STUDIES OF TWO CLASS GRAVITATIONAL LENSES

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy
in the Faculty of Engineering and Physical Sciences

2005

Thomas David York

School of Physics and Astronomy

Th 26382 ✓

Contents

Total word count: 44912

List of Tables

List of Figures

Abstract

This work describes observations of two strong gravitational lenses found in the JVAS/CLASS radio surveys, JVAS B0218+357 and CLASS B0631+519, together with modelling of the mass distribution in B0631+519 and in a third object, CLASS B0128+437.

CLASS B0631+519 was observed with the VLBA and MERLIN at 1.7 GHz. The MERLIN data show that the lensed source consists of a doubly-imaged flat-spectrum region located between two steep-spectrum lobes, one doubly-imaged and the other quadruply-imaged. The VLBA observations were compared with data taken at the same frequency three years earlier. No compelling evidence for variability was found above the 2 σ level. The VLBA data also allowed the flat-spectrum core to be identified in the sub-components making up the flat-spectrum region.

NICMOS and ACS observations of B0631+519 were used to establish the lens galaxy position, and to estimate the redshift of the lens. The ACS data confirm that there are two galaxies present in this system, although one appears to have little effect on the lensed source. The near-infrared NICMOS data show an Einstein ring, which will be of use in constraining more complex mass models.

The gravitational lens JVAS B0218+357 is nearly a perfect target for attempts to measure Hubble's constant, H_0. Measuring the lens galaxy's location relative to the lensed images is a prerequisite for this objective. Data taken with the ACS indicate a galaxy position 60 mas west of the brightest image of the source. The resulting value of H_0 is 68 ± 6 km s^{-1} Mpc^{-1}. However, the position is sensitive to the spiral arms of the lens. When they are masked out, the measured galaxy position shifts by 16 mas to the south-west, giving $H_0 = 56^{+12}_{-15}$ km s^{-1} Mpc^{-1}. The optical position measurement can also be combined with constraints from VLBI observations of the lens; without masking of the spiral arms, the combined constraints produce $H_0 = 70\pm5$ km s^{-1} Mpc^{-1}, or 61 ± 7 km s^{-1} Mpc^{-1} with masking.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

T. York

School of Physics and Astronomy

University of Manchester

Jodrell Bank Observatory

Macclesfield

Cheshire

SK11 9DL

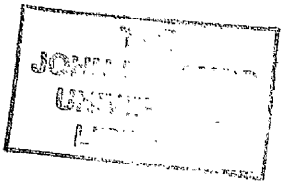
U.K.

September 2005

Copyright

The author

The author obtained an MPhys (Hons) degree from the University of Manchester in July 2001. He joined the Jodrell Bank radio astronomy PhD programme in September 2001.



Publications

Refereed Papers

York T., Jackson N.J., Browne I.W.A., Skelton J.E., *The Hubble constant from the gravitational lens CLASS B0218+357 using the Advanced Camera for Surveys* , 2005, MNRAS, 357, 124

York T., Jackson N., Browne I.W.A., Koopmans L.V.E., McKean J.P., Norbury M.A., Biggs A.D., Blandford R.D., de Bruyn A.G., Fassnacht C.D., Myers S.T., Pearson T.J., Phillips P.M., Readhead A.C.S., Rusin D., Wilkinson P.N., *CLASS B0631+519: Last of the Cosmic Lens All-Sky Survey lenses*, 2005, MNRAS, 361, 259

Conference Contributions

Finding the galaxy position in CLASS B0218+357, contributed talk to ANGLES network meeting, Bonn 2005

Acknowledgements

I am grateful to Professor A. Lyne for making available the facilities at Jodrell Bank Observatory. I received financial support from a PPARC research studentship grant. I would like to thank my supervisor, Dr. Neal Jackson, and my advisor, Prof. Ian Browne, for their patience, insight and assistance during the course of this work. Their help has been invaluable to me. I'd also like to thank Dr Andy Biggs, Dr Olaf Wucknitz, and Dr Leon Koopmans for their assistance and collaboration over the last few years. I'd like to thank the people who help to keep Jodrell Bank Observatory running, especially the admin staff, Janet, Pat and Sarah, and the MERLIN team. You're all great people.

Thanks to the many and various people who I've had to share a house with (and who've had to share a house with me); Martin Norbury, John McKean, David Champion, Iain Brown, Mike Keith, Sanhita Joshi, and the others, and thanks too to the many people I ran into by virtue of working in the same building as they did, including Stuart Lowe, Chris Wareing, Barnabas Rugyendo, Megan Argo and Matt Strong. You're great people too.

I'd like to thank my family for their support and love, including, but not limited to, my parents, Pam and Cliff, my sister Nella, my two Grans, Connie and Sheila, and my very great aunts, Margie and Shirley. And finally Cyril and David, for their love. I'm not sure where I'd be without you all.

There is something fascinating about science.
One gets such wholesale returns of conjecture
out of such a trifling investment of fact.

Mark Twain, in *Life on the Mississippi*

Conventions and Abbreviations

The conventional abbreviations for SI units and astronomical quantities are used.

The following abbreviations have been used in this thesis:

ACS – Advanced Camera for Surveys

AGN – Active Galactic Nucleus/Nuclei

AIPS – Astronomical Image Processing System

CCD – Charge Coupled Device

FWHM – Full Width Half Maximum

GMC – Giant Molecular Cloud

HI – Neutral Hydrogen

HST – Hubble Space Telescope

IGLOO – Interactive Gravitational Lens Optimisation Olgorithm

IRAF – Image Reduction Analysis Facility

ISM – Interstellar Medium

MERLIN – Multi-Element Radio-Linked Interferometer Network

NICMOS – Near Infra-red Camera and Multi Object Spectrometer

NOAO – National Optical Astronomy Observatory

NRAO – National Radio Astronomy Observatory

PSF – Point Spread Function

VLA – Very Large Array

VLBA – Very Long Baseline Array

VLBI – Very Long Baseline Interferometry

WFPC – Wide Field Planetary Camera (2)

WHT – William Herschel Telescope

Chapter 1

Introduction

1.1 Gravitational Lensing

The path of a light ray may be bent in a gravitational field. This follows from the equivalence principle, which states that an accelerating frame of reference cannot be distinguished from a uniform gravitational field by a physical experiment. In an accelerating frame, a light ray travelling perpendicularly to the direction of acceleration would be seen to curve, and so by the equivalence principle we expect the same in to happen a gravitational field.

When a distant source lies behind a nearer mass, such as a large galaxy, the light from the source may be distorted or multiply imaged by the galaxy's gravitational influence. This phenomenon is known as gravitational lensing, and it can be a powerful tool with which to investigate astrophysics. The effects on the light from the source can be inverted to determine the mass distribution of the lens; gravitational lensing is thus one of the few tools available to us that is directly sensitive to mass.

1.2 Historical perspective

The bending of light by massive objects through their gravitational effects was first extensively discussed (using Newton's theory of gravity) by Soldner (1801). Soldner correctly determined the gravitational deflection of light due to the gravitational field of a massive body according to Newton's theory, but made a number of mistakes in the derivation (Trumpler, 1923). Einstein computed the deflection angle using special relativity (SR) in 1911 (Einstein, 1911), and obtained the same result as Soldner. However, the introduction of the general theory of relativity (GR) permitted a more accurate calculation to be made, which predicted a deflection angle of twice the Newtonian value (Einstein, 1916).

To test the predictions of GR for the magnitude of the bend angle, the deflection of starlight close to the Sun was studied during the the solar eclipse of 1919 (to allow stars to be easily detected during daytime). Stars close to the solar limb were shifted during the course of the eclipse, producing a change in their separations compared to those measured six months earlier. The deflection angle produced by the Sun was found to be in agreement with the prediction of GR (Dyson, Eddington & Davidson, 1920). Einstein (1936) further predicted that a star acting as a gravitational lens could form a ring-shaped image of a background source, given excellent alignment. This result had already been found by Chwolson (1924), although he did not quantify the magnitude of the effect. The small predicted image separation of a few milli-arcseconds meant that this possibility did not attract much attention at the time due to the observational difficulties involved.

However, Zwicky (1937) pointed out that galaxies could act as gravitational lenses, producing multiple images separated by much larger distances than the cases of star-by-star lensing (known today as microlensing) described by Chwolson and Einstein. Further work by Refsdal (1964) showed that such lenses could be

used to determine the value of Hubble's constant, H_0. By the time the first lens was discovered by Walsh, Carswell & Weymann (1979), there were a variety of astrophysical applications known for the gravitational lens effect.

The first gravitational lens to be discovered, Q0957+561, was found by Walsh, Carswell & Weymann in 1979 (Walsh, Carswell & Weymann, 1979). The lens consists of two quasar images separated by 6 arcsec, and was identified after optical follow-up observations were made on targets in a 966 MHz radio source survey made with the 76-metre Lovell telescope at Jodrell Bank observatory (Porcas et al., 1980). Q0957+561 exhibits complex lensed structure at radio frequencies (Haschick et al., 1981). A variety of dedicated searches for strong gravitational lenses followed the discovery of Q0957+561; the largest systematic searches were the MIT-Greenbank (MG) survey (Bennett et al., 1986; Hewitt et al., 1989) which operated at 5 GHz with the Very Large Array (VLA) and discovered 5 lenses in a sample of several thousand sources, and the Hubble Space Telescope (HST) snapshot survey (Bahcall et al., 1992), which found 5 lenses in a sample of 502 sources.

Today, over 80 strong gravitational lens systems are known. Such systems have been used to probe the mass distribution inside galaxy clusters (eg; Kneib et al. 1993) and galaxies (eg; Muñoz, Kochanek & Keeton 2001) at redshifts beyond the reach of dynamics studies, to constrain cosmological models (eg; Chae 2003), to examine the ISMs of distant lenses (eg; Mediavilla et al. 2005) and to measure Hubble's constant (eg; Kochanek 2002).

1.3 JVAS/CLASS

The work in this thesis involves lenses discovered in the JVAS/CLASS surveys. The JVAS (Jodrell Bank/VLA Astrometric Survey; Patnaik et al. 1992; Browne et al. 1998) was originally designed to measure the positions of compact radio

sources that would be suitable for use as phase calibrators for interferometers in the northern hemisphere, especially MERLIN (Multi-Element Radio-Linked Interferometer Network; Thomasson 1986). Five new gravitational lenses were identified in this survey, and the previously known lens MG J0414+534 (Hewitt et al., 1992) was also rediscovered. Following the success of the JVAS as a lens search, a more advanced radio lens search was designed with statistical uses in mind, to be known as CLASS (Cosmic Lens All-Sky Survey; Myers et al. 2003; Browne et al. 2003). CLASS discovered a further 16 radio lenses, and became the largest statistically well-defined gravitational lens survey to date.

To make the rejection of candidates as straightforward as possible when relying on radio imaging alone, CLASS concentrated on sources with relatively flat radio spectra (as did JVAS, for different reasons). Flat-spectrum sources are generally a result of self-absorbed synchrotron emission in the centres of active galactic nuclei (AGN) and cannot be resolved even by high-resolution radio observations. Steep-spectrum sources tend to be extended, and would be harder to identify or reject as lenses (Kochanek & Lawrence, 1990). Flat-spectrum lens candidates can be rejected by increasing the resolution of the observations; if the surface brightness of the suspected images differ greatly from each other then the candidate is not a lens, because lensing conserves surface brightness.

The source selection criteria in CLASS were originally based on the 87GB catalogue (Gregory & Condon, 1991) combined with various low-frequency radio surveys (the Texas 365 MHz survey, Cotton 1976, and the WENSS 325 MHz survey, Rengelink et al. 1997), but later phases of CLASS were able to use the 1.4 GHz NVSS catalogue (Condon et al., 1998) combined with the GB6 5 GHz catalogue (Gregory et al., 1996). Sources selected using the later catalogues form the complete CLASS sample (which includes most of the sources from earlier CLASS phases, and all the JVAS sources). The criteria used to select sources for

the complete sample were

- Declination in the range $0° \leq \delta \leq 75°$)
- Galactic latitude $|b| \geq 10°$
- 5 GHz flux density (S_5) $\geq$ 30 mJy
- Spectral index between 1.4 GHz and 5 GHz ($\alpha^5_{1.4}$) ≥ -0.5, where $S \propto \nu^\alpha$

The first item restricts CLASS to the northern sky (a similar survey has been started in that part of the southern sky accessible to the VLA; Winn et al. 2000). The second requirement ensures that the Galactic plane is excluded from the survey, since at low frequencies confusion problems arise there. The third criterion ensures that the sources can be easily detected in snapshot observations with the VLA, and the fourth item selects sources with relatively flat spectra; the convention shown for the sign of the spectral index ν will be used throughout this thesis.

CLASS sources were observed with the VLA at 8.4 GHz in snapshot mode, with typical integrations lasting 30-40 seconds per source. The resulting datasets were mapped with an automatic script (Myers et al., 2003). Sources with morphologies inconsistent with multiply-imaged, flat-spectrum radio cores were rejected. The surviving sources were passed to MERLIN and the VLBA for observation at higher angular resolution. A source was selected for high-resolution follow-up with MERLIN and the VLBA[1] if it matched these criteria:

- Multiple unresolved radio components present, having mutual separations of between 300 mas and 15 arcsec
- A total 8.4 GHz flux ($S_{8.4}$) $\geq$ 20 mJy

[1]VLBA: Very Long Baseline Array

- A flux density ratio between the weakest and strongest components of less than 10:1

The statistically well-defined CLASS sample is a subset of the complete CLASS sample, and contains 13 of the total of 22 lenses found in the JVAS/CLASS. The complete CLASS sample consists of 13783 sources, of which 8958 form the statistically well-defined sample. The majority (2418) of the sources present in the complete sample that are not included in the statistically complete sample were rejected because they were found to have 8.4 GHz fluxes less than 20 mJy. It is the statistically complete sample that can be used to constrain cosmological models (Chae, 2003; Chae et al., 2004). CLASS lenses have also been used to constrain models of galaxy mass density profiles (eg; Muñoz, Kochanek & Keeton 2001) and to measure Hubble's constant (eg; Wucknitz, Biggs & Browne 2004; York et al. 2005). CLASS is likely to remain the largest statistically well-defined lens survey until the advent of next-generation radio instruments such as the Square Kilometer Array, as discussed in Kochanek (2003) and Koopmans, Browne & Jackson (2004).

1.4 This work

This thesis describes research into strong gravitational lensing that the author undertook from 2001 to 2005. The focus of the work has been on two radio lenses discovered in the JVAS/CLASS: B0218+357 (Patnaik et al., 1992; O'Dea et al., 1992; Patnaik et al., 1993) and B0631+519 (York et al., 2005). New observations were made of both lenses. Another lens, CLASS B0128+437, was investigated with a lens modelling code written as part of the work, but no new data were taken.

The work on CLASS B0631+519 involved analysis of 1.7 GHz VLBA data

taken at two different epochs, and of 1.7 GHz data taken by MERLIN. The VLBA data were combined with existing 5 GHz VLBA data to allow the source's flat-spectrum core to be identified with one of the observed radio components, and also to search for variability at 1.7 GHz. The MERLIN data were used to discriminate between two possible arrangements of the radio components in the lensed source; it is now clear that the background radio source has three distinct radio components, two of which are doubly-imaged and the other quadruply-imaged.

An R-band image taken by the William Herschel Telescope (WHT), and F555W, F814W and F160W images taken by the Advanced Camera for Surveys (ACS) and the Near Infrared Camera and Multi-Object Spectrometer (NICMOS) on board the HST were used to establish the lens system's brightness, the lens galaxy's position and other properties, and to estimate the redshift of the source. The HST images resolved questions surrounding the apparent presence of two lens galaxies detected through spectroscopic means by McKean et al. (2004). The HST images also indicated the presence of a near-infrared Einstein ring in B0631+519, very similar to that seen in JVAS B1938+666 (Patnaik et al., 1992; King et al., 1998). The optical lens galaxy position was used to constrain mass models of the system; the extra degrees of freedom made available by fixing the lens position were used to introduce an external shear to the model. An excellent fit to existing MERLIN data was obtained using a simple parametric lens model.

JVAS B0218+357 is a very bright radio lens, having a flux density at 1.4 GHz of over 1.7 Jy (Condon et al., 1998). B0218+357 consists of two images of a quasar core separated by only 334 mas. This system also contains the smallest known Einstein ring visible at radio frequencies, and has a measured time delay of 10.5 ± 0.4 days (Biggs et al., 1999). The lens galaxy is a spiral. The small image separation and the faintness of the lens compared to the bright quasar images

makes it difficult to measure the position of the centre of the lens. This position is degenerate with Hubble's constant for isothermal lens models (Lehár et al., 2000), so measuring the galaxy position would also produce a value for Hubble's constant.

The work described in this thesis on B0218+357 is an optical measurement of the lens galaxy position using the ACS. The data confirmed earlier suggestions that the separation between the images is lower in the optical than in the radio, a curious situation that seems to be due to differential absorption by a molecular cloud in the lens galaxy (Henkel et al., 2005). The ACS data also show that the lens galaxy is nearly face-on. The optical galaxy position was measured by subtracting PSFs from the quasar images and looking for the position about which the remaining flux was most symmetric. The resulting position was then combined with existing models of the mass distribution in B0218+357 that were produced by O. Wucknitz. The combined data fix the value of the Hubble constant.

The structure of this thesis is as follows. Chapter 2 discusses the theoretical principles of gravitational lensing and discusses ways of inverting the lens mapping, that is, of deducing the distribution of mass responsible for an observed lens configuration, as well as describing some of the properties of lens galaxies that lensing can be used to investigate. A computer code is developed to invert JVAS/CLASS lenses, and is tested on the CLASS lens B0128+437. Chapter 3 reports the details of the work undertaken on CLASS B0631+519, and chapter 4 describes the measurement of the galaxy position in JVAS B0218+357. Finally, chapter 5 concludes, and discusses possible future work on both lenses.

Bibliography

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

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

Biggs A.D., Browne I.W.A., Helbig P., Koopmans L.V.E., Wilkinson P.N., Perley R.A., 1999, MNRAS, 304, 349

Browne I.W.A., Wilkinson P.N., Patnaik A.R., Wrobel J.M., 1998, MNRAS, 293, 257

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

Chae K.H., 2003, MNRAS, 346, 746

Chae K.-H., Chen G., Ratra B., Lee D.-W., 2004, ApJ, 607, L71

Chwolson O., 1924, Astron. Nachr., 221, 329

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

Cotton W.D., 1976, ApJS, 32, 467

Dyson F.W., Eddington A.S., Davidson C.R., 1920, Mem. R. Astron. Soc., 62, 291

Einstein A., 1911, Annalen der Physik, 35, 898

Einstein A., 1916, Annalen der Physik, 49, 769

Einstein A., 1936, Science, 84, 506

Gregory P.C., Condon J.J., 1991, ApJS, 75, 1011

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

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

Henkel C., Jethava N., Kraus A., Menten K.M., Carilli C.L., Grasshoff M., Lubowich D., Reid M.J., 2005, A&A, 440, 893

Hewitt J.N., Burke B.F., Turner E.L., Schneider D.P., Lawrence C.R., Langston G.I., Brody J.P., 1989, LNP, 330, 147

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

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

Koopmans L.V.E., Browne I.W.A., Jackson N.J., 2004, New Astron. Rev., 48, 1085

Kochanek C.S., Lawrence C.R., 1990, AJ, 99, 1700

Kochanek C.S., 2002, ApJ, 578, 25

Kochanek C.S., 2003, ASPC, 291, 245

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

Lehár J., et al., 2000, ApJ, 536, 584

McKean J.P., Koopmans L.V.E., Browne I.W.A., Fassnacht C.D., Lubin L.M., Readhead A.C.S., 2004, MNRAS, 350, 167

Mediavilla E., Muñoz J.A., Kochanek C.S., Falco E.E., Arribas S., Motta V., 2005, ApJ, 619, 749

Muñoz J.A., Kochanek C.S., Keeton C.R., 2001, ApJ, 558, 657

Myers S.T., et al., 2003, MNRAS, 341, 1

O'Dea C.P., Baum S.A., Stanghellini C., Dey A., van Breugel W., Deustua S., Smith E.P., 1992, AJ, 104, 1320

Patnaik A.R., Browne I.W.A., Wilkinson P.N., Wrobel J.M., 1992, MNRAS, 254,655

Patnaik A.R., Browne I.W.A., King L.J., Muxlow T.W.B., Walsh D., Wilkinson P.N., 1993, MNRAS, 261, 435

Porcas R.W., Urry C.M., Browne I.W.A., Cohen A.M., Daintree E.J., Walsh D., 1980, MNRAS, 191, 607

Refsdal S., 1964, MNRAS, 128, 307

Rengelink R., Tang Y., de Bruyn A.G., Miley G.K., Bremer M.N., Röttgering H.J.A., Bremer M.A.R., 1997, A&AS, 124, 259

Soldner J., 1801, in Bode J.E., ed, Berliner Astronomisches Jahrbuch f ur das Jahr 1804, pp. 161-172

Thomasson P., 1986, QJRAS, 27, 413

Trumpler R., 1923, PASP, 35, 185

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

Winn J.N., Schechter P.L., Hewitt J.N., Patnaik A., Lovell J.E.J., 2000, AAS, 32, 1578

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

York T., et al., 2005, MNRAS, 361, 259

Zwicky F., 1937, PRL, 51, 290

Chapter 2

Gravitational lensing theory

2.1 Introduction

This chapter introduces some of the physics of strong gravitational lensing. The review articles by Blandford & Narayan (1992), Courbin, Saha & Schechter (2002) and Kochanek et al. (2004), and the monograph by Schneider, Ehlers & Falco (1992) can be consulted for additional detail. This chapter also describes the implementation of a computer code designed to model strong JVAS/CLASS lenses. The code is applied to the lens system CLASS B0128+437[1]

The main astrophysical objective in analysing any particular strong gravitational lens system is to use the information provided by the multiple images to infer both the mass distribution in the lens and the surface brightness distribution of the lensed source. This task is greatly simplified by making the following assumptions:

- Deflection angles produced by the lensing mass distribution are small. Strong lenses exhibit image separations of a few arc-minutes at most (in the case of lensing by a galaxy cluster), so this condition is fulfilled.

[1]The modelling work on CLASS B0128+437 has been partially published in (Biggs et al., 2004).

- Gravitational fields involved in lensing are weak, such that the Newtonian potential Φ of the lens is much smaller than c^2 (c will represent the speed of light throughout this thesis).

- Deflection of a light ray by a lens takes place at a single point along the ray path, on an imaginary plane positioned at the redshift of the lens. This is the so-called "thin lens" approximation. In reality the ray is bent continuously as it travels from source to observer. The scale over which a galaxy-mass lens has an appreciable effect on the path of the ray might be several tens of kpc, while the total length of a light path from source to observer is measured in Gpc, and so the condition holds for galaxy lenses. Systems in which there are multiple lens galaxies at different redshifts are usually modelled by representing each galaxy by its own lens plane.

- No significant source or lens motion with respect to the observer - the peculiar velocities of the lens and source with respect to the observer are assumed to be small compared to the speed of light. The effects of large lens or source velocities have been analysed by Wucknitz & Sperhake (2004).

- The angular scale of the lens is assumed to be large enough that wave phenomena such as diffraction do not significantly alter the interpretation of results from lensing. Diffraction effects are discussed further by Benson & Cooke (1979). This assumption holds for all known gravitational lenses.

The assumption of small angles allows use of the linearised angle formula,

$$x = D\theta, \tag{2.1}$$

to convert observed angular distances (θ) to linear distances (x) on a plane a distance D from the observer. Since there is no single inertial frame that contains

the lensed source, the lens galaxy and the observer all at rest, D in cosmological-scale lensing is generalised to an angular diameter distance, a quantity explicitly defined by the requirement that equation 2.1 remains valid for objects separated by cosmological distances. In a purely matter-dominated universe ($\Omega_M > 0$ and $\Omega_\Lambda = 0$), the angular diameter distance expressed as a function of redshift, $D(z)$, has the form

$$D = \frac{2c}{H_0} \frac{(1 - \Omega_M - \sqrt{1 + \Omega_M z})(1 - \sqrt{1 + \Omega_M z})}{{\Omega_M}^2 (1 + z)^2}, \tag{2.2}$$

where D has units of Mpc if Hubble's constant at the present epoch (H_0) is given in $\mathrm{km\,s^{-1}\,Mpc^{-1}}$and c is given in $\mathrm{kms^{-1}}$. It can be useful to separate angular diameter distance from its dependence on H_0, by defining the unscaled angular size distance d:

$$d = H_0 D. \tag{2.3}$$

For arbitrary values of the cosmological parameters Ω_M and Ω_Λ the angular diameter distances must be calculated numerically (Kayser, Helbig & Schramm, 1997). Angular diameter distances as functions of redshift for several cosmologies are plotted in figure 2.1.

For strong gravitational lenses in which there is a single lens galaxy, there are three distinct angular size distances that are involved. The distance from the observer to the lens (D_l), from observer to source (D_s), and from lens to source (D_{ls}). Over cosmological distances in an expanding universe,

$$D_{ls} \neq D_s - D_l. \tag{2.4}$$

All three distances must usually be calculated numerically in an arbitrary cosmology.

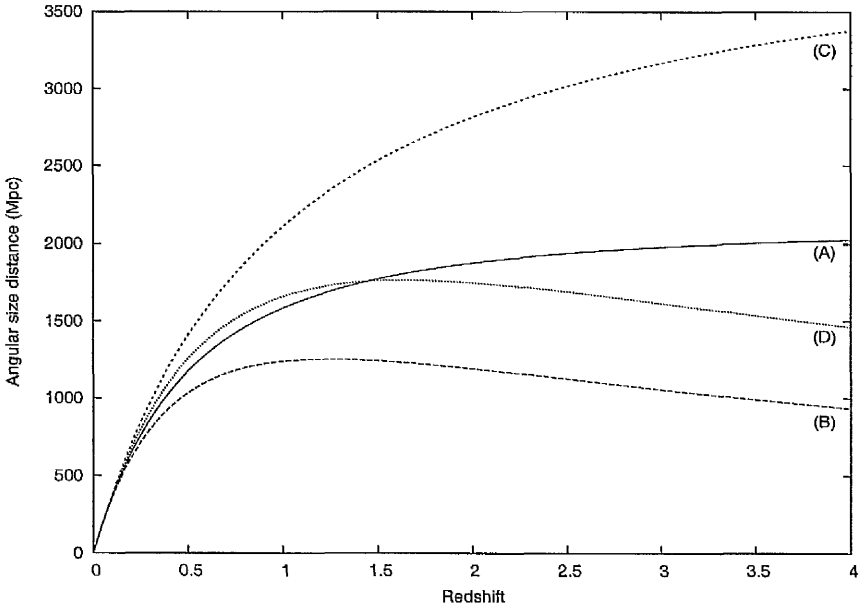


Figure 2.1: Angular diameter distance plotted as a function of redshift for: (A) an empty universe (Ω_M, Ω_Λ) = (0, 0); (B) an Einstein/de-Sitter universe (1, 0); (C) a universe containing only vacuum energy (0, 1); and (D) the present consensus cosmology, (0.27, 0.73). Angular diameter distances are inversely proportional to Hubble's constant, which has been set at 71 $\mathrm{km\,s^{-1}\,Mpc^{-1}}$for these plots.

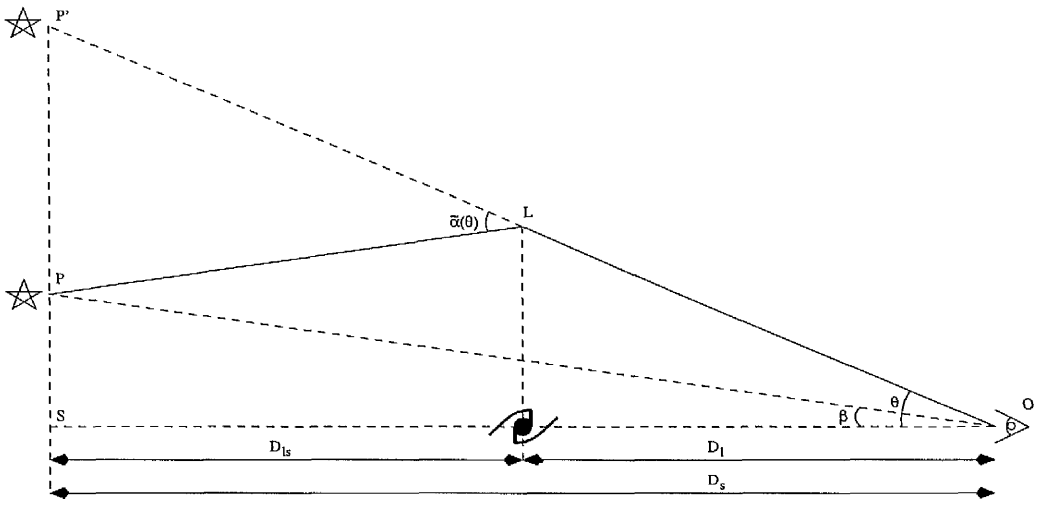


Figure 2.2: The geometry of a gravitational lens system. In order for the observer to see an image of the source, the bend angle $\tilde{\alpha}(\theta)$ must satisfy the geometry shown in this figure. Distances SP and PP' sum to distance SP'. When equation 2.1 is used to express these three lengths in terms of the angular diameter distances and the angles β and θ, equation 2.5 is recovered.

2.2 The lens equation

Figure 2.2 shows a schematic view of a strong gravitational lens system in which the image of a distant background source (such as an active galaxy) is lensed by a galaxy in between the source and observer. Since the thin lens assumption is made here, the lens, source and lensed images are regarded as lying on separate planes, all of which are normal to the line of sight. The angular position vectors[2] $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ are angular positions on the sky of the lensed image and the background source respectively, while $\tilde{\boldsymbol{\alpha}}(\boldsymbol{\theta})$ is the bend angle produced at the lens plane by the gravitational influence of the lensing galaxy. All three vectors are referenced to the optical axis which passes through the centre of the lens galaxy, denoted by line OS in figure 2.2.

Comparing similar triangles in figure 2.2 and making use of the small angle approximation yields

[2]A 2D angular position $\boldsymbol{\theta}$ will have components (θ_1, θ_2).

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{ls}}{D_s}\tilde{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}), \tag{2.5}$$

which is known as the lens equation, and introduces the normalised bend angle $\boldsymbol{\alpha}(\boldsymbol{\theta})$. General relativity provides the link between $\tilde{\boldsymbol{\alpha}}$ and the gravitational potential Φ of the lens,

$$\tilde{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = \frac{2}{c^2}\int \nabla_{2D}\Phi(r,l)dl, \tag{2.6}$$

where r is a 2D linear coordinate perpendicular to the ray path and l is a coordinate measuring distance along the ray path. ∇_{2D} represents the linear gradient operator over a 2D plane perpendicular to the ray path. The factor of 2 distinguishes the predictions of Newtonian gravity and general relativity. It is possible to set up a 2D potential function, the lensing potential $\psi(\boldsymbol{\theta})$, from which the bend angle can be calculated via a gradient operator,

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

in which the 2D gradient operator is defined over sky direction $\boldsymbol{\theta}$ rather than by linear position on the lens plane. It can then be shown that the lensing potential and the Newtonian potential are related by

$$\psi(\boldsymbol{\theta}) = \frac{2}{c^2}\frac{D_{ls}}{D_l D_s}\int \Phi(r,l)dl. \tag{2.8}$$

The Newtonian potential is connected to the 3D mass density $\rho(r,l)$ through Poisson's equation for gravitation,

$$\nabla_{3D}{}^2\Phi(r,l) = 4\pi G\rho(r,l). \tag{2.9}$$

Where the thin lens approximation holds, the lensing properties of a given mass distribution depend on the projected surface mass density rather than directly on ρ. Integrating Poisson's equation along the ray path and defining the

projected surface mass density $\Sigma(\boldsymbol{\theta})$ by

$$\Sigma(\boldsymbol{\theta}) = \int \rho(r, l) dl \tag{2.10}$$

gives

$$\nabla^2 \psi(\boldsymbol{\theta}) = \frac{8\pi G}{c^2} \frac{D_{ls} D_l}{D_s} \Sigma(\boldsymbol{\theta}), \tag{2.11}$$

where the Laplacian is defined over angular direction on the sky. We define the critical surface mass density, Σ_c,

$$\Sigma_c = \frac{c^2 D_s}{4\pi G D_{ls} D_l}. \tag{2.12}$$

A lens which contains no projected density greater than the critical density will cause image distortion, but not multiple imaging. Σ can then be normalised to the critical density; the dimensionless surface mass density κ, known as the convergence, is just

$$\kappa = \frac{\Sigma}{\Sigma_c}, \tag{2.13}$$

and is related to the potential by

$$\nabla^2 \psi(\boldsymbol{\theta}) = 2\kappa(\boldsymbol{\theta}), \tag{2.14}$$

and

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi} \int \kappa(\boldsymbol{\theta}') \ln |\boldsymbol{\theta} - \boldsymbol{\theta}'| \, d^2\theta'. \tag{2.15}$$

The lensing potential is useful when discussing gravitational lenses in terms of Fermat's principle.

2.3 Fermat's principle

Gravitational lensing can be understood through Fermat's principle (Schneider, 1985; Blandford & Narayan, 1986), i.e. the principle that images form along ray paths that are extrema of the light-ray travel time. Such points correspond to stationary points of a time delay surface defined over the immediate neighbourhood of the lens (where the small angle approximation is in force). The time delay towards direction $\boldsymbol{\theta}$, $\tau(\boldsymbol{\theta})$, is given by

$$\tau(\boldsymbol{\theta}) = \frac{1+z_l}{cH_0}\frac{d_l d_s}{d_{ls}}\left(\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\beta})^2 - \psi(\boldsymbol{\theta})\right), \tag{2.16}$$

where z_l is the source redshift. The dependence on H_0 has been made explicit by the use of the unscaled angular diameter distances d_l, d_s and d_{ls}. The first term on the RHS represents the delay due to the difference in geometrical length between the lensed and unlensed paths, and is zero for the direct path from source to observer. The second term is due to the gravitational retardation of photons (the Shapiro delay) and depends directly on the projected gravitational potential of the lens galaxy. The lens potential is typically a monotonically increasing function (towards zero at infinite radius), so the time delay at the centre of the lens due to the Shapiro delay alone is a maximum. Neither term on its own can produce multiple images, as can be seen by setting the lens potential in equation 2.16 to zero everywhere. The resulting time delay surface is purely parabolic and has a single minimum at the source position (figure 2.3). There is then only one image of the source. The Shapiro delay term alone would give rise to a single image at the lens centre, which is a time delay maximum.

At stationary points on the time delay surface, the condition

$$\nabla\tau(\boldsymbol{\theta}) = 0 \tag{2.17}$$

is satisfied. Substituting the expression for the time delay into this condition,

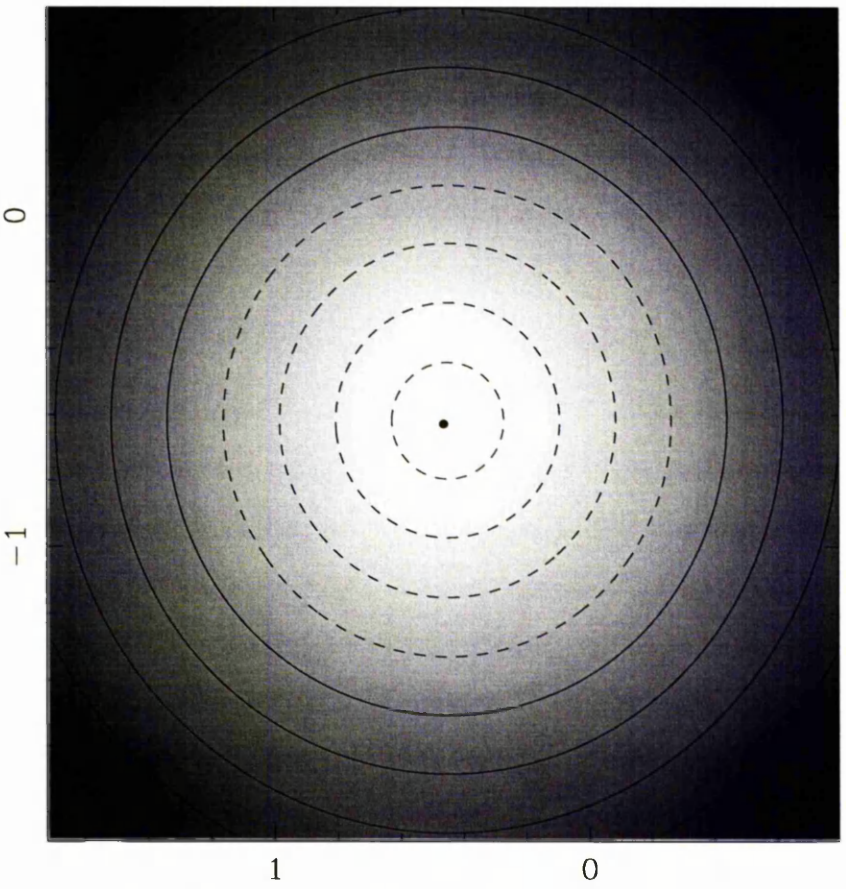


Figure 2.3: Time-delay surface in the absence of a lensing mass distribution. Darker colours represent later arrival times. The surface is a smooth parabola centred on the direction towards the unlensed source position. The filled circle located at the minimum marks the position of the single image that would be seen. The image position and the source position are coincident. Contouring is for clarity only.

we find that at any valid lensed image position $\boldsymbol{\theta}_i$ a relation exists such that

$$\boldsymbol{\beta} = \boldsymbol{\theta}_i - \nabla\psi(\boldsymbol{\theta}_i), \tag{2.18}$$

which is the lens equation for the i^{th} lensed image in a multiple-image system. Given an expression for $\psi(\boldsymbol{\theta}_i)$ the lens mapping can be locally inverted (turned into a mapping from source to lens plane), but it is not usually possible to globally invert the lens equation for a general lens model because for the more complicated models the expression for the bend angle is complicated, transcendental or both. If the lens model has circular symmetry, it is possible to solve the lens equation graphically for simple models with the aid of a diagram first devised by Young (Young et al., 1980). In this type of diagram the bend angle α is plotted against the radial image coordinate θ. From the lens equation (equation 2.18), images will be formed at points where

$$\alpha = \theta - \beta, \tag{2.19}$$

where β is the radial source coordinate. By plotting the LHS (which is fixed by the chosen lens model) and the RHS (which depends only on the source position) on the same (θ,α) coordinates, solutions of the lens equation can be identified as intersections of the LHS and RHS curves. The Young diagram is shown for a simple mass model in figure 2.3.

2.4 Image magnification

Gravitational lensing conserves surface brightness, but can alter the apparent solid angle spanned by a source on the sky. The magnification μ is defined as the ratio of solid angles subtended by an image (ω_{img}) and by the source (ω_{src}),

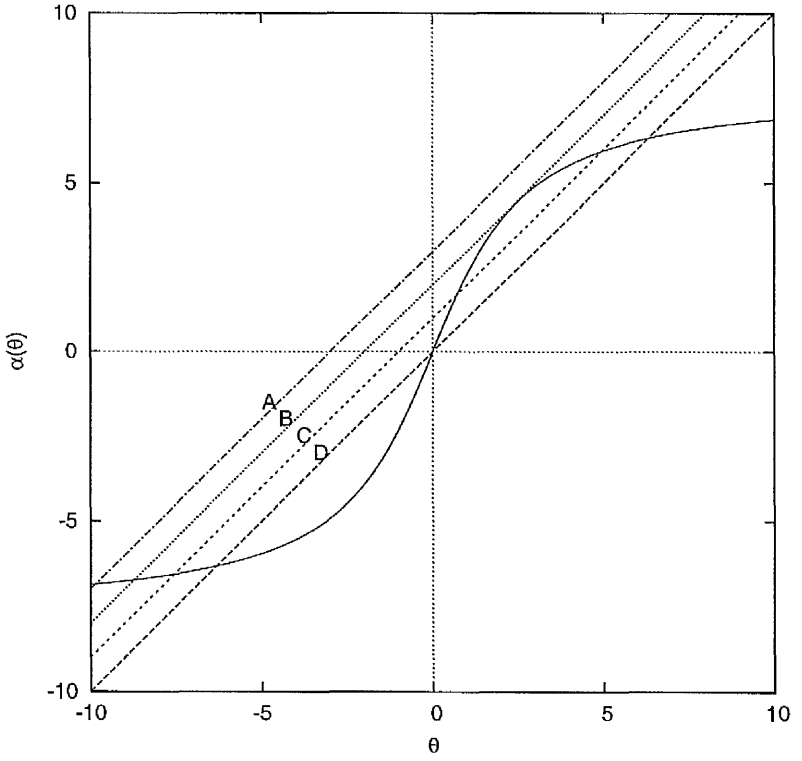


Figure 2.4: The Young diagram for a generic lens model. The quantity θ is the radial image coordinate (which may take on negative values). α is the radial bend angle. The curve marks the bend angle of a simple lens model. The straight lines A, B, C and D represent the RHS in equation 2.19 for three different source positions. Line A only intersects the bend angle curve at one location, and so one image is formed. Line B touches the curve at one point, forming a pair of merging images, while it crosses a second point (the location of the original image). Line C intersects the curve at three positions; two of the images are located very close to one another, while the original image remains on the opposite side of the lens. As the source moves towards the lens position, (line D), the new images separate until one is located at the centre of the lens and the other is equidistant with the original image. The axis units are arbitrary.

$$\mu = \frac{\omega_{img}}{\omega_{src}}. \tag{2.20}$$

Most of the lensed images will have magnifications greater than unity. The only exception is the case of an image close to the centre of a lens galaxy; this image can be demagnified. The total magnification for the lens is the sum of the magnifications of the individual images, and is always greater than unity for a galaxy or cluster mass distribution. The total magnification of the lens is important when using lens survey data to constrain cosmological parameters (Turner, Ostriker & Gott, 1984).

For an unresolved source the magnification (demagnification) effect appears as an increase (decrease) in flux density of the lensed image relative to the unlensed source, so that the magnification can be defined in terms of the image and source flux densities,

$$\mu = \frac{s_{img}}{s_{src}}, \tag{2.21}$$

where s_{img} is the lensed image flux density and s_{src} is the flux density of the source in the absence of lensing. The magnification produced by a gravitational lens can be more generally represented by a magnification tensor M,

$$M = \frac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{\beta}}. \tag{2.22}$$

M is the Jacobian matrix of the mapping from source to lens plane. It is easier to evaluate M's inverse A:

$$A = \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\theta}} = \frac{\partial \boldsymbol{\theta}}{\partial \boldsymbol{\theta}} - \frac{\partial \boldsymbol{\alpha}}{\partial \boldsymbol{\theta}}. \tag{2.23}$$

Clearly, A is the Jacobian matrix of the lens- to source-plane mapping. The components of A are given by

$$A_{ij} = \delta_{ij} - \psi_{,ij}, \tag{2.24}$$

where $\delta_i j$ is the Kronecker delta function and $\psi_{,ij}$ represents the second derivatives of the lensing potential ψ with respect to the ith and jth axes. This can be expressed in matrix form as

$$A = I - U, \tag{2.25}$$

where I is the identity matrix and U is a matrix formed from second derivatives of the lens potential (the Hessian matrix of the potential function). The magnification μ is given by the inverse of the determinant of A,

$$\mu^{-1} = |A|\,. \tag{2.26}$$

U is symmetric for thin-lens problems involving a single lens plane (Blandford & Narayan, 1986; Chae, Mao & Augusto, 2001), so A is also symmetric and the lens mapping can therefore displace or shear images but not rotate them. The effect of A can be seen by writing it as a sum of trace-free and diagonal matrices by defining the shear components γ_1 and γ_2,

$$\gamma_1 = \frac{1}{2}\left(\psi_{,11} - \psi_{,22}\right).\gamma_2 = \psi_{,12} = \psi_{,21} \tag{2.27}$$

The shear can also be represented by a magnitude γ and angle θ_γ,

$$\gamma_1 = \gamma\cos(2\theta_\gamma); \gamma_2 = \gamma\sin(2\theta_\gamma). \tag{2.28}$$

By making use of the relation between convergence and the lensing potential (equation 2.14), A can be written as

$$A = (1-\kappa)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \gamma_1 & \gamma_2 \\ \gamma_2 & -\gamma_1 \end{pmatrix}. \tag{2.29}$$

Equation 2.29 shows that convergence scales a source's image isotropically, while shear distorts it. A circular source will be magnified along two different axes (the eigenvectors of A) by factors inversely proportional to the eigenvalues of A, $\lambda_{\pm}$, which are

$$\lambda_{\pm} = 1 - \kappa \pm \gamma, \tag{2.30}$$

and the magnification, being the determinant of M, is the inverse of the product of these eigenvalues,

$$\mu^{-1} = (1 - \kappa)^2 - \gamma^2. \tag{2.31}$$

At points where either eigenvalue vanishes, the determinant of A is zero and the magnification is formally infinite, because the determinant is equal to the product of the eigenvalues; such points define a critical curve. Critical curves separate regions where either eigenvalue (and hence the magnification) changes sign. Images can be classified by the signs of the eigenvalues, which are known as partial parities (the sign of the magnification is known as the total parity). A negative partial parity indicates that the image is reversed or flipped along the direction of the corresponding eigenvector compared to the source. The partial parities of an image can be used to infer the nature of the time delay surface at that point. If both are positive, the image is a minimum, if both are negative, the image is a maximum, and if one is positive and the other is negative then the image is at a saddle point. In the absence of a lensing mass distribution the image formed of the background source has positive total parity, since it is a minimum in the time delay surface.

The projections of the critical curves on the source plane are known as caustics. Caustics divide regions of differing image multiplicities from one another. An infinitesimal source that crosses a caustic will gain or lose two additional images,

one of which will be a time delay minimum and the other a saddle point. A general theorem exists that proves that any transparent mass distribution whose density declines with radius must produce an odd number of images from any single source (Dyer & Roeder, 1980; Burke, 1981). Most observed lenses have an even number of images because one image (known variously as the "odd", "central" or "core" image) will form at the maximum in the lens potential, which will be close to the centre of the lens galaxy. Near the centre of a lens the convergence is maximised while the shear is minimised, and as a result the magnification of the odd image relative to the source is very much less than unity (the image is demagnified). This, and the possible confusing effect of any low-level radio emission from the lensing galaxy, makes detection of the odd image difficult; at present, the only known lens system in which it is likely that the odd image has been detected is PMN J1632-0033 (Winn, Rusin & Kochanek, 2004). Detection of the odd image strongly constrains the mass distribution in the central regions of the lens. However, the non-detection of odd images can still be used as a constraint on mass models (Rusin & Ma, 2001; Keeton, 2003; Winn, Rusin & Kochanek, 2004).

Most lenses give rise to three- or five-image configurations; the non-detection of the odd image in most has resulted in such systems being referred to as "doubles" or "quads" respectively. Where there are multiple lens galaxies or multiple lensed sources involved, it is possible for more than five images to be seen. Examples include CLASS B1933+503 (a total of ten images and three sources; Sykes et al. 1998) and CLASS B1359+154 (six images of a single source due to three lensing galaxies; Rusin et al. 2001). Figure 2.5 shows the time-delay surfaces, images and critical curves produced by a single galaxy-scale deflector whose projected mass distribution has circular symmetry. A circularly symmetric lens can produce a maximum of three images. The eigenvectors of the lens mapping point

in the radial and tangential directions. The innermost ("radial") critical curve corresponds to the vanishing of the radial eigenvector, and the outermost ("tangential") critical curve marks the vanishing of the tangential critical curve. The tangential critical curve maps to the innermost (tangential) caustic, which is a degenerate point at the centre of the lens; the radial critical curve maps to the outermost (radial) caustic. A source moving across the source plane first crosses the outer caustic, which creates a new pair of images from the opposing point of the radial critical curve. As the source approaches the inner caustic the innermost (odd) image moves closer to the centre of the lens while the other two images approach the tangential critical curve and become increasingly stretched. When the source is positioned directly behind the centre of the lens (i.e. on the tangential caustic) the odd image is located at the centre of the lens and is greatly demagnified, while the remaining two images merge into an Einstein ring that follows the tangential critical curve. The existence of the Einstein ring can also be inferred from the circular symmetry.

Lensing galaxies typically have some degree of ellipticity when seen in projection on the sky, which alters the time delay surface as shown in figures 2.6-2.8. Up to five images can be produced in this configuration. Again, there exist two critical curves, the tangential (outermost) and the radial (innermost), that map to the tangential (innermost) and the radial (outermost) caustics respectively, as in the case of the circular lens. The tangential caustic has an astroid shape. When the tangential caustic is crossed, two additional images are formed on top of the two created by crossing the radial caustic, for a total of five images. When a (small) source is positioned directly behind an elliptical deflector, an observer will see five discrete images rather than a continuous Einstein ring, since the bend angles only point towards the centre of the lens near the symmetry axes of the ellipse (figure 2.6, bottom panel). A final five-image case exists where a lens is

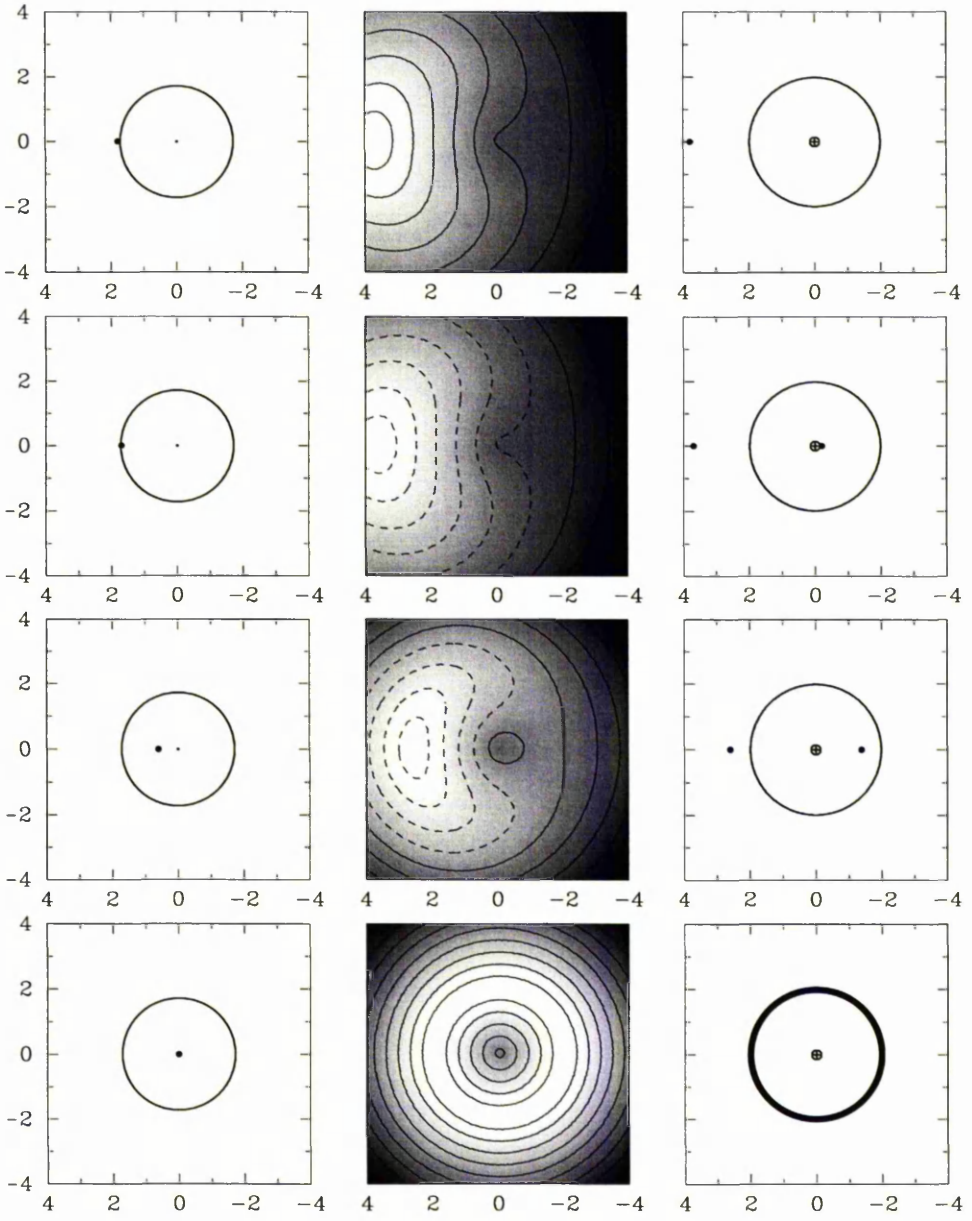


Figure 2.5: Lens systems generated by a circularly symmetric projected mass distribution. The surface is centred on the deflector. Starting at the top, each successive row shows the lens with the source moved closer to the deflector's centre, culminating in the formation of an Einstein ring. The left-hand column shows the caustics (curves) and source position (dot). The central column shows the time delay surface as a greyscale image where darker regions correspond to later arrival times; the surface is contoured for clarity. The right-hand column shows the image positions (dots) and critical curves.

very elliptical (or external shear is very high), and causes two of the cusps of the tangential caustic to protrude beyond the radial caustic. This is known as a "naked cusp" configuration, and it produces a group of three images on one side of the lens and a single image on the opposite side.

Groups of images close to critical curves should exhibit well-defined relations between their magnifications. A merging pair of point-like images formed by a source close to a fold caustic should have equal magnifications, while the sum of the (signed) magnifications of the close triplet of images produced by a naked cusp should be zero (Blandford & Narayan, 1986; Mao, 1992). In practice, these relations are not often observed to hold (Bradač et al., 2002, 2004). They are valid only if the potential can be expanded around the caustic as a Taylor series, i.e. they hold if the potential is smooth on scales of the close image separation. Thus, one possible cause of the magnification relations breaking down is the presence of mass substructure in the lens (on mass scales of less than 10^9 $M_\odot$). Hence, the magnification relations potentially allow the predictions of CDM structure formation models to be tested directly. Kochanek & Dalal (2004) investigate possible causes for magnification ratio discrepancies that are rooted in the lens galaxy environment (e.g. differential absorption), but conclude that the most likely explanation is that low mass halos are present close to the affected images, since the discrepancies are seemingly sensitive to the parities of the lensed images, affecting the brightest minimum less than the brightest saddle point (Witt, Mao, & Schechter, 1995).

Some alternative explanations of anomalous magnification ratios survive at present; Metcalf (2005) suggests that intergalactic CDM structure may also be responsible for producing flux ratio anomalies. Pen & Mao (2005) point out that this possibility can potentially be tested by looking for rotation of strongly lensed images that exhibit VLBI structure, since multi-plane lenses can rotate images

while single-plane lenses can only shear them.

2.5 Hubble's constant from lensing

In a gravitational lens system the total travel time along a ray is not known absolutely. However, the time delays between images are well-defined and, in principle, measurable. The delay between the i^{th} and j^{th} images, $\tau_{i,j}$, is

$$\tau_{i,j} = \frac{1+z_d}{cH_0}\frac{d_d d_s}{d_{ds}}\left(\frac{1}{2}\left|\nabla\phi(\boldsymbol{\theta}_i)\right| - \frac{1}{2}\left|\nabla\phi(\boldsymbol{\theta}_j)\right| - \phi(\boldsymbol{\theta}_i) + \phi(\boldsymbol{\theta}_j)\right). \tag{2.32}$$

In lens systems where the time delays $\tau_{i,j}$ have been measured, all redshifts are known and a reliable mass model has been used to obtain $\phi(\boldsymbol{\theta}_i)$ and its derivatives, the above equation can be solved for H_0, giving a one-step determination of the Hubble constant. That lensing could be used to constrain H_0 was first noted in Refsdal (1964), and was further discussed by Cooke & Kantowski (1975). The method is described as being one-step because it does not rely on the successive distance calibrations that form the basis of the cosmological distance ladder approach employed in the Hubble Key Project Freedman et al. (2001). The advantages of strong lensing in estimating H_0 are that the method is relatively direct and that calibration is relatively simple; the lensing determination of H_0 also samples cosmological distances rather than the local distances involved in the Key Project. The disadvantages are that the method relies on information that we cannot discover independently of lensing for cosmologically distant lenses, i.e. the mass distribution for any particular lens system. Lenses in which the mass distribution is well constrained are therefore most useful when attempting to measure H_0 with strong lensing techniques. There are also degeneracies in the method, for instance, adding a "mass sheet" having a constant convergence in

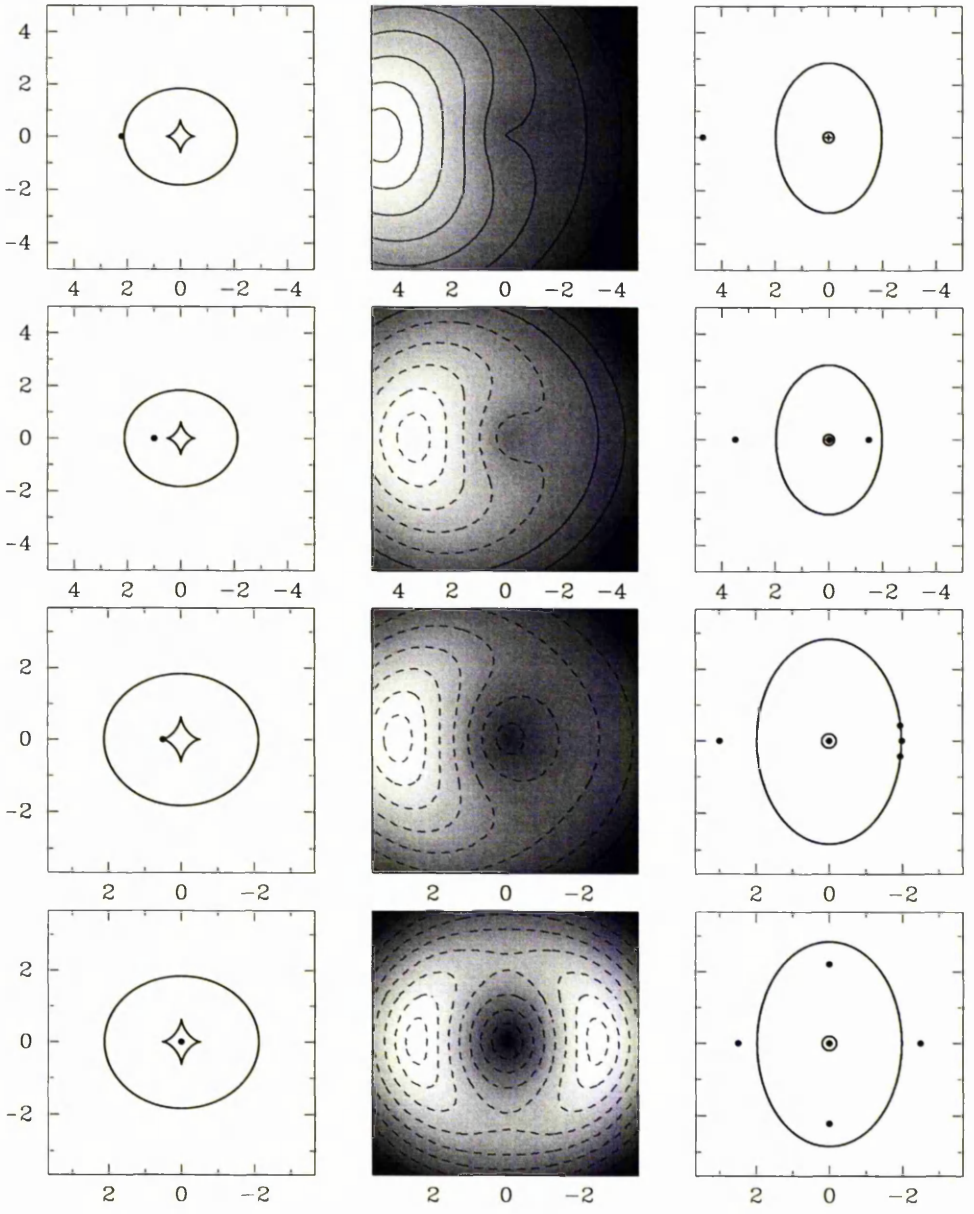


Figure 2.6: Lens systems produced by a mass distribution having elliptical symmetry (or a strong quadrupole) when projected on the sky. The columns are arranged as in figure 2.5. The major axis of the projected mass distribution used is aligned with the vertical axis; the ellipticity of the mass distribution is roughly 0.3. In this series of images the source is moved closer to the centre of the lens (along the minor axis of the lens) in each successive row. At each caustic crossing, an additional pair of images becomes visible. The image scale has been artificially expanded on rows showing five image systems to preserve clarity, but the model's mass scale remained the same for all source positions.

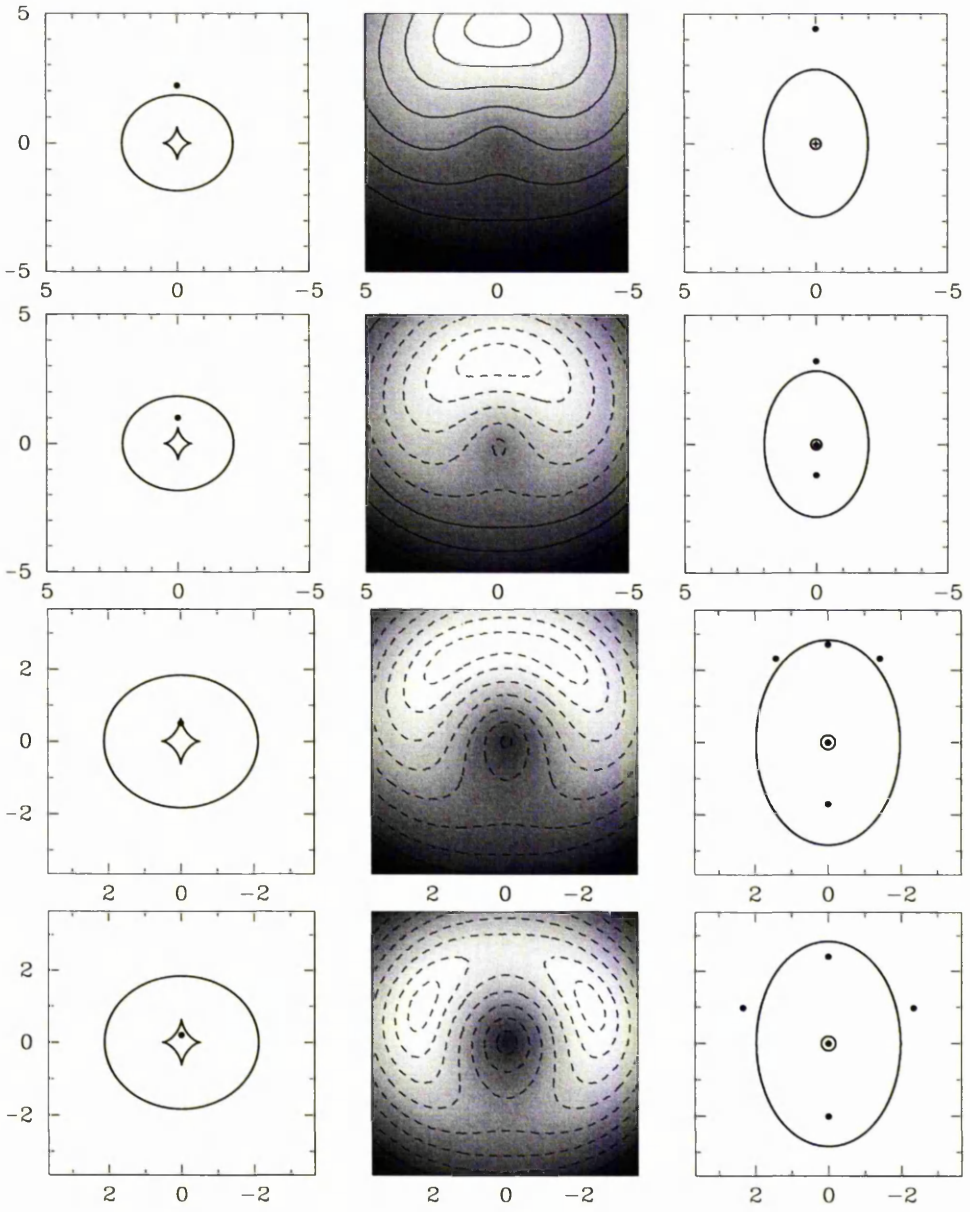


Figure 2.7: (Continued) Lens systems produced by a mass distribution having elliptical symmetry (or a strong quadrupole) when projected on the sky. The columns are arranged as in figure 2.5. In this case the source approaches the centre of the lens along the major axis of the lens. The image configurations formed are similar to those in the previous figure, but the source position lies on the same side of the lens centre as the majority of the images, the opposite from the case for a source approaching along the minor axis. Again, the image scale has been expanded to show the five image case clearly.

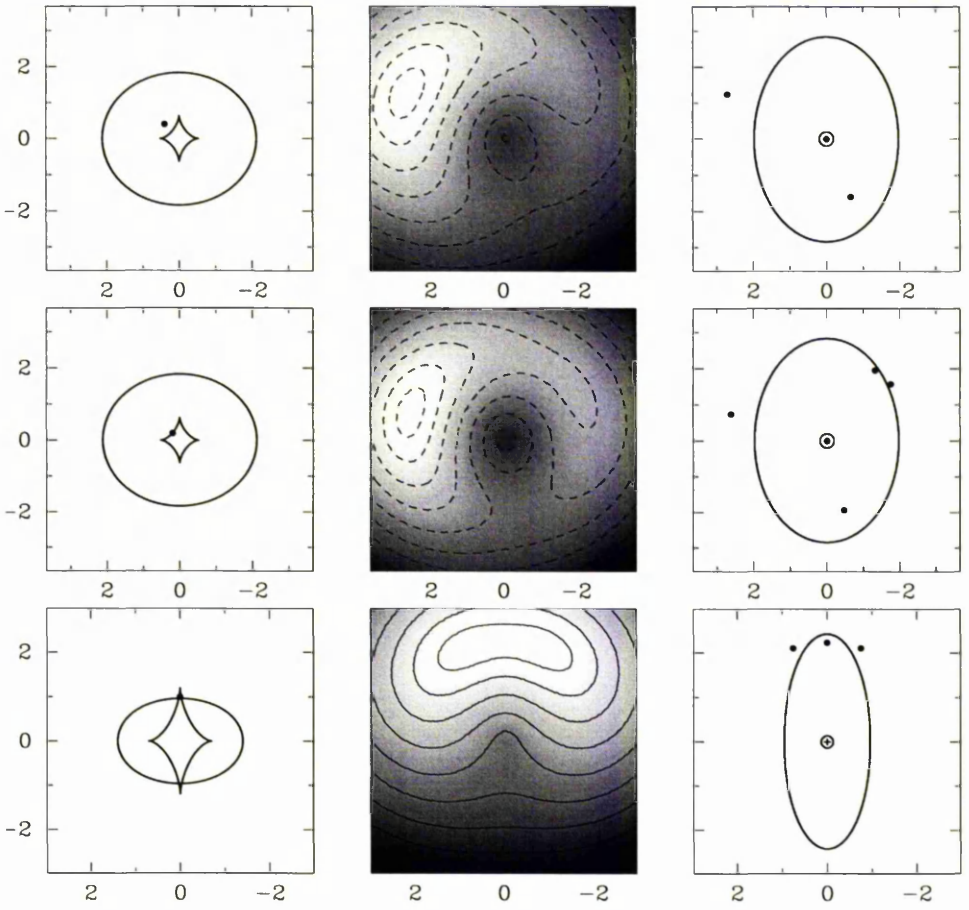


Figure 2.8: (Continued) Lens systems produced by a mass distribution having elliptical symmetry (or a strong quadrupole) when projected on the sky. The columns are arranged as in figure 2.5. The top two rows show the image configurations formed when the source approaches the centre of the lens by through a fold caustic; in the three image case the outer images do not line up with the centre of the lens, and in the five image case two of the images are seen close to one another across the tangential critical curve, while the other two images are located some distance away. The bottom row shows a rare configuration known as a "naked cusp" system; the tangential caustic extends beyond the radial caustic. Sources inside the tangential caustic but outside the radial caustic can form a triplet of images with no counter-image on the other side (and hence, to conserve parity, no odd image either). For all other source locations the behaviour is as shown on the previous pages. A naked cusp configuration can be created by a very elliptical projected mass distribution, or by a distribution of moderate ellipticity assisted by a correctly aligned external shear from other bodies beyond the Einstein radius of the lens.

front of the lens can rescale the time delays without affecting any of the lens observables, although the density required to introduce significant systematic errors in this way is not insignificant (Falco, Gorenstein & Shapiro, 1985; Gorenstein, Shapiro & Falco, 1988). Many of the degeneracies can be broken by including information from other sources, such as dynamical studies of the lenses (Treu & Koopmans, 2002).

Currently, there are less than ten cases of galaxy lensing for which the time delays are known. The first estimates of H_0 made using strong lensing were based on observations of Q0957+561 (Walsh, Carswell & Weymann, 1979), although that lens suffers greatly from degeneracies in the modelling (Chae, 1999). A joint analysis of 5 time-delay lenses by Kochanek (2002) found that values of $H_0 > 70\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ required mass distributions having constant mass to light ratios; by contrast, isothermal mass distributions produced low values of around $50\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$. Other lenses modelled with isothermal mass distributions have produced intermediate values of H_0, for example Fassnacht et al. (2002) measure a value for H_0 of 61-65 $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ in CLASS B1600+434, and for some lenses a high value of H_0 is favoured, for instance Koopmans et al. (2003) measure $75^{+7}_{-6}\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$for CLASS B1608+656.

2.6 Lens models

In this section I summarise two approaches to modelling mass distributions in gravitational lenses. A discussion of source-modelling techniques is left to the next section.

2.6.1 Parametric models

The reason for modelling a gravitational lens is usually to obtain information on the unknown lensing potential $\psi(\boldsymbol{\theta})$ using the observational constraints provided by multiple imaging and distortion. Since the degrees of freedom provided by the data are few in number for most lenses, it is common to adopt a simple functional form for $\psi(\boldsymbol{\theta})$ based on *a priori* information, and attempt to use lensing data to constrain its free parameters. A thorough catalogue of parametric models for gravitational lensing exists (Keeton, 2001); here I will discuss several models that are of interest.

There are many different parametric mass profiles available for modelling strong gravitational lenses, some of which are motivated by analytic galaxy dynamics (the singular isothermal models; Kormann, Schneider, & Bartelmann 1994), some by CDM simulations (the NFW profile; Navarro, Frenk & White 1996), some by the light profiles of galaxies (the Hernquist profile, a dynamical version of the de Vaucouleurs light profile; Hernquist 1990) and some, in part, by all of these (cusped profiles; Muñoz, Kochanek & Keeton 2001). Amongst the simplest models are the point mass and the singular isothermal models. The point mass model's lensing potential is

$$\psi(\boldsymbol{\theta}) = b^2 \ln|\boldsymbol{\theta}|, \tag{2.33}$$

where b is the Einstein radius of the lens, which is given by

$$b = \sqrt{\frac{4GM}{c^2}\frac{D_{ls}}{D_l D_s}}, \tag{2.34}$$

where M is the mass of the lens. The above formula can be generalised to any circularly symmetric projected mass distribution by regarding M as the total mass enclosed within the Einstein radius. Lensing can therefore be used to measure the mass of distant galaxies by measuring the image separation and

calculating the angular diameter distances given known source and lens redshifts and a cosmological model. The point mass forms a maximum of two images. Although the point mass model does not resemble galaxies, it can be used to investigate the possible effects of central super-massive black holes on strong lenses, especially on the odd image. Mao, Witt & Koopmans (2001) placed a point mass model at the centre of another parametric model representing the main bulk of a lens galaxy. They found that the black hole could suppress the formation of the third image for masses above a critical limit, or create an additional image (for a total of four images) for masses below the limit.

A useful model that can represent lens galaxies is the power law

$$\psi(\boldsymbol{\theta}) = b\,|\boldsymbol{\theta}|^{\eta}\,, \tag{2.35}$$

where η is the logarithmic slope of the potential. This profile was used by Rusin & Ma (2001) in their investigation of odd image constraints on mass profiles. The model has circular symmetry. The bend angle is

$$\alpha(\boldsymbol{\theta}) = \frac{b\eta}{\theta^{(1-\eta)}}\hat{\boldsymbol{\theta}}. \tag{2.36}$$

The Young diagram for various values of η is shown in figure 2.6.1.

In general, for logarithmic slopes given by $\eta > 1$, a lens of this type will produce one or three images. For logarithmic slopes given by $\eta <= 1$, the odd image does not form, and either one or two images will be seen. The dividing value of $\eta = 1$ corresponds to an isothermal mass distribution, which is named because it can be derived analytically by treating mass in a galaxy as an isothermal ideal gas in hydrostatic equilibrium. The isothermal profile forms the basis of a widely used set of lens models, which, in the case of spherical symmetry, is known as the singular isothermal sphere (SIS; Kormann, Schneider, & Bartelmann 1994). The SIS surface mass density distribution is given by

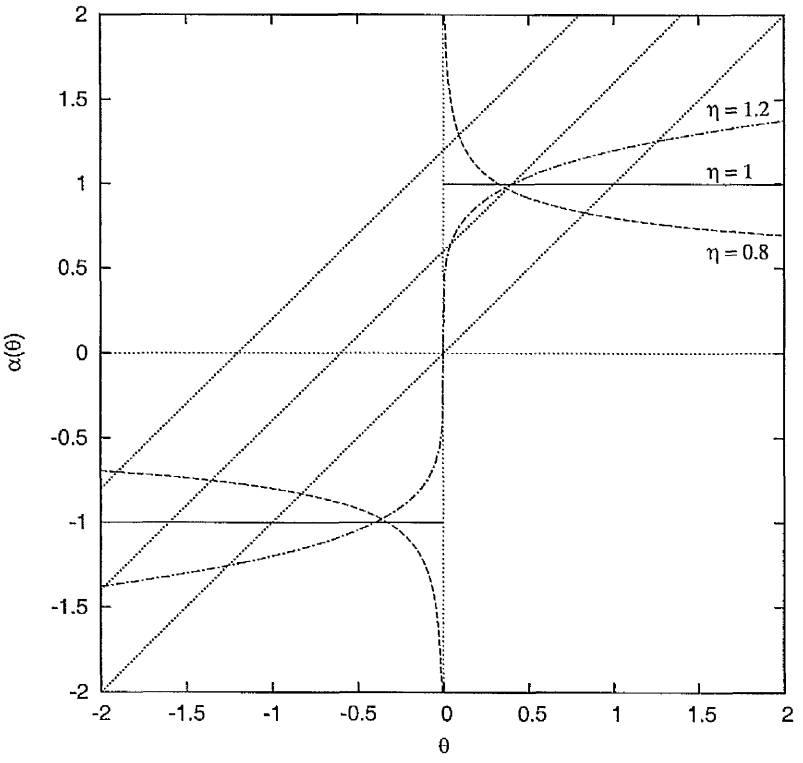


Figure 2.9: Young diagram for the model described by equation 2.35. The value of the power law index η is marked next to the relevant bend angle curve. For values of $\eta > 1$, either one or three images of the source are formed. For $\eta <= 1$, either one or two images are formed, and no odd image is seen because the bend angle diverges as it approaches the origin. The 45deg lines show three different source positions. The axis units are arbitrary.

$$\kappa(\boldsymbol{\theta}) = \frac{b}{2\,|\boldsymbol{\theta}|}. \tag{2.37}$$

where the normalisation differs from equation 2.35 in that b is the Einstein radius of the lens, i.e. the radius of the ring image formed when the source is placed directly behind the centre of the lens. For the singular isothermal models, the Einstein radius (and hence b) can be related to a dynamical property of the lens, the velocity dispersion σ_{iso} of the matter in the isothermal halo:

$$b = 4\pi\left(\frac{\sigma_{iso}}{c}\right)^2 \frac{D_{ls}}{D_s}. \tag{2.38}$$

σ_{iso} does not have to be the same as the velocity dispersion determined using light from the visible matter in galaxies, but in practice it seems to be very similar (Kochanek, 1994). The SIS produces a constant deflection angle away from the centre of the lens,

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = b\frac{\boldsymbol{\theta}}{|\boldsymbol{\theta}|}. \tag{2.39}$$

It is unusual to see the SIS used alone in modelling an observed lens, because real lenses have some degree of ellipticity. The elliptical variant of the SIS, the singular isothermal ellipsoid (SIE) is described in detail by Kormann, Schneider, & Bartelmann (1994). The SIE can produce either two or four images depending on the location of the source, because the odd image is again suppressed by the isothermal slope. The SIS/SIE has several virtues; its derivation is physical and analytic; it is fast to compute, and an isothermal mass profile appears on average to resemble lens galaxies (Cohn et al., 2001; Koopmans & Treu, 2003; Rusin, Kochanek & Keeton, 2003; Winn, Rusin & Kochanek, 2003). However, an SIS/SIE lens will never have a visible odd image; the odd image is completely demagnified. Furthermore, since the surface mass density of the isothermal profile declines proportionally to radius, the total mass of an SIS/SIE is unbounded. A

variety of models have appeared that extend the SIS/SIE, including the non-singular isothermal sphere/ellipsoid (NIS/NIE) which incorporates a constant-density core. The surface mass density of the NIS is

$$\kappa(\boldsymbol{\theta}) = \frac{b}{2\sqrt{|\boldsymbol{\theta}|^2 + s^2}}, \tag{2.40}$$

where s is the radius of the constant-density core (the "core radius"). The NIS/NIE was used for the plots in figures 2.5 and figures 2.6-2.8 because it allows the formation of an odd image despite having an isothermal profile, while remaining relatively simple to compute. The NIS/NIE can be further generalised by allowing the logarithmic slope of the surface mass density distribution to vary away from isothermality, producing a model known as the softened power-law elliptical mass distribution (SPEMD). The expressions for the lensing properties (potential and derivatives) of a SPEMD must be computed numerically; Barkana (1998) provides fast numerical solutions for the SPEMD. Neither the NIS/NIE nor the SPEMD derive from a physical argument, unlike the SIS/SIE (although the SIS/NIS was thought at one point to have empirical support from observations of galaxies; Lauer 1985). Additionally the NIS/NIE still has an infinite mass (the SPEMD may have a finite mass if its logarithmic mass slope is made steeper than isothermal).

Muñoz, Kochanek & Keeton (2001) introduced a new class of "cusped" lens models that replace the constant density core of the NIS/NIE with a power law, motivated by HST studies of the centres of galaxies (Faber et al., 1997) and by work on CDM simulations that indicated cuspy mass profiles in the centres of lenses (Navarro, Frenk & White, 1996). This inner power law changes over to an outer power law past a break radius. The outer power law is typically chosen to be steeper than isothermal so as to produce a lens model with a finite mass. The expression for the surface mass density is unfortunately quite complicated

and unintuitive (as are the expressions for most of the lensing properties), but the spherically symmetric 3D density distribution $\rho(r)$ introduced by Muñoz, Kochanek & Keeton (2001) can be expressed as

$$\rho(r) = \frac{\rho_0}{r^\gamma} \frac{a^n}{\left(r^2 + a^2\right)^{(n-\gamma)/2)}}, \tag{2.41}$$

where r is the (3D) radius to the point of interest, a is the break radius, n is the outer power-law index, γ is the inner power-law index, and ρ_0 sets the mass scale for the model. The model's lensing properties must be computed numerically. Chae (2002) provides fast solutions for the cusp models based on Fourier expansions, including support for the elliptical case. The cusped mass model can resemble many popular lens models when the appropriate parameters are fixed, including the isothermal models and the SPEMD. The main utility of the cusped model is then that it provides a smooth parametrized transition between various, hitherto quite separate, lens models. In situations where the extra degrees of freedom would be unnecessary, however, the computational penalties incurred by evaluating the cusp models mean that more specialised models can still be very useful.

The environments of lens galaxies can affect the lensing observables, principally by producing an additional shear. This shear must be incorporated into a realistic lens model; large shear galaxies can be represented directly using models such as the SIS, or the average effect can be reduced to a simple external shear model,

$$\psi(\boldsymbol{\theta}) = \frac{1}{2}\gamma_{ext}|\boldsymbol{\theta}|^2 \cos 2(\phi - \phi_\gamma), \tag{2.42}$$

where γ_{ext} is the external shear strength, ϕ is the polar angle coordinate in the (θ_1,θ_2) system, and ϕ_γ is the shear angle. This external shear is often misaligned with the internal shear produced by the ellipticity of the lens galaxy, and so to

successfully fit lens models an external shear is often essential. Keeton, Kochanek & Seljak (1997) indicate that the expected level of shear produced by large-scale structure is about 3%. The most extreme lenses seldom require an external shear strength of $\gamma_{ext} > 0.2$.

2.6.2 Non-parametric models

A parametric approach to lensing mass distributions will incorporate some assumptions that may not be true; for instance, that the lens has a constant mass to light ratio, or that lenses have a constant-density core. A failure of such models to fit the data could indicate a failure of these assumptions (and lead to a falsification of the model as a representation of the lensing mass distribution), but it could also represent a failure to allow the models sufficient freedom. For instance – as a trivial example – it will not be possible to satisfactorily model an elliptical lens with a singular isothermal sphere (a circularly symmetric lens model), even though a key assumption of the model profile (the isothermality) might be correct. The solution for this example is to give the parametric model greater freedom, by using a singular isothermal ellipsoid. However, some authors have argued that this lack of freedom is a general feature of the parametric approach that should be tackled by using a more flexible representation for the lensing mass distribution (Williams & Saha, 2000) and/or the source light distribution. The lensing potential (or the projected mass distribution) is represented as the sum of a set of N basis functions $F_i(\boldsymbol{\theta})$ and a set of corresponding weights a_i,

$$\psi(\boldsymbol{\theta}) = \sum_{i=1}^{N} a_i F_i(\boldsymbol{\theta}), \tag{2.43}$$

where the basis functions can be, for example, pixels (Saha & Williams, 1997; Williams & Saha, 2000), terms of a multipole expansion (Trotter, Winn & Hewitt, 2000) or angular Fourier components (Evans & Witt, 2003). However, unless

the number of basis functions is appropriately truncated, a non-parametric lens model can have many free parameters that the data will not constrain. Some *a priori* information can be used to reduce the freedom of the model, for instance, by requiring that successful models should have a positive surface mass density Williams & Saha (2000), but typically the number of degrees of freedom is still large and negative. There is also no known way to reject models that do not correspond to physically plausible 3D mass distributions. Thus, some types of non-parametric model may assign the lensing mass distribution more freedom than is physically possible.

Williams & Saha (2000) take an ensemble approach to non-parametric modelling. They find many acceptable pixellated mass distributions that reproduce the lensing observables, and use the ensemble of models to draw conclusions about the range of values allowed for interesting quantities, such as H_0. Williams & Saha (2000) have used this approach to measure H_0 with the lens systems PG1115+080 (Weymann et al., 1980) and B1608+656 (Myers et al., 1995), finding $H_0 = 61 \pm 11\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$. Williams & Saha (2000) argue that the ensemble approach to non-parametric modelling is useful in establishing solid limits on the permitted values of interesting quantities, such as H_0.

Evans & Witt (2003) consider scale-free lens models that have arbitrary angular structure but a simple radial power-law profile,

$$\psi(\boldsymbol{\theta}) = r^{\eta} F(\phi), \tag{2.44}$$

where r and ϕ are the components of $\boldsymbol{\theta}$ in polar coordinates and η is the logarithmic slope of the potential against radius. Evans & Witt (2003) describe the angular part of the potential $F(\phi)$ with a Fourier series ($F(\phi)$ is periodic). They find that in the isothermal case ($\eta = 1$), the image positions and magnifications are linearly related to the Fourier components describing the lens model. This

allows rapid inversion and fitting of lens models. The degrees of freedom of the model can be controlled by adding or removing terms in the Fourier series; the Evans & Witt (2003) approach therefore also differs from the ensemble approach in that it can optionally have a non-negative number of degrees of freedom. Evans & Witt (2003) apply their technique to show that some lenses previously considered to have flux density ratios distorted by substructure may not have been allowed sufficient freedom in the angular part of their potential. However, they find that the CLASS lens B1422+231 cannot be satisfactorily modelled with smooth potentials, in agreement with previous work on the system (Mao & Schneider, 1998; Bradač et al., 2002).

Parametric models treat a galaxy as a smooth mass distribution, and do not attempt to represent mass substructure. Since the lens potential allows superposition, mass substructure can be simulated with parametric models by adding in small perturbations to a smooth model; for example, Fassnacht et al. (1999) added a potential perturbation when modelling the lens CLASS B2045+265. Koopmans (2005) separates the smooth and non-smooth parts of the potential, and represents the former with a standard parametric model. Modelling mass substructure has become increasingly important in lenses such as CLASS B0128+437 (Biggs et al., 2004), where VLBI observations provide enough constraints to show deviations from a smooth parametric mass model.

2.7 Modelling resolved versus unresolved lenses

Most lenses discovered in the JVAS/CLASS show radio structure consisting of a limited number of point-like images of the central regions of the source, typically an AGN. The constraints on the lens model are limited in such systems. For a system with N unresolved images generated by a single lensed source, there

are $4(N-1)$ independent constraints on the mass model derived from the relative image positions, the image flux density ratios and the time delays. An additional two constraints can be obtained by fixing the lens galaxy position with the aid of optical/near-infrared data. For a double, this produces a maximum of 10 constraints, while for a quad there can be 14 constraints available. In practice, most systems do not have measured time delays and many do not have established galaxy positions. As a result, a double image lens usually provides barely sufficient constraints on a simple isothermal model with ellipticity or an external shear. In some systems the core regions magnified by the lens can be resolved into radio sub-components using VLBI (so-called "VLBI substructure"), and this multiplies the number of constraints available substantially; for instance, the CLASS lens B0128+437 (Phillips et al., 2000) has four images, each of which can be resolved into three sub-components, giving a total of twelve lines of sight through the lens system (ignoring extended emission from the source's host galaxy; Biggs et al. 2004). Biggs et al. (2004) use these additional lines of sight to identify the possible presence of mass substructure in the lens. For many JVAS/CLASS lenses the modelling must be more conservative due to the limited constraints available, restricting itself to measuring the total mass enclosed by the lens and the position angle and strength of any external shear that might be present.

Some JVAS/CLASS lenses exhibit extended structures at scales comparable to the size of the Einstein radius of the lens, usually at radio or near-infrared wavelengths. The JVAS lens B0218+357 exhibits a clear Einstein ring at 1.4 GHz (Patnaik et al., 1992; O'Dea et al., 1992; Patnaik et al., 1993; Biggs et al., 2001), while JVAS B1938+666 (Patnaik et al., 1992; King et al., 1998) shows an infrared Einstein ring. Although both systems show radio images of compact AGN cores, modelling using only those neglects the substantial number of constraints that

the extended emission makes available. The first attempt to make use of the extra constraints was the "Ring Cycle" algorithm (Kochanek et al., 1989). The method works by making use of the fact that images of the same point on the source should have the same surface brightness as each other, since lensing conserves surface brightness. The method projects a radio map of the lens back to the source plane, and associates each projected pixel with the nearest source pixel. Some source pixels will be associated with multiple image pixels (i.e. will be multiply imaged) and for those the method computes a fit statistic that is proportional to the dispersion between the fluxes of the corresponding image pixels.

The ring cycle's main drawback lay in the fact that no map of the sky derived from observation is a true surface brightness map. Instead the true sky surface brightness distribution is convolved with an instrumental response (the point-spread function, or PSF, in optical parlance, or the "dirty beam" in radio). To tackle this for radio data, Kochanek & Narayan (1992) combined the ring cycle with the CLEAN algorithm, which is used to deconvolve radio images produced by interferometers with partially filled apertures (Högbom, 1974; Clark, 1980). The hybrid algorithm, known as LensCLEAN, was applied to various resolved lenses. It was later modified by Wucknitz (2004), who eliminated several biases and defects in the original LensCLEAN approach. This improved LensCLEAN method has been applied to JVAS B0218+357 (Wucknitz, Biggs & Browne, 2004), to invert the lens and determine H_0.

A complementary approach to Kochanek & Narayan (1992) was taken by Wallington, Kochanek, & Narayan (1996), who combined the ring cycle with the Maximum Entropy Method (MEM; Bryan & Skilling 1980; Skilling & Bryan 1984), to produce LensMEM. LensMEM represents the source as a pixellated light distribution and employs a regularising term to enforce smoothness on the best-fit source model. The method consisted of three nested cycles of non-linear

optimisation; one for the source, one for the mass distribution and one for the strength of the regularisation. The resulting method was computationally intensive. Warren & Dye (2003) replaced the non-linear regularisation term used by Wallington, Kochanek, & Narayan with a linear term that allowed the the source cycle to be eliminated and replaced with matrix inversion. Warren & Dye also noted that in some cases the multiple imaging afforded by lensing would allow the regularisation term to be removed. The resulting fast method for modelling of resolved lenses was described by Warren & Dye as "semi-linear" lens inversion (since the lens equations are still generally non-linear in terms of the mass distribution parameters, but the inconvenience caused by the source parameters is removed). Koopmans (2005) has recently introduced another technique based on semi-linear lens inversion, known as gravitational imaging; in this method, the lensing potential is split into a smooth parametric term and a pixellated term; this allows evidence of CDM substructure to be identified more directly.

2.8 Modelling typical JVAS/CLASS lenses using a new code

As shown in the previous section, there are now several effective techniques available for the inversion of the lens mapping, that is, for the determination of the mass distribution responsible for an observed set of lensed images. In the case of lenses discovered during the CLASS and JVAS surveys, the constraints consist mainly of a set of point-like images having measured flux densities and positions. The lens theory described in the previous pages was applied in a new lens modelling code designed to model JVAS/CLASS lenses. The main requirement was the ability to fit complicated parametric mass models to the constraints available in typical JVAS/CLASS lenses.

2.8.1 Inverting the lens equation

It is often necessary in modelling a lens to invert the lens equation; that is, to find the (possibly multiple) images of a given source position. In other words, given the bend angle function and a source position $\boldsymbol{\beta}$, we wish to find all the roots of a vector function $\mathbf{T}(\boldsymbol{\theta})$, which is defined by

$$\mathbf{T}(\boldsymbol{\theta}) = \nabla\tau(\boldsymbol{\theta}) = \boldsymbol{\theta} - \boldsymbol{\beta} - \boldsymbol{\alpha}(\boldsymbol{\theta}). \tag{2.45}$$

It is not possible to accomplish this analytically for a general lens equation because the lens equations will usually be non-linear, so numerical methods must be employed. First, the roots of the lens equation (the image positions) must be bounded. This can be done by dividing the image plane into regions (a tiling) and projecting the boundaries of those regions back to the source plane. Regions will contain an image if their projections contain the source. Where there are multiple images, multiple projected regions will contain the source. Typically, the regions used are triangular, since triangles remain triangles under the lens mapping (Schneider, Ehlers & Falco, 1992). To avoid multiple images ending up in a single region, either the tiling can be made denser or an adaptive technique can be used to place more triangles at points where we might expect to see several images close together. Since images merge across critical curves, we can repeatedly subdivide any triangles that cross critical curves to isolate roots efficiently (Keeton, 2001). A triangle crosses a critical curve if the magnification (equation 2.31) changes sign between any of the vertices. An example of the tiling subdivision implemented in the code is shown in figure 2.8.1.

With the roots bounded on the source and lens planes, the lens equation can be locally inverted for each root by applying a numerical root-finding technique, such as Newton's method (Press et al., 1992). The vector form of Newton's method for the lens problem is

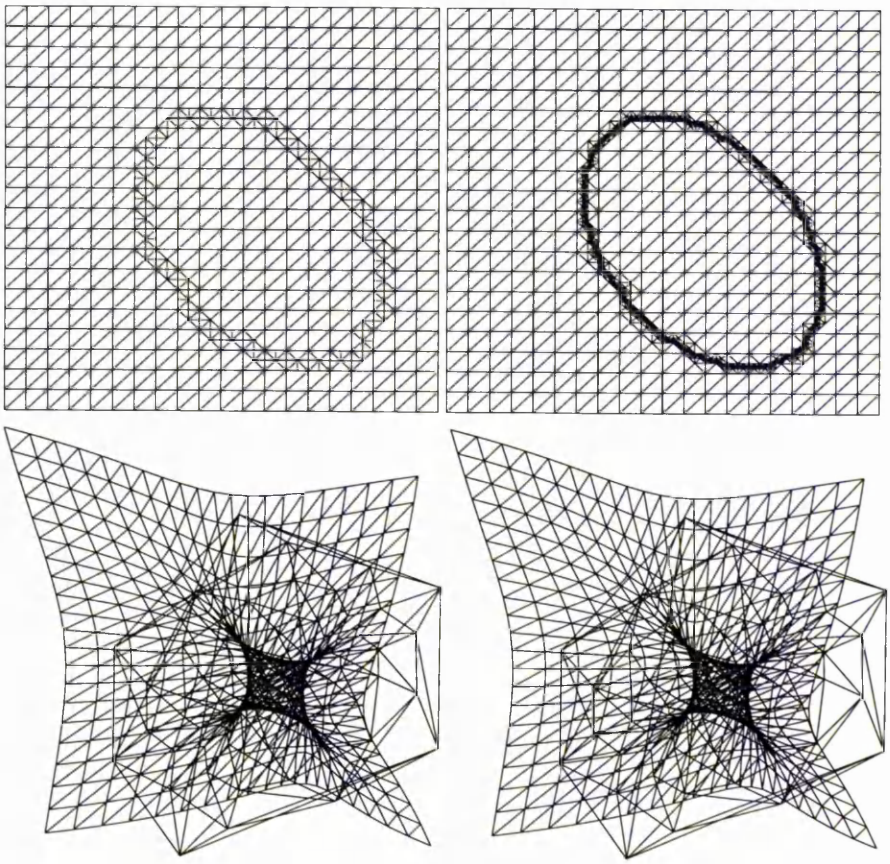


Figure 2.10: Subdividing the image plane into triangles. The tilings for a representative lens model are shown at increasing levels of subdivision from left to right. The top row of images shows the tiling on the lens plane, while the bottom row shows the same tiling projected back to the source plane. The increased resolution around critical curves is clearly evident, but is harder to see on the source plane because the subdivided tiles lie very close to the caustic and are shrunk in a direction perpendicular to it.

$$\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n - M(\boldsymbol{\theta}_n)\left(\boldsymbol{\theta}_n - \boldsymbol{\beta} - \boldsymbol{\alpha}(\boldsymbol{\theta}_n)\right), \tag{2.46}$$

where $\boldsymbol{\theta}_n$ is the best estimate for the image position, $\boldsymbol{\theta}_{n+1}$ is an improved estimate, and $M(\boldsymbol{\theta}_n)$ is the magnification matrix defined in equation 2.22. The improved image positions derived from iteratively applying the above equation can then be used in calculating the image plane χ^2 with equation 2.50. It is also desirable to subdivide tiles iteratively around the centres of lens galaxies, since the odd image will form close to such points for non-isothermal profiles.

The inversion procedure implemented in the code is as follows. First, a regular tiling of triangles is constructed on the lens plane. The current mass model is sampled once at each triangle vertex (so even though four triangles might share a single vertex, that vertex is still only sampled once). Each sample provides the 2D bend angle $\boldsymbol{\alpha}$ and the inverse magnification tensor A at the sample position. Optionally the potential is also returned, although it is not used in the inversion (and for some models, such as the cusped mass models, calculating the lensing potential is extremely involved and slow).

Once all the tiling vertices have been sampled, each triangle is checked to see if the magnification has a consistent sign on all three vertices. If the magnification changes sign across a triangle, then that triangle is lying across a critical curve and must be subdivided if the images of sources lying close to caustics are to be properly identified. The final subdivision pattern used is shown in figure 2.11.

Since the initial tiling is quite coarse (usually the lens is spanned by 20 or 40 triangles along each axis) the code could miss some tricky subdivisions, as shown in figure 2.12. To combat this while avoiding the increased demands of a greater initial resolution, the code will "speculate" and subdivide all triangles at least once, whatever the signs of the magnification at their vertices. If the vertices of the new triangles still all have the same magnification, the subdivision

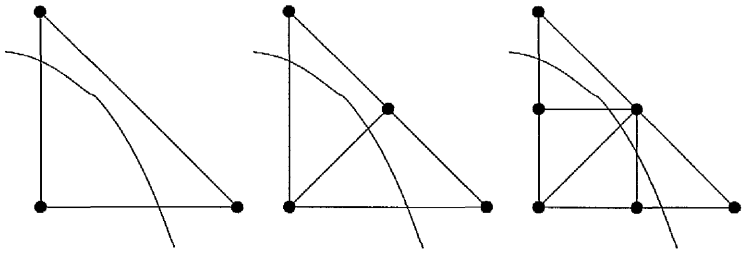


Figure 2.11: Subdivision of a triangle spanning a critical curve. At each subdivision step (of which two are shown here) the triangle is divided into two new triangles with the addition of a new vertex on one of the existing edges of the triangle. This subdivision method will not produce degenerate triangles on the lens plane.

is abandoned and the original triangle restored. This is usually sufficient to completely eliminate undivided tiles that still contain critical curves.

The number of subdivision loops and the initial tiling resolution are under user control. The code uses the positions of the observed images to decide what the "plate scale" of the tiling should be (although this can also be adjusted by the user). Once the lens plane is thoroughly subdivided, the sampled bend angles are used to project the triangle vertices back to the source plane. The code keeps track of the connections between each lens plane triangle and its counterpart on the source plane.

The source position is tested against all source plane triangles, and for those that enclose the source the corresponding triangles on the lens plane are flagged. Each source/image plane triangle pair is then checked for a solution to the lens equation using Newton's method as described before. Typically the method converges very rapidly for images at minima or saddle-points of the time-delay surface (within five iterations). For an odd image (a maximum) the rate of convergence depends on the slope of the mass profile, and increased subdivision may be necessary to prevent the solution diverging.

Profiling of the code confirms that most of the computer time consumed by the inversion step is spent calculating the potential and derivatives for all but the

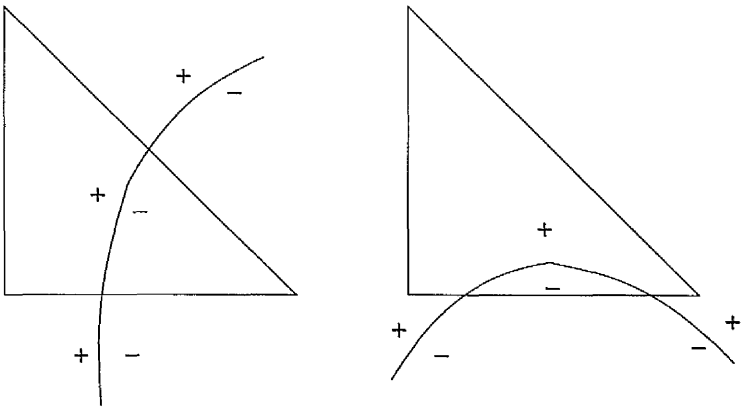


Figure 2.12: In the case on the left, the triangle will be subdivided because the critical curve separates one of the triangle's vertices from the other two. As the ± signs indicate, the isolated vertex will have negative total parity while the other two vertices will have positive total parity. In the case on the right, the triangle will not be subdivided, even though the critical curve enters the triangle. This can lead to discontinuities in the critical curves where some triangles have been subdivided but their immediate neighbours have not.

most trivial mass models. The system can find all the images produced by an SIE lens in roughly five milliseconds on a typical desktop PC, with four subdivision levels applied to an initial 20x20 triangle tiling.

2.8.2 Mass models

The code allows an arbitrary combination of mass models to be used. This is possible because the potential, and hence its derivatives, obey the principle of superposition. The lensing potential generated by a sum of mass models is just the sum of the individual potentials generated by each mass model. The code currently supports the following mass models:

- A point mass lens
- The SIS/SIE/NIS/NIE family of models described by Kormann, Schneider, & Bartelmann (1994)
- An external shear

- A constant-convergence mass sheet
- A linear convergence mass gradient
- An elliptical potential (eg. Blandford & Kochanek 1987)
- The SPEMD models of Barkana (1998), using public numerical codes described in that paper
- The cuspy models of Muñoz, Kochanek & Keeton (2001), using the numerical solutions of Chae (2002)

The addition of new models is simplified by the modular arrangement of the code.

2.8.3 Evaluating a lens model

There are many possible ways to evaluate the ability of a given lens model to reproduce the lens data (Kochanek, 1991). When the images are (as in most CLASS and JVAS lenses) effectively point-like, the general procedure is to take the N observed image positions $\boldsymbol{\theta}_i$ $(i = 1...N)$ and project them back to the source plane by applying the lens equation (equation 2.18), giving new positions $\boldsymbol{\beta}_i$. Since the lens model will not be perfectly correct, and given that the errors in the observed image positions σ_i will be non-zero, the back-projected source positions will not all fall at the same point on the source plane. Some form of fit statistic is computed from the degree of mismatch between model and data, and the lens parameters are then altered in a way that will bring the lens model into closer agreement with the data. The mismatch might be estimated using source plane variables, lens plane variables or a mixture of them. This is the basic algorithm behind parametric fitting.

A simple statistic that quantifies the goodness of fit for the image positions is to compute

$$\chi^2 = \sum_{i \neq j} (\beta_i - \beta_j)^2, \tag{2.47}$$

and minimise it (Kayser et al., 1990). This is known as source-plane minimisation (SPM) because the computation depends entirely on quantities that are defined on the source plane. SPM is extremely fast to compute. However, SPM does not minimise the errors between the model and the data in the image plane, which is where all the observables are located. It can also permit models that do not reproduce the image properties (such as parity) correctly. Kochanek (1991) identified this problem and proposed a new fit statistic, known as pseudo-image plane minimisation (PPM),

$$\chi^2 = \sum_{i=1}^{N} |\mu_i|^2 (\beta_i - \beta_m)^2, \tag{2.48}$$

where μ_i is the magnification at the ith image position and β_m is a modelled source position, which is given optimally by a magnification-weighted average of the back-projected image positions,

$$\beta_m = \frac{\sum_{i=1}^{N} |\mu_i|^2 \beta_i}{\sum_{i=1}^{N} |\mu_i|^2}. \tag{2.49}$$

PPM can be seen as mapping error vectors in the source plane to errors in the lens plane and computing the degree of mismatch there. It is therefore more accurate than SPM, but can produce a misleading statistic if the model is not almost correct, especially around bright images where μ is large. Both SPM and PPM were created as ways to avoid image plane minimisation (IPM) (Schneider et al., 1988). IPM computes the mismatch between data and model on the image plane,

$$\chi^2 = \sum_{i=1}^{N} \frac{(\theta_i - \theta_i')^2}{\sigma_i^2}, \tag{2.50}$$

where σ_i is the position error in the ith image position, and $\boldsymbol{\theta}_i'$ is the corresponding predicted position, which can only be found by inverting the lens equation. Inversion is very time-consuming, so IPM is much slower than the other methods, but it does compute the mismatch between data and model in terms of observable quantities. Thus, it can be more accurate than the other methods when faced with very bad models, and it has a simple meaning. Finding the images also allows us to immediately reject models that do not reproduce the caustic structures correctly, since such models will usually predict incorrect parities and image multiplicities.

The flux densities of the images can be used as constraints by defining an appropriate statistic. A sensible method is to take the modelled source flux density and compare it to the flux densities of the lensed images by making use of the predicted magnifications μ_i at each image location,

$$\chi^2 = \sum_{i=1}^{N} \frac{(f_i - \mu_i f_{src})^2}{\rho_i^2}, \tag{2.51}$$

where the ρ_i are the errors in the flux densities of the images f_i, and f_{src} is the flux density of the modelled source. f_{src} can be calculated in several ways, but the code follows the advice of Keeton (2001) and uses a least-squares calculation to determine the optimum model source flux density from the flux densities of the back-projected images,

$$f_{src} = \frac{\sum_{i=1}^{N} \frac{f_i \mu_i}{\rho_i^2}}{\sum_{i=1}^{N} \frac{\mu_i^2}{\rho_i^2}}. \tag{2.52}$$

It is possible to use a variety of other observables to constrain lens models, such as the orientations of resolved images such as radio jets, especially when looking for signs of mass substructure in lens galaxies Metcalf (2002); Dalal & Kochanek (2002); Mao et al. (2004); Kochanek & Dalal (2004). Often such additional constraints are weak compared to the image positions and fluxes (Bradač

et al., 2002).

2.8.4 Model fitting

Finding the optimum lens model given observational constraints requires us to locate the global minimum of a non-linear multidimensional function. It is usually necessary to employ non-linear minimum finding techniques when general lens models are involved (as discussed before, sometimes linear inversion is possible for specific types of model; Evans & Witt 2003). The code described here uses the downhill simplex method (Press et al., 1992). This method constructs a simplex (a type of polygonal shape that must have $N + 1$ vertices in a space of N free parameters), each vertex of which represents a slightly different set of values for the free parameters of the lens model. The fit statistic is evaluated at all vertices and the simplex may then flip or contract vertices around the vertex having the lowest value of the statistic. As a result, the simplex tends to flip its way down χ^2 gradients until it reaches a minimum in the χ^2 surface, about which it will shrink until a predefined convergence limit is reached. The simplex method is very prone to getting stuck in local minima, and hence it is usual to repeat the fitting process several times, starting from different initial vertices each time. The simplex method can be made more likely to find the global minimum by combining it with simulated annealing (Press et al., 1992), but convergence can become slower.

2.8.5 Application to CLASS lenses

The code has been applied to the lens system CLASS B0128+437 mentioned in section 2.7. In this system, 5 GHz VLBI observations (Biggs et al., 2004) have resolved the four images visible to MERLIN (Phillips et al., 2000), known as A, B, C and D, into three VLBI sub-components each (A1-A3, B1-B3, etc).

The surface brightnesses of components B1-B3 differ from those in the other three images, and Biggs et al. (2004) suggest that scattering in the lens galaxy is responsible. The effect on B1-B3 is frequency dependent so a scattering effect is more likely than microlensing by mass substructure in the lens. The image separation is only 0.54 arcsec, so at present there is no published lens galaxy position for B0128+437. However, there remain more constraints available from the VLBI data than is typical for CLASS radio lenses.

Figure 2.13 shows the optimum model predictions for B0128+437 when using an SIE plus external shear to model the MERLIN image positions and flux densities. The model has little difficulty in reproducing the observables, having a χ^2 of 1.7 for 2 degrees of freedom. However the external shear strength γ and the ellipticity of the lens ϵ are both unusually large, being 0.22 and 0.6 respectively. Phillips et al. (2000) found that a single SIE did not fit the data well, and Norbury (2002) first noted the significant external shear strength required to fit the MERLIN data successfully. For the MERLIN observations, then, the modelling is merely in agreement with what is already known and serves only to establish that the results of the new code are consistent with previous work[3].

Figure 2.14 shows the optimum SIE + shear model for the VLBI data of Biggs et al. (2004). The position errors on the compact sub-components are formally about ± 0.1 mas for images A1-3, C1-3 and D1-3. Position errors for images B1-3 are hard to determine because of the very smooth appearance of the relevant VLBI map, so they have been estimated at ± 5 mas. The flux densities were not used in fitting models to the VLBI data. The predicted positions of images B1-3 are more than 30 mas from the observed positions. There is also a clear discrepancy between the model and the VLBI data at images C1-3, in that the model predicts a much more distorted arrangement of images C1-3 than is

[3]The code has also been checked against the *lensmodel* (Keeton, 2001) package for various test cases.

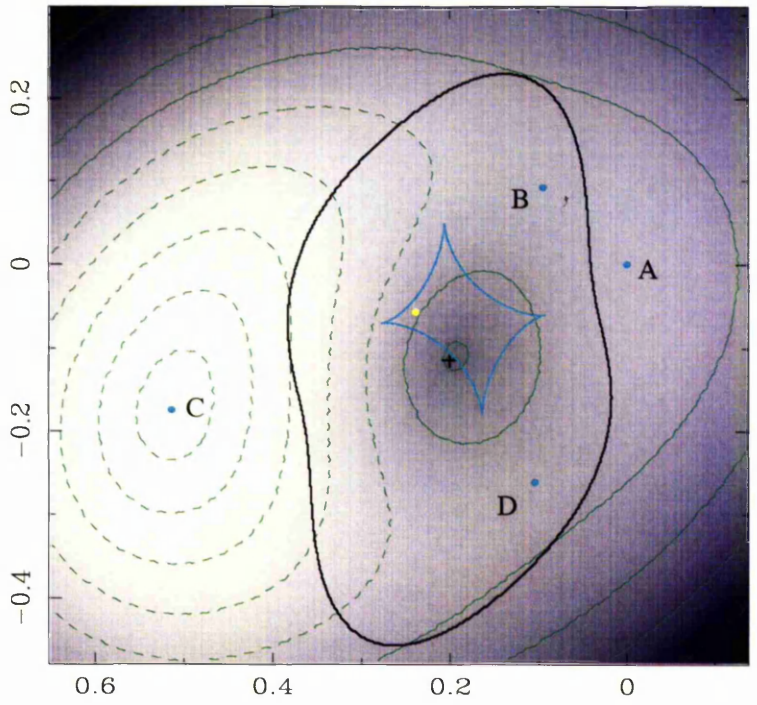


Figure 2.13: Lens modelling of CLASS B0128+437 MERLIN data using the code described in this work. An SIE with external shear was used for the lens model. Four images are observed, indicated by the filled cyan circles. The image positions predicted by the lens model are represented by filled red circles, but they lie under the cyan circles in this plot. The cross near the centre of the image marks the centre of mass of the SIE. The yellow dot is the modelled source position. The cyan curve marks the tangential caustic, and the black curve is the corresponding critical curve. The greyscale represents the time delay surface, which is contoured in green for clarity. The model fits the data well, with a χ^2 of 1.9 for 2 degrees of freedom. The χ^2 figure is dominated by the contribution from the flux density ratios.

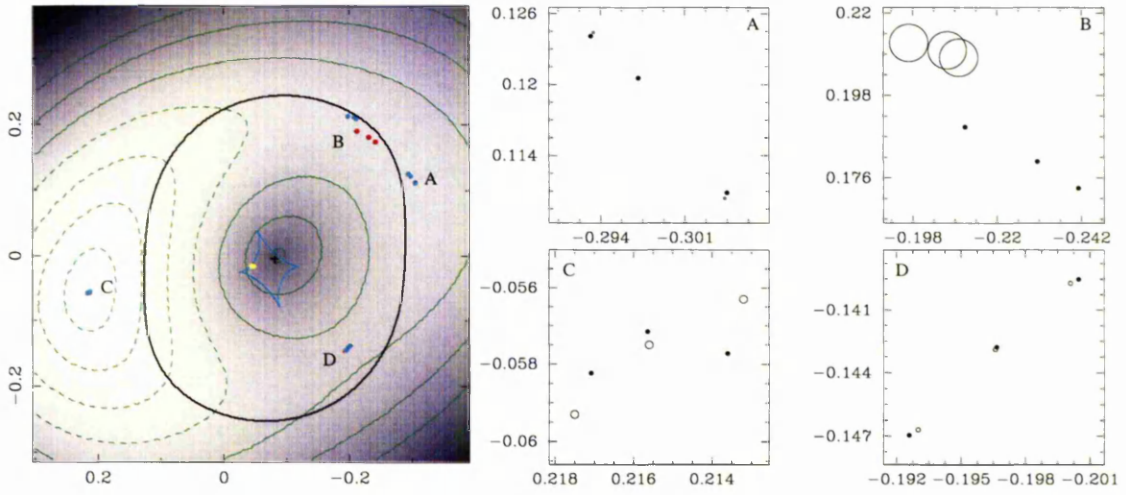


Figure 2.14: This figure shows the SIE+shear model applied to the VLBI data presented in Biggs et al. (2004). The large panel shows the entire lens model, while the smaller panels are expanded views showing the VLBI substructure and the model predictions. The observed images are marked by open circles with radii showing the 1σ position error estimates provided to the code, roughly ± 0.1 mas for images A1-3, C1-3 and D1-3, and ± 5 mas for images B1-3. Although the model reproduces the gross configuration of the images, the detail views show that the model cannot match the sub-component positions in image C.

observed. The overall fit to the VLBI position data is bad; the reduced χ^2 for this model is 40.

For the model in figure 2.15, the estimated position errors were increased to ± 1 mas for images A1-3, C1-3 and D1-3, while the position errors for images B1-3 remained at ± 5 mas. The predicted positions for B1-3 in this model are much closer to the observed positions, but the discrepancy associated with the images C1-3 appears to have got worse. As would be expected, inflating the positional errors improved the reduced χ^2 value substantially, decreasing it to 1.9. The majority of this figure is contributed by the mismatched distortion in images C1-3.

The constraints provided by images B1-3 can be suppressed entirely in the code, so that the data on B1-3 are used to check predicted image multiplicities and parities, but do not contribute to the fit statistic. As a result the penalty on

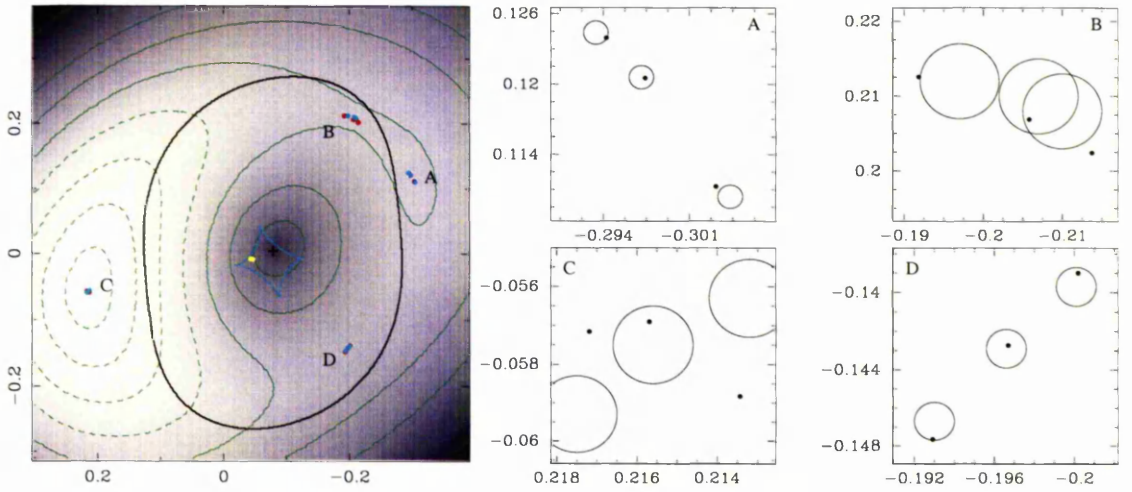


Figure 2.15: The SIE+shear model applied to the VLBI data again. This time, the estimated position errors for images A1-3, C1-3 and D1-3 were increased to ±1 mas. Although the resulting optimum model can fit the positions of images B1-3 better than the previous model, the disagreement with images C1-3 has increased.

very discrepant predictions for images B1-3 is lightened. Such a model, as shown in figure 2.16 reproduces the structure of images C1-3 very well indeed, but the predicted positions of B1-3 are over 100 mas from the observed locations. The reduced χ^2 (without the contribution from B1-3) is about 10. It appears that the SIE+shear model cannot fit the positions of images B1-3 and C1-3 simultaneously.

The above conclusions were published as part of Biggs et al. (2004), and are in agreement with conclusions derived from two other lens modelling packages, *lensmodel* (Keeton, 2001) and IGLOO. Biggs et al. also applied the method of Evans & Witt (2003) to the VLBI data as a way around the possible lack of freedom permitted by the simplistic SIE+shear model used above, but the results did not differ in substance from those found using the parametric non-linear approach. Trying to fit all four images simultaneously with a smooth model leads to a highly distorted model galaxy. One possible explanation for these difficulties could be a compact mass close to images B or C that is responsible for introducing an additional deflection, altering the positions of either image. This

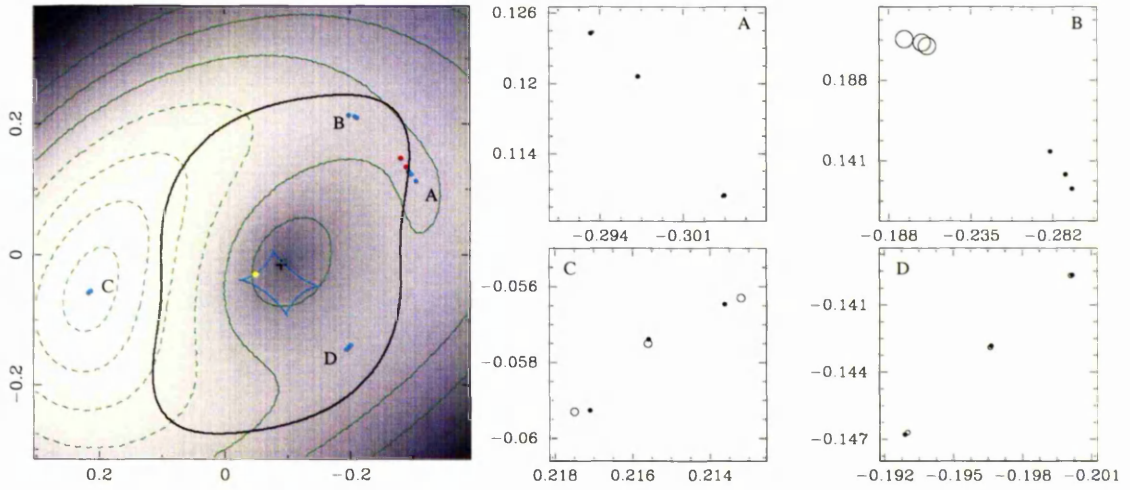


Figure 2.16: In this fit the estimated position errors for images A1-3, C1-3 and D1-3 were decreased to ±0.1 mas again, but image B was ignored by the code (see text). The optimum model produces a much better fit to images C1-3, but the fit to images B1-3 is even worse than in the previous models, with the predicted positions being over 100 mas from the observational constraints.

is not surprising if mass substructure close to images A or B is present, as the Evans & Witt method models an explicitly scale-free galaxy; that is, the lensing potential is separated into a product of angular and radial functions centred on the lens, and the lines of constant convergence all have the same shape independent of radius. A mass excess close to B or C would break this assumption, and as a result would need many higher-order Fourier components to model effectively. As was previously noted, similar behaviour to the B0128+437 case was demonstrated by Evans & Witt in the case of JVAS B1422+231, which many authors have argued is likely to contain mass substructure based on the flux ratio relation violations seen in the system (Mao & Schneider, 1998; Bradač et al., 2002).

The modelling shows that attempting to fit the position of images C1-3 leads to a very discrepant position for B, while attempting to fit image B does not alter the predicted positions of image C1-3 much but does distort them. The simplest approach would therefore be to add mass substructure of the order of 10^8 $M_\odot$ close to image C in model 1. Attempting to move image B instead would require

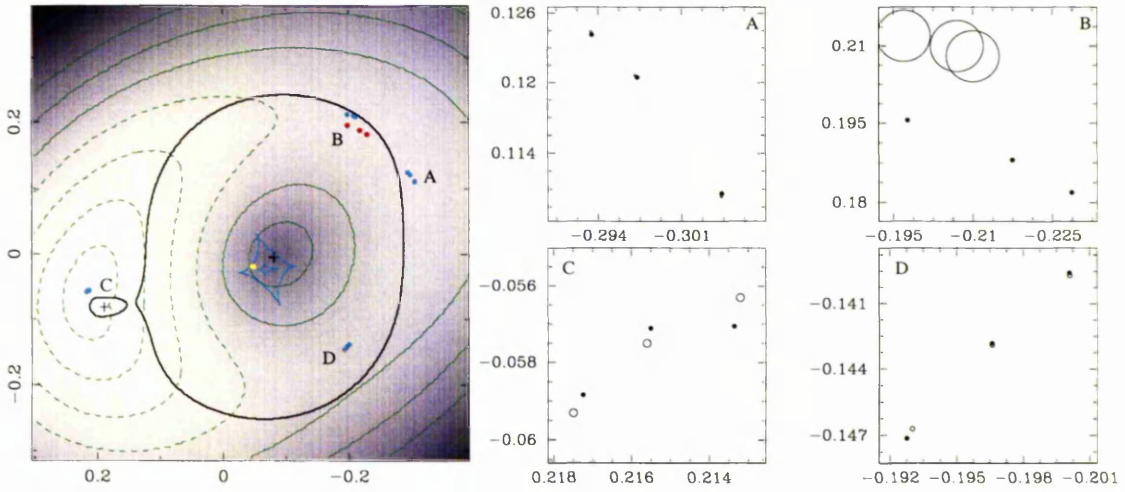


Figure 2.17: SIE+shear model as constrained in figure 2.14, with the addition of a point mass. The centre of the point mass is marked by the cross close to images C1-3. An additional critical curve formed by the point mass is visible, although since the source lies outside the corresponding caustic, no additional images are formed.

a much larger mass ($\sim 10^9$ $M_\odot$) because the deflection required is greater.

Figure 2.17 shows an attempt to improve the SIE+shear fit to images C1-3 by placing a point mass model nearby. The position errors used were ± 0.1 mas for images A1-3, C1-3 and D1-3, and ± 5 mas for images B1-3 (which were allowed to contribute to the fit statistic). The optimum fit had a reduced χ^2 of 28, which is a slight improvement over the equivalent model with no point mass (the number of degrees of freedom is 11 for the base SIE+shear model; adding a point mass reduces this to 8). The fit to B1-3 and C1-3 has improved, but neither the distortion of C1-3 nor the anomalous deflection for B1-3 have been eliminated. The point mass has a critical radius of about 14 mas, which corresponds to a mass at the lens on the order of 10^8 $M_\odot$.

Figure 2.18 shows the SIE+shear model with an NIS placed north of image B. The core radius of the NIS was set larger than its critical radius to suppress the formation of additional images. The optimum model has a reduced χ^2 of 43, comparable to or worse than the same statistic for the SIE+shear model without

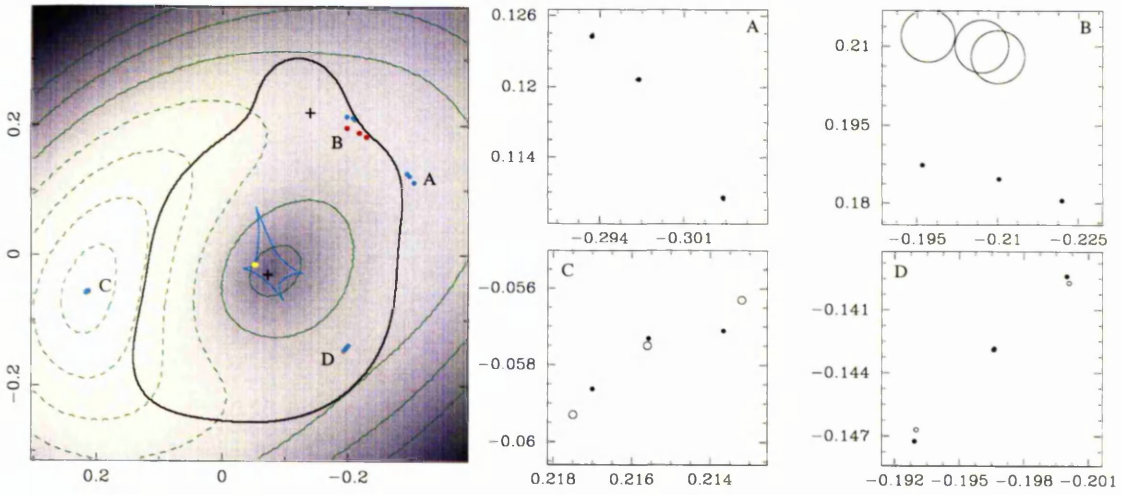


Figure 2.18: SIE+shear model as constrained in figure 2.14, with the addition of a non-singular isothermal sphere (NIS). The centre of the NIS is marked by the cross north of images B1-3. No additional critical curve have been formed because the NIS core radius is larger than its critical radius. The fit to images B1-3 is actually worse than in the previous model, because B cannot be significantly displaced by a nearby mass without affecting A as well.

substructure. The critical radius of the NIS is about 47 mas, which corresponds to a mass enclosed within the Einstein radius of over 10^9 $M_\odot$.

The addition of compact substructure on a mass scale of approximately 10^8 $M_\odot$ near image C can improve the fit of an SIE+shear model to the 5 GHz VLBI data for B0128+437. Placing similar substructure near image B does not produce any improvement in the model. It is likely, therefore, that the poor fits of smooth models to B0128+437 are caused by a mass perturbation in the vicinity of image C. To further explore the mass substructure in B0128+437 it will probably be necessary to give the models more freedom, since a single point mass cannot completely correct the model predictions. This can potentially be done by the gravitational imaging method of Koopmans (2005), but an ability to adapt the resolution of the pixellated potential around regions of interest will probably be essential, because the VLBI maps of each image are separated from one another by hundreds of empty beams. It is also not clear what effect the scattering of

image B will have on a method that regards the (deconvolved) radio maps as true representations of the surface brightness distribution on the sky.

Bibliography

Barkana R., 1998, ApJ, 502, 531

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

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

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

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

Blandford R.D., Narayan R., 1992, ARA&A, 30, 311

Bradač M., Schneider P., Steinmetz M., Lombardi M., King L.J., Porcas R., 2002, A&A, 388, 373

Bradač M., Schneider P., Lombardi M., Steinmetz M., Koopmans L.V.E., Navarro J.F., 2004, A&A, 423, 797

Benson J.R., Cooke J.H., 1979, ApJ, 227, 360

Bryan R.K., Skilling J., 1980, MNRAS, 191, 69

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

Chae K.-H., 1999, ApJ, 524, 582

Chae K.-H., Mao S., Augusto P., 2001, MNRAS, 326, 1015

Chae K.-H., 2002, ApJ, 568, 500

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

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

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

Courbin F., Saha P., Schechter P.L., 2002, Courbin F., Minniti D., eds, Gravitational Lensing: An Astrophysical Tool, Lecture Notes in Physics, vol. 608, p.1

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

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

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

Faber S.M., et al., 1997, AJ, 114, 1771

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

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

Fassnacht C.D., Xanthopoulos E., Koopmans L.V.E., Rusin D., 2002, ApJ, 581, 823

Freedman W.L., et al., 2001, ApJ, 553, 47

Gorenstein M.V., Shapiro I.I., Falco E.E., 1988, ApJ, 327, 693

Hernquist L., 1990, ApJ, 356, 359

Högbom J.A., 1974, A&AS, 15, 417

Kayser R., Surdej J., Condon J.J., Kellermann K.I., Magain P., Remy M., Smette A., 1990, ApJ, 364, 15

Kayser R., Helbig P., Schramm T., 1997, A&A, 318, 680

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

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

Keeton C.R., 2001, astro, arXiv:astro-ph/0102341

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

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

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

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

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

Kochanek C.S., 1994, ApJ, 436, 56

Kochanek C.S., 2002, ApJ, 578, 25

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

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

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

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

Koopmans L.V.E., 2005, astro, arXiv:astro-ph/0501324

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

Lauer T.R., 1985, ApJ, 292, 104

Mao S., 1992, ApJ, 389, 63

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

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

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

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

Metcalf R.B., 2005, ApJ, 622, 72

Muñoz J.A., Kochanek C.S., Keeton C.R., 2001, ApJ, 558, 657

Myers S.T., et al., 1995, ApJ, 447, L5

Navarro J.F., Frenk C.S., White S.D.M., 1996, ApJ, 462, 563

Norbury, M.A., 2002, Ph.D. thesis (University of Manchester)

O'Dea C.P., Baum S.A., Stanghellini C., Dey A., van Breugel W., Deustua S., Smith E.P., 1992, AJ, 104, 1320

Patnaik A.R., Browne I.W.A., Wilkinson P.N., Wrobel J.M., 1992, MNRAS, 254, 655

Patnaik A.R., Browne I.W.A., King L.J., Muxlow T.W.B., Walsh D., Wilkinson P.N., 1993, MNRAS, 261, 435

Pen U.-L., Mao S., 2005, 2005, astro, arXiv:astro-ph/0506053

Phillips P.M., et al., 2000, MNRAS, 319, L7

Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P., 1992, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, Cambridge (UK) and New York (2nd ed)

Refsdal S., 1964, MNRAS, 128, 307

Rusin D., Ma C.-P., 2001, ApJ, 549, L33

Rusin D., et al., 2001, ApJ, 557, 594

Rusin D., Kochanek C.S., Keeton C.R., 2003, ApJ, 595, 29

Saha P., Williams L.L.R., 1997, MNRAS, 292, 148

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

Schneider D.P., Turner E.L., Gunn J.E., Hewitt J.N., Schmidt M., Lawrence C.R., 1988, AJ, 95, 1619

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

Skilling J., Bryan R.K., 1984, MNRAS, 211, 111

Sykes C.M., et al., 1998, MNRAS, 301, 310

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

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

Turner E.L., Ostriker J.P., Gott J.R., 1984, ApJ, 284, 1

Wallington S., Kochanek C.S., Narayan R., 1996, ApJ, 465, 64

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

Warren S.J., Dye S., 2003, ApJ, 590, 673

Weymann R.J., et al., 1980, Nature, 285, 641

Williams L.L.R., Saha P., 2000, AJ, 119, 439

Winn J.N., Rusin D., Kochanek C.S., 2003, ApJ, 587, 80

Winn J.N., Rusin D., Kochanek C.S., 2004, Nature, 427, 613

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

Wucknitz O., 2004, MNRAS, 349, 1

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

Wucknitz O., Sperhake U., 2004, PhRvD, 69, ID: 063001

Young P., Gunn J.E., Oke J.B., Westphal J.A., Kristian J., 1980, ApJ, 241, 507

Chapter 3

CLASS B0631+519

3.1 Introduction

This chapter[1] concerns the last CLASS lens to be reported in the literature, B0631+519. I will summarise the discovery of the lens, recap previously published observations (Phillips, 2001; Norbury, 2002), and then present and discuss new radio and optical/near-infrared observations and show how the new data has improved our understanding of this lens system. The new radio observations presented in this chapter consist of a 1.7 GHz MERLIN observation and two epochs of VLBA observations at 1.7 GHz made on 1999 December 10 and 2002 December 01. The new optical and near-infrared observations were made with the ACS and NICMOS on board the HST, roughly covering the Johnson V, I and H filters, and with the ground-based WHT, covering Johnson R.

[1]The work presented in this chapter has been published as *CLASS B0631+519: Last of the Cosmic Lens All-Sky Survey lenses*, in MNRAS volume 361, page 259 (2005)

Date observed	Instrument	ν [GHz]	τ	σ [μJy beam^{-1}]	θ_B [mas]
1994 March 05	VLA	8.4	16 sec	410	234
1999 July 02	VLA	15	354 sec	360	142
1999 August 03	VLA	8.4	510 sec	86	236
1999 December 10	VLBA	1.7	1.4 hr	130	8
2001 March 21	VLBA	5	1.7 hr	82	3.6
2001 July 07	MERLIN	5	19 hr	120	52
2002 December 01	VLBA	1.7	7.8 hr	64	9
2003 March 01	MERLIN	1.7	20 hr	62	173

Table 3.1: Radio observations made of B0631+519 to date. ν is the frequency of the observation, τ is the integration time, σ is the noise in the residuals after subtraction of a CLEAN model or Gaussian model components, and θ_B is the (geometric) mean beam size.

3.2 Discovery of B0631+519 in CLASS

All radio observations made of B0631+519 to date are summarised in table 3.1. The first CLASS observations of B0631+519 were made in 1994, and it was identified as a strong lens candidate following the re-calibration and re-analysis of CLASS data (Phillips, 2001) during CLASS phase IV. Following its discovery, VLA observations of B0631+519 were made at 8.4 GHz and 15 GHz, as described in Phillips (2001). MERLIN and VLBA observations at 5 GHz are described in Norbury (2002).

The discovery map produced from the 1994 March 05 CLASS snapshot observation is shown in figure 3.1. Two radio components were detected with a separation of 1.15″. The brighter, north-western component (labelled A1) has a flux density of 38±2 mJy, while the fainter, south-eastern) component (labelled B) has a flux density of 5.4±0.4 mJy. There is a south-western extension to the brighter component (A2).

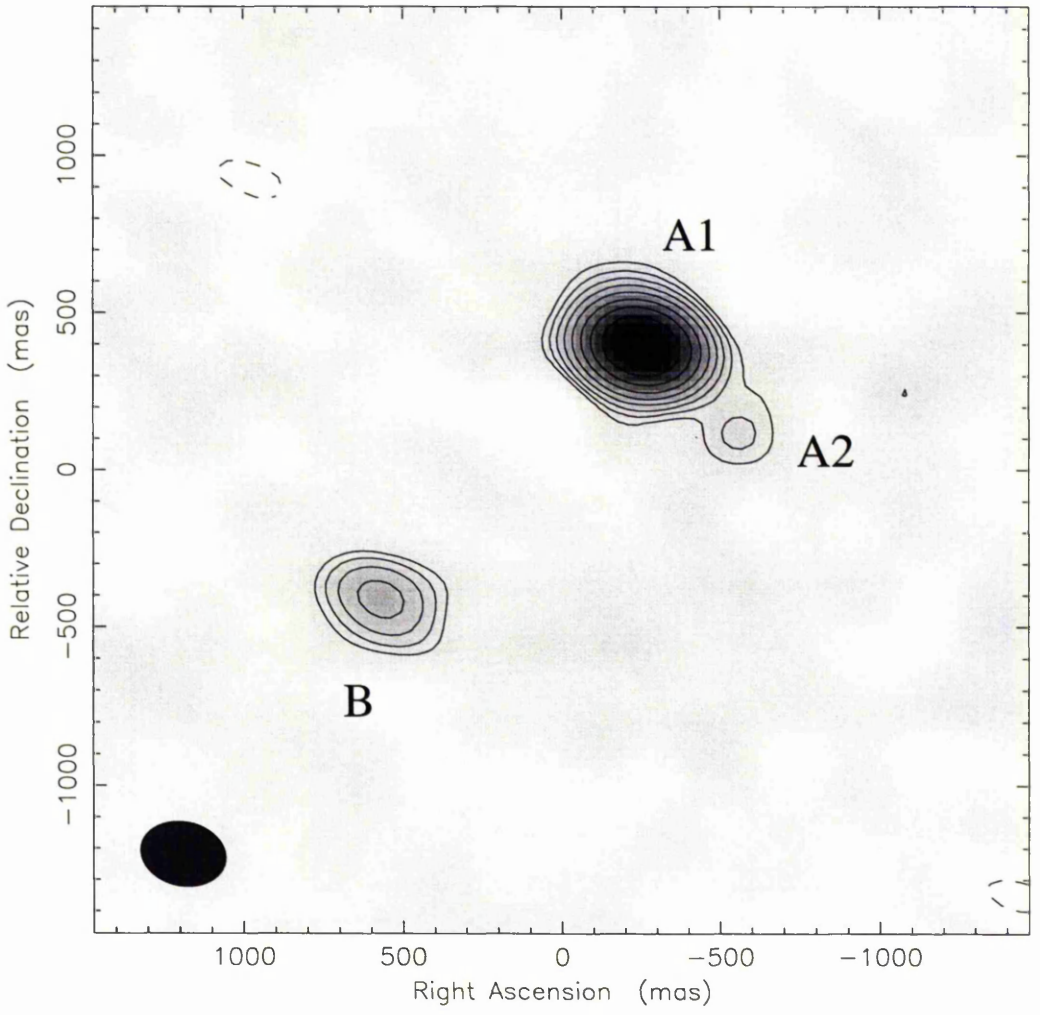


Figure 3.1: Discovery map for CLASS B0631+519, made from the 8.4 GHz VLA snapshot data taken on 1994 March 05. The data are uniformly weighted. The beam is sized 268×204 mas with the major axis having a position angle of +77.4°. The RMS noise level is 410 μJy beam^{-1}. The lowest contours are plotted at -3 and 3 times the RMS noise level; successive contours represent a multiplicative increase in surface brightness by a factor of 1.5, up to the highest contour level of 31.7 mJy beam^{-1}. Component A1 is located at RA: 06^h 35^m 12.31402^s, Dec: +51° 57' 01.8026″(J2000.0) as measured by the VLBA.

3.3 Previous radio observations of B0631+519

3.3.1 VLA

Phillips (2001) described VLA observations of B0631+519 made at 8.4 GHz on 1999 August 03, and at 15 GHz on 1999 July 02. The deconvolved 8.4 GHz map is shown in figure 3.2. These 8.4 GHz observations had a surface brightness sensitivity roughly 5 times greater than that of the discovery snapshot, due to the longer exposure time of 510 seconds. A region of extended emission stretching down to the south-west of A2 is clearly visible in the deconvolved map. It terminates in the region of elevated surface brightness labelled X. Some extended emission is visible on the opposite side of A1 from X, (labelled Y in figure 3.2), although the detection is marginal (about 2.3 σ). Component B remained unresolved in all VLA maps.

I have compared the total flux density detected in the 1999 observations (42±2 mJy) with that detected in the discovery map, and found that the discovery map was 10% brighter, despite the additional emission from the extended structures around A2 and X. The total flux density due to the compact components A1 and B decreases by 13% from the discovery map to the later observations. This discrepancy might be accounted for by uncertainty in the absolute amplitude calibration between the 1994 and the 1999 observations, but could also indicate that the source is variable. The mean size of the synthesized beam varies by less than 3 mas between the two (uniformly weighted) deconvolved maps, or by 2% in area, so it is unlikely that 10% of the emission has been resolved out in one map with respect to the other. Most importantly, the A1:B flux density ratio also changes by 16% between the two observations, which is hard to reconcile with a global mis-calibration of the amplitude scale. The signal-to-noise ratios of both A1 and B were at least 10 in all VLA observations, and the array was in

'A' configuration on both dates.

Phillips (2001) describes a further observation of B0631+519 made with the VLA at 15 GHz on 1999 July 02. A deconvolved, naturally weighted map is shown in figure 3.3. The total flux density detected at 15 GHz is 34±3 mJy, and the A1:B flux density ratio is 7.5:1. Both A1 and B are unresolved by the VLA at this frequency. A2 is not detected, suggesting that it has a steeper spectral index than A1 and B.

The VLA data reported by Phillips (2001) were sufficient to establish B0631+519 as a possible gravitational lens candidate. The candidate met all criteria for inclusion in the statistically well-defined CLASS sample. However, the nature of component A2 and the extended emission seen at 8.4 GHz was unclear. The CLASS methodology (Myers et al., 2003; Browne et al., 2003) is to check that a pair of possibly lensed images have the same surface brightness by observing them with MERLIN. If one image is resolved out while the other remains compact, the candidate is rejected.

3.3.2 MERLIN

Norbury (2002) reports MERLIN observations of B0631+519 at 5 GHz. B0631+519 was observed using MERLIN on 2001 July 07, over a period of 19 hours. A deconvolved map is shown in figure 3.4. Component B has been resolved into two components (B1 and B2) and components A1 and A2 are again detected. A1 and B1 are compact at 50 mas resolution, while A2 and B2 appear slightly resolved. A hypothesis that requires all four components to be images of a single source is not supported by the evidence. Also, the arrangement of the images is not typical of a quad; if B1/B2 are taken to be a merging pair produced by a source close to a fold caustic marking the five-image region, then A1 and A2 should be very widely separated. A hypothesis in which two sources were each doubly imaged

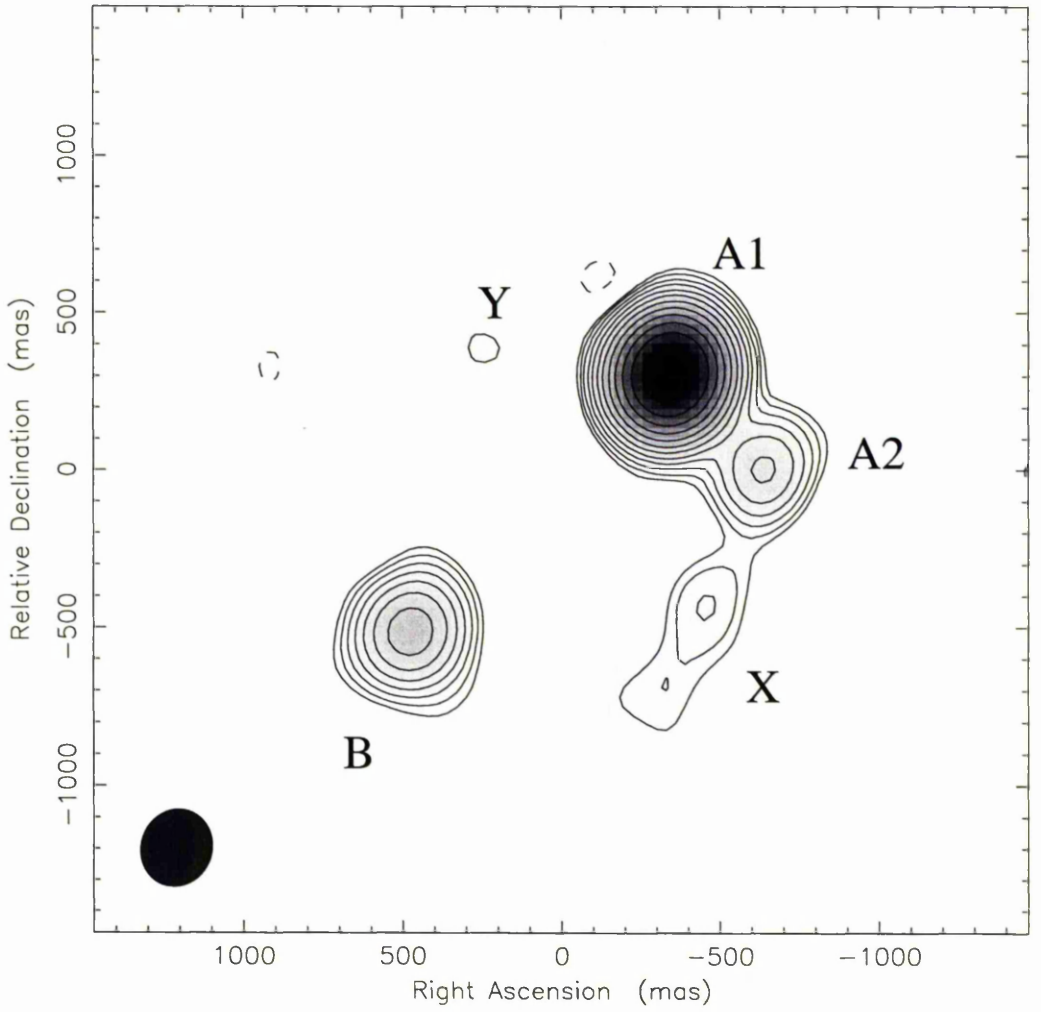


Figure 3.2: Radio map of the VLA 8.4 GHz data taken on 1999 August 03. These observations are roughly a factor of five times more sensitive than the discovery image, with an RMS noise in the residual map of 86 μJy beam^{-1}. The data are naturally weighted. The synthesized beam is sized 250×223 mas with the major axis oriented along a position angle (PA) of −17.5°. Contours are plotted as in figure 3.1; the highest contour level is 22.8 mJy beam^{-1}.

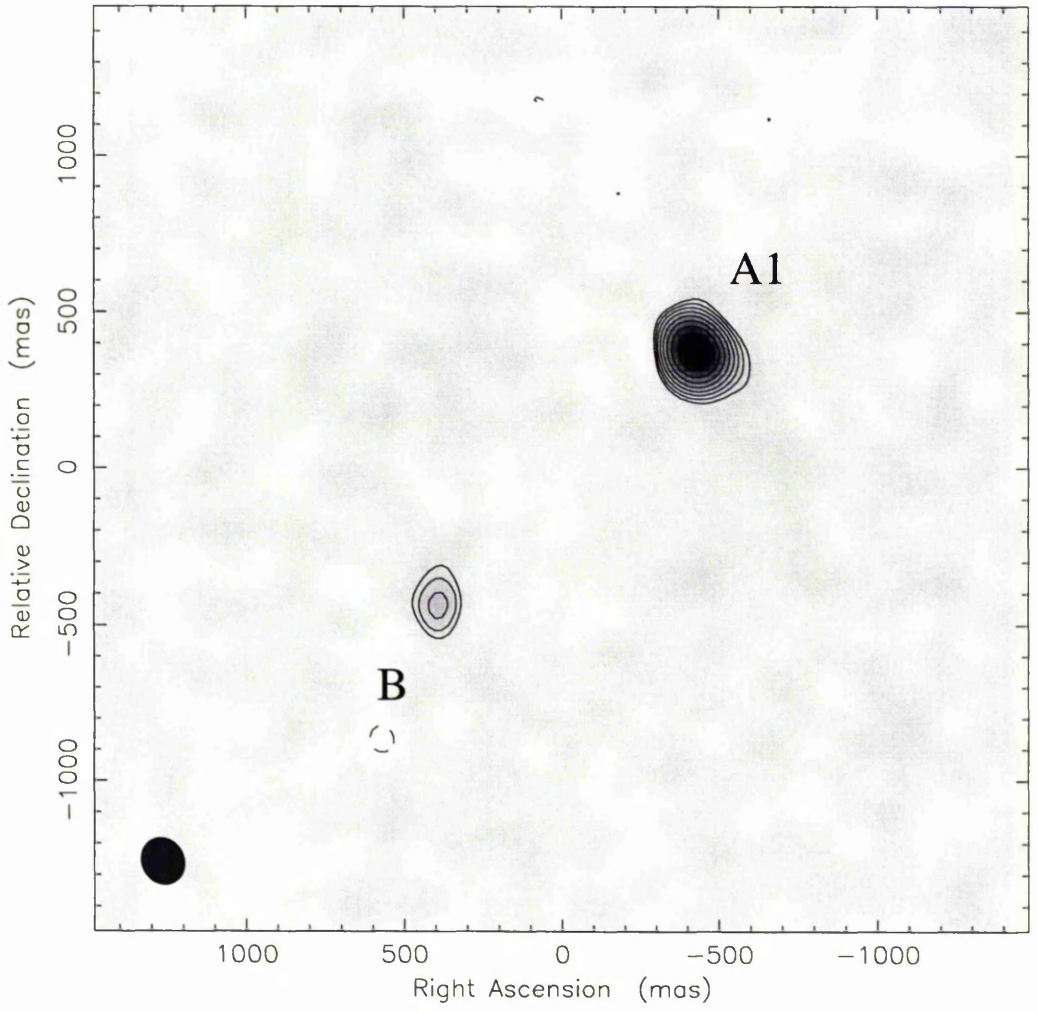


Figure 3.3: 15 GHz VLA data taken on 1999 July 02 (Phillips, 2001). The synthesized beam is 152×132 mas oriented with the major axis at a PA of +26°, and the RMS noise in the residuals is 360 μJy beam^{-1}. The data are naturally weighted. Contours are plotted as for figure 3.1; the highest contour value is 27.8 mJy beam^{-1}.

was therefore adopted by Norbury (2002).

3.3.3 VLBA

Norbury (2002) also reported observations made at 5 GHz with the VLBA, made on 2001 March 21.

Maps of images A1 and B1 are shown in figure 3.5. The resolution achieved was 3.6 mas, sufficient to split A1 into four sub-components spread over an angle of about 20 mas. These sub-components are labelled A1a to A1d. B1 remained unresolved. No emission from the A2, B2, X, Y or Z components was detected down to the 5σ limit of 390 μJy beam^{-1}, indicating that these components have probably been resolved out. Norbury (2002) concluded that the lens hypothesis had not been falsified, and that two possible models of the source structure could be tested by future observations. These models are discussed further in section 3.6.

The sum of the flux densities of components A1a to A1d (46$\pm$2 mJy) is consistent with the MERLIN 5 GHz flux of component A1 to within the measurement errors, which implies both that A1 did not vary by more than 5% over the 4 months separating the observations and also that no radio structure seen in A1 by MERLIN was resolved out by the VLBA, despite the factor of $\sim$14 increase in resolution. However, another possible explanation is that an increase in source flux density (due to intrinsic variability) cancelled out any loss of emission due to resolution effects. The same conclusions and caveats apply to the MERLIN observations of B1 at 5 GHz.

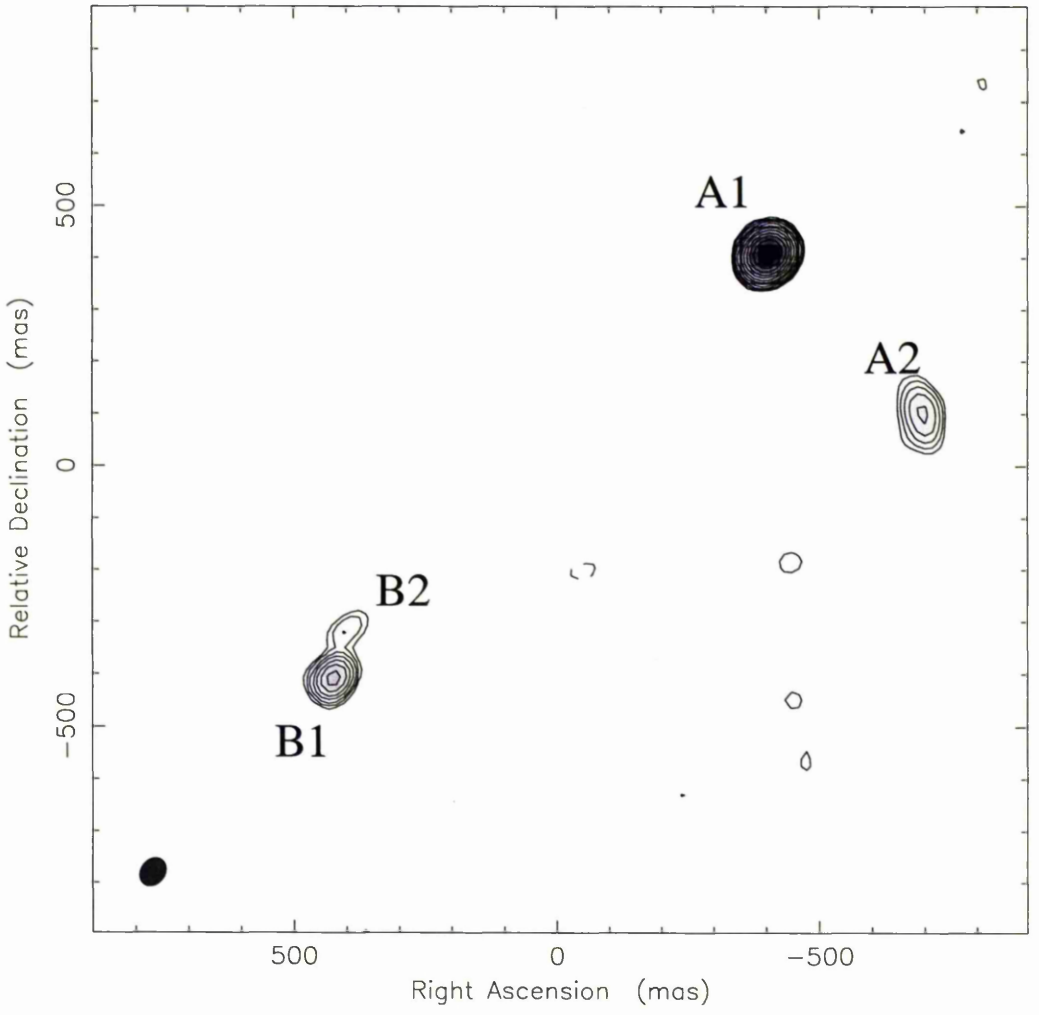


Figure 3.4: Radio map of the 5 GHz MERLIN observations taken on 2001 July 07 Norbury (2002). The data were naturally weighted. The beam size is 58×46 mas with the beam's major axis being oriented at a PA of −34°. The RMS noise in the residuals is 120 μJy beam^{-1}. Contours are plotted at levels of (-3, 3, 4.5, 6.75, 10.1, 15.2, 22.8, 34.2, 51.3, 76.9, 115, 173, 259) times the RMS noise.

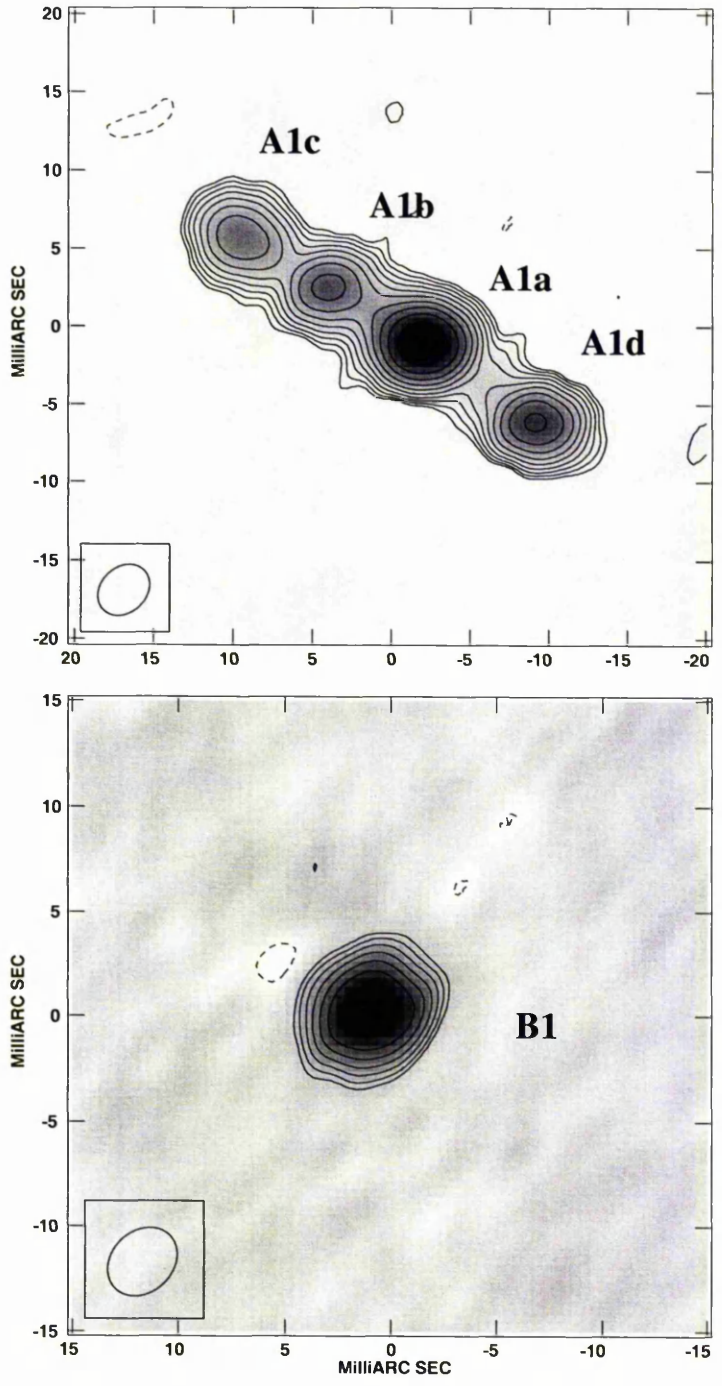


Figure 3.5: Radio map of the 5 GHz VLBA observations taken on 2001 March 21, showing image A (top) and image B (bottom). The data were uniformly weighted. The beam-size is 3.8×3.4 mas oriented at PA +144° and the RMS noise is 82 μJy beam^{-1}. Contours are plotted at levels of (-3, 3, 4.5, 6.75, 10.1, 15.2, 22.8, 34.2, 51.3) times the RMS noise.

3.4 New radio observations

In this section I present and discuss new observations of B0631+519 made with MERLIN at 1.7 GHz and with the VLBA at 1.7 GHz. The MERLIN data strongly constrain the lensed radio source morphology and the image parities in the system compared to previous observations. The MERLIN data, in combination with flux density measurements from other wavelengths, demonstrate that there are three distinct regions of emission in the source.

The 1.7 GHz VLBA data on image A, when compared to the corresponding 5 GHz VLBA data, identify the flat-spectrum radio core in the source. Since there are two epochs of 1.7 GHz VLBA data available, separated by 3 years, they also allow a crude test for variability in the lensed images.

3.4.1 1.7 GHz MERLIN

B0631+519 was observed with MERLIN on 2003 March 01 at a frequency of 1.7 GHz, and for an integration time of 20 hours. The resurfaced Lovell MkIa telescope was included in the network, for a total of 7 telescopes and 21 baselines. Phase referencing was carried out using the source J0636+5009, using a cycle time of 3 minutes on the source for every one minute spent on the phase calibrator. The radio source 3C286 was used to set the flux density scale, and B0552+398 was used calibrate the telescope gains. The observations were phase referenced to the source JVAS J0642+5247. The data were initially phase and amplitude calibrated using the MERLIN D-programs and AIPS[2]. Some flagging was necessary, with the Knockin telescope (and associated baselines) suffering occasional outages during the course of the observations. Mapping and self-calibration were performed using the Caltech difference mapping program DIFMAP (Shepherd, 1997). An RMS

[2]Astronomical Image Processing System

Component	Frequency [GHz]	Position offsets [mas]		Flux density [mJy]
		ΔRA	ΔDec.	
A1	1.7	+0±1	+0±1	66.4±6.6
	5	+0±1	+0±1	46.9±4.7
A2	1.7	−293±1	−313±1	19±1.9
	5	−293±3	−308±3	4.2±0.4
B	1.7	+809±1	−800±1	11.8±1.2
B1	5	+828±1	−816±1	5.4±0.5
B2	5	+801±8	−722±8	1.3±0.1
X	1.7	−40	−880	11±1
	5	-	-	<0.6
Y	1.7	+595±6	+67±6	2.7±0.3
	5	-	-	<0.6
Z	1.7	+691±5	−1016±5	5.0±0.5
	5	-	-	<0.6

Table 3.2: Radio components detected in B0631+519 with MERLIN. All position offsets are relative to A1. X is a large arc that is not well described by a single gaussian component; the position quoted is an estimate for the centre of the feature. The integrated flux density of X was derived by summing within the 3σ contour.

noise level in the deconvolved map of 62 μJy beam^{-1}was reached. The deconvolved radio map is shown in figure 3.6. The map shows emission to both west and east of image A1. These regions of emission are low-frequency counterparts to components X and Y seen in the VLA data. B1 also develops an extension to the west, which is labelled Z in the figure.

Component positions and flux densities from the MERLIN observations are shown in table 3.2. The component positions and flux densities for the 5 GHz data described in Norbury (2002) are also displayed, since they were used in modelling the lens (section 3.6). I have assumed a 10% absolute calibration error on flux densities from MERLIN. Position errors are estimated from the beam size divided by the peak SNR.

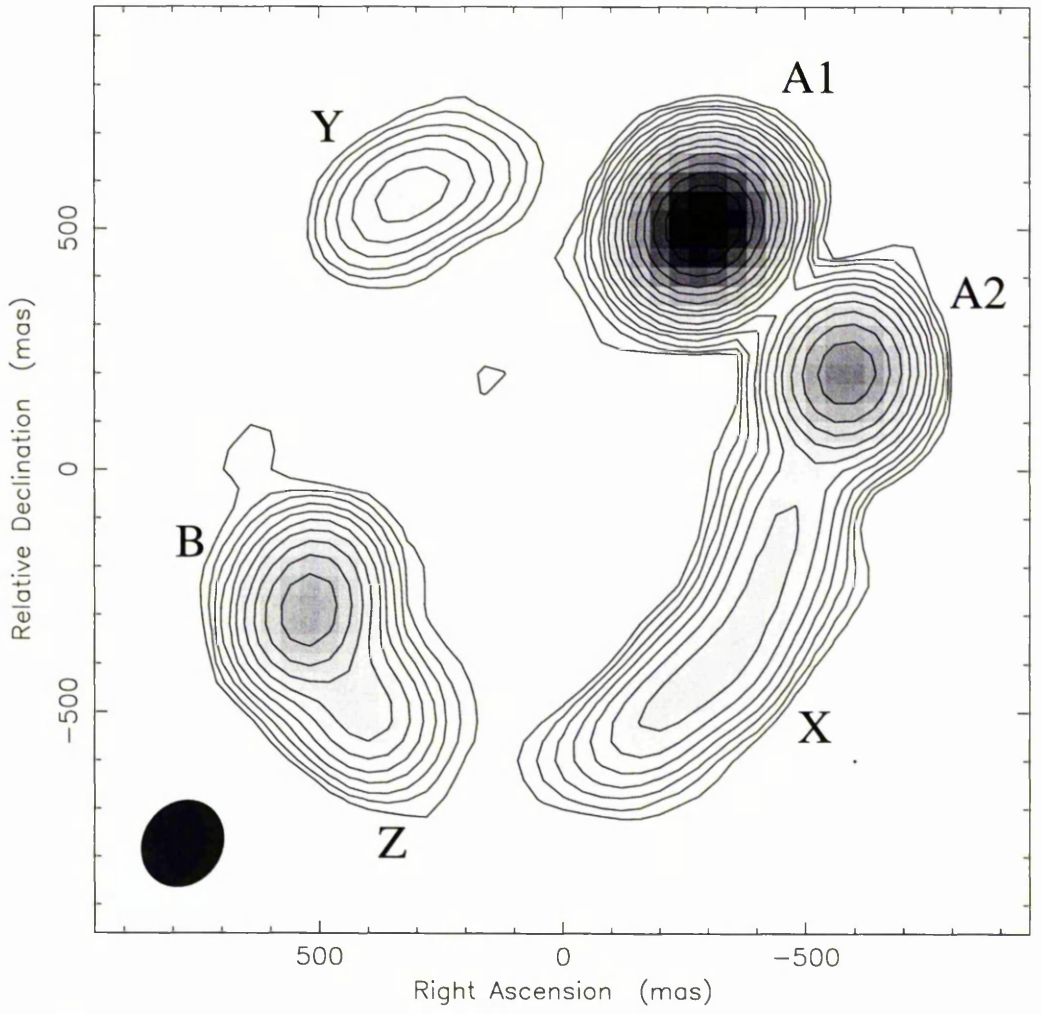


Figure 3.6: Radio map of the 1.7 GHz MERLIN observations taken on 2003 March 01. The data were naturally weighted. The synthesized beam is sized 186×161 mas at PA −31.7°. Contours are plotted at factors of (-3, 3, 4.5, 6.75, 10.1, 15.2, 22.8, 34.2, 51.3, 76.9, 115, 173, 259, 389, 584, 876) times the RMS noise.

3.4.2 1.7 GHz VLBA

Two epochs of phase-referenced VLBA data were obtained at 1.7 GHz, on 1999 December 10 and 2002 December 01. The phase reference source used was J0631+5311. The 1999 data was taken in two polarisations via 4 IFs, 2 IFs to each polarisation; the integration time was 1.4 hours. The 2002 data were taken over 7.8 hours and recording only the left-hand circular polarisation over 4 IFs. In both epochs, 2-bit sampling was used, and correlation at the VLBA correlator in Socorro subdivided each IF into 16 0.5 MHz wide channels and averaged the data into 2 sec integrations. The phase centre was taken halfway between images A1 and B1 to minimise bandwidth smearing.

The data from both epochs were reduced in AIPS following standard VLBA recipes. CLEAN imaging Högbom (1974) and self-calibration were carried out using the AIPS task SCIMG with natural weighting. The CLEANed maps of A1, A2 and B1/B2 from the two epochs are shown in figure 3.7. A1 appears resolved but (unsurprisingly) smoothed compared to the 5 GHz VLBA map, while B1 remains unresolved. A2 is resolved into an arc-like feature of low surface brightness that extends over 80 mas. B2 is partially resolved into an extended source. No emission was detected from components X, Y or Z. The flux densities of A1, A2 and B2 agree between the two epochs to within the measurement errors. The flux density of B1 decreases by about 9% between 1999 and 2002; this decrease is not quite significant at the 2σ level given the measurement errors of about 5% for B1's flux density. The measurements of the A1/B1 image separation from the two epochs are consistent to within 0.5 mas, while the A1-B2 separation measurements agree to within 2 mas.

Flux densities and positions of the VLBA components are shown in table 3.3. The 5 GHz components are also shown to permit comparison with the 1.7 GHz data. The stated errors on flux densities assume a 5% random error in the

absolute flux density scale calibration at the VLBA. For A2 and B2 at 1.7 GHz, I have doubled the formal error to reflect the uncertainty in the flux densities of these resolved components; their flux densities were obtained by summing within the 3σ contour. Positions for A1, B1 and B2 were determined from the best-fit gaussians, restricted to the compact north-eastern feature in the case of A1 at 1.7 GHz. The position given for A2 is that of the brightest pixel.

Comparison of the VLBA maps at 1.7 GHz and 5 GHz suggests that component A1c is too bright at 1.7 GHz to be the flat-spectrum core, and that A1b and A1d are too faint. To check which of the four VLBA sub-components of A1 correspond to the flat-spectrum radio core in the lensed source, I reduced the 1.7 GHz VLBA data of 2002 a second time, this time employing uniform weighting of the visibilities to improve resolution at the expense of surface brightness sensitivity. I registered the resulting 1.7 GHz map with the 5 GHz map by using the position of component B1 as a reference (since B1 is largely unresolved in both maps). To overlay B1 in this way required the 1.7 GHz image to be translated by 1.7 mas south and 8.9 mas east, because slightly different positions were used for the phase reference source J0631+5311 between the 5 GHz and 1.7 GHz observations. The translated, uniformly weighted 1.7 GHz map is shown overlaid on the 5 GHz map in figure 3.8. Sub-component A1a appears to be the flat-spectrum core, while the peak in the 1.7 GHz map corresponds to component A1c. Therefore we apparently see both approaching and receding jets in the source AGN.

3.4.3 Radio spectra from continuum data

The radio observations provide enough information to plot two- or three-point radio spectra of the various components and thereby examine the lens hypothesis for B0631+519. All the lensed images of a source should have similar radio

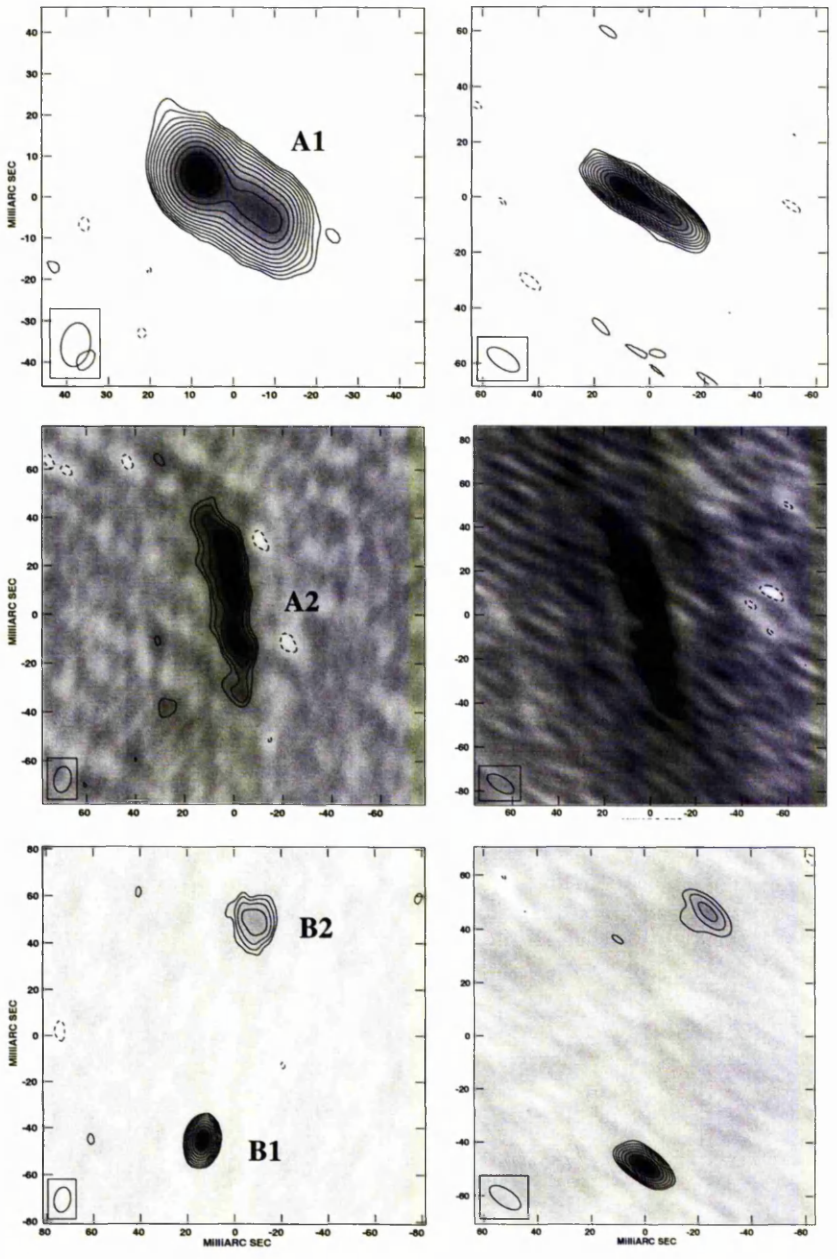


Figure 3.7: Radio maps of the 1.7 GHz VLBA observations. The data were naturally weighted. The column on the left shows the maps from 1999 December 10, while the column on the right shows the maps from 2002 December 01. The synthesized beam in the 1999 observations is sized 11×6 mas, oriented with the major axis at a PA of +53°. The RMS noise for the 1999 observations is 130 μJy beam^{-1}. The beam for the 2002 observations is an elliptical gaussian sized 11.2×7.5 mas oriented with its major axis at a PA of +177°. The RMS noise is 64 μJy beam^{-1}. In both images, contours are plotted at levels of (-3, 3, 4.5, 6.75, 10.1, 15.2, 22.8, 34.2, 51.3, 76.9, 115, 173) times the RMS noise.

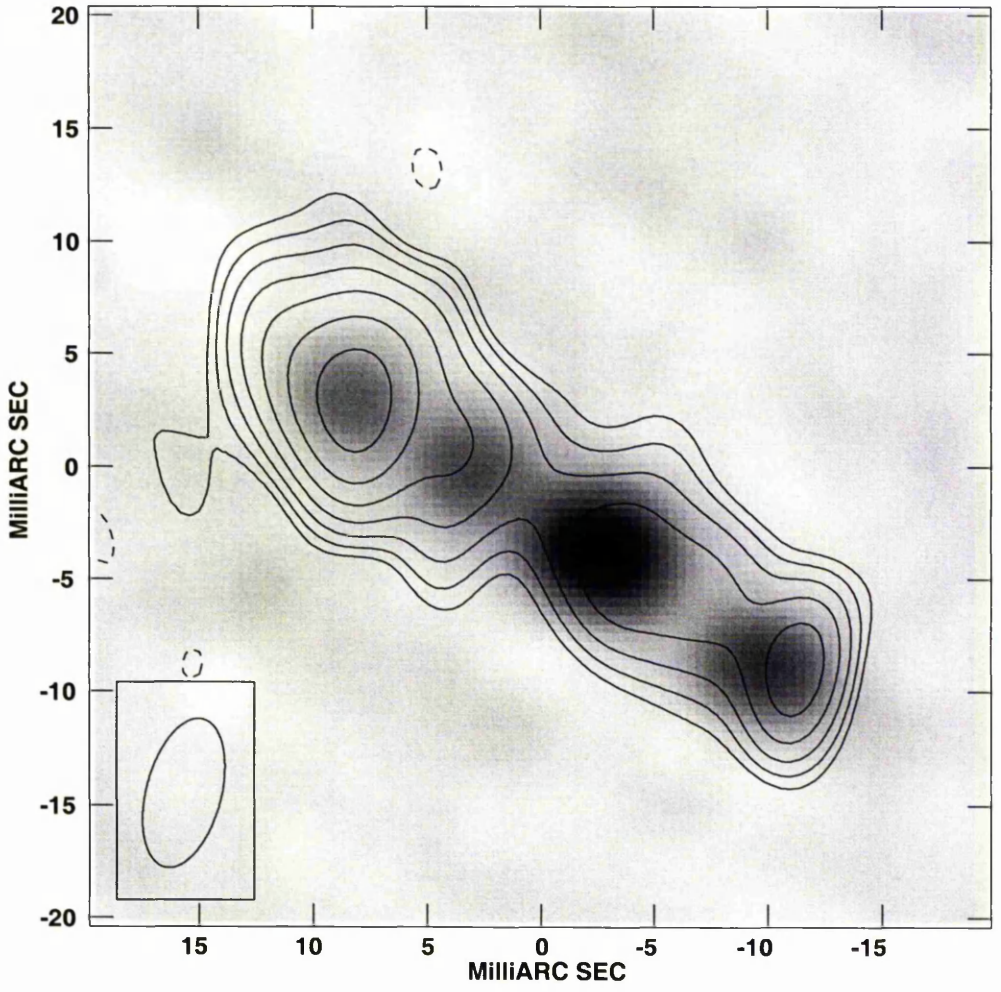


Figure 3.8: Left: the 1.7 GHz VLBA observations from 2002 December 01 plotted with the 5 GHz data from the same instrument. The 5 GHz data are shown in greyscale, while the 1.7 GHz data are shown contoured (with the same contour levels as in figure 3.7). The 1.7 GHz data have been imaged with uniform visibility weighting, which tends to increase resolution at the expense of surface brightness sensitivity. The 1.7 GHz beam is sized !!×!! mas, and is oriented at a PA of +!!°. Right: a spectral index map of the overlay. Broken contours represent negative values of the spectral index ν, while solid contours represent positive values. The central component has the flattest spectrum.

Component	Frequency [GHz]	Position offsets [mas]		Flux density [mJy]
		ΔRA	ΔDec.	
A1	1.7	+0±0.1	+0±0.1	51.4±3
A1a	5	+0±0.1	+0±0.1	20.5±1
A1b	5	+5.7±0.1	+3.5±0.1	8.8±0.4
A1c	5	+11.1±0.1	+6.6±0.1	7.3±0.4
A1d	5	−7.2±0.1	−5.0±0.1	9.2±0.5
A2	1.7	−300	−306	6.6±0.7
	5	-	-	<0.39
B1	1.7	+812.5±0.1	−827.5±0.1	5.6±0.3
	5	+822.5±0.1	−821.5±0.1	5.1±0.3
B2	1.7	+790.7±0.4	−732.9±0.5	2.2±0.2
	5	-	-	<0.39

Table 3.3: Model components for the radio observations of B0631+519 with the VLBA. The 1.7 GHz position offsets are given relative to the surface brightness peak in A1. For the 5 GHz data, position offsets are taken relative to A1a at RA: $06^h\ 35^m\ 12.31390^s$, Dec: +51° 57' 01.7949″(J2000.0).

continuum spectra, provided that the lens galaxy ISM does not scatter or absorb the source light, and that the source flux density is constant (Kochanek & Dalal, 2004).

The component flux densities for B0631+519 are plotted as a function of frequency in figure 3.9. The plot shows that A1 and B1 have similar radio spectra, as do A2 and B2. Possibly X, Y and Z also have similar radio spectra, although at 8.4 GHz component Y is a marginal detection and component Z is a non-detection (I have plotted a 5σ upper limit). The plot suggests that there are three separate regions of emission in the lensed radio source revealed by the radio observations, two of which are doubly-imaged and one of which is probably quadruply imaged (generating images Y, Z and the arc X). The geometry of the source will be discussed further in section 3.6.

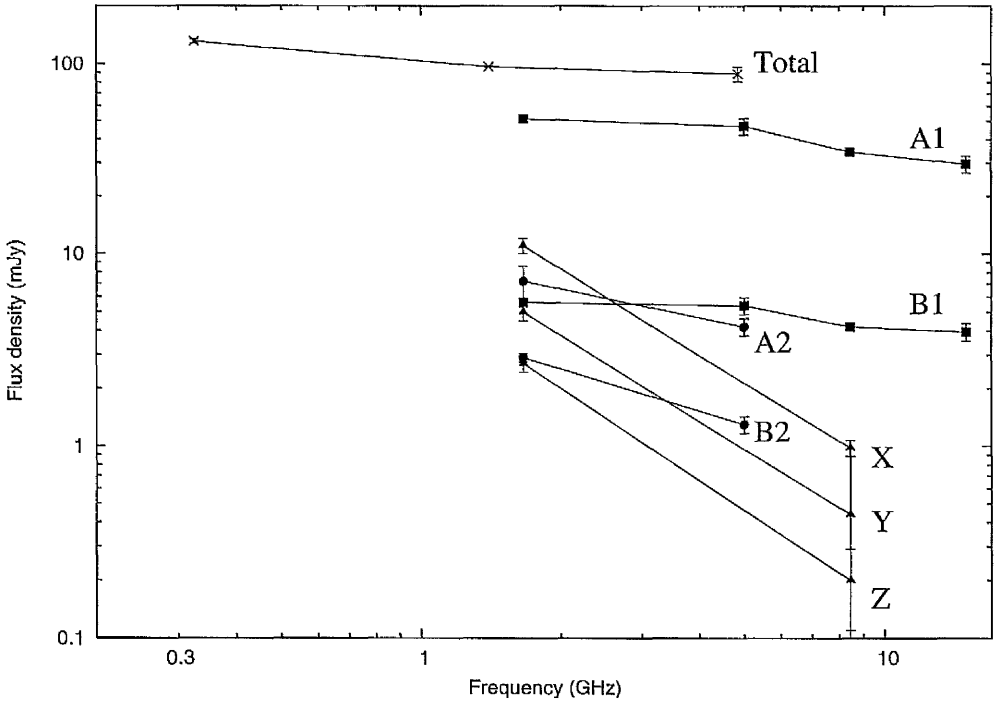


Figure 3.9: Plot of the radio spectra of the various radio components identified in B0631+519. The total flux density is based on measurements at 327 MHz from WENSS (Rengelink et al., 1997), at 1.4 GHz from NVSS (Condon et al., 1998) and at 5 GHz from the GB6 survey (Gregory et al., 1996).

3.5 Optical observations

Typically lens galaxies are radio-quiet ellipticals, so optical/infrared imaging is necessary to establish the lens galaxy position when the lenses are identified through radio searches. The position of the lens galaxy with respect to the lensed images can be used to constrain mass models, and in some systems unique information on the lensed source is available through optical imaging, for instance where an optical or infrared Einstein ring is visible as in CLASS B1938+666 (King et al., 1997, 1998). B0631+519 has been observed at optical/near-infrared wavelengths with the WHT through an R-band filter and with the HST through F555W, F814W and F160W filters. Details of these observations are listed in table 3.4. The HST optical data offer constraints on lens models for this system, in the form of the galaxy position relative to the lensed images and the lens position angle. The WHT data were used to extract an R-band magnitude for B0631+519. The HST near-infrared data show an Einstein ring, presumably the lensed host galaxy of the source. The information available from this resolved ring can constrain mass models once the lens galaxy light is subtracted.

3.5.1 WHT Observations

B0631+519 was observed with the 4.2 metre William Herschel Telescope (WHT) on 2002 February 03, through an R-band filter. The observations were made by the WHT's service observing programme. The detector used was a 1024×1024 pixel TEK CCD, with a plate scale of 0.108 arcsec pixel^{-1} and a total field of view of 111 × 111 arcsec. The seeing ranged from 0.7 to 0.9 arcsec, and the total integration time on B0631+519 was 1800 seconds taken in three exposures of 600 seconds each. The data were debiased, flat-fielded and dark-subtracted using the *Starlink* package CCDPACK.

Absolute astrometric calibration of the CCD-image was carried out using star

Instrument	Detector	Date observed	Filter	τ [sec]	θ_B [arcsec]	μ [arcsec pixel^{-1}]
HST	ACS/WFC	2003 August 19	F555W	2236	0.05	0.05
HST	ACS/WFC	2003 August 19	F814W	2446	0.08	0.05
HST	NICMOS/NIC2	2004 February 15	F160W	2560	0.13	0.075
WHT	Aux. port CCD	2002 February 03	R	1800	0.7-0.9	0.108

Table 3.4: Optical and IR observations made of B0631+519 to date. τ is the integration time, θ_B is the FWHM of the PSF, and μ is the pixel scale of the detector.

positions from the US Naval observatory catalogue A1.0. Five stars were used to establish plate centre, scale and rotation. The astrometric solution had a reported RMS uncertainty of 0.1" arcseconds, or approximately one pixel.

The WHT image is shown in figure 3.11. Since the maximum image separation in B0631+519 is about 1.16 arcseconds, the lens galaxy and lensed images remained blended, although the image of B0631+519 appears extended. Star profiles in the field appear slightly elliptical, being elongated by about 20-25% along PA +30°, possibly as a result of tracking errors reported in the observation log. B0631+519 itself is elongated by 30-40% along PA -30°. The data was therefore only used to measure the integrated magnitude of the lens plus source.

Photometry was undertaken with *Starlink*'s GAIA program using observations of Landolt standard stars 93 407 and 100 280 (Landolt, 1992) to account for atmospheric extinction and to calibrate the brightness scale in magnitudes. First, the counts received from the standard stars were measured using Source Extractor fits (Bertin & Arnouts, 1996). Kron optimal apetures were used (Kron & Shane, 1974). The measured counts C within each aperture were converted to instrumental magnitudes m_{inst}, using

$$m_{inst} = -2.5 \log_{10} \left(\frac{C}{\tau} \right), \tag{3.1}$$

where τ is the integration time in seconds. The Landolt magnitudes (m_{cal}) can be related to the instrumental magnitudes through

$$\Delta = m_{cal} - m_{inst} = \kappa\alpha + z_p, \tag{3.2}$$

where κ models atmospheric absorption, α is the air mass at which the observations are taken, and z_p is a zero point that relates the telescope magnitude system to the Johnson system (strictly, to the system used by Landolt). Landolt 93 407 was observed at two slightly different air masses, so there are three

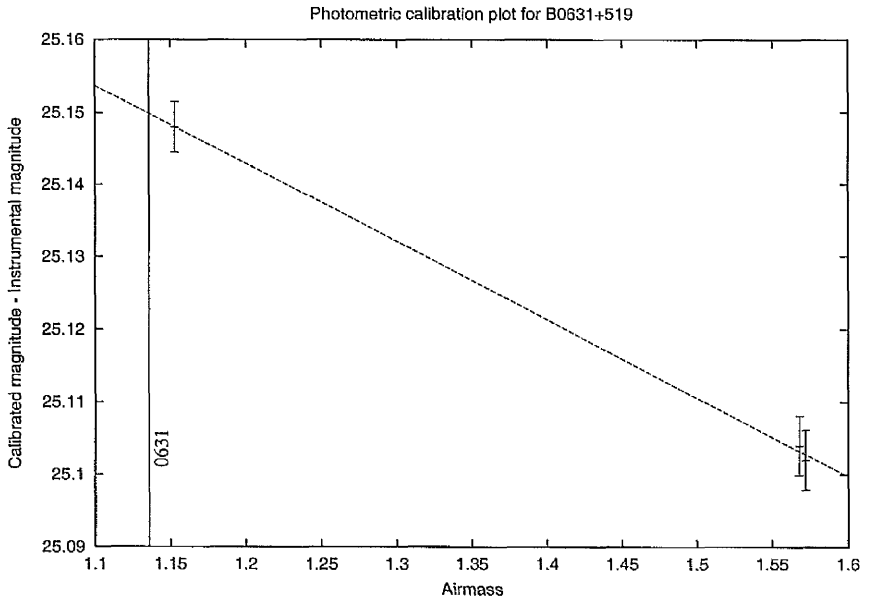


Figure 3.10: Calibration of zero point z_p and atmospheric absorption coefficient κ for the WHT observations of B0631+519. The best fit values found from least-squares fitting were $z_p = 25.262 \pm 0.005$ and $\kappa = -0.101 \pm 0.003$. The error estimates are likely to be unreliable because two of the points are nearly identical. The line marked "0631" marks the air mass at which B0631+519 was observed, $\alpha = 1.135$.

measurements of Δ available for three different values of air mass. κ and z_p were therefore established by least squares fitting, as shown in figure 3.10.

The instrumental magnitude of B0631+519 was measured in a Kron aperture having a radius of 4.29″and an ellipticity of 0.25. After applying the air mass correction and zero point derived from the Landolt standards, the R-band magnitude of B0631+519 is found to be 21.37±0.01; the quoted error estimate is the Poisson error expected in the recorded counts, but does not represent zero point uncertainties or other calibration error estimates.

For consistency with the later HST imaging, the R-band magnitude was corrected for galactic extinction - the HST magnitudes were corrected for galactic extinction because the lens and lensed images were well-resolved, and it was desirable to compare the lens galaxy colours with those of other lenses. Galactic

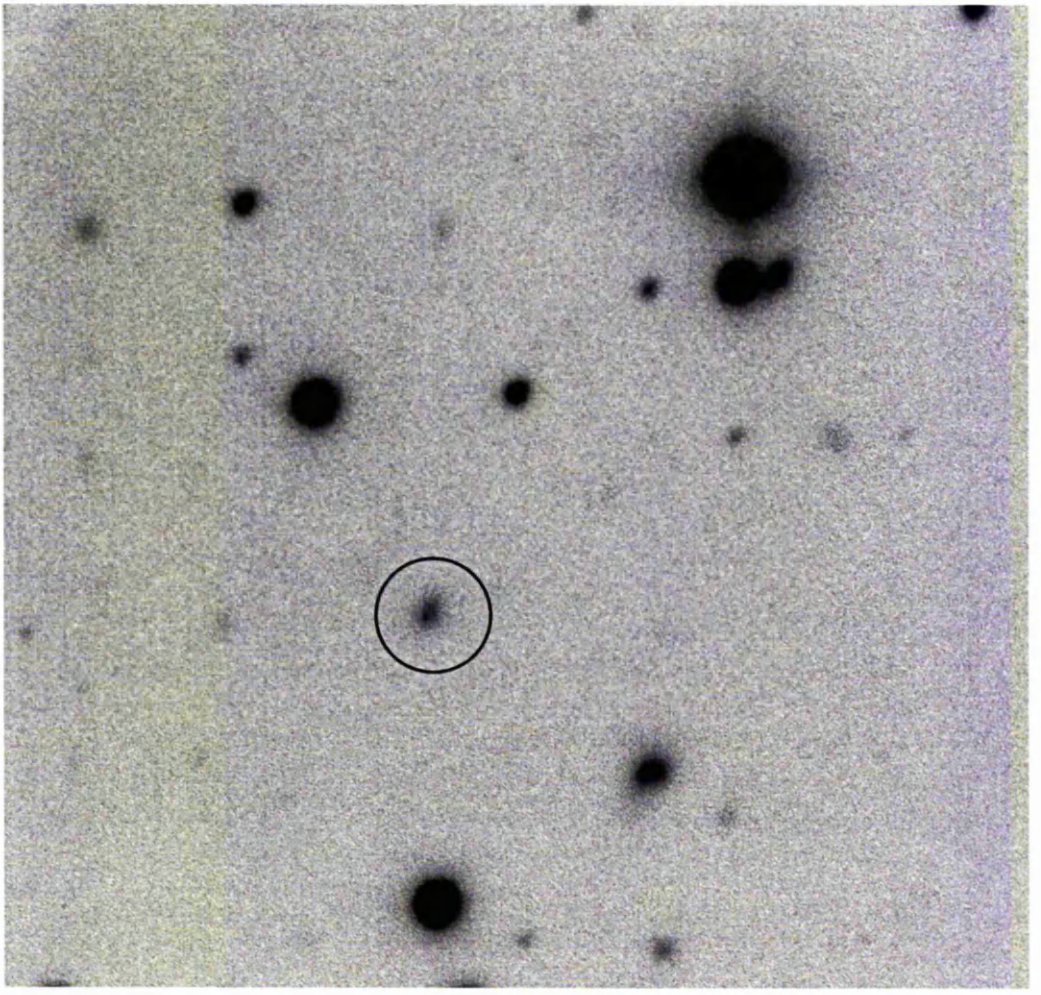


Figure 3.11: R-band image of B0631+519 from the WHT. The field of view is 111 × 111 arcsec. The lens is circled. North is up and east is left.

extinction towards the lens in the R-band was estimated from the dust maps of Schlegel, Finkbeiner & Davis (1998) at 0.25 magnitudes. The total R-band magnitude of the lens system (corrected for galactic and atmospheric extinction) was then found to be 21.1±0.1. The quoted error is believed to account sufficiently for calibration, Poisson and extinction model errors.

3.5.2 HST optical and near-IR Observations

B0631+519 was observed with the *Hubble Space Telescope* (HST) on 2003 August 19 and 2004 February 15 as part of HST proposal 9744, "HST Imaging of Gravitational Lenses" (PI: C.S. Kochanek). Images were taken using the Advanced Camera for Surveys (ACS) Wide Field Channel (WFC) through the F555W and F814W filters. The system was also observed with the NIC2 camera on NICMOS through the F160W filter in the MULTIACCUM observing mode. Details of the optical observations are shown in table 3.4. The ACS observations used a two-point dither pattern. Both pointings in the pattern were split into two exposures to assist with cosmic ray rejection. The two pointings were separated by 0.73 arcsec. The NICMOS data used a four-point dither pattern that placed B0631+519 at the centre of each quadrant of NIC2 in turn. The wideband HST filters F555W, F814W and F160W approximately correspond to Johnson V, I and H respectively.

The ACS and NICMOS data were processed through the OTFR (On-The-Fly Recalibration) facility of the MAST HST archive. The OTFR pipeline carried out the first stages of data reduction, i.e. debiasing, dark-subtracting and flat-fielding the ACS/WFC and NICMOS data. The OTFR pipeline was not used for distortion correction or combination of dithered data. I used the *multidrizzle* script (Koekemoer et al., 2002) from within *PyRAF* to correct the ACS data for distortion and to combine dithered pointings into a single output frame, the shifts between frames being obtained from application of cross-correlation through the *crossdriz* task in IRAF. Some masking was applied prior to sky-subtraction to cover a very large and bright diffraction spike from a star just beyond the detector's edge (tentatively identified as TYC3389, which has a visual magnitude of 7.5). The *multidrizzle* script also performed cosmic-ray rejection. The NICMOS data were combined using *pydrizzle* to interpolate the data onto an output image

with a specified pixel scale of 0.075 arcsec, using the standard distortion solution for NIC2. Cosmic rays in the NICMOS data were identified and blanked by the OTFR pipeline itself. The IRAF task *pedsub* was used to remove some residual dark signal that was evident in the output from the pipeline. All reduced images were aligned to the local celestial coordinate axes as part of the drizzling process.

The HST images (figure 3.12) reveal two galaxies at the lens system position, surrounded by arcs of emission from the lensed source at longer wavelengths; the galaxies are labelled L1 and L2. L2 and L1 are detected in the F555W data but no emission from the lensed images is seen. L1 appears to be an elliptical galaxy with a smooth profile while L2 appears to have an irregular surface-brightness profile. The F814W image clearly shows L1. Some emission from the source in the form of two lensed arcs is detected. L2 is still detected but is considerably fainter than L1, and is partially confused with one of the arcs so measurements of its brightness are difficult. L1 apparently contributes most of the lensing mass since the arcs are centred about L1, not L2. The NICMOS image again shows L1 surrounded by two arcs of extended emission, consistent with the lens hypothesis. L2 is not cleanly detected, and may not be visible at such long wavelengths. The lensed arcs are clearly visible.

The ACS images have a large enough field of view to show the environment of B0631+519. The nearest extended object is located 4.4 arcsec away at a PA of -72°, and has a total $I_{mag} \sim 22.3$ in a 4 arcsec diameter aperture. The object's light distribution is not smooth and is difficult to classify. A cutout of the F814W ACS image, sized 10×10 arcsec, is shown in figure 3.13.

I carried out surface-brightness fitting on L1 using a Sérsic light profile (Sérsic, 1968) convolved with appropriate PSFs produced by TinyTim (Krist, 1993). The Sérsic profile is a generalisation of the de Vaucouleurs light profile (de Vaucouleurs, 1948), and has the form,

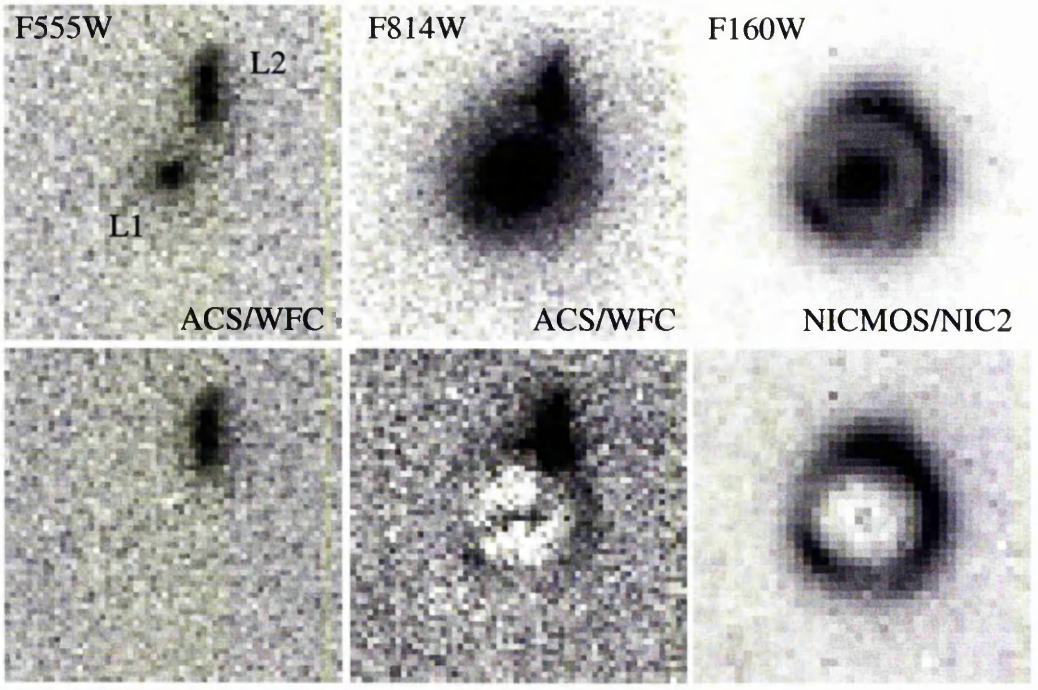


Figure 3.12: HST images of B0631+519. The top row of images show the lens before subtraction of L1, while the bottom row shows the residuals after subtraction of the best-fit Sérsic profile. North is up and east is left. The images are 3×3 arcsec in size, and are centred on L1.

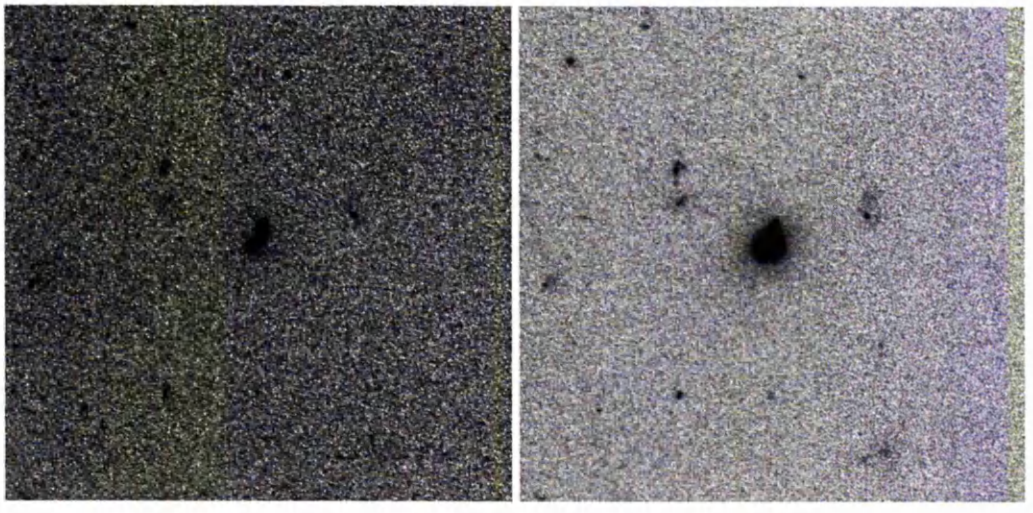


Figure 3.13: F555W/F814W images of the surroundings of B0631+519. North is up and east is left. The F555W image is on the left. Both images are 10×10 arcsec in size, and are centred on B0631+519.

Parameter	Filter		
	F555W (V)	F814W (I)	F160W (H)
Offset in RA [arcsec]	+0.36±0.08	+0.38±0.08	+0.36±0.08
Offset in Dec. [arcsec]	−0.50±0.08	−0.53±0.08	−0.50±0.08
Total magnitude	22.7	20.0	17.7
Half-light radius [arcsec]	0.54	0.56	0.55
Ellipticity	0.09	0.17	0.12
Position angle	−48	−53	−57
Power-law index	0.29	0.28	0.30

Table 3.5: Parameters of the best-fit Sérsic profiles fitted to L1 in the three HST images. A power-law index of 0.25 corresponds to a de Vaucouleurs surface-brightness profile. The total magnitudes are specified in the Johnson photometric system, corrected for galactic extinction. Position offsets are given relative to image A1.

$$s(r) = s_0 \exp\left[-(r/r_0)^{\nu}\right], \tag{3.3}$$

where $s(r)$ is the surface brightness at radius r, s_0 and r_0 are scale factors and ν is a power-law index; the Sérsic model becomes a de Vaucouleurs profile when ν is set to 1/4. The Sérsic profile parameters were varied to minimise the sum of squared residuals between the data and the model. When fitting L1's surface-brightness profile I masked out L2 and the lensed images to avoid disturbing the fit. L2 and the lensed arcs do not have easily parametrizable surface light distributions, so I did not attempt to separate them in the F814W or F160W images, and measured only the sum of light remaining after subtraction of L1. In the F555W image the arcs were not detected, so I used GAIA to fit an elliptical aperture to L2 (after L1 had been subtracted). The centroid of L2 was found to be −0.33 arcsec east and +0.74 arcsec north of L1. The axis ratio of the best-fit aperture was 0.56, with the major axis having a position angle −1.1°east of north. Table 3.5 shows the properties of the Sérsic profiles subtracted from L1.

The ACS and NICMOS images following subtraction of L1 are shown in figure

Source	Brightness			
	V	R	I	H
Total	22.1±0.1	21.1±0.1	19.9±0.1	17.3±0.1
L1	22.8±0.1		20.0±0.1	17.8±0.1
L2	23.0±0.1			>23
L2+images			22.3±0.1	
Images	>25			18.3±0.1

Table 3.6: Results of photometry on B0631+519. The figures reported in this table were measured within a circular aperture of 4 arcsec diameter. For the total luminosity of the subtracted Sérsic models see table 3.5. The WHT R-band magnitude was calculated using a Kron elliptical aperture 4.3 arcsec in (major axis) diameter. All figures have been corrected for galactic and, in the case of the WHT, atmospheric extinction. Non-detections are reported as lower limits on magnitude at the 5σ level.

3.12. The Sérsic profile is a reasonable fit to the HST data, leaving maximum residuals of 3% and 4% of L1's peak brightness in the case of the F814W and F160W images respectively. The residuals are undetectable within the noise in the case of the F555W image, which has the lowest signal-to-noise ratio of all the HST observations.

To measure the magnitudes of L1, L2 and the lensed images, I carried out aperture photometry on B0631+519. I used a circular aperture 4 arcsec in diameter to measure the flux of the lens system as a whole, of the Sérsic model for L1, and finally of whatever light remained after subtraction of the Sérsic model from the data. The total counts in the aperture (after background subtraction) were converted into the HST magnitude system (STMAG) following the procedure outlined in Pavlovsky et al. (2005). The measurements were then converted to the Johnson magnitude system using the relevant published zero points for HST magnitude systems. Galactic extinction was estimated from Schlegel, Finkbeiner & Davis (1998). The results are shown in table 3.6. The differences in L1 magnitudes between tables 3.5 and 3.6 are due to the fact that the values given in the

latter are sums of the flux within finite apertures, while those in the former are equivalent to sums within infinite apertures.

The lens galaxy position relative to the lensed images, when known, provides two constraints on simple parametric lens models under the assumption that the centre of (smoothly distributed) light and the centre of mass for a galaxy are coincident. Since the lens galaxy in B0631+519 is apparently radio-quiet, the radio and HST data must be overlaid as accurately as possible to establish the lens position. The lensed images at near-infrared and optical wavelengths exhibit no obvious point-like components that could be unambigiously aligned with the radio structure. It is necessary therefore to rely on absolute astrometric calibration based on stars visible in the field to align the optical and near-infrared images with radio data.

I used the 2MASS[3] catalogue with *Starlink*'s GAIA to establish an astrometric calibration for the ACS images (F555W and F814W filters). The astrometric solutions were based on twelve star positions and had reported RMS errors of ± 0.08 arcsec in RA and declination, although examination of objects around the field indicated positional agreement between the two calibrated images to within 0.05 arcsec. In the case of NICMOS the relatively small field of view (19.5×19.5 arcsec) meant that only one star near to B0631+519 was imaged. The centre of this star was used as a reference to fix the absolute astrometry for the image. I estimate the astrometric error in the resulting solution for the drizzled NICMOS image at 0.08 arcsec, since the star's position on the sky was derived from the astrometrically calibrated ACS F555W image. The rotation of the image was checked by calculating the position angle of the vector from the field star to L1's position. This angle agreed with the same quantity measured from the ACS images to within 0.05°. Figure 3.14 shows the 1.7 GHz MERLIN map overlaid on the ACS and NICMOS images.

[3]See: http://pegasus.astro.umass.edu

Fitting Sérsic profiles to L1 gave best-fit pixel positions for the galaxy centre in all three HST filters. I converted these pixel positions into celestial coordinates using the absolute astrometric solutions from the HST images. The sky position of image A1 (from the 1.7 GHz VLBA observations) was then subtracted to leave L1's offset on the sky from A1. These offsets are reported in table 3.5. The agreement between the offsets from the F555W and F160W images (which effectively share the same astrometric solution) and the disagreement between F555W and F814W suggests that the error in the galaxy position is dominated by the random error in the astrometric solutions for F555W and F814W; error contributions from the galaxy subtraction procedure and from the distortion solutions for each camera are negligible by comparison. The mean offset from A1 to L1 (based on the ACS data only) is $+0.37\pm0.08$ arcsec in RA and -0.52 ± 0.08 arcsec in declination.

3.5.3 Redshift-dependent properties of B0631+519

McKean et al. (2004) have obtained an optical spectrum of B0631+519 with the W. M. Keck telescope. They report the discovery of two distinct galaxies at two different redshifts. The first is an emission line galaxy at z=0.0896±0.0001; and the second, which is the dominant source of emission in the spectrum, is an elliptical galaxy at z=0.6196±0.0004. These spectral classifications are consistent with the morphologies, colours and flux densities of the two foreground galaxies detected in the HST imaging. Therefore, the z=0.0896 emission line galaxy is L2, and the z=0.6196 elliptical is L1. McKean et al. (2004) refer to these galaxies in redshift order as G1 and G2 respectively; the name order was swapped to L1 and L2 because the HST imaging demonstrates that L1 (G2) is the primary lensing galaxy, something that was not obvious from the spectroscopy alone.

McKean et al. (2004) also discuss the tentative detection of a single emission

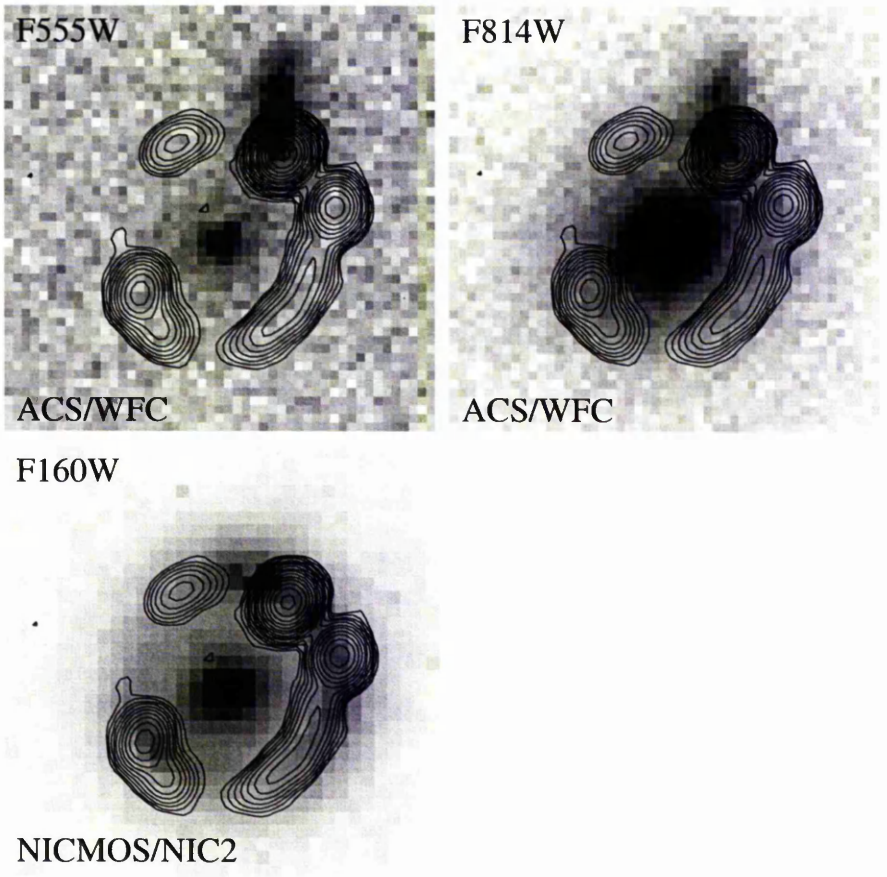


Figure 3.14: The MERLIN 1.7 GHz radio data (contoured) overlaid on the NICMOS and ACS images (greyscale). The ACS images cover 2.45×2.45 arcsec, while the NICMOS image covers 2.475×2.475 arcsec. All images are centred on L1. North is up and east is left.

line from the lensed source at 5040Å. Assuming that the line is either due to Lyα or Mg II emission, they suggest 3.14 or 0.80 for the source redshift, respectively. However, further spectroscopy will need to be carried out to test these possibilities. There is insufficient (unambiguous) photometric information on the lensed images to attempt to choose between these possibilites by using photometric redshift techniques, but, in the absence of further spectroscopy, it is possible to attempt to constrain the source redshift by determining the rest-frame luminosity L of L1 and converting it into a velocity dispersion σ using the fundamental plane of elliptical galaxies (Djorgovski & Davis, 1987; Dressler et al., 1987). This approach has been used before by Tonry & Kochanek (2000) to estimate the source redshift in CLASS B1938+666.

I converted the apparent magnitudes of L1 and L2 as measured by the HST (after translation into Johnson filters V, I and H, and following application of extinction corrections) into absolute magnitude measured in Johnson B using

$$M_B = m_Q - 5\log\left(\frac{D_L(z)}{10pc}\right) - K_{BQ}(z), \tag{3.4}$$

where M_B is the predicted absolute magnitude of the object in B, m_Q is the magnitude of the object in some filter Q, $D_L(z)$ is the luminosity distance[4] to the object at redshift z and $K_{BQ}(z)$ is the K-correction from filter Q to B (Hogg, 2002). I adopt the redshifts of McKean et al. (2004) for L1 and L2 (z_{L1} and z_{L2}). The K-corrections were calculated using the SYNPHOT package in IRAF, together with template spectra for L1 and L2 from Bruzual & Charlot (1993) and Kinney et al. (1996). To calculate K_{BQ} for a particular template, SYNPHOT was first used to redshift the template to z_{L1} or z_{L2} as appropriate. The redshifted spectrum was then multiplied by the transmittance curve of Q (representing the

[4]All cosmological distances in this work were calculated assuming a flat universe with $\Omega_m = 0.27$ and $\Omega_\Lambda = 0.73$.

filter in question) and the transmitted flux summed over wavelength to get the expected apparent magnitude of the template galaxy when seen through the filter, m'_Q. The same method was used to determine the expected apparent magnitude of the un-redshifted template when seen through a Johnson B filter, m'_B. The K-correction is then

$$K_{BQ}(z) = m'_Q - m'_B. \tag{3.5}$$

The normalisation of the template is irrelevant as long as it remains the same for both m'_Q and m'_B.

The K-correction for L1 was obtained from the model elliptical galaxy spectra of Bruzual & Charlot (1993), who provide multiple templates based on various star formation histories and initial mass functions (IMFs). I selected the template that produced the lowest scatter in the predicted rest-frame B magnitudes for L1 amongst the three filters (F555W, F814W and F160W). That is, for each template in turn I calculated the K-corrections for all three filters, and then used equation 3.4 to calculate the corresponding three values of M_B. The template that produced minimum scatter between the three values of M_B was accepted. The optimum template gave an RMS scatter amongst the three filters of ~ 0.02 mag, and was generated assuming a Salpeter IMF (Salpeter, 1955) with mass range from 0.1 to 125 $M_\odot$, in which star formation took place >5 Gyr ago during a burst lasting 1 Gyr.

The K-correction for L2 was estimated using the starburst galaxy spectrum templates of Kinney et al. (1996). Since there are several templates provided, each having differing amounts of internal extinction, I again selected the template that produced the lowest scatter in the predicted rest-frame B magnitudes for L2, but between the F555W and F814W filters only (because L2 is not detected in F160W). This optimum template had $0.15 < \mathrm{E(B\text{-}V)} < 0.25$ and the resulting

scatter was $\sim$0.01 mag.

Applying equation 3.4 to the HST photometry I find a rest-frame magnitude $M_B = -20.57 + 5\log h \pm 0.1$ for L1, and $M_B = -13.95 + 5\log h \pm 0.1$ for L2, where I use $H_0 = 100h$ km s^{-1} Mpc^{-1}. I follow Rusin et al. (2003) and adopt a characteristic rest-frame absolute magnitude in B (M_*), of $-19.7 + 5\log h$. For L1 I find a luminosity L of $2.2 \pm 0.2 L_*$, and for L2 I find $L = 5.0 \pm 0.5 \times 10^{-3} L_*$ (where L_* is the local value).

The fundamental plane of ellipticals is defined by

$$\log R_e = \alpha \log \sigma_c + \beta < \mu >_e + \gamma, \tag{3.6}$$

in which R_e is the effective radius (the radius containing half the total light) of the galaxy in kpc, σ_c is the central velocity dispersion of the luminous matter, $< \mu >_e$ is the galaxy's effective surface brightness in mag arcsec^{-2}, and α, β and γ are empirically determined constants. The effective surface brightness is the mean surface brightness within the effective radius, and is given by

$$< \mu >_e = m_Q - K_{BQ}(z) + 2.5\log 2\pi + 5\log r_e - 10\log(1+z), \tag{3.7}$$

in which r_e is the effective radius in arcseconds, and m_Q is the total magnitude of the object (i.e. the magnitude within an infinite aperture). I have used the total magnitude derived from the Sérsic fits, on the assumption that the extrapolation introduces minimal additional error compared to the uncertainties in K-corrections and magnitude system zero points.

I adopt the values for α, β and γ found by Bender et al. (1998) for local ellipticals, $\alpha = 1.25$, $\beta = 0.32$, $\gamma = -8.895$ for $H_0 = 50$ km s^{-1} Mpc^{-1}. The fundamental plane remains valid at redshifts comparable to that of L1 (eg Treu et al. 1999), as long as it is corrected for luminosity evolution. Treu et al. (1999) determine the intercept of the fundamental plane as a function of redshift; at the

redshift of L1, the correction to the local value is ∼0.4 magnitudes. Assuming that L1 lies precisely on the corrected fundamental plane, I use the results from the Sérsic fits to estimate $\sigma_c = 223^{+46}_{-39}\ h^{-0.8}$.The same value can be arrived at by using the local fundamental plane intercept and correcting the surface brightness defined in equation 3.7 using the conclusions of Treu & Koopmans (2004).

It is possible to make an assumption of isothermality for the lens galaxy L1 in order to estimate the source redshift. Lens galaxies are typically found to have mass profiles within 15% of isothermal within the radius of the lensed images (Rusin et al., 2002; Koopmans & Treu, 2003; Rusin, Kochanek & Keeton, 2003; Treu & Koopmans, 2004; Rusin & Kochanek, 2004). For an isothermal galaxy the (dynamical) velocity dispersion can be related to a function of the source redshift (Kormann, Schneider & Bartelmann, 1994), because the velocity dispersion determines the angular separation of lensed images, $\Delta\theta$, through

$$\Delta\theta = 8\pi\left(\frac{\sigma_{iso}}{c}\right)^2 \frac{D_{ls}}{D_s}, \tag{3.8}$$

where σ_{iso} is the velocity dispersion of the isothermal mass (not necessarily equal to σ_c), and $d_{\rm ls}$ and $d_{\rm s}$ are the angular diameter distances from the lens to the source and from the observer to the source respectively. In figure 3.15 I have used this relation with the observed image splitting to plot the expected σ_{iso} as a function of source redshift for $h = 0.7$, assuming that σ_c and σ_{iso} are approximately equal (Kochanek, 1994). I also plot the two source redshifts suggested by McKean et al. (2004), which between them conveniently span much of the plausible source redshift space. The estimated velocity dispersion of L1 implies that the source redshift z_s is less than 1.8 at 95% confidence, given the assumption of an isothermal mass distribution in L1[5].

[5]In the published paper York et al. (2005), the source redshift is constrained to be greater than 1.2. This is only the case when no luminosity evolution correction is applied.

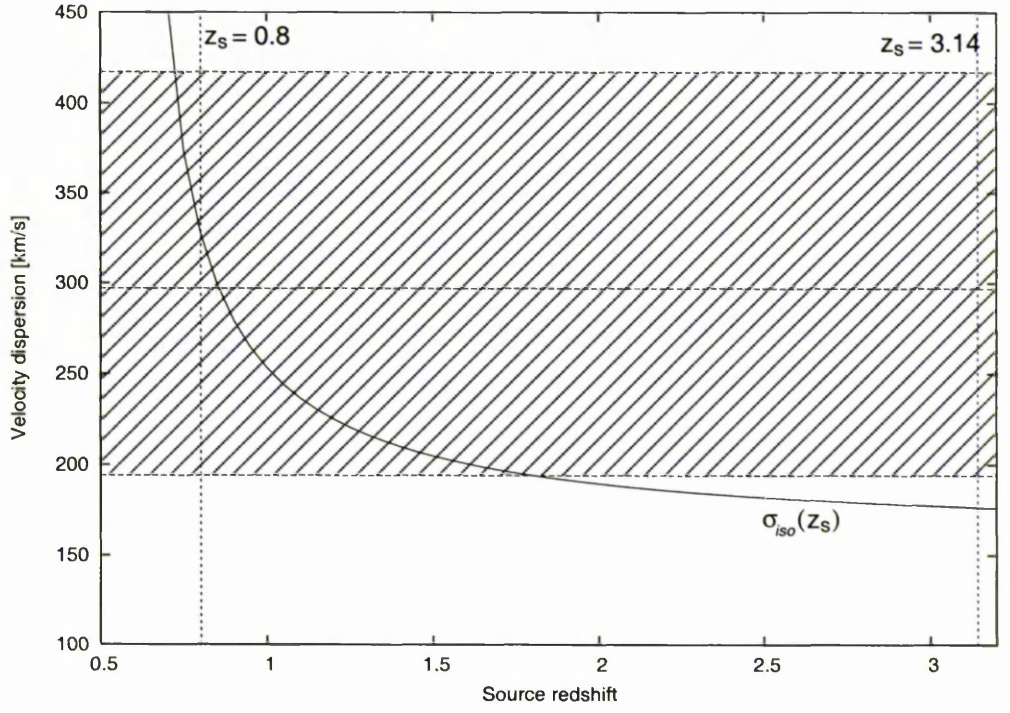


Figure 3.15: Velocity dispersion of L1 (σ_{iso}) as a function of source redshift (z_s), assuming an isothermal mass distribution in the lens. The shaded area represents the velocity dispersion derived from the fundamental plane at the 95% confidence level when $h = 0.7$. Assuming an isothermal mass distribution, the method gives $0.7 < z_s < 1.8$ with 95% confidence.

Further spectroscopy will be essential if the radio source proves to be time-variable, since then B0631+519 would be a potential target for attempts to measure the Hubble constant. A spectroscopic source redshift would also be valuable for normalising mass models of this system.

3.6 Lens modelling of B0631+519

In this section I discuss parametric models of the mass distribution in B0631+519, using the code described in chapter 2. Models were constrained with the 5 GHz MERLIN data on A1, A2, B1 and B2 because these images are relatively compact. I use the optimum mass model to examine L1's mass to light ratio and the geometry of the lensed source.

3.6.1 Mass modelling using the 5 GHz MERLIN data

The fit statistic used is the image-plane minimisation quantity described in chapter 2 (equations 2.50 and 2.51); it is defined over N images having angular positions $\boldsymbol{\theta}_i$ $(i = 1...N)$ and flux densities s_i by

$$\chi^2 = \sum_{i=1}^{N} \frac{|\boldsymbol{\theta}_i - \boldsymbol{\theta}_i'|^2}{{\sigma_{pi}}^2} + \sum_{i=1}^{N} \frac{|s_i - s_i'|^2}{{\sigma_{fi}}^2}, \tag{3.9}$$

where the primed quantities are values predicted by the lens model being evaluated and the σ_{pi} and σ_{fi} quantities account for errors in positions and flux densities respectively. The code assumes that the observed images are unresolved, which is clearly the case for A1 and B1 in the MERLIN 5 GHz data. A2 and B2 are slightly resolved but are still well-represented by Gaussian models. Components A1a, A1b, A1c and A1d will become useful if future high-resolution VLBI observations are able to resolve B1 into corresponding images. Without observed

counterparts in B, the VLBA sub-components in A1 cannot yet be used to constrain the lens model. Images X, Y and Z are very extended and would be best treated using tools such as LENSMEM (Wallington, Kochanek & Narayan, 1996) or LENSCLEAN (Kochanek & Narayan, 1992; Wucknitz, 2004).

Because A1 and B1 do not arise from the same source as A2 and B2, this system has fewer degrees of freedom immediately available than for a single quadruply imaged source. I ignored any contribution L2 might make to the lens effect and modelled L1 as a singular isothermal ellipsoid (SIE; Kormann, Schneider & Bartelmann 1994), whose position, ellipticity and position angle were allowed to vary freely. The constraints available for this simple lens model were then four sets of image positions and flux densities. The free parameters were the positions and flux densities for two sources, the lens position, the lens mass scale and the position-angle and ellipticity of the mass distribution. The number of degrees of freedom, ν, is therefore 1. I allowed 20% errors on image flux densities to account for possible source variability, and used the positions and position errors derived from the MERLIN 5 GHz data for A1, A2, B1 and B2 (because the VLBA observations resolve A2 and B2). The model parameters are shown in table 3.7, and the lens is illustrated in figure 3.16. The optimum model has a reduced χ^2 of 0.41, with roughly equal contributions from flux densities and image positions.

I fixed the galaxy position to that determined optically in section 3.5, and added an external shear to the model. This SIE+shear model was then fitted to the MERLIN image positions and flux densities. The optimum model had a reduced χ^2 of only 0.07 for 1 degree of freedom, suggesting that the random errors on the constraints were over-estimated. Decreasing the flux density errors from 20% to 5% results in a reduced χ^2 of 0.5 which is dominated by the contributions from the image positions. The optimum model parameters do not change significantly when the flux density errors are reduced. The optimum model parameters

Model	SIE	SIE+shear
Lens offset in RA, g_x [arcsec]	$+0.398\pm0.023$	+0.37 (fixed)
Lens offset in Dec., g_y [arcsec]	-0.519 ± 0.005	−0.52 (fixed)
Critical radius, b [arcsec]	0.603 ± 0.002	0.605 ± 0.003
Ellipticity, ϵ	0.13 ± 0.03	0.07 ± 0.04
Lens PA, θ_ϵ	$-57.1\pm0.1°$	$-28\pm12°$
Shear strength, γ		0.03 ± 0.01
Shear PA, θ_γ		$-81\pm0.5°$
χ^2/ν	0.41/1	0.07/1

Table 3.7: Lens models for B0631+519. Lens position offsets are given with respect to image A1, and position angles are measured east of north. The errors in the lens parameters were estimated using Monte Carlo simulations, as shown in figures 3.18 and 3.19.

are listed in table 3.7 and the model is illustrated in figure 3.17. The position of the best SIE-only lens is consistent with the optical lens position within the errors. The model-predicted time delay between A1 and B1 is 10.2 h^{-1} days with the SIE model or 8.5 h^{-1} days with the SIE+shear model.

The errors quoted on the model parameters in table 3.7 were computed by a Monte Carlo approach. 1000 realisations of the data were constructed in which the values of the observables (image positions and fluxes) were chosen using normally distributed random numbers, whose means and variances were set according to the measured values and variances of the observables. These realisations were then fed to the lens modelling code, which varied one model parameter at a time to achieve an optimum fit, while holding the rest of the model fixed to their best known values. The resulting models were saved and a histogram constructed for each variable parameter, which could then be used to determine reasonable error ranges (some of the parameters have very non-gaussian distributions). These histograms are shown in figures 3.18 and 3.19.

The modelling confirms that the radio components can be understood as

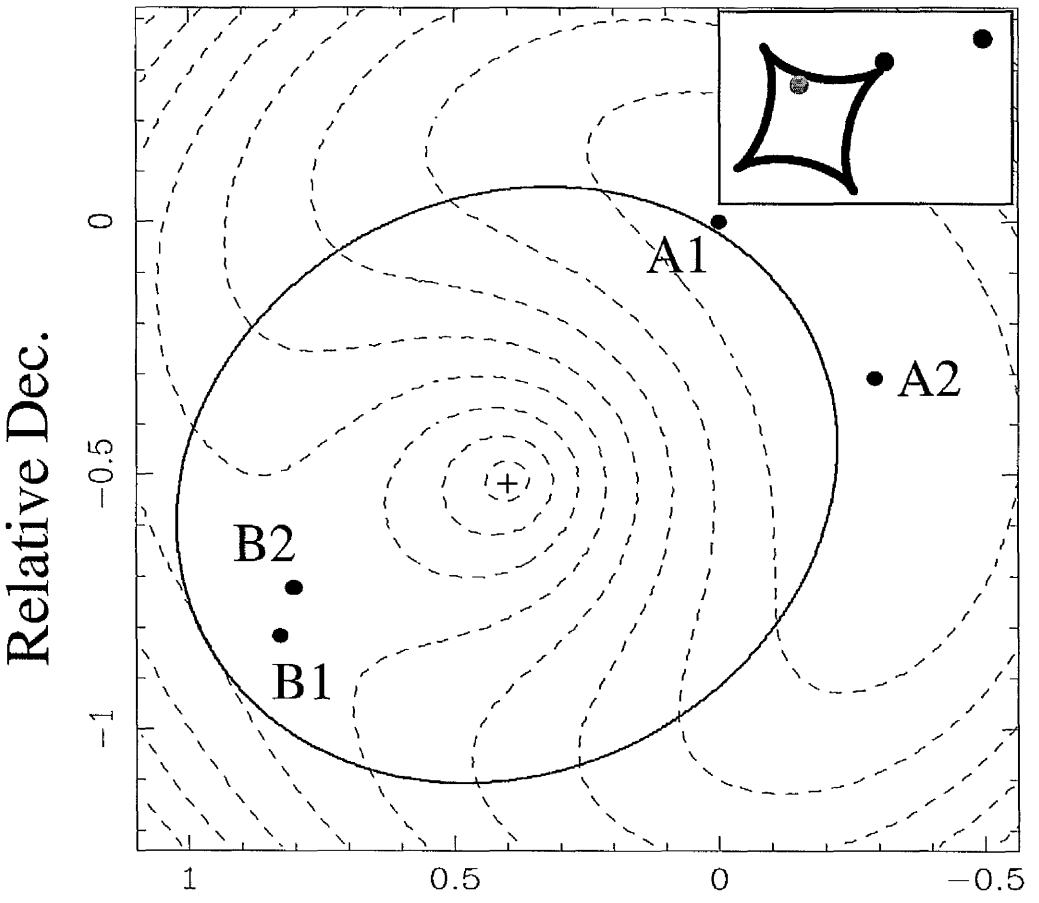


Figure 3.16: This figure shows the best SIE model. The ellipse is the tangential critical curve. The dashed contour lines represent the time delay surface of the lens relative to image A1, and begin at 1.23 h^{-1} days (for the first contour surrounding A1/A2) and increase in steps of 2.05 h^{-1} days. The lensed images are represented by filled circles. The inset diagram shows the source plane; the astroid curve is the tangential caustic, and filled circles represent the modelled source positions. The estimated position of the source responsible for images X, Y, Z and the arc X is marked by a light grey filled circle.

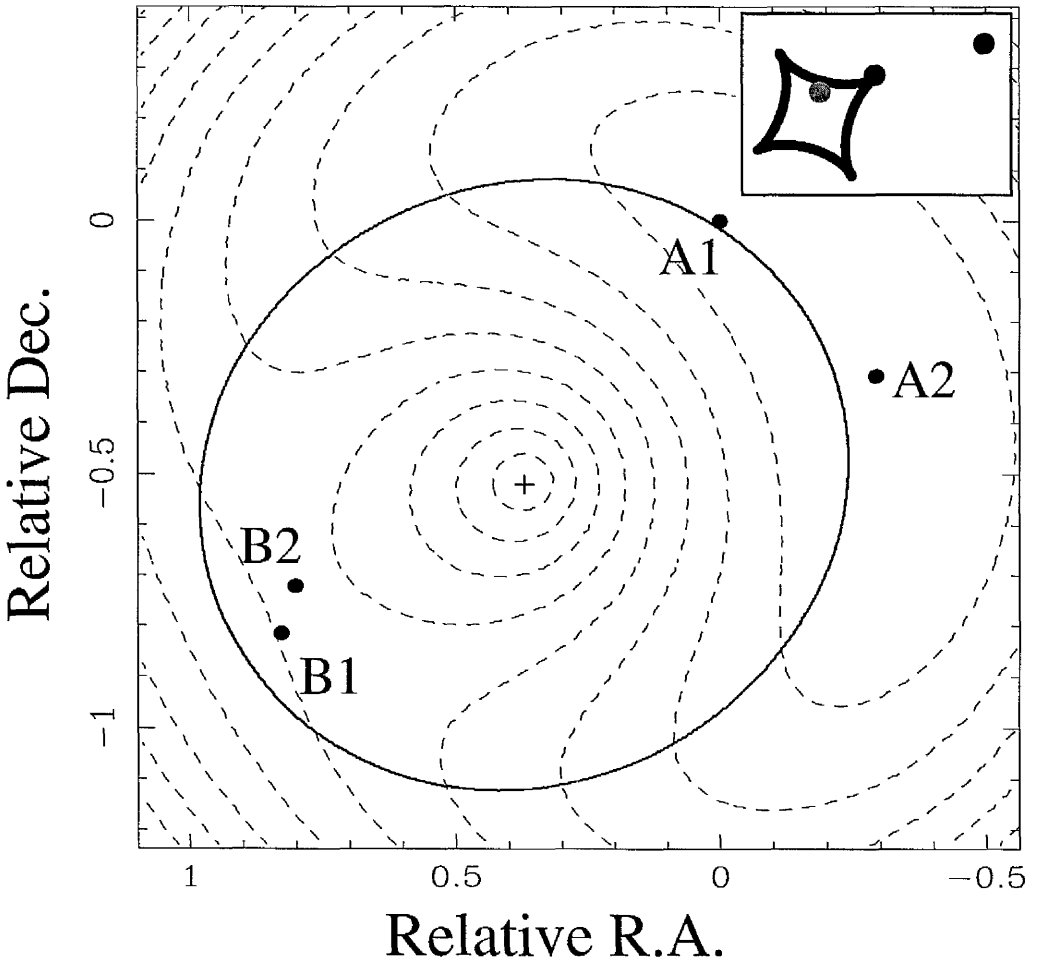


Figure 3.17: Lens models applied to B0631+519. This is the best SIE+shear model found. The lens galaxy position is fixed by the measured optical/IR galaxy position. The first contour surrounding A1/A2 represents 1.11 h^{-1} days, and the contours increase in steps of 1.88 h^{-1} days. All other conventions are the same as in figure 3.16.

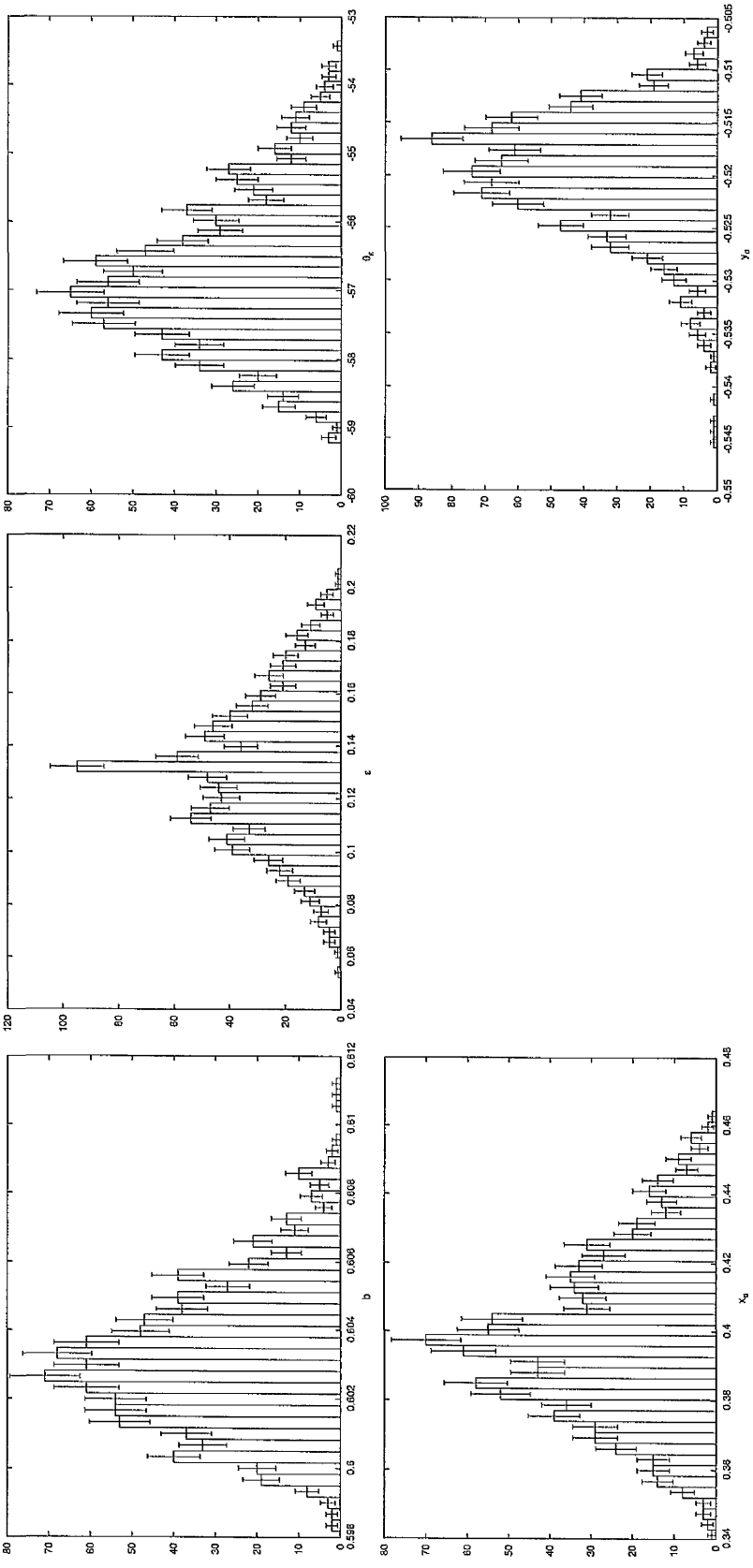


Figure 3.18: Results of a Monte Carlo approach to estimating errors on lens model parameters. Each histogram is the product of 1000 realisations of the data. The model used here is an SIE with no external shear but a free choice of galaxy position. The parameter symbols are the same as in table 3.7.

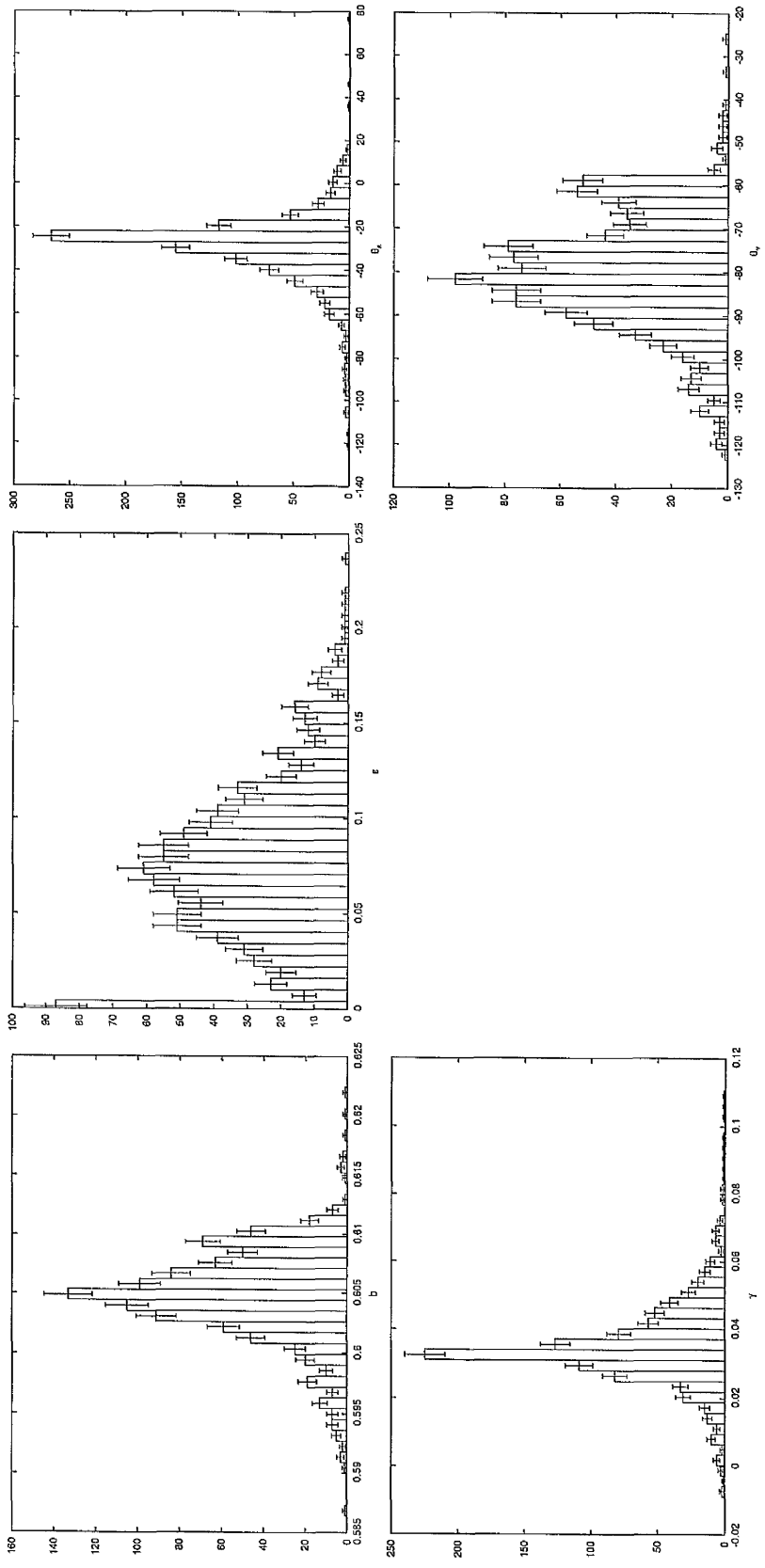


Figure 3.19: The model used to generate these histograms is an SIE with an external shear. The SIE's position is held fixed to the value estimated from optical data.

lensed images of two distinct sources. A1 and B1 are then images of a compact radio core having a flat (synchrotron self-absorbed) spectrum, while A2 and B2 arise from a more extended region of emission to the west of the core, probably a radio lobe. Images A1 and A2 have positive parity, while images B1 and B2 have negative parity. Incorporating images X, Y and Z is more difficult because they are all relatively extended compared to the A and B images, and the radio spectra indicate that these images arise from a third source. Back-projecting the positions of the Y and Z images (from the 1.7 GHz MERLIN data) to the source plane, along with the brightest point on arc X, produces an approximate source position of $+0.41''$east and $-0.49''$north relative to A1 (marked on figures 3.16 and 3.17). The model thus suggests that the X, Y and Z images are produced by a counter-lobe on the eastern side of the source that happens to lie inside the 4-image caustic. With only the VLA data, showing X and Y but no Z, a model involving emission only on the western side of the source's core might also be plausible.

3.6.2 Mass to light ratio of L1

It is possible to estimate the B-band mass-to-light ratio of L1 in the region enclosed by the lensed images. The mass M between the lensed images of a source (in the case of a circularly symmetric lens) can be related to the image separation $\Delta\theta$ and the angular diameter distances D_l (observer to lens), D_s (observer to source) and D_{ls} (lens to source) by

$$M = 3.074 \times 10^7 \, \frac{D_l D_s}{D_{ls}} \, (\Delta\theta)^2, \tag{3.10}$$

where M is given in terms of $M_\odot$, $\Delta\theta$ is given in arcseconds, and the angular size distances D_l, D_s and D_{ls} are all given in Mpc. The estimate is significantly affected by the unknown source redshift (but not by assumptions concerning

the distribution of mass inside the radius of the images). The ratio of angular diameter distances in equation 3.10 decreases by a factor of ~3.5 when the source redshift is increased from 0.8 to 3.14. I therefore estimate the lensing mass using both of the suggested redshifts for the source listed in McKean et al. (2004), since they conveniently span most of the range of plausible source redshifts. The separation between images A1 and B1 is 1.162±0.003 arcsec, which at the redshift of L1 corresponds to a linear distance of $5.6h^{-1}$ kpc. From equation 3.10 I estimate the mass of L1 within the images at $2.17 \pm 0.01 \times 10^{11}h^{-1}\ M_{\odot}$ using a source redshift of $z_s = 0.8$ or $6.27 \pm 0.03 \times 10^{10}h^{-1}\ M_{\odot}$ using $z_s = 3.14$. Combining these estimates with the B-band luminosity derived from the Sérsic profile fits (corrected for a 1.16 arcsec diameter aperture) produces estimates for the mass-to-light ratio of L1 in the B filter of $15 \pm 2h\ M_{\odot}/L_{\odot}$ using $z_s = 0.8$, or $4.3 \pm 0.4h$ $M_{\odot}/L_{\odot}$ using $z_s = 3.14$. These estimates apply to the central region of L1 encircled by the lensed images (i.e. that within a mean radius of 2.8 h^{-1} kpc from L1's centre). Both figures are comparable with M/L estimates from local ellipticals (Gerhard et al., 2001), although the figure for $z_s = 3.14$ is relatively low given $h = 0.7$.

3.7 Conclusions

Together the radio and optical/near-infrared observations presented here show that CLASS B0631+519 can be understood as a gravitational lens system. A simple SIE+shear mass model successfully reproduces features of the radio observations. The mass modelling done so far indicates that the background radio source probably consists of three radio components in a classic double-lobed core-jet arrangement. The eastern lobe is located inside the tangential caustic for the system and is lensed into the arc X and the more compact images Y and Z. The core region is doubly imaged into A1 and B1, while the western lobe is lensed

into images A2 and B2. The magnification provided by the lens and the high resolution of the VLBA has allowed identification of the flat-spectrum AGN core and revealed jet components from both the approaching and the receding jets.

Comparison of the total A1 flux density to the flux density of B1 (both measured by the VLBA at 5 GHz) suggests an A1:B1 magnification ratio of ~9:1, and hence that sub-components in B1 corresponding to those in A1 would span about 2 mas. Resolving the expected sub-components in B1 is thus well within the capabilities of VLBI (Very Long Baseline Interferometry). Identifying the corresponding components in B1 will help constrain the logarithmic slope of the mass distribution.

The wealth of constraints available means that this system will benefit from application of deconvolution/inversion techniques that can utilise the extended images seen in the radio data. Having a good mass model for this system will allow us insights into the mass distributions of ellipticals at redshifts of 0.6, information that is not readily available in any other way. The hints of radio source variability mean that it may be possible to determine a radio time delay and use it to measure the Hubble constant.

This lens system is relatively unusual in having two galaxies at different redshifts present along the line of sight to the source, a situation that also occurs in CLASS B2114+022 (King et al., 1999; Augusto et al., 2001) and in PKS 1830-211 (Pramesh Rao & Subrahmanyan, 1988). In B2114+022 the effect of the nearer lens is much more significant than in B0631+519 (Chae, Mao & Augusto, 2001). In PKS 1830-211 the nearer lens was detected through its HI absorption (Lovell et al., 1996) at a redshift of 0.19 (the main lensing galaxy has a redshift of 0.886). The optical interpretation of 1830-211 is presently unclear, although the main lens is certainly a face-on spiral (Winn et al., 2002; Courbin et al., 2002).

Although L2 appears to have a minor influence on the lensing properties of

B0631+519, the radio emission from the lensed source must pass through L2's ISM. Searches for molecular absorption lines may be worthwhile to provide extra information on the nature of this galaxy. The mass distribution in L2 may well be constrainable with further non-parametric modelling work. Indeed, further work by L.V.E. Koopmans[6] has already shown that a non-parametric approach to the near-infrared data can reveal the perturbing effect of L2 on the smooth lensing properties of the system. The near-infrared source centre deduced from the non-parametric inversion can be projected back to the source plane, and is consistent with the radio source positions within the errors.

The main observational priorities in this system are to establish the source redshift and to check for radio variability using an instrument such as the WSRT. Ground-based optical monitoring for a time delay is probably impractical, because only the radio observations have detected emission from a compact source (images A1 and B1). The extended emission seen in the HST F814W and F160W images is likely to be from the lensed AGN's host galaxy. If the source is variable at radio frequencies, a source redshift together with a detailed mass model would make this lens a good system with which to measure Hubble's constant.

Even if the source is not variable, this system still offers many more constraints on the lens galaxy mass profile than are typically available from gravitational lenses. To refine and extend these constraints, the best option is to request low-frequency (327 MHz or 408 MHz) observations with the VLBA to examine the extended steep-spectrum radio structure revealed in the 1.7 GHz MERLIN map. Yet more constraints are available from the subcomponents in image A1 that were identified in the 5 GHz VLBA data. These should have counterparts in image B1, distributed on roughly $1/8^{th}$ the angular scale of those in A1. To attain the required resolution, global VLBI observations of this lens are needed. VLBI sub-components in B1 would allow us to further constrain mass models for this

[6]Described in detail in York et al. (2005), §4.2

system, improving our knowledge of L1's mass profile.

Bibliography

Augusto P., et al., 2001, MNRAS, 326, 1007

Bender R., Saglia R.P., Ziegler B., Belloni P., Greggio L., Hopp U., Bruzual G., 1998, ApJ, 493, 529

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

Browne I.W.A., Wilkinson P.N., Patnaik A.R., Wrobel J.M., 1998, MNRAS, 293, 257

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

Bruzual A.G., Charlot S., 1993, ApJ, 405, 538

Chae K.-H., Mao S., Augusto P., 2001, MNRAS, 326, 1015

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

Courbin F., Meylan G., Kneib J.P., Lidman C., 2002, ApJ, 575, 95

da Silva Neto D.N., Andrei A.H., Assafin M., Vieira Martins R., 2005, A&A, 429, 739

de Vaucouleurs G., 1948, AnAp, 11, 247

Djorgovski S.G., Davis M., 1987, ApJ, 313, 59

Dressler A., Lynden-Bell D., Burstein D., Davies R.J., Faber S.M., Terlevich R.J., Wegner G., 1987, ApJ, 313, 42

Gerhard O., Kronawitter A., Saglia R.P., Bender R., 2001, AJ, 121, 1936

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

Högbom J.A., 1974, A&AS, 15, 417

Hogg D.W., 2002, astro, arXiv:astro-ph/0210394

King L.J., Browne I.W.A., Muxlow T.W.B., Narasimha D., Patnaik A.R., Porcas R.W., Wilkinson P.N., 1997, MNRAS, 289, 450

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

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

Kinney A.L., Calzetti D., Bohlin R.C., McQuade K., Storchi-Bergmann T., Schmitt H.R., 1996, ApJ, 467, 38

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

Kochanek C.S., 1994, ApJ, 436, 56

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

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

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

Koekemoer A.M., Fruchter A.S., Hook R.N., Hack W., 2002, in Arriba S., ed, The 2002 HST Calibration Workshop: Hubble After the Installation of the

ACS and the NICMOS Cooling System. Space Telescope Science Institute, Baltimore, MD, p. 339

Koopmans L.V.E., 2005, preprint (astro-ph/0501324)

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

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

Krist J., 1993, in Hanisch R.J., Brissenden R.J.V., Barnes J., eds, ASP Conf. Ser. Vol. 52, Astronomical Data Analysis Software and Systems II. Astron. Soc. Pac., San Francisco, p. 536

Kron G.E., Shane C.D., 1974, Ap&SS, 27, 233

Landolt A.U., 1992, AJ, 104, 340

Lovell J.E.J., et al., 1996, ApJ, 472, L5

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

McKean J.P., Koopmans L.V.E., Browne I.W.A., Fassnacht C.D., Lubin L.M., Readhead A.C.S., 2004, MNRAS, 350, 167

Myers S.T., et al., 2003, MNRAS, 341, 1

Norbury, M.A., 2002, Ph.D. thesis (University of Manchester)

Patnaik A.R., Browne I.W.A., Wilkinson P.N., Wrobel J.M., 1992, MNRAS, 254,655

Pavlovsky, C., et al. 2005, "ACS Data Handbook", Version 4.0, (Baltimore: STScI)

Phillips, P.M., 2001, Ph.D. thesis (University of Manchester)

Pramesh Rao A., Subrahmanyan R., 1988, MNRAS, 231, 229

Press W.H., Flannery B.P., Teukolsky S.A., 1986, Numerical Recipes. Cambridge University Press

Rengelink R., Tang Y., de Bruyn A.G., Miley G.K., Bremer M.N., Röttgering H.J.A., Bremer M.A.R., 1997, A&AS, 124, 259

Rusin D., Norbury M., Biggs A.D., Marlow D.R., Jackson N.J., Browne I.W.A., Wilkinson P.N., Myers S.T., 2002, MNRAS, 330, 205

Rusin D., Kochanek C.S., Keeton C.R., 2003, ApJ, 595, 29

Rusin D., et al., 2003, ApJ, 587, 143

Rusin D., Kochanek C.S., 2004, preprint (astro-ph/0412001)

Schlegel D.J., Finkbeiner D.P., Davis M., 1998, ApJ, 500, 525

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

Sérsic J.L., 1968, Atlas de Galaxias Australes, Observatorie Astronómico de Córdoba, Argentina

Shepherd M.C., 1997, in Hunt G., Payne H.E., eds, ASP Conf. Ser. Vol. 125, Astronomical Data Analysis Software and Systems VI. Astron. Soc. Pac., San Francisco, p. 77

Thomasson P., 1986, QJRAS, 27, 413

Tonry J.L., Kochanek C.S., 2000, AJ, 119, 1078

Treu T., Stiavelli M., Casertano S., Møller P., Bertin G., 1999, MNRAS, 308, 1037

Treu T., Koopmans L.V.E., 2004, ApJ, 611, 739

Wallington S., Kochanek C.S., Narayan R., 1996, ApJ, 465, 64

Warren S.J., Dye S., 2003, ApJ, 590, 673

Wilkinson P.N., Browne I.W.A., Patnaik A.R., Wrobel J.M., Sorathia B., 1998, MNRAS, 300, 790

Winn J.N., Kochanek C.S., McLeod B.A., Falco E.E., Impey C.D., Rix H.-W., 2002, ApJ, 575, 103

Wucknitz O., 2004, MNRAS, 349, 1

Chapter 4

Observations of B0218+357 with the ACS

4.1 Purpose of the work

This chapter describes new observations of the gravitational lens B0218+357 made using the Advanced Camera for Surveys (ACS), an optical camera that was installed on the Hubble Space Telescope (HST) in 2002[1]. B0218+357 has been described as an ideal system in which to measure the Hubble constant. The time delay in the system has been measured to 5% precision, the mass model is very well constrained because the lensed radio source is resolved, and the external shear present is minimal due to the relatively isolated location of the lens.

However, the value of H_0 derived from lensing measurements is degenerate with the lens galaxy position in B0218+357 (Lehár et al., 2000). Since the lens is not radio-loud, an optical measurement is the only direct way to measure the position of the centre. The low image separation and the relative faintness of the lens galaxy compared to the lensed images makes such a measurement difficult.

[1]The work presented in this chapter has been published as *The Hubble constant from gravitational lens CLASS B0218+357 using the advanced camera for surveys* in MNRAS volume 357, pages 124-134 (2005)

The ACS has better optical performance than previous HST cameras used to image B0218+357, and far better resolution than is consistently available from ground-based instruments, even with adaptive optics. This work involves using the ACS to measure the galaxy position in B0218+357, and hence obtain an estimate of H_0.

4.2 The gravitational lens CLASS B0218+357

B0218+357 was identified as a likely strong lens candidate in the JVAS survey (Patnaik et al., 1992). Observations by O'Dea et al. (1992) and Patnaik et al. (1993) confirmed that at radio frequencies the system consisted of two lensed images of a flat-spectrum source, and an Einstein ring. The image separation is only 334 mas, which is the lowest separation presently known for an Einstein ring in a strong lens system.

Further radio imaging resolved both the A and B images into core-jet structures (Patnaik, Porcas & Browne, 1995; Kemball, Patnaik & Porcas, 2001; Biggs et al., 2003), as well as revealing more details of the Einstein ring (Biggs et al., 2001). The combined constraints from the core-jet structure and the ring together strongly constrain mass models. A and B lie at different distances from the galaxy, and with the Einstein ring constraints permit determination of both the angular structure of the lensing mass (Wucknitz, 2004; Wucknitz, Biggs & Browne, 2004) and the mass-radius relation for the lens.

The lensed source, a radio loud FR-I source, has a redshift of 0.944 (Cohen, Lawrence & Blandford, 2003). The lensed source is very bright at radio frequencies, the total flux density seen by MERLIN at 5 GHz being over 1.2 Jy. Biggs et al. (1999) and Cohen et al. (2000) looked for correlated variations in the two quasar core images and found time delays of 10.4±0.5 days (95% confidence) and $10.1^{+1.5}_{-1.6}$ days (95% confidence) respectively. Since the source is both variable and

multiply imaged, this lens system can be used to measure Hubble's constant at the present epoch, H_0.

Browne et al. (1993) used the WHT to obtain an optical spectrum of B0218+357 with a high signal-to-noise ratio. Previous work was unable to determine the galaxy type or redshift unambiguously (O'Dea et al., 1992; Stickel & Kuhr, 1993), although O'Dea et al. (1992) did find that the lensed source contributed a majority of the optical light in the system compared to the lens galaxy. Based on their WHT spectrum, Browne et al. (1993) proposed that the lens galaxy is a spiral at a redshift of 0.6847, a conclusion supported by HI 21-cm absorption measurements (Carilli, Rupen & Yanny, 1993) which measured a redshift of 0.68466±0.00004. The spiral nature of the lens was also suggested by the substantial difference in rotation measure between the two lensed images in radio observations (Patnaik et al., 1993), and by HI absorption (Carilli, Rupen & Yanny, 1993). This hypothesis was later supported further by measurements of the optical colour of the lens, determined by ground-based imaging (Grundahl & Hjorth, 1995), and later still by HST/NICMOS imaging (Jackson, Xanthopoulos & Browne, 2000).

In the optical, the small image separation combined with the relative dimness of the lensing galaxy makes it difficult to separate galaxy light from the background source light and so find the lens position. Unfortunately H_0 is degenerate with the lens position for isothermal or nearly-isothermal mass models (Lehár et al., 2000). The accuracy of the measurement of H_0 in B0218+357 is therefore limited by the accuracy with which the lens position can be determined. Also, existing optical and near-infrared images of the lensed source suggest that both image separation and brightness ratio are in conflict with the values seen at radio frequencies. The A image, which is further from the lens galaxy, is fainter than the B image (Grundahl & Hjorth, 1995), despite being a factor $\sim$3 brighter in the radio. Evidence from molecular line studies showed absorption of the radio

emission from the background radio source by species in the lensing galaxy including CO, HCO+, HCN (Wiklind & Combes, 1995), formaldehyde (Menten & Reid, 1996) and water (Combes & Wiklind, 1997), suggesting that a molecular cloud in the lens was intercepting some of the emission from image A. The lensing galaxy appears close to face-on, a conclusion deduced from its symmetrical appearance in optical images. This conclusion is consistent with the small velocity line-width of the absorption lines (e.g. Wiklind & Combes 1995). The new HST/ACS observations described in this chapter confirm that the lens is a nearly face-on spiral.

4.3 B0218+357 and Hubble's constant

CLASS B0218+357 offers numerous advantages over other lenses in the quest to estimate H_0 using gravitational lensing:

1. The observational constraints are numerous and have been extensively studied.

2. The radio source is relatively bright (a few hundred mJy at GHz frequencies) and variable at radio frequencies. As a result, time delay measurements have a small error (Biggs et al., 1999; Cohen et al., 2000).

3. The system is at relatively low redshift. This means that the derived values for H_0 will not depend strongly on the cosmological model used to determine angular diameter distances.

4. The lens is an isolated single galaxy and there are no field galaxies nearby to contribute to the lensing potential (Lehár et al., 2000).

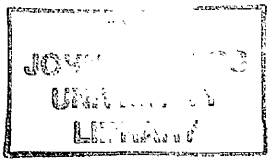
5. The degeneracy between H_0 and the slope of the mass profile has been broken by VLBI observations of the core-jet radio structures (Wucknitz,

Biggs & Browne, 2004).

Although most lenses have at least one of these desirable properties, CLASS B0218+357 is the only one known so far which has all of them. The only disadvantage is the problematic galaxy position, which has resulted in this system being excluded from consideration by many authors (eg. Schechter 2001; Kochanek 2003). However, B0218+357 has been investigated extensively at radio wavelengths with the aim of determining H_0. Biggs et al. (1999) used the time delay and an SIE lens model to deduce a value for Hubble's constant of 69^{+13}_{-19} km s^{-1} Mpc^{-1} (95% confidence). It should be noted, however, that the error bars on the assumed position for the lensing galaxy with respect to the lensed images were over-optimistic and hence their quoted error on H_0 is too small. Cohen et al. (2000) also observed B0218+357 with the VLA, and measured a value for the time delay of $10.1^{+1.5}_{-1.6}$ days. This corresponds to an H_0 value of 71^{+17}_{-21} km s^{-1} Mpc^{-1} (95% confidence), the larger error bars in Cohen et al. measurement being due to their use of a more general model for the source variability, although they used the same model for the lensing effect as Biggs et al.. The error bars do not take into account any systematic error associated with the uncertain galaxy position.

Lehár et al. (2000) used the then existing constraints to model the CLASS B0218+357 system with an SIE. They found that the implied value of H_0 was degenerate with the position of the centre of the lensing galaxy, with a change of about 0.7 km s^{-1} Mpc^{-1} (about 1 per cent) in the value of H_0 for every 1 mas shift in the central galaxy position. Their uncertainty on the position derived from HST infra-red observations using the NICMOS camera is approximately $\pm$30 mas, and suggests a value for H_0 of only 20$\pm$20 km s^{-1} Mpc^{-1}.

Recently, using a modified version of the LensClean algorithm, Wucknitz, Biggs & Browne (2004) have been able to constrain the lens position from radio



data of the Einstein ring. With the time delay from Biggs et al. (1999), they obtain for isothermal models a value of $H_0 = 78 \pm 6$ km s^{-1} Mpc^{-1} (95% confidence). They use VLBI results from other authors (Patnaik, Porcas & Browne, 1995; Kemball, Patnaik & Porcas, 2001) as well as their own data (Biggs et al., 2003) to constrain the radial profile of the mass distribution using the VLBI substructure and obtain $\eta \sim 1.04$, very close to an isothermal profile ($\eta = 1$ for an isothermal model, as defined in chapter 2).

In the following sections, I discuss a measurement of the galaxy position that was made using data from the Advanced Camera for Surveys (ACS). This position was then combined with lens models constrained using VLBI data.

4.4 The data

4.4.1 The Advanced Camera for Surveys

Since the lens system is only 334 mas across, ground-based telescopes using adaptive optics would be required to resolve the lensed images from the galaxy. Adaptive optics operate on the same principle as phase referencing in radio astronomy; the atmosphere introduces phase errors in the wavefront arriving from a target, and the adaptive optics system uses a reference wavefront to determine phase corrections to the target wavefront. Typically the reference wavefront comes from either a bright star close to the target, or from an artificial guide star generated using a laser. When B0218+357 was considered for an observing programme to locate the lens galaxy, accessible instruments having adaptive optics all lacked laser guide stars, and hence a bright, nearby star was required for any possible observations. The closest star to B0218+357 is 16^{th} magnitude and is approximately 55″away. The most suitable instrument for these observations was therefore the Hubble Space Telescope (HST).

The instrument chosen for the observations was the Advanced Camera for Surveys (ACS; Clampin et al. 2000). This instrument has a much greater throughput than the WFPC2. However, the ACS is an off-axis instrument and experiences significant geometric distortion of its field of view. The distortion must be corrected by software, since introducing corrective optics into the ACS would reduce its advantages relative to WFPC2 (i.e. would decrease the sensitivity of the instrument).

Although observations in the blue end of the spectrum would have maximised angular resolution, the likely morphological type of the lensing galaxy (Sa/Sab) meant that asymmetry due to star formation in the spiral arms could have caused problems in the deconvolution process. Hence observations at wavelengths longer than 4000Å in the rest frame of the galaxy were desirable. At the redshift of the lens this dictated the use of Johnson I band, i.e. the F814W filter on the HST. This filter has a broad frequency response, with a bandwidth of 250 nm. At the redshift of the lens galaxy in B0218+357, the central wavelength of the F814W filter is equivalent to 483 nm.

The ACS has two optical/near-IR channels (cameras), the Wide Field Channel (WFC) and the High Resolution Channel (HRC). The HRC's pixel response exhibits a diffuse halo at longer wavelengths due to scattering within the CCD. As a result, roughly 10% of the flux from a point source will be scattered into this halo at 8000Å, making it more difficult to resolve the lensing galaxy from the lensed images. We selected the WFC for use in our observing programme since it does not suffer from this effect (due to a correction made to the manufacturing process after the HRC was fabricated). Unlike the HRC, the WFC moderately under-samples the HST point-spread function (PSF) at 8000Å. To counter this effect, we employed a strategy of dithered pointings. By dithering exposures (that is, shifting successive HST pointings by a fractional number of pixels according

to a pattern), the sampling of the point spread function can be improved and some of the effects of under-sampling removed when the separate exposures are combined (Hook, Pirzkal & Fruchter, 1999).

4.4.2 Dithering and drizzling

The dithering process complicates data reduction, in that the shifts between images can be significant; images cannot be directly combined without special care. Multiple exposures in general can simplify the application of image processing techniques; for example, cosmic ray removal/rejection in a single long exposure is a complicated problem in feature detection and recognition, whereas when multiple exposures are available the problem becomes a much easier one involving only averaging and simple comparisons. Multiple exposures also suffer less from cosmic rays in the first place (fewer pixels of any individual exposure will be affected by cosmic rays), and in the case of the HST multiple exposures are a practical necessity because a single orbit lasts only 90 minutes.

The advantages of dithering are that detector defects (such as dead columns, or permanent hot or cold pixels) do not fall on the same place in the entire multiple exposure set (so that all pixels in the output image contain some information from the source), and, crucially, that it allows undersampled images to be restored some way towards full nyquist sampling. For example, the pixel size of the ACS/WFC is roughly 50 mas. At an observing wavelength of 814 nm, the HST's 2.4 m mirror gives a diffraction limited PSF spanning about 70 mas, so the WFC undersamples the PSF by a factor of about 30%. Dithered images separated by non-integer numbers of pixels can be combined to partially or fully recover the nyquist-sampled resolution of the image, depending on the number, quality and spacing of the dithered images.

Ideally, the resolution could be recovered by simply interlacing the dithered

images. The images would need to be separated by precise fractions of a pixel; if the sampling density was to be doubled over that of the input dithers, the dither steps would be half-pixels. However, pointing errors will tend to prevent the demanded dither steps being exactly reproduced, and in a camera with significant geometric distortion (as in the ACS) interlacing is not possible directly. Alternatively, if enhanced sampling is not a concern, the images can be shifted by the appropriate dither step so that they are all aligned, and then the stack of transformed images is averaged to produce a final output image. However, this method is useless if improved sampling is required, and shifting undersampled images can result in blurring and causes problems with bad pixels and cosmic rays.

There are currently two linear techniques reported in the astronomical literature that are capable of recovering a near nyquist-sampled image from multiple dithered images in the face of pointing errors, bad pixels and detector distortion. The first, due to Lauer, involves combining the dithered images in fourier space (Lauer, 1999). The second method, known as drizzling, was first described by Fruchter & Hook (1997) and involves treating the input and output images as grids, with the pixels of the input grid contributing flux to pixels in the output grid based on their area of overlap. Lauer's method has some advantages compared to drizzling (primarily that it has a solid theoretical basis in Fourier theory and has no arbitrary parameters that need to be fixed by the user, whereas drizzling is a comparatively *ad hoc* approach). However, drizzling can handle distortion correction in a straightforward fashion (an important point in dealing with such a distorted camera as the ACS), and is supported as part of the standard HST data reduction pipeline (as the *drizzle* task). Visual inspection of the example images presented in Lauer (1999) indicates that the output quality from drizzling can be comparable to that available from Lauer's method. The optical/near-infrared

data used in the analysis of both B0218+357 and B0631+519 has been processed using the "drizzle" method of image combination, via the IRAF package of the same name.

The process of drizzling, as implemented by the *drizzle* task, works as follows. Dithered exposures are related to an output image through a (possibly nonlinear) transformation, which can correct distortion, apply a uniform rotation and account for input image translation due to dithering. The output image usually has a finer plate scale than the input images. The input exposures and the output image are treated as grids formed by pixel corners. The transformed input exposures are mapped on to the output frame, and the transformed pixels are shrunk around their new centres by a user-determined scale factor (known in drizzle terminology as the "pixfrac", symbol p). This is done so as to try and minimise blurring introduced by the drizzling process. The best value of p must be determined through trial-and-error; smaller values reduce blurring but they also make it more likely that an output pixel will receive flux contributions from a small number of input pixels, increasing noise. The extreme case in which $p = 0$ corresponds to interlacing. Finally, the flux from each input pixel is divided up amongst all the output pixels that it overlaps, in proportion to the area of overlap. Thus drizzling preserves total flux.

In the work on B0218+357 we used two distinct four-point dither patterns, alternating between them over the course of the observations. The dithered data were then processed with *drizzle*. The dither patterns used are shown in figure 4.1 (see also Mutchler & Cox 2001 for more information on HST dither patterns).

The WFC has a field of view of $202'' \times 202''$, and a plate scale of approximately 50 mas pixel^{-1} (dependent on the geometric distortion). The distortion introduced by the off-axis location of the instrument is corrected for by the ACS data reduction packages (e.g. Blakeslee et al. 2003). The WFC detector consists

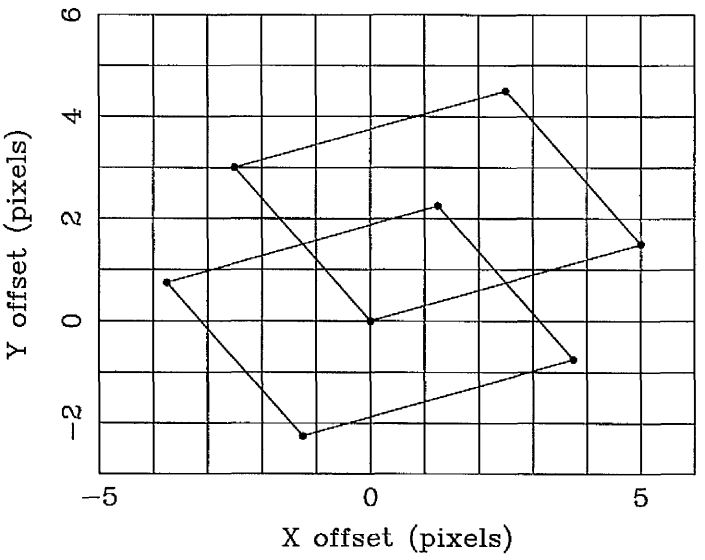


Figure 4.1: 4+4 dither pattern used for the observations of 0218+357. The pattern provides steps to the level of $\frac{1}{4}$-pixel.

of two mosaiced CCDs separated by an inter-chip gap of 50 pixels. Both CCDs have an active area of 2048×4096 pixels. Since we used a gain of unity, saturation in the images occurs near the 16-bit analogue to digital conversion limit of 65,000 e^- $pixel^{-1}$ rather than at the WFC's full well point of 85,000 e^- $pixel^{-1}$. We determined the exposure time required on B0218+357 through simulation.

The full programme of B0218+357 observations was carried out over the period from August 2002 to March 2003. Details of the observing dates and exposure times are shown in table 4.1. The total available telescope time was split into 7 visits on B0218+357, 6 of which provided 2 hours integration time on the science target. The remaining visit (16) was designed to permit the programme to be salvaged in the unlikely case that the observing pattern chosen for visits 10-15 proved to be both inappropriate and uncorrectable. This visit suffered from increased observing overheads relative to the other visits and provided an integration time of 1 hour, 22 minutes on B0218+357 as a result. In order to deconvolve the images we required an ACS/WFC PSF, so two short visits (1 and 2 in table 4.1) were dedicated to observing two Landolt standard stars (Landolt

1992). Following McLure et al. (1999), observations were taken with several different exposure times to allow the construction of a composite PSF that would have good signal-to-noise in both core and wings whilst avoiding saturation of the core. Standards were selected to be faint enough not to saturate the WFC chip on short integration times, and to have the same $V-I$ colour, to within 0.2 magnitudes, as the lensed images in the B0218+357 system. The exposure times on the standard stars ranged from 0.5 seconds to 100 seconds each, the longest exposures each being split into two 50 second exposures to simplify cosmic ray rejection.

4.5 Reduction of the ACS data

4.5.1 Basic calibration

The uncalibrated data produced by automatic processing of raw HST telemetry files by the OPUS pipeline at the Space Telescope Science Institute (STScI) were retrieved along with flat fields, superdarks and other calibration files.

The data were processed through the ACS calibration pipeline, *CALACS* (Pavlovsky et al., 2002), which runs under NOAO's IRAF software. The *CALACS* pipeline de-biased, dark-subtracted and flat-fielded the data, producing a series of calibrated exposures. The pipeline also combined the CR-SPLIT exposures in visits 1, 2 and 16 to eliminate cosmic rays. The calibrated exposures were in general of acceptable quality for use in the next stage of reduction, except for visit 15 in which there was some contamination of the images by stray light, probably from a WFPC2 calibration lamp (R. Gilliland and M. Sirianni, private communication). It is possible that this defect can be corrected in the future using the techniques employed by Williams et al. (1996) to remove stray light from some

Visit no.	Target	Observation date	Exposure time	Dither pattern	File name root
10	CLASS B0218+357	2003 February 28	20×360 sec	4+4	j8d410
11	CLASS B0218+357	2003 March 01	20×360 sec	4+4	j8d411
12	CLASS B0218+357	2003 January 17-18	20×360 sec	4+4	j8d412
13	CLASS B0218+357	2003 March 06	20×360 sec	4+4	j8d413
15	CLASS B0218+357	2003 March 11	20×360 sec	4+4	j8d415
14	CLASS B0218+357	2002 October 26-27	20×360 sec	4+4	j8d414
16	CLASS B0218+357	2002 September 11	12×360 sec	4+4	j8d416
			8×75 sec	4C	
1	92 245	2002 October 17-18	8×0.5 sec	4+4	j8d401
			8×8 sec	4+4	
			1×360 sec	-	
2	PG0231+051B	2002 August 25	8×0.5 sec	4+4	j8d402
			8×8 sec	4+4	
			8×50 sec	4C	

Table 4.1: Log of HST observations. All observations were taken with the ACS using the F814W filter, corresponding to I band. A 4+4 dither pattern refers to an eight-point dither consisting of two nested parallelograms, whereas 4C refers to a four-point dither parallelogram with the exposure at each point split into two for explicit cosmic-ray rejection.

HDF exposures, but we have not attempted to deconvolve the contaminated visit.

The calibrated exposures were fed to the next stage of reduction, based around the *dither* package (Fruchter & Hook, 2002), and the STSDAS packages *pydrizzle* (Hack, 2002) and *multidrizzle* (Koekemoer et al., 2002). These tools clean cosmic rays, remove the ACS geometric distortion and "drizzle" the data on to a common output image (Fruchter & Hook, 2002). The drizzle method projects the input images on to a finer grid of output pixels. Flux from each input pixel is distributed to output pixels according to the degree of overlap between the input pixel and each output pixel. To successfully combine dithered images into a single output image, knowledge of the pointing offsets between exposures is needed. The expected offsets are determined by the dither pattern used, but the true offsets might vary from those expected due to thermal effects (Mack et al., 2003) within single visits. To determine the true pointing offsets between dithered exposures, we cross-correlated the images. Since the WFC at I-band moderately undersamples the HST's PSF, and since most of the features detected in the images were extended rather than stellar, we opted for pixel-by-pixel cross-correlation rather than comparisons of positions of stars between different pointings, to maximise our use of the available information.

4.5.2 Drizzling

Images in each visit were drizzled on to a common distortion-corrected frame and then pairs of these images were cross-correlated. The two-dimensional cross-correlations show a Gaussian near their centres. The shift between pairs of images is measured by fitting a Gaussian function to the peak in the cross-correlation, and the estimated random error in the shift is derived from the position error given by the Gaussian fit. For our data, the random errors in the measured shifts ranged from 0.8 to 2.5 mas. The RMS scatter between corresponding pointings

within a visit was typically less than 10 mas, or 20% of a single WFC pixel. We fed these shifts to the *multidrizzle* script, which carried out the drizzling of visits to common, undistorted output frames. As part of the drizzling process, we opted to decrease the output pixel size from 50 mas square (the natural size of the undistorted output pixels) to 25 mas square. To avoid blurring and "holes" in the output image, the input pixels were shrunk to 70% of their nominal size before being drizzled on to the output frames (Fruchter & Hook, 2002). We used a gaussian drizzling kernel to slightly improve resolution and reduce blurring. At the end of this process each visit, except for visit 15, provided us with a output image with improved sampling compared to the individual input exposures. The deconvolution of these images is described in Section 4.6.

Since deconvolution depends greatly on the accuracy of the PSF model, we have produced a number of different PSFs. Unfortunately the Landolt standard star (Landolt 92 245) observed in visit 1 was resolved into a 0.5″ double by the ACS/WFC, so we have concentrated on extracting a PSF from visit 2 (Landolt PG0231+051B). The calibrated exposures were combined using *multidrizzle*. Saturated pixels were masked. The resulting PSF suffered from serious artifacts consisting of extended wings approximately 80 mas up and down the chip from the central peak, possibly due to imperfect removal and combination of pixels which were affected by bleeding of saturated columns. We believe these extended wings to be artifacts since they rotate with the telescope rather than the sky, and are not present in stars in other visits.

In lieu of ideal standard star PSFs, we have created per-visit PSFs by averaging field stars together. These field star PSFs do not suffer from the artifacts present in the visit 2 PSF, although they still produce residuals after subtraction of approximately 10%. Such residuals interfere with attempts to deconvolve the lens, and so we have allowed for this error when performing the data analysis

described in Section 4.6. The residual error increases linearly with the counts, rather than with their square root, since we assume it has a gaussian distribution.

4.5.3 Noise in the ACS images

To deconvolve the images an excellent noise model is essential. The noise model should produce an accurate prediction of the uncertainty in the counts reported by a pixel, given some details about the detector. All the single-pixel noise models described in this chapter can be written in the form

$$\sigma^2 = \alpha * \rho + \beta, \tag{4.1}$$

where α and β are constants that depend on the details of the noise model, ρ is the count rate (usually in electrons per second) recorded in the image pixel, and σ is the corresponding estimated noise level (which is in the same units as ρ). The ACS noise can be modelled using the standard CCD noise model (Howell, 1992) as a starting point. This model takes into account several different sources of noise:

1. Noise generated in the CCD electronics during read-out ("read noise")
2. Noise due to counts picked up from the sky ("sky noise")
3. Noise arising from the non-zero temperature of the CCD ("dark current")
4. Noise added by the finite accuracy of the digitisation process

The standard noise model is represented by the so-called CCD equation (Merline & Howell, 1995). The noise in a single pixel (σ_e) that intercepts C photons from the source and S photons from the sky, in a CCD having a read noise of R electrons and a dark current of D, is given by

$$\sigma_e^2 = \left(R^2 + S + D + G^2\sigma_f^2\right) + C, \tag{4.2}$$

where G is the gain factor used to convert electrons to ADUs, and σ_f represents the digitisation noise in the ADC, and has a constant value of 0.289. The read noise and digitisation noise are shot noise sources rather than counting (poisson) errors, and so their contributions to the variance σ_e are squared. The remaining sources of noise contribute as poisson errors. When the number of counts C is large enough, the CCD equation can be simplified to

$$\sigma_e = \sqrt{C}, \tag{4.3}$$

which is the obvious result that in the limit of high signal to noise ratio, CCD noise obeys simple Poisson statistics. There are several other complications that need to be addressed before the noise model correctly represents the noise properties seen in the reduced, drizzled ACS data. First, the ACS data were drizzled onto a finer pixel grid than the original CCD data. The noise models must therefore be re-expressed in terms of the counts in the (smaller) output pixels rather than the original input pixels. Second, multiple input images were drizzled on to a single output grid, thus reducing the noise through averaging. Let q represent the size of pixels in the drizzled output image in terms of the size of the input pixels (so that drizzling from an input frame of 50 mas pixels to an output frame of 25 mas pixels gives $q = 0.5$), and let N be the number of input exposures drizzled on to the output image. Then the final noise equation for N averaged, dithered exposures, giving the single-pixel noise in electrons per second, σ_{eps}, is

$$\sigma_{eps}^2 = \frac{q^4}{N\tau^2}\left[\left(R^2 + S + D + G^2\sigma_f^2\right) + \frac{\rho\tau}{q^2}\right]. \tag{4.4}$$

This noise estimation equation was tested using simulations of the drizzling

process. For each simulation run, a 20x20 grid of empty pixels was prepared to represent the input frame. The output frame consisted of a 20x20 grid whose pixels were considered to have a quarter of the area of the input pixels (i.e. a pixscale of 0.5). The initially empty input pixel grid was processed in the following way:

1. To each pixel was added a gaussian noise distribution having a mean of zero and a standard deviation equal to the read noise of the ACS/WFC (about 5 electrons per pixel)

2. Poisson noise was added to represent sky counts. A mean sky level of 40 counts per pixel was adopted.

3. Poisson noise was added to 25 pixels to represent stars. Star noise was scaled to a mean of 36000 counts.

4. The mean sky level was subtracted from all pixels to simulate the sky subtraction undertaken before drizzling.

5. All pixels were converted from units of counts to counts per second by dividing by a nominal exposure time of 360 seconds.

The input pixels were shrunk by 30% before drizzling, corresponding to a pixfrac of 0.7. These settings (pixscale 0.5, pixfrac 0.7) correspond to the settings used to drizzle the B0218+357 ACS data. In every simulation run, 20 such input frames were prepared and drizzled on to the output frame. The simulations ignored bias frame removal, flat-fielding, bad pixels and cosmic rays. No digitisation noise or dark current was modelled, since the noise is dominated by poisson counting errors for our data. The digitisation noise is significant only for very large gain figures (the ACS images of B0218+357 have unity gain). The dark current in the ACS/WFC is less than 0.01 electrons per second, and for an

exposure time of 360 seconds would contribute a noise level of only 1 electron per pixel.

At the end of each simulation, the mean flux and RMS were computed for the input and output frames. Pixels containing mainly background were separated from those containing mainly starlight, and the mean and RMS were computed for both sets of pixels. The resulting four points of comparison between the noise model and the simulations are summarised in figure 4.5.3 for a series of 1000 simulations. These plots show that the analytic noise model performs as expected under ideal drizzling conditions.

The predictions of the model for the background noise (in the absence of sources) were checked against blank sky in the ACS observations of B0218+357. The measured RMS agrees with the model to within 4% for all processed visits.

4.6 Analysis of the position data and results

4.6.1 General remarks

Figure 4.3 shows the ACS image of the CLASS B0218+357 system from the combined dataset, consisting of all science visits on B0218+357 excluding visit 15. To produce this combined image, separate visits were related through restricted linear transformations (rotation and translation) based on the positions of unsaturated stars common to all images, and then co-added. In locating the lensing galaxy, however, we did not use this combined image, preferring to work with the separate visits. The two compact images (A and B) can be distinguished as can the lensing galaxy which lies close to B. The spiral arms of the galaxy are clearly seen confirming the earlier deductions that the lensing galaxy is a spiral (Wiklind & Combes, 1995; Carilli, Rupen & Yanny, 1993). The spiral arms

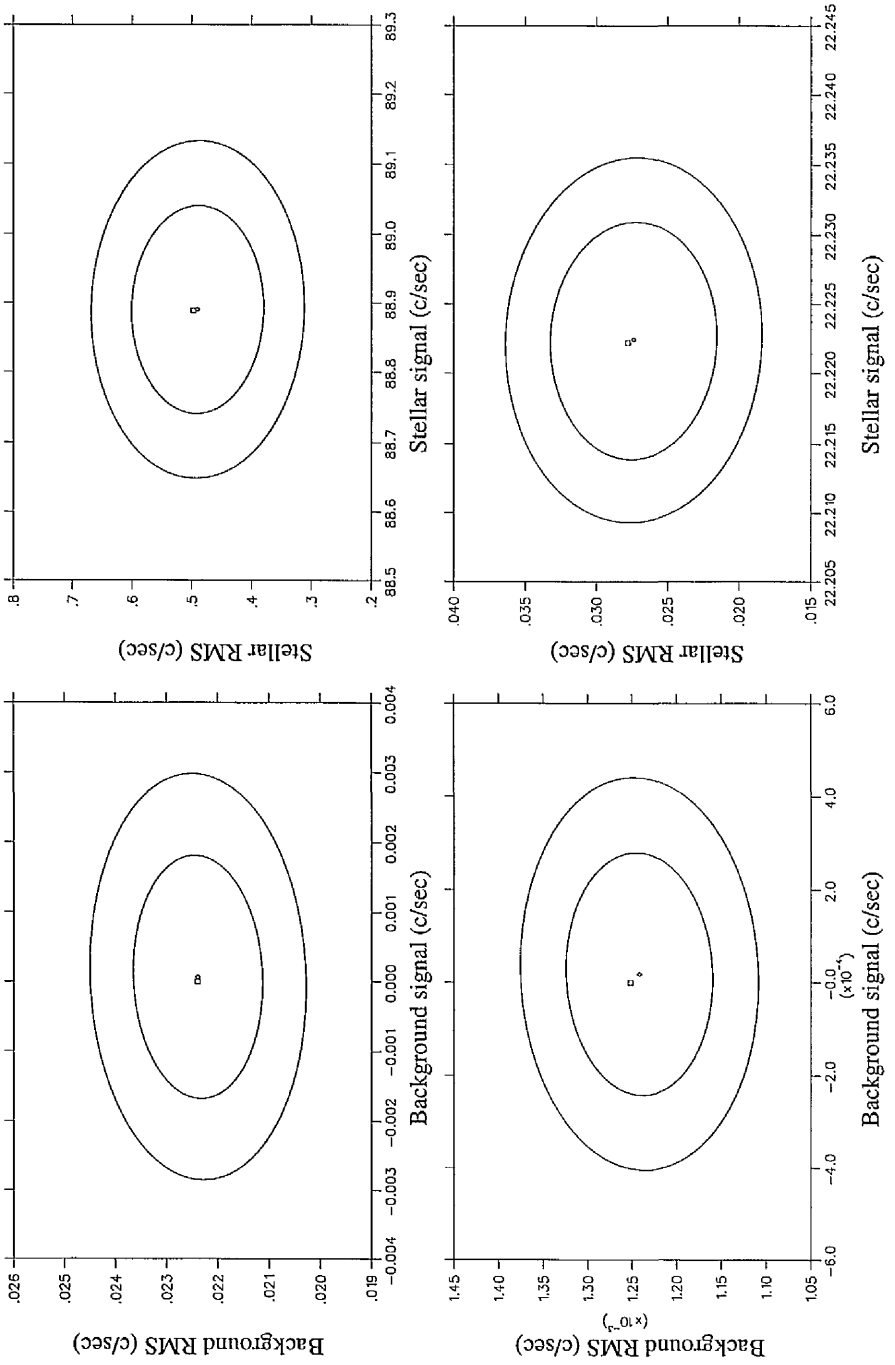


Figure 4.2: Analytic predictions for the mean signal and RMS noise in drizzled data compared to 1000 simulated drizzled images. The inner ellipse in each frame bounds 68% of the simulation results, the outer ellipse, 95%. Diamonds mark the centre of the ellipses (i.e. mean results of the simulations). The squares mark the predictions calculated from equation 4.4. The model and the averaged simulation results agree to within 0.11 σ.

appear to be smooth and regular and there is no sign of significant clumping associated with large-scale star formation. The galaxy appears almost exactly face-on. We deduce this by assuming a galaxy position close to B and comparing counts between pixels at 90 degree angles from each other about the assumed centre. Examination of the residuals reveals no sign of ellipticity.

The core of the lensing galaxy is strongly blended with B (figure 4.3) and is relatively weak. The peak surface brightness of image B apparently exceeds that of the galaxy by a factor of about 30-50. Thus the determination of an accurate position for the galaxy is a challenging task. Before discussing the process in more detail below, we outline the various steps we go through to obtain a galaxy position. They are:

1. For each visit we measure the positions of the A and B images.

2. We subtract PSFs from these positions.

3. Using PSF-subtracted data we look for the galaxy position about which the residuals left after PSF subtraction appear most symmetric. We do not subtract a galaxy model from the images. This approach finds the centre of the most symmetric galaxy consistent with the data.

4. We compute the mean and variance of the galaxy positions found from the individual visits.

4.6.2 Analysis procedure

Our reduced ACS/WFC images are over 8450x8500 pixels in size, and cover 202″x202″on the sky. We cut out a region of 128x128 pixels (3.2″x3.2″) centred on the lens and analyse this to make fitting computationally practical and to isolate the lens from other objects on the sky.

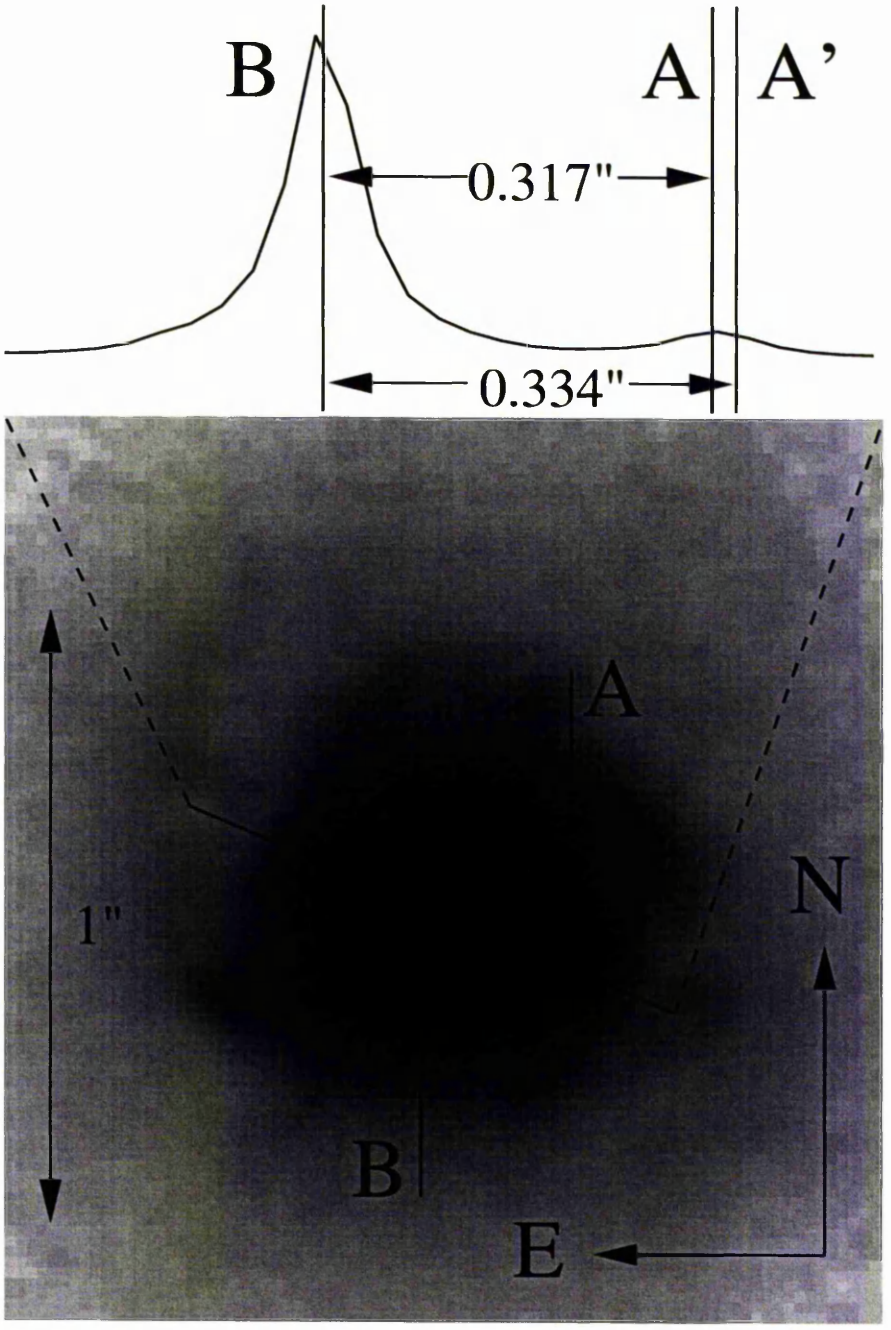


Figure 4.3: Combined ACS image of B0218+357. The lensed images are both visible; the brighter image, B, is close to the centre of the lensing galaxy. The spiral arms of the galaxy are clearly visible. The plot above the image shows a one dimensional slice passing through images A and B. The best-fit positions of A and B on this slice are marked, along with A', the position of A expected from the radio image separation (334 mas). The separation between A and B in the optical image is 317±4 mas (2σ).

Since the images, particularly image B, have much higher surface brightnesses than the galaxy, their positions can be located relatively accurately by subtracting parametric or empirical PSFs from the data and minimising the residuals. Fits were carried out on circular regions centred on the brightest pixel of each compact image. The regions chosen were 11 pixels in diameter. To avoid any bias arising from the choice of PSF, we used both parametric models (Airy and Gaussian functions) and the field star PSFs in the fits. The field star PSFs were consistently better fits to both A and B than the parametric models, Gaussians being insufficiently peaked and Airy functions having diffraction rings that were too prominent. A linear sloping background was modelled along with the PSF in order to take account of the flux due to the galaxy. Typically the various methods agreed on positions to within a tenth of a pixel (2.5 mas).

The separation of A and B determined by optical PSF fitting is consistently less than the radio separation of 334 mas. We find that the mean image separation in the optical is 317±2 mas ($1\,\sigma$) when the field star PSFs are used, and 315±4 mas when Gaussian PSFs are used. The corresponding result for Airy function PSFs is 311±10 mas. These values are mean separations taken over the six processed visits on B0218+357. We have checked the plate scale of drizzled images against stars listed in the US Naval Observatory's B1.0 catalogue, and find that the nominal drizzled plate scale of 25 mas pixel^{-1} is consistent with that catalogue.

An anomalous optical image separation has been suggested before in the case of B0218+357, starting with ground-based optical imaging by Grundahl & Hjorth (1995) and again by Hjorth (1997). Jackson, Xanthopoulos & Browne (2000) used NICMOS imaging to find an image separation of 318±5 mas, in agreement with our result from field star PSFs. We hypothesise that this low separation may be a result of the high, and possibly spatially variable, extinction in the region of

A. We suggest that some of the image A optical emission arises from the host galaxy rather than from the AGN which dominates the B image emission. Thus the centroid of A may not be coincident with the AGN image. Image A may be obscured completely and the emission seen could be due to a large region of star formation associated with an obscuring giant molecular cloud. This view is supported by the recent work of Henkel (2005), who deduce the existence of a large, warm (35K) molecular cloud from molecular line studies using the Effehlsberg 100 metre radio telescope. In view of this probability, and the fact that B is consistently much brighter than A in the optical, we measure galaxy positions as offsets from our measured position for B. Thus, although A's position may be distorted in the optical it does not directly influence our measurements of H_0.

Having determined positions and fluxes for A and B, we created an image containing two PSFs as a model for the flux from the lensed images alone (the "model image"). A model image was made separately for each visit and subtracted from each observed image, leaving a residual image which contained only the galaxy plus subtraction errors. Figure 4.4 shows a typical image with A and B subtracted. The model allowed us to keep track of how much PSF flux was removed from each pixel in producing the residual images.

Initially, we attempted to fit parametric models to the data. However, the low surface brightness of the lens galaxy and the artefacts produced by subtracting the bright quasar images meant that we could not find a stable parametric model. We opted to use the criterion of maximum symmetry in the residuals as a fit statistic, rather than attempting to fit a parametric model, such as an exponential disk profile, to the light distribution of the galaxy. The symmetry fit statistic is expressed as

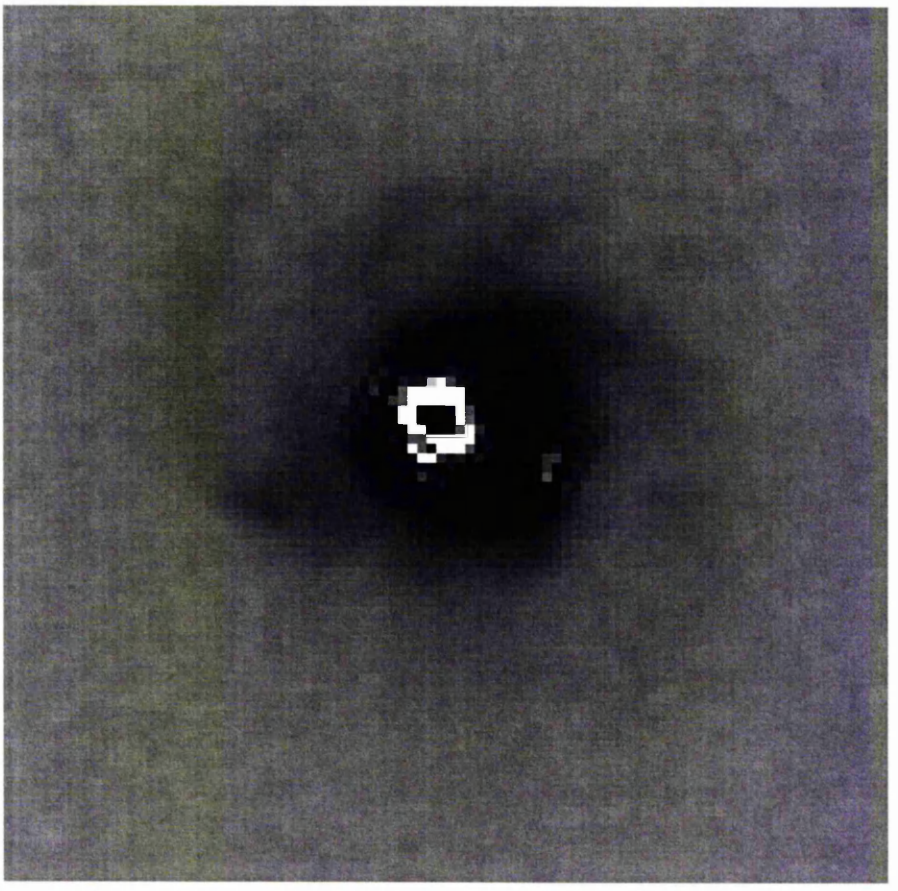


Figure 4.4: A visit image with A and B subtracted after fitting fluxes and positions. The residuals near the centre of each subtracted image show maxima of approximately 20% of the unsubtracted light.

$$\chi^2 = \sum_P \frac{|s(\mathbf{r}) - s(\mathbf{r}')|^2}{\sigma(\mathbf{r})^2 + \sigma(\mathbf{r}')^2}, \tag{4.5}$$

in which $s(\mathbf{r})$ is the count-rate at image pixel position $\mathbf{r} = (x, y)$ in counts per second, $\sigma(\mathbf{r})$ is the estimated noise at $\mathbf{r}$, P is the set of pixels included in the fit (see below) and $\mathbf{r}'$ is the reflection of $\mathbf{r}$ around the galaxy position $\mathbf{g}$, given by simple geometric considerations as

$$\mathbf{r}' = 2\mathbf{g} - \mathbf{r}. \tag{4.6}$$

The random error for each pixel in the image, $\sigma(\mathbf{r})$, can be estimated from the ACS noise model (equation 4.4), but to suppress the effect of the artefacts left behind by PSF subtraction, we added a term to model the assumed random error in the PSF. The estimate of the noise is then given by

$$\sigma(\mathbf{r})^2 = {\sigma_{eps}}^2 + \mu^2 s_p(\mathbf{r})^2, \tag{4.7}$$

where μ is the assumed fractional error in the PSF and $s_p(\mathbf{r})$ is the count-rate from A and B alone (stored in the model image). The set of pixels (P) included in the calculation of this χ^2 figure can bias the fit if it is ill-chosen. When $\mathbf{r}'$ falls outside the boundaries of the image, the pixel $\mathbf{r}$ is considered to contribute nothing to the χ^2 statistic and the pixel is not included in the set P. Such pixels therefore do not contribute to the number of constraints available and as a result do not increase the number of degrees of freedom in the fit. Alternative treatments can introduce bias; for instance, if these pixels are assigned large χ^2 values the fitting program is biased towards placing the galaxy in the geometrical centre of the image. If the same pixels are considered to contribute zero towards the χ^2 statistic but are still counted as part of the set P, they increase the number of degrees of freedom in the fit and bias the fit to positions away from the image's geometrical centre. To avoid these possibilities we do not count degrees of freedom

from pixels whose reflection about the galaxy centre ends up outside the image boundaries.

The symmetry criterion is non-parametric and has the advantage of minimizing the assumptions that are imposed on the data; the use of a particular distribution as a function of radius in any case usually contains an implicit assumption of symmetry. Using the symmetry criterion on its own is in principle robust whether or not the galaxy has a central bulge, and should also be unaffected if the galaxy contains a bar. The symmetry criterion will also hold for galaxies having moderate inclinations to the line of sight. For a circularly symmetric galaxy, the main effect of a small deviation away from a face-on orientation will be to render the observed image slightly elliptical. The basic symmetry criterion is that points and their reflections about the true centre of the galaxy's image should have the same flux (to within the measurement errors). Therefore whether the galaxy's image has circular or slightly elliptical isophotes is unimportant because in both cases the same isophote passes through both point and reflection. This argument breaks down for spirals with significant inclinations as absorption is likely to become important and destroy any symmetry present in the image.

The main violation of the symmetry criterion that remains after suppressing the PSF errors is produced by the spiral arms of the lens. The centre of maximum symmetry could be displaced by spiral arms if they are not themselves symmetric about the galaxy centre. To analyse the possible effect on H_0 of spiral arms we have made two sets of fits. In the first set we used all the data and did not apply any masking. In the second set we masked off the most obvious spiral arms by using an annular mask centred on image B with an inner radius of $0.375''$ and an outer radius of $0.875''$. Regions within the inner radius or beyond the outer radius of the mask were left free to contribute to the symmetry fit. We implemented

the symmetry fits so that masked pixels did not contribute to either the number of degrees of freedom or to the χ^2 value.

The PSF error (μ in equation 4.7) can cause a systematic change in galaxy position when varied between 0.05 and 0.15 (5 to 15%). An increase of μ from 0.05 to 0.15 can increase H_0 by up to 10 $\mathrm{km\,s^{-1}\,Mpc^{-1}}$. For values above 0.15 the systematic change in galaxy position is small compared with the random error. We estimate the PSF error separately for each visit, by taking the range of the highest and lowest residuals and dividing that range by the peak count-rate of image B. We find values between 0.07 (visit 11) and 0.19 (visit 10) for μ. For the other visits, the estimated value for μ is found to be 0.12. In the remainder of the paper we do not allow μ to vary freely but fix it to these estimated values.

4.6.3 Extraction of the galaxy position

In applying the symmetry criterion to 0218 we calculated the symmetry χ^2 statistic (i.e. that of equation 4.5) for a grid of galaxy positions extending 20 mas east to 100 mas west of B, and from 80 mas south to 50 mas north of B. The spacing between adjacent grid points is 5 mas. We present these grids in figures 4.5 and 4.6 for visits 10, 11, 12, 13, 14 and 16. The grids are shown both with and without masking of spiral arms.

In figure 4.7 we show the per-pixel χ^2 contributions between the two cases of no masking and masked spiral arms for the optimum galaxy position, together with the contributions when no masking is used and zero PSF error is applied. With zero PSF error it is clear that the residuals from the subtraction of A and B dominate the χ^2 measure. The effect of a non-zero PSF error is to suppress the A/B subtraction residuals and cause the spiral arms to dominate the χ^2 measure, unless they are masked. Because it is unclear which position (masked or unmasked) best represents the mass centre of the lensing galaxy, we report

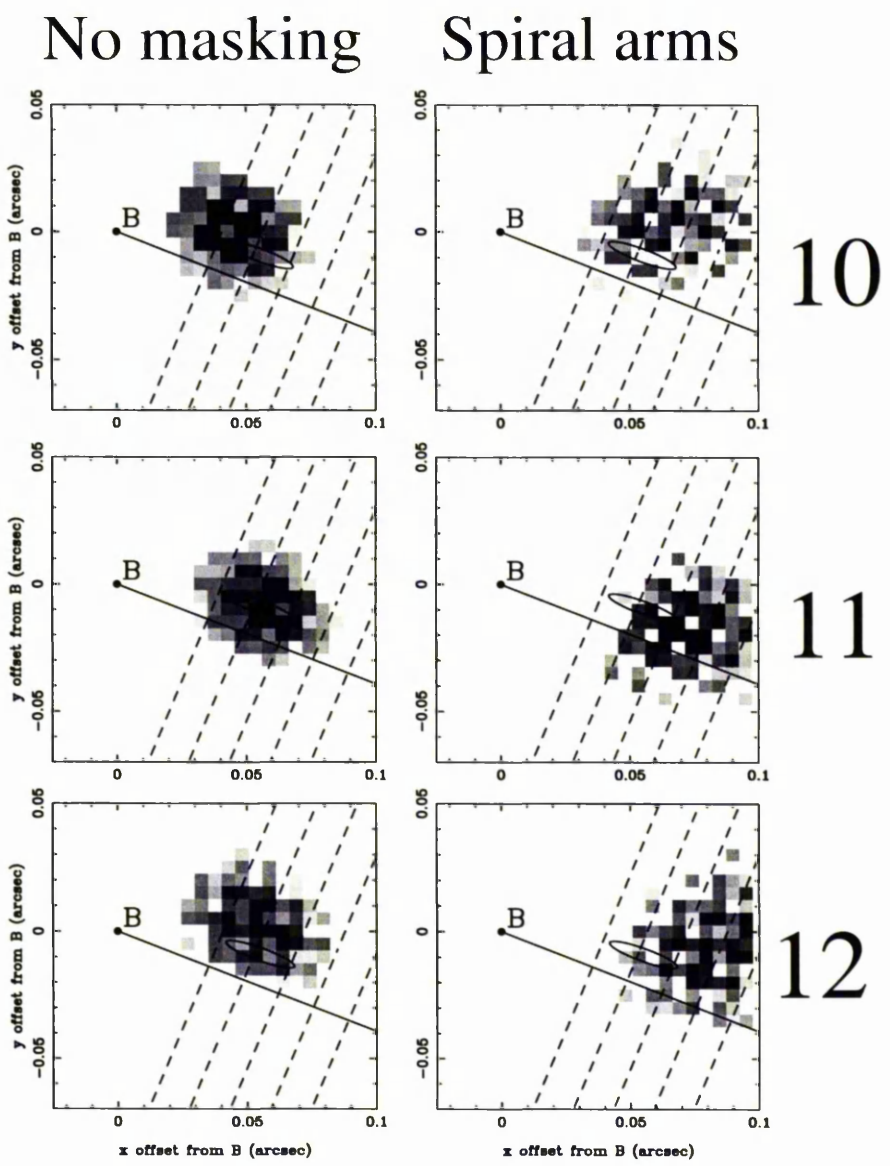


Figure 4.5: χ^2 grids for the galaxy position. The visit number is shown to the right of each pair of grids. In this image, visits 10-12 are plotted. The right-hand plot for each visit shows the effect of masking out the spiral arms, whilst the left-hand plot shows the χ^2 grid when no masking is applied. The position of B is marked, as is a line pointing towards the (radio) A component. The ellipse represents the position of the galaxy centre found by Wucknitz, Biggs & Browne (2004) using LensClean modelling of the Einstein ring, and the dotted lines represent H_0 of (90,80,70,60,50)km s^{-1} Mpc^{-1} from left to right, assuming an isothermal model. The axes are RA/Dec offsets from the position of B, expressed in arc-seconds. The RA offset is given with west as positive.

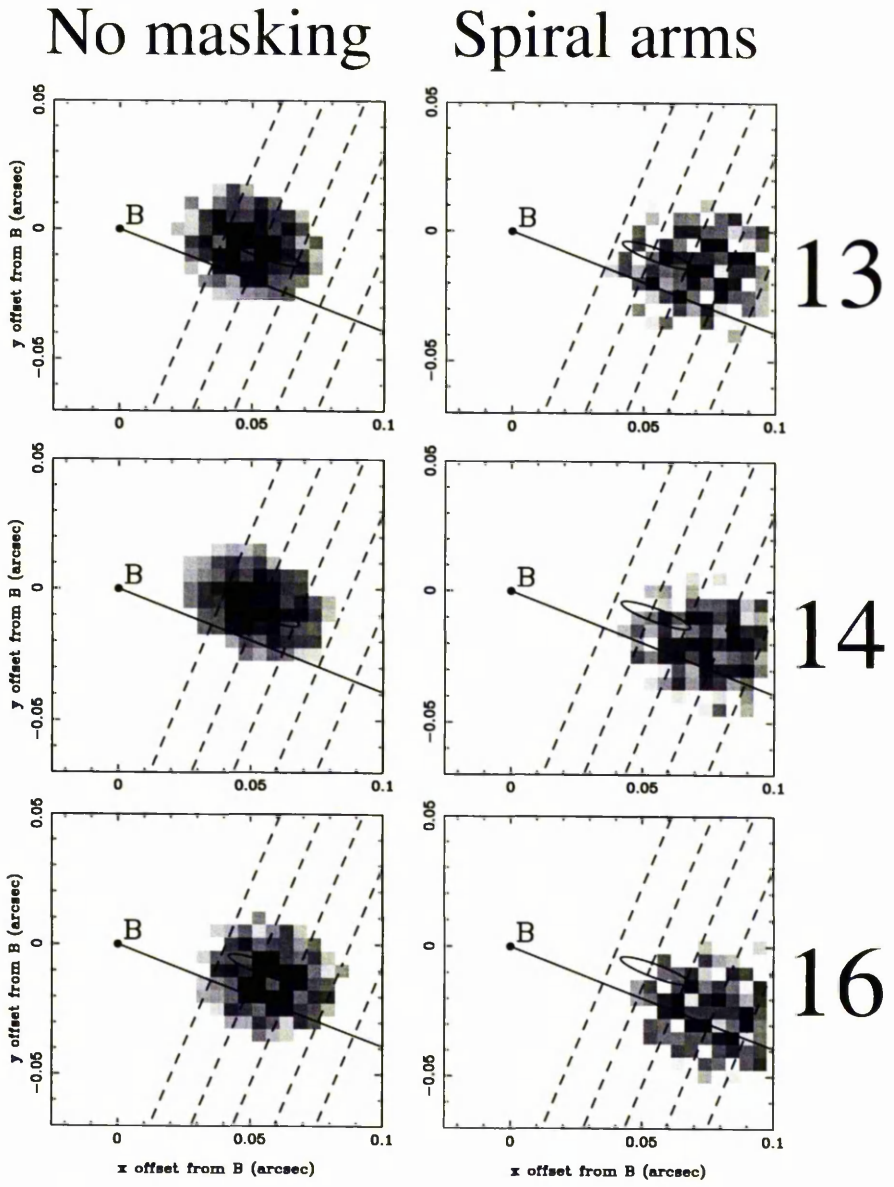


Figure 4.6: χ^2 grids for the galaxy position in visits 14-16. The conventions remain as in figure 4.5.

Visit	Centre (No masking)		Centre (Spiral arms masked)	
	$\Delta\alpha$	$\Delta\delta$	$\Delta\alpha$	$\Delta\delta$
10	+50	+6	+70	+12
11	+60	−4	+69	−18
12	+59	+9	+84	+8
13	+54	−2	+72	−5
14	+59	+0	+76	−16
16	+61	−6	+79	−14
Mean	+57±2	+1±2	+75±2	−6±5

Table 4.2: Derived optical centre of the galaxy, expressed as offsets in mas from the measured optical position of B. RA offsets are given with west as positive.

both masked and unmasked galaxy positions (and hence estimates for H_0) on equal terms.

Deriving errors on position from the individual visits is difficult because the symmetry χ^2 increases very rapidly away from the minimum. An error measure derived from the shape of the minimum for a single visit implies a spuriously high accuracy for the galaxy position. It is likely that the number of degrees of freedom in the fit is over-estimated and that many pixels do not contribute any useful information to the fit statistic, since the drizzling process introduces correlations between neighbouring drizzled pixels. However, the scatter between positions derived from different visits is large. We therefore estimate errors on the galaxy position by taking ellipses that enclose 68% and 95% of the measurements from all visits to define our $1\,\sigma$ and $2\,\sigma$ confidence levels. Figure 4.8 shows the 95% confidence ellipses on the galaxy position for both sets of fits, as well as the position derived from LENSCLEAN applied to VLA data by Wucknitz, Biggs & Browne (2004).

We note that the symmetry fit would produce a χ^2 of unity for an image containing only background noise. The positions obtained from the symmetry fits were combined with the extra constraints available from the VLBI substructure

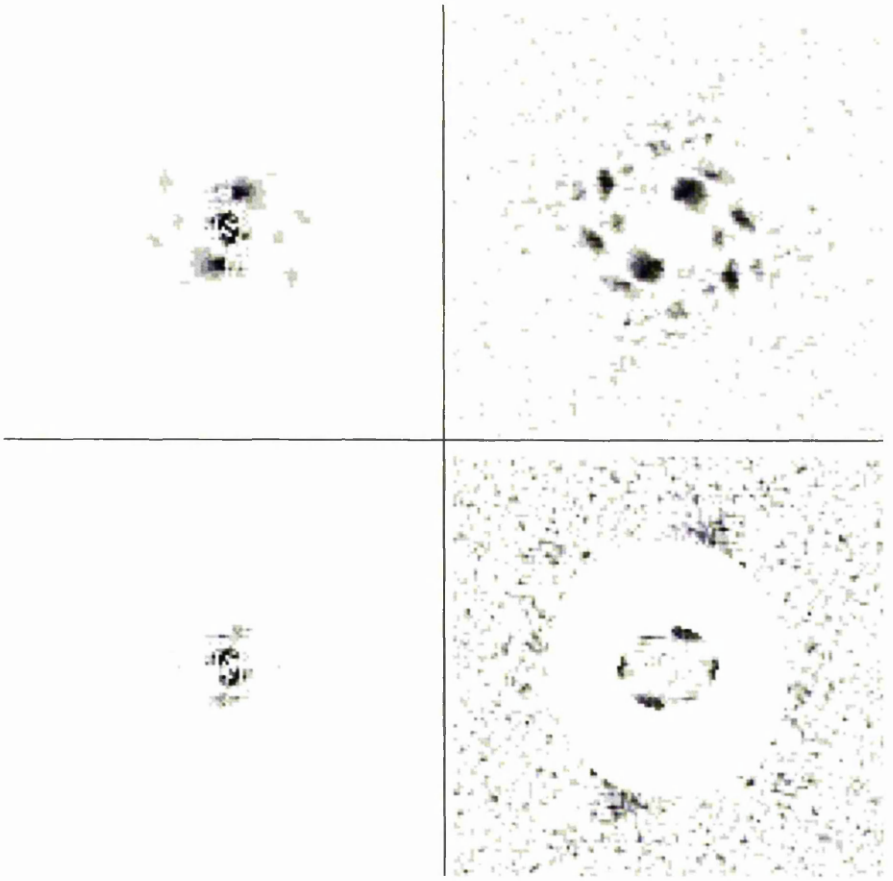


Figure 4.7: These images show the contribution of each pixel to the symmetry χ^2 for visit 11, using the best-fit optical galaxy positions. The top-left image shows the per-pixel χ^2 contributions when no masking is used and no PSF error is used. The image on the top-right shows the contributions when no masking is applied and a 7% PSF error is used. The bottom-left image shows the contributions when spiral arms are masked, with no PSF error. The bottom-right image shows the χ^2 contributions when the prominent spiral arms are masked and a 7% PSF error is used. All images are 128 pixels (3.2″) in both width and height.

described in Patnaik, Porcas & Browne (1995) and Kemball, Patnaik & Porcas (2001), which were used to constrain mass models by Wucknitz (2004). The optical galaxy position was combined with the models of Wucknitz (2004) by adding χ^2 values for the galaxy position to the χ^2 values from the lens models. However, the χ^2 values taken from the symmetry fitting grid have too many degrees of freedom, and so we assume that the optical position minimum is parabolic and form a new χ^2 statistic based on our 68% and 95% confidence ellipses. We define the new χ^2 statistic to have a value of 2.31 on our 68% confidence ellipse, and a value of 5.99 on the 95% ellipse, and sum this statistic with that from the lens modelling. We emphasize that the scatter between visits dominates the random error budget for our measurement of H_0.

Combining the VLBI and optical constraints shows that the best-fit galaxy position and the optical galaxy position are not coincident, as shown in table 4.3. The galaxy position shifts by up to 13 mas between the optical fit and the optical+VLBI fit. The shapes of the confidence regions are also altered. However, the value of H_0 is not very sensitive to this mainly northerly shift.

4.7 Extraction of H_0

The general relation between the time delay $\Delta t_{i,j}$ between the i^{th} and j^{th} images, the Hubble constant H_0 and the lens model, parametrised by the potential ψ, is given by

$$c\,\Delta t_{i,j} = \frac{1+z_l}{H_0}\frac{d_l\,d_s}{d_{ls}}(\phi_i - \phi_j) \quad , \tag{4.8}$$

where z_l is the redshift of the lens, d_l and d_s are the angular size distances to the lens and source, respectively, d_{ls} is the angular size distance to the source measured from the lens, and ϕ_i is the scaled time delay at the position of the i^{th} image ($\boldsymbol{\theta}_i$),

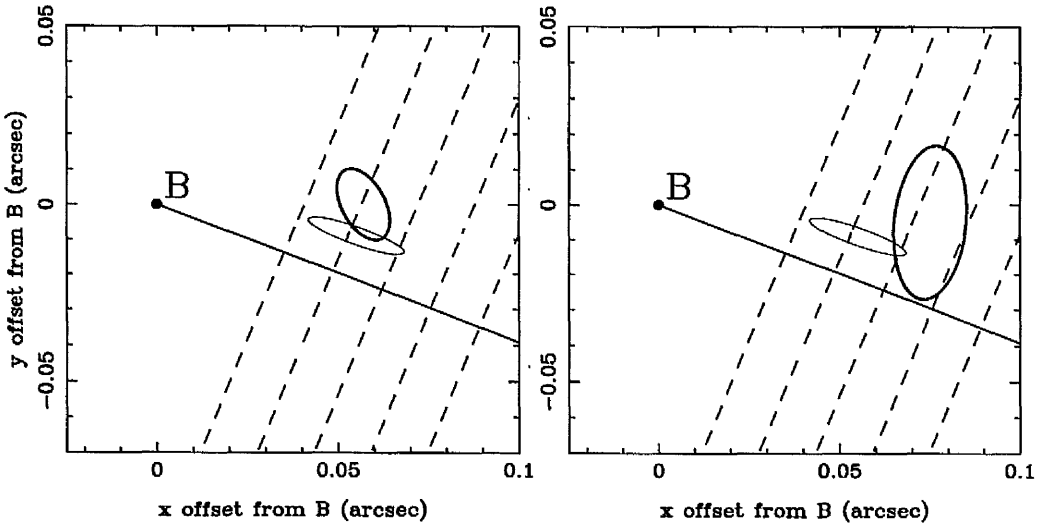


Figure 4.8: The optical galaxy position compared to that determined by Wucknitz, Biggs & Browne (2004) using LensClean. The error ellipses are 95% confidence regions. The left-hand plot is for the case with no masking. The right-hand plot represents the case with masking of spiral arms. The position of B is marked, as is a line pointing towards the A component. The dotted lines are contours of H_0 in the strictly isothermal case, and correspond (from right to left) to $H_0 = (90,80,70,60,50)$ km s^{-1} Mpc^{-1}. The positions determined from optical data are shown bolded.

Data used	Masking	Mass profile	Position (mas)		H_0	H_0	η	Ellipticity
			$\Delta\alpha$	$\Delta\delta$	(isothermal)	(variable η)		(of potential)
Optical	None	-	+57	+1	79±7	68±6	$1.13^{+0.07}_{-0.09}$	0.08±0.03
Optical	Spiral	-	+75	−5	66±9	56^{+12}_{-15}	1.16±0.19	0.05±0.04
VLBI+Optical	None	Isothermal	+60	−13	74±5	-	-	0.05±0.02
VLBI+Optical	Spiral	Isothermal	+74	−19	64±7	-	-	0.03±0.02
VLBI+Optical	None	Variable	+60	−12	-	70±5	1.05±0.03	0.04±0.02
VLBI+Optical	Spiral	Variable	+74	−18	-	61±7	1.05±0.04	0.04±0.02

Table 4.3: Lens galaxy positions and the corresponding values of H_0 and the mass profile slope η. The optical positions are derived from the ACS images only. The "VLBI+Optical" positions incorporate constraints from the LENSCLEAN-based lens modelling of Wucknitz, Biggs & Browne (2004). The "Mass profile" column indicates what mass profile was assumed when combining the VLBI and optical constraints. H_0 values are given in $\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Position offsets are referenced to image B, and RA offsets are given taking west as positive. All errors are quoted at 95% confidence.

$$\phi_i = \frac{1}{2}|\nabla\psi(\boldsymbol{\theta}_i)|^2 - \psi(\boldsymbol{\theta}_i) \quad . \tag{4.9}$$

The angular size distances are normalized in these equations, since they do not include factors of H_0. For general isothermal models without external shear the relation becomes particularly simple and can be written as a function of the image positions alone, without explicitly using any lens model parameters (Witt, Mao & Keeton, 2000):

$$\phi_i = \frac{1}{2}|\boldsymbol{\theta}_i - \boldsymbol{\theta}_0|^2 \tag{4.10}$$

Here $\boldsymbol{\theta}_0$ is the position of the centre of the lens. External shear γ changes ϕ_i by a factor between $1 \pm \gamma$ depending on the relative direction, typically resulting in similar factors for the value deduced for H_0. A general analysis for power-law models with external shear can be found in Wucknitz (2002).

Using the recipe described in previous sections our lens position translates to a Hubble constant of $\mathrm{H}_0 = 79 \pm 7\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ in the shearless isothermal case[2] without masking, and to $66 \pm 9\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ with masking.

Estimates of external shear and convergence from nearby field galaxies and large scale structure are of the order 2 per cent (Lehár et al., 2000) and would affect the result only to the same relative amount, sufficiently below our current error estimate to allow us to neglect these effects.

The value of the Hubble constant we derive depends on the slope of the mass distribution of the lensing galaxy. In figure 4.9 we show the permitted values of the Hubble constant for different models – isothermal and with a variable η in an elliptical potential model – plotted against measured galaxy position. The

[2]A concordance cosmological model with $\Omega = 0.3$ and $\lambda = 0.7$ and a homogeneous matter distribution is used for the calculation of all distances in this chapter

elliptical power-law potential is parametrised as follows:

$$\psi(\boldsymbol{\theta}) = \frac{\theta_{\mathrm{E}}^{2-\eta}}{\eta} r_{\epsilon}^{\eta}(\boldsymbol{\theta}) \quad , \tag{4.11}$$

$$r_{\epsilon}^{2} = \frac{\theta_x^2}{(1+\epsilon)^2} + \frac{\theta_y^2}{(1-\epsilon)^2} \quad , \tag{4.12}$$

$$\boldsymbol{\theta} = (\theta_x, \theta_y) \quad , \tag{4.13}$$

where θ_E is the Einstein radius of the model, η is the power-law index of the potential's profile and ϵ is the ellipticity of the potential. For details of our modelling procedure the reader is referred to Wucknitz, Biggs & Browne (2004). It is evident that the preferred value of the Hubble constant is somewhat reduced compared to what is obtained by forcing the mass distribution to be isothermal. We also show contours of the radial power law η plotted against galaxy position. The optical lens position gives $\eta = 1.13^{+0.07}_{-0.09}$ $(2\,\sigma)$.

As discussed before, B0218+357 has the advantage of clear substructure in the two images which can be mapped with VLBI. The VLBI data can be used independently to derive the slope of the mass profile of the lens (Wucknitz & Sperhake, 2004). Biggs et al. (2003) and Wucknitz, Biggs & Browne (2004) find a value of η=1.04±0.02. Combining the VLBI constraints with the optical lens position gives a value of η=1.05±0.03 (95% confidence). We therefore adopt this value for the mass profile's logarithmic slope and obtain a Hubble constant of 70±5 $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (95% confidence) for the case with no masking and 61±7 $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (95% confidence) when the spiral arms are masked. The ellipticity of the potential is small (about 0.04) in each case, but the ellipticity of the mass distribution will be about three times this, 0.12. The lens galaxy could be more inclined than it appears, or there could be a bar or similar feature present.

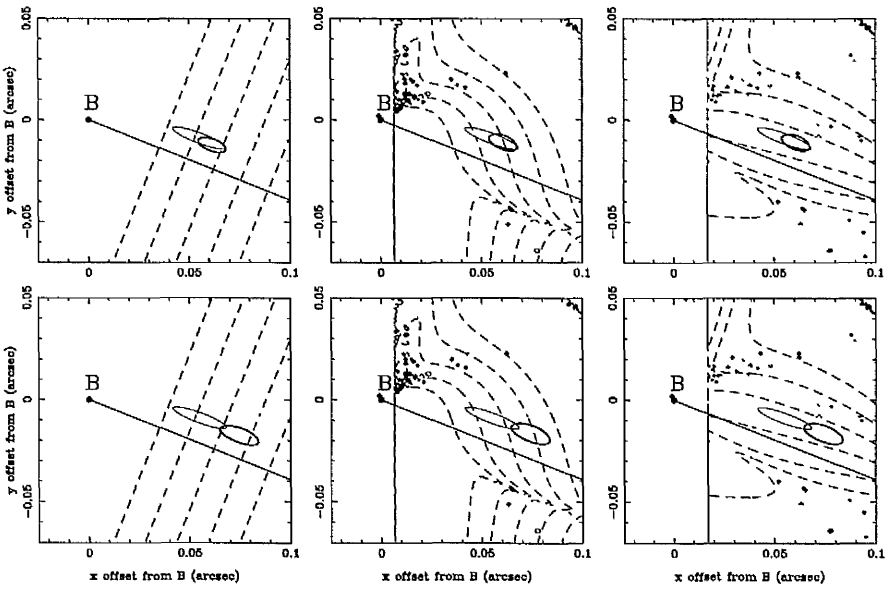


Figure 4.9: The 95% confidence regions for the cases of no masking (top row) and masked spiral arms (bottom row). The plots in the left-hand column show the galaxy position confidence regions and contours of H_0 calculated for isothermal models. The plots of the central column show the confidence regions superimposed over contours of H_0 calculated whilst allowing the logarithmic mass slope η to vary. This reduces H_0 slightly compared to the fixed $\eta = 1$ (isothermal) case. The plots in the right-hand column show the same confidence regions superimposed over contours of η. The confidence regions take both our optical position and the VLBI constraints used by Wucknitz, Biggs & Browne (2004) into account. The 95% confidence error ellipse of Wucknitz, Biggs & Browne is also shown. Contours of H_0 are again at (90,80,70,60,50)km s^{-1} Mpc^{-1} reading from left to right across an image, and contours of η are (0.8,0.9,1.0,1.1,1.2,1.3) reading from bottom to top. The optical+VLBI error ellipses are shown in bold relative to the LensClean ellipse of Wucknitz, Biggs & Browne.

4.8 Conclusions

We have analysed the deepest optical image yet taken of B0218+357 to measure the position of the lens galaxy. We find that simple subtraction of a parametric galaxy model and two point sources is insufficient to constrain the galaxy position, and we confirm earlier suggestions that the image separation in the optical is lower than that in the radio, most probably due to significant extinction around image A. Taking advantage of the symmetric appearance of the lens, we have defined the centre as that point about which the residuals (after subtraction of A and B) are most symmetric. To account for artifacts in our empirical PSF model we have introduced an extra noise term. We have also masked off the most prominent spiral arms to test the effect on H_0. We find that the lens galaxy position is 57±2 mas west and 0±2 mas south of image B when no masking is applied. Combined with the results of Wucknitz, Biggs & Browne (2004), this leads to a value for H_0 of 70±5 $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (95% confidence). When the most obvious spiral arms are masked out, we find an optical galaxy position of 75±2 mas west and −6±5 mas south from image B. This results in a value for H_0 of 61±7 $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (95% confidence) when combined with VLBI constraints.

Further work on this lens will involve increased use of LensClean to further limit the power law exponent η using VLBI constraints. Observations have also been made using the VLA with the Pie Town VLBA antenna, which together with VLBI will further improve the lens model for this system.

Bibliography

Biggs A.D., Browne I.W.A., Helbig P., Koopmans L.V.E., Wilkinson P.N., Perley R.A., 1999, MNRAS, 304, 349

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

Biggs A.D., Wucknitz O., Porcas R.W., Browne I.W.A., Jackson N.J., Mao S., Wilkinson P.N., 2003, MNRAS, 338, 599

Blakeslee J.P., Anderson K.R., Meurer G.R., Benitez N., Magee D., 2003, Astronomical Data Analysis & Software Systems, vol 295, eds H.E. Payne et al., ASP, San Francisco, p. 257

Browne I.W.A., Patnaik A.R., Walsh D., Wilkinson P.N., 1993, MNRAS, 263, L32

Carilli C.L., Rupen M.P., Yanny B., 1993, ApJ, 412, L59

Clampin M., et al., 2000, SPIE, 4013, 344

Cohen A.S., Hewitt J.N., Moore C.B., Haarsma D.B., 2000, ApJ, 545, 578

Cohen J.G., Lawrence C.R., Blandford R.D., 2003, ApJ, 583, 67

Combes F., Wiklind T., 1997, ApJ, 486, L79

Fruchter A., Hook R.N., 1997, in "Applications of Digital Image Processing XX", Tescher A.G., ed, Proc. SPIE Vol. 3164, p. 120-125

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

Grundahl F., Hjorth J., 1995, MNRAS, 275, L67

Hack W.J., 2002, in Bohlender D, et al., eds, "Astronomical Data Analysis Software and Systems XI", ASP Conference Proceedings, Vol. 281., p.197, ASP: San Francisco.

Henkel C., Jethava N., Kraus A., Menten K.M., Carilli C.L., Grasshoff M., Lubowich D., Reid M.J., 2005, A&A, 440, 893

Hjorth J., 1997, in "Proceedings of the Golden Lenses, Hubble's constant and galaxies at high redshift", Workshop held at Jodrell Bank, June 1997

Hook R.N., Pirzkal N., Fruchter A.S., 1999, in ASP Conf. Ser., Vol. 172, Astronomical Data Analysis Software and Systems VIII, eds. D.M. Mehringer, R.L. Plante, D.A. Roberts (San Francisco:ASP), 337

Howell S.B., 1992, in "Astronomical CCD observing and reduction techniques", ASP Conf. Ser., Vol. 23, (San Francisco:ASP)

Jackson N., Xanthopoulos E., Browne I.W.A., 2000, MNRAS, 311, 389

Kemball A.J., Patnaik A.R., Porcas R.W., 2001, ApJ, 562, 649

Kochanek C.S., 2003, ApJ, 583, 49

Koekemoer A.M., Fruchter A.S., Hook R.N., Hack W., 2002, in S. Arriba, ed., "The 2002 HST Calibration Workshop : Hubble after the Installation of the ACS and the NICMOS Cooling System", Proceedings of a Workshop held at the Space Telescope Science Institute, Baltimore, Maryland, Baltimore:STScI, p. 339

Lauer T.R., 1999, P.A.S.P., 111, 1434

Landolt A.U., 1992, AJ, 104, 340

Lehár J., et al., 2000, ApJ, 536, 584

Mack J., et al., 2003, ACS Data Handbook, Version 2.0, Baltimore:STScI

McLure R.J., Kukula M.J., Dunlop J.S., Baum S.A., O'Dea C.P., Hughes D.H., 1999, MNRAS, 308, 377

Menten K.M., Reid M.J., 1996, ApJ, 465, L99

Merline W.J., Howell S.B., 1995, Expt. Astron., 6, 163

Mutchler M., Cox C., 2001, Instrument Science Report ACS 2001-07, Baltimore:STScI

O'Dea C.P., Baum S.A., Stanghellini C., Dey A., van Breugel W., Deustua S., Smith E.P., 1992, AJ, 104, 1320

Patnaik A.R., Browne I.W.A., Wilkinson P.N., Wrobel J.M., 1992, MNRAS, 254, 655

Patnaik A.R., Browne I.W.A., King L.J., Muxlow T.W.B., Walsh D., Wilkinson P.N., 1993, MNRAS, 261, 435

Patnaik A.R., Porcas R.W., Browne I.W.A., 1995, MNRAS, 274, L5

Pavlovsky C., et al., 2002, "ACS Instrument Handbook", Version 3.0, Baltimore:STScI

Schechter P., 2001, in T.G. Brainerd & C.S. Kochanek., eds., "Gravitational Lensing: Recent Progress and Future Goals", ASP Conf. Proc., Vol. 237, San Francisco:ASP

Stickel M., Kuhr H., 1993, A&AS, 101, 521

Wiklind T., Combes F., 1995, A&A, 299, 382

Williams R.E., et al., 1996, AJ, 112, 1335

Witt H.J., Mao S., Keeton C.R., 2000, ApJ, 544, 98

Wucknitz O., 2002, MNRAS, 332, 951

Wucknitz O., 2004, MNRAS, 349, 1

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

Chapter 5

Summary and conclusions

5.1 A new lens modelling code

A computer code was developed for use in modelling JVAS/CLASS radio lenses. Such lenses usually consist of multiple unresolved images of a lensed radio-loud AGN. The code uses a regular triangular tiling to isolate roots of the (non-linear) bend angle equation, which are solved for using the Newton-Raphson method. To ensure merging images are separated, the code subdivides the tiling close to critical curves.

The code was applied to the CLASS gravitational lens B0128+437, an interesting lens system in which the background source consists of three compact radio components, each lensed into four images, A1-3, B1-3, C1-3 and D1-3. The code was used in an attempt to find a smooth model with which the deflection angle near image B and the distortion near image C could be simultaneously reproduced. No such model could be found. The code was then used to test various simplified models of the mass substructure in the system. A point mass model was placed close to images C1-3 to try and improve the best fit; this approach gained some success but the model predictions and the observations were still inconsistent. A second attempt involved placing a non-singular isothermal

sphere near image B, but it produced no improvement in the mass model. Future work on this lens system will involve increasing the freedom of the mass model, perhaps by application of a multi-resolution version of the recently introduced gravitational imaging technique.

5.2 Observations of JVAS B0218+357

The lens system JVAS B0218+357 has been observed using the Advanced Camera for Surveys on HST. The images have sufficient sensitivity to show the spiral arms of the lens galaxy and demonstrate clearly that the lens is almost face-on. The lens has a very low surface brightness relative to the images of the background quasar. The ACS data add to previous evidence that the image separation varies with wavelength; the ACS data support an image separation of 317±4 mas, while the same quantity in the radio is 334 mas. The F814W flux ratio between the images is anomalous; image B is between 10 and 14 times brighter than image A in the ACS data, while at 8.4 GHz image A is roughly three times brighter than image B. This is in agreement with previous imaging work. The most likely cause of the anomalous optical flux ratio and image separation is then a giant molecular cloud obscuring image A, although it is surprising that image B is not absorbed as much as image A despite being located closer to the centre of the lens than image A. Evidence from molecular line studies support this hypothesis, however.

To establish the offset between the lens galaxy centre and the quasar images (the only features visible at both radio and optical wavelengths), it was necessary to subtract PSFs from the data that were placed to model the lensed quasar images. Parametric models were initially used to model the smooth light distribution of the lens galaxy, but the large dynamic range between the lensed images

and the galaxy, the presence of the spiral arms and residuals left over from subtraction of the PSFs from the quasar images made a stable fit impossible. To avoid the need for a parametric model the centre of the galaxy was taken to be that point about which the residuals (after subtraction of the quasars) were most symmetric. A simple weighting scheme was also used to reduce the influence of pixels affected by the subtraction of the quasars. The resulting method produced a galaxy position offset from image B by 57 mas west and 1 mas north.

This position was combined with existing lens models based on radio data to fix Hubble's constant, giving $H_0 = 70 \pm 5$ km s^{-1} Mpc^{-1}(95% confidence). The random errors were estimated from the scatter in the lens galaxy positions obtained from separate exposures. A source of systematic error has been identified in the measurement. The symmetry method is sensitive to the disposition of the spiral arms of the lens. If the arms are masked out, the galaxy position changes by 19 mas compared to the case without masking. This reduces the estimated value of Hubble's constant by 9 km s^{-1} Mpc^{-1}. The spiral arms make up only a minor part of the total lens mass (which should be dominated by the lens galaxy's dark matter halo), but they have a substantial influence on the lens's apparent centre of light because they are the dominant source of asymmetry in the lens galaxy image after subtraction of the quasar images.

It is unlikely that any improvements can be expected in the quality of the ACS data, which is already far higher than that available on other lenses. The possible areas of improvement include better characterisation of the ACS PSF and the recovery of visit 15 using the techniques pioneered for the Hubble Deep Field. It is not clear how the effect of the spiral arms on the optimum galaxy position may be reduced, given the low surface brightness of the lens.

5.3 Observations and modelling of B0631+519

In the case of B0631+519, this work has shown that three distinct regions of the lensed radio source can be identified using new data from the VLBA and MERLIN at 1.7 GHz. One of the regions is quadruply imaged while the other two are doubly-imaged. Previous observations were unable to distinguish between two different candidate morphologies. The host galaxy is a confirmed elliptical and now has a measured position offset relative to the lensed radio images. A prediction has been made for bounds on the redshift of the lensed source. The work has also shown that a simple lens model can fit the radio data, and has given a plausible position for the quadruply imaged source. The presence of a near-infrared Einstein ring and the large number of constraints available at radio wavelengths will make this an excellent system in which to examine the mass profiles of early-type lens galaxies.

Future investigations could determine the influence of L2 on the lensing properties of the system[1], especially through the use of molecular line absorption studies and polarisation studies. VLBI observations of image B, made with a view to detecting counterparts of images A1a-d, would allow the components of the magnification tensor to be determined at two points, which, as discussed in chapter 2, is sufficient to allow the logarithmic slope of the radial part of the surface mass density to be estimated.

Source variability would make this lens an excellent system with which to measure Hubble's constant. The likely time delay between images A1 and B1 is on the order of a week, which is within the grasp of radio monitoring campaigns (for instance, B0218+357 has a time delay of similar length). The evidence for source variability in B0631+519 is weak, based on the data collected so far. Monitoring

[1]The near-infrared data can be inverted through gravitational mass imaging to show that the faint irregular galaxy G2 alters alters the lens magnification away from that predicted by the simple SIE model by about 10% (L.V.E. Koopmans, private communication)

with an instrument such as WSRT would allow source variability to be clearly identified, if present. Further spectroscopy would be useful in establishing both the velocity dispersion of the lens and the lensed source's redshift. This step will be essential if the source is found to be time-variable and becomes the target of a monitoring campaign.