ASTROPHYSICAL APPLICATIONS OF THE EINSTEIN RING GRAVITATIONAL LENS, MG1131+0456

by

GRACE HSIU-LING CHEN

Bachelor of Science, San Francisco State University, 1990

SUBMITTED TO THE DEPARTMENT OF PHYSICS
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1995

Signature of Author: _____

Department of Physics
May 1995

Certified by: _____

Jacqueline N. Hewitt
Thesis Supervisor

Accepted by: _____

George F. Koster
Chairman, Graduate Committee

ASTROPHYSICAL APPLICATIONS OF THE EINSTEIN RING GRAVITATIONAL LENS, MG1131+0456

by

GRACE HSIU-LING CHEN

ABSTRACT

MG1131+0456 has been observed extensively with the Very Large Array telescope, and its complex morphology has provided the opportunity for investigating three important astrophysical problems: (1) the mass distribution in the lens galaxy, (2) the structure of the rotation measure distribution in the lens, and (3) the possibility of using the system to constrain the Hubble parameter.

We determine the mass distribution of the lens galaxy by modeling the multifrequency VLA images of the system. Among the models we have explored, the profile of the surface mass density of the best lens model contains a substantial core radius and declines asymptotically as $r^{-1.2\pm0.2}$. The mass inside the lens enclosed by the average ring is very well constrained. If the source is at redshift of 2.0 and the lens is at redshift of 0.5, the mass of the lens enclosed by the ring radius is $\sim 2.0 \times 10^{11}$ solar mass.

We develop a new technique to determine the structure of the rotation measure distribution in the lens. We have detected variations in the rotation measure distribution in the lens galaxy of MG1131, and the structure of the variations indicates that the magnetic field in the lens galaxy is likely to change on a scale of $\sim 1''$.

The VLA observations of MG1131+0456 at 8 and 15 GHz taken at several epochs show that the compact components in the system are weakly variable. Results of Monte Carlo simulations show that a time delay should be measurable if approximately 100 observations are obtained. Hence, one should seriously consider MG1131+0456 for measuring the Hubble parameter.

Table of Contents

Acknowledgement

I could not have completed this thesis without the following people's support and love ....

First of all, I would like to thank my thesis advisor, Prof. Jacqueline N. Hewitt, for giving me the opportunity to conduct researches in the field of gravitational lensing and the support and encouragement she has been given me in the last five years. I would also like to thank Prof. Christopher S. Kochanek, from whom I have learned how to attack problems critically, and Dr. Edward Fomolont, who was always there for giving me valuable advice on analyzing the VLA data.

There are no words big enough to express my gratitude for the the love my parents, sister, and brother have given me. Without their support and their beliefs on my ability, I would have left MIT four years ago, and this thesis would have never been written.

I would also like to thank Charlie K., Sam, Debbie, Chris, Max, Andre, Lori, Cathy, John E., John C., Frowny, Ann C., Ann S., and Prof. Bernard Burke for making the MIT radio group a pleasant place to work in. I would like to thank Debbie, Charlie, and Cathy, especially, for taking the time to proofread this thesis carefully.

To Charlie Collins, I would have gone insane by now without the laughters your son, Michael, and your wife have brought during our routine lunches on Monday. Thank you all!

Finally, I would like to thank a very special friend, *my Bär*, for his unmeasurable love and care. I will always treasure the time we have spent together here.

1. Introduction

1.1 History of Gravitational Lenses

The idea that a light ray can be deflected by the presence of a massive object was first considered by Sir Isaac Newton in the early 1700s. One hundred years later, based on Newtonian gravitational theory, Soldner calculated the deflection of the light ray from its straight path when it passes by a massive object (Soldner 1804). After another one hundred years, Einstein performed the same calculation using his theory of general relativity and obtained a result that is twice as the value predicted by Soldner (Einstein 1915). In 1919, Dyson set up an experiment to measure the deflection of a starlight passing by the limb of the sun during a solar eclipse and confirmed Einstein's prediction (Dyson 1919). Later in that year, Lodge (1919) introduced the interesting idea that it is possible for a gravitational potential to be strong enough to create more than one image of a background source. Chwolson (1924) and Einstein (1936) both applied Lodge's idea of multiple imaging to nearby stars and concluded that the angular separation between the images would be too small to be resolved by ground-based optical telescopes and that the phenomenon of multiple imaging would not be observable. It did not take the ingenious Zwicky long to realize that not only stars but also distant galaxies could act as lenses (Zwicky 1937). In that case, the angular separation between a pair of images would be detectable, and furthermore, if detected, the gravitational lens system could be used for numerous astrophysical applications. In Zwicky's 1937 paper, he stated:

> The discovery of images of nebulae which are formed through the gravitational fields of nearby nebulae would be of considerable interest for a number of reasons.
>
> (1) It should furnish an additional test for the general theory of relativity.
>
> (2) It would enable us to see nebulae at distances greater than those ordinarily reached by even the greatest telescopes....
>
> (3) ...Observations of the deflection of light around nebulae may provide the most direct determination of nebular masses ....

Zwicky's pioneering ideas led the field of gravitational lensing to extragalactic astronomy.

After 1937, the field remained dormant for roughly thirty years. New applications, such as the use of gravitational lenses for distance measurements (Klimov 1963; Liebes 1964; Refsdal 1964, 1966) and as probes to determine the stellar composition (Chang & Refsdal 1979) were considered in 1960's and 1970's. However, the field remained as an esoteric playground for relatively few theorists during that period. The situation changed in 1979 after the discovery of the first gravitational lens system (Walsh, Carswell, & Weymann 1979). As part of a project identifying the optical counterparts to sources in a radio survey at 996 MHz (Cohen et al. 1977), Walsh et al. (1979) discovered a system, named Q0957+561, consisting of two compact components both at redshifts of $z \sim 1.4$. The similarities between their spectra and the presence of a lensing galaxy at $z \sim 0.36$ (Stockton 1980; Young et al. 1980) verified the lensing hypothesis. After the discovery, many began to show tremendous interest in the field of lensing, and numerous papers concerning both observational and theoretical work appeared. The second gravitational lens system was discovered only one year later (Weymann et al. 1980). In 1986, a second type of lens, a long circular arc with angular radius $\sim 25''$ in Abell 370, was discovered by Soucail et al. (1987), and in 1987, a third type of lens, a ring (known as MG1131+0456), was discovered by Hewitt et al. (1988) as part of the radio lens survey project conducted at MIT (Hewitt 1986; Lehár 1991).

At the present time, there are more than 20 confirmed lenses, and more are waiting to be discovered. The field of gravitational lensing is no longer considered an "esoteric field" but one of the most active fields in extragalactic astronomy.

1.2 Theory of Gravitational Lensing

Since the purpose of this section is to help the readers to understand the content

of this thesis, presenting a complete discussion of the theory of gravitational lensing would not be appropriate. Instead, only the basic notions, concepts, and theories of gravitational lensing relating to this thesis will be discussed. A complete treatment of gravitational lensing theory is presented in the book, *Gravitational Lenses* (Schneider, Ehlers, & Falco 1992). Several excellent review papers on the subject (Blandford & Kochanek 1987; Canizares 1987; Blandford & Narayan 1992) are also available in the literature. These sources are recommended for readers seeking a more thorough discussion of this subject.

1.2.1 Gravitational Lensing Optics

In the most simple configuration, there are only a few elements in gravitational lensing optics: a source, S, a massive object acting as a lens, L, and an observer, O. Figure 1.1 displays a schematic of the optics. The gravitational potential of the massive object (L) causes the light ray emitted from the source to be deflected from its straight path by an amount α', so the observer sees an image formed at an angular location $\mathbf{x}$ instead of $\mathbf{u}$. If the gravitational potential of the lens is known, the amount of the bending, α ($\alpha = \mathbf{x} - \mathbf{u}$) or $\alpha' = (D_{os}/D_{ol})\alpha$ can be calculated exactly from Einstein's theory of general relativity. In an astrophysical setting, stars, globular clusters, galaxies, and galaxy clusters are all capable of producing lensing effects.

An elegant way to understand image formation in gravitational lensing is to utilize Fermat's principle. This idea was fully developed by Schneider (1985) and Blandford & Narayan (1986). Let us refer back to the optical geometry laid out in Figure 1.1. When the lens is absent, the light travels on the path **SO**; and when the lens is present, it travels on the path **SPO**. The light path in each case is the null geodesic of the metric connecting the observer and the source, and the light travel time can be calculated using the theory of general relativity. In a weak field approximation, the difference

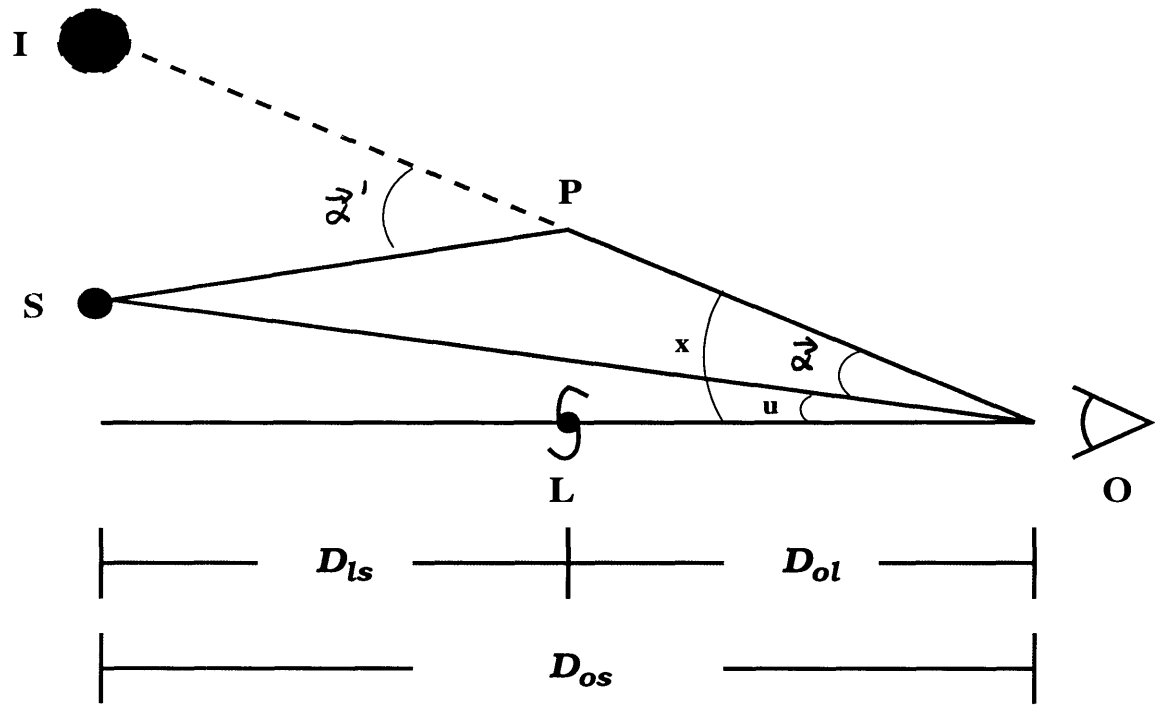


Figure 1.1 – Basic geometry of gravitational lensing optics. A light ray from a source S at redshift z_s is deflected due to the presence of the lens L. Assuming that the lens is thin compared to the total path length, the amount of deflection can be described by a single parameter α'. Because of the deflection, an observer sees the image of the source appearing at an angular position $\mathbf{x}$ instead of the true source position $\mathbf{u}$. D_{ij} are the angular diameter distances between the source, lens, and observer.

in the light travel time between these two paths, known as the time delay, has the form,

$$t = (1+z_l)\Big[\frac{D_{os}D_{ol}}{2cD_{ls}}(\mathbf{x}-\mathbf{u})^2 - 2\int_{light\ path}\frac{\Phi}{c^3}ds\Big]. \qquad (1-1)$$

D_{ij} are the angular diameter distances between objects i and j, c is the speed of light, z_l denotes the redshift of the lens, and Φ is the three-dimensional Newtonian potential of the lens. In general, D_{ij} depends on the model of cosmology. In a Friedman cosmology,

$$D_{ij} = \frac{2c}{\Omega_o H_o}\frac{(G_iG_j + \Omega_o - 1)(G_j - G_i)}{(1+z_i)(1+z_j)^2}$$

$$G_i = (1 + \Omega_o z_i), \qquad (1-2)$$

where Ω_o is the the density parameter, z_i is the redshift of the object i, and H_o is the Hubble constant*. Each term in Eqn (1-1) represents a different physical process. The first term accounts for the extra path on which light needs to travel due to the presence of the lens, and the second term accounts for the delay in time when an object passes through a gravitational potential well as predicted by general relativity.

The three-dimensional Newtonian potential, Φ, satisfies the Poisson equation,

$$\nabla^2 \Phi = 4\pi\rho, \qquad (1-3)$$

where ρ is the mass density of the lens. It is useful to define a rescaled two-dimensional Newtonian potential

$$\phi = 2\frac{D_{ls}}{D_{os}D_{ol}} \int_{light\ path} \frac{\Phi}{c^2} ds \qquad (1-4),$$

which also has a similar Poisson equation as Eqn (1-3),

$$\nabla^2 \phi = \frac{8\pi D_{os} D_{ol} \Sigma}{D_{ls}} = \frac{2\Sigma}{\Sigma_c}, \qquad (1-5)$$

where Σ is the surface mass density and $\Sigma_c = D_{ls}/4\pi D_{os} D_{ol}$ – the critical surface density – is the minimum surface density a homogeneous lens must have in order to produce multiple images (see discussion in Section 1.2.2). Expressing the time delay in terms of ϕ, we find

$$t = \frac{(1+z_l) D_{os} D_{ol}}{c D_{ls}} \left[\frac{1}{2}(\mathbf{x} - \mathbf{u})^2 - \phi\right]. \qquad (1-6)$$

Now we are ready to apply Fermat's principle. Fermat's principle states that light travels on a path that is stationary on the time surface, which means that images are formed on the extrema of the time surface. Applying this idea to Eqn (1-6), the location of the image $\mathbf{x}$ corresponding to the source $\mathbf{u}$ can be found at

$$\mathbf{u} = \mathbf{x} - \nabla\phi(\mathbf{x}), \qquad (1-7)$$

* The Hubble constant is assumed to be $H_o = 100h$ km Mpc^{-1} s^{-1}.

and the deflection angle α is

$$\alpha = \nabla\phi. \qquad (1-8)$$

The gravitational deflection does not add emissivity to the light bundle or absorb light from it, so the surface brightness of a bundle of light stays unchanged through gravitational lensing. However, the cross section of the bundle is altered by the deflection. Because of the change in the cross section, the image can appear larger (magnified) or smaller (demagnified) than the source size. The Jacobian of the transformation between $\mathbf{u}$ and $\mathbf{x}$ computes the ratio of the elemental areas in the two planes, so the image magnification can be evaluated effectively from the Jacobian. The magnification tensor, M_{ij}, can be calculated as

$$M_{ij} = \frac{\partial x_i}{\partial u_j} = [(\delta_{ij} - \partial_{ij})\phi(\mathbf{x})]^{-1}, \qquad (1-9)$$

and the total magnification is

$$M = \det |M_{ij}| .$$

For a system containing a single deflector, the magnification tensor is symmetric, so the matrix can be diagonalized into

$$M_{ij} = \begin{pmatrix} \frac{1}{\rho_1} & 0 \\ 0 & \frac{1}{\rho_2} \end{pmatrix},$$

where ρ_1 and ρ_2 are both real. The diagonalized matrix provides an explicit description of the shape and size of the image. The shape of the image can be constructed by compressing or expanding the shape of the source by an amount $1/\rho_1$ and $1/\rho_2$ along the principal axes. Because the eigenvalues are real, no rotation is involved. In general, the diagonalized matrix can be written in the form of

$$M_{ij} = \begin{pmatrix} \frac{1}{\kappa+\gamma} & 0 \\ 0 & \frac{1}{\kappa-\gamma} \end{pmatrix}.$$

The quantity $\kappa = 1 - \Sigma/\Sigma_c$ governs the expansion or contraction along the principal axes and is known as the *convergence*. The quantity γ governs the amount of shear along the principal axes, so it is naturally called the *shear term* in lensing.

1.2.2 Image Multiplicities

From the structure of the lens equation (Eqn 1-7), it is clear that there is only one source corresponding to each image, but the converse is not true. In other words, depending on the potential of the lens, the mapping between **u** and **x** can be one-to-many, so a source can be lensed into more than one image.

The counting of images is rather simple; it is just the number of solutions which satisfy the relation

$$\alpha = \nabla\phi(\mathbf{x}) = \mathbf{x} - \mathbf{u}.$$

One interesting characteristic of gravitational lensing is that the image multiplicity changes as a function of the source location. This concept can be most easily understood with graphical methods. For simplicity, we assume that ϕ has circular symmetry and that $\nabla\phi$ is as shown by the curve in Figure 1.2. (The purpose of assuming circular symmetry is to reduce the lens equation to one dimension, so **x-u** is a straight line with its intercept on the abscissa representing the value of **u**. The result can be generalized to all other cases.) Depending on the source location, there can be one or more solutions. For the case demonstrated here (see Figure 1.2), there is (are) one (three) solution(s) if the location of source is at $|u| > u_{mult}$ ($|u| < u_{mult}$). The regions in which a source can be multiply imaged (i.e. $|u| < u_{mult}$) are called the multiply imaged regions, and regions in which a source can only be singly imaged (i.e. $|u| > r_{mult}$) are called the singly imaged regions. When the source is crossing from a multiply imaged region to a singly imaged region, two images merge, magnify infinitely, and then disappear. The locus of the merging in the image plane is known as the *critical line*, and the associated locus in the source plane is called the *caustic line*. In general, caustic lines divide the source plane into regions of image multiplicity differing by ± 2. Of course, not all potentials are not strong enough to produce multiple images. The term *weak lensing* refers to the cases where the potentials are

incapable of multiple imaging whereas *strong lensing* refers to the other cases.

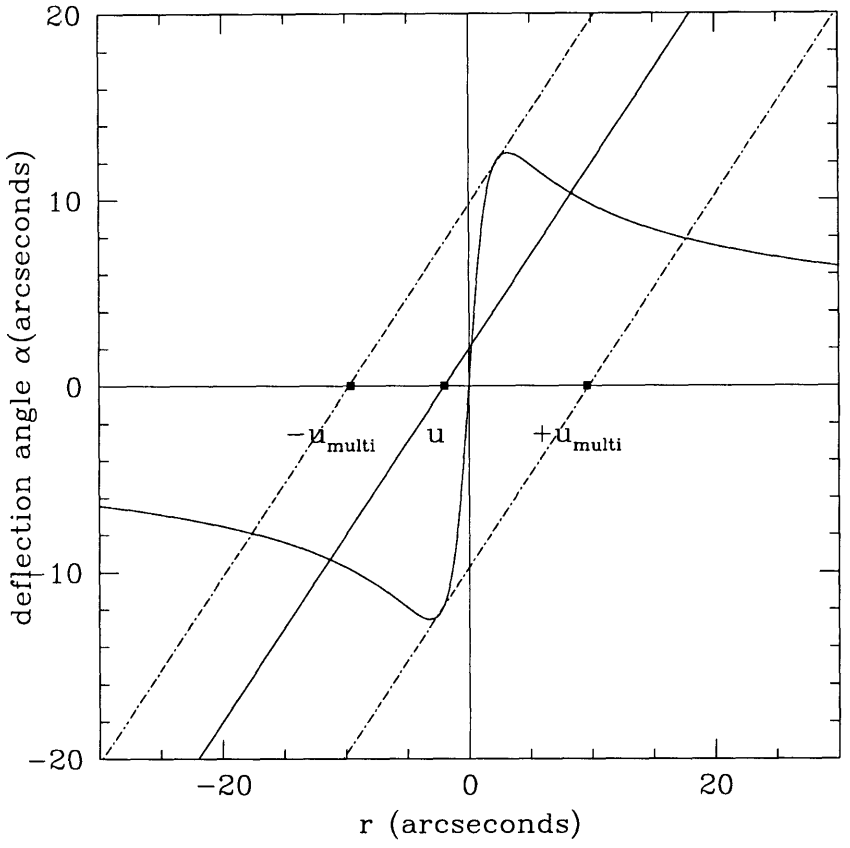


Figure 1.2 — A plot demonstrating the concept of multiple imaging in gravitational lensing. For a given source at position **u**, the number of image corresponding to this source is just the number of solutions satisfying the lens equation $\alpha = \nabla\phi(\mathbf{x}) = \mathbf{x} - \mathbf{u}$. The solid curve represents $\nabla\phi(\mathbf{x})$. The solid straight line represents the function **x - u** whose intercept on the abscissa denotes the value of **u**. When the source is located in the multiply imaged region ($|u| < u_{mult}$), there are three solutions to the lens equation, hence three images. On the other hand, when the source is located in the region ($|u| > u_{mult}$), there is only one solution, hence one image. The boundary of the multiply imaged region is marked by the two dash-dotted lines.

1.2.3 Simple Lens Models

The simplest type of gravitational potential is that of a point mass. Light traveling past a point mass M is deflected by an amount $\alpha' = 4GM/c^2 D_{ol} x$ (where $\alpha = [D_{ol}/D_{os}]\alpha'$), where x is the image position (in arc seconds) relative to the point mass. Figure 1.3 shows the deflection angle as a function of x produced by a point

mass with M = $10M_{\odot}$ located at a distance 10 pc from the observer. The straight line in the figures represents the function $\alpha = \mathbf{x} - \mathbf{u}$ when the source is at 20 pc away from the observer. It is clear from Figure 1.3 that the multiply imaged region for a point mass lens extends to infinity, and sources at all positions, except at the optical axis, will always be lensed into two images. When the source is on the optical axis, instead of two images, a ring of images with an angular radius of $\theta_E = \sqrt{4GMD_{ls}/c^2 D_{os} D_{ol}}$ appears. The ring and θ_E are generally referred to as the Einstein ring and the Einstein radius. For a source located at any other position u, two images appear at $x_{\pm} = (u \pm \sqrt{u^2 + 4\theta_E^2})/2$. The flux magnifications associated with $x_{\pm}$ are $M_{\pm} = [2 \pm u/\sqrt{(u^2 + \theta_E^2)} \pm \sqrt{(u^2 + \theta_E^2)}/u]/4$.

One other frequently used simple lens model is the singular isothermal sphere model. The surface mass density distribution in a singular isothermal sphere lens has the form $\Sigma = \sigma^2/2GD_{ol}x$, where σ is the velocity dispersion of the objects in the lens galaxy. The simplicity of the potential and its ability to explain the flat rotation curve of spiral galaxies make it one of the "beloved" models for describing the potentials of galaxies or clusters of galaxies. The two-dimensional potential of this model is $\phi = br$, where $b = 4\pi D_{ls}\sigma^2/D_{os}c^2$ (in arc seconds). The most peculiar characteristic of this potential is the constant deflection angle (see Figure 1.4). As in the case of the point mass potential, a source on the optical axis is imaged into a ring with $\theta_E = b$. However, multiple imaging is only permitted when the source is located at $|u| < b$. In the case of multiple imaging, because of the constant deflection, the image separation between a pair of images is always $2b$, and the geometric part of the time delay between this pair of image is zero.

1.3 Applications of Gravitational Lensing

1.3.1 Mass Distribution of the Lens

The most direct application of gravitational lensing is to determine

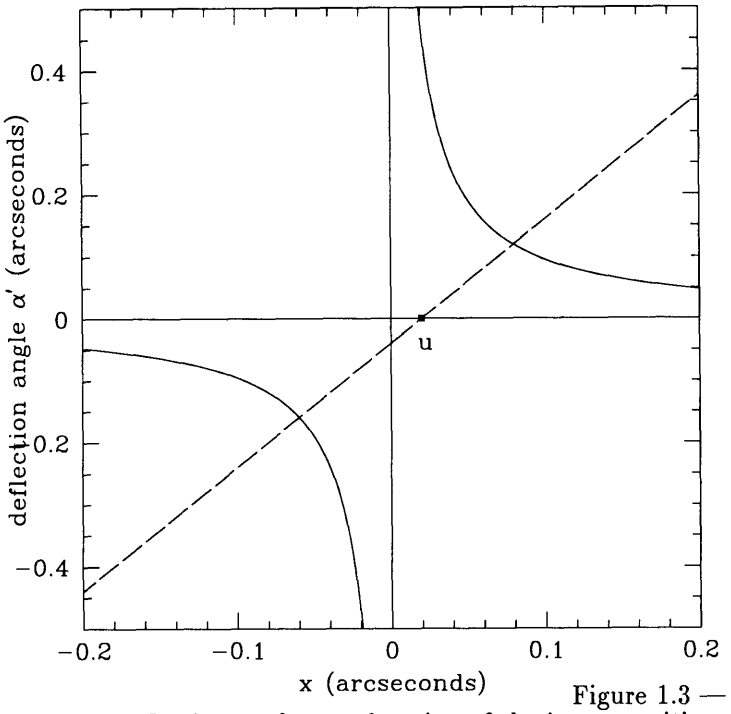


Figure 1.3 — The solid curve represents the deflection angle as a function of the image position x for a point mass model. The point mass is assumed to have a mass $M = 10M_{\odot}$ and a position 10pc away from the observer. The dashed line represents the function $\alpha =$ **x - u** when the source is at 20 pc away from the observer. The interceptions between the solid curves and the dashed line represent the locations of the images.

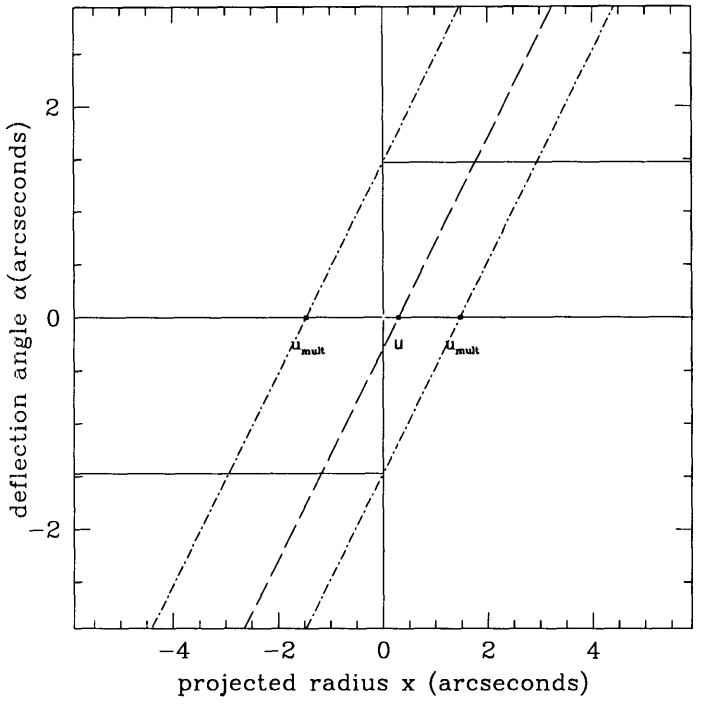


Figure 1.4 — The horizontal lines represent the constant deflection angle of a singular isothermal sphere lens model. The dashed line shows the function **x - u**. The boundary of the multiply imaged region is marked by the two dash-dotted lines. $u_{mult} = b$.

the mass distribution of the lens and to investigate the amount of dark matter embedded in the lens. The traditional method of determining the mass distribution of a galaxy or a cluster of galaxies is to model the velocity field, but there are two major problems with this method. The first is the need to find a luminous dynamical tracer. This often limits the determination of the mass distribution to regions that are dominated by luminous matter. The second is that only the line of sight component of the velocity field is detectable; the lack of other components makes the determination of the mass distribution very difficult. Both limitations, hence, reduce the effectiveness of this technique. Since gravitational lensing depends only on the gravitational potential of the lens galaxy, lenses should provide new and more direct determinations of the matter distribution.

The possibility of using gravitational lensing to investigate the structure of the mass distribution in galaxies was realized even before the discovery of the first gravitational lens system (Bourassa & Kantowski 1975). Most lensed systems used for this kind of work consist of two or four unresolved images. Because of the paucity of constraints and the restriction of the images to a limited range of radii from the center of the lens, these systems, having little or no constraints on the radial distribution of matter, cannot distinguish between models with more complex structure (Kochanek 1991a; Wambsganss & Paczyński 1994). In the case where an extended background source is lensed into images located in regions over a wide range of radii from the lens, the number of constraints is greatly increased. these systems should give a better determination of the mass distribution of the lens. The radio rings (see Patnaik 1994 for a review) and the cluster arcs (see Soucial & Mellier 1994 for a review) belong to this class of lens.

1.3.2 Cosmography

The idea of measuring H_o using gravitational lenses was first recognized by Refsdal in 1964. The time delay between a pair of images depends on the angular diameter distance D_{ij} and the two-dimensional lens potential ϕ (see eqn 1-6). Since D_{ij} is inversely proportional to H_o, the time delay is also inversely proportional to H_o. Thus, a lens system with reliable measurements of the time delay, z_s, and z_l and a well constrained lens model provides an opportunity for one to determine H_o with a method differing from the traditional ones (Jacoby et. al. 1991). Unfortunately, it is difficult to find a system with all the needed ingredients. Today, H_o has only been determined from one gravitational lens system, Q0957+561, and the H_o determination from this system is still uncertain because of the uncertainties of the time delay measurements (Vanderriest et al. 1989; Schild & Cholfin 1986; Schild 1990; Lehár et al. 1992; Press, Rybicki, & Hewitt 1991a, 1991b) and of the lens model (Falco, Gorenstein & Shapiro 1991; Kochanek 1991b; Bernstein, Tyson, & Kochanek 1993; Dahle, Maddox, & Lilje 1994). In theory, if the redshifts of the lens and the source are known, the lens systems composed of both extended structure (which provides strong constraints on the lens model) and compact components (from which the time delay can be measured) would be more suitable for this kind of work. Several groups are currently monitoring other gravitational lenses (van Ommen et al. 1995; Moore & Hewitt 1995), with the goal of measuring the time delays in these systems. Efforts are underway to estimate H_o from lensed systems with extended emission.

Testing the existence of the cosmological constant with gravitational lensing statistics is another important application of lenses. The frequency of lensing depends on the volume of the universe, so the cosmological constant has a strong impact on the number of lenses which we expect to find. It is well known that flat cosmologies dominated by a cosmological constant predict a much larger lensing frequency than

the standard Friedman-Robertson-Walker (FRW) cosmology (Turner 1990; Fukugita & Turner 1991; Fukugita et al. 1992). Using optical survey data, Fukugita & Turner (1991) and Fukugita et al. (1992) found that cosmologies with $\lambda_o \geq 0.95$ are inconsistent with the observed frequency of lensing. The probability distribution of the lens redshifts also depends on the cosmological constant (Kochanek 1992). Using the distribution of the known lens redshifts, Kochanek (1992) concludes that a flat universe consisting of a large cosmological constant is not consistent with the observed lens redshift distribution.

1.4 Einstein Ring Lens System, MG1131+0456

MG1131+0456, the first Einstein ring gravitational lens system detected (Hewitt et al. 1988), was mapped as part of the MG gravitational lens search project (Hewitt 1986; Lehár 1991). The original 5 and 15 GHz VLA maps of this system revealed an unusual morphology consisting of an elliptical ring, two compact components, and one low surface brightness extended component detectable at 5 GHz. It is believed that an extended component located directly behind the lens is imaged into the ring, one compact component located slightly away from the optical axis is doubly imaged, and one other extended component located far away from the optical axis is only singly imaged. In addition to the radio observations, Hewitt et al. (1988) also obtained an R band optical image and a spectrum of this source with the National Optical Astronomy Observatory's 4-m telescope on Kitt Peak, hoping to find an optical identification. They found a faint optical object ($m_R \sim 22$) with a featureless spectrum near the radio position of MG1131. No lines were detected in the spectrum, so the redshifts of neither the lens or the source could be determined.

A year after the discovery of MG1131, Kochanek et al. (1989) developed a sophisticated inversion algorithm, the Ring Cycle, to model the 15 GHz map from Hewitt et al. (1988) under the assumption that the flux densities in the image represented

the true surface brightness of the source. This assumption has a severe shortcoming since the resolution of even the 15 GHz maps has a perceptible effect on the structure of the image, and the modeled ring image always appears to be much less elliptical than the observed one. The 5 GHz image simply could not be modeled because of the breakdown of the assumption. Nevertheless, Kochanek et al. found a best fit elliptical isothermal model for the lens. The inverted source exhibited structures similar to that observed in typical extragalactic radio sources - a core, two lobes, and a possible jet.

In 1991, Hammer et al. (1991) used the ESO/NTT telescope to obtain optical images (through the B, V, R, and Gunn i filters) and spectroscopic observations with spectral coverage between 5000 $\AA$ $\sim 9300\AA$ at the position of MG1131. The system was clearly detected in V, R, and the Gunn i bands, but was too faint to be detected in the B band. The object seen in the optical images appears more extended than an elliptical galaxy. After subtracting the brightness profile of an elliptical galaxy from the image, they found a ring-like residual similar to the radio ring, which they interpreted as the lensed image of the optical emission from the radio source. However, their claim has not been confirmed by others. The source they detected also has a featureless spectrum, in agreement with Hewitt et al.'s finding.

MG1131 has also been detected at 2.2 μm and 1.2 μm (Annis 1992, Larkin et al. 1994), and its infrared structure is complex. At 2.2 μm, both extended diffuse emission (shaped like an elliptical halo) and two prominent compact components are visible (Annis 1992; Larkin et al. 1994); however, only the diffuse emission is visible at 1.2 μm (Larkin et al. 1994). After subtracting the galaxy profile estimated using the 1.2 μm image from the 2.2 μm image, Larkin et al. found that the positions and the orientation of the infrared compact components agreed with the compact components detected at radio wavelengths. Both Annis and Larkin et al. detected two fainter galaxies near MG1131+0456 and an excess of galaxies within 20″ of MG1131.

Therefore, it is possible that there is a cluster in the field.

1.5 Outline

The goal of this thesis is to use MG1131 as a means to (1) determine the mass distribution in the lens galaxy, (2) study the rotation measure structure in the lens, and (3) discuss the possibility of determining the Hubble parameter from this system. The organization is as follows. Chapter 2 describes the multi-frequency radio observations of MG1131 and the calibration procedures. Chapter 3 presents the analysis of the continuum and the linearly polarized emission detected at 5, 8, 15, and 22 GHz and a discussion of the nature of components in the system. Chapter 4 describes the methodology used for modeling MG1131. Chapter 5 reports the results of modeling the 5, 8, and 15 GHz images and discusses the mass distribution of the lens galaxy inferred from the best lens model. Chapter 6 describes a new method of studying the structure of rotation measure in high redshifted galaxies using gravitational lenses and the results when this method is applied to MG1131. Chapter 7 discusses a variability study of the compact components in MG1131 and the possibility of using it to determine the Hubble parameter. A summary is presented in Chapter 8

References

Annis, J. A. 1992, Ap. J., **391**, L7.

Bernstein, G. M., Tyson, J. A. & Kochanek, C. S. 1993, Astron. J., **105**, 816.

Blandford, R. D. & Kochanek, C. S. 1987, *in Dark Matter in the Universe*, eds. Bahcall, J., Piran, R. & Weinberg, S., World Scientific, pp. 134.

Blandford, R. D. & Narayan, R. 1986, Ap. J., **310**, 568.

———————————— 1992, Ann. Rev, Astron. Astrophys., **30**, 311.

Bourassa, R. R. & Kantowski, R. 1975, Ap. J., **195**, 13.

Canizares, C. R. 1987, *in Observational Cosmology*, eds. Hewitt, A., Burbidge, G. &

Fang, L.-Z., IAU Symposium 124, pp. 729.

Chang K. & Rafsdel, S. 1979, Nature, **282**, 561.

Chwolson, O. 1924, Astro. Nachrichten, **221**, 329.

Cohen, A. M., Porcas, R. W., Browne, I. W. A., Daintree, E. J., & Walsh, D. 1977, Mem. R. Astro. Soc., **84**, 1.

Dahle, H., Maddox, S. J. & Lilje, P. B. 1994, Ap. J., **435**, L79.

Dyson, F. W. 1919, Observatory, **42**, 389.

Einstein, A. 1915, Preuss. Adak. Wiss. Berlin, Sitzber., **47**, 831.

———— 1936, Science, **84**, 506.

Falco, E. E., Borenstein, M. V. & Shapiro, I. I. 1991, Ap. J., **372**, 364.

Fukugita, M. & Turner, E. L. 1991, M.N.R.A.S, **253**, 99.

Fukugita, M., Futamase, T., Kasai, M., & Turner, E. L. 1992, Ap. J., **393**, 3.

Jacoby, G. H., Branch, D., Ciardullo, R., Davies, R. L., Harris, W. E., Pierce, M. J., Pritchet, C. J., Tonry, J. L. & Welch, D. 199?, Publications of the Astronomical Society of the Pacific, **104**, 599.

Hammer, F., Le Févre, R., Angonin, M. C., Meylan, G., Smette, A. & Surdej, J. 1991, Astron. Astrophys., **250**, L5.

Hewitt, J. N., 1986, *A Search for Gravitational Lensing*, Ph. D. thesis, Massachusetts Institute of Technology.

Hewitt, J. N., Turner, E. L., Schneider, D. P., Burke, B. F., Langston, G. I. & Lawrence, C. R. 1988, Nature, **333**, 537.

Klimov, Y. G. 1963, Sov. Phys. Doklady, **8**, 119.

Kochanek, C. S., Blandford, R. D., Lawrence, C. R. & Narayan, R. 1989, M.N.R.A.S., **238**, 43.

Kochanek, C. S. 1992, Ap. J., **397**, 381.

———— 1991a, Ap. J., **373**, 354.

———— 1991b, Ap. J., **382**, 58.

Lehár 1991, *The Time Delay in the Double Quasar 0957+561 and a Search for Gravitational Lenses*, Ph. D. thesis, Massachusetts Institute of Technology.

Lehár, J., Hewitt, J. N., Roberts, D. H. & Burke, B. F. 1992, Ap. J., **384**, 453.

Larkin, J. E., Matthews, K., Lawrence, C. R., Graham, J. R., Harrison, W., Jernigan, G., Lin, S., Nelson, J., Neugebauer, G., Smith, G., Soifer, B. T. & Ziomkowski, C. 1994, Ap. J., **420**, L9.

Liebe Jr., S. 1964, Phys. Rev., **133**, B835.

Lodge, O. 1919, Nature, **104**, 354.

Moore, C. B. & Hewitt, J. N. 1995, submitted to Ap. J.

Patnaik, A. R. 1994, *in Gravitational Lenses in the Universe*, Proceedings of the 31st Liege International Astrophysics Colloquium, eds. Surdej, J., Fraipont-Caro, D., Gosset, E., Refsdal, S. & Remy. M., Liege University, 311.

Press, W. H., Rybicki, G. B., & Hewitt, J. N. 1991a, Ap. J., **385**, 404.

——————————————————— 1991b, Ap. J., **385**, 416.

Rafsdel, S. 1964, M.N.R.A.S., **128**, 307.

———— 1966, M.N.R.A.S., **134**, 315.

Schild, R. E. & Cholfin, B. 1986, Ap. J., **300**, 209.

Schild, R. E. 1990, Astron. J., **100**, 1771.

Schneider, P. 1985, Astron. Astrophys., **143**, 413.

Schneider, P., Ehlers, J. & Falco, E. E. 1992, *Gravitational Lenses*, Springer, New York.

Soldner, J. 1804, *Berliner Astro. Jahrb. 1804*, pp. 161.

Soucail, G., Fort, B., Mellier, Y. & Picat, J. P. 1987, Astron. Astrophys, **172**, L14.

Soucial, G. & Mellier, Y. 1994, *in Gravitational Lenses in the Universe*, Proceedings of the 31st Liege International Astrophysics Colloquium, eds. Surdej, J., Fraipont-Caro, D., Gosset, E., Refsdal, S. & Remy. M., Liege University, 595.

Stockton, A. 1980, Ap. J., **242**, L141.

Turner, E. L. 1990, Ap. J., **365**, L43.

Vanderriest, C., Schneider, J., Herpe, G., Chevreton, M., Moles, M. & Wlérick, G. 1989, Astron. Astrophys., **215**, 1.

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

van Ommen, T. D., Jones, D. L., Preston, R. A. & Jauncey, D. L. 1995, JPL Astrophysics Preprint.

Wambsganss, J. & Paczyński, B. 1994, Astron. J., **108**, 1156.

Weymann, R. J., Latham, D., Angel, J. R. P., Green, R. F., Liebert, J. W., Turnshek, D. A., Turnshek, D. E. & Tyson, J. A. 1980, Nature, **285**, 641.

Young, P., Gunn, J. E., Kristian, J., Oke, J. B., & Westphal, J. A. 1980, Ap. J., **241**, 507.

Zwicky, F. 1937, Phys. Rev., **51**, 288.

2. High Angular Resolution Radio Observations

2.1 Two Element Interferometry

Figure 2.1 displays the geometry of a distance source S with a size A_S and two radio antennas located at p_1 and p_2 receiving signals from this source. The coordinate system is chosen so that the plane formed by (x,y) is parallel to the plane tangent to the celestial sphere in the direction of the source S. Let $\mathbf{r_1}$, $\mathbf{r_2}$, and $\mathbf{r_s}$ be the distances to the antennas and the source, respectively. Then, the distances from the source to the antennas are $\mathbf{R_1} = \mathbf{r_s} - \mathbf{r_1}$ and $\mathbf{R_2} = \mathbf{r_s} - \mathbf{r_2}$. At any instant, each source element dS at position $\mathbf{r_s}$ produces an electric field $E_1(r_s, t)$ at p_1 and $E_2(r_s, t)$ at p_2, so the total electric field received by these two antennas due to the whole source is

$$e_{p1}(t) = \int_{A_S} E_1(\mathbf{r_s}, t)dxdy; \quad e_{p2}(t) = \int_{A_S} E_2(\mathbf{r_s}, t)dxdy. \qquad (2-1).$$

We can relate E_1 and E_2 with the complex electric field amplitude in the source $\mathcal{E}(r_s, t)$ by

$$E_j(\mathbf{r_s}, t) = \mathcal{E}(r_s, t - \frac{R_i}{c})\Big[\frac{e^{-2\pi i\nu(t-R_i/c)}}{R_j}\Big] \qquad (2-2)$$

where $j = 1$ or 2. The coherence function of the electric field at positions $\mathbf{r_1}$ and $\mathbf{r_2}$ is

$$\begin{aligned}\Gamma(\mathbf{r_1}, \mathbf{r_2}, \tau) &= \frac{1}{2T}\int_{-T}^{T} e_{p1}(t)e^*_{p2}(t-\tau)dt \qquad (2-3)\\ &= \langle e_{p1}(t)e^*_{p2}(t-\tau)\rangle.\end{aligned}$$

The two antennas form a two-element interferometer when the signals received by them are correlated. The correlation is done by multiplying the signals received by the two antennas and averaging the product over a finite time period. The output of the correlation is normally referred to as the *visibility*, denoted as $V(\mathbf{r_1}, \mathbf{r_2})$. Compared

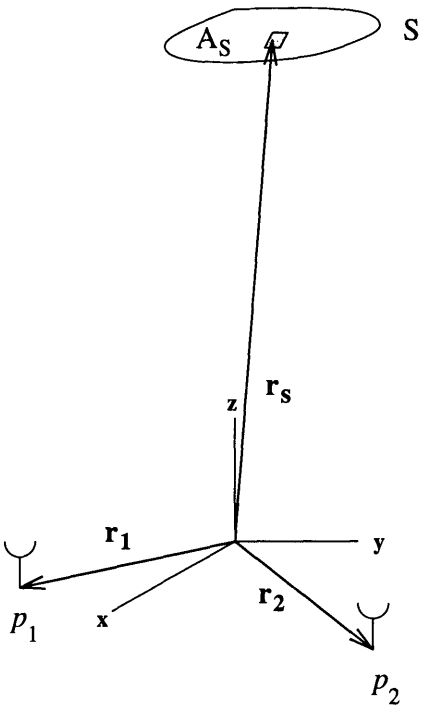


Figure 2.1 – A simple schematic showing the locations of the source and the two antennas. The coordinate is chosen such that the plan formed by (x, y) is parallel to the plane tangent to the celestial sphere in the direction of the source S.

with the coherence function described above, the visibility is the same as the mutual coherence function $\Gamma(\mathrm{r}_1, \mathrm{r}_1, \tau = 0)$, so

$$\mathrm{V}(\mathbf{r_1}, \mathbf{r_2}) = \langle e_{p1}(t) e^*_{p2}(t) \rangle. \qquad (2-4)$$

Assuming that the source is incoherent (i.e., the radiation from any source element dS_m is statistically independent of that from any other element dS_n), then

$$\begin{aligned} \mathrm{V}(\mathbf{r_1}, \mathbf{r_2}) &= \int_{A_S} \langle E(r_s, t - \frac{R_1}{c}) E^*(r_s, t - \frac{R_2}{c}) \rangle A_S \frac{e^{2\pi i \nu (R_1 - R_2)/c}}{R_1 R_2} dx dy \\ &= \int_{A_S} \langle E(r_s, t) E^*(r_s, t - \frac{R_2 - R_1}{c}) \rangle A_S \frac{e^{2\pi i \nu (R_1 - R_2)/c}}{R_1 R_2} dx dy. \end{aligned}$$

If the quantity $(R_1 - R_2)/c$ is much smaller than the reciprocal receiver bandwidth of the telescope, we can approximate $\langle E(r_s, t) E^*(r_s, t - (R_2 - R_1)c) \rangle$ by $\langle E(r_s, t) E^*(r_s, t) \rangle$,

so

$$\begin{aligned} \mathrm{V}(\mathbf{r_1}, \mathbf{r_2}) &= \int_{A_S} \langle E(r_s,t) E^*(r_s,t) \rangle A_S \frac{e^{2\pi i \nu (R_1 - R_2)/c}}{R_1 R_2} dx dy \\ &= \int_{A_S} B(x,y) \frac{e^{2\pi i \nu (R_1 - R_2)/c}}{R_1 R_2} dx dy, \end{aligned}$$

where $B(x,y)$ is the brightness distribution of the source. Assuming that the source is far away from the antenna, (i.e., $\mathbf{r_s} \ll \mathbf{r_1}$ or $\mathbf{r_2}$), then

$$\begin{aligned} R_i &= |\mathbf{r_s} - \mathbf{r_i}| \\ &\simeq r_s [1 + \frac{1}{2}\left(\frac{r_i}{r_s}\right)^2 - \mathbf{r_i} \cdot \hat{\mathbf{s}}], \end{aligned}$$

where $\hat{\mathbf{s}} = \mathbf{r_s}/r_s$, and i is either 1 or 2. Then,

$$R_1 - R_2 \simeq r_s \left[\frac{r_1^2 - r_2^2}{2 r_s^2} - \frac{(\mathbf{r_1} - \mathbf{r_2}) \cdot \hat{\mathbf{s}}}{r_s} \right].$$

If we define $\delta = (r_1^2 - r_2^2)/2R$ and $\mathbf{r_1} - \mathbf{r_2} = \mathbf{b}$ (the baseline separation), then

$$V(\mathrm{r_1}, \mathrm{r_2}) = e^{2\pi i \nu \delta / c} \int B(x,y) \frac{e^{2\pi i \nu \mathbf{b} \cdot \hat{s}/c}}{r_s^2} dx dy. \qquad (2-5)$$

The wavefront of the source reaches one antenna at a time $\mathbf{b} \cdot \hat{\mathbf{s}}/c$ later than the other, so this term is general referred to as the *geometrical delay*, τ_g. In terms of τ_g, the visibility is

$$V(\mathrm{r_1}, \mathrm{r_2}) = e^{2\pi i \nu \delta / c} \int B(x,y) \frac{e^{2\pi i \nu \tau_g}}{r_s^2} dx dy. \qquad (2-6)$$

For a small source, we can define the coordinate system such that $\hat{\mathbf{s}} = \hat{\mathbf{s}}_\mathbf{o} + \sigma$, where $\hat{\mathbf{s}}_\mathbf{o}$ is a unit vector with coordinates (0, 0, 1), and σ is a vector with small magnitude. By definition, $1 = |\hat{\mathbf{s}}^2| = |\hat{\mathbf{s}}_\mathbf{o}^2| = |s_o^2| + 2\hat{\mathbf{s}}_\mathbf{o} \cdot \sigma + \sigma \cdot \sigma \simeq 1 + 2\hat{\mathbf{s}}_\mathbf{o} \cdot \sigma$, so $\hat{\mathbf{s}_\mathbf{o}}$ and σ are nearly perpendicular. Hence, we can write σ as $\sigma = (l, m, 0)$, where l and m are both small. We can also express $\mathbf{b}$ as $\mathbf{b} = \lambda(u, v, w)$ (in this coordinate system). Together, $\mathbf{b} \cdot \hat{\mathbf{s}} = ul + vm + w$, so

$$V(\mathbf{r_1}, \mathbf{r_2}) = e^{2\pi i \nu \delta / c} \int B(x,y) \frac{e^{2\pi i (ul + vm + w)}}{r_s^2} dx\ dy.$$

Since $dx\ dy/r_s^2 = dl\ dm$,

$$V(u,v,w) = e^{2\pi i(\nu\delta/c+w)} \int B(l,m) e^{2\pi i(ul+vm)} dl\ dm.$$

For $r_s \ll |\mathrm{b}|/\lambda$, $\delta \simeq 0$, so we can neglect the phase term due to δ and approximate $V(u,v,w)$ as

$$V(u,v,w) = e^{2\pi iw} \int B(l,m) e^{2\pi i(ul+vm)} dl\ dm. \qquad (2-7)$$

We can redefine the quantity $V' = e^{2\pi iw} V$‡, so that

$$V'(u,v) = \int B(l,m) e^{-2\pi i(ul+vm)} dl\ dm. \qquad (2-8)$$

In summary, the brightness distribution of the sky, $B(l,m)$, and the mutual coherence function relative to the phase tracking center, $V'(u,v)$, form a Fourier transformation pair. For an unresolved point source with a known flux f_S located in the direction of $\hat{\mathbf{s}}$, the modified visibility of this source is simply $V'(u,v) = f_S$ if the phase tracking center is defined to be at the position of the source (i.e., $\hat{\mathbf{s}} = \hat{\mathbf{s}}_\mathbf{o}$).

Since every receiver has a finite bandwidth, $\Delta\nu$, the output of the correlation is an averaged quantity over the finite bandwidth. This causes the measured visibility function, $V^m(u,v)$, to differ from the true visibility, $V(u,v)$, by

$$\begin{aligned} V^m(u,v) &= \int_{\nu_o-\Delta\nu/2}^{\nu_o+\Delta\nu/2} V(u,v) d\nu \qquad (2-9) \\ &= \int_{\nu_o-\Delta\nu/2}^{\nu_o+\Delta\nu/2} \int B(x,y) \frac{e^{-2\pi i\nu\tau_g}}{R^2} dx\ dy \\ &= \Delta\nu \frac{\sin \pi\Delta\nu\tau_g}{\pi\Delta\nu\tau_g} V(u,v). \end{aligned}$$

In short, when $\tau_g \neq 1$, the finite bandwidth of the receiver has an effect of reducing the visibility amplitude. To void degrading the signal by less than 1%, the band width of the receiver must be limited to $\Delta\nu\tau_g \ll 0.018$.

‡ $V'(u,v)$ is the mutual coherence function relative to the direction $\hat{\mathbf{s}}_\mathbf{o}$, which is usually called the phase tracking center.

2.2 Observations

MG1131 was observed extensively with the Very Large Array telescope (VLA)* (see Table 2.1). The instrument consists of 27 steerable antennas, each 25 meters in diameters, arranged in a three-armed array (Thompson et al. 1980). Each pair of antennas in the VLA, separated by a distances **b**, forms a two-element interferometer and measures points in the visibility function $V'(u,v)$. The brightness distribution of the sky is then reconstructed by taking the Fourier transform of the measured visibility (see Eqns 2-5 and 2-6). The angular resolution of the telescope is roughly λ/b_{max}, with λ being the observing wavelength and b_{max} being the largest antenna separation in the array. The mobility of the VLA antennas allows them to be placed in different array configurations, providing various angular resolution options for radio astronomers. At the present time, there are four array configurations, labeled A, B, C, and D. A, being the largest array, has an angular resolution typically in the range of $0\farcs1 \sim 1''$, and D, the smallest array, has an angular resolution in the range of $3'' \sim 30''$. The system is extremely sensitive. With 1 hour of integration time, the thermal noise of the system can be as low as $28, 20, 72,$ and 125μJy† for 5, 8, 15, and 22 GHz observations respectively (Perley 1992).

The observations were carried out over a time period from September 1987 through March 1994 in the A, B, and C configurations at frequencies 5, 8, 15, and 22 GHz. For all observations, the observing bandwidth was 100 MHz, split into two adjacent 50 MHz bands (known as IF1 and IF2). Both senses of circular polarization (RCP and LCP) were correlated, so the four Stoke's parameters (I, Q, U, and V) can be obtained. We took the 15 and 22 GHz observations to study the nature of the compact

* The VLA is part of the National Radio Astronomy Observatory, which is operated by Associated Universities, Inc., under cooperative agreement with the National Science Foundation.

† 1 Jy $= 10^{-26}$ W Hz^{-1} m^{-2}

Table 2.1: List of Observations Used

Date	ν_{IF1} (GHz)	ν_{IF2} (GHz)	array config.	VLA Beam[1]	measurements[2]
20 Sep 87	14.915	14.965	A	$0''.12 \times 0''.11@ + 15°$	V
	4.835	4.885	A	$0''.33 \times 0''.32@ - 13°$	α, LM, RM
02 Nov 88	14.915	14.965	A	$0''.14 \times 0''.12@ - 3°$	V
	8.415	8.465	A	$0''.21 \times 0''.19@ - 20°$	V, RM
	4.835	4.885	A	$0''.33 \times 0''.33$	RM
15 Dec 88	14.915	14.965	A	$0''.12 \times 0''.12$	V
	8.415	8.465	A	$0''.20 \times 0''.19@ - 40°$	V
27 Jan 89	8.415	8.465	A	$0''.21 \times 0''.20@ - 21°$	V
26 Jun 91	8.415	8.465	A	$0''.19 \times 0''.19$	V
26 Jul 91	8.415	8.465	A	$0''.19 \times 0''.18@ - 25°$	α, LM, RM
17 Oct 91	8.415	8.465	A/B	$0''.58 \times 0''.24@ + 89°$	V
19 Nov 91	8.415	8.465	B	$0''.63 \times 0''.59@ - 35°$	V
10 Dec 91	8.415	8.465	B	$0''.59 \times 0''.42@ + 11°$	V
20 Dec 91	14.915	14.965	B	$0''.38 \times 0''.39@ - 44°$	α
10 Jan 92	8.415	8.465	B/C	$0''.62 \times 0''.47@ + 3°$	V
24 Oct 92	8.415	8.465	A	$0''.23 \times 0''.21@ + 29°$	V
05 Nov 92	22.435	22.485	A	$0''.09 \times 0''.09$	α, LM
05 Nov 92	14.915	14.965	A	$0''.12 \times 0''.12$	α, LM, V
06 Nov 92	8.415	8.465	A	$0''.20 \times 0''.20$	V
10 Dec 92	8.415	8.465	A	$0''.21 \times 0''.20@ + 5°$	V
13 Feb 93	8.415	8.465	A/B	$0''.66 \times 0''.24@ + 88°$	V
22 Mar 93	8.415	8.465	B	$0''.66 \times 0''.57@ - 2°$	V
10 Mar 94	8.415	8.465	A	$0''.25 \times 0''.22@ - 77°$	RM
	4.960	5.020	A	$0''.38 \times 0''.35@ - 70°$	RM
	4.535	4.485	A	$0''.42 \times 0''.38@ - 70°$	RM

1 FWHM of the major and minor axes of the synthesized beam and its position angle measured north to east.

2 V=variability, LM= lens modeling, RM=rotation measure,α=spectral index

components, the 8 GHz observations in both A and B configurations to investigate the variability of the system, and the 5 GHz observations to obtain the structural information of the rotation measure in the lens. At each frequency, the dataset with the best quality was chosen for detailed lens modeling. Table 2.1 summarizes the date, central frequencies of both IF's, array configuration, the full width half maximum (FWHM) of the elliptical Gaussian synthesized beam, and the scientific purpose of each observation.

2.3 Calibration and Mapping

In principle, the observed visibility function $\tilde{V}_{ij}(t)$ is different from the true visibility $V'_{ij}(t)$ by (neglecting the effects due to the finite bandwidth)

$$\tilde{V}_{ij}(t) = g_{ij} V'_{ij} + n_{ij},$$

where g_{ij} is the baseline-based complex gain of the instrument, n_{ij} is the noise on each baseline, and the subscript ij denotes the baseline formed by the ith and the jth telescopes. Hence, one can obtain V'_{ij} from the observation only if g_{ij} is calibrated. When we observe an unresolved source with flux f_S, $V'_{ij} = f_S$ (see discussion in Section 2.1). In this case, if the noise is small, the amplitude of the gain, $|g_{ij}|$, can be determined from the observed visibility, $\tilde{V}_{ij}$, by

$$|g_{ij}| \simeq \frac{|\tilde{V}_{ij}|}{f_S},$$

and the phase of the the complex gain is simply the phase of $\tilde{V}_{ij}$. In other words, one can use the observations of an unresolved source with known flux to calibrate the complex gain, g_{ij} (see Perley, Schwab, & Bridle 1988; Thompson, Moran, & Swanson 1986). In principle, the amplitude and the phase of the gain can vary as a function of time, so it is common to observe the calibrator (i.e. the unresolved source)

several times throughout the entire observation to track the amplitude and phase variations. The g_{ij}s in our observations were calibrated using this strategy. The flux densities of all our observations were calibrated to the flux of 3C286 estimated using the scale determined by Perley (the VLA Calibration Manual 1990 Edition; Baars et al., 1988). The complex antenna gains were calculated from short observations of 1055+018 for observations in 1987 and 1988, and 1148-001 for the rest. Both are VLA calibrators near MG1131 (see VLA Calibrator List). The assumed positions for 1055+018 and 1148-001 (J2000 RA DEC coordinate system) were $10^h\ 58^m\ 29^s.605$ $+01°\ 33'\ 58''.81$ and $11^h\ 50^m\ 43^s.8709$ $-00°\ 23'\ 54''.210$ respectively. The uncertainties in these positions were on the order of $0''.01$ (see the VLA calibrator manual).

The polarization state of the emission can be determined from the Stoke's parameters, I, Q, U, and V; these parameters can be obtained by correlating the RCP (and also LCP) signals received by one antenna with the RCP and LCP signals received by the other. Let $V(R_1, R_2)$, $V(R_1, L_2)$, $V(L_1, R_2)$, and $V(L_1, L_2)$ by the visibility formed by correlating the RCP and LCP signals received by the two antennas. For a perfect system,

$$V(\mathrm{R}_1, \mathrm{R}_2) = I + V.$$

$$V(\mathrm{L}_1, \mathrm{L}_2) = I - V,$$

$$V(\mathrm{R}_1, \mathrm{L}_2) = (Q + iU)e^{-2i\chi_p},$$

and

$$V(\mathrm{R}_1, \mathrm{R}_2) = (U - iU)e^{2i\chi_p},$$

where χ_p is the parallactic angle of the observation. In reality, a small amount of right-circular polarization can show up in the left hand channel (and vice versa) due to the imperfection of the antenna-feeds. Because of this leakage, the voltages measured from the RCP and LCP channels (denoted as v_R and v_L) is different from

the true electric fields (denoted as E_R and E_L) by

$$v_R = E_R e^{-i\chi_p} + D_R E_L e^{i\chi_P}$$

and

$$v_L = E_L e^{i\chi_p} + D_L E_R e^{-i\chi_P},$$

where D_R and D_L represent the leakages to the RCP and LCP channels, respectively. Because of the leakages,

$$V(\mathrm{R}_1, \mathrm{R}_2) = I + V,$$

$$V(\mathrm{L}_1, \mathrm{L}_2) = I - V,$$

$$V(\mathrm{R}_1, \mathrm{L}_2) = e^{-2i\chi_p}(Q + iU) + (D_{\mathrm{R1}} D^*_{\mathrm{L2}})I,$$

and

$$V(\mathrm{R}_1, \mathrm{R}_2) = e^{2i\chi_p}(Q - iU) + (D_{\mathrm{L1}} D^*_{\mathrm{R2}})I.$$

Therefore, the leakage terms must be calibrated so that the Stoke's parameters can be computed correctly.

The simplest way to calibrate the leakage terms, D_R and D_L, is to use the observations of an unpolarized source. In that case, Q and U are both zero, so D_R and D_L can be easily determined from $V(R_1, L_2)$ and $V(L1, R_2)$. However, for the observation spans enough parallactic angle coverage, it is possible to solve the polarization property of the source and leakage terms simultaneously from the observation. Since the polarization properties of the phase calibrators we used were unknown, we had to calibrate D_R and D_L using the latter method. This restricted us to perform polarization calibration only on the datasets with sufficient parallactic angle coverage. For these datasets, the systematic phase differences between the right and left hand circular polarizations were determined from the observations of 3C286 assuming that the electric field of its linearly polarized emission points at 33° measured from north to east.

The calibrated uv data points were Fourier transformed to construct a brightness distribution map. The straight Fourier transform of the calibrated uv data points produces a *dirty map* (I_D) which is different from the true sky brightness distribution (I) because of the number of measured visibility points in each observation is limited. In general, $I_D = I \otimes B_D + N$. B_D is the Fourier transform of the sampling function $\tilde{B}_D$ in the u-v plane ($\tilde{B}_D(\mathbf{u}) = \sum_j w_j \delta(\mathbf{u} - \mathbf{u_j})$, where j's are the sampled visibilities and w_j is the statistical weight at that jth visibility point) and is often called the dirty beam. N is the thermal noise in the map. Because $\tilde{B}_D$ is irregular and discrete, the structure of B_D consists of slowly decreasing side lobe patterns which make the interpretation of the dirty maps very difficult. Thus, the map can only be useful after deconvolving B_D from I_D. At the present time, several non-linear image restoration methods – such as CLEAN (Högbom 1974; Clark 1980), Maximum Entropy Methods (MEM) (Cornwell 1984; Cornwell & Evans 1985; and Narayan & Nityananda 1986), and the Non-Negative Least Square Methods (NNLS) (Briggs et al. 1994) – are employed to deconvolve B_D from I_D. Among these methods, CLEAN is the most commonly used algorithm. A detailed description of CLEAN is presented in Chapter 4. All data presented in this thesis were processed using CLEAN.

To improve the quality of the maps, the 5 and 8 GHz datasets were subjected to two or three iterations of self-calibration, solving only the phases of the gain (see Pearson & Readhead 1984 for a review). Since a confusing source at $11^h31^m58^s.207 + 04°54'44''.89$ (J2000 RA DEC coordinate system) was found in all datasets, this source was included in the self-calibration model. No self-calibration was applied to the 15 and 22 GHz datasets because the signal-to-noise ratios in the these maps were insufficient.

For datasets that allowed polarization calibration, maps of Q and U stokes parameters [$Q = P\cos(2\chi)$ and $U = P\sin(2\chi)$, where P is the total polarization intensity and χ is the position angle of the electric field vector] were computed. At each pixel,

the intensity (P), position angle of the electric field vector (χ), and percentage polarization (π) of the linearly polarized emission were calculated from the Q and U maps. Two types of depolarization can occur due to instrumental effects. The first is due to averaging over the bandpass and is usually called bandpass depolarization. This can happen when the orientation of the electric field vector varies as a function of frequency. In that case, averaging the polarization over the entire bandpass would result in a reduction of the polarization strength. The second type of depolarization is due to averaging over the finite angular resolution of the telescope and is often referred to as beam depolarization. This effect takes place when the χ varies on a spatial scale that is much smaller than the beam width. As a result of spatial averaging, the observed polarization strength would appear weaker. Using a top hat model for the beam (cf. Dreher, Carilli & Perley 1987), the degree of beam depolarization can be estimated by $\sin\eta/\eta$, where η is the change in χ over the beam width. The amount of depolarization is less than 5% if the change in χ is less than 28° over the beam. To eliminate the bias due to beam depolarization, we compared values of χ in the region surrounding each pixel and searched for those in which the value of χ varied by 30° or more over the beam. These pixels were excluded from the polarization data.

All datasets were calibrated and mapped using the software in the National Radio Astronomy Observatory's Astronomical Image Processing System (AIPS) package.

References

Baar, J. W. M., Genzel, R., Pauliny-Toth, I. I. K. & Witzel, A. 1977, Astron. Astrophys., **61**, 99.

Briggs, D. S., Davis, R. J., Conway, J. E., & Walker, R. C. 1994, VLBA Memo, # 697.

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

Clark, B. G. 1989, *in Synthesis Imaging in Radio Astronomy*, volume 6, 1.

Cornwell, T. J. 1984, in Indirect Imaging, eds. J. A. Roberts, Cambridge University Press (Cambridge, England), pp. 291.

Cornwell, T. J. & Evans, K. F. 1985, Astron. Astrophys., **143**, 77.

Högbom, J. 1974, Ap. J. S., **15**, 417.

Dreher, J. W., Carilli, C. L. & Perley, R. A. 1987, Ap. J., **316**, 611.

Narayan, R. & Nityananda, R. 1986, Ann. Rev. Astron. Astrophys., **24**, 127.

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

Perley, R. A., Schwab, F. R. & Bridle, A. H. 1989 (eds.), *Synthesis Imaging in Radio Astronomy*, volume 6

Perley, R. A. 1992, *Very Large Array Observational Status Summary*, Charlottesviell: National Radio Astronomy Observatory.

Thompson, A. R., Clark, B. G., Wade, C. M. & Napier, P. J. 1980, Ap. J. S., **44**, 151.

Thompson, A. R., Moran, J. M. & Swenson, G. W., Jr. 1986, *"Interferometry and Synthesis in Radio Astronomy"*, John Wiley & Sons, New York.

3. Radio Morphology and Nature of MG1131+0456

Much of the material presented in this chapter is taken from the paper, *"Multi-frequency Radio Images of the Einstein Ring Gravitational Lens MG1131+0456,"* by Chen & Hewitt (1993).

3.1 Maps of Total Intensity and Polarized Emission

The contour levels in Figures 3.1-3.4 represent the total intensity of MG1131 at 5, 8, 15, and 22 GHz. The synthesized beam at each frequency is shown in the lower left corner. All maps display features recognized by Hewitt et al. (1988). At 5 GHz, the source consists of a radio ring, two unresolved components (A and B), and one low surface brightness extended component in the southwest direction (C). At 8 GHz, in addition to A, B, C, and the ring, a new component (D) at the center of the ring is detected. Although there is a hint of D in the 5 GHz map, it is not cleanly resolved from the ring. At 15 GHz, the surface brightnesses of C and D are below the detection limit ($\sim$ 120 μJy/beam), so only A, B, and the ring are present in the map. The angular resolution at 15 GHz (0.″12) is fine enough to reveal substructures in the compact components and the ring; A and B are resolved into a possible core-jet structure, and the ring, marginally resolved, has a gap in the northwest quadrant. At 22 GHz, only the compact components are detectable with a significant signal-to-noise ratio ($> 10\sigma$, $1\sigma = 200\mu$Jy). The positions of the compact components in the 22 GHz map coincide with the 15 GHz peak positions of A and B.

Given the complex and extensive data reduction procedure described in Chapter 2, one might suspect that the presence of D at 8 GHz is produced solely by an artifact of the procedure. We conducted experiments to test for this possibility. Throughout the whole process, the CLEAN algorithm is the most likely source of artifacts. As a test, we created a map in which CLEAN was only permitted to

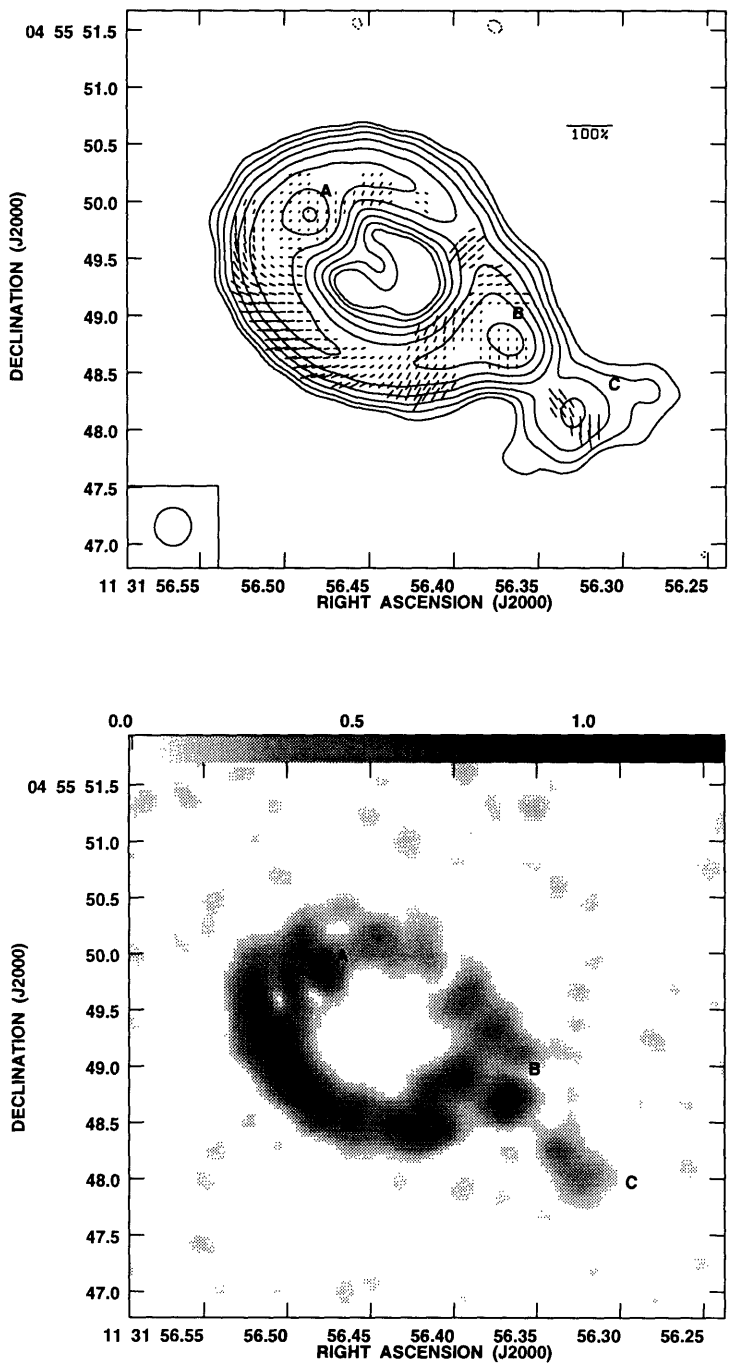


Figure 3.1 – 5 GHz images. (a)Total intensity image. The contour levels are -1, 1, 2, 4, 8, 16, 32, 64, and 95% of the peak brightness of the map (19.9 mJy/beam). The vectors superimposed on the contour levels represent the electric field vectors of the polarized emission. The lengths of the vectors are proportional to the fractional polarization according to the scale shown in the upper right hand corner. The half-power contour of the beam is plotted in the lower left corner. (b) Grey-scale image of the polarized intensity ($\sqrt{Q^2 + U^2}$). The bar on the top represents the polarization intensity from 0.0 to 1.5 mJy.

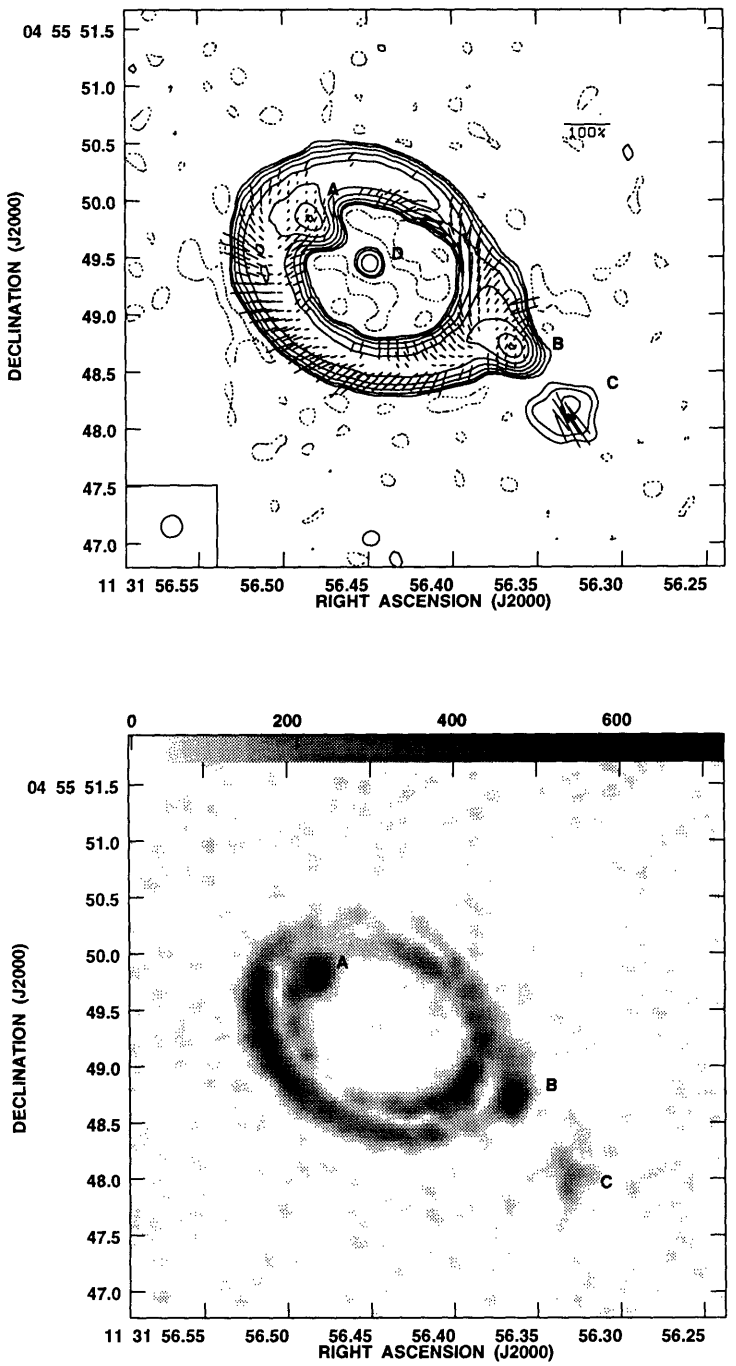


Figure 3.2 – 8 GHz images. (a)Total intensity image. The contour levels are -1, 1, 2, 4, 8, 16, 32, 64, and 95% of the peak brightness of the map (6.7 mJy/beam). The vectors superimposed on the contour levels represent the electric field vectors of the polarized emission. The lengths of the vectors are proportional to the fractional polarization according to the scale shown in the upper right hand corner. The half-power contour of the beam is plotted in the lower left corner. (b) Grey-scale image of the polarized intensity ($\sqrt{Q^2 + U^2}$). The bar on the top represents the polarization intensity from 0 to 700 μJy.

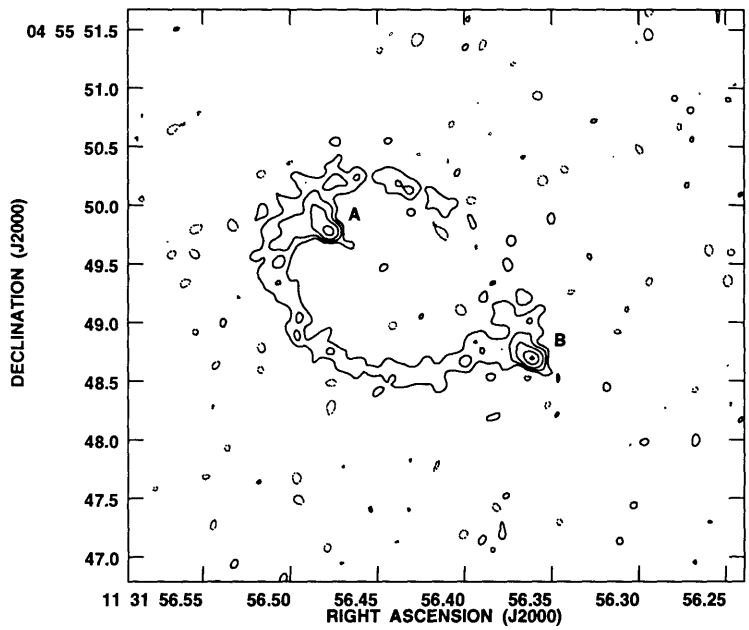


Figure 3.3 – 15 GHz total intensity image. The contour levels are -8, 8, 16, 32, 64, and 95% of the peak (4.2 mJy/beam).

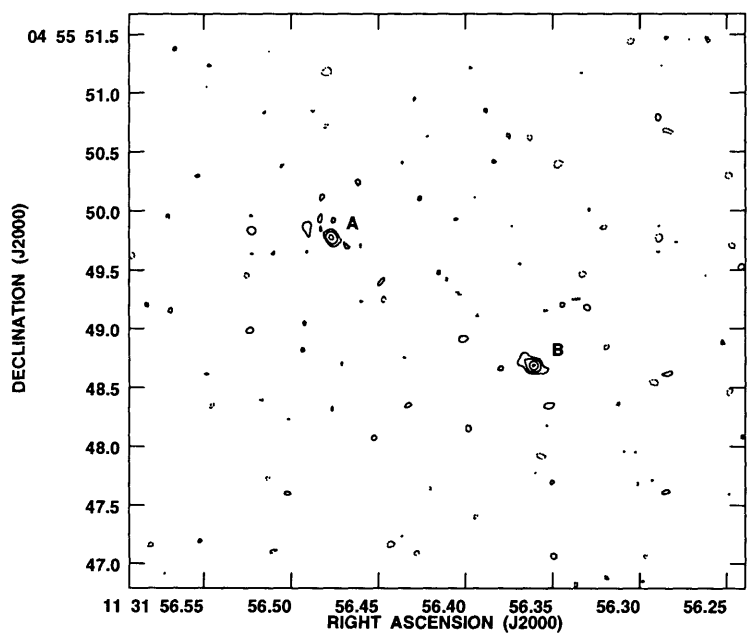


Figure 3.4 – 22 GHz total intensity image. The contour levels are -16, 16, 32, 64, and 95% of the peak (3.8 mJy/beam).

search for clean components in regions *excluding* the center of the ring. Any real source inside of the ring not CLEANed should be accompanied by a dirty beam pattern; we looked for the existence of this beam pattern around the central component. In our test map, the faint central object remained and was surrounded by a correct dirty beam pattern. Thus, it is unlikely that D is an artifact. To test the accuracies of the flux density detected in D, we conducted two other experiments. In the first experiment, we segmented the 8 GHz July 91 data into several one-hour-long datasets and reduced these test datasets with the same procedures outlined in Section 2.2. The purpose of this experiment was to test the consistency of the flux density detected in D with changing uv coverage and noise. In all test datasets, the central object was detected, and the flux densities (S_D) obtained from the test datasets varied slightly from one to another, yielding a mean and a standard deviation of $\overline{S_D} = 0.32 \pm 0.06$ mJy. For the second experiment, we subtracted D from the maps, added a 1 mJy artificial point source at the center of the ring to each of the one-hour-long test datasets, and reduced these datasets in the same manner. The purpose of this experiment was to test for any bias in the flux density detected in D due to the complex data reduction procedures. The 1 mJy artificial point source was recovered in all datasets, indicating that no significant flux bias was introduced. Thus, we are confident that the flux density found for D was not corrupted by our complex data reduction routines.

Linearly polarized emission in MG1131 was detected at only 5 and 8 GHz. The vectors in Figures 3.1(a) and 3.2(a) represent the fractional strength and position angle of electric field vector (χ) of the polarized emission, and the grey scales in Figures 3.1(b) and 3.2(b) represent the total polarization intensity (P). Table 3.1 lists the fractional polarization strength for all components at 5 and 8 GHz. At both frequencies, the fractional strength and the position angle of the electric field vector detected in A closely resemble the ones detected in B. These features further strengthen the

interpretation of their being images of the same object. No polarization in D was detected at either frequency. The thermal noise in the 8 GHz total polarization intensity map was $\sim 12\mu$Jy, and a signal $> 60\mu$Jy would clearly be detected. Thus, the polarization strength of D must be $< 60\mu$Jy. Using this number and the flux density of D (0.32 ± 0.06 mJy), we estimated the upper limit on the percent polarization in D to be 19%. The total polarization intensity at 8 GHz had an unusual feature [see Figure 3.2(b)]. A gap, corresponding to the unpolarized emission of the ring, was present and separated the ring into an inner ring and outer ring.

Table 3.1: Percent Polarization

ν(GHz)	A(%)	B(%)	C(%)	D(%)	ring(%)
4.860	5	5	30	not detected	30
8.440	10	10	60	not detected	30

3.2 What is Component D?

The roles of components A, B, C, and the ring seem clear. A and B are images of a compact source (possibly the core in a typical extragalactic radio source); the ring is the lensed image of an extended source (possibly one of the radio lobes in a typical double-lobe extragalactic radio source) located directly behind the lens; and C is the image of another extended source (possibly the other lobe) located further away from the optical axis. What, then, is D? In the framework of gravitational lensing, there are three possible hypotheses for explaining the existence of D. It can be (1) the second image of C, (2) the radio emission from the lens galaxy, or (3) the odd image of the system. We consider these hypotheses separately.

3.2.1 Is D the Lensed image of C?

In a singular isothermal sphere lens model, every source in the multiply imaged region produces an image pair with equal image separation (see discussion in Section 1.3). In MG1131, the image separations between C and D (denoted by CD) and between A and B (denoted by AB) are similar. The image separation between A and B estimated from the 22 GHz map is 2.″045, and the separation between C and D estimated from the 8 GHz map is 2.″15. If the potential of the lens galaxy can be described by a singular isothermal sphere model, we would expect AB and CD to have the same angular separation. Thus, the close agreement between the two separations seems to support the hypothesis that C and D are images of the same source. However, the polarization properties in C and D do not support this hypothesis.

Gravitational lensing should not alter the intrinsic properties of the source because a gravitational field acts on radiation like a non-dispersive medium. This implies that the direction and degree of polarization of different images arising from the same source can only differ when their respective ray bundles have been affected differently by the intervening matter. At 8 GHz, C is 60% polarized whereas the polarization strength of D is $< 19\%$ (see Table 3.1). Thus, to support the hypothesis that C and D are images of the same source, an external mechanism must be present to depolarize D at least from 60% to 19%. What kind of physical processes can produce such a large depolarization?

When a magnetized plasma is present between the emission region and the observer, the linearly polarized radiation from the source of emission will undergo Faraday rotation by an amount

$$\Delta\chi = RM\ \lambda^2; \quad RM = 8.1 \times 10^5 \int nB_{\|} dL, \qquad (3-1)$$

where n is the electron density in the plasma (in cm^{-3}), $B_{\|}$ is the magnetic field projected onto the line of sight (in Gauss), and L is the path on which the light

travels (in pc). RM is commonly called the rotation measure. Foreground Faraday rotation can result in some degree of depolarization if the telescope has a large enough bandpass, leading to a bandpass depolarization, or a large enough beam size, leading to a beam depolarization. As discussed in Chapter 2, the bandpass depolarization is caused by a rapid rotation of the plane of the linearly polarized emission as a function of frequency. Given RM, the degree of bandpass depolarization across the observing wavelength band of $\lambda_1 < \lambda < \lambda_2$ can be calculated by (cf. Hennessy, Owen & Eilek 1987),

$$\frac{\pi}{\pi_0} = \left| \frac{\int_{\lambda_1}^{\lambda_2} \exp[i(RM)\lambda^2]\, d\lambda}{\int_{\lambda_1}^{\lambda_2} d\lambda} \right|, \qquad (3-2)$$

where π and π_o represent the observed and intrinsic fractional polarization strength. Using the observing bandpass at 8 GHz (3.534 to 3.586 cm), we calculated π/π_o as a function of RM (see the solid curve in Figure 3.5). To depolarize D from 60% to <19%, a $RM > 150{,}000$ rad/m^2 is needed (see the dashed line in the same figure). This is two orders of magnitude larger than the rotation measure seen in Cygnus A (Dreher et al. 1987), the largest one ever detected. Thus, it is extremely unlikely that the light ray in component D would have encountered an external screen with rotation measure > 150,000 rad/m^2.

Similar to the discussion in Chapter 2, the variations of RM on spatial scales smaller than the angular resolution of the telescope can result in some depolarization. Using a top hat model for the beam (cf. Dreher et al. 1987) and assuming that the intrinsic position angle of the polarized emission does not vary across the beam, we can estimate the degree of depolarization due to the variation of RM across the beam by

$$\frac{\pi(\lambda)}{\pi_o} = \frac{|\sin(\Delta RM\ \lambda^2)|}{\Delta RM\ \lambda^2}, \qquad (3-3)$$

where ΔRM is the change of rotation measure across the beam. Figure 3.6 shows π/π_o as a function of ΔRM. To depolarize D from 60% to 19% by beam depolarization

alone, a $\Delta RM \simeq 1,800$ rad/m^2 over 0.″20 (the beam size at 8 GHz) is needed. Assuming that the lens galaxy is at redshift of $z_l = 0.5$, the angular resolution at 8 GHz (0.″02) corresponds to a linear scale of ~ 0.5 Kpc. The value of RM and the variation of RM over the 0.″20 scale required to depolarize D would still be large if we account for both effects simultaneously. If we assume no absorption occurs between the emitting region and the observer, the bandpass and beam depolarizations are the most likely sources of external depolarization at radio frequencies. Since it is extremely unlikely that the absence of polarization of D is due to the bandpass

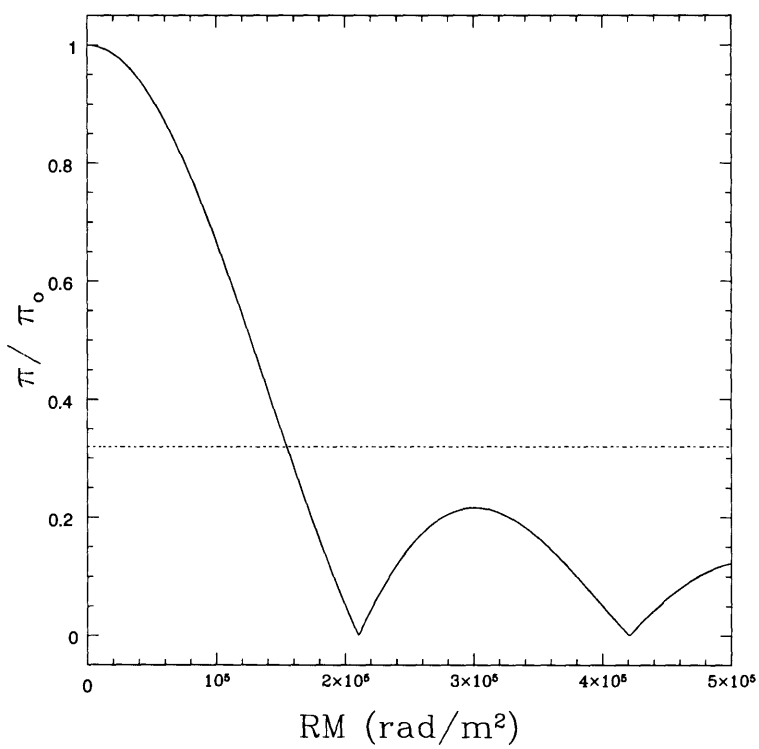


Figure 3.5 – The bandpass depolarization function (π/π_o) as a function of RM. This function is calculated using the observing bandpass at 8 GHz (3.534 to 3.586 cm). The dotted line represents the degree of depolarization needed to make C and D be the images of the same source, giving a lower limit to the RM needed.

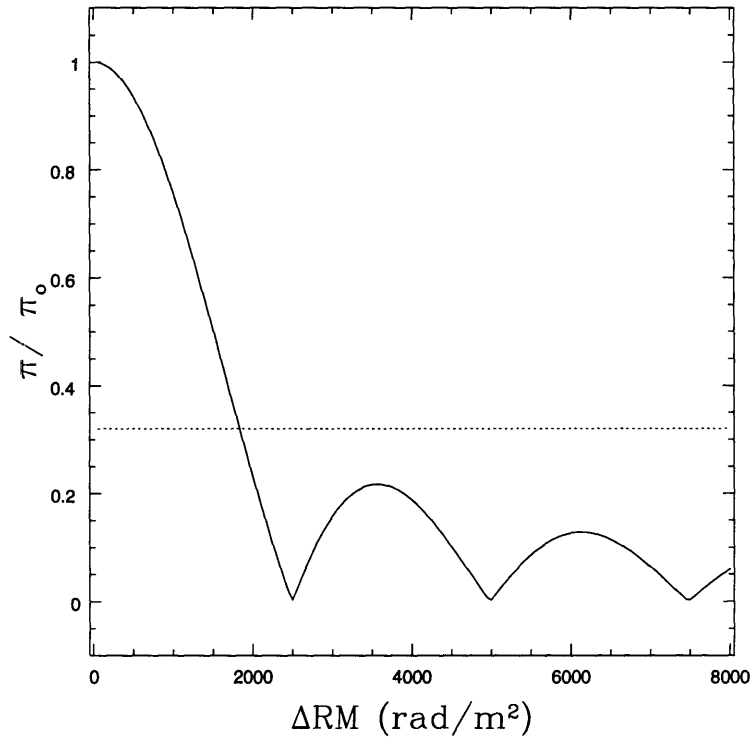


Figure 3.6 – The beam depolarization function (π/π_o) as a function of ΔRM for the top hat model. The dotted line represents the degree of depolarization required to make C and D be the images of the same source.

depolarization, the beam depolarization, or the combination of both, the polarization differences between C and D cannot be produced by external mechanisms alone. Therefore, we conclude that C and D cannot be images of the same source.

3.2.2 Is D the Radio Image of the Lens Galaxy?

It is clear from the symmetry of the system that the position of the lens must be near the center of the ring. The component D, being at the center of the ring, undoubtedly fulfills the positional requirement for being the radio image of the lens galaxy. However, the lens galaxy is not the only object expected to be found at that position; the odd image of the system is also expected to appear there. A

more qualitative and interesting way to investigate the validity of this hypothesis is to compare the flux density detected in D with the expected flux density at 8 GHz from regular galaxies (by that, we mean non-active ellipticals and spirals). If the flux density in D is much larger than that is expected from regular galaxies, it is almost certain that the hypothesis is false unless the lens galaxy of MG1131 is peculiar. If the flux density of D falls in the range expected from regular galaxies, we can estimate the likelihood of D being the radio image of the lens galaxy by considering the flux density in D and the radio luminosity function of regular galaxies. Since elliptical galaxies are most likely to produce strong lensing (Fukugita & Turner 1991) and to have detectable radio emission, examining only the elliptical galaxies should give us a conservative estimate. We, therefore, consider only elliptical galaxies in the following discussion.

The radio luminosity function, $\varphi(L)dL$, describes the number of galaxies per unit volume of space with radio luminosity between L and $L + dL$. Since the radio luminosities of elliptical galaxies correlates strongly with their optical luminosity (i.e. an elliptical galaxy with a brighter optical magnitude is more likely to be a strong radio emitter), it is useful to introduce a bivariate form of the radio luminosity function, $\varphi(M, L)$, which depends on both the optical magnitude M and the radio power L. In short, $\varphi(M, L)$, a subset of $\varphi(L)$, is the radio luminosity function determined only from the sample of galaxies with optical magnitudes between M and $M + dM$.

In 1977, Auriemma et al. combined sources from four surveys and compiled a local bivariate radio luminosity function of elliptical and S0 galaxies at 1.415 GHz. All sources in the sample have $z < 0.1$. The optical observations of most of these sources were taken in the p band, so the bivariate radio luminosity function was determined using their absolute photomagnitudes (M_p). A bivariate radio luminosity function was obtained in each of the four intervals of optical magnitudes M_p: (-22, -21), (-21, -20), (-20, -19), and (-19, -18). The number density of elliptical and S0 galaxies in

Table 3.2: Number density of ellipticals and S0

M_p† range	N* (10^{-4} Mpc^{-3})
$-22 \sim -21$	0.139
$-21 \sim -20$	1.333
$-20 \sim -19$	5.180
$-19 \sim -18$	17.78?

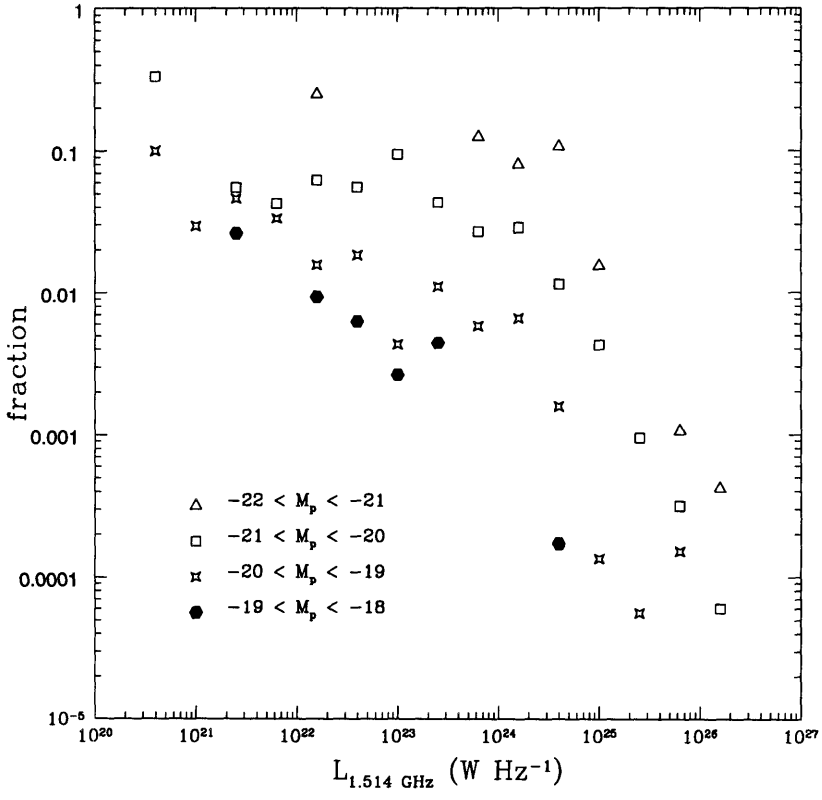


Figure 3.7 – The fractions of galaxies having radio luminosity $L_{1.415\ GHz}$ in each of the optical magnitude intervals are shown. The lower left corner shows the symbols representing data in each of the optical intervals.

† Absolute photomagnitudes
* Number density.

each optical magnitude interval is listed in Table 3.2. The fractions of galaxies having radio luminosity $L_{1.4\ GHz}$ in each of the optical magnitude intervals are plotted in Figure 3.7.

To utilize Auriemma et al.'s bivariate radio luminosity function, we need to estimate the radio luminosity of the lensing galaxy at 1.415 GHz and its absolute optical photomagnitude. At 8 GHz, the integrated flux density of D is 0.32 ± 0.06 mJy. Assuming the spectral index of D to be ~ 0.7, we estimated the flux density of D at 1.415 GHz to be 1.12 mJy. If we further assume that the flux density detected in D comes solely from the lens galaxy and that the redshift of the lens is in a reasonable range ($0.4 < z_l < 0.9$), the luminosity density of the lens galaxy at 1.4 GHz should be

$$L_{1.4\ GHz} \sim \frac{1.9 \times 10^{23} \sim 1.1 \times 10^{24}}{h^2}\ \mathrm{W\ Hz^{-1}},$$

where h is the dimensionless parameter in the Hubble constant. The range of redshifts for the lensing galaxy is chosen according to the probability distribution found by Kochanek (1992). The luminosity calculated above is consistent with the luminosity normally seen in big ellipticals (see Figure 3.7); thus, the flux density detected in D is reasonable for large ellipticals.

The optical counterpart of MG1131 has an apparent visual magnitude of $m_v = 22.66 \pm 0.25$ (Hammer et al. 1991), so we assume that as the visual magnitude of the lens galaxy. Of course, it is not clear that all the optical emission is due solely to the lensing galaxy. If the lensed images are also visible at the optical frequencies, the intrinsic optical magnitude of the lens galaxy would be fainter than what is assumed. Nevertheless, since a fainter galaxy is less likely to be a radio source, our assumption is conservative. For a galaxy with an apparent magnitude m_i (where i = B, V, R, etc. filter band) at redshift of z_i, its absolute magnitude M_i is,

$$M_i = m_i - 5.0 \log \frac{R}{10} - K_i(z_i),$$

where R is the luminosity distance in pc and K_i is the K correction* at i band. Using the K_v correction calculated from the spectral energy distribution tabulated by Coleman, Wu & Weedman (1990), we estimated the limit on M_v of the lensing galaxy in the above redshift range ($0.4 < z_l < 0.9$) to be

$$M_v > -22.78.$$

Using

$$M_p = M_v + 1.25$$

(see Colla et al. 1975), we estimated the absolute photomagnitude of the lens galaxy to be $M_p > -21.53$.

From Auriemma et al.'s work, the number density of ellipticals with $-22 < M_p < -21$ and radio luminosity $L_{1.415\ GHz} \approx 1.9 \times 10^{23} \sim 1.1 \times 10^{24}/h^2$ W Hz^{-1} is 1.12×10^{-6}Mpc^{-3}. The total number density of elliptical galaxies with $-22 < M_p < -21$ is 1.39×10^{-5} Mpc^{-3} (see Table 3.2). This indicates that only $\sim 8\%$ of the elliptical galaxies with $-22 < M_p < -21$ would have radio luminosities as large as what was observed. For fainter galaxies (i.e. $M_p > -21$), the fraction would be even smaller. Thus, the probability that D is entirely due to the radio emission of the lensing galaxy is moderately small ($\leq 8\%$). Of course, we cannot rule out the possibility that the radio emission from the lens galaxy contributes only a fraction of the flux density detected in D. In that case, the probability would be larger than 8%.

3.2.3 Is D the Odd Image of the System?

The third hypothesis — is D the odd image of the system? — can only be addressed with the help of the lens models; thus, we will discuss the possibility of this hypothesis in Chapter 5.

* The K correction accounts for the effects of cosmological redshifts.

3.3 Spectral Index Measurements of All Components

Synchrotron radiation, produced by electrons spiraling around a magnetic field, is believed to be the main source of the non-thermal radio emission detected in most extragalactic radio sources. In a constant magnetic field, an ensemble of electrons with a power law energy distribution,

$$n(E)dE = n_o E^{-p} dE \qquad (3-4)$$

produces synchrotron radiation with a spectrum (in the optically thin region)

$$S(\nu) \propto \nu^{-\alpha}; \quad p = 1 - 2\alpha, \qquad (3-5)$$

where $n(E)$ is the number density of electrons, E is the energy, S is the flux density of the source, and ν is the frequency. The spectral indices of different components in a typical extragalactic radio source are different. The typical spectral index of the emission from a radio lobe is in the range of $1.3 > \alpha > 0.5$. By comparison, the spectrum of a core is much flatter, with a spectral index typically < 0.5.

The spectral indices of A and B were determined from the flux densities measured from the 15 and 22 GHz Nov 92 datasets. We chose these two datasets to determine α_A and α_B for two reasons: (1) the flux contamination from the ring was small in

Table 3.3: Flux densities of A and B

ν(GHz)	A(mJy)	B(mJy)
14.940	3.18 ± 0.19	3.90 ± 0.23
22.460	2.60 ± 0.30	3.60 ± 0.40

these two maps (see Figures 3.3 and 3.4), so the flux densities of A and B can be more accurately determined, and (2) these two observations were taken on the same date, so effects due to flux variability were eliminated. The integrated flux densities of A and B at 15 and 22 GHz obtained from these two maps (see Table 3.3) yield a $\alpha_A = 0.47 \pm 0.28$ and $\alpha_B = 0.17 \pm 0.28$. Inspection of the 15 and 22 GHz maps shows that A and B are resolved, and they appear more extended in the 15 GHz image; thus, the flux densities of the compact components were most likely to be overestimated (rather than underestimated). This error would be larger at 15 GHz than at 22 GHz. In addition, the atmospheric phase variation is greater at 22 GHz, which systematically lowers the flux density. Both biases caused the apparent spectra to appear steeper than the actual ones, so the spectral indices calculated here should be interpreted as upper limits to the true values. Assuming that A and B are images of the same source and that the spectral variability of this source is small, we can estimate the spectral index of the source, α_s, using the values of α_A and α_B. The mean value of α_A and α_B is 0.32 ± 0.20, so the spectral index of the source is likely be on this scale.

The spectral indices of A and B are consistent with those normally found in the cores of classical core-lobe radio sources. If A and B were indeed images of the core of the source, they would likely be variable at some level. Indeed, the peak flux densities of A and B measured from the 15 GHz Nov 92 dataset were different from the ones observed by Hewitt et al. (1988), indicating a small degree variability in the source. A monitoring project for the purpose of investigating the level of variability in A and B was carried out (Hewitt, Chen, & Messier 1995). The result of this project will be discussed in Chapter 7.

The spectral index distribution across the ring was also computed. The 15 GHz Dec 90 B configuration dataset and the 5 GHz Sep 87 A configuration dataset had similar uv coverage, so they were chosen for the determination of the spectral index

distribution across the ring. The 5 GHz map was convolved to the exact resolution of the 15 GHz map and shifted so that the peak brightnesses between the two maps coincided before the spectral index was computed. Regions near the compact components were masked to eliminate any effects due to variability. The resulting spectral index map of the ring is presented in Figure 3.8. It is clear that the spectral indices of the ring span a wide range, from 0.5 to 1.2. These values are typical for radio lobes. Thus, our finding further supports the interpretation of the ring being a lensed image of a radio lobe.

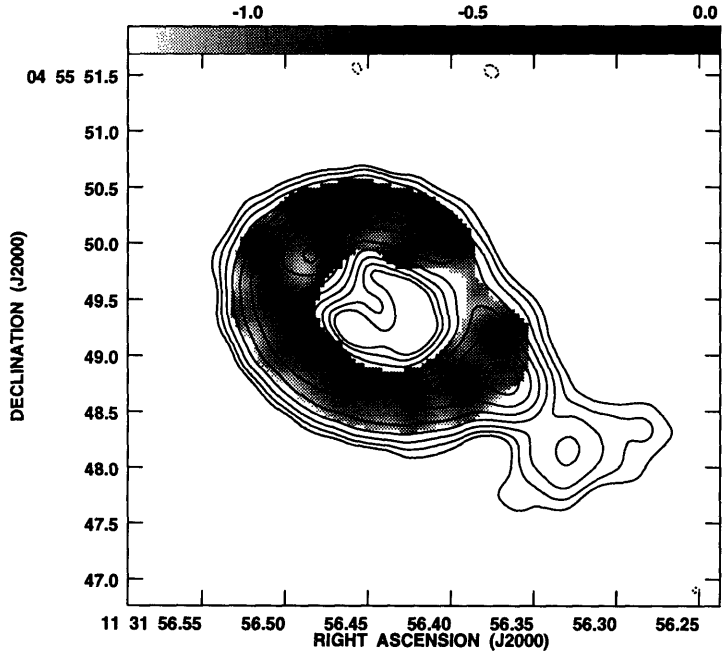


Figure 3.8 – The grey-scale map of the spectral index distribution across the ring. Regions near the compact components are masked to eliminate any effects due to variability.

3.4 Depolarization Structure

For optically thin synchrotron emission, the radiation is mostly linearly polarized. Assuming a uniform magnetic field without the presence of a thermal plasma, the degree of polarization in synchrotron radiation is independent of frequency, and it is given by (see Vershuur & Kellermann 1988)

$$\pi = \frac{p+1}{p+\frac{7}{3}},$$

where p is the power law index of the energy distribution of the electron [see Eqn (3-4)]. However, if a thermal plasma is present in the emission region, or between that region and the observer, the position angle of the linearly polarized emission will undergo Faraday rotation [see Eqn (3-1)], resulting in a frequency dependent reduction of polarization strength. In the case where the synchrotron emitting region and the plasma causing the Faraday rotation coexist, the plane of linearly polarized emission from the rear of the emitting region can be rotated by a large angle before emerging from the front end of the region. The contribution from both sides results in a wavelength dependent polarization strength. This situation was considered by Burn (1966). He showed that for a uniformly emitting slab of depth L containing a uniform magnetic field and a thermal plasma of electron density n, the observed degree of polarization $\pi(\lambda)$ is

$$\pi(\lambda) = \frac{\pi_o \sin(RM\lambda^2)}{RM\lambda^2},$$

where π_o is the fractional polarization strength when the thermal plasma is absent. In the case where the plasma is present between the emitting region and the observer, beam depolarization or bandpass depolarization can occur (see discussion in Section 3.2). Both effects are wavelength dependent. We note that the effects of bandpass or beam depolarization can be reduced when using a telescope with better resolution and smaller bandwidth.

The depolarization ratio measures the relative fractional polarization strength between two frequencies and can reveal useful information on the structure of the source. The depolarization ratio is defined as

$$DP(\lambda_1, \lambda_2) = \frac{\pi(\lambda_1)}{\pi(\lambda_2)},$$

where λ is the wavelength and $\lambda_1 > \lambda_2$. A smaller DP value means that the fractional polarization strength is weaker at the longer wavelength. We computed the DP structure in MG1131 using the polarized emission detected at 5 and 8 GHz. The error in DP, assuming quadrature summation of noise (cf. Dreher, Carilli & Perley 1987), is given by

$$\frac{\delta D}{D} = \left[(\frac{\delta \pi_1}{\pi_1})^2 + (\frac{\delta \pi_2}{\pi_2})^2\right]^{1/2}.$$

We calculated DP only in regions of the maps with signal-to-noise ratios greater than 8, so the error in DP is at most $\sim$18%. The calculation was performed on two pairs of observations. We first carried out the calculation using the 5 and 8 GHz Nov 88 datasets to avoid the effects due to possible variability. In addition, since the quality of the 5 GHz Sep 87 and 8 GHz Jul 91 datasets was much higher, we also performed the same DP calculations on these two datasets. We found no differences between these two sets of calculations; therefore, we restrict further discussion to the results obtained from the Sep 87-Jul 91 pair.

Figure 3.9 shows the DP map obtained from the Sep 87-Jul 91 pair. Inspection of the DP map indicates that MG1131 exhibits considerable structure in the depolarization ratio, ranging from 0 to higher than unity. Components A and B have a DP of 0.5; component C has a $DP < 0.4$; and there is a very large variation of DP around the ring. Recent radio observations of high-power, double lobe, single jet extragalactic radio sources showed that these sources display an asymmetry in their depolarization structure. The lobe on the jet side has fractional polarization strength that is independent of frequency whereas the lobe in the counter jet side has

polarization strength decreasing with increasing wavelength (Laing 1988; Garrington et al 1988, 1991). This feature is named the Laing-Garrington Effect after their discoveries. The observed depolarization asymmetry is consistent with the finding of the x-ray halo surrounding large ellipticals (see Fabbiano 1989 for a review) and the

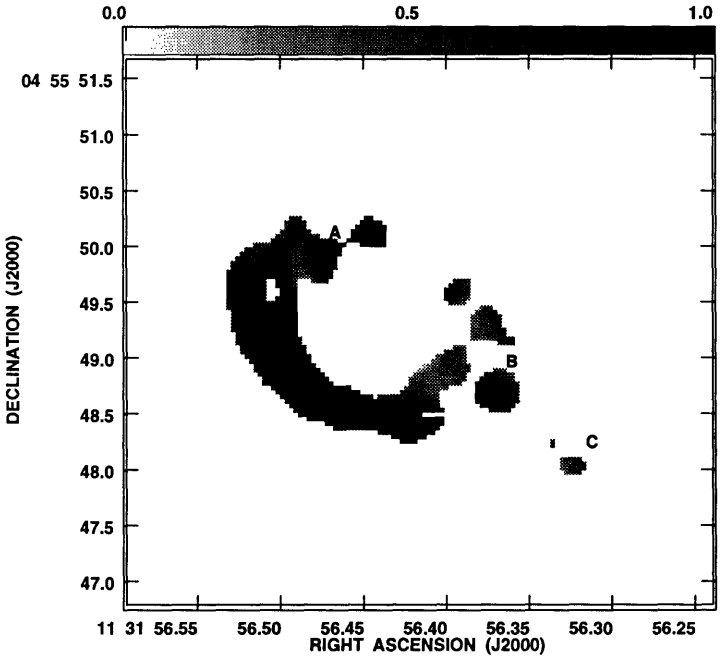


Figure 3.9 – The grey-scale image of the depolarization structure.

interpretation that the one-sidedness of jets in extragalactic objects is due to relativistic Doppler beaming. In this scenario, the lobe on the jet side, being closer to the observer, is seen through less of a Faraday screen and hence has a lower Faraday dispersion (Burn 1966), leading to smaller degree of depolarization. On the other hand, the lobe on the counter jet side, being away from the observer, is seen through more of the Faraday screen, so the Faraday dispersion it experiences is larger, resulting into a larger degree of depolarization. If we assume that the environment around the

background source of MG1131 is similar to what is typically found in high powered extragalactic radio sources, a simple model of the source structure can be constructed to explain the depolarization feature detected in MG1131. Component C, experiencing the most depolarization, is the farthest from the observer and thus is the image of the radio lobe associated with the counter-jet; those parts of the ring experiencing a depolarization ratio ranging from 0.5 to 1 are images of the collimated material (a jet?) between the core and the lobe on the jet side. Based on these assumptions, we are able to reconstruct the variation in the depolarization detected in MG1131 from a simple model consisting of a lens with a generic elliptical potential and a radio source with a core, a one-sided jet, and two radio lobes. The structure of the background source in the model and the reconstructed depolarization variation in MG1131 are displayed in Figure 3.10. Our simple model suggests that MG1131 is intrinsically a "normal" high-power extragalactic radio source (see also Kochanek et al. 1989). We will compare this simple model with the results obtained from the detailed lens modeling (see Chapter 5).

References

Auriemma, C., Perola, G. C., Ekers, R., Fanti, R., Lari, C., Jaffe, W. J. & Ulrich, M. H. 1977, Astron. Astroph., **57**, 41.

Burn, B. F. 1966, M.N.R.A.S., **133**, 67.

Coleman, G. D., Wu, C. & Weedman, D. W., 1980, Ap. J., **43**, 393.

Chen, G. H. & Hewitt, J. N. 1993, Astron. J., **106**, 1719.

Dreher, J. W., Carilli, C. L. & Perley, R. A. 1987, Ap. J., **316**, 611.

Fabbiano, G. 1989, Ann. Rev. Astron. Astrophys., **27**, 87.

Fukugita, M. and Turner, E. L., 1991, M.N.R.A.S., **253**, 99.

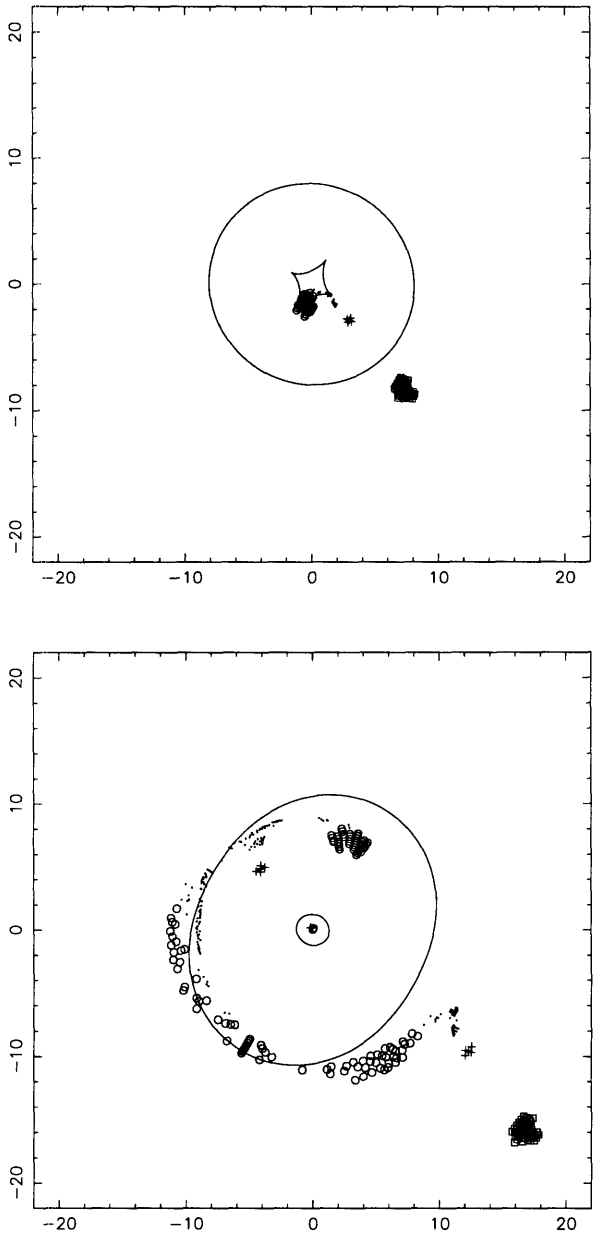


Figure 3.10 – A model that reproduces the depolarization in MG1131. The lens is assumed to have an elliptical potential, and the source consists of a core (represented by crosses), a one-sided jet (represented by dots), a lobe on the counter-jet side (represented by squares), and a lobe on the jet side (represented by circles). (a) The image of the source. The location of the source with respect to the caustics (solid lines) in the source plane are shown. (b) The lensed image. The solid lines are the critical curves. The image of the core (the crosses) correspond to component A and B; the image of the jet (the dots) correspond to the parts of the ring with depolarization ratio ranging from 0.5 to 1; the image of the lobe on the jet side (the circles) correspond to the parts of the ring with unity depolarization ratio; and the image of the lobe on the counter-jet side (the squares) corresponds to component C.

Garrington, S T., Leahy, J. P., Conway, R. G., and Laing, R. A. 1988, Nature, **331**, 147.

Garrington, S. T., Conway, R. G., & Leahy, J. R. 1991, M.N.R.A.S., **250**, 171.

Hammer, F., Le Fèvre, O., Angonin, M. C., Meylan, G., Smette, A. & Surdej, J. 1991, Astron. Astrophys., **250**, L5.

Hennessy, G. S., Owen, F. N. & Eilek, J. A. 1989, Ap. J., **347**, 144.

Hewitt, J. N., Turner, E. L., Schneider, D. P., Burke, B. F., Langston, G. I. & Lawrence, C. R. 1988, Nature, **333**, 537.

Hewitt, J. N., Chen, G. H., Messier, M. D. 1995, Astron. J., in press.

Vershuur, & Kellerman, K. I. 1988 (eds.), *Galactic and Extragalactic Radio Astronomy.*

Kochanek, C. S., Blandford, R. D., Lawrence, C. R. & Narayan, R. 1989, M.N.R.A.S., **238**, 43.

Kochanek, C. S., 1992, Ap. J., 384, 1.

Laing, R. A. 1988, Nature, **331**, 149.

4. Methodology of the Lens Modeling

Much of the material presented in this chapter and the next chapter is taken from the paper, *"The Mass Distribution of the Lens Galaxy in MG1131+0456,"* by Chen, Kochanek, & Hewitt (1995).

The complex and extended structure in MG1131 makes the system one of the best candidates (among the known gravitational lenses) for studying the mass distribution of the lens galaxy. Since the lens galaxy is most likely to be at a large redshift (though the redshift of the lens galaxy has not yet been determined), the results of our lens modeling can give important insight into the mass distribution of elliptical galaxies.

Our previous analysis showed that different components in MG1131 have different spectral indices (see discussion in Section 3.3), so maps at different frequencies emphasize different features in the system. To maximize the number of available constraints, the lens modeling was conducted using maps at several frequencies. We selected the observation with the best quality at each available frequency (5, 8, 15, and 22 GHz) for modeling. Table 4.1 summarizes the characteristics of each selected map (see Figures 3.1 - 3.4 for the morphology). We chose not to model fully the 22 GHz data because only the compact components (A and B) were detected with high signal-to-noise ratio at this frequency.

In this chapter, we concentrate on the description of methods employed in the modeling. The results of the modeling will be presented in the next chapter.

4.1 Multipole Expansion of the Potential

The general solution to the Poisson equation of the rescaled two-dimensional potential [see Eqn (1-5)] is

$$\phi(\mathbf{x}) = \frac{2}{\Sigma_c} \int \Sigma(\mathbf{x}') \ln \mid \mathbf{x} - \mathbf{x}' \mid d^2\mathbf{x}'. \qquad (4-1)$$

Table 4.1: Summary of the Selected Dataset

frequency GHz	Synthesized Beam[a]	noise μJy/pixel	peak mJy	pixel size ″/pixel	A[b] (±0.″10)	B[b] (±0.″10)	D[b]
5	0.″33 × 0.″32 @ − 13°	60	19.9	0.04	(−0.″44, 0.″43)	(1.″20, −0.″76)	confused
8	0.″19 × 0.″19 @ − 24°	35	6.7	0.04	(−0.″49, 0.″42)	(1.″21, −0.″75)	(0.″00, 0.″00)
15	0.″12 × 0.″12 @ − 31°	130	4.2	0.02	(−0.″43, 0.″32)	(1.″31, −0.″76)	not detected
22	0.″098 × 0.″086@ − 50°	200	3.8	0.02	(−0.″43, 0.″32)	(1.″31, −0.″76)	not detected

a FWHM of the major and minor axes of the synthesized beam, and its position angle measure north to east.

b The positions are given relative to $11^h31^m56.448^s + 04°55'49.''45$, the position of D at 8 GHz. The 5 and 8 GHz positions are determined using the Q maps, and the 15 and 22 GHz positions are determined using the I maps.

Like electrostatic (or magnetostatic) potentials, the gravitational potential can be expanded into multipoles. The term $\ln \mid \mathbf{x} - \mathbf{x}' \mid$ can be expanded into

$$\ln \mid \mathbf{x} - \mathbf{x}' \mid = \frac{1}{2}\ln(r_>) - \sum_{m+1}^{\infty} \frac{1}{m}\left(\frac{r_<}{r_>}\right)^m \cos\left[m(\theta - \theta')\right], \qquad (4-2)$$

where $r_>$ ($r_<$) is the larger (smaller) of $\mid \mathbf{x} \mid$ and $\mid \mathbf{x}' \mid$ and θ (θ') is the position angle of $\mathbf{x}$ ($\mathbf{x}'$). Combining the expression in Eqn (4-1) and (4-2), we can expand the two-dimensional potential into multipoles as

$$\phi(\mathbf{x}) = \phi_o(\mathbf{x}) + \sum_{m=1}^{\infty} \phi_m(\mathbf{x}), \qquad (4-3)$$

where

$$\phi_o(\mathbf{x}) = \frac{2\pi}{\Sigma_c}\int_0^{\infty} \Sigma(\mathrm{r}')\ln(r_>)\ r'dr', \qquad (4-4)$$

and

$$\phi_m(\mathbf{x}) = \frac{2}{m\Sigma_c}\int_0^{\infty}\int_0^{2\pi} \Sigma(\mathbf{x}')\left(\frac{r_<}{r_>}\right)^m \cos\left[m(\varphi - \varphi')\right] r'dr'd\varphi'. \qquad (4-5)$$

The *monopole term* is the term associated with $m = 0$; the *dipole term* is associated with $m = 1$; the *quadrupole term* is associated with $m = 2$; and so forth. If we define the origin of the coordinates to be at the center of the mass of the system (i.e. if we define $\int\int \Sigma(\mathbf{x}')\mathrm{x}'d\mathbf{A} = 0$ and $\int\int \Sigma(\mathbf{x}')\mathrm{y}'d\mathbf{A} = 0$), the dipole term of the potential outside the mass distribution vanishes (see discussion in §2.4 of Binney & Tremaine 1987). Thus, the effect of the dipole term can be eliminated by a coordinate transformation if the region of interest is outside of the matter distribution.

The surface mass distribution, $\Sigma(\mathbf{x})$, can also be expanded into multipoles as

$$\Sigma(\mathbf{x}) = \Sigma_o(\mathrm{r}) + \sum_{m=1}^{\infty}\left[\Sigma_{cm}(\mathrm{r})\cos m\theta + \Sigma_{sm}(\mathrm{r})\sin m\theta\right]. \qquad (4-6)$$

Then,

$$\begin{aligned}\phi_o(\mathbf{x}) &= \frac{2\pi}{\Sigma_0}\int_0^{\infty} \Sigma(\mathrm{r}')\ln(r_>)\ r'dr' \qquad (4-7)\\ &= \frac{2\pi}{\Sigma_c}\Big[\ln(r)\int_0^{r} \Sigma_0(\mathrm{r}')\mathrm{r}'d\mathrm{r}' + \int_r^{\infty} \Sigma_0(\mathrm{r}')\ln(\mathrm{r}')\mathrm{r}'d\mathrm{r}'\Big],\end{aligned}$$

and

$$\phi_m(\mathbf{x}) = C_{cm}(\mathrm{r})\cos m\theta + C_{sm}(\mathrm{r})\sin(m\theta), \qquad (4-8)$$

where

$$\begin{aligned} C_{\substack{cm\\sm}}(\mathrm{r}) &= \frac{-2\pi}{\Sigma_c m}\int_0^r \Sigma_{\substack{cm\\sm}}\mathrm{r}'\left(\frac{r_<}{r_>}\right)^m \mathrm{r}'d\mathrm{r}' \qquad (4-9)\\ &= \frac{-2\pi}{\Sigma_c m}\left[\mathrm{r}^{-m}\int_0^r \Sigma_{\substack{cm\\sm}}(\mathrm{r}')\mathrm{r}'^{m+1}d\mathrm{r}' + \mathrm{r}^m\int_r^\infty \Sigma_{\substack{cm\\sm}}(\mathrm{r}')\mathrm{r}'^{1-m}d\mathrm{r}'\right]. \end{aligned}$$

For a surface mass distribution that is nearly circularly symmetric, Σ_{cm}(or Σ_{sm}) terms are small, so the series converges quickly.

4.2 Potentials Used for the Modeling

For a lens with a circularly symmetric surface mass distribution, the two dimensional potential of the galaxy is also circularly symmetric since the higher order multipoles in ϕ are zero [see Eqn 4-9]. In that case, the lens equations [Eqn (1-7) and (1-8)] reduce to equations in one-dimension, and the Einstein ring produced by such a lens must be a circular ring. The ring in MG1131 is rather elliptical. The ellipticity (i.e., the definition of the ellipticity is $e_{ring} = 1 - b/a$, where b and a are the semi-minor and semi-major axis of the ellipse) of the ring estimated from the 8 GHz image (see Figure 3.2) is 0.18 ± 0.02. Thus, the geometry of the ring indicates that the gravitational potential responsible for producing MG1131 cannot be circularly symmetric.

If we assume that the surface mass density of the lens can be described by a sum of the two lowest order terms in the expansion (i.e. the monopole and the quadrupole terms), we can write the surface mass distribution as

$$\Sigma(\mathbf{x}) = \Sigma_o(\mathrm{r}) + \Sigma_2(\mathrm{r})\cos 2(\theta - \theta_\gamma),$$

where θ_γ is the position angle of the quadrupole term. If we further assume that the mass responsible for the quadrupole term is well outside the region of interest, then

the two-dimensional potential can be written as

$$\phi(\mathbf{x}) = \phi_0(\mathbf{x}) + \frac{\gamma}{2}\mathrm{r}^2 \cos 2(\theta - \theta_\gamma),$$

where γ is

$$\gamma = \frac{\pi}{\Sigma_c} \int_{r_{min}}^{\infty} \frac{\Sigma_2(\mathrm{r}')}{\mathrm{r}'} d\mathrm{r}'$$

and ϕ_o is the monopole of the potential [see Eqn (4-7) for the full mathematical expression]. The quadrupole term breaks the circular symmetry and introduces "shearing effects" in lensing (see discussion in Section 1.2). Since the mass producing the shear is outside the region of interest, this type of quadrupole term is generally referred to as the external shear term. We assume that the mass distribution of the lens producing MG1131 has such a quadrupole term.

In general, there are two approaches to modeling the monopole term (or any other term). We can (1) parameterize the mass distribution or (2) parameterize the two-dimensional potential and determine the mass distribution from the best fit potential. The advantage of the first approach is that we can directly model the mass distribution. The problem with this approach is the difficulty in obtaining an analytical form for the two-dimensional potential [i.e., the integrals in Eqn (4-7) for most cases do not have a simple closed form], even for a relatively simple surface mass distribution. As a result, the lens equations and the magnifications can only be solved and evaluated numerically, making the lens modeling difficult and time consuming. In comparison, the second method offers an opposite advantage and disadvantage. If we parameterize the potential instead of the mass distribution, the potential is guaranteed to have an analytical expression, so the lens equations can be solved and the magnifications can be evaluated much more easily. The disadvantage of this method is that the mass distribution determined from the best fit model could be non-physical.

We modeled the monopole of the lens galaxy in MG1131 using both approaches.

For the first model, we assumed that the mass distribution of the lens has a constant mass-to-light ratio, so we parameterized the surface mass density Σ_0(r) by a de Vaucouleurs function (de Vaucouleurs 1948). For the second model, we assumed that the monopole of the potential can be described by a central core radius (s) and a power-law slope (α) as $\phi_o = b^{2-\alpha}(r^2 + s^2)^{\alpha/2}/\alpha$; we named this the α model. The properties of each potential are discussed separately in the following two sections.

For the rest of this thesis, the two-dimensional potential and the surface mass distribution of the lens galaxy are expressed in polar coordinates (r, θ) centered at the position of the lens. The angular coordinate θ is measured from west to north, and the coordinates (x, y) represent a right-handed coordinate system with $x = r\cos\theta$ and $y = r\sin\theta$.

4.2.1 de Vaucouleurs Model

The projected surface brightness $I(r)$ of a typical elliptical galaxy can be described by de Vaucouleurs' empirical law

$$I(r) = I(R_e)e^{-7.67[(r/R_e)^{1/4}-1]} \qquad (4-10)$$

(de Vaucouleurs 1948), where R_e is the effective radius enclosing half of the total light. If we assume that the mass-to-light ratio (M/L) in the galaxy is constant and that Σ_o(r) $\propto e^{-7.67[(r/R_e)^{1/4}-1]}$, then the mass interior to a radius r is M($< r$) = $MF(r/R_e)/2$, where M is the total mass of the galaxy, F is the function

$$F(r/R_e) = \Big[\int_0^{r/R_e} uI(u)du\Big]\Big[\int_0^1 uI(u)du\Big], \quad 0 \le F \le 2 \qquad (4-11)$$

(Maoz & Rix 1993), and $I(x)$ is the surface brightness at radius $x = r/R_e$. The deflection equation produced by this monopole in an external shear is

$$\mathrm{u} = \mathrm{x} - bR_eF[r/R_e]\frac{\mathbf{x}}{r^2} - \gamma\begin{pmatrix}\cos 2\theta_\gamma & \sin 2\theta_\gamma \\ \sin 2\theta_\gamma & -\cos 2\theta_\gamma\end{pmatrix}\begin{pmatrix}x \\ y\end{pmatrix}, \qquad (4-12)$$

where the characteristic deflection scale is $b = [D_{ls}/(D_{ol}D_{os})][2GM/(c^2R_e)]$. Expressed in terms of R_e and b, the total mass of the galaxy is

$$\begin{aligned} \mathrm{M} &= \frac{D_{os}D_{ol}}{D_{ls}}\frac{c^2R_eb}{2G} \\ &= 2\pi\Sigma_cR_eb \\ &= 1.5h^{-1}R_eb\Big[\frac{D_{os}D_{ol}}{2r_HD_{ls}}\Big] \times 10^{12}M_\odot, \end{aligned} \tag{4-13}$$

where R_e and b are in arcseconds and $r_H = c/H_o$ is the Hubble radius. The mass interior to a radius r can be expressed in term of the critical surface mass density, R_e, and b as

$$\mathrm{M}(< r) = \pi\Sigma_cR_ebF[\frac{r}{R_e}].$$

There are altogether six free parameters in this model: the scaling factor b, the lens position (x_l, y_l), the strength of the shear γ, the orientation of the shear θ_γ, and the effective radius R_e. For a fixed value of R_e, the scaling factor b is fixed very precisely by the average radius of the ring. For a ring with a small ellipticity, the average radius of the ring $\langle r\rangle$ is approximately $\langle r\rangle \approx \sqrt{bR_eF[\langle r\rangle/R_e]}$. γ is constrained by the ellipticity of the ring, and θ_γ is constrained by the position angle of the ring. The mass enclosed by the average ring radius is $\mathrm{M}(< \langle r\rangle) = \pi\Sigma_cR_ebF[\langle r\rangle/R_e] = \pi\Sigma_c\langle r\rangle^2$. Thus, the mass interior to the average ring radius should be independent of the lens parameters.

4.2.2 α Model

The two-dimensional potential of the α model is parameterized as

$$\phi = \frac{b^{2-\alpha}(r^2+s^2)^{\alpha/2}}{\alpha} + \frac{1}{2}\gamma r^2\cos 2(\theta-\theta_\gamma), \tag{4-14}$$

where b is the tangential critical radius of the lens if $s = 0$, and s is the core radius (Blandford & Kochanek 1987). The exponent α determines how rapidly the density declines with radius. If $\alpha = 1$, the model has the same effects as ones in an isothermal

sphere model (for $r \gg s$, $\phi \propto r$, surface density $\Sigma \propto 1/r$) while if $\alpha = 0$, the potential reduces to a Plummer model (for $r \ll s$, $\phi \propto \ln r$, $\Sigma \propto 1/r^4$). The surface mass density of the α model is

$$\Sigma = \frac{1}{2}\Sigma_c \nabla^2 \phi = \frac{1}{2}\Sigma_c b^{2-\alpha}(r^2+s^2)^{(\alpha-2)/2}\left[\frac{2s^2+\alpha r^2}{r^2+s^2}\right], \qquad (4-15)$$

and the mass inside the projected radius r is

$$\begin{aligned} \mathrm{M}(<r) &= \pi\Sigma_c b^{2-\alpha} r^2 (r^2+s^2)^{(\alpha-2)/2} \qquad (4-16) \\ &= 7.4h^{-1}b^{2-\alpha}r^2(r^2+s^2)^{(\alpha-2)/2}\left[\frac{D_{os}D_{ol}}{2r_H D_{ls}}\right] \times 10^{11} M_\odot \end{aligned}$$

(b and s are in arcseconds).

We can derive several useful scalings of the lensing properties of the model that explain the parameters determined when we fully model the system. The shape and size of the ring are set by the semi-major and semi-minor axes of the tangential critical line

$$r_\pm^2 = b^2(1 \mp \gamma)^{2/(2-\alpha)} - s^2, \qquad (4-17)$$

which is approximately equal to $r_\pm \simeq b[1 \pm (2-\alpha)^{-1} - \beta^2/2]$ if $\gamma \ll 1$ and $\beta = s/b \ll 1$. The average ring radius is $\langle r \rangle = (1/2)(r_+ + r_-) \simeq b(1-\beta^2/2)$ and the average ring ellipticity is $e = 1 - r_-/r_+ \simeq 2\gamma/(2-\alpha)$. If the core radius is small ($\beta \ll 1$) then the average ring radius determines the parameter b, and the ellipticity of the ring determines the dimensionless shear γ and the exponent α. For a circular lens the average density inside the tangential critical lines is equal to the critical density Σ_c, so the mass interior to the ring or the critical radius, $\mathrm{M}(<\langle r \rangle) = \pi \langle r \rangle^2 \Sigma_c$, is independent of the lens parameters.

The value of the core radius is constrained by the flux of the central (or "odd") image and the size of the multiply imaged region at the ring radius. The flux of the central image depends on the magnification at the center of the lens ($r = 0$)

$$M_0 = \left[1 - \left(\frac{b}{s}\right)^{2-\alpha}\right]^{-2}, \qquad (4-18)$$

and $M_0 \simeq \beta^{2(2-\alpha)}$ for $\beta \ll 1$. Since b is largely fixed by the diameter of the ring, the core radius controls the flux of the central image. A larger core radius produces a brighter central image. There is no simple analytic form describing the size of the multiply imaged region, but the trends can be understood from the behavior of the bending angle and the peak deflection of the lens. The monopole deflection of the lens is $b^{2-\alpha} r (r^2 + s^2)^{(\alpha-2)/2}$, and the peak of the deflection is $\propto b^{2-\alpha} s^{\alpha-1}$ at the location $r^2 = s^2/(1-\alpha)$. When $\alpha < 1$, a decreasing core radius s increases the peak deflection and expands the multiply imaged region.

4.2.3 Qualitative Analysis of the α model

The analytical nature of the α model enables us to predict the outcome of the full lens modeling with relatively few constraints. There are altogether seven parameters in the α model: b, x_l, y_l, γ, θ_γ, s, and α. The geometry of the ring fixes the position of (x_l, y_l) reasonably accurately. For models with fixed values of α and s, the relative positions and flux densities of the compact components (A and B) allow us to solve for b, γ, and θ_γ. The requirement that the ellipticity of the tangential critical line must match the measured ellipticity of the ring gives us the last piece of information needed to find a relationship between α and s.

If we assume that the compact components are not variable or that their variability is small, we can ignore the effects due to the time delay (see Section 1.2) and use the flux densities of the compact components measured from the same map for the calculations of b, γ, and θ_γ. We used the positions and fluxes of the compact components derived from the 22 GHz map since they are best isolated from the extended ring at that frequency. Table 4.1 lists the positions of A and B at this frequency. The 22 GHz fluxes of A and B are 2.6 mJy and 3.6 mJy respectively (see Table 3.3). The ellipticity of the ring estimated from the 8 GHz image is $e_{ring} = 0.18 \pm 0.02$.

For fixed values of α and s, we used the relative positions and the flux densities

of A and B to solve for b, γ, and θ_γ and calculated the ring ellipticity corresponding to this set of lens parameters. Figure 4.1 shows contours of the ring ellipticity as a function of α and s. The ellipticity varies monotonically with α and s, and increasing values of the exponent or the core radius lead to smaller ellipticities. Our analytical calculations predict two important results. First, among models which fit the constraints of the compact components, only models with $\alpha < 0.9$ can produce an ellipticity as large as what is observed, so only models with $\alpha < 0.9$ can successfully fit MG1131.

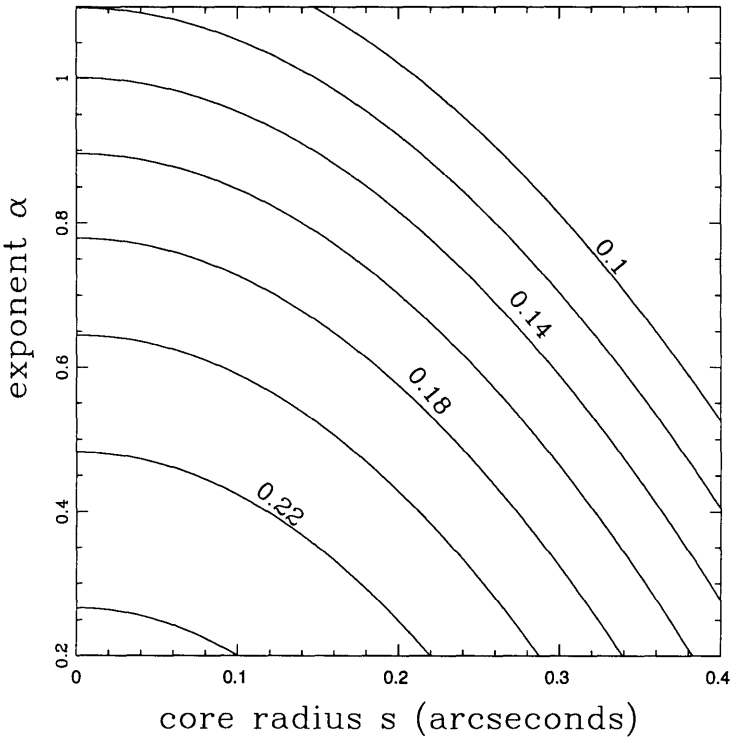


Figure 4.1 – A plot of the predicted ellipticity of the ring as a function of α and s. The contour levels show the predicted ellipticity from 0.1 to 0.24 with an increment of 0.02. The measured ellipticity of the ring is 0.18 ± 0.02.

Second, when $\alpha \leq 0.9$, there is a restricted range of core radii consistent with the constraints for each value of α, so only models with the core radius within this range can fit the system. We recall that the Kochanek et al. (1989) modeled MG1131 using the $\alpha = 1$ model (albeit with a different elliptical structure) and produced rings that

were more circular than those in the data, consistent with this qualitative picture. Our previous calculation assumed that the compact components are not variable. In reality, we detected a small level of variability in both A and B (see discussion in Chapter 7), and we would like to check whether our prediction would be affected if we take this variability into account. We performed the same calculation including the variability we detected in A and B and the time delay estimated from our best lens model. Adding the variability shifts the ellipticity contour lines in Figure 4.1 upward (if the flux density of the compact source is increasing in time) or downward (if the flux density of the compact source is decreasing in time) very slightly. Thus, the variability of A and B has virtually no effect on our prediction.

4.3 A Summary of the LensClean Algorithm

The finite resolution of the telescope causes the observed brightness distribution to differ from the true distribution, and this difference is particularly significant for sources that are extended and complex. The assumption that the observed flux densities in the image represent the true surface brightness can have a fatal consequence in modeling extended lenses. Thus, it is essential to develop a lens modeling method which takes the instrumental effects into account. The newly developed lens inversion algorithm, LensClean (Kochanek & Narayan 1992), is one such algorithm. The basic principle in the LensClean algorithm is to apply the image reconstruction technique used in CLEAN to lens inversion. The technique is rather powerful. It can simultaneously determine the underlying structure of the source and the best set of parameters for the lens. In this section, we present a short summary of the LensClean algorithm. Since the basic philosophy of LensClean is the one in CLEAN, we will first give a summary of CLEAN. Readers should consult the original paper (see Kochanek & Narayan 1992) for a more detailed description of LensClean. We modeled MG1131 using this algorithm.

4.3.1 CLEAN

As discussed in Chapter 2, a straight Fourier transform of the observed visibility function produces a brightness distribution I_D which differs from the true distribution I by $I_D = I \otimes B_D + N$. B_D is the Fourier transform of the sampling function in the uv plane ($\tilde{B}_D(\mathbf{u})$), and N is the noise in the map. The sampling function $\tilde{B}_D(\mathbf{u})$ can be represented as $\tilde{B}_D(\mathbf{u}) = \sum_j w_j \delta(\mathbf{u} - \mathbf{u_j})$, where the j's are the sampled visibilities and w_j is the statistical weight of the visibility point j. Since the sampling is discrete and irregular, the structure of B_D contains large number of sidelobes extending to infinity, which causes the interpretation of the dirty map I_D to be very difficult or sometimes impossible. To make use of the data obtained from the VLA (or any other radio interferometer), it is necessary to develop techniques which deconvolve B_D from I_D. In 1974, Högbom introduced the CLEAN algorithm which provided one solution for deconvolving B_D from I_D (Högbom 1974). The idea is as follows. Suppose that the source of interest is an unresolved point source with a flux density S. The dirty map of this source is $I_D = S \otimes B_D + N = SB_D + N$. Since B_D is completely known — it is just the Fourier transform of the sampling function, the sidelobes of B_D in the dirty map can be completely removed by subtracting a copy of the dirty beam multiplied by S at the position of the point source. The true brightness distribution of the source can be recovered by replacing the original map by one which consists of a point source with a flux S at the position of the peak in I_D. By doing so, we "deconvolve" the dirty beam pattern from the dirty map I_D and "reconstruct" the brightness distribution which closely resembles the sky distribution I. If we assume that any extended source can be represented by a collection of point sources, we can iteratively find all the point sources representing the extended source with the technique described above. We can then construct a modeled image for the sky brightness distribution I_c (called the CLEAN map) by putting all the modeled point

sources at the correct positions with the correct flux densities. These modeled point sources are generally called the CLEAN components. In summary, the algorithm is as follows (cf. Cornwell & Braun 1989):

(1) Find the position and the flux density of the peak in the current dirty map I_D.

(2) Subtract from the dirty image a copy of B_D multiplied by the peak flux density and a factor γ_g at the position of the peak. Add a point source with the same flux density and the same position to the CLEAN map, I_c. The factor γ_g, known as the loop gain (less than unity), is introduced to prevent numerical instabilities.

(3) Return to step (1) and hunt for the next brightest point source unless the residuals in the current map are below or comparable to the estimated noise in the map.

(4) Accumulate all the point sources in I_c and smooth the image by convolving it with a CLEAN beam which is usually an elliptical Gaussian fitted to the central lobe of the dirty beam. Then add in the residuals left in the dirty map.

Although the algorithm was introduced as a heuristic method, Schwarz (1978) was able to find a mathematical basis behind such an image reconstruction method. He found that the flux density and the position which CLEAN picks at each iteration are the ones that minimize the root mean square of the residuals in the dirty map. The mean square residuals after subtracting a component of flux S at location $\mathbf{x_o}$ are

$$\begin{aligned} R^2 &= \int d^2\mathbf{u} \mid \tilde{I}_D - S\tilde{B}_D \exp(i\mathbf{u}\cdot\mathbf{x_o}) \mid^2 \\ &= \int d^2\mathbf{u} \mid \tilde{I}_D \mid^2 - 2SI_D \otimes B_D + S^2 \int d^2\mathbf{u} \mid \tilde{B}_D \mid^2 \end{aligned} \qquad (4-19)$$

(Kochanek & Narayan 1992). Minimizing R^2 with respect to the position or the flux density S indicates that the optimal CLEAN component to select is the one with the position at the maximum of $(I_D \otimes B)$ and flux at the maximum of $I_D \otimes B / \int d^2\mathbf{u} \mid \tilde{B}_D \mid^2$. If all the visibility points have equal weights (e.g. $w_j = 1$ whenever

the point is sampled and 0 when the point is not sampled), the location of the maximum of $(I_D \otimes B)$ is at the peak of the I_D, and the maximum of $I_D \otimes B / \int d^2\mathbf{u} |\ \tilde{B}\ |^2$ is proportional to the peak of the I_D.

4.3.2 LensClean

LensClean, in short, utilizes the iterative and minimization philosophy in CLEAN and applies it to lens inversion. The major difference between LensClean and CLEAN is that LensClean takes the lensing effects into consideration when deciding which is the next optimal CLEAN component to select. Like CLEAN, LensClean models the structure of the source (in the source plane) as an ensemble of point sources. Given a lens model and a source with a flux S^i and position $\mathbf{u}^i$ in the source plane, we can calculate the positions $\mathbf{x}_j^i$ and the magnification M_j^i of all the n^i images corresponding to that source. The subscript j denotes the j^{th} image of the source, where $1 \leq j \leq n^i$. We can also calculate the mean square residuals in the dirty map after subtracting the images corresponding to a point source in the source plane with flux S^i and position $\mathbf{u}^i$. The mean square residuals are

$$R^2 = \int d^2\mathbf{u}\ |\ \tilde{I}_D - S^i \sum_{j=1}^{n^i} M_j^i \tilde{B} \exp\left(-i\mathbf{u}\cdot\mathbf{x}_j^i\right)\ |^2 \qquad (4-20)$$

(Kochanek & Narayan 1992). LensClean finds the next most optimal $\mathbf{u}^i$ and S^i by minimizing R^2 with respect to the position and flux. When such a source is found, LensClean deposits a point source at the position $\mathbf{u}^i$ with S^i in the source map and subtracts the corresponding n^i images with the correct M_j^i and correct positions $\mathbf{x}_j^i$ from the dirty map. After all source components are found, LensClean calculates the root mean square (rms) of the flux density in the residual image. The goodness of fit of the current lens model to the observed image is evaluated based on these rms residuals, and the best fit lens model is determined by minimizing these rms residuals as a function of lens parameters.

Since the dirty beam extends to infinity, it is computationally expensive to compute the expression in Eqn (4-20) to find the next optimal source clean component; hence, it is not practical to use the dirty map for the lens inversion. Instead, since the CLEAN beam (normally a Gaussian fitted to the central lobe of the dirty beam) drops off rather quickly, it is more practical to use the CLEAN map and the CLEAN beam for the inversion. Our lens modeling results were obtained using the CLEAN maps.

4.4 Optimizing the Lens Parameters in Multi-Dimensional Space

Obtaining the optimal set of parameters in multi-dimensional space numerically is complicated by the possibility that the algorithm can mistakenly treat one of the local minima as the global minimum, which causes the program to converge to a false solution. Among the parameters in our lens models, the structural parameters (R_e for the de Vaucouluers model and α and s for the α model) have the greatest uncertainties. We wanted to understand the dependency of the error surface on these parameters to ensure that the optimal parameters are found at the global minimum rather than a local minimum, so we performed the minimization with care. Our minimization procedure is:

(1) For a model with fixed structural parameters (R_e or α and s), we tested a series of models optimizing the variables b, θ_γ, and γ on a grid of fixed lens positions x_l and y_l. The lens position is reasonably well determined from the image geometry, so we examined a region approximately 0.″80 by 0.″64 centered on the position of component D. The initial values of the variables were set to fit the positions and fluxes of the compact components (A and B). The optimization of the remaining parameters was rapid and well determined. The best solutions on the grid of positions were used as the initial data for a final optimization to determine the best model for the current values of the structural parameters.

(2) We did a series of these models with varying length scales R_e or s (at fixed α) to determine the best fit scales and their allowed ranges.

(3) For the α model we repeated these procedures for the range from $\alpha = 0.3$ to $\alpha = 1$. Outside this range the residuals begin to rise steeply.

(4) Once we have isolated reasonable models for the 8 GHz image of the lens, we repeated the procedures for the 5 GHz and 15 GHz images.

All the intermediate solutions and their residuals were kept as a database for estimating the errors in the parameter estimates (see discussion in Section 4.5).

4.5 Error Estimation and Goodness of Fit

LensClean makes a χ^2 fit of the reconstructed image to the input image, so the fundamental measure of the error is the mean square difference between the two images. To construct the χ^2 statistics, we must understand the number of degrees of freedom, N_{dof}, the portions of the map that contribute to the error, and the noise level. The noise level is at least the measured noise in the map σ_0 (rms noise per pixel), although systematic errors due to the use of the CLEAN map as the initial image may lead to a higher effective noise level (see discussion in Section 5.1.2). Let N_{pix} be the number of pixels of size Δx in the map, N_{mult} be the number of multiply imaged pixels in the map, σ_{tot} be the rms residual per pixel over the entire map, and σ_{mult} be the rms residual per pixel over the multiply imaged region. The definition of χ^2 and N_{dof} depends on the effects of finite resolution and the subtraction of the residuals from the original CLEAN map.

The original CLEANing of the map produces a component map and a residual map. The components are convolved with a Gaussian CLEAN beam, and the residuals are added to produce the final CLEAN image. If the CLEAN is deep enough, the residual map is uncorrelated with the dirty beam and represents random or systematic noise in the measured visibilities. When we run LensClean on the CLEAN map

using the CLEAN beam, we can subtract most of the original residuals because of the differences between the compact CLEAN beam and the complicated dirty beam. Thus, LensCleaning the CLEAN map produces negligible residuals in the singly imaged region of the lens. In the multiply imaged region, the noise at the location of the different images is uncorrelated, so we cannot subtract all the residuals.

The number of independent data points is not the number of pixels, but the number of independent beam elements in the map. This is most easily seen by taking the "data" to be the gridded Fourier components of the map. Then we can define a χ^2 in the Fourier space as

$$\chi^2 = \sum_{i=1}^{N_{data}} \frac{(\tilde{I}_{ij} - \tilde{I}_{ij}^m)^2}{\sigma_f^2} \qquad (4-21)$$

where N_{data} is the number of Fourier components, σ_f is the noise associated with each Fourier component, and $\tilde{I}_{ij}$ and $\tilde{I}_{ij}^m$ are the measured and modeled Fourier components. The number of cells occupied by the CLEAN beam in Fourier space is

$$N_{data} = \frac{(\Delta x)^2}{2\pi\sigma_b^2} N_{pix} \qquad (4-22)$$

where σ_b is the width of the Gaussian beam. The number of independent data points is the same as the area of the map $N_{pix}\Delta x^2$ divided by the effective area of the Gaussian beam $2\pi\sigma_b^2$.

We use Parseval's theorem to relate residuals and errors in the Fourier and map planes,

$$\sum(\tilde{I}_{ij} - \tilde{I}_{ij}^m)^2 = N_{pix}\sigma_{tot}^2 \qquad \text{and} \qquad N_{data}\sigma_f^2 = N_{pix}\sigma_o^2 \qquad (4-23)$$

where σ_{tot}^2 is the pixel-to-pixel residual after LensClean. Thus the χ^2 statistic for the total map is

$$\chi_{tot}^2 = N_{data}\frac{\sigma_{tot}^2}{\sigma_o^2} = N_{pix}\frac{(\Delta x)^2}{2\pi\sigma_b^2}\frac{\sigma_{tot}^2}{\sigma_o^2}, \qquad (4-24)$$

and the χ^2 statistic for just the multiply imaged regions of the map is

$$\chi_{mult}^2 = N_{mult}\frac{(\Delta x)^2}{2\pi\sigma_b^2}\frac{\sigma_{mult}^2}{\sigma_o^2}. \qquad (4-25)$$

To evaluate the significance of the χ^2, we must also determine the number of degrees of freedom. There are M parameters in the lens model, but for the moment they are not important in estimating N_{dof} because M is small for both models. When we use LensClean to model both the lens potential and the structure of the source (in the source plane), the source structure is specified by a large number of source components, so the number of source components is the main contributor to N_{dof}. Suppose the image we use has "infinite" resolution in the sense that no images of one source point are correlated with the images of another source point by the effects of the beam. Then, a singly imaged source has one measured flux in the image, a triply imaged source has three measured fluxes, and a quintupally imaged source has five measured fluxes. In each case we fit one source flux parameter leaving no degrees of freedom for the singly imaged source, two for the triply imaged source, and four for the quintupally imaged source. If we fit the entire image plane in which fraction f_1 is singly imaged, f_3 is triply imaged, and f_5 is quintupally imaged ($f_1 + f_3 + f_5 = 1$), the number of degrees of freedom is $N_{dof} = N_{data}(1 - f_1 - f_3/3 - f_5/5) - M \simeq (2/3)N_{mult}$ (Wallington & Kochanek 1994). Thus, the number of degrees of freedom is proportional to the area that is multiply imaged.

With the addition of a beam that correlates nearby images, this overestimates the number of degrees of freedom in the model. For example, the Plummer model ($\alpha = 0$) with a fixed ring radius produces a larger and larger multiply imaged region as the core radius is reduced. In the limit that $s \to 0$, the multiply imaged region becomes infinite, suggesting that $N_{dof} \to \infty$. With finite resolution this result is clearly incorrect. Most of the large multiply imaged region consists of sources that have one image at a large radius from the lens center with nearly unit magnification, and two strongly demagnified images in the core of the lens. Consider the case where the potential has a singular core. Let there be a pair of sources, each with one image at a large radius and one image in the core. If we simultaneously fit two such image

pairs, the previous approximation gives two degrees of freedom because we fit four image fluxes with two source fluxes. However, the beam makes the two inner images fully correlated (i.e., they are the same point) so we are really fitting three image fluxes with two source fluxes giving only one degree of freedom. For N such image pairs, all N images in the core are correlated, so there are only $N+1$ image fluxes available to model N source fluxes, leaving only one degree of freedom. Thus with the addition of a finite beam size, the number of degrees of freedom is not simply proportional to N_{mult} because many of the multiple image systems act as if they were singly imaged.

The tangential critical line effectively separates the inner images from the outermost images and suggests the correct formulation for estimating the number of degrees of freedom for multiple imaging and finite resolution. Suppose we have a symmetric lens that generates only one or three images, and we assign the source flux to fit the images outside the tangential critical line exactly. All the residuals from this procedure are inside the tangential critical line, and the number of degrees of freedom is the number of resolution elements inside the tangential critical line. This model generalizes to simultaneously fitting all the images and is approximately correct when we add the five image region. Thus the correct estimate for the number of degrees of freedom in the models is

$$N_{dof} = \frac{A_{tan}}{2\pi\sigma_b^2} - M, \qquad (4-26)$$

where A_{tan} is the area inside the tangential critical line of the lens. A_{tan} has the desirable property of being fixed for all reasonable lens models, because all reasonable lens models must have the same average tangential critical line to be able to fit the ring. Thus $A_{tan} = \pi\langle r\rangle^2$ where $\langle r\rangle$ is the average ring radius, with corrections that are second order in the dimensionless shear, γ. The approximation fails when the size of the region inside the tangential critical line is comparable to the size of the

CLEAN beam. Note that a good inversion should have an rms residual for the whole map smaller than the intrinsic noise by the factor $(A_{tan}/A_{map})^{1/2}$, and an rms residual for the multiply imaged region smaller than the intrinsic noise by the factor $(A_{tan}/A_{mult})^{1/2}$.

Given a χ^2 and N_{dof}, and assuming the errors are Gaussian and uncorrelated, we expect $\chi^2 = N_{dof}$ for a good inversion with the value of χ^2 distributed as a χ^2 distribution with N_{dof} degrees of freedom. Since $N_{dof} \gg 1$, the expected standard deviation of the χ^2 from N_{dof} is approximately $(2N_{dof})^{1/2}$. In practice, we tend to find larger values of χ^2 both because of the real limitations in the models and because of systematic error. One source of systematic error is the interpolation and reconstruction error due to CLEAN (see discussion in Section 5.1.2). A second systematic difficulty is that LensClean needs to use an automatic stopping criterion to decide when to stop CLEANing. Like normal uses of CLEAN on extended sources, it is possible to subtract more flux in the map by varying the total number of components in use. We stop the procedure when it can no longer reduce the peak residual, so there is still subtractable flux in the map when the procedure stops. We can compensate for the uncertainties in the true noise level and the stopping criterion by renormalizing the errors so that $\chi^2_r = N_{dof}\chi^2/\chi^2_{min}$, where χ^2_{min} is the smallest measured value of χ^2. This will systematically *underestimate* the statistical significance of differences among models, and systematically *overestimate* the error estimates on model parameters. In short, it provides a reasonably well defined method of making conservative error estimates.

The limits on the variations of a single parameter about a minimum χ^2_{min} in the χ^2 distribution are determined by the variations in $\Delta\chi^2 = \chi^2 - \chi^2_{min}$, which is expected to follow a χ^2 distribution with one degree of freedom. The formal 68% confidence interval is given by the range of values that produces $\Delta\chi^2 \leq 1$. In practice (see below) we find this criterion gives unrealistically small error estimates. A more

conservative error estimate is found by renormalizing the χ^2 of all the intermediate trial models in the database and collecting all trial models with $\Delta\chi^2_r \leq 15.1$ (the formal 99.99% confidence level). The error bars on the parameters are set to be the largest deviation from the best fit value. The error bars should not, however, be considered true 99.99% confidence level error bars because this assumes that the errors follow a Gaussian distribution with N_{dof} number of degree of freedom. A more realistic assessment might be to consider them to be two standard deviation error bars. In practice, the correct way to estimate the significance of the errors would be with Monte Carlo simulations of the data, but Monte Carlo error estimation requires many realizations, each as time consuming as the original inversion, making it computationally impractical.

References

Binney, J. & Tremaine, S. 1987, *Galactic Dynamics*, Princeton University Press, Princeton, New York.

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

Chen, G. H. & Hewitt, J. N. 1993, AJ, **106**, 1719

Chen, G. H., Kochanek, C. S. & Hewitt, J. N. 1995, Ap. J., in press.

Cornwell, T. J. & Braun, R. 1989, *in Synthesis Imaging in Radio Astronomy*, Volume 6, eds. R. A. Perley, F. R. Schwab & A. H. Bridle, Astronomical Society of the Pacific, pp. 167.

de Vaucouleurs, G. 1948, Ann d'Astrophys., **11**, 247

Hewitt, J. N., Turner, E. L., Schneider, D. P., Burke, B. F., Langston, G. I. & Lawrence, C. R. 1988, Nature, **333**, 537

Högbom, J. A. 1974, A&AS, **15**, 417

Kochanek, C. S. 1995, Ap. J., in press.

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

Kochanek, C. S., Blandford, R. D., Lawrence, C. R. & Narayan, R. 1989, MNRAS, **238**, 43

Maoz, D. & Rix H.-W. 1993, ApJ, **416**, 425

Schwars, U. J. 1978, Astron. Astrophys., **65**, 345.

Wallington, S. & Kochanek, C. S. 1994, ApJ, in press.

5. Results of the Lens Modeling

5.1 8 GHz Result

The 8 GHz map has the best combination of resolution and signal to noise. The intrinsic rms noise per pixel in this map is 35 μJy, and the FWHM of the synthesized beam is 0.″19 (see Table 4.1). The number of resolution elements inside the tangential critical line is 65, so the numbers of degrees of freedom are $N_{dof} = 59$ for the de Vaucouleurs model and $N_{dof} = 58$ for the α model.

5.1.1 de Vaucouleurs Models

Figure 5.1 shows the rescaled χ^2, χ^2_r, estimated from both the total (solid line) and multiply imaged region (dashed line) as a function of the effective radius, R_e (when other parameters are optimized); the $\Delta\chi^2_r = 15.1$ limit from the minimum of χ^2_r is represented by the dotted line. Models with R_e in the range $0.''64 \lesssim R_e \lesssim 0.''96$ have χ^2_r below the $\Delta\chi^2 = 15.1$ limit. Table 5.1 summarizes the values of the lens parameters for several different de Vaucouleurs models. The best fit model has $R_e = 0.''83 \pm 0.''13$, $\sigma_{tot} = 39.5$ μJy, $\sigma_{mult} = 55.6$ μJy, and a peak residual of 342 μJy. We find the same best fit effective radius whether we use χ^2_{tot}, χ^2_{mult}, or even just the peak residual. Figure 5.2 shows the reconstructed image, the residuals, and the inferred source of the best fit model. Although the reconstructed image is very similar to the original image and the peak residuals are only 5% of the original peak of the map, there clearly are significant residuals. The best fit model has $\chi^2_{tot} = 825$ and $\chi^2_{mult} = 703$. These χ^2 values are formally 60 (!) standard deviations from the target value $N_{dof} = 59$. If we understand the χ^2 statistics and if the estimates are not dominated by unknown systematic errors, then we can completely reject the de Vaucouleurs + external shear model for the mass distribution in the lens galaxy of MG1131. The best test of whether we do understand the errors is to see whether we can find models with

significantly lower χ^2 values.

The error bar associated with the effective radius is determined by the point where $\Delta\chi^2_r = 15.1$ (see the discussion in Section 4.5). This is an extremely conservative way of determining parameter errors. In the χ^2 parameter estimation, a one standard deviation change in the parameter corresponds to $\Delta\chi^2 = 1$ whereas our standard error bar is defined using $\Delta\chi^2_r = 15.1$. We can illustrate the conservatism of our parameter errors by noting that the error bar on R_e using $\Delta\chi^2_r = 1$ in the rescaled statistic is 0.″05, and the error bar using $\Delta\chi^2 = 1$ in the unrescaled statistic is 0.″01.

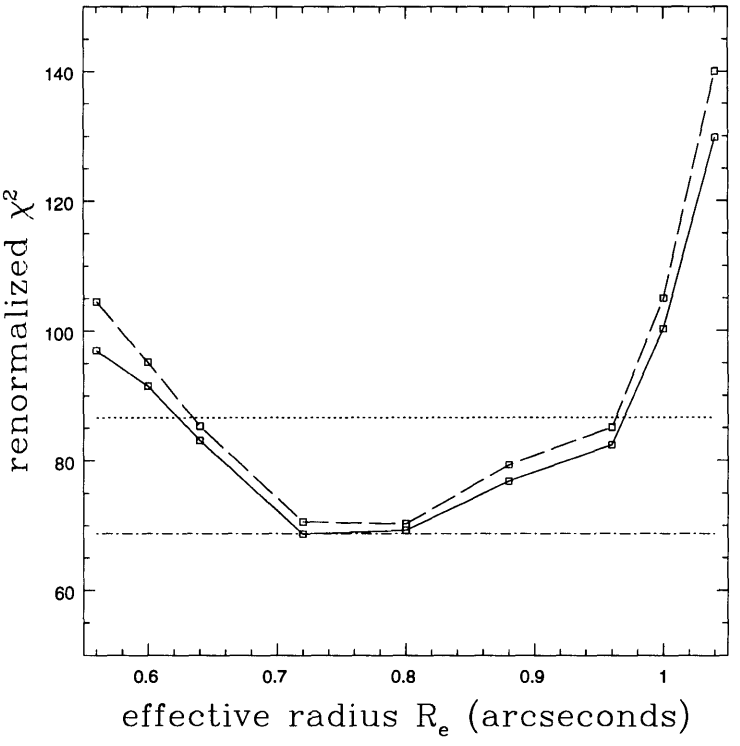


Figure 5.1 – The renormalized χ^2 of the 8 GHz de Vaucouleurs models as a function of the effective radius R_e. To ease the comparison between the de Vaucouleurs models and the α models, the renormalization constant used here is the same as the one used in the 8 GHz α models. The dotted-dashed line represents the one standard deviation increase in χ^2_r from the best α model. The solid line shows the renormalized χ^2_{tot}, and the dashed line shows the renormalized χ^2_{mult} of the de Vaucouleurs models. The dotted line shows $\Delta\chi^2_r = 15.1$ from the best de Vaucouleurs model in the renormalized χ^2 statistics.

Table 5.1: The 8 GHz de Vaucouleurs Models

R_e	$\Delta\sigma$	χ^2_{tot} [a]	χ^2_{mult} [a]	b	x_l [b]	y_l [b]	γ	θ_γ
$''$	μJy			$''$	$''$		degress	
0.56	436	1244	1100	1.094 ± 0.012	-0.017 ± 0.006	0.003 ± 0.008	0.164 ± 0.006	-26.0 ± 0.4
0.60	461	1173	1002	1.053 ± 0.005	-0.016 ± 0.007	0.002 ± 0.011	0.158 ± 0.006	-26.1 ± 0.7
0.64	446	1066	899	1.032 ± 0.015	-0.011 ± 0.012	0.010 ± 0.013	0.156 ± 0.008	-26.5 ± 0.9
0.72	393	880	743	0.995 ± 0.013	-0.003 ± 0.010	-0.009 ± 0.010	0.143 ± 0.010	-25.5 ± 1.0
0.80	366	889	741	0.950 ± 0.011	-0.001 ± 0.011	-0.005 ± 0.017	0.141 ± 0.009	-25.9 ± 1.2
0.88	390	987	837	0.922 ± 0.010	0.011 ± 0.10	-0.014 ± 0.010	0.129 ± 0.006	-25.9 ± 1.6
0.96	404	1056	896	0.893 ± 0.008	0.019 ± 0.009	-0.017 ± 0.013	0.128 ± 0.011	-26.1 ± 2.0
1.00	467	1285	1106	0.879 ± 0.018	0.023 ± 0.011	-0.017 ± 0.025	0.120 ± 0.007	-25.5 ± 1.3
1.04	519	1664	1475	0.855 ± 0.012	0.025 ± 0.011	-0.020 ± 0.011	0.114 ± 0.007	-24.9 ± 1.7
0.83 ± 0.013	342	825	703	0.941 ± 0.034	0.005 ± 0.014	-0.010 ± 0.011	0.135 ± 0.015	-25.9 ± 0.6

a $N_{dof} = 59$.
b The position of the lens is given relative to the position of component D at 8 GHz

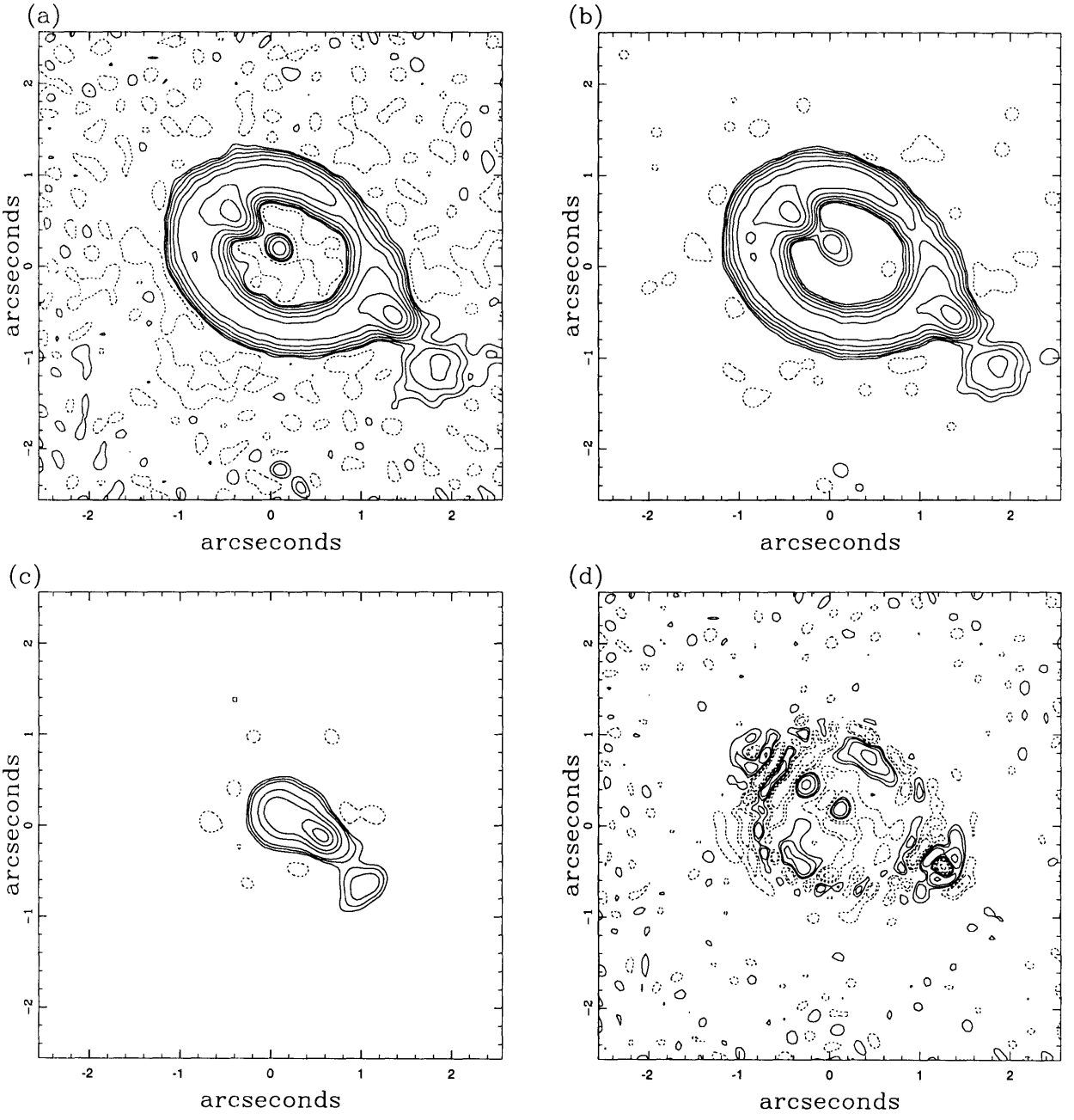


Figure 5.2 – (a) The observed image, (b) the reconstructed image, (c) the inferred source, and (d) the residual map obtained from the best 8 GHz de Vaucouleurs model. The contour levels in the figures are: (a) and (b) -1, 1, 2, 4, 8, 16, 32, 64, 128, and 256 × 35 μJy, (c) -1, 1, 2, 4, 8, 16, 32, 64, and 95% × 155 μJy, and (d) -8, -4, -2, -1, 1, 2, 4, and 8 × 35 μJy. In a perfect reconstruction, (a) and (b) should be identical.

Both are significantly smaller than our standard error estimate of 0.″13. The large difference between the χ^2 and χ^2_r statistics is caused by the large rescaling of the errors needed to make the de Vaucouleurs models a good fit to the data.

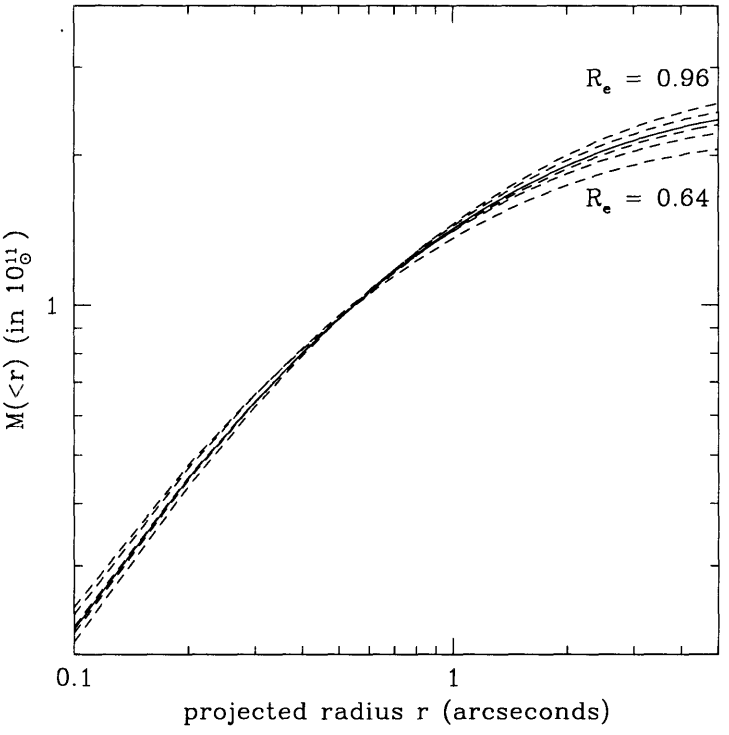


Figure 5.3 – The integrated mass distribution as a function of projected radius r for various de Vaucouleurs models. The mass is calculated assuming $z_s = 2.0$ and $z_l = 0.5$ in an Einstein de Sitter universe. For other values of redshifts, multiply the mass by $4.55 D_{os} D_{ol}/(2 r_h D_{ls})$, where D_{ij} is the angular diameter distance between i and j and r_h is the Hubble radius. The solid line shows the mass profile for the best de Vaucouleurs model ($R_e = 0.''83$), and the dashed line shows the mass profile for models with $R_e = 0.''64, 0.''72, 0.''80, 0.''88$, and $0.''96$.

The total mass of the galaxy in the best fit model is $\mathrm{M} = 1.17 h^{-1} [D_{ol} D_{os}/(2 r_h D_{ls})] \times 10^{12} M_\odot$. The value of M depends on the redshifts of the source and the lens. Since both z_s and z_l are unknown, we can not evaluate the quantity in the brackets. However, if we assume that the source is at $z_s = 2.0$ and that the lens is at $z_l = 0.5$, then the quantity in the bracket would be 0.22 in an Einstein de Sitter universe, and the total mass would be $\mathrm{M} = 2.57 \times 10^{11} h^{-1} M_\odot$. Figure 5.3 shows the profile of the integrated mass distribution as a function of the projected radius r for models with

R_e in the range $0\farcs64 \lesssim R_e \lesssim 0\farcs96$. As expected, the mass interior to the average ring radius $\langle r \rangle \sim 0\farcs9$ is rather insensitive to the model parameters. All models have $M(\langle r \rangle) \sim 1.4 \times 10^{11}_{\odot}$ if $z_s = 2.0$ and $z_l = 0.5$. Figure 5.4 shows the monopole deflection angles as a function of the projected radius r for these models. As expected, the deflection angles at the average ring radius for all models are $\sim 0\farcs9$.

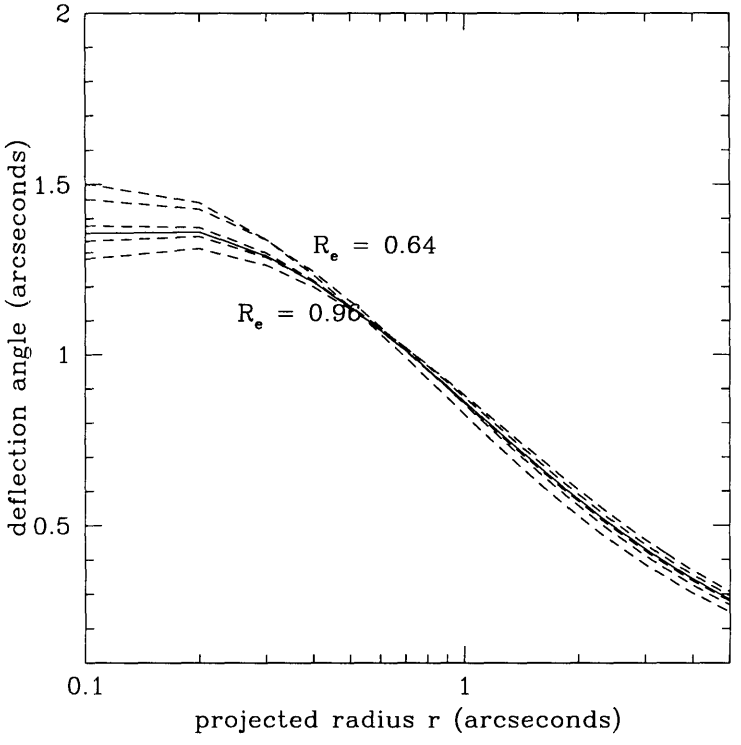


Figure 5.4 – The monopole deflection as a function of the projected radius r for various de Vaucouleurs models. The solid line shows the profile for the best de Vaucouleurs model ($R_e = 0\farcs83$), and the dashed line shows the profile for models with $R_e = 0\farcs64, 0\farcs72, 0\farcs80, 0\farcs88$, and $0\farcs96$.

Hammer et al. (1991) determined the best fit effective radius to their optical data to be $R_e \sim 0\farcs5$, but no uncertainty was given. Our lens modeling results strongly rule out the lens model with a mass distribution having an $R_e = 0\farcs5$. If the uncertainty in Hammer et al.'s estimate were small, our result ($R_e = 0\farcs83 \pm 0\farcs13$) would indicate that the mass distribution does not follow the light and that there is "more" dark mass at larger radii from the center of the galaxy. However, our best fit R_e could still

be consistent with the optical estimate if the uncertainties in the optical estimate are large enough. New optical data, with a better estimate of the uncertainties in the R_e fitted to the optical profile, is essential in making a more quantitative comparison between the modeled mass profile and the profile of the optical light. We would like to bring out an interesting result found by Kochanek (1995). He modeled another ring system, MG1655+134 (Langston et al. 1989), for the purpose of determining the mass distribution of the lens galaxy in that system. His result indicated that the best fit de Vaucouleurs mass model also has an effective radius R_e much larger than the optical estimates, which is in parallel with what we find.

5.1.2 α Models

The qualitative analysis presented in the last chapter (see discussion in Section 4.2.3) predicts that only models with α and s in a restricted range can successfully fit the system. We compare this prediction with the results obtained from the full lens modeling. Figure 5.5 gives a typical example of how χ_r^2 varies as a function of s, at a fixed α (α is set to 0.6 in this example), when other parameters (i.e., b, x_l, y_l, γ, and θ_γ were optimized). The dotted line represents the $\Delta\chi_r^2 = 15.1$ limit. The model with $s = 0\farcs16$ has the smallest χ_r^2, and only models with $0\farcs08 \lesssim r \lesssim 0\farcs24$ have χ_r^2 below the $\Delta\chi_r^2 = 15.1$ limit. Figure 5.6 shows the optimal s and the range of s with $\Delta\chi_r^2$ below the 15.1 line as a function of α. To compare with the analytical prediction, we superimpose the results obtained from the full lens inversion onto the analytical estimates (see Figure 5.7). For models that can simultaneously fit the constrains of the compact components and the ellipticity requirement (i.e. models with $\alpha < 0.9$), the analytical predictions agree with the results obtained from the full inversion.

Table 5.2 summarizes the models derived from the 8 GHz image. The lens position is fixed and independent of the other lens parameters because the lens position is

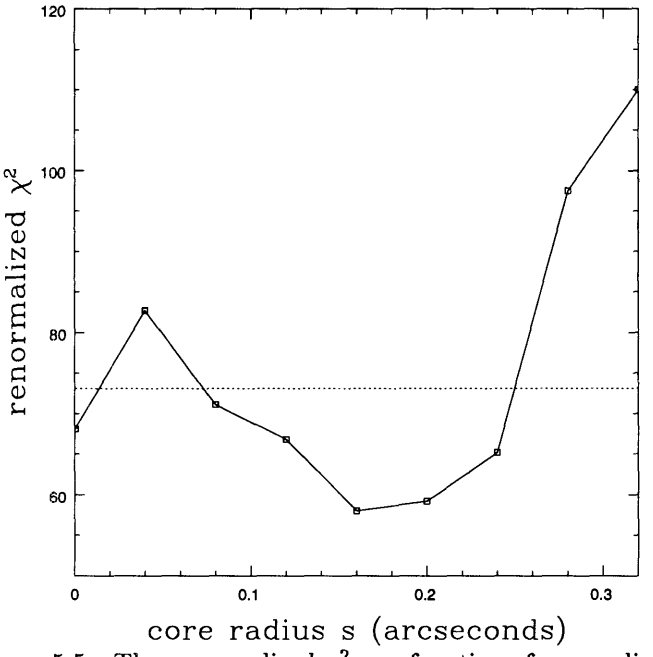


Figure 5.5 – The renormalized χ^2 as a function of core radius s for $\alpha = 0.6$. This figure demonstrates how the limits on s, at a fixed value of α, are determined. The dashed line shows $\Delta\chi^2 = 15.1$ in the renormalized χ^2 statistics. The range of the core radii under the dashed line determines the limits on s. This method is applied to every value of α to determine the limits on s.

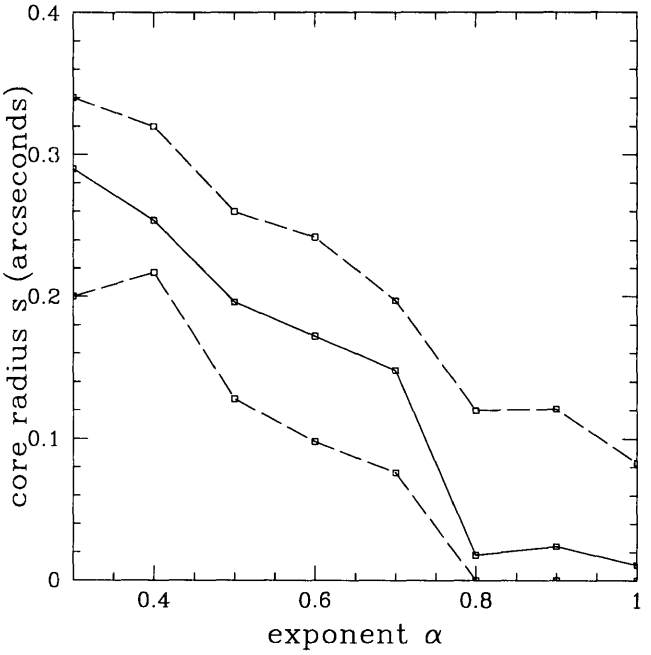


Figure 5.6 – The optimized value of the core radius s (solid line) and its limits (dashed line) as a function of α.

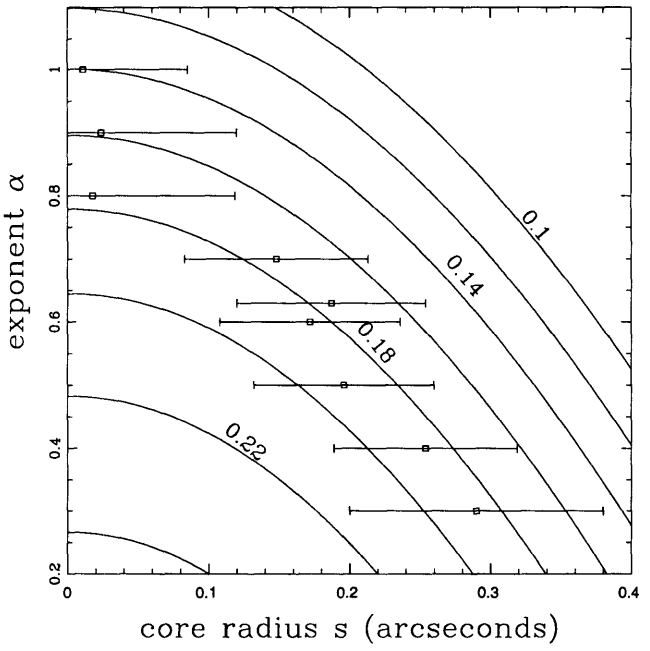


Figure 5.7 – The points and their error bars show the optimized values of the core radii s and their limits obtained from the full inversion. The contour levels represent the predicted ellipticity obtained from the analytical estimates (see §4.2.3). For models with $\alpha < 0.9$, the analytical predictions agree with the results obtained from the full inversion.

strongly constrained by the geometry of the ring. Similarly, the position angle of the shear is model independent because the tangential critical line of the lens must have its major axis at right angles to the major axis of the ring as discussed by Kochanek et al. (1989). None of the models has a χ^2 value close to $N_{dof} = 58$. The best fit model has a $\chi^2_{tot} = 740$ and a $\chi^2_{mult} = 598$. Although these χ^2 values are much larger than N_{dof}, they are significantly smaller than the best fit de Vaucouleurs model ($\chi^2_{tot} = 825$ and $\chi^2_{mult} = 703$). If we rescale the errors so that the best fit α model has $\chi^2_r = N_{dof}$, then the de Vaucouleurs models are 0.63 (0.94) standard deviations worse in the total (multiple image) error estimates. If we do not rescale the errors, the difference between the χ^2 in de Vaucouleurs model and in the α models of is much larger – 8 (10) standard deviation. This change of the differences in the χ^2 between

Table 5.2: 8 GHz α models

α	$\Delta\sigma$	χ^2_{tot} [a]	χ^2_{mult} [a]	b	x_l [b]	y_l [b]	γ	θ_γ	s
	μJy			"	"		degrees	"	
0.3	398	955	797	0.951 ± 0.039	0.015 ± 0.011	-0.007 ± 0.018	0.133 ± 0.005	-26.5 ± 1.9	0.290 ± 0.090
0.4	386	804	646	0.933 ± 0.022	0.001 ± 0.022	-0.007 ± 0.018	0.137 ± 0.009	-25.7 ± 1.7	0.254 ± 0.065
0.5	344	792	658	0.922 ± 0.025	0.004 ± 0.010	-0.011 ± 0.018	0.132 ± 0.006	-25.8 ± 1.8	0.196 ± 0.064
0.6	343	744	611	0.919 ± 0.008	0.009 ± 0.010	-0.006 ± 0.007	0.125 ± 0.003	-26.2 ± 0.7	0.172 ± 0.068
0.7	346	768	628	0.919 ± 0.015	-0.001 ± 0.012	0.002 ± 0.019	0.123 ± 0.008	-26.5 ± 1.5	0.148 ± 0.065
0.8	421	880	768	0.907 ± 0.016	0.001 ± 0.007	0.008 ± 0.025	0.113 ± 0.006	-27.2 ± 2.0	0.018 ± 0.101
0.9	366	951	832	0.912 ± 0.018	-0.015 ± 0.018	-0.001 ± 0.023	0.111 ± 0.007	-25.6 ± 1.6	0.024 ± 0.096
1.0	411	1178	1022	0.918 ± 0.011	-0.020 ± 0.023	0.017 ± 0.022	0.108 ± 0.010	-27.2 ± 2.5	0.026 ± 0.074
0.63 ± 0.23	357	740	598	0.924 ± 0.017	-0.002 ± 0.007	-0.002 ± 0.010	0.124 ± 0.012	-25.9 ± 1.3	0.187 ± 0.067

a $N_{dof} = 58$.
b The position of the lens is given relative to the position of component D at 8 GHz

the rescaled and non-rescaled χ^2 statistics shows the conservatism inherent in the rescaling process.

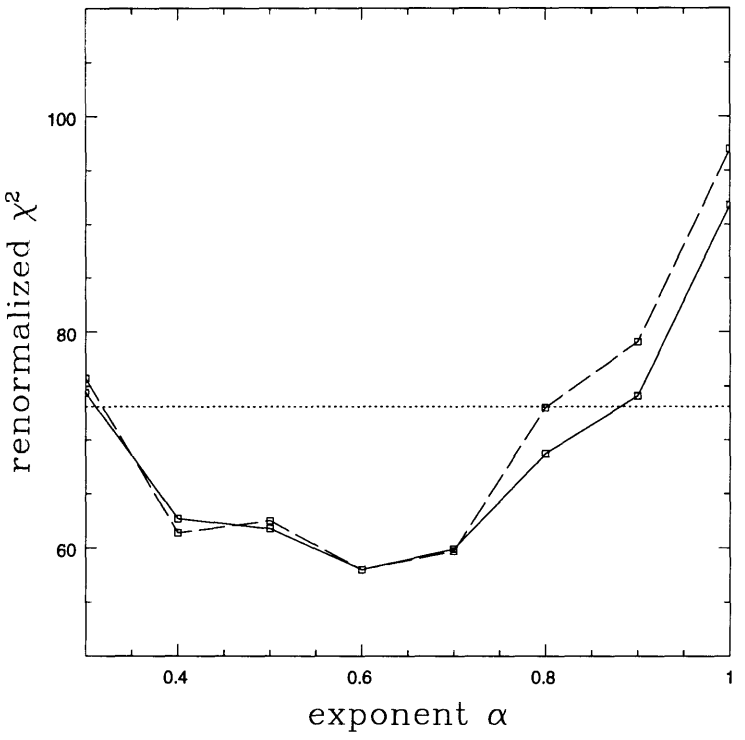


Figure 5.8 – The renormalized χ^2 of the 8 GHz α models at the optimum value of the core radius. The solid line shows the renormalized χ^2_{tot}, and the dashed line shows the renormalized χ^2_{mult}. The horizontal dotted line represents $\Delta\chi^2 = 15.1$ in the renormalized χ^2 statistics.

Figure 5.8 shows the two renormalized χ^2 statistics as a function of the exponent α after the core radius is optimized. With the renormalization, we find that models with $0.4 \lesssim \alpha \lesssim 0.8$ are within $\Delta\chi^2_r = 15.1$ of the minimum. The best fit value of α is $\alpha = 0.63 \pm 0.23$. The isothermal model ($\alpha = 1$) lies outside the permitted range, making this the first example of a lens whose radial mass distribution apparently cannot be modeled by a quasi-isothermal potential. We should again note the conservatism of this range estimate: using normal one standard deviation errors with $\Delta\chi^2 = 1$ on the rescaled (unrescaled) value of the χ^2 gives errors on the best fit value of α of

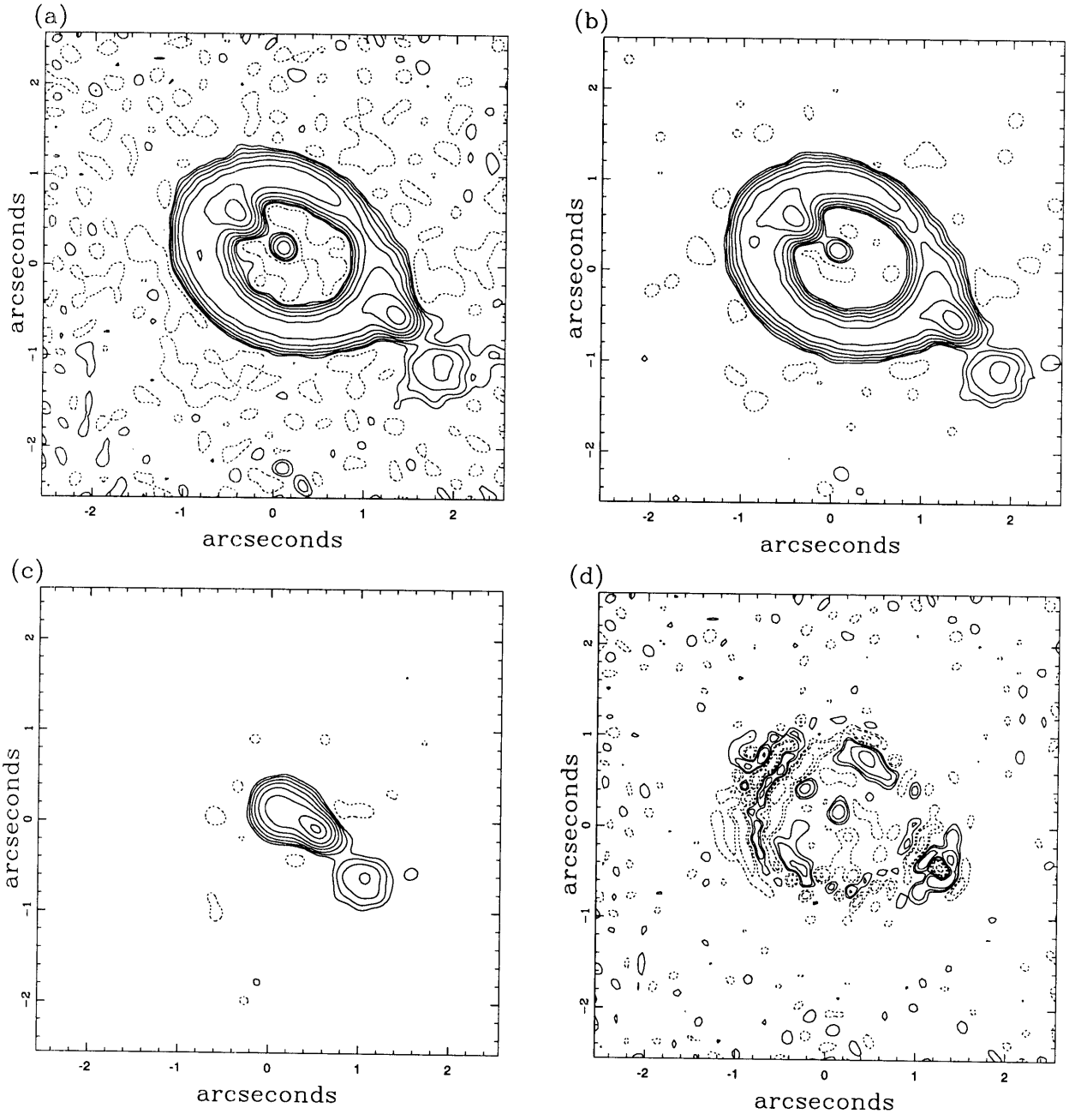


Figure 5.9 – (a) The observed image, (b) the reconstructed image, (c) the inferred source, and (d) the residual map obtained from the best 8 GHz α model. The contour levels in the figures are: (a) and (b) -1, 1, 2, 4, 8, 16, 32, 64, 128, and 256 $\times$ 35μJy, (c) -1, 1, 2, 4, 8, 16, 32, 64, and 95% $\times$ 155μJy, and (d) -8, -4, -2, -1, 1, 2, 4, and 8 $\times$ 35μJy. In a perfect inversion, (a) and (b) should be identical.

± 0.06 (± 0.02). The peak residual in the best fit map is 357 μJy, and all the acceptable models ($0.4 \lesssim \alpha \lesssim 0.8$) have peak residuals smaller than 421 μJy/pixel. Figure 5.9 shows the residual map, the inferred source, and the reconstructed image for the best fit model given in Table 5.2. The reconstruction is a plausible extragalactic radio source consisting of a bright core, two radio lobes and, possibly, a short jet. Since the residuals are relatively small, the reconstructed image and the observations cannot easily be distinguished by eye. The reason why the χ^2 is large can be clearly seen from the large residuals.

Table 5.3 summarizes some of the physical properties of the models such as the mean ring radius, the ring ellipticity, the mass interior to the ring, and the time delay between the compact components. Models with $0.4 \lesssim \alpha \lesssim 0.8$ give roughly the same ellipticity (~ 0.17) which matches the value measured from the ring (0.18 ± 0.02). As expected, the good models track the band of ellipticities consistent with the ellipticity of the ring (see Figure 5.7). Note that the best fit isothermal model derived from fitting the ring gives an estimated ring ellipticity of 0.194 while the analytic estimate derived from fitting the compact components in Section 4.2.3 predicts that an isothermal model fit to the compact components would give an ellipticity of 0.14. In the inversion, the dominant constraint is the structure of the ring, so the isothermal models obtained from the full inversion are influenced much more strongly by the constraints of the ring rather than the constraints of the compact components. In contrast, the analytical estimates were obtained by fitting only the constraints of the compact components. Kochanek et al. (1989) modeled the 15 GHz map of MG1131 (in which the dominant constraints were the compact components) using an elliptical isothermal model for the lens galaxy. We recall that Kochanek et al.'s models (1989) always produced a ring that was too circular, consistent with our understanding of the behavior of the isothermal model. We compute the mass profile (see Figure 5.10) and the monopole deflection angle (see Figure 5.11) as a function of the projected

radius r assuming that $z_s = 2.0$ and $z_l = 0.5$. As one would expect, the mass interior to the ring and the deflection angle at $\langle r \rangle$ are accurately determined and insensitive to the lens models. For $z_s = 2.0$ and $z_l = 0.5$, $\mathrm{M}(< \langle r \rangle) \sim 1.4 \times 10^{11}_{\odot}$, similar to the mass predicted by the de Vaucouleurs models.

Table 5.3. Physical Properties of the 8 GHz α Models

α	$\langle r \rangle$[a]	e[b]	$\mathrm{M}(\langle r \rangle)$[c]	$\Delta\tau$[c]
	$''$		$h^{-1} \times 10^{11} M_{\odot}$	h^{-1} years
0.3	0.914	0.159	1.36	0.183
0.4	0.907	0.170	1.34	0.175
0.5	0.909	0.169	1.35	0.171
0.6	0.912	0.170	1.35	0.163
0.7	0.917	0.178	1.37	0.154
0.8	0.916	0.173	1.37	0.147
0.9	0.922	0.184	1.38	0.135
1.0	0.929	0.194	1.40	0.124
0.63 ± 0.23	0.914	0.172	1.36	0.159

a The average ring radius.

b The ellipticity of the ring is defined as $e = 1 - b/a$, where b and a are the semi-minor and semi-majo axes of the ring.

c The massses and the time delays listed in the table are calculated assuming that $z_s = 2.0$ and $z_l = 0.5$ in an Einstein de Sitter universe. To correct for other combination of redshifts, multiply the mass by $4.55 D_{ol} D_{os} / 2 r_h D_{ls}$ and the $\Delta\tau$ by $3.03(1 + z_l) D_{os} D_{ol} / 2 r_h D_{ls}$.

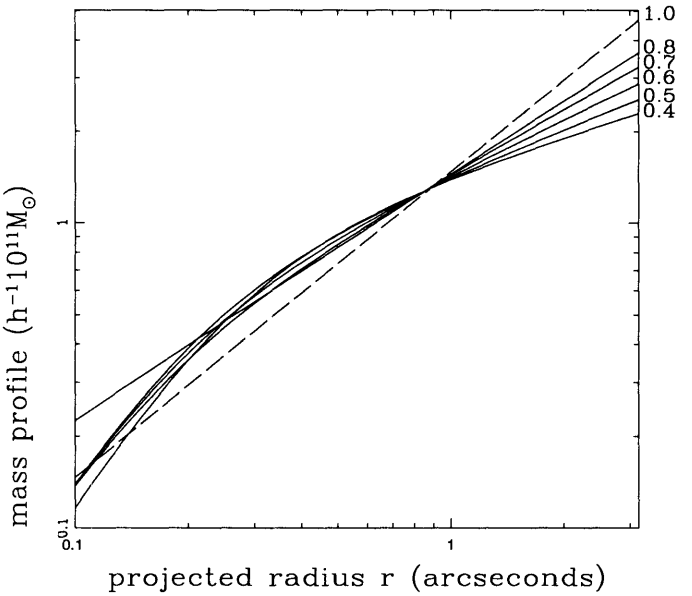


Figure 5.10 – The monopole mass distribution of the acceptable 8 GHz α models as a function of projected radius (solid lines). The values of α are labeled on the right hand border of the figure. The mass is calculated assuming $z_s = 2.0$ and $z_l = 0.5$ in an Einstein de Sitter universe. For other values of redshifts, multiply the mass by $4.55 D_{os} D_{ol}/(2 r_h D_{ls})$, where D_{ij} is the angular diameter distance between i and j. The mass distribution of the best fit isothermal model is also shown for comparison (dashed line).

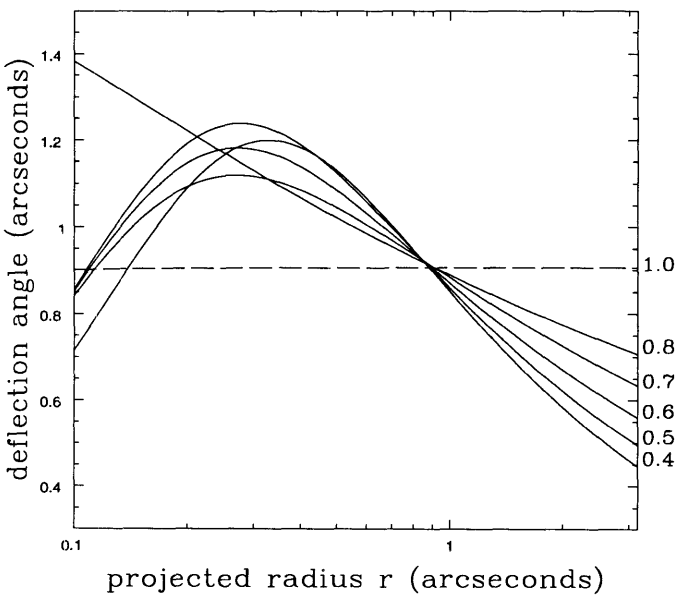


Figure 5.11 – The monopole deflection angle of the acceptable α models as a function of the projected radius r. The values of α are labeled on the right hand border of the figure.

The predicted time delay varies slightly from one model to the other (see Table 5.3), with the largest variation of $\pm 9\%$ between models with $0.4 < \alpha < 0.8$. The formal one standard deviation errors on the time delay are $\pm 1\%$ for $\Delta\chi^2 = 1$ and $\pm 4\%$ for $\Delta\chi^2_r = 1$.

The formal value of the χ^2_{mult} of the best α model ($\chi^2_{mult} = 598$) is about 50 (!) standard deviations from the target of $N_{dof} = 58$. Such a poor fit to the data requires some justification. Part of the problem is the use of the CLEAN map as the initial image. All types of non-linear image restoration methods interpolate between the irregularly sampled visibilities in the Fourier space, and it is highly possible that the interpolation introduces some artifacts into the restored image. For gravitational lens systems, the unmeasured visibilities are correlated with each other because of the global constraints of lensing - different parts of the images are originated from the same source. Since all image restoration methods fill in the missing visibilities "locally" without accounting for the global constraints of lensing, systematic errors are introduced by using any reconstructed map as the initial image for lens modeling.

Changing the loop gain parameter, γ_g, in CLEAN (see Section 4.3.1) effectively changes how CLEAN interpolates between the measured visibilities. Thus, we can examine some of the systematic errors introduced by CLEAN by comparing the results obtained from images reconstructed with different γ_g. We made seven 8 GHz images each with a different γ_g (see Table 5.4 for the list of γ_g). Examining the maps, we found that the basic structure of the ring was preserved when we vary γ_g, but some of the small scale structures were not. The distribution of the CLEAN components in the CLEAN map becomes more clumpy when the γ_g is larger. We first LensCleaned all these maps using the parameters in the $\alpha = 0.6$ model. Table 5.4 shows the σ_{tot} and the peak residual $\Delta\sigma$ in each case. Since the model was optimized using the map with $\gamma_g = 0.1$, it is not surprising that this map has the smallest residuals. The largest residual is 17% larger, and the average residual is 9% higher than the residual

of the map with $\gamma_g = 0.1$. The magnitude of the residuals suggests that there is a systematic error of about 20 μJy per pixel rms in addition to the intrinsic noise of 35 μJy per pixel when the CLEAN map is used as the initial image.

In addition, we also wanted to investigate whether the converged lens parameters would change if we used a different CLEANed map as the initial image. To do so, we optimized the lens model to fit each of the seven CLEAN maps. Table 5.5 summarizes the result. The converged models are all very similar. The scatter in the parameters amounts to 0.3% in the critical radius, 0.″003 in the lens position, 1.6% in the dimensionless shear, and 0.4° in the angle of the shear. These systematic errors are smaller than the statistical errors we derived previously (see Table 5.2), so these tests suggest that the results of the inversions are not dominated by the systematic

Table 5.4: The First Gain Experiment

gain	σ^a(μJy/pixel)	$\Delta\sigma^b$(μJy)
0.20	41.7	362
0.15	42.7	382
0.10	37.6	343
0.08	41.1	404
0.06	38.8	363
0.05	40.1	376
0.02	44.1	416
mean	40.9 ± 0.22	378 ± 25

a The root mean square flux in the residual map
b The peak flux in the residual map

Table 5.5: The Second Gain Experiment - Reoptimized Models

gain	b	x_l	y_l	γ	θ_γ	σ^a	$\Delta\sigma^b$
	$''$	$''$	$''$		degrees	μJy/pixel	μJy
0.20	0.923	0.007	-0.004	0.125	-26.6	40.7	368
0.15	0.927	0.009	-0.014	0.124	-25.4	41.7	356
0.10	0.919	0.009	-0.006	0.125	-26.2	37.6	343
0.08	0.921	0.008	-0.006	0.123	-26.3	40.4	393
0.06	0.921	0.010	-0.008	0.124	-26.1	38.7	363
0.05	0.922	0.008	-0.007	0.125	-25.9	39.6	375
0.02	0.919	0.002	-0.005	0.127	-25.9	40.0	342
mean	0.922 ± 0.003	0.008 ± 0.003	-0.007 ± 0.003	0.125 ± 0.001	-26.1 ± 0.4	39.8 ± 1.3	363 ± 18

a Root mean square flux in the residual map.
b Peak flux in the residual map.

introduced by using the CLEAN maps as the initial images for lens modeling.

Another systematic effect that causes the χ^2 to remain higher than expected is introduced by the automated stopping criterion required by LensClean. When we modeled the system using LensClean, we terminated the program when the peak residual could no longer be reduced further. This is a very good criterion for deciding when it is no longer profitable to pursue the current model. It does, however, mean that the procedure stopped before the mean square residuals used in the χ^2 statistics were truly minimized. If we take the best fit model and use 30,000 components (instead of the 1500 used with the standard stopping criterion), the χ^2_{tot} drops to 336 from 740 (and χ^2_{mult} drops to 315 from 598). For a comparison, using 30,000 components reduces the χ^2 estimates of the best de Vaucouleurs model to χ^2_{tot} drops to 430 from 825 (and χ^2_{mult} drops to 407 from 703). Both χ^2 values of the best de Vaucouleurs model are still considerably larger than the ones from the best α model. Thus, a different stopping criterion reduces the absolute residuals, but leaves the differences between models unaffected. If we include the additional 20μJy per pixel systematic error introduced due to the use of the CLEAN map as the initial image, we can further reduce the χ^2 of the best α model to $\chi^2_{tot} = 253$ (and $\chi^2_{mult} = 237$), which is still much larger than required for a good fit to the data. In the end, even though the reconstructed image closely resembles the observed image, our best fit model is not totally consistent with the data. We will discuss the implications of this result in Section 5.5.

5.1.3 Can the Nature of D Bias the Result?

The core radius plays a variety of roles in the lens models. It controls the flux of the central component D, the size of the multiply imaged region, and the fraction of component C that is multiply imaged. We can examine whether the core radius is determined by the need to fit the flux of the central component or by the need to

correctly fit other constraints of the system by comparing the results obtained from the maps including and excluding D as part of the lensed images. Since the nature of D is ambiguous (see discussion in Chapter 3), we can also check the effects of misinterpreting the nature of D on the lens modeling by looking at models with D subtracted.

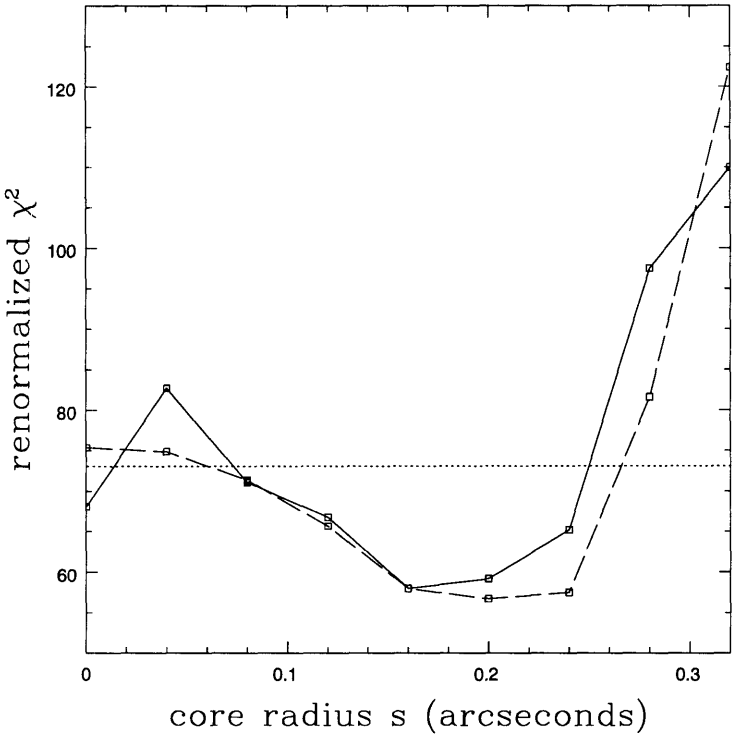


Figure 5.12 – The result of the experiment testing the effect of D. The dashed line shows the renormalized χ^2_{tot} as a function of the core radius at a fixed α ($\alpha = 0.6$) when D is subtracted before lens modeling. The results of the equivalent lens models when D is not subtracted are also shown (the solid line). The similarity of the statistics shows that the lens models are not affected by the assumed nature of D. The dotted line shows $\Delta\chi^2_r = 15.1$ in the renormalized χ^2 statistics.

A modified map was made by subtracting D from system. We modeled this map at a fixed value of $\alpha = 0.6$ to redetermine the limits on the core radius. Since the main purpose of this experiment was to understand the dominant constraints on the core radius, examining a single value of α should be sufficient. Figure 5.12 shows the χ^2_r as a function of core radius when D is subtracted (dashed line). For comparison,

we also plot the results when D is not subtracted (solid line). The structures of both curves and the limits on s obtained from both cases are rather similar, demonstrating two important points. First, the core radius of the potential is largely constrained by the location of the caustics and the geometry of the ring, rather than by the flux of the central component. This means that simply fitting the core radius to produce the flux of the central component can produce a qualitatively and quantitatively worse model than fitting the core radius to get the best average fit to the extended structure. Second, any misinterpretation of component D does not significantly bias the results of our current models, because the flux density of D is not a major contributor to the constraints.

Our results indicated that the constraints from the geometry of the system require the best fit model to have a substantial core radius, so there must be a central lensed image. The flux density of the central lensed component predicted by our best fit 8 GHz lens model is 266 μJy, which is within the one standard deviation error of the measured flux density of D (320 ± 60 μJy). Therefore, we believe that most, if not all, of the flux of D is a lensed image and not emission from the lensing galaxy (see also the discussion in Chapter 3).

5.2 5 and 15 GHz Model

Gravitational lenses are achromatic so the models that fit the 8 GHz data should also fit the 5 GHz and 15 GHz data. We model the 5 and 15 GHz maps using the same procedure outlined in Chapter 4. However, we explored only the α models (because they have the greatest potential for producing a different result) and limited the analysis to models that produce acceptable fits to the 8 GHz data (i.e., model with α in the range $0.4 \lesssim \alpha \lesssim 0.8$). The target χ^2 values for the 5 and 15 GHz maps are 15 and 153 respectively. The two questions we want to examine are whether the

Table 5.6: 5 GHz α models

α	$\Delta\sigma$	χ^2_{tot} [a]	χ^2_{mult} [a]	b	x_l [b]	y_l [b]	γ	θ_γ	s
	μJy			"	"		degrees	"	
0.4	639	449	392	0.936 ± 0.003	0.020 ± 0.016	0.007 ± 0.011	0.145 ± 0.011	-25.9 ± 1.1	0.258 ± 0.004
0.5	605	407	363	0.921 ± 0.005	0.022 ± 0.012	0.004 ± 0.014	0.133 ± 0.015	-25.8 ± 1.7	0.196 ± 0.004
0.6	682	473	418	0.921 ± 0.005	0.026 ± 0.020	0.005 ± 0.012	0.130 ± 0.011	-26.1 ± 1.7	0.162 ± 0.018
0.7	707	531	481	0.916 ± 0.009	0.022 ± 0.017	0.009 ± 0.022	0.125 ± 0.012	-26.4 ± 2.6	0.120 ± 0.029
0.8	739	541	526	0.909 ± 0.006	0.019 ± 0.022	0.015 ± 0.012	0.115 ± 0.009	-26.5 ± 1.8	0.018 ± 0.004
0.5 ± 0.3	605	407	363	0.921 ± 0.016	0.022 ± 0.012	0.004 ± 0.015	0.133 ± 0.018	-25.8 ± 0.6	0.196 ± 0.058

a $N_{dof} = 15$.
b The position of the lens is given relative to the position of component D at 8 GHz

best fit 8 GHz model is also the best fit model for the images at the other two frequencies and whether the model is a good fit at the other two frequencies. Table 5.6 summarizes the converged models for the 5 GHz map, and Figure 5.13 shows the renormalized χ^2 as a function of α. The dotted line represents the $\Delta\chi^2_r = 15.1$. The best fit model has $\alpha = 0.5 \pm 0.3$. All parameters of the models $0.4 \lesssim \alpha \lesssim 0.8$ obtained from the 5 GHz map are consistent with the ones obtained from the 8 GHz map. The absolute values of the χ^2 - 407 (363) in the total (multiply) image region - are again much larger than the target value of $N_{dof} = 15$ by 71 (63) standard deviations. The peak residuals in the map are only $< 3.8\%$ of the peak brightness. However, this is still considerably larger than expected because the 5 GHz map has a rather large

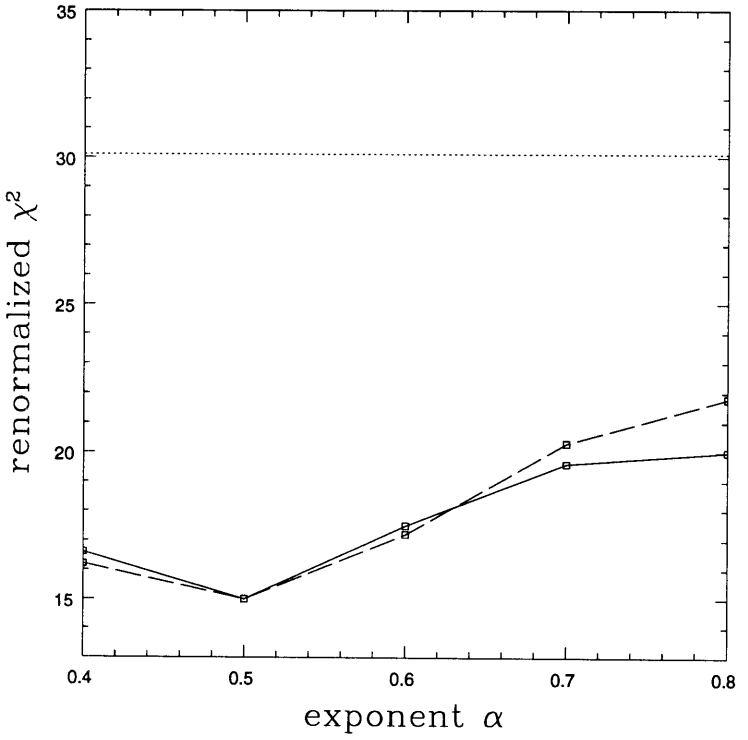


Figure 5.13 – The renormalized χ^2 as a function of α for the 5 GHz α models. The solid line shows the renormalized χ^2_{tot}, and the dashed line shows the renormalized χ^2_{mult}. The dotted horizontal line shows $\Delta\chi^2_r = 15.1$.

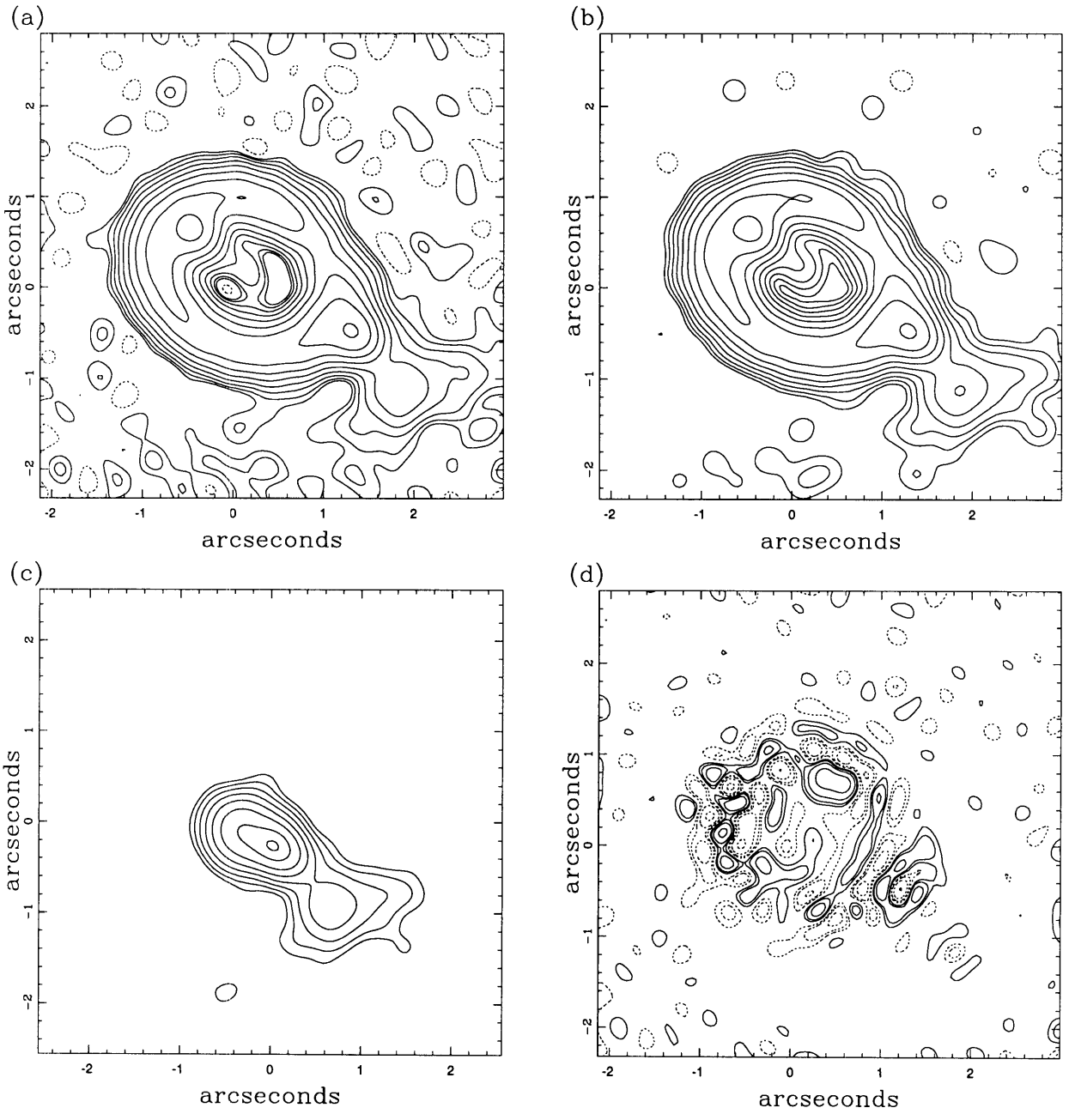


Figure 5.14 – (a) The observed image, (b) the reconstructed image, (c) the inferred source, and (d) the residual map of the best 5 GHz α model. The contour levels are: (a) and (b) -1, 1, 2, 4, 8, 16, 32, 64, 128, and 256 $\times$ 60 μJy, (c) -1, 1, 2, 4, 8, 16, 32, 64, and 95% $\times$ 155μJy, (d) -8, -4, -2, -1, 1, 2, 4, and 8 $\times$ 60μJy.

signal-to-noise ratio. Figure 5.14 shows the residual map, the inferred source, and the reconstructed image for the best fit 5 GHz α model. We again find a model that to the eye looks perfectly acceptable, but can be formally ejected by the goodness-of-fit criteria.

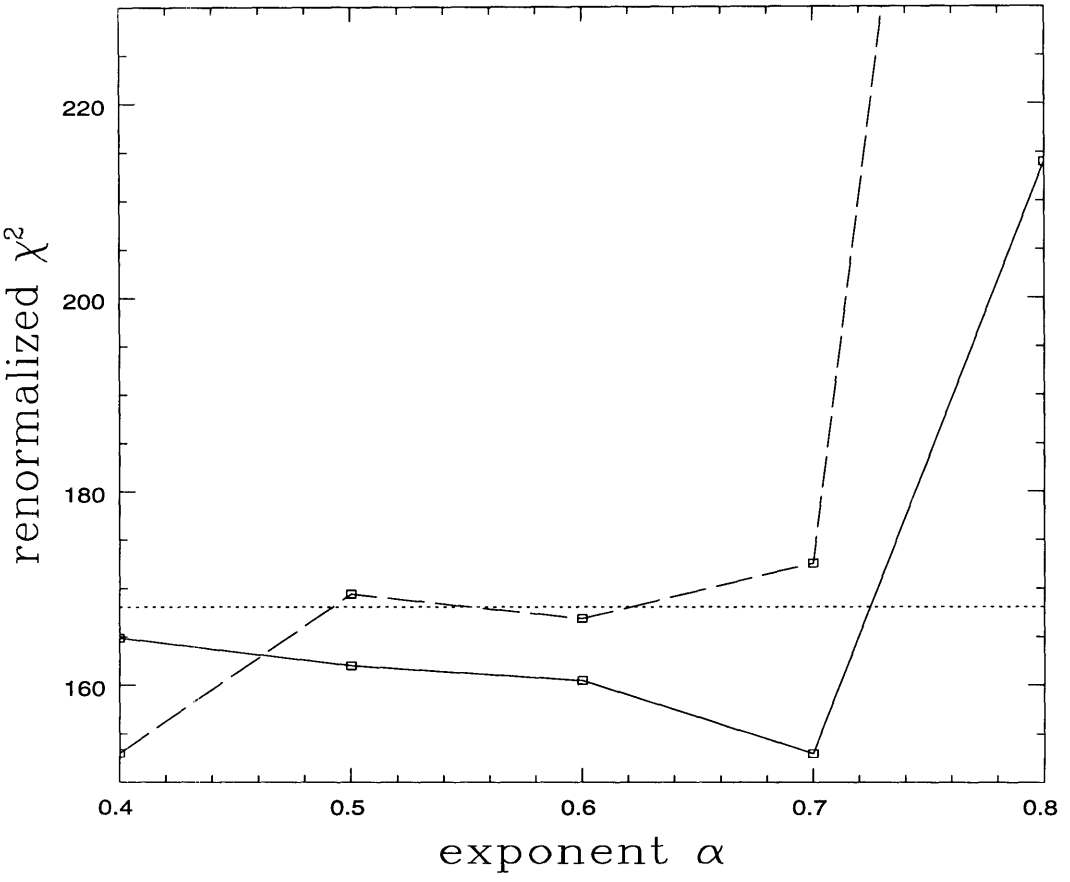
All converged lens parameters are consistent with the ones in the 8 GHz models given the estimated uncertainties. Since there are no strong compact sources in the map that could be used as position standards, the modeled lens positions are expressed relative the absolute position of component D detected at 8 GHz. The lens position is systematically shifted by 0.″02 in the x coordinate and 0.″01 in the y coordinate. The 5 and 8 GHz observations were taken on different dates and calibrated with different phase calibrators, leading to positional uncertainties on that order. Thus, the shifts are likely due to the problems in registrating the two images. The systematic uncertainties in registering the different maps are larger than the uncertainties in the lens positions from the inversions of the independent maps. This highlights the difficulty of attempting simultaneous models of several different maps using LensClean (Kochanek & Narayan 1992) since one would have to introduce two additional free parameters in the model to correct for the positional offsets between these images.

Table 5.7 summarizes the converged models for the 15 GHz map, Figure 5.15 shows the residuals as a function of α, and Figure 5.16 shows the residual map, the inferred source, and the reconstructed image of MG1131 at this frequency. Once again we find that the lens parameters are consistent with those found in the 8 GHz modeling. The absolute values of the χ^2 statistic - 1047 (427) in the total (multiply) image region - are again much higher than the target value of $N_{dof} = 153$.

5.3 Two Nearby Galaxies

Two fainter galaxies near the MG1131 position were detected by both Annis

(1992) and Larkin et al. (1994). The two galaxies are both approximately 3 arcseconds from the lens position. The redshifts of these two galaxies are currently unknown. If they are at the same redshift as the lens, they, being so close to the system, could make significant contributions to the formation of the lensed image. The structure of our lens models does not explicitly include the effects due to these two galaxies, so we would like to investigate whether our model parameterization is adequate to account for the effects due to these two galaxies.



Figures 5.15 – The renormalized χ^2 as a function of α for the 15 GHz α models. The solid line shows the renormalized χ^2_{tot}, and the dashed line shows the renormalized χ^2_{mult}. The dotted horizontal line shows $\Delta\chi^2_r = 15.1$.

Table 5.7. 15 GHz α models

α	$\Delta\sigma$	χ^2_{tot} [a]	χ^2_{mult} [a]	b	x_l [b]	y_l [b]	γ	θ_γ	s
	μJy			$''$	$''$		degrees	$''$	
0.4	559	1128	379	0.938 ± 0.004	0.024 ± 0.004	0.004 ± 0.003	0.133 ± 0.006	-26.6 ± 0.4	0.254 ± 0.003
0.5	597	1107	420	0.922 ± 0.008	0.021 ± 0.006	0.002 ± 0.005	0.132 ± 0.011	-25.8 ± 0.8	0.196 ± 0.005
0.6	421	1059	403	0.925 ± 0.005	0.049 ± 0.008	-0.015 ± 0.007	0.115 ± 0.011	-26.5 ± 0.6	0.172 ± 0.008
0.7	467	1047	428	0.914 ± 0.008	0.022 ± 0.004	0.003 ± 0.003	0.125 ± 0.002	-26.5 ± 1.1	0.148 ± 0.005
0.8	540	1465	902	0.907 ± 0.010	0.021 ± 0.010	0.002 ± 0.010	0.113 ± 0.009	-27.2 ± 2.2	0.018 ± 0.010
0.7 ± 0.2	467	1047	427	0.914 ± 0.010	0.022 ± 0.027	0.003 ± 0.018	0.125 ± 0.010	-26.5 ± 0.6	0.145 ± 0.048

a $N_{dof} = 153$.
b The position of the lens is given relative to the position of component D at 8 GHz

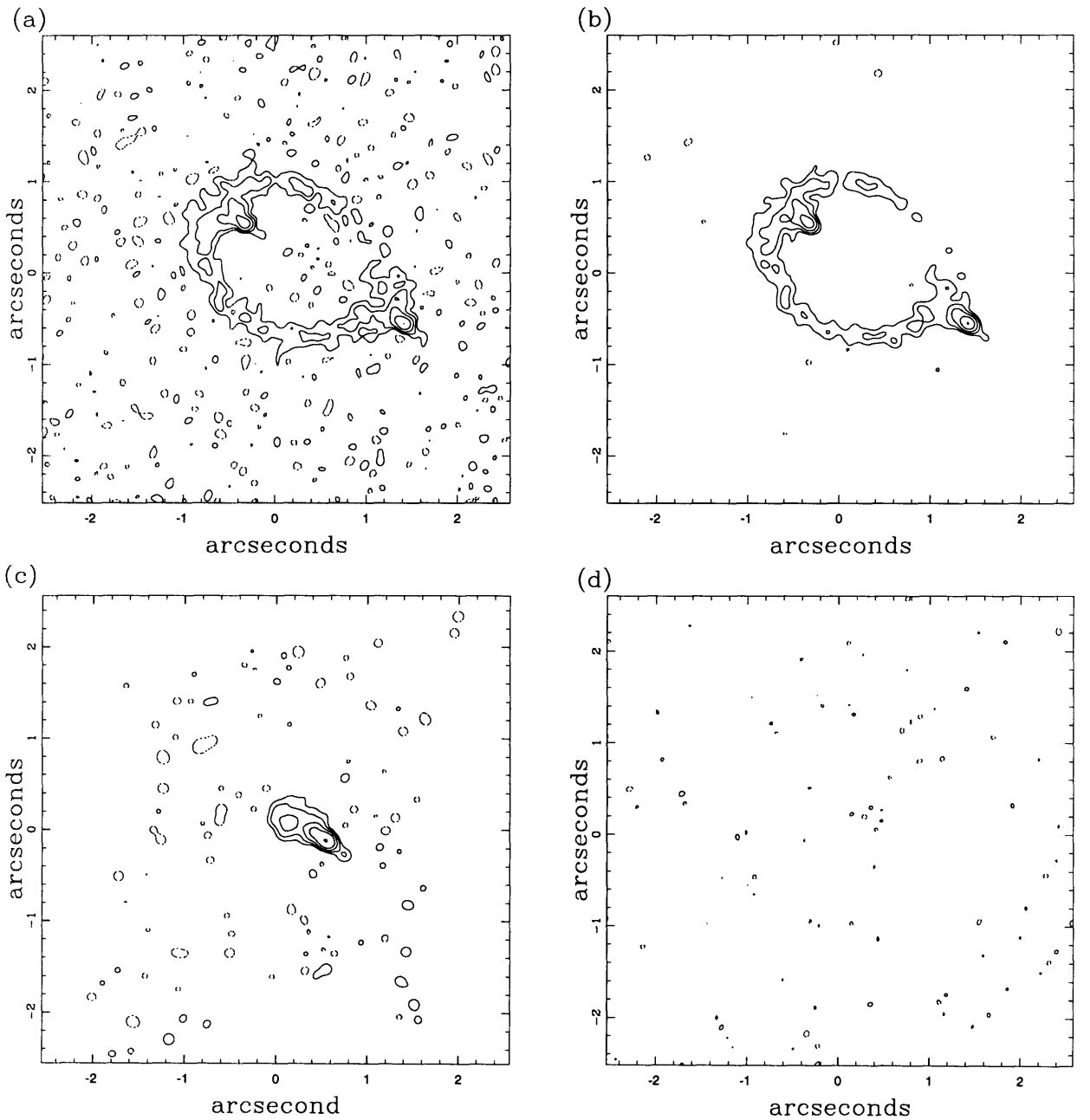


Figure 5.16 –(a) The observed image, (b) the reconstructed image, (c) the inferred source, and (d) the residual map of the best 15 GHz α model. The contour levels are: (a) and (b) -2, 2, 4, 8, 16, 32, 64, 128, and 256 $\times$ 60 μJy, (c) -4, 4, 8, 16, 32, 64, and 95% $\times$ 62 μJy, (d) -4, -2, -1, 1, 2, and 4 $\times$ 130μJy.

Annis (1992) named the galaxies C and D, but we rename them G1 and G2, respectively, to prevent confusion with the radio components in the system. We will call the primary lens galaxy L. Relative to the best fit lens position, G1 is at (-0.″09, -3.″15) and G2 is at (2.″04, 1.″89). The K magnitudes of L, G1, and G2 are 16.8, 20.8 and 21.2 respectively (Larkin et al. 1994). Assuming G1 and G2 are at the same redshift as L, the magnitude differences and the Faber-Jackson (1976) relation predict the ratios of the critical radii of G1 and G2 to L to be $b_{G1}/b_L = 0.16$ and $b_{G2}/b_L = 0.13$. This means that G1 and G2 have a perturbative effect on the lens model, and we can expand the potentials of G1 and G2 as a power series centered on the primary lens galaxy L. The expansion of the perturbing potential of each galaxy to second order in b_G/r_G is

$$\phi_e = \text{constant} + a_1 x + a_2 y + \frac{1}{2}\kappa_e r^2 + \frac{1}{2}\gamma_e r^2 \cos[2(\theta - \theta_e)]. \qquad (5-1)$$

The constant and linear $(a_1 x + a_2 y)$ terms have no effects on the model, so the first terms that modify the lens model are the convergence term $(\kappa_e r^2/2)$ and the shear term $(\gamma_e r^2 \cos[2(\theta - \theta_e)]/2)$. If the perturbing galaxies have the same monopole structure as the lens galaxy and a small core radius $(s/r_G \ll 1)$, then the convergence and shear from one of the two galaxies are

$$\kappa_G = \frac{\alpha_G}{2}\left(\frac{b_G}{r_G}\right)^{2-\alpha_G} \qquad \text{and} \qquad \gamma_G = \left(1 - \frac{\alpha_G}{2}\right)\left(\frac{b_G}{r_G}\right)^{2-\alpha_G}, \qquad (5-2)$$

where α_G and b_G are the monopole lens parameters associated with the external galaxy, and r_G is the distance from the center of the principle lensing galaxy to the center of the external galaxy. If we continue the expansion of the effects of the perturbing galaxies, the next order terms are smaller by another factor of b_G/r_G compared to the convergence and shear terms. Thus, they would change the deflections of rays by at most 0.″002. This is smaller than the changes in the deflections of rays produced by our typical errors on parameters (see Table 5.2), so, within our current modeling ability, it is an unmeasurable perturbation.

The convergence term ($\kappa_e r^2/2$) has no observable consequences in a lens model; it simply rescales the parameters of the models (Alcock & Anderson 1985, 1986, Gorenstein, Falco & Shapiro 1988). In general, the true critical line of the lens galaxy $b_L = (1 - \kappa_e)b$ where b is the parameter fitted in the lens models. Since b_{G1}/b_L and b_{G2}/b_L are both small, the convergence produced by G1 and G2 at the ring is also small. If the mass distributions in G1 and G2 were to follow the isothermal sphere model, the convergence would be $\kappa_e \simeq 0.05$. If the mass distributions of these two perturbing galaxies are more centrally concentrated, the convergence due to the two external galaxies would be even smaller. Thus the convergence from the nearby galaxies introduces only a small rescaling of the lens parameters.

The structure of the shear term produced by the two galaxies is identical to the external shear we use in the models. The superposition of three separate external shears for the main lens galaxy and the two perturbing galaxies is simply an external shear model with a different ellipticity and position angle. Thus, the quadrupole structure of our models should be adequate in accounting for the primary effects produced by G1 and G2. In fact, the quadrupole structure we assumed is in many ways a better model for the effects of the perturbing galaxies than for the primary lens.

The orientation of the perturbing shear [see Eqn (5-2)], which depends only on the mass ratio and relative positions of G1 and G2, has $\theta_e = -23°$. This is only 3° from the position angle of the shear in our lens models (see Table 5.2) - a remarkable coincidence! Could G1 and G2 be the main cause of such an elliptical ring (if they both are at the same redshift as the lens)? The strength of the shear depends somewhat on the form of the monopole, with $\gamma_e = 0.030$ for an isothermal model and $\gamma_e = 0.003$ for a point mass model (assuming the Faber-Jackson relation). Both values are much smaller than the shear of $\gamma \simeq 0.13$ needed to fit the ring in our models. Therefore, unless our estimate of their masses relative to L is in error, G1 and G2 do not produce

the shear required to model the ring in MG1131.

The coincidence between the modeled position angle of the external shear and that computed for G1 and G2 prompted us to explore scenarios in which the shear can be dominated by the external galaxies. Any such scenario would require increasing the influence of G1 and G2 relative to the principle lensing galaxy. In Figure 5.17 we show the effects on the critical lines and on the inferred shear at the ring radius when the luminosity difference between G1 and L (and also G2 to L) is reduced by factors of 25 and 100. We do this for both point mass and isothermal lens models, rescaling the parameters of the primary lens galaxy to account for the convergence introduced by G1 and G2. We see that the observed K band luminosity ratios must be in error by a factor of 25 for the isothermal models before the shear from G1 and G2 is large enough to match that required by the models. For the point mass model, a factor of $\sim$ 100 in error is needed to make G1 and G2 be the main contributors to the shear needed to fit the ring.

If the Faber-Jackson (1976) relation holds, then changing the relative lensing strength of G1/G2 and L must come either from reducing the luminosity of L or increasing the intrinsic luminosity of G1 and G2. The luminosity of L is inferred by subtracting the contributions from the lensed images of the compact radio components A and B. The optical residual is lumpy and ill-formed (Larkin et al. 1994) and Hammer et al. (1991) claim to see the ring in which case more of the flux of L might be from the source rather than the lens. Nonetheless, it is hard to see how the estimates could be off by a factor of 25 in luminosity. The infrared images of Larkin et al. (1994) also indicate that the MG1131 lens galaxy is dusty. If extinction by L is dimming the light from G1 and G2, the amount of dust required is large, and preserving the position angle of the external shear would require the unlikely coincidence that G1 and G2 have the same extinction. However, the position angle is not a sensitive function of the mass ratio between G1 and G2, and

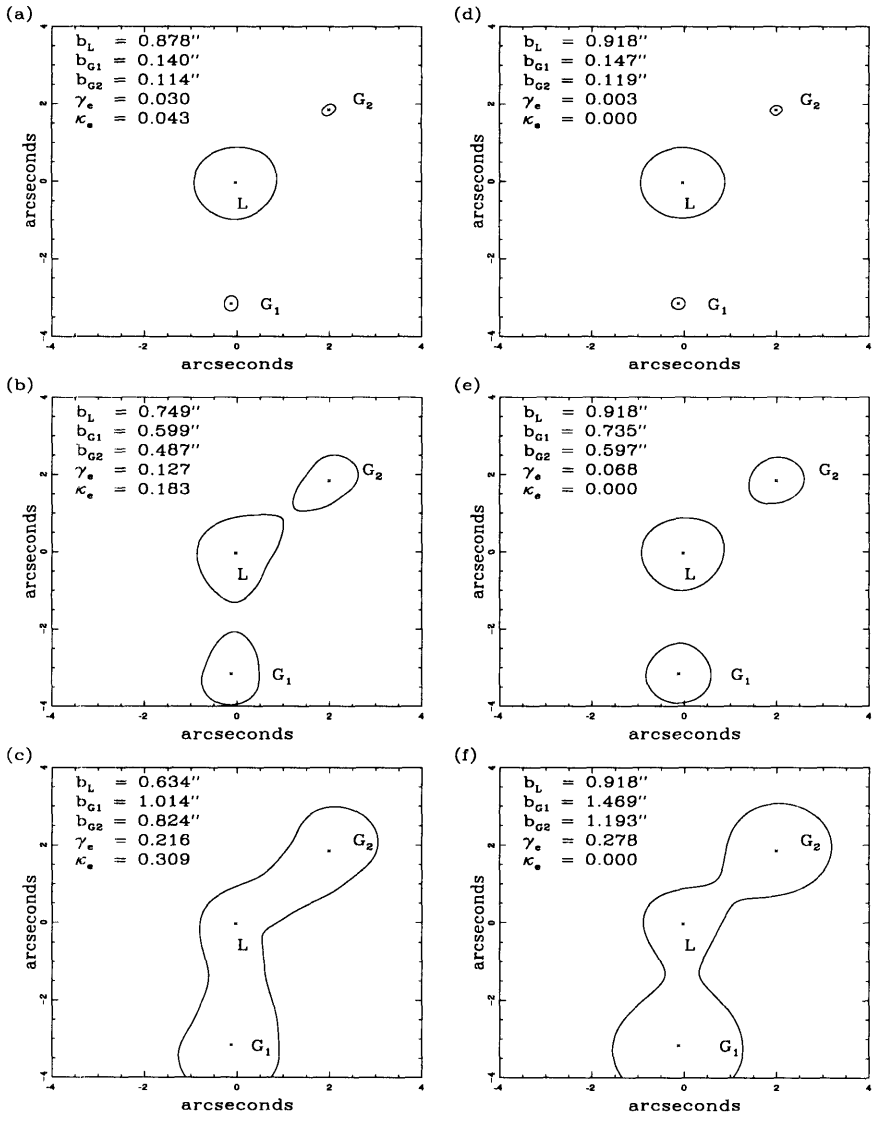


Figure 5.17 – The critical lines produced by various combinations of b_L, b_{G1}, and b_{G2} in an isothermal lens potential ((a) to (c)) and a point mass lens potential ((d) to (f)). The values of b_L, b_{G1}, and b_{G2} are chosen so that, by using the Faber-Jackson relation, the ratios of the intrinsic luminosity between G1 and L (also G2 and L) are exactly as observed in (a) and (d), 25 times as observed in (b) and (e), and 100 times as observed in (c) and (f). In all cases, the value of b_L is chosen so that $b \sim 0\farcs918$ (the best fit value of the isothermal sphere model). The magnitudes of the shear (γ_e) and the convergence (κ_e) due to the presence of G1 and G2 are also listed. If the masses of G1 and G2 are large enough, the magnitudes of the shear from G1 and G2 will exceed the value required to fit the ring ($\gamma = 0.13$).

it would differ by less than 10° for magnitude differences as large as 1.8. Finally, if G1 and G2 were at a different redshift than L, then the estimates would also be incorrect. It is very hard to evaluate this possibility because of the complete absence of redshift information on L, G1, G2, and the source. A surrounding cluster could easily generate the required shear, but if the shear is generated by a cluster there is no reason for it to be aligned with the predicted shear from G1 and G2.

5.4 Conclusion

Among the models we explored, there is a clear best model fitting the 5, 8, and 15 GHz images. In all cases, the peak residuals of the best fit model are smaller than 5% of the original peak of the map, and the reconstructed images are visually almost identical to the images observed. The best fit model is the α model, (which has a monopole potential with $\phi \propto (r^2 + s^2)^{\alpha/2}$) with $\alpha = 0.6 \pm 0.2$. The surface density profiles of the best fit model asymptotically declines as $r^{-1.4\pm0.2}$. Assuming that the external shear well represents the angular structure of the lensing potential, we find that an isothermal lens ($\alpha = 1$) with a surface density profile that asymptotically declines as r^{-2} is inconsistent with the data - the first example in which a lens cannot be modeled by the isothermal profile. The isothermal models fail to fit the lens because they cannot simultaneously fit the constraints from the compact components and the ring given the fixed quadrupole structure of the lens.

In the α models the core radius of the lens is a strong function of the asymptotic exponent α. Models with steeply declining density distributions $\alpha \lesssim 1$ require finite core radii that are a reasonable fraction of the ring radius. The dominant constraint on the core radius is not, in fact, the central image in the lens but the need to fit the extended structure of the ring and to locate properly the multiply imaged region. While the central image is a visually appealing source of constraints on the core radius, models that set the core radius to fit the central image can catastrophically

fail to fit the ring. The flux density of the central component of our best fit model is within one standard deviation of the measured flux of component D. Within the limits of our overall goodness of fit, we can rule out the hypothesis that component D is emission by the lens galaxy.

The mass interior to the average radius of the ring is insensitive to lens models. Assuming that $z_s = 2.0$ and $z_l = 0.5$, the integrated mass inside the average ring radius is $M(< \langle r \rangle) \sim 1.4 \times 10^{11}_{\odot}$. The accuracy of measurement of H_0 depends on the accuracies of both the time delay measurements and the lens modeling. Although we cannot determine the actual value of the time delay without knowing the redshifts (for a lens at $z_l = 0.5$ and a source at $z_s = 2.0$ in a Einstein-DeSitter cosmology, our best α model predicts a the delay that is approximately $58h^{-1}$ days, see Table 5.3), we can estimate the uncertainties in the time delay due to the models. The formal one standard deviation uncertainty in the time delay is $\pm 1\%$ for $\Delta\chi^2 = 1$ or $\pm 4\%$ for $\Delta\chi^2_r = 1$. We estimate a maximum uncertainty of $\pm 9\%$ from the limits when $\Delta\chi^2_r = 15.1$. The key observational problem in MG1131 is measuring the redshifts of the lens and the source.

We also fit de Vaucouleurs (1948) models to the 8 GHz image. We found that the best fit effective radius was $R_e = 0\farcs83 \pm 0\farcs13$, larger than the $R_e = 0\farcs5$ estimated from optical images of the lens. Unfortunately, we can not make further comparison between the modeled mass distribution and the light distribution due to the lack of uncertainties in the optical estimates. New optical data on the morphology of the image, with estimates of the uncertainties on fitted parameters like the effective radius would allow more quantitative comparisons of the fitted mass model to the optical properties of the lens. The best fit α models are considerably better than the best de Vaucouleurs models. The two models differ by 10 standard deviations in the χ^2, or one standard deviation in the rescaled χ^2_r.

The two galaxies G1 and G2, located about 3″ from the lens, perturb the lens

model, primarily by adding an additional external shear. The external shear produced by the two galaxies is only 3° from the axis of the external shear fitted in the model. This is a remarkable coincidence. The magnitude of the shear produced by G1 and G2 is almost an order of magnitude too small to produce the ellipticity of the ring. We estimated that the mass/velocity ratios estimated from the luminosities of the lens galaxy, G1, and G2 must be in error by a factor of 25 or more before G1 and G2 have enough mass relative to the primary lens galaxy to produce the observed ellipticity. There appears to be no plausible scenario to make the errors that large, so the alignment seems to be only a remarkable coincidence.

Although our models reconstruct the lensed image rather well, the χ^2 value of even the best fit model (~ 50 standard deviation above the $\chi = N_{dof}$) is still much larger than the N_{dof}. When we take the systematic effects caused by using the CLEAN map as the initial map for the inversion and by using the finite number of CLEAN components in LensCleaning the map, we can halve the value of the χ^2. However, even with these corrections, the χ^2 value is still much larger than expected. The large χ^2 value for the best fit model means that there is still substantial room for improvement in the models. The models we have tried assume that the ellipticity of the ring is generated solely by the matter well outside the ring and that the surface densities of the lens are circularly symmetric. Since the apparent isophotes of elliptical galaxies exhibit ellipse-like features (see Jedrzejewski 1987, for a review), it is very likely that potentials with circular symmetric surface densities are not sufficient in representing the true mass distribution of the lens galaxy. We suspect that a lens with true ellipsoidal isodensities plus an external shear would be the best model to try next. It is possible that an isothermal radial distribution will be consistent with the data if the angular structure of the model is changed.

References

Alcock, C. & Anderson, N. 1985, ApJ, **291**, L29

———————— 1986, ApJ, **302**, 43

Annis, J. A. 1992, ApJ, **391**, L17

de Vaucouleurs, G. 1948, Ann d'Astrophys., **11**, 247

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

Gorenstein, M, V., Falco, E. E. & Sharpiro, I. I. 1988, ApJ, **327**, 693

Hammer, F., Le Fèvre, O., Angonin, M. C., Meylan, G., Smette, A. & Surdej, J. 1991, A&A, **250**, L5

Kochanek, C. S. 1995, Ap. J., in press

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

Kochanek, C. S., Blandford, R. D., Lawrence, C. R. & Narayan, R. 1989, MNRAS, **238**, 43

Langston, G.I., Schneider, D.P., Conner, S., Carilli, C.L., Lehár, J., Burke, B.F., Turner, E.L., Gunn, J.E., Hewitt, J.N. & Schmidt, M. 1989, AJ, **97**, 1283

Larkin, J. E., Matthews, K., Lawrence, C. R., Graham, J. R., Harrison, W., Jernigan, G., Lin, S., Nelson, J., Neugebrauer, G., Smith, G., Soifer, B. T. & Ziomkowski, C. 1994, ApJ, **420**, L9

6. Probing the Structure of the Rotation Measure in Highly-Redshifted Galaxies with Gravitational Lenses

Valuable information on the Galactic **B** field structure has been obtained from studies of Faraday rotation in the linearly polarized light emitted from extragalactic radio sources when it passes through the Galaxy (e.g. Simard-Normandin & Kronberg 1980; Sofeu & Fujimoto 1983; Vallée & Bignell 1983; Simonetti, Cordes, & Spangler 1984, 1986; Lazio, Spangler & Cordes 1990). If there exists a source emitting linearly polarized light behind a galaxy of interest, the observer, at the appropriate position to detect the light passing through the galaxy, can get information about the **B** field in the galaxy by measuring the Faraday rotation the light ray underwent. The difficulty in this method is to to find a system in which there exists a light source behind the galaxy. Gravitational lenses are such systems. Most of the gravitational lens systems detectable at radio wavelengths are partially linearly polarized, so they can be used to study the structure of the rotation measure distribution in the lens galaxy. If the electron density distribution of the lens galaxy is known, the determination of the rotation measure in the lens galaxy (RM_L) would be valuable for understanding the magnetic field in the high-redshifted galaxy. In this chapter, we investigate the possibility of using the RM determined from MG1131 to probe the structure of the rotation measure distribution in the lens galaxy.

6.1 Determination of the Rotation Measure in MG1131+0456

In theory, if the intrinsic position angle (χ_o) of the polarized emission is not wavelength dependent and if there are no ambiguities in the observed position angle of the linearly polarized light (χ_{obs}), observations at two wavelengths should be sufficient for determining the rotation measure in any foreground Faraday screen. Let χ^1_{obs} and χ^2_{obs} be the observed position angles at λ_1 and λ_2 respectively. Then, the rotation measure in the foreground screen is simply $RM = (\chi^2_{obs} - \chi^1_{obs})/(\lambda_1^2 - \lambda_2^2)$. In reality,

because of the $\pm n\pi$‡ ambiguities in the observed position angle, the RM determined from observations at only two wavelengths is highly unreliable (see Figure 6.1 for illustration). The effects due to the $n\pi$ ambiguities can be removed if the χ_{obs} used to for the determination are sampled at sufficiently many wavelengths.

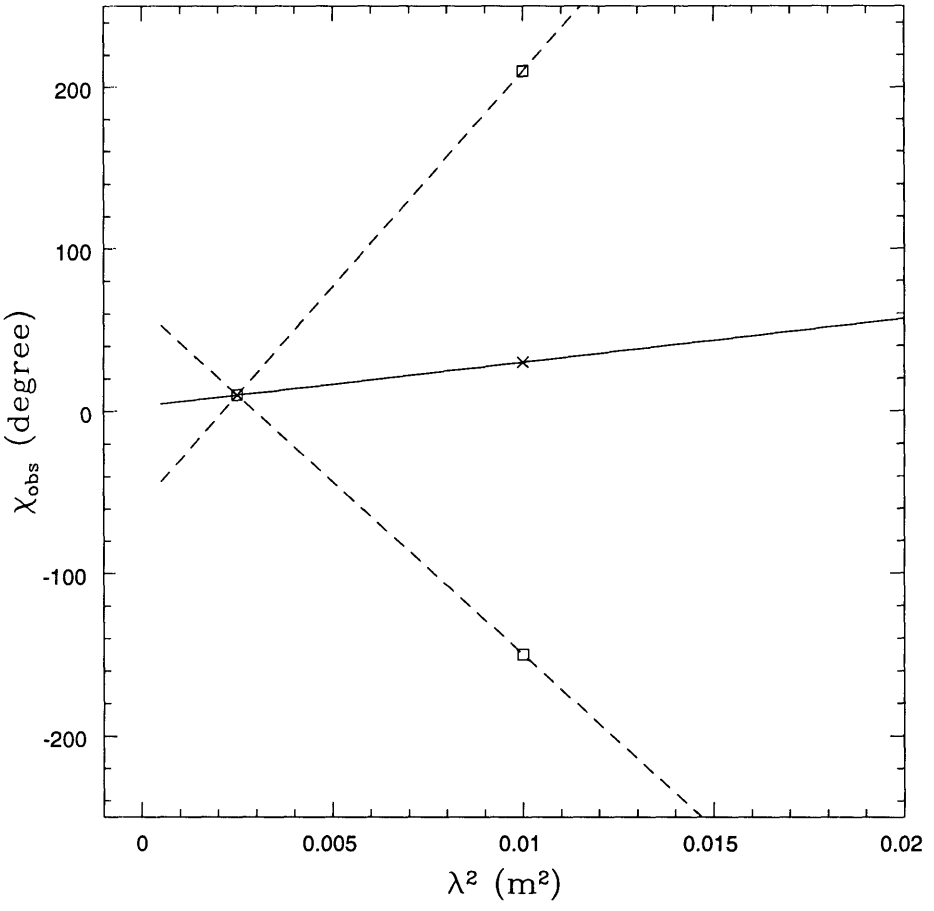


Figure 6.1 – The plot demonstrates the problems caused by the $n\pi$ ambiguities in the observed χ_{obs}. Let χ_1 and χ_2 be the two observed χ_{obs}'s (represented by the crosses). In principle, the slope of the straight line formed by χ_1 and χ_2 (shown by the solid line) is the RM in the foreground Faraday screen. However, since we can not distinguish between χ_2 (or χ_1) from $\chi_2 + \pm n\pi$ (or $\chi_1 + \pm n\pi$), there will be more than one RM capable of fitting the two data points (see the dashed lines).

‡ In practice, χ_{obs} is determined from the Q and U maps, where $Q = P\cos(2\chi_{obs})$ and $U = P\sin(2\chi_{obs})$. P is the total polarization intensity. The $n\pi$ ambiguities in χ_{obs} arise because both Q and U remain unchanged if we replace χ_{obs} by $\chi_{obs} + n\pi$.

We determined RM in MG1131 from the observations taken in March 1994 since they gave the best wavelength coverage. We treated the two IF's (see Chapter 2) in the 5 GHz band as two separate observations, so we had measurements of χ_{obs} at five wavelengths: 6.69, 6.62, 6.04, 5.98, and 3.55 cm. Before computing RM in the system, we used Monte Carlo simulations to examine whether the wavelength coverage provided by these datasets is sufficient to remove the effects due to the $\pm n\pi$ ambiguities. For the simulation, we first determined the Q and U flux densities of the source at each sampled wavelength (6.69, 6.62, 6.04, 5.98, and 3.55 cm) assuming that the intrinsic position angle of the polarized emission (χ_o^{true}) is 30°, the rotation measure is $RM^{true} = 200$ rad/m^2, and the total polarization strength of the source is 1 Jy at all wavelengths. Then, for each realization, we added Gaussian noise to each of the predetermined Q and U flux densities, giving the quantities which we named Q_{obs} and U_{obs}. The amplitude of the noise is chosen according to the signal-to-noise ratio (SNR) (i.e., the amplitude of the noise equals the amplitude of the signal divided by SNR). We found the best RM and χ_o fitted to Q_{obs} and U_{obs} by minimizing the quantity

$$\chi^2 = \sum_{i=1}^{N} \frac{(Q_{obs} - Q(\lambda))^2}{\sigma_Q^2} + \sum_{i=1}^{N} \frac{(U_{obs} - U(\lambda))^2}{\sigma_U^2}, \qquad (6-1)$$

$$Q(\lambda) = P\cos[2(\chi_o + RM\lambda^2)], \qquad (6-2)$$

and

$$U(\lambda) = P\sin[2(\chi_o + RM\lambda^2)]. \qquad (6-3)$$

The optimal RM and χ_o were found by searching through a two-dimensional space in the ranges of $0 < \chi_o < \pi$ and $-3000 < RM < 3000$ rad/m^2. We suspected that the question of whether the wavelength coverage is sufficient depends on the SNR, so we made three sets of simulations, assuming SNR = 2, 5, and 10. Each set of simulations contained 500 realizations.

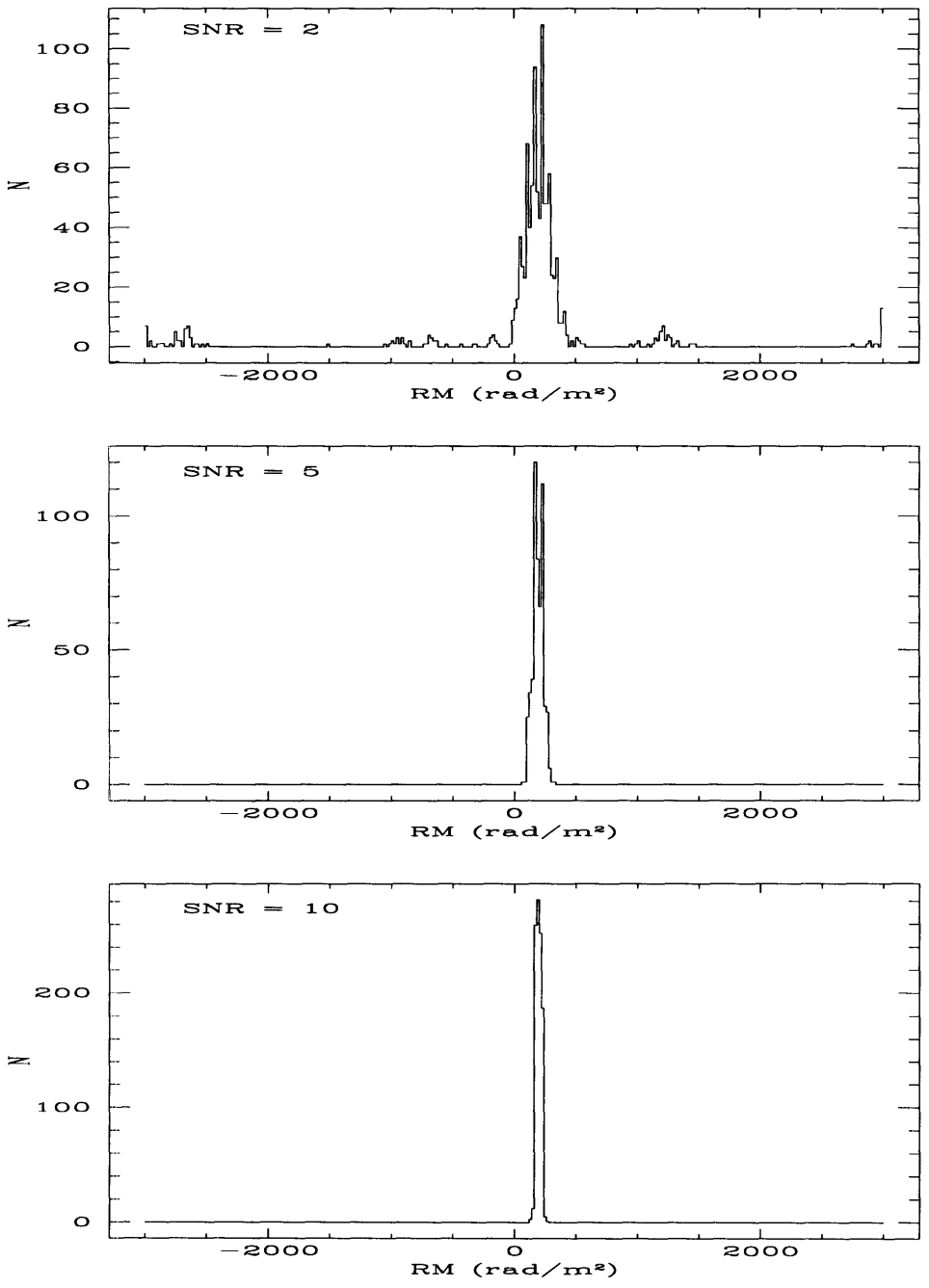


Figure 6.2 – The distribution of best fit RM obtained from the simulations generated assuming different SNR. The true value of the rotation measure is $RM^{true} = 200$ rad/m^2. When SNR = 5 or 10, there are no ambiguities in the best fit RM, and the values of the best fit RM agree with the true value. However, for SNR = 2, ambiguities in the determined RM appear. This indicates that our wavelength coverage is not sufficient to resolve the problems due to the $n\pi$ ambiguities for data points with SNR = 2.

The distribution of the best fit RM of each set of simulations is plotted in Figure 6.2. In the case where SNR equals 5 or 10, the best fit RM are all clustered around RM^{true} (200 rad/m^2), and the peak of the distribution is exactly at RM^{true}, indicating that our wavelength coverage is sufficient to resolve the problem caused by the $n\pi$ ambiguities if the data points have SNR> 5. We used the half-width-half-maximum (HWHM) of the distribution to estimate the uncertainty in the best fit RM. For SNR = 5, the uncertainty is $\pm$50 rad/m^2; for SNR = 10, the uncertainty is $\pm$25 rad/m^2. The situation is different for SNR = 2: although most of the best fit RM still cluster around the RM^{true} (200rad/m^2), there are substantial numbers of points appearing around 1300, -900, 3000, and -2500 rad/m^2, which correspond to the first and second order (i.e. $n = \pm 1$ and $n = \pm 2$) ambiguities. Similar results were obtained when different values of RM^{true} and χ_o^{true} were assumed. Our simulations confirm that the condition of "sufficient" wavelength coverage is indeed dependent on the SNR and that our wavelength coverage is only sufficient for data points with SNR> 5. Therefore, we excluded data points with SNR<5 from the determination of RM.

All datasets were calibrated and mapped with the procedures outlined in Chapter 2. Figures 6.3 - 6.7 show the position angle of the polarized emission in MG1131 at all five frequencies. We found the best fit RM by minimizing the χ^2 in Eqn (6-1). All maps were convolved to the resolution of the 6.69 cm map before the map of RM was computed. Since only data points with SNR> 5 were included in the RM calculation, the largest uncertainties in the determined RM should be $\pm$50 rad/m^2. A large variations in RM (ranging from -500 rad/m^2 to 500 rad/m^2) are seen in MG1131. Figure 6.8 shows a grey-scaled map of the distribution of RM in MG1131.

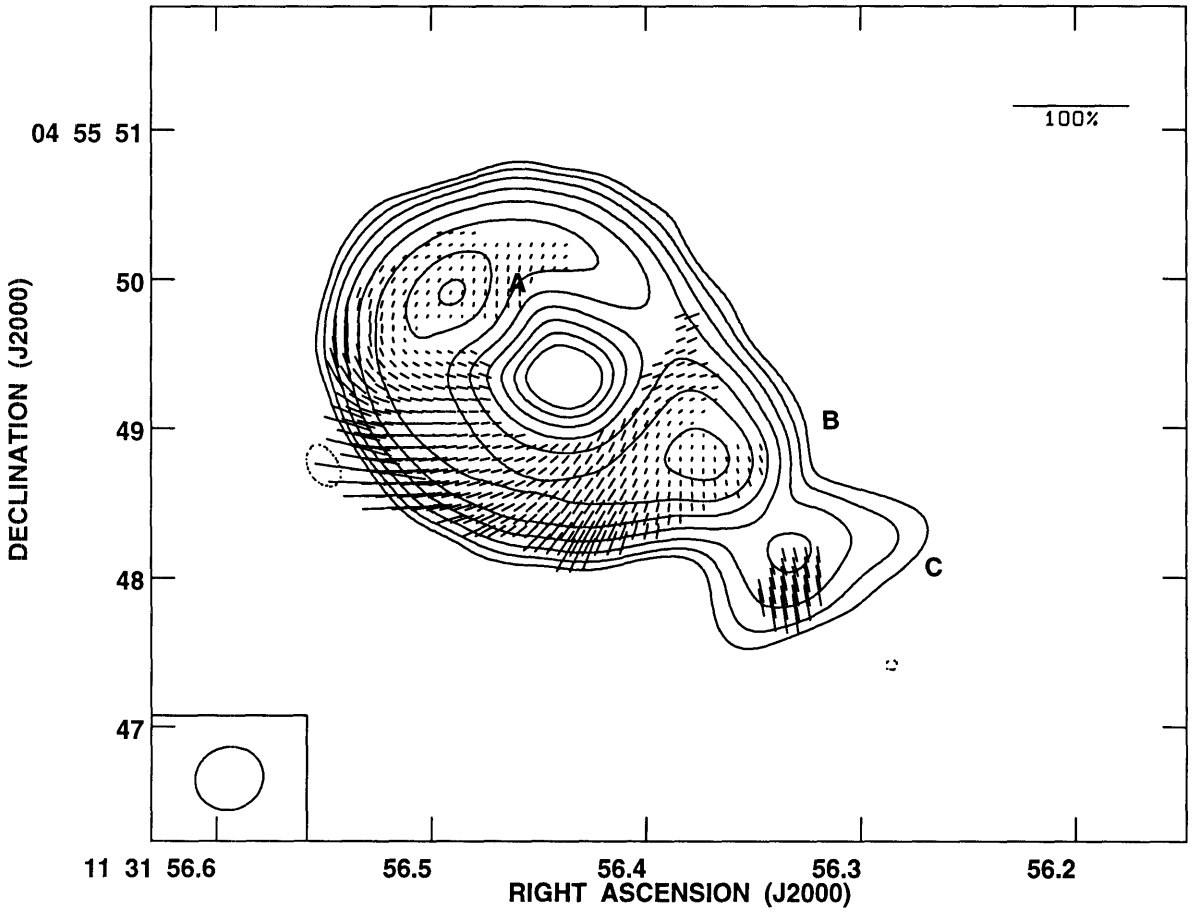


Figure 6.3 : The electric field vectors of the linearly polarized emission of MG1131 at 6.69 cm. The lengths of the vectors indicate the percent polarization strength of the emission. A scale of 100% is shown in the upper right hand corner. We superimpose the contour plot of MG1131 for the purpose of displaying the locations of the components in the system. The contour levels are -1, 1, 2, 4, 8, 16, 32, 64, and 95% of the peak.

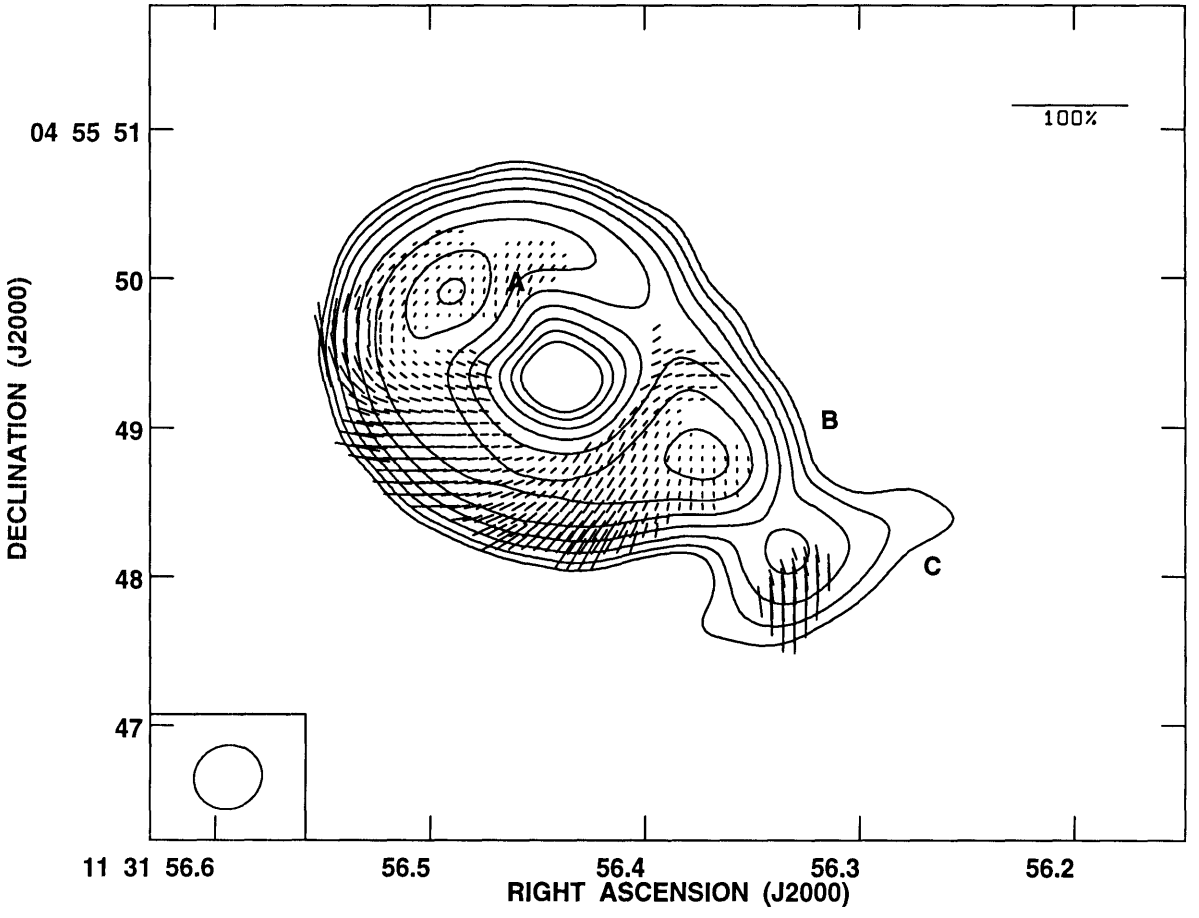


Figure 6.4 : The electric field vectors of the linearly polarized emission of MG1131 at 6.62 cm.

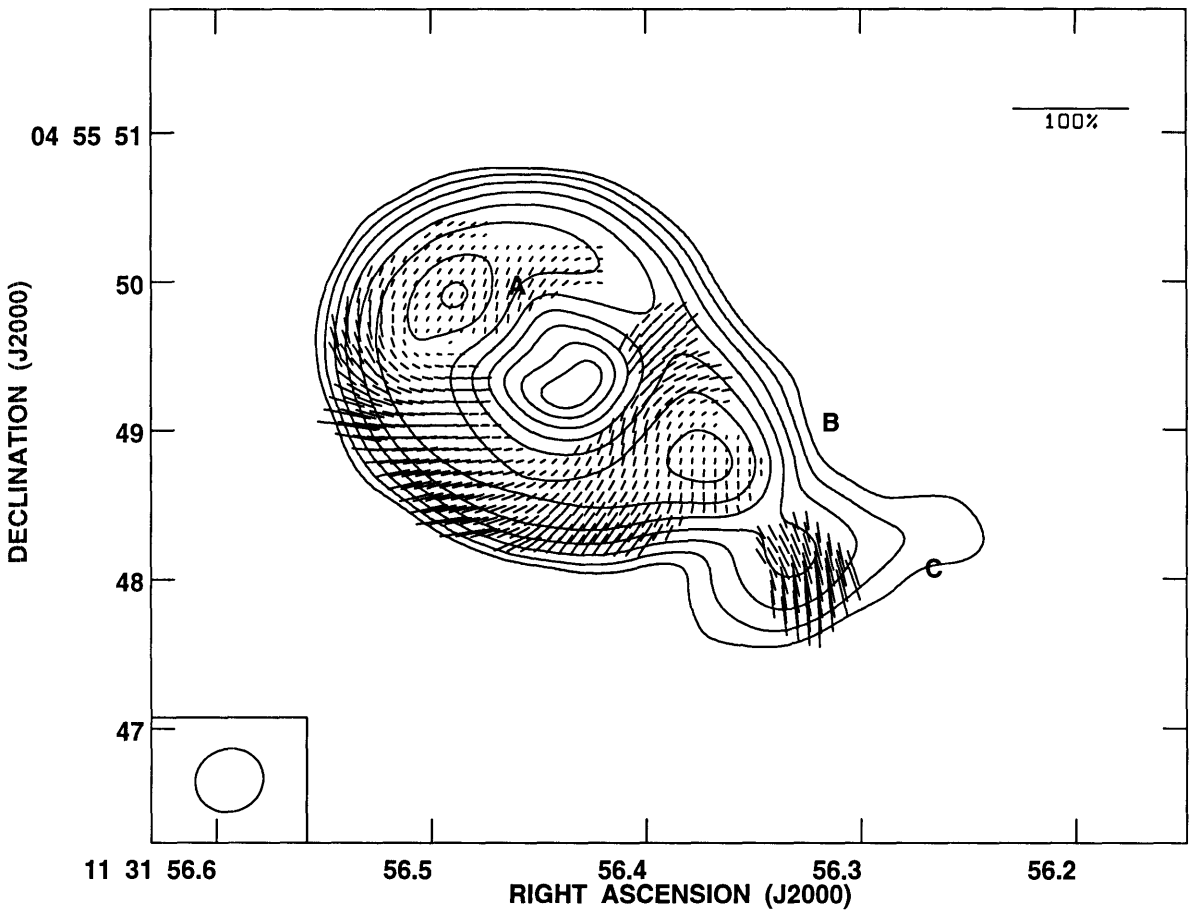


Figure 6.5 : The electric field vectors of the linearly polarized emission of MG1131 at 6.04 cm.

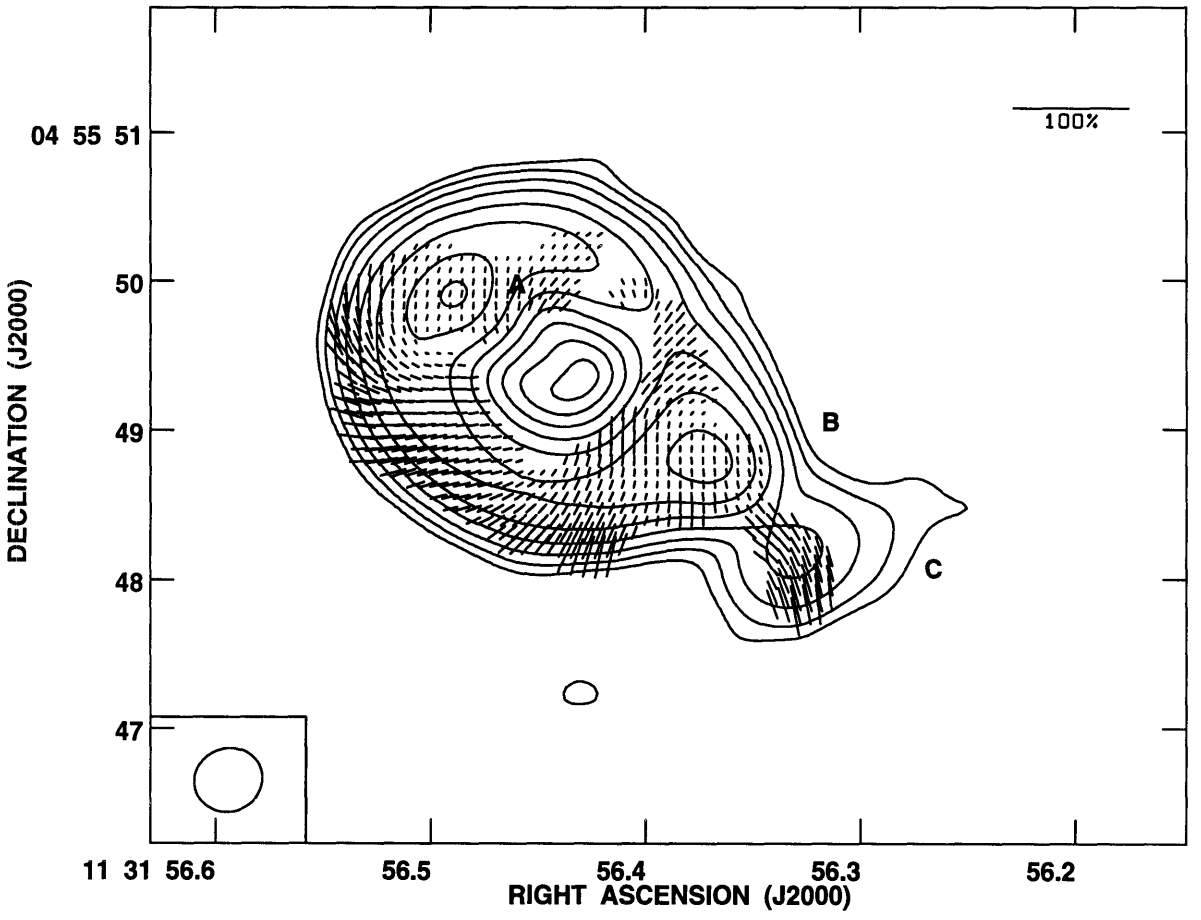


Figure 6.6 : The electric field vectors of the linearly polarized emission of MG1131 at 5.98 cm.

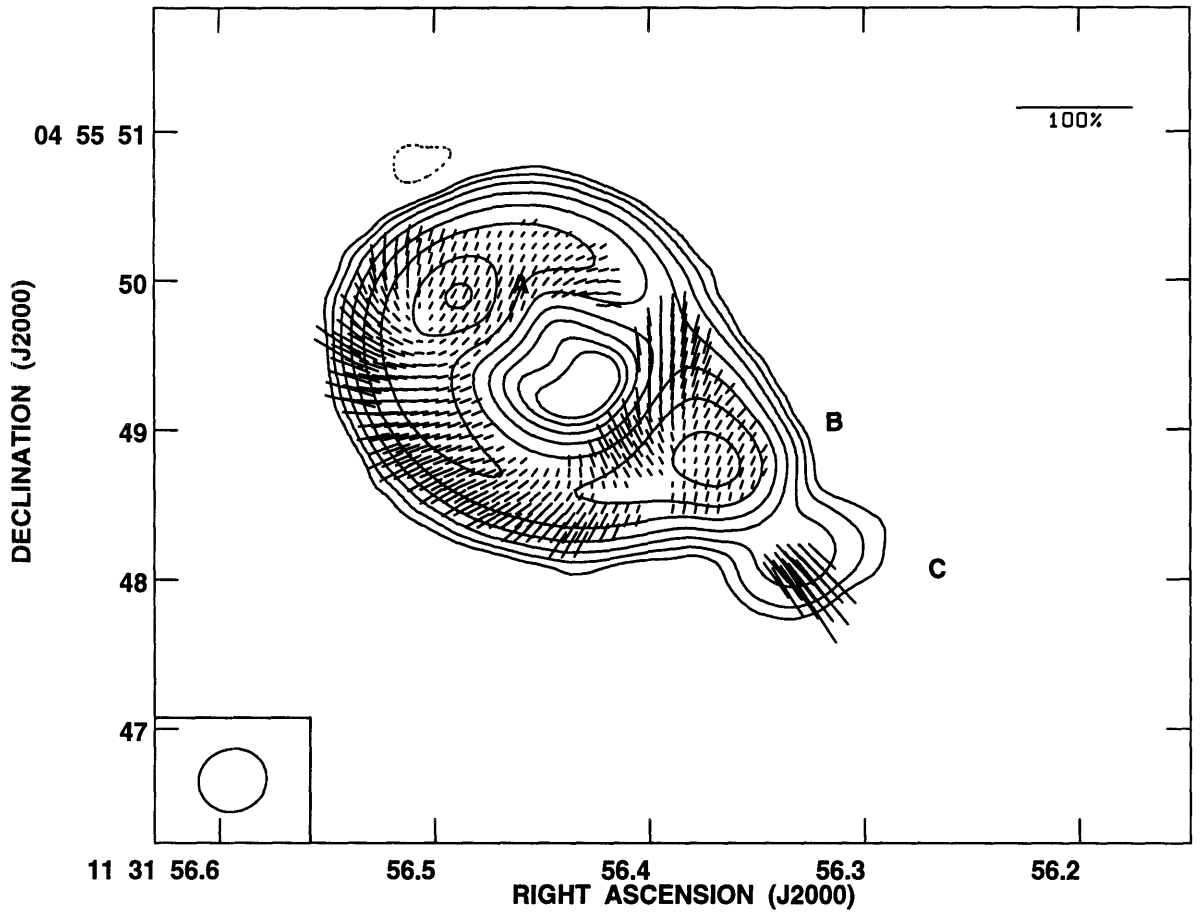


Figure 6.7 : The electric field vectors of the linearly polarized emission of MG1131 at 3.35 cm.

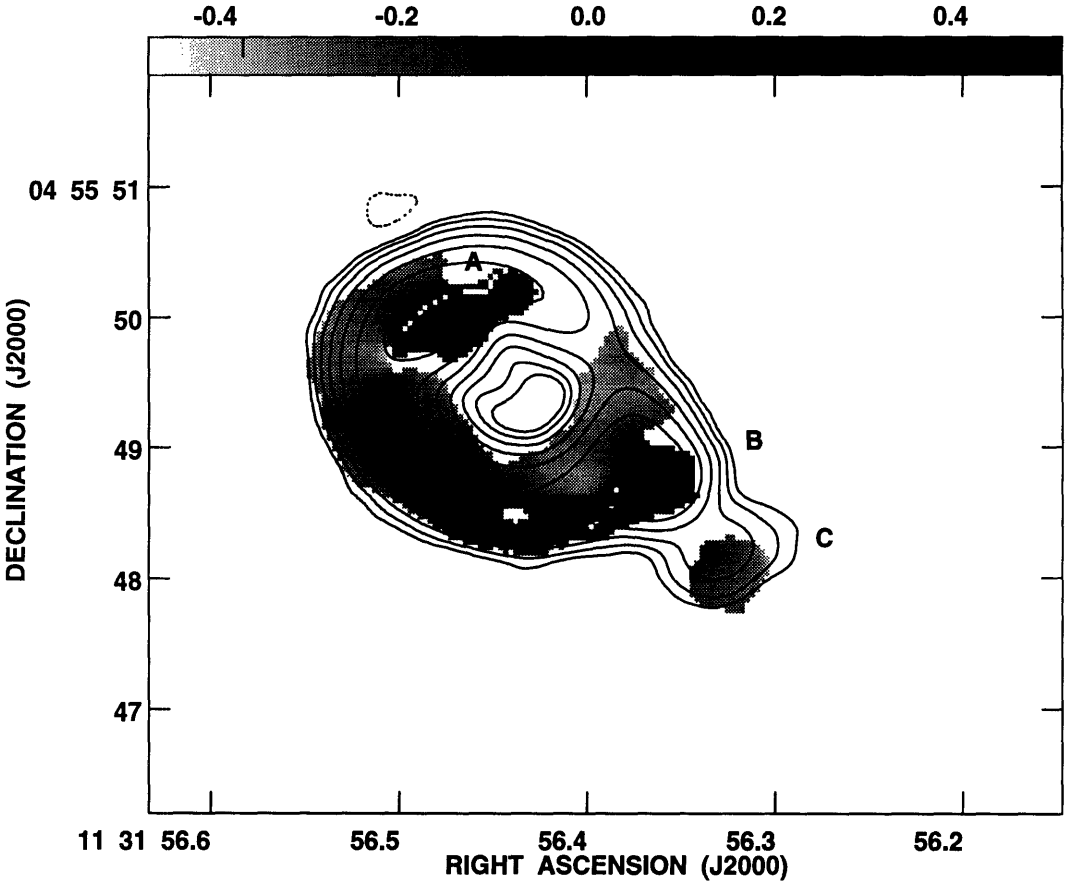


Figure 6.8 : The grey scale *RM* map determined from the polarization angles detected in MG1131. The scale on the top shows the grey scale in units of 1000 rad/m^2.

6.2 Physical Considerations and the ΔRM Measurements

In most known multiply imaged systems that are lensed primarily by a single galaxy, the images are formed at locations roughly 1″ from the the center of the lens. For a lens at $z_l = 0.5$, 1″ is $\sim$ 7 kpc. This means that the light rays of the background source in these systems must pass through the lens galaxy. If the source emits linearly polarized light, and if the lens galaxy has a magnetic field, when the light passes through the lens galaxy, its position angle must be changed due to Faraday rotation. Of course, the lens galaxy is not the only Faraday screen the light ray might encounter. There are four other possible Faraday screens: (1) the plasma within the emission region of the background source, (2) the plasma in the halo surrounding the background source, (3) the plasma in the intergalactic medium (IGM), and (4) the plasma in our Galaxy. Unfortunately, the only measurable quantity is the combined effect due to all screens.

Since the images of the same source should have the same intrinsic properties, we can eliminate the effects due to the plasma within the emission region of the source if we only consider the differences of the rotation measure detected in a pair of images of the same source (denoted as ΔRM). The calculation of ΔRM requires both the measurement of RM in the system and the knowledge of which part of image arising from the same source, so we determined ΔRM in MG1131 by first using our best lens model, the $\alpha = 0.63$ lens model (see Chapters 4 and 5), to identify pairs of images arising from the same source and then calculating the differences in the rotation measure between these pairs. We excluded component C from the calculation because the source of C is only singly imaged. Figure 6.9 shows $|\Delta RM|$ as a function of $|\Delta r|$ (the angular separation between a pair of images) on a log-log scale. The values of $|\Delta RM|$ range from 4 to 400 rad/m^2, and there is an interesting feature in $|\Delta RM|$ vs. $|\Delta r|$. When $\Delta r < 1\farcs5$, $|\Delta RM|$ increases monotonically as a

function of Δr; however, when $\Delta r > 1''.5$, $|\Delta RM|$ drops rapidly as a function of $|\Delta r|$.

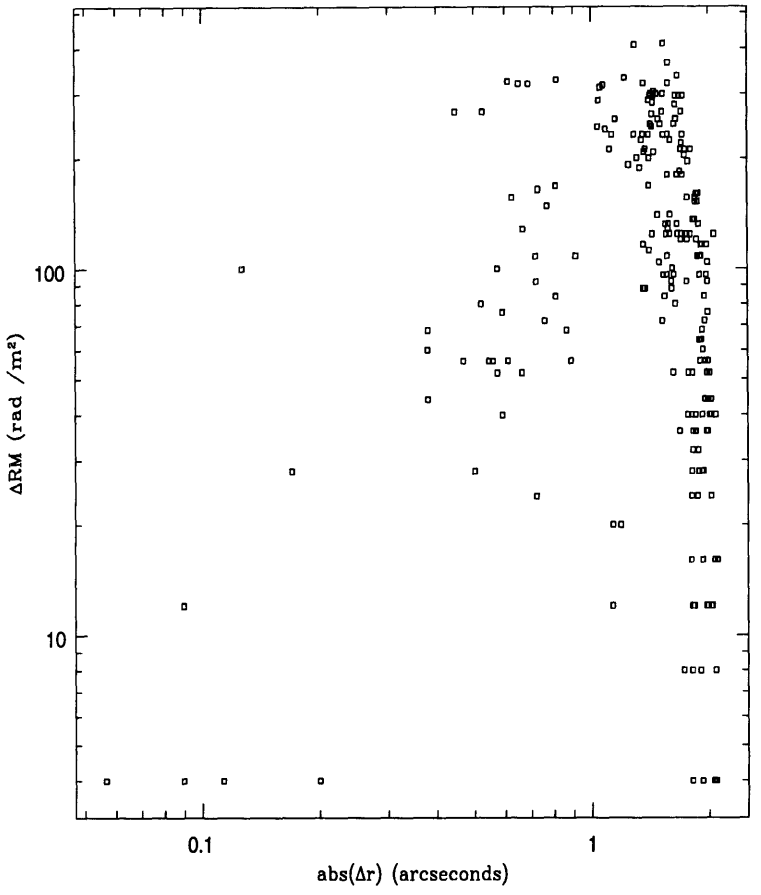


Figure 6.9 – A log-log scale plot of $|\Delta RM|$ vs. Δr (the angular separation between a pair of two images) determined from the $\alpha = 0.63$ lens model (see Chapters 4 and 5).

All other screens (i.e. the plasma around the source, the plasma in the IGM, and the plasma in our Galaxy) are external to the source, so, in general, the ΔRM we determined contains the contributions from all other screen. In other words,

$$\Delta RM = \Delta RM_S + \Delta RM_{IGM} + \Delta RM_L + \Delta RM_{Galaxy}, \qquad (6-4)$$

where ΔRM_s, ΔRM_{IGM}, ΔRM_L, and ΔRM_{Galaxy} are the difference in the rotation measure experienced by the light rays of a pair of images when they pass through

each medium. We examine the contribution due to the Galaxy, the plasma around the source, and the IGM to the measurement of ΔRM separately in the next following sections.

6.3 Effects due to the Galaxy

The rotation measure of our Galaxy has been intensively studied (see Beck & Kronberg 1989 for a review) so we can estimate the quantity ΔRM_{Galaxy} in Eqn (6-4). Figure 6.10 shows the ray geometry of a pair of images when their light rays enter the Galaxy. Let $\delta\theta$ be the angular separation between the two light rays and r be the distance between the observer to the top (or the bottom) of the Galactic disk in the direction of the source (as indicated in the figure). The linear separation between the two rays when they encounter our Galaxy is

$$\delta d = 0.00484\Big(\frac{r}{\text{kpc}}\Big)\Big(\frac{\delta\theta}{''}\Big) \text{ pc.}$$

To estimate the scale of δd, we need to know the typical values of r and $\delta\theta$. The value of r is difficult to estimate because it depends on the Galactic coordinates of the source. Assuming that the Galactic disk height is $\sim$ 0.1 kpc, then $r < 1$ kpc for a source at a galactic latitude $b > 3°$. The image separations detected in most of the multiply imaged gravitational lenses are $\sim 2''$. Thus, if we assume that $r \sim 1$ kpc and that $\delta\theta \sim 2''$, the linear separation between the rays when they encounter the Galaxy (δd) is roughly 0.01 pc. The change of the rotation measure in our Galaxy in a scale of 0.01 pc is $\ll 10$ rad/m^2 (Simonetti & Cordes 1986;

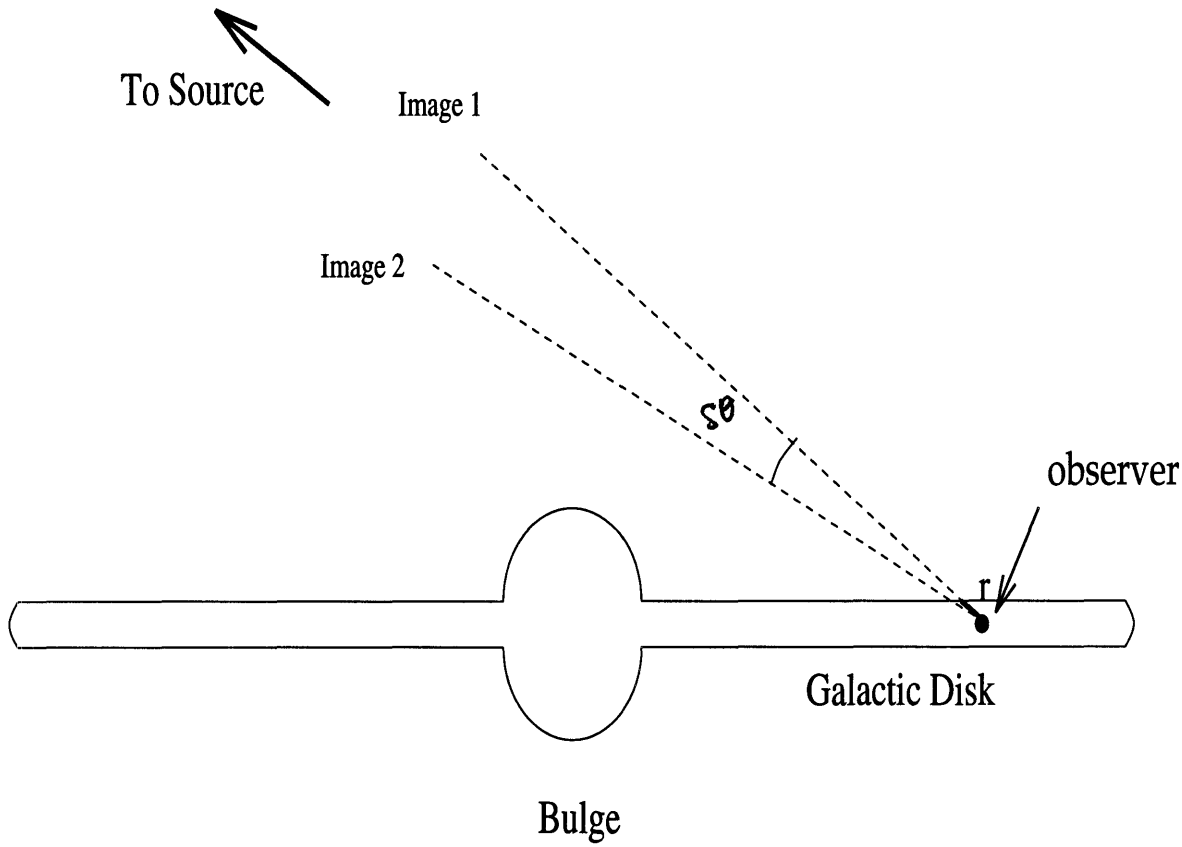


Figure 6.10: The geometry of the light rays from a pair of images entering our Galaxy. r is the distance from the observer to the outermost part of the Galaxy in the direction to the source, and $\delta\theta$ is the angular separation between the two light rays.

Laoz, Spangler & Cordes 1990); hence,

$$\Delta RM_{Galaxy} \ll 10 \text{ rad/m}^2 \qquad (6-5),$$

which is much smaller than the ΔRM detected in MG1131.

6.4 Effects due to the Plasma Around the Source

If the source is embedded in a halo† containing magnetized plasma, the halo

† Recent X-ray observations of luminous elliptical galaxies reveal the presence of a halo, consisting primarily of hot gas, surrounding each galaxy (see Fabbiano 1989 for a review). Since most of the host galaxies in the powerful double-lobe extragalactic radio sources are giant ellipticals (Kellermann & Owen 1988), it is likely that each powerful double-lobe radio source is surrounded by such a halo.

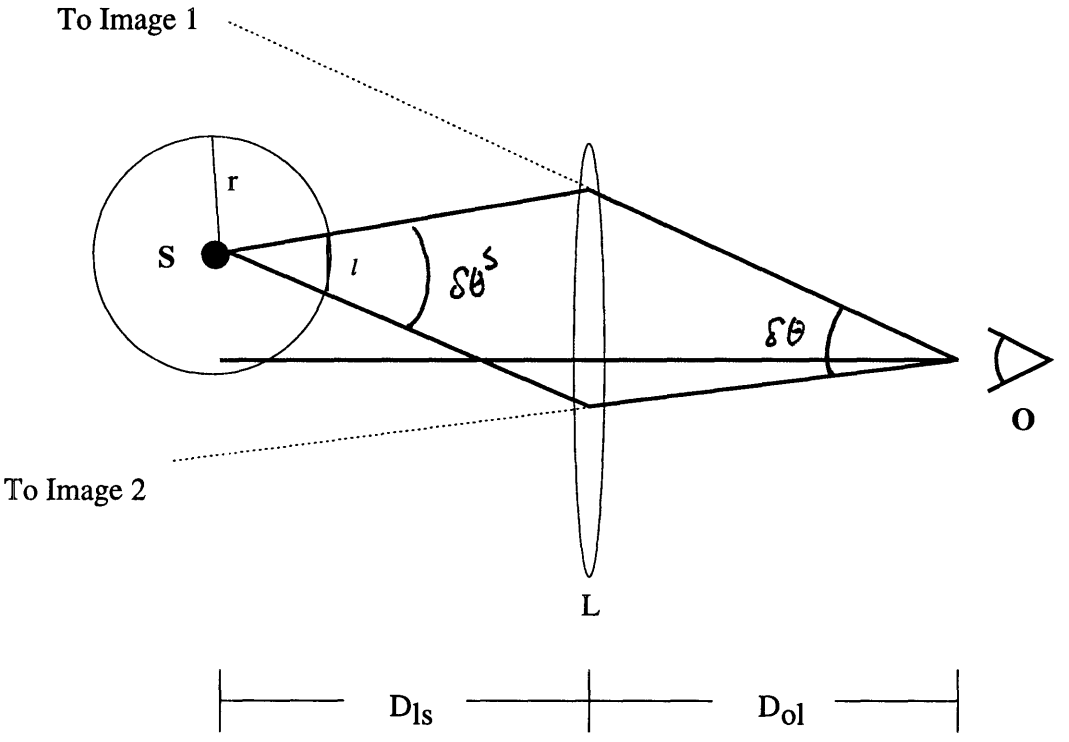


Figure 6.11 – A simple schematic showing the ray geometry of a source embedded in the halo being lensed into two images. r is the radius of the halo, $\delta\theta$ is the angular separation between the two images detected by the observer, $\delta\theta^s$ is the angular separation between the two light rays when they leave the source, and l indicates the linear separation between the two light rays when they leave the halo. Due to gravitational deflection, $\delta\theta^s \neq \delta\theta$.

becomes a natural Faraday foreground screen. Figure 6.11 shows the ray geometry for a point source embedded in a halo. Consider two rays which leave the source and are focused onto the observer O. Let $\delta\theta^s$ be the angular separation between the two rays leaving the source, and $\delta\theta$ be the angular separation detected by the observer O. Because of the gravitational deflection, $\delta\theta^s \neq \delta\theta$. Using the small angle approximation, we can relate $\delta\theta^s$ and $\delta\theta$ by

$$\delta\theta^s = \frac{D_{ol}}{D_{ls}}\delta\theta,$$

where D_{ij} is the angular diameter distance between i and j. Let r be the radius of the halo and l be the linear distance between the two light rays when they leave the

halo (as indicated in the figure). Then, the linear separation between these two rays when they leave the halo is

$$l = 0.00484 \frac{D_{ol}}{D_{ls}} \left(\frac{r}{\text{kpc}}\right) \left(\frac{\delta\theta}{''}\right) \text{ pc.}$$

If we assume that the redshifts of the source and the lens are $z_s = 2.0$ and $z_l = 0.5$ and that r is between 40 and 70 kpc#, then the value of l is in the range of 0.28 to 0.48 pc, which is an extremely small linear scale. This brings out an important point: by measuring the difference in RM between a pair of images in a gravitationally lensed system, we are probing the changes of RM in the halo on an extremely small scale.

We can evaluate the contribution of the halo to ΔRM (in Eqn 6-4) by asking whether it is possible for the RM in the halo to vary on such a small scale. When the source is not lensed, the telescope can only resolve the source structure that is larger than the angular resolution of the telescope. Thus, the observer can only detect the "averaged effect" of the source structure that is smaller than the resolution size. The resolution of most radio telescopes, when they are used to observe extragalactic sources, is much larger than 0.4 pc. Thus, if RM in the halo indeed varies on a scale of 0.4 pc, the apparent polarization strength of the source would be much weaker than the intrinsic polarization strength because of beam depolarization (see discussion in Chapter 3). We attempted to estimate the effects of beam depolarization using Monte Carlo simulations. For the simulations, we generated a one-dimensional array of χ_S, representing the position angle of the light after it leaves the halo. The simulations were done by assuming that the intrinsic polarization angle of the light is coherent on a scale much larger than 0.4 pc and that the position angle of χ_S at position i differs from the one at position $i + 0.4$ pc by $\delta\chi = n\delta RM_{0.4\text{ pc}}$, where $\delta RM_{0.4\text{ pc}}$ is the variation in RM in 0.4 pc and n is either $+1$ or -1. The value of n (i.e., either $+1$ or -1) was chosen randomly. The apparent fractional polarization strength π

The radii of most x-ray halos detected fall within this range (see Fabbiano 1987).

was determined by computing the vector average over the entire array of χ_S. The length of the array is set by the telescope resolution. Since the angular resolution power of the VLA in the A configuration (used to observe extragalactic sources) is on the order of kilo parsecond, we simulated the string of χ_S assuming that the angular resolution of the telescope is 1 kpc. It is obvious that the beam depolarization ratios, π/π_o (where π_o is the intrinsic fractional polarization strength of the source), depend strongly on $\delta RM_{0.4\text{ pc}}$, so we made three sets of simulations, each with a different $\delta RM_{0.4\text{ pc}}$: 10, 50, 100 rad/m^2. Each set of simulations contained 500 realizations, and for each realization, we calculated the beam depolarization ratio π/π_o. Since the effects of Faraday rotation also depend on the observing wavelength, we generated three groups of simulations assuming observing wavelengths of 3.6, 6, and 18 cm. Figures 6.12 - 6.14 and Table 6.1 show the distribution of the beam depolarization ratio, π/π_o, and its mean value, $\overline{\pi/\pi_o}$, as a function of $\delta RM_{0.4\text{ pc}}$ and the observing wavelength.

Table 6.1: The Mean Beam Depolarization Ratio, $\overline{\pi/\pi_p}$

$\delta RM_{0.4\ pc}$	3.6 cm	6 cm	18 cm
10	0.97	0.78	0.11
50	0.55	0.20	0.02
100	0.29	0.10	0.00

The amount of rotation is proportional to λ^2, so it is not surprising that the effect of beam depolarization is the most (least) significant at 18 (3.6) cm. However, the severe depolarization at 18 cm even with a very small $\delta RM_{0.4\text{ pc}}$ is still striking! For $\delta RM_{0.4\text{ pc}} = 10$ rad/m^2, the apparent polarization strength at 18 cm is only 11% of

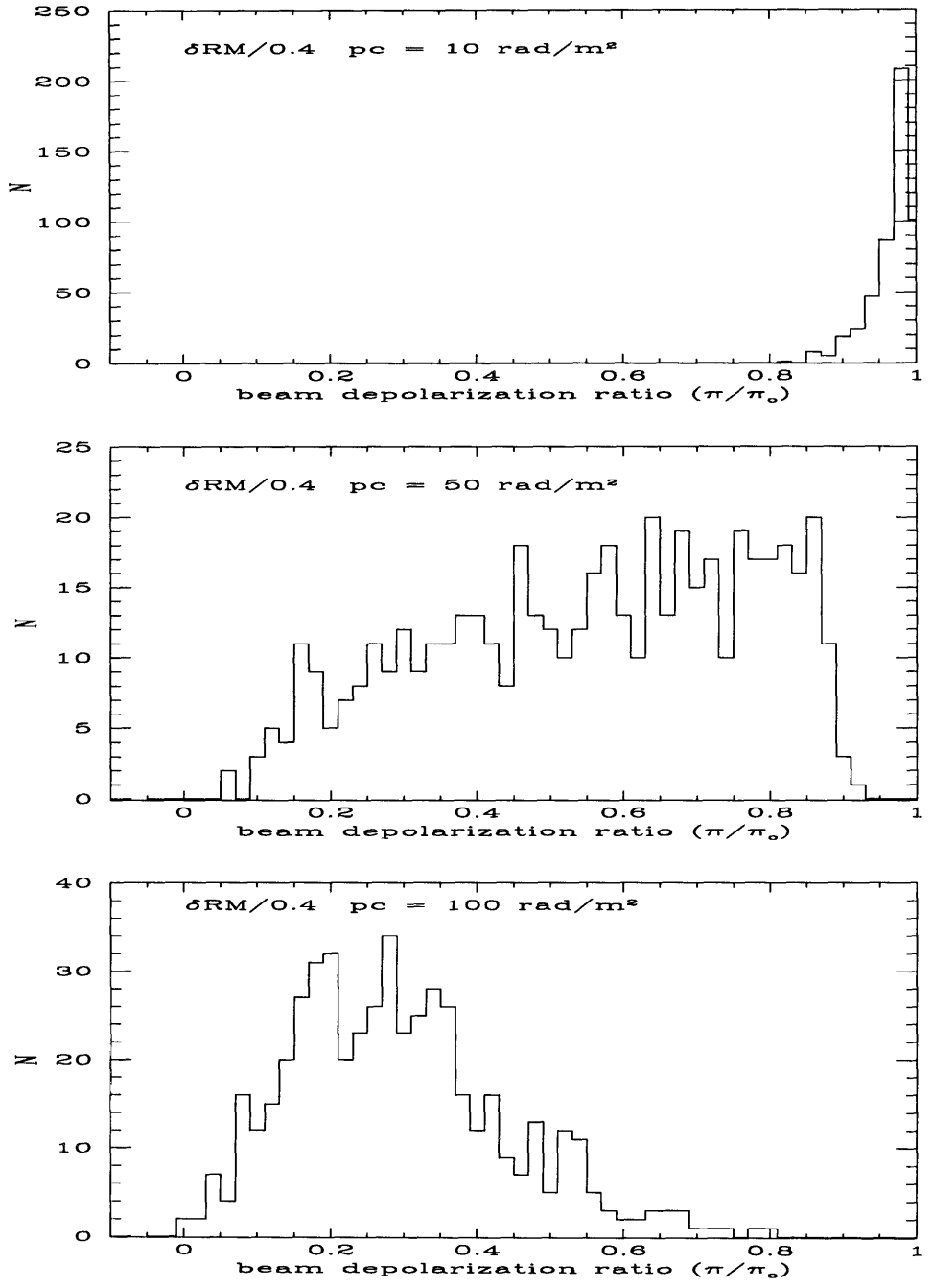


Figure 6.12 – The distribution of π/π_o at 3.6 cm for three different $\delta RM/0.4$ pc. The value of $\delta RM/0.4$ pc is shown in the upper left corner. Each set of simulations contains 500 realizations.

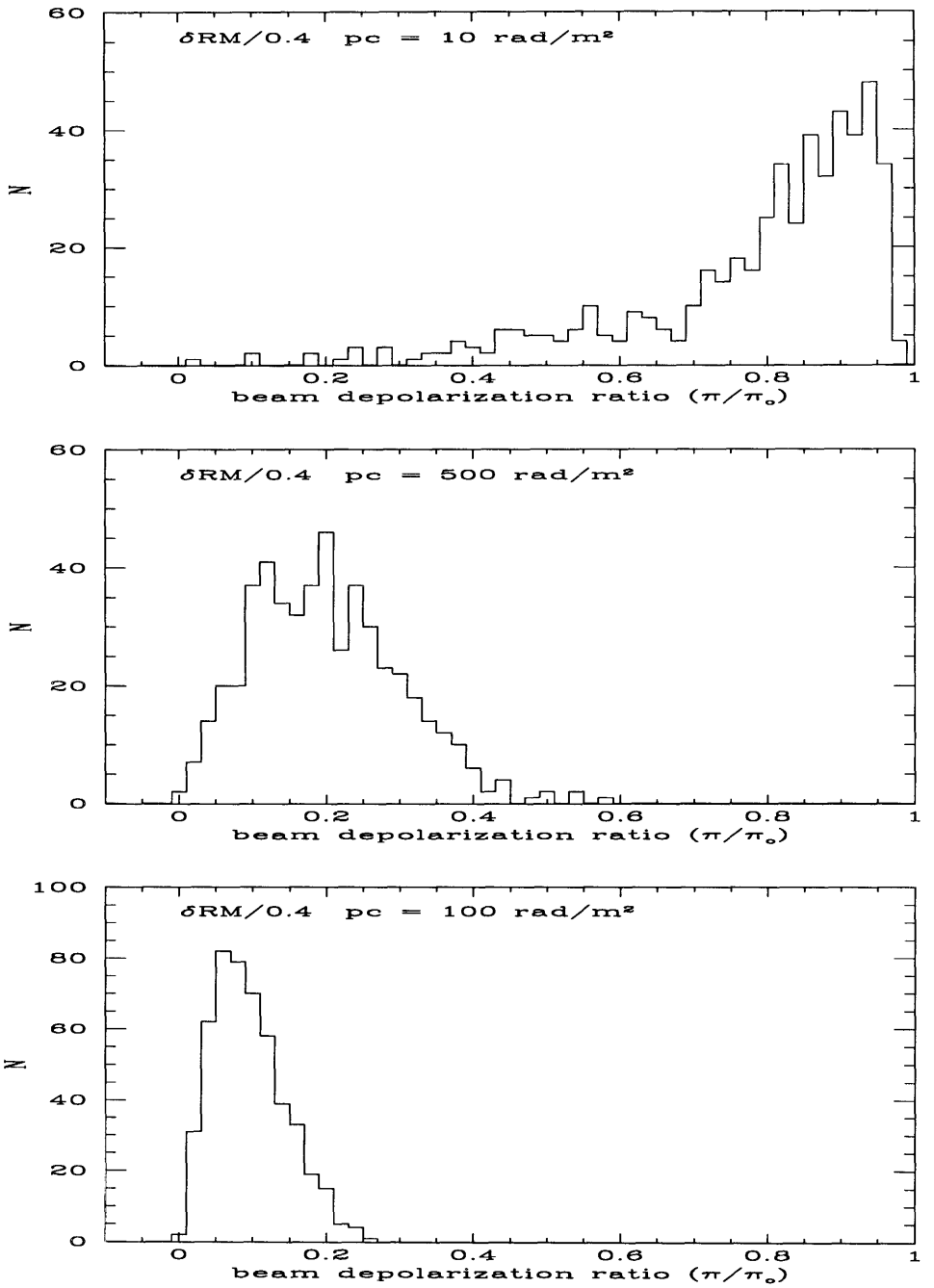


Figure 6.13 – The distribution of π/π_o at 6 cm.

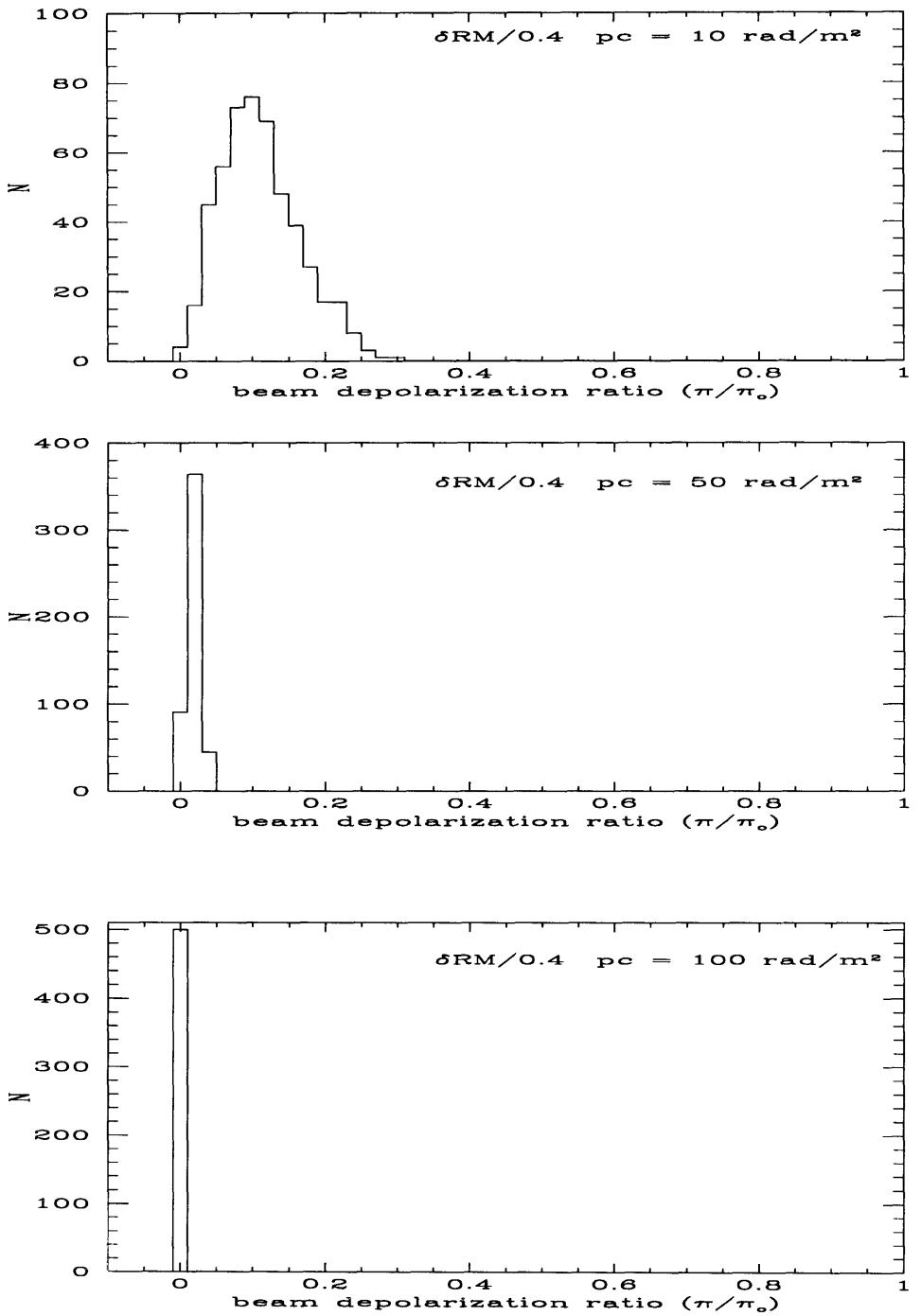


Figure 6.14 – The distribution of π/π_o at 18 cm.

the intrinsic polarization strength. Furthermore, if $\delta RM_{0.4\ \mathrm{pc}}$ were larger than 10 rad/m^2, the polarization strength of the source at 18 cm would not be detectable because of the severe beam depolarization. The outcome of our simulation indicates one important result. If there is a significant change of RM in the halo on a scale ~ 0.4 pc, the observed polarization strength of most extragalactic radio sources at 18 cm would be severely beam depolarized. The observed polarization strength of a typical extragalactic radio source at radio frequencies is quite substantial (for example, the fractional polarization strength of a typical radio lobe is ~ 0.3). Thus, to be consistent with the radio observations of extragalactic radio sources, $\delta RM_{0.4\ \mathrm{pc}}$ in the halo must be ≤ 10 rad/m^2. In other words, our simulation indicates that the contribution of the halo around the source to ΔRM is roughly

$$\Delta RM_S \leq 10 \mathrm{rad/m}^2, \qquad (6-6)$$

which is, again, much smaller than the ΔRM detected in MG1131.

6.5 Effects due to the Intergalactic Medium

The presence of the Lyman α clouds (see Sargent 1988 for a review) confirms the existence of the intergalactic medium (IGM), so it is entirely possible for a light ray leaving the source at a moderate redshift to encounter one or more components of the intergalactic medium before reaching the observer. The contribution due to the IGM is rather difficult to estimate because we do not know very much about it. Nevertheless, we can make a crude estimate of the IGM contribution to our ΔRM measurements from the measurements of RM in other extragalactic radio sources.

In 1981, Simard-Normandin, Kronberg, and Button (1981) obtained RM for 555 extragalactic radio sources for the purpose of mapping the RM distribution of the sky. Their studies should give us some information on the rotation measure due to the IGM (denoted as RM_{IGM}). To avoid confusion due to our Galaxy, we considered only

the RM measurements they obtained for sources with a Galactic latitude $b > \pm 60°$; there are altogether 99 sources with $b > \pm 60°$ in their sample. Their studies showed that the typical values of RM for these sources are between ~ -40 and 40 rad/m^2. If we assume that most of the rotation measure experienced by the light rays in these sources is caused by the presence of the IGM, the statistics of the rotation measure detected in these sources should allow us to place a limit on the contribution of the IGM to our ΔRM measurements. Of course, this assumption is not entirely correct since other agents (i.e., the Galaxy, the plasma around the source, and the plasma within the emission region) can also contribute to the Faraday effects. However, it allows us to place a *conservative estimate* on the contribution due to the IGM. The rotation measures in the sources with $b > \pm 60°$ in their sample give a mean rotation measure, $\overline{RM}$, of ~ -1.11 rad/m^2 and a standard deviation from the mean, $\sigma_{\overline{RM}}$, of ~ 25 rad/m^2, indicating that the RM_{IGM} (conservatively) is likely to be on the order of 25 rad/m^2. Therefore, we estimate that

$$\Delta RM_{IGM} \sim 25 \text{ rad/m}^2,$$

which, again, is much smaller than the ΔRM we detected in MG1131 (see discussion in Section 6.2).

6.6 Discussion

Our previous analysis suggested that $\Delta RM_{Galaxy} \ll 10$ rad/m^2, $\Delta RM_S \leq 10$ rad/m^2, and $RM_{IGM} \sim 25$ rad/m^2. All are much smaller than the ΔRM we detected in MG1131 (ranges from 4 to 400 rad/m^2). Combining the effects due to the Galaxy, the halo around the source, the IGM, and the uncertainties in the RM measurements (see discussion in Section 6.1) in quadrature gives a value of ~ 60 rad/m^2, which is, again, much smaller than the ΔRM detected in MG1131. Hence, we believe that the ΔRM detected in this system is most likely to be due to the variation of

the rotation measure in the lens galaxy (RM_L). In Figure 6.15, we superimposed the limit of 60 rad/m^2 (the dashed line) on the log-log plot of the $|\Delta RM|$ vs. Δr.

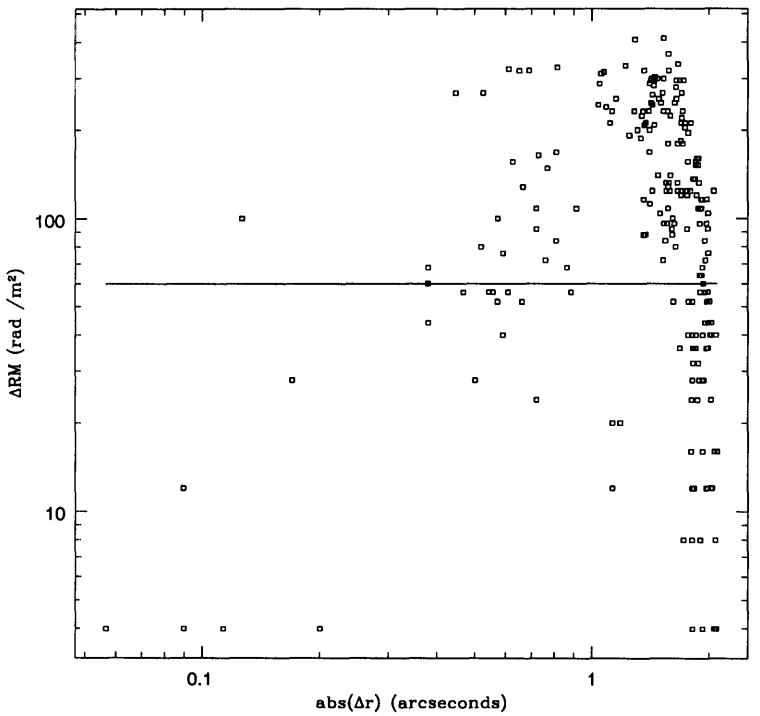


Figure 6.15 – The 60 rad/m^2 limit (the solid line) is superimposed on the $|\Delta RM|$ vs. Δr plot (on a log-log scale).

We can only determine the ΔRM in conjunction with the lens model. As discussed in Chapters 4 and 5, although our best lens model can produce an image that closely resembles the observed image, our model is still not completely consistent with all the constraints in the system. Then, is it possible that the structure we have detected in ΔRM is an artifact caused by the imperfection of the lens model? If this is true, the structure of ΔRM should depend on the model employed since the characteristics of the imperfections in each model would be different. Thus, we can examine this possibility by investigating the differences in the structure of ΔRM when a different lens model is used for the calculation. We proceeded with the investigation by repeating the same procedures for determining ΔRM but with a different lens model,

the optimal isothermal model fitted to the system (see discussion in Chapter 5). We chose the isothermal model for this experiment because it gave a significantly worse fit to the system than our best lens model. Figure 6.16 show the $|\Delta RM|$ as a function of Δr determined when the isothermal model was used on a log-log scale. For comparison, we superimpose the results obtained when the $\alpha = 0.63$ lens model was employed for the calculation. It is clear that the results obtained using either model are rather similar. The results of this experiment let us believe that the structure of ΔRM is not dependent on the lens model employed, so they are not likely to be artifacts of the imperfection of the lens model.

6.7 Conclusion

We have shown that most of the ΔRM detected in MG1131 is due to the variation of the rotation measure in the lens galaxy. Our measurements of ΔRM obtained from MG1131 indicate that the rotation measure in the lens galaxy varies on a scale roughly 200 to 400 rad/m^2 between two points separated by $\sim 1''$, suggesting that the magnetic field in the lens galaxy is most likely to have variation on a scale of $1''$, also. It appears that the ΔRM measurements are not sensitive to the lens model employed. The structure of ΔRM as a function of Δr we obtained is important since it provides the first observational constraint for modeling the magnetic field in high-shifted galaxies. We believe that the same method should be applied to other gravitational lenses with linearly polarized emission. Collectively, the information obtained from these lenses would lead a new horizon in understanding the magnetic field structure in high-redshifted galaxies.

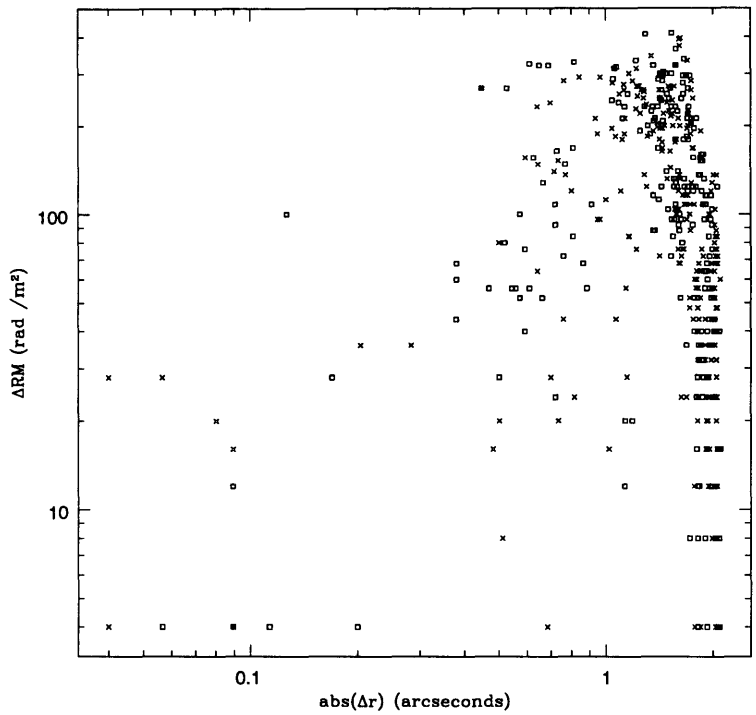


Figure 6.16 – A log-log scale plot of $|\Delta RM|$ vs. Δr. The crosses represents the data points obtained when the isothermal lens model is used for the calculation, and the squares represents the ones obtained when the $\alpha = 0.63$ lens model is used.

References

Beck, R. & Bronberg, P. P. (eds.), *Galactic and Intergalactic Magnetic Fields*, IAU Symposium 140, Kluwer Academic Publishers.

Fabbiano, G. 1989, Ann. Rev. Astron. Astrophys., **27**, 87.

Laoz, T. J., Spangler, S. R. & Cordes, J. M. 1990, Ap. J., **363**, 515.

Kellermann, K. I. & Owen, F. N. 1988, *in Galactic and Extragalactic Radio Astronomy*, eds. ?? ?? Verschuur & K. I. Kellermann, pp. 563.

Kronberg, P. P. & Perry, J. 1982, Ap. J., **263**, 518.

Sargent, W. L. W. 1988, *in QSO Absoption Lines*, Space Telescope Scinece Institute Symposium Series, eds. J. C. Blades, D. Turnshek & C. A. Norman, pp. 1.

Simard-Normandin, M. & Kronberg, P. P 1980, Ap. J., **242**, 74.

Simard-Normandin, M., Kronberg, P. P, & Button, S. 1981, Ap. J.S., **45**, 97.

Simonetti, J. H., Cordes, J. M. & Spangler, S. R. 1984, Ap. J., **284**, 126.

Simonetti, J. H. & Cordes, J. M. 1986, Ap. J., **206**, 735.

Sofue, Y. & Fujimoto, M. 1983, Ap. J., **265**, 722.

Vallée, J. P. & Bignell, R. C. 1983, Ap. J., **272**, 131.

7. The Possibility of Measuring the Hubble Parameter from MG1131+0456

Much of the material presented in this chapter is taken from the paper, *"Variability in the Einstein Ring Gravitational Lens MG1131+0456,"* by Hewitt, Chen, & Messier (1995).

One important astrophysical application of gravitational lenses is the determination of the Hubble parameter, H_o (see discussion in Section 1.4.2). The measurement of H_o (from any lensed system) is rather difficult, for it requires three pieces of information about the lensed system which are difficult to obtain: (1) the redshifts of the lens and the source, (2) the time delay of the system (i.e., the difference in the light travel time of a pair of images arising from the same source) and (3) a reliable model describing the gravitational potential of the lens. Since the redshifts of most extragalactic sources are determined from their optical spectral lines, systems in which the lens galaxy and source both emit strongly at optical wavelengths are more likely to yield measurable redshifts. The time delay in the lensed system is only measurable by comparing the features in the light curves (i.e., the variations of the flux density as a function of time) of a pair of images arising from the same source, so the time delay can be obtained only from a lensed system containing a multiply imaged non-quiescent source. Since compact sources are more likely to be variable, lensed systems containing compact components have a higher probability of providing measurable time delays. As discussed in Chapters 1 and 4, reliable lens models can be obtained from systems containing extended structures. Thus, an ideal system for the application of determining the Hubble parameter is one consisting of both compact components and extended structure. MG1131 is one such system; in this chapter, we examine the possibility of using it to measure H_o.

7.1 Variability of the Compact Components

To use MG1131 for the determination of H_o, the time delay of the system must be known, so we first investigate whether MG1131 contains a multiply imaged variable source (because the time delay of the system is only measurable from the light curves of such a pair). We do not expect the ring or the component C in the system to vary since both the observational properties (i.e. the spectral and polarization characteristics of the ring and component C; see Chapter 3) and the structure of the source inferred from of the best lens model suggest that they are the lensed images of radio lobes, which are known to be quiescent sources. On the other hand, the spectral indices of the compact components A and B are consistent with the spectral indices of radio cores (some of which are known to be variable), so they are likely to display flux variations. We thus focus the investigation on the compact components, A and B.

We examined the variability of A and B by monitoring the 8 GHz flux densities of A and B detected by the VLA from Nov 88 to Mar 93 (see Table 2.1 for a complete list of these observations). All 8 GHz datasets used for the investigation were acquired when the VLA was in either the A or the B configuration. The system was also observed at 15 GHz at several different epochs (see Table 2.1). Although these datasets were taken primarily for other scientific purposes, the flux densities of A and B measured from these datasets also provide information on their variability. Thus, we included the 15 GHz datasets in the investigation as well. The analyses and results obtained from both frequencies are presented in the following sections.

7.1.1 8 GHz

7.1.1.1 Data Analysis

There are altogether fourteen 8 GHz datasets in our sample (see Table 2.1). Eight were taken in the A array configuration, three were taken in the B array configuration,

and three were taken in a mixed array configuration (the array between A and B or between B and C configuration). All datasets were calibrated and mapped using the procedures outlined in Chapter 2.

The most straightforward way to determine the variability of any source is to measure its absolute flux density as a function of time (i.e., the light curve). At 8 GHz, the apparent flux densities of A and B are contaminated by the flux of the ring because (1) the angular resolution at 8 GHz is not sufficient to resolve the structure of the ring from the compact components and (2) the flux density of the ring is prominent at this frequency. As a result, the flux densities at the positions of A and B measured from the map (f_A^m and f_B^m) differ from their true flux densities (f_A and f_B) by $f_A^m = f_A + f_A^{ring}$ and $f_B^m = f_B + f_B^{ring}$, where f_A^{ring} and f_B^{ring} are the flux densities contributed by the ring at the positions of A and B. The measurements of f_A and f_B are only possible if f_A^{ring} and f_B^{ring} , which can be estimated only from a model describing the structure of the ring, are known. For MG1131, because none of the observations contains sufficient information to accurately construct the model for the structure of the ring, we cannot measure the absolute flux densities of A and B. However, if we assume that the flux densities contributed by the ring at the positions of A and B are the same in every observation (i.e. f_A^{ring} and f_B^{ring} do not change from one observation to another), we can measure the variation in the flux densities of the compact components equally well by comparing the measured flux densities (f_A^m and f_B^m) to a reference flux standard containing the same amount of f_A^{ring} and f_B^{ring}. The relative flux densities of A and B in any observation can be obtained by measuring the residual fluxes at the positions of A and B after subtracting the reference map from the program map. In principle, if the ring were a quiescent source, there should be no flux in the residual map except at the positions of A and B, and the variability of the compact components are indicated by these residual fluxes around the positions of A and B.

There are two technical difficulties associated with this method. The first is that we can only perform the map subtraction if all maps are perfectly aligned. When we observe a source with the VLA, the absolute position of the source can differ from one observation to another because of the calibration errors and the positional uncertainties of the phase calibrator. Thus, we must manually align the maps before performing the map subtraction. The other problem is that the reconstruction of the ring can be different from one image to another due to the differences in the visibility sampling and the calibration errors. The differences in the ring reconstruction can cause f_A^{ring} and f_B^{ring} to change from one image to another, thereby corrupting our measurements. We can align the maps and reduce the differences in the ring reconstruction with a technique known as "cross-calibration", introduced by Mason (1986). The main concept of the technique is as follows. Make a model of the source, containing information on both the position and structure, from the datasets with the most complete uv coverage and the fewest calibration errors. Then, calibrate all other datasets (after they have been self-calibrated) with this model. This step aligns the maps and minimizes the differences in the structure of the source reconstructed from the measured visibilities. We adopted this technique as part of our data analysis procedure.

For the purpose of obtaining information on the variability in the compact components, we analyzed the 8 GHz datasets using the following procedure, which are similar to the procedures used by Mason (1986) and Lehár et al. (1992):

(1) Compute one further iteration of phase corrections by "cross-calibrating" the dataset with respect to a common source model. We chose the CLEAN component model of the reference map as the common source model.

(2) Convolve the maps (including the reference map) with a standard Gaussian beam. The beam size was chosen so that all the data sets would have the resolution of the largest beam. The beam was $0\farcs227 \times 0\farcs227$ for the A array maps and $0\farcs67 \times 0\farcs67$

for the B array maps. Since the angular resolutions of the datasets taken in the mixed arrays were closer to the resolution of the B array, we treated these datasets as the B array datasets.

(3) Subtract the reference map from the program maps. The contamination from the ring is strongly dependent on the resolution of the map, so we chose two maps, one for each array configuration, as the references. The Nov 88 the Mar 93 datasets were selected to be the reference maps for the A array datasets and the B array datasets, respectively.

(4) Measure the integrated residual flux at the position of A (or B) by summing the residual flux densities within a box of size three times the beam size centered at the position of A (or B). We denote the measurements by f_A^r (or f_B^r).

7.1.1.2 Error Estimates

There are four possible sources of errors in the f_A^r and f_B^r measurements: (1) the thermal noise of the receivers, (2) the errors in the flux density scale determined from the observations of the flux calibrator, 3C286, (3) the systematic errors due to the differences in the ring reconstruction in each map (even after cross-calibration), and (4) the systematic effect due to the procedure of cross-calibration. The thermal noise of the receiver is expected to contribute roughly 40 μJy/pixel to each of the images, and the estimated errors in the flux density scale calibrated by 3C286 is about 2%. Although all maps were cross-calibrated, this procedure only minimizes the differences in the ring reconstruction. As a result, there are residual differences in the structure of the ring left in each map, which could add significant errors to the measurements of f_A^r and f_B^r. It is not possible to estimate analytically the systematic errors due to either the imperfect ring reconstruction or the cross-calibration since the errors are caused by procedures involving several non-linear iterative steps, so we estimate them numerically.

We estimated the effects due to the residual differences in the ring reconstruction by investigating the differences in f_A^r and f_B^r obtained from datasets observed on the same day but with different visibility sampling and calibration errors. The purpose of restricting the experiment to datasets taken on the same day is to avoid confusion caused by possible intrinsic variability. To obtain the test sample for this numerical experiment, we segmented a long observation into several smaller datasets and used them as the test datasets. We prepared two groups of test samples, one for each array configuration. For the A (B) array, the sample was generated from the Jun 91 (Dec 91) observation. We calibrated, mapped, and analyzed each test dataset independently with the procedures outlined in Chapter 2 and Section 7.1.1.1. The f_A^r and f_B^r were measured from each test dataset. The standard deviation from the mean of f_A^r (and also f_B^r) was about 0.12 mJy for Group A and 0.70 mJy for Group B, indicating that the level of systematic errors caused by the imperfect ring reconstruction are at that level.

We next examined the systematic effects due to cross-calibration. As discussed in the previous section, the purpose of cross-calibrating all datasets to a common source model is to minimize the differences in the reconstructed structure of the ring. We suspected that the procedure for doing this would also have the undesirable effect of removing or reducing the signals due to the intrinsic source variations, so we conducted a numerical experiment to examine this possibility. We took the same test samples as mentioned above and modified all test datasets by adding 1 mJy of flux to both A and B to simulate the flux variation. If cross-calibration did not remove the signal of variability, the measurements of the residuals at A and B obtained from the modified test datasets would be different from the measurements obtained from the original test datasets by exactly 1 mJy. We checked for this difference. We analyzed all the modified test datasets in each group with the same procedure, measured the residual fluxes at the positions of A and B, and compared them with

the ones obtained from the original test datasets. The differences between the f_A^r (and also f_B^r) obtained from the modified datasets and those from the original test datasets were much less than 1 mJy. The mean difference obtained from the A array test datasets was 0.93 mJy, and that obtained from the B array test datasets was 0.87 mJy. Both indicate that the cross-calibration has the systematic effect of reducing the signal of variability by ~10%.

In summary, combining the errors due to the thermal receiver noise, the uncertainty in the flux density scale, and the imperfect ring reconstruction in quadrature gives conservative estimates of 0.2 mJy for the A array data and 0.8 mJy for the B array data. Figures 7.1 and 7.2 show the residual maps of observations in both A and B array configurations; the contour levels represent -3, -2, -1, 1, 2, and 3 times the estimated noise. Our numerical experiments showed that the cross-calibration reduces the variability signal by approximately 10%, so our analysis lead to a conservative estimate of the source variability.

7.1.1.3 Evidence for variability

Except for the Jul 91 observation, the residual flux densities of the ring, after the map subtraction, are all well below the estimated flux errors. The residual map in the Jul 91 dataset shows significant negative residuals at the position of the ring, indicating that the flux density of the ring detected in Jul 91 is significantly lower than the ring flux densities detected in Nov 88 dataset. A large fraction of the visibilities in the Jul 91 observation were measured when the source was at low elevation. These visibilities were more likely to contain larger calibration errors, so we suspected that the calibration errors in these data points were the cause of the large negative ring residual in the difference map. To test our hypothesis, we conducted a numerical experiment. We segmented the Jul 91 data into three smaller datasets, containing the visibilities acquired at the beginning, the middle, and the end of the Jul 91

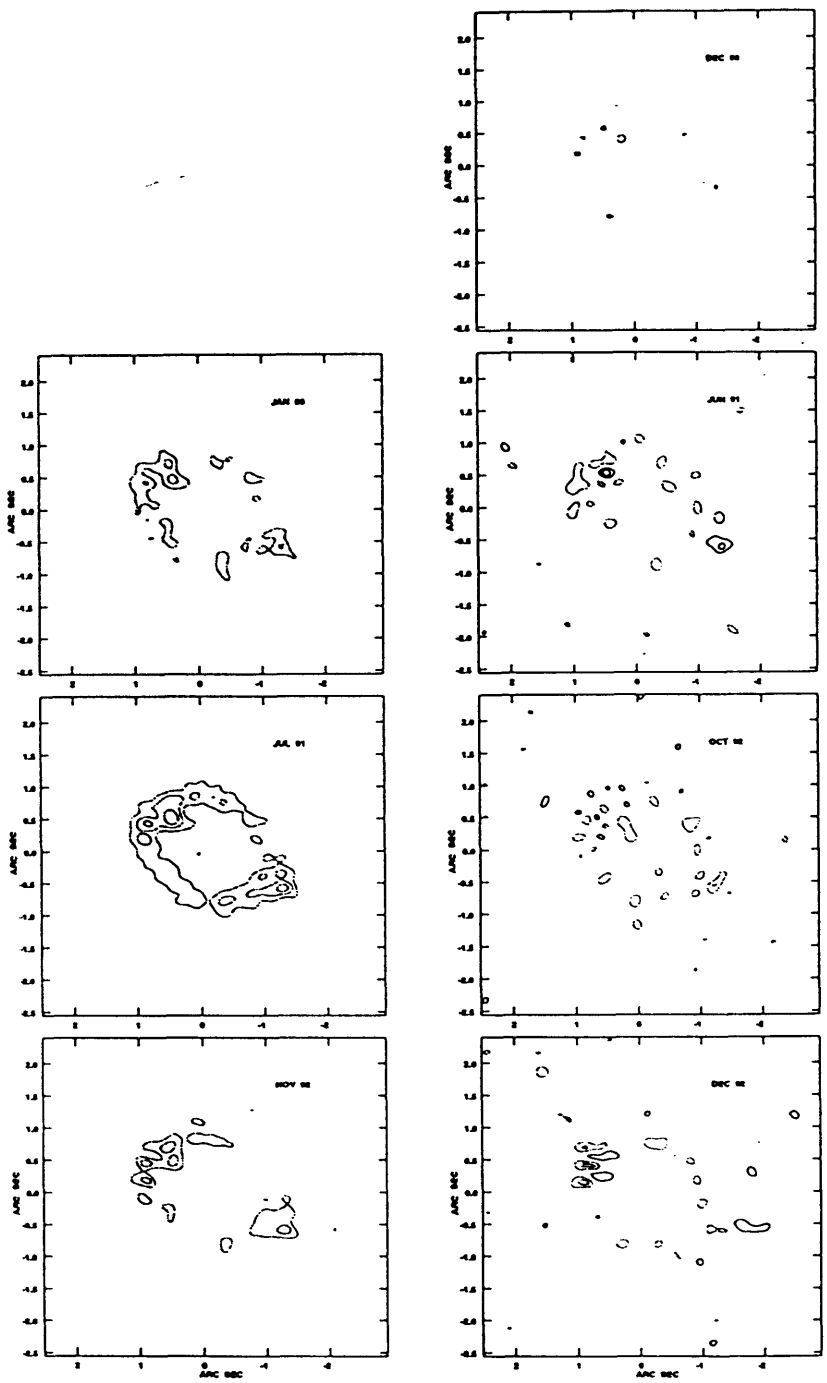


Figure 7.1 – Contour plots of the residual maps of the 8 GHz A array datasets. The Nov 88 dataset was used as the reference standard. Contour levels are -3, -2, -1, 1, 2, and 3 × 0.2 mJy.

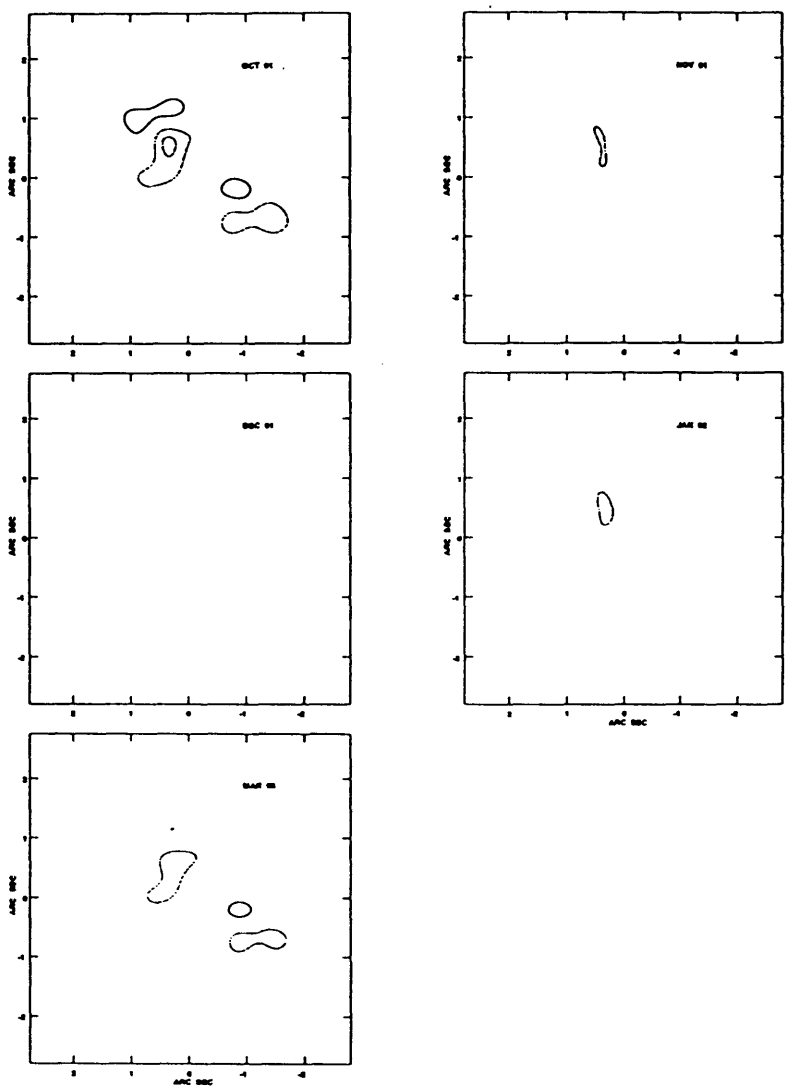


Figure 7.2 – Contour plots of the residual maps of the 8 GHz B array datasets. The Mar 93 dataset was used as the reference standard. Contour levels are -2, -1, 1, and 2 × 0.8 mJy.

observing run. Transit occurred near the end of this observation, so the dataset containing visibility acquired at beginning of the run should have the largest calibration errors. Each dataset was calibrated and mapped with the standard procedures. As expected, the flux density of the ring reconstructed from the visibilities measured at the beginning of the run was the lowest, and the flux density of the ring reconstructed from the visibilities measured at the end of the run was the highest, which confirmed our hypothesis. Unfortunately, these calibration errors cannot be corrected easily, so we excluded Jul 91 dataset from any further analyses.

Among the A array datasets, only the Jun 91 dataset gives clear evidence of flux variation in A and B (see Figure 7.1). The residual of the Jun 91 dataset has a well subtracted ring and two point-like residuals of 540 and 470 μJy at the positions of A and B, respectively. However, none of the other A array datasets has residuals significantly above the estimated measurement error. Given our estimates of the measurement errors, we do not see any evidence for variability in any of the B array datasets, indicating that the flux variations in these components must be less than 0.8 mJy (the measurement error). Considering all the observations, we conclude that the variability of the compact components is rather weak, with flux densities varying on a scale of 500 μJy over 20 months.

7.1.2 15 GHz

MG1131 was observed with the VLA in the A configuration at 15 GHz at six different epochs (see Table 2.1 for the complete list). At 15 GHz, the atmospheric phases fluctuate rapidly and irregularly. Although the observations were set up so that the phase calibrator was observed at least once within the coherent time scale of the atmospheric phase variation, we still were not able to track the atmospheric phase variation throughout the observations completely. This added additional phase errors to the measured visibilities. We could not perform self-calibration to correct

these phase errors because the signal-to-noise ratios in the 15 GHz maps were not sufficient for such an operation. We therefore restrict further discussion to the four observations in which the atmospheric phases were most stable: Sep 87, Nov 88, Dec 88, and Nov 92. Images constructed from the data taken on those dates are shown in Figure 7.3. The root-mean-square noise in all four maps is within a factor of two of the expected receiver noise, indicating that the residual phase errors in all four datasets were small.

At 15 GHz, the flux densities of A and B measured from the map should not be severely contaminated by the flux densities of the ring (as they were in the 8 GHz maps) because (1) the resolution at 15 GHz is much higher than at 8 GHz and (2) the flux density of the ring is rather low at this frequency. Thus, we took a different approach to analyzing the 15 GHz data. Instead of measuring the flux densities of A and B relative to some reference standard, we measured the absolute flux densities. The differences in the uv coverage between observations should have little effect on our measurements because, again, the ring flux is low and we are primarily interested in the flux densities of the compact components. The flux of each compact component (either A or B) was measured by summing the flux densities within a box of $0\farcs12 \times 0\farcs12$ centered at the position of that component. Since the size of the box is only comparable to the size of the beam at FWHM, the measured flux densities (and hence the variability) are underestimated. We chose this conservative measure to minimize the confusion due to the surrounding extended emission. The results are presented in Table 7.1.

Since MG1131 could not be self-calibrated at 15 GHz, we suspected that the uncertainties in our measurements were caused mainly by the residual phase errors. We carried out numerical tests to estimate the uncertainties due to these phase errors At 15 GHz, both phase calibrators (1055+018 and 1114-001) are unresolved on all baselines in the A configuration, so we could calculate their flux densities by considering

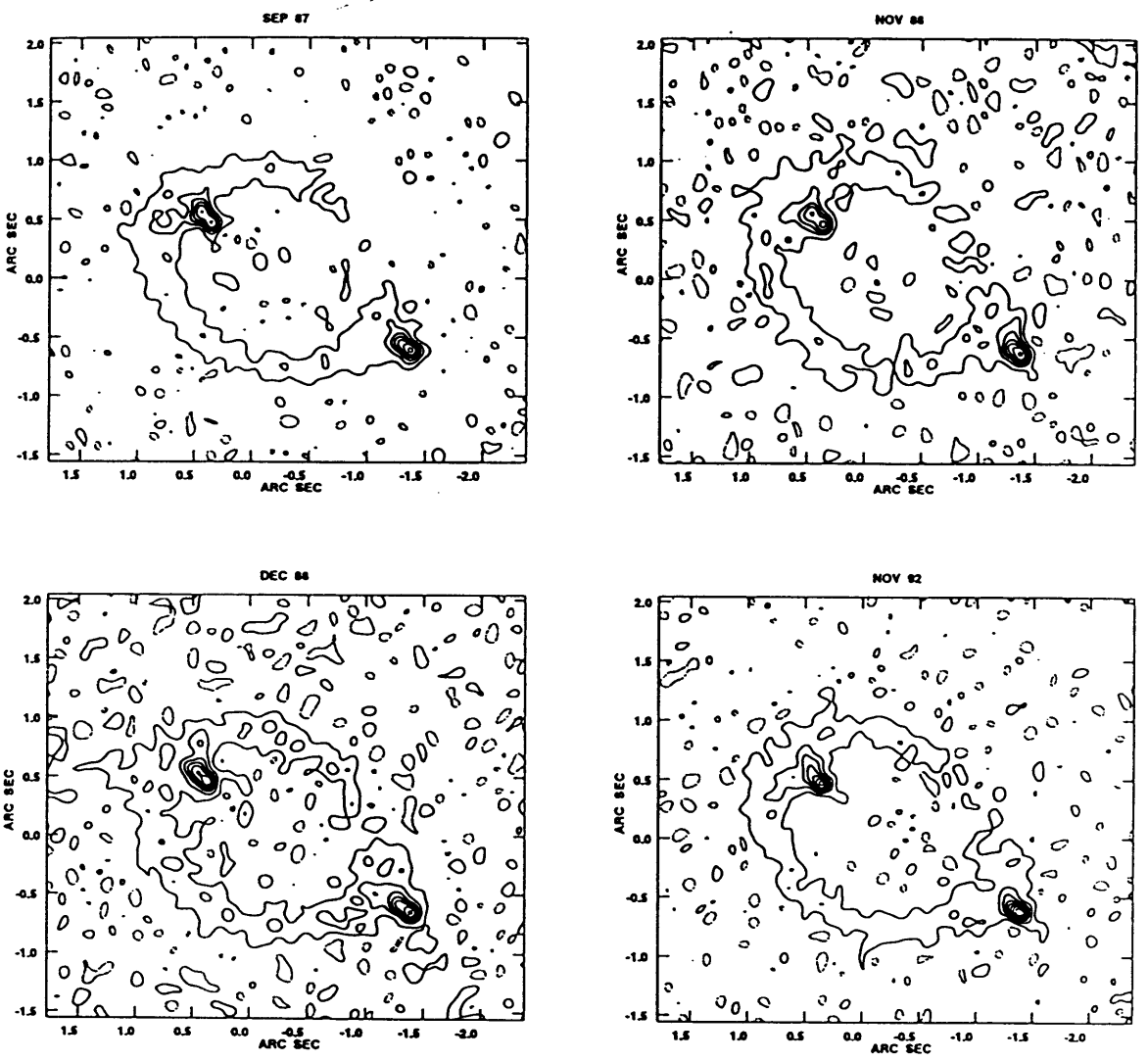


Figure 7.3 – Plots of the 15 GHz images. Contour levels are -0.25, 0.25, 0.75, 1.25, 1.75, 2.20, 2.75, 3.25, and 3.75 mJy.

Table 7.1: Flux density measurements of A and B at 15 GHz

Date	A (mJy)	B (mJy)
20 Sep 87	2.73 ± 0.11	3.23 ± 0.13
2 Nov 88	2.08 ± 0.17	3.07 ± 0.26
15 Dec 88	2.71 ± 0.11	3.08 ± 0.12
6 Nov 92	3.18 ± 0.19	3.90 ± 0.23

only the visibility amplitudes. The atmospheric phase errors should have no effect to this calculation. On the contrary, when the calibrators are mapped, both visibility amplitudes and phases are used, so the flux densities of the calibrators measured from the map must be influenced by the atmospheric phase errors. Thus, we can estimate the effects of phase errors on the measured flux densities by comparing the flux density of the calibrator calculated from the visibility amplitudes with that measured from the map. We examined how the flux densities measured from the map differed from the calculated values if the atmospheric phases were calibrated once every τ minutes. The experiment was conducted using four different values of τ: 2, 10, 20, and 30 minutes, and we performed the same numerical experiment on all four observations. Table 7.2 tabulates the calculated and measured flux densities as a function of τ for each observation.

For point sources (such as the phase calibrators), phase errors simply cause the flux measured from the map to be lower than the actual value. At 15 GHz, the atmospheric phase variation is rapid, so datasets with less frequent calibration should contain more phase errors. Hence, it is not surprising that the measured flux densities decreases as τ increases (see Table 7.2). However, for sources with complicated

Table 7.2: The measured and calculated flux densities of the calibrator as a function of τ

τ (min)	Sep 87 (mJy)	Nov 88 (mJy)	Dec 88 (mJy)	Nov 92 (mJy)
2	4.07	4.27	4.20	0.89
10	4.03	4.13	4.20	0.85
20	3.96	4.01	4.08	0.85
30	3.60	3.47	3.85	0.83
calculated	4.15	4.48	4.37	0.90

structures, the phase errors cause a transfer of brightness from one component to another. We can crudely take the level of flux lost in the calibrator due the phase errors as the level of uncertainty, caused by the residual phase errors, in the flux densities measured for A and B. In our observations, the phase calibrator was observed once every 10 minutes, so we use the results of the numerical experiments with $\tau = 10$ to estimate the uncertainties in our measurements. For $\tau = 10$ min, the phase errors caused the measured flux density of the phase calibrator to be lower than the calculated one by 3%, 8%, 4%, and 6% for the Sep 87, Nov 88, Dec 88, and Jan 92 datasets respectively, so we estimated that the phase errors caused the flux densities of A and B measured from the Sep 87, Nov 88, Dec 88, and Jan 92 datasets to be uncertained by the same degree. Combining this uncertainties with the errors in the flux density scale determined from the observations of the flux calibrator (roughly 2%; see Section 7.1.1.2), we estimated that the flux densities of A and B measured from the Sep 88, Nov 88, Dec 88, and Jan 92 are uncertain by 4%, 8%, 4%, and 6%, respectively (see the error bars in Table 7.1). It is clear from Table 7.1 that the flux

variations in A and B are larger than our estimated measurement errors, indicating that the source is variable.

7.2 Is the Time Delay in MG1131+0456 Measurable?

Our analyses of both the 8 and 15 GHz datasets indicated that the source lensed into components A and B is weakly variable, at a level of ~ 500 μJy over 20 months. We could not construct the light curves of A and B from our current measurements at either 8 or 15 GHz because the time sampling in these datasets was too sparse; hence, it is not possible to determine the time delay of the system from our datasets. However, for the purpose of investigating the possibility of using MG1131 to measure H_o, the more interesting question to address is whether the time delay in MG1131 can, in principle, be measured from A and B. If the time delay is indeed measurable, then the next important question to address is how one must plan the observations to determine the time delay of the system, given that the flux variations in the compact components are as weak as detected. We attempt to answer these two questions with simulations.

Let N be the total number of observations available for constructing the light curves, δt be the time between two observations, and δS be the uncertainty in the flux density from each measurement. Assuming that the system has a true time delay τ_{true} and that the flux ratio between the pair of images is $R_{GL} = 0.8$, we simulated the light curves of A and B with a level of variability agreeing with our observations. In principle, it is only possible to generate synthetic light curves with characteristics consistent with the observations if the measurements are adequate to provide a quantitative description of the properties of the light curves. Unfortunately, our current flux measurements of A and B are insufficient to provide such information. However, it was shown by Hughes, Aller, & Aller (1992) that the light curves of most variable radio sources display similar characteristics. They studied 51 variable radio

sources, measured their light curves at centimeter wavelengths, and found that the structure functions * of the light curves of most sources in their sample could be fitted by a power-law model of the form

$$V(\tau) = V_1 \tau^{\alpha}.$$

The values of the power-law index α were clustered around unity with a mean of $\alpha = 1.09 \pm 0.34$. If the source lensed into components A and B is like other variable radio sources, then its structure function is most likely to have a power-law index of 1; thus, we assumed that the structure function of A and B can be expressed by a power-law model with a power law index $\alpha = 1$. The value of V_1 was determined by fitting this model to the flux densities of component B measured at 15 GHz. The best fit value is $V_1 = 2 \times 10^{-4}$ [mJy]^2/day. Figure 7.4 shows the fit. This model predicts that the fluxes of A and B change by roughly 0.4 mJy in one year (i.e. $\langle [S(t) - S(t+\tau)]^2 \rangle^{1/2} \simeq 0.4$ mJy), which is consistent with our measurements. We therefore simulated the light curves of A and B assuming a structure function of $V(\tau) = 2 \times 10^{-4} \tau$ [mJy]^2/day.

We suspected that the possibility of accurately measuring the time delay in the system depends on the size of the measurement error, so we made two groups of simulations, each with a different δS. For the first group, we assumed that $\delta S = 0.2$ mJy, corresponding to the errors we estimated for the current measurements; for the second group, the δS was assumed to be 0.1 mJy. It is also likely that the

* The characteristics of the source variability are most commonly described by its first order structure function $V(\tau)$ of the light curves $S(t)$, where the structure function is defined as

$$V(\tau) = \langle [S(t) - S(t+\tau)]^2 \rangle / 2$$

(cf. Simonetti, Cordes & Spangler 1984). The angle brackets denote the expectation value.

level of accuracy in measuring the time delay also depends on the total number of observations, N, and the time between measurements, δt, so for each group, we made three sets of simulations with different combinations of N and δt. They were: (1) N = 100 and δt = 10 days, (2) N = 50 and δt = 50 days, and (3) N = 50 and δt = 20 days. Each set of simulations contained 1000 realizations (i.e. 1000 pairs of synthesized light curves). The τ_{true} in each pair of light curves was chosen randomly between 0 and 200 days.

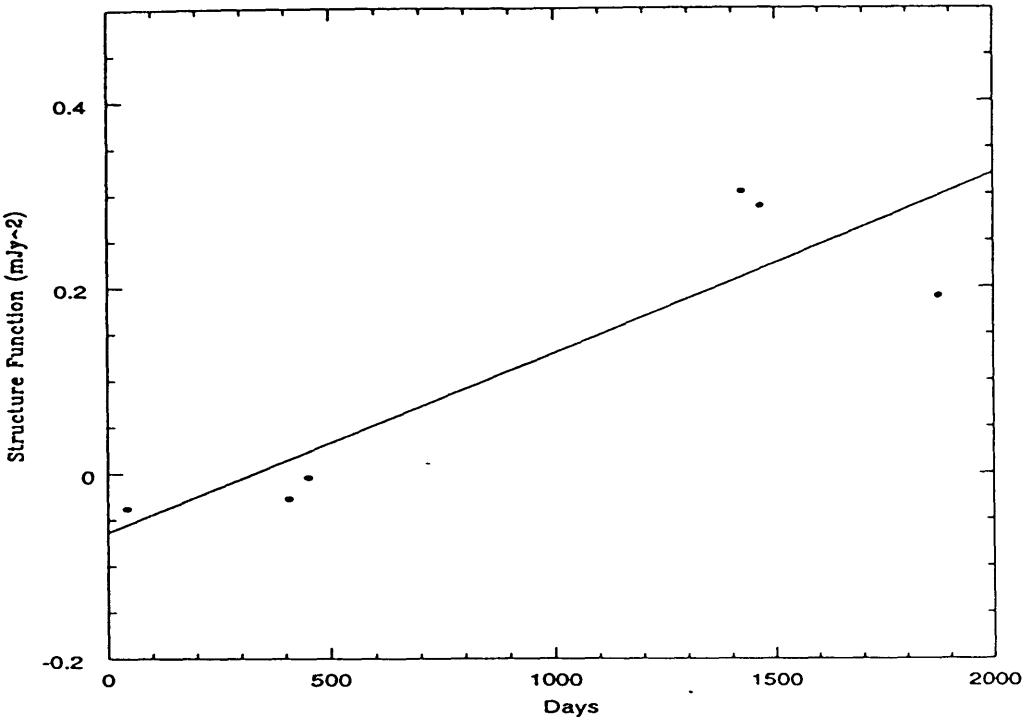


Figure 7.4 – The data points show the structure function of the variable source $[S(t) - S(t+\tau)]^2/2$ as a function of τ calculated from the 15 GHz flux measurements of the component B. The solid line shows the model fitted to the data points, $V(\tau) = 2 \times 10^{-4}\tau$ [mJy]^2/day

We determined the time delay between each pair of synthesized light curves using the χ^2 statistical analysis introduced by Press, Rybicki, & Hewitt (1992). For each pair, a trial time delay was assumed, the light curve for component B was shifted by

the trial time delay and scaled by R_{GL}. The two light curves were then combined, giving a data vector

$$y_i = y(t_i); \qquad i = 1, 2, ..., 2N.$$

We adopted a χ^2 statistic of

$$\chi^2 = (\mathbf{y} - \alpha\mathbf{E})^{\mathrm{T}}\mathbf{A}(\mathbf{y} - \alpha\mathbf{E}),$$

where $\mathbf{E}$ is the column vector $(1, 1, 1, 1, ..., 1)^{\mathrm{T}}$, $\mathbf{A}$ is the inverse of the covariance matrix calculated from the structure function and the assumed measurement error, and

$$\alpha = \frac{\mathbf{E}^{\mathrm{T}}\mathbf{A}}{\mathbf{E}^{\mathrm{T}}\mathbf{A}\mathbf{E}}\mathbf{y}.$$

[See Press, Rybicki, & Hewitt (1992) for a more detailed discussion of this technique and a justification for using this χ^2 statistics.] In other words, the χ^2 measures the goodness-of-fit of the combined light curve to the structure function $V(\tau)$ of the light curve. The calculation was carried out for each value of the trial delay τ, producing $\chi^2(\tau)$. The value of τ which gives the smallest χ^2 value was taken as the best fit time delay of the system, τ_{fitted}. For each realization, τ_{fitted} was found by searching through the trial delays in the range of ± 300 days. For most cases, despite the small level of variability in the light curves, there was an unambiguous global minimum in the $\chi^2(\tau)$ curve, indicating a clear time delay detection. We refer to these cases as "good" experiments. Nevertheless, there were several cases in which no global minimum could be found in the $\chi^2(\tau)$ curve. We viewed these as failed experiments. For every "good" experiment, we computed the difference between the fitted and the true delays $\Delta = \tau_{fitted} - \tau_{true}$. Table 7.3 tabulates the number of good experiments, N_{good}, and the mean and standard deviation of Δ for each of the six sets of simulations.

We can summarize the results of the simulations into five points. First, despite that the level of variability in the compact components is weak, the time delay in the

Table 7.3: Results of the simulations

N	δt (days)	δS (mJy)	N_{good}	$\langle\Delta\rangle$ (days)	$\sigma(\Delta)$ (days)
100	10	0.2	927	0.1	24
50	10	0.2	490	2.2	38
50	20	0.2	795	0.5	34
100	10	0.1	1000	0.6	7
50	10	0.1	805	1.7	29
50	20	0.1	981	1.9	13

system should be measurable if enough observations are given. Second, decreasing the measurement errors can significantly improve the likelihood of determining the time delay of the system unambiguously. Third, given the same measurement error, the success rate of determining the time delay is higher for the experiment containing more observations. Fourth, if the number of observations (N) is fixed, the time delay is more likely to be measurable from the experiment which has a larger total time span. In other words, the dataset with the largest Nδt would have a higher probability of revealing the time delay of the system. Fifth, in all cases, $\langle\Delta\rangle$ is close to zero, indicating that there is no significant bias in the time delay determined from observations.

7.3 Conclusion: The Prospect of Measuring the Hubble Parameter

We have found a multiply imaged variable source in the system from which the time delay of the system can be determined. Although the level of variability in this source is small, our simulations indicated that the time delay would be measurable by monitoring the flux variations in A and B (i.e., the images of this source) if the

monitoring project contained about 100 observations, taken once every 10 or 20 days. For a source as complicated as MG1131, obtaining an image with sufficient quality would require at least 30 minutes of integration time on source. This means that one would need at least 50 hours of observing time over 3 years to get an unambiguous time delay measurement of this system. Although the amount of observing time for measuring the time delay is somewhat demanding, it is not unreasonable.

As discussed at the beginning of this chapter, the time delay measurement is not the only element needed in the determination of H_o. A reliable lens model is also required. The extended structure in this system provides very tight constraints on the lens model. Although we have used two sophisticated lens potentials to model this system, we still could not find a model that is completely consistent with all the constraints in the system (see discussion in Chapters 4 and 5). The strong constraints imposed by the structure of the system should, we hope, eliminate (or reduce) the possibility of finding degenerate lens models for the system, hence, degenerate values of H_o. Of course, the redshifts of the system are also required for measuring H_o, and both z_s and z_l are currently unknown. The lens is detectable at both optical and infrared wavelengths, and the compact components of the system are visible at the infrared wavelength (Larkin et al. 1994; Annis 1992; and Hammer et al. 1991). Thus, it should be possible to determine z_s and z_l from the optical or the infrared observations.

In summary, although several pieces of information about this system are still needed before it can be used to measure H_o, we believe that these pieces of information are obtainable. We thus conclude that MG1131 should be seriously considered as a probe for measuring H_o.

References

Annis, J. A. 1992, Ap. J., **391**, L17.

Hammer, F., Le Févre, O., Angonin, M. C., Meylan, G., Smette, A. & Surdej, J. 1991, A&A, **250**, L5

Hewitt, J. N., Chen, G. H. & Messier, M. D. 1995, Astron. J., in press.

Hughes, P. A., Aller, H. D. & Aller, M. F. 1992, Ap. J., **396**, 469.

Larkin, J. E., Matthews, K., Lawrence, C. R., Graham, J. R., Harrison, W., Jernigan, G., Lin, S., Nelson, J., Neugebrauer, G., Smith, G., Soifer, B. T. & Ziomkowski, C. 1994, ApJ, **420**, L9

Lehár, J., Hewitt, J. N., Roberts, D. H. & Burke, B. F. 1992, Ap. J., **384**, 453.

Mason, C. R. 1986, Ap. J., **302**, L27.

Press, W. H., Rybicki, G. B. & Hewitt, J. N. 1992, Ap. J., **385**, 404.

Simonetti, J. H., Cordes, J. M. & Spangler, S. R. 1984, Ap. J., **284**, 126.

8. Summary

We have modeled the gravitational potential of the lens galaxy in MG1131+0456 using the total intensity maps of the system at 5, 8, and 15 GHz. Among the models we explored, we found a clear best model fitted to all radio images. The profile of the surface mass density of the best fit lens model has a substantial core radius ($0\farcs19 \pm 0\farcs07$) and declines asymptotically as $r^{-1.4\pm0.2}$. The images reconstructed from the best fit model closely resemble the observed images. The structure of the source inferred by the this model consists of a bright core, two radio lobes, and, possibly, a short jet; these features are consistent with the variations of depolarization ratios detected in the system. The flux of the central component predicted by the model is within one standard deviation of the measured flux of component D, so we can rule out the hypothesis that component D is the radio image of the lens galaxy within the limits imposed by our overall goodness of fit problem. As expected, the mass interior to the average radius of the ring ($\sim 0\farcs9$) can be determined very precisely and independently of the models used: $M(<\langle r\rangle) \sim 6.4\ h^{-1} D_{ol} D_{os}/(2 r_H D_{ls}) \times 10^{11} M_\odot$. Unfortunately, since both z_s and z_l are unknown, we cannot convert the mass into real physical units at this time.

Although our best lens model produces images which resemble the observed images closely at all frequencies, the χ^2 of the model is still much higher than expected, indicating that the model is not completely consistent with all the observational constraints in the system. This means that there is substantial room for improvement. The models which we have explored assume that the ellipticity of the ring is generated solely by the matter well outside the ring and that the surface density of the lens is circularly symmetric. Since the apparent isophotes of elliptical galaxies are known to exhibit ellipse-like feature, it is likely that potentials with circularly symmetric surface densities are not sufficient in representing the true mass distribution of the

lens galaxy. Thus, we believe that adding a non-circularly symmetric term to the surface mass density inside the lens would be the next step to take for improving the lens model. However, because of the presence of the two external galaxies, it would be important to keep the external shear field to account for the lensing effects due to these two galaxies. Together, we think that a model consisting of a lens with true ellipsoidal isodensity and an external shear would be the best model to try next.

We have shown that the structure of the rotation measure distribution in the lens galaxy can be obtained by measuring the difference in the rotation measure between a pair of images arising from the same source. We applied this technique to MG1131+0456. We found evidence that the rotation measure in the lens galaxy varies on the order of 200 to 400 rad/m^2 on a scale $\sim 1''$, indicating that the magnetic field in the lens galaxy is likely to vary on the same spatial scale. At the present time, there is essentially no information regarding the structure of the magnetic field in highly-redshifted galaxies, so our measurements on the structure of the rotation measure distribution in the lens galaxy, giving the first observational constraints for modeling the magnetic fields in these galaxies, are important. We believe that the same method should be applied to other gravitational lenses with linearly polarized emission. Collectively, the results obtained from these lenses should lead a new direction in understanding the magnetic field structure in high-redshifted galaxies.

We found that components A and B in the system are variable but only on the level comparable to the current measurement error. We made Monte Carlo simulations to examine the possibility of measuring the time delay by monitoring the flux densities of A and B. The results of the simulations indicated that, giving the current level of measurement error, the time delay would be measurable if the monitoring project contained about 100 observations, taken once every 10 to 20 days. For a source as complicated as MG1131+0456, each VLA observation requires at least 30 minutes of integration time. Hence, to yield a time delay measurement, one needs at least 50

hours of VLA time over 3 years to carry out the monitoring project. However, if the measurement error is reduced, fewer observations would be needed for determining the time delay. The main source of error of the flux density measurements at 8 GHz is the confusion from the ring emission, and that at 15 GHz is the residual atmospheric phase errors. Thus, observations with higher resolution taken at lower frequencies, such as low frequency VLBI or VLBA observations, may provide measurements with smaller errors.

One major observational problem of the system is that both z_s and z_l are not known. It is not possible to measure the Hubble parameter or to express the mass in the lens galaxy in real physical units without these values. The lens is detectable at both optical and infrared wavelengths, and the compact components of the system are visible at the infrared wavelengths. Thus, given enough observing time, one should be able to measure both z_s and z_l.

We have shown that the time delay in MG1131+0456 is measurable if enough observations are available. Therefore, with new optical or infrared observations, yielding z_s and z_l, and better radio monitoring observations, yielding the time delay measurement, MG1131+0456 has the potential to provide a reliable measurement of the Hubble parameter.