there is another brightness function Z which corresponds to unmeasured visibilities in the *uv*-plane then $B_d * Z = 0$. Therefore, any $I + \alpha Z$ will be a solution where α is a multiplicative constant and Z is referred to as the *invisible distribution*. This implies there is no unique solution to the linear deconvolution problem and hence, non-linear methods are needed to find an optimum solution.

The two widely used non-linear methods for deconvolution are and Maximum Entropy Method (MEM). MEM reconstructs the surface brightness from the cross-correlation of the signals by model fitting. The model predicted brightness distribution (B') is fitted to the *uv*-data by maximizing a function f(B') called the *entropy* of the brightness distribution. The entropy function is defined such that it has a unique solution when maximized. A thorough analysis of MEM and comparison with is given in Nityananda & Narayan (1982). Both the methods have qualitative and quantitative

differences in the reconstructed brightness distributions. If the source structure is point like, is preferred computationally whereas MEM gives faster and better results for extended sources. However, careful data processing is needed for MEM to work. The method is implemented in the data reduction of both lenses (except see sect. 5.3.1) hence a step-by-step description of the method is given below.

3.5.1 Deconvolution: CLEAN Algorithm

This iterative algorithm by Högbom (1974) uses the *a priori* information of the dirty beam to discard the side-lobes and retain the real components in the dirty map. For instance it is not likely that the sky brightness distribution shows the same complex features as the side-lobes of the dirty beam. Such constraints help to limit the number of plausible models that will fit the observed features in the dirty map.

The steps performed in are,

1. The flux density and position of the intensity peak in the dirty map are determined
2. A factor γ ($\leq$ 1) called the loop gain times the peak strength times the beam is subtracted from the peak position of the dirty map. Typically, $\gamma \sim 1\%-5\%$ of the peak
3. This peak position and flux of the so called component is stored
4. Steps 1 to 3 are repeated until no flux density peak higher than a threshold level (for example close to the rms) is found. The clean component subtracted dirty map is now referred to as the residual map
5. The components are convolved with a beam (generally an elliptical Gaussian beam with same Full-Width at Half-Maximum (FWHM) as that of the central lobe of the dirty beam)
6. The convolved components are added to the residual map to create a smoothed map

Further development to this algorithm is the Clark algorithm which unlike works in both the image and *uv*-plane. It consists of two parts. a) Finding the components by accepting any point in the dirty map (with an intensity above a threshold fraction of the image peak) greater than the highest side-lobe of the dirty beam. b) Fourier transforming the components altogether to the visibility plane using FFT, convolving with the sampling function, transforming back to the image plane to subtract them from the dirty map. Cotton & Schwab developed it further by altering the second part of the Clark algorithm. Here, the FT of the components are simply subtracted from the un-gridded visibility data which eliminates any errors that result from the process of gridding. It has been implemented in 's[4] Astronomical Image Processing Software ( ) used for data reduction of the lens systems in this thesis.

[4]NRAO-National Radio Astronomy Observatory

4 Luminous substructure in the lens system MG 2016+112

Numerical simulations of hierarchical structure formation for a CDM cosmology have predicted a large population of sub-halos within the main halos of galaxies (e.g., Moore et al. 1999). However, the observed satellite mass fraction of the Milky Way and other nearby galaxies, have defied these predictions by an factor of 10-100 (e.g., Klypin et al. 1999). It could be that the sub-halos do exist but are unobserved since they are not luminous. For example, many satellites of the Milky Way were discovered recently which has made the problem of missing satellites less severe (Martin et al. 2007; Simon & Geha 2007). A tool to test the existence of the missing sub-halo fraction would be through gravitational lensing which is sensitive to matter irrespective of whether it is dark or luminous.

With the increasing popularity of the applications of gravitational lensing, more and more lens systems were discovered. Although most of the lens systems could be well fitted with a smooth model of the mass distribution, the flux ratios are often very poorly fitted, for instance, MG 0414+0534 (Lawrence et al. 1995), B1422+231 (Patnaik et al. 1992), B2045+265 (Fassnacht et al. 1999). This is the problem of the anomalous flux ratios. Mao & Schneider (1998) noticed that the discrepant flux ratios of the lensed images of B1422+123 could be evidence of substructure (i.e. low mass halos $\sim 10^6$ $M_\odot$) for example, in the form of globular clusters in the vicinity of the main deflector. Subsequently, Metcalf & Madau (2001) carried out simulations to demonstrate that small-scale structures ($\lesssim 10^9$ $M_\odot$) indeed affect the magnifications of the lensed images without significantly distorting the image positions. Furthermore, Dalal & Kochanek (2002) devised a method to predict the abundance of the satellites of the lensing galaxy and applied this to a sample of quadruple lens systems. Their predictions of the satellite mass fraction ($0.6 \lesssim f_{sat} \lesssim 7$ per cent) were found to be consistent with the CDM predictions.

It was soon suspected that substructure could also produce astrometric perturbations in the lensed images at an observable level. However, very few such cases are known because the background source should have fine structure at milli-arcsecond resolution (this is only possible to detect in the radio). Secondly, as shown by Metcalf & Madau (2001), the astrometric perturbations ($\sim$10 mas) in the lensed images can be detected for substructure with a mass of $\gtrsim 10^8$ $M_\odot$ and which has a line-of-sight alignment with one of the image features. For example, the milli-arcsecond scale structure found in the lensed images of CLASS B0128+437 could not be fitted with a smooth mass model or a model with higher order multipoles of the lens potential (see Biggs et al. 2004). A second (and the most extreme) case of an astrometric anomaly is that of one of the first lenses discovered, MG 2016+112 which is investigated in this chapter.

The gravitational lens system MG 2016+112 has intrigued astronomers since its discovery. It showed a puzzling three image configuration (A, B and C) when observed in the radio. Optical/infrared imaging and spectroscopy confirmed A and B as lensed images. However, the third component C eluded any of the following classification a) a lensed image b) a lensing galaxy at the same redshift as that of the main lensing galaxy or c) a lensing galaxy at a much higher redshift (i.e. a second lens plane). Subsequent high-resolution radio observations revealed component C to be consisting of *a pair* of merging images straddling the critical curve. Such images are expected to be mirror symmetric and have equal magnification. Hence, MG 2016+112 came to be known as a quadruple image lens system.

Interestingly, asymmetry in the positions of the fine structure of the merging mirror images was found, violating the prediction from gravitational lensing theory. Any violation observed in the flux densities of the lensed images could be a result of intrinsic source variability or propagation effects like absorption or scattering. An anomaly in the astrometry of the lensed images however, has to be of gravitational origin. Such cases are much stronger evidence of CDM substructure than the anomalous flux ratio cases since no other known phenomenon can introduce astrometric asymmetry in the lensed images. The mass model results have claimed that MG 2016+112 has luminous substructure and the problem of the astrometric asymmetry in the pair of merging images can be resolved (Kochanek private communication, see also Kochanek et al. 2004; Chen et al. 2007).

The work presented here is divided into the following sections. In section 4.1, the background on MG 2016+112 is described with the help of earlier work on this lens system. In section 4.2, new multi-frequency high-resolution observations with VLBI are presented to study the spectra of the fine structure resolved in all of the lensed images. In section 4.3, the new observational constraints are used to make a better mass model and revisit the substructure problem. The discussion of the results and comparisons with previous work are described in section 4.4. Section 4.5 summarizes and concludes the results of the new observations and the mass models of the lens system.

4.1 Background

Like the blind men groping an elephant from various sides and drawing different conclusions with every further exploration, astronomers grappled to solve the mystery of the lens system MG 2016+112 with every new multi-wavelength observation. In this section, an up-to-date viewpoint of this system is presented.

4.1.1 The Lensed Images A and B

MG 2016+112 was the first gravitational lens system discovered with a systematic search for lenses. It was found in the MIT-Green Bank 6-cm survey (1981; Bennett et al. 1986) which was made with the 91 m transit telescope of the NRAO. Three compact components (A, B and C), which form nearly a right-angled triangle, were found with the VLA (see Fig. 4.1; Lawrence et al. 1984). The separation between images A and B is about 3.4 arcsec. Component C lies 2 arcsec away to the south-east of B and is the strongest of all. The overall integrated spectrum of this lens system is that of a Gigahertz Peaked

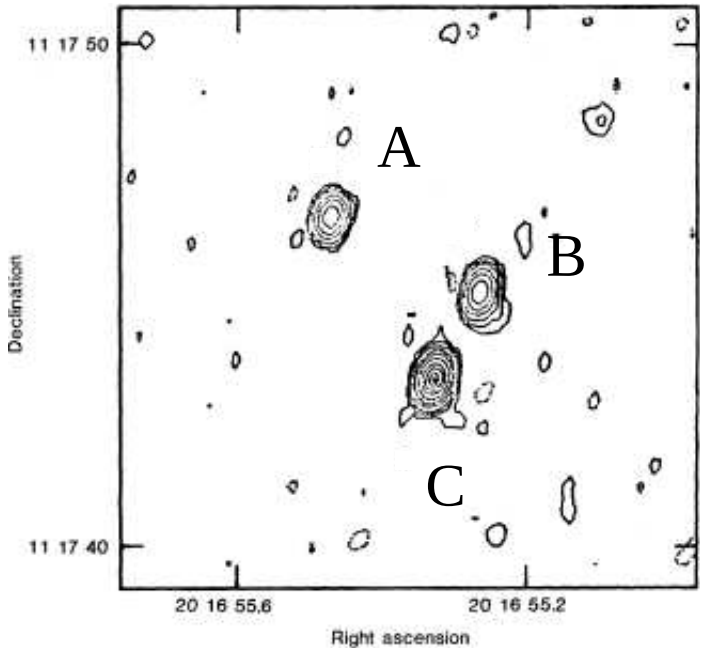

Figure 4.1: MG 2016+112 discovered in the radio at 5 GHz using the VLA reveals three unresolved components A, B and C (Lawrence et al. 1984).

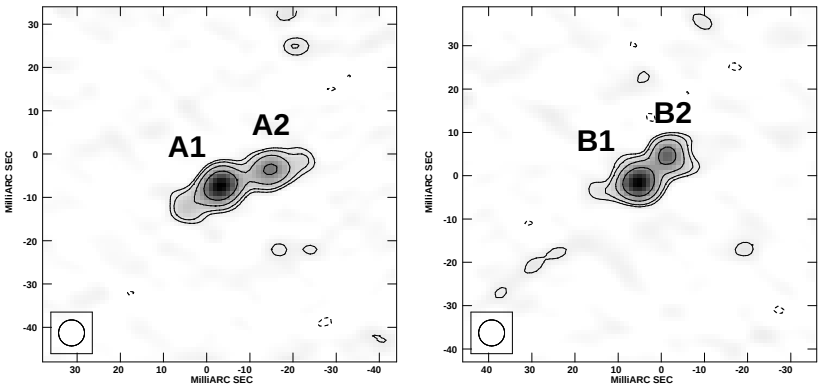

Figure 4.2: Each lensed image A and B is resolved into two components with the 5 GHz European VLBI Network (EVN) observations (Koopmans et al. 2002b).

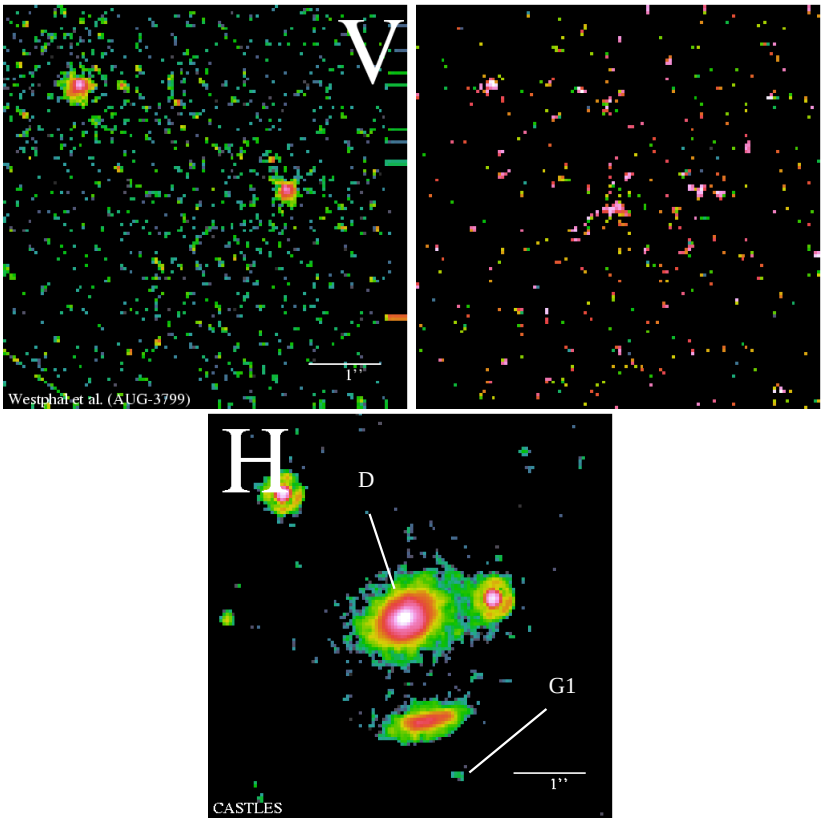

Figure 4.3: *HST* images of MG 2016+112 in *V, I* and *H* bands. Components A and B are visible in all of the three bands. Although C is detected in all bands, clear arc-like emission is seen only in the near infrared (*H* band). In the other bands C is fainter by a magnitude compared to A and B. The lensing galaxy D is visible only in the infrared and appears elliptical. Courtesy: CfA-Arizona Space Telescope LEns Survey (CASTLES; http://cfa-www.harvard.edu/castles/Individual/MG2016.html)

Spectrum source. Component B showed a radio spectrum similar to component A and the flux density ratio is B:A∼1, whereas component C has a much flatter spectrum. Hence, A and B were found to be consistent with being lensed images while the nature of C was not understood. The first 18 cm VLBI observations suggested that there was more than one component in images A and B (Heflin et al. 1991). These were clearly resolved into two components in A (A1 and A2) and B (B1 and B2) (see Fig. 4.2; Koopmans et al. 2002b).

The optical counterparts of A and B showed compact emission with an apparent magnitude of 22.5 in the photometric *r* band (see Fig. 4.3 for the *HST* image). Optical spectroscopy of A and B showed only narrow emission lines (Lyα, N V, Si IV−O IV, C IV and He II) at a redshift of $z = 3.273$ (Lawrence et al. 1984). This further confirmed that A and B are the lensed images of the same background source. The optical and radio flux density ratios of the components are found to be similar as expected. Furthermore, Elíasdóttir et al. (2006) recently showed that the optical emission of images A and B do not suffer from extinction from dust in the lensing galaxy.

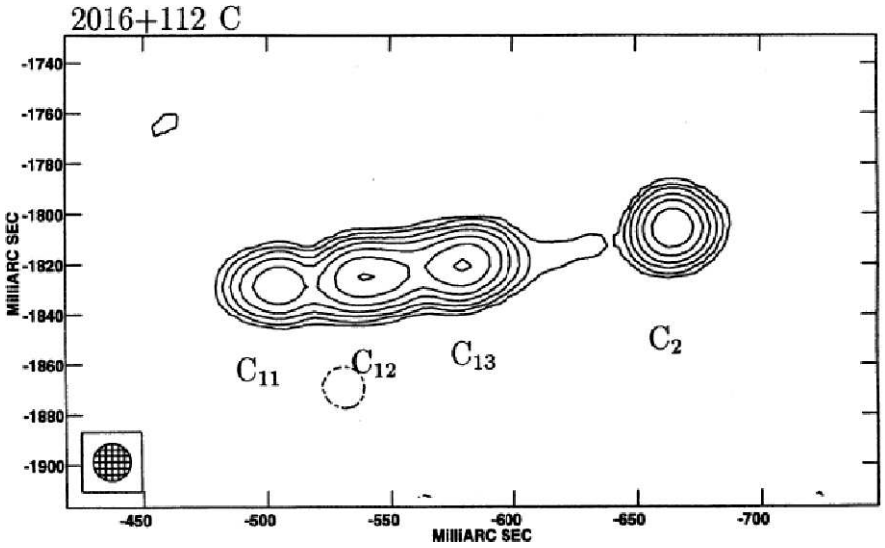

Figure 4.4: Region C was clearly resolved into four compact components in EVN 18 cm observations (Garrett et al. 1996).

4.1.2 The Lens

A giant red galaxy D was found in the optical lying close to the centroid of A, B and C (Schneider et al. 1985, see Fig. 4.3 which shows the *HST H* band image). Based on the Ca II absorption lines found in the spectrum of D, Schneider et al. (1986) deduced a redshift of 1.01. The *HST V*(F555W) and *H* band (F160W) magnitudes are 25.12±0.93 and 18.46±0.09, respectively. The axial ratio is found to be 0.57±0.01 with a position angle of 59±2 deg. The stellar velocity dispersion of D is 328±32 km s^{-1} (Koopmans & Treu 2002a). From the surface brightness profile and metallicity of D, it was interpreted as a massive elliptical galaxy consisting of old and metal-rich stars. This radio-quiet massive galaxy is by far the most distant galaxy-scale lens observed.

4.1.3 Region C

The third radio component (C) in MG 2016+112 deserves special attention and hence it is summarized here. Since its discovery, region C has eluded any simple interpretation. It was a long standing belief that C or at least a part of it was a lensing galaxy at a different or the same redshift as D (see e.g., Lawrence et al. 1984; Lawrence et al. 1993; Nair & Garrett 1997). The detection of Ly-α emission at $z = 3.273$ in region C, called C′, was believed to be the third lensed image (Schneider et al. 1986). Thus, region C was regarded as a composite structure which included a lensing galaxy C and a third lensed image C′.

Region C was found to be faint and diffuse in the optical, with an apparent magnitude of 23. The *K* band imaging presented in Lawrence et al. (1993) showed extremely red emission from component C and a photometric estimate of its redshift suggested that it is a galaxy at a much higher redshift than D. However, the red arc-like emission resolved by deep *V* and *I* band Keck images could be better explained if associated with the host galaxy of the lensed quasar (e.g., see the *HST H* band image in Fig. 4.3; Benítez et al.

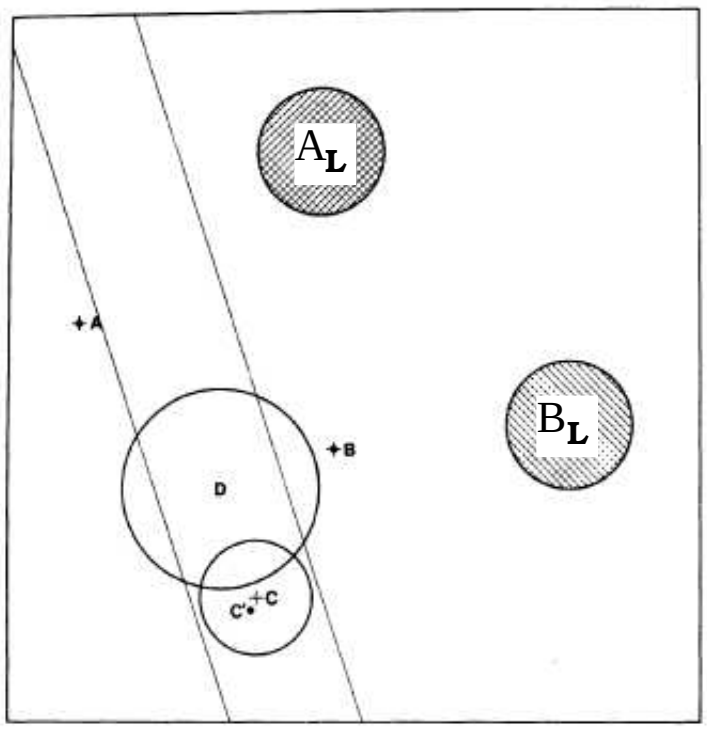

Figure 4.5: The hatched circles A1 and B1 represent the ionized clouds associated with the background quasar, responsible for strong Ly$-\alpha$ emission lines. The crosses mark the radio positions, the dots mark the point-like sources of narrow emission lines from A, B and C′. The light circles indicate extended optical emission (Schneider et al. 1986).

1999).

5 GHz MERLIN observations first resolved component C into two features, an east-west extended C_1 and a compact C_2 (Garrett et al. 1994a). Over the next two years using the better resolution of the European VLBI Network (EVN), the extended feature C1 was resolved into three more components (C_{11}, C_{12} and C_{13}) approximately along on east-west direction (see Fig. 4.4; Garrett et al. 1996). Although these were believed to be three lensed images, their identity was not certain, while C_2, still a compact component, was likely to be a third lensed image due to its association with the source of Ly-α emission, C′.

Yamada et al. (2001) studied the optical spectrum of region C and found Lyα, N V, C IV, He II and C III] emission lines redshifted to 3.273, as found in images A and B. No contamination from an object at a different redshift was found in the spectrum of C. However, the emission-line flux ratios of B and C were not in agreement. Nevertheless, region C was believed to be consistent with being a lensed image of part of the background quasar, unlike B which is the lensed image of the whole quasar. The differences in the emission line ratios could be reconciled if the ionization parameters[1] are different.

[1]The ionization parameter is an indicator of the number of ionizing photons available per atom in a given medium.

4.1.4 The Lensed Quasar

In addition to the complex morphology of the lensed images, the nature of the background source was another puzzle. Since only narrow emission lines are observed from the background source, Yamada et al. (2001) concluded that it is an obscured AGN. The radio power of this source is found to be too weak to be a radio-loud quasar at a redshift of 3.273. Moreover, images A and B showed compact emission in the radio and optical. This is typical of quasars which dominate the emission from their host galaxy whereas the emission from galaxies is extended. Hence, the background source is referred to as an obscured radio-quiet quasar. This would classify the source as one of the highest redshift type II quasars known. Time delay monitoring of the lensed quasar was undertaken with the VLA for more than a decade. No strong evidence of intrinsic variability was found, and hence no time delay could be determined for this lens system (Haarsma et al. 2000).

Two low surface brightness emission line regions (A_1 and B_1) were detected about 3 arcsec north-west and west of A and B respectively (see Fig. 4.5; Schneider et al. 1986). To avoid any confusion, these are labeled as A_L and B_L throughout. No radio counterparts or any continuum emission was found at these positions. Schneider et al. (1987) found that A_L and B_L also have a redshift of $z = 3.273$. Hence, the emission was not from the foreground. Assuming A_L and B_L are double images like A and B, then the emission lines found in A_L should have been detected in B_L. The Ly-α emission line was detected in A_L but only marginally detected in B_L. The difference in their spectra suggested that A_L and B_L are not lensed images of the same region of a background source at a redshift of 3.273.

4.1.5 The Lens Environment

The properties of the lens galaxy D were found to be similar to a Brightest Cluster Galaxy (BCG) at a lower redshift. Therefore, the field was observed in the X-ray to look for a cluster around the giant elliptical D. The associated cluster was regarded as dark since no optical/infrared emission was found from the cluster (e.g., Schneider et al. 1985). Observations with the Advanced Satellite for Cosmology and Astrophysics (ASCA) indicated diffuse X-ray emission from the lens system associated with the alleged dark cluster (Hattori et al. 1997). After the launch of *Chandra* in 1999, Chartas et al. (2001) observed the lens system with better resolution. The X-ray emission was found to be from the lensed images (A and B) of the background quasar and discrete sources in the lens field (see Fig. 4.6). Nevertheless, a 3 σ upper limit on the 2-10 keV luminosity of the cluster was quoted as 1.7×10^{44} erg s^{-1} by Chartas et al. (2001). Furthermore, an upper limit on the mass from Chartas et al. (2001) and luminosities of galaxies measured by Benítez et al. (1999, see below) were used to derive an upper limit on the mass-to-light ratio of the cluster. The limit was placed within a radius of 800 h_{50}^{-1} kpc of $M/L_V < 190 h_{50} (M/L_V)_\odot$. It is consistent with being an average to low mass cluster in comparison with low redshift massive clusters and studies of several groups/clusters (Hoekstra et al. 2002; Girardi et al. 2002). The absence of diffuse emission in the *Chandra* observations might indicate that the cluster of galaxies is not virialized yet. Furthermore, deep optical and infrared observations by Benítez et al. (1999) found a red sequence of galaxies and Soucail et al. (2001) spectroscopically confirmed an over-density of six galaxies at the redshift of the lens. It was noted that none of the X-ray point source positions of Chartas et al. (2001)

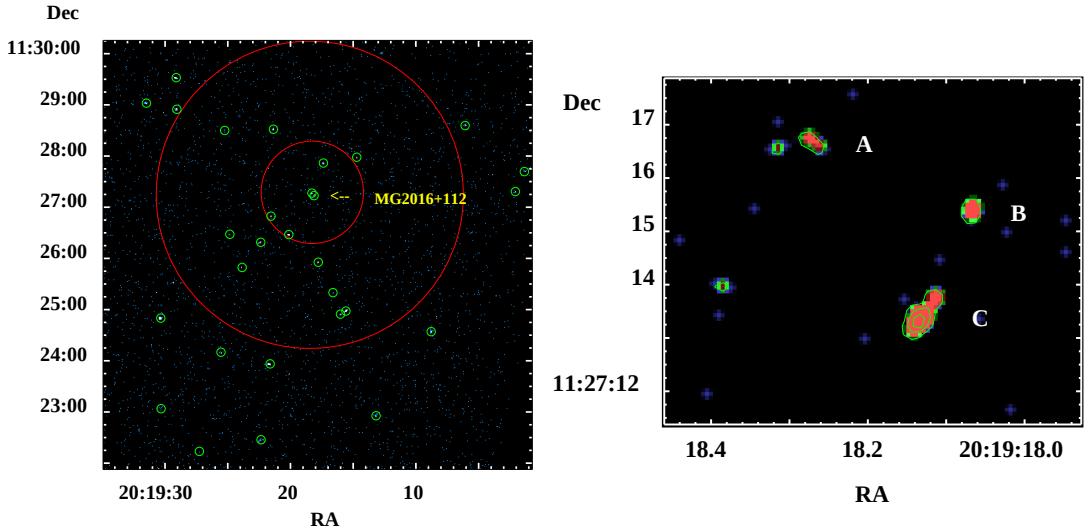

Figure 4.6: *Left*: An 8'×8' field around MG 2016+112 in the X-ray shows several point sources but no clear diffuse emission from the cluster. The inner red circle (size ∼ 1′) and outer red circle (size ∼ 3′) were inspected with *ROSAT* HRI and *ASCA*, respectively. *Right*: Emission from the individual components A, B and C is weak but detected by the *Chandra* X-ray Observatory (Chartas et al. 2001).

matched with the optical positions of Soucail et al. (2001). Clowe et al. (2001) found an over-density of ten galaxies in the lens plane within ∼ 53″ of MG 2016+112 based on photometric estimates from Keck imaging. Another group of fainter objects about ∼ 6.3″ to the north-west of the lens was also found. Note that this nearly overlaps with the position of the Lyα emission region B_L. A weak lensing analysis of the field indicated a 3 σ signal about 64″ to the north-west of the lensing galaxy. This weak lensing signal was consistent with a singular isothermal sphere mass model with a velocity dispersion $\sigma = 970$ km s^{-1}. However, due to obscuration from the stars in the field of the peak of the weak lensing signal, it was not possible to detect any galaxies there with certainty.

In the *HST H* band image, three nearby galaxies are found G1, G2 and G3. The dwarf galaxy G1 to the south-west of MG 2016+112 (see Fig. 4.3) is known to have the same redshift as the primary lens D (Koopmans & Treu 2002a) whereas no redshifts are known for the two line-of-sight nearby galaxies G2 and G3.

4.1.6 History of Mass Models

Over a period of a quarter century several interesting models for this lens system were put forward. These models are summarised in this section.

Following the discovery of the lens system, Narasimha et al. (1984) presented two possible mass models a) a galaxy at the position of C and a cluster or b) a galaxy at the position of C and another galaxy at a different redshift. Based on further observational results a new model was made by Narasimha et al. (1987). The model consisted of two lensing galaxies C and D and a cluster at the same redshift. MG 2016+112 was described as a five image lens system (A, B, two images in C and one near D).

Since no cluster was detected in the optical, the model of Narasimha & Chitre (1989) included a dark matter halo instead of a cluster and the lens galaxies at D and C. The constraints were A, B, C′, A_L and B_L as the lensed images of the background source. The background source was assumed to have a finite size (i.e. not point-like) such that it lies on the caustic (see Fig. 4.7). The mass model predicted that A and B are the double images of the core of the background source whereas a pair of merging images with a high magnification in the tangential direction are formed at the position of C. Furthermore, regions A_L and B_L were shown to be singly imaged and correspond to different regions in the background source.

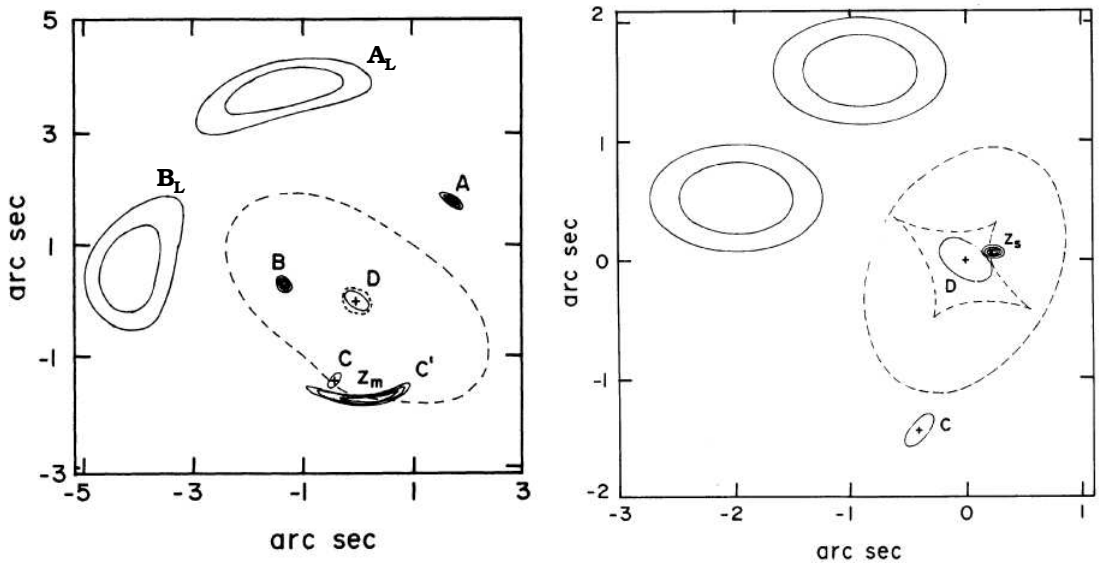

Figure 4.7: The image configuration in the image plane and the source configuration in the source plane both show the position of the lensing galaxies D and C. Z_m from the image plane maps back to Z_s in the source plane. The dotted lines are the critical curves in the image plane which map to the caustics in the source plane. The singly imaged A_L and B_L map to the two elliptical rings in the source plane. Note that the convention is North is up and east is right (Narasimha & Chitre 1989).

Langston et al. (1991) used a single screen with a single elliptical lens D in the mass model. All components (A, B and C) seen in radio (3.6 cm) and in *K* band data were associated with the background quasar and explained qualitatively by the model. Region C was predicted to show an east-west extended flat-spectrum structure and a central compact steep-spectrum component. Nair & Garrett (1997) discussed two models with two lenses, C and D, in two lensing planes and dark matter halo(s). However, note that these models are not determined from a fit to the observational data. Spectral analysis and/or time delay measurements would confirm either of the models.

Similar to the scenario proposed by Narasimha & Chitre (1989), a recent model by Koopmans et al. (2002b, hereafter K02) described MG 2016+112 as a quadruply imaged lens system. However, a single screen consisting of the primary lens D, two singular isothermal sphere (SIS) mass distributions and an external shear was considered instead

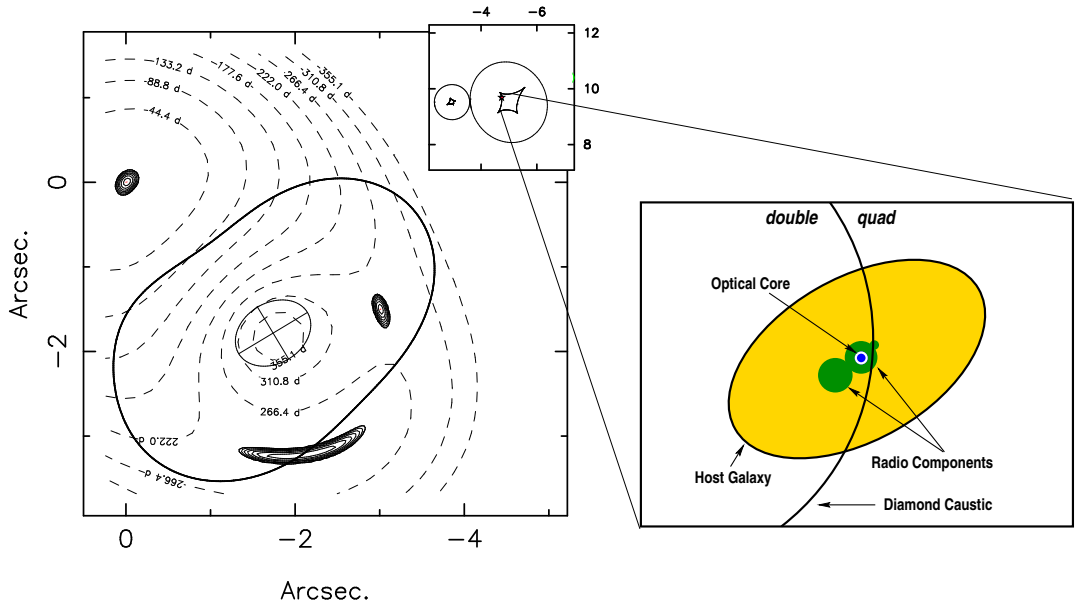

Figure 4.8: The dashed lines show the time-delay surface countours for MG 2016+112 in the image plane. The ellipse with a cross marks the primary lens D. The thick solid curve shows the critical curve in the image plane. The corresponding caustics in the source plane can be seen in the inset in the top-right corner. A star marks the position of the source in the caustic. On the right, the cartoon shows the source structure lying on the tangential caustic (Koopmans et al. 2002b). see text for details.

of multiple lens planes. According to the K02 scenario, the background source with a core-jet structure is situated on the tangential caustic, see Fig. 4.8. As known from lensing theory, the number of images changes by two as the source crosses the caustic (see section 1.2.5). The two compact components in the radio are doubly imaged which are visible in images A and B. The optical core is coincident with one of the two radio components and is also doubly imaged (see Fig. 4.3), whereas part of the radio counter-jet and the host galaxy to the right of the caustic are quadruply imaged. Thus, only this part of the counter-jet close to the caustic is highly magnified and constitutes the complete ~ 200 mas long region C in the radio. Images close to the critical curve are known to have high magnification but the K02 model predicts an unusually high magnification (~ 300) in region C. Similarly, the emission from the host galaxy is found in the infrared in the form of a tangentially magnified arc as expected in lensing.

An important point that needs to be discussed here is the asymmetry in the separations of the merging pair of components found in region C. This anomaly is believed to be due to a dwarf galaxy G1 to the south of C (Kochanek et al. 2004). Chen et al. (2007) have made a mass model which takes into account the satellite G1. The model suggested a velocity dispersion for the satellite of ~ 95 km s^{-1} at 95 per cent confidence.

4.2 New Observations in the Radio

Radio observations of MG 2016+112 are presented in this section. These include simultaneous MERLIN and global VLBI observations of MG 2016+112 at both 1.7 and 5 GHz, and with the High Sensitivity Array (HSA) at 8.4 GHz.

4.2.1 MERLIN Observations

The MERLIN array was used to observe MG 2016+112 simultaneously with the global VLBI observations at 1.7 and 5 GHz. The purpose was to have a double check on the calibration of the global experiment and to be aware of any loss of flux in the global VLBI observations. The observations were scheduled on 2002 February 25 at 1.7 GHz and on 2001 November 17 at 5 GHz. Both of the experiments had the same observational setup. The data were taken in a single IF of 15 MHz bandwidth and subdivided into 15 channels. The flux calibrator (3C 286) and the point source calibrator (2134+004) were observed for a few minutes (one-two scans). The observations nodded between the lens system (MG 2016+112) and the phase calibrator (B2029+121). Further observational details can be found in Table 4.1.

Table 4.1: Details of the observation scans

Type of experiment	Total observation time (hrs)	Integration time (s)	Scan length (min) Target	Phase calibrator
MERLIN				
1.7 GHz	14	8	3.7	1.7
5 GHz	14	8	3.7	0.8
VLBI				
1.7 GHz	17	2	4.5	2.4
5 GHz	17	1	4.5	1.5

Using the standard MERLIN data reduction program, the data were corrected for any non-closing errors and the amplitude calibration was performed. At 1.7 GHz, the amplitude on the Mk II telescope baselines was down on a few scans and on a few others the Lovell Telescope was off-source. These scans were deleted from the data. The 5 GHz data generally looked good. With the phase referencing technique, the fringe fitting solutions determined for B2029+121 were applied to calibrate the phase information on MG 2016+112. Several iterative runs of the task were performed to determine the phase and amplitude solutions. These solutions were interpolated using the task and were applied to the lens system. A phase referenced map was made, subsequently. Using the phase-referenced map as an initial model, phase self-calibration was performed on the lens system. The resulting map was used as a model for a new iteration of calibration of the data and the cycle was repeated until the final map showed no significant improvement.

Fig. 4.9 shows the maps of lensed images of MG 2016+112 at 1.7 and 5 GHz. The weighting of the *uv*-data can be chosen between natural and uniform to give either better

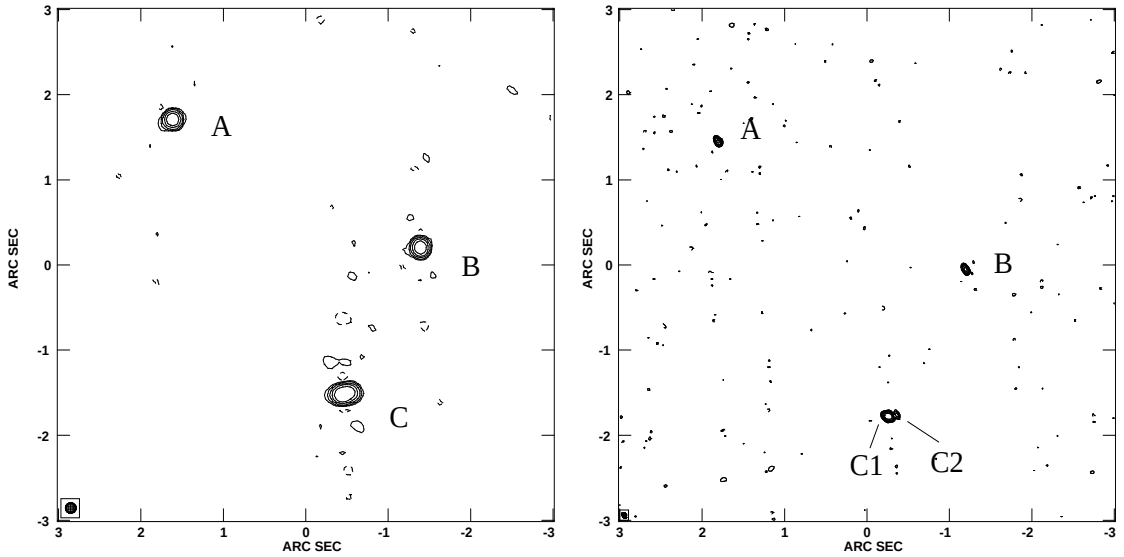

Figure 4.9: *Left*: The MERLIN image at 1.7 GHz of MG 2016+112 shows three lensed components. A and B are compact and C is slightly extended in the east-west direction. The map is restored with a beam of size 0.21×0.13 arcsec^2. *Right*: The higher resolution 5 GHz MERLIN image shows that components A and B are still compact but component C has resolved into a brighter extended component C1 and a compact component C2. The map is restored with a beam of size 70.7×43.5 mas^2. The noise in the maps are 0.27 mJy beam^{-1} at 1.7 GHz and 0.18 mJy beam^{-1} at 5 GHz. The contour levels are at (-3, 3, 6, 12, 24, 48) × the rms noise in the respective maps.

Table 4.2: MERLIN components of MG 2016+112 at 1.7 GHz. The '–' implies that the integrated flux density could not be determined.

Components	RA (mas)	Dec (mas)	S_{peak} mJy beam^{-1}	S_{total} mJy
A	0±0.5	0±0.5	34.6±0.3	34.9±0.5
B	−3005.8±0.5	−1503.9±0.5	34.8±0.3	35.6±0.5
C1a	−2045.4±0.9	−3246.2±0.9	13.4±0.3	–
C1b	−2093.8±1.4	−3221.5±1.4	42.9±0.3	62.8±0.6

Table 4.3: MERLIN components of MG 2016+112 at 5 GHz.

Components	RA (mas)	Dec (mas)	S_{peak} mJy beam^{-1}	S_{total} mJy
A	0.0±0.5	0.0±0.5	11.7±0.2	12.1±0.3
B	−3003.2±0.5	−1502.0±0.5	13.1±0.2	13.1±0.3
C1a	−2048.3±0.5	−3230.5±0.5	15.4±0.2	20.0±0.4
C1b	−2090.8±0.5	−3225.6±0.5	13.1±0.2	13.2±0.3
C2	−2171.5±1.5	−3212.8±1.5	3.1±0.2	3.0±0.3

resolution or sensitivity at the cost of the other. It is possible however, to optimize the weighting by setting the parameter to 0. Gaussian model fitting of the components in the image plane was done with the task . The results are presented in Tables 4.2 and 4.3. Lensed images A and B are found to be compact, and were fitted by a single Gaussian at both frequencies. The slightly extended component C was fitted with two Gaussian components at 1.7 GHz and with three at 5 GHz. Note that the components of C1 that were detected at 1.7 and 5 GHz may not be the same. These components are identified by comparing the model fitting of the high resolution VLBI images (see Section 4.2.2). The relative flux densities of components B, C1 (C1a+C1b) and C2 with respect to A for the MERLIN 5 GHz imaging are consistent with previous measurements by Garrett et al. (1994a). However, the absolute flux densities of the lensed images in the new observations are less by 23 per cent which might be due to errors in the amplitude calibration of either the previous or new data sets.

4.2.2 Global VLBI Observations

The earlier 1.7 GHz EVN observations revealed fine structure in region C. Also, the predictions of the spectra of the components in region C (Garrett et al. 1996) could be verified with a high resolution spectral analysis. This would also rule out some mass models which have different predictions for the spectra of the components in region C. Therefore, high resolution global VLBI observations were undertaken at 1.7 and 5 GHz on 2002 February 25 and 2001 November 17, respectively.

Since MG 2016+112 has a low flux density, phase referenced observations are vital in determining the phase corrections on all baselines. A strong source within 2 deg of the lens system (e.g., B2029+121 with a total flux density ~ 0.9 Jy), was used as the phase calibrator for these observations. Therefore, the antennas had to nod between the lens system and the phase calibrator regularly throughout the observing run. The scan lengths for each observation of the target (MG 2016+112) and the calibrator (B2029+121), the correlator integration time and the total time of the observations are listed in Table 4.1. The data were taken in four IFs at 1.7 GHz and in two IFs at 5 GHz. Each IF had a bandwidth of 8 MHz and was further divided into 16 channels.

The antennas used for the 1.7 GHz observations were Effelsberg (EB), Jodrell Bank (JB), Medicina (MC), Onsala (ON), Torun (TR) and the 10 antennas of the VLBA. In ad-

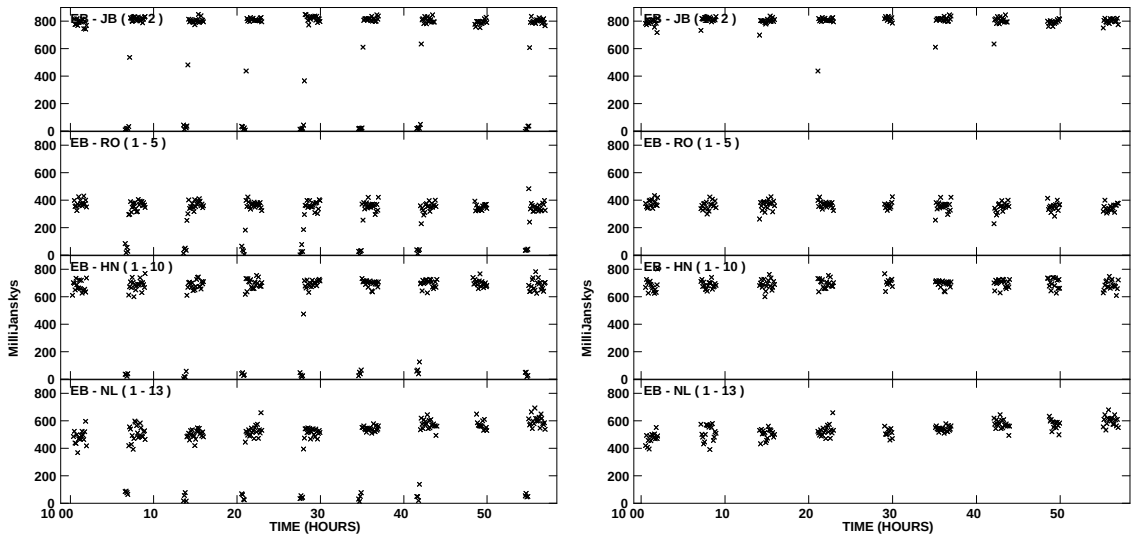

Figure 4.10: A few scans on four baselines before and after editing the data on the phase calibrator are shown. The first few seconds on every scan were deleted.

dition, the phased VLA (Y) and the 70 m dishes of Robledo and Goldstone were included for better sensitivity. Unfortunately, the data on the tapes were lost for Arecibo (AR) and Westerbork (WB) at the VLBA correlator. At 5 GHz, all of the VLBA antennas and EB, JB, MC, ON, TR, AR, Y and WB were used.

Editing

At 1.7 GHz the amplitudes of the data on the phase calibrator were edited prior to the calibration to remove any obvious bad scans or baselines with bad data. For example, all of the scans on all of the baselines were found to have bad data in the first few seconds for the phase calibrator B2029+121 (see Fig. 4.10). This was probably due to the antennas not being on-source at the beginning of the scans. It was difficult to deduce if this had also affected the lens system because it has low SNR as compared to the calibrator and complicated structure, due to which the data appears noise-like. Therefore, the task
was used to delete the first 0.25 mins of all of the scans on the calibrator.

Calibration

There were some errors introduced in the signal amplitude during the digitization of the measured signal. The correction factors for these errors can be determined by checking the mean value of the auto-correlated signal and adjusting these to unity. The cross-correlated data were corrected for these errors using the task in the multi-frequency observations presented here. The data were also corrected for the change in the parallactic angle with the task . This was done to account for the apparent change in the position angle of the source as the source moves across the sky for an antenna with an altitude-azimuth mount. After the data were corrected for the change in the parallactic angle, the *a priori* amplitude calibration was performed. To determine the system temperature of the

phased VLA during the observations, the amplitude level of the phase calibrator was set to the known values obtained from the list of calibrators at different frequencies. The task was used to generate the system temperature and the gain values for each antenna.

The data for the calibrator were fringe fitted on the longest scan to determine the phase, delay and rate solutions for a single IF initially, and then for all of the IFs combined, which were then applied to the whole data. Using the task , the large data set was divided into single source files to carry out self-calibration. The weights associated with the data were calibrated. Since averaging the data across the bandwidth can degrade the fidelity of the map features, the data channels were not averaged during splitting, in spite of the large number of visibilities. The phase calibrator was mapped by performing phase self-calibration and then both phase and amplitude calibration. The phase and amplitude corrections from the best calibrated map of B2029+121 were applied to the target MG 2016+112. Subsequently, the data on the lens system were phase self-calibrated with data integrated over a 3 and 2 min long solution interval for 1.7 and 5 GHz data sets, respectively.

Imaging

When mapping sources in a wide field (e.g., a few arcseconds wide with a resolution of sub-milliarcsecond), the target sources within the field can have a large spatial separation. In order to save the computing time of mapping a single wide field, multiple small fields centred on the target sources were mapped. Thus, three windows centred at the position of A, B and region C were ed while mapping MG 2016+112. To get the optimum combination of resolution and sensitivity, the weighting scheme in the maps is chosen between uniform and natural by setting the parameter to 0. At 1.7 GHz, the weighting was chosen to produce an optimal balance between resolution and sensitivity, whereas at 5 GHz the resolution was good enough, so only the sensitivity had to be up-weighted, hence, natural weighting was applied.

The lensed images A and B

Fig. 4.11 shows lensed images A and B, and the pair of merging images in C at 1.7 GHz. From the high resolution observations, images A and B both show a rich structure. The series of components in A and B are non-collinear. As known from gravitational lensing, such non-collinear structure in the lensed images can show the expected opposite parity. This is clearly demonstrated by the opposite curvature in the strings of components of A and B. In the earlier EVN 5 GHz observations (Koopmans et al. 2002b), images A and B showed two components each with some hint of structure (see Fig. 4.2). Since components A1-B1 and A2-B2 are compact and have high flux densities, these were detected in the earlier EVN 5 GHz observations.

In the new global VLBI 1.7 GHz data presented here, images A and B are found to have three and two new components, respectively. Model fitting of the images was carried out using Powell's minimization routine outside AIPS (kindly provided by S. More, see Appendix B) and the results are given in Table 4.4. All of the components are numbered in decreasing order of their total intensity. Image A is fitted with five Gaussian components and image B is fitted with only four Gaussian components. Since image B is the

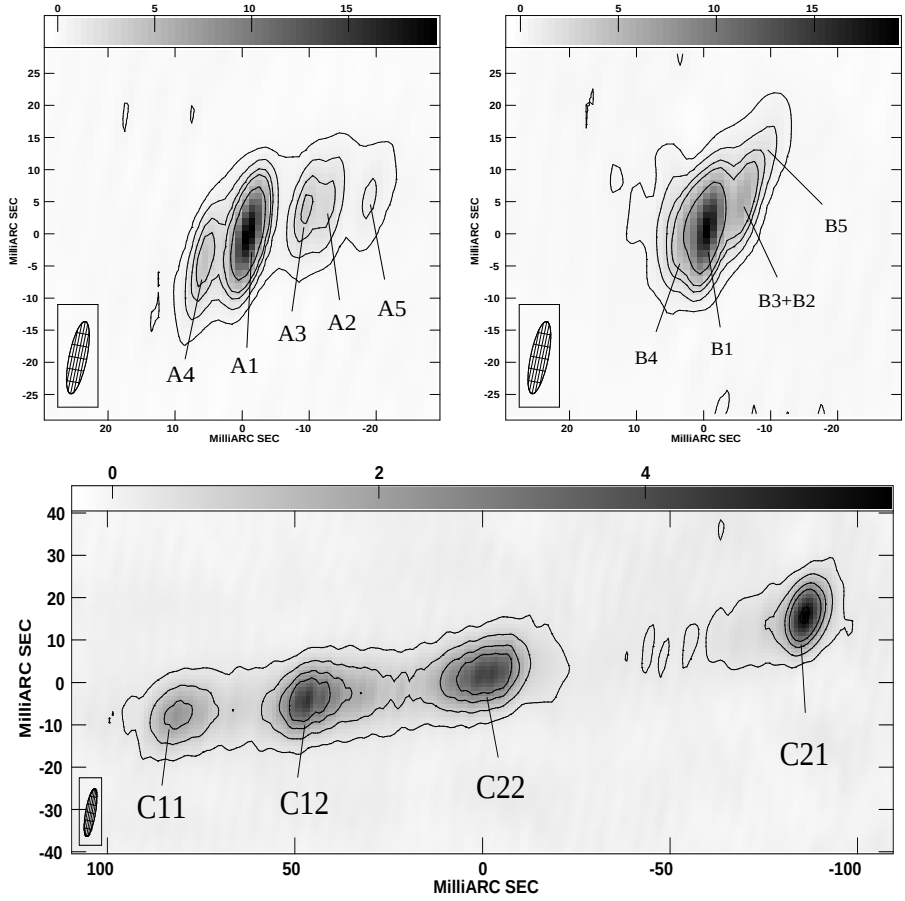

Figure 4.11: The global VLBI images of lensed images A and B, and region C at 1.7 GHz with a weighting between uniform and natural. The rms noise in the maps is 0.08 mJy beam^{-1} and the contours are (-3, 3, 6, 12, 24, 48) × the rms noise in the map. The size of the restoring beam is 11.1×2.6 mas^2 and the position angle is −10.58 deg.

counterpart of image A with a relative magnification ~ 1, image B is also expected to have five components. The component B3+B2 is identified as the composite of the unresolved components B3 and B2. Furthermore, if B3 has a higher peak flux density than B2 (like its counterpart A3 in image A), then the peak position of B3+B2 will be closer to the peak position of B3. It is also noted from the comparison of fitted components of images A and B that the fitted components of image B do not represent the counterparts of image A.

Fig. 4.12 shows the high resolution global VLBI images at 5 GHz. Image A clearly shows the five components which were not well-resolved at 1.7 GHz. Here, image B also shows the expected five-component structure. Component 2 in images A and B has a higher surface brightness than component 3, whereas at 1.7 GHz, component 3 has a higher surface brightness than component 2 in image A. The model fitting was again carried out with Powell's minimization routine. Images A and B are fitted with five components each. The results of fitting Gaussian model components are presented in Table 4.5. The errors at all of the frequencies were determined based on the principles described in Fomalont (1999, see Appendix C).

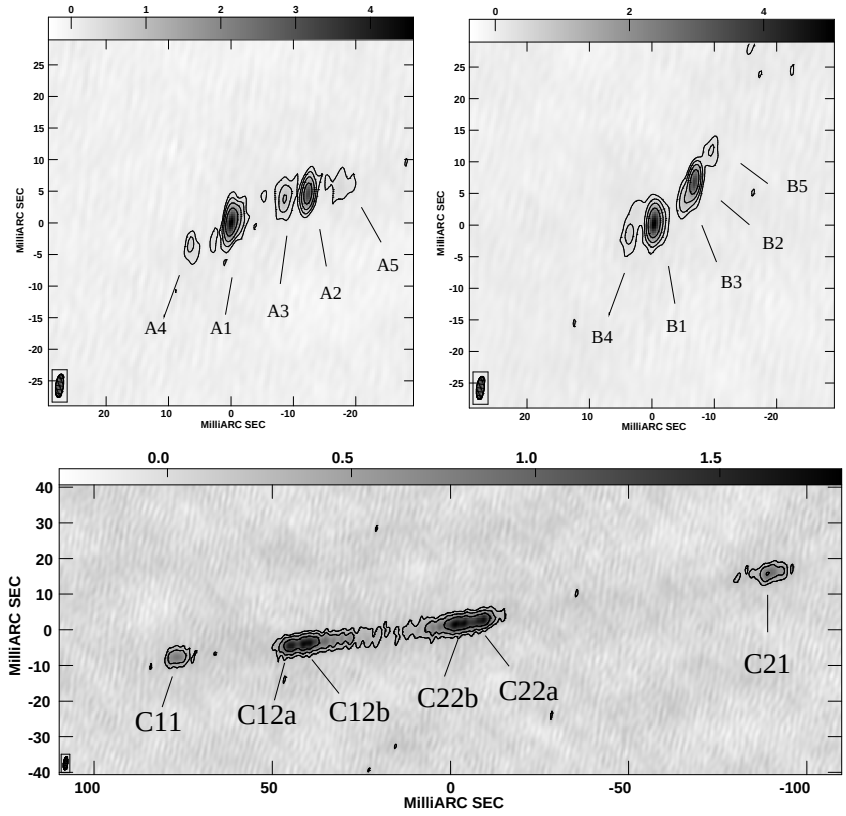

Figure 4.12: The global VLBI images of lensed images A and B, and region C at 5 GHz with natural weighting applied. The rms noise in the maps is 0.06 mJy beam^{-1} and the contours are (-4, 4, 8, 16, 32) times the rms noise in the map. The size of the restoring beam is 3.7×1.2 mas^2 and the position angle is −7.54 deg.

New Nomenclature for components of region C

The pair of merging images in region C are referred to as C1 (the east pair) and C2 (the west pair) (see Fig. 4.12 at 5 GHz). The components have been numbered ascendingly going inwards from the outside. The outer components are labelled as C11 and C21 for the east and west pair, respectively. The elongated components are labelled as C12 and C22, which are further resolved into several components. Each of these components are labelled as a, b, c etc. going inwards from the outside, for example, C12a, C12b, etc. in C1 and C22a, C22b, etc. in C2. Note that this labeling convention is different from that used in Garrett et al. (1996).

Region C

Region C was previously known to have four components and all of these were detected in the 1.7 GHz imaging (see Fig. 4.11). The outer two components on either side (C11 and C21) are fitted with two Gaussian model components each. However, due to the low resolution and low SNR, these are identified as a single component each, and their centroid (flux-density weighted) positions are reported here. The inner pair of elongated

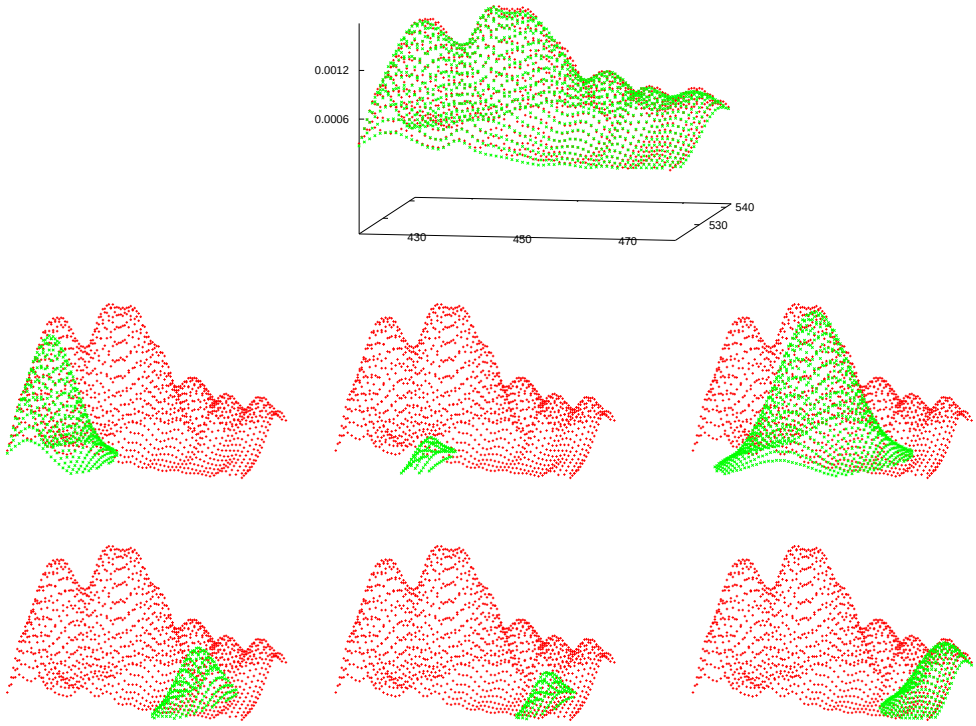

Figure 4.13: A multi-component Gaussian fit to the component C12 at 5 GHz in the image plane. Model fitting is done using Powell's minimization routine outside of AIPS (see Appendix B). The z−axes shows surface brightness in mJy beam $^{-1}$. The results of the fit are presented in Table 4.5.

components were each fitted by three Gaussian model components. The peak positions, the peak flux densities and the total intensities from the Gaussian model fitting are given in Table 4.4. Note that the components a, b and c found in component C12 may not be the counterparts of C22 for three reasons. Firstly, these components are not well-resolved. Secondly, if these are a pair of merging images straddling the critical curve, then the magnification gradient is rapidly changing on either sides of the critical curve. Ideally, it should change similarly on either sides. However, practically, this may not be observed. Thirdly, the ratio of component separations of the west pair C2 (C21-C22) and east pair C1 (C11-C12) is asymmetric indicating different stretching on either sides. It is likely that the dwarf G1 is distorting the expected mirror symmetry in the merging images. Hence, its contribution to the mass model will be tested in section 4.3

At 5 GHz in region C of Fig. 4.12, all of the four components are detected, and are further resolved in some cases. Therefore, centroid positions were determined for most of the components by fitting multiple Gaussian components. Here, component C11 is fitted with one Gaussian. C12 was modeled with six Gaussian components which are shown in Fig. 4.13. The surface brightness distribution from the image is plotted in red and the Gaussian model components are plotted in green. The six surface plots below show the individual model fits of 6 components to the observed surface brightness distribution. C22 was fitted with seven Gaussian components, similarly. The component C21 is extended, and hence, fitted with three unresolved components. Three centroid positions

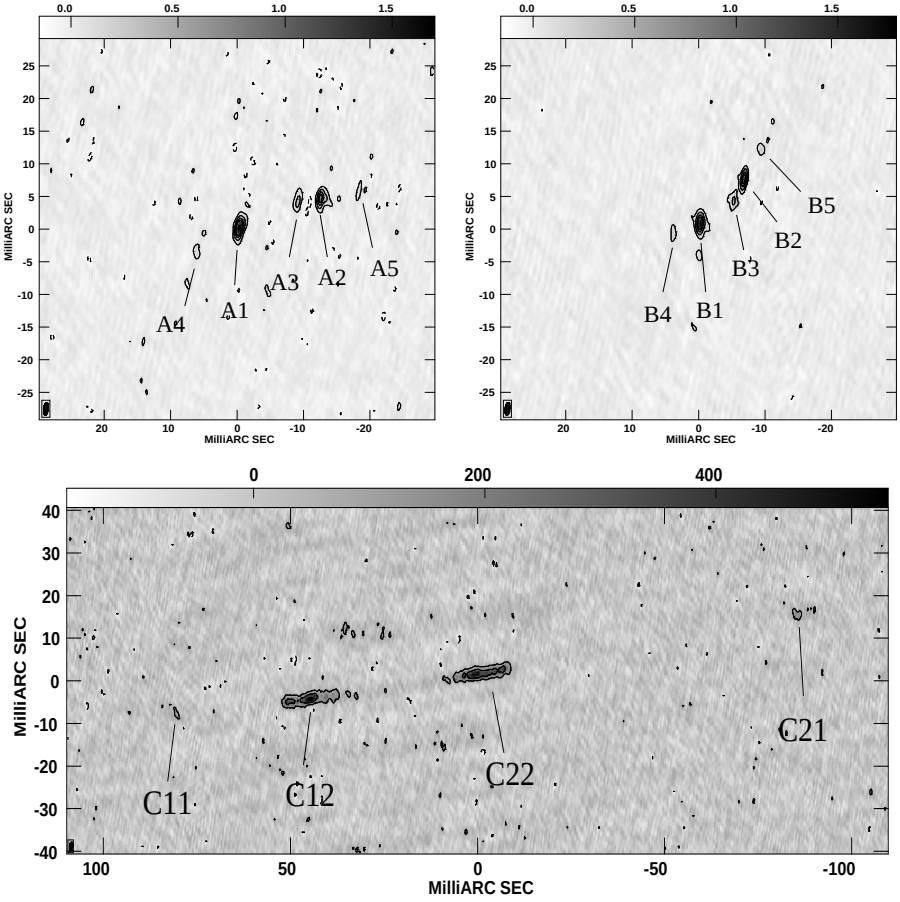

Figure 4.14: The HSA images of A, B and region C at 8.4 GHz with a weighting between uniform and natural. The rms noise in the maps is 33 μJy beam^{-1} and the contours are (−3, 3, 6, 12, 24, 48) times the rms noise in the map. The size of the restoring beam is 1.9×0.7 mas^2 and the position angle is −7.02 deg.

were determined for each of the elongated components C12 and C22, and one centroid position for C21. The peak positions, flux densities and total intensities for A and B are presented in Table 4.5. For region C, the centroid positions with 1 mas uncertainties, peak flux densities and the total flux densities are also given in Table 4.5.

4.2.3 High sensitivity Array (HSA) Observations

Further high resolution observations at 8.4 GHz were needed to independently confirm the series of components detected in the lensed images at 5 GHz, and to better determine the spectra of lensed images A and B. Since the background quasar has a steeply falling spectrum, observations at a higher frequency demanded increased sensitivity. Moreover, to carry out a spectral analysis of the finely resolved structure, high frequency and high resolution imaging was needed. Therefore, MG 2016+112 was observed with the HSA at 8.4 GHz. The HSA included the following large antennas: the 305 m-Arecibo (Ar), 100 m-Effelsberg (Eb), 100 m-Green Bank Telescope (GBT) and phased VLA (Y), in

Table 4.4: Peak positions, peak and total flux densities of components of A, B and C at 1.7 GHz for the global VLBI observations.

Components	RA (mas)	Dec (mas)	S_{peak} mJy beam^{-1}	S_{total} mJy
A1	0.00±0.02	0.00±0.02	19.7±1.6	24.3±2.5
A2	−12.2±0.1	4.5±0.1	2.4±0.5	2.8±0.7
A3	−8.6±0.1	3.8±0.1	3.5±0.6	4.5±0.9
A4	6.2±0.1	−3.8±0.1	4.1±0.6	6.5±1.2
A5	−18.0±0.3	5.5±0.3	1.1±0.3	1.7±0.6
B1	−3005.74±0.02	−1503.63±0.02	11.5±0.1	10.9±1.6
B3+B2	−3011.0±0.1	−1498.0±0.1	3.1±0.6	2.9±0.8
B4	−3004.5±0.1	−1502.6±0.1	8.5±1.0	21.9±2.8
B5	−3012.5±0.2	−1495.9±0.2	2.5±0.5	4.2±1.0
c11	−2013.6±0.2	−3233.3±0.2	2.1±0.4	11.6±2.3
c12a	−2045.4±0.1	−3229.9±0.1	4.1±0.6	15.8±2.4
c12b	−2053.2±0.3	−3228.7±0.3	1.5±0.4	4.6±1.2
c12c	−2061.0±0.3	−3228.2±0.3	1.5±0.4	5.7±1.4
c21	−2178.8±0.1	−3209.6±0.1	4.9±0.7	18.4±2.1
c22a	−2098.3±0.2	−3221.3±0.2	1.7±0.4	7.2±1.7
c22b	−2092.2±0.3	−3223.9±0.3	3.1±0.5	17.5±3.0
c22c	−2083.5±0.6	−3227.3±0.6	0.8±0.3	3.3±1.2

addition to the 10 VLBA antennas.

The observations were made on 2006 April 30 and lasted for about 7.5 hrs. The right hand and left hand circular polarization data were recorded together in four IFs, each with 8 MHz bandwidth and 16 channels. No cross polarization was performed. The data were correlated with the VLBA correlator using an integration time of 1 s to reduce time averaged smearing. Arecibo had a power failure and problems with software which allowed observations for only 1.5 hr from the 3 hr window available on-source. The data were reduced by following a similar technique to that used to reduce the global 1.7 and 5 GHz data. The images were weighted to obtain an optimum combination of sensitivity and resolution in the maps.

The images of A, B and region C are shown in Fig. 4.14. The HSA imaging at 8.4 GHz confirmed the detection of 3 new components in A and B in addition to the two components known from the earlier observations. The model fitting of the images A and B was done using the task . A and B were both fitted with 5 Gaussian components each. The lensed images of MG 2016+112 are weak at 8.4 GHz (mJy level). Hence, the measured flux densities of the faint components 4 and 5 in images A and B are not robust. Moreover, the pair of merging images in C have a low SNR. The total flux densities of these components were therefore measured by integrating the flux densities of the pixels above 3 σ confidence (where σ is the rms map noise). The positions of these components were found from the peak flux density positions. The peak flux densities, total intensities

Table 4.5: Peak positions, peak and total flux densities of components of A, B and C at 5 GHz for the global VLBI observations. For the C region except C11, the centroids positions and the total flux densities are given.

Components	RA (mas)	Dec (mas)	S_{peak} mJy beam^{-1}	S_{total} mJy
A1	0.00±0.02	0.00±0.02	4.4±0.6	6.5±1.0
A2	−12.20±0.02	4.44±0.02	3.3±0.5	3.6±0.7
A3	−8.5±0.1	3.7±0.1	0.8±0.2	1.8±0.6
A4	6.6±0.1	−3.7±0.1	0.6±0.2	0.9±0.4
A5	−17.4±0.3	5.7±0.3	0.4±0.2	1.6±0.7
B1	−3005.95±0.02	−1503.94±0.02	4.9±0.6	7.1±1.1
B2	−3012.48±0.02	−1497.20±0.02	3.2±0.5	3.0±0.6
B3	−3011.5±0.1	−1499.4±0.1	1.0±0.3	2.2±0.6
B4	−3002.0±0.1	−1505.7±0.2	0.7±0.2	1.5±0.5
B5	−3015.1±0.2	−1492.7±0.2	0.4±0.2	0.7±0.3
c11	−2012.3±0.2	−3234.2±0.2	0.8±0.2	4.7±1.4
c12a	−2043.8±1.0	−3231.0±1.0	1.5±0.3	4.8±1.1
c12b	−2049.1±1.0	−3230.3±1.0	1.7±0.3	9.1±0.3
c12c	−2056.3±1.0	−3229.6±1.0	0.7±0.2	4.6±1.1
c21	−2179.4±0.1	−3210.4±0.1	0.9±0.2	5.2±1.0
c22a	−2097.6±1.0	−3210.4±1.0	1.3±0.3	8.0±1.4
c22b	−2091.3±1.0	−3224.9±1.0	1.5±0.3	7.3±1.3
c22c	−2086.2±1.0	−3225.5±1.0	0.9±0.2	5.0±0.9

and the peak positions of the components are presented in Table 4.6. The errors were determined using the same method as before.

Fifth lensed image

To estimate an upper limit on the flux density of a possible fifth image located in the vicinity of the lensing galaxy D, a fourth sub-field was mapped along with the sub-fields centred on images A, B and C. The sub-field centred on the primary lens D at 1.7, 5 and 8.4 GHz was ed using the task . The sizes of the fields mapped were $0.51'' \times 0.51''$ at 1.7 GHz, $0.3'' \times 0.3''$ at 5 GHz and a field of size $0.2'' \times 0.2''$ at 8.4 GHz (e.g., see Fig. 4.15). The fields were naturally weighted in order to achieve maximum sensitivity. However, no radio emission from a fifth component or from the lensing galaxy was found. The flux density limits (5σ-level) are 0.41, 0.18 and 0.10 mJy beam^{-1} at 1.7, 5 and 8.4 GHz, respectively. From the models with small constant-density core radius highly demagnified images are predicted (Narasimha et al. 1986). Thus, the fifth lensed image is not expected to be detected in the observations presented here.

Table 4.6: Peak positions, peak and total flux densities of components of A, B and C at 8.4 GHz for the global VLBI observations.

Components	RA (mas)	Dec (mas)	S_{peak} mJy beam^{-1}	S_{total} mJy
A1	0.0±0.1	0.0±0.1	1.6±0.1	2.6±0.1
A2	−12.3±0.1	4.5±0.1	0.9±0.1	1.4±0.1
A3	−8.8±0.1	3.9±0.1	0.3±0.1	0.5±0.1
A4	6.4±0.4	−3.7±0.4	0.2±0.1	0.3±0.1
A5	−17.9±0.2	5.3±0.2	0.1±0.1	0.1±0.1
B1	−3005.5±0.1	−1503.7±0.1	1.7±0.1	2.8±0.1
B2	−3012.1±0.1	−1497.0±0.1	1.4±0.1	1.8±0.1
B3	−3010.6±0.4	−1500.3±0.4	0.4±0.1	0.6±0.1
B4	−3001.5±0.2	−1505.9±0.2	0.1±0.1	0.3±0.1
B5	−3014.6±0.2	−1492.5±0.2	0.2±0.1	0.3±0.1
c11	−2012.3±0.3	−3233.3±0.3	0.2±0.1	0.4±0.1
c12	−2047.8±0.3	−3229.6±0.3	0.6±0.1	7.5±0.7
c21	−2177.8±0.5	−3209.9±0.5	0.2±0.1	1.2±0.3
c22	−2092.8±0.3	−3223.7±0.3	0.5±0.1	8.5±1.0

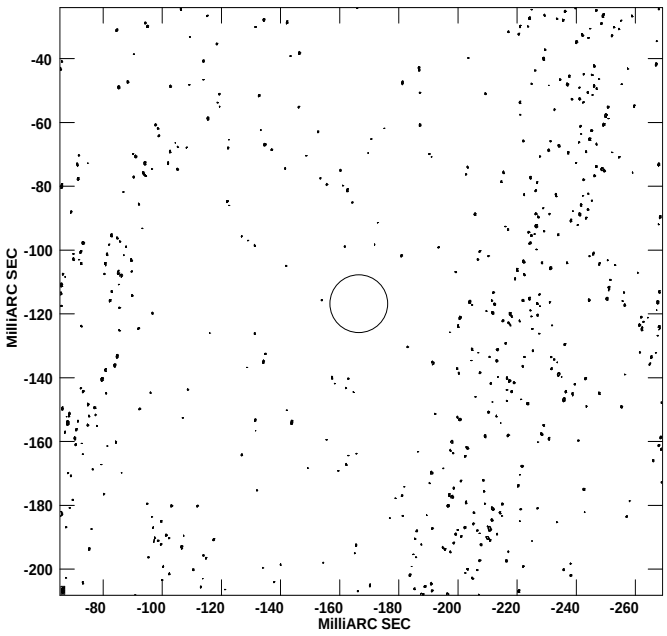

Figure 4.15: The field of view at 8.4 GHz centred on the position of galaxy D taken from the optical data.

4.3 Mass Models

In this section, various mass models for the lens potential of MG 2016+112 are tested with increasing complexity. Given the large number of observational constraints and the un-

certainty of the region C components, different combinations of observational constraints are also tested. The aim is to find the simplest scenario for the mass distribution and the background source that will fit the positions of the five radio components observed in images A and B, the flux density ratio of (at least) component A1 and B1, and the positions of the four components in region C. Firstly, the predictions of the K02 mass model are tested for its consistency with the new observations. Next, two-galaxy mass models are studied which take into account all of the observed constraints and the dwarf companion to the lens. Lastly, three-galaxy models are also tested and the predictions of various mass models are further discussed.

4.3.1 The K02 Mass Model

The multi-frequency high resolution observations were conducted with the aim of testing the K02 mass model. As a sanity check, the K02 mass model was first reconstructed from the K02 constraints. In Koopmans et al. 2002b, the source component 2 was shown to be situated on a caustic (see Fig. 4.8). The positions of A1-B1 as doubly imaged and A2-B2-C12b-C22b as quadruply imaged components along with the flux density ratio (S_B/S_A) were used. The mass model consisted of the main galaxy D (SIE), a mass distribution M1 (SIS) which contributes to the convergence coming from the environment as detected in the weak lensing analysis of Clowe et al. (2001), and a mass distribution M2 (SIS) due to another physically nearby over-density of galaxies found spectroscopically by Soucail et al. (2001). The reconstructed K02 mass model could reasonably reproduce the K02 data constraints with . Note that the K02 constraints did not include the outermost pair of components (C11 and C21). Hence, the astrometric anomaly issue was not addressed in K02.

In light of the rich core-jet structure found in the high resolution observations presented in this chapter, the reconstructed mass model of K02 was used to predict all of the lensed counterparts of the newly found components in images A and B using 5 GHz data from Table 4.5. Fig. 4.16 shows the observed and predicted image positions for all components in A, B and C. Here, components 3, 4 and 5 of image A were used to predict the positions of their counterparts. Components 4, 1, 3 and part of 2 are doubly imaged whereas component 5 and part of component 2 (c, o and e) are quadruply imaged. The labels c, o and e correspond to the merging pair of components C11-C21, C12a-C22a and C12b-C22b respectively[2].

Component 5 is predicted to have four images which are referred as A5, B5, C1-5 and C2-5. C1-5 and C2-5 have opposite parity and are predicted to be about 100 mas on either side of region C. Their magnification relative to A5 (or B5)[3] is larger by a factor of ~ 10. In gravitational lensing, the surface brightness of the lensed images is conserved. Therefore, the quadruple counterparts of A5 and B5 are expected to be 10 times larger in solid angle.

At 1.7 GHz, the flux densities of the components are higher than at other frequencies and the resolution of the global VLBI observations can be lowered enough such that

[2]The labels c, o and e have no significance and are merely used for labeling convenience

[3]Since the flux density ratio of image A and B is ~ 1 and the same is almost true between C1 and C2, the components of A and C2 will be taken as the representative components in the discussion about magnification for the sake of simplicity.

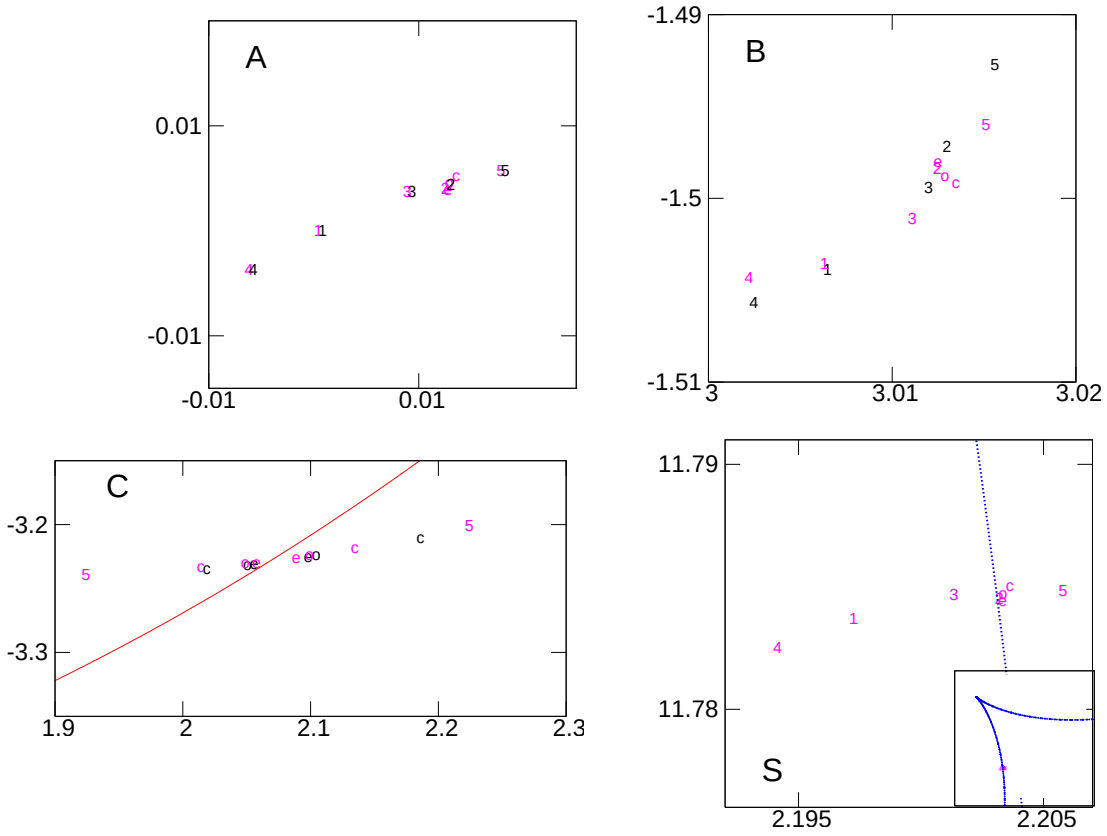

Figure 4.16: All of the counterparts of components of lensed image A and those of C1 as predicted by the K02 mass model. The inset in the panel 'S' shows the source positions on the tangential caustic which is magnified and shown in the panel. The labels in black are observed positions and those in magenta are predicted by the model. Note that the scales are different in each panel.

a component ten times larger than for example, A5 should be detected. However, no such components were detected close to the expected positions of the third (C1-5) and fourth (C2-5) counterparts. To be certain that these observations were capable of detecting these components, fake Gaussian components with a total flux density of 10 mJy and size 0.01×0.01 arcsec^2 were introduced in the *uv*-data at the model predicted positions of C1-5 and C2-5. Using the modified *uv*-data set, new maps were made. It was found that the sensitivity of these data were not sufficient to detect C1-5 and C2-5.

The MERLIN 5 GHz observations did have the right resolution and sensitivity to detect the expected components in region C. The panel on the left of Fig. 4.17 shows the map of the real data at 5 GHz with the C region showing the known four components without any sign of C1-5 (or C2-5) components. Here too, the fake Gaussian components test was performed as before. Fig. 4.17 (right) shows the same observations with the fake components inserted. The reproduced fake components can be seen on either side of the C region. The non-detection of component 5 in the C region in the real MERLIN data indicates either that component 5 is not quadruply imaged and the K02 scenario is incorrect, or that the relative magnification of component 5 in region C is lower than what

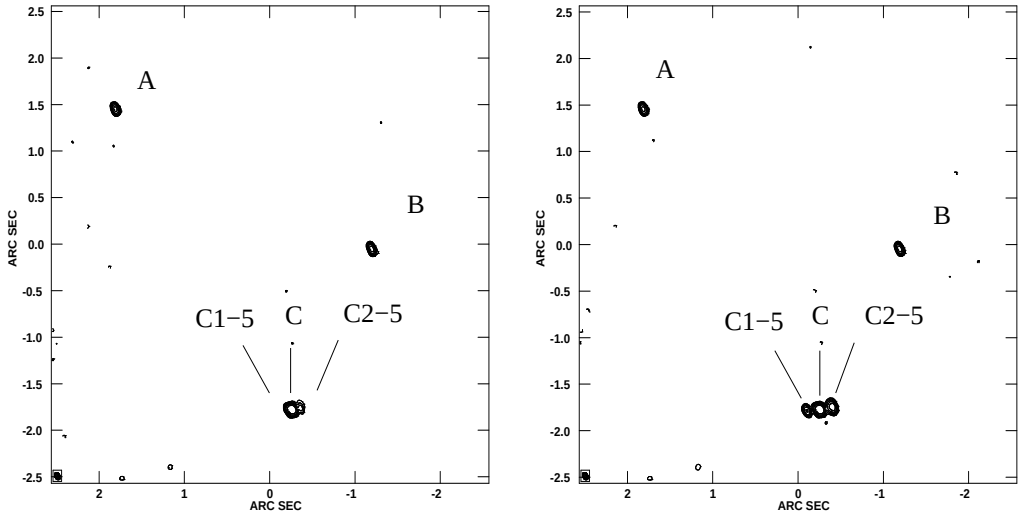

Figure 4.17: MERLIN 5 GHz images restored with a beam of size 0.07×0.05 arcsec^2. The image on the left shows A, B and C with no sign of the new components in the vicinity of region C. When fake components with a surface brightness as predicted from the model, are inserted in the *uv*-data, then these components could be recovered in the image.

is predicted, and the K02 scenarios needs to be modified to take this into account.

It is noted, that irrespective of the newly found discrepancy of the component 5 predictions with the K02 mass model, the K02 mass model was originally not complete since it did not fit the positions of the outer components (C11 and C21) on either side of region C.

4.3.2 Constraints and Priors

The constraints on the positions of the lensed images are taken from the new high resolution VLBI data at 5 GHz (see Table 4.5). The mass models are not sensitive to an astrometric shift of the components which are $\lesssim 1$ mas. Thus, the astrometric uncertainties of all of the components are chosen to be 1 mas in spite of the higher precision of the observations (except see section 4.3.3). For all of the mass models, the flux density ratio (S_{B1}/S_{A1}) of component 1 is constrained to 1.09±0.22 from Table 4.5.

The lensed images A and B show opposite parity and the same five components. Therefore, all of the five components are at least doubly imaged. The components in region C appear to consist of two opposite parity merging images with each showing a pair of compact and extended source components. The emission line spectra of region C indicate the same redshift as that of A and B. Moreover, it is unlikely that the components of region C correspond to a physically distinct source from the lensed quasar, which happen to lie close together to give the expected four-image configuration of a single object. Hence, region C is almost certainly related to the same background source, but it is not clear to which part of the components in A (or B) or anything that is unseen in A (or B).

Since the main lensing galaxy D and the other line-of-sight nearby galaxies are detected in the optical, their positions, measured with respect to the compact optical com-

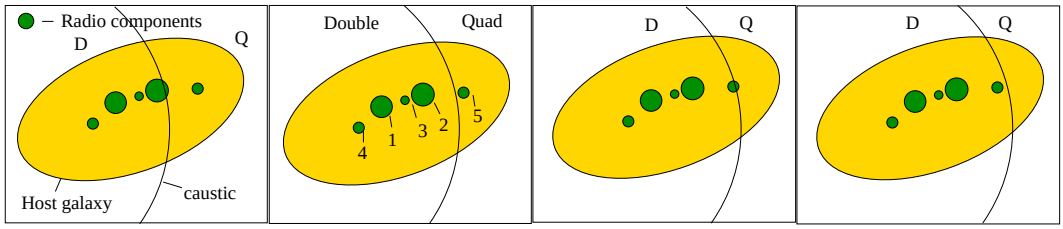

Figure 4.18: Various scenarios A, B, C and D (from left) to test the high-resolution structure observed in the lensed images.

ponent A, are taken from the CASTLES lens database. These positions also provide constraints for the mass models. The ellipticity of the galaxy D in the optical is found to be $\sim 0.43\pm0.01$ with a position angle of -59 deg. Since the galaxy D is a giant elliptical with a stellar velocity dispersion of ~ 328 km s^{-1}, it has the largest contribution to the image splitting and the ellipticity of the halo. Thus, the ellipticity and position angle of the halo is not expected to be much different from the stellar surface brightness profile and a prior $\sim 0.43\pm0.01$ was introduced on the ellipticity and the position angle was fixed to -59 deg for the halo of D. These constraints ensure that the χ^2 minimization does not converge to any unreasonable mass models with an otherwise lower χ^2 as compared to a desired model. For example, mass models of lens systems with both a high ellipticity and shear and the right combination of their individual position angles are capable of producing six or eight lensed images as shown in Keeton et al. (2000a). Furthermore, the images are formed along a semi-circle configuration in such cases. This peculiar configuration has not been observed amongst the lens systems discovered so far. Since MG 2016+112 shows only four images (A, B, C1 and C2) of the background source such mass models are not considered.

4.3.3 A Two-Galaxy model

For the two-galaxy mass model, the dominant elliptical galaxy D is chosen as a SIE and the dwarf galaxy G1, also at the same redshift as D, is chosen as a SIS for all of the scenarios. The angular separations of D and G1 are $(-1.740, -1.782)$ and $(-2.499,-4.037)$ arcsec relative to the optical core of A[4]. The uncertainties on the separations along each axis are 3 mas for D and 28 mas for G1. Including the dwarf G1 is essential because of its proximity to the C region which shows the asymmetrically separated pair of opposite parity features. An external shear is also included in the mass model.

Fig. 4.18 shows different scenarios of the background source straddling the tangential caustic that are investigated for the two-galaxy model. The four possible scenarios are, Scenario A - the caustic goes through source component 2, Scenario B - caustic goes between source components 2 and 5 such that it grazes the source component 5, Scenario C - the caustic goes through source component 5 and Scenario D - the caustic is situated beyond the source component 5.

[4]The optical core is assumed to be coincident with the radio component A2 owing to its flatter spectrum as compared to the other radio components (see section 4.4.1).

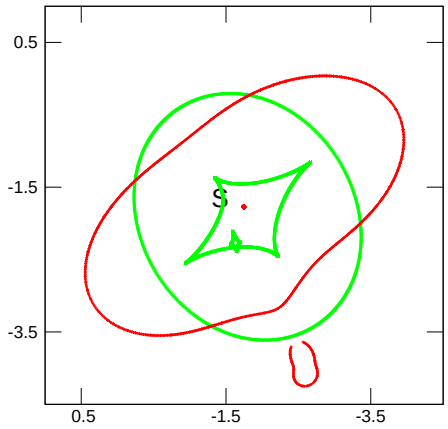

Figure 4.19: The critical curves (red) and the caustics (green) for the two-galaxy model. The source S would shift infinitesimally across the caustic for the different scenarios thereby, changing the image positions and their magnifications slightly.

Scenario A

Constraints: In this scenario, the caustic goes through the source component 2 as was also the case for the K02 scenario[5]. The source components 4, 1, 3 and a part of 2 fall in the doubly imaged region, whereas the rest of the part of 2 and component 5 are quadruply imaged (see Fig. 4.18). The inner components of C (i.e. labels - o and e) are chosen as the counterparts of A2 and an unseen A2e to the west, respectively, while the outer components (i.e. label - c) are associated with some unseen component A2c. The uncertainties for all known components are chosen to be 1 mas, whereas for the unseen components it is chosen to be 5 mas so that they do not contribute significantly to the χ^2. An exception to this are the uncertainties of the counterparts of component A5 in region C i.e. C1-5 and C2-5. Their uncertainties are chosen to be very high (i.e. 10^4 mas) since their positions might be affected due to the substructure in an unexpected way. A total of 41 constraints exist for this mass model. There are 14 free source positions and four free parameters, that is, the Einstein radii of the two galaxies, the shear and its position angle. The degrees of freedom (dof) are 23.

Results/Predictions: The best fitting model has a reduced $\chi^2 \sim 3.5$. The largest contribution to the total χ^2 comes from the image positions ($\chi^2 \sim 60$, most of which is from components of region C) and the second largest is from the galaxy positions (total $\chi^2 \sim 18$; most of which is due to the dwarf G1). The model predicted Einstein radii[6] of the galaxies D and E were found to be 1.570 and 0.143 arcsec, respectively. The recovered ellipticity of D was 0.42. The shear was found to be 10 per cent with a position angle of -41.5 deg measured East of North. The critical curves and the caustics for the two-galaxy model can be seen in Fig. 4.19. The observed and fitted image positions are shown in the

[5]Note that the scenario A and the scenario of K02 are the same. The difference lies in the data constraints that were used to make the mass models. Hence, they predict slightly different image positions and magnifications.

[6]see section 2.2.4 for the definition of Einstein radius in case of a SIE

top panel of Fig. 4.20. The relative magnification of the components in region C, with respect to A2-A2o, is very high. For example, $S_e/S_{A2e} \sim 10^3$ and $S_c/S_{A2c} \sim 10^2$. Hence, the counterparts of image C in images A and B are predicted to be 100-1000 times fainter (at a level of $\sim$10 μJy) and 10-30 times smaller in size. These counterparts can not be detected with the sensitivity of the observations undertaken at any of the three frequencies presented here.

Component 5 is predicted to have two images in C which are referred to as C1-5 and C2-5. The predicted flux density ratio is $S_{C2-5}/S_{A5} \sim 15$ and a size that is at least four times larger than that of A5. Since C1-5 and C2-5 are not detected in the MERLIN observations, this scenario is ruled out.

Scenario B

Constraints: In this scenario, although the caustic is situated between component 2 and 5, it is closer to component 5 such that C11-C21-A5-B5 are its four lensed images whereas the inner elongated pair of components 'o' and 'e' are associated with unseen components A5o and A5e to the east of the A5 detected. Thus, components 4, 1, 3 and 2 will be doubly imaged as shown in the second panel of Fig. 4.18. The number of constraints are 41 and free parameters remain the same, thereby, giving a dof of 23.

Results/Predictions: Not surprisingly, the reduced χ^2 is 3.6 because this scenario is a slight modification of scenario A and is not expected to modify the global mass model significantly. Furthermore, the individual χ^2 contributions and the best-fitting model parameters are also similar. However, the predictions are expected to change here. The positions of the components in images A and B are fitted within 1 mas except for components 4 and 5, which are fitted within 1.5 mas. The observed and model predicted image positions are shown in the bottom panel of Fig. 4.20. The relative magnifications of the C components with respect to A are predicted to be $S_e/S_{A5e} \sim 10^3$ and $S_c/S_{A5} \sim 200$. Since the inner components of C ('o' and 'e') have very high magnification, their counterparts in A and B would be unseen which is consistent with the observations. However, the counterparts of components 'c' at the position of A5 (or B5) are predicted to be $\sim$100 times fainter whereas the observed relative magnification is $\lesssim$10. Due to this inconsistency, this scenario is not acceptable either.

Scenario C

Constraints: Here, the caustic goes through component 5 such that the inner elongated pair ('o-e') of region C is associated with A5 and an unseen component to the west (A5e) which are both quadruply imaged. The pair of components ('c') is associated with an unseen component A5c which would lie further to the north-west. Components 4, 1, 3 and 2 are doubly imaged. The total number of constraints and free parameters are the same as before, hence the dof is 23.

Results/Predictions: The best-fitted model parameters are the same as before. The reduced χ^2 is 3.6 and the individual χ^2 contributions are similar to the previous scenarios. The doubly imaged components 1, 2 and 3 are fitted within the 1 mas uncertainties whereas

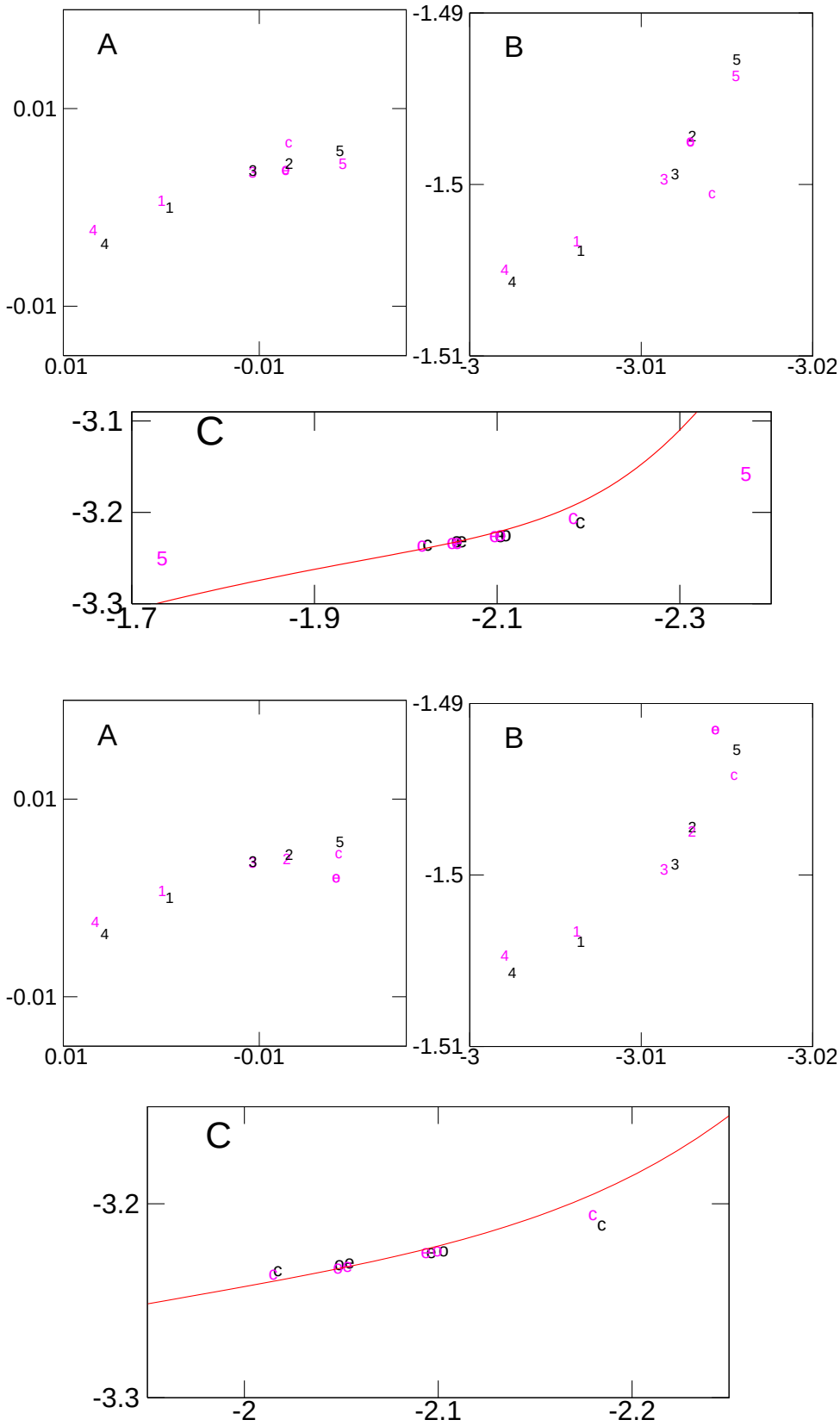

Figure 4.20: In scenario A - components 2-o and 2-e are associated with each other and scenario B - components 5-c are associated with each other. The labels in black are the observed positions and those in magenta are predicted by the model. The x- and y-axes represent relative Right ascension (RA) and relative Declination (Dec). Note that the separations are in arcsec and the scales are different in each panel.

component 4 is fitted within 1.5 mas. The counterparts of c, o and e in A (or B) are predicted within 1.5 mas to the west of the peak position of A5. The relative magnification of the C region is predicted to be 100-1000 times higher than their counterparts in A5 (or B5). Therefore, the corresponding components in A and B would have flux densities of the order 5-50 μJy which is much below the noise level of the global 5 GHz observations. Moreover, it is certainly not possible to resolve these components which would have sizes 10-30 times smaller. Nevertheless, since these are undetected they are consistent with our observations. This scenario, although, consistent with the new observations presented here, may or may not stand true with better observational constraints in the future.

Scenario D

Constraints: In this scenario, the caustic is situated to the west of component 5 such that components 4, 1, 3, 2 and 5 are all doubly imaged and the components of C are associated with some undetected components in A and B. The total number of constraints is 45. There are 16 free source positions and 4 free parameters and hence, the dof is 25.

Results/Predictions: The reduced χ^2 for the best-fitting model is $\sim$ 3.4 and the model parameters are similar to those given in scenario A. The doubly imaged components 1, 2 and 3 are fitted within their uncertainties (1 mas) whereas components 4 and 5 are fitted within 1.5 mas. The c, o and e components are predicted to have four images. Their counterparts in A and B are expected to have a relative demagnification of 100-1000 times. Such components would not have been detected in the multi-frequency observations presented here. Therefore, this scenario is also consistent with the observations. In fact, observations that have 1000 times better sensitivity are needed to test this scenario.

No constraints on the position of the lens mass components

The cluster associated with the main galaxy D is believed to be a proto-cluster which is not centrally concentrated yet. The conclusion that the cluster is not virialized is inferred from the absence of any diffuse X-ray emission (Chartas et al. 2001). Therefore, the optical position of the BCG (galaxy D) may not be coincident with the centre of the cluster.

Mass models with no constraints on the positions of the lensing galaxy were hence tested for different scenarios. The following results are described for scenario D. Initially, the position of the main galaxy D was allowed to be free. The best fitting model shifted the galaxy D by $\sim$60 mas to the west. The reduced χ^2 of 1.5 was mainly improved by better fitting the image positions. The total χ^2 of the galaxy position has not changed significantly which arises here solely due to the dwarf G1. The best fitting parameters are similar to the mass model with the position of D constrained. The Einstein radii are 1.551 and 0.14 arcsec for D and G1, respectively. The fitted ellipticity is 0.43 and the shear is 11 per cent with a position angle of 41.8 deg. The results of the mass modeling are summarized in Table 4.7.

4.3.4 A Three-Galaxy Model

In spite of the high shear ($\sim$ 10 per cent) for the SIE+SIS mass model, the image positions could not be reproduced satisfactorily. Given the dense environment around the lensing galaxy and a handful of line-of-sight nearby galaxies, the galaxy G2 (SIS) was added to the two-galaxy model to account for some effects from the environment. This three-galaxy model was tested for scenario A.

Scenario A

Constraints: A galaxy G2 which was detected in the optical/infrared imaging at (-5.749, 1.767) arcsec relative to image A, was added to the two-galaxy model in scenario A. In this scenario, the caustic is situated on component 2. The number of constraints are therefore 27. However, the number of free parameters increased by one due to the Einstein radius of the third galaxy, hence the dof reduced to 22.

Results/Predictions: Introducing a third galaxy G2 does not result in any significant change to either the reduced χ^2 ($\sim$ 3.6) or the best-fitted model parameters. The Einstein radii of D, G1 and G2 of the best-fitting model are 1.565, 0.133, 0.075 arcsec, respectively. The ellipticity of D is 0.42, with its position angle fixed to -59 deg. The shear has a strength of 9.6 per cent and a position angle -40.5 deg with hardly any improvement over the two-galaxy model. Note that the K02 mass model, which also has two mass components that are different from the ones used here, in addition to galaxy D, also finds a shear of $7-12$ per cent with a position angle -51 to -24 deg. As a result, the predicted image positions and relative magnifications are still of the same order which demand that the counterpart of component 5 be visible in region C. Since it is not found in the VLBI/MERLIN observations presented here, the situation in scenario A is not improved by introducing a third galaxy at the position of G2. The results of the mass modeling are summarized in Table 4.7.

4.4 Results and Discussion

In this section, the results from the new observations and the new mass models made with the new constraints are discussed. These results are further compared with the previous models for the lens system.

4.4.1 The Radio Spectra

The left panel of Fig. 4.21 shows the radio spectra of A, B and C between 1.7 and 8.4 GHz. Images A and B have similar spectra and flux densities. This is consistent with previous multi-epoch and multi-frequency observations. The radio spectrum of region C is found to be flatter relative to images A and B. Furthermore, the flux density of region C is significantly higher as expected for highly magnified images near the critical curve.

The panel in the middle of Fig. 4.21 shows the spectra of all five of the components in image A. All of the components have steep spectra between 1.7 and 5 GHz, except component A2. The spectra of all of the components further steepen between 5 and

Table 4.7: Results of the mass models and the fitted parameters. The subscripts 'ip' and 'if' denote the χ^2 due to the image positions and flux densities, respectively. The subscript 'gp' denotes the χ^2 due to the positions of all of the galaxies in a model whereas 'oth' denotes the χ^2 contribution from other priors e.g., ellipticity. See text for further details.

Model–Scenario	χ^2_{tot}	χ^2_{ip}	χ^2_{if}	χ^2_{gp}	χ^2_{oth}	Model parameters	Comments/Status
SIE+SIS+shear							
Scen A	81.0	60.0	1.9	18.0	0.3	b_D=1.570, P_D=(−1.746, −1.777), e_D=0.42, γ_s= (0.10, 41.5), b_{G1}=0.143 , P_{G1}=(−2.572, −3.972)	predicts extra components in region C – not found in data (Not acceptable)
Scen B	83.4	60.6	1.9	20.8	0.05	b_D=1.568, P_D=(−1.746, −1.776), e_D=0.43, γ_s= (0.10, 41.5), b_{G1}=0.146, P_{G1}=(−2.575, −3.976)	relative magnification of component 5 component C21–do not match data (Not acceptable)
Scen C	83.5	61.5	1.9	20.0	0.07	b_D=1.569, P_D=(−1.746, −1.776), e_D=0.43, γ_s= (0.10, 41.5), b_{G1}=0.145, P_{G1}=(−2.573, −3.977)	relative magnifications are consistent with the data presented here (acceptable – needs further investigation)
Scen D	85.0	62.6	1.9	20.3	0.04	b_D=1.569, P_D=(−1.746, −1.776), e_D=0.43, γ_s= (0.10, 41.5), b_{G1}=0.146, P_{G1}=(−2.573, −3.979)	relative magnifications are consistent with the data presented here (acceptable – needs further investigation)
position of Gal D free – Scen D	41.0	21.3	3.1	16.3	0.2	b_D=1.551, P_D=(−1.799, −1.779), e_D=0.43, γ_s= (0.11, 41.8), b_{G1}=0.144, P_{G1}=(−2.581, −3.959)	An offset of ∼ 60 mas from the optical position is significantly large (Not acceptable)
SIE+SIS+SIS+shear							
Scen A	79.5	59.0	1.9	18.0	0.7	b_D=1.551, P_D=(−1.799, −1.779), e_D=0.42, γ_s= (0.09, 40.5), b_{G1}=0.146, P_{G1}=(−2.574, −3.964) b_{G2}=0.075, P_{G2}=(−5.749, 1.767)	Adding G2 does not improve the model fit to the data (Not acceptable)

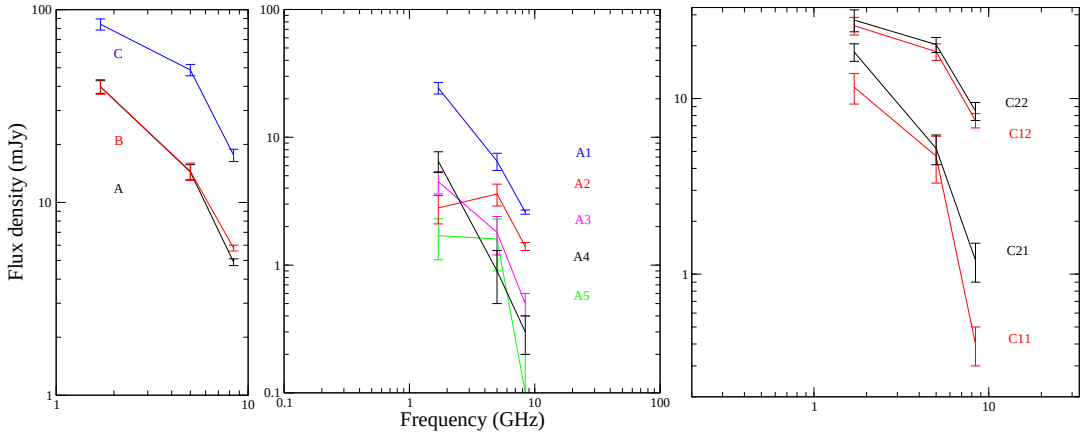

Figure 4.21: Total spectra of A, B and C and the spectra of the five and four components of A and C, respectively using the global VLBI and HSA observations.

8.4 GHz. Note that component A5 also shows a steep spectrum like the other components but has larger uncertainties. Since component A2 is flatter than the other components, with a turnover in the spectrum close to 5 GHz, it is most likely the radio core of the quasar. Thus, it is doubly imaged similar to the optical core.

The panel on the right of Fig. 4.21 shows the spectra of the four components in region C. The merging images are expected to belong to the same part of the background quasar and hence, they should have similar spectra. The components of the inner pair (C12-C22) show similar spectra and the outer pair (C11-C21) also have similar spectra within the uncertainties. The inner pair of elongated components of C (C12-C22) have flatter spectra than the outer pair of components (C11-C21). Since the components of the inner pair also have higher flux densities, these dominate the spectrum of region C at low resolutions, for example, as found in the MERLIN imaging presented in section 4.2.1.

The spectral similarity of component A2 with that of the elongated component C12 in region C might appear to favor scenario A. However, the non-detection of the expected counterpart of A5 in region C disfavors this possibility, as was shown in section 4.3. Furthermore, any association between the components of C and those of image A (or B), on the basis of their spectra, should not be trusted because the components in C1 (or C2) probably correspond to a region which is about one tenth the size of any of the components found in A (or B), and may not have spectra similar to the spectrum of the component as a whole.

4.4.2 Relative Magnification of the Images

The relative magnification (μ_r) of the lensed images, that is, the magnification of an image (B) with respect to another image (A), is simply the ratio of individual magnifications μ_B and μ_A (see Eq. 1.20). Ideally, the surface brightness of the lensed images is conserved, hence, the observed flux density ratios of the images should be equal to the ratio of their solid angles. This can be expressed as,

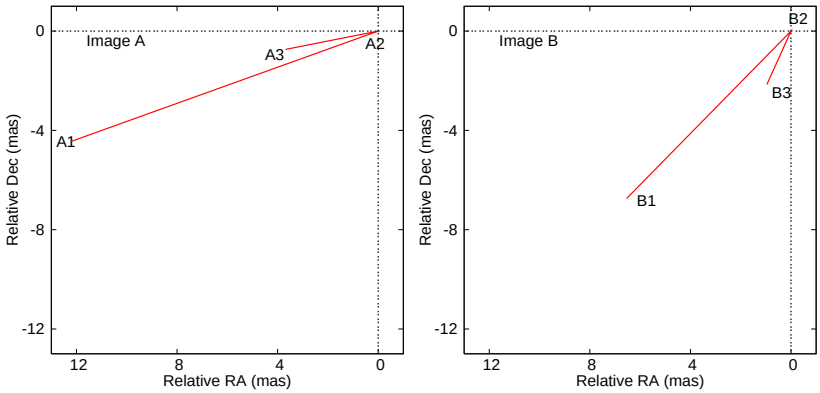

Figure 4.22: Vectors A1 and A3 are defined w.r.t A2, chosen as the origin. Similarly, vectors B1 and B3 are defined w.r.t B2, chosen as the origin. The mapping of the triangle A1-A2-A3 to B1-B2-B3 is defined by a matrix called the relative magnification matrix.

$$\mu_r = \frac{\mu_B}{\mu_A} = \frac{S_B}{S_A} = \frac{\omega_B}{\omega_A} = |det(M_{ij}^{AB})|, \tag{4.1}$$

where the relative magnification matrix (M_{ij}^{AB}) is defined as,

$$\boldsymbol{I_B} = M_{ij}^{AB} \boldsymbol{I_A}\,. \tag{4.2}$$

Here, the vectors $\boldsymbol{I_A}$ and $\boldsymbol{I_B}$ correspond to vectors in images A and B, respectively. This formulation can be applied to real gravitational lens systems (e.g., Garrett et al. 1994b; Jin et al. 2003) provided the lensed images show a rich structure of non-collinear features.

In the lensed images A and B of MG 2016+112, there are five non-collinear components. It is simpler to define the magnification matrices for a set of three non-collinear components instead of attempting to fit for all of the components simultaneously. The combinations are chosen such that at least one out of the three components is a bright compact component. A total of five sets of triplets are investigated. For every triplet, the matrices are defined three times such that every component is chosen once as the origin. For example, 1-2-3 is a triplet of components 1, 2 and 3. A matrix mapping of vectors (***A1***, ***A3***) of image A to the vectors (***B1***, ***B3***) of image B, is defined with component 2 as the origin (see Fig. 4.22). Therefore, the vector ***B1*** is mapped to ***A1*** through a matrix, which is defined as,

$$\boldsymbol{B_1} = \begin{pmatrix} B_{1x} \\ B_{1y} \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} A_{1x} \\ A_{1y} \end{pmatrix} = M_{ij}^{12} \boldsymbol{A_1}. \tag{4.3}$$

Next, with component 3 at the origin, a matrix mapping between the vectors (***A1***, ***A2***) and (***B1***, ***B2***) is defined. Finally, a matrix mapping with component 1 as the origin is defined between the images.

Shuffling the components at the origin one by one and re-calculating the matrix is expected to give the same result and is performed to test the repeatability of the method used here. Furthermore, the relative magnification matrix for the images A and B are

calculated both at 5 and 8.4 GHz, which again are expected to be consistent. Unless there is any frequency dependent structure in the images, the flux density ratios (i.e. the determinants) for the counterparts of a component should be same. Turning it around, an expensive method to find out any frequency dependent structure would be to compare the relative magnification matrices.

The positions of the components of images A and B are taken from Tables 4.5 and 4.6, after applying the appropriate shift of origin for each case. The uncertainties are chosen as 0.1 mas for all positions throughout for simplicity. Nevertheless, the determinants of the magnification matrices are found to be consistent between 5 and 8.4 GHz for all of the sets to within the uncertainties (see Table 4.8). Furthermore, the determinants are consistent between any two sets. The negative values of the determinants confirms the opposite parity of image B with respect to image A.

4.4.3 Substructure

The prediction of Dalal & Kochanek (2002), for the satellite mass fraction of the total mass of a typical lens galaxy halo, is $0.006 < f_{sat} < 0.07$ at 90 per cent confidence which is consistent with CDM predictions. The lens system B2045+265 has the most discrepant flux ratios in the radio amongst the known flux ratio anomalous cases. McKean et al. (2007) have detected a satellite within the Einstein radius of the lens system which is evidence of luminous substructure. However, the satellite is not confirmed to be at the redshift of the lens. Also, the mass distribution of the satellite must be highly elongated to satisfy the observed properties of the lensed images.

Similarly, Ros et al. (2000) have found that a previously known object near to the lensed images of MG J0414+0534 when included in the mass model, fits the positions and flux densities of the lensed images of the radio core well. On the other hand, Trotter et al. (2000) have taken into account only the astrometric constraints and investigated higher order multipoles of the potential by Taylor expansion. In conclusion, they find that neither the object X nor an over-density of galaxies to the south-west have a significant contribution to the image distortions.

In MG 2016+112, the lensed images A and B show similar spectra. The expected parity and the surface brightness of the images is also found to be conserved. In region C, the components in the inner pair, C12 and C22 show similar spectra and flux densities. The outer pair of components C11-C21 show similar spectra, albeit the flux densities are different. Moreover, component C21 is notably displaced from the position expected for a merging image pair. With a smooth mass model of SIE+shear centred at D, it is possible to fit the positions of the components of images A and B only. Fitting the components of region C resulted in a very poor fit with a reduced χ^2 of about 180. In MG 2016+112, a satellite galaxy (G1) has been spectroscopically confirmed to be at the redshift of the lensing galaxy D (Koopmans & Treu 2002a). Including this satellite G1 in the mass model significantly improved the fit to all of the astrometric constraints (reduced $\chi^2 = 3.5$).

Additionally, the relative magnification matrix mapping could be used to test the presence of substructure near the lensed images A and B. Note that it is not possible to carry out the test for region C due to two reasons, a) the C components straddle the critical curve, and due to the high magnification gradient the linearity of the mapping does not hold true and b) since the features in C are almost along a straight line, the errors intro-

Table 4.8: The determinants of the relative magnification matrices for the components of lensed images A and B of MG 2016+112.

Components	1-2-3		1-3-5		1-4-5	
	5 GHz	8.4 GHz	5 GHz	8.4 GHz	5 GHz	8.4 GHz
1	−1.08±0.34	−1.31±0.30	−1.29±0.22	−1.08±0.13	−1.02±0.15	−0.79±0.11
2	−1.09±0.35	−1.32±0.31				
3	−1.10±0.35	−1.31±0.31	−1.29±0.22	−1.08±0.13		
4					−1.02±0.15	−0.78±0.11
5			−1.28±0.22	−1.08±0.13	−1.02±0.15	−0.78±0.11

Components	2-4-5		2-3-4	
	5 GHz	8.4 GHz	5 GHz	8.4 GHz
1				
2	−1.30±0.27	−0.85±0.12	−0.91±0.20	−1.18±0.21
3			−0.91±0.20	−1.18±0.21
4	−1.29±0.27	−0.85±0.12	−0.90±0.20	−1.18±0.21
5	−1.29±0.27	−0.84±0.12		

duced would be too high to estimate the determinant reliably. Thus, despite the spectroscopic confirmation of the luminous substructure that clearly affects the position of the C21 component, its contribution cannot be tested independent of any assumptions about the mass model. Nevertheless, the determinant of the relative magnification matrix between images A and B is found to be in agreement with the observed flux density ratios of the images. Furthermore, it holds true for the 5 and 8.4 GHz data sets, which have high enough resolution to resolve the components unlike the 1.7 GHz data set. Therefore, it can be concluded that no significant substructure is found in the immediate vicinity of either the image A or image B components.

4.4.4 Comparison Between Mass Models

For MG 2016+112, the mass model of Koopmans et al. (2002b) was made with A1-B1 and A2-B2-C12-C22 as the constraints without taking into account the asymmetric separations in the merging pair C1 and C2. The counterparts of C11-C21 were predicted to be lying to the north-west of A2-B2. Subsequently, a mass model with galaxy D and a nearby dwarf galaxy G1 was claimed to have fitted all of the data from Koopmans et al. (2002b), including the problematic asymmetric positions (Kochanek private communication, see Kochanek et al. 2004). Using the same data, the mass model of Chen et al. (2007) was made with SIE+SIS+shear and a multipole term of order 4 for the main galaxy. The best fitting mass model had a reduced χ^2 of 3 (Dalal private communication).

However, the matter did not close here since the new observations presented in this chapter detected component 5 in images A and B at the expected position to the north-west, but the observed flux density was too high compared to what was predicted from the mass model for A5 to be the counterpart of C11-C21. Therefore, mass models like SIE+shear, SIE+SIS+shear or SIE+SIS+SIS+shear that assume the caustic to be situated either on component 2 (Scenario A) or between components 2 and 5 (Scenario B, see Fig. 4.18), will either predict two counterparts of A5-B5 in region C that will have a magnification more than A5-B5, and should have been detected, or predict a high relative magnification for C11/A5, inconsistent with the observed flux density ratio.

One of the ways to get around this problem is to have the caustic situated on component 5 (Scenario C) or to the north-west of component 5 (Scenario D). Now, any smooth mass model that fits the astrometric constraints will predict relative magnifications which will not be verifiable using the observations presented here. Thus, this would resolve the inconsistencies between the predictions and observations. However, the models presented here have not successfully fitted the high-resolution astrometric constraints (reduced $\chi^2 \sim 3.5$) and hence, these are an approximate representation of the true mass model. Note that the models presented here, and that of Chen et al. (2007) are equally good on the basis of goodness of fit of the models.

An alternative to the smooth mass models which include one or more luminous clumps of substructure, would be to test a model with higher order multipoles of the gravitational lens potential which may not require additional substructure. Such attempts have been made previously by Evans & Witt (2003) and Congdon & Keeton (2005) for the anomalous flux ratio cases, and by Biggs et al. (2004) for B0218+437 which is another likely case of an astrometric anomaly. Perhaps, higher order effects are giving rise to tiny distortions that could not be fitted with a smooth mass model and one or two luminous

sub-halos.

Velocity dispersion of the elliptical D

The main galaxy D is fitted with an isothermal ellipsoid and there is no simple relation between the Einstein radius and the velocity dispersion for a SIE, whereas the mass within the Einstein radius, the Einstein radius and the velocity dispersion have a simple relation for a SIS. Therefore, a SIS equivalent of the mass of a SIE within the critical curve is used to find the SIS equivalent Einstein radius, which is then used to find the predicted velocity dispersion. For a circularly symmetric profile and very small (~ 0) projected separation between the source and the lensing galaxy, the mass within the Einstein radius is related by,

$$M = \frac{b^2 c^2}{4G} \frac{D_d D_s}{D_{ds}}. \tag{4.4}$$

Furthermore, for an isothermal sphere, the Einstein radius (b) can be related to the velocity dispersion by

$$b = \frac{4\pi\sigma^2}{c} \frac{D_{ds}}{D_s}. \tag{4.5}$$

The velocity dispersion is estimated from the best-fitting Einstein radius of the two-galaxy models. Given any of the scenarios for a SIE+SIS+shear model, the Einstein radius of D is ~1.57 arcsec[7] which corresponds to $\sigma = 343$ km s^{-1}. The velocity dispersion from the model is in agreement with the central stellar velocity dispersion 328 ± 32 km s^{-1} measured by Koopmans & Treu (2002a), and with the predicted velocity dispersion of 320-342 km s^{-1} from the Koopmans et al. (2002b) mass model. Treu & Koopmans (2002) showed that the mass density profile within the Einstein radius of MG 2016+112 is isothermal (i.e. slope $\gamma = 2.0 \pm 0.1$) using the combined mass estimates from stellar dynamics and strong lensing, which is consistent with the assumption of isothermality in the mass models presented here.

A limit on the mass of the satellite G1

For the SIE+SIS+shear models, the Einstein radius of the satellite galaxy (i.e. the SIS component) is found to be ~ 0.14 arcsec. Kochanek et al. (2004) have quoted the mass of G1 as 1 per cent of the mass of the main galaxy D. Since the Einstein radius $b \propto \sqrt{M}$, the SIE+SIS+shear model predicted mass fraction of the dwarf galaxy is 0.8 per cent, which is consistent with that found by Kochanek et al. (2004). However, it is still higher than the upper limit of ~0.5 per cent placed by simulations of satellites within the range of typical image separations (Mao et al. 2004). The predicted velocity dispersion of G1 is 99 km s^{-1}. This is consistent with the results of Chen et al. (2007), who have found the velocity dispersion to be $87 < \sigma < 101$ km s^{-1} for the satellite G1.

The mass models of Kochanek et al. (2004) and Chen et al. (2007) could neither be reconstructed from the old data nor be verified with the new observations since their data constraints and mass model parameters have not been presented. Even though similar

[7] Note the convention implemented in for Einstein radius of SIE, see section 2.2.4

mass models are investigated here using the new observations, only an *approximate* mass model with consistent values of the parameters like the Einstein radius or velocity dispersion could be found.

4.5 Conclusions

Multi-frequency high-resolution radio observations of the gravitational lens MG 2016+112 were conducted to carry out a spectral analysis and to find a mass model for the complex structure in the lensed images. Radio maps made with simultaneous MERLIN and global VLBI observations at both 1.7 and 5 GHz were presented. Subsequently, HSA observations at 8.4 GHz were undertaken to carry out a spectral study of the components at high resolution. In addition to the two previously known components in images A and B, three new components were detected in the observations presented here. The observations with the HSA proved crucial in the confirmation of the new components. A total of five components are now found in images A and B. No more new components are detected above 33 μJy within a region of size 0.21×0.21 arcsec^2 centred at images A and B from the HSA imaging at 8.4 GHz. A 5σ upper limit was placed on the peak surface brightness of an odd image in the vicinity of the lens D, or radio emission from D, of 0.18 mJy beam^{-1} at 8.4 GHz.

The overall radio spectra and the flux densities of the component in A and B were found to be similar. The flux density ratio of images A and B were consistent with the determinant of the relative magnification matrix. Therefore, there is no significant substructure or any other effects that might affect the flux densities of the images. In region C, the morphology and spectra of C11-C21 and C12-C22 were found to be similar, as expected for lensed images. Furthermore, the observed flux densities of the C2 pair are found to be higher than the C1 pair at all frequencies, which could be due to the proximity of the satellite galaxy G1 to the C2 pair with a positive parity (Keeton 2003b). The identification of components in region C with those in image A (or B), on the basis of their spectra, cannot be done because the highly magnified components of region C correspond to extremely small regions in either the detected (4, 1, 3, 2 or 5) or undetected components of images A and B.

Several mass models with more than one mass component in a single lens plane were investigated for four scenarios. In these scenarios, components of region C were constrained as the lensed counterparts of different parts of the components of images A and B, and the consequences of doing so were assessed. The mass models tested here with scenarios A and B predicted relative magnifications of the images, that were inconsistent with the observations. The predictions of scenarios C and D were consistent with the observations presented here. Note that the predictions of the mass models of Koopmans et al. (2002b), Kochanek et al. (2004) and Chen et al. (2007) are not consistent with the new observations because of the detection of component 5. A SIE+SIS+shear model with the satellite galaxy G1 (SIS) found at the same redshift as the lensing galaxy D (SIE) improved the fit (reduced $\chi^2 = 3.5$) to the astrometric constraints significantly, as stated previously by Kochanek et al. (2004). However, a model with an even more complex mass distribution than that which has been tested here, which is causing small scale deviations in the positions of the C components, is required.

5 B2108+213: a massive radio-loud lens in a galaxy group

5.1 Introduction

There are now over a hundred examples of galaxy-scale lens systems known. These systems have been found from systematic searches around potential lens galaxies (e.g. Fassnacht et al. 2004; Bolton et al. 2006a; Moustakas et al. 2007) and from large surveys of the lensed source parent population (e.g. Bahcall et al. 1992; King et al. 1999; Browne et al. 2003; Myers et al. 2003). Serendipitous discoveries of gravitational lensing from deep high-resolution optical and infrared imaging have also been made (e.g. Fassnacht et al. 2006b). Typically, the image separations for these gravitational lens systems are between 0.5–1.5 arcsec, with only a handful having image separations $\gtrsim 3$ arcsec (see Fig. 5.1; Walsh et al. 1979; Lawrence et al. 1984; Wisotzki et al. 1993; Muñoz et al. 2001; Sluse et al. 2003; Oguri et al. 2004; McKean et al. 2005; Bolton et al. 2006b; Inada et al. 2006). Since the enclosed mass of a lens system is proportional to the square of the image separation (i.e. $M_E \propto \Delta\theta^2$ for a circularly symmetric mass distribution; Kochanek 1991), such wide image separation lens systems (i.e. $\gtrsim 3$ arcsec) may be due to halos which are an order of magnitude more massive than those of typical lens galaxies.

As such, wide image separation lens systems could be used to probe the matter distribution at the top end of the mass function for galaxy-scale structures. Alternatively, the wide image separation could be due to the lens galaxy being in an over-dense environment, for example, in a group or cluster of galaxies. Recent imaging and spectroscopic surveys of the local environments of lens galaxies have found many to be members of larger structures (Fassnacht & Lubin 2002; Fassnacht et al. 2006a; Momcheva et al. 2006; Williams et al. 2006; Auger et al. 2007a). However, the total contribution of the group or cluster to the lensing mass distribution is thought to be no more than ~ 5 per cent for systems with image separations of ~ 1 arcsecond (Momcheva et al. 2006; Auger et al. 2007a). Studies of gravitational lens systems with larger image separations may show an enhanced lensing mass distribution which has been boosted by the environment. An extreme example of this is cluster lensing which can produce image separations much larger than 10 arcsec (e.g. Oguri et al. 2004; Inada et al. 2006).

The gravitational lens system CLASS B2108+213 has two radio-loud lensed images separated by 4.56 arcsec (McKean et al. 2005). This image separation immediately identifies B2108+213 as an excellent opportunity to study a mass regime between the typical galaxy and cluster-scales. New high-resolution radio imaging of B2108+213 is presented with the twin aims of determining the nature of the third radio component and finding

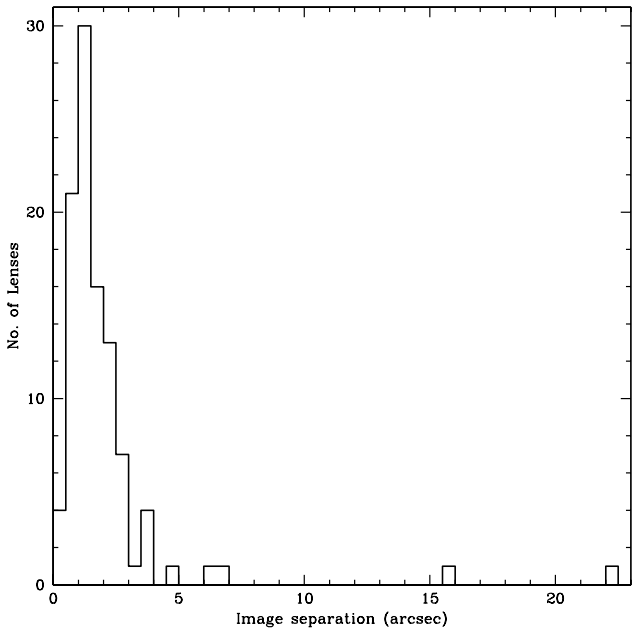

Figure 5.1: A histogram of the distribution of image separations for lens systems. Arcsecond-lenses are found in abundance.

additional observational constraints for the lensing mass model from the two lensed images. 1.4 GHz imaging with MERLIN is presented to show the low surface brightness emission from this system. High resolution imaging of the lensed images at 1.7 GHz with the VLBA and at 5 GHz with global VLBI are also presented. These new data show extended jet emission from both lensed images on mas-scales which are used to test new mass models for B2108+213. Finally, the results and conclusions are presented.

5.2 Discovery in the Radio

B2108+213 was first found in 8.46 GHz imaging with the VLA as can be seen in Fig. 5.2. This showed compact emission from three components viz. A, B and C. The discovery was made during the Cosmic Lens All-Sky Survey (CLASS; Myers et al. 2003; Browne et al. 2003). CLASS is a survey to study radio-loud gravitational lens systems. It comprises compact flat-spectrum radio sources. The complete sample consists of 11 685 sources which were selected to have a flux density ≥ 30 mJy at 5 GHz. These sources were selected from the Green Bank Survey (GB6) catalogue (Gregory et al. 1996) at 5 GHz and the NVSS [1] catalogue (Condon et al. 1998) at 1.4 GHz such that the spectral index $\alpha^5_{1.4} > -0.5$ (where $S_\nu \propto \nu^\alpha$). The primary aim of this survey was to find galaxy-scale lenses i.e. those with image separations of the order of a few arcseconds. The sources for which the integrated flux density of the images at 8.46 GHz was ≥ 20 mJy were further selected as the lens candidates for the refined CLASS sample. This resulted in the discovery of a total of 22 gravitational lens systems. Since B2108+213 did not satisfy the

[1] National Radio Astronomy Observatory Very Large Array Sky Survey

selection criteria, it did not belong to the statistically well defined sample. Nevertheless, it is the widest separation lens system discovered in CLASS.

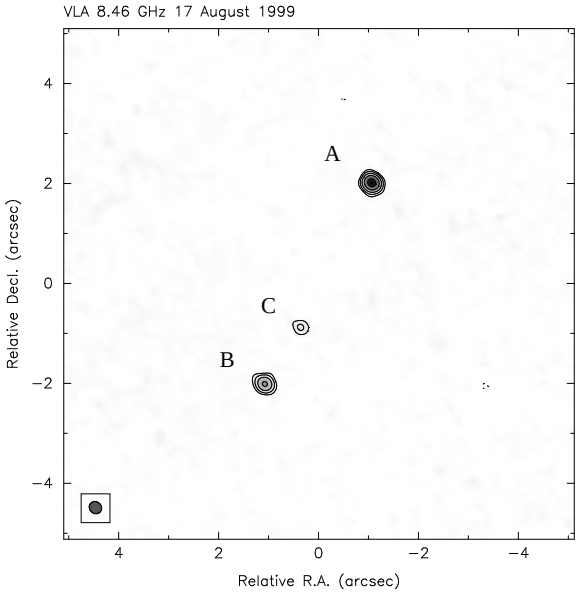

Figure 5.2: A 30 s snapshot image of B2108+213 at 8.46 GHz using the VLA shows three components A, B and C (McKean et al. 2005). The three components are not collinear.

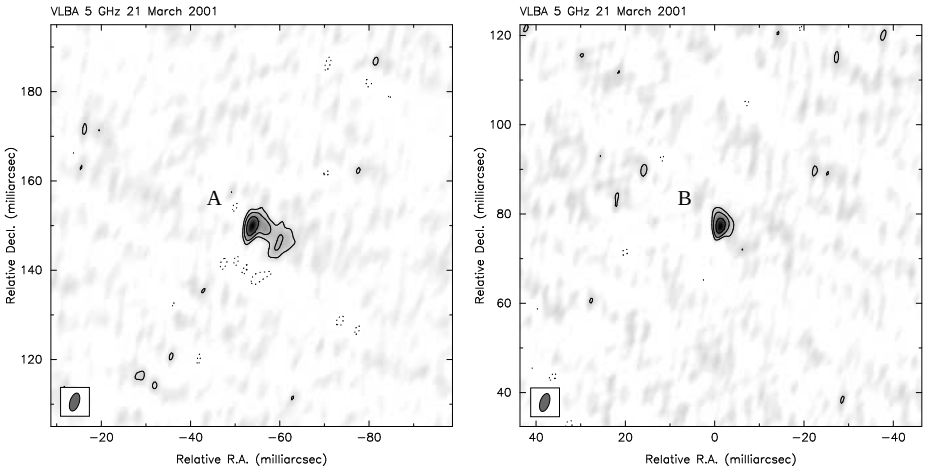

Figure 5.3: VLBA 5 GHz maps detected emission from A and B. Further structure was resolved in image A. No emission was detected from C (McKean et al. 2005).

Follow-up observations of B2108+213 with the Multi-Element Radio Link Interferometric Network (MERLIN) at 5 GHz found the radio spectra of the two lensed images (A and B) to be similar ($\alpha \sim 0.15$ between 5 and 8.46 GHz), with a flux-ratio of $S_B/S_A \sim 0.5$. High resolution radio imaging with the Very Long Baseline Array (VLBA) at 5 GHz found the surface brightness of images A and B to be consistent with gravitational lensing (see Fig. 5.3). A feature extending from the south-west of the compact component in A was

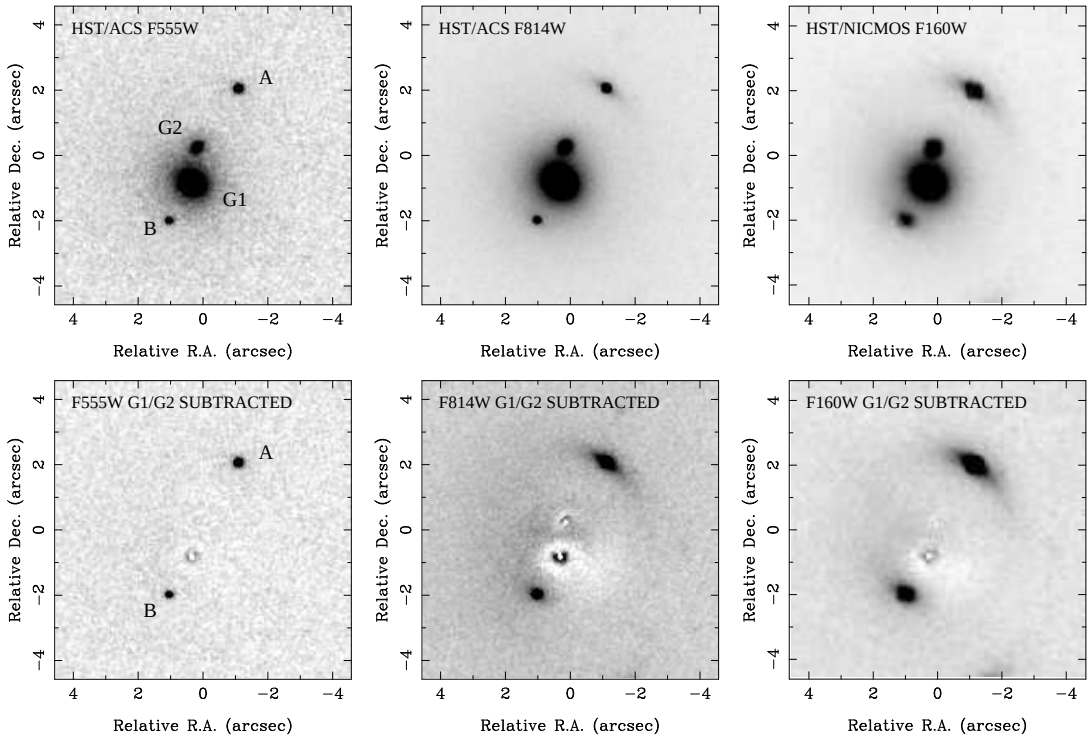

Figure 5.4: *HST V, I* and *H*-band images of B2108+213 clearly show lensed images A and B and a giant elliptical galaxy G1 with a companion G2 within the Einstein radius of the system. The lower panels show G1+G2 subtracted residual images. Note that some extended emission of the background quasar is visible in image A (McKean et al. 2005).

also detected and B showed a single component. Component C was detected and found to be resolved with the deep MERLIN observations whereas the VLBA observations could not detect C because the emission was perhaps resolved out.

5.2.1 Optical/Infrared

The optical and infrared imaging with the *HST* show compact emission from the lensed images A and B. The optical flux density ratio (B:A) is similar to that observed in the radio (see Fig. 5.4). The spectra of A and B are found to be similar which is consistent with A and B being lensed images of the same background source. Since the spectra of the lensed images are featureless, the background source is believed to be a BL Lac type quasar (McKean et al. in prep). Thus, the spectroscopic redshift of the quasar is not known because measuring the redshift with featureless spectrum becomes difficult.

Not surprisingly, a massive elliptical galaxy (G1) is found at the expected position of the lens. Moreover, the third radio component (C) detected with the VLA and MERLIN imaging is coincident with G1. A companion galaxy (G2) is also found within the Einstein radius of the system. However, there is no evidence of emission from G2 in the radio. The residuals are shown in the lower panels of Fig. 5.4 after subtracting a de Vaucouleurs profile for both the galaxies. An asymmetry found in the residual images near G1 is believed due to G1 having a disturbed and complex morphology. Using spectroscopy, the

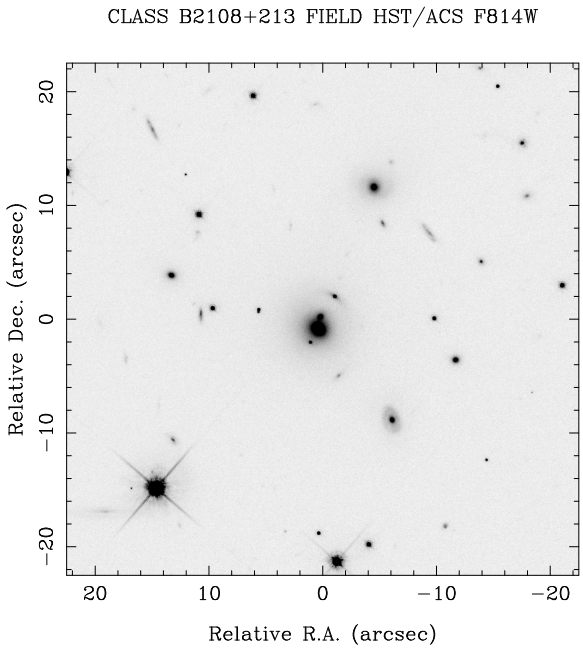

Figure 5.5: *HST I*-band image with a field of view of size 223×223 kpc^2 shows several galaxies in proximity to G1 (center) indicating the presence of a galaxy group (McKean et al. 2005).

redshift for G1 was determined to be 0.365 (McKean et al. in prep). The spectrum of G1 has no strong emission lines and is consistent with an early type galaxy. The redshift of G2 is unknown because the angular separation between G1 and G2 is ∼ 1.1 arcsec. A spectroscopic survey of the environment around the main lensing galaxy has found at least 40 group/cluster galaxies at the same redshift as G1 (z = 0.365, see Fig. 5.5). The galaxy velocity dispersion is 450 km s^{-1}. About 4 galaxies are found within a radius of 12 arcsec (∼ 60 kpc) from G1. Also, the central stellar velocity dispersion of G1 was found to be 360 km s^{-1}. These new data confirm that G1 is a massive galaxy in a dense group/cluster environment (McKean et al. in preparation).

5.2.2 Recent X-ray Analysis

B2108+213 was recently included by Fassnacht et al. 2007 in their sample of moderate redshift ($0.3 < z < 0.6$) galaxy groups. X-ray observations were done with *Chandra* of this system. The field seems to show two main sources of compact emission, one centred at the lens system and the other to the east. The high resolution of *Chandra* could also resolve the compact emission from the AGN of the lensed quasar and the lensing galaxy G1 (see Fig. 5.6) which was masked out to map the diffuse emission. The elongated low surface brightness X-ray contours are stretched along an east-west direction and form a bridge between the two compact sources of emission. The field shows multi-component diffuse emission. Although one of the low surface brightness regions (with a size of ∼2′×1′) is found to be roughly centred on the lens system, the centroid of the overall diffuse emission is offset from the position of the brightest group galaxy (G1). The group members show a

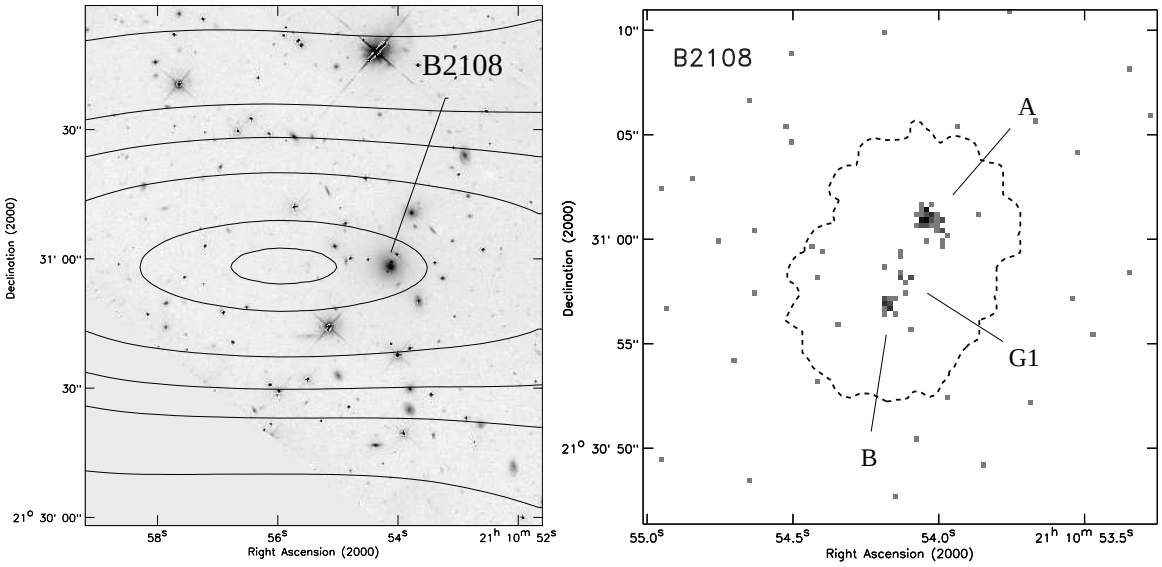

Figure 5.6: The brightest non-stellar object in the grey-scale field of the *HST* image is B2108+213. The overlaid X-ray contours show diffuse emission which might be due to the interaction within the groups of galaxies (Fassnacht et al. 2007, see text). Compact X-ray emission is seen from A, B and the lensing galaxy G1 which was masked out to determine the extended emission.

non-Gaussian velocity dispersion profile indicating a dynamically disturbed system. Their estimated group velocity dispersion is 470±50 km s−1.

5.2.3 Existing Mass Model

A mass model for the system was proposed by McKean et al. (2005) which incorporated both G1 and G2 as singular isothermal spheres with an external shear component. This model successfully reproduced the observed positions and flux-ratios of the lensed images and required only a small contribution from the field (the external shear was ∼ 2 per cent). However, given the limited number of observational constraints that were provided from the two lensed images, this model has no degrees of freedom.

Moreover, lens theory predicts that an extended mass distribution should produce an additional lensed image near the centre of the lens potential (e.g. Rusin & Ma 2001; Keeton 2003a). The detection of such odd images is extremely rare for galaxy-scale systems because the image magnification tends to zero as the inner density profile approaches isothermal ($\rho \propto r^{-2}$). However, the non-detection of core lensed images can place a very strong lower limit to the density profile of the lensing mass distribution (e.g. Rusin & Ma 2001; Keeton 2003a; Boyce et al. 2006a; Zhang et al. 2007). There is currently only one accepted detection of a core lensed image, PMN J1632−0033, whose lensing galaxy has a global power law slope of $\gamma \sim 1.91 \pm 0.02$ (Winn et al. 2002, 2004).

The nature of the third radio component of B2108+213 is not entirely clear. The position of C is very close to that of the main lensing galaxy as expected for the odd image. However, the flux-density of component C appears to be too large to be the core

lensed image; the flux-ratio is $S_C/S_A \sim 0.1$. Therefore, new mass models are also made to determine the status of C as either the third lensed image or the AGN of the lensing galaxy.

5.3 New Observations in the Radio

In this section, new radio imaging of B2108+213 with MERLIN, the VLBA and global VLBI are presented.

5.3.1 MERLIN 1.4 GHz Observations

B2108+213 was observed with MERLIN at 1.4 GHz in two runs. The first set of observations were carried out for 14 h on 2005 March 19 and 20 with all of the MERLIN antennas except for the Lovell telescope. The second observing run, which included all of the antennas, lasted for 14 h and 8 h on 2005 April 15 and 16, respectively. 3C286 and OQ208 were used as the flux density and polarization calibrators. A switching cycle of ~ 2 and ~ 6 min was used between the phase calibrator (B2103+213) and the lens system. The data were taken in a single IF and divided into 15 channels of 1 MHz width each. The data were taken in both the right and left hand circular polarizations. The initial editing was done with the routine and the initial flux-density calibration with the routine using the MERLIN pipeline. Most of the further data reduction and mapping was done using the MERLIN automated pipeline within AIPS. This work was done in collaboration with Dr. Tom Muxlow, Jodrell Bank Observatory (JBO).

The MERLIN imaging detected a combination of compact and very low surface brightness extended emission. The total intensity image was initially cleaned with the AIPS task to subtract the higher surface brightness compact radio components (A, B and C). The residual image, which contained the low surface brightness emission, was then deconvolved with the maximum entropy routine with a starting model of the central part of the residual image smoothed with a circular Gaussian of full width at half maximum (FWHM) of 2 arcsec. Finally, the high surface brightness features were restored with the AIPS task into the low surface brightness map and the combined total intensity map was produced.

In the left panel of Fig. 5.7 shows the total intensity map made from the combined datasets. Components A, B and C in this image are restored with a 0.252×0.165 arcsec^2 beam. For the first time extended emission is found on either side of component C spreading over an area of $\sim10\times2$ arcsec^2. It appears as though the extended jet emission originates from component C. The positions and the flux densities obtained from fitting Gaussian model components to A, B and C with the task in AIPS are given in Table 5.1. The total flux-density for the extended structure, measured by integrating over the region within the 3σ limit, is ~ 70 mJy. Earlier measurements from the NVSS catalogue suggest a flux density of ~ 53 mJy for this system. The difference in these two estimates might be due either to difference in calibration or the maximum entropy method used for mapping the extended emission. The map of linear polarized emission shows only image A, with a polarized flux density of 0.6 ± 0.1 mJy (see the right panel of Fig. 5.7). The rms noise of the polarized image is ~ 0.1 mJy beam^{-1}. Therefore, these observations were not sensitive

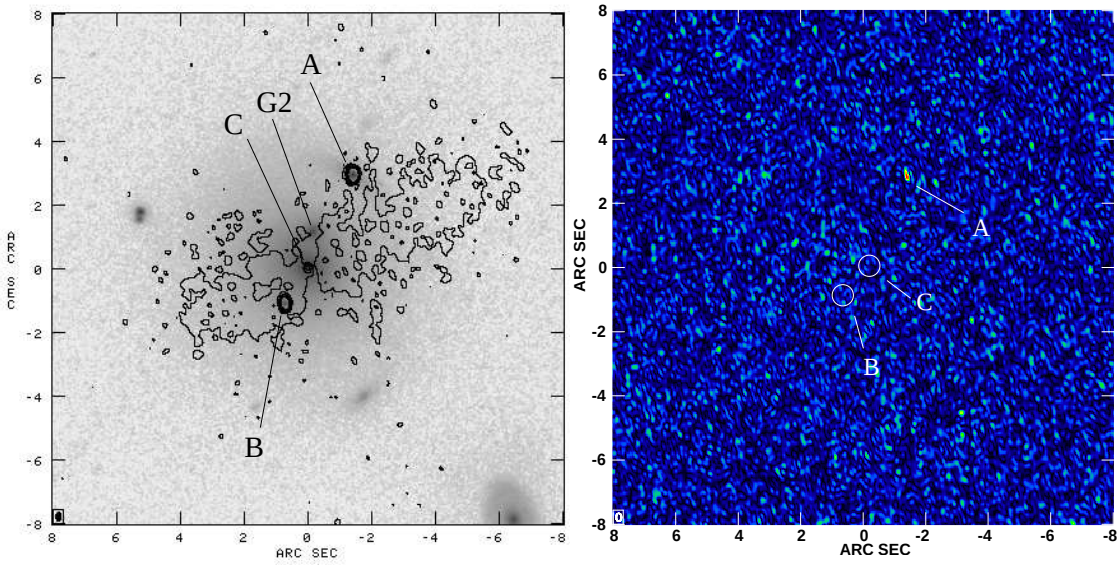

Figure 5.7: Left: HST *I*-band (F814W) image of B2108+213 in grey-scale overlaid with the contours at 1.4 GHz using MERLIN. The lensed images A and B are coincident in optical and radio. The core component C coincident with Galaxy G1 is accompanied by low surface brightness emission extending on either side in the radio. Galaxy G2 does not show any counterpart in the radio. The contours are (−3, 3, 6, 12, 24, 48, 96, 192)×0.03 mJy beam^{-1} (rms in the map). North is up and east is to the left. The MERLIN map was produced in collaboration with Tom Muxlow. Right: The polarized emission at 1.4 GHz with MERLIN is seen from image A only. The polarized emission from B and C, if any, is below the noise level.

enough to detect any polarized emission from image B, assuming that the polarized flux ratio between A and B is same as the flux ratio.

Table 5.1: The flux densities and positions of A, B and C from fitting Gaussian model components to the MERLIN 1.4 GHz map. The separations are measured relative to A.

Comp.	RA (mas)	Dec (mas)	S_{peak} (mJy beam^{-1})	S_{total} (mJy)
A	0.0±0.2	0.0±0.2	16.3±0.8	16.4±0.8
B	2141.3±0.4	−4026.7±0.4	7.1±0.5	7.7±0.4
C	1434.4±6.0	−2915.6±6.0	0.5±0.1	0.6±0.1

5.3.2 VLBA 1.7 GHz Observations

B2108+213 was observed on 2002 June 19 with the VLBA at 1.7 GHz. The aim of this observation was to image any low surface brightness extended jet emission from the lensed images and to detect component C on mas-scales. The observing run lasted for

7.5 h in total, with ~ 5.2 h spent observing B2108+213. The calibrator B2103+213 was used for phase-referencing, with a 5 min time cycle between the lens system (3.5 min) and the calibrator (1.5 min). The data were taken in the left-hand circular polarization through 4 IFs, each with 8 MHz bandwidth. The aggregate bit rate was 128 Mb s^{-1} with 2 bit sampling. All of the VLBA antennas were available with the exception of Los Alamos. The data were correlated at the VLBA correlator where each IF was divided into 16×0.5 MHz channels and averaged over 2-s time intervals. For wide-field imaging with interferometers, bandwidth smearing of the *uv*-data can cause radial smearing of the map features which reduces the peak flux-density of a component far away from the phase centre (see Chapter 3). To limit the effect of bandwidth smearing, the data were correlated at two positions centred on images A and B. Note that for the set-up used here bandwidth smearing effects are significant[2] at a radius of $\gtrsim$ 11 arcsec, which is well outside the maximum lensed image separation (4.56 arcsec).

Both the A and B data sets were inspected in . Since the data set B had a better correlation, only this data set was reduced and analyzed further. The data were amplitude calibrated from the system temperature and gain curve values for each antenna, and corrected for the change in the parallactic angle. The data for B2103+213 were phase and amplitude self-calibrated and the solutions were applied to the lens system B2108+213. The mapping of B2108+213 was carried out using , without further frequency or time averaging. While mapping a wide field of view, a less time consuming technique is to map only the regions of interest. Thus, during the imaging of B2108+213 three sub-fields centred on A, B and C were mapped. An iterative process of cleaning and phase-only self-calibration, using a 9 min solution interval, was used to map the lens system.

In the upper panel of Fig. 5.8, the naturally weighted maps of lensed images A and B, and of radio component C. Image A shows a compact core and a radio-jet extending to the south-west which was fitted by a three component Gaussian model. Previous VLBA imaging at 5 GHz by McKean et al. (2005) found the jet to extend over 10 mas in scale, whereas the new deeper 1.7 GHz map shows more emission extending beyond 20 mas from the radio core. Image B shows a single core component with a hint of jet emission to the north-east. Image B is also fainter than image A, which is consistent with gravitational lensing where the surface brightness of the lensed images is conserved. A single component Gaussian fit successfully modeled the emission from image B. The radio component C was detected for the first time at mas-scales. A single component Gaussian fit shows that the emission from C is compact (deconvovled FWHM is 5.3 mas). The results of fitting Gaussian model components using the AIPS task are listed in Table 5.2. Note that the extended emission detected in the MERLIN imaging has been resolved out here.

5.3.3 Global VLBI 5 GHz Observations

In order to better resolve the north-east extension in image B and to determine the spectral index of component C, a global VLBI observation of B2108+213 at 5 GHz was undertaken. The lens system was observed on 2006 February 17 with the Effelsberg, Jodrell

[2]For a 10 per cent loss in the measured peak flux density of a point source.

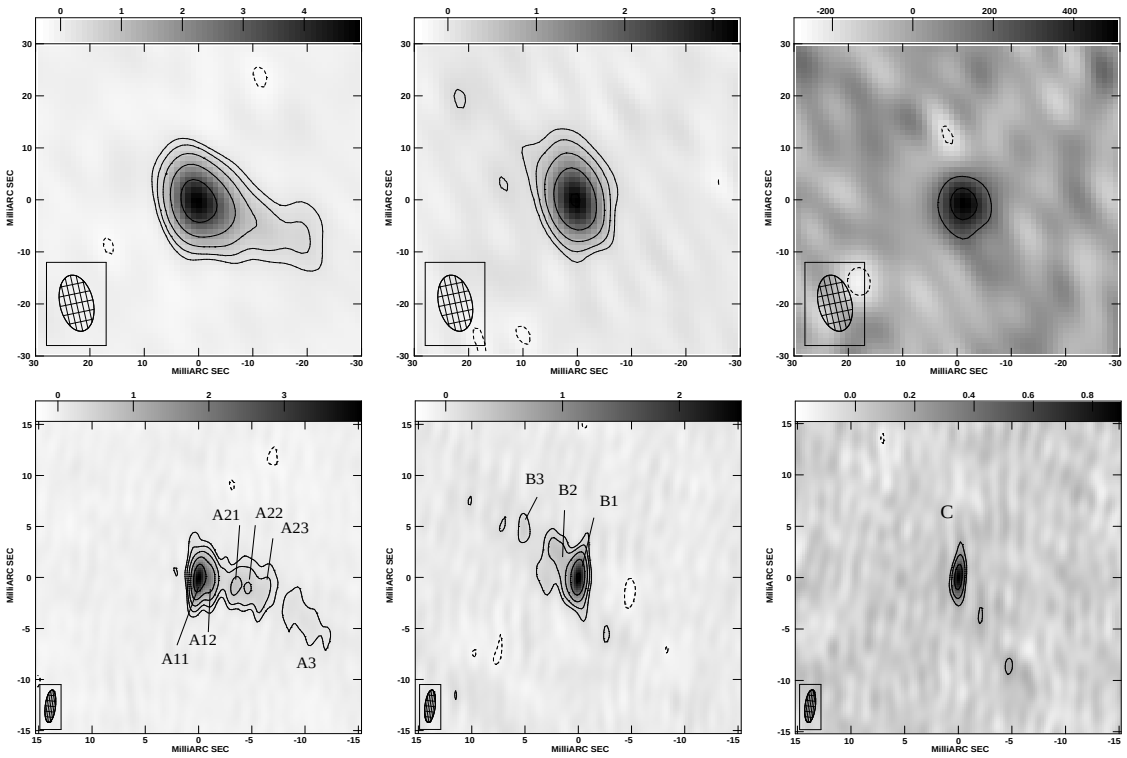

Figure 5.8: The top panel shows VLBA 1.7 GHz maps of B2108+213 (A, B and C) restored with a beam size of 10.9×6.1 mas^2 and position angle of 12.18°. The bottom panel shows Global VLBI 5 GHz maps restored with a beam size of 3.19×0.97 mas^2 and position angle of −6.17°. Component A shows extended structure to the south-west direction at both frequencies. Component B shows north-east extension, which is better resolved at 5 GHz. Component C which is coincident with the main lensing galaxy, is compact at both frequencies. The contours for all the maps are (−3, 3, 6, 12, 24, 48) × σ_{map} given in Table 5.2 for the 1.7 GHz dataset and is 0.05 mJy beam^{-1} for the 5 GHz dataset. North is up and east is left. Grey-scales are in mJy beam^{-1}, except for the 1.7 GHz image of component C which is in μJy beam^{-1}.

Table 5.2: Positions, flux densities and rms map noise of the components at 1.7 GHz.

Component	RA (mas)	Dec (mas)	S_{peak} (mJy beam^{-1})	S_{total} (mJy)	σ_{map} (mJy beam^{-1})
A1	0.0±0.1	0.0±0.1	3.0±0.2	4.6±0.2	0.070
A2	3.0±0.2	−0.9±0.2	2.3±0.1	2.4±0.1	0.070
A3	−12.2±0.5	−5.5±0.5	0.6±0.1	1.7±0.1	0.070
B	2135.9±0.1	−4030.5±0.1	3.3±0.2	4.3±0.2	0.065
C	1429.0±0.5	−2889.9±0.5	0.5±0.1	0.5±0.1	0.065

Bank Mk2, Westerbork, Medicina, Torun, Noto, Green Bank and 10 VLBA antennas. The 14.5 h long observation was taken in both the right and left-hand circular polarizations, through 4 IFs each with 8 MHz bandwidth and a 256 Mb s^{-1} bit rate. Here, the 5 min time cycle between the lens and the calibrator (B2103+213) was divided into 3 and 2 min scans, respectively for phase-referencing. The correlation was carried out at JIVE[3] where the data in each IF were divided into 16×0.5 MHz channels and time-averaged over 1-s intervals. The maximum field-of-view defined by bandwidth and time-averaged smearing is $\sim$ 8 arcsec with this setup. Therefore, a single correlation of the data with the phase centre at the mid-point between lensed images A and B was obtained.

The data were reduced using in a similar manner to the 1.7 GHz VLBA observations. The Green Bank Telescope had severe problems throughout the observing run, hence it was removed completely from the dataset. All baselines with Torun were discarded because of a poor amplitude calibration. Throughout the data reduction process all of the antennas were given equal weights to stop the large antennas (like the 100 m Effelsberg telescope) dominating the *uv*-dataset. This avoids higher side-lobes and a deterioration of the image quality. The data were reduced by adopting a procedure similar to that followed at 1.7 GHz. The phase calibrator (B2103+213) was mapped initially to determine the phase corrections which were applied to B2108+213. Subsequently, the data for B2108+213 were self-calibrated with a 3-min long solution interval, and ed using . The naturally weighted global VLBI maps of B2108+213 are presented in the lower panel of Fig. 5.8.

Image A shows the same core-jet structure observed at 1.7 GHz and previously at 5 GHz (McKean et al. 2005). These new Global VLBI data clearly resolve image A into three main components A1, A2 and A3. Component A1 is made up of two compact sub-components A11 and A12. Component A2 appears to be an extended knotty jet feature, which is further divided into three sub-components, A21, A22 and A23. The third component (A3) of image A is a faint jet feature which bends flux density towards south. Components A1 and A2 were fitted with two and three elliptical Gaussian models respectively. Since it is not possible to fit more than 4 components simultaneously in , Powell's method was used for model fitting. Other AIPS tasks that can deal with more than four Gaussian at once work in the *uv*-plane. However, given the wide field and multiple regions of interest in the target source, each with a complicated structure, the best option was to work in the image plane. Powell's minimization method can be used to work in the image plane, like the AIPS task . Fig. 5.9 shows the observed (in red) and the model (in green) surface brightness distributions of components A1 and A2. Component A3 could not be well fitted with multiple Gaussian models. Therefore, the flux density was measured by summing all of the surface brightness emission within the 3σ boundary. The position of A3 was obtained from the surface brightness peak. Note that due to the different angular resolutions and frequency dependent structure, the components (A1, A2 and A3) at 1.7 GHz do not correspond to those observed at 5 GHz – this is simply a naming convention.

The map of image B shows the expected core-jet structure to the north-east. B1 is identified as the core, and B2 and B3 as the counterparts to the jet features detected from the 5 GHz imaging of A. The emission from image B was fitted with three elliptical

[3]Joint Institute for VLBI in Europe

Table 5.3: Positions and flux-densities of the fitted Gaussian components for the 5 GHz data. Component A3 was not fitted with a two-dimensional Gaussian (see Section 5.3.3 for details).

Component	RA (mas)	Dec (mas)	S_{peak} (mJy beam^{-1})	S_{total} (mJy)
A11	0±0.1	0±0.1	3.57±0.43	3.81±0.47
A12	−1.1±0.1	0.0±0.1	1.85±0.31	2.63±0.34
A21	−3.5±0.1	−0.6±0.1	0.72±0.20	1.18±0.21
A22	−4.7±0.1	−0.5±0.1	0.30±0.13	0.32±0.13
A23	−6.0±0.1	−1.3±0.1	0.45±0.16	1.42±0.17
A3	−9.8±0.5	−3.9±0.5	0.38±0.15	0.43±0.22
B1	2135.2±0.1	−4030.0±0.1	2.31±0.34	2.90±0.37
B2	2137.0±0.3	−4028.0±0.3	0.45±0.16	1.08±0.17
B3	2140.5±0.4	−4025.2±0.4	0.29±0.13	0.32±0.13
C	1430.7±0.1	−2888.1±0.1	0.86±0.10	0.86±0.10

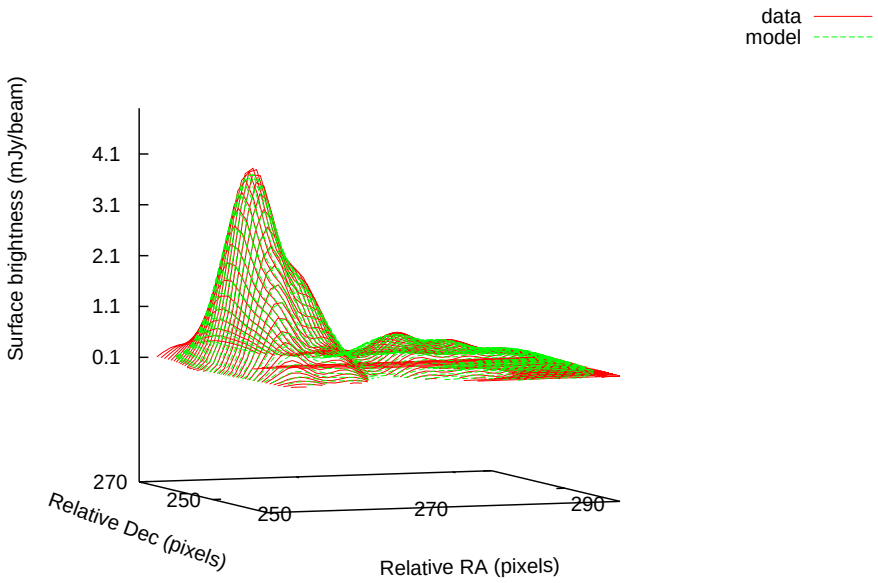

Figure 5.9: A plot of surface brightness distribution of components A1 and A2 together in image A at 5 GHz overlaid with the best-fitting model. By using Powell's minimization routine the data is fitted in the image plane with a multiple Gaussian component model to measure the flux densities and peak intensity positions of the components.

Gaussian components, using the same minimization method as for image A. Note that the small feature to the south of component B2 was excluded during the fitting process. Component C was fitted with a compact elliptical Gaussian model (deconvolved FWHM is 0.8 mas) at 5 GHz. There is no evidence of any collimated jet emission towards the extended lobes detected in the 1.4 GHz MERLIN observations. The fitted positions and flux densities of the Gaussian model components are presented in Table 5.3. The errors for the flux densities of A and B were calculated based on the principle described in Fomalont (1999) (see Appendix C), and the errors on the positions were taken from the AIPS task

.

5.4 Mass Models

Mass models for the B2108+213 lens system are now investigated. The current best model has two singular isothermal spheres for galaxies G1 and G2, and an external shear (McKean et al. 2005). This model reproduces the observed positions and flux densities of the lensed images, but has no degrees of freedom due to the limited number of observational constraints. Using the new high resolution VLBI observations presented here, several cases are dealt with. For instance, whether the image configuration can be better explained by a single lensing galaxy (G1), or by two galaxies (G1 and G2). Furthermore, whether component C can be a core lensed image is also tested. Finally, the influence of the environment on the mass model is investigated. The lens modeling was carried out using the publicly available code (Keeton 2001).

5.4.1 Case 1: C as the Active Nucleus of Lensing Galaxy G1

The position of radio component C was used as the position for the lens galaxy G1 in this case. As can be seen from Table 5.4, the position obtained for component C from the VLBI imaging is not consistent with the position measured for G1 from the *I*- and *H*-band imaging (F814W and F160W, respectively) with the *Hubble Space Telescope*. This is almost certainly due to G1 having a complicated surface brightness profile, particularly in the infrared. However, the position for component C is consistent with the *V*-band (F555W) imaging. Therefore, the position of G2 and the flux ratio between G1 and G2, which are used in the models, are taken from the *V*-band data.

As expected from lensing, the three main components of image A can be identified as the counterparts of the components in image B. However, image A is further resolved and shows several sub-components in A1 and A2. The compact sub-component A11 is probably the core, hence it was used as the position for A1. As component A2 is much more extended, a flux density weighted position was determined from the sub-components ($\Delta\alpha = -3.8\pm0.2$ mas, $\Delta\delta = -0.7\pm0.2$ mas relative to A11). For component A3 and all of the components of lensed image B, the positions from Table 5.3 were used. The flux densities for only A1 (i.e. A11+A12) and B1 were used for the modeling. The total number of constraints to the lens models provided from the lensed images is 13.

First, the lens as a single galaxy G1, with a singular isothermal sphere (SIS) density profile and an external shear is considered. This model has 9 parameters (1 for the Einstein radius, 2 for the shear and position angle, and 6 parameters for the three source positions)

Table 5.4: The relative positions of both the galaxies from A and the magnitudes measured with the *HST* in *V, I* and *H* bands (McKean et al. 2005).

Band	Filter	Comp (mas)	RA (mas)	Dec (Vega)	Magnitude
V	F555W	A	0±1	0±1	22.56±0.15
		B	2131±3	−4032±4	23.73±0.17
		G1	1427±5	−2890±5	19.91±0.11
		G2	1273±5	−1786±5	22.59±0.15
I	F814W	A	0±1	0±1	20.96±0.15
		B	2131±2	−4032±3	21.97±0.17
		G1	1428±2	−2888±2	17.63±0.11
		G2	1275±2	−1786±2	20.61±0.15
H	F160W	A	0±1	0±1	18.14±0.15
		B	2131±2	−4032±3	18.86±0.17
		G1	1421±2	−2888±2	15.65±0.11
		G2	1263±1	−1795±2	18.36±0.15

and 4 degrees of freedom (dof). The total χ^2_4 of the model is 7.6, where the subscript refers to the dof, and the results are given in Table 5.5. This simple model fits the data well, and requires a large Einstein radius of 2.18 arcsec and an external shear of 0.05 at a position angle of 107.3 deg. The uncertainties on the fitted model parameters that are given in Table 5.5 are from the 1σ confidence levels (see Press et al. 1992). Fig. 5.10 shows χ^2 as a function of each of the parameters that were varied for this model. When fitting a single parameter, the 1σ confidence level is obtained for the χ^2 corresponding to $\chi^2 - \chi^2_{min}$ =1.

Now the contribution of G2, if included as part of the lens potential, is tested. Hence, galaxy G2 is introduced as a companion lensing galaxy at the same redshift as G1. A SIS centred on the optical position of G2 was added to the SIS+shear model for G1. For an isothermal sphere, the Einstein radius (b) can be related to the velocity dispersion i.e. $b \propto \sigma^2$ and hence, the mass. Also, from the Faber-Jackson relation (Faber & Jackson 1976), the luminosity is proportional to the velocity dispersion ($L \propto \sigma^4$). Therefore, the ratio of the optical luminosities of G2 and G1 (see Table 5.4) are used to constrain their mass ratio by equating $b_{G2}/b_{G1} = 0.29\pm0.05$. Here, the luminosities are calculated assuming that G2 is at the same redshift as G1 and the error on the ratio takes into account the errors on the apparent magnitudes only. Note that including G2 in this way does not increase the degrees of freedom. The total χ^2_4 of the model is 6.95 which is only marginally better than the single SIS+shear model used for G1. The fitted parameters are given in Table 5.5. Including G2 lowers the Einstein radius of G1 to 1.71 arcsec and reduces the shear to 0.04.

Note that if the ratio of the masses is not fixed and the Einstein radius of G2 is left as a free parameter, then the Einstein radius of G1 is lowered further to 0.76 arcsec and the shear is reduced to 0.01. Also, the Einstein radius of G2 increases to 1.5 arcsec. Here, the simplest model that can be fitted needs to have the main lensing galaxy positioned

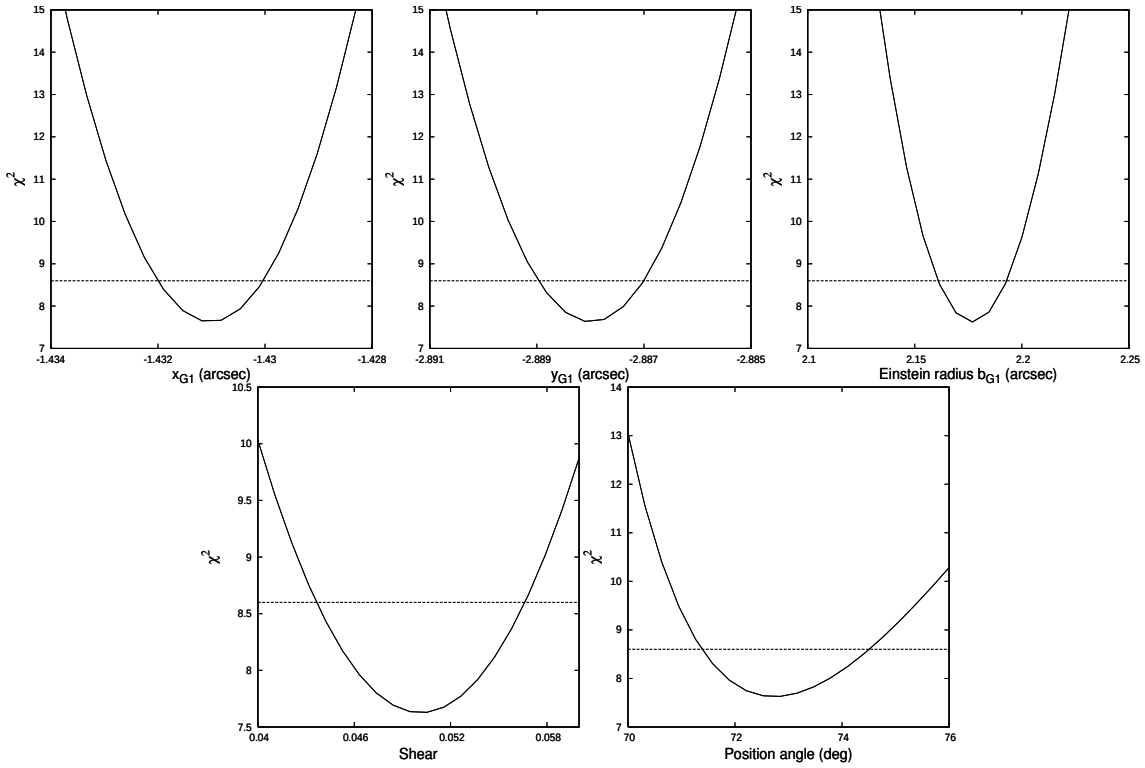

Figure 5.10: The χ^2 as a function of the parameters for the SIS+shear model are plotted. The dashed horizontal lines are for $\chi^2 = \chi^2_{min}$+1. The corresponding values give the 1-σ confidence level for the respective parameters.

co-linearly between the two lensed images. Since G2 is closest to this, it is taken as the main perturber and the shear is reduced. The reduced χ^2 of this model is 0.9, which is lower than for the case when the mass ratio is fixed (reduced χ^2 = 1.7) and for the single galaxy model (reduced χ^2 = 1.9). In Fig. 5.11, the χ^2 plot of the Einstein radii of G1 and G2 is shown. There is a clear degeneracy between the Einstein radii of G1 and G2. However, since G1 is clearly the dominant mass clump of the system as shown from the optical data (see Figs. 5.4 and 5.5), the prior from the optical luminosity is used for the SIS+SIS+shear model to break this degeneracy and to produce a physically plausible model.

So far, only an external shear was used to account for any mass ellipticity in the model. It is clear that some ellipticity is required to account for the non-collinearity between the lensed images and lensing galaxy G1. The ellipticity in the model could be due to G1 having an elliptical halo. Hence, the external shear was replaced with an elliptical mass distribution for G1, while a SIS is fixed at the position of G2. The dof of this singular isothermal ellipsoid (SIE)+SIS model is 4, and the resulting reduced χ^2 is 1.6. This model fits the image positions slightly better than previous models without changing the fitted parameters significantly (see Table 5.5). The ellipticity of G1 is found to be 0.135 at a position angle of 105.5 deg. The ellipticity of the surface brightness profile of G1, as measured from the *HST* data, is 0.14 at a position angle of 57 deg (McKean et al. 2005). The offset between the position angle of the halo and the light distribution of G1 suggests

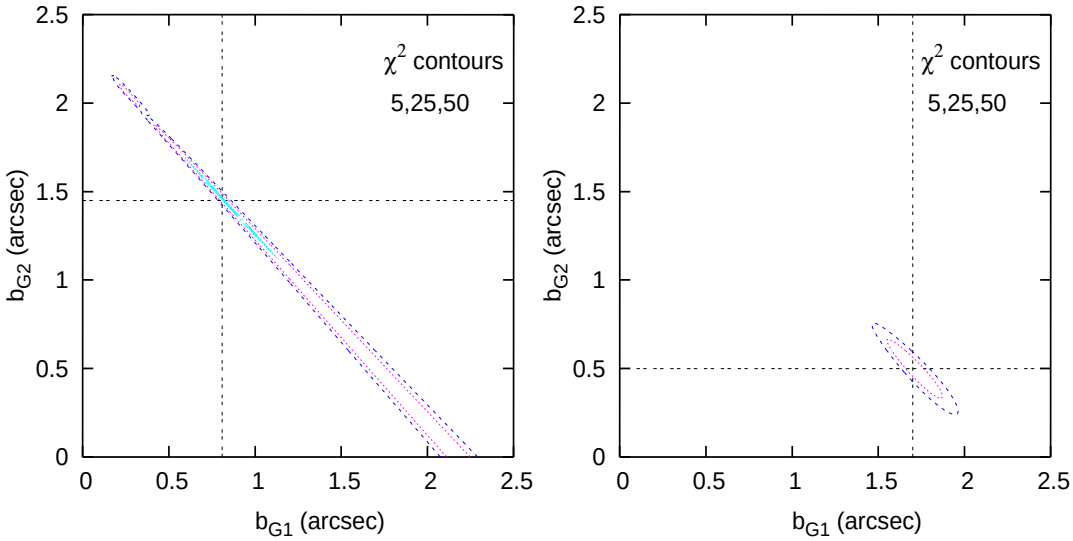

Figure 5.11: A degeneracy between the Einstein radii of G1 and G2 for the SIS+SIS+shear model is found (left). Without any constraint on the masses of the galaxies, the best fit model requires G2 to be four times more massive than G1, which contradicts the optical luminosities. Introducing a constraint on the mass ratio between G1 and G2 breaks this degeneracy (right). The intersection point of the dotted lines marks the minimum in both the plots.

that the environment may also be affecting the shape of the lensing potential. However, since there are only two lensed images of B2108+213 it is not possible to constrain both an elliptical potential for G1 and an external shear. Moreover, ellipticity and shear (and also the presence of another lensing galaxy, G2) will produce similar observable effects. Thus, the actual origin of the asymmetry between the lensed images and the lensing galaxy G1 is difficult to verify. For instance, using a SIE+SIS+shear as a toy model artificial data were generated. The data were modified arbitrarily and a SIE+SIS+shear model was fitted to these data by varying shear and ellipticity. Fig. 5.12 shows the degeneracy between different values of shear and ellipticity which combine to give equal χ^2 values.

5.4.2 Case 2: C as the Core Lensed Image

Extended mass distributions are expected to produce odd number of lensed images with the oddth image near the position of the lensing galaxy (Dyer & Roeder 1980; Burke 1981). However, the mass density distribution close to the centre of the lensing galaxy will affect the magnification of the core lensed (odd) image. For example, an isothermal density profile will completely demagnify the core lensed image. Since galaxies are known to have global density profiles close to isothermal (e.g. Koopmans et al. 2006b), searches for core lensed images tend to concentrate on asymmetric double image systems (e.g. Boyce et al. 2006a; Zhang et al. 2007), that is, lens systems where the A:B flux ratio is greater than 10:1. This is because in these cases the magnification of the core lensed image is highest and these systems offer the best possible chances of detection. The only known example of a galaxy-scale lens system with a core lensed image is PMN J1632−0033 (Winn et al. 2004), which is an asymmetric double with a flux ratio of

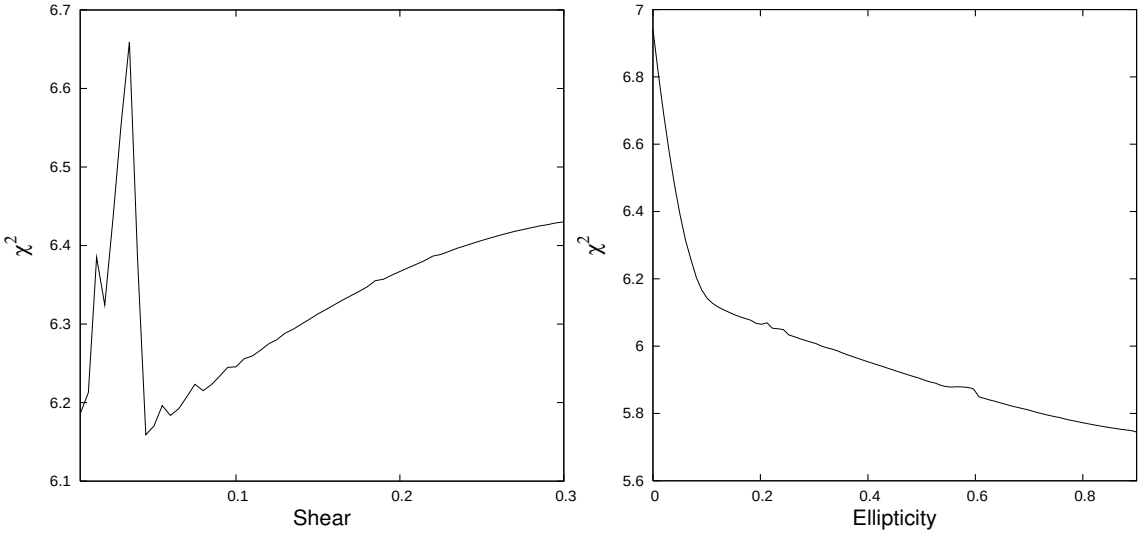

Figure 5.12: The χ^2 function for the given range of shear and ellipticity values of a SIE+SIS+shear does not show a well-behaving function with a minimum. The χ^2 increases, as the shear is increased or as the ellipticity is decreased which implies they are degenerate.

15:1. For B2108+213, the flux ratio of image A to B is about 2:1, so it seems unlikely that component C could be a core lensed image. However, given the possibility that the lens potential may be made up of several galaxies, the overall density profile might be shallower than isothermal. Therefore, different mass models to establish the nature of component C are now explored.

Throughout this section the position of galaxies G1 and G2 are taken from the *V*-band data (McKean et al. 2005) and the mass ratio is fixed. The positions and flux densities of A, B and C are taken from the global VLBI 5 GHz data (Table 5.3). For the isothermal models presented in Section 5.4.1, no core lensed image is produced. Therefore, mass distributions that deviate from isothermal close to the centre of the lensing galaxy G1 are investigated by inserting a core radius at the inner part of the halo.

A non-singular isothermal Sphere (NIS) halo is used for G1 which has an isothermal mass profile with a flat density core (i.e. $\rho \propto r^0$). The second lensing galaxy G2 is too far away for its core properties to significantly affect the magnification of the central image. Therefore, an SIS profile is chosen for galaxy G2. An external shear is also included in the model. This new model has three additional constraints from the position and flux density of component C and 10 free parameters to fit, thereby giving 6 dof. The model fits the position of component C as a core lensed image well, but cannot fit the flux density, resulting in a very high total χ^2_6 of 88.6. The best fitting core radius is 0.1 mas ($\equiv$ 0.5 pc) and gives a flux density of 10^{-6} mJy for the core lensed image, which is still several orders of magnitude fainter than the observed flux density of component C. This model fails because a much higher core radius is required to fit the flux density of C, but this will be at the expense of fitting the position of component C and the positions and flux densities of the lensed images A and B. Therefore, it seems very unlikely that the third radio component is a core lensed image because the observed flux-density is just too high.

Table 5.5: The fitted parameters with the respective 1-σ uncertainties for the single lensing galaxy case (SIS+shear), and the two lensing galaxy cases (SIS+SIS+shear and SIE+SIS) with the mass ratio of G2 and G1 fixed. The angular separations (relative to image A1) and Einstein radii are in milli-arcsecond. The position angle of the shear and ellipticity is measured in degrees east of north. The superscript '*' refers to the values of the parameters resulting from the fitted models.

Name	Fitted values		
	SIS+shear	SIS+SIS+shear	SIE+SIS
G1 (α, δ)	1431±1	1431±1	1431±1
	−2888±1	−2888±1	−2888±1
G2 (α, δ)		1273±5	1273±5
		−1786±5	−1786±5
G1 Einstein radius	2177±16.0	1711±38	1736±38
G2 Einstein radius		498±39	506±39
External Shear	0.050±0.007	0.038±0.005	
Position angle	$107.3^{+1.3}_{-1.8}$	$105.6^{+1.5}_{-2.0}$	
Ellipticity			0.135±0.017
Position angle			$105.5^{+1.6}_{-2.1}$
Source 1 (α, δ)*	989.4	1059.0	1057.6
	−2110.3	−2058.8	−2021.6
Source 2 (α, δ)*	987.7	1057.6	1056.2
	−2110.1	−2058.3	−2021.1
Source 3 (α, δ)*	985.4	1055.7	1054.4
	−2111.3	−2059.1	−2021.7
S_{BA} (S_{B1}/S_{A1})*	0.64	0.63	0.63

5.4.3 Limits on the Density Profile of G1

For dark matter dominated structures like galaxy clusters, the inner density profile of the mass distribution tends to be quite shallow (i.e. $\gamma \sim 1$–1.5). This is based on the results from numerical simulations (e.g. Navarro, Frenk & White 1996; Moore et al. 1998) and from gravitational lensing (e.g. Sand et al. 2005). In this section, an attempt is made to place constraints on the density profile of the lensing halo, with the aim of determining whether it is consistent with an isothermal mass profile, or a shallower dark-matter dominated profile. To do this, the core lensed image is assumed not to be coincident with the component C. An upper limit to the flux-density of a core lensed image is then derived, using the 3σ flux-density limit determined from the rms map noise ($S_C \leq 150\ \mu$Jy) from the 5 GHz data.

A group halo with a power-law density profile, $\rho \propto r^{-\gamma}$, centred on the position of G1 is introduced. In addition, a core radius of 0.5, 50, 150 and 250 pc is incorporated. The model includes an external shear. The VLBI 5 GHz data for lensed images A and B are used to constrain the model. The Einstein radius, external shear and position angle are

then optimized for density profiles with γ between 1.4 and 2.3, and for each value of the core radius. In Fig. 5.13, the relative magnification between lensed image A and a possible core lensed image C as a function of power-law density profile slope is shown. An upper limit to the flux-ratio of a possible core lensed image and image A is also marked. First, it is clear that steeper density profiles result in a more demagnified core lensed image, as expected. Also, increasing the core radius from 0.005 to 250 pc, increases the magnification of the core lensed image. Interestingly, the upper limit to the relative magnification requires that the density profile of any group halo be steeper than ~ 1.5.

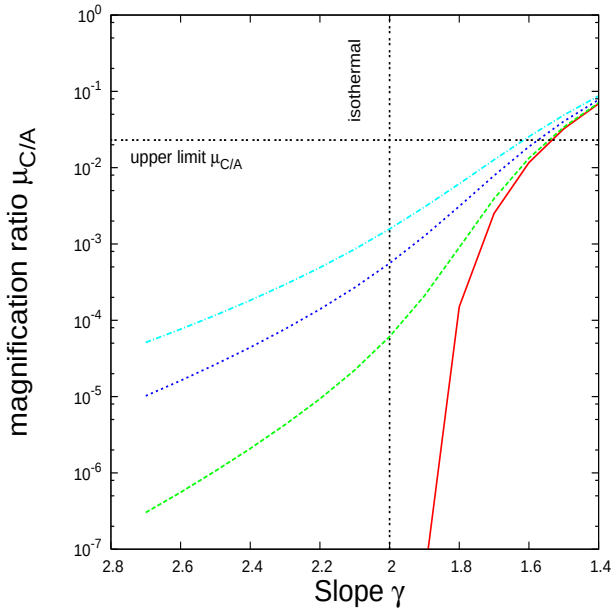

Figure 5.13: The relative magnification of image A and a possible core lensed image C as a function of density power-law slope for a single lensing galaxy (G1) in the absence of a supermassive black hole. The solid (red) line is for a core radius of 0.005 pc. The dashed (green), dotted (blue) and dashed-dotted (cyan) lines are for core radii of 50, 150 and 250 pc, respectively. An upper limit to the relative magnification between images A and C is shown.

The presence of a radio-loud AGN implies that there is also a supermassive black hole at the centre of the lensing galaxy G1. The presence of a black hole is expected to increase the number of core lensed images to two, or possibly demagnify a core lensed image completely, depending on the mass of the black hole (Mao, Witt & Koopmans 2001; Rusin, Keeton & Winn 2005). Using the stellar velocity dispersion of the lensing galaxy G1 ($\sigma_v = 360$ km s^{-1}; McKean et al. in preparation) and the known correlation between black hole mass and the stellar velocity dispersion of the host galaxy, the black hole associated with the AGN within G1 is found to have a mass of $\sim 10^9$ M$_\odot$. Including this black hole as a point mass at the centre of the G1 halo results in any core lensed image being completely destroyed.

The mass models which have G1 as the only lensing galaxy (SIS), or also include G2 as a companion lensing galaxy (SIS) to G1 (SIS or SIE), fitted the positions of the lensed images well, but failed to recover the flux-ratio; the observed flux ratio is $S_B/S_A = 0.45$

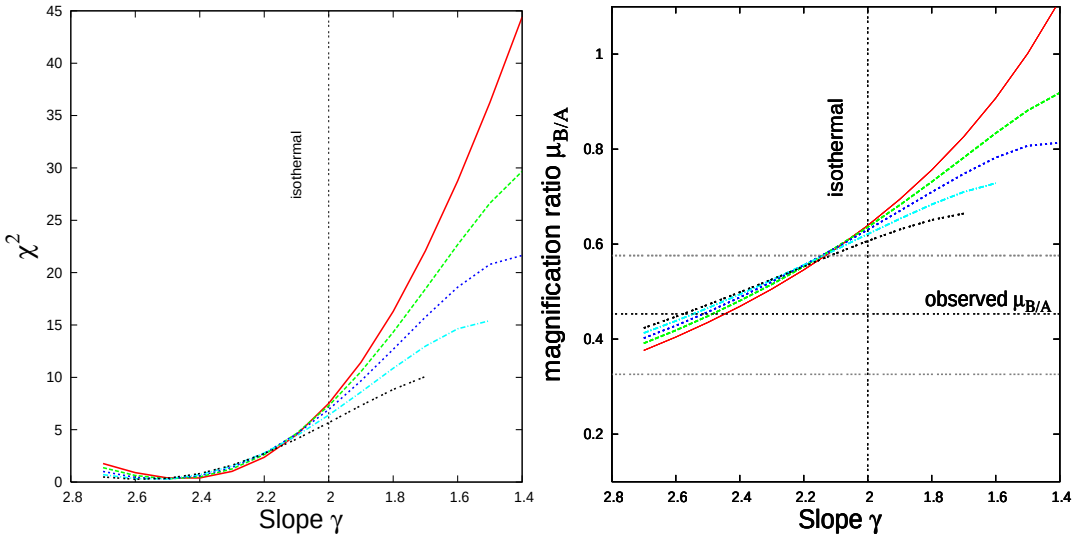

Figure 5.14: The χ^2 (left) and the relative magnification between images A and B (right) as a function of density power-law slope for a single lensing galaxy (G1). The observed flux-ratio at 5 GHz is shown with a conservative 20 per cent uncertainty in the flux-densities of images A and B. The solid (red) line is without the inclusion of galaxy G2. The dashed (green), dotted (blue), dashed-dotted (cyan) and double-dotted (black) lines are for models that include G2 as a SIS with an Einstein radius of 0.3, 0.5, 0.7 and 0.9 arcsec.

at 5 GHz whereas the modeled flux-ratio is ∼ 0.64 (see Table 5.5). The optical flux ratios are 0.47, 0.40 and 0.23 for the *V*, *I*, and *H*-band data from the Table 5.4, respectively. This may suggest substructure (see section 5.5.1) or that the overall density profile of the lens system differs from isothermal, possibly due to the surrounding environment. Therefore, a different density profile is investigated to test whether it is a better fit to the data. A variable power-law density profile is used again for G1 in this spherical mass model. Since an isothermal mass profile is no longer used for G1, it is not possible to fix the ratio of the Einstein radii of G1 and G2 from their optical luminosities via the Faber–Jackson relation. Therefore, G2 is included as a singular isothermal sphere with a fixed Einstein radius of 0, 0.3, 0.5, 0.7 and 0.9 arcsec to test different scenarios. An external shear is also included. This model has 9 parameters and 13 constraints. In the left panel of Fig. 5.14, the model χ^2 is shown as a function of power-law density slope and for different Einstein radii for G2. It is clear that those models with a steeper density profile are a better fit to the data. However, increasing the Einstein radius of G2 also improves the fit, but not to the extent as to justify shallower mass profiles for G1. Note that for the cases where the Einstein radii of G2 is 0.7 and 0.9 arcsec, shallow mass profiles for G1 (i.e. $\gamma \leq 1.7$ and 1.5, respectively) produce four lensed images due to the introduction of a new minimum and saddle point in the time-delay surface. Therefore, these models can be ruled out.

In the right panel of Fig 5.14, the predicted flux-ratio of images A and B is shown as a function of the density profile of G1 and for the same set of Einstein radii for G2. The observed flux-ratio at 5 GHz, with a 20 per cent error on the flux densities of images A and B, is also shown. It is clear that shallow density profiles are inconsistent with the flux ratio of A and B, if the halo is fixed at the position of G1 (assuming the flux densities

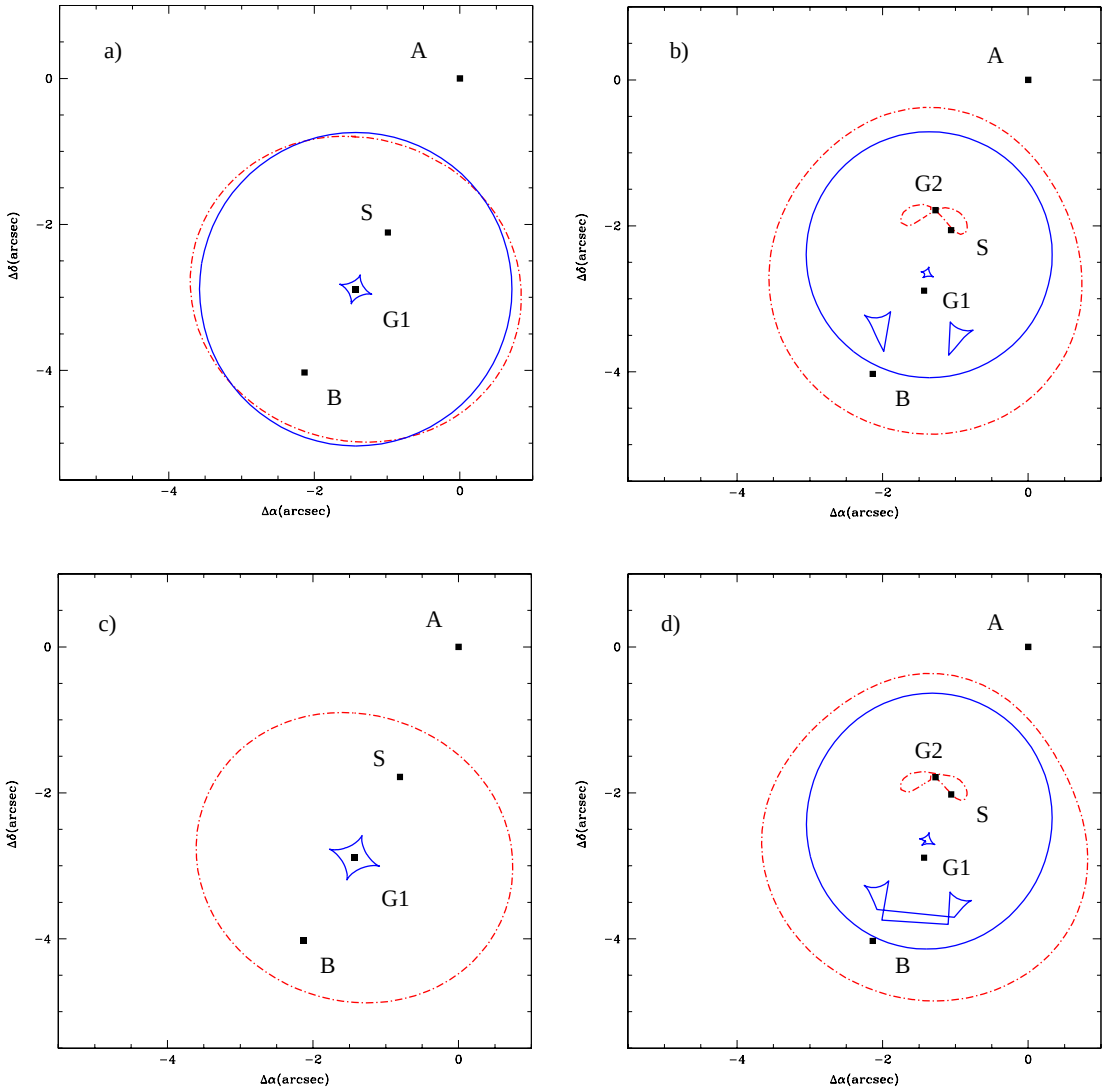

Figure 5.15: Critical curves (dashed-dotted red lines) and caustics (solid blue lines) for a single galaxy G1 are shown on the left and including the effect of galaxy G2 are shown on the right. The figures in the panel represent the following cases: a) SIS+shear b) SIS+SIS+shear c) power law for G1 (with a slope of $\gamma = 2.45$) + shear and d) SIE+SIS. Also, in panel (c) the radial caustic is not shown owing to its extremely large (~400 arcsec) extent which is expected for profiles steeper than isothermal.

are not significantly affected due to variability and a time delay). Furthermore, models with shallower profiles (i.e. $\gamma \leq 1.7$) require the flux-density of B to be close to that of A, which is not consistent with the observed flux-ratio. The best fitting models require the density slope for G1 to be steeper than isothermal (i.e. $\gamma = 2.45^{+0.19}_{-0.18}$, for a single spherical mass model), and are fairly insensitive to the Einstein radius of G2.

Fig. 5.15 shows the critical curves in the image plane and the caustics in the source plane for four lens models; SIS+shear, SIS+SIS+shear, SIE+shear and power-law for G1+shear. The filled squares represent the observed positions of the images and galax-

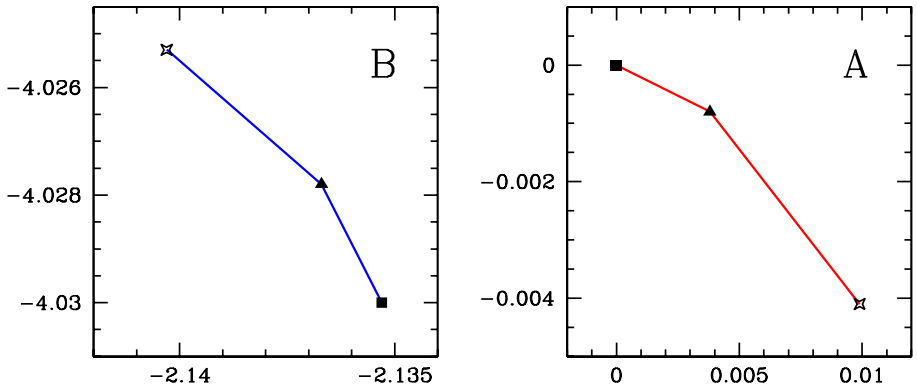

Figure 5.16: Counterpart components of the lensed image A in lensed image B are identified with similar symbols. The opposite parity of image B can be seen with respect to image A. The angular separations on x-axis (relative RA) and y-axis (relative Dec) are in units of mas.

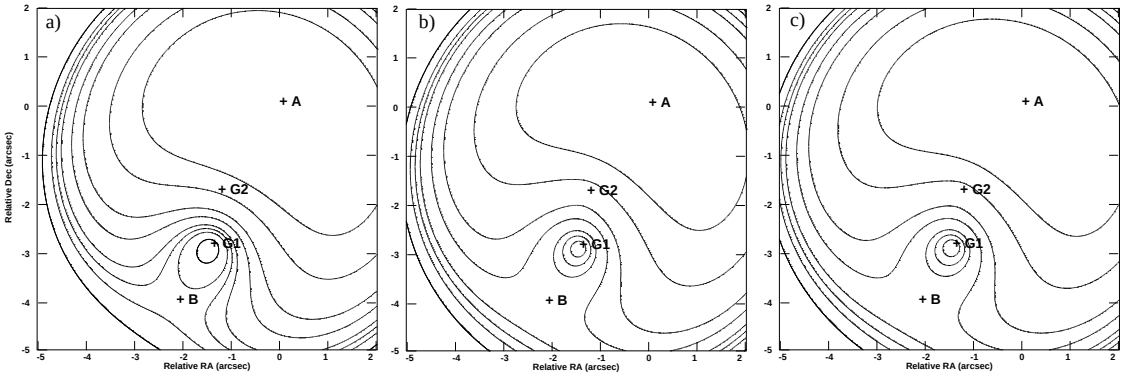

Figure 5.17: Time delay surface plots for a) SIS+shear b) SIS+SIS+shear and c) SIE+SIS models. The levels correspond to -256,-192,-128,-96,-64,-48,-32,-24 and -16 in units of days with respect to the central maximum near G1. The predicted time delays between A and B for the respective models are 251.4, 191.3 and 199.7 days assuming a source redshift of 0.74 (see 5.5.4).

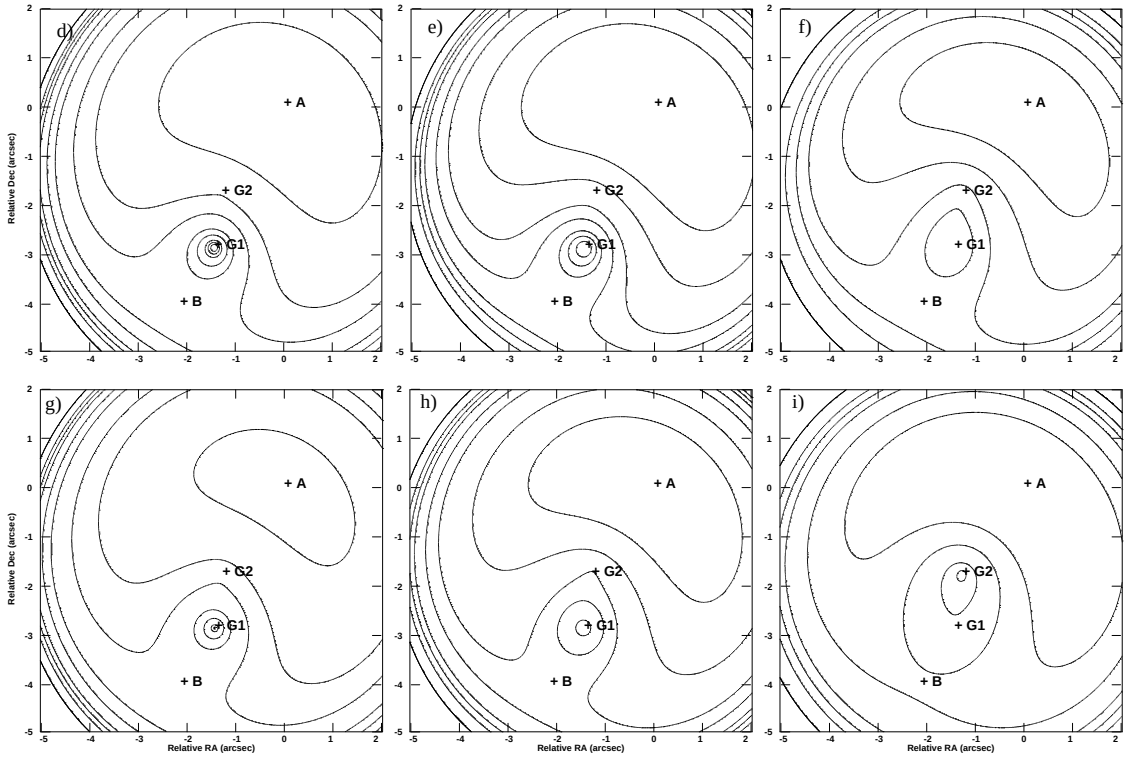

Figure 5.18: Time delay surface plots for models with decreasing power law profiles for G1 (i.e. γ = 2.7, 2.0, 1.4; from left to right) and increasing Einstein radii for G2 (b_{G2} = 0.5 and 0.9 arcsec; from top to bottom). The time delay contour levels are the same as shown in the previous figure.

ies, and the predicted position of the background source from each of the models. The background source is located between the radial caustic (blue circle) and the tangential caustic (blue astroid) as expected for a two image configuration. Note that the radial critical curves are not produced due to a singularity in the density profile. The tangential critical curve (red circle) is an approximate estimate of the image separation. It also predicts the parity of the images depending on their location. Image A has positive parity i.e. the parity is same as that of the background source. On the other hand, image B shows negative parity i.e the parity is opposite of that of the background source. This lensing effect can be verified in images which have more than two non-collinear components as is the case here (see Fig. 5.16).

B2108+213 has not been monitored for variability and no time delay measurements are known yet. Nevertheless, the fitted mass models can be used to predict time delays for the images. The time delay contours are shown in Fig. 5.17 and Fig. 5.18. However, since the redshift of the background source is not known, it is assumed to be at z = 1.5. Therefore, the predicted time delays are for an illustrative purpose only. The time delay surfaces showed in Fig. 5.17 are for SIS+shear, SIS+SIS+shear and SIE+SIS mass models. The time delay surfaces in the upper panel of Fig. 5.18 are for a power-law profile for G1 (i.e. γ = 2.7, 2.0, 1.4) and an SIS with an Einstein radius of 0.5 arcsec for G2. The lower panel shows the time delay surfaces for the same models but with a larger Einstein

radius of 0.9 arcsec for G2. The upper panel (or the lower panel) indicates that for a given Einstein radius of G2, the topology of the time delay surface is affected significantly for a profile which is shallower than isothermal for G1. Also, the central maximum starts shifting towards G2. The panels (f) and (i) indicate that as G2 becomes more massive (i.e. has a larger Einstein radius) for a given profile ($\gamma = 1.4$) for G1, the location of the central maximum is again pulled towards G2. This introduces an additional pair of images and changes the flux ratios of the images. Since these models are not consistent with the data, they are ruled out.

5.5 Discussion

The aims of this chapter were to extract information from the high resolution structure in the lensed images, to determine the nature of the third radio component and to investigate the B2108+213 lensing potential, with particular emphasis on the contribution of the second lensing galaxy and the group environment. The extent to which these aims have been met is now discussed.

5.5.1 Flux-Density Ratio of the Images

As was shown for MG 2016+112 in section 4.4.2, a relative magnification matrix can be defined for the lensed images of B2108+213. Well resolved and non-collinear structure allows the determination of the relative magnification matrix with a higher accuracy. In the case of B2108+213, it is possible to calculate the magnification matrix but a large error is expected due to the uncertainties in the astrometric measurements of features with low SNR and because of the uncertainties from the small degree of non-collinearity of the image features.

The first step in calculating the magnification matrix is to identify the counterpart components in the two lensed images. Since both of the images (A and B) have clearly identifiable core-jet structure, these are used. Now, consider the vectors $\boldsymbol{A_1}$ (A_{1x}, A_{1y}), $\boldsymbol{A_3}$ (A_{3x}, A_{3y}), $\boldsymbol{B_1}$ (B_{1x}, B_{1y}) and $\boldsymbol{B_3}$ (B_{3x}, B_{3y}) where $A_2(0,0)$ and $B_2(0,0)$ are at the origin as shown in Fig. 5.19. The solid lines of image A are mapped on to the dashed lines of image B. Mathematically, the mapping is given by the relative magnification matrix as

$$\boldsymbol{B_1} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} A_{1x} \\ A_{1y} \end{pmatrix} = M_{ij}^{12} \boldsymbol{A_1} \,. \tag{5.1}$$

The vector $\boldsymbol{B_3}$ is mapped to $\boldsymbol{A_3}$ and is defined similar to Eq. 5.1. The elements of the matrix can then be determined by solving Eq. 5.1 for vectors $\boldsymbol{B_1}$ and $\boldsymbol{B_3}$ simultaneously.

Thus, the magnification matrix is found to be, $\begin{pmatrix} -0.42 & -0.31 \\ -0.56 & 0.17 \end{pmatrix}$ and the determinant is -0.27 (the negative sign here implies the images have opposite parity). Monte Carlo realizations of the image positions using the astrometric uncertainties have been used to sample the probability distribution of the determinant given the data. The median of this distribution and the 68 per cent confidence interval is given by -0.26 ± 0.28 (see Appendix D). The observed flux-density ratio is $S_B/S_A = 0.45 \pm 0.13$ and the error is derived from the 20 per cent uncertainty on the flux-densities of A and B. Although, they appear to

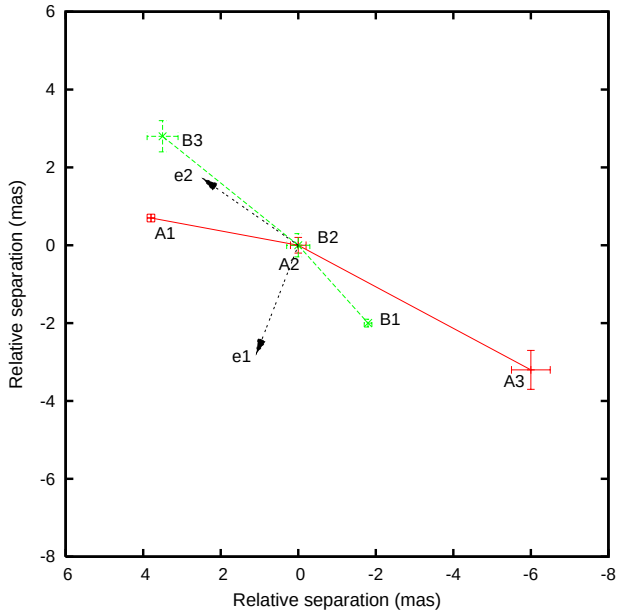

Figure 5.19: The vectors of images A (red) and B (green) with their respective origins A2 and B2 are shown to be overlapped. The arrows (black) correspond to the eigenvectors e1 and e2 of the relative magnification matrix. The eigenvalue of e1 is 0.39 and that of e2 is −0.63. The vectors of image A are decomposed along the eigenvectors and scaled by the eigenvalues to obtain the respective vectors in image B.

be consistent, the estimate of the determinant is not very precise as anticipated from the uncertainties involved.

5.5.2 The Lensing Potential

The new 1.7 and 5 GHz data presented here have found compact and extended emission from the two lensed images A and B which is consistent with gravitational lensing. The surface brightnesses of the two lensed images are the same as expected for gravitational lensing. Furthermore, since three non-collinear sub-components were detected in the 5 GHz data of images A and B, the expected parity reversal has also been observed. In Fig. 5.20, the spectral energy distributions of images A and B in the radio are shown. These spectra have been constructed using the data presented here and from the data presented by McKean et al. (2005). It is clear that both lensed images have very similar flat radio spectra. Note that the MERLIN 1.4 GHz flux-densities are much higher than the VLBA 1.7 GHz flux-densities for images A and B. This is either due to a calibration error in the MERLIN data, or more likely, that extended jet emission from the two lensed images has been resolved out by the high resolution 1.7 GHz imaging.

Isothermal mass models which have only G1 as the lensing galaxy or which include the companion lensing galaxy G2, are found to fit the data well. However, the two lens galaxy model appears to be a slightly better fit. Both of these lens models are a good fit to the positions of the lensed images, but do not fit the flux-ratio well. This could be due to variability and a time delay in the radio-loud lensed images, or may indicate that

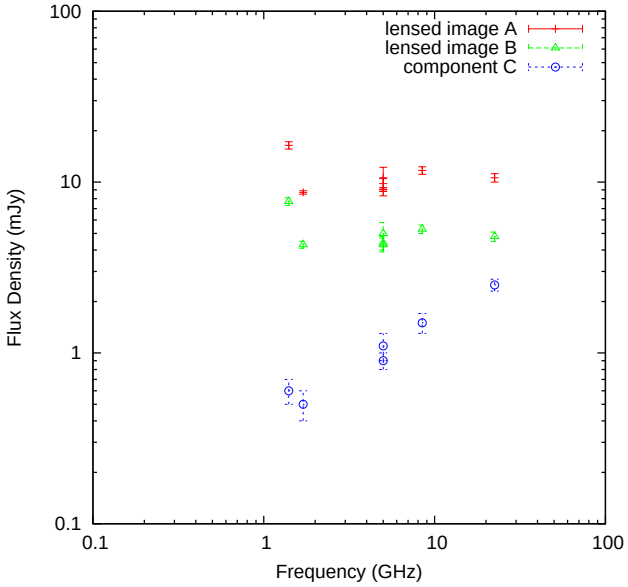

Figure 5.20: The radio spectra of lensed images A and B, and component C. The 1.4, 1.7 and 5 GHz data points are from this paper. The rest of the data points are taken from McKean et al. (2005).

a more complex mass model including additional group galaxies is required. Although B2108+213 has not been monitored for variability there are 3 independent observations carried out with MERLIN and the VLBA at 5 GHz to a low surface-brightness limit. As can be seen from Fig. 5.20 the flux-densities from these observations are consistent to within their uncertainties. It seems likely that the group environment is playing a role in the image splitting of B2108+213. The dark matter dominated halos, that is, those with shallow density profiles ($\gamma \leq 1.5$) seem not to be consistent with the observed flux-ratios of the lensed images, or the non-detection of a core lensed image (in the absence of a supermassive black hole). In fact, density profiles that are steeper than isothermal are preferred. Even though the density profiles of isolated galaxies appear to be consistent with isothermal (e.g. Koopmans et al. 2006b), it is predicted that galaxies undergoing an interaction can have density profiles that are steeper than isothermal for $\lesssim 0.5$ Gyr after the initial interaction, before returning to the original isothermal state (Dobke, King & Fellhauer 2007). Given the close proximity of galaxy G2, this scenario is certainly consistent with the lens system B2108+213. However, further observations will need to be carried out to confirm that the density profile of G1 is steeper than isothermal. In particular, measuring the redshift and stellar velocity dispersion of the companion galaxy G2, coupled with the source redshift, will allow the inner density profile of the lensing mass distribution to be determined. Also, if the background source yields a time-delay in the future, then the density profile can be found by assuming the Hubble constant (e.g. Dobke & King 2006; Auger et al. 2007b; Read, Saha & Macció 2007). Since all of the mass models tested here require an external shear of $\sim$ 4–5 per cent, it is likely that additional group/cluster galaxies are contributing to the overall potential.

5.5.3 A Radio-loud Lensing Galaxy

It seems almost certainly the case, for several reasons, that the compact emission from radio component C is due to a radio-loud AGN hosted within the main lensing galaxy G1. First, as can be clearly seen from Fig. 5.20, the radio spectrum of component C is rising from 1.4 to 22.46 GHz, whereas the radio spectra of the lensed images A and B are flat. The spectral index of component C is $\alpha_{1.4}^{22.46} = 0.51 \pm 0.01$. The spectral difference between component C and the two lensed images could be due to free-free absorption from the inter-stellar medium of the lensing galaxy (e.g. as in the case of PMN 1632-0033; Winn et al. 2004). However, the effect of free-free absorption is known to be a strong function of frequency, and the spectrum of component C shows no curvature and is constantly rising from 1.4 to 22.46 GHz. Second, the radio emission is coincident with the optical position of the lensing galaxy G1, as expected for an AGN. Third, although it is possible to fit a core lensed image at the position of component C, the flux-density of component C is much too large to be a third lensed image. Note that this would be exacerbated if component C has been affected by free-free absorption.

The classification of component C as the core of an AGN within G1 is also consistent with the extended lobe emission detected on either side of the lensing galaxy from the MERLIN 1.4 GHz imaging. Moreover, the diffuse lobed morphology of the extended emission is as expected for a Fanaroff-Riley type I radio source (FR I; Fanaroff & Riley 1974). There are no highly collimated jets or hotspots which are typically seen in the more powerful FR II sources. The total 1.4 GHz rest-frame luminosity of the radio emission associated with G1 is $L_{1.4} \sim 10^{25}$ W Hz^{-1} (this assumes a spectral index of $\alpha = -0.8$ for the lobe emission). Extrapolating this luminosity to 178 MHz gives $L_{0.178} \sim 10^{24}$ W Hz^{-1} sr^{-1}, which is below the FR I–FR II luminosity divide of $L_{0.178} \sim 10^{25}$ W Hz^{-1} sr^{-1} (Fanaroff & Riley 1974). Finally, the absolute *R*-band magnitude of the lensing galaxy G1 is $M_R = -24$. From the correlation between the rest-frame 1.4 GHz luminosity and the absolute *R*-band magnitude of the host galaxy (see Fig. 5.21; Ledlow & Owen 1996), which divides sources into FR I and FR II, the G1 radio source is found in the FR I region. Although component C appears to be unresolved in the 1.4 GHz MERLIN and 1.7 GHz VLBA imaging, the presence of the extended jet emission may explain why component C was slightly extended in the MERLIN 5 GHz map presented by McKean et al. (2005).

There are now two gravitational lens systems from the CLASS survey with a known radio-loud lensing galaxy; the other is CLASS B2045+265 (Fassnacht et al. 1999; McKean et al. 2007). This gives a fraction of 11±8 per cent for lens galaxies with a radio-loud AGN from CLASS. This agrees closely with deep radio imaging of optically selected gravitational lens candidates from the Sloan Digital Sky Survey (∼10 per cent; Boyce et al. 2006b).

It is possible to probe different lines-of-sight by searching for differences in the properties of lensed images, which should be identical in the absence of variability. Since image separations are typically only ∼ 0.5–1.5 arcsec, it is the lensing galaxy which is mostly being probed. This technique has been most successful at optical wavelengths where the dust extinction along the lines-of-sight to each of the lensed images has been used to test galactic extinction laws in high redshift lens galaxies (Falco et al. 1999; Wucknitz et al. 2003; Elíasdóttir et al. 2006). At radio wavelengths, free-free absorption of one (or more) of the lensed images has probed the ISM of the lensing galaxy (e.g. Winn et al.

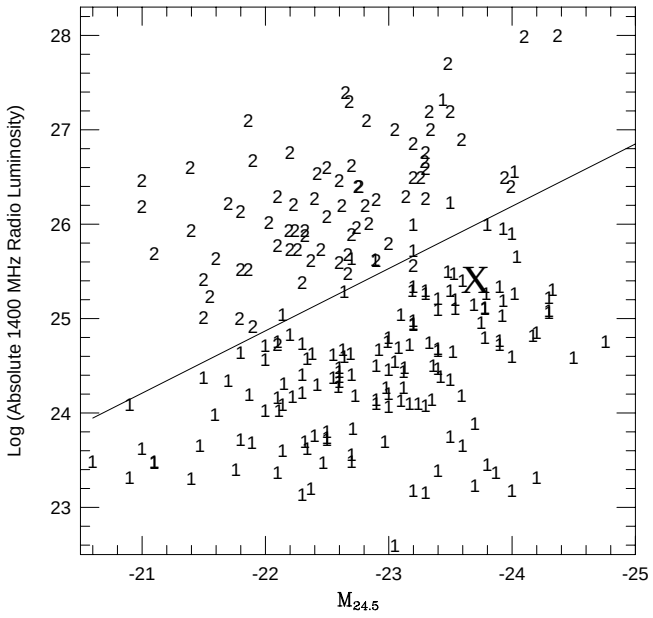

Figure 5.21: The radio-optical correlation was found for the population of FR I and FR II sources by Ledlow & Owen 1996 in their sample of host galaxies. The 'X' marks the approximate location of the lensing galaxy G1 of B2108+213 using their cosmology (H_0 = 75 km s^{-1} Mpc^{-1},q_0 = 0). Note that a different cosmology is used in this thesis. As a result, the difference in the magnitudes is $\sim$ 0.3. However, it does not affect the qualitative conclusion of B2108+213 as an FR I source.

2003; Mittal et al. 2007). B2108+213 has an unique situation of a radio lobe from an AGN within the lensing galaxy passing in front of one of the lensed images (see Fig. 5.7). Therefore, B2108+213 can be used to probe the composition of an FR I radio jet. As can be seen from Fig. 5.20, the flux-ratio between images A and B is almost constant from 1.4 to 8.46 GHz[4]. As such, there is no evidence of free-free absorption of the flux-density of image B. Another possible propagation effect is Faraday depolarization/rotation of the emission from lensed image B as it passes through the magnetized plasma of the radio-jet. However, the MERLIN data presented here were not deep enough to establish whether or not image B has been depolarized. The background source is found to be polarized by $\sim$ 4 per cent at 1.4 GHz from the data obtained for lensed image A. Therefore, further radio imaging should be carried out to determine if the properties of image B have been affected by the radio jet.

5.5.4 Estimating the Source Redshift

Since the spectrum of the background quasar of B2108+213 is found to be featureless, the determination of a spectroscopic redshift has been evasive. Therefore, an indirect method to estimate the redshift is attempted in this section.

The redshift of the background source can be estimated provided the stellar velocity

[4]The 22.46 GHz value has been determined by assuming the flux-ratio is unchanged at 22.46 GHz - see McKean et al. (2005) for details.

distribution and the Einstein radius of the lensing galaxy are known. For an isothermal mass distribution, the Einstein radius (b) of the lens is known to have a simple relation with the velocity dispersion (σ) given by

$$b = \frac{4\pi\sigma^2}{c}\frac{D_{ls}}{D_s} \tag{5.2}$$

where D_{ls} is the angular diameter distance between the lens and source and D_s is between the observer and the source. The angular diameter distance between redshifts z_1 and z_2 is expressed as

$$D_{ang}(z_1, z_2) = \frac{a(z_2)}{a(z_1)} f_k[w(z_1, z_2)] \tag{5.3}$$

which relates the angular sizes of objects at a redshift z_2 as seen from a redshift z_1 to their physical sizes. Here $a(z)$ is the scale factor, f_k (w) is a function of the curvature k and $w(z)$, $w(z_1, z_2)$ is the comoving distance (see Appendix D for details) given by

$$w(z_1, z_2) = \frac{c}{H_0}\int_{a(z_1)}^{a(z_2)} \frac{da}{a^2 E(a)} \tag{5.4}$$

E(a) is obtained from Friedmann's equation.

$$E(a) = \sqrt{\Omega_{r0}a^{-4} + \Omega_{m0}a^{-3} + \Omega_{\Lambda 0} + \Omega_K a^{-2}} \tag{5.5}$$

where $\Omega_{m0} = 0.3$ and $\Omega_{\Lambda 0} = 0.7$ and other factors can be neglected from the WMAP results.

Now, using the Einstein radius = 1.711±0.038 from Table 5.5, σ = 360±30 km s^{-1} (McKean et al. in prep) and the redshift of the lens (z_l=0.365), the source redshift z_s is calculated by finding the root of R(z_s) which is defined as

$$R(z_s) = \frac{bc}{4\pi\sigma^2} - \frac{D_{ls}}{D_s} \tag{5.6}$$

Thus, the solution of Eq. 5.6 gives a redshift of $z = 0.74$ for the source. Instead of solving the root of the Eqn. 5.6, a range of velocity dispersions can be calculated as a function of redshift as shown in Fig. 5.22. The redshift corresponding to the measured velocity dispersion (σ = 360 km s^{-1}) then gives the source redshift (i.e. $z = 0.74$).

5.6 Conclusions

The new high resolution MERLIN and VLBI imaging of the gravitational lens system B2108+213, whose wide image separation is consistent with a massive lensing galaxy or possibly a group of galaxies, have been presented. The VLBI imaging at 1.7 and 5 GHz found extended emission in the lensed images whose surface brightness and parities are consistent with gravitational lensing. Using the new constraints provided from the two lensed images, mass models for the B2108+213 lens potential were tested. The properties of the lensed images were found to be consistent with either a single massive lensing galaxy, or a two galaxy lens model which accounts for a nearby companion to the main lensing galaxy. In these cases the companion G2 was always represented by a singular

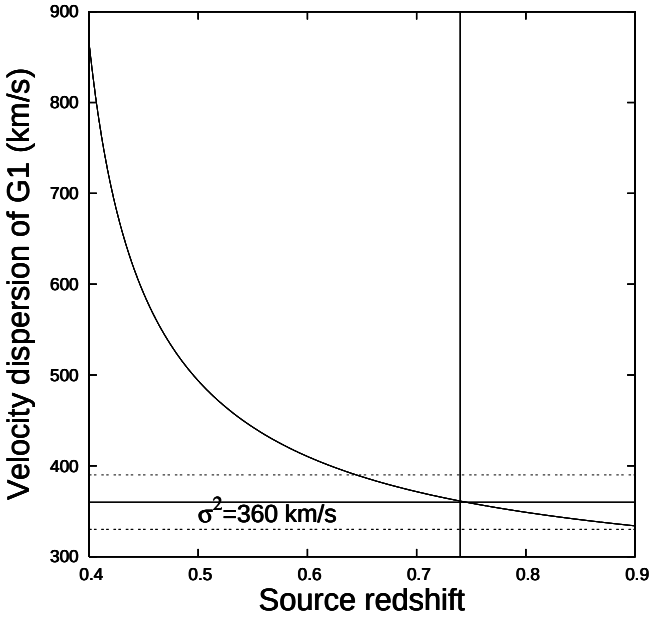

Figure 5.22: The velocity dispersion of an isothermal profile for G1 as a function of redshift. The measured velocity dispersion of 360 km s^{-1} implies a source redshift of 0.74. The double-dotted lines are the errors on the velocity dispersion measurement.

isothermal sphere whereas the main galaxy was modeled as both an isothermal sphere and ellipsoid. Steeper than isothermal density profiles for the main lensing galaxy were a better fit to the data. Further models which include additional group galaxies will be tested when a spectroscopic survey of the B2108+213 local environment has been completed.

Emission from the third radio component, which is coincident with the optical position of the main lensing galaxy, was found for the first time at mas-scales at both 1.7 and 5 GHz. Furthermore, MERLIN imaging at 1.4 GHz detected extended low surface brightness emission on either side of the third radio component. This jet emission has a morphology and luminosity which is consistent with an FR I type radio source. Attempts to model the radio core of the third radio component as a core lensed image failed because it was not possible to fit the observed flux-density (to within several orders of magnitude) by using a core radius and/or a variable power-law density profile for the main lensing galaxy. Therefore, the third radio component is definitely due to emission from an AGN embedded within the main lensing galaxy.

6 Summary

In this thesis, a tale of two wide separation lens systems was described[1]. The lens systems with large image separations are found to lie in dense environments like groups or clusters (e.g., Keeton et al. 2000b; Oguri 2006). Amongst many other applications, lensing is used as a probe of the sub-halo population (e.g., Dalal & Kochanek 2002; Shin & Evans 2008). Furthermore, lensing is used to study the mass distribution of the lensing galaxies and contribution of the lens environments (e.g., Momcheva et al. 2006; Auger 2008). With the above applications in mind, the research carried out on two lens systems MG 2016+112 and B2108+213 is summarized in this chapter.

MG 2016+112

MG 2016+112 was observed both at 1.7 and 5 GHz using MERLIN and VLBI simultaneously and at 8.4 GHz using the HSA. The observations at each frequency were carried out at separate epochs. From the observations of MG 2016+112 presented here, three new components were found in images A and B. All of the components (two components in each A and B, and four components in region C) known from previous data were also detected in these observations. The total spectra of A and B are similar whereas the spectrum of C is different as suggested from previous data. The spectra of all of the components in A (and B) were found to be steep between 5 and 8.4 GHz. On the other hand for image A, between 1.7 and 5 GHz the spectra were steep with the exception of component A2 which has a much flatter spectrum than the rest of the four components[2]. The spectra of the inner pair of extended components in region C were found to be flatter than the spectra of the outer pair of components as known from previous observations.

The new observational constraints were used to test numerous mass models. One of the aims of conducting high resolution and high sensitivity observations was to test the mass model of Koopmans et al. (2002b, henceforth, K02). The K02 model predicted component 2 and any new components to its north-west as quadruply imaged. Since component 5 in image A and B was detected to the north-west of component 2, it should be quadruply imaged. According to the prediction of K02 mass model, a pair of images of component 5 should be situated in region C on either side of the critical curve. This pair was predicted to have ∼10 times higher flux density than their counterparts in images A

[1] Admittedly, the research work presented here does not involve as interesting a plot as was conceived by the author of *A Tale of Two Cities*. Nevertheless, like the two lead baryonic characters with contrasting personalities and sharing one love in the book, the two lens systems under scrutiny (which are much more than just baryons) with dissimilar properties allow to probe the mysterious (dark) matter in the Universe.

[2] In image B, it was not possible to resolve B3 and B2. Hence, the spectrum of the individual component B2 could not be determined.

and B. The simultaneous observations with MERLIN had the apt resolution and sensitivity to detect such components. An inspection of the 5 GHz MERLIN maps suggested no new components in region C. A test was performed on the MERLIN data to ensure that the observations had the desired resolution and sensitivity to detect the expected pair of images of component 5. Artifical Gaussian components were introduced in the data with the expected position and size, assuming that lensing conserves the surface brightness of these components and the components have circular shapes. The artificial components were reproduced in the MERLIN images presented here. This led to the conclusion that the K02 model was not acceptable.

Next, new mass models were investigated that with the aim of fitting the new data especially the asymmetric separation of components in region C (which was ignored in the K02 model). The components in region C are merging pairs of partial images of the background quasar and their counterparts in the components of images A and B are not well-identified. Thus, various scenarios were tested by associating the pair in region C with different regions in images A and B. These scenarios were tested for SIE+SIS+shear and SIE+SIS+SIS+shear models. The acceptable scenarios include those where region C was associated with either component 5 or an unseen component to the north-west of component 5. The best reduced χ^2 was ∼3.5 for the SIE+SIS+shear model, which included the main lensing galaxy D (SIE) and the luminous satellite G1 (SIS). Although G1 was found to be largely responsible for causing the asymmetric separation of the components in region C, an even more complex model is perhaps required to fit the data better (i.e. expected reduced $\chi^2 \sim 1$).

The relative magnification matrix mapping is linear when inspected over small regions in the image plane and certainly holds true for regions far from the critical curves. Images A and B satisfy these conditions and the non-collinear components in images A and B were used to calculate the relative magnification matrix mapping between image A and image B. In the absence of substructure in the vicinity of these images, the determinant of the matrix mapping should equal the observed flux density ratio of the images. No detectable substructure was found in images A and B. An upper limit on the mass of the substructure near images A and B could be found from the Einstein radius (i.e. half of the mean component separation) and assuming a point mass for the substructure. Thus, for an Einstein radius $\theta_E = 3$ mas, a substructure in image A or B must be $\lesssim 10^7\,M_\odot$ at the lens redshift to produce no observable effects. From the SIE+SIS+shear model, the velocity dispersion of the lensing galaxy D was estimated as 343 km s^{-1} consistent with the measured central stellar velocity dispersion (Koopmans & Treu 2002a). Furthermore, the satellite galaxy G1 was estimated to have a mass of $\sim 10^{10}\,M_\odot$ and a velocity dispersion of ∼100 km s^{-1} which is consistent with the estimates of Chen et al. (2007).

Future work: Perhaps the blind men have come to share a common perspective, until the devil reveals its other dark sides, that is, radio astronomers have come to a conclusion that in the lens plane of MG 2016+112 the luminous substructure is certainly required but one should expect the unexpected. The source of uncertainties in the mass models presented here are unclear. Nevertheless, further complicated mass models for example, by introducing a gradient in the mass or higher order multipoles in the potential should be tested. On the other hand, the background quasar in MG 2016+112 is a type II quasar and is not well-studied. Its complicated morphology, rich multi-wavelength structure lensed at different spatial scales and low flux density in spite of being lensed has made it difficult to

decipher the properties of this high redshift quasar. Further deep observations at multiple wavelengths are needed to investigate the morphological and evolutionary stage of this type II quasar in order to compare it with other high (or similar) redshift quasars.

B2108+213

Radio observations of B2108+213 at 1.7 and 5 GHz using VLBI and at 1.4 GHz using MERLIN, carried out at different epochs, were presented. Three radio components A, B and (for the first time using VLBI) C were detected in all observations. Both of the lensed images A and B showed core-jet structure with the expected opposite parity in the features although with low degree of non-collinearity whereas component C was found to be compact. Component C was known to have a spectrum different from those of images A and B. The spectral analysis from the new observations presented here were in agreement with the previous results. From the MERLIN imaging at 1.4 GHz, low surface brightness emission centred at component C, and extended on either side of C, was detected. Therefore, the third radio component was identified as the AGN of the lensing galaxy at a redshift $z = 0.365$. The extended emission was found to be the radio lobes of the AGN. The estimate of the radio power from this region, minus the emission from A and B, was found to be consistent with the lensing galaxy being a FR I type radio galaxy. This is the first case where a massive elliptical lensing galaxy has an AGN with an FR I-like radio lobes.

It was not clear whether the compact radio component of C is a third lensed image or an AGN embedded within the lens galaxy G1 due to its proximity to the optical position of G1. Although the spectral analysis and the new MERLIN imaging presented here suggested that component C corresponds to an AGN of the lensing galaxy, new mass models were tested to confirm this conclusion. Profiles with constant-density cores were investigated which can produce odd images but with a demagnified flux density of the odd image. No mass model could simultaneously fit the position and the flux density of component C. Either the predicted odd image was situated much farther from the lens galaxy G1 than the position of component C or the flux density was lower by several orders as compared to the observed flux density of component C. Thus, both the mass model predictions and the observations confirmed the status of C as an AGN within the lens galaxy G1.

From optical imaging, the lens galaxy G1 showed a close companion G2 to the line-of-sight. The lens galaxy G1 was not only known to belong to a group of galaxies but was also known to be the Brightest Group Galaxy (BGG). A recent X-ray analysis has indicated a pair of actively interacting galaxy-groups with the lens galaxy of B2108+213 as the BGG of one of the groups (Fassnacht et al. 2007). The constraints from previous observations (McKean et al. 2005) were not sufficient to test various mass models. The new observational results presented here provide enough constraints and hence, several mass models were tested. Most of the image splitting (~4.6 arcsec) was found to be due to the main galaxy G1 with an isothermal profile. Since the flux ratios were not fitted well, profiles other than isothermal for G1 were tested. The best-fit profile was found to be for a slope of $\gamma = 2.45^{+0.19}_{-0.18}$. From recent simulations, lens galaxies interacting with their companions were shown to have a steepened profile shortly after their initial interaction ($\lesssim$0.5 Gyr) before returning to an isothermal state (Dobke et al. 2007). Furthermore, an

analysis of fifteen lens galaxies from the SLACS also suggested that those lens galaxies which were best fitted with a steeper than isothermal profile had a close companion (Auger 2008). Thus, the results presented here suggested that lens galaxy G1 was interacting with the companion galaxy G2.

The background quasar is known to have a featureless spectrum and hence, believed to be a BL Lac type quasar. Thus, there is no redshift measurement for the lensed quasar. Using the Einstein radius of G1 as 1.711 arcsec from mass models presented here and the stellar velocity dispersion of 360 km s^{-1} from McKean et al. (in prep.), the source redshift was estimated to be $z = 0.74$.

Future work: The results of the new mass models suggested that lens galaxy G1 has a steeper than isothermal profile due to the interaction with the companion galaxy G2. This could be tested by measuring the redshift and velocity dispersion of G2, and the redshift of the background quasar. Furthermore, a measurement of the relative time delay of the images could also constrain the density profile.

Since polarization is not affected by lensing and B2108+213 has the unique situation of one lensed image (B) seen through the radio lobe whereas the other lensed image (A) is not contaminated due to any radio emission from the lens galaxy, multi-frequency polarization observations of B2108+213 must be performed using the extended Very Large Array (eVLA). These new observations may allow the polarization of each image to be measured which could then be compared to find any relative Faraday rotation between the two lensed images due to the magnetic field from the radio lobes.

A Cosmological concepts and terms

Observations suggest that the large scale properties of the Universe are invariant under translations and rotations. The translation symmetry is a manifestation of the homogeneity while the rotational symmetry is a manifestation of the isotropy in the Universe. The metric in such a spacetime can be written as

$$ds^2 = c^2dt^2 - a^2(t)\left[d\chi^2 + f_k^2(\chi)\,(d\theta^2 + \sin^2\theta d\phi^2)\right]. \tag{A.1}$$

This is called the Friedmann-Robertson-Walker (FRW) metric. Here, c is the speed of light and $a(t)$ (henceforth, a) is the time dependent *scale factor*. The scale factor is normalized such that its value today ($t = t_0$) is unity. t is the time measured by the observers which are stationary at coordinates denoted by (χ, θ, ϕ). The coordinates θ and ϕ are the angular coordinates while χ is the coordinate in a radial direction. The factor $f_k(\chi)$ depends upon the geometry of the Universe which determines the sign of the parameter k. The results from cosmic microwave background experiments suggest that the Universe is flat which implies $k = 0$. In a flat Universe the factor $f_0(\chi) = \chi$ and the metric equation is,

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

The spatial part of this metric is Euclidean and is scaled by the scale factor.

A *comoving distance* is that distance between two points in space which does not change as a function of time. The corresponding coordinates (χ, θ, ϕ) are called *comoving coordinates*. The *physical distance* x between two points in space is then time dependent because $x = \chi\, a$ [1] and the scale factor is time dependent. As a result, the physical separation between any two points today ($t = t_0$) is the *comoving distance*.

Einstein's field equations relate the metric to the matter content in the Universe by,

$$G_{\mu\nu} = \frac{8\pi G}{c^2} T_{\mu\nu}\,, \tag{A.3}$$

where G is the gravitational constant, $G_{\mu\nu}$ is the Einstein tensor while $T_{\mu\nu}$ describes the energy density of the Universe. The labels μ and ν run over the time index (0) and the three spatial indices (1, 2 and 3). The homogeneity and isotropy of the Universe also implies that the matter content can be described in terms of a perfect fluid with energy density $\rho(t)$ and pressure $p(t)$. The energy momentum tensor is diagonal and is given by $(\rho\, c^2, -p, -p, -p)$[2] which has only two independent parameters. The two independent Einstein's equations are

$$-3\frac{\ddot{a}}{a} = 4\pi G\,(\rho + 3\,p)\,, \tag{A.4}$$

[1] Assuming a comoving distance in the radial direction (i.e. $\theta = 0$ and $\phi = 0$) for simplicity.

[2] Only the diagonal elements are given here for simplicity

and

$$\frac{\ddot{a}}{a} + 2\,(\dot{a}/a)^2 = 4\pi G(\rho - p)\,. \tag{A.5}$$

Rearranging Eq. A.4 and A.5 gives the first and second order differential equations in terms of the scale factor and are expressed as

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho \tag{A.6}$$

and

$$\left(\frac{\ddot{a}}{a}\right) = -\frac{4\pi G}{3}(\rho + 3\,p)\,. \tag{A.7}$$

These are called Friedmann's equations. The *Hubble parameter* is defined as

$$H \equiv \left(\frac{\dot{a}}{a}\right), \tag{A.8}$$

and represents the rate of expansion. Assuming conservation of energy, an equation governing the time evolution of the energy density ρ and pressure p can be found as

$$\frac{\partial}{\partial t}\rho + 3\frac{\dot{a}}{a}(\rho + p) = 0\,. \tag{A.9}$$

Let the equation of state which relates the energy density ρ with pressure p to be

$$p = w\rho\,. \tag{A.10}$$

For various species in the Universe, the equation of state parameter is found to assume different values, that is, $w = 0$ for nonrelativistic matter, $w = 1/3$ for radiation and $w = -1$ for dark energy. The total energy density is then a sum of the energy densities of the three species,

$$\rho = \rho_m + \rho_r + \rho_\Lambda\,. \tag{A.11}$$

Substituting p from the equation of state in the time evolution equation (Eq. A.9) then gives the variation of the energy density ρ_s of any of the species 's', independently, as a function of the scale factor,

$$\rho_s = \rho_{s,0}a^{-3(1+w)}\,, \tag{A.12}$$

where $\rho_{s,0}$ is the energy density of a species at $t = t_0$. The matter energy density falls as a^{-3}, since the number density of matter decreases due to the expansion of space. The energy density of photons falls as a^{-4}, which reflects a combined effect of a decrease in their number density as a^{-3} and a decrease in their momentum as a^{-1} due to the expansion. Photons of wavelength λ_1 emitted at scale factor a_1 that arrive today ($t = t_0$) have a wavelength λ_0 which is related to the scale factor by

$$\frac{\lambda_0}{\lambda_1} = \frac{1}{a_1}. \tag{A.13}$$

Hence the redshift z_1 is related to a_1 by

$$z_1 = \frac{\lambda_0 - \lambda_1}{\lambda_1} = \frac{1 - a_1}{a_1}\,. \tag{A.14}$$

Therefore, the scale factor is conveniently expressed as a function of redshift,

$$a = \frac{1}{1+z}\,. \tag{A.15}$$

The energy density $\rho_{s,0}$ of species 's' today ($t = t_0$) is often expressed in terms of a dimensionless cosmological parameter Ω_s defined as

$$\Omega_s = \frac{8\pi G}{3H_0^2}\rho_{s,0}\,, \tag{A.16}$$

where H_0 is the Hubble parameter for $t = t_0$ commonly called the *Hubble constant*. The evolution of the scale factor can be calculated by using Eq. A.6 and A.12 in terms of the cosmological parameters. For a flat Universe this is then,

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0^2\left(\frac{\Omega_m}{a^3} + \frac{\Omega_r}{a^4} + \Omega_\Lambda\right) = H_0^2 E(a)\,, \tag{A.17}$$

where the cosmological parameters are constrained from the CMB observations. Recent results from the Wilkinson Microwave Anisotropy Probe (WMAP) data release suggest that $\Omega_m = 0.258 \pm 0.030$, $\Omega_\Lambda = 0.742 \pm 0.030$ and $h = 0.719^{+0.026}_{-0.027}$ (Dunkley et al. 2008). Here, h is a dimensionless parameter generally used to express the Hubble constant as $H_0 = 100\,h$ km s^{-1} Mpc^{-1}. The scale factor as a function of time can be found out by integrating Eq. A.17 numerically.

A.1 Distances in cosmology

The simplest distance that can be found from the metric is the comoving distance χ_1 to a redshift of z_1 where the scale factor is a_1. The path of light in the Universe is described by null geodesics and a light ray which propagates radially has $d\theta = 0$ and $d\phi = 0$, then

$$a\,d\chi = -c\,dt = -c\,\frac{da}{\dot{a}}\,. \tag{A.18}$$

Using Eq. A.17 and integrating Eq. A.18 gives

$$\chi_1 = \frac{c}{H_0}\int_{a_1}^{1} \frac{da}{a^2\left(\Omega_m a^{-3} + \Omega_r a^{-4} + \Omega_\Lambda\right)^{1/2}}\,. \tag{A.19}$$

For the purpose of gravitational lensing, physical distances at different redshifts have to be routinely converted into angles as seen at a given redshift. Hence, the most useful distance measure is the *angular diameter distance*. The physical size of an object, ξ_1, at redshift of z_1 is related to the angular size of the object, θ_1, measured today (i.e. $z = 0$) as

$$\xi_1 = D_{\rm ang}(0, z_1)\,\theta_1\,. \tag{A.20}$$

Consider an observer at redshift $z = 0$ to be situated at the centre of a circle such that the radius of the circle is the comoving distance χ_1. The angle subtended at the observer by

the circle is 2π and the physical size of the circle at redshift z_1 is $2\pi\chi_1 a_1$ (assuming a flat Universe). Consider an angular element ξ_1 on the circle with a physical size θ_1, then

$$\frac{\xi_1}{2\pi\,\chi_1\,a_1} = \frac{\theta_1}{2\pi}\,. \tag{A.21}$$

Comparing Eq. A.20 and A.21, the angular diameter distance of the circle as seen from $z = 0$ is given by

$$D_{\rm ang}(0, z_1) = a_1\chi_1\,. \tag{A.22}$$

Similarly, it can be shown that the angular distance between redshifts z_1 and z_2 is

$$D_{\rm ang}(z_1, z_2) = a_2\,(\chi_2 - \chi_1)\,. \tag{A.23}$$

Another commonly used distance measure is the *luminosity distance*. The luminosity distance D_l to a redshift z_1 is defined via the flux F received today (i.e. $z = 0$) from a source with an intrinsic luminosity L at redshift z_1 by

$$F = \frac{L}{4\pi D_l^2}\,. \tag{A.24}$$

The luminosity distance and the angular diameter distance are related by the Etherington relation as

$$D_l(z) = (1 + z)^2 D_{\rm ang}(z)\,. \tag{A.25}$$

B Powell's minimization method

Various methods are known in order to perform chi-squared minimization (a technique to find a best-fitting model) (Press et al. 1992). One such method is Powell's minimization which was used in the Gaussian model fitting and is described below. Powell's method uses a set of direction vectors to probe the gradient in the *n*-dimensional parameter space, starting with an initial guess and minimizes the χ^2 along these directions to reach the minimum in the parameter space.

An example

Consider a three dimensional parameter space (p_1, p_2, p_3) for simplicity. Let the initial guess be (p_{10}, p_{20}, p_{30}) for the first iteration, then the χ^2 is minimized say, along p_1 in the parameter space. Let the least χ^2 thus found be at (p_{11}, p_{20}, p_{30}). Now, choosing (p_{11}, p_{20}, p_{30}) as the initial point, the χ^2 is minimized along parameter p_2 and let (p_{11}, p_{21}, p_{30}) be the point for which a least χ^2 is obtained. Once again starting with (p_{11}, p_{21}, p_{30}), the χ^2 function is now minimized along p_3 such that the minimum occurs at (p_{11}, p_{21}, p_{31}). This marks the end of the first iteration that estimates a direction of descent. Let $\bar{e}_1$ be the vector joining the initial guess point (p_{10}, p_{20}, p_{30}) and the final point (p_{11}, p_{21},p_{31}) in the three dimensional parameter space.

Similar to the first iteration, the χ^2 is minimized individually but along p_2 and p_3, and the new direction vector $\bar{e}_1$ (instead of p_1) in the second iteration. Let the final minimum be obtained at (p_{12}, p_{22}, p_{32}) and the vector joining the minimum to (p_{11}, p_{21},p_{31}) be $\bar{e}_2$. In the third iteration, the minimization is carried out along the vectors $\bar{e}_1$, $\bar{e}_2$ and p_3. Every new vector is directed towards a steeper descent. The minimization is continued till the $\Delta\chi^2$ between the previous and current iteration is a small number. To ensure that a global minimum is reached, the above described process is repeated with a different set of initial guess parameters.

In summary, the above example when generalized to an *n*-dimensional parameter space will work as follows. Starting with a point as an initial guess in the parameter space, the χ^2 function is minimized in *n* steps along one vector direction at a time. At the end of the first *n*-step iteration, a direction of descent is found by the vector $\bar{e}_1$ connecting the first best fit parameters to the initial guess value of the parameters. Next, minimizing along $\bar{e}_1$ and any $(n-1)$ parameters such that the minimization is always along one vector direction at a time. At the end of the second *n*-step iteration, a new vector $\bar{e}_2$ is determined by connecting the second best fit parameters to the first best fit parameters. The third iteration is continued by minimizing along vectors $\bar{e}_1$ and $\bar{e}_2$, and any of the $n-2$ parameters. The iterations are repeated till the χ^2 values from the last and the current iterations are not

significantly different. This method converges to the local minimum and needs a good initial guess of the parameters.

With Powell's minimization method, multiple Gaussian components can be fitted to data and the fitted parameters are supplied without any uncertainties arising either from the data or from the model fitting procedure. There are other methods (e.g. Levenberg-Marquardt method) that output the errors however, Powell's method is simpler to implement. Hence, Powell's method was used and the uncertainties on the fitted parameters were determined based on the method outlined in Fomalont (1999) for error determination.

C Analytical approximations to estimate errors

The procedure of model fitting has uncertainties associated with the output quantities. From the model fitting to features in interferometric radio images, the uncertainties are dependent on the signal-to-noise of the detection. An analytical approximation to the uncertainties of the parameters based on the approach outlined by Fomalont (1999, Lobanov priv. comm.). Let the following be the modeled parameters such that S_{peak} is the peak flux density, S_{tot} is the total flux density and let the σ_{rms} be the post-fit rms noise of the residual map[1]. The uncertainty in the modeled peak flux-density is then given by,

$$\sigma_{peak} = \sigma_{rms}\left(1 + \frac{S_{peak}}{\sigma_{rms}}\right)^{1/2} . \tag{C.1}$$

Incorporating the uncertainties from the calibration of the data (σ_c) gives the total uncertainty in the peak flux-density as

$$\sigma_p = \left(\sigma_{peak}^2 + \sigma_c^2\right)^{1/2} . \tag{C.2}$$

The uncertainty in the modeled total flux-density is

$$\sigma_{tot} = \sigma_p\left(1 + \frac{S_{tot}^2}{S_{peak}^2}\right)^{1/2} . \tag{C.3}$$

Let d be the size of a component[2], then the uncertainty in the modeled peak position is

$$\sigma_d = \frac{\sigma_{rms}\, d}{2\, S_{peak}} . \tag{C.4}$$

The uncertainties determined using the above equations are true to first order. The difference between the true errors and analytically estimated errors are less than a factor of two.

[1] Instead of the post-fit rms, the off-source rms was taken for the calculations in this thesis.

[2] For components with sizes smaller than the beam, the beam size is taken as d and for components with sizes larger than the beam, the diameter of a circular approximation of the extent of the component is taken as d.

D Errors using a Monte Carlo approach

Given a set of observables, O_0, there is a set of n parameters denoted by $\boldsymbol{a_0}$ which minimizes the χ^2 of the observable data set. The observables cannot be measured with infinite precision and have certain distribution associated with them. The confidence limits that can be placed on the parameters are dependent on the uncertainties of the observables. Thus, there are probabilities associated with the set of parameters inferred from the observables. Monte Carlo methods are used to assign confidence limits on the parameters.

Hypothetical datasets ($O_1, O_2, O_3, \ldots$) can be generated by drawing out random numbers with a distribution associated with the observables. The procedure of fitting (e.g., a χ^2 minimization) can be repeated on the hypothetical datasets to obtain the respective set of best-fit parameters $\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}, \ldots$ The resulting sets of parameters ($\boldsymbol{a}$) are distributed in the n-dimensional parameter space. Projections of this distribution place confidence limits in the parameter space of interest. For example, confidence limits can be determined for a single parameter, error ellipses can be given for a set of two parameters and analogus confidence regions in the higher dimensional parameter space of interest can be given.

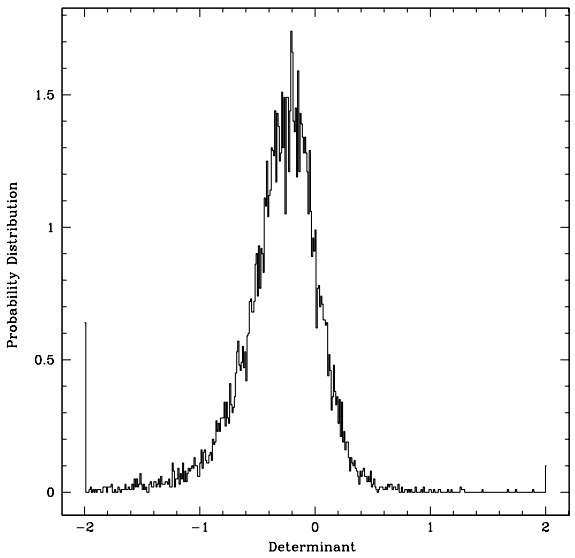

Figure D.1: The mean value of the determinant of the relative magnification matrix is 0.27 with a 1σ value of ~ 0.2.

The aforementioned approach was used in this thesis to calculate the error on the determinant of the relative magnification matrix, given the observable positions of the images. Each image position was associated with an error which was assumed to be normally distributed in its position. Ten thousand hypothetical positions of each image were drawn from a Gaussian distribution centred on the observed position and a width equal to 1 sigma error. The relative magnification matrix was calculated using each hypothetical dataset. The distribution of the determinant of the relative magnification matrix inferred from these datasets is shown in Fig. D.1. The mean and the width of this distribution was reported as the mean and the error of the determinant of the relative magnification matrix.

Bibliography

Alard C., 2008, ArXiv e-prints, 803, arXiv:0803.2295

Auger M. W., Fassnacht C. D., Abrahamse A. L., Lubin L. M., Squires G. K., 2007, AJ, 134, 668

Auger M. W., Fassnacht C. D., Wong K. C., Thomspon D., Matthews K., Soifer B. T., 2007b, in press (arXiv:0710.1650)

Auger M. W., 2008, MNRAS, 383, L40

Bahcall J. N., Maoz D., Doxsey R., Schneider D. P., Bahcall N. A., Lahav O., Yanny B., 1992, ApJ, 387, 56

Bahcall N. A., Cen R., Davé R., Ostriker J. P., Yu Q., 2000, ApJ, 541, 1

Bartelmann M., Huss A., Colberg J. M., Jenkins A., Pearce F. R., 1998, A&A, 330, 1

Belokurov V., et al., 2007, ApJ, 654, 897

Benítez N., Broadhurst T., Rosati P., Courbin F., Squires G., Lidman C., Magain P., 1999, ApJ, 527, 31

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

Biggs A. D., Browne I. W. A., Wilkinson P. N., 2001, MNRAS, 323, 995

Biggs A. D., Browne I. W. A., Jackson N. J., York T., Norbury M. A., McKean J. P., Phillips P. M., 2004, MNRAS, 350, 949

Biggs A., 2005, Future Directions in High Resolution Astronomy, 340, 433

Binney J., Tremaine, S., 1987, Princeton, NJ, Princeton University Press, 1987, 747 p.

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

Blanton M. R., et al., 2003, ApJ, 592, 819

Bolton A. S., Burles S., Koopmans L. V. E., Treu T., Moustakas L. A., 2006, ApJ, 638, 703

Bolton A. S., Moustakas L. A., Stern D., Burles S., Dey A., Spinrad H., 2006, ApJL, 646, L45

Born M., Wolf E., 1959, London: Pergamon Press, 1959

Boyce E. R., Winn J. N., Hewitt J. N., Myers S. T., 2006, ApJ, 648, 73

Boyce E. R., Bowman J. D., Bolton A. S., Hewitt J. N., Burles S., 2006, ApJ, 640, 62

Browne I. W. A. et al., 2003, MNRAS, 341, 13

Burke W. L., 1981, ApJL, 244, L1

Burke, B. F., 1990, Gravitational Lensing, 360, 127

Caon N., Capaccioli, M., D'Onofrio M., 1993, MNRAS, 265, 1013

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

Coles J., 2008, ArXiv e-prints, 802, arXiv:0802.3219

Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, AJ, 115, 1693

Congdon A. B., Keeton C. R., 2005, MNRAS, 364, 1459

Chartas G., Bautz M., Garmire, G., Jones C., Schneider D. P., 2001, ApJ, 550, L163

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

Clowe D., Trentham N., Tonry J., 2001, A&A, 369, 16

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

Diemand J., Zemp M., Moore B., Stadel J., Carollo C. M., 2005, MNRAS, 364, 665

Diemand J., Kuhlen M., Madau P., 2007, ApJ, 657, 262

Dobke B. M., King L. J., 2006, A&A, 460, 647

Dobke B. M., King L. J., Fellhauer M., 2007, MNRAS, 377, 1503

Dunkley J., et al., 2008, ArXiv e-prints, 803, arXiv:0803.0586

Dyer C. C., Roeder R. C., 1980, ApJL, 238, L67

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

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

Faber, S. M., & Jackson, R. E. 1976, ApJ, 204, 668

Falco E. E. et al., 1999, ApJ, 523, 617

Fanaroff B. L., Riley J. M., 1974, MNRAS, 167, 31P

Fassnacht C. D., et al., 1999, AJ, 117, 658

Fassnacht C. D., Lubin L. M., 2002, AJ, 123, 627

Fassnacht C. D., Moustakas L. A., Casertano S., Ferguson H. C., Lucas R. A., Park Y., 2004, ApJL, 600, L155

Fassnacht C. D., Gal R. R., Lubin L. M., McKean J. P., Squires G. K., Readhead A. C. S., 2006a, ApJ, 642, 30

Fassnacht C. D. et al., 2006b, ApJ, 651, 667

Fassnacht C. D., Kocevski D. D., Auger M. W., Lubin L. M., Neureuther J. L., Jeltema T. E., Mulchaey J. S., McKean J. P., 2007, arXiv:0711.2066

Fomalont E. B., 1999, in Talyor G. B., Carilli C. L., Perley R. A., eds, Synthesis Imaging in Radio Astronomy II. Astron. Soc. Pac., San Francisco, p. 301

Garrett M. A., Muxlow T. W. B., Patnaik A. R., Walsh D., 1994, MNRAS, 269, 902

Garrett M. A., Calder R. J., Porcas R. W., King L. J., Walsh D., Wilkinson P. N., 1994, MNRAS, 270, 457

Garrett M. A., Porcas R. W., Nair S., Patnaik A. R., 1996, MNRAS, 279, L7

Gavazzi R., Treu T., Rhodes J. D., Koopmans L. V. E., Bolton A. S., Burles S., Massey R. J., Moustakas L. A., 2007, ApJ, 667, 176

Girardi M., Manzato P., Mezzetti M., Giuricin G., Limboz F., 2002, ApJ, 569, 720

Gregory P. C., Scott W. K., Douglas K., Condon J. J., 1996, ApJS, 103, 427

Haarsma D. B., Hoekema K. J., Hewitt J. N., Langston G. I., 2000, Bulletin of the American Astronomical Society, 32, 1578

Hattori M., et al., 1997, Nature, 388, 146

Heflin M. B., Gorenstein M. V., Lawrence C. R., Burke B. F., 1991, ApJ, 378, 519

Högbom J. A. 1974, A&AS, 15, 417

Hoekstra H., Franx M., Kuijken, K., van Dokkum P. G., 2002, MNRAS, 333, 911

Inada N. et al., 2006, ApJL, 653, L97

Jin C., Garrett M. A., Nair S., Porcas R. W., Patnaik A. R., Nan R., 2003, MNRAS, 340, 1309

Jing Y. P., Suto Y., 2002, ApJ, 574, 538

Kauffmann G., White S. D. M., Guiderdoni B, 1993, MNRAS, 264, 201

Kazantzidis S., Kravtsov A. V., Zentner A. R., Allgood B., Nagai D., Moore B., 2004, ApJ, 611, L73

Keeton C. R., Mao S., Witt H. J., 2000, ApJ, 537, 697

Keeton C. R., Christlein D., Zabludoff, A. I., 2000, ApJ, 545, 129

Keeton C. R., 2001, arXiv:astro-ph/0102340v1

Keeton C. R., 2003, ApJ, 582, 17

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

Keeton C. R., Zabludoff A. I., 2004, ApJ, 612, 660

King L. J., et al., 1998, MNRAS, 295, L41

King L. J., Browne I. W. A., Marlow D. R., Patnaik A. R., Wilkinson P. N., 1999, MNRAS, 307, 225

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

Knebe A., Wießner V., 2006, Publications of the Astronomical Society of Australia, 23, 125

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

Kochanek C. S., 1991, ApJ, 373, 354

Kochanek C. S., et al., 2000, ApJ, 543, 131

Kochanek C.S., Schneider P., Wambsganss J., 2004, Part 2 of Gravitational Lensing: Strong, Weak & Micro, Proceedings of the 33rd Saas-Fee Advanced Course, G. Meylan, P. Jetzer & P. North, eds. (Springer-Verlag: Berlin)

Koopmans L. V. E., Treu T., 2002, ApJ, 568, L5

Koopmans L. V. E., Garrett M. A., Blandford R. D., Lawrence C. R., Patnaik A. R., Porcas R. W. 2002, MNRAS, 334, 39

Koopmans L. V. E., Treu T., Fassnacht C. D., Blandford R. D., Surpi G., 2003, ApJ, 599, 70

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

Koposov S., et al., 2007, ArXiv e-prints, 706, arXiv:0706.2687

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

Langston G., Fischer J., Aspin C., 1991, AJ, 102, 1253

Lawrence C. R., Schneider D. P., Schmidet M., Bennett C. L., Hewitt J. N., Burke B. F., Turner E. L., Gunn J. E., 1984, Science, 223, 46

Lawrence C. R., Neugebauer G., Matthews K., 1993, AJ, 105, 17

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

Ledlow, M. J., & Owen, F. N. 1996, AJ, 112, 9

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

Mao S., Witt H. J., Koopmans L. V. E., 2001, MNRAS, 323, 301

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

Marlow D. R., Browne I. W. A., Jackson N., Wilkinson P. N., 1999, MNRAS, 305, 15

Martin N. F., Ibata R. A., Chapman S. C., Irwin M., Lewis G. F., 2007, MNRAS, 380, 281

McKean J. P. et al., 2005, MNRAS, 356, 1009

McKean J. P. et al., 2007, MNRAS, 378, 109

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

Mittal R., Porcas R., Wucknitz O., 2007, A&A, 465, 405

Momcheva I., Williams K., Keeton C., Zabludoff A., 2006, ApJ, 641, 169

Moore B., Governato F., Quinn T., Stadel J., Lake G., 1998, ApJL, 499, L5

Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, ApJ, 524, L19

Moustakas L. A. et al., 2007, ApJL, 660, L31

Muñoz J. A. et al., 2001, ApJ, 546, 769

Myers S. T. et al., 2003, MNRAS, 341, 1

Nair S., Garrett M. A., 1997, MNRAS, 284, 58

Narasimha D., Subramanian K., Chitre S. M., 1984, ApJ, 283, 512

Narasimha D., Subramanian K., Chitre S. M., 1986, Nature, 321, 45

Narasimha D., Subramanian K., Chitre S. M., 1987, ApJ, 315, 434

Narasimha D., Chitre S. M., 1989, AJ, 97, 327

Narayan R., Bartelmann M., 1996, ArXiv Astrophysics e-prints, arXiv:astro-ph/9606001

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

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

Nityananda R., Narayan R., 1982, Journal of Astrophysics and Astronomy, 3, 419

Norberg P., et al., 2002, MNRAS, 332, 827

Oguri M. et al., 2004, ApJ, 605, 78

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

Oguri M., 2006, MNRAS, 367, 1241

Oguri M., 2007, ApJ, 660, 1

Patnaik A. R., Browne I. W. A., Walsh D., Chaffee F. H., Foltz C. B., 1992, MNRAS, 259, 1P

Patnaik A. R., Browne I. W. A., King L. J., Muxlow T. W. B., Walsh D., Wilkinson P. N., 1993, MNRAS, 261, 435

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

Perley R. A., Schwab F. R., Bridle A. H., 1989, Synthesis Imaging in Radio Astronomy, 6,

Perlmutter S., et al., 1999, ApJ, 517, 565

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Cambridge: University Press, —c1992, 2nd ed.

Read J. I., Saha P., Macciò A. V., 2007, ApJ, 667, 645

Readhead A. C. S., Wilkinson P. N., 1978, ApJ, 223, 25

Ros E., Guirado J. C., Marcaide J. M., Pérez-Torres M. A., Falco E. E., Muñoz J. A., Alberdi A., Lara L., 2000, A&A, 362, 845

Rubin V. C., Burstein D., Ford W. K., Jr., Thonnard N., 1985, ApJ, 289, 81

Rusin D., Ma C., 2001, ApJL, 549, L33

Rusin D., Keeton C. R., Winn J. N., 2005, ApJL, 627, L93

Salucci P., Burkert A., 2000, ApJ, 537, L9

Sand D. J., Treu T., Ellis R. S., Smith G. P., 2005, ApJ, 627, 32

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

Saha P., Coles J., Macciò A. V., Williams L. L. R., 2006, ApJ, 650, L17

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

Schneider D. P., Lawrence C. R., Schmidt M., et al., 1985, ApJ, 294, 66

Schneider D. P., Gunn J. E., Turner E. L., et al., 1986, AJ, 91, 991

Schneider D. P., Gunn J. E., Turner E. L., et al., 1987, AJ, 94, 12

Schneider P., Ehlers J., Falco E. E., 1992, Gravitational Lenses, Springer-Verlag. Berlin

Shin E. M., Evans N. W., 2008, MNRAS, 297

Simon J. D., Geha M., 2007, ApJ, 670, 313

Sluse D. et al., 2003, A&A, 406, L43

Smoot G. F., et al., 1992, ApJ, 396, L1

Soucail G., Kneib J.-P., Jaunsen A. O., et al., 2001, A&A, 367, 741

Spergel D. N., et al., 2003, ApJS, 148, 175

Taylor G. B., Carilli C. L., Perley R. A., 1999, ASP Conf. Ser. 180: Synthesis Imaging in Radio Astronomy II, 180

Thompson A. R., Moran J. M., Swenson G. W., 1986, New York, Wiley-Interscience, 1986, 554 p.

Thompson A. R., 1999, Synthesis Imaging in Radio Astronomy II, 180, 11

Tremaine S. et al., 2002, ApJ, 574, 740

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

Treu T., Koopmans L. V. E., 2004, ApJ, 611, 739

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

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

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

Williams K. A., Momcheva I., Keeton C. R., Zabludoff A. I., Lehár J., 2006, ApJ, 646, 85

Winn J. N. et al., 2002, AJ, 123, 10

Winn J. N., Rusin D., Kochanek C. S., 2003, ApJ, 587, 80

Winn J. N., Rusin D., Kochanek C. S., 2004, Nature, 427, 613

Winn J. N., et al., 2004, AJ, 128, 2696

Wisotzki L., Koehler T., Kayser R., Reimers D., 1993, A&A, 278, L15

Wucknitz O., Wisotzki L., Lopez S., Gregg M. D., 2003, A&A, 405, 445

Wucknitz O., 2004, MNRAS, 349, 1

Yamada T., Yamazaki S., Hattori M., Soucail G., Kneib J.-P., 2001, A&A, 367, 51

York T., Jackson N., Browne I. W. A., Wucknitz O., Skelton J. E., 2005, MNRAS, 357, 124

Zensus J. A., Diamond P. J., Napier P. J., 1995, ASP Conf. Ser. 82: Very Long Baseline Interferometry and the VLBA, 82

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

Zwicky F., 1937, ApJ, 86, 217

Acknowledgments

Finally, the day has arrived when I can hold others responsible for my work. Let me begin by acknowledging the support of my supervisors Prof. Peter Schneider and Prof. Johannes Schmid-Burgk. I want to thank for giving critical comments and useful suggestions. I would also like to thank Prof. Anton Zensus for being supportive and providing educating work environment in the VLBI group of MPIfR. I am thankful to the IMPRS which encouraged interaction amongst students and scientists, and gave me the opportunity to improve upon my soft skills. I would like to express my gratitude towards Prof. Ian Browne and the ANGLES team because of whom it was possible to get acquainted with the research scientists and the research, specifically, in Gravitational Lensing.

I am extremely thankful to Richard Porcas with whom I interacted since the time I joined this research institute. I hope I have absorbed at least some of the vast knowledge and experience he has in Radio astronomy, Lensing and various things in life including cigars and puzzles. My special and heart-felt thanks to Dr. John McKean who has been there almost like a supervisor and a friend. I really appreciate his encouragement and support throughout especially during the last phase of my PhD and not to mention his great sense of humour which always cheered me up. I am very much grateful to Chuck Keeton, Leon Koopmans, Edward Boyce and Neal Dalal for their immediate respones to all my queries and the discussions I had with them, and in particular, Olaf Wucknitz for clarifying doubts and for giving intellectual remarks from time to time.

I acknowledge the technical support provided by Walter Alef and the intellectual support provided by Alan Roy and Silke Britzen for involving me in the seminar activities. I am grateful to Andrei Lobanov for his advice, suggestions and witty remarks; Dharam Vir Lal for various discussions and timely assistance; Eduardo Ros for overall encouragement and his cheerful personality; Yuri Kovalev for sharing his knowledge and providing several suggestions scientific and otherwise; and introducing me to Russian language. Furthermore, I want to thank Dharam and Babita for being there like a family.

I appreciate very much the support of my friends, Thushara particularly for being my local Guardian, Poornnima and Suchita for general support, tasty food and interesting discussions, Girisan for supplying me with dozens of movies. My ex-office-mates Krisztina, Rupal, Violette, Sang-Sung who have helped me from time to time in several ways for which I am grateful to them. My colleagues Peter, Manolis, Nadia, Nicola, and many more from the new IMPRS generation have been a happy bunch to be with. Thanks for the Fiddlers' nights.

I really thank my family for their unconditional love and support I received especially Dad who never encouraged me to watch movies and insisted to focus on my work. I want to thank my friends back in India and elsewhere too especially, Raj who bore with my unwavering complaints about the harsh realities of life. I would like to mention that

it was an enjoyable and great learning experience for me. In the end and above all, I am indebted to someone who worked harder than me and endured my mood swings. “I cannot imagine the work presented in this thesis and myself without you, Surhud”.