GRAVITATIONAL LENS MASS RECONSTRUCTION

Simon Dye

Presented for the degree of Doctor of Philosophy at
The University of Edinburgh
1999

This thesis is solely my own composition, except where specifically indicated in the text.

Simon Dye,
July 1999.


10
S. Dye
12/99

Contents

Constants and Units used in this Thesis

G	Newton's gravitational constant	6.67×10^{-11}	$\mathrm{Nm^2kg^{-2}}$
c	Speed of light in a vacuum	3.0×10^{8}	$\mathrm{ms^{-1}}$
$\mathrm{H_0}$	Hubble's constant	$100h$	$\mathrm{kms^{-1}Mpc^{-1}}$
$\mathrm{M_\odot}$	Solar mass	1.99×10^{30}	kg
$\mathrm{L_\odot}$	Solar luminosity	3.83×10^{33}	erg $\mathrm{s^{-1}}$
h	Planck's constant	6.63×10^{-34}	Js
k_B	Boltzmann's constant	1.38×10^{-16}	erg K−1
Mpc	Mega parsec	3.1×10^{22}	m
eV	Electron Volt	1.60×10^{-19}	J

ACKNOWLEDGEMENTS

Attempting to thank everyone who has contributed to this work is almost an impossibility. I apologise now for the uncountable number of people who inevitably I cannot mention through lack of space (or lack of a good memory !).

My first acknowledgement undoubtedly goes to my supervisor Andy Taylor. Chosen some six weeks after first arriving at the IfA in January 1996 only to discover the imminent and permanent departure of my intended supervisor Lance Miller, Andy quickly began to impress me with his sharp intellect, unwavering and sanguine joviality and his sink-like Guinness drinking capability. He has made my stay at Edinburgh both an enlightening and an entertaining one. Not once has he turned me away from his office claiming a lack of time, not once has he failed to conjure up some piece of wizardry from his bag of mathematical tricks to solve any problem I've ever presented him with and not once has his level of interest or enthusiasm in my work dwindled. Andy, I thank you.

Secondly, I would like to thank Eduard Thommes for his assistance towards the photometric work presented in Chapter 5. Without Eduard's help either in Edinburgh or during my stay at the MPIA in Heidelberg, circumventing the deluge of problems this work constantly presented would have been an even more arduous task. So long Eduard and thanks for all the rice (I jest not).

Staying on the academic side of things, I must thank all members of staff at the IfA who have helped me in one way or another. Thanks also to Klaus Meisenheimer at the MPIA for his support throughout my stay there and during composition of our A1689 paper. I would also like to extend my gratitude to Eelco van Kampen for being generally helpful and a source of amusing yet surreal humour. Finally, thanks to Alan Heavens and Richard Ellis for trawling through this thesis prior to my viva.

On a less academic level, I send a huge thank you to my parents for their moral (and financial !) support. Its nice to know there's always someone looking out for you, be it in the form of ensuring you've consumed a sufficient volume of greens or the occasional sending of a tenner through the post. Hugs and thanks to Charlie for

not only providing me with additional moral support, but also her cottage which I used as an idyllic writer's retreat during the composition of some of this work. I must mention my uncle, Roy Jackson, for his probing questions and general interest in my work throughout its duration.

Last but probably not least, thanks to those who provided my social (and sometimes office) entertainment. Most notably, I refer to the original lower terminal room posse comprising Richard Knox, Alison Stirling and Elese Archibald who collectively and patiently bore the brunt of my perhaps over zealous supply of practical jokes. A quick acknowledgement to Peter Watts and Rob Smith, both of whom I have had the priviledge of sharing a flat *and* an office with and both of whom have caused me a mixture of emotions with their guitar playing. Finally, hi to Matthew Horrobin who never seems short of ideas or tenacity when it comes to going for a drink, Jo McAllister who on numerous occasions made me run around a squash court until I nearly collapsed and Kenton D'Mellow who never seemed to know what day it was (and probably still doesn't - its Tuesday today Kent!). May the force be with you all.

Abstract

A highly desirable technique sought after by cosmology is one which enables the accurate mass measurement of rich galaxy clusters. From observations of their abundance and primarily their mass, clusters give strong constraints on the density parameter of the Universe, models of structure formation and normalisation of the power spectrum of density fluctuations. Gravitational lensing provides such a technique. Prevailing over X-ray temperature and virial velocity methods known to be problematic, lensing permits determination of cluster masses independent of dynamical state.

This thesis concentrates mainly on the exploitation of the magnification properties of lenses rather than those of shear analysis which relies upon the quantification of galaxy image distortions. Magnification allows absolute mass measurements, breaking the sheet-mass degeneracy experienced by shear. To this extent, a theoretical analysis of the geometrical magnification of angular separations between galaxies lying behind a lensing cluster is performed. This sees application to the cluster Abell 1689 using V and I band observations to select background galaxies based on their V-I colour. The distribution of source number counts in the observed field of view results in the determination of a radial mass profile and a mass map for Abell 1689. This predicts a projected mass interior to $0.24h^{-1}$Mpc of $M(< 0.24h^{-1}\mathrm{Mpc}) = (0.50 \pm 0.09) \times 10^{15}h^{-1}\mathrm{M}_{\odot}$.

A new method of directly determining accurate, self-consistent lens mass and shear maps in the strong lensing regime from magnification is presented. The method relies upon pixellization of the surface mass density distribution which generates a simple, solvable set of equations. The concept of pixellization is also directed at shear analysis to give rise to a simplified method of application. Through use of cluster models,

the method is verified before the magnification data from the colour-selected number counts is input to compute a self-consistent mass map of Abell 1689.

The property of lens magnification to amplify observed background source fluxes is investigated. Using an independent set of observations in nine optimally chosen filters, photometric redshifts of objects lying in the field of Abell 1689 are calculated. In addition to providing an unambiguous distinction between cluster members, foreground objects and background sources this also enables computation of the source luminosity function. Comparison of this with the distribution of luminosities in an observed offset field quantifies the lens-induced flux magnification to arrive at an independent mass profile measurement of Abell 1689. A projected mass interior to $0.25h^{-1}$Mpc of $M(< 0.25h^{-1}\mathrm{Mpc}) = (0.48 \pm 0.16) \times 10^{15} h^{-1}\mathrm{M}_{\odot}$ is found.

Chapter 1

Introduction

The nature and origin of the Universe has been the subject of much speculation since intelligent, reasoning human beings first inhabited the Earth's surface and began to look up into the sky. The route to the development of a successful theory capable of explaining observed cosmological phenomena has proved to be a tortuous one with many pitfalls along the way. It is only in this century that mankind has achieved the bulk of its understanding of our cosmos, largely aided by the rapid development of technology. Only now are we equipped with a sufficiently advanced arsenal of instrumentation to begin to satisfy the questions that have been left unanswered all this time.

This chapter is divided into two main parts. The first gives a review of modern physical cosmology, mainly emphasising those areas relied upon by later chapters. A short summary on the notion of dark matter is given. The second part of this chapter discusses the importance of studying galaxy clusters, namely the way in which they provide constraints on various cosmological parameters and structure formation models. A brief outline and comparison of mass derivation methods alternative to those of gravitational lensing is presented.

1.1 Cosmology

Cosmology is the branch of astronomy concerned with the origin, properties and evolution of our Universe. Unlike the apparatus of a laboratory experiment which allows direct interaction, astronomers are forced into the more passive role of relying upon observations of the Universe for the formulation of physical laws which govern it.

The currently accepted Big Bang model of the Universe owes its reputation to several supporting observations. Two of the most important astronomical discoveries of this century which have provided the strongest evidence in favour of the Big Bang are the discovery of the cosmic microwave background radiation (CMBR) by Penzias and Wilson in 1965 and the discovery of Universal expansion by Hubble in 1929. The validity of Hubble's law (see Section 1.1.3) supports the case of an expanding Universe whereas the CMBR is almost indisputable evidence of a hot, dense beginning. Measurements by the COBE (COsmic microwave Background Explorer) satellite launched in 1989 showed that the CMBR is isotropic to one part in 10^5. This is in excellent agreement with inflation (Guth 1981) which predicts a highly isotropic universe on scales far larger than the size of regions in causal contact.

The Copernican Principle asserts that we are not privileged observers. Assuming this holds true, our isotropic Universe must therefore be a homogeneous one. The condition of homogeneity and isotropy on large scales is a statement of the Cosmological Principle; a fundamental concept in cosmology which gave rise to the development of the Robertson-Walker metric in 1935 (see Section 1.1.2). Modern observations of the distribution of mass from redshift surveys certainly seem to validate this statement (eg. Tadros et al 1999). In the light of this and similar evidence gathered by astronomers over the years, cosmology has proceeded by extrapolating the properties and laws of the observable Universe to the Universe as a whole.

1.1.1 Newton's Law of Gravity

The strongest force of nature on large scales is that due to gravity. It is therefore not surprising that gravity plays a central role in any theory used to model the evolution and dynamics of the Universe. Sir Isaac Newton in the 17th Century was responsible

for the first description of gravity with his 'law of universal gravitation'. This quantified the attractive force between two objects with masses m_1 and m_2 separated by a distance $r = |\boldsymbol{r}|$ as

$$\boldsymbol{F} = \frac{Gm_1m_2\boldsymbol{r}}{r^3}. \tag{1.1}$$

Newton also established a framework for mechanics, one of the most important results of which was his 'second law of motion' which stated that a body of 'inertial mass' m feels an acceleration $\boldsymbol{a} = \boldsymbol{F}/m$ if acted upon by a force $\boldsymbol{F}$. This inertial mass turns out to be exactly the same as the gravitational mass affecting the magnitude of the gravitational pull; a fact known as the 'weak equivalence principle'. Using this as the basis for his 'strong equivalence principle' which postulates that physics in freely falling and inertial frames is identical, Einstein developed his General Theory of Relativity. The General Theory of Relativity provides the best description of gravity to date.

1.1.2 The Friedmann Equations

In the special theory of relativity, the *interval* between two points in space-time with co-ordinates (t, x, y, z) and $(t + \mathrm{d}t, x + \mathrm{d}x, y + \mathrm{d}y, z + \mathrm{d}z)$ is defined as

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - (\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2). \tag{1.2}$$

This quantity is invariant under a co-ordinate transformation and equals zero if the two co-ordinates it spans are linked by a ray of light. The interval may be written in terms of the *metric tensor*, g_{ij}, such that

$$\mathrm{d}s^2 = g_{ij}\mathrm{d}x^i\mathrm{d}x^j \tag{1.3}$$

where $x^0 \equiv ct$ and x^1, x^2, x^3 are the spatial co-ordinates x, y, z. The metric tensor completely describes the geometry of space-time. For the flat geometry described by special relativity (referred to as Minkowski space-time), the metric is evidently the diagonal matrix diag$[1, -1, -1, -1]$. To allow for a description of curved space-times, a more general metric is needed. It can be shown from simple geometric considerations (Weinberg 1972) that the most general space-time metric describing a homogeneous

and isotropic universe is that of the *Robertson-Walker* (RW) metric,

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a(t)^2\left[\frac{\mathrm{d}r^2}{1-Kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi)\right], \tag{1.4}$$

where the spherical polar, co-ordinates (r, θ, ϕ) have been used. These co-ordinates are 'comoving co-ordinates' which, as the name implies, move with the expansion of the Universe. Physical distances are decomposed into a time-independent comoving co-ordinate and a time-dependent, dimensionless *scale factor*, $a(t)$. This is defined by the ratio of proper distances (see Section 1.1.8),

$$a(t) = \frac{R(t)}{R(t_0)} \tag{1.5}$$

where t_0 is the current epoch. The co-ordinate r is referred to as the 'comoving angular diameter distance'. K is a free parameter which controls the space curvature of the Universe depending on its acquired value. The comoving co-ordinates can be arbitrarily scaled such that K takes on one of three values:

- $\boldsymbol{K = 1}$ Space is 'closed', meaning that a trajectory of fixed θ and ϕ will always end where it started from. Geometry is described by the surface of a 3-sphere[1] and thus the internal angles of a triangle add up to $> 180°$.
- $\boldsymbol{K = 0}$ Space is flat in which case Euclidean geometry applies.
- $\boldsymbol{K = -1}$ Space is 'open' and infinite in extent. The geometry of such a space is described by the 4 dimensional equivalent of the surface of a saddle in 3 dimensions. The internal angles of a triangle add up to $< 180°$.

The geometry of space-time is determined by matter in the Universe. The fundamental equation resulting from Einstein's General Theory of Relativity relates the curvature of space to the energy density and momentum of matter in the Universe though the *energy-momentum tensor*, T_{ij} (eg. Peacock 1999),

$$\frac{8\pi G}{c^4}T_{ij} = R_{ij} - \frac{1}{2}Rg_{ij} - \Lambda g_{ij}. \tag{1.6}$$

Here, R_{ij} is the *Ricci tensor* which holds spatial curvature information and R is the *Ricci scalar* formed from the contraction of the metric tensor and the Ricci tensor.

[1] A 3-sphere is the four dimensional equivalent of the surface of a sphere in 3 dimensions.

Λ is the *cosmological constant* which, if greater than zero, corresponds to a repulsive force acting against the pull of gravity on the Universe and vice versa if negative.

Using the RW metric, the solutions to Einstein's equation (1.6) are given by the *Friedmann equations*:

$$\dot{a}^2 = \frac{8\pi G}{3}\rho a^2 - Kc^2 + \frac{\Lambda c^2 a^2}{3} \tag{1.7}$$

$$\ddot{a} = -\frac{4}{3}\pi G\left(\rho + \frac{3p}{c^2}\right)a + \frac{\Lambda c^2 a}{3}. \tag{1.8}$$

In these equations, a is the scale parameter of equation (1.5), ρ is the volume density of matter in the Universe and p is the matter 'pressure'. This pressure is effectively a correction for relativistic particles and is negligible in the Newtonian limit. It originates from the thermal motion of particles in the same way that the classical definition of pressure stems from the collision of gas molecules with a surface. p can be thought of as a flux density of momentum[2].

The Friedmann equations govern the dynamics of the Universe given its matter content and the contribution from Λ. They are not independent of each other; equation (1.8) can be derived from equation (1.7) by taking into consideration the evolution of density and therefore pressure.

A similar result can be obtained from Newtonian arguments. By considering the acceleration of a test particle on the boundary of a sphere due to the mass it encloses, an analogous form of equation (1.8) can be derived[3]. Integrating this gives the Newtonian equivalent of equation (1.7). The difference between the 'Newtonian Friedmann' and the full Friedmann equations is the Λ term, the relativistic pressure term in equation (1.8) and the interpretation of K as Universal curvature which do not arise from Newtonian considerations.

[2]The relativistic correction for density is $\rho \to \rho(1 + v^2/c^2)$ which combined with $p = \rho v^2/3$ gives $\rho \to \rho + 3p/c^2$.

[3]This only holds true if spherical symmetry is assumed. If the collapsing region is ellipsoidal for example, GR predicts a different result from that predicated by Newtonian gravity.

1.1.3 The Hubble Parameter

The uniform expansion of the Universe as a result of the Big Bang has the consequence that, in the absence of gravitational influences, any two galaxies in the Universe recede from each other. Because the expansion is uniform, larger separations give rise to larger recession velocities. This was first discovered by Edwin Hubble in 1929. In plotting galaxy recession velocities (away from us) against their distance (determined from Cepheid variable stars which vary on a timescale dependent on their intrinsic brightness), he noticed a correlation characterised by

$$v = H_0 R \tag{1.9}$$

This is *Hubble's Law* which relates an object's proper distance R (see Section 1.1.8) with its recession velocity v. H_0 is referred to as the Hubble constant. As the Universe evolves, this changes as the rate of expansion changes and so the subscript '0' denotes its value at the current epoch (this is a convention adopted hereafter). More generally, equation (1.9) is expressed as,

$$\dot{R}(t) = H(t)R(t) \tag{1.10}$$

where $H(t)$ is the *Hubble parameter* and the dot denotes time differentiation. Using the result of equation (1.5), $H(t)$ in terms of the scale factor is

$$H(t) = \frac{\dot{a}(t)}{a(t)}. \tag{1.11}$$

The value of the Hubble constant is presently a matter of much debate although recent estimates seem to scatter about a value of $H_0 \sim 65\text{kms}^{-1}\text{Mpc}^{-1}$ with an uncertainty of about 30% (eg Freedman et al 1998). Knowledge of the Hubble constant is required for the calculation of many cosmological quantities and so to compensate for its uncertainty, it is parameterised by the dimensionless quantity h defined as

$$h = \frac{H_0}{100\text{kms}^{-1}\text{Mpc}^{-1}}. \tag{1.12}$$

1.1.4 Redshift

Just as acoustic waves experience a Doppler-shift when a source moves with respect to an observer, the light from objects receding from us due to Universal expansion is

shifted to lower frequencies. Objects further away which recede more quickly therefore appear more red and are said to be 'redshifted'. Redshift is defined by

$$z = \frac{\lambda_o - \lambda_e}{\lambda_e} \tag{1.13}$$

where λ_o and λ_e is the observed and emitted wavelength respectively.

An object's redshift can be expressed in terms of its recession velocity. Suppose a light-emitting object moves with a velocity $\boldsymbol{v}$ relative to an observer. If, in the rest frame of the body, wavecrests are separated by a time t, measured to be t' in the rest frame of the observer, the Lorentz transformation for time stipulates that

$$t' = \frac{t - (-\boldsymbol{v}).\boldsymbol{x}/c^2}{\sqrt{1 - v^2/c^2}} \tag{1.14}$$

where $\boldsymbol{x}$ is the vector in the body's rest frame between wavecrests. The magnitude of $\boldsymbol{x}$ is simply the emitted wavelength of light, $\lambda_e = ct$. Similarly, $\lambda_o = ct'$ so that equation (1.14) becomes

$$\frac{\lambda_o}{\lambda_e} = \frac{1 + (v/c)\cos\theta}{\sqrt{1 - v^2/c^2}} = 1 + z. \tag{1.15}$$

If the body recedes radially, $\boldsymbol{x}$ and $\boldsymbol{v}$ are parallel so that equation (1.15) gives

$$1 + z = \left(\frac{1 + v/c}{1 - v/c}\right)^{1/2}. \tag{1.16}$$

Cosmological Redshift

Using the RW metric, redshift due to the expansion of the Universe can be expressed in terms of the scale parameter. Consider a light ray emitted from a source with constant comoving co-ordinates such that two subsequent wavecrests are emitted at proper times t_e and $t_e + \Delta t_e$ and observed at times t_o and $t_o + \Delta t_o$. As stated in Section 1.1.2, the interval for a light ray equals identically zero so that the RW metric for a radially travelling beam[4] is therefore integrated to give

$$\int_{t_e}^{t_o} \frac{c\,\mathrm{d}t}{a(t)} = \int_0^r \frac{\mathrm{d}r}{\sqrt{1 - Kr^2}} = \int_{t_e + \Delta t_e}^{t_o + \Delta t_o} \frac{c\,\mathrm{d}t}{a(t)}. \tag{1.17}$$

[4]The radial direction is chosen for simplicity, without loss of generality.

The right hand side of this expression is equal to the left hand side since the object has constant comoving co-ordinates. If Δt_o and Δt_e are small, then one can write

$$\frac{\Delta t_e}{a(t_e)} \simeq \frac{\Delta t_o}{a(t_o)} \tag{1.18}$$

and because $\Delta t_o / \Delta t_e = \lambda_o / \lambda_e = 1 + z$, this can be written

$$1 + z = \frac{a(t_o)}{a(t_e)}. \tag{1.19}$$

This is believed to be the main contribution to redshifts measured in practice; ie. that due to the recession of objects from us owing to Universal expansion. Superimposed on top of these cosmological redshifts are perturbations due 'peculiar velocities' which result from the action of local gravitational fields on objects. Gravitational fields also cause redshifts in another, more direct way. Einstein's general theory of relativity describes how light undergoes a change in wavelength if observed at a location at a different gravitational potential than the location it started from. If this change in gravitational potential is $\Delta\phi$, then the redshift is

$$z = \frac{\Delta\phi}{c^2}. \tag{1.20}$$

This effect becomes more important at larger distances. An appreciation of this comes from realising that from the point of view of a photon travelling towards us, we sit at the centre of a sphere containing mass and hence at the bottom of a potential well. Ignoring the mass outside this sphere means that the measured redshift of the photon is lessened by the increase in energy it experiences in falling into the potential well (see Bondi 1947).

1.1.5 The Density Parameter

The Friedmann equation (1.7) can be written such that ρ allows for all contributions to the density, in which case the term involving the cosmological constant no longer explicitly appears:

$$\frac{\dot{a}^2}{a^2} = \frac{8\pi G\rho}{3} - \frac{Kc^2}{a^2}. \tag{1.21}$$

This is often referred to as the energy equation of the Universe since, broadly speaking, it describes the relationship between the Universe's kinetic and potential energy. Using

the result of equation (1.11) this becomes

$$\frac{Kc^2}{a^2} = H^2\left(\frac{\rho}{\rho_c} - 1\right) \tag{1.22}$$

where,

$$\rho_c = \frac{3H^2}{8\pi G}. \tag{1.23}$$

ρ_c is called the *critical density* of the Universe for reasons now explained.

If $\rho = \rho_c$, then equation (1.22) implies that $K = 0$ and therefore that space is flat (see Section 1.1.2). Ignoring the contribution to ρ from Λ means that the density evolves like $\rho \propto a^{-3}$ (lengths $\propto a^{-1}$) and hence from equation (1.21), $\dot{a} \propto a^{-1}$. This means that the the rate of expansion of the Universe tends asymptotically to zero. If $\rho < \rho_c$ then $K < 0$ and equation (1.21) predicts that $\dot{a}$ is always positive; the Universe expands forever. If $\rho > \rho_c$ then $K > 0$ and the Universe initially expands before collapsing back on itself. Including the effects of Λ confuses matters since its contribution to ρ does not evolve with time.

Clearly, the ratio of ρ/ρ_c determines whether the Universe is closed, open or critical. This is an important ratio and thus is given the status of a fundamental cosmological parameter called the *density parameter*,

$$\Omega = \frac{\rho}{\rho_c} = \frac{8\pi G\rho}{3H^2}. \tag{1.24}$$

Since this quantity depends on H and ρ which both evolve with time, so too does Ω. Its value at the current epoch is denoted as usual with the suffix '0'. Recent estimates show a relatively wide range of values although most suggest a sub-critical universe with $0.2 < \Omega_0 < 1.0$ (eg. Coles & Ellis 1997).

1.1.6 Evolution of Ω and H

Contributions to ρ in equation (1.21) come from matter ($\rho_m \propto a^{-3}$), radiation ($\rho_r \propto a^{-4}$) and Λ (ρ_Λ constant). Since these three contributions evolve differently, their combination to form the overall density is

$$\rho = \rho_\Lambda + \rho_m + \rho_r = \rho_{\Lambda 0} + \rho_{m0}\left(\frac{a_0}{a}\right)^3 + \rho_{r0}\left(\frac{a_0}{a}\right)^4 \tag{1.25}$$

where $a_0 \equiv a(t_0)$. Two manipulations now need to be made in order to determine the evolution of H. The first is brought about by defining

$$\Omega_{[\Lambda 0, m0, r0]} = \frac{8\pi G}{3H_0^2}\rho_{[\Lambda 0, m0, r0]} \tag{1.26}$$

to give from equation (1.25)

$$\frac{8\pi G\rho}{3} = H_0^2\left[\Omega_{\Lambda 0} + \Omega_{m0}(1+z)^3 + \Omega_{r0}(1+z)^4\right]. \tag{1.27}$$

The second involves taking the current epoch version of equation (1.22) and rescaling with the scale factor to give

$$\frac{Kc^2}{a^2} = H^2(\Omega - 1) = H_0^2(\Omega_0 - 1)(1+z)^2. \tag{1.28}$$

The time evolution of the Hubble parameter is then given by substituting equations (1.27) and (1.28) into equation (1.21);

$$H^2 = H_0^2\left[\Omega_{\Lambda 0} + \Omega_{m0}(1+z)^3 + \Omega_{r0}(1+z)^4 + (1-\Omega_0)(1+z)^2\right]. \tag{1.29}$$

This is an important relationship. As the next section shows, equation (1.29) is used to form the connection between distance and redshift.

Evolution of Ω is simply determined by substituting for H in equation (1.28) using the result of equation (1.29);

$$\Omega - 1 = \frac{\Omega_0 - 1}{\Omega_{\Lambda 0}(1+z)^{-2} + \Omega_{m0}(1+z) + \Omega_{r0}(1+z)^2 - \Omega_0 + 1}. \tag{1.30}$$

Inspection shows that as $z \to \infty$, the right hand side of this equation $\to 0$ and so at early times $\Omega \to 1$. If a cosmological constant exists, then at late times when $a \to \infty$ (in which case z effectively $\to -1$), Ω tends towards 1.

1.1.7 The Deceleration Parameter

A useful parameter in cosmology is the deceleration parameter, q, which provides a measure of the rate of change of the Universe's expansion at a given epoch. This is defined in terms of the scale factor as

$$q = -\frac{\ddot{a}a}{\dot{a}^2}. \tag{1.31}$$

From the Friedmann equation (1.8) and the result of equation (1.11), this can be written in terms of the contributions to Ω:

$$q = \frac{4\pi G\rho}{3H^2} + \frac{4\pi Gp}{c^2H^2} - \frac{\Lambda c^2}{3H^2} = \frac{\Omega_m}{2} + \Omega_r - \Omega_\Lambda. \tag{1.32}$$

Determinations of the current epoch value q_0 have mainly come from the use of supernovae as standard candles. The latest estimates seem to indicate that q_0 is negative (eg. Riess et al 1998, Perlmutter et al 1998) which implies an accelerating Universe. Some refinement to these measurements may still be necessary however.

1.1.8 Cosmological Distances and Volumes

A useful definition of distance is essential in any cosmological theory. This is especially true of gravitational lensing which is based on geometrical arguments as Section 2.1 shows. For this reason, a careful consideration of the notion of distance in our Universe is presented here for the particular cases of proper, comoving, luminosity and angular diameter distances.

Proper Distance

The *proper distance* between two points is the distance simultaneously measured by a chain of rulers spanning both points at time t. Taking one point as the origin of the co-ordinate system so that only the r co-ordinate need be considered means that from the RW metric with $\mathrm{d}t = 0$ for simultaneity, the proper distance is

$$d_{pr}(t) = \int_0^r \frac{a(t)\mathrm{d}r}{\sqrt{1-Kr^2}} = a(t)f(r) \tag{1.33}$$

where

$$f(r) \equiv \int_0^r \frac{\mathrm{d}r}{\sqrt{1-Kr^2}} = \begin{cases} \sin^{-1}(r) & \text{for} \quad K = +1 \\ r & \text{for} \quad K = 0 \\ \sinh^{-1}(r) & \text{for} \quad K = -1 \end{cases} . \tag{1.34}$$

To express this in terms of redshift for practicality, the radial part of the RW metric is combined with equations (1.19) and (1.11) to give

$$\frac{\mathrm{d}r}{\sqrt{1-Kr^2}} = \frac{c\,\mathrm{d}z}{a_0H}. \tag{1.35}$$

Substituting for H using equation (1.29) and then integrating shows that

$$a_0 f(r) = \frac{c}{H_0} \int_0^z \left[\Omega_{\Lambda 0} + \Omega_{m0}(1+z)^3 + (1-\Omega_0)(1+z)^2\right]^{-1/2} \mathrm{d}z \equiv \frac{c}{H_0} S \qquad (1.36)$$

which is clearly the proper distance at the current epoch. S is defined for later use. The contribution from Ω_r has been neglected here since it is only significant at very early times when the energy content of the Universe was mainly in the form of radiation. This so called 'radiation dominated era' soon changed to a 'matter dominated era' as the Universe expanded and cooled. This change-over happened at a very high redshift (ie. $z \sim 1000's$), much higher than the redshifts considered in this thesis and so Ω_r is ignored hereafter. The cosmological model which assumes that the Universe contains a critical density of only dust so that $p = \Omega_\Lambda = 0$ and $\Omega_m = 1$ is referred to as the 'Einstein-de-Sitter model'.

The following analytical solution for equation (1.36) exists in the case of $\Omega_\Lambda = 0$:

$$a_0 f(r) = \frac{2c}{H_0} \frac{\Omega_{m0} z + (\Omega_{m0} - 2)(\sqrt{1+\Omega_{m0} z} - 1)}{\Omega_{m0}^2 (1+z)}. \qquad (1.37)$$

This is *Mattig's equation* (Mattig 1958).

Comoving Distance

Using the result of equations (1.28), (1.34) and (1.36), the comoving radial distance D(z) may be written:

$$D(z) = a_0 r = \begin{cases} \frac{c}{H_0\sqrt{\Omega_0 - 1}} \sin\left(S\sqrt{\Omega_0 - 1}\right) & \text{for} \quad \Omega_0 > 1 \\ \frac{c}{H_0} S & \text{for} \quad \Omega_0 = 1 \\ \frac{c}{H_0\sqrt{1-\Omega_0}} \sinh\left(S\sqrt{1-\Omega_0}\right) & \text{for} \quad \Omega_0 < 1 \end{cases} . \qquad (1.38)$$

Note that equation (1.28) ensures a physical solution by forcing the argument of all square roots above to be positive through the value of K.

Luminosity Distance

The proper distance is of little practical interest since it is impossible to simultaneously measure the distance elements between two points. A more useful definition of distance

is that of the *luminosity distance* chosen to obey the Euclidean result that an object's flux scales as the inverse square of its distance from the observer. Denoting L as the power of a source emitted at time t and l as the flux received at time t_0, the luminosity distance is defined as

$$D_l = \left(\frac{L}{4\pi l}\right)^{1/2}. \tag{1.39}$$

If the source lies at a comoving radial co-ordinate r then the surface area of the sphere which passes the observer at time t_0 is $4\pi(a_0 r)^2$. The emitted photons experience the dual effect of being redshifted by the Universe's expansion and having their arrival rate reduced due to time dilation. These two effects combine to give an observed source power of $(a/a_0)^2$ times the power output at time t and hence l in terms of L is

$$l = \frac{L}{4\pi D^2}\left(\frac{a}{a_0}\right)^2. \tag{1.40}$$

with D as defined by equation (1.38). Combining equations (1.39) and (1.40) therefore gives the luminosity distance as

$$D_l = \frac{a_0}{a}D(z) = D(z)(1+z). \tag{1.41}$$

Angular Diameter Distance

Instead of the requirement that fluxes adhere to the inverse square law, an alternative distance can be defined which preserves the Euclidean behaviour that an object's angular size varies as the inverse of its separation from an observer. Such a distance is called an *angular diameter distance.*

Suppose an object lying at the comoving radial co-ordinate r at time t has a proper diameter of D_{pr}. If the angle subtended by D_{pr} in the θ direction is $\Delta\theta$ then the RW metric gives

$$D_{pr} = ar\Delta\theta. \tag{1.42}$$

Using the definition above, the angular diameter distance of the source is then given as

$$D_a = \frac{D_{pr}}{\Delta\theta} = ar = \frac{D(z)}{1+z}. \tag{1.43}$$

Angular diameter distances are the distances used in gravitational lensing for reasons made apparent in Chapter 2. Depending on the cosmological parameters chosen,

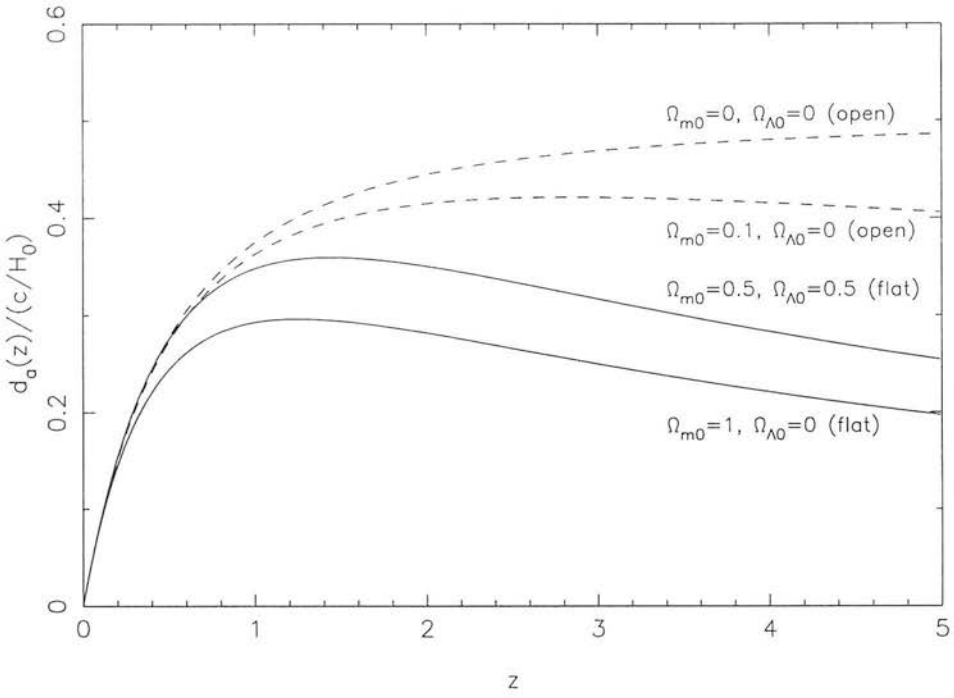

Figure 1.1: Variation of angular diameter distance with redshift for four different cosmological models. Dashed lines indicate an open universe.

angular diameter distances do not necessarily increase with increasing redshift. Figure 1.1 illustrates this for four different cosmological models. In a universe with $\Omega_{m0} = 1, \Omega_{\Lambda 0} = 0$, D_a begins increasing with redshift but then reaches a maximum at $z \simeq 1.2$ before falling off. Equation (1.42) shows that this therefore implies that the angular size of an object can appear to increase with distance beyond $z \simeq 1.2$. This unintuitive behaviour is a result of gravitational focusing of light rays due to the curvature of space induced by the Universe's mass content. Figure 1.1 shows that stronger focussing occurs in a universe with a higher value of Ω_{m0}.

Comoving Volume

Calculation of comoving volumes is necessary for many cosmological applications. One such application is the determination of luminosity functions such as that of Section 5.3.

The comoving volume element of an observed solid angle $\Delta\omega$ is given by

$$\mathrm{d}V = \Delta\omega(a_0 r)^2 \mathrm{d}(a_0 r). \tag{1.44}$$

From equation (1.36), the comoving volume between the redshifts z_1 and z_2 in a matter dominated universe ($\Omega_r = 0$) is therefore,

$$V = \Delta\omega \frac{c}{H_0} \int_{z_1}^{z_2} \frac{D^2(z)\,\mathrm{d}z}{[\Omega_{\Lambda 0} + \Omega_{m0}(1+z)^3 + (1-\Omega_0)(1+z)^2]^{1/2}}. \tag{1.45}$$

1.1.9 The Power Spectrum of Density Fluctuations

The power spectrum is relied upon at several points throughout the work presented here. This section is therefore included entirely for completeness.

The variation in density at a given point $\boldsymbol{x}$ in a volume V is commonly parameterised by the 'density contrast', δ, defined as

$$\delta(\boldsymbol{x}) = \frac{\rho(\boldsymbol{x}) - \overline{\rho}}{\overline{\rho}} \tag{1.46}$$

where ρ is the mass density at $\boldsymbol{x}$ and $\overline{\rho}$ is the average density in V. The *autocorrelation function* of the density contrast is

$$\xi(\boldsymbol{r}) = \langle \delta(\boldsymbol{x})\delta(\boldsymbol{x}+\boldsymbol{r}) \rangle \tag{1.47}$$

where the angular brackets denote the average over all positions in V. For a distribution of discrete objects, $\xi(\boldsymbol{r})$ is interpreted as the excess probability of a randomly chosen object having a neighbour at $\boldsymbol{r}$ from it.

It is beneficial to consider mass fluctuations as a superposition of plane waves since they evolve independently of each other while the fluctuations are still linear. It is therefore advantageous to work in Fourier space. The density contrast can be expressed as a Fourier series,

$$\delta(\boldsymbol{x}) = \sum_{\boldsymbol{k}} \delta_{\boldsymbol{k}} \exp(i\boldsymbol{k}.\boldsymbol{x}) = \sum_{\boldsymbol{k}} \delta^*_{\boldsymbol{k}} \exp(-i\boldsymbol{k}.\boldsymbol{x}) \tag{1.48}$$

with $\boldsymbol{k}$ acquiring only discrete values due to the periodic boundary conditions of the finite volume V. Substituting both normal and complex conjugate versions of δ into equation (1.47) gives

$$\xi(\boldsymbol{r}) = \left\langle \sum_{\boldsymbol{k}} \sum_{\boldsymbol{k}'} \delta_{\boldsymbol{k}} \delta^*_{\boldsymbol{k}'} e^{i(\boldsymbol{k}'-\boldsymbol{k}).\boldsymbol{x}} e^{-i\boldsymbol{k}.\boldsymbol{r}} \right\rangle. \tag{1.49}$$

The fact that

$$< e^{i(\boldsymbol{k}-\boldsymbol{k}').\boldsymbol{x}} >= \delta^D(\boldsymbol{k}-\boldsymbol{k}') \tag{1.50}$$

where δ^D is the Dirac delta function gives, in passing to the limit where $V \to \infty$ so that the summation in (1.49) becomes an integral,

$$\xi(\boldsymbol{r}) = \frac{1}{(2\pi)^3} \int P(k).e^{-i\boldsymbol{k}.\boldsymbol{r}} \mathrm{d}\boldsymbol{k} \tag{1.51}$$

This defines the power spectrum

$$P(k) \equiv \left\langle |\delta_{\boldsymbol{k}}|^2 \right\rangle \tag{1.52}$$

which in taking the inverse Fourier transform of equation (1.51) shows that

$$P(k) = \int \xi(\boldsymbol{r}) e^{i\boldsymbol{k}.\boldsymbol{r}} \mathrm{d}\boldsymbol{r}. \tag{1.53}$$

The power spectrum is evidently the Fourier transform of the autocorrelation function, a result known as the *Wiener-Khintchine theorem*. For convenience, $P(k)$ is typically parameterised as a power law (over certain ranges of k) such that $P(k) \propto k^n$ with the *spectral index* n often a function of k.

1.1.10 Magnitudes & The K-correction

The apparent magnitude of an object is defined as

$$m = c - 2.5 \lg l \tag{1.54}$$

where c is a constant of normalisation and l is the measured flux defined as the power received per unit area in a given bandwidth. Since l is inversely proportional to the square of luminosity distance (Section 1.1.8), one can define an absolute magnitude M as being the apparent magnitude an object would have at a distance of 10pc;

$$m - M = 5 \lg \left[\frac{D_l(pc)}{10} \right] \tag{1.55}$$

Determination of M for nearby stars and galaxies is a straightforward process once m and D_l have been measured, however for more distant objects, the effects of redshift must be taken into consideration. There are four ways in which redshift affects the measured flux within a given bandwidth. The first two were mentioned in Section 1.1.8 and these are that time dilation and photon reddening reduce the flux by a factor of $(1 + z)^2$. The third is that an object's spectral energy distribution is

stretched by a factor of $(1+z)$ so that if $L(\lambda)$ is the power output at wavelength λ, $L(\lambda) \rightarrow L(\lambda/(1+z))$. This stretching means that the number of photons per unit wavelength interval is reduced by a factor of $(1+z)$. If photons are selected within a certain bandwidth then the total number received is less by a factor of $(1+z)$ than that which would have been detected in the absence of redshift. This is the fourth effect and is clearly not applicable if the total or *bolometric* flux is considered.

Consider the emission of photons from a source at time t lying at the comoving radial co-ordinate r. From Section 1.1.8, the surface of the sphere described by the photons at time t_0 when they are detected is $4\pi D^2$. Allowing for all four redshift effects above, the measured flux l_m in a waveband described by a filter with transmission $T(\lambda)$ is therefore,

$$l_m = \frac{\int_0^\infty T(\lambda)L(\lambda/(1+z))\mathrm{d}\lambda}{4\pi D^2(1+z)^3} = \frac{\int_0^\infty T(\lambda)L(\lambda/(1+z))\mathrm{d}\lambda}{4\pi D_l^2(1+z)}, \tag{1.56}$$

using the result of equation (1.41). In order to calculate the absolute magnitude of the source in the rest frame, the following flux is required:

$$l_{req} = \frac{\int_0^\infty T(\lambda)L(\lambda)\mathrm{d}\lambda}{4\pi D_l^2}. \tag{1.57}$$

From equation (1.54), the difference between the measured and required absolute magnitude is

$$M_{req} - M_m = 2.5 \lg \left[\frac{l_m}{l_{req}}\right] = K(z). \tag{1.58}$$

This correction is referred to as the *K-correction*, usually written as

$$K(z) = 2.5 \lg \left[\frac{\int_0^\infty T(\lambda)L(\lambda/(1+z))\mathrm{d}\lambda}{\int_0^\infty T(\lambda)L(\lambda)\mathrm{d}\lambda}\right] - 2.5 \lg(1+z). \tag{1.59}$$

Equation (1.59) shows that both the redshift and the form of $L(\lambda)$ must be known in order to correctly calculate an object's $K(z)$. In many cases, the form of $L(\lambda)$ is unknown however as Section 5.3.1 discusses, a good approximation is that $K \propto z$.

1.2 Dark Matter

Determination of the matter content of the Universe is vital for the evaluation of its contribution to Ω_0. One only has to look as far as the planets to conclude that not

all mass is luminous. Studies must therefore rely upon indirect means of detecting mass, such as its gravitational influence on luminous mass or, as the main theme of this thesis, light itself.

1.2.1 Evidence of Dark Matter

From the careful observation of stars in the solar neighbourhood in the 1920's, Jan Oort predicted that more mass than that visible was required to explain the large stellar motions perpendicular to the galactic plane. At the time, this 'missing mass' was attributed to the distribution of interstellar gas and dust (although doubts still remain as to whether this completely explains the phenomenon). Using a similar approach except with galaxies (see Section 1.4.1), Zwicky (1933) calculated a *mass to light ratio* (M/L) of $\sim 50h\mathrm{M}_{\odot}/\mathrm{L}_{\odot}$ for the Coma cluster. Given that galaxies were then thought to consist mainly of stars with a M/L of $\sim 1h\mathrm{M}_{\odot}/\mathrm{L}_{\odot}$, this implied that the Coma cluster must be about 50 times the summed mass of the galaxies themselves.

Zwicky's result is often regarded as the first evidence of dark matter. Although on a smaller scale, a later measurement of the rotation curve of our neighbouring galaxy M31 by Rubin & Ford (1970) supported this result. They found that the orbital velocity of stars at a large radius did not fall off as $r^{-1/2}$, but remained constant out to the largest radius they were able to observe. This was confirmed in other spiral galaxies and thus immediately gave rise to the notion that spirals must be embedded in a dark matter halo.

Modern observations seem to indicate that M/L ratios and hence the abundance of dark matter, increase with increasing object scale. For example, recent measurements of the Coma cluster give a M/L of close to $400h\mathrm{M}_{\odot}/\mathrm{L}_{\odot}$ in comparison to our local group with a M/L of $\sim 100h\mathrm{M}_{\odot}/\mathrm{L}_{\odot}$ (Binney & Tremaine 1987). On the scale of galaxies, typical estimates for spirals lie at M/L$\sim 30h\mathrm{M}_{\odot}/\mathrm{L}_{\odot}$. The orbital motion of stars in ellipticals is somewhat less ordered than that in spirals and so masses must be determined from the dispersion of measured galaxy velocities. This gives a M/L for the central regions in ellipticals of $\sim 10h\mathrm{M}_{\odot}/\mathrm{L}_{\odot}$.

1.2.2 Contributions to Ω_0

The critical M/L of the Universe (ie. that required to give $\Omega_0 = 1$ from mass contributions only) can be deduced from knowledge of the local luminosity density, j_0. Denoting the M/L of the local Universe as Υ enables the mass density to be written as

$$\rho_{m0} = j_0 \Upsilon. \tag{1.60}$$

From the definition of the density parameter with $\Omega_0 = 1$, this gives a critical M/L of

$$\Upsilon_c = \frac{\rho_c}{j_0}. \tag{1.61}$$

Loveday et al (1992) measure a luminosity density from nearby galaxies of $j_0 = (1.2 \pm 0.3) \times 10^8 h \mathrm{L}_\odot \mathrm{Mpc}^{-3}$ which therefore means that,

$$\Upsilon_c = (2200 \pm 500) h \mathrm{M}_\odot / \mathrm{L}_\odot. \tag{1.62}$$

With a typical M/L of $\sim 400 h \mathrm{M}_\odot / \mathrm{L}_\odot$, this rough calculation shows that rich clusters are insufficiently massive to provide a critical density. Assuming for an instant that the M/L of rich clusters is a fair representation of the M/L of the Universe[5] means that matter makes a contribution of $\Omega_{m0} = 400/2200 \simeq 0.2$ to the total density parameter.

1.2.3 The Nature of Dark Matter

Comparison of the observed abundance of light elements (Deuterium, Helium, Lithium) with the predictions made from the theory of big bang nucleosynthesis (Alpher et al 1948) constrain the amount of baryonic material in the Universe to be

$$0.011 h^{-2} < \Omega_{baryon} < 0.015 h^{-2}. \tag{1.63}$$

(Coles & Lucchin 1997). Given the abundance of matter from observations of rich clusters, this clearly implies the existence of non-baryonic material.

[5] One could argue that given the observed trend indicating an increasing M/L on larger scales, this is actually an underestimation.

There are currently several postulates as to the nature of this non-baryonic material and these fall into one of two categories:

- **Hot Dark Matter (HDM)** The term 'hot' refers to the fact that the particles which make up the dark matter are relativistic and therefore have a substantial pressure. Their velocities, being close to the speed of light, prevent them from clumping together and they are expected to be distributed smoothly throughout the Universe. The gravitational effect of this smooth distribution is to smooth out baryonic fluctuations by pulling them apart. This has the consequence of a 'top-down' formation history in which the largest structures formed first. The most common candidate particle for HDM is the massive neutrino with a mass of ~ 10 eV.

- **Cold Dark Matter (CDM)** This type of matter is referred to as 'cold' since its constituent particles are non-relativistic. Their relatively low speeds allow them to coagulate to form large amounts of small scale structure. This gives rise to a 'bottom-up' formation scenario in which larger structures form later. A typical example of a CDM particle is the axion, not present in the standard model of particle physics.

Determinations of the power spectrum of density fluctuations in the Universe (see Section 1.1.9) indicate that neither type of dark matter completely explains the observed fluctuations on all scales. HDM fails to predict a large enough power on small scales whereas CDM has trouble accounting for the observed large scale fluctuations. This has led to models which involve a mixture of hot and cold matter types, often called 'warm dark matter' models (see for example Borgani et al 1996). Observations of large scale structure such as galaxy clusters can constrain this mixture (Kofman et al 1996). A more detailed discussion of clusters now follows.

1.3 Galaxy Clusters

Galaxies are not randomly distributed across the sky. Their positions are correlated and there exist regions where their surface number density is noticeably higher than

average. This is clearly seen in the APM galaxy survey (Maddox et al 1990) of ~ 3 million optically selected galaxies shown in Figure 1.2. Such groups consist of a wide ranging number of galaxies from a few up to anywhere from a hundred to a thousand closely separated members forming *galaxy clusters*, seen here as the darkest patches.

Figure 1.2: The APM survey of ~ 3 million optically selected galaxies contained within an area of ~ 4300 square degrees centred on the south galactic pole (Maddox et al 1990). Empty patches occur where foreground stars, dwarf galaxies and globular clusters are masked out. A darker shading represents a higher number density. Observations were carried out in the blue band in the magnitude range $17 < b_j < 20.5$.

The most prominent galaxy cluster seen is the Virgo cluster. Its central region has a diameter of $\sim 7^\circ$ with its main body extending over an area of roughly $15^\circ \times 40^\circ$. Even in as early as the 18th century, Sir William Herschel noted that 1/3rd of the galaxies observed at the time were contained in the Virgo cluster despite it only covering 1/8th of the sky.

1.3.1 Properties of Clusters

Having established that a cluster is some sort of galactic congregation, the question, 'How is a galaxy cluster defined ?' immediately arises. The answer is that there is not a

standard definition although criteria are usually based upon similar characteristics. In terms of physical properties, a rough guide is that clusters are gravitationally bound, virialised systems with a mass of $\sim 10^{14}\mathrm{M}_{\odot}$ within a region of $1h^{-1}\mathrm{Mpc}$ in radius. In terms of their identification from optical observations, clusters are typically defined as being regions where the number density of galaxies within a particular aperture exceeds some threshold value above the mean.

A widely cited example of a cluster catalogue is that of Abell (1958). The Abell catalogue contains 4076 clusters, 2712 of which are north of $-20°$ declination and away from the galactic plane. Abell defined clusters by three criteria:

- The cluster must contain at least 50 galaxies within a magnitude range $m_3 \leq m \leq m_3 + 2$, where m_3 is the magnitude of the third brightest galaxy in the cluster.
- These galaxies must all lie within a radius of $R_a = 1.5h^{-1}\mathrm{Mpc}$, R_a being referred to as the Abell radius.
- The estimated cluster redshift (see Section 1.1.4) must be within a range of $0.02 \leq z \leq 0.2$.

Clusters were also assigned a 'richness' depending on the number of galaxies the cluster contained. The poorest clusters were assigned a richness of 0 with richer clusters being assigned higher numbers. A similar, yet more relaxed set of criteria were used by Zwicky et al (1961-1968) to identify 9134 clusters in the same area of sky as Abell, although this catalogue is not as widely used.

In the late 1960's when X-ray telescopes became available, it was discovered that clusters are powerful emitters of X-ray radiation. Typical X-ray luminosities are found to lie between $10^{43} - 10^{45}$ erg s^{-1} making them the most luminous X-ray sources in the sky (see the review by Sarazin 1986). As the resolution of telescopes improved, it became apparent that X-ray emission from clusters is extended rather than point-like and that the spectra are best described by the bremsstrahlung process[6] from a

[6]Bremsstrahlung radiation occurs when energetic electrons are decelerated in passing close to atomic nuclei.

hot, dilute plasma typically $10^7 - 10^8$K in temperature with a density of $\sim 10^{-3}$ particles per cm^3. Making the assumption that this intracluster gas is in hydrostatic equilibrium with the gravitational potential of the cluster enables estimation of the total mass (see Section 1.4.2). Mass estimations made in this way agree reasonably well (ie. within a factor of ~ 2, see Lubin & Bahcall 1993) with those from masses derived from galaxy velocity dispersions discussed in Section 1.4.1.

1.3.2 The Cosmological Importance of Clusters

Estimation of the cosmological parameters that best describe our Universe is a fundamental challenge to modern observational cosmology. Galaxy clusters provide not only an extremely useful tool for the study of these parameters, but also a unique means of attempting to understand the formation of large scale structure. Their importance stems largely from their youth, having formed relatively late in the formation history of the Universe.

Cluster Formation Rates as a Constraint on Ω

As Richstone et al (1992) noted, the rate of formation of clusters heavily depends on the density of the Universe. In a low density universe, the expectation is that cluster formation and indeed growth of structure begins earlier than in a high density universe. This has the dual consequence that in a low density universe, a higher abundance of clusters is anticipated and that their present day rate of formation should be lower.

Observed abundances of clusters, or more specifically, knowledge of the cluster mass function[7] can therefore provide a strong constraint on the matter density of our Universe (eg. Fan et al 1997, Viana & Liddle 1996, Eke et al 1996). Accurate measurement of cluster masses thus plays a crucial role in the application of these abundance arguments. For example, Eke et al (1996) predict that the number of clusters per comoving volume with a mass greater than $3.5 \times 10^{14} h^{-1} \mathrm{M}_\odot$ in a universe with $\Omega_0 = 0.3$ should be 15 times the number in a universe with $\Omega_0 = 1$.

[7] ie. the comoving number density of clusters of a given mass

The current belief concerning the formation of structure in the Universe is that tiny perturbations in the density field of the early Universe gave rise to gravitational instabilities causing those perturbations to grow. These small perturbations continued growing and coalescing to form clouds of matter which condensed to form stars in turn forming galaxies and the larger structures observed today[8]. In this so called 'hierarchical formation model', clusters are expected to form at the highest peaks of the early density field. Their large separations of typically 10-30 times that of galaxy separations make them prime candidates for tracing large scale structure since a relatively small sample can be used to probe large distances. This combines with the fact that the positions of clusters in the present day Universe are probably close to their initial formation positions enabling application of linear structure formation theory. The findings of Watkins (1997), who measured the peculiar velocities of several clusters to be typically 500 kms^{-1} or less, supports this view.

Mapping the distribution of matter contained within a cluster, such as that considered in Section 3.2.6 also lends valuable information to theories of structure formation. The ability to measure a cluster's dark matter content (see Section 1.2) enables investigation into the phenomenon of 'biassing', that is, how stringently luminous mass traces the distribution of dark matter (see Mann et al 1998 and references contained therein). Detection of substructures in clusters gives an indication of their age (Richstone et al 1992). The argument used here is that once a cluster has collapsed, low contrast substructures of at least $\sim 20\%$ of the total cluster mass with a density comparable to that of the cluster will be mixed away in a time scale of ~ 1 dynamical time[9]. The existence of substructure in a cluster therefore suggests that a cluster has recently formed. Measurement of the fraction of clusters which exhibit such structures can therefore uncover their formation history and put constraints on cosmological models.

[8]This is in contrast to a fragmentation model in which the largest structures form first.

[9]A dynamical time is defined as being one quarter of an object's orbital period about some central mass (Binney & Tremaine 1987).

Normalisation of the Power Spectrum from Cluster Abundance

A great volume of work has focussed on using the abundance of clusters to constrain the power spectrum of density fluctuations in the Universe (see Section 1.1.9). Various observations of large scale structure give the shape of the power spectrum on different scales (eg. Landy et al 1996, Peacock & Dodds 1994) but its amplitude is difficult to obtain (Tadros et al 1999 show one particular means of achieving this). Clusters provide an ideal means of normalising the power spectrum as the following discussion shows.

The theory of inflation predicts that large scale structure seen today was seeded by quantum fluctuations in the density field of the very early Universe. Quantum theory states that these fluctuations are Gaussian-distributed which leads to the conclusion that fluctuations in the density field measured today on a large enough scale should also be distributed in a Gaussian fashion. The probability distribution function for the density contrast, δ (see Section 1.1.9), averaged over some volume with a characteristic scale R on large scales can therefore be written as,

$$p(\delta, R) = \frac{1}{\sqrt{2\pi}\,\sigma_R} \exp\left(-\frac{\delta^2}{2\sigma_R^2}\right). \tag{1.64}$$

This becomes a poor approximation on smaller scales where the distribution of mass is less smooth causing the distribution to be skewed to smaller values of δ. Of course, the scale at which the Gaussian nature of the distribution becomes inappropriate increases as one looks further back into the Universe by virtue of the simple fact that matter was more homogeneously distributed at earlier times. The variance σ_R^2 is determined by windowing the power spectrum with $W_R^2(k)$; the Fourier transform of the volume within which δ is averaged in equation (1.64),

$$\sigma_R^2 = \int \mathrm{d}^3k P(k) W_R^2(k). \tag{1.65}$$

Clearly, the value of σ_R^2 is directly proportional to the power spectrum $P(k)$. If σ_R^2 can therefore be measured via an independent method then the amplitude of $P(k)$ is known.

Consider the gravitational collapse of a spherically over-dense region to form a rich cluster. In an expanding universe, this collapse can only occur if the average density

contrast in that spherical region is greater than a critical value of $\delta_c = 1.68$ (eg. White et al 1993). According to the Press-Schechter formalism of structure formation (Press & Schechter 1974), the fraction of mass in the Universe contained in objects characterised by their *linear* scale R is

$$F_c = \int_{\delta_c}^{\infty} \mathrm{d}\delta p(\delta, R) \tag{1.66}$$

with p given by equation (1.64). From the measurement of cluster masses and the number density of clusters, F_c can be independently determined from

$$F_c = \frac{M_c n_c}{\overline{\rho}}, \tag{1.67}$$

where M_c is the mass of a typical cluster characterised by its linear scale R, n_c is the measured spatial number density of clusters and $\overline{\rho}$ is the average density of the Universe. Equating equations (1.66) and (1.67) shows that σ_R can be determined in terms of terms of $\overline{\rho}$ if the quantities M_c and n_c are known. Accurate cluster mass and abundance determinations are therefore important to constrain the amplitude of $P(k)$ through equation (1.64).

The fact that R in the Press-Schechter formalism is the linear scale means that clusters probe scales comparable to their initial collapse size. In the early Universe when matter was much more homogeneously distributed, this means that a typical cluster must have had to collect material from a much larger volume. Making the assumption that such a cluster collapsed from a spherical volume of homogeneous material to form a mass typically seen today of say $\sim 6 \times 10^{14} h^{-1} \mathrm{M}_{\odot}$ (eg. White et al 1993) within an Abell radius gives R;

$$\begin{aligned} M_c &= \frac{4\pi}{3} R^3 \rho_{cr} \Omega_{m0} \sim 6 \times 10^{14} h^{-1} \mathrm{M}_{\odot} \\ \Rightarrow R &\simeq 8 h^{-1} \Omega_{m0}^{1/3} \mathrm{Mpc}. \end{aligned} \tag{1.68}$$

where ρ_{cr} is the critical density of the Universe parameterised by the mass density parameter Ω_{m0} (see Section 1.1.5). Evidently, clusters therefore provide a means of normalising the power spectrum on a scale of 8 Mpc via the rms of density fluctuations on this scale, σ_8. Fan et al (1997) show that if the power law $\sigma_R \propto M^{-\alpha}$ is adopted, then the integrated cluster mass function $n(>M)$ behaves like

$$\ln n(>M) \propto \sigma_8^{-2} M^{-2\alpha}. \tag{1.69}$$

Expressed more qualitatively, this states that determination of σ_8 is far more reliant upon accurate knowledge of mass than the number density of clusters and thus provides further motivation for the work carried out in this thesis.

As a quick exercise, a value for n_c can be readily estimated from Abell's catalogue. The redshift range within which all Abell clusters lie is $0.02 \leq z \leq 0.2$ encompassing a volume of $\sim 3.5 \times 10^8 h^{-3}\mathrm{Mpc}^3$ (Bartelmann 1996). With such a small upper redshift limit, this volume is relatively insensitive to cosmological parameters (see Section 1.1.8). Of the 2712 Abell clusters, there are 1894 with a richness of ≥ 1. Considering only these, this yields a value of

$$n_c = \frac{1894}{3.5 \times 10^8} h^3 \mathrm{Mpc}^{-3} \simeq 5 \times 10^{-6} h^3 \mathrm{Mpc}^{-3}. \tag{1.70}$$

Using the value of M_c in equation (1.68) allows the observed cluster mass fraction to be calculated from equation (1.67) to give

$$F_c \simeq 8 \times 10^{-3} \Omega_{m0}^{-1} \tag{1.71}$$

This shows that clusters are rare objects in a universe where $\Omega_0 \simeq 1$. Equating this observed fraction with that from equation (1.66) therefore gives that

$$\sigma_8 = \begin{cases} 0.68 & (\Omega_{m0} = 1) \\ 1.38 & (\Omega_{m0} = 0.2) \end{cases} . \tag{1.72}$$

This rough calculation is consistent with the more detailed analysis considered in Viana & Liddle (1996), Eke et al (1996) and Pen (1998).

1.4 Cluster Masses from Virial and X-Ray Temperature Measurements

Section 1.3.2 explained the importance of clusters as large scale structures and hence justified the motivation behind measurement of their mass. This section discusses the two main methods of obtaining cluster mass other than that of lensing which is postponed for a more detailed discussion in Chapter 2.

1.4.1 Virial Masses

The virial theorem asserts that for a system of self-gravitating masses to be in dynamical equilibrium, the time average of twice the total kinetic energy of all the masses with respect to the centre of mass of the system must equal the time average of the negative of the gravitational potential energy for the system (eg. Tremaine & Lee 1987). The dynamical time of a cluster of galaxies is far too large for a time averaged kinetic and potential energy to be calculated and thus galaxy motions can only ever be observed at a particular epoch. The virial theorem must therefore be more correctly written,

$$2T + U \simeq 0 \tag{1.73}$$

where T and U are the kinetic and potential energy respectively. This becomes less of an approximation when more galaxies are involved in the calculation (Limber & Mathews 1960).

Making the assumption that clusters are in a state of dynamical equilibrium enables computation of their mass through measurement of the dispersion of galaxy velocities. The total kinetic and total potential energy can be written as

$$\begin{aligned} T &= \frac{1}{2}\sum_{i=1}^{N} m_i v_i^2 \\ U &= -\frac{G}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\frac{m_i m_j}{r_{ij}}, \end{aligned} \tag{1.74}$$

where the ith galaxy has mass m_i, velocity v_i with respect to the cluster centre of mass and lies a distance r_{ij} from galaxy j. The double summation in the relation for the potential energy here acts over all galaxy pair combinations. Defining the mass weighted velocity dispersion as

$$\sigma_v^2 = < v^2 > = \frac{\sum_{i=1}^{N} m_i v_i^2}{M}, \tag{1.75}$$

where M is the total cluster mass in galaxies shows that the virial theorem in equation (1.73) may be written as

$$M = \frac{R_g \sigma_v^2}{G}. \tag{1.76}$$

Here, the gravitational radius R_g has been defined as

$$R_g = 2M^2 \left(\sum_{i=1}^{N}\sum_{j\neq i}^{N} m_i m_j / r_{ij} \right)^{-1}. \tag{1.77}$$

In practice, observations can only directly allow projected velocities and galaxy separations to be measured. To take this into account, the velocity dispersion in equation (1.76) must be replaced with the dispersion of projected velocities, σ_{vr} and the inter-galaxy separation with its projected quantity r'_{ij}. Limber and Mathews (1960) showed that assuming spherical symmetry,

$$\begin{aligned} T &= \frac{3}{2}\sum_{i=1}^{N} m_i v_{ri}^2 \\ U &= -\frac{G}{\pi}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\frac{m_i m_j}{r'_{ij}}, \end{aligned} \tag{1.78}$$

so that equation (1.76) becomes

$$M = \frac{3\pi R'_g \sigma_{vr}^2}{2G} \tag{1.79}$$

with R'_g given by equation (1.77) evaluated with the projected inter-galaxy separation.

Equation (1.79) therefore gives the virial mass of the cluster in galaxies once σ_{vr} and R'_g are known. σ_{vr} is directly measurable. Studies show (eg. Sarazin 1986) that the distribution of velocities in a typical cluster is roughly Gaussian. R'_g however, requires knowledge of the galaxy masses (or at least their relative mass). One means of calculating R'_g is to assume that a cluster galaxy's luminosity is proportional to its mass. However, studies of galactic rotation curves do not show a strong correlation between luminosity and mass (eg. Rubin et al 1982). An even simpler alternative is to assume that all galaxies have the same mass (Giradi et al 1997, Heisler et al 1985) so that,

$$R'_g \simeq 2N^2 \left(\sum_{i=1}^{N}\sum_{j\neq i}^{N} 1/r'_{ij}\right)^{-1}. \tag{1.80}$$

Through N-body simulations, Heisler et al (1985) have shown that this approximation continues to give reasonably accurate cluster mass estimations when galaxy masses are dispersed according to realistic mass distributions.

The reliability of the virial theorem was tested by Bahcall & Tremaine (1981) who found it to be both inefficient and biassed towards under-predicting masses of gravitationally bound systems by factors of up to 10 or more. As spherical symmetry is assumed in the allowance for projection effects, their findings indicated that virial based mass determinations become less accurate, the more elliptical galaxy orbits

become. As an alternative, they conceived the 'projected mass estimator' (see Heisler et al 1985 for details) which was found to give more accurate mass measurements;

$$M = \frac{32}{\pi G} < v_r^2 r' > \tag{1.81}$$

where v_r signifies the radial velocity.

The calculations shown in this section, however, assume that a cluster's mass lies solely in its galaxies. The fact that clusters are observed to have a hot inter-galactic plasma, thought to originate from young stars in a cluster's galaxies, shows that this is not the case. Furthermore, gravitational lensing analyses of clusters such as that discussed later in this thesis, show evidence of inter-galactic dark matter. Another limitation to the virial theorem comes from the assumption that clusters are in a state of equilibrium. The existence of substructure, known to be present in most clusters from optical and X-ray measurements (eg. Bird 1993, 1994), has the consequence that equilibrium assumptions are not valid. In fact, investigations show that even low levels of contamination from substructure can cause severe effects on estimations of a cluster's kinematical properties (Bird 1994, Beers et al 1990).

1.4.2 X-Ray Temperature Masses

As Section 1.3.1 noted, clusters typically harbour a hot and yet dilute inter-galactic plasma, thought to originate as ejected gas from stars in their galaxies. This gas is most likely heated by supernovae to high temperatures which results in the emission of X-rays due to the bremsstrahlung process. Assuming that this gas is in hydrostatic equilibrium so that the inward pull on the gas due to its self-gravity is exactly balanced by the force due to outward pressure, p, gives for the radial gradient of this pressure (eg. Binney & Tremaine 1987),

$$\frac{\mathrm{d}p}{\mathrm{d}r} = -\frac{GM(r)\rho(r)}{r^2}. \tag{1.82}$$

where $M(r)$ is the mass of gas contained within a sphere of radius r and ρ is the gas density. Using the ideal gas law,

$$p = \frac{\rho k_B T}{m} \tag{1.83}$$

with T and m being the gas temperature and molecular mass respectively allows equation (1.82) to be written as

$$M(r) = -\frac{k_B T r}{G\overline{m}} \left(\frac{\mathrm{d} \ln \rho}{\mathrm{d} \ln r} + \frac{\mathrm{d} \ln T}{\mathrm{d} \ln r} \right). \tag{1.84}$$

Here, $\overline{m}$ is the mean molecular mass of the gas. The X-ray temperature T is measured by fitting the observed spectrum to a bremsstrahlung spectrum. Its radial profile is particularly difficult to measure with the limited resolution of current X-ray detectors[10] although evidence suggests that cluster cores are approximately isothermal. This allows the temperature gradient term in equation (1.84) to be omitted for the determination of core masses.

The next step usually taken is to assume a radial *mass* and *gas* density profile, ρ_m and ρ respectively, in accordance with the King (1966) model for a self gravitating isothermal sphere,

$$\begin{aligned} \rho_m(r) &\propto \left(1 + \left(\frac{r}{r_c}\right)^2\right)^{-3/2} \\ \rho(r) &\propto \left(1 + \left(\frac{r}{r_c}\right)^2\right)^{-3\beta/2} \end{aligned} . \tag{1.85}$$

The cluster core radius r_c and the parameter β are obtained by fitting the observed flux profile to (Sarazin 1986),

$$S_X(r) \propto \left(1 + \left(\frac{r}{r_c}\right)^2\right)^{-3\beta+1/2} \tag{1.86}$$

which is expected if the gas density follows the form in equation (1.85). Physically, β corresponds to the ratio of the specific kinetic energy of the total mass to the specific kinetic energy of the gas in a cluster (Cavaliere & Fusco-Femiano 1976),

$$\beta = \frac{\overline{m}\sigma_v^2}{k_B T} \tag{1.87}$$

where σ_v is the velocity dispersion of the previous section. Values of β obtained by fitting the flux profile in this way are typically $\sim 2/3$ (eg. Lewis et al 1999, Edge & Stewart 1991, Jones & Forman 1984) which implies that the gas is more energetic than the matter in clusters. This is in disagreement with determinations of β calculated

[10] This should be possible however with the *XMM* space-borne X-ray telescope due to be launched at the end of 1999

directly from equation (1.87) using the measured temperature and velocity dispersion. The problem is referred to as the β *discrepancy* since for equilibrium conditions to hold true, a value of $\beta = 1$ should be observed. One simple suggestion by Bahcall & Lubin (1994) is that the discrepancy is due to an inadequately assumed mass profile.

Making the assumption that isothermal conditions do hold true in the cluster gives the mass profile from substitution for ρ in equation (1.85) into equation (1.84),

$$M(r) = \frac{3\beta k_B T}{\overline{m}G} \frac{r^3/r_c^2}{1 + r^2/r_c^2}. \tag{1.88}$$

1.4.3 X-Ray, Virial & Lensing Mass Comparisons

The details of measuring cluster mass by exploiting the phenomenon of gravitational lensing are left until Chapter 2. For the purposes of comparison however, a discussion of the results of such evaluations is presented in this section.

The determination of cluster mass using virial and X-ray methods, as the previous sections have shown, make assumptions based on spherical symmetry and equilibrium conditions within clusters. Studies have shown not only that substructure is present within most observed clusters, but that this substructure can give rise to severe distortions in a cluster's assumed kinematical properties (Bird 1994, Beers et al 1990). X-ray measurements are also limited to probing only relatively small central regions of clusters which host the X-ray emitting gas.

Another complication is that of *cooling flows* which act to destroy hydrostatic equilibrium. The rate of gas cooling through the bremsstrahlung process in a cluster is proportional to the local electron density (Binney & Tremaine 1987). This means that gas can cool quicker than it is heated at small radii where the electron density is higher and hence flow into the cluster centre. This flow is usually sufficiently slow in comparison to the sound speed that hydrostatic equilibrium is approximately maintained, however in some circumstances it may not be (Allen 1998). Clusters which exhibit cooling flows however are generally virialized systems. Compensating for the effects of cooling flows often means that reasonable agreement between X-ray and lensing mass can be acheived.

The technique of measuring cluster mass with gravitational lensing has the advantage that it assumes nothing about the internal conditions of a cluster. As Chapter 2 explains, masses are inferred from the amount of deflection a light ray experiences in passing a cluster; a measure that is independent of its internal motions. That is not to say that lensing goes completely without its own set of assumptions however. In order to calculate the size of light deflections, one typically measures distributions of the ellipticities or the flux of galaxies lying behind a lens. Assumptions therefore have to be formed about the intrinsic distribution of these properties. In measuring ellipticities, the assumption usually taken is that the average intrinsic ellipticity of galaxies over a large enough sample is zero whereas the distribution of intrinsic galaxy fluxes must be determined by other means. Unlike the assumptions associated with X-ray and virial mass measurements, those required by lensing are readily verifiable by observing areas of sky in which there is a lack (or at least a negligible amount) of lensing. Furthermore, any deviations from or uncertainties in these assumptions can be unambiguously quantified and encompassed in an error analysis of the mass result.

Studies into the comparison of cluster mass determination methods show that in general, X-ray and virial mass measurements tend to give consistent results (eg. Lewis et al 1999, who found that $M_{vir}/M_{x-ray} = 0.96 \pm 0.07$ with a sample of 14 clusters). This is reflected in findings which confirm the behaviour that

$$\sigma \propto \sqrt{T} \tag{1.89}$$

as predicted by equation (1.87). Alternatively, lensing tends to give larger masses than those due to X-ray or virial predictions on average. Consistency could be argued, however, given the relatively large uncertainties involved. For example, the analysis of ~ 30 clusters by Wu & Fang (1997) found that

$$M_{lens} = (1.42 \pm 0.99) M_{vir} = (2.23 \pm 1.15) \beta_{fit}^{-1} M_{x-ray}. \tag{1.90}$$

where β_{fit} is the value of β obtained by fitting the X-ray flux profile in equation (1.86). A later result by the same author (Wu et al 1998) made the conclusion that in fact mass calculations from weak lensing are in agreement with X-ray and virial masses whereas strong lensing tends to predict masses larger by factors of $\sim 2-4$ (the distinction between weak and strong lensing is clarified in Section 2.1.5).

Such arguments of course depend on the specific structure of the cluster under scrutiny. A cluster composed of two masses lying along the line of sight for example, would be expected to give a larger virial mass than the lensing mass which would be in turn expected to be larger than the X-ray temperature mass due to projection effects (Reblinsky & Bartelmann 1999, Giradi et al 1997, Miralda-Escudé & Babul 1995). Other differences might arise from the possibility of clusters being supported by a non-thermal pressure such as that due to magnetic fields (Ensslin et al 1998, Loeb & Mao 1994).

The debate as to which is the most effective means of estimating cluster masses is an ongoing one. The method provided by gravitational lensing is the most recent, motivated strongly by modern developments in telescope and computational technology to enable both observation of deeper, clearer images and more complex image analysis. As the work of Reblinsky & Bartelmann (1999) using N-body simulations distinctly demonstrates, gravitational lensing predicts cluster masses far more accurately than those from virial estimates. Opinions seem to be converging on the fact that given its non-parametric approach and limited, quantifiable assumptions, gravitational lensing is the strongest contender for the accurate measurement of cluster mass.

1.5 Thesis Layout

The primary aim of this PhD has been to develop and apply mass reconstruction techniques to lensing clusters. In particular, this work has concentrated on the exploitation of the lens magnification effect described in Chapter 2, first discussed as a serious contender for mass reconstruction by Broadhurst et al (1995).

The work presented here naturally falls into three separate areas. Following on from Chapter 2 which details the theory of gravitational lensing and current mass reconstruction techniques, Chapter 3 sees the application of one of these methods. In this chapter, the gravitational effect of the cluster Abell 1689 on the number density of observed background galaxies is investigated and a mass measurement is made (Taylor & Dye et al 1998). Chapter 4 looks at a new approach to mass reconstruction. This new method, which relies upon pixellization of the observed field of view, allows very

accurate, self-consistent mass determinations (Dye & Taylor 1998). Finally, Chapter 5 studies the lens induced effect on the luminosity function of galaxies lying behind Abell 1689. In collaboration with the Max Planck Institute for Astronomy in Heidelberg, this work utilises the powerful technique of photometrically evaluating object redshifts to provide an independent mass measure of the cluster (Dye et al 1999).

Chapter 2

Mass Reconstruction Theory

Before beginning a detailed account of the lens mass reconstruction techniques used in this thesis, an appreciation of the underlying theory of gravitational lensing is required. This chapter therefore lays down the key elements of lensing theory in the first half as a foundation for the mass reconstruction methods covered in the latter half.

The latter half is broken down into two further sections to coincide with the fact that reconstruction techniques are generally divided into two categories; those using shear information and those using magnification data. Section 2.2 gives an account of two different shear methods. Although the second half of Chapter 4 uses measurements of shear, Section 2.2 is presented more for reasons of completeness than because it sees application in this thesis.

Section 2.3 contains a more detailed discussion concerning mass reconstruction using magnification data since this is the main theme throughout the PhD work presented here. Two independent ways of measuring magnification are given in this section, each of which are applied in the separate Chapters 3 and 5.

2.1 Lensing Theory

Interest in lensing dates back as far as 1919 when Eddington (1919) first confirmed Einstein's prediction that the gravitational influence of matter deflects light. By measuring the positions of stars close to the Sun during the eclipse of 1919, Eddington showed that its mass caused a deflection of ~ 1 arcsec. The first bona fide gravitational lens was discovered by Walsh et al (1979) who showed from spectral measurements that the quasar Q0957+561 exhibited a double image. Since then, a multitude of lensed systems have been discovered (eg. see the review by Narayan 1998), demonstrating the use of gravitational lensing as an important and highly versatile cosmological probe.

2.1.1 Lensing by Point Masses

The propagation of light in arbitrary curved spacetimes is a complicated theoretical problem. Fortunately, for most cases of astrophysical interest, it can be assumed that the geometry of the Universe is closely described by the Robertson-Walker metric and that matter inhomogeneities which cause lensing are no more than local perturbations. It therefore follows that in the majority of lensing scenarios studied, the path of light from a distant source is well approximated by unperturbed source-lens and lens-observer photon paths with the deflection occurring in the vicinity of the lens.

This is depicted for the ideal case of a point mass in Figure 2.1 which shows a photon path subject to a deflection with an overall bend angle $\hat{\alpha}$. This bend angle is calculated by evaluating the integral of the acceleration perpendicular to the photon path given by the perpendicular gradient of the lens gravitational potential Φ:

$$\hat{\alpha} = \frac{2}{c^2} \int \nabla_{\perp} \Phi \mathrm{d}l \tag{2.1}$$

The integral here acts along the unperturbed photon path, an approximation called the *Born Approximation*. With $\Phi = GM/|\boldsymbol{x}|$, $\boldsymbol{x}$ being the vector from the point mass to a position on the photon path, the result of equation (2.1) is (Schneider 1992),

$$\hat{\alpha} = \frac{4GM}{c^2 \xi} = \frac{2\mathrm{R_s}}{\xi} \tag{2.2}$$

where M is the mass of the point mass and ξ is the *impact parameter* defined in Figure 2.1. Note that the deflection angle is twice the inverse of the impact parameter in units

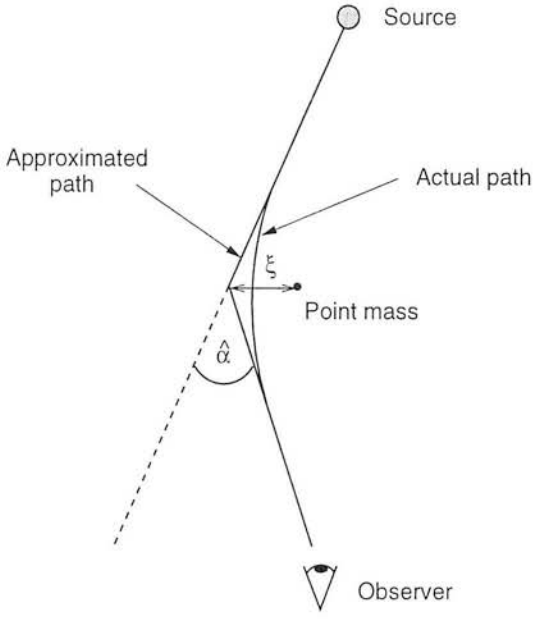

Figure 2.1: Deflection of light through angle $\hat{\alpha}$ by a point mass.

of the Schwartzschild radius R_s. Equation (2.2) holds true provided $\xi \gg R_s$. This result from general relativity, predicting a deflection angle twice that derived from Newtonian considerations, was verified by Eddington during the 1919 solar eclipse.

2.1.2 The Thin Screen Approximation

If light deflection occurs over a distance which is short in comparison to the total length of the photon path then the deflecting mass distribution is said to be a *thin lens*. In such circumstances, the geometrical depth of the mass distribution will be a small fraction of the observer-source distance. This allows application of the *thin screen approximation* in which the mass is considered to effectively lie in a sheet, orthogonal to the line of sight.

Thick lenses prove to be much more complicated to analyse although can often be separated into two or more thin lenses. Further complications arise in practice due to multiple deflections along the line of sight. Fortunately, evidence such as the results of studies into the lensing of the CMBR (eg. Bernardeau 1998) show that this is not a major concern when measuring the lensing effects of massive objects such as clusters. The large scale deflection caused by a typical rich cluster on the whole, completely dominates any small scale deflections due to intervening mass.

The relatively compact nature of most clusters means that the thin screen approximation may be applied. This is advantageous since a simple generalisation of the deflection angle due to a point mass can be made. Projecting the volume mass density ρ onto an orthogonal sheet which lies in the *lens plane* gives the projected surface mass density Σ at a point $\boldsymbol{\xi}$ in the lens plane as

$$\Sigma(\boldsymbol{\xi}) = \int \rho(\boldsymbol{\xi}, z)\mathrm{d}z \tag{2.3}$$

where z is a co-ordinate along the line of sight. The deflection angle at $\boldsymbol{\xi}$ now becomes the sum of all deflections due to individual mass elements in the plane, namely,

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2} \int \frac{\Sigma(\boldsymbol{\xi}')(\boldsymbol{\xi}' - \boldsymbol{\xi})}{|\boldsymbol{\xi}' - \boldsymbol{\xi}|^2} \mathrm{d}^2\xi' \tag{2.4}$$

which is a two component vector.

2.1.3 The Lens Equation

The formal description of a gravitationally lensed system is based on simple geometry. Consider Figure 2.2 which illustrates an observer O at a distance $\mathrm{D_d}$ from the lens plane L which in turn lies a distance $\mathrm{D_{ds}}$ from a source S. Defining the line $\mathrm{OS_0}$ as the optical axis, the position of the source in the source plane is described by the vector $\boldsymbol{\eta}$ from the optical axis. Similarly, the position of the intersection of the light from the source with the lens plane is given by the vector $\boldsymbol{\xi}$. It is assumed, as before, that $\hat{\boldsymbol{\alpha}}$ is the deflection seen at the lens plane by the light ray.

The geometry in Figure 2.2 shows that the vector $\boldsymbol{\xi}$ can be expressed as,

$$\boldsymbol{\xi} = \boldsymbol{\eta} + (\hat{\boldsymbol{\alpha}} - \boldsymbol{\theta})\mathrm{D_{ds}}. \tag{2.5}$$

Noting that $\boldsymbol{\theta} = \boldsymbol{\xi}/\mathrm{D_d}$, equation (2.5) becomes,

$$\boldsymbol{\eta} = \frac{\mathrm{D_s}}{\mathrm{D_d}}\boldsymbol{\xi} - \mathrm{D_{ds}}\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}). \tag{2.6}$$

While equation (2.6) has been derived using Euclidean geometry, it can also hold true in curved spacetimes. This is the case if the distances used in its derivation satisfy the Euclidean result that two objects with a physical separation s subtending an observed angle θ are a distance $s\theta$ away. Distances which adhere to this condition

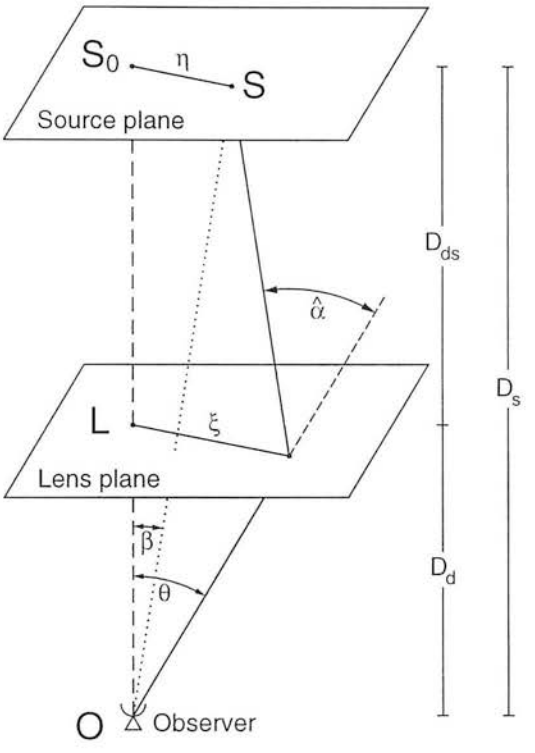

Figure 2.2: A typical lensing geometry with the observer at O, source at S and lens at L. The path taken by the light ray is shown by the solid line which undergoes deflection $\hat{\boldsymbol{\alpha}}$.

are angular diameter distances (see Section 1.1.8). Of course this result uses the approximation that $\sin\theta \simeq \theta$ for small θ. In a typical lensing scenario with a lens of diameter ~ 5Mpc at a redshift of $z = 0.2$, this works out to be in error by less than one part in 10^6.

Simplification to equation (2.6) can be made if the physical lengths $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are replaced by the angular variables,

$$\boldsymbol{\beta} = \frac{\boldsymbol{\eta}}{\mathrm{D_s}} \quad , \quad \boldsymbol{\theta} = \frac{\boldsymbol{\xi}}{\mathrm{D_d}} \tag{2.7}$$

so that the deflection angle may be replaced by its scaled counterpart,

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{\mathrm{D_{ds}}}{\mathrm{D_s}}\hat{\boldsymbol{\alpha}}(\mathrm{D_d}\boldsymbol{\theta}) = \frac{1}{\pi}\int_{\mathrm{lens}} \mathrm{d}^2\theta' \kappa(\boldsymbol{\theta}') \frac{\boldsymbol{\theta} - \boldsymbol{\theta}'}{|\boldsymbol{\theta} - \boldsymbol{\theta}'|^2} \,. \tag{2.8}$$

The dimensionless surface mass density or *convergence*, κ, is given by,

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(\mathrm{D_d}\boldsymbol{\theta})}{\Sigma_{\mathrm{CR}}}, \quad \Sigma_{\mathrm{CR}} = \frac{c^2}{4\pi G}\frac{\mathrm{D_s}}{\mathrm{D_{ds}}\mathrm{D_d}} \tag{2.9}$$

which defines the critical surface mass density $\Sigma_{\rm CR}$. In a very qualitative sense, $\Sigma_{\rm CR}$ is a characteristic value which distinguishes strong lensing from weak lensing. Although complicated by the effects of lens shear, it is generally true that surface mass densities with $\Sigma \simeq \Sigma_{\rm CR}$ demonstrate strong lensing behaviour. The term 'strong lensing' here and indeed throughout the rest of this thesis is taken to describe the lensing which occurs in the vicinity of a critical line (see Section 2.1.6).

Re-writing equation (2.6) in terms of the scaled deflection angle and the quantities $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ gives,

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}). \tag{2.10}$$

This important equation is commonly referred to as the *lens equation*. It describes the mapping between positions of objects in the source plane and observed image positions in the lens plane. If the deflection angle is small, an invertible, one-to-one mapping between the lens and source co-ordinates exists in which case the lensing is weak. Strong lensing occurs when the deflection angle is larger and a non-unique mapping results in the formation of multiple images.

2.1.4 Deflection Potential

Using the identity $\boldsymbol{\nabla} \ln|\mathbf{x}| = \mathbf{x}/|\mathbf{x}|^2$ enables equation (2.8) to be written in terms of the deflection potential ψ, such that

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \boldsymbol{\nabla}\psi(\boldsymbol{\theta}), \quad \psi(\boldsymbol{\theta}) = \frac{1}{\pi}\int_{\rm lens} d^2\theta' \kappa(\boldsymbol{\theta}') \ln|\boldsymbol{\theta} - \boldsymbol{\theta}'| \tag{2.11}$$

Furthermore, application of the identity

$$\nabla^2 \ln|\mathbf{x}| = 2\pi\delta^2(\mathbf{x}) \tag{2.12}$$

to equation (2.11) shows that ψ and κ are related via Poisson's equation:

$$\nabla^2\psi = 2\kappa \tag{2.13}$$

For this reason, lensing is often regarded as a two dimensional potential theory.

2.1.5 Image Magnification and Distortion

Section 2.1.3 made the distinction that strong lensing occurs in the near vicinity of critical lines. Strong lensing is responsible for the arcs and multiple images observed in many lensed systems. Less dramatic image distortions form due to weak lensing effects. Weak shear introduces a bias into the distribution of image ellipticities and orientations which can be calculated in the statistical mean (see Section 2.2.2). Image distortion is an inevitable consequence of lensing and one which plays a vital role in all lens-mass reconstruction techniques.

Image distortion can be quantified by the Jacobian matrix $\mathcal{A}$ of the mapping obtained from the lens equation,

$$\mathcal{A} \equiv \frac{\partial \beta_i}{\partial \theta_j} = A_{ij} = \delta_{ij} - \psi_{ij} \tag{2.14}$$

where $\psi_{ij} \equiv \partial^2\psi/\partial\theta_i\partial\theta_j$ and the result of equation (2.11) has been implemented. Since the Laplacian of ψ from equation (2.13) is twice the convergence, κ in terms of ψ is

$$\kappa = \frac{1}{2}(\psi_{11} + \psi_{22}). \tag{2.15}$$

There are two other linear combinations of ψ_{ij} which are important. These are the components of the *shear tensor*:

$$\gamma_1 = \frac{1}{2}(\psi_{11} - \psi_{22}), \qquad \gamma_2 = \psi_{12} = \psi_{21}. \tag{2.16}$$

The Jacobian matrix in equation (2.14) can be consequently written as

$$\mathcal{A} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix} \tag{2.17}$$

Realizing that the above matrix is a rotation from a diagonal frame, the effect of this mapping on the image of a circular source may be investigated. The rules governing the transformation of cartesian tensors (eg. Goodbody 1982) allow $\mathcal{A}$ to be written as a rotation of this diagonal frame through an angle ϕ:

$$\begin{aligned} \mathcal{A} &= \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} 1 - \kappa - \gamma & 0 \\ 0 & 1 - \kappa + \gamma \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \\ &= (1 - \kappa) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \gamma \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix}. \end{aligned} \tag{2.18}$$

The meaning of shear and convergence is now made intuitively clear. Convergence alone causes an isotropic focusing of light rays leading to an isotropic magnification of the source. Shear introduces anisotropy into the mapping with magnitude $\gamma = ({\gamma_1}^2 + {\gamma_2}^2)^{1/2}$ causing the image of a circular source to be distorted into an ellipse with orientation ϕ. The value 2ϕ appears as a trigonometric argument due to the fact that the two orthogonal states of the shear, whose magnitudes are γ_1 and γ_2, are separated by an angle of 45 degrees (see Figure 2.3). This is a direct consequence of the quadrupolar nature of gravitational fields. The elongation, ϵ, of this ellipse is given by the ratio of eigenvalues of $\mathcal{A}$:

$$\epsilon = \frac{\text{major axis length}}{\text{minor axis length}} = \frac{1-\kappa+\gamma}{1-\kappa-\gamma}. \tag{2.19}$$

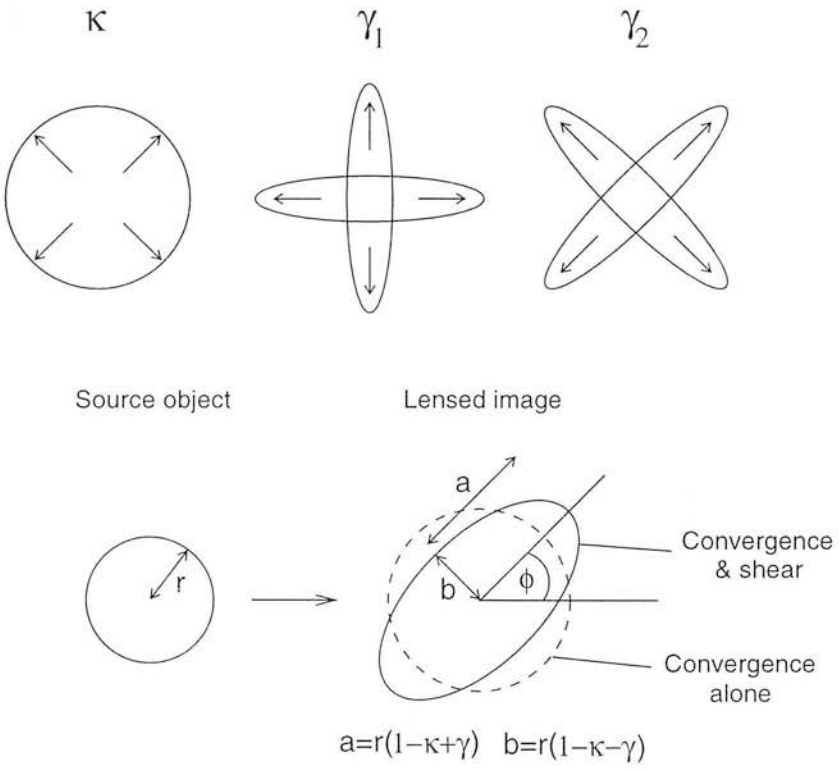

Figure 2.3: The effect of convergence and shear on the lensed image of an intrinsically circular object. *Top*: Separated contributions from κ and the components of shear to the distortion. *Bottom*: The combined effect of κ and γ.

The overall magnification, μ, resulting from this mapping is given by the inverse of the determinant of the Jacobian matrix which from equation (2.17) is

$$\mu = \frac{1}{|(1-\kappa)^2 - \gamma^2|}. \tag{2.20}$$

Elliptical distortion of a circular image holds true for any type of lens mass distribution apart from a *sheet mass* distribution in which all material lies in a sheet of constant surface density. Since κ in this case is therefore constant, it must be true that $\gamma = 0$ from equations (2.15) and (2.16) and hence the image of a circularly shaped background galaxy would remain circular. The size of such a galaxy (its solid angle subtended on the sky), however, would be magnified by a factor of $(1 - \kappa)^{-2}$ as seen from equation (2.20).

Surface Brightness Conservation

The fact that gravitational lensing is an achromatic process and that photons are neither created nor destroyed signifies that the surface brightness of a lensed object remains unchanged. The surface brightness, S, of an object relates to its flux f as

$$S = \frac{f}{\omega} \tag{2.21}$$

where ω is the solid angle it subtends on the sky. It therefore follows that in a lensing process in which the solid angle of an object is increased by a factor μ, the observed flux of the object must also be increased by the same factor to keep the surface brightness constant. Although at first sight this appears to defy energy conservation, integration of the flux measured at all points on a sphere encompassing the source and lens would give exactly the same value as if no lens were present. Flux magnification is an important result of lensing and forms a key role in the work outlined in this thesis.

2.1.6 Caustics and Critical Lines

The magnification equation (2.20) predicts that infinite magnification results if the condition $1 - \kappa = \pm\gamma$ holds true for any given point in the lens plane. Whilst this is true in theory, real magnifications remain finite because real sources are extended and the magnification averaged over an extended source is always finite. Infinite magnification never exists in practice because the geometrical arguments used in deriving the magnification equation break down in this limit. One therefore has to employ the results of wave optics which always yields finite magnifications.

The closed curves lying in the lens plane defined by $|\mathcal{A}| = 0$ are known as *critical lines*. These curves may also be mapped back onto the source plane via the lens equation to form another set of curves referred to as *caustics*. Source objects lying close to caustics generally have greatly magnified images. Sources which lie within the region bounded by caustics are observed to have multiple images. In the hypothetical scenario of an extended source moving across a caustic towards the centre of a lens, three stages are observed: 1) The source brightens as it approaches the caustic. 2) As the source crosses the caustic, another image appears at some point on the critical line. 3) The newly formed second image splits into two images, one of opposite parity, as the source continues moving and the original image of the source dims considerably (see Schneider, Ehlers & Falco 1992 for further discussion). This process forms the basis of the 'odd number theorem' which states that the number of images formed by a lens is always odd, provided that the surface mass density is bounded and non-singular so that the deflection angle is continuous and does not diverge. An odd number of images is not seen in reality if two images are merged into one or if one of the images is dimmed to make it undetectable.

Figure 2.4 demonstrates the odd number theorem graphically for an arbitrary circularly symmetric lens. In this graph, the deflection angle $\alpha(\theta)$ is plotted against the radial co-ordinate θ from the lens centre. The lens equation stipulates that images form when $\theta - \beta = \alpha(\theta)$. Values of θ at intersections of α with the straight line $\theta - \beta$ therefore correspond to the radial co-ordinates of the image positions. The mass distribution in Figure 2.4 shows that a source at the radial co-ordinate β in the source plane has three images at the radii θ_1, θ_2 and θ_3. This source lies inside the critical lines at $\theta = \theta_c$.

2.1.7 Einstein Radius

One final noteworthy definition is that of the *Einstein radius*. Consider a circularly symmetric lens with an arbitrary mass profile $M(\theta)$. The lens equation gives

$$\beta = \theta - \frac{\mathrm{D_{ds}}}{\mathrm{D_d D_s}} \frac{4GM(\theta)}{c^2\theta}. \tag{2.22}$$

Due to the rotation symmetry of such a lens system, a source lying exactly on the

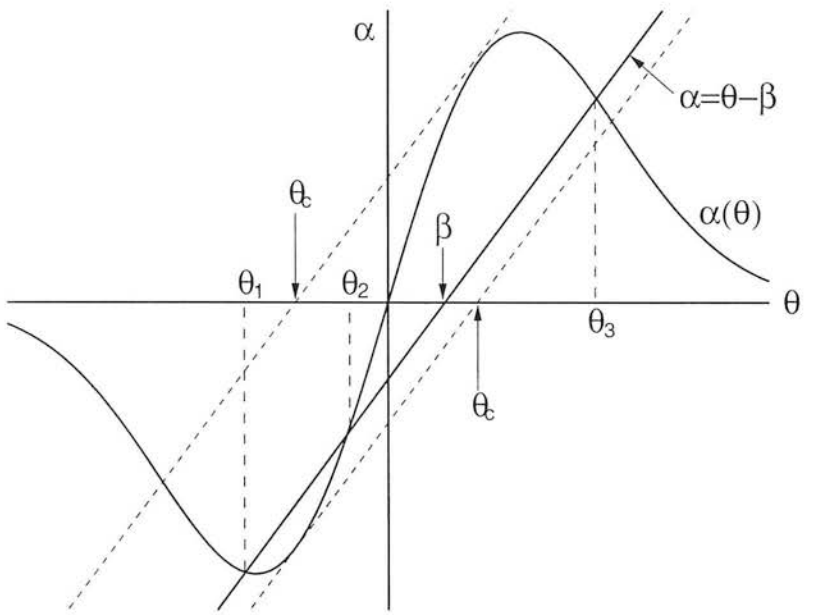

Figure 2.4: The three intersections at θ_1, θ_2 and θ_3 between the curve $\alpha(\theta)$ and the straight line $\theta - \beta$ correspond to the radii of images resulting from a source placed at a radius β. Sources placed beyond the critical lines at θ_c are not multiply imaged as the straight line $\theta - \beta$ makes only one intersection with α.

optical axis (ie. $\beta = 0$) is imaged as a ring centred on $\theta = 0$ if the lens is supercritical (ie. $\Sigma > \Sigma_{\rm CR}$ somewhere in the lens). Setting $\beta = 0$ in equation (2.22) gives the radius of the ring:

$$\theta_{\rm E} = \left(\frac{4GM(\theta_{\rm E})}{c^2} \frac{\rm D_{ds}}{\rm D_d D_s} \right)^{1/2}. \tag{2.23}$$

This is the Einstein radius of the lens which scales as $M^{1/2}$.

With the necessary aspects of the fundamental theory of lensing now detailed, the following section explains their use in the development of mass reconstruction techniques.

2.2 Lens Mass Reconstruction using Image Shear

Observational attempts to obtain κ and hence lens mass can be crudely divided into two categories: those utilizing shear information and those directly measuring magnification. Although not strictly independent of each other (as they are both directly

obtained from the deflection potential discussed in Section 2.1.5), they are commonly treated so in practice. The information obtained from each is generally complementary and recent work has seen a hybridisation of both in an attempt to improve mass estimates (see for example Bartelmann et al 1996).

The following discussion and indeed that of the next chapter refers in particular to the mass reconstruction of clusters. These represent the largest known gravitationally bound objects in the Universe and therefore ease the practical difficulties associated with detecting mass in this way. This short section specifically details mass reconstruction using lens induced image distortions. Reconstruction from magnification is left until the following section.

2.2.1 Mass Contained within a Giant Arc

A simple means of estimating a lower bound on the mass of a supercritical cluster is provided by calculating the projected mass within a circle traced out by a giant arc. Giant arcs form when a source lies very close to the optical axis of the lens system. The Einstein radius of the system is therefore approximately equal to the radius of the arc. By combining equations (2.9) and (2.23), the mass contained within one Einstein radius and hence a good approximation to the mass contained within the arc is given by

$$M(\theta_{\rm E}) = \frac{{\rm D_d D_s}}{{\rm D_{ds}}} \frac{c^2}{4G} \theta_{\rm E}^2 = \pi (\theta_{\rm E} {\rm D_d})^2 \Sigma_{\rm CR}. \tag{2.24}$$

Expressed more qualitatively, the average surface density within a radius of $\theta_{\rm E}$ from the cluster centre is simply equal to the critical surface density. If the redshift of the cluster and the source object are known then $M(\theta_{\rm E})$ can be evaluated. This technique of course is only applicable to supercritical clusters with nearly circularly symmetric mass distributions. Although knowledge of the mass is limited to the inner arc region, Fort and Mellier (1994) note that this mass shows a robust agreement with the mass calculated from the cluster's internal velocity dispersion.

2.2.2 Shear Analysis

Lens shear introduces a bias to the distribution of orientations and ellipticities of background galaxy images. The intrinsic variation of individual galaxy orientations and ellipticities produces a random element which can be statistically removed by averaging over a large enough sample. Fortunately, the scale over which typical cluster lensing potentials vary is much larger than the average separations of background galaxies readily observed and so this is quite easily achieved.

The shear field is related to the potential field through equation (2.16). Using this and equation (2.11) shows that the relationship between κ and γ is:

$$\begin{aligned}\gamma_1(\boldsymbol{\theta}) &= \frac{1}{\pi}\int_{\text{lens}} \mathrm{d}^2\theta' \kappa(\boldsymbol{\theta}') \frac{(\theta_y-\theta'_y)^2-(\theta_x-\theta'_x)^2}{[(\theta_x-\theta'_x)^2+(\theta_y-\theta'_y)^2]^2} \\ \gamma_2(\boldsymbol{\theta}) &= \frac{2}{\pi}\int_{\text{lens}} \mathrm{d}^2\theta' \kappa(\boldsymbol{\theta}') \frac{(\theta_y-\theta'_y)(\theta_x-\theta'_x)^2}{[(\theta_x-\theta'_x)^2+(\theta_y-\theta'_y)^2]^2}\end{aligned} \tag{2.25}$$

where θ_x and θ_y are the x and y components of $\boldsymbol{\theta}$ respectively. Defining a complex shear,

$$\gamma_c(\boldsymbol{\theta}) = \gamma_1(\boldsymbol{\theta}) + \mathrm{i}\gamma_2(\boldsymbol{\theta}) \tag{2.26}$$

and the function

$$\mathcal{D}(\boldsymbol{\theta}) = \frac{\theta_y^2-\theta_x^2-2\mathrm{i}\theta_x\theta_y}{|\boldsymbol{\theta}|^4} \tag{2.27}$$

shows that the relationship between convergence and shear is a convolution:

$$\gamma_c(\boldsymbol{\theta}) = \frac{1}{\pi}\int_{\text{lens}} \mathrm{d}^2\theta' \mathcal{D}(\boldsymbol{\theta}-\boldsymbol{\theta}')\kappa(\boldsymbol{\theta}') \tag{2.28}$$

This may be inverted via, for example, Fourier methods to give (Kaiser & Squires 1993)

$$\kappa(\boldsymbol{\theta}) = \frac{1}{\pi}\int_{\text{lens}} \mathrm{d}^2\theta' \mathrm{Re}[\mathcal{D}^*(\boldsymbol{\theta}-\boldsymbol{\theta}')\gamma_c(\boldsymbol{\theta}')] + \kappa_0 \tag{2.29}$$

where Re[ ] takes the real part of γ_c and the complex conjugate $\mathcal{D}^*$. The constant κ_0 illustrates the important effect that using shear, only differential lens mass may be determined. In other words, κ can only be calculated up to an overall constant. This phenomenon is commonly referred to as the *sheet mass degeneracy*. It is a direct result of the fact stated in Section 2.1.5 that homogeneous sheets of mass produce no shear. As will be discussed in later Sections concerning magnification analysis, this problem

can be circumvented to allow an absolute measure of κ and hence determination of κ_0.

Equation (2.29) forms the crux of what has become known as the 'Kaiser-Squires' mass reconstruction method. In order to proceed, the shear field must be measured. Several elaborate techniques exist to do this (Bonnet & Mellier 1995, Kaiser, Squires & Broadhurst 1995). Essentially, the background galaxies in an image are gridded into bins and their average ellipticity is calculated for each bin. Using equation (2.19), a quantity g called the *reduced shear* can be defined,

$$< g >= \left\langle \frac{\epsilon - 1}{\epsilon + 1} \right\rangle = \left\langle \frac{\gamma}{1 - \kappa} \right\rangle \tag{2.30}$$

where $<>$ denotes taking the average in a particular bin. The fact that only this ratio of $\gamma/(1-\kappa)$ can be obtained from measuring ellipticities re-iterates the point that the method is insensitive to isotropic image expansions. This is seen by multiplying the Jacobian in equation (2.17) by some constant λ so that the resulting transformation in κ and γ is

$$1 - \kappa \rightarrow \lambda(1 - \kappa), \quad \gamma \rightarrow \lambda\gamma \tag{2.31}$$

which has no effect on the measured quantity in equation (2.30). Ideally of course, the direct shear is required if the Kaiser-Squires method is to be used. In the weak lensing limit such that $\kappa \ll 1$, equation (2.30) shows that the observed, reduced shear is approximately equal to the direct shear required in equation (2.29).

In general, γ in equation (2.29) can be replaced by the term $g(1-\kappa)$ which then yields an iterative equation for κ. As shown in Seitz and Schneider (1995), this integral equation can be solved in a few iteration steps. With this correction taken into account, κ is then no longer only determined up to an overall constant as in equation (2.29) but becomes subject to the degeneracy shown in equation (2.31). Although in theory λ can take on any value, it is restricted in practice to ensure that κ is always positive. This similarly provides a lower limit on the mass.

2.3 Lens Mass Reconstruction using Magnification

The use of flux magnification as a tool for the reconstruction of lens mass forms the underlying theme of this thesis. First suggested by Broadhurst, Taylor and Peacock (1995, BTP hereafter) in the weak lensing regime, they showed that the sheet mass degeneracy can be broken by measuring the lensed distortion of the joint magnitude-redshift distribution of background galaxies. Clustering of these background galaxies introduces correlations in redshift space, unlike magnitude space. Separation of this joint distribution into a redshift and a magnitude distribution is therefore necessary for a treatment by maximum likelihood analysis. In light of this, the section presented here distinguishes mass reconstruction using redshift from that using magnitude.

The effect of magnification to cause a geometrical amplification may also provide a means of mass reconstruction. Although not considered in this thesis, one possibility involves measurement of the shift in the distribution of galaxy sizes. The amplification of inter-source spacing forms the second possibility in which a dilution of the projected background number density can be directly related to the mass. In all cases of magnification analysis, measurements must be compared with the observed properties of unlensed field galaxies.

Section 2.3.1 quantifies the effect of lens magnification on the redshift distribution of background source galaxies. This is not directly applied in this thesis but leads to the integrated number count technique of Section 2.3.2 which is applied in Chapter 3. Magnification of galaxy magnitudes, or more specifically their luminosity function, is presented in Section 2.3.3. This is applied in Chapter 5. Finally, several methods to convert magnification to κ are covered in Section 2.3.4.

2.3.1 Reconstruction from Lensed Redshift Distributions

Representing the observed number density of galaxies in the magnitude-redshift space interval $\mathrm{d}m\mathrm{d}z$ by the bivariate distribution $N(m, z)$, the effect of placing a lens in front of such a distribution is twofold. Firstly, flux magnification produces a translation in the magnitude distribution (see Section 2.1.5). Secondly, a dilution of the surface number density of objects at a particular magnitude occurs as a result of the isotropic

expansion of not only image sizes, but also their angular separation. Labelling this translated distribution as $N'(m,z)$, its relation to the untranslated distribution is thus

$$N'(m,z) = \frac{N(m + 2.5 \lg \mu(z), z)}{\mu} \tag{2.32}$$

The observed function $N(m,z)$ is related to the luminosity function $\phi(L,z)$ via

$$\phi(L,z)\mathrm{d}L\mathrm{d}V(z) = N(m,z)\mathrm{d}m\mathrm{d}z \tag{2.33}$$

where L is the luminosity and $\mathrm{d}V(z)$ is the redshift dependent volume interval. Integrating over apparent magnitude, the number density of objects at a particular redshift, $n_0(z)$, is obtained:

$$n_0(z)\mathrm{d}z = \mathrm{d}V(z) \int_{L_{min}}^{\infty} \phi(L,z)\mathrm{d}L = \Phi[L_{min}, z]\mathrm{d}V(z) \tag{2.34}$$

where L_{min} is the lower limit on the luminosity set by the flux limit of the survey and the integral of the luminosity function is denoted by Φ. Lensing effectively reduces L_{min} via flux magnification so that the effect on the redshift distribution through equation (2.32) gives the lensed redshift distribution:

$$n'(z)\mathrm{d}z = \mu^{-1}\Phi[L_{min}/\mu, z]\mathrm{d}V(z). \tag{2.35}$$

Approximating Φ as a power law,

$$\Phi \propto L^{-\beta(z)} \quad , \quad \beta(z) = -\frac{\mathrm{d}\ln\Phi[L,z]}{\mathrm{d}\ln L} \tag{2.36}$$

allows equation (2.35) to be expressed as,

$$n'(z)\mathrm{d}z = \mu^{\beta(z)-1} n_0(z)\mathrm{d}z \tag{2.37}$$

Equation (2.37) demonstrates the fact that the magnification effect is a competitive one. Geometric magnification (ie. amplification of galaxy image size and spacing) competes with flux magnification which results in the detection of extra galaxies as faint objects are made brighter and brought into the survey flux limits. The winning effect is determined by the value of β. Three possibilities exist: 1) $\beta < 1$ in which case the geometric effect dominates and the surface number density of galaxies is reduced relative to the expected number. 2) $\beta = 1$ so that $n' = n_0$ and no lensing effect is seen. 3) $\beta > 1$ in which case the flux magnification effect dominates and more galaxies are observed.

In their analysis, BTP use the R band counts of Metcalf et al (1994) via a model luminosity function to generate the functions

$$\begin{aligned} n_0(z) &= 11.7z^{1.63}\exp[-(z/0.51)^{1.79}] \\ \beta(z) &= 0.15 + 0.6z + 1.1z^{3.2}. \end{aligned} \tag{2.38}$$

Using these functions in conjunction with equation (2.37) for background galaxies shows the effect of lensing on the redshift distribution $n(z)$. Figure 2.5 plots the quantity $\mathrm{dP}(z)/dz = n'(z)/\int_0^\infty n'(z)dz$ for five different magnifications with a lens redshift of $z_l = 0.2$. The plot shows that at small redshifts, angular scattering of images by the lens dominates and the number of objects drops. At large redshifts however, the increase in the total number of observable galaxies dominates. Notice that a node exists where these effects cancel.

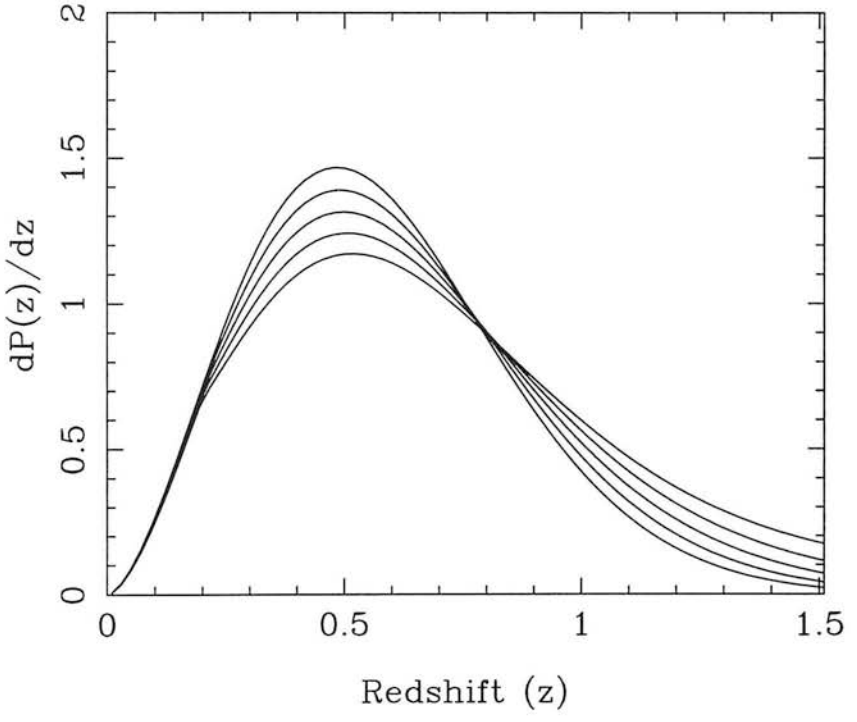

Figure 2.5: The expected lens distortion of a redshift distribution given by (2.38) for a lens redshift of $z_l = 0.2$. Five magnifications are shown ranging from 1 to 2 in steps of 0.25. Higher magnifications produce a deficit of objects at low z and an increase at high z.

Having detailed the effects a lens has on a redshift distribution, the next step is to devise a means of extracting the characteristics of the lens causing such effects. In the considerations taken thus far, several assumptions have been made which must be addressed for practical purposes.

The first assumption is that knowledge of the intrinsic redshift distribution of objects behind the lens has been obtained to enable comparison with the measured lensed distribution. Obviously, the intrinsic background object distribution cannot be measured directly. The Cosmological Principle must therefore be relied upon to assert that the universe is homogeneous and isotropic. This implies that a redshift distribution measured from regions of sky in which no lensing is present will necessarily be the same (once corrections such as our motion with respect to the microwave background are allowed for) as the intrinsic distribution of the objects lying behind the lens. In other words, from a practical point of view, the observed lensed distribution must be compared with the average obtained from a set of unlensed observed areas of sky.

This leads to the second assumption that the effect can be measured and is not overwhelmed by noise. Several sources of noise must be allowed for: 1) Shot noise from the limited number objects available for analysis. 2) Measurement errors on object redshifts. 3) Galaxy clustering. The first two sources of noise rely upon the the depth of the observations made and the accuracy and sensitivity of the instruments used. The third is a cosmological source of uncertainty and can be estimated relatively easily if the clustering power spectrum at the redshift of interest is known. The required rms, σ is then just the following integral in Fourier space (BTP):

$$\sigma^2 = \int \Delta(k)|W(k)|^2 \frac{\mathrm{d}k}{k} \tag{2.39}$$

where $W(k)$ is the Fourier transform of the spatial bin containing the objects and $\Delta(k) = (2\pi^2)^{-1}P(k)k^3$, $P(k)$ being the clustering power spectrum of Section 1.1.9.

Likelihood Analysis

Armed with information about the practicalities of performing such redshift distribution measurements, a method of estimating a lens' surface density can now be tailored. The obvious choice of tactic is the maximum-likelihood method. Dividing the redshift axis up into q bins, the required likelihood function is

$$\mathcal{L}(\mu) \propto \prod_{1}^{q} \mathrm{P}[n|\lambda(\mu, z)] \tag{2.40}$$

where $P[n|\lambda]$ is the probability of measuring n objects given an expected content of λ. An estimate of the lens magnification μ and hence its surface density (see Section 2.3.4 or Chapter 4) is therefore obtained by maximizing $\mathcal{L}$ with respect to μ. A measure of the uncertainty in this maximum is also readily obtained from the width of the likelihood curve between the two points corresponding to

$$\ln \mathcal{L} = \ln \mathcal{L}_{max} - 0.5\Delta\chi^2 \tag{2.41}$$

where $\Delta\chi^2$ is the change in χ^2 appropriate for the desired confidence level (eg. for a single maximum likelihood parameter as is the case here, 1σ errors correspond to $\Delta\chi^2 = 1$). Another estimate of the 1σ errors can be obtained by assuming that the likelihood distribution is Gaussian or at least near-Gaussian so that

$$\sigma_\mu \simeq \left\langle -\frac{\partial^2 \ln \mathcal{L}}{\partial \mu^2} \right\rangle^{-1/2} \tag{2.42}$$

The size of the redshift bins must be chosen so that the statistical independence assumed in equation (2.40) is valid. If it were the case that the redshift distribution of background objects was purely Poissonian, then it would be possible to make the bins infinitesimally small and have only one object per bin. Because of galaxy clustering however, this is not the case. It is therefore necessary to choose a bin width which is larger in radial extent than the coherence length of clustering. Adjacent bins are always subject to the effects of clustering, although this effect is reduced by using larger bin sizes. Very large bins however, produce a large uncertainty in the likelihood analysis and therefore an optimum bin size is one which reduces likelihood errors as much as possible whilst ensuring that the correlation coefficient between adjacent bins is a negligible value.

In order to carry out the likelihood analysis, it is necessary to define the probability $P[n|\lambda(\mu, z)]$ in equation (2.40). A useful model to adopt for this is the lognormal model (see for example, Coles & Jones 1991). Justification of the lognormal model not only comes from observations dating back to Hubble, who realised that the surface density of galaxies was lognormally distributed, but also from the fact that it modifies the canonical Gaussian density distribution to correct the inaccurate assignment of positive probabilities to negative densities.

The lognormal model works by constructing a new density contrast, δ', from a

density contrast, δ, described by the Gaussian distribution,

$$\mathrm{P_G}(\delta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\frac{\delta^2}{\sigma^2}\right) \tag{2.43}$$

such that,

$$1 + \delta' = A \exp\delta \tag{2.44}$$

The normalisation coefficient A is calculated by imposing mass conservation such that,

$$\langle 1 + \delta' \rangle = 1 = A \langle \exp\delta \rangle = A \int_{-\infty}^{\infty} \mathrm{d}\delta\, \mathrm{P_G}(\delta) \exp\delta \tag{2.45}$$

Since

$$\int_{-\infty}^{\infty} \mathrm{d}\delta\, \mathrm{P_G}(\delta) \exp\delta = \exp\left(\frac{\sigma^2}{2}\right) \tag{2.46}$$

the normalisation factor is therefore,

$$A = \exp\left(-\frac{\sigma^2}{2}\right) \tag{2.47}$$

and hence the lognormal density perturbation is given by,

$$1 + \delta' = \exp(\delta - \sigma^2/2). \tag{2.48}$$

To incorporate finite numbers of galaxies, the assumption is made that the observed number of galaxies, n, is Poisson distributed with an expected number, $\lambda(\mu, z)$ which is in turn subject to the lognormal fluctuations described. The probability $\mathrm{P}[n|\lambda(\mu, z)]$ required by the likelihood analysis is then expressed in terms of the Gaussian and Poisson distributions by the compound distribution,

$$\mathrm{P}[n|\lambda(\mu, z)] = \int_{-\infty}^{\infty} \mathrm{d}\delta\, \mathrm{P_P}[n|\lambda \exp(\delta - \sigma^2/2)]\, \mathrm{P_G}(\delta) \tag{2.49}$$

where the Poisson distribution is,

$$\mathrm{P_P}[n|\lambda] = \frac{e^{-\lambda}\lambda^n}{n!}. \tag{2.50}$$

The expected number of objects in each bin, λ, is the number anticipated given a particular magnification. From equation (2.37) this is therefore simply

$$\lambda = n'(z)\Delta z = \lambda_0 \mu^{\beta(z)-1} \exp\left(\frac{\sigma^2}{2}\right). \tag{2.51}$$

Here, Δz is the redshift bin width and the normalisation factor appears explicitly for simplification of equation (2.49) upon substitution. The quantity λ_0 is therefore defined as

$$\lambda_0 = n_0(z) \exp\left(-\frac{\sigma^2}{2}\right) \Delta z \tag{2.52}$$

so that the complete expression for $\mathrm{P}[n|\lambda(\mu, z)]$ is,

$$\mathrm{P}[n|\lambda(\mu, z)] = \frac{[\lambda_0 \mu^{\beta(z)-1}]^n}{n!} \int_{-\infty}^{\infty} \mathrm{d}\delta \exp\left[-\frac{\delta^2}{2\sigma^2(z)} - \lambda_0 \mu^{\beta(z)-1} e^{\delta} + \delta n\right] \tag{2.53}$$

The integral in this expression must be evaluated numerically. Once a suitable model or fit for $n_0(z)$, $\beta(z)$ and $\sigma(z)$ has been achieved for the unlensed galaxy population, equation (2.53) can be used with equation (2.40) to find the maximum-likelihood magnification. The surface mass density may then be derived by using one of the methods detailed in Section 2.3.4 or by the pixellized reconstruction method described in Chapter 4.

2.3.2 Integrated Number Counts

In practice, it is not always possible or feasible to measure a sufficient number of redshifts for galaxies behind a lensing cluster in order to obtain a reliable $n'(z)$ required by equation (2.37). Observations will only typically give the integrated number of objects visible behind the lens in which case modifications to the analysis detailed so far must be made.

The modification comes in the form of an approximation. By observing objects in more than one optical pass-band, it is possible to apply colour selection criteria to choose objects known to lie beyond a particular redshift (see chapter 3). The majority of objects selected in this fashion will also be subject to an upper redshift limit owing to the flux limit of the observation. If the redshift limits are chosen in such a way to ensure that $\beta(z)$ varies slowly over this interval, integration of equation (2.37) over z can be performed by assuming β is constant to give,

$$n' = \mu^{\beta-1} n_0. \tag{2.54}$$

These integrated number densities are subsequently used to replace their redshift dependent versions in equation (2.53) for the likelihood determination of μ.

Observations show that the value of β is dependent on the colour classification of galaxies observed (Broadhurst 1995). As equation (2.54) shows, β controls whether the number of objects observed behind a lens is enhanced or decreased relative to the intrinsic number expected. From observations of far field galaxies in the V and I passbands, Broadhurst notes that the luminosity function of objects selected with the criterion $V - I < 1.0$ yields a value of $\beta \simeq 1.0$. The magnification effect would therefore not be noticed in such galaxies. Alternatively, Broadhurst finds that faint blue far field galaxies selected by $V - I > 1.5$ give a value of $\beta \simeq 0.4$ and hence would show a clear depletion in numbers, relative to the expected number (see Section 3.2.1).

A prediction of this effect can be made for an isothermal sphere, characterised by

$$\kappa \propto \frac{1}{\theta} \quad , \quad \gamma = \kappa \tag{2.55}$$

(see Appendix A.2) where θ is the radial distance from the centre of the sphere. Substituting this into equation (2.20) gives

$$\mu = \left| 1 - \frac{\theta_c}{\theta} \right|^{-1} \tag{2.56}$$

where θ_c is the radius of the critical line. Using this with equation (2.54) shows how n'/n_0 varies with θ:

$$\frac{n'}{n_0} = \left| 1 - \frac{\theta_c}{\theta} \right|^{1-\beta} \tag{2.57}$$

Counting the number of lensed galaxies in annuli centred on a cluster and comparing with the expected number would therefore show the behaviour described by equation (2.57). Figure 2.6 illustrates this for different values of β.

2.3.3 Reconstruction from Lensed Luminosity Functions

In a similar fashion to the way in which magnification can be obtained from redshift distributions, the maximum-likelihood methodology may be applied to luminosity functions. Because of the effects of galaxy clustering, a sufficiently large bin size was chosen to obtain the statistical independence required by the formation of the likelihood function in equation (2.40). In the case of magnitude space, each galaxy is assumed to be drawn at random from the luminosity distribution. Assuming that galaxy luminosities are independent of environment, each galaxy magnitude therefore

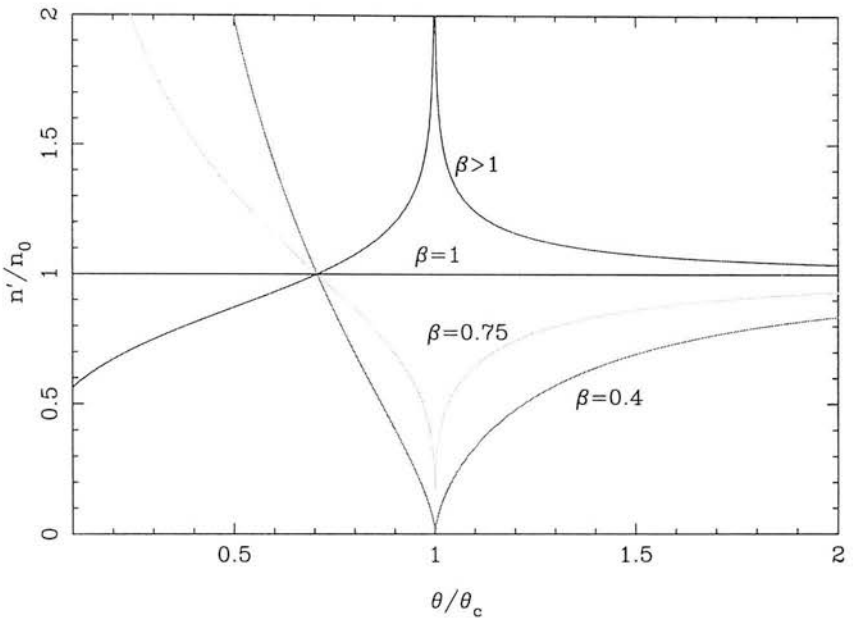

Figure 2.6: Prediction of radial variation of number counts of objects in annuli centred on an isothermal sphere for different values of β. θ_c is the critical line radius.

has complete independence from the others in the sample. The bins in magnitude space can be subsequently made infinitesimally small so that each bin contains only one galaxy. Equation (2.40) is therefore modified to the following:

$$\mathcal{L}(\mu) \propto \prod_i \mathrm{P}[M_i|\mu, z_i] \tag{2.58}$$

where M_i is the absolute magnitude and z_i is the redshift of the ith galaxy observed. The index i runs from 1 to the total number of galaxies. Similar to equation (2.32) in Section 2.3.1, the shift in the absolute magnitude distribution at each redshift compared with the intrinsic luminosity function $\phi(M, z)$ can be modeled by the probability density

$$\mathrm{P}[M|\mu, z] = \frac{\phi(M + 2.5 \lg \mu(z), z)}{\int \phi(M + 2.5 \lg \mu(z), z) \mathrm{d}M} \tag{2.59}$$

which allows the likelihood equation (2.58) to be evaluated once a form for $\phi(M, z)$ has been obtained. In practice, the redshift dependency of μ is brought about directly by the redshift dependence of κ (see equation 5.34). As was the case with redshift distributions, the luminosity function or magnitude distribution must be determined from observations of unlensed regions of the sky. The usual functional form chosen for $\phi(M, z)$ is the Schechter function (Schechter 1976), a result derived directly from

the Press-Schechter (1974) model for mass formation in the Universe:

$$\phi(M,z) = \phi^*(z)10^{0.4(M_*-M)(1+\alpha)} \exp\left[-10^{0.4(M_*-M)}\right] \tag{2.60}$$

where ϕ^*, M_* and α are determined by fitting to the observed magnitude distribution (see Section 5.3.3). The maximum likelihood magnification and its error are evaluated in exactly the same manner discussed in Section 2.3.1.

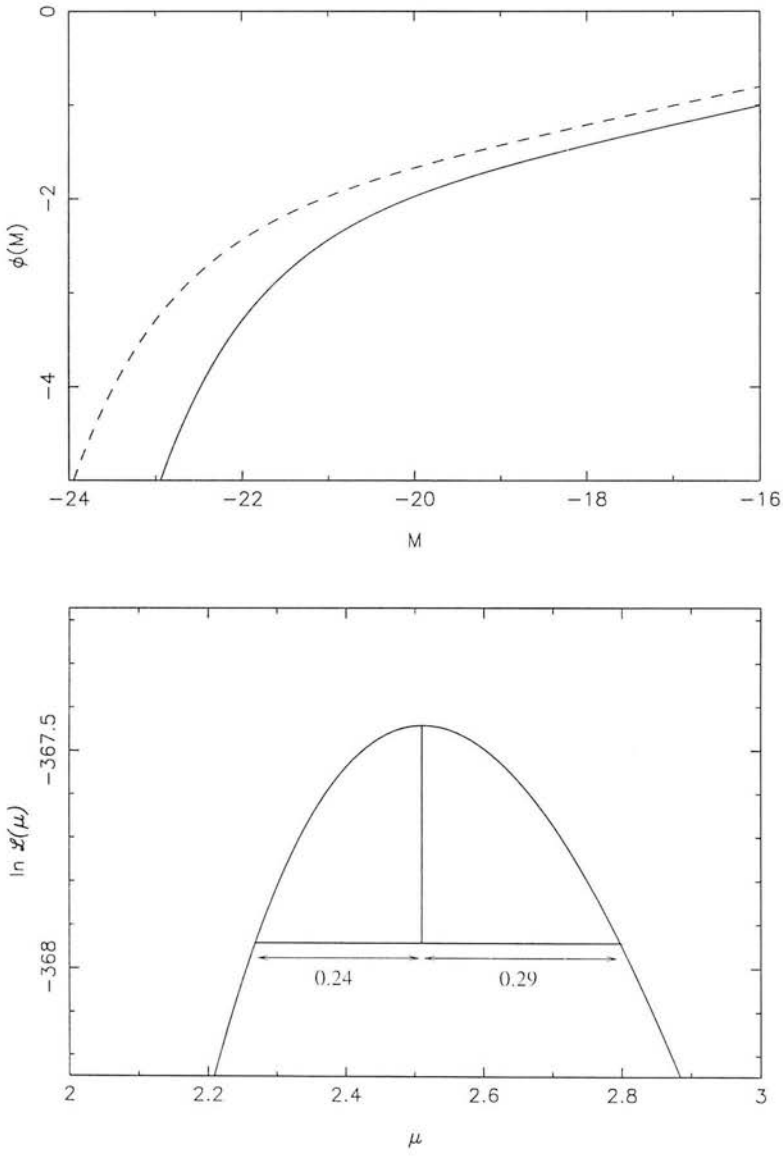

Figure 2.7: Maximum likelihood determination of lens magnification using luminosity functions. *Top*: A Schechter function ($M_* = -21$, $\alpha = 1.5$, solid line) is lensed with $\mu = 2.5$ (dashed line). *Bottom:* Maximum likelihood analysis calculates $\mu = 2.51^{+0.29}_{-0.24}$.

Figure 2.7 shows the results of maximum likelihood determination of μ using a

simulated luminosity function. The solid line in the top half of the figure is the intrinsic luminosity function from which 500 objects all lying at a fixed redshift are randomly associated magnitudes. As a simple proof of concept, this distribution is then lensed by a sheet-like distribution of mass. Since all objects lie at the same redshift and the magnification of the sheet is constant across it, all object fluxes are amplified by the same factor. This results in a simple translation of the luminosity function shown by the dashed line in the top plot of Figure 2.7. Of course in reality, the distribution of source redshifts gives a non-trivial transformation of magnitudes and hence a simple translation such as this is unlikely.

In practice, the form of the intrinsic luminosity function of galaxies lying behind the lens is known from independent observations of un-lensed regions of sky. Calculation of the probability in equation (2.59) required by the likelihood analysis is then possible, assuming this form also holds true for the lensed galaxies. In this demonstration, the form of the intrinsic luminosity function is already defined as the Schechter function described by $M_* = -21$ and $\alpha = -1.5$ and so the maximum likelihood estimator is completely defined.

The bottom half of Figure 2.7 shows the maximum likelihood curve formed by plotting equation (2.58) against μ. The input amplification applied by the sheet lens of $\mu = 2.5$ is calculated by the likelihood as $\mu = 2.51^{+0.29}_{-0.24}$. These errors are determined from the width of the curve at $\ln(\mathcal{L}_{max}/\mathcal{L}) = 0.5$ (see Section 2.3.1). This serves to demonstrate the point made in Section 2.3.1 that likelihood distributions are generally not Gaussian but may be approximated so.

2.3.4 Conversion of μ to κ

Having obtained a measure of magnification, the next step is to determine the surface mass density responsible for the lensing. Rather than work with the real surface mass density, it is more convenient to work with the scaled quantity κ. This can be easily transformed to its equivalent real value using equation (2.9).

Transforming magnification data into κ is a non-trivial exercise. Equation (2.20) in Section 2.1.5 shows why. In order to calculate κ from μ, the shear field must be

known. There are several ways in which this problem may be dealt with. Some of these are listed here. Only the last two, are used in this thesis; that of the κ estimator and that employing the axially symmetric solution.

Direct Measurement of Shear

The most obvious way is by simply measuring the shear field through observations of galaxy ellipticities as described in Section 2.2.2. Although using this measured shear field directly with magnification information in equation (2.20) would yield κ, this is not the most efficient means of doing so. This is because as equations (2.15) and (2.16) show, κ and γ are not completely independent quantities, but are fundamentally related via the deflection potential. This fact can be exploited to considerably improve the signal to noise of the reconstruction as shown by Bartelmann et al (1996). Their method determines the least-χ^2 fit to the values of the potential measured in bins on a rectangular grid by minimizing the quantity

$$\chi^2 = \sum_j \frac{1}{\sigma_g^2(j)} \left[g_i(j) - \hat{g}_i(j)\right]^2 + \frac{1}{\sigma_r^2(j)} \left[r(j) - \hat{r}(j)\right]^2 \tag{2.61}$$

where g_i are the two components of the reduced shear given in equation (2.30), with summation being implied over index i and $r = \mu^{-1}$. The quantities with hats are calculated from the potential field being fit and those without are the measured quantities. The summation here is over all grid bins and is weighted by the variances of g and r from the data.

Bartelmann et al (1996) have applied this technique to simulated galaxy clusters and typically find that they can reconstruct their cluster mass to within $\pm 10\%$ of the correct value. While this method clearly works, it requires measurement of the shear field which is not always possible with images observed for the purpose of extracting magnification information. This is because poor seeing acts to smooth away the coherent lens-induced ellipticities of images of especially faint galaxies but does not drastically alter the ability to detect and measure a galaxy's brightness.

Another technique which can be used to overcome the lack of shear information is the iterative method illustrated in Figure 2.8 (see also, for example van Kampen 1998). Starting with an initial guess for κ (by the estimator method described later for example), an initial estimate of the deflection potential is formed. This is achieved by using either equation (2.11) or more efficiently, in Fourier space using the Fourier transform of equation (2.15) to give[1],

$$\psi(\boldsymbol{k}) = -\frac{2\kappa(\boldsymbol{k})}{k_1^2 + k_2^2} \tag{2.62}$$

where k_1 and k_2 are the components of $\boldsymbol{k}$. $\psi(\boldsymbol{k})$ is then used to calculate the components of shear in Fourier space by using the Fourier transform of equation (2.16),

$$\begin{aligned} \gamma_1(\boldsymbol{k}) &= -\frac{1}{2}(k_1^2 - k_2^2)\psi(\boldsymbol{k}) \\ \gamma_2(\boldsymbol{k}) &= -k_1 k_2 \psi(\boldsymbol{k}) \end{aligned} \tag{2.63}$$

Fourier transforming back into real space, γ is then used with equation (2.20) to form a new estimate of κ which completes the iterative cycle. This is repeated until sufficient convergence has been reached. Unfortunately, the iterative technique only works for lenses which are sub-critical (Dye & Taylor 1998). The method is extremely unstable when critical lines are introduced and divergence occurs as a result.

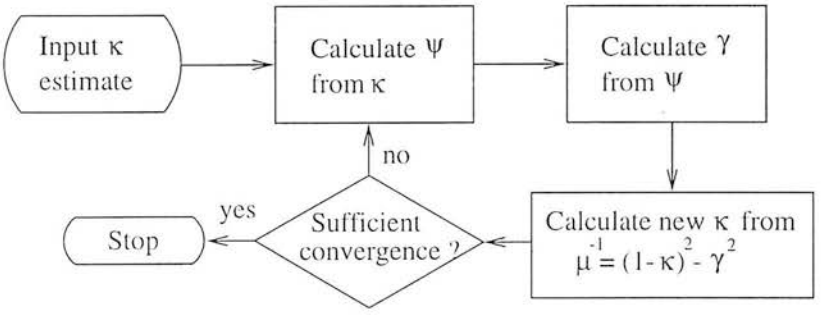

Figure 2.8: Iterative procedure to calculate κ given only information about μ.

Chapter 4 demonstrates a more sophisticated iterative technique developed by Dye & Taylor (1998) which does not suffer from divergence and does not require shear information.

[1]Here, the result that the Fourier transform of $\mathrm{d}f(x)/\mathrm{d}x$ is $-ikf(k)$ is used.

κ Estimators

A very quick and simple means of estimating κ is by use of the κ *estimator* (Taylor et al 1998, T98 hereafter and also van Kampen 1998). The principle elaborates on the idea first presented by BTP that in weak lensing regimes where $\kappa \ll 1$ and $\gamma \ll 1$, the terms κ^2 and γ^2 can be neglected in the magnification equation (2.20) so that a weak estimation of κ is given by

$$\kappa_{weak} = \frac{1}{2}(1 - \mu). \tag{2.64}$$

Because this approximation is extremely inaccurate for values of approximately $\kappa > 0.1$, T98 introduced two estimators for κ which can be used in the strong lensing regime.

The idea relies upon the fact that realistic cluster mass distributions lie somewhere between that of an isothermal sphere and that of a homogeneous sheet of matter. Sheet-mass distributions were discussed in Section 2.1.5 and shown to be characterised by

$$\kappa = constant \quad , \quad \gamma = 0. \tag{2.65}$$

The mass distribution of an isothermal sphere as covered in Section 2.3.1 is similarly characterised by equation (2.55). Using these two results for γ in equation (2.20) therefore produces two estimates for κ,

$$\begin{aligned} \kappa_+ \equiv \kappa(\gamma = 0) &= 1 - \mathcal{P}\mu^{-1/2} \\ \kappa_- \equiv \kappa(\gamma = \kappa) &= \frac{1}{2}(1 - \mathcal{P}\mu^{-1}) \end{aligned} \tag{2.66}$$

where $\mathcal{P}$ is a parity which flips from being $+1$ outside regions bounded by a critical line to -1 inside. The *sheet estimator*, κ_+, provides an upper bound on κ and a lower bound is provided by κ_-, the *isothermal estimator*. Before a critical line is crossed, it is generally true that $\kappa_{weak} \gg \kappa_+ > \kappa_{true} > \kappa_-$.

Figure 2.9 shows a scatter plot of inverse magnification against κ for a simulated cluster in a cold dark matter universe (van Kampen & Katgert 1997). The κ_{weak} estimator demonstrates its inaccuracy after $\kappa > 0.1$, overestimating κ by up to a factor of 2. The κ_+ estimator is seen to act as a strong upper bound until the critical line is crossed. This is as expected since in a real cluster, $\gamma \geq 0$ always holds true so

that the upper limit provided by $\kappa_+ \equiv \kappa(\gamma = 0)$ cannot be crossed. κ_- only provides a reliable lower limit for $\kappa > 0.2$ however. van Kampen (1998) finds that for the most massive simulated clusters, this limit must be extended to $\kappa > 0.4$. This stems from substructure in the outskirts of clusters where κ can be low enough to result in the condition that $\kappa < \gamma$. This exceeds the lower limit set by $\kappa_- \equiv \kappa(\gamma = \kappa)$.

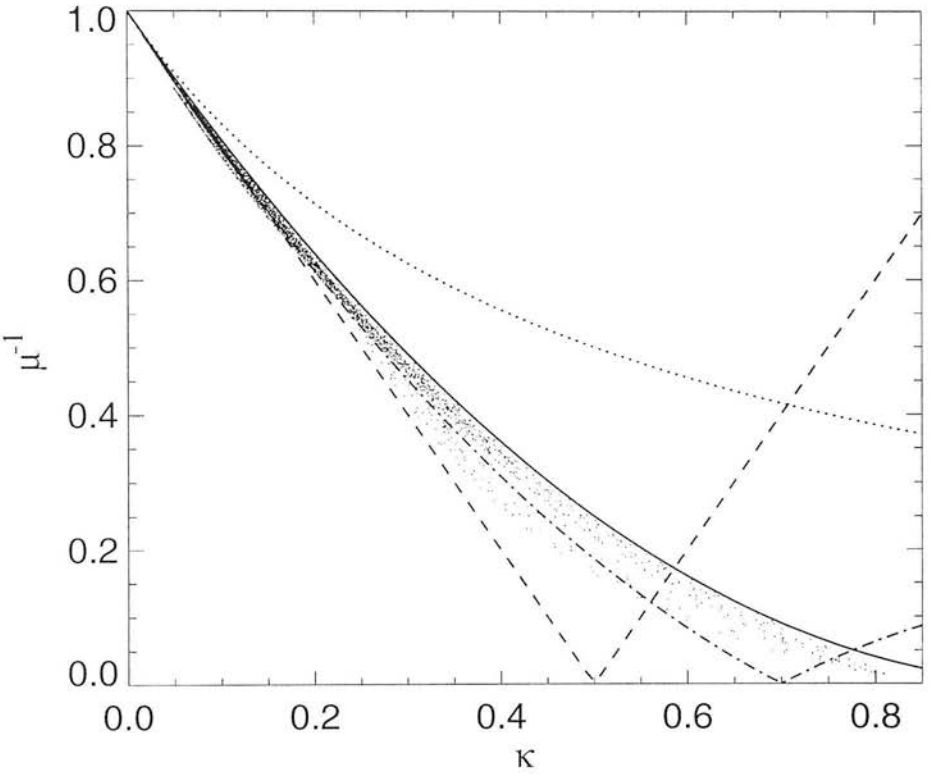

Figure 2.9: Variation of inverse magnification with κ for each pixel of a simulated cluster. Superimposed onto the plot are shown the estimators κ_+ (solid), κ_- (dash), κ_{weak} (dot) and κ_c with $c = 0.7$ (dot-dash).

The fourth estimator, κ_c, shown in Figure 2.9 is a parabolic approximation motivated by cluster simulations and the fact that it has an invertible $\mu(\kappa)$ relation (van Kampen 1998). Behaving essentially as an average of κ_+ and κ_-, κ_c is chosen to vary as $\gamma \propto \sqrt{\kappa}$. It is convenient to express the constant of this proportionality as

$$\gamma = |1 - c| \sqrt{\frac{\kappa}{c}} \tag{2.67}$$

so that the inverse magnification is simply given by

$$\mu^{-1} = |(\kappa_c - c)(\kappa_c - 1/c)| \,. \tag{2.68}$$

The constant c, can be adjusted to vary the two critical lines which occur when $\kappa = c$ and $\kappa = 1/c$.

Figure 2.10 graphically illustrates how the introduction of a second critical line has the consequence that four solutions to equation (2.68) for κ exist. This is only true for values of μ^{-1} which give real solutions. The possibility can arise that when attempting to estimate κ in the region between both critical lines (the shaded area in the plot) μ^{-1} is larger than the maximum value permitted by the parabolic estimator in this region. If the critical line region has been incorrectly defined or if μ^{-1} is in error, this can occur. In this case, the estimator breaks down and the estimators κ_+ and κ_- must be used instead to provide an upper and lower bound on κ.

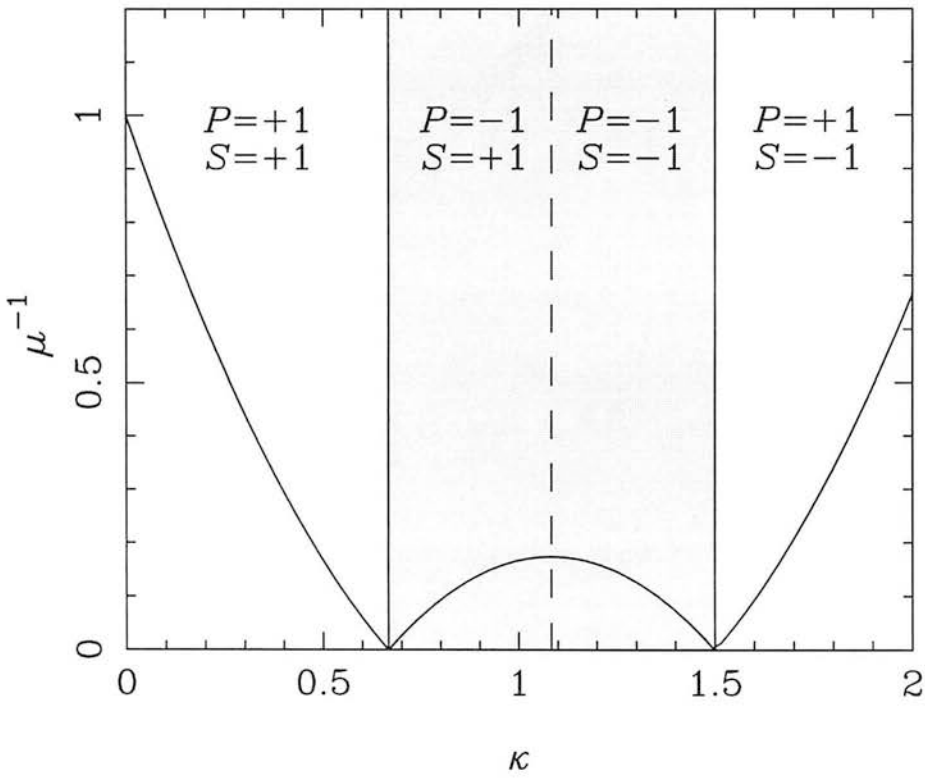

Figure 2.10: The estimator κ_c showing parities $\mathcal{P}$ and $\mathcal{S}$. The grey area between the two critical lines can give non-physical results for κ if μ^{-1} is too large.

The four solutions for κ_c from equation (2.68) are given by

$$\kappa_c = \frac{1}{2c}\left((c^2+1) - \mathcal{S}\sqrt{(c^2+1)^2 - 4c^2(1-\mathcal{P}\mu^{-1})}\right) \tag{2.69}$$

The parity $\mathcal{P}$ has the same function as in equation (2.66) except now, it can flip twice as the outer and then the inner critical lines are crossed. $\mathcal{P}$ appears as a result of

the modulus operator in equation (2.68). The second parity $\mathcal{S}$ is due to the parabolic nature of equation (2.68) and is the sign of $(c^2+1)/2c - \kappa$. Figure 2.10 shows the various parity flips described. Note that κ_c becomes κ_+ when $c = 1$.

Axially Symmetric Lens Reconstruction

If one makes the assumption that a lens is axially symmetric, then a non-local functional relationship between κ and γ is fixed. Armed with magnification information, a self-consistent κ and γ profile can be calculated with an iterative procedure.

Axial symmetry allows the geometry described in Section 2.1.3 to be simplified to one dimension. This is explicitly seen by considering the scaled deflection angle in equation (2.8). By symmetry, the impact vector $\boldsymbol{\theta}$ may be restricted to lie along the positive θ_1 axis in the lens plane such that $\boldsymbol{\theta} \equiv (\theta, 0)$. Similarly, using the polar co-ordinates θ' and ϕ to write

$$\boldsymbol{\theta}' \equiv (\theta' \cos\phi, \theta' \sin\phi) \tag{2.70}$$

with $\theta \geq 0$ and $\theta' \geq 0$ shows that the components of the deflection angle may be written as,

$$\begin{aligned} \alpha_1(\theta) &= \frac{1}{\pi}\int_0^\infty \theta' \mathrm{d}\theta' \kappa(\theta') \int_0^{2\pi} \mathrm{d}\phi \frac{\theta - \theta'\cos\phi}{\theta^2 + \theta'^2 - 2\theta\theta'\cos\phi} \\ \alpha_2(\theta) &= \frac{1}{\pi}\int_0^\infty \theta' \mathrm{d}\theta' \kappa(\theta') \int_0^{2\pi} \mathrm{d}\phi \frac{-\theta'\sin\phi}{\theta^2 + \theta'^2 - 2\theta\theta'\cos\phi}. \end{aligned} \tag{2.71}$$

The restriction enforced on the direction of the impact vector combined with the axial symmetry means that α_2 vanishes. The inner integral of equation (2.71) for α_1 adheres to

$$\int_0^{2\pi} \mathrm{d}\phi \frac{\theta - \theta'\cos\phi}{\theta^2 + \theta'^2 - 2\theta\theta'\cos\phi} = \begin{cases} 2\pi/\theta & \text{for} \quad \theta' < \theta \\ 0 & \text{for} \quad \theta' > \theta \end{cases} \tag{2.72}$$

(see Appendix A.1) and thus the matter contained within a disk of radius θ about the centre of mass contributes to the deflection angle at $\boldsymbol{\theta}$ as if it were all located at that centre. Matter outside this disk has no influence on the deflection in a similar manner to the effect of gravitational forces from spherical mass distributions. Equation (2.71) therefore becomes

$$\alpha(\theta) \equiv \alpha_1(\theta) = \frac{2}{\theta}\int_0^\theta \theta' \mathrm{d}\theta' \kappa(\theta') \equiv \theta\,\overline{\kappa}(\theta) \tag{2.73}$$

where $\overline{\kappa}(\theta)$ is the mean value of κ within a circle of radius θ.

To arrive at an expression which relates κ to γ, the lens equation of Section 2.1.3 must be employed. Combining equation (2.10) with the result of equation (2.73) gives

$$\boldsymbol{\beta} = \boldsymbol{\theta}(1 - \overline{\kappa}(\theta)) \tag{2.74}$$

so that the Jacobian of the transformation is

$$\mathcal{A} \equiv \frac{\mathrm{d}\boldsymbol{\beta}}{\mathrm{d}\boldsymbol{\theta}} = \begin{pmatrix} 1-\overline{\kappa} & 0 \\ 0 & 1-\overline{\kappa} \end{pmatrix} - \frac{1}{\theta}\frac{\mathrm{d}\overline{\kappa}}{\mathrm{d}\theta}\begin{pmatrix} \theta_1^2 & \theta_1\theta_2 \\ \theta_1\theta_2 & \theta_2^2 \end{pmatrix}. \tag{2.75}$$

Comparing the components in equation (2.75) with those in equation (2.17) shows that the components of γ are therefore

$$\begin{aligned} \gamma_1 &= \frac{1}{2\theta}\left(\theta_1^2 - \theta_2^2\right)\frac{\mathrm{d}\overline{\kappa}}{\mathrm{d}\theta} \\ \gamma_2 &= \frac{\theta_1\theta_2}{\theta}\frac{\mathrm{d}\overline{\kappa}}{\mathrm{d}\theta} \end{aligned} \tag{2.76}$$

and hence

$$\gamma^2 = \gamma_1^2 + \gamma_2^2 = \frac{\theta^2}{4}\left(\frac{\mathrm{d}\overline{\kappa}}{\mathrm{d}\theta}\right)^2. \tag{2.77}$$

From equation (2.73),

$$\frac{\mathrm{d}\overline{\kappa}}{\mathrm{d}\theta} = \frac{2}{\theta}(\kappa - \overline{\kappa}) \tag{2.78}$$

which upon substitution into equation (2.77) relates γ and κ via

$$\gamma^2 = (\kappa - \overline{\kappa})^2. \tag{2.79}$$

The next step is to include amplification. This is simply achieved by substituting equation (2.79) into equation (2.20) to give

$$\mu = \left|(1-\overline{\kappa})(1-2\kappa+\overline{\kappa})\right|^{-1}. \tag{2.80}$$

The surface mass density and shear can now be solved through series solution by dividing the surface mass density into consecutive annuli of equal width. Labelling the innermost annulus (a disk) with the index $n = 1$ and segregating $\overline{\kappa}$ into an inner term, η_{n-1}, and a surface term gives,

$$\overline{\kappa}_n = \eta_{n-1} + \frac{2}{n+1}\kappa_n. \tag{2.81}$$

η_{n-1} is given by the discretisation of equation (2.73),

$$\eta_{n-1} = \frac{2}{n(n+1)} \sum_{m=1}^{n-1} m\kappa_m. \tag{2.82}$$

Writing equation (2.80) in terms of quantities relating to the n^{th} shell and using the parity of Section 2.3.4 gives

$$\mathcal{P}\mu_n^{-1} = (1 - \overline{\kappa}_n)(1 - 2\kappa_n + \overline{\kappa}_n). \tag{2.83}$$

Using this with equation (2.81) gives, upon rearrangement,

$$\begin{aligned} \kappa_n &= \frac{(n+1)}{4n}\{n + 1 - (n-1)\eta_{n-1} - \\ &\quad \mathcal{R}[(n - 1 - (n+1)\eta_{n-1})^2 + 4n\mathcal{P}\mu_n^{-1}]^{1/2}\}. \end{aligned} \tag{2.84}$$

This is the axially symmetric series solution for κ. As in Section 2.3.4, the image parity $\mathcal{P}$ flips about critical lines. $\mathcal{R}$ is a parity similar in nature to the parity $\mathcal{S}$ of Section 2.3.4 which permits higher solutions of κ.

The freedom of choice for both parities in equation (2.84) is usually constrained in practice by the observed position of critical lines. This leaves only one degree of freedom; choice of the parameter η_0. This is in fact limited to some degree by the requirement that non-physical solutions should be avoided. Expressed mathematically, this means that the condition $\eta_0^2 \geq \mathcal{P}\mu^{-1}$ must be satisfied for a given magnification profile. Combination of equation (2.79) and equation (2.81) shows that $\eta_0 = \gamma_1$ so that the freedom is on the shear within the first annulus. The choice of γ_1 has only a small effect on the overall κ profile obtained. This is demonstrated in Section 3.2.5 where the axially symmetric solution is applied to Abell 1689.

Chapter 3

Magnification of Source Number Counts by Abell 1689

This chapter sees application of the integrated number count technique of Section 2.3.2 to the cluster Abell 1689 (RA,Dec.[J2000]: $13^h11^m34^s, -01°21'54''$). The work presented here forms the predominant result of the publication by Taylor & Dye et al (1998). Studies into the mass of the cluster Abell 1689 from X-ray temperature measurements (Yamashita 1994), virial velocity determinations (Teague et al 1990, TCG hereafter) and shear analysis (Tyson & Fischer 1995, Kaiser 1996) all conclude that it is one of the most massive clusters in the known Universe. This therefore makes it a prime target for the verification of the magnification method.

3.1 Data Acquisition

Observation and basic reduction of the data contained in this Chapter was performed by T.J. Broadhurst and N. Benítez of the University of California, Berkeley in collaboration with the Institute for Astronomy, Edinburgh.

3.1.1 Observation of A1689

Images of A1689 were taken in the V and I bands during a run in February 1994 at ESO's NTT 3.6m telescope. Each band was observed for a total integration time of 6000 seconds in a field of view covering ~ 35 square arcminutes centred on the cluster (see figure 3.1). Seeing was similar in both bands, with FWHM of $0.8''$ and a CCD pixel scale of $0.34''$. The passbands and exposures were chosen such that the cluster E/S0 galaxies would be bluer than a good fraction of the background, requiring much deeper imaging in the bluer passband for detection. The cluster was observed to a limiting magnitude of I=24.

The images were de-biased and flattened with skyflats using standard IRAF procedures. Large scale gradients in brightness of a few percent remained after the flat-fielding process. These were removed by correcting with a flat field obtained from the images themselves by masking objects and smoothing the result. This gave a homogeneous photometry across the field with a 1σ deviation of 0.1 magnitudes.

3.1.2 Object Detection

Determination of object magnitudes in each of the co-added V and I band images was made possible through use of software called *SExtractor* (Bertin & Arnouts 1996). SExtractor is essentially an object extraction algorithm which operates on FITS images to produce a catalogue of detections. An *object* in SExtractor is defined as a set of connected pixels which exceed a user-defined threshold above the background. By calculating the second order moments of the light distribution in these pixels, SExtractor fits an ellipse to every detected object and outputs elongation and orientation to the catalogue.

The integrated fluxes within the fitted ellipses of all detected objects are then calculated and converted to magnitudes. A calibration offset must usually be applied to these magnitudes since SExtractor makes the assumption that the images were observed in photometric conditions, that instrument transmission is 100% and that the detector gain is unity (ie. the number of electrons output by the CCD circuitry is equal to the number of photons striking the CCD). Calibration was not necessary in

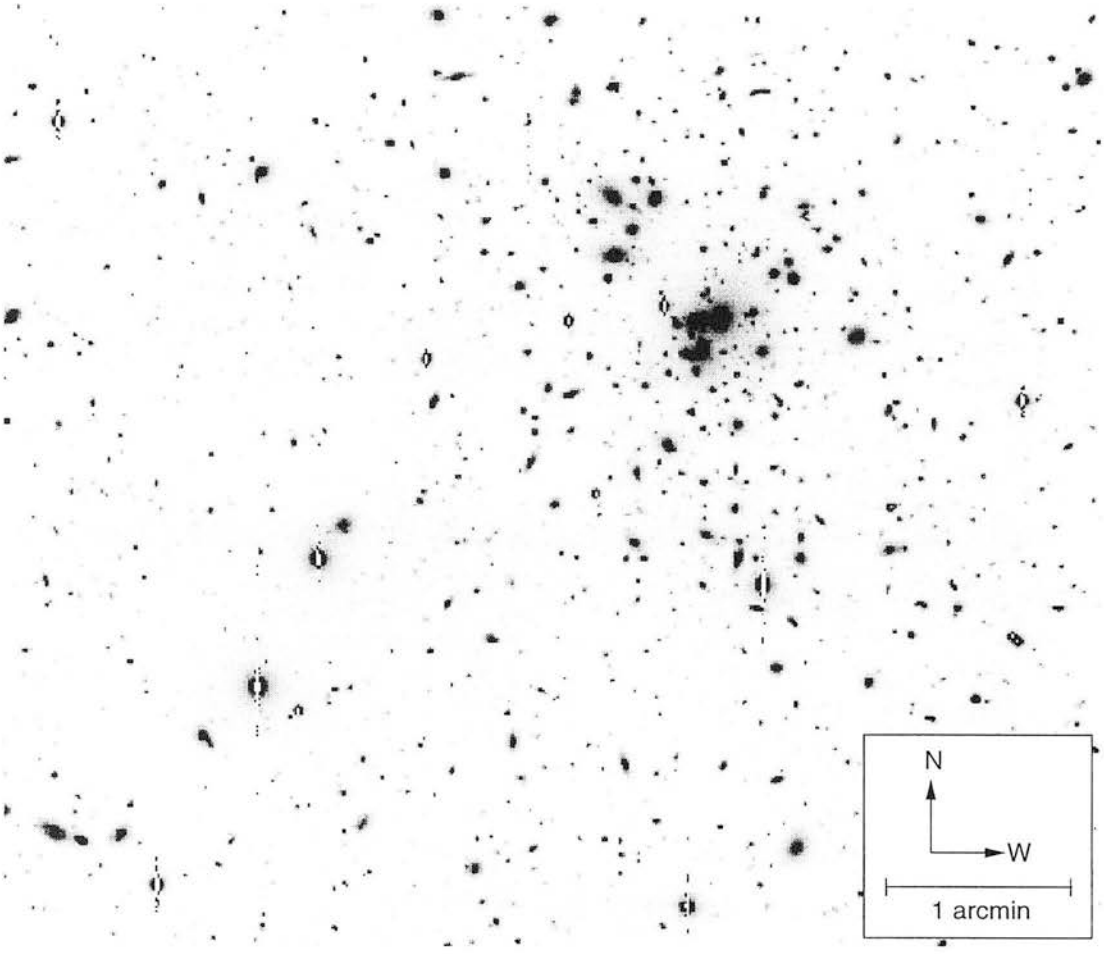

Figure 3.1: Post-reduction I band image of Abell 1689. The cluster centre lies in the north-west quadrant. Note that all stars, determined from their high brightness and low FWHM, have had their centres and vertical leaking removed. A logarithmic grey scale is used to highlight faint objects relative to the bright cluster members.

this case since only the difference between the V − I colour of the background objects and the V − I colour of the cluster galaxies was required for the purpose of object selection (see Section 3.1.3).

Objects (consisting of 8 or more connected pixels) in the V and I band images of A1689 were detected with a threshold of 2σ above the background noise to yield a total number of approximately 4000 objects in each. The inevitable consequence of choosing a threshold as low as 2σ means that roughly 5% of background pixels are above this limit and hence several spurious objects are detected. The majority of these spurious objects were removed from the object catalogue of each frame by performing a co-incidence match between the V and I object catalogues. By comparing

the catalogue of matched objects to the images, an optimal separation tolerance of about 2 pixels was chosen. Several spurious objects still remained after the coincidence match however. As Figure 3.2 shows, a large number of noisy objects were detected around the edges of large extended objects, these areas being above the detection threshold. These remaining objects were removed by hand by comparing with the images in each band.

Figure 3.2 shows the central parts of A1689 in both bands and the objects detected in each. The matched objects in this part of the cluster are shown in the bottom part of Figure 3.2, each having an associated V and I magnitude. Approximately 3000 objects in total were matched, demonstrating that $\sim 25\%$ of the objects detected at the 2σ threshold were spurious.

3.1.3 Object Selection

To measure the lensing effect on the number density of background objects, distinction of cluster and foreground objects was necessary. This was achieved through the identification of objects on a colour-magnitude diagram formed by plotting the $V - I$ colour against I magnitude for every object.

Figure 3.3 shows the colour-magnitude diagram obtained for the 3000 objects detected in the field of A1689. The cluster objects were identified by their strong E/SO colour sequence which forms the well defined horizontal band across the diagram at roughly $V - I = 1.5$. The sharp upper edge of this band represents the reddest galaxies in the cluster. Assuming that the cluster galaxies are of a similar nature to the field galaxies, this means that objects which lie redward of the cluster E/SO sequence must be redder because they lie at higher redshifts. In other words, objects which lie in this part of the colour-magnitude diagram represent a population of background objects. These will be referred to as the *red background galaxy population.*

The red background galaxy population is used throughout this chapter to measure the lensing effect of A1689. The reasons for this are twofold. Firstly, as explained immediately above, one can unambiguously separate the population from the cluster galaxies and hence minimal contamination of the sample is expected. The second

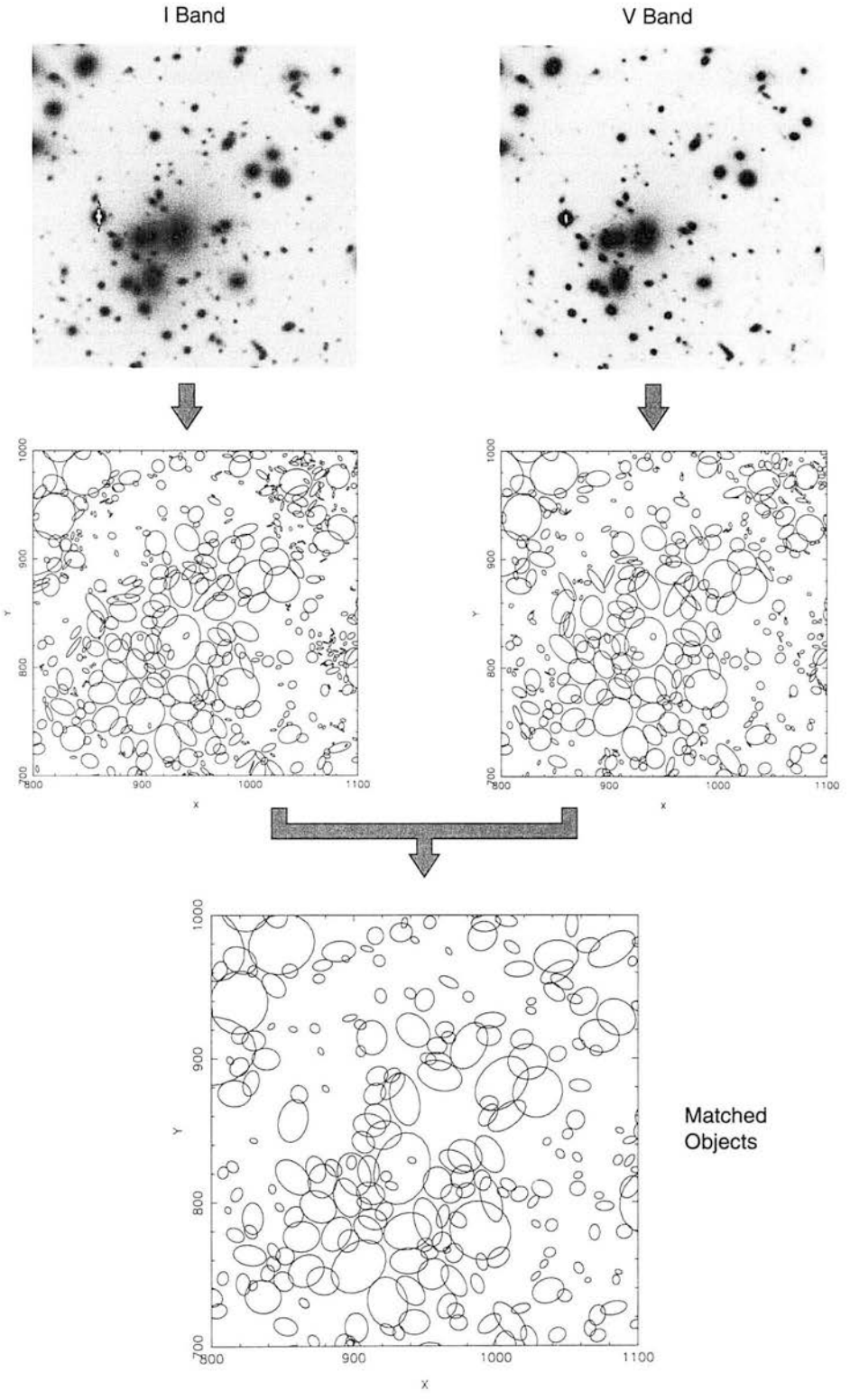

Figure 3.2: V and I band images and object detections of the central region of A1689. The V and I objects were matched using a position coincidence test with an optimal separation tolerance of 2 pixels to produce the bottom plot.

reason, given in Chapter 2, is that field galaxies with $V - I > 1.5$ have a shallow number count slope which maximises the lensing signal detected.

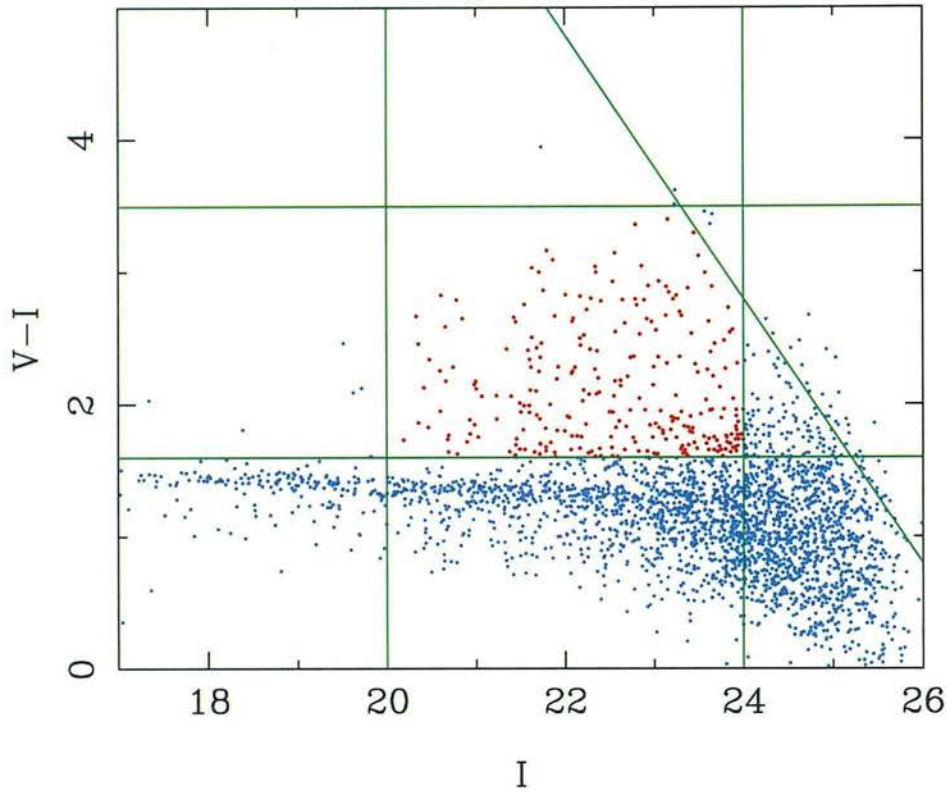

Figure 3.3: Colour-magnitude diagram of the 3000 matched objects around A1689. Note the well defined horizontal cluster sequence. Although not strictly necessary, a rough calibration was provided from the cluster magnitudes of Gudehus & Hegyi (1991).

Figure 3.3 shows the limits used to identify objects. The value of $V - I$=1.6 was selected as the lower cut-off for the red background galaxies. The I band magnitude was limited to $20 < I < 24$ and the V band magnitude limit of $V < 28.0$ was enforced to ensure completeness of the sample. A further colour cut of $V - I < 3.5$ was applied where the reddest galaxies are expected to cut off given the survey magnitude limits. Anything redward of this limit was considered unreliable (possibly a bad magnitude determination by the extraction software or incorrect association of objects between the both bands during the matching process).

To isolate the bright cluster members and foreground contaminants, all objects with $V - I < 1.6$ and $I < 22$ were selected. Isolation of the faint cluster objects is made impossible by the fact that these objects are located in the same region of

colour-magnitude space as the expected location of faint blue background galaxies (ie. roughly $I > 22$ and $V - I < 1.0$). Fortunately, this is not a major concern because the cluster and foreground identified objects are ultimately only used to estimate the obscuration of the red background objects (see Section 3.1.4). The faint and hence small cluster objects therefore only play a small part in the obscuration compared with the larger and brighter members.

Since the identification of cluster members is important to remove contamination of the background sample, the colour selected candidates were cross-checked with the photometric redshift data presented in Chapter 5. The photometrically identified cluster objects were in good agreement with the colour selection.

3.1.4 Masking

Having identified the various populations on the colour-magnitude diagram, a mask to eliminate areas obscured by cluster and foreground objects (the *mask objects*) was produced. This was necessary to quantify the effects of obscuration bias in the surface density of background object counts.

Some experimentation was required to choose the most suitable parameters output by SExtractor for creation of the mask. Although SExtractor outputs the semi-major and semi-minor axes of detected objects, these are computed from the maximum and minimum FWHM of the object profiles and as such are typically too small to represent the true obscuration (see Figure 3.4). Attempts at scaling these to match image object sizes proved unfeasible due to the lack of a global scale factor. An alternative means of obtaining ellipse sizes was achieved from using the detected number of pixels lying inside an object's lowest isophote. Representing this quantity as n_p and setting it as the area of an ellipse with semi-major and semi-minor axes a and b respectively gives

$$a = \sqrt{\frac{n_p}{\pi(1-e)}}, \quad b = \sqrt{\frac{n_p}{\pi}(1-e)} \tag{3.1}$$

where the ellipticity, e, is defined here as

$$e = \frac{a-b}{a}. \tag{3.2}$$

(Note that this definition used by SExtractor differs from the more common definition:

$e = (a - b)/(a + b)$.) The ellipses generated in this way were found to be far more reliable as demonstrated in the bottom right-hand section of Figure 3.4.

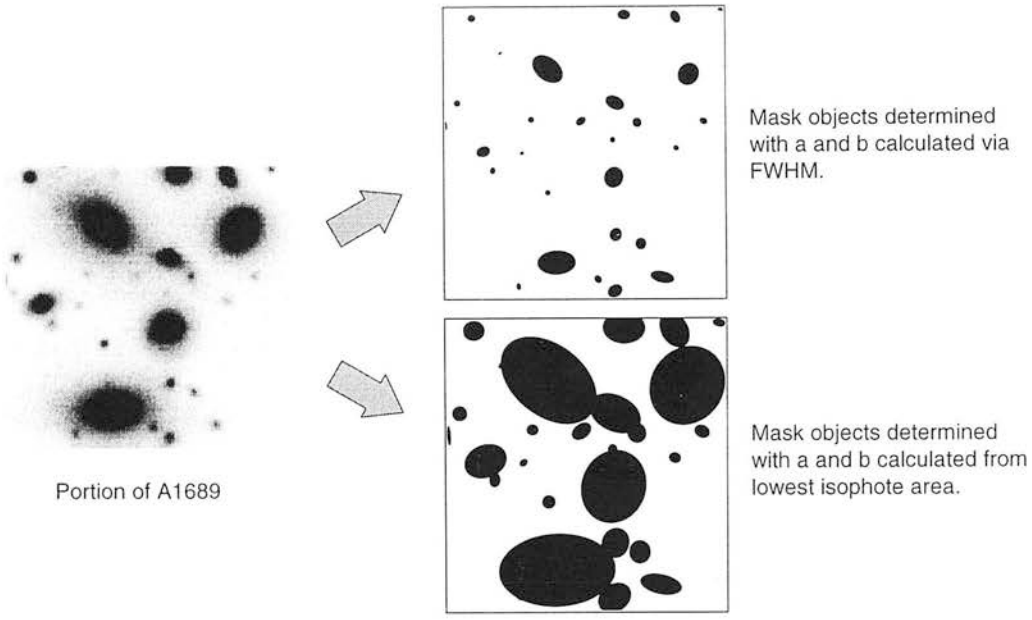

Figure 3.4: Comparison of objects defined using semi-major and semi-minor axis information directly output from SExtractor with that calculated from the area contained within an object's lowest isophote. The latter allows the most accurate mask definition.

Figure 3.5 shows the final mask used together with the red background galaxies plotted as open circles. Some of the larger stars not properly accounted for by SExtractor were manually added. The concentric rings shown here are the annular bins used in Section 3.2.1 to measure the radial background number counts.

3.2 Results

3.2.1 Number counts

In Section 2.3.2, it was noted that the luminosity function of field galaxies selected by $V - I > 1.5$ is sufficiently flat to cause the number density of such objects under magnification to be reduced. It is therefore expected that a plot of the radial variation in number counts about the centre of A1689 will show a larger depletion in areas where the magnification is higher as predicted by equation (2.54).

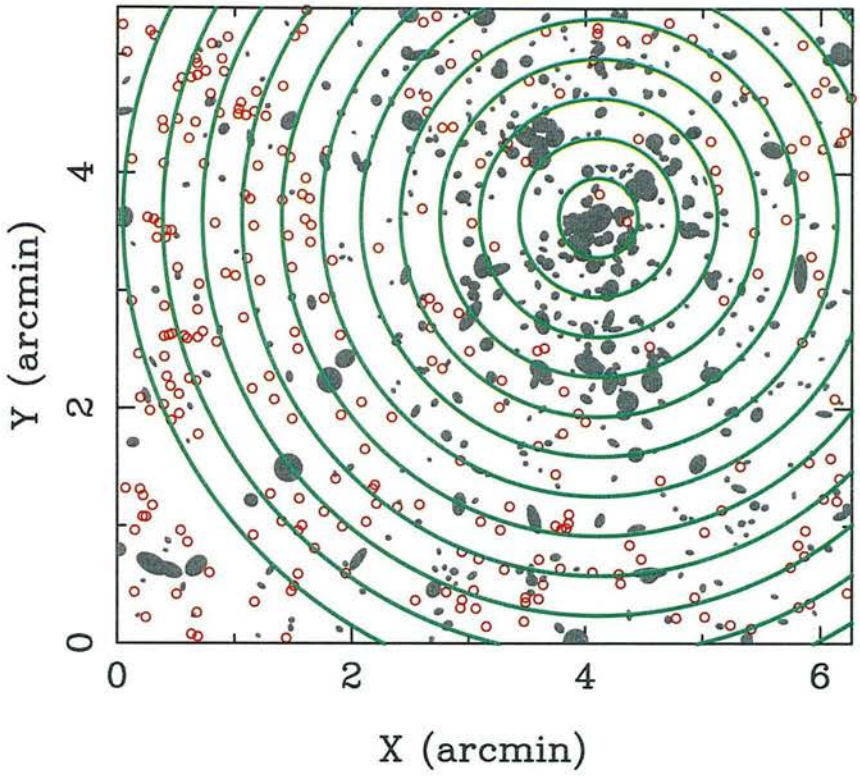

Figure 3.5: A1689 field showing mask (grey ellipses), red background galaxies (open red circles) and annular bins used to measure radial number counts (green).

Counting the number of red background sources in each of the annuli shown in Figure 3.5 enables the plot shown in Figure 3.6 to be drawn. As anticipated, the number of galaxies is reduced in comparison with the number of field galaxies expected in the absence of lensing. Section 3.2.2 describes how the intrinsic background number density is derived to enable the expected number in a given bin, n_0 to be calculated taking obscuration by mask objects into account. In the second bin from centre, there are no detected galaxies at all which suggests that infinite magnification and hence a critical line exists at this radius. This is in agreement with several faint arcs present in the image in this area.

The dashed line plotted in Figure 3.6 shows the variation in number counts predicted for an isothermal lens adjusted so that the critical line radius matches the position of the second radial bin. As the plot shows, near-isothermal behaviour is exhibited by A1689, apart from a large depletion in counts at a radius of 2.4 arcmin and an over dense region at a radius of approximately 3.75 arcmin. The over-dense region is most likely a result of clustering of the background objects. The error bars

plotted are based purely on Poisson statistics of the expected number and therefore do not take clustering into account (unlike the mass plots which follow).

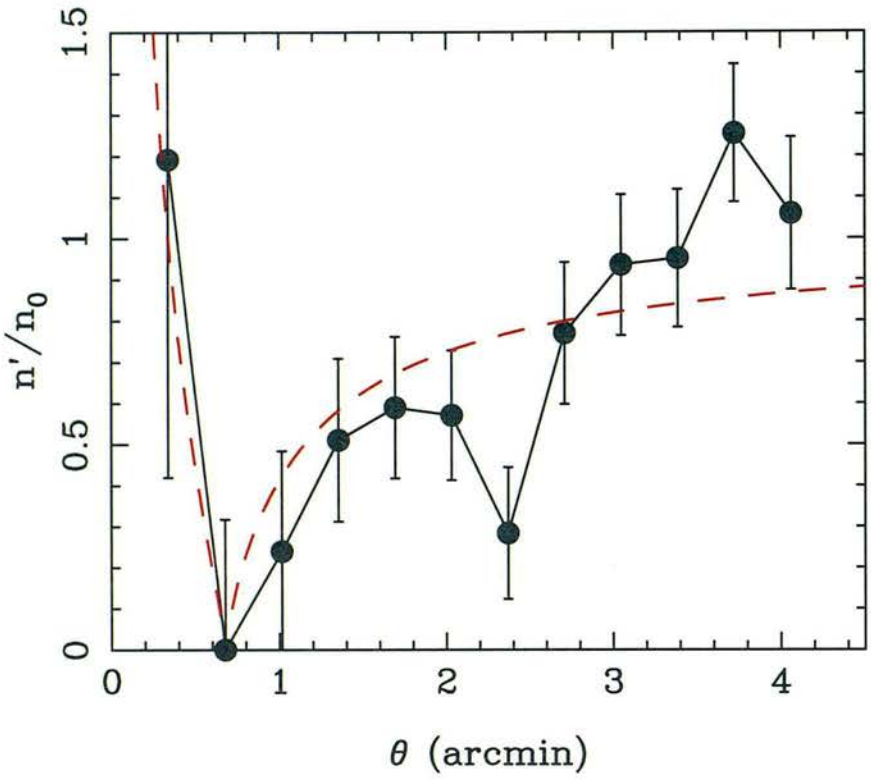

Figure 3.6: Radial number counts of objects around A1689 expressed as a fraction of the expected number. The error bars are derived from Poisson statistics. The dashed line shows the behaviour of an isothermal lens adjusted to match the critical line position.

One explanation for the depletion of counts seen at $\theta \simeq 2.4$ is that the cluster has an extended structure with a large mass at this radius. Referring to the image shown in Figure 3.1, this is vaguely reflected in the distribution of light from the cluster galaxies with an above average contribution coming from the large concentration of objects to the north-east of the cluster centre. An alternative hypothesis is that this is the first glimpse of a second critical line (see Taylor & Dye et al 1998 for a more detailed account). Since the critical radius of a given cluster scales the same way as the radius of an Einstein ring, ie. proportional to D_{ds}/D_s, a population of background galaxies lying at a high redshift would have a larger critical radius than a population at a low redshift. If there are two distinct populations of galaxies, one at low redshift and the other at high redshift, there could in principle be two critical lines. This phenomenon is thought to have been observed with the lensing cluster Cl0024+1654

(Fort, Mellier & Dantel-Fort 1997). In the absence of further evidence for a second high redshift population, only the single critical line model is considered here.

3.2.2 Normalisation of Background Number Counts

Of major importance to the lens magnification method is the normalisation of the background galaxy population. One means of providing this normalisation is to assume negligible cluster mass at the edge of the observed field and normalise to the number density of objects there. This defeats one of the main advantages of the magnification method in enabling the computation of absolute masses. One would be effectively using the same normalisation as the shear method such that the estimated mass could only ever be determined relative to the mass at the edges of the field.

Instead, the normalisation adopted in this work, that being the same normalisation used to provide the ratio of n'/n_0 in Figure 3.6, comes from the data of Smail et al (1995). These deep observations in the V, R and I passbands with the Keck telescope have a limiting magnitude of $R \approx 27$. A χ^2 straight line fit to the total corrected[1] differential galaxy count rate in the I-band gives

$$\log_{10} n = (0.271 \pm 0.009)\mathrm{I} - 1.45 \tag{3.3}$$

over the range $20 < \mathrm{I} < 24$, where n is per magnitude per square degree. Applying the colour criteria used in Section 3.1.3 to this data, one finds that the red galaxy population with $\mathrm{V} - \mathrm{I} > 1.6$, can be well approximated by

$$\log_{10} n(\mathrm{red}) = (0.0864 \pm 0.0187)\mathrm{I} + (2.12 \pm 0.41) \tag{3.4}$$

over the range $20 < \mathrm{I} < 24$. Figure 3.7 shows the magnitude distribution for the full dataset and for the red-selected galaxy population and the best-fit lines. Integrating the fit for the red galaxies yields a total count rate of $n = 12.02 \pm 3.37$ galaxies per square arcminute in the range $20 < \mathrm{I} < 24$. Re-writing equation (2.36) in terms of magnitudes, the slope parameter β is

$$\beta = 2.5 \, \frac{\mathrm{d} \log_{10} n(m)}{\mathrm{d}m} \tag{3.5}$$

[1]Smail et al corrected for completeness of their number counts using Monte Carlo simulations in which artificial galaxies were added to the data for re-application of their object detection algorithm.

which from equation (3.4) implies that $\beta = 0.216 \pm 0.047$. This is the value of β which is used in the subsequent analysis. Section 3.2.8 discusses the impact on the final mass estimate due to this uncertainty in β.

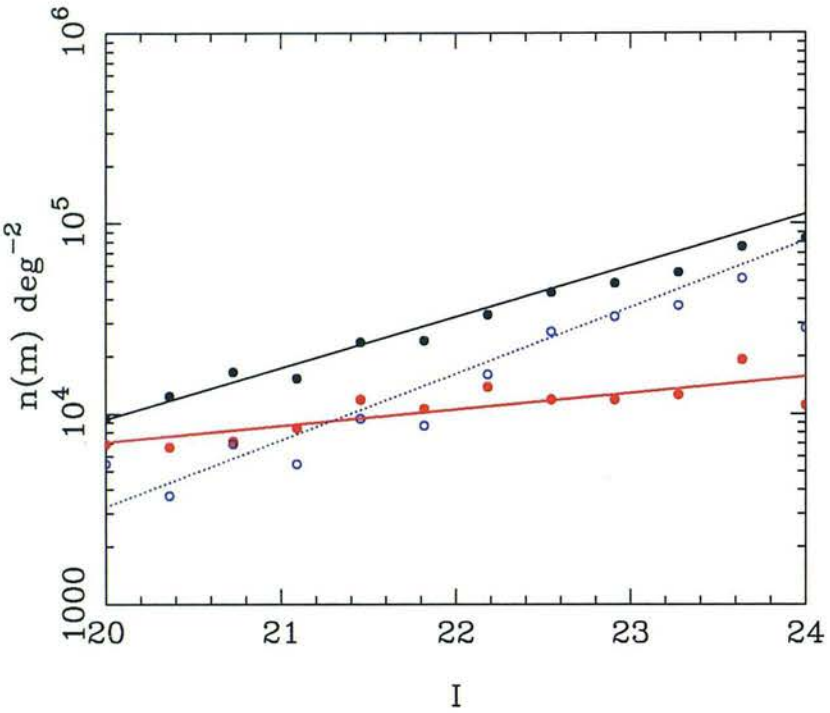

Figure 3.7: Magnitude distribution of all I-band galaxies (black dots), the red selected galaxies (red dots) and the blue background galaxies (blue open dots) from the Keck data of Smail et al. (1995). The lines are the best fits to the data.

The blue galaxies in the Keck sample, selected by $V - I < 1.0$, were also counted to verify the results of Broadhurst (1995) discussed in Chapter 2. The number counts, over the same range as the red counts, were found to be fitted by

$$\log_{10} n(\text{blue}) \approx 0.35\text{I} - 3.49 \tag{3.6}$$

resulting in $\beta = 0.88$. This is close to the lens invariant $\beta = 1$ and in good agreement with the findings of Broadhurst. Within the magnitude range $23 < \text{I} < 24$, chosen to minimize cluster contamination when applied to the A1689 data (see Figure 3.3), the number density of blue galaxies was found to be $n_0(\text{blue}) = 15.5$ galaxies per square arcminute.

Using this value of n_0(blue), the radial variation in number counts of blue galaxies

in the cluster field was plotted. Figure 3.8 shows the counts measured in the same annular bins illustrated in Figure 3.5. As anticipated, the blue galaxy counts do not show any obvious signs of a magnification signal. Deviations from the invariant value of $n/n_0 = 1$ mainly arise from contamination by faint cluster members (see Section 3.1.3). This suspicion is strengthened by the fact that the largest deviation occurs at a low radius where there is a higher number density of cluster objects. The lack of a significant change in radial blue counts also rules out the possibility that the depletion seen in the red counts is an effect of obscuration by dust.

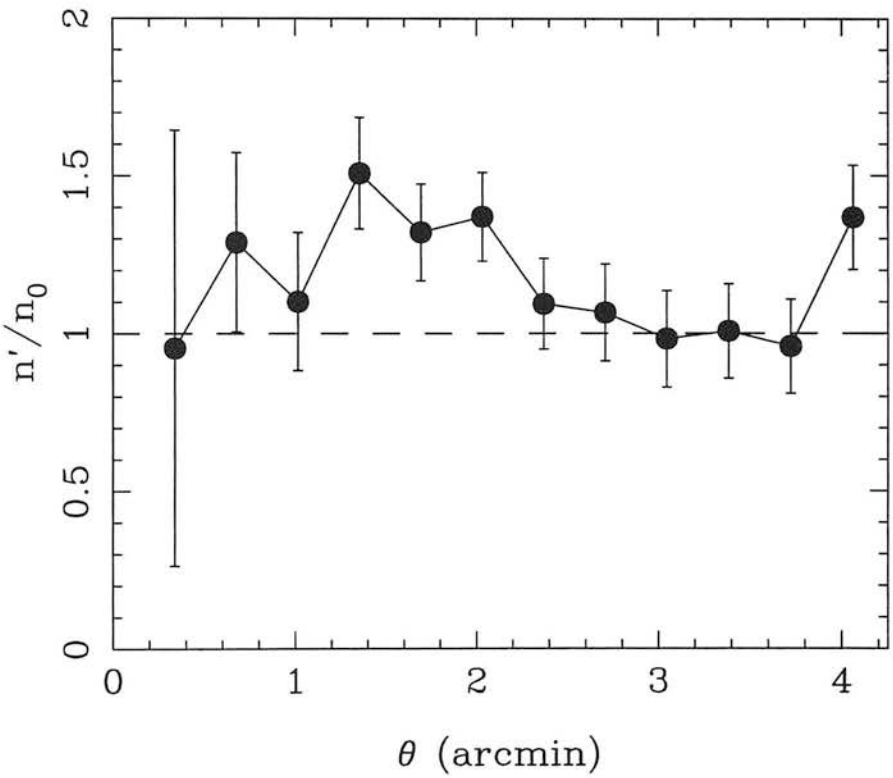

Figure 3.8: The radial profile of blue galaxy counts behind A1689. The background count density is $n_0 = 15.5$ galaxies per square arcminute. Deviations from the invariant value (shown by the dashed line) arise mainly from contamination by faint blue cluster members.

3.2.3 Clustering properties of the background population

With the number density and slope of the background galaxies established, the final piece of information required for the likelihood analysis of Chapter 2 is the amplitude of clustering of the background galaxies, ie. $\sigma(z)$ in equation (2.53).

The amplitude of clustering of I-band galaxies and its dependence on redshift was estimated from the results of Le Févre et al (1996) using the Canada-France Redshift Survey (CFRS). Le Févre et al. find that there is little difference between the clustering properties of red and blue populations of galaxies for $z > 0.5$, implying that the populations were well mixed at this epoch. Their findings lead to the result that the clustering variance of I-band galaxies scales according to (Taylor & Dye 1998),

$$\sigma^2(z) = 10^{-2} z^{-2.8} (\theta/1')^{-0.8} \tag{3.7}$$

within an observed circular area of radius θ(arcmin). To calculate the variance in a given annulus, θ is simply chosen to be the radius of a circle with the same area as the annulus.

Since the likelihood analysis in this chapter is concerned only with quantities integrated over redshift, a value of $\sigma(z)$ averaged over the redshift distribution of background sources must be obtained. In practice, the range of values of σ calculated for a realistic range of redshifts has a negligible effect on the maximum likelihood result for κ. The error on κ however is affected although only a weak dependence is exhibited; the change in width of the likelihood curves for κ in each annulus of Figure 3.5 over the redshift range $z = 0.4 - 1.0$ was found to be less than 15% in every case.

Crampton et al (1995), again using CFRS data, show that galaxies selected by $V - I > 1.6$ have a median redshift of $z \approx 0.8$. In light of the above findings, this redshift was used in equation (3.7) to provide the value of σ in the subsequent analysis.

3.2.4 Reconstruction of radial κ profile

From the radial number counts presented in Section 3.2.1, application of the likelihood analysis in Chapter 2 allowed calculation of the radial variation of A1689's magnification shown in Figure 3.9.

The errors in Figure 3.9 correspond to a confidence level of 1σ and are derived from the width of the likelihood curve as described by equation (2.41). Shot noise and background galaxy clustering are therefore incorporated. The dashed line is the behaviour of the same isothermal sphere model as that used in Figure 3.6. The underdensity at $\theta = 2.4$ in the number counts is manifested as the peak seen at the same

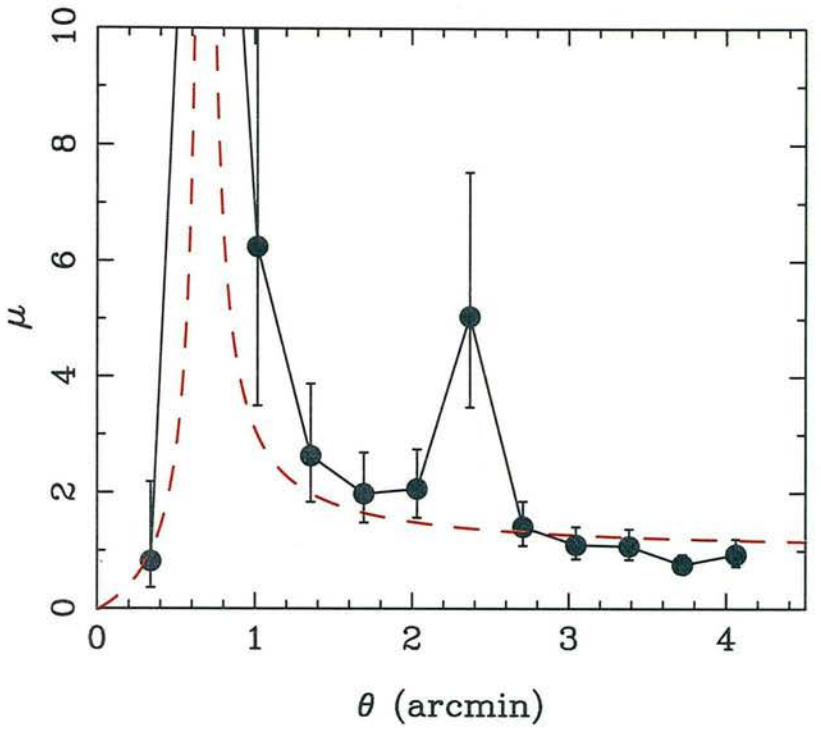

Figure 3.9: Radial variation of A1689's maximum-likelihood-derived magnification. The errors correspond to a confidence level of 1σ and allow for clustering and shot-noise. The dashed line shows the behaviour of the isothermal model in Figure 3.6.

radius in the magnification. It can be seen that this peak is approximately 2σ above the magnification predicted by the isothermal sphere. The bin at $\theta \simeq 0.7$ which coincides with A1689's critical line radius has an infinite magnification as expected (in fact, the analysis returns an arbitrarily large value which explains why the lines drawn from neighbouring data points are not completely vertical).

Application of the κ estimators in equation (2.66) enables an upper and lower bound for κ to be plotted. The parity required by both estimators was flipped about the bin at $\theta \simeq 0.7$ (the party in this bin is not a concern since $\kappa_+ = 1$ and $\kappa_- = 0.5$ regardless of parity when the magnification is infinite). The resulting profile is shown in Figure 3.10 where the red shaded area represents the region in which κ is bounded. Shot noise and clustering contribute to the uncertainty which corresponds to the area shaded in orange. The solid curve shows the same isothermal model.

Away from the cluster centre, the isothermal and sheet estimators give a consistent

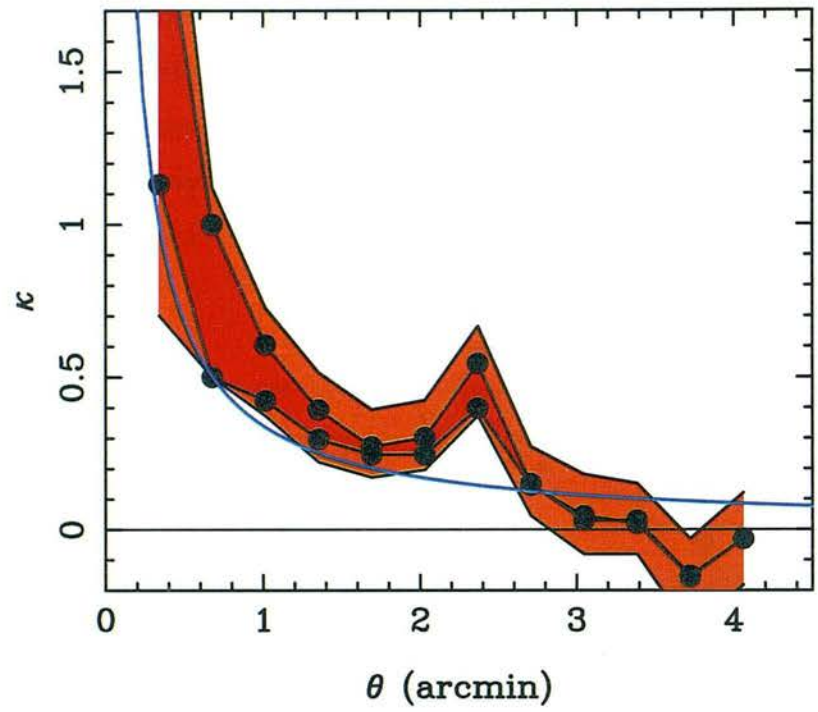

Figure 3.10: Radial profile of κ for A1689. The area shaded in red is the region in which κ is bounded while the orange area represents 1σ uncertainties from shot noise and clustering. The blue curve is the isothermal sphere model normalised to the critical line at $\theta = 0.75'$.

result for κ although this becomes dominated by noise after $\theta \simeq 3'$. Within the region $\theta < 2.8'$, a significant mass detection is measured. The peak at $\theta \simeq 2.4'$ again comes from the peak found in the magnification at this radius. No evidence of this peak is found in the data of Kaiser (1996); correlation introduced by the shear method may have suppressed such a structure. Repeating the procedure for annuli centred on different positions about the peak of the cluster light distribution gave a weaker, less significant mass profile in each case. This would be expected if the peak of the mass density were associated with that of the light.

3.2.5 Axially Symmetric Lens Solution for Abell 1689

With the radial magnification data of the previous section, the axially symmetric (non-local) solution for κ was computed. The parities in equation (2.84) were set to $\mathcal{R} = \mathcal{P} = 1$ in every annulus except the first which, lying inside the critical line, was

assigned $\mathcal{P} = -1$. With an initial shear chosen as $\gamma_1 = 0.3$, the profile plotted in Figure 3.11 was obtained.

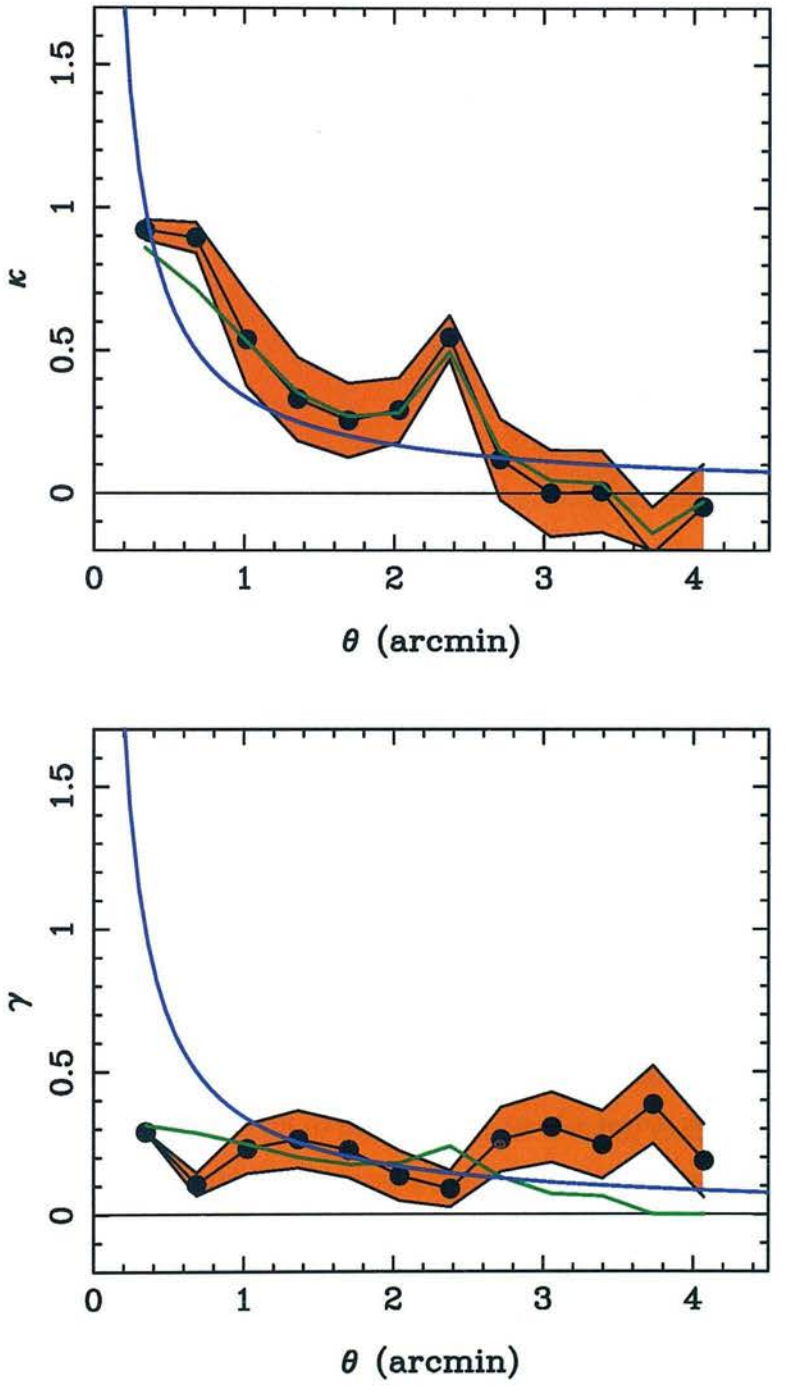

Figure 3.11: Radial profiles of κ and γ calculated by solving the axially symmetric lens equation (2.84. In both plots, 1σ errors are represented by the orange area, the blue curve is the isothermal model and the green line is the local parabolic estimator, κ_c, with $c = 0.7$.

The choice of $\gamma = 0.3$ for the first annulus has a fairly insensitive effect on the

resulting profiles, only affecting the first two annuli. The uncertainty on the first shell is small because the result is chosen a priori. The act of averaging in the series solution means that errors do not strongly propagate to higher annuli. Again, a peak in the κ distribution is seen at $\theta = 2.4'$, this time with the shear accounted for.

Estimation of γ shows a somewhat flat profile with $\gamma \simeq 0.25 \pm 0.15$ over most of the range. A slight increase beyond $\theta = 2.4'$ is due to the spike in the κ profile at that radius. This increase is not measured in the angle averaged measurements of γ by Kaiser (1996), where the mean shear is $\gamma = 0.15 \pm 0.05$. Errors in both profiles were derived from the propagation of errors from the magnification data.

3.2.6 Reconstruction of κ in 2D

Instead of binning the red background galaxies in annuli to obtain a radial profile, one can apply a rectangular grid to the field and count the number of galaxies in each grid bin. Using the same likelihood analysis, the magnification can be determined for every bin in the field. Applying one of the κ estimators then allows the two dimensional distribution in κ to be visualised.

The choice of grid dimensions determine the overall resolution and signal to noise of the κ distribution. Grid bins which are too small give rise to shot noise domination whereas grid bins which are too large produce a poor resolution. For the data presented here, an optimal size of 10×10 bins was chosen. The two dimensional distribution of red background galaxy number counts is shown in Figure 3.12. The underlying 10×10 resolution was Gaussian smoothed with a scale length of 0.35 arcmin onto a finer grid. The figure shows that a readily apparent lack of counts is seen in the vicinity of the critical line located by the white dashes.

The corresponding κ distribution produced with the sheet estimator is shown in the top half of Figure 3.13. Reversing the parity of bins lying interior to the critical line shown in Figure 3.12, the 10×10 grid was Gaussian smoothed onto a finer grid with a smoothing scale of 0.35 arcmin. In the lower half of Figure 3.13, the distribution of '1σ' maximum likelihood errors is shown. These are reasonably constant across the field.

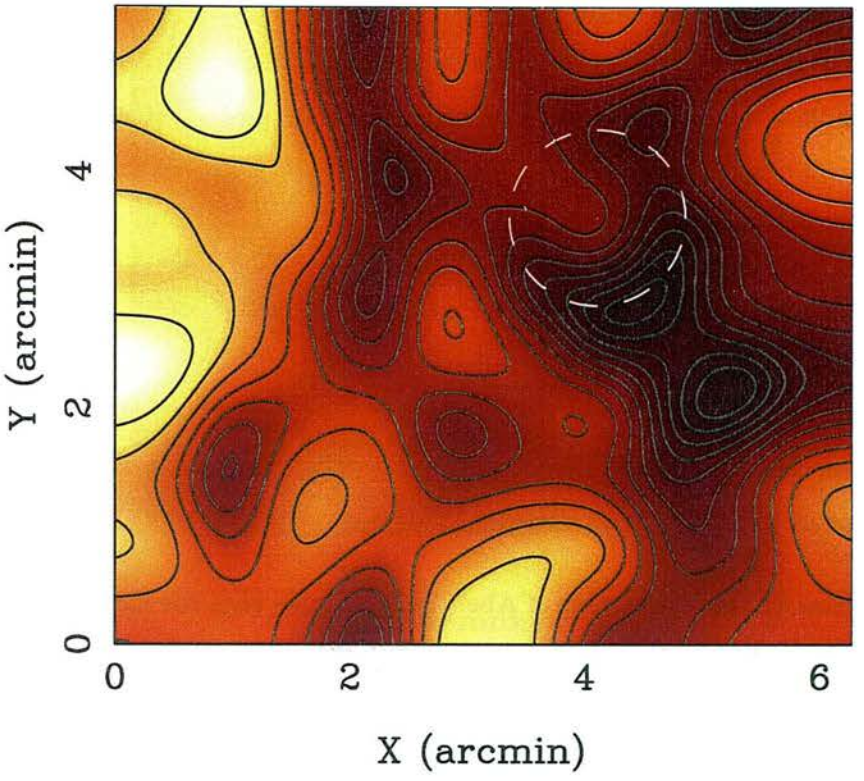

Figure 3.12: Distribution of red background galaxies behind A1689 smoothed with a smoothing scale of 0.35 arcmin. Darker shading signifies a lower density of counts. The contours are linearly spaced by $\Delta n = 1.45$ objects per square arcminute with a maximum and minimum number density of 23.0 and 1.1 obj/sq. arcmin respectively. The dashes show the approximate position of the critical line. Orientation is the same as that in Figure 3.1.

The peak in κ located at the cluster centre has a value of $\kappa \simeq 1.4$, slightly less than that predicted from the local radial profile due to the effects of smoothing. The structure seen extending from the cluster centre to the south-west has a value of $\kappa \simeq 0.9$ which when compared to the same vicinity in the error map shows that it is at least a 3σ detection.

Given that X-ray observations of A1689 with the ROSAT satellite in the $0.5 - 2$ keV band (Wang & Ulmer 1997, Daines et al 1998) show a relatively compact and smooth 2D flux distribution, this mass extension is a little surprising. There are several possible explanations for this. One is that the lens detected structure truly is a clump of dark matter, not followed by the X-ray emitting cluster gas. An alternative might be that the source number counts attributed to the structure are simply a particularly rare underdensity occurring as a result of clustering. Deeper observations would clearly

answer this conundrum. A shear analysis performed by Kaiser (1996) on the same data used in this chapter also hints at a substructure towards the south-west although this is a little tenuous. Claims of a consistent detection of substructure are therefore withheld.

3.2.7 Transforming κ to Mass Surface Density

The definition of κ in equation (2.9) indicates that in order to obtain the real mass surface density Σ, the value of $\Sigma_{\rm CR}$ is required. The angular diameter distance to the lens and source galaxies must therefore be known.

Spectroscopic measurements of Abell 1689 indicate that the mean redshift of cluster members is $\overline{z} = 0.184$ (TCG). In converting this to an angular diameter distance, a particular cosmology must be assumed. Fortunately, this does not pose much of a dilemma since as BTP have shown, the cosmological model chosen has only a small effect on the final derived surface mass density. An Einstein-de-Sitter universe is therefore assumed for simplicity so that the angular diameter distance of an object is related to its redshift using equations (1.36) and (1.43) to give

$$D_a(z) = \frac{2c}{H_0}\left(\frac{1-(1+z)^{-1/2}}{1+z}\right). \tag{3.8}$$

The angular diameter distance between two objects at redshifts z_1 and z_2 in this case is

$$D_a(z_1,z_2) = D_a(z_2) - \frac{1+z_1}{1+z_2}D_a(z_1) = \frac{2c}{H_0}\left(\frac{(1+z_1)^{-1/2}-(1+z_2)^{-1/2}}{1+z_2}\right). \tag{3.9}$$

Substituting these relations into equation (2.9) shows that the mass surface density is

$$\Sigma = \kappa\frac{cH_0}{8\pi G}\frac{(1-(1+z_S)^{-1/2})(1+z_L)}{(1-(1+z_L)^{-1/2})((1+z_L)^{-1/2}-(1+z_S)^{-1/2})} \tag{3.10}$$

where z_S and z_L are the redshifts of the background galaxies and the cluster respectively. The redshift of the background galaxies is again taken to be $z \approx 0.8$ from the work by Crampton et al (1995). Equation (3.10) with $z_L = 0.184$ and $z_S = 0.8$ therefore yields

$$\Sigma = 5.9 \times 10^{15}\kappa[h{\rm M}_\odot{\rm Mpc}^{-2}]. \tag{3.11}$$

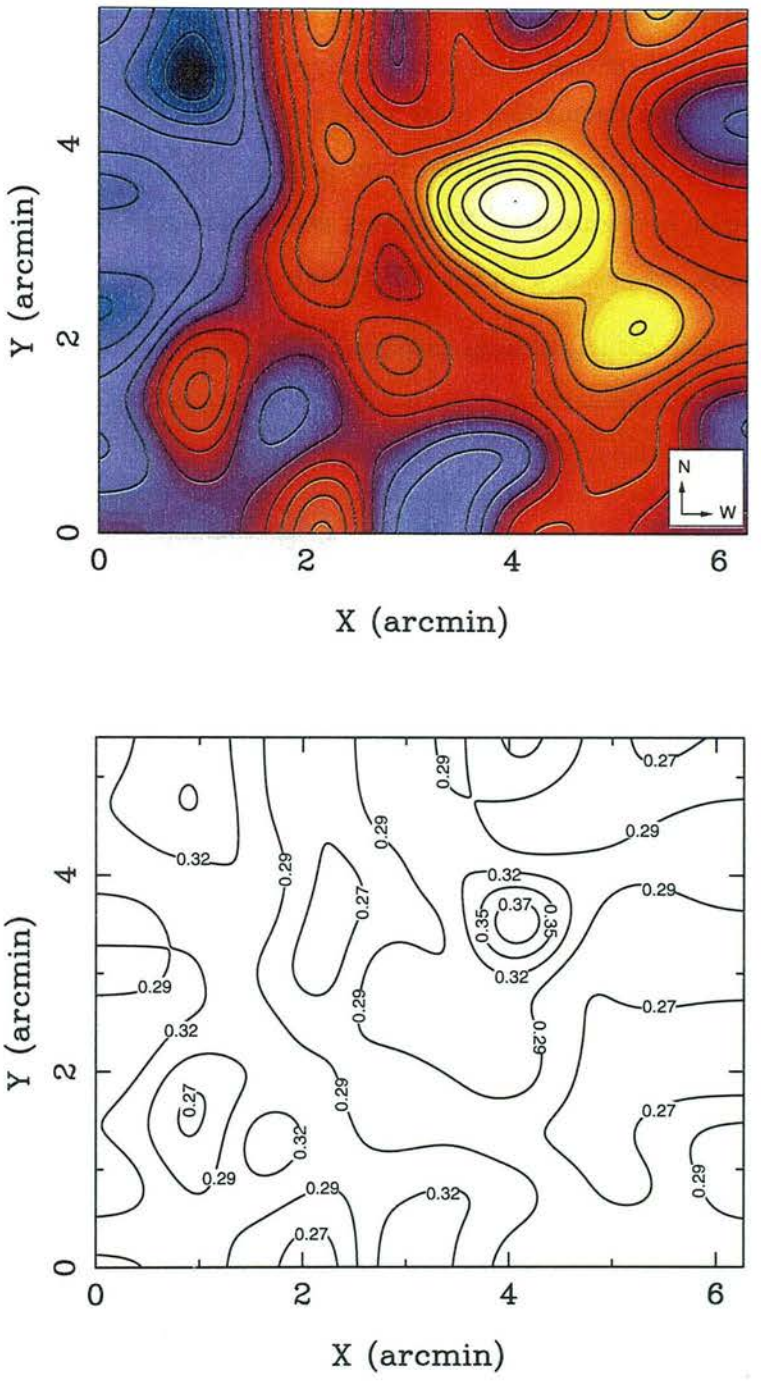

Figure 3.13: *Top*: Reconstructed surface mass density of Abell 1689 from the red background galaxy population using the sheet estimator. Light regions are high density. A parity flip about the critical line in Figure 3.12 has been applied. The maximum surface mass density is $\kappa = 1.4$, at $(4.02', 3.41')$, consistent with the peak in the light distribution. The minimum is $\kappa = -0.4$. A linear spacing of $\Delta\kappa = 0.12$ separates the 15 contours. The map is Gaussian smoothed with a smoothing length of 0.35 arcmin. *Bottom*: 1σ error distribution on κ.

3.2.8 Projected Mass

Figure 3.14 shows the cumulative mass profile of A1689 calculated from the axi-symmetric solution. Errors are derived from propagation of the 1σ uncertainty on the κ profile shown in Figure 3.11. The green line is the cumulative mass calculated using the parabolic estimator with $c = 0.7$. Clearly, good agreement is seen.

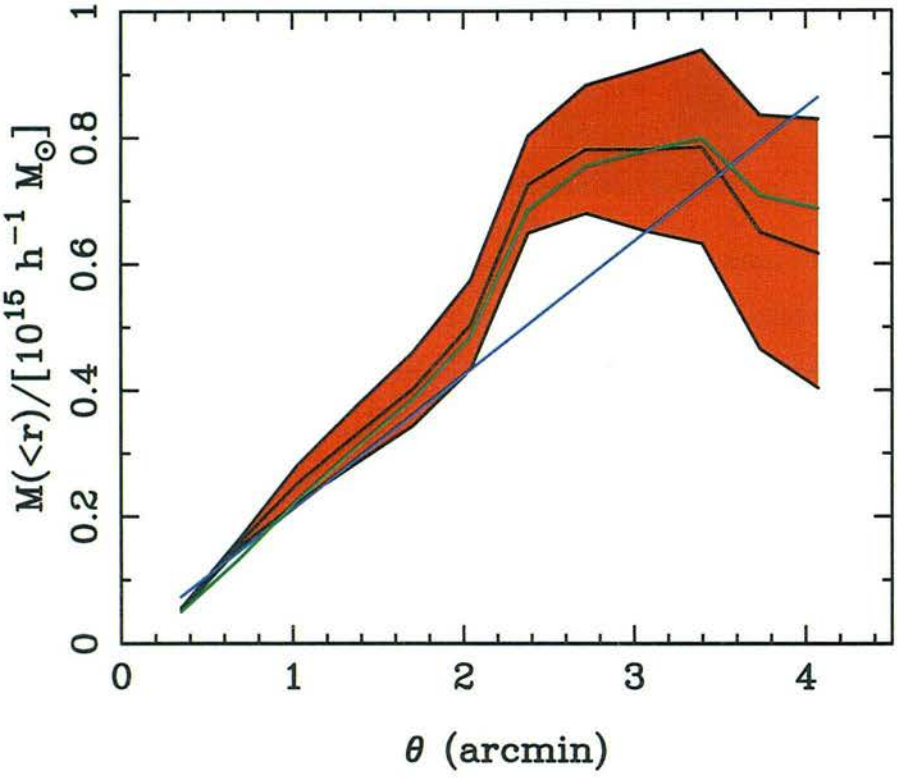

Figure 3.14: The cumulative mass profile of Abell 1689, calculated from the axi-symmetric solution. Orange shading shows the 1σ uncertainty propagated from errors in the κ profile of Figure 3.11. The green line shows the cumulative mass estimated from the parabolic estimator with $c = 0.7$. The isothermal fit to the critical line, similar to the shear results of Kaiser (1996) and Tyson & Fischer (1995) is plotted as the blue line.

Using the axi-symmetric solution, a projected mass interior to $0.24h^{-1}$Mpc ($\equiv 2.1'$ at the cluster redshift) of

$$M_{2d}(< 0.24h^{-1}\mathrm{Mpc}) = (0.50 \pm 0.09) \times 10^{15} h^{-1}\mathrm{M}_{\odot} \tag{3.12}$$

is predicted. Furthermore, this mass scales as

$$M_{2d}(< R) \simeq 3.5 \times 10^{15} (R/h^{-1}\mathrm{Mpc})^{1.3} h^{-1}\mathrm{M}_{\odot} \tag{3.13}$$

for $R < 0.32h^{-1}$Mpc, similar to that for an isothermal sphere which scales as $M \propto R$. Beyond $R = 0.32h^{-1}$Mpc, noise begins to dominate and hence the only obtainable result is the upper bound that $\kappa < 0.1$.

Error Analysis

Excluding errors computed from the likelihood analysis, three other sources of error must be taken into consideration. These are the uncertainty on the number count slope, β, the background count normalisation, n_0 and the redshift assumed for the red background galaxy population.

Error propagation shows that the combined fractional uncertainty in Σ is

$$\left(\frac{\delta\Sigma}{\Sigma}\right)^2 = 0.14\sigma_{z_S}^2 + \left(\frac{1-1/\kappa}{2(\beta-1)n_0}\right)^2 \sigma_{n_0}^2 + \left(\frac{\ln(1-\kappa)(1-1/\kappa)}{\beta-1}\right)^2 \sigma_\beta^2. \tag{3.14}$$

The uncertainty in source redshift may be estimated from the width of the CFRS red galaxy distribution by Lilly et al (1995). This gives a value of $\sigma_{z_S} \simeq 0.4$. Using the value of $\kappa = 0.5$, when combined with the errors for β and n_0 gives

$$\frac{\delta\Sigma}{\Sigma} = 0.24. \tag{3.15}$$

Including the errors due to shot noise and clustering from the likelihood analysis, the total uncertainty in the projected mass estimate above is therefore 30%.

3.2.9 Comparison and Discussion of Results

The work of Tyson & Fischer (1995) using shear analysis concludes that out to a radius of $R = 0.4h^{-1}$Mpc, the projected mass of A1689 follows an isothermal profile given by,

$$M_{2D}(< R) = (1.8 \pm 0.1) \times 10^{15}(R/h^{-1}\mathrm{Mpc})h^{-1}\mathrm{M}_\odot. \tag{3.16}$$

The mass contained within $R = 0.24h^{-1}$Mpc they therefore predict as

$$M_{2D}(< 0.24h^{-1}\mathrm{Mpc}) = (0.43 \pm 0.02) \times 10^{15}h^{-1}\mathrm{M}_\odot. \tag{3.17}$$

Kaiser (1996), using shear information measured from the images presented in this chapter finds again that A1689 is well fitted by an isothermal profile, very similar to

that by Tyson & Fischer, to give

$$M_{2D}(< 0.24h^{-1}\mathrm{Mpc}) = (0.43 \pm 0.04) \times 10^{15} h^{-1}\mathrm{M}_{\odot}. \tag{3.18}$$

These results are in good agreement with the projected mass interior to $R = 0.24h^{-1}$Mpc of $M_{2D} = (0.50 \pm 0.09) \times 10^{15} h^{-1}\mathrm{M}_{\odot}$ calculated in this chapter.

Using the CDM ($\Omega_0 = 1$, $\sigma_8 = 0.54$) N-body cluster simulations of van Kampen (in preparation), projected mass can be transformed to an equivalent 3D mass, line of sight velocity dispersion and X-ray temperature. These transformations allow approximate comparison of the result of this work with the results from methods of mass determination other than lensing.

A measurement of the line of sight velocity dispersion of A1689 is made by TCG who find

$$\sigma_v(< 1.5h^{-1}\mathrm{Mpc}) = 2355^{+238}_{-183}\ \mathrm{kms}^{-1}. \tag{3.19}$$

From the simulations, the projected mass of this chapter transforms to a dispersion of

$$\sigma_v(< 1.5h^{-1}\mathrm{Mpc}) = 2200 \pm 500\ \mathrm{kms}^{-1}, \tag{3.20}$$

in good agreement with the results of TCG. This is most likely an overestimate of the true dispersion since the simulations take into consideration the effects of superposition of clusters, infall along filaments and interlopers. den Hartog & Katgert (1996) attempt to remove contamination from interlopers using the TCG data and arrive at a velocity dispersion of $\sigma_v = 1860\mathrm{kms}^{-1}$. This agrees with σ_v calculated from the projected mass assuming an isothermal sphere which from equation (A.5) is $\sigma_v = 1645 \pm 150\mathrm{kms}^{-1}$. An explanation for this discrepancy is that A1689 is composed of two or more clumps of matter along the line of sight, a possibility hinted towards by the TCG data. Taylor & Dye et al (1998) note that A1689 is well fitted by two isothermal spheres, one at a redshift of $z = 0.18$ with $\sigma_v = 1500\mathrm{kms}^{-1}$ and one at $z = 0.20$ with $\sigma_v = 750\mathrm{kms}^{-1}$ to give a total projected dispersion of $2300\mathrm{kms}^{-1}$.

A final comparison may be made with X-ray temperature studies of A1689. Mushotzky & Scharf (1997), using the ASCA X-ray satellite measure the X-ray temperature of A1689 as $T_X = 9.0^{+0.4}_{-0.3}$keV. This compares with the results of Yamashita (1994) who measures a temperature of 8.7keV and 7.2keV from the ASCA and Ginga satellites

respectively. These temperatures may be transformed into the quantity M_{500} using the following relationship derived by Evrard et al (1996) from cluster simulations,

$$M_{500} = 1.11 \times 10^{15} \left(\frac{T_X}{10\mathrm{keV}} \right)^{3/2} h^{-1}\mathrm{M}_{\odot}. \tag{3.21}$$

where M_{500} is the 3D mass within a radius defined by an over-density of $500\rho_c$. Taking the result by Mushotzky & Scharf, this predicts a mass of

$$M_{500} = (0.95 \pm 0.16) \times 10^{15} h^{-1}\mathrm{M}_{\odot}. \tag{3.22}$$

Transforming the projected lensing mass of this chapter to M_{500} using the simulations by van Kampen gives

$$M_{500} = (1.60 \pm 0.65) \times 10^{15} h^{-1}\mathrm{M}_{\odot}, \tag{3.23}$$

implying an X-ray temperature of $T_X = 12.7 \pm 3.4$keV. This is somewhat higher than the directly measured quantities.

The overall measured temperature of a system of clumps superimposed along the line of sight tends to be dominated by the temperature of the hottest clump. If the velocity dispersion of the largest isothermal sphere in the two-clump model for A1689 proposed by Taylor & Dye et al is taken, this transforms to an X-ray temperature of $T_X \simeq 7.0$keV, consistent with the lower measurement by Yamashita. This provides further evidence that A1689 is not an isolated mass. Bartelmann & Kolatt (1998) proceed a step further and suggest that such discrepancies between X-ray and lensing measurements enable calculation of cluster morphology.

3.3 Summary

This chapter has demonstrated the viability of cluster mass reconstruction using the integrated number count technique of Section 2.3.2. By applying colour cuts to V and I band data of objects in the field of Abell 1689, the background population of red galaxies has been selected. These sources are shown to exhibit a dilution in surface number density in regions of high magnification due to their relatively flat number count slope. The effects of non-linear source clustering and shot noise have been taken into consideration for the computation of a projected mass profile and 2D mass map of A1689.

Having derived an analytical relationship to provide a self-consistent mass and shear profile by assuming an axially symmetric mass distribution, a projected mass interior to $R = 0.24h^{-1}\mathrm{Mpc}$ of

$$M_{2D}(< 0.24h^{-1}\mathrm{Mpc}) = (0.50 \pm 0.09) \times 10^{15} h^{-1}\mathrm{M}_{\odot} \tag{3.24}$$

has been calculated. This is in good agreement with mass estimates resulting from shear analysis. A comparison with X-ray temperature and virial measurements suggests that A1689 is not isolated but probably composed of two or more structures along the line of sight.

Chapter 4

Self-Consistent Mass Reconstruction

4.1 Introduction

The production of self-consistent mass and shear maps is the ultimate aim of mass reconstruction using magnification information. Stated more quantitatively, the requirement is to devise a method which enables the calculation of mass and shear maps given an observed map of magnification such that at any point in the lens plane, equation (2.20) is satisfied.

The method of reconstruction using κ estimators in Section 2.3.4 dealt with this issue by assuming that the mass distribution of a real lens lies somewhere between that of a homogeneous sheet and an isothermal sphere. Since both of these mass distributions have analytical forms which relate γ in terms of κ, equation (2.20) allowed κ to be expressed in terms of magnification alone. The two relationships obtained in this way, (the κ estimators), provided an upper and lower bound on the mass value in a given area of the lens plane to produce either radial profiles or 2D distributions.

Section 2.3.4 showed that if a lens can be assumed to be axially symmetric, then a self-consistent solution can be obtained but only to arrive at a radial mass profile. Attempts to calculate a self-consistent 2D mass distribution using an iterative tech-

nique based on the derivation of κ and γ from the deflection potential only converge for sub-critical distributions. As section 2.3.4 discussed, rapid divergence occurs if critical lines are present.

The first half of this Chapter describes a method capable of computing a self-consistent solution for γ and κ in 2D with the presence of critical lines using only magnification data (see Dye & Taylor 1998). The second half demonstrates how the method's underlying theme can be applied to galaxy ellipticity measurements to provide a simplification to shear analysis.

4.2 Self-Consistent Magnification Analysis

The method is based on pixellization of the κ distribution as suggested by AbdelSalam et al (1998) who used it to estimate the mass of the cluster Abell 370. Their technique necessitates the measurement of multiple image positions of lensed objects which act as constraints on their mass solution. This proves to be a robust method but is limited to clusters which exhibit multiple image systems and requires good quality data. Furthermore, without knowledge of object redshifts, identification of matching multiple objects is difficult.

The basic, underlying pixellization methodology used by AbdelSalam et al is modified here for the incorporation of magnification data. The use of magnification data with this method ensures that its application is not limited to clusters which exhibit multiple images and can be used with data observed under poorer conditions.

4.2.1 Theory

Derivation of the method begins by dividing up the deflection potential in the lens plane into a rectangular grid of pixels. The continuous deflection potential of equation (2.11) is then pixellized such that

$$\psi(\boldsymbol{\theta}) \simeq \sum_m \psi_m(\boldsymbol{\theta}), \quad \psi_m(\boldsymbol{\theta}) = \kappa_m \Gamma_m(\boldsymbol{\theta}) \tag{4.1}$$

where the summation acts over all pixels, $\psi_m(\boldsymbol{\theta})$ is the contribution to the potential at $\boldsymbol{\theta}$ from pixel m and κ_m is the mean value of κ in that pixel. Comparison of equation

(4.1) with equation (2.11) shows that the term $\Gamma_m(\boldsymbol{\theta})$ is therefore given by

$$\Gamma_m(\boldsymbol{\theta}) = \frac{1}{\pi}\int_m \mathrm{d}^2\theta' \ln|\boldsymbol{\theta}-\boldsymbol{\theta}'|. \tag{4.2}$$

The integral here only acts over pixel m, explaining the origin of the approximation in equation (4.1). In order to arrive at this result, it has been assumed that κ varies only slowly over the pixel so that it may be treated as a constant and taken outside of the integral in equation (2.11). In practice, this condition is enforced anyway by the resolution of the gridded magnification data.

The next step is to derive a relationship for the pixellized shear. This is straightforwardly achieved through application of equation (2.16) to the pixellized potential in equation (4.1):

$$\begin{aligned}\gamma_1(\boldsymbol{\theta}) &= \frac{1}{2}\left(\partial_1^2-\partial_2^2\right)\sum_m \kappa_m\Gamma_m(\boldsymbol{\theta}) \\ \gamma_2(\boldsymbol{\theta}) &= \partial_1\partial_2\sum_m \kappa_m\Gamma_m(\boldsymbol{\theta})\end{aligned} \tag{4.3}$$

where $\partial_i \equiv \partial/\partial\theta_i$. For the purpose of simplification, the following quantities are defined,

$$\begin{aligned}a_m(\boldsymbol{\theta}) &= \frac{1}{2}\left(\partial_1^2-\partial_2^2\right)\Gamma_m(\boldsymbol{\theta}) \\ b_m(\boldsymbol{\theta}) &= \partial_1\partial_2\Gamma_m(\boldsymbol{\theta})\end{aligned} \tag{4.4}$$

and the summation convention is adopted so that the shear components may be written as

$$\gamma_1(\boldsymbol{\theta}) = \kappa_m a_m(\boldsymbol{\theta}), \quad \gamma_2(\boldsymbol{\theta}) = \kappa_m b_m(\boldsymbol{\theta}). \tag{4.5}$$

Equation (4.5) is an important result which demonstrates the main advantage of pixellization. The shear is now expressed in terms of the pixellized κ distribution and the quantities a_m and b_m which are readily evaluated. For square pixels, a_m and b_m have the following analytical form (see Appendix A.3):

$$\begin{aligned}a_m(\boldsymbol{\theta}) &= \frac{1}{\pi}\arctan\left[\frac{r_2^2-r_1^2}{(r_1^2+r_2^2)^2-1/4}\right] \\ b_m(\boldsymbol{\theta}) &= \frac{1}{2\pi}\ln\left[\frac{(r_1^2+r_2^2+1/2)^2-(r_1+r_2)^2}{(r_1^2+r_2^2+1/2)^2-(r_1-r_2)^2}\right]\end{aligned} \tag{4.6}$$

where r_1 and r_2 are the components of the vector $\boldsymbol{r} = \boldsymbol{\theta}-\boldsymbol{\theta}_m$ from position $\boldsymbol{\theta}$ to the centre of pixel m at $\boldsymbol{\theta}_m$. Figure 4.1 shows the geometry for clarification.

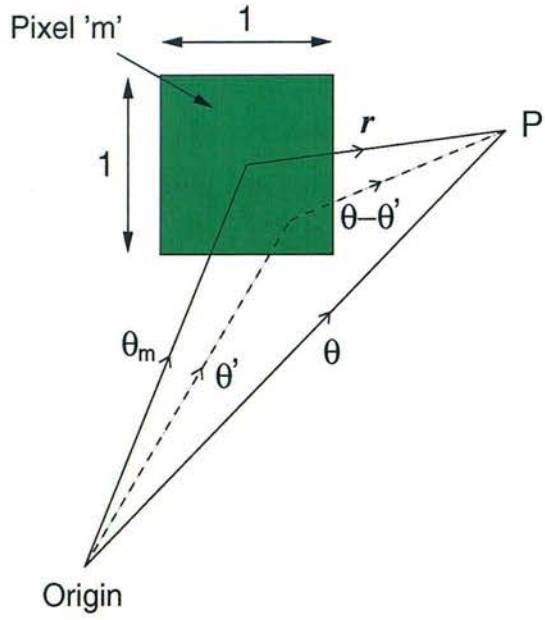

Figure 4.1: Clarification of the geometry used to calculate the contribution of the potential at point P from pixel m. The centre of pixel m is described by the vector $\boldsymbol{\theta}_m$ from the origin and by the vector $\boldsymbol{r}$ from point P. Integration over the pixel is carried out with respect to the variable vector $\boldsymbol{\theta}'$ (or $\boldsymbol{\theta} - \boldsymbol{\theta}'$). The units of length are chosen such that pixels measure 1 unit on a side.

The current formalism allows the determination of individual pixel contributions towards quantities at any continuous point P in the lens plane described by $\boldsymbol{\theta}$. This is a more general treatment than typically required in practice. Knowledge of quantities such as the shear or the potential are only usually required at points on the grid. The quantity $\boldsymbol{\theta}$ can therefore be pixellized in the same fashion as $\boldsymbol{\theta}_m$ such that $\boldsymbol{\theta} \Rightarrow \boldsymbol{\theta}_n$ with $\boldsymbol{\theta}_n$ locating only pixel centres. This means that the quantities $a_m(\boldsymbol{\theta})$ and $b_m(\boldsymbol{\theta})$ in equation (4.4) can be represented as the square matrices a_{mn} and b_{mn} with each index running from 1 to the total number of pixels.

In analogy with the quantities a_{mn} and b_{mn} for γ, one might expect to derive a similar result for κ. This can be attempted by applying equation (2.13) to the pixellized potential in equation (4.2):

$$c_{mn} = \frac{1}{2\pi}\nabla^2 \int_m \mathrm{d}^2\theta' \ln|\boldsymbol{\theta}_n - \boldsymbol{\theta}'| \tag{4.7}$$

where ∇^2 acts on $\boldsymbol{\theta}_n$. The identity $\nabla^2 \ln|\mathbf{x}| = 2\pi\delta^2(\mathbf{x})$ used in section 2.1.4 means

that

$$c_{mn} = \left\{ \begin{array}{ll} 1 & \text{if } m = n \\ 0 & \text{if } m \neq n \end{array} \right\} = \delta_{mn}. \tag{4.8}$$

In other words, c_{mn} is simply the Kronecker delta function. This is exactly as one would have hoped since the value of κ in any given pixel should not depend upon κ in any other pixel.

To bring magnification into the derivation, the pixellized versions of κ and γ are substituted into equation (2.20) so that,

$$\begin{aligned} \mathcal{P}\mu_n^{-1} &= (1-\kappa_n)^2 - \gamma_n^2 \\ &= (1-\kappa_n)^2 - (\kappa_m a_{mn})^2 - (\kappa_m b_{mn})^2 \end{aligned} \tag{4.9}$$

with summation implied over index m. Defining N as the total number of pixels in the grid means that equation (4.9) represents a set of N simultaneous equations which can be alternatively expressed as a vector equation. This is more readily seen by defining the $N \times N \times N$ matrix $\mathbf{G}$ whose elements are given by

$$\mathrm{G}_{pqn} = \delta_{pn}\delta_{qn} - a_{pn}a_{qn} - b_{pn}b_{qn} \tag{4.10}$$

to enable equation (4.9) to be written as

$$\mathbf{1} - 2\boldsymbol{\kappa} + \boldsymbol{\kappa}\mathbf{G}\boldsymbol{\kappa}^t - \boldsymbol{\mu}^{-1} = 0. \tag{4.11}$$

Here, $\boldsymbol{\mu}^{-1}$ is the N-dimensional vector of pixellized inverse magnification parity-signed values, $\boldsymbol{\kappa}^t$ is the transpose of the vector $\boldsymbol{\kappa}$ of pixellized values of κ and $\mathbf{1}$ is the vector $(1, 1, 1, \cdots)$. Once again, $\mathcal{P}$ is the image parity of Section 2.3.4.

Equation (4.11) is the first main result of this chapter. Given a measured distribution of inverse magnification values, κ can now be solved without the need for any assumptions concerning γ or the lens mass distribution. Having solved for κ, the corresponding shear distribution can be subsequently calculated from equation (4.5).

4.2.2 Application to Cluster Models

Verification of the method is obtained in this section through its application to two idealised cluster models. Starting with a predetermined κ distribution, the corresponding shear distribution is derived using the Fourier methods discussed in Section

2.3.4. To remove the effects of the Fourier transform's implicitly assumed periodic boundary conditions, the γ distribution is windowed to select only the relatively unaffected central region. The size of this window is determined by the accuracy required; a smaller window removes more of the boundary and hence a more accurate distribution remains.

Applying the same window to the initial κ distribution then allows the magnification distribution to be calculated using equation (2.20). The magnification is used in equation (4.11) which is then solved for κ. This solution for the κ distribution is finally used to determine the shear distribution from equation (4.5) so that a comparison can be made between the initial, predetermined κ and γ and the κ and γ computed using the pixellated method.

For both the truncated isothermal sphere and dumb-bell models which follow, sufficient accuracy was obtained using an initial Fourier grid of 512×512 pixels windowed to a grid size of 32×32 pixels.

Truncated Isothermal Sphere Model

The initial pixellated mass was distributed according to

$$\kappa = \frac{c}{r + r_0} \tag{4.12}$$

where r is the radial distance from the centre of the sphere, r_0 is the degree of truncation and c is a constant. The value of c determines the size of the critical line and indeed whether one exists. A larger value of c essentially gives a larger critical line radius. If c is smaller than a certain value dependent on the value of r_0, no critical line is present. (This is not true of a pure isothermal sphere in which case $r_{crit} = 2c$.)

Choosing c to result in a critical line which encompassed a central 2×2 pixel region, the distributions of κ and γ in the top third of Figure 4.2 were generated. After calculation of the inverse magnification from these and a reversal of parity in the inner critical line pixels, a solution for κ was obtained. The solved κ distribution together with the associated γ distribution is shown in the middle third of Figure 4.2. Both the predetermined and solved distributions shown here were Gaussian smoothed onto a finer grid from their underlying 32×32 grid of pixels with a smoothing scale

of one pixel width. This is purely to aid visual comparison. The residuals plotted in the bottom third of Figure 4.2 show the difference between the predetermined and the solved distributions as percentages of $(\kappa_{\rm init} - \kappa_{\rm solved})/\kappa_{\rm init}$. The contours were generated directly from the unsmoothed grid.

The residuals show that κ has been solved to an accuracy of less than one percent over most of the grid[1]. This is a negligible error in comparison to the noise typically encountered in practice from background clustering of objects, shot noise (Taylor & Dye et al 1998) and the uncertainties resulting from use of the κ estimator method (see van Kampen 1998). The recovered shear distribution is more affected although still fares better than γ calculated from uncorrected Fourier techniques which is typically $\sim 30\%$ worse.

The main contribution to these residuals is from boundary effects which arise from attempting to recover a nonlocal shear in a finite area. Much work has been carried out into the removal of such effects (see Squires & Kaiser 1996 and Seitz & Schneider 1995 for example) which, as has been demonstrated, have little impact on the recovered κ. In light of this, treatment of these boundary effects has not been considered in this work.

Dumb-bell Model

To test the method with a more general, asymmetric mass distribution, a dumb-bell model consisting of two offset, truncated isothermal spheres of different mass was used. Proceeding in exactly the same manner as with the isothermal model, the inverse magnification distribution was computed from the predetermined κ and γ distributions shown in the top third of Figure 4.3. After reversal of parity in the inner critical line regions, equation (4.11) was solved for κ. The solved κ and γ distributions are shown together with their residuals in the lower two thirds of Figure 4.3. As in the case of the isothermal model, the residuals between the predetermined and solved κ are typically less than one percent over most of the field. Once again, γ fares worse with a central two thirds of the field yielding residuals smaller than 10%.

[1] Analysis shows that this is true regardless of the grid dimensions.

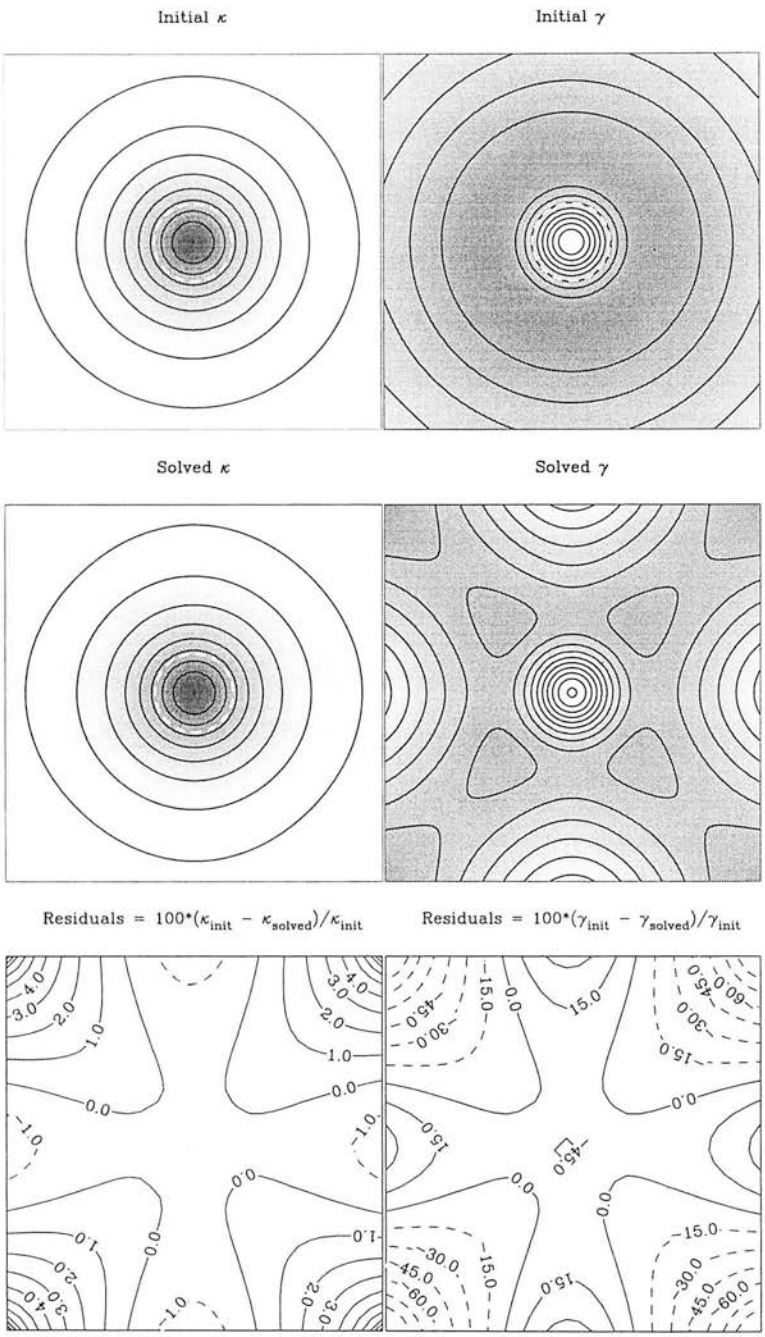

Figure 4.2: Truncated Isothermal Sphere Model: The predetermined (initial) κ and γ used to form the magnification distribution from which the solved κ and γ are derived. The distributions are smoothed from the 32×32 grid onto a 128×128 grid. White dashes show the position of the critical line. Contours are linearly spaced and set at the same levels in both κ plots and in both γ plots. Residuals are expressed as percentages of $(\kappa_{\rm init} - \kappa_{\rm solved})/\kappa_{\rm init}$.

4.2.3 Practical Considerations

The set of N simultaneous equations represented by equation (4.11) were solved iteratively using the Powell method as provided by the 'C05PCF' NAG library routine. A

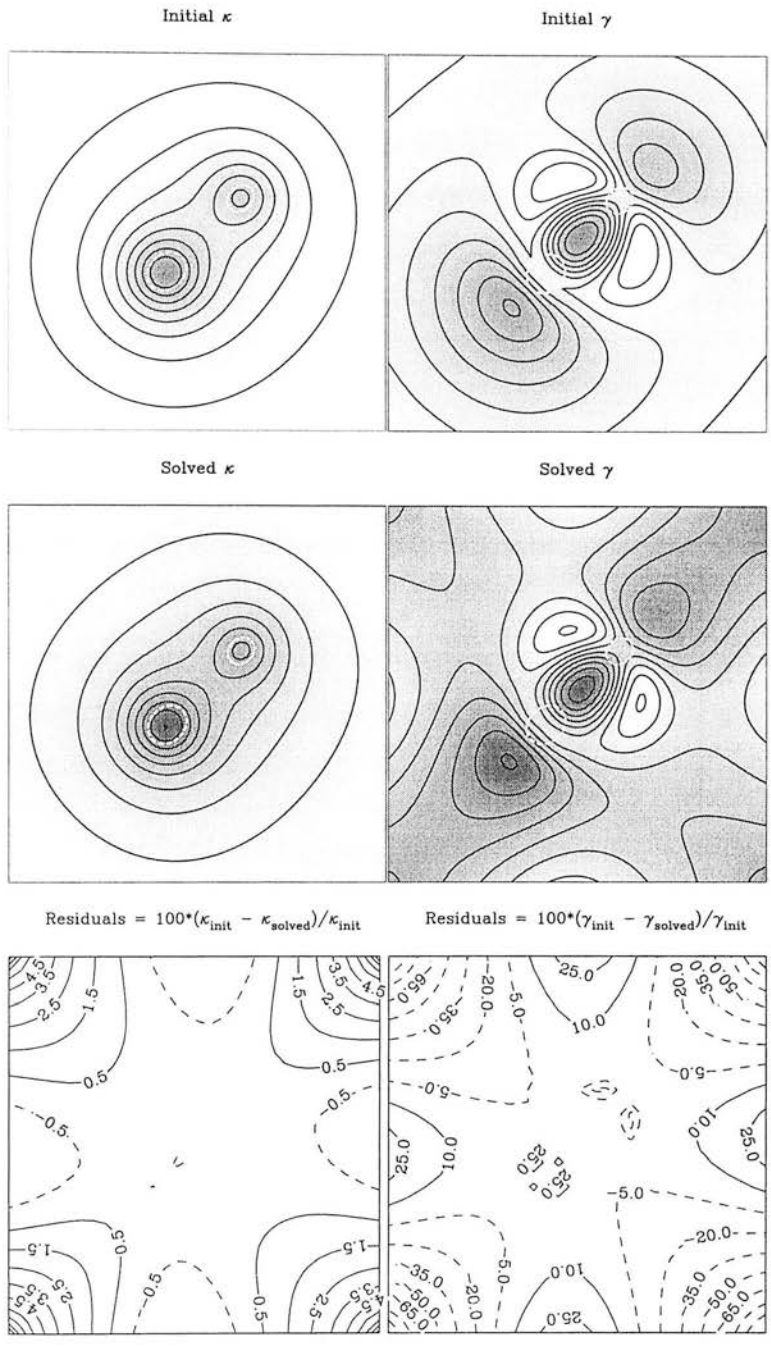

Figure 4.3: Dumb-bell Model: Predetermined (initial) κ constructed from two offset, truncated isothermal spheres of different mass. The critical lines are shown as white dashes. Distributions are smoothed from the 32×32 grid onto a 128×128 grid. Linearly spaced contours are set at the same levels for κ and at the same levels for γ.

solution for particularly fine grids can prove computationally intensive, especially on the amount of CPU memory used which scales as N^4. For typical grid resolutions used in practice however, this tends not to be a major concern. Figure 4.4 illustrates the

convergence of the solution to the isothermal model in Section 4.2.2 with $N = 1024$. Plotted against each iterative step is the RMS of equation (4.11). As the figure shows, rapid convergence occurs in the first few cycles of the iteration. The 30 steps shown took approximately one minute of CPU time on a 400Mhz DEC-Alpha workstation.

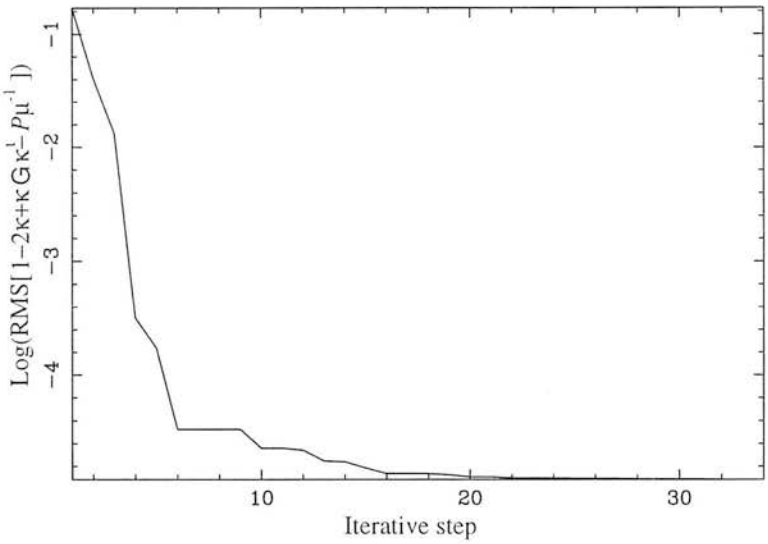

Figure 4.4: Convergence of κ solution for the isothermal lens in Section 4.2.2. Plateaus are attributed to the inner workings of the 'C05PCF' NAG library routine.

Since the Powell algorithm is an iterative process, an initial estimate of the solution is required to start from. The choice of this initial estimate proves to play no noticeable part in the solution obtained which acts to demonstrate the robust nature of this method. Distributions from all three κ estimators of Section 2.3.4 and even a completely flat distribution were used as initial estimates, all four giving the same converged result.

One feature of this method is that correct choice of pixel parity (especially for low grid resolutions) is important in order to achieve a sensible solution. The inappropriate assignment of parities to pixels manifests itself in a manner dependent on whether a pixel is wrongly assumed to lie inside or outside a critical line. If a pixel is assumed to lie inside a critical line when in actual fact it does not, κ in that pixel is overestimated. In the reverse situation, κ is either underestimated or the pixels immediately surrounding the incorrectly assigned pixel are overestimated.

Figure 4.5 demonstrates the effect of incorrect parity assignment for an isothermal sphere. The left hand side of this figure shows the consequence of attempting to reconstruct a sub-critical distribution assuming that the central pixel lies inside a critical line. The reconstructed value of $\kappa = 1.0$ in this central pixel compared to the value of $\kappa = 0.7$ in the predetermined central pixel shows that κ has been overestimated. The value of κ in the surrounding pixels is within one percent of the solution for κ which would have been obtained if the critical line was assumed absent.

The right hand side of Figure 4.5 shows the reverse situation. The predetermined distribution has a critical line which encompasses the central pixel with $\kappa = 1.3$. The pixels immediately to the north, south, east and west of this pixel have the value $\kappa = 1.1$. Attempting to reconstruct κ assuming the absence of a critical line results in the overestimated value of $\kappa = 1.2$ in these surrounding pixels. In this case, the reconstructed central pixel has $\kappa = 1.3$, agreeing with the predetermined value and hence the solved distribution is flatter than it should be.

The error in κ resulting from bad pixel parity assignment depends, as one would expect, on the severity of the mismatch. In situations where mass distributions are very nearly critical or only just critical, the incorrect assumption of the presence or absence of a critical line does not dramatically modify the reconstructed results. More extreme mismatches however, give larger discrepancies which could in principle be used to determine the accuracy of critical line positions. This would inevitably prove too ambiguous with real data and so critical line positions should always be determined from the observed image directly.

To ensure that the method does not break down with noisy data, Gaussian noise was added to the amplification. Errors in κ resulting from noise in the inverse amplification were found to propagate as one would expect from equation (4.11). For an isothermal lens, the expected result derived from simple error propagation of equation (2.20),

$$\delta\kappa = \frac{\delta\mu}{2\mu^2} \tag{4.13}$$

was recovered, indicating that pixellization does not lead to spurious noise properties.

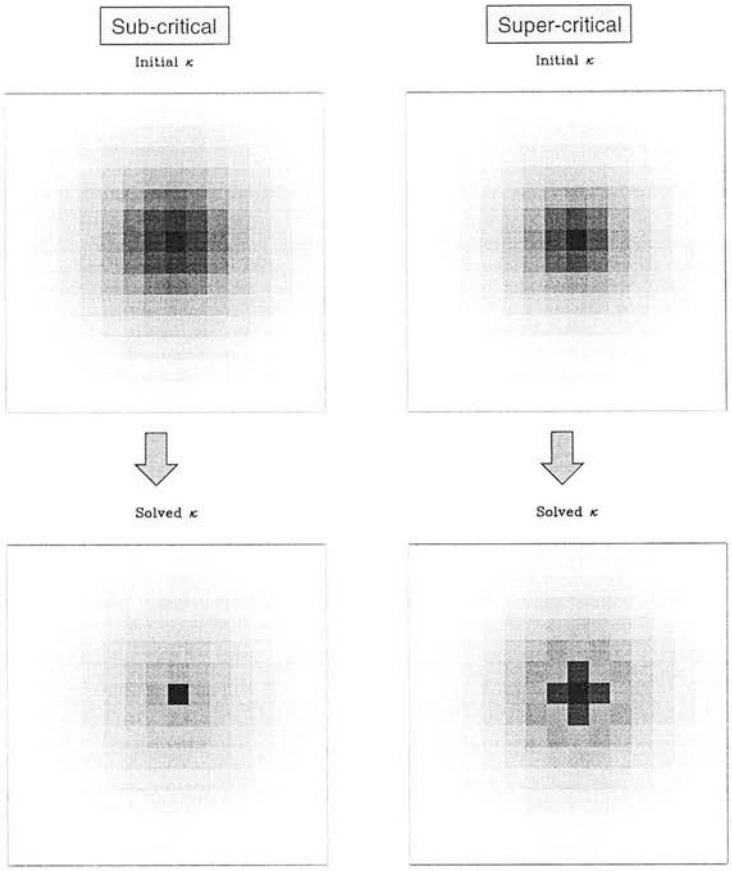

Figure 4.5: Effects of incorrect pixel parity assignment. Reconstruction of the predetermined, sub-critical isothermal sphere model on the left hand side results in an overestimation of κ by 40% in the central pixel under the incorrect assumption that this central pixel lies within a critical line. On the right hand, κ has been overestimated in the pixels immediately north, south, east and west by 10% due to the incorrect assumption that the central pixel is not inside a critical line.

4.2.4 Application to Abell 1689

The method was applied to the magnification data presented in Chapter 3 for the lensing cluster A1689. Using the same 10×10 grid of pixels and the same parity change in the 2×2 inner critical line region as before, the κ and γ distribution shown in Figure 4.6 was obtained.

Comparison with the mass density map illustrated in Figure 3.13 which was produced with the sheet κ estimator shows very similar structure as one would have expected. The value of κ at the peak calculated here is approximately 10% lower than

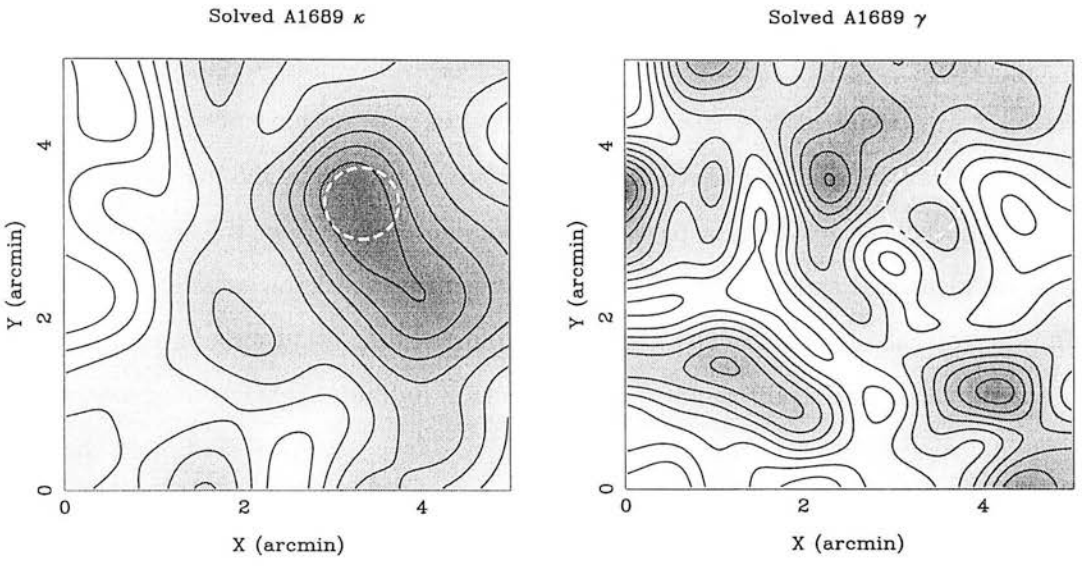

Figure 4.6: Self-consistent reconstruction of κ and γ for Abell 1689 using the magnification data presented in Chapter 3. The critical line is shown by the white dashes. Both distributions are Gaussian smoothed onto a finer grid. Maximum κ and γ values are $\kappa_{\max} = 1.2$ and $\gamma_{\max} = 0.5$ respectively. The 10 contours are linearly spaced by $\Delta\kappa = 0.15$, $\Delta\gamma = 0.15$.

the peak value in Figure 3.13 since the sheet estimator over-estimates κ as discussed in Section 2.3.4. This result has little effect on the total integrated mass of A1689 presented in Section 3.2.8 obtained with the axially symmetric 1D solution. The γ distribution is shown for completeness although undoubtedly suffers from boundary effects typically encountered in the models of Section 4.2.2.

4.3 Shear Analysis

Having shown that pixellization enables the accurate reconstruction of surface mass density from magnification data, the next logical step is to attempt to obtain a similar success with shear data. This section details how this is achieved by providing a simplified and pixellized alternative to the ellipticity equation presented by Kaiser (1995, K95 hereafter).

4.3.1 Derivation of Pixellized Ellipticity Equation

Section 2.2.2 discussed how shear analysis extracts information about the mass distribution of an intervening lens from the measurement of lens-induced distortion of background galaxy images. In particular, it was noted how only the ratio $\gamma/(1-\kappa)$ could be measured in bins across the field of view through knowledge of the average ellipticity of galaxies contained within those bins. These ellipticities are in practice obtained by calculating the quadrupole moments of individual galaxy images to arrive at the ellipticity parameters e_{ij} (Valdes, Tyson & Jarvis 1983). K95 showed that the ellipticity parameters are related to surface mass density and shear via,

$$e_{ij} = \frac{\gamma_{ij}}{1-\kappa}, \quad \gamma_{ij} = \begin{pmatrix} \gamma_1 & \gamma_2 \\ \gamma_2 & -\gamma_1 \end{pmatrix}. \tag{4.14}$$

One way of solving this for κ in the weak lensing regime is to follow the approach of Kaiser & Squires (1993) as outlined in Section 2.2.2 or the generalization of this to the strong regime by K95. It might be expected that an alternative way of proceeding would be to pixellize equation (4.14) and use the result of equation (4.5) to give,

$$e_{(1)n} = e_{(1)n}\kappa_n + \sum_m a_{mn}\kappa_m, \quad e_{(2)n} = e_{(2)n}\kappa_n + \sum_m b_{mn}\kappa_m \tag{4.15}$$

where $e_{(i)} = \gamma_i/(1-\kappa)$ and summation is shown for clarity. Equations (4.15) are essentially matrix equations which can therefore be rearranged using matrix inversion to give

$$\boldsymbol{\kappa} = \mathbf{e}_{(1)}\mathbf{H}_{(1)}^{-1} = \mathbf{e}_{(2)}\mathbf{H}_{(2)}^{-1}. \tag{4.16}$$

In this equation, $\mathbf{e}_{(1)}$ and $\mathbf{e}_{(2)}$ are the vectors formed from the ellipticity parameters and the elements of the square matrix $\mathbf{H}$ are given by

$$\mathrm{H}_{(1)mn} = e_{(1)n}\delta_{mn} + a_{mn}, \quad \mathrm{H}_{(2)mn} = e_{(2)n}\delta_{mn} + b_{mn}. \tag{4.17}$$

Unfortunately, this does not prove to be a viable means of reconstructing κ. In calculating matrix condition numbers, one finds that the matrix $\mathbf{H}_{(1)}$ is singular and the matrix $\mathbf{H}_{(2)}$ is ill-conditioned. It is perhaps not too surprising that this is the case. As equation (4.16) shows, this method allows determination of κ independently from each ellipticity parameter. The shape and orientation of a galaxy can only be

uniquely described by both ellipticity parameters and so one would expect that the same should follow for the determination of κ.

The criterion that both ellipticity parameters should be used for a determination of κ can be satisfied by applying the result of K95 that

$$\partial_i \kappa = \partial_j \gamma_{ij}, \tag{4.18}$$

to equation (4.14) instead of its direct pixellization. This gives

$$\partial_i \kappa = e_{ij} \partial_j (1-\kappa) + (1-\kappa) \partial_j e_{ij} \tag{4.19}$$

and making the substitution that $\partial_i \kappa = -\partial_i (1-\kappa)$ and $\partial_j = \delta_{ij} \partial_i$ allows this to be written,

$$(\delta_{ij} + e_{ij}) \partial_j \ln(1-\kappa) + \partial_j e_{ij} = 0. \tag{4.20}$$

Since $B^{-1}_{ki} B_{ij} = \delta_{kj}$ for any square matrix $\mathbf{B}$, multiplying the equation above by $(\delta_{ki} + e_{ki})^{-1}$ yields

$$\partial_k \ln(1-\kappa) + (\delta_{ki} + e_{ki})^{-1} \partial_j e_{ij} = 0. \tag{4.21}$$

Relabelling indices and using the fact that the right hand side of this equation is the derivative of a logarithm gives the result that

$$\partial_i \ln(1-\kappa) + \partial_j \ln(\delta_{ij} + e_{ij}) = 0. \tag{4.22}$$

The logarithm of the matrix in this equation is interpreted using the result that for a square matrix $\mathbf{B}$ and the identity matrix $\mathbf{I}$, the following Taylor expansion holds true

$$\ln(\mathbf{I} + \mathbf{B}) = \mathbf{B} - \frac{1}{2}\mathbf{B}^2 + \frac{1}{3}\mathbf{B}^3 - \cdots \tag{4.23}$$

in analogy with the result if $\mathbf{B}$ and $\mathbf{I}$ had been scalar quantities. Applying this Taylor expansion to the logarithm in equation (4.22) and noting from the definition of the ellipticity parameters in equation (4.14) that

$$(e_{ij})^2 = (e_1^2 + e_2^2)\delta_{ij} = e^2 \delta_{ij}, \quad (e_{ij})^3 = e^2 e_{ij}, \quad \cdots \tag{4.24}$$

gives

$$\begin{aligned} \ln(\delta_{ij} + e_{ij}) &= e_{ij} - \frac{1}{2} e^2 \delta_{ij} + \frac{1}{3} e^2 e_{ij} - \cdots \\ &= \left(\sum_{n=0}^{\infty} \frac{e^{2n+1}}{2n+1} \right) \frac{e_{ij}}{e} - \left(\sum_{n=1}^{\infty} \frac{e^{2n}}{2n} \right) \delta_{ij} \\ &= \frac{1}{2} \ln\left(\frac{1+e}{1-e} \right) \frac{e_{ij}}{e} - \frac{1}{2} \ln(1-e^2) \delta_{ij}. \end{aligned} \tag{4.25}$$

The logarithmic substitutions made here for the infinite series are only valid provided that $e < 1$. Inserting equation (4.14) into the magnification equation (2.20) shows that

$$\mu^{-1} = |(1-\kappa)^2(1-e^2)|. \tag{4.26}$$

Hence the parity changes when $e^2 > 1$. Since e_{ij} and e_{ij}^{-1} are observationally indistinguishable and flip from one to another whenever there is a parity change (Kaiser 1995), the requirement that $e < 1$ is satisfied by simply inverting the ellipticity matrix when a critical line is crossed.

Substituting the result of equation (4.25) into equation (4.22) and rearranging then gives

$$\ln(1-\kappa) = \frac{1}{2}\ln(1-e^2) - \frac{1}{2}\partial^{-2}\partial_i\partial_j\frac{e_{ij}}{e}\ln\left(\frac{1+e}{1-e}\right). \tag{4.27}$$

where ∂^{-2} is the inverse Laplacian in 2D. An expression for the term $\partial^{-2}\partial_i\partial_j e_{ij}$ in this equation is derived directly by operating on equation (4.14):

$$\begin{aligned}\partial^{-2}\partial_i\partial_j e_{ij} &= \partial^{-2}(\partial_1^2 - \partial_2^2)e_1 + 2\partial^{-2}\partial_1\partial_2 e_2 \\ &= (\partial_1^2 - \partial_2^2)\partial^{-2}e_1 + 2\partial_1\partial_2\partial^{-2}e_2.\end{aligned} \tag{4.28}$$

The next step is to apply pixellization so that as in Section 4.2.1 the following transformations take place:

$$\begin{aligned}(\partial_1^2 - \partial_2^2)\partial^{-2}e_1(\boldsymbol{\theta}) = (\partial_1^2 - \partial_2^2)\frac{1}{2\pi}\int \mathrm{d}^2\theta' e_1(\boldsymbol{\theta}')\ln|\boldsymbol{\theta} - \boldsymbol{\theta}'| \rightarrow \\ e_{(1)m}(\partial_1^2 - \partial_2^2)\frac{1}{2\pi}\int_m \mathrm{d}^2\theta' \ln|\boldsymbol{\theta}_n - \boldsymbol{\theta}'| = e_{(1)m}a_{mn}\end{aligned} \tag{4.29}$$

and

$$\begin{aligned}\partial_1\partial_2\partial^{-2}e_2(\boldsymbol{\theta}) = \partial_1\partial_2\frac{1}{2\pi}\int \mathrm{d}^2\theta' e_2(\boldsymbol{\theta}')\ln|\boldsymbol{\theta} - \boldsymbol{\theta}'| \rightarrow \\ e_{(2)m}\partial_1\partial_2\frac{1}{2\pi}\int_m \mathrm{d}^2\theta' \ln|\boldsymbol{\theta}_n - \boldsymbol{\theta}'| = \frac{1}{2}e_{(2)m}b_{mn}.\end{aligned} \tag{4.30}$$

where again, summation over index m is implied in both. Inserting these results into the pixellized version of equation (4.27) and then rearranging gives the final result:

$$\kappa_n = 1 - (1-e_n^2)^{1/2}\exp\left[-\frac{1}{2}(s_{(1)m}a_{mn} + s_{(2)m}b_{mn})\right] \tag{4.31}$$

where,

$$s_{(i)m} = \frac{e_{(i)m}}{e_m}\ln\left(\frac{1+e_m}{1-e_m}\right), \quad i = 1, 2. \tag{4.32}$$

Equation (4.31) is the second main result of this chapter. κ is now expressed entirely in terms of the measured ellipticity parameters and a_{mn} and b_{mn} of Section 4.2.1. As with other shear reconstruction methods however, the surface mass density obtained in this way is not absolute. Section 2.2.2 discussed how this is a result of the fact that the ellipticity parameters, or more specifically, the quantity $\gamma/(1-\kappa)$ is invariant to isotropic magnification resulting from homogeneous sheet-like distributions of mass. It is readily seen that if reconstruction of such a sheet mass distribution is attempted using equation (4.31), then since $e = 0$ everywhere, $\kappa_n = 0$ for any pixel n.

In practice, normalization comes from stipulating that $\kappa \geq 0$ everywhere. Observations in which clusters are larger than the field of view therefore only offer a lower bound to the absolute quantity of mass present. Of course, the results of magnification analysis can be employed to provide an absolute normalization of mass.

4.3.2 Application to Dumb-bell Model

This section sees the application of the pixellated ellipticity reconstruction technique to a dumb-bell cluster model similar to that presented in Section 4.2.2. In the exact same fashion as in Section 4.2.2, γ is calculated using Fourier methods from the generated κ distribution. The distribution of ellipticity parameters is then calculated from equation (4.14) which is subsequently used in equation (4.31) to reconstruct κ.

Figure 4.7 shows the generated mass distribution and the distortion field. Each ellipse in this distortion plot represents the observed shape that an intrinsically circular object would have if its image were to lie at that point in the lens plane. Combination of equation (4.14) with the result from Section 2.1.5 that $\gamma_1 = \gamma \cos 2\phi$ and $\gamma_2 = \gamma \sin 2\phi$ shows that the orientation of each ellipse is given by

$$\phi = \frac{1}{2} \arctan\left(\frac{e_2}{e_1}\right). \tag{4.33}$$

The ratio of semi-major to semi-minor axis or the elongation of each ellipse is given by equation (2.19) which in terms of e is

$$\epsilon = \frac{a}{b} = \frac{1+e}{1-e}. \tag{4.34}$$

Results of the reconstruction of κ are shown in Figure 4.8. Normalization of κ

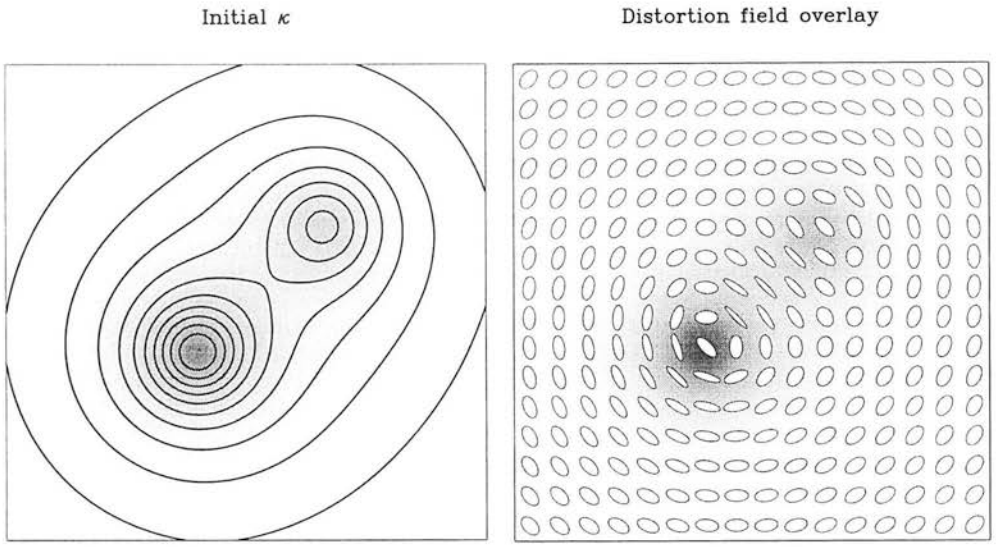

Figure 4.7: Generated dumb-bell model: The κ distribution is constructed from two truncated isothermal spheres of different mass. The distortion field is illustrated by plotting the apparent shape of an intrinsically circular background object.

was provided by setting the average of the two peak values to the average of both peaks in the generated distribution. In this plot, the contours are set at the same levels as those in Figure 4.7. The residuals show the reconstruction errors are again dominated by boundary effects. These spread further into the distribution than the boundary effects encountered with the pixellated magnification solution. This results in the reconstruction being limited to an accuracy of approximately 10% across most of the field of view.

4.4 Summary

A method for directly calculating accurate, self-consistent surface mass density and shear distributions from the lens amplification and critical line positions has been presented. This has been demonstrated with the isothermal sphere and dumb-bell cluster models in reconstructing the surface density to within a percent over most of the field of view. Reconstruction of the shear pattern has a lower fractional accuracy of a few tenths due to boundary effects. The method has been applied to the magnification data of Chapter 3 to enable computation of the 2D surface mass density and shear distribution of Abell 1689.

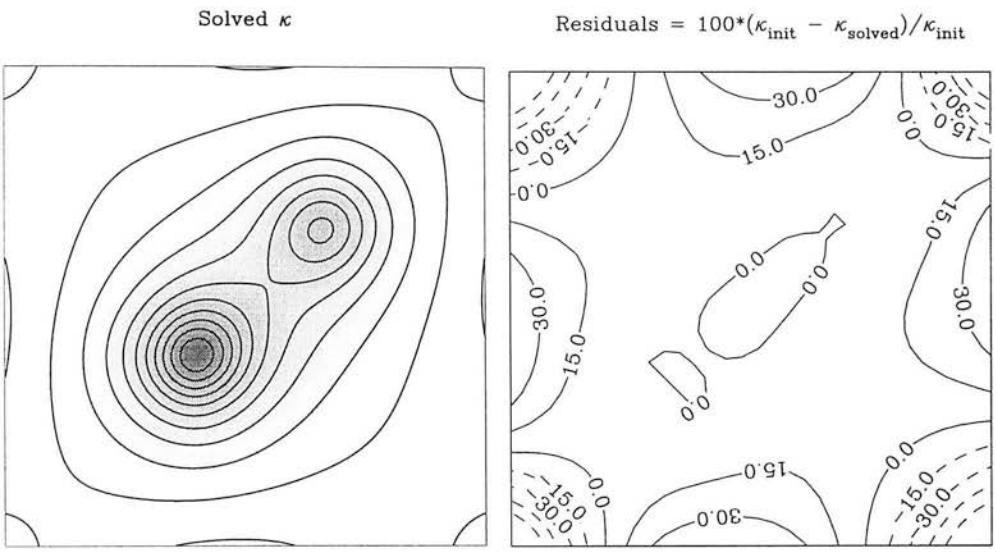

Figure 4.8: Reconstructed κ from the ellipticity parameters. Contours are at the same levels as in Figure 4.7. The residuals show that reconstruction is limited to an accuracy of approximately 10% across most of the field of view.

A simplified solution to the problem of estimating surface mass density from galaxy ellipticities has also been presented. This approach puts the calculation of surface mass from shear and magnification on an equal footing. Investigation of the combined analysis is left for future work.

Chapter 5

The Mass of Abell 1689 from Luminosity Function Analysis

The technique of reconstructing cluster mass by measuring its lensing effect on the intrinsic distribution of background object luminosities is applied to Abell 1689 in this chapter. Using an independent set of observations of the cluster, photometric redshifts and magnitudes determined for all objects detected in the field of view enable calculation of the source luminosity function. By comparing this to observations of an unlensed offset field, the method outlined in Section 2.3.3 gives a mass measurement of A1689.

The work in this chapter was carried out in collaboration with the Max-Planck Institute for Astronomy in Heidelberg (referred to hereafter as 'the MPIA'). Throughout this chapter, all reduction and photometric analysis was performed using the 'MPIAPHOT' software package (Meisenheimer & Röser 1996). MPIAPHOT was written at the MPIA as an extension to ESO's image reduction and analysis software, MIDAS (Munich Data and Analysis Software).

5.1 Observations of A1689

5.1.1 Data Acquisition

Observations of Abell 1689 were made in 9 different filters over a combined period of 19 nights at the Calar Alto 3.5m telescope in Spain. These 19 nights were divided between 4 separate observing runs as listed below:

Dates	Observer(s)
28/05/95 - 04/06/95	K. Meisenheimer[1], E.M. Thommes[1,2] & A.N. Taylor[2]
22/02/96 - 24/02/96	K. Meisenheimer[1] & J.A. Peacock[2]
14/06/96 - 16/06/96	A.N. Taylor[2] & S. Dye[2]
01/06/97 - 05/06/97	S. Dye[2]

[1] MPIA, Heidelberg. [2] IfA, Edinburgh.

A total of 12 hours worth of data with sufficiently good seeing was observed during this time. Table 5.1 lists the total integration time observed with each filter. The I band filter (826 nm) data is the data observed with ESO's 3.5m NTT presented in Chapter 3. Figure 5.1 shows the transmission of all filters used.

Filter: $\lambda_c/\Delta\lambda$ (nm)	t_{int}(s)	No. Filter Subsets	Use
826/137 (I-band)	6000	1	Global SED
774/13	6800	2	H_α at $z = 0.18$
703/34	4100	2	Background z
614/28	7700	2	Background z
572/21	6300	2	Background z
530/35	3300	1	Background z
480/10	4200	1	4000Åat $z = 0.18$
466/8	4800	1	Ca H,K at $z = 0.18$
457/96 (B-band)	6000	2	Global SED

Table 5.1: The observations of Abell 1689 in all 9 filters detailing the total integration time in each filter, t_{int}. The number of filter subsets gives the number of different nights observed with that filter.

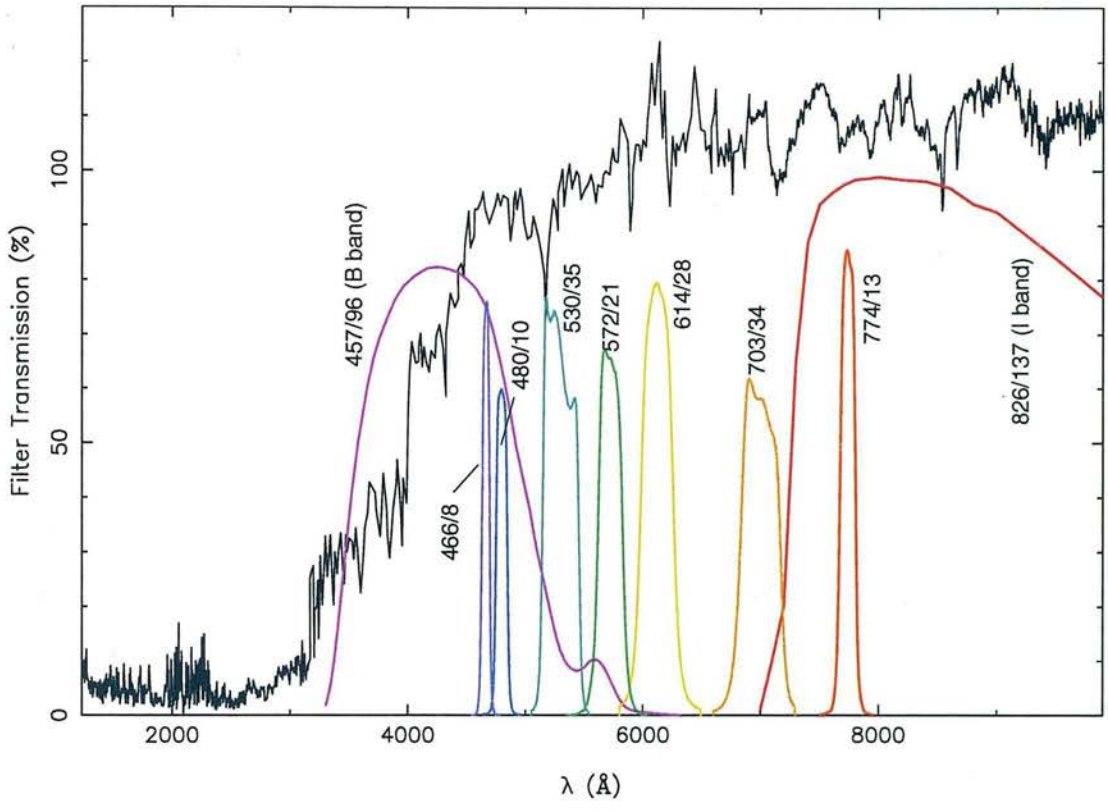

Figure 5.1: Transmission curves of filters used for observations of Abell 1689 at Calar Alto. The uncalibrated spectrum of the central dominant elliptical galaxy in A1689 is plotted for reference (see Section 5.2.3).

To clarify the terminology used throughout this chapter, a set of images of Abell 1689 observed on the same night in a particular filter will be referred to as a 'filter subset'. The term 'filter set' refers to the total collection of cluster images observed in a specific filter. As Table 5.1 shows, observations of A1689 were comprised of either one or two filter subsets.

This specific range of filters was chosen to optimize distinction between foreground, cluster and background objects. The specific use of each filter is listed in Table 5.1. The known redshift of A1689 of $z \simeq 0.18$ allowed the narrow band filters to identify cluster objects by searching for spectral features such as the 4000Åbreak, the H_α line and the Ca H&K lines. The filters 530/35, 572/21, 614/28 and 703/34 enabled division of the background galaxies into redshift bins between $z = 0.2$ and $z = 0.8$. Discrimination of background and foreground objects was provided by measurement of the global spectral energy distribution (SED) in the I and B broadband filters.

5.1.2 Image Reduction

Flat fields for each filter subset were constructed from typically four or five dusk sky flats observed specifically for that subset. Cosmic rays and stars were eliminated from the combined flat fields by median filtering the individual flats. Each image belonging to a particular filter subset was then bias-subtracted and flattened with its corresponding flat field according to

$$f_i(x,y) \rightarrow \frac{f_i(x,y) - b_i}{f_{flat}(x,y)/ < f_{flat} >} \tag{5.1}$$

where $f_i(x,y)$ is the flux of the raw A1689 image i in the pixel at (x,y) and $f_{flat}(x,y)$ is the flux in the same pixel in the flat field. The flat field is normalised by dividing it by the mean flux per flat field pixel $< f_{flat} >$. The bias, b_i, is calculated as the mean value of flux in pixels in the overscan region of the image (generated by the CCD hardware by reading out extra lines of pixels which are not physically present on the chip). Use of flat fields specific to each night ensured that optimal flat fielding was achieved. All images were flattened to an accuracy of $\pm 1\%$ variation in background flux across the field of view. After the initial stage of flat fielding, some images exhibited a large scale gradient in flux. These images were subsequently more accurately flattened using a fitted second order 2D polynomial provided by the MIDAS function 'FIT/FLAT'.

With anywhere from 4 to 9 images per filter subset, cosmic ray removal was performed using the pixel rejection algorithm built into the MPIAPHOT command 'COSMIC/MEDIAN'. Prior to application of this algorithm, all images within a filter subset required alignment with each other. Alignment was obtained using two MPIAPHOT functions. The first, 'FIND/OBJ', was used to detect objects in each image and create a table of positions and integrated fluxes. Having generated tables for every image in the filter subset, the second function 'FIND/MOVE' was used to match objects between each table and calculate the translation required to align the images. Translations were then recorded in the header of each image for the pixel rejection algorithm.

The pixel rejection algorithm works by first calculating a median image M from all images f_i within a filter subset such that for every pixel (x,y),

$$M(x,y) = \text{median}\{f_i(x,y) | i = 1,n\}. \tag{5.2}$$

The variance of pixel values within a circle centred on every pixel (x_0, y_0) in this median image is then calculated via,

$$\sigma^2(x_0, y_0) = \frac{1}{m-1} \sum_{(x,y)|x^2+y^2 \leq R^2} (M(x,y) - M(x_0, y_0))^2, \tag{5.3}$$

where m is the number of pixels contained within the circle of radius R. This variance is used as a threshold for the rejection of pixels in each filter subset image. Denoting cosmic-corrected images as c_i, every pixel (x, y) in each image is attributed a value according to

$$c_i(x,y) = \begin{cases} f_i(x,y) & \text{if} \quad f_i(x,y) - M(x,y) \leq n_{hi}\sigma(x,y) \\ M(x,y) & \text{if} \quad f_i(x,y) - M(x,y) > n_{hi}\sigma(x,y) \end{cases}. \tag{5.4}$$

If a pixel in a particular image is affected by a cosmic ray causing the value of that pixel to lie a user defined n_{hi} standard deviations away from its median value, then it is replaced by its median value. If such a pixel replacement occurs, surrounding pixels contained within the same radius R as used in the variance calculation above are checked and replaced with the value $cs_i(x, y)$ according to,

$$cs_i(x,y) = \begin{cases} c_i(x,y) & \text{if} \quad c_i(x,y) - M(x,y) \leq n_{lo}\sigma(x,y) \\ M(x,y) & \text{if} \quad c_i(x,y) - M(x,y) > n_{lo}\sigma(x,y) \end{cases}. \tag{5.5}$$

This allows for any influence the cosmic ray may have had on surrounding pixels. Setting the user defined quantity n_{lo} lower than n_{hi} allows the less affected surrounding pixels to be detected and replaced with the median value. For the cosmic ray replacement of the Calar Alto images, the values $n_{hi} = 4.0$ and $n_{lo} = 2.5$ were used.

Figure 5.2 shows an almost true colour image of the post-reduced observations of A1689. This image was constructed by superimposing data from the 614nm (red), 530nm (green) and 457nm (blue) filters. The intensity scale is logarithmic to highlight faint structures such as the faint arc, just visible to the lower left of the cluster centre.

5.2 Photometric Analysis

5.2.1 Flux Conversion

Construction of the luminosity function which follows later in this chapter required object fluxes in terms of the number of photons per unit time per unit area. The

Figure 5.2: Post-reduced colour image of Abell 1689 observed with the 3.5m CA telescope. A logarithmic intensity scale highlights faint structures such as the faint arc to the lower left of the cluster centre.

'fluxes' assigned to pixels in the raw Calar Alto images were simply the values output by the telescope's analogue to digital conversion hardware.

The electric charge stored in each CCD pixel as a result of incident photons during an image exposure is read at the end of an exposure as a voltage. The CCD hardware converts these analogue voltages into digital units (DUs) which are the values assigned to pixels in the raw image. Since DUs are directly proportional to the number of photons which have struck a given pixel, a simple conversion factor enables the number of incident photons at that pixel to be calculated. This conversion factor typically

changes between detectors. Of the four separate observing runs made at Calar Alto, three different detectors and hence three different conversion factors with a value of around $2-3$ were in operation.

Clearly, knowledge of the number of incident photons per DU for a given detector is essential for the correct calculation of object fluxes. To obtain fluxes in terms of the number of photons per unit time per unit area, once DUs are converted to the number of photons, the total integration time for each image is used to calculate the rate of photon arrival. The last step is to calculate the area within which the photons in each pixel are collected. This is a function of the size of the telescope mirror, the optics in the instrument used to focus the image onto the CCD and the area covered by one CCD pixel.

Using the MPIAPHOT function 'PREP/CCD', image fluxes were automatically converted from units of DUs to units of number of photons per unit time per unit area. This command was configured specifically for the FREDUK (prime focal reducer) instrument used to obtain all of the A1689 images observed at Calar Alto.

5.2.2 Object Tables

Instead of summing together images in each filter set, photometric evaluation was carried out on images individually. In this way, the mean weighted flux $I^{(b,m)}$ of an object m observed in a filter set b was calculated as,

$$I^{(b,m)} = \frac{\sum_i \frac{I_i^{(b,m)}}{\left(\sigma_i^{(b,m)}\right)^2}}{\sum_i \frac{1}{\left(\sigma_i^{(b,m)}\right)^2}} \tag{5.6}$$

with an error given by

$$\overline{\sigma}^{(b,m)} = \left(\sum_i \frac{1}{\left(\sigma_i^{(b,m)}\right)^2} \right)^{-1/2} . \tag{5.7}$$

The quantity $\sigma_i^{(b,m)}$ is the standard deviation of background pixel values surrounding object m in image i. Background pixels are segregated by applying a cut to the histogram of fluxes in pixels which lie within a box centred on the object. This cut removes the high flux pixels belonging to the object itself and any other neighbouring

objects within the box.

The method of obtaining object fluxes with MPIAPHOT involves integrating the flux of all pixels inside a fixed aperture centred on each object (see Section 5.2.3). The centre of each object is provided by a 'mark table' of object co-ordinates which accompanies each image. The integrated flux of each object is then systematically written into this accompanying table. Since fluxes of a given object between all images in a particular filter set are ultimately combined, the identity of objects between images must be maintained. This is ensured by simply listing the same objects in the same order in every mark table.

The production of a mark table for each image would not be viable using the MIDAS 'FIND/OBJ' command independently on images simply because a different number of objects would be detected in each (due to, for example, the random variation between objects with brightnesses close to the detection thresholds and image offsets) and objects would also be ordered differently. Ideally, a mark table produced by detecting objects in a deeper image would solve these problems if it were used for all images.

Fortunately, the I band data used in this Chapter is deeper than all filters observed at Calar Alto. The table of objects generated by SExtractor in Chapter 3 was therefore used as the underlying mark table for every image. The depth of the I band image meant that there were more objects in the mark table than the number detectable in the Calar Alto images. In cases where I band objects were too faint to be detected in the Calar Alto images, the photometric analysis of the Calar Alto data (which follows later in this chapter) returned fluxes integrated over blank image regions. These 'null objects' were easily removed from the final object catalogue by virtue of their large flux scatter and approximately zero fluxes.

Image Distortion

A complication with using the I band SExtracted data table was that images observed with the Calar Alto 3.5m telescope are radially distorted, unlike the I band images taken with the NTT. In order to obtain correct alignment with the Calar Alto images,

the I band mark table therefore had to be radially distorted in the same fashion as the Calar Alto image distortion.

Previous work by the MPIA has shown that this radial distortion can be fitted by,

$$\Delta r = ar^2 + br^4 \tag{5.8}$$

where Δr is the radial outward shift experienced by an object lying in the image at a radius r from the optical axis. The parameters a and b are obtained in the following way:

- The I band mark table is first scaled, rotated and translated to overlay a typical Calar Alto image. This is done using the 'FIND/MOVE' command which compares a table of positions of objects in the Calar Alto image provided by the 'FIND/OBJ' command with the mark table. At this stage, the alignment is only approximate since the radial distortion cannot be accounted for by 'FIND/MOVE'.

- Objects in the I band mark table are matched with objects in the Calar Alto table. The translations required to shift each I band object onto its matching Calar Alto object are recorded in the I band mark table.

- Values of Δr versus r calculated as

$$\begin{aligned} \Delta r &= \sqrt{(x' - x_0)^2 + (y' - y_0)^2} - r \\ r &= \sqrt{(x - x_0)^2 + (y - y_0)^2} \end{aligned} \tag{5.9}$$

  are plotted for each object where (x, y), (x', y') and (x_0, y_0) are the co-ordinates of the original position of the object in the I band mark table, the shifted position and the optical axis respectively. A plot something similar to that in the top half of Figure 5.3 is obtained.

- The co-ordinates of the optical axis and the global scaling of object positions are adjusted manually until the plot of Δr versus r exhibits a minimum at $r = 0$ and a minimal spread of points along the curve as shown in the bottom half of Figure 5.3. An inaccurate global scaling of object positions has the effect of producing a minimum in Δr versus r away from $r = 0$. This is also true of an inaccurately located optical axis which causes the increased spread of points.

- Finally, the parameters a and b are calculated by manually fitting equation (5.8) to the plot of Δr versus r.

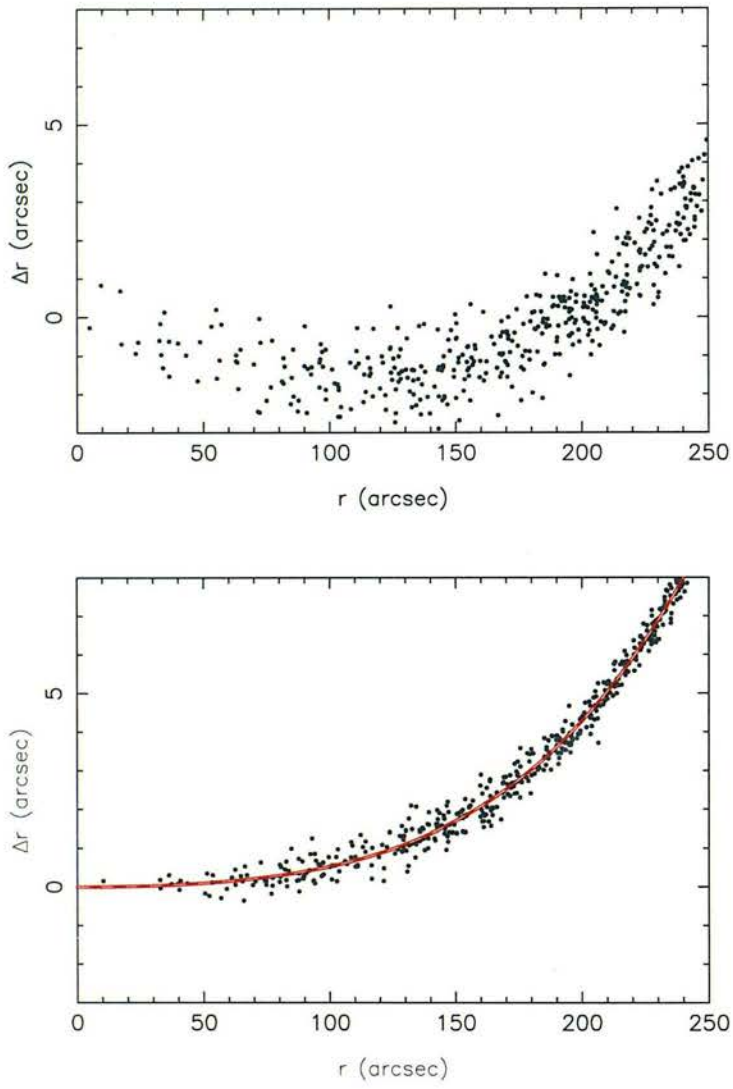

Figure 5.3: Distortion fitting of the Calar Alto 3.5m telescope images. Δr is calculated for every object as the radial distance of the object from the optical axis in the I band mark table subtracted from its radial distance in the image. If the I band mark table is over-scaled and not centred correctly on the optical axis, a minimum and a large scatter of points occurs in a plot of Δr vs r as shown in the top plot. The bottom plot shows how correct scaling and positioning of the mark table minimizes the scatter of points and removes the minimum for manual fitting of equation (5.8). The dashed line was plotted with the manually fitted parameters $a = 3.5 \times 10^{-5}$ and $b = 1.8 \times 10^{-9}$.

With values of $a \sim 10^{-5}$ and $b \sim 10^{-9}$ being typical for the Calar Alto 3.5m images, equation (5.8) shows that at the edge of the field of view corresponding to $r \simeq 200$

arcsec with the FREDUK instrument, a distortion of $\Delta r \simeq 5$ arcsec occurs. Correct distortion compensation was therefore essential to achieve a trustworthy photometric analysis.

After applying the distortion to the I band mark table, copies were made for other images regardless of filter as long as location of the cluster centre was roughly the same for each (with a tolerance of $\sim$ 30 arcsec). In some images, the cluster centre was located too far away from the majority to give good alignment with the distorted table. Separate distorted I band mark tables were produced for each of these. To ensure optimal alignment, each distorted table was finely scaled, rotated and translated onto its corresponding image, again using the 'FIND/MOVE' and 'FIND/OBJ' facilities.

5.2.3 Photometry

As Section 5.2.2 discussed, MPIAPHOT calculates object fluxes by summing the fluxes of pixels within a fixed aperture centred on each object. Using a fixed aperture means that the seeing in all images must be the same to determine the correct colours. Since image seeing can vary considerably even between two images observed on the same night, this requirement can only be realistically met by degrading the seeing of each image to the worst seeing measured among all images. Rather than inefficiently degrade the seeing of an entire image, MPIAPHOT degrades only inner aperture regions by Gaussian convolution. The FWHM of the convolving Gaussian is therefore dependent on the original seeing of the image and is calculated automatically from this.

The seeing of each image was determined with MPIAPHOT by averaging the FWHM of all stars in the image (selected by their low FWHM and high brightness). The worst seeing was found in the 466/8 filter image set with an average measured seeing of 2.10 arcsec. Flux integrations for every object in every frame across all filters were then performed by MPIAPHOT which degraded each image to a seeing of 2.10 arcsec. The local background sky count was automatically subtracted from each integrated object by the same histogram technique used for the flux error estimation detailed in Section 5.2.2. In the central parts of the cluster, this automatic back-

ground removal proved to overestimate background counts due to the high number density of objects there. For this reason, a circular region of 80 arcsec in diameter centred on the cluster was defined within which the average background count was manually determined for each image. MPIAPHOT was then instructed to subtract this manually determined constant background level from all flux integrations within this area instead of executing an automatic subtraction.

To allow for changing weather conditions, images within each filter set were normalised to an arbitrarily chosen image in that filter set. Normalisation was conducted by scaling the object fluxes of each image, apart from the reference image, so that the average flux of the same stars in each image was equal. This ensured correct calculation of the weighted flux and flux error from equation (5.6) and (5.7). The weighted flux and its error were later scaled to their calibrated photometric values.

In order to obtain photometric fluxes, the usual line of approach is to observe 'standard objects' with known, calibrated spectra. Integration of a standard's spectrum, multiplied with the transmission efficiency of the filter and instruments used in the observation, allows calculation of the expected photometric flux for that standard in that filter. Comparing this with the actual flux measured for the standard gives the conversion factor required to scale the flux of all objects to their photometric values. Under ideal circumstances, a calibrated spectrum of either a cluster member or another object in the field of view would enable direct photometric calibration of each image. Unfortunately, such a calibrated spectrum was unavailable although an uncalibrated spectrum of the large central elliptical galaxy in A1689 was obtained (Pickles & van der Kruit 1991). Calculation of the relative flux of objects between filters was therefore possible, but not calculation of the photometric flux. These fluxes were later calibrated using observations of a standard star located in another field (see Section 5.2.4).

The spectrum of Pickles & van der Kruit (1991) is shown in Figure 5.4. The range of wavelengths covered by the filters exceeded the wavelength range covered by this spectrum and so between the wavelengths 124 nm to 316 nm and 690 nm to 990 nm, a standard elliptical spectrum was used. This was redshifted to the measured redshift of $z = 0.183$ of the central elliptical galaxy (Teague, Carter & Gray 1990).

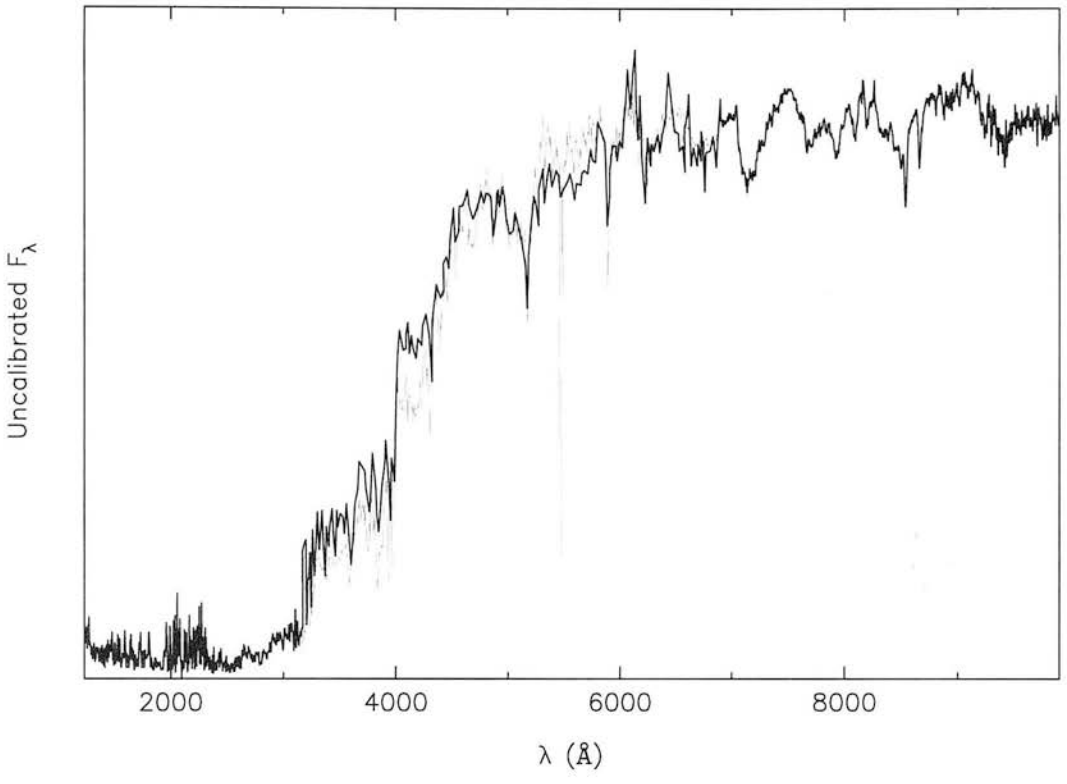

Figure 5.4: The uncalibrated spectrum of the large central elliptical galaxy in Abell 1689. Between 3160Å< λ <6900Å, the published spectrum of Pickles & van der Kruit was used. The regions λ <3160Å and λ >6900Å were taken from a standard elliptical galaxy spectrum (which is also shown between 3160Å< λ <6900Å as a grey line).

For comparison, this standard elliptical spectrum is also drawn in grey between 316 nm and 690 nm in Figure 5.4.

Denoting the spectral energy of the large central elliptical galaxy in A1689 at wavelength λ in units of energy per unit area per unit time per unit wavelength as $F_s(\lambda)$ allows its observed count rate $I_{b,s}$, expressed as the number of photons per unit time per unit area in a given filter b to be written (eg. Thommes, 1996)

$$I_{b,s} = k_b \int \mathrm{d}\lambda \frac{E(\lambda)T_b(\lambda)F_s(\lambda)\lambda}{hc}. \tag{5.10}$$

In this relation, the function $T_b(\lambda)$ describes the transmission efficiency of the filter (see Figure 5.1) while $E(\lambda)$ is the combined filter-independent efficiency of the detector and telescope optics. Figure 5.5 shows how the efficiency of the Calar Alto optics and the TEK7 CCD used on two of the observing runs depend on wavelength. The photometric scale factor, k_b is a quantity which relates the predicted value of $I_{b,s}$

calculated from the integral on the right hand side of equation (5.10) to the measured value of $I_{b,s}$. Without the transmission characteristics of the atmosphere being taken into consideration in equation (5.10), k_b absorbs the wavelength dependent nature of the weather at the time of observation across the wavelength range spanned by filter b. For this reason, under rapidly changing weather conditions, the value of k_b is strictly only applicable to the time the observation was made and the filter it was observed with. Since the functions $T_b(\lambda)$, $E(\lambda)$ and $F_s(\lambda)$ are known and $I_{b,s}$ is measured, k_b can be directly calculated. Without correct calibration of $F_s(\lambda)$, only the relative values of k_b for each filter are important.

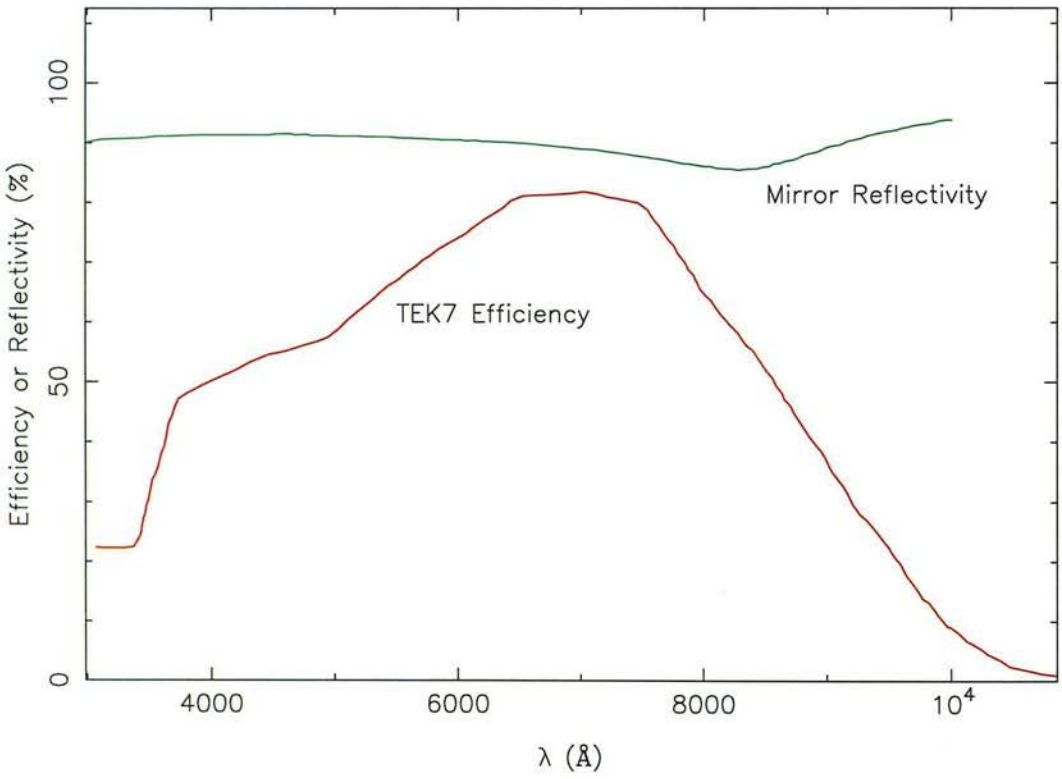

Figure 5.5: The reflectivity of the Calar Alto 3.5m telescope plotted with the efficiency of the TEK7 CCD chip used during two of the observing runs.

Suppose an object m with a known spectrum $F_m(\lambda)$ is observed in filter b to have a count rate $I_{b,m}$. If the spectrum is calibrated, the photometric intensity $F_{b,m}$ of the object in that filter can be calculated by evaluating,

$$F_{b,m} = \int \mathrm{d}\lambda T_b(\lambda) F_m(\lambda). \tag{5.11}$$

Since none of the objects observed have known spectra however, calculation of the

photometric intensity directly using equation (5.11) is not possible. The equivalent of equation (5.10) for object m shows that only the quantity

$$\int \mathrm{d}\lambda E(\lambda) T_b(\lambda) F_m(\lambda) \lambda = \frac{hc\, I_{b,m}}{k_b} \tag{5.12}$$

is known if k_b has been determined. If the filter is sufficiently narrow such that the spectrum of an object, the detector efficiency and the efficiency of the telescope optics do not vary considerably across the filter width, then the following approximation from equation (5.12) and (5.11) can be made,

$$F_{b,m} \simeq T_b(\lambda_{b0}) F_m(\lambda_{b0}) \Delta\lambda_b = \frac{hc\, I_{b,m}}{k_b \lambda_{b0} E(\lambda_{b0})} \tag{5.13}$$

where $\Delta\lambda_b$ and λ_{b0} is the width and central wavelength of the filter respectively.

Photometric Redshifts

Table 5.1 shows that not all of the observed filters were narrow band filters. The approximation in equation (5.13) was therefore not used. An alternative means of obtaining photometric intensities was employed by using software written and developed as part of the PhD work by C. Wolf (Wolf 1998) at the MPIA. The primary motivation for using this approach, outlined below, was that photometric redshifts were provided in addition to intensities.

The software functions by fitting model spectra to the set of calibrated photon count rates measured for each galaxy across all filters. Expressed more quantitatively, equation (5.11) was applied for each filter to a library of template spectra to arrive at a set of scaled filter counts for each spectrum. Galaxies were then allocated library spectra by finding the set of library colours which best fit the measured galaxy colours.

The spectral library was formed from the template galaxy spectra of Kinney et al. (1996). A regular grid of galaxy templates was generated, varying in redshift along one axis from $z = 0$ to $z = 1.6$ in steps of $\Delta z = 0.002$ and ranging over 100 spectral types from ellipticals, through spirals to starbursts along the other.

The set of photometric errors given by equation (5.7) for an individual galaxy across all filters gives rise to an error ellipsoid in colour space. Using the size and location of these error ellipsoids, probabilities of each library entry causing the observed

sets of colours for each galaxy were then calculated as

$$p(\boldsymbol{q}|z,s) = \frac{1}{\sqrt{(2\pi)^n |V|}} \exp\left(-\frac{1}{2}\sum_{j=1}^{n}\frac{[q_j - Q_j(z,s)]^2}{\sigma_j^2}\right) \tag{5.14}$$

where n is the number of colours, σ_j comes from propagation of the error given by equation (5.7) and $V \equiv \mathrm{diag}(\sigma_1^2, ..., \sigma_n^2)$. Each galaxy's position vector in colour space, $\boldsymbol{q} \equiv (q_1, ..., q_n)$ is compared with the colour vector $\boldsymbol{Q}$ of the library spectrum with a given redshift z and type s. Finding the maximum probability corresponding to the closest set of matching colours therefore immediately establishes redshift and galaxy type. An assessment of the uncertainty in this redshift is subsequently obtained directly from the distribution of the probabilities associated with neighbouring library spectra.

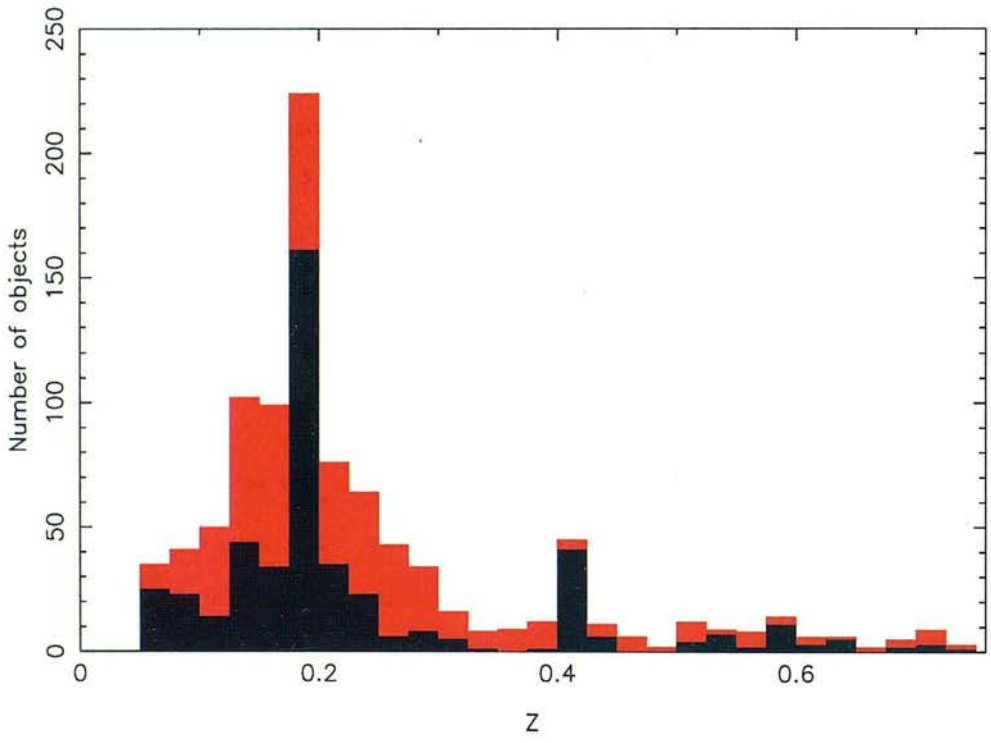

Figure 5.6: Redshift distribution of the 958 objects photometrically evaluated in the field of A1689. The red histogram plots all 958 redshifts whereas the black histogram plots only the 470 redshifts with a 1σ error in redshift of less than 0.05. The peak at $z \simeq 0.18$ is due to the cluster galaxies.

Determination of the photometric redshifts and magnitudes for the data presented in this chapter was carried out by C. Wolf at the MPIA. The analysis successfully returned the redshifts together with the apparent and rest-frame absolute magnitudes

(assuming an Einstein-de-Sitter universe) of 958 correctly classified objects in the observed field of view of Abell 1689. Figure 5.6 shows the distribution of these 958 objects. 470 of these objects have a 1σ error in redshift of less than 0.05. With the chosen range of filters, only redshifts within $0.05 \leq z \leq 0.80$ could be considered reliable and hence only this range is plotted[1]. The cluster galaxies are clearly discernable as the peak at $z \simeq 0.18$.

The feature at $z \simeq 0.4$ is most likely real and not an artifact of the photometric method. Such artifacts occur due to 'redshift focusing' when particular redshifts are measured more accurately than others. Where the uncertainty is larger, galaxies can be randomly scattered out of redshift bins, producing under-densities and corresponding over-densities where the redshift measurement is more accurate. This effect depends on the details of the filter set, being more common when fewer filters are used, but can be modelled by Monte Carlo methods.

The top half of Figure 5.7 shows the results of one realisation of such a Monte Carlo test for redshift focusing. The plot indicates how accurately the method reproduces redshifts of spectra scaled to I = 20 with photometric noise levels taken from the A1689 filter set. Each point represents a single library spectrum. Reproduced spectral redshifts, $z_{\rm phot}$, were determined by calculating colours through application of equation (5.11) to the library spectra with redshifts $z_{\rm lib}$. These colours were then randomly scattered by an amount determined from the filter-specific photometric error measured in the A1689 data before application of the redshift estimation method outlined above. The bottom half of Figure 5.7 shows the same plot generated using spectra scaled to I = 21 with the same photometric error taken from the A1689 data.

The accuracy of reproduced redshifts at I = 20 is clearly better than those at I = 21 where photometric noise is more dominant. There is a lack of any sign of redshift focusing in the vicinity of $z \simeq 0.4$ leading to the statement made previously that the feature seen in the A1689 data is probably real. The I = 21 plot which corresponds approximately to the sample magnitude cut of B = 23.7 (see Section 5.4.1) shows that input library redshifts of $z < 0.05$ begin to be pushed beyond $z = 0.05$. This only

[1] This upper redshift limit is decided by the stipulation that the 4000Åbreak must lie in or blue-ward of the second reddest filter in the set.

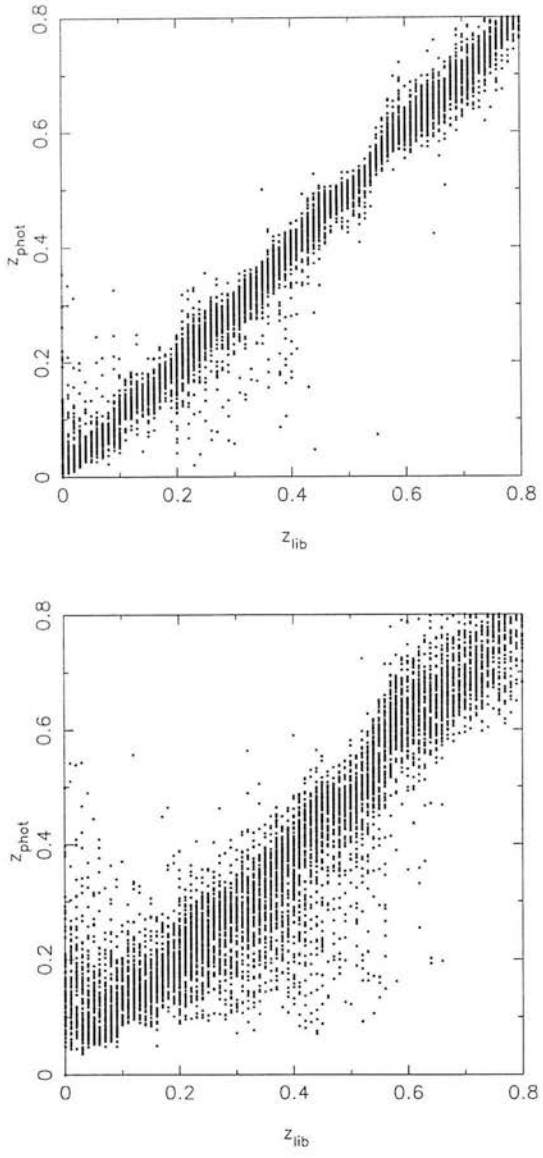

Figure 5.7: A single Monte Carlo realisation showing the accuracy of the photometric redshift evaluation method. Input library spectra with redshifts $z_{\rm lib}$ are scaled to I = 20 (top) and I = 21 (bottom) and subsequently used to calculate sets of colours using the A1689 filterset. These colours are randomly scattered by the filter-specific photometric errors measured in the A1689 data before calculating the reproduced redshifts $z_{\rm phot}$.

marginally affects the overall redshift distribution and yet partly explains the lack of galaxies at $z < 0.05$ in the A1689 redshifts of Figure 5.6. It is worth emphasising here that the significance of the peak at $z \simeq 0.18$ attributed to the cluster galaxies is far in excess of any effects of redshift focusing.

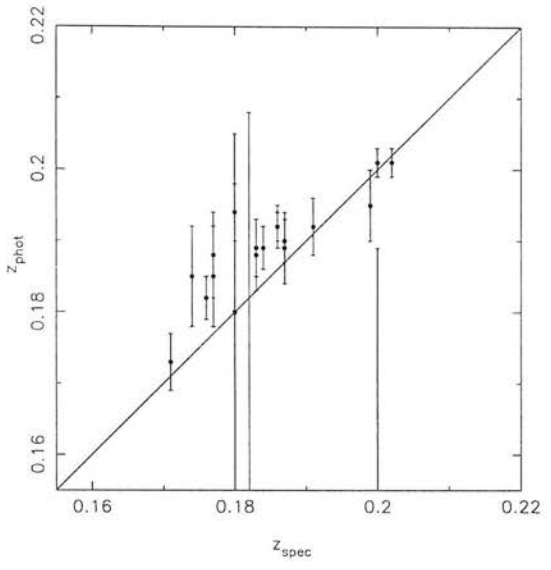

Figure 5.8: Comparison of the photometric redshifts estimated in the cluster Abell 1689 with spectroscopically determined redshifts. The distribution shows slight non-Gaussianity in the error distribution. The mean redshift of the cluster determined spectroscopically is $z = 0.185$ (Teague, Carter & Gray 1990), while the mean photometric redshift is $z = 0.189 \pm 0.005$.

Figure 5.8 shows a comparison of the photometrically determined redshifts around the peak of the redshift distribution of Abell 1689, with spectroscopically determined redshifts (Teague, Carter & Gray, 1990). The filter set was selected primarily to distinguish the cluster members, hence at higher redshift Monte Carlo estimates of the redshift uncertainty must be relied upon (see Section 5.4.2).

Abell 1689 lies in a region of sky where there is a very low level of galactic dust. The redshifts are therefore not affected by this source of contamination. However, dust in the cluster itself is another concern. The effects of reddening by cluster dust were modeled by the MPIA who found that although magnitudes are slightly affected, the redshifts experience only an insignificant difference.

5.2.4 Intensity Calibration

The magnitudes returned from the photometric redshift analysis were not calibrated since as mentioned in the previous section, an uncalibrated spectrum of the central galaxy in A1689 was used to obtain only the correct *relative* flux between filters.

Calibration of these magnitudes was made possible using the Oke (1990) standard star 'G60-54' which was observed during each night at the Calar Alto telescope.

In order to obtain calibrated magnitudes, absolute photometric scale factors were required for each filter. These absolute values were provided in exactly the same manner as the relative values obtained from the uncalibrated spectrum of the large elliptical galaxy in the centre of A1689. Using the MPIAPHOT software, the integrated flux $I_{b,s}$ of the standard in each filter was evaluated from the observations. MPIAPHOT was then used to calculate the expected photometric flux using the integral on the right hand side of equation (5.10) with the calibrated spectrum of the standard shown in Figure 5.9. This enabled determination of absolute photometric scale factors for the filters the standard was observed with.

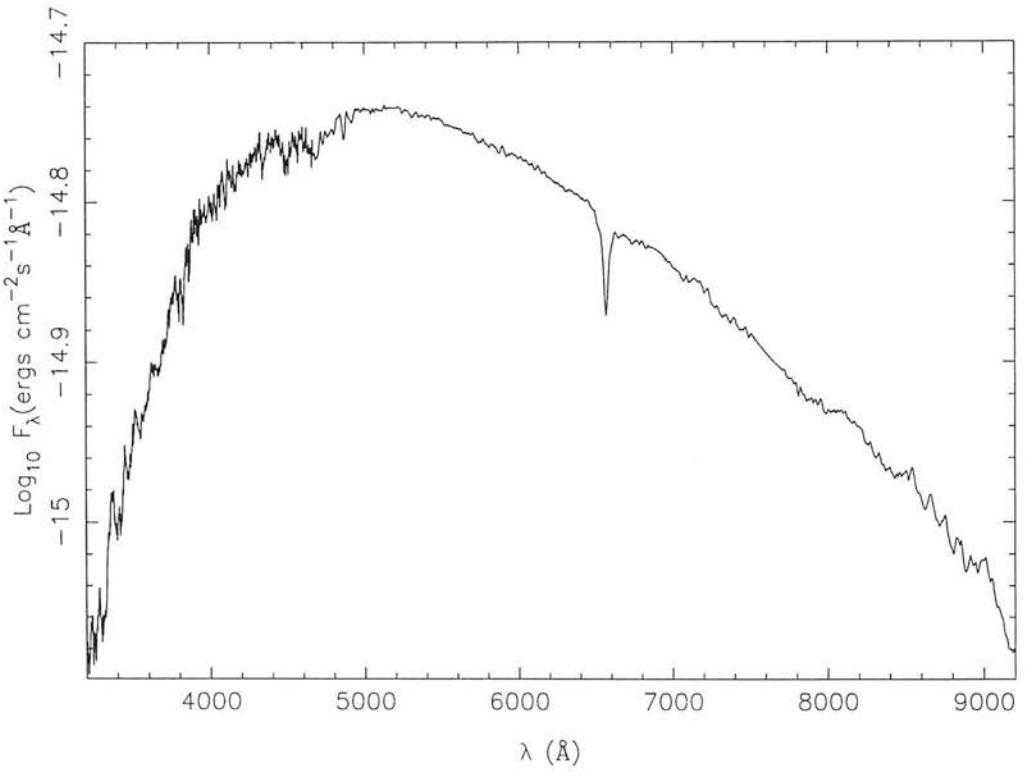

Figure 5.9: The calibrated spectrum of the standard star G60-54 (Oke 1990) used for magnitude calibration.

The uncalibrated photometric intensities $F_{b,m}$ were calculated from equation (5.11) which used a fitted spectrum $F_m(\lambda)$ scaled according to the value of the uncalibrated photometric scale factors. Calibrated intensities were therefore simply obtained by multiplying the uncalibrated intensities by the ratio of the uncalibrated to the absolute

photometric scale factor. The correctly calibrated magnitudes $M_{b,m}^{cal}$ were subsequently obtained using,

$$M_{b,m}^{cal} = M_{b,m} + 2.5\lg(k_b'/k_b) = M_{b,m} + \Delta M_b \tag{5.15}$$

where k_b' is the absolute photometric scale factor. Equation (5.15) defines the calibration magnitude offset ΔM_b required in each filter.

The standard star was observed in 6 filters. Values of k_b'/k_b and the required magnitude calibration offset ΔM defined in equation (5.15) for these 6 filters is shown in Table 5.2. The 1σ error $\sigma_{\Delta M}$ in this table was calculated using

$$\sigma_{\Delta M} = \frac{2.5}{k\ln 10}\sqrt{\sigma_k^2 + \left(\frac{k}{k'}\right)^2 \sigma_{k'}^2} \tag{5.16}$$

derived from propagation of the errors σ_k and $\sigma_{k'}$ on k and k'. The errors σ_k and $\sigma_{k'}$ are in turn the result of propagation of the uncertainty in flux as prescribed by equation (5.7) of the standard star and the large elliptical galaxy in the centre of A1689. As discussed in Section 5.2.3, the photometric scale factor depends on the weather conditions at the time of the observation and so to minimise the inaccuracy of the calculated magnitude offsets, the uncalibrated photometric scale factors were taken from A1689 images observed as near in time as possible to the standard star observations. This difference in time is also listed in Table 5.2.

The fact that the six filters in Table 5.2 required approximately the same calibration offsets in magnitude demonstrates that both the original uncalibrated and the calibrated photometry was carried out correctly and that the spectra used in each were also accurate. The small deviations from the average magnitude offset arise due to changing weather conditions and slight differences in the position of the object mark tables between filters. The two magnitude offsets for the 466/8 filter were calculated using the same A1689 image and two different standard star images observed at different times. Although a time of over 1.5 hours elapsed between these observations, the magnitude offsets are in close agreement with each other suggesting steady weather conditions. Comparing this with the 572/21 filter however shows that the converse is true; a standard star image observed a mere 4 minutes later results in a pair of offsets differing more than in the case of the 466/8 filter. This could be due to more unsettled weather conditions during the 572/21 observations although it is difficult to distinguish this from a slightly misaligned object mark table.

Filter: $\lambda_c/\Delta\lambda$ (nm)	k_b'/k_b	ΔM	$\sigma_{\Delta M}$	Δt_{obs} (min)
774/13	7.66	2.21	0.05	3
703/34	7.05	2.12	0.02	6
614/28	6.98	2.11	0.02	2
572/21	6.67	2.06	0.03	3
	7.52	2.19	0.04	7
480/10	6.31	2.00	0.04	3
466/8	7.24	2.15	0.04	4
	7.31	2.16	0.06	93

Table 5.2: Photometric scale factor ratios and magnitude offsets required for calibration of images observed in 6 of the 9 filters. The magnitude offset ΔM is defined in equation (5.15). Δt_{obs} gives the elapsed time between observations of the standard star and the A1689 image used to determine k_b. The 1σ error $\sigma_{\Delta M}$ is calculated using equation (5.16).

Despite these small discrepancies, the calibration offset was deemed sufficiently precise to calculate an average value and apply it to the apparent and absolute magnitudes returned from the photometric analysis. An average magnitude offset of

$$\Delta M = 2.11 \pm 0.01 \tag{5.17}$$

calculated by weighting the 8 offsets in Table 5.2 with their associated errors was applied to the magnitudes of all objects throughout all of the 9 filters. The error here was calculated as a standard error on the mean.

Using the fact from equation (2.32) that a lens induced magnification μ translates magnitudes according to $M \rightarrow M + 2.5 \lg \mu$ shows through error propagation that the error on the magnitude offset corresponds to an error in magnification of $\pm 0.01\mu$. Later sections in this chapter show that this is a negligible source of error.

5.3 Luminosity Function Analysis

Section 2.3.3 discussed how application of maximum likelihood theory to an observed distribution of object magnitudes enables the determination of lens induced magnification. As equation (2.59) shows, this can only be accomplished once the distribution of the unlensed magnitudes of those objects is known. A similar predicament to that of the number count analysis in Section 3.2.2 is again encountered. In the same way that an offset field was used to gain knowledge of the unlensed number density of background objects in the number count analysis, an offset field was used in this analysis to determine the unlensed magnitude distribution of objects behind A1689.

The offset field was observed to a depth of $B \simeq 24.5$ as part of the Calar Alto Deep Imaging Survey (CADIS) conducted by the MPIA (Thommes et al 1999). Data for this survey was observed in 16 filters from the B to the K band with the 2.2m telescope at Calar Alto. Using exactly the same methods outlined above, photometric redshifts and apparent and rest-frame absolute magnitudes were determined for all objects in all filters. To ensure a fair comparison with the A1689 data however, only the CADIS B band observations were considered in this analysis.

Although evolutionary traits of the luminosity function could have been investigated in principle (see for example, Lilly et al 1995, Ellis et al 1996), this was not considered in the work presented here. Instead, all galaxies between the redshift limits of interest were used to calculate a total luminosity function. This was primarily to ensure a sufficiently large number of objects in the A1689 sample which must have the same redshift selection as that of the offset field for a fair comparison in the likelihood analysis (Section 5.4.2).

5.3.1 The CADIS B Band Luminosity Function

An estimate of the luminosity function of galaxies in the CADIS B band was provided initially using the canonical $1/V_{max}$ method introduced by Schmidt (1968) for the study of quasar evolution. The quantity V_{max} is computed for each galaxy as the comoving volume within which the galaxy could lie and still remain in the redshift and magnitude limits of the survey. For an Einstein-de-Sitter universe, this volume is

calculated as,

$$\begin{aligned} V_{max} &= \left(\frac{c}{H_0}\right)\delta\omega \int_{max(z_l, z_{m_{min}})}^{min(z_u, z_{m_{max}})} \mathrm{d}z \frac{D^2(z)}{(1+z)^{3/2}} \\ &= 4\left(\frac{c}{H_0}\right)^3 \delta\omega \left[\frac{1}{1+z} - \frac{2}{\sqrt{1+z}}\right]_{max(z_l, z_{m_{min}})}^{min(z_u, z_{m_{max}})} \end{aligned} \tag{5.18}$$

where $\delta\omega$ is the solid angle of the observed field of view and $D(z)$ is the comoving radial distance defined in equation (1.38). The upper limit of this integral is set by the minimum of the upper limit of the redshift interval chosen, z_u, and the redshift at which the galaxy would have to lie to have an apparent magnitude of the faint limit of the survey, $z_{m_{max}}$. Similarly, the maximum of the lower limit of the chosen redshift interval, z_l, and the redshift at which the galaxy would have to lie to have an apparent magnitude of the bright limit of the survey, $z_{m_{min}}$, forms the lower limit of the integral. This lower integral limit plays a non-crucial role when integrating over large volumes originating close to the observer where the volume element makes only a relatively small contribution to V_{max}.

The redshifts $z_{m_{max}}$ and $z_{m_{min}}$ are calculated for each object by finding the roots of

$$M - m_{[max/min]} - 5\lg h_0 + 5\lg[D(z)(1+z)] - K(z) + 42.39 = 0 \tag{5.19}$$

where M is the absolute magnitude of the object, $m_{[max/min]}$ is the appropriate survey limit and $K(z)$ is the K-correction (see Section 1.1.10). Although the K-correction for each object at its actual redshift was known from its apparent magnitude and absolute rest-frame magnitude, the redshift dependence of this K-correction was not. In principle, this redshift dependence could have been calculated directly for each object using its best fit spectrum returned from the photometric analysis, however these spectra were unfortunately not available. As such, an approximated K-correction was used instead.

To demonstrate the validity of this approximation, the K-correction for the standard elliptical galaxy spectrum plotted in grey in Figure 5.4 was calculated. Using a hypothetical top hat B band filter centred on $\lambda = 650$nm and 170nm in width, the K-correction plotted in Figure 5.10 was computed through application of equation (1.59). This plot shows that $K(z)$ is approximately proportional to z and so a good

approximation to $K(z)$ is obtained using

$$K(z) = \frac{K(z_0)}{z_0} z \tag{5.20}$$

where $K(z_0)$ is the K-correction of the object at its actual redshift z_0.

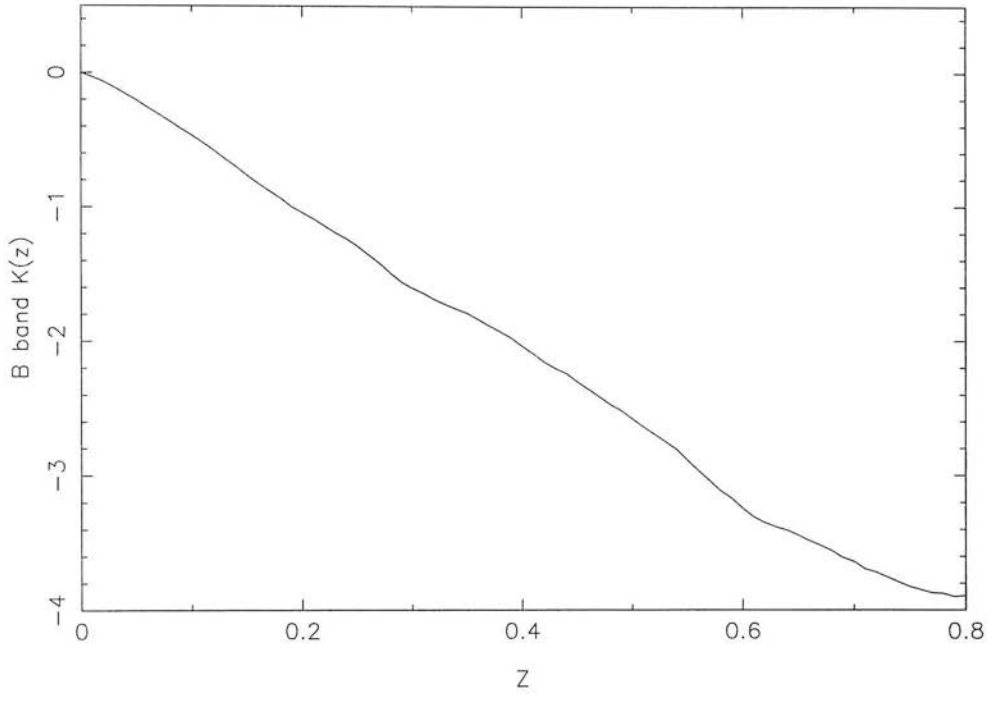

Figure 5.10: K-correction calculated using the standard elliptical galaxy spectrum plotted in Figure 5.4 with a top hat B band filter centred on $\lambda = 650$nm and 170nm in width.

Of course, the proportionality observed in equation (5.20) has only been proven to be applicable to standard elliptical galaxies observed in the B band. Coinciding with the redshift range $0 \leq z \leq 0.8$ forced upon the A1689 observations here, Lilly et al (1995) find that the K-correction in the B band for elliptical galaxies is also roughly proportional to redshift. Furthermore, they show that this is also the case for spiral and irregular galaxies. Based on this evidence, the approximation in equation (5.20) was therefore adopted for $K(z)$.

Once V_{max} has been calculated for all objects, the luminosity function ϕ at the rest-frame absolute magnitude M in bins of width $\mathrm{d}M$ is then computed from,

$$\phi(M)\mathrm{d}M = \sum_i \frac{1}{V_{max,i}} \tag{5.21}$$

where the sum acts over all objects having magnitudes between $M - \mathrm{d}M/2$ and $M + \mathrm{d}M/2$.

Figure 5.11 shows the luminosity function of B band magnitudes from the CADIS offset field which has a solid viewing angle of $\delta\omega = 100$ arcmin^2. To match the selection of objects lying behind A1689, only objects within the redshift range $0.3 \leq z \leq 0.8$ were chosen. A further restriction on the apparent B magnitude of $m_B \leq 24.5$ was applied for completeness of the sample (see Section 5.3.2), yielding a total of 371 objects. The data points in Figure 5.11 are centred on bins chosen to maintain an equal number of objects in each. The 1σ errors shown here were calculated using Monte Carlo simulations. 1000 realisations were performed by randomly scattering individual object redshifts in accordance with their associated errors provided in the CADIS dataset (the NAG routine G05DDF was used to generate normally distributed errors). For each realisation, the V_{max} of each object was recalculated using the re-sampled redshift. The resultant standard deviation of the distribution of values of ϕ for each bin given by equation (5.21) was then taken as the error. In this particular case, no consideration was given to the magnitude errors or the propagation of the redshift error into object magnitudes. Section 5.3.3 discusses this further. Finally, the solid curve in Figure 5.11 shows the Schechter function determined with the maximum likelihood technique discussed in Section 5.3.3.

5.3.2 Sample Incompleteness

Determination of the faint magnitude limit beyond which a sample becomes incomplete is essential for the correct calculation of any luminosity function. It is self evident from the previous section that an incorrect evaluation of the limiting magnitude m_{max} would result in larger values of V_{max} and hence a biassed luminosity function. As Section 5.5.4 details, this affects calculation of the lens mass.

An estimation of the incompleteness of the CADIS B band sample can be obtained using the V/V_{max} statistic (Schmidt 1968). In this ratio, V_{max} is calculated as before whereas V is the comoving volume described by the observer's field of view from the same lower redshift limit in the integral of equation (5.18) to the redshift of the object. If a sample of objects is unclustered, exhibits no evolution (ie. a systematic change in intrinsic luminosity with redshift) and is complete, the position of each object in its associated volume V_{max} will be completely random. If this is the case, then the

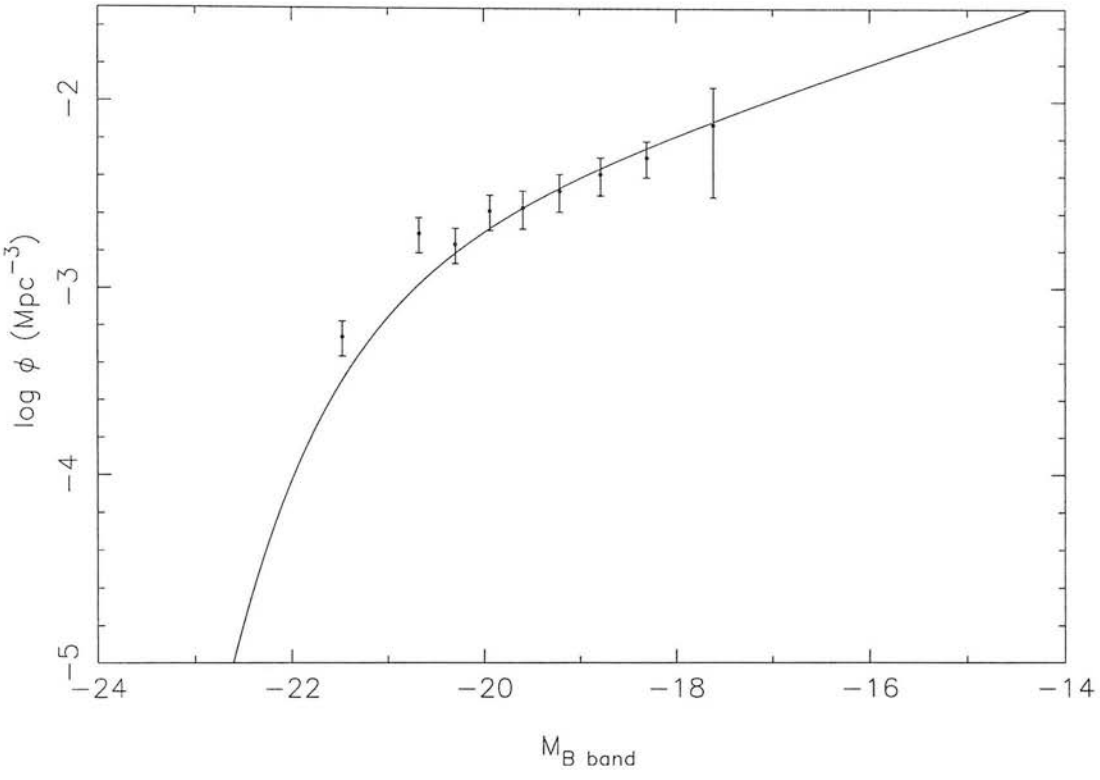

Figure 5.11: The CADIS B band object luminosity function calculated with the $1/V_{max}$ formalism. Errors account only for errors in redshift. Points lie at bin centres, the widths of each chosen to hold the same number of objects. There are 371 objects in total selected by the redshift limits $0.3 \leq z \leq 0.8$ and the apparent B magnitude $m_B \leq 24.5$. The solid line is the Schechter function determined in Section 5.3.3.

distribution of the V/V_{max} statistic over the range 0 to 1 will be uniform and the average value of V/V_{max} across all objects will be equal to 0.5.

If the sample is affected by evolution such that more intrinsically bright objects lie at the outer edges of the V_{max} volume, then V/V_{max} is biassed towards values larger than 0.5. The reverse is true if a larger number of brighter objects lie nearby. If the sample is incomplete at the limiting apparent magnitude chosen, estimations of V_{max} will be on average too large and will cause V/V_{max} to be biassed towards values less than 0.5 as shown by the histogram in the top left hand corner of Figure 5.12.

The requirement that the average V/V_{max} should have a value of 0.5 subject to the conditions outlined above is of course only applicable in the large number limit. In practice, shot noise combines with the effect of clustering to confuse the situation.

An estimation of the error in the average V/V_{max} value is therefore important for the purposes of estimating the limiting apparent magnitude of a sample. Consider the theoretical variance $\sigma^2_{V/V_{max}}$ of one V/V_{max} measurement,

$$\sigma^2_{V/V_{max}} = \int_0^\infty \mathrm{d}\left(\frac{V}{V_{max}}\right)\left(\frac{V}{V_{max}} - \left\langle\frac{V}{V_{max}}\right\rangle\right)^2 P\left(\frac{V}{V_{max}}\right) \tag{5.22}$$

where P is the normalised distribution of values of V/V_{max} and angular brackets denote the mean value. For an ideal sample with no evolution or clustering, P is a uniform distribution over the range 0 to 1 and $< V/V_{max} >= 0.5$ which gives

$$\sigma^2_{V/V_{max}} = \int_0^1 \mathrm{d}\left(\frac{V}{V_{max}}\right)\left(\frac{V}{V_{max}} - \frac{1}{2}\right)^2 = \frac{1}{12}. \tag{5.23}$$

The standard deviation, $\sigma_{<V/V_{max}>}$, of the average $< V/V_{max} >$, is then calculated from the combination of errors on each of the V/V_{max} measurements. In this hypothetical case, these are all equal and are given by equation (5.23) so that

$$\sigma^2_{<V/V_{max}>} = \sum_{i=1}^{N}\left(\frac{\partial < V/V_{max} >}{\partial (V/V_{max})_i}\right)^2 \sigma^2_{(V/V_{max})_i} = \frac{1}{12N}. \tag{5.24}$$

In order to arrive at an apparent magnitude limit for the CADIS field, values of $< V/V_{max} >$ were calculated for different applied limiting magnitudes and plotted as shown in Figure 5.12. The grey region in this plot corresponds to the 1σ errors described by equation (5.24) which lessen at the fainter limiting magnitudes due to the inclusion of more objects. Clustering in the CADIS field adds extra noise and so these errors are an underestimate of the true errors.

Without knowledge of the effects of clustering, a limiting magnitude of $m_B \leq 24.5$ gives a value of $< V/V_{max} > \simeq 0.5$. This is in agreement with the apparent magnitude limit at which the number counts begin to fall beneath that measured by deeper surveys. The histogram in Figure 5.13 shows how the number of objects varies with magnitude for the CADIS B band. The solid curve plotted over this shows the scaled B band counts measured by Lilly et al (1991) to a depth of $m_B \simeq 26$. This clearly begins to depart from the histogram at $m_B \simeq 24.5$ which is thus taken to be the limiting magnitude for the CADIS observations. Effects due to this choice are investigated later.

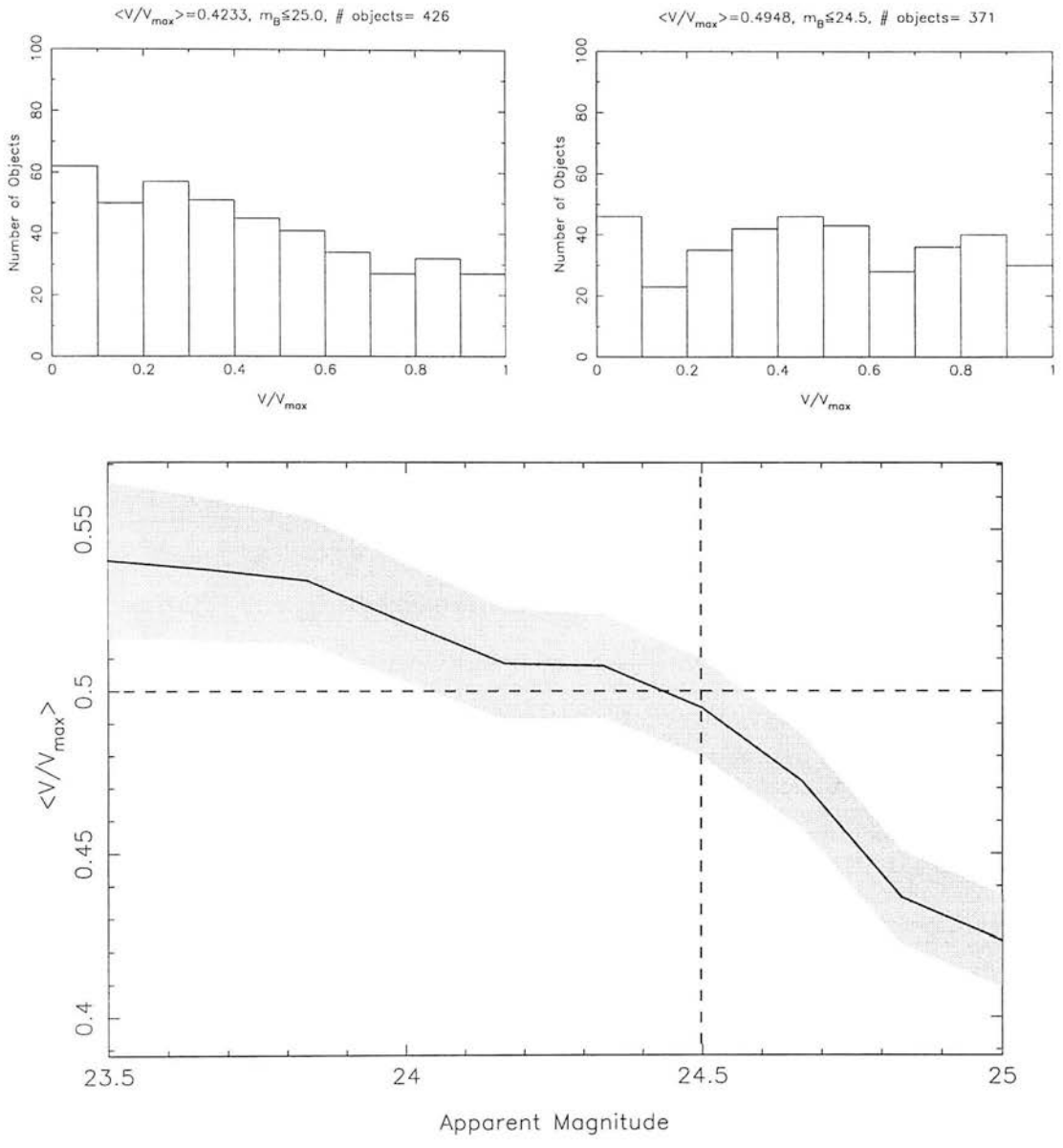

Figure 5.12: The main graph shows the variation of the mean V/V_{max} with limiting apparent B magnitude for the CADIS field. The grey region corresponds to the 1σ errors described by equation (5.24) which are an underestimate due to the unconsidered effects of galaxy clustering. The top left plot shows the distribution of V/V_{max} for too faint a limiting magnitude of $m_B \leq 25$ which results in a bias towards values less than 0.5. The top right plot is the distribution with the chosen $m_B \leq 24.5$

5.3.3 Parameterisation of the Luminosity Function

To perform the likelihood analysis outlined in Section 2.3.3, a parametric form of the luminosity function was required. This parametric form was provided by a Schechter

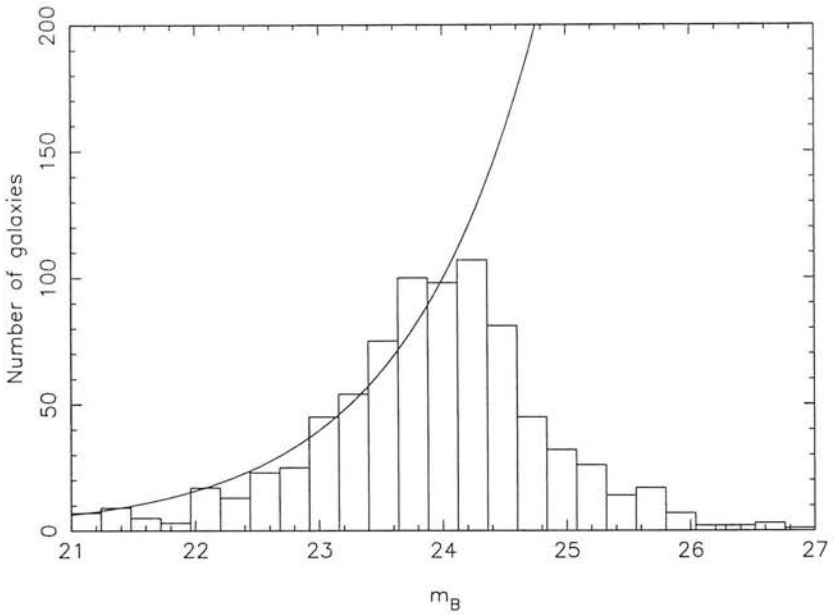

Figure 5.13: Variation of CADIS galaxy counts with apparent B band magnitude shown by the histogram. The curve shows the deeper B band counts of Lilly et al (1991) scaled for the CADIS field of view.

luminosity function (Schechter 1976):

$$\phi(M) = \phi^* 10^{0.4(M_*-M)(1+\alpha)} \exp\left[-10^{0.4(M_*-M)}\right]. \tag{5.25}$$

The parameters M_* and α were determined using the maximum likelihood method of Sandage, Tammann & Yahil (1979, referred to as STY hereafter). In much the same way as the probability in equation (2.59) was formed, the STY method forms the probability p_i that a galaxy i has an absolute magnitude M_i,

$$p_i \equiv p(M_i|z_i) \propto \frac{\phi(M_i)}{\int_{max(M_{min}(z_i),M_1)}^{min(M_{max}(z_i),M_2)} \phi(M)\mathrm{d}M} \tag{5.26}$$

where $M_{max}(z_i)$ and $M_{min}(z_i)$ are the absolute magnitude limits corresponding to the apparent magnitude limits of the survey at a redshift of z_i. Conversion of these apparent magnitude limits includes the K-correction by use of equation (5.19) with z set to z_i. A further restriction is placed upon the integration range by imposing another set of magnitude limits $M_1 < M < M_2$ which for the CADIS data were set at the maximum and minimum absolute magnitudes found in the sample.

The likelihood function in this case is a two dimensional function of the Schechter parameters M_* and α formed from the product of all probabilities p_i. The best fit M_*

and α are therefore found by maximizing the likelihood function,

$$\ln \mathcal{L}(M_*, \alpha) = \sum_{i=1}^{N} \left\{ \ln \phi(M_i) - \ln \int_{max(M_{min}(z_i), M_1)}^{min(M_{max}(z_i), M_2)} \phi(M) \mathrm{d}M \right\} + c_p \qquad (5.27)$$

with the constant c_p arising from the proportionality in equation (5.26). An estimate of the errors on M_* and α are calculated by finding the contour in α, M_* space which encompasses values of α and M_* lying within a particular confidence level from values giving the maximum likelihood $\mathcal{L}_{max}$. In log-likelihood space as described by equation (5.27), this contour lies at $\ln \mathcal{L}_{max} - 0.5\Delta\chi^2$ where $\Delta\chi^2$ depends upon the desired confidence level.

The top left hand plot of Figure 5.14 shows these 1σ and 2σ confidence regions for the CADIS sample, which for two degrees of freedom correspond to a $\Delta\chi^2$ of 2.30 and 6.17 respectively. This plot highlights the degeneracy typically encountered between values of M_* and α (see for example, Lin et al, 1997). For any given Schechter-like distribution of magnitudes, any uncertainty in the magnitude at which the 'knee' of the distribution lies as characterised by M_*, is compensated by a change in the slope of the Schechter function at fainter magnitudes. If this knee is forced to brighter magnitudes, a steeper faint magnitude slope is required to maximize the likelihood function and hence a more negative value of α results.

The errors on M_* and α obtained from the likelihood contours only represent a level of uncertainty to the fit of the Schechter function to the absolute magnitudes. With this method, no consideration is given to the error in each object's magnitude or redshift. These uncertainties were taken into account using Monte Carlo simulations. In this case, both object redshifts and magnitudes were randomly scattered in 1000 realisations. Redshift and apparent magnitude errors were available from the CADIS dataset as derived by the photometric analysis software. Errors in absolute magnitude were then calculated from the propagation of uncertainties in redshift and apparent magnitude through equation (5.19) to give,

$$\sigma^2_{M_i} = \left(\frac{K_i}{z_i} - \frac{5}{\ln 10} \frac{1 - 0.5(1+z_i)^{-1/2}}{1 + z_i - (1+z_i)^{1/2}} \right)^2 \sigma^2_{z_i} + \sigma^2_{m_i} \qquad (5.28)$$

for each object with redshift z_i and apparent magnitude m_i. Here, the K-correction given by equation (5.20) has been used such that the quantity $K_i \equiv K(z_i)$ is calculated from equation (5.19) using m_i, M_i and z_i as they appear in the CADIS dataset. The

typical ratio of apparent magnitude to redshift error was found to be $\sigma_{m_i}/\sigma_{z_i} \sim 1\%$ due to the relatively unprecise nature of photometric redshift determination.

After scattering magnitudes and redshifts in this way, each of the 1000 re-sampled realisations of the original data set produced were then used to generate individual likelihood distributions. The scatter of maximal values of M_* and α for each of these distributions is shown in the plot at the top right corner of Figure 5.14. It is not surprising that a similar degeneracy to that seen in the original likelihood contours is also seen here. To include the effects of this scatter in the likelihood contours, the log likelihood distributions were then co-added to form one, effectively convolved distribution.

The lower section of Figure 5.14 shows the contours obtained as a result of this convolution. As expected, inclusion of the redshift and magnitude errors to the original likelihood errors has resulted in enlargement of the 1σ and 2σ confidence regions. To obtain the overall errors on M_* and α, projections of the contours are made onto their respective axes. As seen in Figure 5.14, the resulting maximum likelihood M_* and α together with their errors were found to be

$$M_* = -20.94^{+0.47}_{-0.64}\,, \quad \alpha = -1.45^{+0.25}_{-0.23}\,. \tag{5.29}$$

The Schechter function described by these parameters is shown in Figure 5.11 by the solid line. (As a note of interest, χ^2 fitting of the $1/V_{max}$ data points in Figure 5.11 predicts a slightly less steep faint-end slope than that determined with the STY method. This is in disagreement with the work of Willmer (1997) who showed that fitting Schechter functions to simulated data using the $1/V_{max}$ method generally gives a steeper faint-end slope.)

Unlike the $1/V_{max}$ method of estimating luminosity functions, the STY method is incapable of yielding absolute quantities since the ratio of terms in equation (5.26) is independent of the normalisation constant ϕ^*. Although not essential for the likelihood analysis here, ϕ^* was calculated by normalising to the $1/V_{max}$ luminosity function over the range of magnitudes $M_1 < M < M_2$,

$$\phi^* = \frac{\sum_{i=1}^{N} 1/V_{max,i}}{\int_{M_1}^{M_2} \phi(M)\mathrm{d}M}. \tag{5.30}$$

The summation in the numerator here runs over all galaxies in the sample. Using the

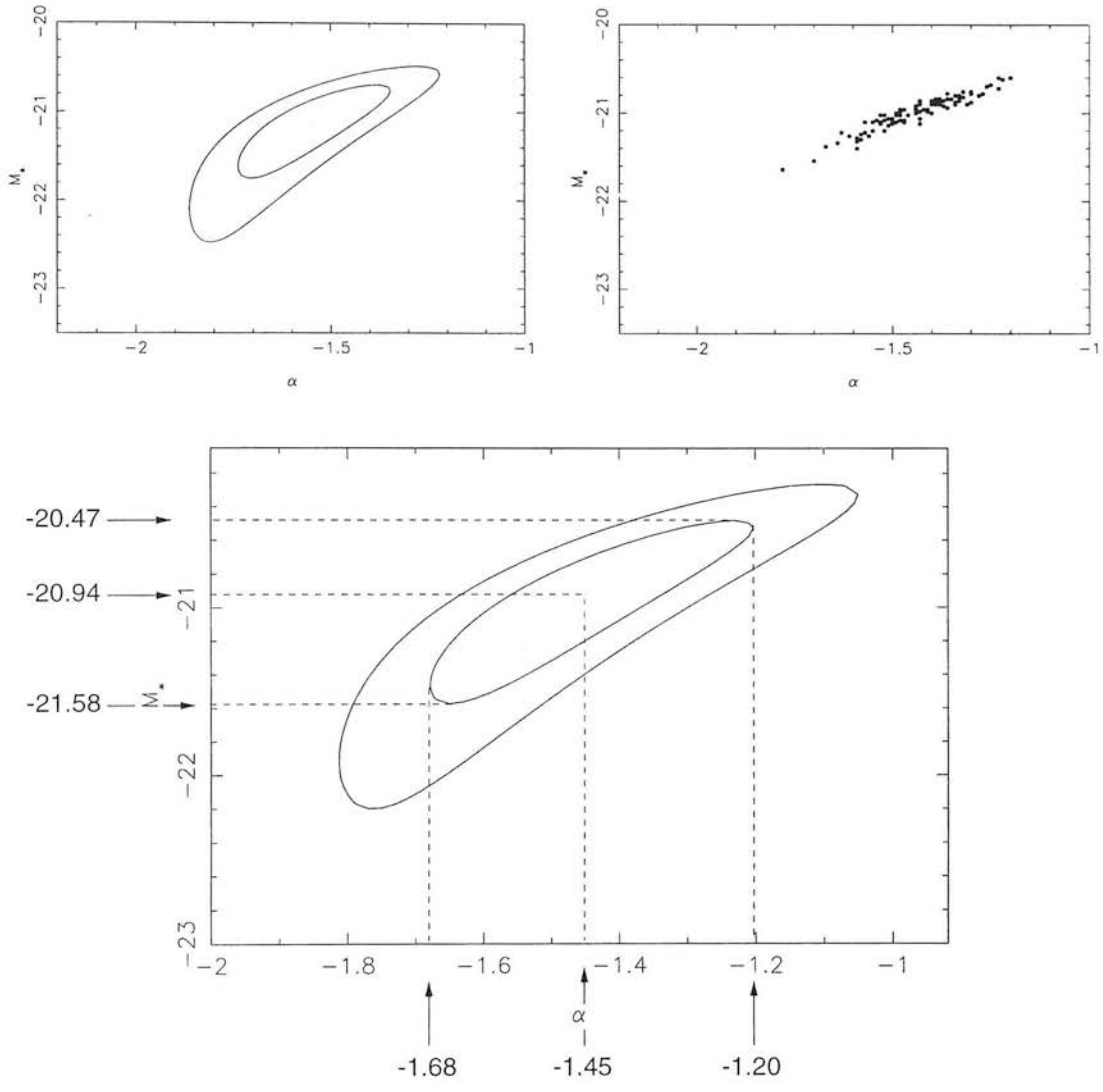

Figure 5.14: Determination of the Schechter parameters M_* and α for the CADIS data. The top left plot shows the maximum likelihood contours corresponding to 1σ and 2σ confidence levels without regard for redshifts or magnitude errors. The scatter in the maximum likelihood M_* and α resulting from these errors is shown in the top right plot. Convolving the top two plots gives the overall 1σ and 2σ confidence levels shown in the bottom plot.

values of M_* and α from (5.29) for the CADIS data gives a normalisation constant of

$$\phi^* = 2.05 \times 10^{-3}. \tag{5.31}$$

5.4 Lens Mass Determination

With knowledge of the distribution of absolute object magnitudes over the redshift range $0.3 < z < 0.8$, the statistical search for a sign of lens-induced magnitude brightening was made possible. In a similar manner to the case of the number count analysis in Chapter 3, the CADIS luminosity function was assumed and used as a fair measure of the intrinsic source magnitude distribution.

5.4.1 A1689 Sample Incompleteness

Obtaining a complete sample of background objects in the A1689 field of view is of the same importance as ensuring completeness of the CADIS sample. As Section 5.5.4 shows, the limiting faint magnitude chosen for the sample noticeably influences lens mass determinations.

Estimation of the incompleteness of the A1689 sample was carried out in exactly the same fashion as with the CADIS sample. The $< V/V_{max} >$ statistic for all galaxies lying within $0.3 < z < 0.8$ was plotted with limiting apparent B magnitude. The upper plot in Figure 5.15 shows this variation with the errors given by equation (5.24) plotted in grey. Once again, these errors are an underestimate of the true errors and so a limiting magnitude of $m_B \leq 23.7$ was chosen to give a total of 152 background objects with $< V/V_{max} >= 0.478$. This limit was also in agreement with that predicted directly from the number counts in a similar fashion to that of Section 5.3.2. The lower plot in Figure 5.15 shows the magnitude-redshift distribution of all 958 objects in the overall A1689 dataset. The selected 152 background objects used for the analysis which follows are highlighted by the dashed box. Cluster members are again clearly discernable as the peak at $z \simeq 0.18$

5.4.2 Determination of κ

In the number count analysis of Chapter 3, the lack of redshift data forced the approximation that all objects were assumed to lie at the same redshift. With a fixed lens and source redshift, the surface mass density κ, being effectively independent of

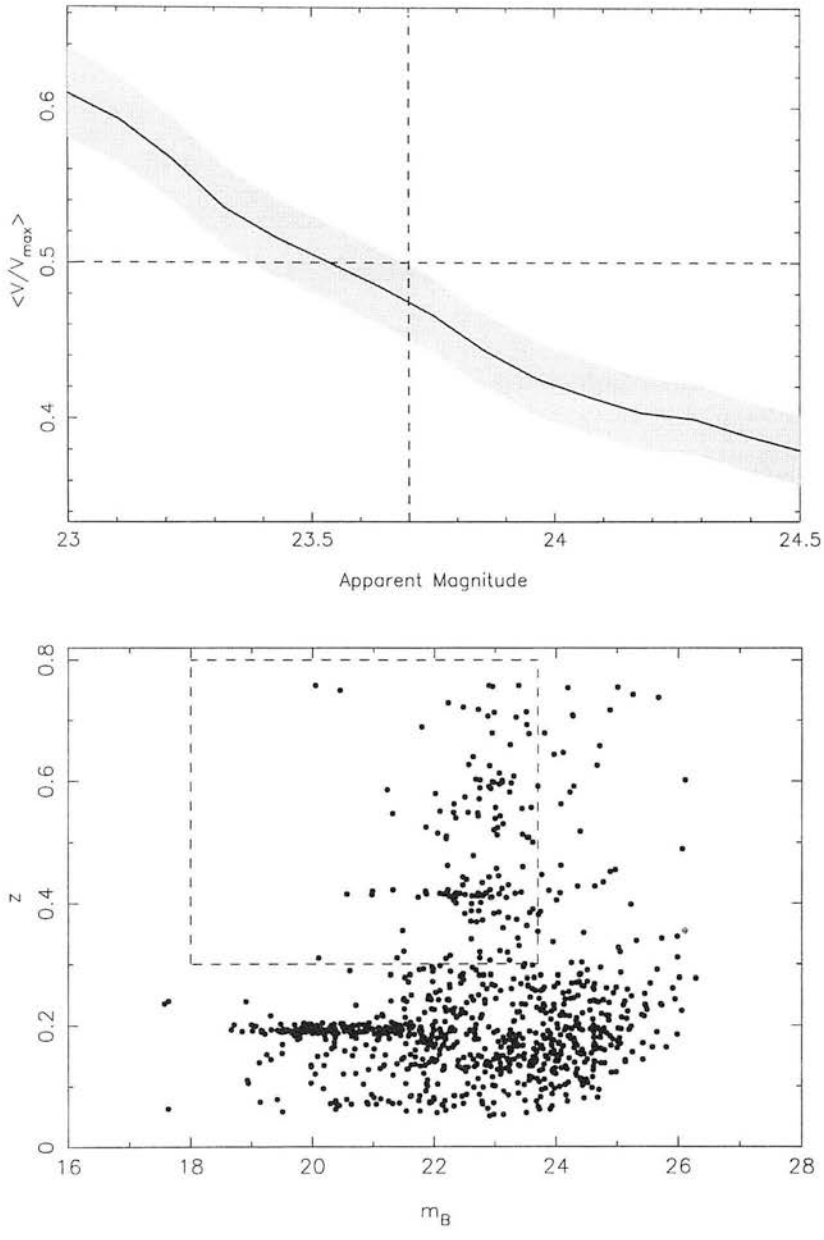

Figure 5.15: The top plot shows variation of the $< V/V_{max} >$ statistic with limiting apparent B band magnitude for the A1689 sample. 152 objects were selected by the redshift and apparent magnitude limits highlighted by the dashed box in the bottom plot.

redshift, could therefore be used as a likelihood parameter for the real lens surface mass density Σ in equation (2.9). The situation here is changed however since the redshift of all objects is known. κ must therefore be replaced by its redshift dependent form, $\kappa(z_S)$ where z_S refers to the source redshift.

Evidently, $\kappa(z_S)$ can not be used directly as a likelihood parameter for Σ. It is

therefore necessary to separate $\kappa(z_S)$ into two parts, one containing all z_S dependency and the other containing only lens redshift terms. This is easily accomplished from rearrangement of equation (3.10) to give

$$\kappa(z_S) = \frac{8\pi G}{cH_0}\Sigma\left[\frac{(1+z_L)^{1/2}-1}{(1+z_L)^2}\right]\left[\frac{(1+z_S)^{1/2}-(1+z_L)^{1/2}}{(1+z_S)^{1/2}-1}\right] = \kappa_\infty f(z_S) \qquad (5.32)$$

where z_L is the fixed lens redshift and $f(z_S)$ is a function which simplifies the notation by defining,

$$\kappa_\infty = \frac{8\pi G}{cH_0}\Sigma\left[\frac{(1+z_L)^{1/2}-1}{(1+z_L)^2}\right]. \qquad (5.33)$$

This parameterisation of $\kappa(z_S)$ conveniently defines the quantity κ_∞ (BTP) which is used as the likelihood estimator for Σ in this case.

Maximum likelihood determination of κ_∞ was therefore performed using equation (2.59) of Section 2.3.3 with $\phi(M)$ described by the Schechter function stipulated by the parameters in (5.29) above. Instead of using the amplification μ, a redshift dependent version of the κ estimators in equations (2.66) and (2.68) were used by writing

$$\mu(z_S) = \begin{cases} |1-2\kappa_\infty f(z_S)|^{-1} & \text{isothermal} \\ |(\kappa_\infty f(z_S)-c)(\kappa_\infty f(z_S)-1/c)|^{-1} & \text{parabolic} \\ (1-\kappa_\infty f(z_S))^{-2} & \text{sheet} \end{cases}. \qquad (5.34)$$

Note that the likelihood estimator formed from the probability in equation (2.59) involves the bivariate function $\phi(M,z)$. Only the magnitude dependence of this function is considered here of course since no consideration has been given to the evolutionary aspects of the CADIS luminosity function (see the introduction to Section 5.3).

Error Sources

Three sources of error on the value of κ_∞ were taken into consideration:

1 The maximum likelihood error obtained as the width of the likelihood distribution at $\ln\mathcal{L}_{max} - 0.5\Delta\chi^2$. All object magnitudes and redshifts were taken as presented directly in the A1689 data while assuming the Schechter parameters from (5.29).

2 The uncertainty of the Schechter parameters M_* and α from the likelihood analysis of the CADIS offset field.

3 The redshift and magnitude uncertainties of individual objects in the A1689 data, derived from the photometric analysis.

The latter two items in this list were dealt with simultaneously using Monte Carlo simulations once again. 1000 realisations were performed in which values of M_* and α were drawn at random from the convolved likelihood distribution shown in the lower part of Figure 5.14. For each realisation, redshifts and absolute magnitudes of objects were scattered in accordance with their associated photometric errors. The standard deviation of the scattered values of κ produced in this way was then added in quadrature to the uncertainty of the maximum likelihood error obtained from item one of the list above to give the overall error.

The magnitude calibration error of $\Delta M = 0.01$ discussed in Section 5.2.4 was ignored. Inspection of the form of the Schechter function in equation (2.60) shows that a systematic magnitude offset is exactly equivalent to an error in M_*. As (5.29) shows, the 1σ error in M_* is ~ 0.6 which completely overwhelms any uncertainty in magnitude calibration.

Source Parity

Interpretation of the results returned from the likelihood analysis requires a careful consideration of parity before κ_∞ can be evaluated. Since all source objects were assumed to lie at a fixed redshift in the work of Chapter 3, the critical line position about which parities were flipped remained fixed. Parity must be dealt with differently when a distribution of source redshifts exists since the critical line radius scales with source redshift as $\theta_c \propto D_a(z_L, z_S)/D_a(z_S)$ (see Section 2.1.7). Each source has its own critical radius and therefore a different critical value of κ_∞ about which its likelihood distribution is reflected. This is seen from inspection of the estimators in equation (5.34).

The contribution to the maximum likelihood distribution of κ_∞ in a given region of the field of view from each source was thus determined in the following manner:

- The critical line radius for the source was calculated from its redshift. This was

achieved using the proportionality noted above, normalised to the arc observed to the south-east of the cluster (lower left of cluster centre in Figure 5.2). With a radius of $\theta \simeq 0.85'$ and a redshift of $z = 0.8$ (Tyson & Fischer 1995), this arc gives the critical line radius as

$$\theta_c = 1.24 \, \frac{0.92(1+z_S)^{1/2} - 1}{(1+z_S)^{1/2} - 1} \tag{5.35}$$

for Abell 1689 using the results of equations (3.8) and (3.9) for an Einstein-de-Sitter universe.

- The source parity was determined from the spatial location of the object with respect to the region enclosed by its critical line.

- Applying equation (2.59) with the fitted CADIS Schechter function, the probability distribution for κ_∞ was calculated. Being reflected about its critical κ_∞, this distribution typically has two equal peaks. Depending on the source parity chosen, one of the peaks is taken as the correct solution and extrapolated over the whole κ_∞ range searched by the likelihood analysis (see Figure 5.16).

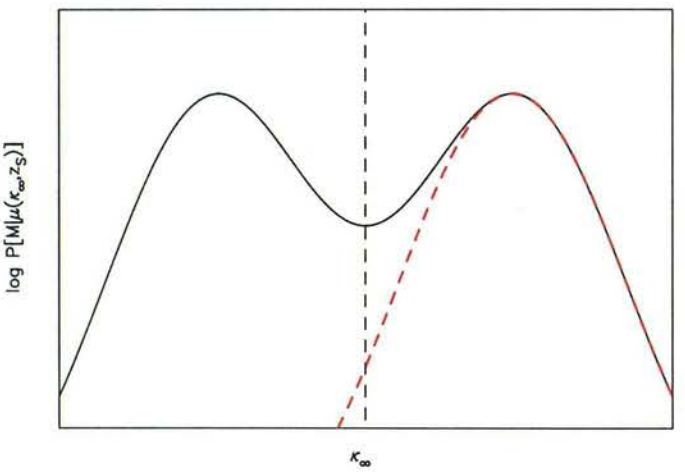

Figure 5.16: Illustration of an extrapolated probability distribution peak for a source at redshift z_S. The extrapolated curve is the source's contribution to the overall likelihood distribution. The vertical dashed line locates the critical κ_∞.

Each source probability distribution calculated in this way was combined with the distribution obtained for all of the other sources in the region under consideration to arrive at the overall likelihood distribution for that region. The final mass calculated proves to have a negligible sensitivity to the normalisation given by the arc redshift.

All masses in the following section were calculated assuming a range of arc redshifts from $z = 0.8 - 1.2$ and no noticeable change in the results occurred.

5.5 Results

5.5.1 Radial Mass Profile

The annular binning of objects about the cluster centre seen in the number count analysis of Section 3.2.4 was applied to the luminosity function method of this chapter. The relatively small number of background objects contained in the A1689 sample however was unfortunately insufficient to allow calculation of a profile of the same resolution as the number count profile.

Apart from the effects of shot noise, this limitation results from the simple fact that bins which are too narrow do not typically contain a large enough number of intrinsically bright objects. This has the effect that the knee of the Schechter function assumed in the likelihood analysis is poorly constrained. As equation (2.59) shows, a large uncertainty in M_* directly results in a large error on the magnification and hence on κ_∞. Experimentation with a range of bin widths quickly showed that in order to achieve a tolerable precision for κ_∞, bins had to be at least ~ 1.1 arcmin in width. With the observed field of view, this gave a limiting number of merely three bins, illustrated in the lower half of Figure 5.17.

The upper half of Figure 5.17 shows the κ profile of Section 3.2.4 overlaid by data points calculated in the three bins using the luminosity function method. The upper data points correspond to the sheet estimator while the lower points are due to the isothermal estimator. The 1σ error bars plotted here were calculated taking all three contributions listed in Section 5.4.1 into account (with 1000 Monte Carlo simulations). Values of κ_∞ were converted to values of κ for the purpose of comparison with the results of Section 3.2.4 using,

$$\kappa = \frac{\kappa_\infty}{N_b} \sum_{i=1}^{N_b} f(z_i) = \kappa_\infty < f_b > \tag{5.36}$$

where N_b is the number of galaxies in bin b and $< f_b >$ denotes the average value of $f(z)$ in the bin which was consistently found to be $\simeq 0.57$.

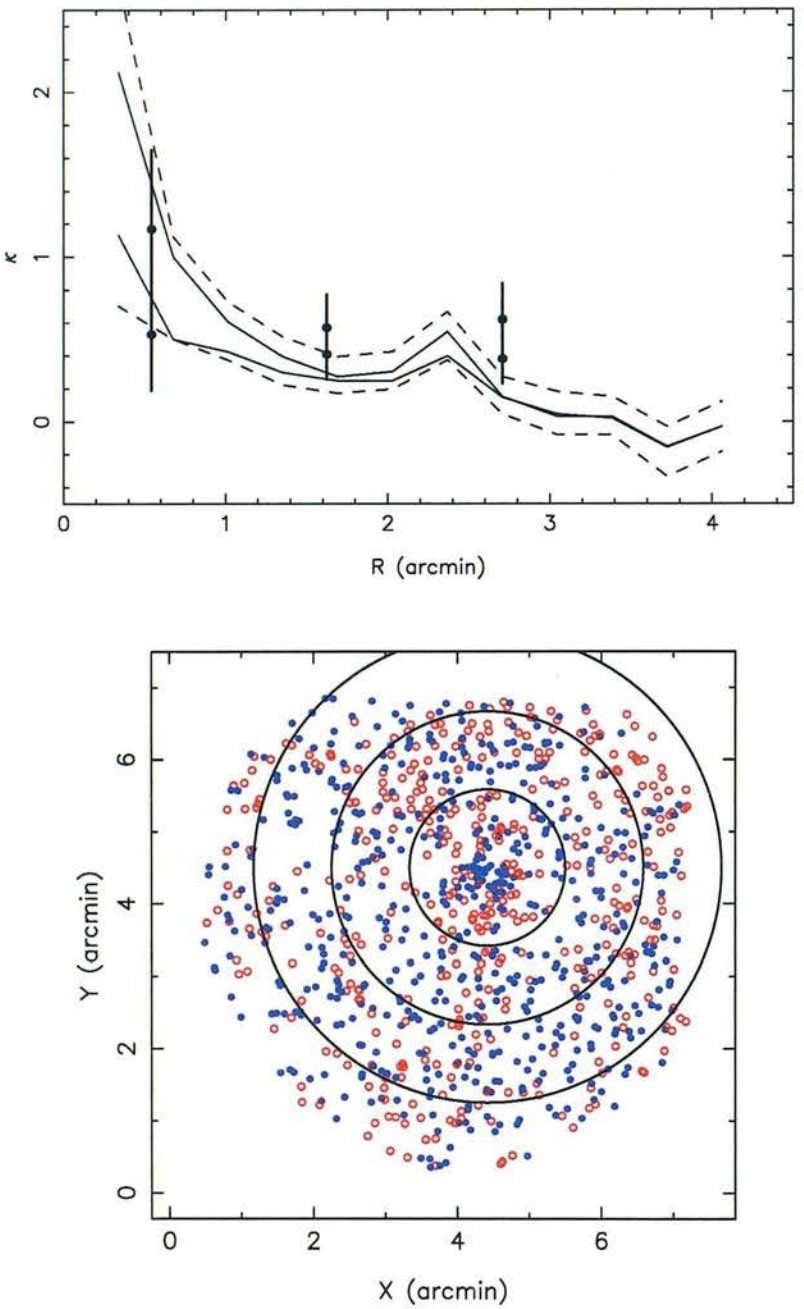

Figure 5.17: *Top*: Comparison of radial κ profiles. Data points show isothermal (lower) and sheet (upper) estimated κ obtained from the luminosity analysis. 1σ error bars are plotted. The solid and dashed lines indicate the profile obtained in Chapter 3 with the number count method. *Bottom*: Spatial location of the annular bins on the A1689 field of view. Open red dots are objects selected by $z > 0.2$ and solid blue dots by $z \leq 0.2$.

Despite the relatively large errors, the data points from the luminosity function analysis show a fall off with radius, in good agreement with the profile derived from

the number count study. It is noticeable however that the profile they indicate is perhaps a little flatter than that derived from the number counts. It appears that more mass is detected at larger radii although this is not particularly significant.

5.5.2 Aperture κ

In addition to the radial profile, the variation of average surface mass density contained within a given radius was calculated. By applying the likelihood analysis to the objects contained within an aperture of varying size, a larger signal to noise was permitted at larger radii where more objects are encompassed. With a small aperture, the same low galaxy count problem is encountered as Figure 5.18 shows by the large uncertainty in this vicinity. In this plot, the parabolic estimator of equation (5.34) is used to obtain κ_∞ which is scaled to κ using equation (5.36). The orange region depicts the 1σ errors, again taking all 3 sources of uncertainty from Section 5.4.2 into account. The solid black line shows the variation of aperture κ calculated by averaging the parabolic estimator profile of Section 3.2.5. The errors on this are shown by the dashed lines but account only for clustering and shot noise without inclusion of the uncertainty of background count normalisation or redshift.

As expected from the results of Section 5.5.1, generally more mass than that predicted from the number counts is seen, especially at large radii. The following section quantifies this for a comparison with the projected mass result of Chapter 3.

5.5.3 Projected Mass

From the values of κ_∞ used to generate the κ profile in Section 5.5.1 and the result of equation (5.33), the cumulative projected masses in Table 5.3 were calculated. Errors were derived from propagation of the errors on the binned values of κ_∞.

These projected masses are in excellent agreement with those of Chapter 3. Since $1' = 0.117h^{-1}\mathrm{Mpc}$ in the cluster frame, the second cumulative mass listed in Table 5.3 gives

$$M_{2d}(< 0.25h^{-1}\mathrm{Mpc}) = (0.48 \pm 0.16) \times 10^{15} h^{-1}\mathrm{M}_\odot \tag{5.37}$$

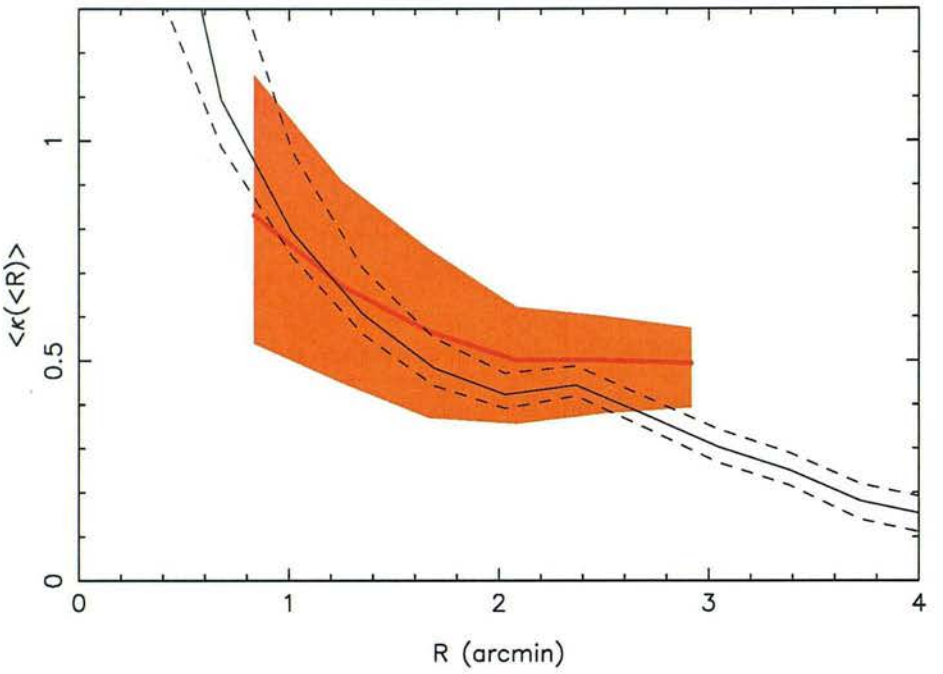

Figure 5.18: Variation of average surface mass density contained within a given radius R (red line). 1σ errors are shown by the orange region. The black solid and dashed lines show the average surface mass density and 1σ errors (due to likelihood analysis only) of the parabolic estimated profile of Section 3.2.5.

Aperture Radius (arcsec)	Aperture Area (Mpc^2)	$M_{2d}(< R)$
65	0.050	$(0.16 \pm 0.09) \times 10^{15} h^{-1} \mathrm{M}_{\odot}$
130	0.202	$(0.48 \pm 0.16) \times 10^{15} h^{-1} \mathrm{M}_{\odot}$
195	0.454	$(1.03 \pm 0.27) \times 10^{15} h^{-1} \mathrm{M}_{\odot}$

Table 5.3: Cumulative projected masses at the radii given by the profile of Section 5.5.1.

which is perfectly consistent with the result from the number count study. The error here is also essentially the same as the 30% error quoted on the result of Chapter 3 in allowing for the effects of uncertainty in the background number count slope, redshift distribution and surface number density. The projected mass contained within 195 arcsec is a little higher than that from the number count work although given the errors involved, this can still be regarded consistent.

5.5.4 Effects of Sample Incompleteness

One final uncertainty not taken into consideration thus far is that due to sample incompleteness. Changing the limiting apparent B magnitude m_{max} in the determination of the CADIS luminosity function directly affects the fitted values of M_*, α and hence the maximum likelihood κ_∞. Similarly, the differing numbers of objects included in the A1689 sample from variations in m_{max} also has an influence on κ_∞.

Table 5.4 quantifies this effect for the CADIS objects. It can be seen that increasing m_{max} (ie. including fainter objects) has little effect on κ_∞ until the limit $m_{max} \simeq 24.5$ is reached. Beyond this limit, κ_∞ starts to fall. Two implications can therefore be made. Firstly, this suggests that the magnitude limit in Section 5.3.2 from the V/V_{max} test, being consistent with the limit here, was correctly chosen. Secondly, κ_∞ is relatively insensitive to the choice of m_{max} if the sample is complete (and not smaller than the limit at which shot noise starts to take effect).

m_{max}	M_*	α	κ_∞(iso)	κ_∞(para)	κ_∞(sheet)
25.5	-20.57	-0.80	$0.61^{+0.03}_{-0.04}$	$0.69^{+0.04}_{-0.06}$	$0.76^{+0.07}_{-0.08}$
25.0	-20.76	-1.10	$0.65^{+0.04}_{-0.04}$	$0.77^{+0.06}_{-0.06}$	$0.84^{+0.08}_{-0.09}$
24.5	-20.94	-1.45	$0.70^{+0.06}_{-0.04}$	$0.85^{+0.08}_{-0.08}$	$0.96^{+0.10}_{-0.10}$
24.0	-21.49	-1.87	$0.74^{+0.03}_{-0.04}$	$0.91^{+0.13}_{-0.12}$	$1.08^{+0.16}_{-0.17}$
23.5	-20.75	-1.53	$0.75^{+0.04}_{-0.04}$	$0.90^{+0.06}_{-0.07}$	$1.10^{+0.10}_{-0.09}$

Table 5.4: Variation of limiting apparent B magnitude m_{max} of the CADIS field and its effect on the Schechter parameters and the resulting value of κ_∞. The apparent magnitude limit of $b = 23.7$ was assumed for the A1689 data in calculating the maximum likelihood κ_∞. *Errors are taken only from the width of the likelihood curves.*

The effect of varying the magnitude limit of the A1689 sample is quantified in Table 5.5. A clear trend is also seen here; as m_{max} is reduced, κ_∞ falls. Assuming linearity, a rough estimate of the uncertainty of κ_∞ given the uncertainty of the sample magnitude limit is given by:

$$\Delta\kappa_\infty = \begin{cases} \sim 0.1\Delta m_{max} & \text{isothermal} \\ \sim 0.2\Delta m_{max} & \text{parabolic} \\ \sim 0.4\Delta m_{max} & \text{sheet} \end{cases} \tag{5.38}$$

m_{max}	κ_∞(iso)	κ_∞(para)	κ_∞(sheet)
24.5	$0.77^{+0.03}_{-0.03}$	$1.03^{+0.11}_{-0.11}$	$1.30^{+0.15}_{-0.13}$
24.0	$0.76^{+0.04}_{-0.04}$	$0.94^{+0.10}_{-0.10}$	$1.12^{+0.13}_{-0.12}$
23.5	$0.69^{+0.06}_{-0.07}$	$0.79^{+0.10}_{-0.11}$	$0.92^{+0.13}_{-0.12}$

Table 5.5: Variation of the maximum likelihood determined κ_∞ with limiting apparent B magnitude m_{max} of the A1689 data. The Schechter parameters of Section 5.3.3 were assumed in the likelihood analysis.

Referring to Figure 5.15, a suitable uncertainty in m_{max} of the A1689 sample of say ± 0.2 magnitudes might be argued. If this were the case, the projected masses of the previous section calculated with the parabolic estimator would have a further error of $\sim 5\%$.

5.6 Signal to Noise Predictions

Including all possible contributions of uncertainty in the calculation of mass, the previous section has shown that even with relatively few galaxies, a significant cluster mass profile can be detected. One can make predictions of the sensitivity of the method with differing input parameters potentially obtained by future measurements. This exercise also serves as an optimisation study, enabling identification of quantities requiring more careful measurement and those which play an insignificant part.

The most convenient means of carrying out this investigation is by application of the reconstruction method to simulated galaxy catalogues. Catalogues were therefore constructed by randomly sampling absolute magnitudes from the Schechter function fitted to the CADIS offset field in Section 5.3.3. Redshifts were assigned to each magnitude by randomly sampling the distribution parameterised by T98 (their equation 22) from the Canada France Redshift Survey (Lilly et al. 1995). A range of catalogues were produced, varying by the number of objects they contained and their distribution of galaxy redshift errors modeled from the A1689 data.

Figure 5.19 shows how the distribution of photometric redshift error, σ_z, correlates with the A1689 B-band apparent magnitude. No significant correlation between σ_z and redshift was found. Catalogue objects were thus randomly assigned redshift errors

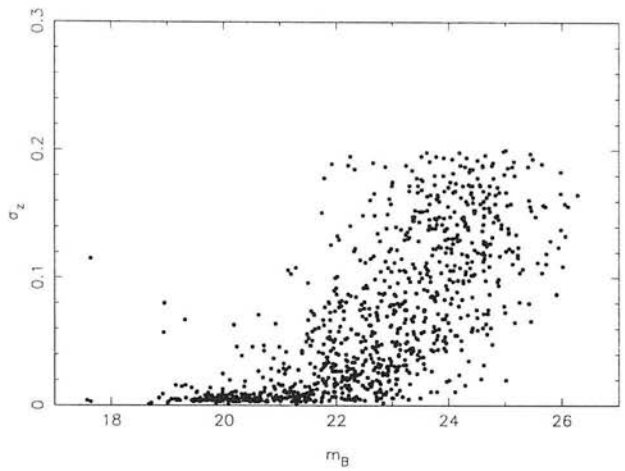

Figure 5.19: Correlation of photometric redshift error with apparent B-band magnitude for the A1689 data. No significant correlation between σ_z and z exists.

in accordance with their apparent magnitude, given by the correlated distribution in Figure 5.19. Different catalogues were generated from different scalings of this distribution along the σ_z axis.

Each catalogue was then lensed with a sheet mass characterised by $\kappa_\infty = 1$ before applying the reconstruction. 1000 Monte Carlo realisations were performed for each reconstruction, scattering object redshifts according to their assigned errors in the same manner as in the reconstruction of A1689. Furthermore, to model the uncertainty associated with the offset field, assumed values of the Schechter parameters M_* and α were once again subject to Monte Carlo realisations. All catalogues were reconstructed assuming sets of Schechter parameters drawn from a range of scaled versions of the distribution shown in Figure 5.14.

The resulting scatter measured in the reconstructed value of κ_∞ for each catalogue and assumed α-M_* scaling was combined with the average maximum likelihood error across all realisations of that catalogue to give an overall error. This total error was found to be well described by,

$$\sigma^2_{\kappa_\infty} = \frac{1 + (2\sigma_z)^2}{n} + (0.12\sigma_{M_*})^2 + (0.37\sigma_\alpha)^2 - 0.18\sigma_{\alpha M_*} \tag{5.39}$$

where n is the number of galaxies, σ_z is the sample average redshift error and σ_{M_*} and σ_α are the projected errors on M_* and α respectively as quoted in equation (5.29).

The quantity $\sigma_{\alpha M_*}$ is the covariance of α and M_* defined by

$$\sigma_{\alpha M_*} = \frac{\int \mathcal{L}(M_*, \alpha)(M_* - \langle M_* \rangle)(\alpha - \langle \alpha \rangle)\ \mathrm{d}M_* \mathrm{d}\alpha}{\int \mathcal{L}(M_*, \alpha)\ \mathrm{d}M_* \mathrm{d}\alpha} \tag{5.40}$$

where the likelihood distribution $\mathcal{L}$ is given by equation (5.27). We find that $\sigma_{\alpha M_*} = 0.039$ for the CADIS offset field. Equation (5.39) is valid for $n \geq 20$ and $0.0 \geq \sigma_z \geq 0.3$.

Equation (5.39) shows that when the number of objects is low, shot noise dominates. With $n \simeq 200$ however, uncertainties from the calibration of the offset field start to become dominant. The factor of 2 in the photometric redshift error term stems from the fact that redshift errors also translate directly into absolute magnitude errors through equation (5.28). Another discrepancy arises when comparing this redshift error with the redshift error contribution of 25% claimed for the A1689 data in Section 5.4.2. This is accounted for by the fact that K-corrections were present in the A1689 data whereas in the simulated catalogues there were not. Equation (5.28) quantifies the increase in magnitude error with the inclusion of K-corrections. This translates to an approximate increase of 20% in the overall error with an average K-correction of -1.0 for the A1689 data.

Emphasis should be placed on the criteria for which equation (5.39) is valid. The predicted overall error rises dramatically when fewer than $\simeq 20$ objects are included in the analysis. Simulations with 15 objects resulted in maximum likelihood errors rising to beyond twice that predicted by simple shot noise. This stems mainly from the effect mentioned in Section 5.5.1, namely the failure of the likelihood method when the knee of the Schechter function is poorly constrained.

Section 6.3 discusses possible measures to take based on the predictions of these simulations in order to optimise a multi-colour study such as this.

5.7 Summary

Photometric redshifts and magnitudes have been determined for objects in the field of Abell 1689 from multi-waveband observations. This has allowed calculation of the luminosity function of source galaxies lying behind the cluster. Comparison of this with the luminosity function obtained from a similar selection of objects in an

unlensed offset field has resulted in the detection of a bias in the A1689 background object magnitudes attributed to lens magnification by the cluster. After a careful consideration of all possible sources of uncertainty, calculation of a significant radial mass profile for A1689, consistent with that obtained in the number count study of Chapter 3 has been calculated. This predicts a projected mass interior to $0.25h^{-1}$Mpc of

$$M_{2d}(< 0.25h^{-1}\mathrm{Mpc}) = (0.48 \pm 0.16) \times 10^{15} h^{-1}\mathrm{M}_{\odot} \tag{5.41}$$

in excellent agreement with the number count result.

This chapter has been primarily devoted to demonstrating the viability of mass reconstruction using the luminosity function method of BTP. The technique is limited mainly by shot noise arising from limited numbers of background galaxies, as the signal-to-noise analysis of Section 5.6 has shown, but is independent of their clustering. Section 6.3 discusses ways of improving the efficiency of the method.

Chapter 6

Summary

This thesis has focussed on the development and application of several gravitational lens mass reconstruction techniques. In particular, attention has been given to those methods which exploit the property of lens magnification since this allows measurement of absolute mass. This is in contrast to methods based upon quantifying image shear which are plagued by the sheet-mass degeneracy (Falco, Gorenstein & Shapiro 1985) limiting determination of lens mass to differential values.

The purpose of this short chapter is to summarize the results and conclusions made in Chapters 3, 4 and 5 in addition to providing a prediction of the role of lensing in the future. The following short section gives a history of the development of methods seen in this work. Section 6.2 compares and contrasts the main results of this thesis with those obtained via alternative methods employed by other researchers in the field. In Section 6.3, consideration is given not only to work which could follow that presented in this thesis but also new lines of research which lensing is beginning to evolve towards.

6.1 A Short History

Several different approaches have been taken throughout the work of Chapters 3, 4 and 5. There are benefits and disadvantages to each.

Reconstruction using the projected number density dilution of background red galaxies as applied to Abell 1689 in Chapter 3 is a simple and an efficient method. However, prior knowledge of the limits on a cluster's shear field must be assumed before mass estimations can be made. While this was shown via agreement with other methods not to pose any major problems, it has the consequence of obtaining less accurate solutions being able to provide essentially only an upper and lower bound. Although this was circumvented in 1D by providing a self-consistent axial solution, a self-consistent solution for γ and κ in 2D remained undiscovered.

The lack of a self-consistent solution in 2D provided motivation for the work in Chapter 4. Previous attempts at generating a self-consistent solution by direct iteration from a measured magnification field (see Section 2.3.4) were known to fail upon introduction of critical lines. By pixelizing the magnification data so that the problem could be interpreted as a matrix inversion calculation, a method for producing self-consistent 2D solutions with the presence of critical lines was derived. This method does not provide an alternative means of measuring magnification but relies upon its measurement prior to application.

An alternative method of measuring magnification came in Chapter 5; the so-called luminosity function method. By determining the increase in observed background source flux, the amplification by Abell 1689 was measured. This important piece of work is the first ever investigation into the flux magnification by lensing. In addition to providing a proof of concept, it enabled verification of the measurements of magnification made in Chapter 3. The method requires measurement of intrinsic source brightnesses and as Chapter 5 showed, this involves a relatively large amount of observation even using more efficient photometric redshift techniques (although see Section 6.3). This disadvantage aside, the method boasts independence of source clustering, unlike the more simplistic number count approach of Chapter 3.

In summary, choice of the most suitable method should be decided by the amount of available observing time. A quick yet less accurate study can be performed with two-filter observations using the number count method whereas a more detailed, clustering noise independent result can be acheived with several filters using the luminosity function method.

The main results of Chapters 3, 4 and 5 are now summarised in the next section.

6.2 Conclusion of Results

6.2.1 Number Count Study

In chapter 3, the viability of cluster mass reconstruction using the integrated galaxy number counts technique of Section 2.3.2 was demonstrated. By applying colour cuts to V and I band data of objects in the field of Abell 1689, the background population of red galaxies was selected. These sources were shown to exhibit a dilution in surface number density in regions of high magnification due to their relatively flat number count slope. The effects of non-linear source galaxy clustering and shot noise were taken into consideration for the computation of a projected mass profile and 2D mass map of A1689.

The derived 2D mass distribution of A1689 was shown to exhibit an extended structure to the south-west of the cluster. Allowing for shot noise and source clustering, the significance of this peak was calculated to be at a 3σ level. Given that X-ray observations of A1689 with the ROSAT satellite in the $0.5-2$ keV band (Wang & Ulmer 1997, Daines et al 1998) show a relatively compact and smooth 2D flux distribution without such an extension, this is a little surprising. Two possible explanations were given. The first was that this truly is a clump of dark matter not followed by the X-ray emitting cluster gas. The second was simply attributed to a rare underdensity of background galaxy counts. Curiously, a shear analysis by Kaiser (1996) on the same data as that of Chapter 3 hinted at this structure although caution is taken in claiming a consistent detection.

Chapter 3 also derived an analytical relationship to provide a self-consistent mass and shear profile by assuming an axially symmetric mass distribution. Using this relationship with the radial magnification distribution, a projected mass interior to $R = 0.24h^{-1}\mathrm{Mpc}$ of

$$M_{2D}(< 0.24h^{-1}\mathrm{Mpc}) = (0.50 \pm 0.09) \times 10^{15} h^{-1}\mathrm{M}_{\odot} \tag{6.1}$$

was calculated. The errors here account for the effects of shot noise and source clus-

tering. Systematics due to number count normalisation from the offset field and the uncertainty of background object redshifts increases this total error to $\sim 30\%$.

Comparison with mass estimates from shear analysis confirms that this result is in good agreement. The shear analysis of A1689 by Kaiser (1996) calculates that $M_{2D}(< 0.24h^{-1}\mathrm{Mpc}) = (0.43 \pm 0.04) \times 10^{15}h^{-1}\mathrm{M}_{\odot}$ while that predicted by Tyson & Fischer (1995) is $M_{2D}(< 0.24h^{-1}\mathrm{Mpc}) = (0.43 \pm 0.02) \times 10^{15}h^{-1}\mathrm{M}_{\odot}$. Section 3.2.9 provides more details and includes a comparison with X-ray and virial measurements.

6.2.2 Pixellization of Magnification and Shear

Chapter 4 derived a method for directly calculating accurate, self-consistent surface mass density and shear distributions from gridded lens amplification values and critical line positions. This was demonstrated with the isothermal sphere and dumb-bell cluster models in reconstructing the surface density to within 1% over most of the field of view. Reconstruction of the shear pattern was demonstrated to have a lower fractional accuracy of a few tenths due to boundary effects. The method was applied to the magnification data of Chapter 3 to enable computation of the 2D surface mass density and shear distribution of Abell 1689. The resulting surface mass density distribution produced by this novel new technique was found to be somewhere between that calculated by the isothermal estimator and that derived by the sheet estimator in the number count study. This is exactly as one would have expected, given that the mass profile deduced in Chapter 3 showed evidence of A1689 being somewhere between that of an isothermal sphere and a sheet mass.

A simplified solution to the problem of estimating surface mass density from galaxy ellipticities was also presented in Chapter 4. This approach puts the calculation of mass from shear and magnification on an equal footing, essential for a reconstruction technique which encompasses both (see Section 6.3.1).

6.2.3 Magnification of the Source Luminosity Function

The effect of lens magnification by Abell 1689 to enhance the flux of background source galaxies and hence distort their luminosity function was investigated in Chapter 5.

This work (see Dye et al 1999) is the first ever application of the luminosity function reconstruction method suggested by Broadhurst, Taylor and Peacock (1995).

Photometric redshifts and magnitudes were determined for all objects in the field of view from a set multi-waveband observations observed with the Calar Alto 3.5m telescope in Spain. Comparison of the background source galaxy luminosity function with that obtained from a similar selection of objects in an unlensed offset field resulted in the detection of a magnification bias in the A1689 background object magnitudes. After a careful consideration of all possible sources of uncertainty, calculation of a significant radial mass profile for A1689, consistent with that obtained in the number count study of Chapter 3 was calculated. This predicted a projected mass interior to $0.25h^{-1}$Mpc of

$$M_{2d}(< 0.25h^{-1}\mathrm{Mpc}) = (0.48 \pm 0.16) \times 10^{15} h^{-1}\mathrm{M}_{\odot} \tag{6.2}$$

perfectly consistent with the number count result.

The luminosity function method is independent of background galaxy clustering since it can be assumed that luminosities are a random sampling of luminosity space. Section 5.6 showed how the method is limited by the number of background galaxies observed although this is easily rectified with deeper observations or use of larger telescopes (see Section 6.3).

6.2.4 Constraints on the Structure of A1689

A comparison of measurements of A1689 using X-ray and virial methods suggests that A1689 is not an isolated cluster. This is reflected in the fact that X-ray temperature measurements of the cluster indicate a lower mass than the lensing mass and that virial analyses measure a larger velocity dispersion than expected from the lensing mass (Giradi et al 1997). This is consistent with the notion of A1689 being comprised of clumps of mass aligned along the line of sight. If this were the case, cluster galaxy velocity measurements would yield a higher spread in their distribution and the X-ray temperature would be dominated by the temperature of only the larger clump.

Miralda-Escudé & Babul (1995) point out that X-ray and lensing mass discrepancies seen in clusters in general cannot be accounted for by cluster gas inhomogeneity,

non-thermal pressure support of the gas or non-isothermality alone. They hypothesise that a combination of all three of these effects could be the cause of such discrepancies. While it is certainly plausible that this may be the explanation in the majority of cases, the largest contribution possible from all three of these effects still falls somewhat short of explaining particularly large mass disagreements.

Section 3.2.9 demonstrated how decomposing A1689 into two aligned clumps approximately explained the relatively small X-ray mass and the large virial mass. Daines et al (1998) lend further support to the 'multi-clump' idea by pointing out that in fact three substructures are required to bring the mass measurements of A1689 into agreement with each other. This, they claim, is supported by the cluster redshift measurements of Teague, Carter & Gray (1990) which shows tentative evidence of three distinct groups of redshift detections. A more detailed spectral study of the cluster would undoubtedly clear this ambiguity.

6.3 Future Work

Lensing is an area of research still very much in its youth. Much of the work carried out in the field serves only to prove the viability of lensing and much of it is devoted to devising new techniques. There are several areas in need of attention. This section aims to highlight some of those and predict the future roles lensing will take.

6.3.1 Optimal Combination of Magnification and Shear

Magnification and shear measurements complement one another. This is apparent by virtue of the simple fact that magnification unlike shear provides absolute mass measurements whereas shear analysis generally allows better image resolution. A method which optimally combines both is therefore a very powerful one.

Attempts to form a hybrid reconstruction method using combined shear and magnification information has received some attention in the last few years although there is much room for improvement. Work by Bartelmann et al (1996, see Section 2.3.4) approach this hybridisation using a least squares method. They use an iterative ap-

proach with quantities being calculated from a grid of values of the underlying lens deflection potential. Another promising means of combining shear and magnification data is via the maximum entropy method (MEM). The MEM was first applied to lensing by Wallington, Narayan & Kochanek (1994) who used it for the reconstruction of 1D image flux profiles. The method has the advantage that it performs well with noisy data. As noted by Bridle et al (1998) who use it to calculate 2D mass distributions, MEM can reproduce cluster structures slightly beyond the field of view.

These efforts are only a start however. A hybrid technique which enables the inclusion of critical lines remains elusive. Another reason for concern is the correlation of errors between shear and magnification data. For example, the uncertainty from shot noise affects both magnification and shear measurements in the same way. Regions with a large number density of background galaxies give rise to a small shot noise contribution to both. Clearly there must exist an optimal way of combining the two measurements. Ideally, inclusion of the other error sources for each method would give rise to complementary errors such that a noisy measurement from one would be compensated for by an accurate measurement from the other. A study to investigate this is required.

6.3.2 Optimisation of Luminosity Function Method

Section 5.6 demonstrated that the dominant source of error on the mass determination of Chapter 5 was shot noise. The most immediate improvement to a multi-colour study such as this would therefore be to increase galaxy numbers. Also noted in Section 5.6 was the fact that only when bins contain $\simeq 200$ objects do offset field uncertainties become important. Observing in broader filters is one way to combat the limit presented by galaxy numbers. The final number of 958 galaxies classified by the photometric redshift analysis of Section 5.2.3 was limited mainly by the data observed in the narrow 466/8 filter. Despite being observed to approximately the same integration time as the 466/8 filter, the much broader I band filter enabled detection of ~ 3000 galaxies. Using broader filters will also inevitably give rise to less accurate photometric redshifts. However as the analysis of the signal-to-noise simulations of Section 5.6 showed, one can afford to sacrifice redshift accuracy quite considerably

before its contribution becomes comparable to that of shot noise.

Deeper observations provide another obvious means of increasing the number of galaxies. The error predictions of Section 5.6 indicate that the expected increase in galaxy numbers using an 8 metre class telescope with the same exposure times as those in Chapter 5 work should reduce shot noise by a factor of ~ 3. Since deeper observations would also reduce redshift and offset field calibration uncertainties to negligible levels, the only source of error would be shot noise. In this case, the signal to noise for κ_∞ from equation 5.39 becomes simply $\kappa_\infty \sqrt{n}$ and hence the mass estimate for A1689 could be quoted with a 9σ certainty.

6.3.3 Weak Lensing Surveys

In exactly the same manner that a foreground cluster gravitationally deflects light from background sources, large scale mass fluctuations can weakly distort the images of distant galaxies. Blandford & Jaroszynski (1981) first gave a quantitative estimate of the amplitude of this effect. Valdes, Tyson & Jarvis (1983) were the first to attempt to measure it although their results were inconclusive due to technical limitations. More recently, Kaiser et al (1998) proved that weak lensing by large scale structure can indeed be detected by analysing the shear pattern of background galaxies in a $15' \times 15'$ image of the supercluster MS0302+17.

In a more statistical approach, Blandford (1990) and Miralda-Escudé (1991) computed the two-point shear correlation function and established how it depends on the power spectrum of density fluctuations in an Einstein-de-Sitter universe. This was generalised to allow for any value of Ω by Villumsen (1996). Bernardeau, van Waerbeke & Mellier (1997) showed how the second and third order moments of κ depend on the cosmological parameters. This provides a mechanism to enable determination of Ω and Λ from observations of cosmic shear. All studies agree that the effects are small; the rms amplitude of distortions is expected to be $\sim 1\%$ with a correlation length of $\sim 1^\circ$. With the advent of new surveys such as that which will be produced by the proposed VISTA telescope, this weak signal will be easily detected. Lensing will thus place valuable constraints on the cosmological parameters which describe our Universe for comparision with the results of NASA's CMBR surveying satellite,

MAP, due for launch in 2000. Both of these will provide excellent foundations for ESA's highly awaited and more ambitious Planck satellite to be launched in 2007.

6.3.4 Outlook

Lensing is fast beginning to establish itself as a valuable and reliable tool. The next few years will more than likely see greater application of lens mass reconstruction methods such as those presented in this thesis. More numerous and more accurate cluster mass measurements are still required to gain a deeper knowledge of large scale structures. These methods will be progressively applied to higher redshift clusters as attempts are made to understand cluster formation in realms where the equilibrium assumptions of X-ray and virial mass techniques become even more invalid. As well as studying highly dense mass fluctuations such as clusters, investigations into weaker large scale mass fluctuations will begin to adorn the pages of various research journals in an attempt to construct a more complete picture of our Universe.

Appendix A

Mathematical Proofs

A.1 Axially Symmetric Lens

The inner integral in equation (2.71) may be re-written as

$$\begin{aligned}\int_0^{2\pi} \mathrm{d}\phi \frac{\theta - \theta' \cos\phi}{\theta^2 + \theta'^2 - 2\theta\theta' \cos\phi} &= \frac{1}{2\theta}\int_0^{2\pi} \mathrm{d}\phi + \frac{1}{2}\int_0^{2\pi} \mathrm{d}\phi \frac{\theta^2 - \theta'^2}{\theta^3 + \theta\theta'^2 - 2\theta^2\theta' \cos\phi} \\ &= \frac{\pi}{\theta} + \frac{\theta^2 - \theta'^2}{2(\theta^3 + \theta\theta'^2)}\int_0^{2\pi} \mathrm{d}\phi \frac{1}{1 - \frac{2\theta^2\theta'}{\theta^3 + \theta\theta'^2}\cos\phi} \qquad \text{(A.1)}\end{aligned}$$

Using the identity

$$\int_0^{2\pi} \mathrm{d}\phi \frac{1}{1 + a\cos\phi} = \frac{2\pi}{\sqrt{1 - a^2}}, \quad a^2 < 1 \qquad \text{(A.2)}$$

allows the integral in equation (A.1) to be evaluated to give

$$\begin{aligned}\int_0^{2\pi} \mathrm{d}\phi \frac{\theta - \theta' \cos\phi}{\theta^2 + \theta'^2 - 2\theta\theta' \cos\phi} &= \frac{\pi}{\theta} + \frac{\theta^2 - \theta'^2}{2(\theta^3 + \theta\theta'^2)} \cdot \frac{2\pi}{\sqrt{1 - \frac{4\theta^2\theta'^2}{(\theta^2 + \theta'^2)^2}}} \\ &= \frac{2\pi}{\theta}. \qquad \text{(A.3)}\end{aligned}$$

If $\theta' > \theta$ then the integral in equation (A.1) evaluated between $0 < \phi < \pi$ is equal to exactly the opposite of its evaluation between $\pi < \phi < 2\pi$ and hence in total equals zero.

A.2 Isothermal Sphere Lens Model

A.2.1 Proof that $\kappa \propto \theta^{-1}$

The volume mass density of an isothermal sphere is given by (see for example Binney & Tremaine 1997),

$$\rho(\theta_{3d}) = \frac{\sigma^2}{2\pi G\theta_{3d}^2} \tag{A.4}$$

where σ is the velocity dispersion and the radius $\theta_{3d}^2 = \theta_1^2 + \theta_2^2 + \theta_3^2$. In order to obtain the surface mass density, the volume mass density must be integrated along the line of sight. Taking the line of sight to be θ_3, the surface mass density is therefore,

$$\begin{aligned}
\Sigma(\theta) &= 2\int_0^\infty \rho(\theta_{3d})\mathrm{d}\theta_3 \\
&= \frac{\sigma^2}{\pi G\theta}\int_0^\infty \frac{\mathrm{d}(\theta_3/\theta)}{1+\theta_3^2/\theta^2} \\
&= \frac{\sigma^2}{\pi G\theta}\left[\arctan\left(\frac{\theta_3}{\theta}\right)\right]_{\theta_3/\theta=0}^{\theta_3/\theta=+\infty} \\
&= \frac{\sigma^2}{2G\theta}
\end{aligned} \tag{A.5}$$

where $\theta^2 = \theta_1^2 + \theta_2^2$. Since $\kappa = \Sigma/\Sigma_{CR}$ from equation (2.9),

$$\kappa = \frac{\sigma^2}{2\Sigma_{CR}G\theta}. \tag{A.6}$$

A.2.2 Proof that $\kappa = \gamma$

This proof is quite trivial having done most of the derivation in Section 2.3.4. One begins by using the result of equation (2.79),

$$\gamma^2 = (\kappa - \overline{\kappa})^2. \tag{A.7}$$

$\overline{\kappa}$ is the mean value of κ within a circle of radius θ given by,

$$\overline{\kappa}(\theta) = \frac{2}{\theta^2}\int_0^\theta \theta'\mathrm{d}\theta'\kappa(\theta'). \tag{A.8}$$

Since the previous section showed that

$$\kappa = \frac{c}{\theta}, \quad c = \frac{\sigma^2}{2\Sigma_{CR}G}, \tag{A.9}$$

one can calculate an expression for $\overline{\kappa}$:

$$\overline{\kappa} = \frac{2}{\theta^2}\int_0^{\theta} \theta' \mathrm{d}\theta' \, \frac{c}{\theta'} = \frac{2c}{\theta} = 2\kappa. \tag{A.10}$$

Substitution into equation (A.7) then gives the required result:

$$\gamma^2 = (\kappa - 2\kappa)^2 = \kappa^2. \tag{A.11}$$

A.3 Pixellized κ

A.3.1 Analytical form for a_m

Derivation of the equation in question for a_m requires evaluation of the following:

$$a_m(\boldsymbol{\theta}) = \frac{1}{2\pi}\left(\frac{\partial^2}{\partial\theta_1^2} - \frac{\partial^2}{\partial\theta_2^2}\right)\int_m \mathrm{d}^2\theta' \ln|\boldsymbol{\theta} - \boldsymbol{\theta}'| \tag{A.12}$$

Operating on the integral with the partial derivatives and defining $x \equiv (\theta_1 - \theta_1')$ and $y \equiv (\theta_2 - \theta_2')$ gives

$$a_m(\boldsymbol{\theta}) = \frac{1}{\pi}\int_{y=r_2-0.5}^{y=r_2+0.5}\int_{x=r_1-0.5}^{x=r_1+0.5} \frac{y^2 - x^2}{(x^2+y^2)^2}\,\mathrm{d}x\,\mathrm{d}y. \tag{A.13}$$

The integral here acts over pixel m. Two assumptions have been made:

- It is assumed that pixels are square. Of course the derivation here may be generalised to account for rectangular pixels but this is rarely needed in practice.
- The units of length have been scaled so that pixels measure one unit on a side. This is purely to simplify the derivation and again can be generalised if necessary.

The vector $\boldsymbol{r}$ locates the centre of the pixel from the position $\boldsymbol{\theta}$ hence the integral is evaluated between $-0.5 \leq r_1 \leq 0.5$ and $-0.5 \leq r_2 \leq 0.5$ (see Figure 4.1 for clarification).

Integrating equation (A.13) gives

$$\begin{aligned} a_m(\boldsymbol{\theta}) &= \frac{1}{\pi}\left[\arctan\left(\frac{y}{x}\right)\right]_{y=r_2-0.5,\ x=r_1-0.5}^{y=r_2+0.5,\ x=r_1+0.5} \\ &= \frac{1}{\pi}\left\{\arctan\left(\frac{r_2-0.5}{r_1-0.5}\right) + \arctan\left(\frac{r_2+0.5}{r_1+0.5}\right)\right. \\ &\quad \left. - \arctan\left(\frac{r_2+0.5}{r_1-0.5}\right) - \arctan\left(\frac{r_2-0.5}{r_1+0.5}\right)\right\} \end{aligned} \tag{A.14}$$

Using the identity,

$$\arctan(x) \pm \arctan(y) = \arctan\left(\frac{x \pm y}{1 \mp xy}\right) \tag{A.15}$$

and grouping together the first and third terms and the second and fourth terms in equation (A.14) enables simplification to,

$$a_m(\boldsymbol{\theta}) = \frac{1}{\pi}\left\{\arctan\left(\frac{1-2r_1}{2(r_1^2+r_2^2-r_1)}\right) + \arctan\left(\frac{1+2r_1}{2(r_1^2+r_2^2+r_1)}\right)\right\}. \tag{A.16}$$

Using the identity in equation (A.15) again produces the required result,

$$a_m(\boldsymbol{\theta}) = \frac{1}{\pi}\arctan\left(\frac{r_2^2-r_1^2}{(r_1^2+r_2^2)^2-1/4}\right). \tag{A.17}$$

A.3.2 Analytical form for b_m

To calculate the relationship for b_m, one needs to evaluate the following:

$$\frac{1}{\pi}\left(\frac{\partial^2}{\partial\theta_1\partial\theta_2}\right)\int_m \mathrm{d}^2\theta' \ln|\boldsymbol{\theta}-\boldsymbol{\theta}'|. \tag{A.18}$$

As in the derivation for a_m, simplification is achieved by defining $x \equiv (\theta_1 - \theta_1')$ and $y \equiv (\theta_2 - \theta_2')$. Equation (A.18) may then be re-written as,

$$\begin{aligned} b_m(\boldsymbol{\theta}) &= \frac{1}{2\pi}\left(\frac{\partial^2}{\partial x\partial y}\right)\int_{y=r_2-0.5}^{y=r_2+0.5}\int_{x=r_1-0.5}^{x=r_1+0.5} \ln(x^2+y^2)\,\mathrm{d}x\,\mathrm{d}y \\ &= \frac{1}{2\pi}\left[\ln(x^2+y^2)\right]_{y=r_2-0.5,\ x=r_1-0.5}^{y=r_2+0.5,\ x=r_1+0.5} \\ &= \frac{1}{2\pi}\ln\left\{\frac{[(r_1-0.5)^2+(r_2-0.5)^2][(r_1-0.5)^2+(r_2-0.5)^2]}{[(r_1-0.5)^2+(r_2+0.5)^2][(r_1+0.5)^2+(r_2-0.5)^2]}\right\}. \end{aligned} \tag{A.19}$$

Again, as in the derivation for a_m, it has been assumed that pixels are square and the units of length are chosen such that pixels measure one unit on a side. Further simplification gives the final result,

$$b_m(\boldsymbol{\theta}) = \frac{1}{2\pi}\ln\left[\frac{(r_1^2+r_2^2+1/2)^2-(r_1+r_2)^2}{(r_1^2+r_2^2+1/2)^2-(r_1-r_2)^2}\right] \tag{A.20}$$

REFERENCES

Abell G.O., 1958, ApJS, 3, 211

AbdelSalam H.M., Saha P., Williams L.L.R., 1998, MNRAS, 294, 734

Allen S.W., 1998, MNRAS, 296, 392

Alpher R.A., Bethe H., Gamow G., 1948, Phys. Rev., 73, 803

Bahcall J.M., Lubin L.M., 1994, ApJ, 426, 515

Bahcall J.M., Tremaine S., 1981, ApJ, 244, 805

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

Bartelmann M., 1996, PhD thesis, Ludwig-Maximilians-Universität München

Bernardeau F., 1998, A&A, 338, 767

Bernardeau F., van Waerbeke L., Mellier Y., 1997, A&A, 322, 1

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

Beers T.C., Flynn K., Gebhardt K., 1990, AJ, 100, 32

Binney J.J., Tremaine S., 1987, 'Galactic Dynamics', Princeton University Press

Bird C.M., 1993, PhD thesis, University of Minnesota and Michigan State University

Bird C.M., 1994, AJ, 107, 1637

Blandford R.D., 1990, Quarterly Journal of the RAS, 31, 305

Blandford R.D., Jaroszynski M., 1981, ApJ, 246, 1

Bondi H., 1947, MNRAS, 107, 410

Bonnet H., Mellier Y., 1995, A&A, 303, 331

Borgani S., Lucchin F., Matarrese S., Moscardini L., 1996, MNRAS, 280, 749

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

Broadhurst T.J., 1995, Proc. of 5th Maryland meeting 'Dark Matter' AIP 336 College Park, eds. S. Holt, C.L. Bennett

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

Cavaliere A., Fusco-Femiano R., 1976, A&A, 49, 137

Coles P., Ellis G.F.R., 1997, 'Is the Universe Open or Closed ?', Cambridge University Press

Coles P., Jones B.J.T., 1991, MNRAS, 248, 1

Coles P., Lucchin F., 1997, 'Cosmology: The Origin and Evolution of Cosmic Structure', J. Wiley & Sons

Crampton D., Le Févre O., Lilly S.J., Hammer F., 1995, ApJ, 455, 96

Daines S., Jones C., Forman W., Tyson J.A., 1998, ApJ, submitted

Dye S., Taylor A.N., 1998, MNRAS, 300, L23

Dye S., Taylor A.N., Thommes E., Meisenheimer K., Wolf C., Peacock J.A., 1999, submitted to MNRAS

Eddington A.S., 1919, Observatory, 42, 119

Edge A.C., Stewart G.C., 1991, MNRAS, 252, 428

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

Ellis R.S., Colless M., Broadhurst T., Heyl J., Glazebrook K., 1996, MNRAS, 280, 235

Ensslin T.A., Ellis R., Kneib J.P., La Borgne J.F., Pelló R., Smail I., Sanahuja B., 1998, MNRAS, 295, 75

Evrard A.E., Metzler C.A., Navarro J.F., 1996, ApJ, 469, 494

Falco E.E., Gorenstein M.V., Shapiro I.I., 1985, ApJ, 437, 56

Fan X., Bahcall N.A., Cen R., 1997, ApJ, 490, L123

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

Fort B., Mellier Y., 1994, A&A Review, 5, 239

Freedman W.L., Mould J.R., Kennicutt R.C., Madore B.F., 1997, IAUS 183E, 14 'Cosmological Parameters and the Evolution of the Universe'

Giradi M., Fadda D., Escalera E., Giuricin G., Mardirossian F., Mezzetti M., 1997, ApJ, 490, 56

Goodbody A.M.,1982, 'Cartesian Tensors', Ellis-Harwood

Gudehus D.H., Hegyi D.J., 1991, AJ, 101, 18

Guth A.H., 1981, Phys. Rev. D., 23, 347

den Hartog R., Katgert P., 1996, MNRAS, 279, 349

Heisler J., Tremaine S., Bahcall J.N., 1985, ApJ, 298, 8

Jones C., Forman W., 1984, ApJ, 276, 38

Kaiser N., 1995, ApJ, 439, L1

Kaiser N., Wilson G., Luppino G., Kofman L., Gioia I., Metzger M., Dahle H., 1998, astro-ph/9809268

Kaiser N., 1996, in 'Gravitational Dynamics', Proc. of the 36th Herstmonceux Conf., eds O. Lahav, E. Terlevich, R.J. Terlevich (Cambridge Uni Press)

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

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

van Kampen E., 1998, MNRAS, 301, 389

van Kampen E., Katgert P., 1997, MNRAS, 209, 327

King I.R., 1966, AJ, 71, 64

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

Kofman L., Klypin A., Pogosyan D., Henry J.P., 1996, ApJ, 470, 102

Kolb E.W., Turner M.S., 1990, 'The Early Universe', Addison Wesley

Landy S.D., Schectman S.A., Lin H., Kirshner R., Oemler A., Tucker D., 1996, ApJ, 456, L1

Le Févre O., Hudon D., Lilly S.J., Crampton D., Hammer F., Tresse L., 1996, ApJ, 461, 534

Lewis A.D., Ellingson E., Morris S., Carlberg R.G., 1999, ApJ, 517, 587

Lilly S.J., Cowie L.L., Gardner J.P., 1991, ApJ, 369, 79

Lilly S.J., Tresse L., Hammer F., Crampton D., Le Févre O., 1995 ApJ, 455, 108

Limber D.N., Mathews W.G., 1960, ApJ, 132, 286

Lin H., Yee H.K.C., Carlberg R.G., Ellingson E., 1997, ApJ, 475, 494

Loeb A., Mao S., 1994, ApJ, 435, L109

Longair M.S., 1984, 'Theoretical Concepts in Physics', Cambridge University Press

Loveday J., Peterson B.A., Efstathiou G., Maddox S.J., 1992, ApJ, 390, 338

Lubin L.M., Bahcall N.A., 1993, ApJ, 415, L17

Maddox S.J., Efstathiou G., Sutherland W.J., Loveday J., 1990, MNRAS, 243, 692

Mann R.G., Peacock J.A., Heavens A.F., 1998, MNRAS, 293, 209

Mattig W., 1958, Astron. Nach., 284, 109

Meisenheimer K., Röser H.J., 1996, MPIAPHOT User Manual, MPIA Heidelberg

Metcalf N., Shanks T., Fong R., Roche N., 1995, MNRAS, 273, 257

Miralda-Escudé J., 1991, ApJ, 380, 1

Miralda-Escudé J., Babul A., 1995, ApJ, 449, 18

Mushotzky R.F., Scharf C.A., 1997, ApJ, 482, L13

Narayan R., 1998, New Astron. Rev., 42, 73

Oke J.B., 1990, AJ, 99, 1621

Peacock J.A., 1999, 'Cosmological Physics', Cambridge University Press

Peacock J.A., Dodds S.J., 1994, MNRAS, 267, 1020

Pen U.L., 1998, ApJ, 498, 60

Pickles A.J., van der Kruit P.C., 1991, A & A(S), 91, 1

Press W.H., Schechter P.L., 1974, ApJ, 187, 425

Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P., 1992, 'Numerical Recipes', 2nd Edition, Cambridge University Press

Perlmutter S. et al, 1998, Nature, **391**, 51

Reblinsky K., Bartelmann M., 1999, A&A, 345, 1

Richstone D., Loeb A., Turner E.L., 1992, ApJ, 393, 477

Riess A.G., et al, 1998, astro-ph/9805201, submitted to ApJ

Rubin V.C., Ford W.K., 1970, ApJ, 159, 379

Rubin V.C., Ford W.K., Thonnard N., Burstein D., 1982, ApJ, 261, 439

Sandage A., Tammann G.A., Yahil A., 1979, ApJ, 232, 352

Sarazin C.L., 1986, Rev. Mod. Phys., 58, 1

Schechter P., 1976, ApJ, 203, 297

Schmidt M., 1968, ApJ, 151, 393

Schneider P., Ehlers J., Falco E.E., 1992, 'Gravitational Lenses', Springer-Verlag

Seitz C., Schneider P., 1995, A&A, 297, 287

Smail I., Hogg D.W., Yan L., Cohen J.G., 1995, ApJ, 449, L105

Squires G., Kaiser N., 1996, ApJ, 473, 85

Tadros H., Ballinger W.E., Taylor A.N., Heavens A.F., Efstathiou G., Saunders W., Frenk C.S., Keeble O., Mcmahon R., Maddox S.J., Oliver S., Rowan-Robinson M., Sutherland W.J., White S.D.M., 1999, MNRAS, 305, 527

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

Teague P.F., Carter D., Gray P.M., 1990, ApJS, 72, 715

Thommes E.M., 1996, Doctoral Thesis, MPIA Heidelberg

Thommes E.M. et al, 1999, 'CADIS Revised Proposal', MPIA Publication

Tremaine S., Lee H.M., 1987, 'Dark Matter in The Universe', eds. J.Bachall, T.Piran, S.Weinberg

Tyson J.A., Fischer P., 1995, ApJ, 446, L55

Valdes F., Tyson J.A., Jarvis J., 1983, ApJ, 271, 431

Viana P.T.P, Liddle A.R., 1996, MNRAS, 281, 323

Wallington S., Narayan R., Kochanek S., 1994, ApJ, 426, 60

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

Wang Q.D., Ulmer M.P., 1997, MNRAS, 292, 920

Watkins R., 1997, MNRAS, 292, 59

Weinberg S., 1972, 'Graviation & Cosmology: Principles and Applications of the General Theory of Relativity', Wiley

White S.D.M., Efstathiou G., Frenk C.S., 1993, MNRAS, 262, 1023

Wolf C., 1999, Doctoral Thesis, MPIA Heidelberg

Wu X., Fang L., 1997, ApJ, 483, 62

Yamashita, 1994, in Recontre de Moriond on 'Clusters of Galaxies', eds F. Durret, A. Mazure, J. Tran Thanh Van

Zwicky F., 1933, Helv. Phys. Acta., 6, 110

Zwicky F., Herzog E., Wild P., Karpowicz M., Cowal C., 1961-1968, 'Catalogue of Galaxies and Clusters', Vol 1-6, (Caltech, Pasadena)

Mon. Not. R. Astron. Soc. **300,** L23–L28 (1998)

Self-consistent gravitational lens reconstruction

Simon Dye* and Andy Taylor*

Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ

Accepted 1998 August 14. Received 1998 August 14; in original form 1998 May 15

ABSTRACT

We present a new method for directly determining accurate, self-consistent cluster lens mass and shear maps in the strong lensing regime from the magnification bias of background galaxies. The method relies upon pixellization of the surface mass density distribution which allows us to write down a simple, solvable set of equations. We also show how pixellization can be applied to methods of mass determination from measurements of shear and present a simplified method of application. The method is demonstrated with cluster models and applied to magnification data from the lensing cluster Abell 1689.

Key words: galaxies: clusters: general – cosmology: theory – gravitational lensing – large-scale structure of Universe.

1 INTRODUCTION

The possibility of reconstructing cluster lens mass distributions from the magnification bias of background galaxies was first suggested by Broadhurst, Taylor & Peacock (1995) and first demonstrated by Taylor et al. (1998, T98 hereafter). They showed how a direct, local measure of the lens convergence, $\kappa = \Sigma/\Sigma_c$, where Σ is the mass surface density and Σ_c is the critical surface density, could be obtained from knowledge of the lens magnification. In this way, one could measure absolute surface mass densities, thereby breaking the 'sheet-mass' degeneracy found in methods based on distortions of background galaxies (Tyson, Valdes & Wenk 1990; Kaiser & Squires 1993; Seitz & Schneider 1995).

Van Kampen (1998) and T98 have shown how one can extend magnification analysis into the strong lensing regime. By making reasonable assumptions about γ, the lens shear, they showed that one could place quite stringent bounds on κ. In addition, T98 found an exact solution for the profile of axisymmetric lenses, although not for more general 2D cases.

Inverse reconstruction methods based on maximum likelihood (Bartelmann et al. 1996) and maximum entropy (Seitz, Schneider & Bartelmann 1998; Bridle et al. 1998) have gone some way towards providing a unification of both shear and magnification information. Until now, however, no direct method using only magnification has existed.

In this letter, we show how to directly compute an accurate, self-consistent 2D distribution of κ and γ in the strong lensing regime from magnification. This direct approach has the advantage over indirect alternatives that uncertainties can easily be determined and the application is much quicker. The method is based on pixellization of the κ distribution, suggested by AbdelSalam, Saha & Williams (1998), who used it to estimate the mass of Abell 370 from multiple images. We generalize the method further and also derive a simplified solution to the problem of estimating mass from shear, based on the approach of Kaiser & Squires (1993).

*E-mail: sd@roe.ac.uk (SD); ant@roe.ac.uk (AT)

2 RECONSTRUCTION OF κ AND γ

T98 showed how to estimate cluster surface mass using the magnification measured from the distortion in background galaxy number counts. Here our problem is to find an accurate method for reconstructing the surface mass density, given the magnification by an arbitrary lens. The inverse magnification factor at a given position in the lens plane is

$$A^{-1} = |(1-\kappa)^2 - \gamma^2|, \tag{1}$$

where κ is the lens convergence and γ is the shear. The shear can be decomposed into two orthogonal polarization states, γ_1 and γ_2, which are related to the lens convergence by

$$\gamma_1 = \tfrac{1}{2}\partial^{-2}(\partial_1^2 - \partial_2^2)\kappa, \quad \gamma_2 = \partial^{-2}\partial_1\partial_2\kappa. \tag{2}$$

where $\partial_i \equiv \partial/\partial\theta_i$ and ∂^{-2} is the 2D inverse Laplacian. The total shear is given by $\gamma^2 = \gamma_1^2 + \gamma_2^2$. One might expect that equation (1) could be solved iteratively by first estimating κ, using this to calculate γ and then updating the estimate of κ using equation (1) again. This proves to be highly unstable in the strong lensing regime, however, rapidly diverging after only a few iterations (Seitz & Schneider 1995).

To find a stable solution to equation (1), we first pixellize the image. Following AbdelSalam et al. (1998), we can now write

$$\gamma_i^n = D_i^{nm}\kappa_m, \quad i = 1, 2 \tag{3}$$

with summation implied over index m and where κ_m and γ_i^n are the pixellized convergence and shear distributions respectively. The

transformation matrices, D_i^{mn}, are

$$D_1^{mn} = \frac{1}{2}(\partial_1^2 - \partial_2^2)\int_m \mathrm{d}^2\theta' \ln|\theta_n - \theta'| = \frac{1}{\pi}\tan^{-1}\left[\frac{x_1^2 - x_2^2}{(x_1^2 + x_2^2)^2 - 1/4}\right], \quad (4)$$

and

$$D_2^{mn} = \partial_1\partial_2\int_m \mathrm{d}^2\theta' \ln|\theta_n - \theta'| = \frac{1}{2\pi}\ln\left[\frac{(x_1^2 + x_2^2)^2 - 2x_1x_2 + 1/4}{(1/2 + x_1^2 + x_2^2)^2 - (x_1 - x_2)^2}\right], \quad (5)$$

with the integration acting over the m^{th} pixel. $x = \theta_n - \theta_m$ is the difference between pixels m and n which are assumed to be square in calculating these analytic expressions. Equation (1) can now be written as the vector equation,

$$\mathbf{1} - 2\boldsymbol{\kappa} + \boldsymbol{\kappa}\mathbf{G}\boldsymbol{\kappa}^{\mathrm{t}} - \mathcal{P}A^{-1} = 0 \quad (6)$$

where A^{-1} is the N-dimensional vector of pixellized inverse magnification values, $\boldsymbol{\kappa}^{\mathrm{t}}$ is the transpose of the vector $\boldsymbol{\kappa}$ of pixellized convergence values and $\mathbf{1}$ is the vector $(1, 1, 1, \ldots)$. The matrix $\mathbf{G}$ is the $N \times N \times N$ matrix

$$G_{pqn} = \delta_{pm}\delta_{qn} - D_1^{pm}D_1^{qn} - D_2^{pm}D_2^{qn}, \quad (7)$$

where δ_{ij} is the Krönecker delta, and summation is only over indices p and q. The parity of the measured inverse amplification $A^{-1}(\theta)$ is handled by $\mathcal{P}$ which flips from being $+1$ outside regions bounded by critical lines to -1 within such regions.

The amplification equation in the form of equation (6) is the first main result of this letter. We can now solve for $\boldsymbol{\kappa}$ numerically (see Section 4) given a measured inverse amplification. Having solved for $\boldsymbol{\kappa}$, the corresponding shear distribution can then be calculated from equation (3).

3 APPLICATION TO CLUSTER MODELS

We apply the method to two types of idealized cluster models. Starting with a predetermined cluster mass density distribution, the corresponding shear distribution is derived using Fourier methods (see, for example, Bartelmann & Weiss 1994). From these, the resulting magnification is calculated from equation (1) and then windowed to remove boundary effects. Using equation (6), we solve for κ. γ is then solved using equation (3). A grid of 32 by 32 pixels is used in both models.

3.1 Truncated isothermal sphere model

We first test the method with a simple truncated isothermal lens model. The pixellated mass distribution is laid down using $\kappa \propto (r + r_0)^{-1}$, where r is the radial distance from the centre of the sphere and r_0 is a constant.

Fig. 1 shows the κ and γ distribution from which the magnification distribution was calculated, the solved κ and γ distribution, and the difference between them. The plotted distributions are smoothed from the underlying grid and the white dashes highlight the critical line of the lens. The residuals are shown as percentage deviations from the true distribution. These are less than one per cent for κ over most of the grid, which is negligible in comparison to the errors typically found in practice from background clustering, shot noise (T98) and the uncertainties resulting from the use of local κ estimators (see van Kampen 1998). The recovered shear distribution is more affected, although it still fares better than γ calculated from uncorrected Fourier techniques. The main contribution to these residuals is from boundary effects arising from trying to recover a non-local shear in a finite area. Since much work has been carried out in the removal of such effects (see Squires & Kaiser 1996 and Seitz & Schneider 1995, for example) which have little impact on the recovered κ, we shall address the problem elsewhere.

3.2 Dumb-bell mass model

The method was also tested with a more general dumb-bell model. Magnification was determined in the same fashion as for the isothermal model, setting a negative parity inside the critical lines, shown by the white dashes in Fig. 2. Once again the residuals between the initial and solved κ are typically less than one per cent, while those for γ are typically 10 per cent and again come mainly from boundary effects.

4 PRACTICAL CONSIDERATIONS

We solve equation (6) with the hybrid Powell method (NAG routine C05PCF). The number of equations needed to solve for κ is equal to the total number of grid pixels, which can prove computationally intensive for especially fine grids. We find that this is not a problem for grid resolutions used to measure magnification bias in practice. The 32×32 grid of pixels used for the models in Section 3 was solved in approximately one minute on an average workstation. The residuals exhibit no noticeable dependence on grid size.

The Powell algorithm is an iterative process and therefore requires an initial estimate of the solution to start from. The choice of the initial estimate turns out to be irrelevant. We have tried a wide range of initial distributions, and even starting from a uniform distribution we arrived at the same final solution.

We have found, however, that the correct choice of pixel parity (especially for low grid resolutions) is essential in order to achieve a sensible result. Inappropriate assignment of parities to pixels manifest themselves, as one would expect, by κ being overestimated when a pixel is wrongly assumed to lie inside a critical line, and underestimated in the reverse situation. This provides a means of checking whether critical line positions have been properly defined by looking for large discontinuities in the κ distribution. Models with dual critical lines requiring dual parity flips have also been tested and we find that κ can be recovered just as well.

Finally, to ensure that the method does not break down with noisy data, we introduced a random noise term to the amplification. Errors in κ resulting from noise in the inverse amplification propagate as one would expect from equation (6). For an isothermal lens we recovered the expected result, $\delta\kappa = \delta A/2A^2$, indicating that pixellization does not lead to spurious noise properties.

5 APPLICATION TO ABELL 1689

We apply the method to the magnification data presented in T98 for the lensing cluster A1689. A 12×12 grid is used as the best compromise between shot noise in galaxy counts per bin and the resolution of the derived κ map. Identification of the critical line was achieved by locating giant arc positions in the observed image.

Fig. 3 shows the solved mass density and shear distribution. Comparison with the mass density map illustrated in T98 (their fig. 6), which was produced with the sheet κ estimator, shows very similar structure. We find that the value of κ at the peak calculated here is approximately 10 per cent lower than the peak value in T98, since the sheet estimator over-estimates κ inside critical line regions. This has little effect on the total integrated mass of

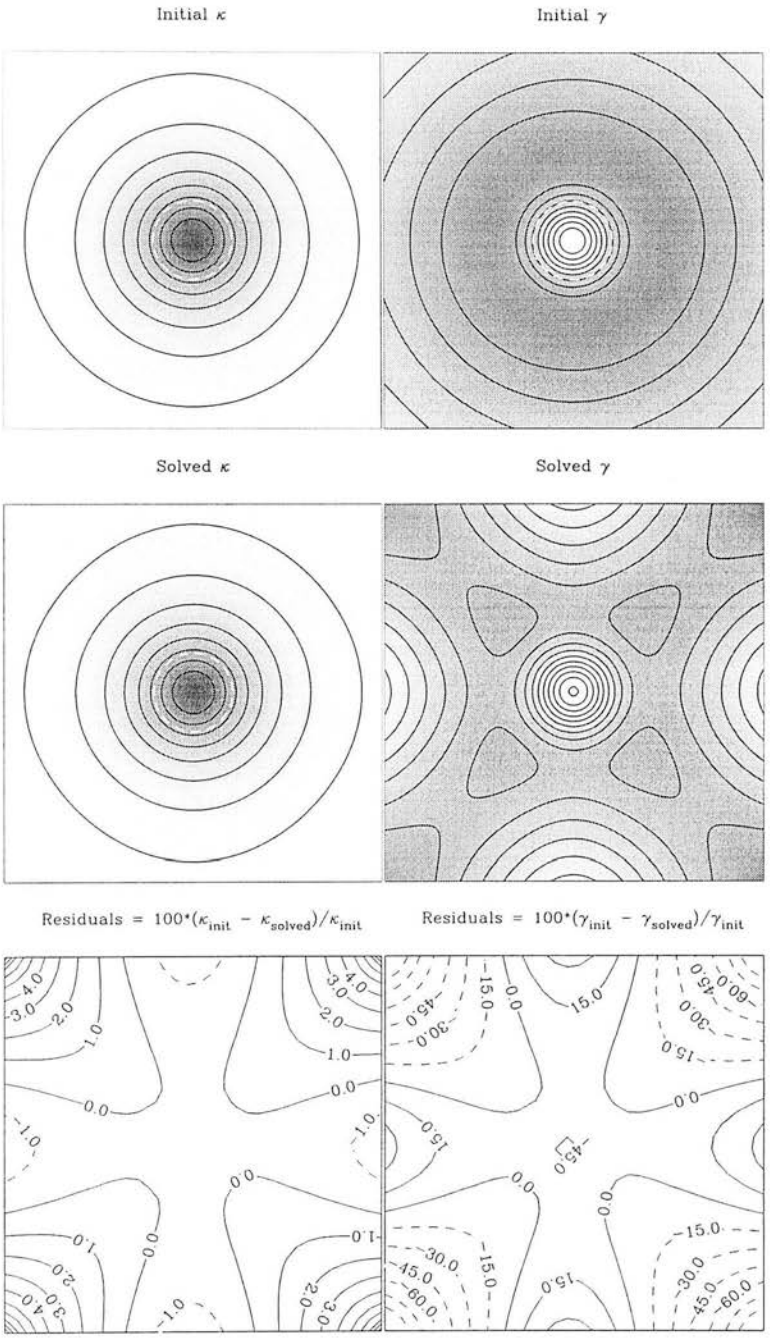

Figure 1. Truncated isothermal sphere model. The initial κ and γ used to form the magnification distribution from which the solved κ and γ are derived. Underlying grid dimensions are 32×32. White dashes show the position of the critical line. Contours are linearly spaced and are set at the same levels in both κ plots and in both γ plots. Residuals are expressed as percentages of $(\kappa_{init} - \kappa_{solved})/\kappa_{init}$.

A1689 found in T98. The γ distribution is shown for completeness, although it undoubtedly suffers from boundary effects typically found in the models.

6 SHEAR ANALYSIS

Having shown that pixellization allows us to accurately reconstruct surface mass densities from magnification data, we now apply it to shear analysis. Shear analysis exploits the idea that a given distribution of images of galaxies lying behind a lensing cluster will, in the statistical mean, have regions of lens-induced correlations in image orientation and ellipticity. Measuring the quadrupole moments of individual galaxy images enables the construction of a map of the ellipticity parameters, e_{ij} (Valdes, Tyson & Jarvis 1983). The ellipticity parameters relate to the surface mass density and shear via

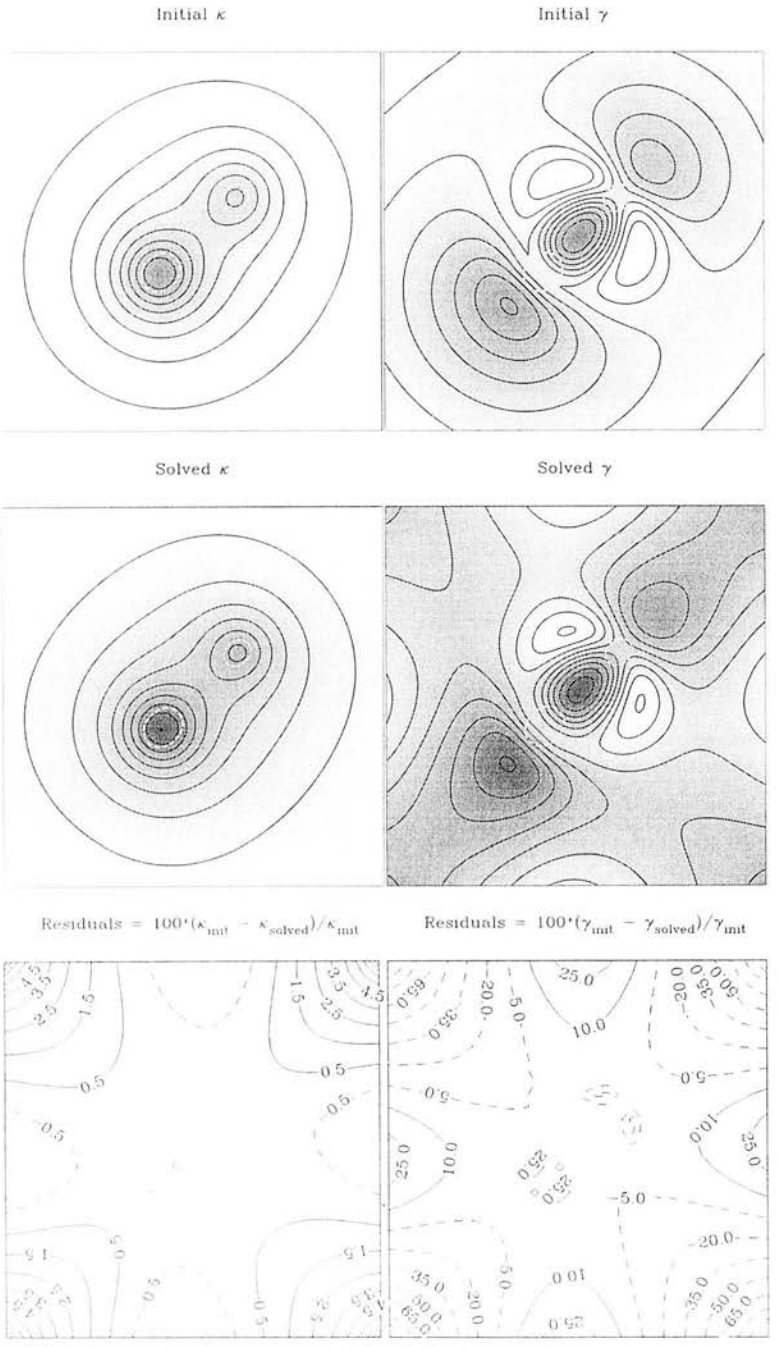

Figure 2. Dumb-bell model: critical lines are shown as white dashes. Underlying grid dimensions are 32×32. Linearly spaced contours are set at the same levels for κ and at the same levels for γ.

(Kaiser 1995, hereafter K95):

$$c_{ij} = \frac{\gamma_{ij}}{1-\kappa}, \quad \gamma_{ij} = \begin{pmatrix} \gamma_1 & \gamma_2 \\ \gamma_2 & -\gamma_1 \end{pmatrix}. \tag{8}$$

One way of solving this for κ in the weak lensing regime is to follow the approach of Kaiser & Squires (1993). Generalizations of this to the strong regime have been made by K95. One would have hoped that an alternative to such approaches would be to pixellize equation (8) and use equation (3) to solve it by matrix inversion. However, the resulting matrix equation is ill-conditioned, since the matrix D_1^{mn} is singular and D_2^{mn} is itself ill-conditioned. Instead, we show a new, simplified expression for the solution to Kaiser's ellipticity equation and then pixellize it.

Starting with the equation (K95),

$$\partial_i \kappa = \partial_j \gamma_{ij} \tag{9}$$

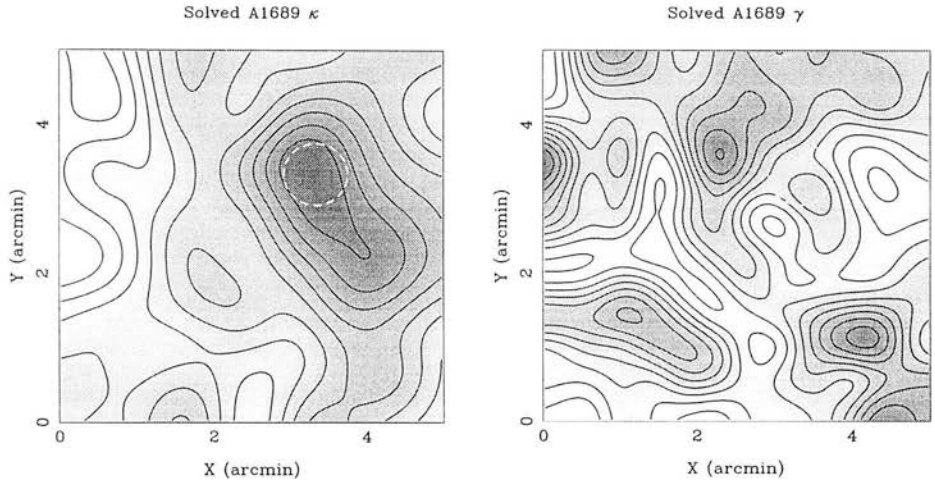

Figure 3. Abell 1689 solved convergence and shear distributions. Darker areas represent a higher distribution density. White dashes show the observed critical line. The plots are smoothed from a 12×12 grid with north up and east to the left.

and using equation (8), one can show that

$$\partial_i \ln(1-\kappa) = -\partial_j \ln(\delta_{ij} + e_{ij}). \tag{10}$$

The term on the right hand side is obtained from the definition,

$$\ln(\mathbf{I}+\mathbf{B}) = \mathbf{B} + \tfrac{1}{2}\mathbf{B}^2 + \tfrac{1}{3}\mathbf{B}^3 + \ldots \tag{11}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{B}$ is an arbitrary square matrix. Using this expansion and collecting even and odd terms

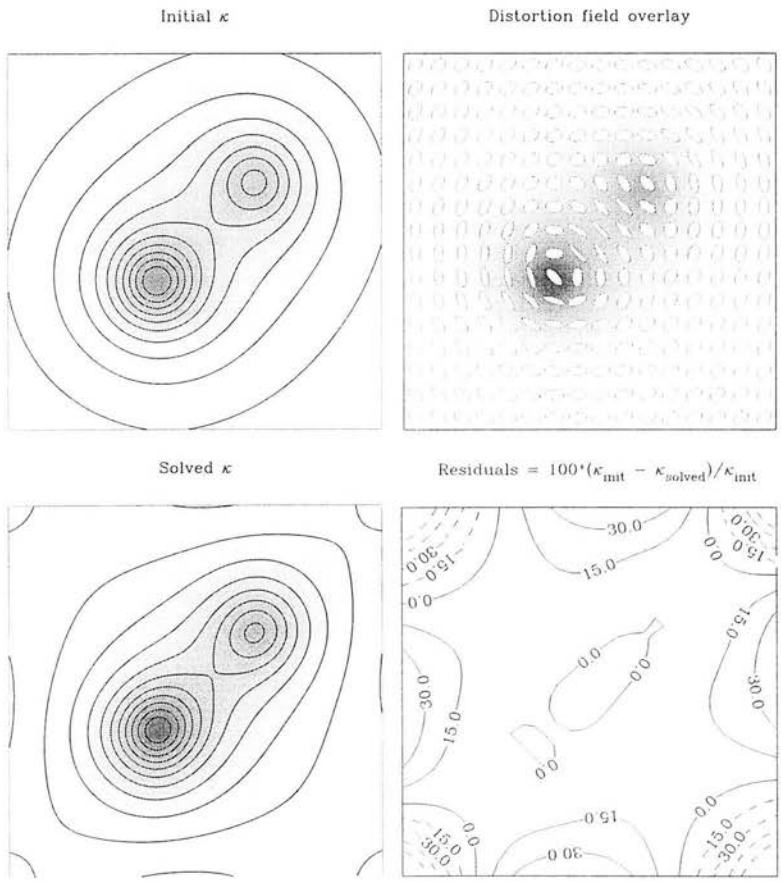

Figure 4. Reconstruction of κ from the ellipticity parameters. Contours are at the same levels in both κ plots. The distortion field is illustrated by plotting the apparent shape of an intrinsically circular background object.

we find

$$\ln(\delta_{ij} + e_{ij}) = -\frac{1}{2}\ln(1 - e^2)\delta_{ij} + \frac{1}{2}\ln\left(\frac{1+e}{1-e}\right)\frac{e_{ij}}{e}, \quad (12)$$

where $e^2 = e_1^2 + e_2^2$ and $e_i = \gamma_i/(1 - \kappa)$. This result requires that $e < 1$. Inserting equation (8) into the magnification equation (1) we find

$$A^{-1} = |(1 - \kappa)^2(1 - e^2)|. \quad (13)$$

Hence the parity changes when $e > 1$. Since e_{ij} and e_{ij}^{-1} are observationally indistinguishable and flip from one to another whenever there is a parity change, we can satisfy the criterion $e < 1$ just by noting the critical line positions and inverting the ellipticity matrix when one is crossed.

Finally, inserting equation (12) into equation (10), and solving for κ we find the pixellized solution is

$$\kappa_n = 1 - (1 - e_n^2)^{1/2} \exp\left[-\frac{1}{2}(D_1^{nm} s_1^m + D_2^{nm} s_2^m)\right] \quad (14)$$

where

$$s_i = \frac{e_i}{e}\ln\left(\frac{1+e}{1-e}\right), \quad i = 1, 2. \quad (15)$$

Equation (14) is the second main result of this letter. We can directly calculate κ given a measured ellipticity field. Fig. 4 shows the results of reconstructing κ using equation (14) for the dumb-bell model. The ellipticity parameters are calculated from equation (8) using the κ and γ distribution. We normalize the reconstructed κ to both peaks in the initial κ distribution. The residuals, again being dominated by boundary effects, show that reconstruction is possible to within approximately 10 per cent across the field of view. This can be alleviated by a larger field of view.

7 SUMMARY

We have outlined a method for directly calculating accurate, self-consistent surface mass density and shear distributions from the lens amplification and critical line positions. The method has been demonstrated with the isothermal sphere and dumb-bell cluster models. We find that it reconstructs the surface density to within one per cent over most of the field of view. The reconstruction of the shear pattern only has a fractional accuracy of a few tenths because of boundary effects. We have applied the method to magnification data from Abell 1689, and have reconstructed its surface mass and shear distribution.

We have also found a simplified solution to the problem of estimating surface mass density from galaxy ellipticities. This approach puts the calculation of surface mass from shear and from magnification on an equal footing, and we shall investigate the combined analysis elsewhere.

ACKNOWLEDGMENTS

SD is supported by a PPARC studentship. ANT is a PPARC research associate. Thanks to our collaborators Tom Broadhurst, Txitxo Benitez and Eelco van Kampen for allowing us to use data from Abell 1689. We also thank Hanadi AbdelSalam for bringing the benefits of pixellization to our attention and for useful discussion.

REFERENCES

AbdelSalam H. M., Saha P., Williams L. L. R., 1998, MNRAS, 294, 734
Bartelmann M., Narayan R., Seitz S., Schneider P., 1996, ApJ, 464, 115
Bartelmann M., Weiss A., 1994, A&A, 287, 1
Bridle S. L., Hobson M. P., Lasenby A. N., Saunders R., 1998, MNRAS, in press (astro-ph/9802159)
Broadhurst T. J., Taylor A. N., Peacock J. A., 1995, ApJ, 438, 49
Kaiser N., 1995, ApJ, 439, L1 (K95)
Kaiser N., Squires G., 1993, ApJ, 404, 441
Seitz C., Schneider P., 1995, A&A, 297, 287
Seitz C., Schneider P., Bartelmann M., 1998, preprint (astro-ph/9803038)
Squires G., Kaiser N., 1996, ApJ, 473, 85
Taylor A. N., Dye S., Broadhurst T. J., Benitez N., van Kampen E., 1998, ApJ, 501, 539 (T98)
Tyson J. A., Valdes F., Wenk R. A., 1990, ApJ, 349, L1
Valdes F., Tyson J., Jarvis J., 1983, ApJ, 271, 431
van Kampen E., 1998, MNRAS, in press

This paper has been typeset from a TEX/LATEX file prepared by the author.

THE ASTROPHYSICAL JOURNAL, 501:539–553, 1998 July 10

GRAVITATIONAL LENS MAGNIFICATION AND THE MASS OF ABELL 1689

A. N. TAYLOR AND S. DYE
Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK; ant@roe.ac.uk, sd@roe.ac.uk
T. J. BROADHURST AND N. BENÍTEZ
Department of Astronomy, Campbell Hall, University of California, Berkeley, CA; tjb@wibble.berkeley.edu, benitezn@wibble.berkeley.edu
AND
E. VAN KAMPEN
Royal Observatory Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, and Theoretical Astrophysics Center, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark; eelco@tac.dk
Received 1997 June 27; accepted 1998 February 17

ABSTRACT

We present the first application of lens magnification to measure the absolute mass of a galaxy cluster: Abell 1689. The absolute mass of a galaxy cluster can be measured by the gravitational lens magnification of a background galaxy population by the cluster gravitational potential. The lensing signal is complicated by the intrinsic variation in number counts resulting from galaxy clustering and shot noise and by additional uncertainties in relating magnification to mass in the strong lensing regime. Clustering and shot noise can be dealt with using maximum likelihood methods. Local approximations can then be used to estimate the mass from magnification. Alternatively, if the lens is axially symmetric we show that the amplification equation can be solved nonlocally for the surface mass density and the tangential shear. In this paper we present the first maps of the total mass distribution in Abell 1689, measured from the deficit of lensed red galaxies behind the cluster. Although noisier, these reproduce the main features of mass maps made using the shear distortion of background galaxies, but have the correct normalization, finally breaking the "sheet-mass" degeneracy that has plagued lensing methods based on shear. Averaging over annular bins centered on the peak of the light distribution, we derive the cluster mass profile in the inner 4′ (0.48 h^{-1} Mpc). These show a profile with a near-isothermal surface mass density $\kappa \approx (0.5 \pm 0.1)(\theta/1')^{-1}$ out to a radius of 2.′4 (0.28 h^{-1} Mpc), followed by a sudden drop into noise. We find that the projected mass interior to 0.24 h^{-1} Mpc is $M(<0.24\ h^{-1}\ \mathrm{Mpc}) = (0.50 \pm 0.09) \times 10^{15}\ h^{-1}\ M_{\odot}$. We compare our results to masses estimated from X-ray temperatures and line-of-sight velocity dispersions, as well as to weak shear and lensing arclets. We find that the masses inferred from X-ray, line-of-sight velocity dispersions, arclets, and weak shear are all in fair agreement for Abell 1689.

Subject headings: galaxies: clusters: individual (Abell 1689) — gravitational lensing

1. INTRODUCTION

The magnitude and distribution of matter in galaxy clusters should in principle provide a strong constraint on cosmological models of structure formation and the mean mass density of the universe. In addition, a direct image of the mass density will tell us much about the relationship between gas, galaxies, and dark matter, and whether light is indeed a fair—if biased—tracer of mass.

Early techniques for estimating the mass in clusters include dynamical methods, from the line-of-sight velocity dispersion of member galaxies, and X-ray temperature measurements. However, these estimates make some strong assumptions about equilibrium conditions in the cluster.

Kaiser & Squires (1993) circumvented this problem by showing that a more direct method of estimating the mass, with no underlying assumptions about the dynamical or thermodynamical state of the cluster, was to measure the shear field in the source distribution of the cluster background (Kaiser & Squires 1993; Tyson, Valdes, & Wenk 1990; Schneider & Seitz 1995). On average, the shear pattern of a population of unlensed galaxies should be randomly distributed. But in the presence of a massive gravitational lensing cluster, the shear field is polarized. Since the shear field is related (nonlocally) to the surface mass density, the shear can be used to estimate the mass distribution—up to an arbitrary constant. The presence of this arbitrary constant, referred to as the "sheet-mass" degeneracy, means that only differential masses can be measured. Shear maps are conventionally normalized to the edge of the observed field, or such that the inferred mass density is everywhere positive, and so represent a lower limit on the mass.

Soon after, Broadhurst, Taylor, & Peacock (1995, hereafter BTP) showed that the sheet-mass degeneracy could be broken by use of the gravitational lens magnification effect. The number and magnitude-redshift distribution of background galaxies is distorted by the gravitational field of the lensing cluster, and in the weak lensing regime this distortion provides a straightforward estimate of the surface mass density. With calibration from offset fields the cluster mass distribution can be properly normalized.

BTP also suggested that a degraded, but much quicker, estimate of the magnification effect could be made from the distortion of angular number counts of background sources. Broadhurst (1995) found evidence for this distortion in the background counts of the cluster Abell 1689, as did Fort, Mellier, & Dantel-Fort (1997) for Cl0024. In this work we apply the methods developed by BTP and extended by Taylor & Dye (1998) in estimating the surface mass density from the distortion of angular counts, including the effects of shot noise and galaxy clustering, and those of van Kampen (1998) in estimating the surface mass density in the strong lensing regime to Abell 1689.

The layout of the paper is as follows. In § 2 we describe the magnification effect itself. In § 3 we describe the effects of shot noise and clustering on estimates of the surface mass density. In § 4 we describe how to estimate the surface mass density in the strong lensing regime using local approximations and introduce a new self-consistent nonlocal solution for axially symmetric lenses. We apply these methods to map out the mass in the cluster Abell 1689 in § 5 and find its profile. Our mass estimate is compared to other estimates in § 6, and our conclusions are presented in § 7.

2. THE MAGNIFICATION EFFECT

The observed number of galaxies seen in projection on the sky is (BTP; Taylor & Dye 1998)

$$n' = n_0 A^{\beta-1}(1+\Theta) , \tag{1}$$

where n_0 is the expected mean number of galaxies in a given area at a given magnitude. Variations in this mean arise from the angular perturbation in galaxy density Θ as a result of galaxy clustering and from gravitational lens magnification. The lens amplification factor is

$$A = |(1-\kappa)^2 - \gamma^2|^{-1} , \tag{2}$$

where

$$\kappa = \frac{\Sigma}{\Sigma_{\rm crit}} \tag{3}$$

is the surface mass density in units of the critical surface mass $\Sigma_{\rm crit}$. The amplitude of the shear field is given by γ, and the background galaxy luminosity function is locally approximated by

$$n(L) \sim L^{-\beta} . \tag{4}$$

The amplification index $\beta - 1$ accounts for the expansion of the background image and for the increase in number as faint sources are lensed above the flux limit.

In the absence of galaxy clustering and finite sampling effects, the background galaxy distribution can simply be inverted, via equation (1), to find the amplification. One can then solve equation (2) to find the surface density, with some realistic assumptions about the shear. In § 4 we discuss various approximations that allow us to do this.

However, given a small resolution scale for the surface amplification, galaxy clustering and finite sampling will in general be an important effect. In § 3 we discuss the effects of intrinsic variation in the distribution of the background galaxy sources.

3. GALAXY CLUSTERING NOISE

The main sources of uncertainty in lens magnification are a result of shot noise, finite sampling, and the intrinsic clustering of the background source population that introduce correlated fluctuations in the angular counts. As we are viewing small angles, the clustering properties of the background source galaxies are not in general linear, unless the depth of background is sufficient to wash out the clustering pattern. As a result, it is not sufficient to make the usual assumption that galaxy clustering can be modeled by a Gaussian distribution.

We can account for the effects of shot noise and nonlinear clustering by modeling the angular counts by a lognormal-Poisson model (Coles & Jones 1991; BTP; Taylor & Dye 1998)—a random point-process sampling of a lognormal density field. The distribution function of source counts is then

$$P(n) = \frac{1}{n!}\langle \lambda^n e^{-\lambda}\rangle , \tag{5}$$

$$= \frac{\lambda_0^n}{n!}\int_{-\infty}^{\infty}\frac{dx}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2}{2\sigma^2} - \lambda_0 e^x - nx\right), \tag{6}$$

where $\lambda = \lambda_0 e^x$ is the local mean density, x is a Gaussian random variable of zero mean and variance σ^2, and $\lambda_0 = n_0 A^{\beta-1} e^{-\sigma^2/2}$ correctly normalizes the counts. The linear clustering variance σ^2 is related to the nonlinear variance by $\sigma^2 = \ln(1+\sigma_{\rm nl}^2)$. We have tested this distribution against available data and find that it is an excellent fit to the distribution of counts in the deep fields. The only parameters are the observed count per pixel n and the variance of the lognormal field. The amplitude of clustering of the density field and its dependence on redshift can be estimated from, e.g., the I-band–selected galaxies in the Canada-France Redshift Survey in the range $17.5 < I < 22.5$ (Le Févre et al. 1996; see § 5.2.3). The quantity required is the variance in a given area of sky, which can be estimated by averaging the observed angular correlation function $\omega(\theta)$ over a given area:

$$\sigma_{\rm nl}^2 = \bar{\omega}(\theta) = \frac{1}{\Omega(\theta)}\int_{\Omega} d^2\theta' \omega(\theta') , \tag{7}$$

where $\Omega(\theta)$ is the area.

Our method of approach is then that discussed by BTP. At each pixel in a map of the source counts, one uses the distribution equation (6) as a likelihood function, $\mathscr{L}(A\,|\,n, \sigma) = P(n\,|\,\sigma, A)$, assuming a uniform prior for the amplification. The surface density is then found from the amplification by making some realistic assumption about the shear and maximizing the likelihood. In § 4 we discuss a number of ways of transforming from the amplification to κ in the strong lensing regime.

4. THE STRONG LENSING REGIME

Transforming from amplification to the surface mass density is potentially nontrivial, as we have no shear information. One could incorporate this from independent measurements of the shear field, but for the present discussion we are interested in developing a completely independent lensing approach. We shall discuss combining shear and magnification elsewhere. In principle, one could generate a first guess for the surface mass density and iterate the amplification equation toward a solution of both surface density and shear. However, given the small field of view and uncertainties introduced by parity changes, this can be an unstable problem. In addition, as the solutions are in general multivalued, we would hope to start from as near to the correct solution as possible. In this section we discuss a number of reasonable approximations for solving the amplification equation (2). These can be regarded as solutions in their own right, or as the first best guess to an iterated solution. We begin by discussing the local approximation methods suggested and tested on simulated clusters by van Kampen (1998). Then, in § 4.2 we present a new self-consistent solution to the amplification equation for κ and γ for an axially symmetric lens.

4.1. *Local Approximations to the Surface Mass Density*

There exist only two local relations between γ and κ that result in a single caustic solution of the amplification equation (2) that is easily invertible (van Kampen 1998): $\gamma = 0$, corresponding to a sheet of matter, and $\gamma = \kappa$, for an isotropic lens. These two relations have corresponding estimators for κ as a function of amplification:

$$\kappa_0 \equiv \kappa(\gamma = 0) = 1 - \mathscr{P} A^{-1/2} , \tag{8}$$

$$\kappa_1 \equiv \kappa(\gamma = \kappa) = \tfrac{1}{2}(1 - \mathscr{P} A^{-1}) , \tag{9}$$

where $\mathscr{P} = \pm 1$ is the image parity.

Let us assume that the surface mass density of the lens is smooth over some scale. In this case, for a sufficiently smooth lens, $\gamma \leq \kappa$ (BTP). The equality holds in the case of an isotropic lens, for instance the isothermal lens. The inequality holds for any anisotropic lens, with the sheet mass at the extreme. For a smooth lens these two estimates bound the true value, $\kappa_1 \leq \kappa \leq \kappa_0$. Before caustic crossing it can also be shown that $\kappa_1 \leq \kappa_0 \leq \kappa_{\rm weak}$ holds, where $A = 1 + 2\kappa_{\rm weak}$ is the weak lensing limit (BTP). Hence the weak lensing approximation will overestimate the cluster mass in the strong regime, usually by a factor of 2 (van Kampen 1998).

In practice, substructure and asphericity of the cluster will induce extra shear (e.g., Bartelmann, Steinmetz, & Weiss 1995), especially in the surrounding low-κ neighborhood, where substructure is relatively more dominant, and filaments make the cluster most aspherical. This means that the lens will not be smooth for small κ, and therefore κ_1 is a lower limit for the true κ only for the central parts of the cluster, in the case where the lens parity is known. Van Kampen (1998) found it to be a good lower limit only for $\kappa > 0.4$ (for the most massive clusters), while for $\kappa < 0.2$, κ_1 is usually fairly close to the true value. For angle-averaged κ-profiles, κ_1 is a good lower limit for $\langle\kappa\rangle_\theta > 0.2$. All this has no bearing on κ_0, which remains a strong upper limit until the first caustic crossing.

A heuristic approximation, motivated by numerical cluster models, that tries to take these cluster lens features into account while still giving an invertible $A(\kappa)$ relation is (van Kampen 1998)

$$\gamma = |1 - c| \sqrt{\frac{\kappa}{c}} , \tag{10}$$

which results in an amplification relation that admits the full four solutions:

$$A^{-1} = |(\kappa - c)(\kappa - 1/c)| , \tag{11}$$

with caustics at $\kappa = c$ and $1/c$. The solution for κ is then

$$\kappa_c = \frac{1}{2c} \left[(c^2 + 1) - \mathscr{S} \sqrt{(c^2 + 1)^2 - 4c^2(1 - \mathscr{P} A^{-1})}\right] . \tag{12}$$

We shall refer to this as the parabolic approximation. Solutions are set by choosing the parities $\mathscr{P}, \mathscr{S} = \pm 1$, where $\mathscr{P}$ is the image parity, and $\mathscr{S}$ is the sign of $[(c^2 + 1)/2c - \kappa]$. Note that the sheetlike solution is recovered by setting $c = 1$.

Figure 1 shows a plot of κ versus the inverse amplification A^{-1} for the three estimators. Also shown is the weak field approximation. The points are taken from a simulated lensing cluster (van Kampen & Katgert 1997) that is of comparable size to A1689. It is clear that κ_0 is a strong bound, at least until a caustic is crossed, and that κ_1 provides a very good bound for $\kappa > 0.2$. The weak field approximation, however, is extremely bad, except in the very weak regime ($\kappa < 0.1$). The parabolic approximation behaves as it is designed to do: it is a good fit between the other two strong lensing estimators for the central parts of the cluster, while also modeling the $\gamma > \kappa$ behavior for small κ. These results are fairly robust over a wide range of clusters and for all realistic values of the cosmological density parameter Ω_0.

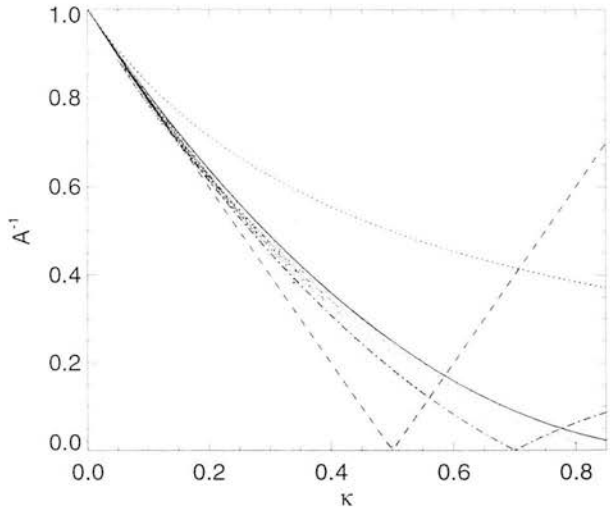

FIG. 1.—Scatter plot of the surface mass density κ vs. the inverse amplification A^{-1} for a simulated cluster in a CDM universe (see van Kampen 1998 for details). The cluster is at a redshift of 0.183, and the background population is at $z = 0.8$. *Solid line:* $\gamma = 0(\kappa_0)_h$ strong lensing approximation. Before caustic crossing this is a hard bound on the locus of points. *Dashed line:* $\gamma = \kappa(\kappa_1)_h$ approximation, which is a good lower bound for $\kappa > 0.2$ for this cluster. The weak lensing approximation (*dotted line*) is seen to be a very bad approximation for $\kappa > 0.1$. The dot-dashed line is a good fit to the simulation for the parabolic approximation $\gamma = |1 - c|(\kappa/c)^{1/2}$, with $c = 0.7$.

4.2. *A Nonlocal Approximation to the Surface Mass Density*

An alternative approach is to assume axial symmetry for the lens. Because this fixes a nonlocal functional relationship between κ and γ (eq. [15]), we can solve the amplification equation (2) for a self-consistent κ and γ profile. Although we shall apply our results to circularly averaged data, these results hold for any self-similar embedded set of contours.

We define a mean surface density interior to a contour by integration over the interior area $\Omega(\theta)$,

$$\bar{\kappa}(\theta) = \frac{1}{\Omega(\theta)} \int_\Omega d^2\theta' \kappa(\theta') . \tag{13}$$

The deflection angle for the axisymmetric lens is

$$\Delta\theta = \theta\bar{\kappa} , \tag{14}$$

and the shear is given by

$$\gamma = \gamma_t = |\kappa - \bar{\kappa}| , \tag{15}$$

where the tangential term γ_t is the only component of shear that is generated. The amplification factor is given by

$$A^{-1} = |(1 - \bar{\kappa})(1 - 2\kappa + \bar{\kappa})| . \quad (16)$$

One can now simultaneously solve for the surface mass density, shear, and amplification by series solution. First, we divide the surface mass into consecutive shells with equal separation (any arbitrary separation can be used; we have chosen a regular separation for convenience). If we split $\bar{\kappa}$ into an interior term, η_{n-1}, and a surface term, then for the nth shell we have

$$\bar{\kappa}_n = \eta_{n-1} + \frac{2}{(n+1)} \kappa_n , \quad (17)$$

where we have defined

$$\eta_{n-1} = \frac{2}{n(n+1)} \sum_{m=1}^{n-1} m\kappa_m . \quad (18)$$

The surface mass density in the nth shell is then given by

$$\kappa_n = \frac{(n+1)}{4n} \left(n + 1 - (n-1)\eta_{n-1} - \mathscr{S}\{[n - 1 - (n+1)\eta_{n-1}]^2 + 4n\mathscr{P}A_n^{-1}\}^{1/2} \right) , \quad (19)$$

where $\mathscr{P}$, $\mathscr{S} = \pm 1$ are again the image parities. The only freedom that we have, for a given amplification profile, is the choice of the shear on the first shell $\gamma_1 = \eta_0$ and the parity. It should be noted that given the amplification and having fixed the parities, one has to ensure that the first γ satisfies $\gamma^2 \geq \mathscr{P}A^{-1}$, in order to avoid unphysical solutions. The nonlocal approximation contains both the sheet and isothermal solutions as specific solutions. The uncertainty on κ and γ can be found by simple error propagation of the uncertainty on the measurement of the amplification.

Having shown in §§ 2, 3, and 4 how, in principle, one can measure the surface mass density from angular number counts, in § 5 we exploit these methods to measure the mass distribution in the lensing cluster Abell 1689.

5. APPLICATION TO A1689

In this section we apply the methods discussed in §§ 2, 3, and 4 to observational data. We begin by describing the data.

5.1. *The Data*

5.1.1. *Data Acquisition and Reduction*

The data were obtained during a run in 1994 February at ESO's NTT 3.6 m telescope, with 10^4 s integration in the V and I bands and covering 70 arcmin^2 on the cluster. Seeing was similar in both bands, with FWHM of 0.″8 and a CCD pixel scale of 0.″34. The EMMI instrument was used throughout. The passbands and exposures were chosen such that the cluster E/S0 galaxies would be bluer than a good fraction of the background, requiring much deeper imaging in the bluer passband for detection. The cluster was observed down to a limiting magnitude of $I = 24$.

The images were debiased and flattened with skyflats using standard IRAF procedures. After this, there remained some large-scale gradients of a few percent, probably caused by some rotation of the internal lens. We additionally corrected each separate exposure with a smoothed version of itself, obtained after masking out the cluster and other bright objects. Following this, we had homogeneous photometry across the field. (A further discussion of the reduction procedure can be found in Benítez et al. 1998.) The zero point was found to be good to 0.1 mag. High humidity on a few nights meant that some of the data were not photometric, so we calibrated with the photometric data. The object detection and classification was performed with SExtractor.

5.1.2. *Separation of Cluster and Background*

To measure the distortion in background counts, we must first separate the background from cluster members and mask off the area that they obscure. Cluster galaxies were identified from the strong cluster E/S0 color sequence, which forms a horizontal band across the color-magnitude diagram, shown in Figure 2. The sharp upper edge of this band represents the reddest galaxies in the cluster. Galaxies redder than this are cosmologically redshifted, and hence they represent a background population. As well as isolating cluster members, this selection should also ensure that any foreground galaxies are removed. Anything redder than $V-I = 1.6$ was selected as a background galaxy. Further color cuts where imposed to ensure completeness of the sample. The range of magnitudes was restricted to $20 < I < 24$, and the V band was limited to $V < 28$. Finally, we also cut at $V-I < 3.5$, where the reddest galaxies cut off.

Since the identification of cluster members is important for removing contamination of the background sample, we also checked our color-selected candidates with new data from a photometric redshift survey of the same field (Dye et al. 1998). We found general agreement with the simpler color selection.

Having identified foreground and cluster members, we produced a mask to eliminate those areas obscured by cluster members that would otherwise bias the mass estimate. To isolate the cluster members for the mask, we selected all the galaxies in the color-magnitude diagram lower than $V-I = 1.6$ and less than $I = 22$. This isolated most of

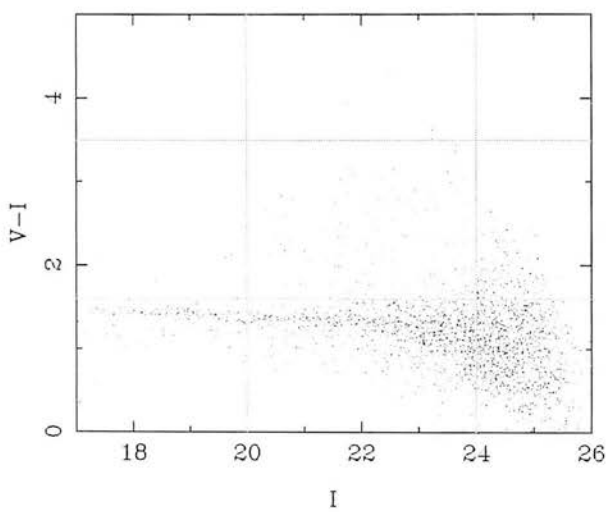

FIG. 2.—Color-magnitude diagram for A1689, overlaid with color cuts used to isolate the cluster members from the background population: $20 < I < 24$, $1.6 < V-I < 3.5$, and $V < 26.8$. The strong horizontal band of galaxies is the cluster E/S0 sequence.

the cluster sequence. We identified the remaining galaxies in the region $V-I < 1.6$ and $I > 22$, $V < 26.8$ as the faint blue background population. It is clear from Figure 2 that the distinction between faint cluster member and faint blue background galaxy is rather vague. However, since the faint cluster members are also the smallest, the masked area is fairly insensitive to the exact division. Figure 3 shows the distribution of cluster galaxies and the red background population. The concentric circles are centered on the peak in the cluster light distribution and show the position of the annuli used to calculate the radial profile in § 5.4.

5.1.3. *Selection by Color*

Once the cluster galaxies have been isolated, the background galaxies may be subdivided into a red and blue population, separated by $V-I = 1.6$. The observed slope of the luminosity function for these two populations for $I > 20$ is $\beta_R = 0.38$ and $\beta_B = 1$ (Broadhurst 1995; we shall do a more accurate fit using our color cuts in § 5.2.2). From equation (1) we expect that the surface density of red galaxies will be suppressed because of the dilation effect, while magnification of the faint blue galaxy population will compensate for the dilation. Hence, selecting by color allows us to identify a population of galaxies with a very flat luminosity function to boost the lensing signal, at the expense of a reduction in galaxy numbers. Simple error analysis shows that the signal-to-noise ratio varies as (Taylor & Dye 1998)

$$\mathrm{S/N} = 2|\beta - 1|\kappa A(1 - \kappa + \gamma'/\kappa')\sqrt{n}(1 + n\sigma^2)^{-1/2}\,, \quad (20)$$

where $' \equiv \partial/\partial R$. While the signal-to-noise ratio is a linear function of the slope of the luminosity function, it only grows with the square root of the galaxy numbers, assuming Poisson statistics. Hence one can get a better signal-to-noise ratio by preselection of the red background population to boost the signal, at the expense of numbers. Equation (20) also shows that one can get a better signal by observing to fainter magnitudes to enhance the surface number density and reduce the contribution from intrinsic clustering simultaneously (see Taylor & Dye 1998 for a more detailed discussion of observing strategies).

There is also a practical reason for favoring the red galaxy population. While the cluster members are unlikely to be redder than the cluster E/S0 sequence, the distinction between faint blue galaxies and cluster members, based on selection from the color-magnitude diagram alone, is somewhat vague. There may be blue cluster members that will contaminate the sample of blue background galaxies. In the absence of redshift information, the blue background population is clearly harder to isolate.

As we have noted, the red population has relatively few faint counts, so that the expansion term in equation (1) dominates, and there is a net underdensity of red galaxies

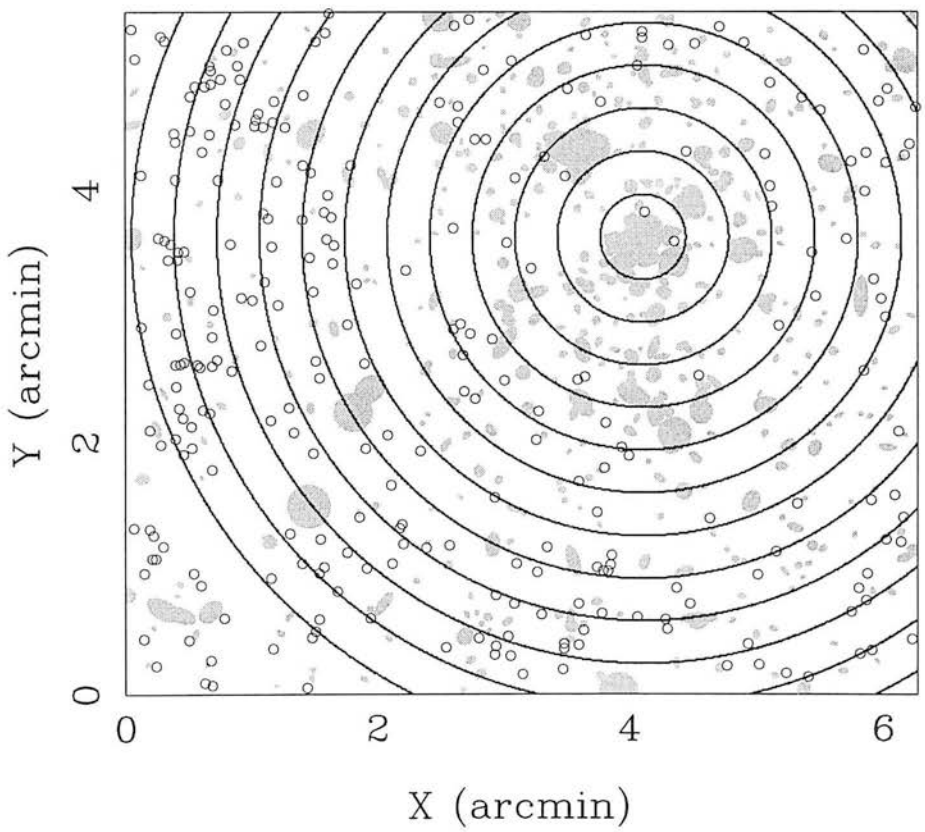

FIG. 3.—Masked region of A1689 (*gray area*). Cluster members were selected using color information (see text) and then masked over so that these regions do not affect the surface density estimate of background sources. The total region masked is about 10% of the area. The background galaxies are also shown as open circles. Superposed are the concentric bins used to calculate the radial profile, centered on the peak in the light distribution. North is up, and east is to the left.

behind the cluster (see Figs. 4 and 7). Conversely, faint blue galaxies are numerous and cancel the expansion effect. As expected, we found that the blue galaxies were uniform across the A1689 field (see Fig. 8). This is a good indicator that it is the magnification effect at work and not some spurious contaminant, for example color gradients across the field or large-scale variations caused by clustering. In addition, it also indicates that the deficit in the red population is not due to dust obscuration or reddening in the cluster, as this would affect both red and blue populations in equal measure.

5.2. *The Distribution of Background Galaxies*

In Figure 4 we show the surface distribution of the red population behind A1689, Gaussian smoothed on a scale of 0.'35. There are 268 background galaxies. The cluster members have been masked out and the masked areas interpolated over. The masked region contributes to only $\approx 10\%$ of the total field. Figure 3 shows the masked region. The cluster center, identified as the peak of the light distribution, is at (4.'1, 3.'6).

The angular size of the cluster scales as

$$R(\theta) = 0.87 D_A(z_c)(\theta/1') \, h^{-1} \text{ Mpc} , \quad (21)$$

where $D_A(z) = 2[1 - (1 + z)^{-1/2}]/(1 + z)$ is the comoving, dimensionless angular distance in an Einstein–de Sitter universe. Hence, at the redshift of Abell 1689, $z_c = 0.183 \pm 0.001$ (Teague, Carter, & Gray 1990), and 1′ is about $0.117 \, h^{-1}$ Mpc.

Figure 4 clearly shows a deficit of galaxies about the central peak in the light distribution at (4.'1, 3.'6). At $\theta = 0.'75$ there is an arc of very underdense number counts to the southwest of the cluster center, marked by a dashed line (The background is somewhat obscured by the cluster mask to the northeast of the cluster center.) This is clear indication of a caustic feature in the background number counts, where the number density drops to zero because of dilation. This exactly corresponds to the radius of the blue arcs observed by Tyson & Fischer (1995) at $\theta = 0.'85$ (see also the radial number counts in § 5.4). This is strong evidence that we have detected the magnification effect in the background counts.

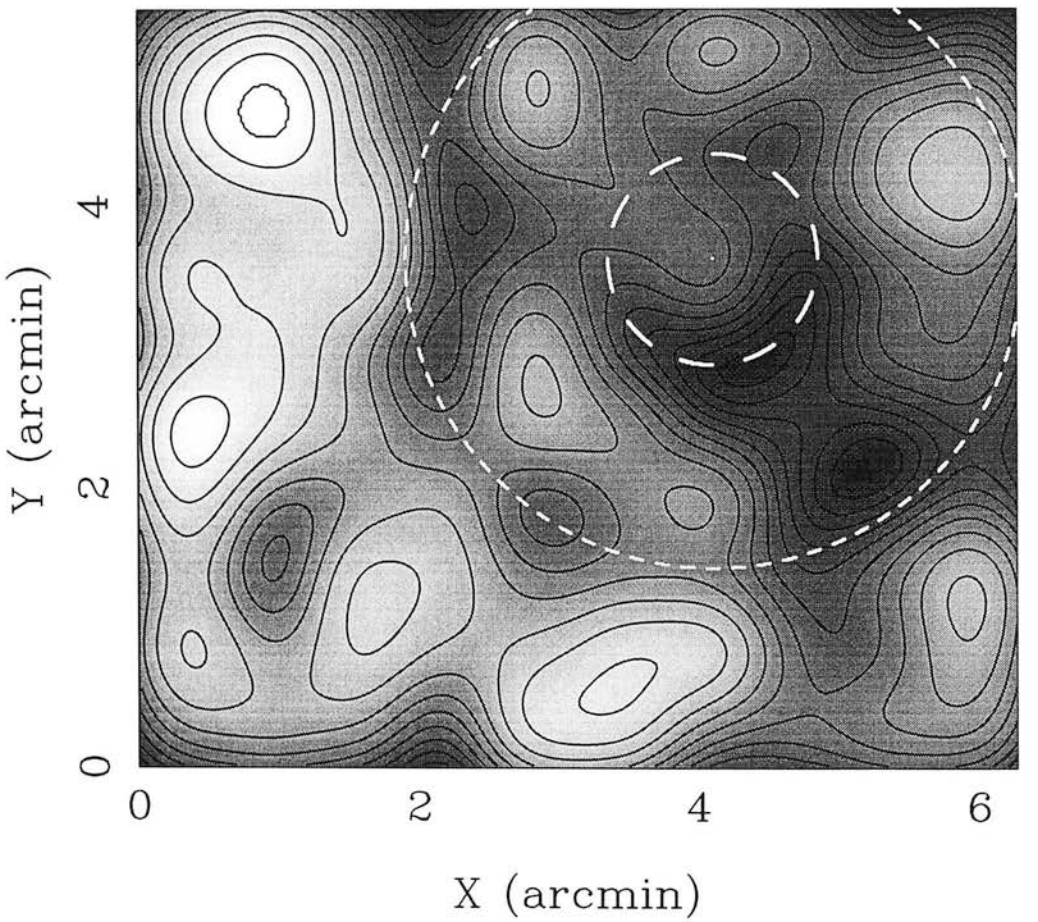

FIG. 4.—Distribution of red *I*-band background sources for Abell 1689. Darker gray areas indicate an underdensity of source counts. The image is Gaussian smoothed with a smoothing scale of 0.'35. The peak of the light distribution is at (4.'1, 3.'6). The maximum density of objects is 23.0 arcmin^{-2}, and the minimum is 1.1 arcmin^{-2}. There are 15 contour lines spaced by $\Delta n = 1.46$ galaxies arcmin^{-2}. A strong caustic feature is seen 0.'75 from the peak (*inner dashed line*), more visible to the southwest, as the other side of the peak is masked over. A second feature is found in the radial profile at 2.'2 (*outer dotted line*). The image is oriented with east to the left and north to the top.

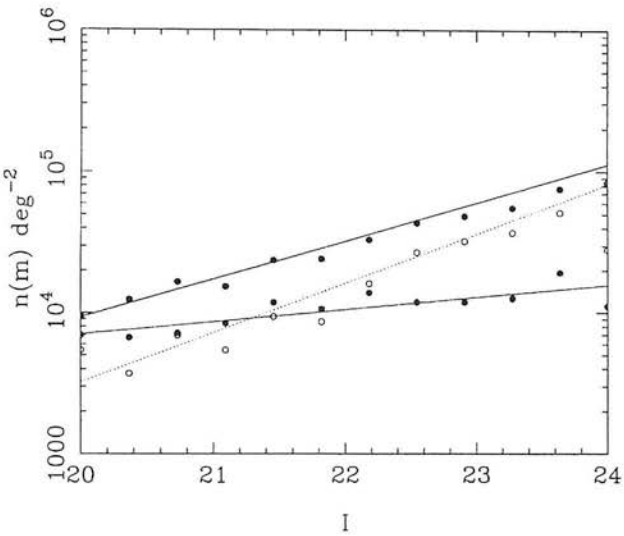

FIG. 5.—Magnitude distribution of all I-band galaxies (*solid dots*), the red-selected galaxies (*gray dots*), and the blue background galaxies (*open dots*). The lines are the best fits to the data.

5.2.1. *The Redshift Distribution of Background Galaxies*

The efficiency of lensing varies with the redshift of the background source (BTP). Therefore it is important to estimate the background redshift distribution. Crampton et al. (1995) find that Canada-France Redshift Survey (CFRS) has a median redshift of $z = 0.56$ for galaxies in the range $17.5 < I < 22.5$. They also show a color-redshift diagram that indicates that the red galaxy population ($V-I > 1.6$) has a median redshift of about $z \approx 0.8$ (Crampton et al. 1995, their Fig. 5). More accurately we can integrate the best-fit Schechter function found by Lilly et al. (1995) for the CFRS red galaxy population. This has parameters $\phi^* = 0.0031 \pm 0.00095$, $\alpha = 1.03 \pm 0.15$, and $M^*(B) = -21$, where $M(B) = I - 5 \log_{10} (D_L/10 \text{ pc}) + 2.5 \log_{10} (1 + z) + k$-correction, where the k-correction is discussed in their paper, and $D_L = (1 + z)^2 D_A(z)$ is the luminosity distance. Lilly et al. found no detectable evolution of the luminosity function of the CFRS red population, and we assume no evolution. Extrapolating to the magnitude range $20 < I < 24$, we find that the redshift distribution can be well fitted by the function

$$\phi(z) = \frac{\alpha z^2}{z_*^3 \, \Gamma(3/\alpha)} \exp\left[-\left(\frac{z}{z_*}\right)^\alpha\right], \tag{22}$$

with $\alpha = 1.8$ and $z_* = 0.78$ to about 5% accuracy over a redshift range of $0.25 < z < 1.5$. The moments of this distribution are

$$\langle z^n \rangle = z_*^n \frac{\Gamma[(3 + n)/\alpha]}{\Gamma(3/\alpha)} . \tag{23}$$

Hence, for the red galaxy population we find that $\langle z \rangle = 0.96$ and $\sigma_z = 0.42$. To simplify the analysis of the lensing properties of the cluster, we shall assume hereafter that the background distribution is at a single redshift of $z = 0.8$ (the mode of the distribution) and has an uncertainty of $\delta z = 0.4$.

As the caustic indicated by the blue arcs coincides with the magnification caustic, we can presume that the galaxy forming the arcs lies at the same redshift as the magnified red background galaxies, $z \approx 0.8$. At present we do not know the redshift of this arc.

5.2.2. *Number Counts of the Background Galaxy Population*

Of major importance to the lens magnification method is the normalization of the background galaxy population. The CFRS is not adequate for this, since their color cuts were in the rest frame $U-V$, rather than in the observed $V-I$. Instead, we have used the Keck data of Smail et al. (1995), who observed deep VRI images down to a limiting magnitude of $R \approx 27$. The total differential galaxy count rate in the I band can be approximated by

$$\log_{10} n = (0.271 \pm 0.009)I - 1.45 \tag{24}$$

over the range $20 < I < 24$, where n is in mag^{-1} deg^{-2}. We have applied our color criteria (see § 5.1) to the Keck data and find that the red galaxy population $V-I > 1.6$ can be well approximated by

$$\log_{10} n(\text{red}) = (0.0864 \pm 0.0187)I + (2.12 \pm 0.41) \tag{25}$$

over the range $20 < I < 24$. Figure 5 shows the magnitude distribution for the full data set and for the red-selected galaxy population and the best-fit lines. Integrating the fit for the red galaxies yields a total count rate of $n = 12.02 \pm 3.37$ galaxies arcmin^{-2} in the range $20 < I < 24$. Since $\beta = 2.5 \, d \log_{10} n/dm$, we find that the Keck data imply $\beta_R = 0.216 \pm 0.047$. This is the value of β that we shall use in the subsequent analysis.

An alternative, although less exact, method of normalization is to assume negligible cluster mass at the edge of the field and to normalize the cluster to this. In general, this would put a lower limit on the mass and is similar to the method used to normalize shear mass maps. In fact, if we do this for A1689, we find a background count rate that is very similar to that given by the Keck data. The error introduced into the final mass estimate by uncertainties in β scales as $\delta\kappa/\kappa \approx \delta\beta/|1 - \beta|$, which for the Keck data results in a fractional error of around 5%.

We have also fitted the blue counts in the Keck sample (Fig. 5). Over the same range as the red counts, we find that $\log_{10} n(\text{blue}) \approx 0.35I - 3.49$, resulting in $\beta_B = 0.88$, close to the lens invariant $\beta = 1$, and a count density between $23 < I < 24$ of $n_0(\text{blue}) = 15.5$ galaxies arcmin^{-2}.

5.2.3. *Clustering Properties of the Background Population*

The amplitude of clustering of I-band galaxies and its dependence on redshift can be estimated from the CFRS (Le Fèvre et al. 1996). Le Fèvre et al. (1996) find that there is little difference between the clustering properties of red and blue populations of galaxies for $z > 0.5$, implying that the populations were well mixed at this epoch. We therefore apply their clustering results directly to our red galaxy population. They fitted their results to a power-law model for the evolving correlation function, $\xi(r) = (r/r_0)^{-\gamma}$, where

$$r_0(z) = r_0(0)(1 + z)^{-(3+\varepsilon)/\gamma} , \tag{26}$$

where $\varepsilon = 1 \pm 1$, and $r_0(z = 0.53) = 1.33 \pm 0.09 \; h^{-1}$ Mpc, and $\gamma = 1.64 \pm 0.05$ is in this section the slope of the correlation function.

The quantity that we require is the variance in a given area of sky, which can be estimated by averaging the observed angular correlation function $\omega(\theta)$ over a given

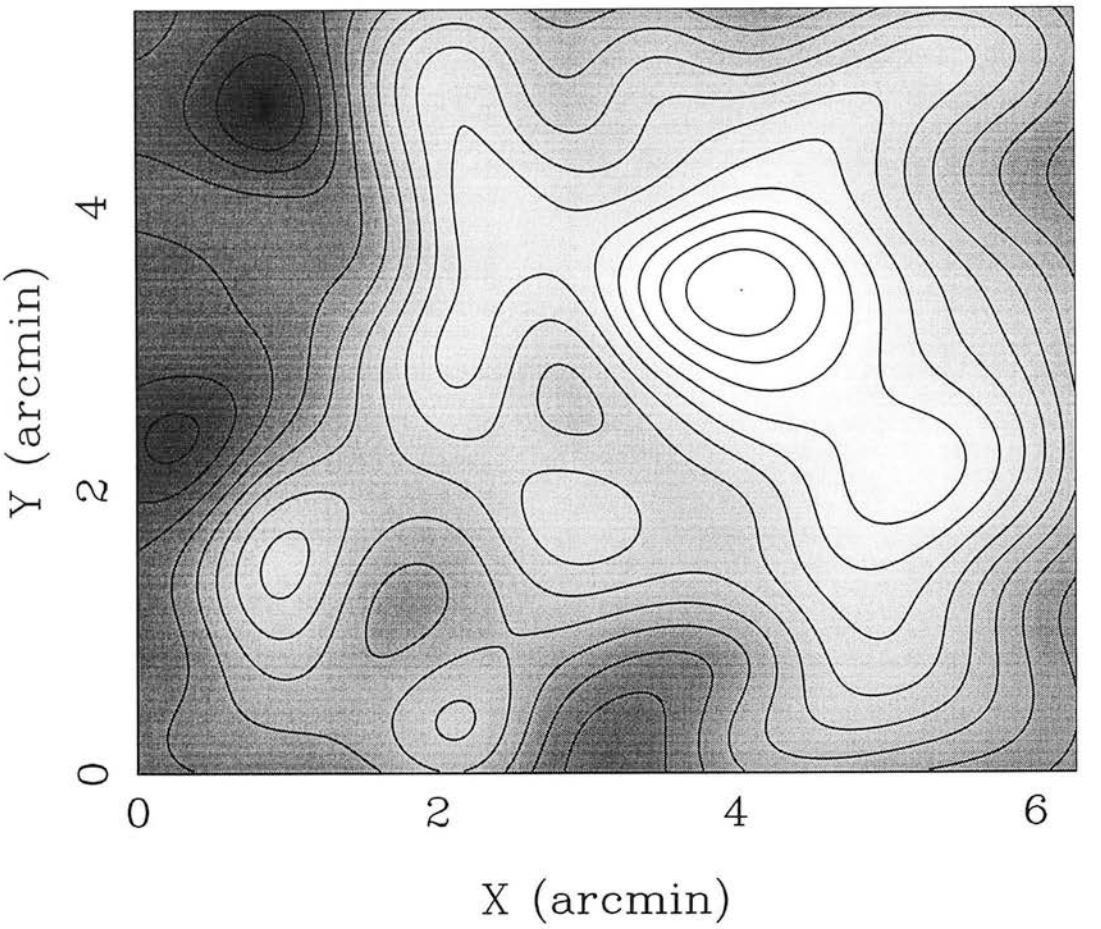

FIG. 6.—Reconstruction of the surface mass density of Abell 1689 from the red background galaxy population, using the nonlinear local sheet approximation ($\gamma = 0$) and a full likelihood analysis in two dimensions. Light regions are high density. Only one caustic line is assumed, at $\theta = 0.'75$ from the peak of the light distribution. The maximum surface density is $\kappa = 1.35$, at (4.'02, 3.'41), consistent with the peak in the light distribution. The minimum surface mass density is $\kappa = -0.47$. There are 15 linearly spaced contours, separated by $\Delta\kappa = 0.12$, and the map is Gaussian smoothed with a smoothing length of $\theta_S = 0.'35$. North is up and east is to the left.

area (eq. [7]). The clustering variance for I-band galaxies then scales roughly as (Taylor & Dye 1998)

$$\sigma^2_{\rm nl} = 10^{-2} z^{-2.8} (\theta/1')^{-0.8} , \qquad (27)$$

where the sampled area is a circle of radius θ, and we have assumed unbiased linear evolution of the density field. The background galaxies are assumed to all lie at $z \approx 1$.

5.3. *Reconstructing the Surface Mass Density*

In Figure 6 we plot the reconstructed surface mass density of Abell 1689 using the nonlinear local sheet approximation κ_0 (see § 4.1), changing parity on the caustic line at $\theta = 0.'75$ (see Fig. 4). The uncertainty on the peak of the mass distribution is somewhat large (see § 5.2), but significant features can be seen around the cluster core. There appears to be an extension to the southwest that is not seen in the cluster galaxy distribution. Interestingly, there also appears to be a loosely connected ridge, about 2.'4 from the peak. We shall discuss this feature further below, but note that the shear mass map derived by Kaiser (1996, Fig. 2) shows similar extensions and ridge, although the extension to the west is not apparent in the shear map. Two underdense regions are also seen to the south and to the east in both maps. While the comparison is only qualitative and the maps are noisy, we find this very encouraging, as these maps are derived from completely independent methods, although the underlying data set is the same.

5.4. *The Mass Profile of Abell 1689*

While the mass maps are suggestive, a more quantitative measure can be made by angle averaging the counts and calculating the mass profile. Figure 7 shows the radial counts about the peak in the light distribution, normalized to the Keck data. The plotted error bars result from only Poisson statistics, although in the mass analysis below we shall take into account the effects of clustering. A general trend is clear and lies close to the prediction for an isothermal lens normalized to the blue arc caustic. This has a surface mass density of $\kappa = 0.375(\theta/1')^{-1}$, corresponding to a virial velocity of 1600 km s^{-1}. Again, it is worth emphasizing that the zero of the number counts at $\theta = 0.'75$ corresponds to the caustic inferred from the blue arcs. The second dip will be discussed in more detail in § 5.4.4. The

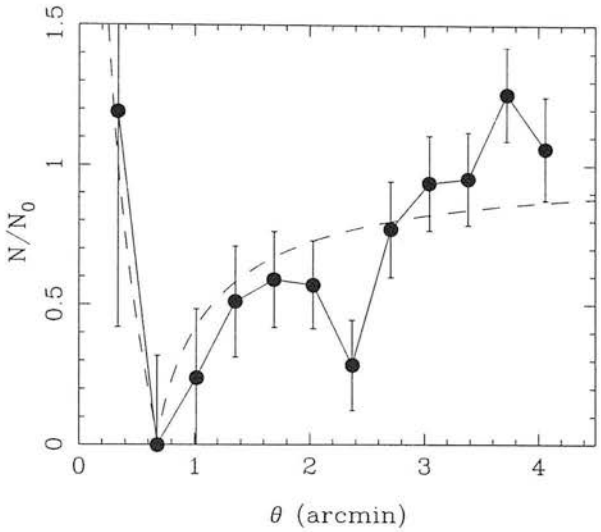

FIG. 7.—Radial profile of red counts behind Abell 1689. The background count density is $n_0 = 12$ objects arcmin^{-2}. Superposed is the profile for an isothermal model, normalized at the caustic radius $\theta = 0.'75$ (*dashed line*).

increase in counts at $\theta = 3.'7$ is likely to be the result of a clustering effect. Table 1 contains the shell radii, red galaxy counts, and total and obscured area of the annuli.

In Figure 8 we show that radial profile for the blue galaxy population, normalized with the Keck data in § 5.2.2. As expected, there is no lensing signal. The slight increase toward the cluster center is caused by contamination from the blue cluster members.

5.4.1. *Local Approximations for the Surface Mass Density*

Figure 9 shows the radial mass profile of the cluster Abell 1689 assuming a single caustic at $\theta = 0.'75$. The inner two

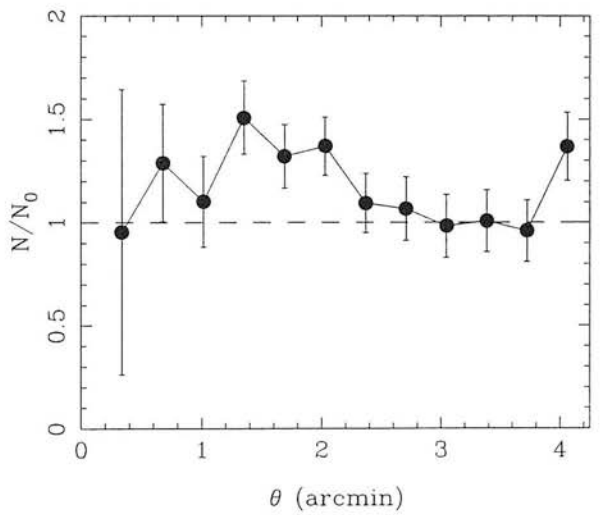

FIG. 8.—Radial profile of blue counts behind Abell 1689. The background count density is $n_0 = 22$ objects arcmin^{-2}. As expected from the nearly lens-invariant slope $\beta_{\rm blue} = 0.88$, the number counts are nearly flat and at large radii tend toward $n/n_0 = 1$. The slight increase toward the cluster center probably results from contamination of the counts by blue cluster members.

solid lines are calculated using the lognormal-Poisson likelihood estimator (eq. [6]) with each of the two single caustic strong lensing approximations (eqs. [8] and [9]). The light shaded region indicates the 1 σ uncertainty owing both to shot noise and to the effects of clustering. The dark shaded region indicates the region between the two extreme estimators. Away from the cluster center these agree and are equal to the weak lensing estimator, but noise effects become dominant. Closer to the cluster center the uncertainty due to the shear increases and becomes dominant at $\theta < 1'$. However, the cluster mass profile is significantly detected between $1' < \theta < 2.'6$. We also appear to see a deviation from an isothermal profile, which is also plotted. When the procedure was repeated with the center of the annuli offset from the peak of the light distribution, the mass profile was weaker and less significant, as one would expect if the peak of the mass density was associated with that of light.

5.4.2. *Nonlocal Approximation for the Mass Density and Shear*

In Figures 10 and 11 we assume axisymmetry and equation (19) to calculate the surface mass density and shear simultaneously. We set $\gamma_1 = 0.3$ for the first shell. The resulting profile is fairly insensitive to this choice, only affecting the first two shells. The uncertainty on the shear in the first shell is small, because this must be chosen a priori. However, averaging over shells means that the errors do not strongly propagate through to higher radii. Again, a mass detection is found between $1'$ and $2.'8$, this time with the shear being accounted for. In this region $\kappa \approx 0.4 \pm 0.15$, which is somewhat higher than that found by the shear estimate of $\kappa = 0.2 \pm 0.1$ (Kaiser 1996). (Note that we quote Kaiser's color-selected sample, where cluster members that may contaminate the shear estimate have been removed. This corresponds to combining our red and blue background populations. This will change the redshift distribution of the background and include some residual blue cluster contamination that may account for the discrepancy.) Also, for the single caustic solution, we see a large spike at $2.'2$, which is not seen in the Kaiser (1996) results. However, the shear method correlates points, which may lead both to the suppression of features and to underestimation of the errors.

Our estimate of the shear field is far more uncertain, with $\gamma_t = 0.2 \pm 0.3$ over most of the range. There is a slight increase beyond $2.'4$ owing to the spike in the surface mass profile at that radius, but the profile is dominated by noise. This increase is not reflected in the angle-averaged measurements of Kaiser (1996), where the mean shear is $\gamma = 0.15 \pm 0.05$.

5.4.3. *Local Approximation for the Surface Mass Density and Shear*

Figures 10 and 11 also show κ and γ estimated from the parabolic solution of § 4.1. We find good agreement between the local and nonlocal approximations for κ, but the the shear profiles are somewhat different, reflecting that one estimator is local and one nonlocal. However, the large uncertainties produced by each estimator mean that we cannot predict the shear profile with much certainty from the available data.

5.4.4. *Two Background Populations?*

An interesting feature of the counts in Figure 7 is the appearance of two pronounced dips, one at $0.'75$ and another one at $2.'2$. While the inner dip has already been

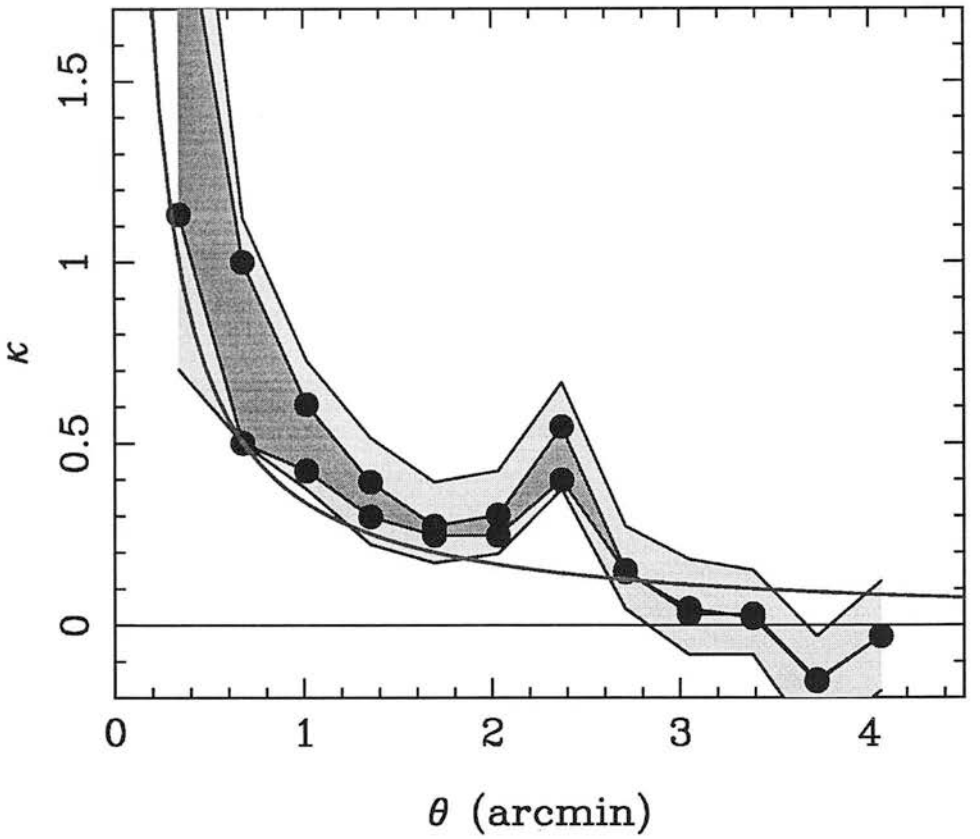

FIG. 9.—Radial profile of surface mass density of cluster Abell 1689. *Dark solid region:* Uncertainty caused by the strong lensing estimators. *Lighter shaded region:* Uncertainty caused by clustering and shot-noise uncertainty in the background population. *Solid line:* A singular isothermal profile, normalized to the caustic feature at $\theta = 0'.75$.

identified with a caustic line, the outer dip is somewhat anomalous. A number of possibilities could account for this. The feature was noted in the mass plot as a low signal-to-noise ridge in the density and can be seen in the number counts as a loosely connected ring about the cluster center. One possibility is that this results from clustering in the background population, combined with a large mass concentration to the southeast of the peak in the light distribution. There are few cluster members in the region of the ridge or the bump, so the effect is not caused by masking.

An alternative is that this is the first glimpse of a second caustic line. In principle, a second caustic can be created by placing the background galaxies at two redshifts, one at low redshift, and one at high redshift (e.g., Fort et al. 1996). The observed number counts would then be given by

$$n'/n_0 = A_1^{\beta-1} + \nu(A_2^{\beta-1} - A_1^{\beta-1}) , \qquad (28)$$

where $A_i = A(f_i)$, with $f_i = \kappa(z_i)/\kappa_\infty = [(1 + z_i)^{1/2} - (1 + z_L)^{1/2}]/[(1 + z_i)^{1/2} - 1]$ (BTP), and $i = 1, 2$ for the two galaxy populations. Here, ν is the fraction of galaxies at redshift z_2. An outer caustic line must be produced by the high-redshift population. If we make this population lie at $z = 0.8$, then the low-redshift population must lie at $z = 0.3$. Both populations are reflecting the same arc: the difference in projected radii is wholly a result of their relative redshifts.

However, this would double the predicted mass from lens magnification, making Abell 1689 a very extreme cluster. In addition, it seems hard to make a caustic line from the high-redshift population for such a massive cluster without forming a second, inner radial caustic. As the strongest arc is tangential and is seen near the inner arc, one would have to conspire to have a nearby galaxy, at $z = 0.3$, lensed and lying at the same projected radii as the radial arc produced by the high-redshift population. This seems highly unlikely.

One could also keep the mass roughly constant and place a second population at $z > 0.8$. This is a possibility, but it does not strongly affect our mass estimate assuming a single caustic solution. In the absence of further evidence for a second high-redshift population, we shall only consider the single caustic model.

5.5. *Mass Estimate of Abell 1689*

5.5.1. *From κ to Mass Surface Density*

Assuming that the background galaxies all lie at the same redshift of $z = 0.8$, and given that the surface density scales as

$$\Sigma = \frac{1}{3}\,\Sigma_0\,\kappa\,\frac{(\sqrt{1+z}-1)(1+z_L)^2}{(\sqrt{1+z_L}-1)(\sqrt{1+z}-\sqrt{1+z_L})} , \qquad (29)$$

where $\Sigma_0 = 8.32 \times 10^{14}\, h\, M_\odot\ \mathrm{Mpc}^{-2}$ is the mean mass per unit area in the universe, then we find that the surface mass density is

$$\Sigma = 5.9 \times 10^{15} \kappa (h\, M_\odot\ \mathrm{Mpc}^{-2}) . \qquad (30)$$

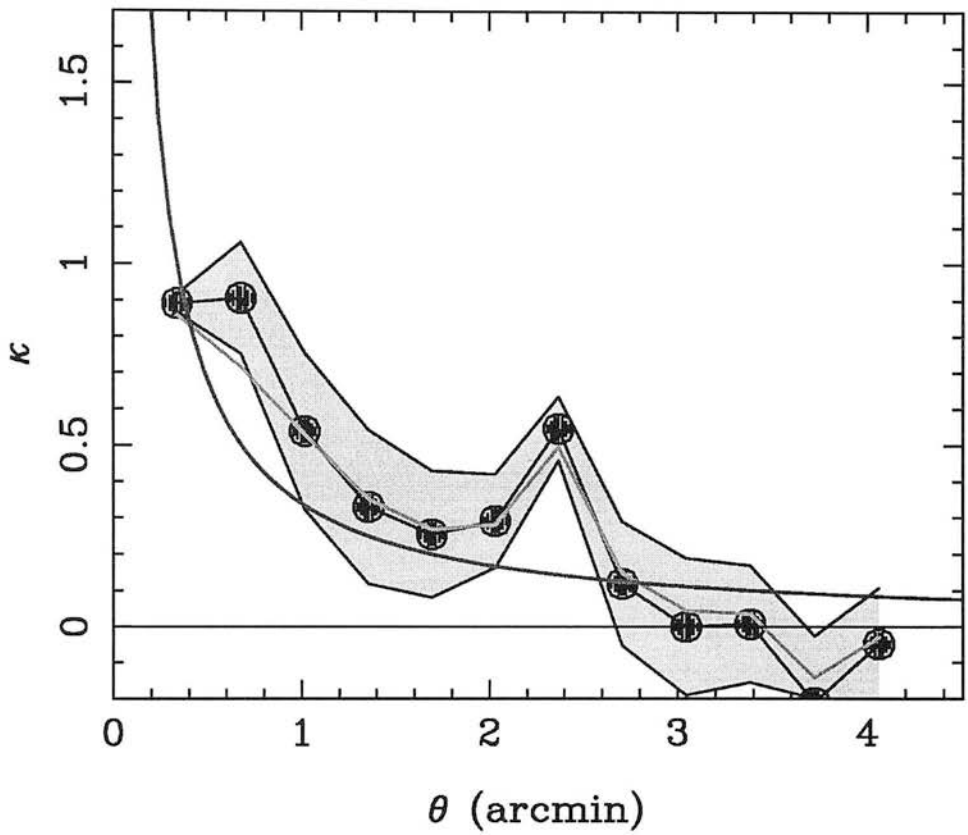

FIG. 10.—Radial profiles of surface mass density κ for A1689 (*solid line with dots*), calculated by solving the axially symmetric lens equation (19). *Shaded regions:* 1 σ errors calculated via error propagation from the uncertainty on the measured amplification profile. *Solid dark line:* A singular isothermal profile normalized to the caustic feature at $\theta = 0'.75$. *Lighter solid line:* Local parabolic estimator κ_c.

Although we have assumed an Einstein–de Sitter universe, these results only depend weakly on cosmology (BTP).

5.5.2. *Uncertainty in the Redshift Distribution*

The error introduced by assuming that the background galaxies lie at the same redshift can be estimated by error propagation and by assuming $\delta z = 0.4$ (see § 5.2.1). Hence, $\delta\Sigma = |\partial\Sigma/\partial z|\,\delta z$ and the fractional uncertainty on the surface mass density owing to the uncertainty in redshift distribution of the background galaxies is $\delta\Sigma/\Sigma = 0.37\ \delta z = 0.148$. The same error is also found in mass estimates based on the shear pattern.

5.5.3. *Uncertainty Arising from Normalization of Background Counts*

Assuming a sheet mass solution (κ_0 in § 4.1), we find that the uncertainty arising from the normalization of the background counts is $\delta\kappa = (|1 - \kappa|/2|\beta - 1|)\delta n_0/n_0$. For A1689 and the red galaxy population, this is $\delta\kappa = 0.15|1 - \kappa|$. For an average $\kappa = 0.5$, the uncertainty is around $\delta\kappa = 0.07$.

5.5.4. *The Cumulative Mass Distribution*

Figure 12 shows the cumulative mass interior to a shell, calculated from both the nonlocal approximation (§ 4.2) and the local parabolic approximation allowing only a single caustic solution (§ 4.1). The uncertainties are treated by error propagation. We find that the two-dimensional projected mass interior to $0.24\ h^{-1}$ Mpc is

$$M_{2D}(<0.24\ h^{-1}\ \mathrm{Mpc}) = (0.50 \pm 0.09) \times 10^{15}\ h^{-1}\ M_\odot\ , \tag{31}$$

and that the two estimators are in good agreement. We find that the projected mass scales as

$$M_{2D}(<R) \approx 3.5 \times 10^{15}(R/h^{-1}\ \mathrm{Mpc})^{1.3}\ h^{-1}\ M_\odot\ , \tag{32}$$

for $R < 0.32\ h^{-1}$ Mpc, similar to that for an isothermal sphere, $M \sim R$. Hence it appears that A1689 has a near-isothermal core. Beyond $R = 0.32\ h^{-1}$ Mpc the lensing signal is lost in background noise, and we can only say that $\kappa \leq 0.1$

Including the uncertainty from the background redshift distribution and the normalization of background counts increases the error to about 30%.

6. COMPARISON WITH OTHER MASS ESTIMATES OF A1689 AND INFERRING THE THREE-DIMENSIONAL MASS DISTRIBUTION

In this section we compare the mass derived from lens magnification to that found from a number of other independent measurements. First, we compare our results to estimates of the mass based on the shear pattern found around A1689 (§ 6.1). The magnification and shear complement each other in that the shear pattern has a higher

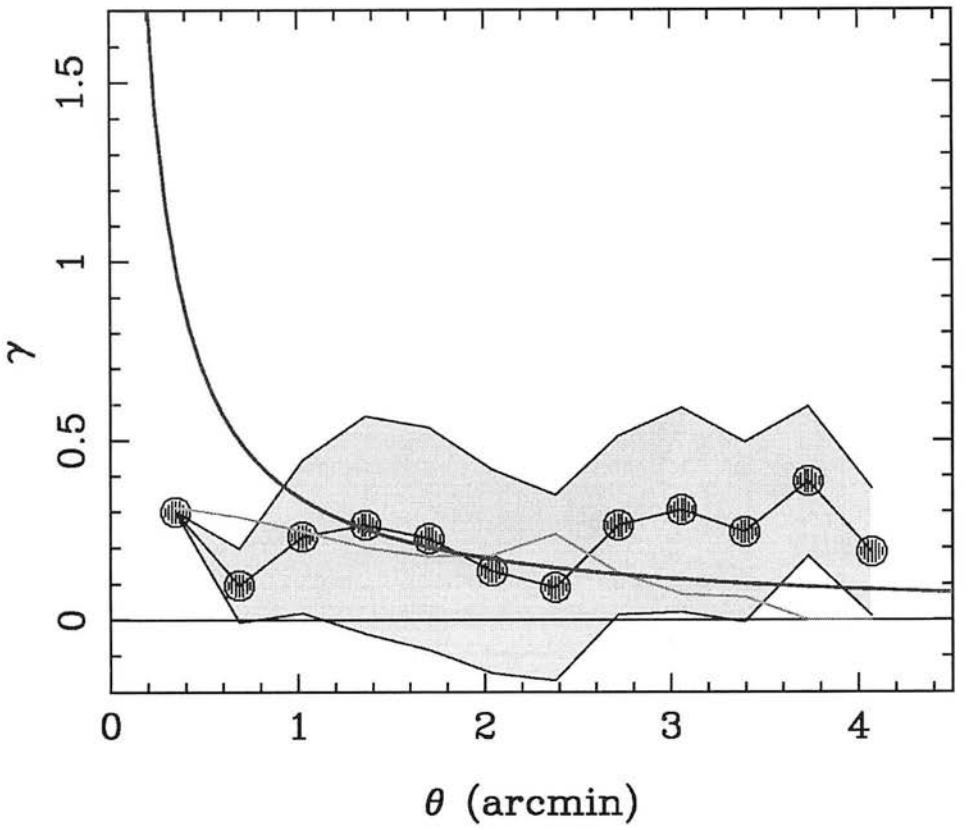

FIG. 11.—Radial profiles of tangential shear γ_t for A1689 (*solid line with dots*), calculated by solving the axially symmetric lens equation (19). *Shaded regions:* 1 σ errors calculated via error propagation from the uncertainty on the measured amplification profile. *Solid dark line:* A singular isothermal profile normalized to the caustic feature at $\theta = 0.'75$. *Lighter solid line:* Local parabolic estimator κ_c.

signal-to-noise ratio, since it is not affected by clustering noise (although with redshift information, the magnification can also be measured free from clustering noise; see BTP), but suffers from the "sheet-mass" degeneracy. We shall combine the magnification and shear pattern elsewhere.

While the lens magnification mass is vital for fixing the total two-dimensional projected mass distribution independently of any assumptions about the dynamical state of the cluster, much information can be gained by combining this with other mass estimates, assuming that these are not strongly biased by their reliance on thermodynamical equilibrium. In this section we describe a method for transforming from the two-dimensional lens mass to other cluster characteristics, such as the line-of-sight velocity dispersion (§ 6.3) and the X-ray temperature (§ 6.4). Discrepancies that arise between these predicted characteristics and the actual measurements can be used to infer information about the mass distribution along the line of sight (Bartelmann & Kolatt 1997). We find that while there is fair agreement between all of the mass estimates when projection effects are taken into account, the agreement is better if the cluster A1689 is composed of two clusters superposed along the line of sight and separated by about $\Delta z = 0.02$.

The transformation from a two-dimensional projected lensing mass to a three-dimensional mass, line-of-sight velocity dispersion and X-ray temperature can be made using either the isothermal model or by using relations found in N-body simulations of clusters. While the former provides a simpler method, one has more freedom with simulations to include or exclude the various projection effects that contaminate measurements of these quantities. In this section we shall use the relations found by van Kampen (1998) from an ensemble of CDM cluster simulations, all with $\Omega_0 = 1$ and $\sigma_8 = 0.54$. These relations are model dependent, but serve to aid comparison between the various mass measurements. We have also provided a table of quantities (Table 2) in which the uncertainties have been calculated by combining the error on the cluster mass with the dispersion found in the deprojection relations.

We begin by comparing the lens magnification mass to the mass determined from the shear field around A1689.

6.1. *Comparison with Arclets and Weak Shear*

Tyson & Fischer (1995) provide mass profiles of A1689 from arclets, another independent estimator of the mass, normalized to the caustic line indicated by the blue arcs. They find that the two-dimensional projected mass within $R = 0.1\,h^{-1}$ Mpc is

$$M_{2D}(<0.1) = (0.18 \pm 0.01) \times 10^{15}\ h^{-1}\ M_\odot\ . \quad (33)$$

They also find that the mass scales like an isothermal sphere out to $0.4\,h^{-1}$Mpc, before turning over to an $R^{-1.4}$ profile. This implies that in the regime that we probe with the magnification the cumulative mass scales like

$$M(<R) = (1.8 \pm 0.1) \times 10^{15}(R/h^{-1}\ \text{Mpc})\ h^{-1}\ M_\odot\ . \quad (34)$$

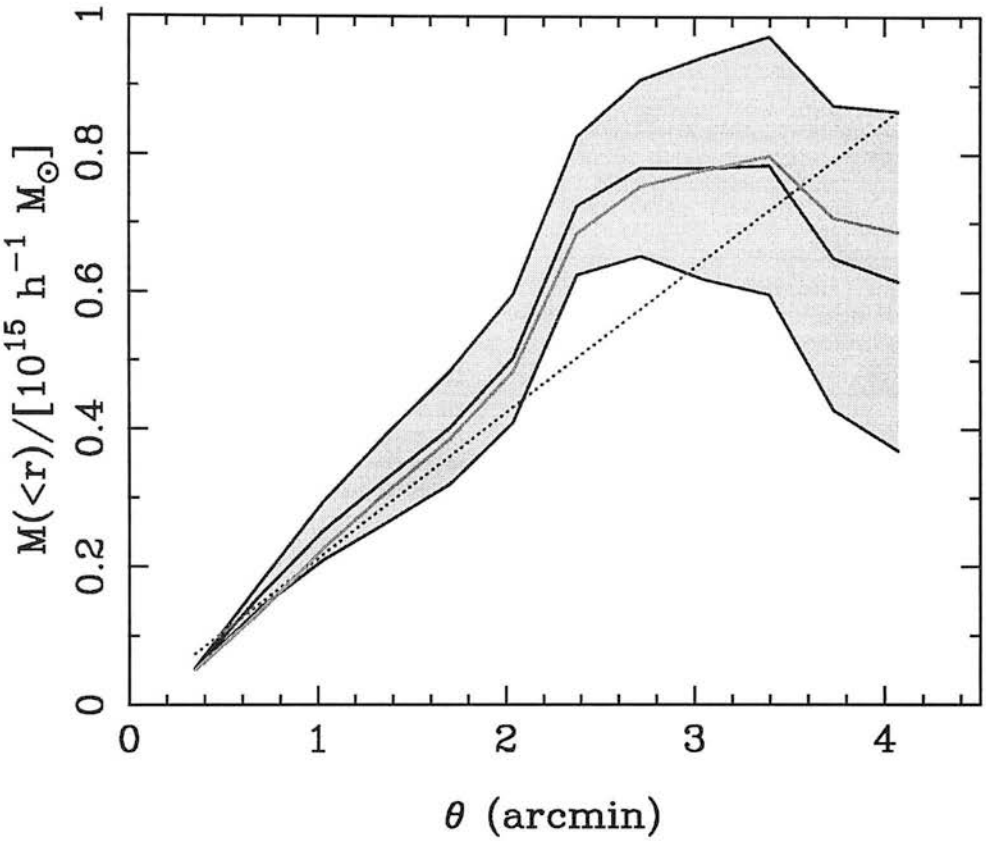

FIG. 12.—Cumulative mass profile of Abell 1689. *Solid dark line and shaded uncertainties:* Estimated using the axisymmetric nonlocal estimator described in § 4.2. *Lighter gray line:* Cumulative mass estimated from the local parabolic approximation κ_c, described in § 4.1. Also plotted is the isothermal fit to the blue arc caustic (*dotted line*), similar to the shear results of Kaiser (1996) and Tyson & Fischer (1995).

This is very close to the profile that we find from lens magnification (eq. [32]). Using this, we scale their results, giving

$$M_{2D}(<0.24) = (0.43 \pm 0.02) \times 10^{15}\ h^{-1}\ M_{\odot}\ , \quad (35)$$

in good agreement with the mass from magnification.

Kaiser (1996) has also calculated κ based on the weak shear method (Kaiser & Squires 1993), using the same data that we have used here for A1689. We noted above that there are qualitative similarities between the weak shear maps and those presented by Kaiser, which is significant, since the methods are independent. The mass density profile found from the shear pattern is also well fitted by an isothermal profile:

$$M_{2D}(<R) = 1.8 \times 10^{15}(R/h^{-1}\mathrm{Mpc})\ h^{-1}\ M_{\odot}\ , \quad (36)$$

with a 10% statistical uncertainty and further 10% systematic error owing to the uncertainty in the redshift distribu-

TABLE 1

PARAMETER VALUES AT VARIOUS ANGULAR RADII

r (arcmin)	N	N/N_0	Annulus Area (arcmin^2)	Obscured Area (arcmin^2)
0.33.....	2	1.19	0.35	0.21
0.67.....	0	0.00	1.08	0.25
1.01.....	4	0.24	1.79	0.40
1.35.....	13	0.51	2.51	0.39
1.69.....	20	0.59	3.23	0.40
2.03.....	23	0.57	3.62	0.26
2.36.....	11	0.28	3.46	0.23
2.70.....	26	0.77	3.08	0.26
3.04.....	32	0.94	3.01	0.16
3.38.....	34	0.95	3.14	0.17
3.72.....	45	1.25	3.18	0.20
4.06.....	31	1.06	2.49	0.06

NOTES.—Angular radius (r in arcminutes), number of red galaxies (N), ratio of galaxies to background (N/N_0), the total area of the annuli, and the area obscured by the mask. The unobscured area is total area − obscured area. The expected number of galaxies in an annuli is $N_0 = n_0 \times$ unobscured area.

TABLE 2

MASS ESTIMATES FOR A1689

Quantity (1)	This Work (2)	Other (3)
$M_{2D}(<0.24)$...	0.50 ± 0.09	0.43 ± 0.02 (Tyson & Fischer 1995)
		0.43 ± 0.04 (Kaiser 1996)
$M_{3D}(<0.5)$....	0.72 ± 0.25	
M_{500}	1.6 ± 0.65	0.95 ± 0.16 (Yamashita 1994)
$\sigma_v(<1.5)$.....	2200 ± 500	2355^{+238}_{-183} (Teague et al. 1990)

NOTES.—Mass estimates for A1689 based on lens magnification (col. [2]) and from other measurements (col. [3]). Masses are given in units of $10^{15}\ h^{-1}\ M_{\odot}$, and velocities are quoted in units of km s^{-1}. Distance are given in h^{-1} Mpc. The other measurements are based on arclets (Tyson & Fischer 1995), the shear pattern (Kaiser 1996), X-ray temperatures (Yamashita 1994), and line-of-sight velocity dispersion (Teague et al. 1990). Also given are the three-dimensional mass estimates from lens magnification.

tion (§ 5.5.2). Compared to our two-dimensional mass, Kaiser's analysis suggests that

$$M_{2D}(<0.28) = (0.43 \pm 0.04) \times 10^{15}\ h^{-1}\ M_{\odot}\ , \quad (37)$$

again in good agreement with that found by the magnification method.

6.2. *The Three-dimensional Mass Estimated from Lensing Alone*

The three-dimensional mass inferred from the two-dimensional projected mass inside a sphere of radius $r = 0.5\ h^{-1}$Mpc is

$$M_{3D}(<0.5) = (0.72 \pm 0.25) \times 10^{15}\ h^{-1}\ M_{\odot}\ , \quad (38)$$

while the mass inside an Abell radius, $r = 1.5\ h^{-1}$Mpc, is

$$M_{3D}(<1.5) = (1.6 \pm 0.6) \times 10^{15}\ h^{-1}\ M_{\odot}\ . \quad (39)$$

These estimates are probably an overestimate of the true three-dimensional mass, since the dispersion in the simulations includes the effect of the alignment of the clusters' principle axis along the line of sight. Given that the inferred three-dimensional mass is so high, A1689 is probably lying at the extreme of such a distribution. In such cases, the three-dimensional mass may be much lower than mass inferred from a two-dimensional projection. We discuss this possibility in the next few sections.

6.3. *Velocity Dispersion of Abell 1689*

The predicted line-of-sight velocity dispersion estimated from the simulations includes the effects of superposition of clusters, infall along filaments, and interlopers, and so tends to predict larger velocities and larger uncertainties than for an isolated cluster. Including these effects into our estimate for Abell 1689, we find

$$\sigma_v(<1.5\ h^{-1}\ \mathrm{Mpc}) = 2200 \pm 500\ \mathrm{km\ s^{-1}} \quad (40)$$

for the line-of-sight velocity dispersion inferred from the two-dimensional lensing mass. A measurement that may also include these effects is given by Teague et al. (1990), who find

$$\sigma_v(<1.5\ h^{-1}\ \mathrm{Mpc}) = 2355^{+238}_{-183}\ \mathrm{km\ s^{-1}}\ , \quad (41)$$

in good agreement with our model. However, both of these values are very high, much higher than the estimate for an isolated isothermal sphere, which for A1689 gives a velocity dispersion of 1645 ± 148 km s^{-1}. This discrepency between lensing mass and the velocity dispersion suggests that A1689 is not a single isolated cluster, but a superposition of smaller clumps that contribute to the total measured velocity dispersion. Den Hartog & Katgert (1996) have tried to take into account interlopers in A1689 and, using the Teague et al. data, find a value of $\sigma_v = 1861$ km s^{-1}.

Following a suggestion of Miralda-Escudé & Babul (1995), we shall assume that A1689 is composed of two superposed isothermal spheres. Placing one cluster at $z = 0.18$ with a velocity dispersion of 1500 km s^{-1} and a second at $z = 0.20$ with a velocity dispersion of 750 km s^{-1}, we find that we can reproduce a total projected velocity dispersion of around 2300 km s^{-1}, in agreement with both observed and simulated values. Figures 4 and 5 of Teague et al. (1990) also provide marginal evidence for a second concentration of galaxies at $z = 0.2$. Furthermore, if we estimate the integrated surface mass of these two clusters,

$$M_{2D}(<R) = 7.38 \times 10^{14} \sigma_{1000}^2 \left(\frac{R}{h^{-1}\ \mathrm{Mpc}}\right) h^{-1}\ M_{\odot}\ , \quad (42)$$

where $\sigma_{1000} = \sigma_v/1000$ km s^{-1}, we reproduce a lensing mass of $M_{2D}(<R) = 2 \times 10^{15}(R/h^{-1}\ \mathrm{Mpc})\ h^{-1}\ M_{\odot}$, in agreement with what we see from lensing. Hence it seems plausible that the lensing mass and velocity dispersion of A1689 can both be explained by a superposition of a rich and a poor cluster.

6.4. *X-Ray Mass Estimates of Abell 1689*

Evrard, Metzler, & Navarro (1996) have found that the mass within the radius defined where the mean cluster density is 500 times the critical density is strongly correlated with the cluster temperature. They fitted this relation from simulations with

$$M_{500} = 1.11 \times 10^{15} \left(\frac{T_X}{10\ \mathrm{keV}}\right)^{3/2} h^{-1}\ M_{\odot}\ , \quad (43)$$

where T_X is the broad-beam temperature, and M_{500} is the three-dimensional mass within a radius defined by an overdensity $500\rho_{\mathrm{crit}}$. This radius is roughly given by $r_{500} = 1.175\ h^{-1}$ Mpc.

X-ray temperatures of A1689 have been measured by both *Ginga* and *ASCA*. Yamashita (1994) has analyzed these data and finds $T = 9 \pm 1$ keV, while Mushotzky & Scharf (1997) find $T = 9.02^{+0.4}_{-0.3}$ keV. Daines et al. (1997) have also recently reanalyzed *ROSAT* PSPC observations and find a mean temperature of $T_X = 10.2 \pm 4$ keV. Note that we are quoting the mean temperature and incorporated the 40% uncertainty in the error estimate, rather than quoting upper limits as Daines et al. do. The major uncertainty in measuring X-ray temperatures here is instrumental, as 10 keV is approaching the limit of *ROSAT*'s sensitivity.

Taking the result of Yamashita and the relation found by Evrard et al., we find that

$$M_{500} = (0.95 \pm 0.16) \times 10^{15}\ h^{-1}\ M_{\odot}\ . \quad (44)$$

Using the simulated scaling relations, we find

$$M_{500} = (1.6 \pm 0.65) \times 10^{15}\ h^{-1}\ M_{\odot} \quad (45)$$

for Abell 1689, implying an X-ray temperature of $T_X = 12.7 \pm 3.4$ keV, within the 1 σ uncertainty of the measured X-ray temperature. Again, if we consider A1689 as a double cluster, the nearer, larger mass concentration would be detected in X-ray, lowering the expected X-ray temperature. From the velocity dispersions we can infer a temperature nearer to $T_X = 0.7$ keV, slightly below, but again in agreement with observations.

In conclusion, although we find a high mass, there is a general consistency between the mass of A1689 estimated from lens magnification and shear. In addition, we find a fair agreement between the lens mass and the line-of-sight velocity dispersion if we take into account projection effects. Modeling A1689 as a double cluster, we find that the velocity dispersion can be much lower, implying two smaller clusters, with the lensing mass a superposition of cluster masses. This hypothesis might also help explain the marginal discrepancy with X-ray temperature.

Finally, A1689 is in projection a highly spherical cluster, in contrast with the majority of clusters, which appear

extended. While this may be a result of its high mass, it is also possible that A1689 has its major axis aligned along the line of sight, pointing toward a second cluster. While much of the evidence on the mass distribution along the line of sight is circumstantial, all of these effects would conspire to give A1689 its impressively massive appearance.

7. DISCUSSION

The absolute surface mass density of a galaxy cluster can be estimated from the magnification effect on a background population of galaxies, breaking the "sheet-mass" degeneracy. To apply this in practice, we have taken into account the nonlinear clustering of the background population and shot noise, both of which contribute to uncertainties in the lensing signal (Taylor & Dye 1998). A further complication is the contribution of shear to the magnification in the strong lensing regime, where the magnification signal is stronger. We have argued that this can be circumvented by approximate methods that can be local, where a relationship between surface mass and shear is assumed (van Kampen 1998), or by a nonlocal approximation where only the shape of the cluster is assumed. Both approximations seem to work well on simulated data.

We have applied these methods to the lensing cluster Abell 1689, using Keck data of Smail et al. (1995) to normalize the background counts and the CFRS results to infer the redshift distribution and clustering properties of our data. Using a $\gamma = 0$ approximation of the surface density in the strong lensing regime, we have reconstructed a two-dimensional mass map for A1689 in the innermost 27 arcmin^2, from where a substantial part of the lensing signal comes. The two-dimensional map has general features that are similar to those seen from shear maps (Kaiser 1996). This is encouraging for both methods, as they are independent determinations of the mass distribution.

For a more quantitative measure, we have binned the data in annuli around the peak in the light distribution and found a significant (5 σ) drop in the number counts, dropping to zero where a caustic is inferred from arcs. Local and nonlocal approximations were used to find the κ profile from the number counts and estimate the shear field. We found these to be quantitatively similar to that found by the shear method.

We have also discussed the possibility of a second population of background galaxies, creating a second dip in the radial number counts and a spike in the mass profile. However, we argued that it is unlikely that there is a second low-z population, as the cluster mass would be improbably high, and if there is a high-z population, it has little effect on our results.

We have calculated a cumulative mass profile for A1689 and find a projected two-dimensional cumulative mass of

$$M_{2D}(<0.24\ h^{-1}\ \mathrm{Mpc}) = (0.50 \pm 0.09) \times 10^{15}\ h^{-1}\ M_{\odot}\ . \quad (46)$$

Such a large mass is very rare in a CDM universe normalized to the observed cluster abundance, and may indicate that A1689 is composed of two large masses along the line of sight and/or filaments connected to the cluster and aligned along the line of sight. This is also implied by the high line-of-sight velocity dispersion, which would be enhanced by merging clusters (Miralda-Escudé & Babul 1995) or by infall from aligned filaments.

We have compared our mass estimates to other estimates available in the literature and find that the lens magnification, shear, arclets, line-of-sight velocity dispersions, and the X-ray temperature mass estimates are all in reasonable agreement, to within the uncertainties at this time.

The results presented here are from 3 hours integration on the 3.6 m NTT. Longer integration times have the combined benefit of increasing the number of background galaxies, and so reducing shot noise, and of reducing the contribution from cosmic variance (eq. [20]; § 5.1.3). Hence, by increasing the exposure time, we can expect to reduce the uncertainty from lens magnification by a factor of 2 or so.

One drawback of this analysis is the contribution of clustering noise to the background counts. This can be removed using redshift information, either from spectroscopy or more efficiently, using photometric redshift information (BTP). We shall explore this elsewhere (Dye et al. 1998).

If our results are extended to other clusters, we can hope to have a good representation of the total mass distribution, gas, and galaxy contents with which to make strong statistical arguments about the matter content of the largest gravitationally collapsed structures in the universe.

A. N. T. thanks the PPARC for a research associateship and the University of Berkeley and the Theoretical Astrophysics Center, Copenhagen for their hospitality during the writing of this paper. E. v. K. acknowledges an European Community Research Fellowship as part of the HCM programme and thanks the University of Edinburgh and ROE for their hospitality. This work was supported in part by Danmarks Grundforskningsfond through its funding of the Theoretical Astrophysics Center. S. D. thanks the PPARC for a studentship. N. B. L. thanks the Spanish MEC for a Ph.D. scholarship, the University of Berkeley for their hospitality and financial support from the Spanish DGES, project PB 95-0041. We thank Ian Smail, who kindly provided us with a copy of his Keck data. We also thank John Peacock, Alan Heavens, Jens Hjorth, Adrian Webster, and an anonymous referee for comments and useful suggestions.

REFERENCES

Bartelmann, M., & Kolatt, T. S. 1998, MNRAS, submitted
Bartelmann, M., Steinmetz, M., & Weiss, A. 1995, A&A, 297, 1
Benitez, N., et al. 1998, in preparation
Broadhurst, T. J. 1995, in AIP Conf. Proc. 336, Dark Matter, ed. S. S. Holt & C. L. Bennett (New York: AIP)
Broadhurst, T. J., Taylor, A. N., & Peacock, J. A. 1995, ApJ, 438, 49
Coles, P., & Jones, B. 1991, MNRAS, 248, 1
Crampton, D., Le Fèvre, O., Lilly, S. J., & Hammer, F. 1995, ApJ, 455, 96
Daines, S., Jones, C., Forman, W., & Tyson, J. A. 1998, ApJ, submitted
den Hartog, R., & Katgert, P. 1996, MNRAS, 279, 349
Dye, S., et al. 1998, in preparation
Evrard, A. E., Metzler, C. A., & Navarro, J. F. 1996, ApJ, 469, 494
Fort, B., Mellier, Y., & Dantel-Fort, M. 1997, A&A, 321, 353
Kaiser, N. 1996, in Proc. 36th Herstmonceux Conf., Gravitational Dynamics, ed. O. Lahav, E. Terlevich, & R. J. Terlevich (Cambridge: Cambridge Univ. Press)
Kaiser, N., & Squires, G. 1993, ApJ, 404, 441
Lilly, S., Tresse, L., Hammer, F., Crampton, D., & Le Fèvre, O. 1995, ApJ, 455, 108
Le Fèvre, O., Hudon, D., Lilly, S. J., Crampton, D., Hammer, F., & Tresse, L. 1996, ApJ, 461, 534
Miralda-Escudé, J., & Babul, A. 1995, ApJ, 449, 18
Mushotzky, R. F., & Scharf, C. A. 1997, ApJ, 482, L13
Schneider, P., & Seitz, C. 1995, A&A, 294, 411
Smail, I., Hogg, D. W., Yan, L., & Cohen, J. G. 1995, ApJ, 449, L105
Taylor, A. N., & Dye, S. 1998, MNRAS, submitted
Teague, P. F., Carter, D., & Gray, P. M. 1990, ApJS, 72, 715
Tyson, J. A., & Fischer, P. 1995, ApJ, 446, L55
Tyson, J. A., Valdes, F., & Wenk, R. A. 1990, ApJ, 349, L1
van Kampen, E. 1998, MNRAS, submitted
van Kampen, E., & Katgert, P. 1997, MNRAS, 209, 327
Yamashita, K. 1994, in Clusters of Galaxies, ed. F. Durret, A. Mazure, & J. T. V. Tran (Gif-sur-Yvette, France: Editions Frontières), 153