Gravitational Lensing as a Probe of Dark Matter, the Distance Scale, and Gravitational Waves in the Universe

by

Rennan Bar-Kana

B.A. Physics and Mathematics, University of Pennsylvania (1993)

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1997

Author .....
Department of Physics
May 2, 1997

Certified by .....
Edmund Bertschinger
Professor of Physics
Thesis Supervisor

Accepted by .....
George F. Koster
Chairman, Physics Graduate Committee

Gravitational Lensing as a Probe of Dark Matter, the Distance Scale, and Gravitational Waves in the Universe

by

Rennan Bar-Kana

Submitted to the Department of Physics
on May 2, 1997, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

We study the theoretical basis as well as a number of observational applications of gravitational lensing. We begin by deriving the general equations governing light propagation using a formulation of general relativity in terms of Hamiltonian dynamics. We specialize these equations to the case of a deflector such as a galaxy, and show how to obtain the standard thin-lens equations of lensing. We then analyze two applications where the more general equations must be used.

The first of these is the deflection of light due to large-scale structure on supergalactic scales. This deflection is typically small, but we show that when present in addition to a lensing galaxy it can have substantial effects. This includes an effect on a major application of lenses; we find that large-scale structure induces an uncertainty in determinations of the Hubble constant, typically of order $5\% - 10\%$. The second application of the equations of motion is to lensing due to gravitational waves. We consider several effects of gravitational waves on lenses but conclude that they are small and unlikely to be observable.

We then turn our attention to a couple observed multiply-imaged quasars, the four-image system PG1115+080 and the double 0957+561. We analyze recently measured intensity variations in the images of PG1115+080 in order to determine relative delays in the arrival times of light among the different image paths. We find a value of $25.0^{+3.3}_{-3.8}$ days (95% confidence) for the largest delay, but an accurate determination of the Hubble constant must await further constraints on lens models. We also find evidence for a magnification gradient across the source, which suggests that the optical emission region of quasars is about 1000 AU. In contrast to PG1115+080, lens models of the double quasar 0957+561 are well constrained due to high-resolution radio observations that detect six source components. We fit a number of models to this data and obtain a Hubble constant of 55 km $\sec^{-1}$ Mpc^{-1}, with a 20% uncertainty at 95% confidence.

Thesis Supervisor: Edmund Bertschinger
Title: Professor of Physics

Acknowledgments

First, I thank my advisor, Prof. Ed Bertschinger, for enriching my knowledge and guiding me in the field of cosmology. His breadth of knowledge and interests allowed me to learn about many different aspects of the field. I thank him for encouraging me to work independently, which prepared me for the future of my career, while at the same time he was always generous with insightful discussions and comments.

I thank Prof. Paul Schechter for initiating my interest in observations of lensing, thus allowing me to experience first-hand the rewards of combining theory with raw data. His unique perspective on the field made conversations always rewarding. I also thank Prof. Jackie Hewitt, the third member of my Thesis committee, for illuminating discussions on various aspects of lensing.

I am grateful to Prof. Alan Guth, whose generous efforts on my behalf helped bring me to MIT. Outside MIT, Prof. Chris Kochanek of the CfA had a profound impact on this thesis, through his scientific work as well as a number of conversations. I continue to be inspired by my undergraduate advisor, Prof. Paul Steinhardt of Penn, who was responsible for bringing me into the field of cosmology.

I have enjoyed working with the students and postdocs in the cosmology group at MIT, especially Matias Zaldarriaga, whose opinions on science and life are always intriguing. I am grateful to Uros Seljak, whose thesis work was a major inspiration for my own work.

Last but certainly not least, I thank my most loving family: My dad for always telling me that I can do it, my mom for providing mental support as well as actual sustenance, and my brother for keeping up the brotherly love despite the intercontinental distance.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The gravitational deflection of light from a source (usually a quasar) by a foreground object (e.g. a galaxy or cluster) can lead to the formation of multiple images of the source. The theory which describes this gravitational lensing is based on a standard "thin-lens" approximation, in which the deflection of light rays is approximated as occuring over a very short distance compared to the total observer-source distance. There are, however, a number of situations which cannot be described by this simple theory.

Gravitational lensing is a special case of the propagation of light in a universe which is homogeneous on a large-scale average, but contains small perturbations to the metric, produced e.g. by galaxies, clusters, and larger structures. In this thesis we apply the well-developed formalism for describing perturbed cosmological models to the problem of gravitational lensing, producing an alternative derivation of the thin-lens description as well as a study of a number of problems whose treatment requires the more general formalism.

Gravitational lensing is also one of the most promising methods of mapping the distribution of matter at cosmological distances. Detailed observations of multiple images of quasars have been used to try to reconstruct the lensing mass distribution. It has also long been recognized that measurements of the time delay between images allow a direct measurement of cosmological distances. In this thesis we analyze the two observed gravitational lenses which have thus yielded measurements of the Hubble constant.

1.1 Gravitational lensing theory

Figure 1-1 shows a standard gravitational lensing diagram, where a light ray leaves the source, is deflected by a lensing matter distribution, and reaches the observer. Since in practice a galaxy deflects light over a short distance compared to the cosmological distances between the observer, lens and source, we can use a thin-lens approximation where light travels in a straight line from the source to the lens, is deflected instantaneously at the lens plane, and then once again travels in a straight line to the observer. Assuming small angles, in a flat universe we can use simple

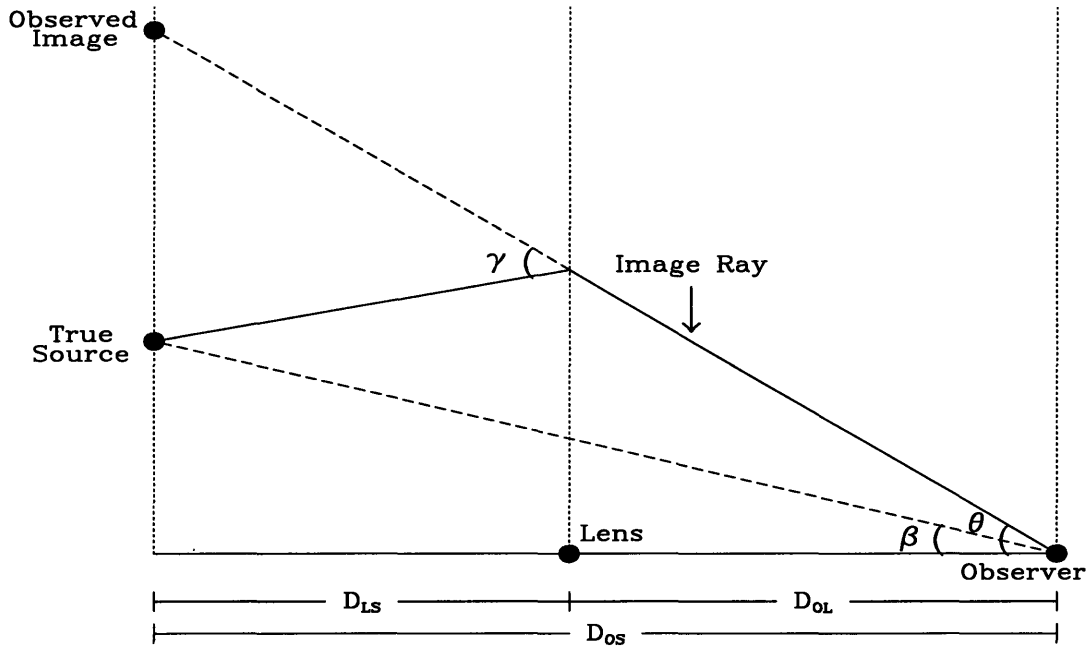

Figure 1-1: Gravitational lens diagram showing an image ray, the image angle θ, source angle β, deflection angle γ, and several angular diameter distances.

geometry to derive the lens equation,

$$\vec{\beta} = \vec{\theta} - \frac{D_{LS}}{D_{OS}}\vec{\gamma} \, . \tag{1.1}$$

In the Newtonian calculation, the deflection angle is given by integrating the deflection along the path. In general relativity, space-time curvature is described by a metric which gives the distances and times measured by observers. In the weak field limit of general relativity, the metric is expressed as $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, where $\eta_{\mu\nu}$ is the flat metric of special relativity,

$$ds^2 = [-c^2 d\tau^2 + d\vec{x} \cdot d\vec{x}\,] \, , \tag{1.2}$$

and Einstein's theory is linearized in the small perturbation $h_{\mu\nu}$. The resulting Einstein equations assume a form similar to Maxwell's equations. Just as in electromagnetism, there are static solutions as well as wave solutions. For a static, Newtonian source the metric is found to be

$$ds^2 = [-(1+2\phi)c^2 d\tau^2 + (1-2\phi) d\vec{x} \cdot d\vec{x}\,] \, , \tag{1.3}$$

in terms of the Newtonian potential ϕ. The geodesic equations in this metric for a photon yield the same equations as for a Newtonian particle but with the deflection

due to a point mass doubled to (Schneider et al. 1992)

$$\vec{\gamma} = \frac{4GM}{c^2}\frac{\vec{\eta}}{\eta^2} , \tag{1.4}$$

where $\vec{\eta} = D_{OL}\vec{\theta}$ measures distance in the lens plane. Under the assumption of an instantaneous deflection, the impact parameter of the light ray in the lens plane is fixed for a given $\vec{\theta}$. Thus for a distribution of matter in the lens plane, the total deflection is the superposition of the deflections from individual mass elements,

$$\vec{\gamma} = \frac{4G}{c^2}\int d^2\vec{\eta}'\ \Sigma(\vec{\eta}')\ \frac{\vec{\eta}-\vec{\eta}'}{|\vec{\eta}-\vec{\eta}'|^2} , \tag{1.5}$$

where Σ is the projected surface mass density of the lens. This essential quantity is given by the volume mass density projected onto the lens plane, since a light ray has approximately the same impact parameter with respect to each such column of mass perpendicular to the lens plane. The deflection angle can be derived from a two dimensional potential,

$$\Psi(\vec{\eta}) = \frac{4G}{c^2}\int d^2\vec{\eta}'\ \Sigma(\vec{\eta}')\ \ln\left(\frac{|\vec{\eta}-\vec{\eta}'|}{D_{OL}}\right) , \tag{1.6}$$

where $\vec{\gamma} = \partial_{\vec{\eta}}\Psi$, and $\nabla^2_\eta\Psi = 8\pi G\Sigma/c^2$. We can define dimensionless variables, $\vec{\alpha} = D_{LS}\vec{\gamma}/D_{OS}$ and $\psi = D_{LS}\Psi/(D_{OL}D_{OS})$. Then the lens equation becomes

$$\vec{\beta} = \vec{\theta} - \vec{\alpha} , \tag{1.7}$$

and the potential satisfies

$$\nabla^2_\theta\psi = 2\frac{\Sigma}{\Sigma_c} , \tag{1.8}$$

in terms of the critical density $\Sigma_c = c^2D_{OS}/(4\pi GD_{OL}D_{LS})$.

We have so far neglected the cosmological expansion of the universe and the possibility of spatial curvature. The metric can be generalized to include these cases, and using general relativity it can be shown (e.g. Schneider et al. 1992) that all the equations above (except for (1.2) and (1.3)) retain the same form if all distances are interpreted as angular diameter distances. This means e.g. that $D_{OL}\vec{\theta}$ is the measured distance in the lens plane corresponding to an angle $\vec{\theta}$ as measured by the observer. In the special case of an Einstein-de Sitter universe, the angular diameter distance between redshift z_i and $z_j > z_i$ is given by

$$D_{ij} = \frac{2c}{H_0}\frac{1}{1+z_j}\left(\frac{1}{\sqrt{1+z_i}} - \frac{1}{\sqrt{1+z_j}}\right) , \tag{1.9}$$

where H_0 is the Hubble constant.

Lensing leads to a magnification of the source. While lensing conserves surface brightness, the area subtended by an image at the observer can be greater than it

would be in the absence of the lens. The magnification is given by this mapping of the area or solid angle,

$$\mu = \det \left| \frac{\partial \vec{\beta}}{\partial \vec{\theta}} \right|^{-1} . \tag{1.10}$$

The lens potential also delays the arrival time of a given light ray. This delay has a geometrical part, due to the longer path traveled, and a gravitational part, the "Shapiro time delay" which describes the slowing down of clocks in gravitational fields. Up to a constant, for small angles the time delay is given by

$$\Delta t = \frac{1+z_L}{c} \frac{D_{OL} D_{OS}}{D_{LS}} \left[\frac{1}{2} (\vec{\theta} - \vec{\beta})^2 - \psi(\vec{\theta}) \right] . \tag{1.11}$$

Fermat's principle implies that light travels on the path of stationary arrival time, and indeed if we set $\partial \Delta t / \partial \vec{\theta} = 0$ at fixed $\vec{\beta}$, we recover the lens equation.

For a given deflection law $\vec{\alpha}(\vec{\theta})$, we can solve the lens equation for each source position and find the multiple image positions. In an observed lens we can measure the various image positions and their relative magnifications, fit a model and thus determine the relative time delays up to the overall factor in equation (1.11). If the lens and source redshifts are measured then this factor depends only on the Hubble constant (and weakly on other cosmological parameters such as the mean density Ω). If the source intensity varies, by monitoring the images we can observe these relative time delays and measure the Hubble constant, as first envisioned by Refsdal (1964, 1966).

The most common formalism used for deriving these basic equations of gravitational lensing is based on Fermat's principle. This approach is developed fully in Schneider et al. 1992; A nice condensed version of gravitational lensing theory is given in Blandford & Kochanek 1989. In chapter 2 of this thesis we present an original derivation of these basic equations of gravitational lensing. We start from the Hamiltonian description of particle (or photon) dynamics. In this presentation, the treatment of the relativistic case is very similar to the classical, Newtonian one. Furthermore, the generalization to an expanding universe is straightforward with this method. We use the equations of motion directly to find both the light path and the time delay, which simplifies some of the key arguments in the derivation of the thin-lens approximation.

1.2 Large-scale structure (LSS)

Our alternative derivation of the equations of lensing also clarifies the precise approximations needed for the thin-lens equations to be valid. When light rays are deflected continuously throughout the path from the source to the observer, more general equations of motion must be used. One example is the deflection due to large-scale inhomogeneities in the universe, accumulated as light rays traverse the cosmological distance from the source to the observer. LSS can be described with a metric similar to the weak field limit given above, except for a scale factor which

accounts for the expansion of the universe. In the longitudinal gauge (Bardeen 1980), the line element is

$$ds^2 = a^2(\tau)[-(1+2\phi)d\tau^2 + (1-2\phi)d\vec{x}\cdot d\vec{x}\ ]\ , \tag{1.12}$$

where we set c = 1. Here τ is the conformal time, related to physical time t through $\tau = \int(dt/a)$, $a(\tau)$ is the expansion factor, and we are using comoving coordinates $\vec{x}$. Redshift in a flat, matter-dominated universe is given by $1/(1+z) = a(\tau) = (H_0\tau/2)^2$. The scalar potential ϕ satisfies the cosmological Poisson equation

$$\nabla^2\phi = 4\pi G a^2 \bar{\rho}\ \delta\ , \tag{1.13}$$

where $\bar{\rho}$ is the mean density of the universe and $\delta = (\rho - \bar{\rho})/\bar{\rho}$ is the local density perturbation. Since we can only try to predict statistical properties of δ and ϕ, the potential is usefully characterized by its Fourier transform $\phi(\vec{k},\tau)$, where $\phi(\vec{x},\tau) = \int d^3k\phi(\vec{k},\tau)e^{i\vec{k}\cdot\vec{x}}$. The potential transform is a random field with ensemble mean and variance $\langle\phi(\vec{k},\tau)\rangle = 0$ and $\langle\phi(\vec{k},\tau)\phi^*(\vec{k}',\tau)\rangle = P_\phi(k,\tau)\delta^3(\vec{k}-\vec{k}')$, where $P_\phi(k,\tau)$ is the power spectrum of the potential at time τ, simply related to the density power spectrum by $P_\phi(k,\tau) = (4\pi G a^2(\tau)\bar{\rho}(\tau))^2\, k^{-4} P_\rho(k,\tau)$. Since the density power spectrum is related to the two point correlation function by a Fourier inverse, the power spectrum can be estimated from galaxy redshift surveys. In addition, inflationary models predict an initially Gaussian density field, whose statistics are completely characterized by the initial power spectrum.

The deflection of light due to LSS is not large enough to produce multiple images, but it may distort the shapes of sources. Such weak lensing has been investigated both analytically (e.g. Miralda-Escudé 1991; Kaiser 1992) and in N-body simulations (Jaroszynski et al. 1991; Blandford et al. 1991). These studies find a shear of order 1%, coherent over a scale of $\approx 1°$, in a flat cold dark matter (CDM) model. This shear may in principle be detected observationally as a coherent distortion of background galaxies, when averaged over a sufficiently large angular field in order to be separated from the random scatter of intrinsic ellipticities (e.g. Mould et al. 1994; Villumsen 1995).

The shear due to LSS can also affect strong lensing (i.e. multiply imaged systems), when it acts in addition to a strong primary lens, a galaxy or cluster near the line of sight to the source. This effect is enhanced compared to weak lensing, because of the small angular separations between multiple images. Also, the higher redshift of quasars compared to faint galaxies increases the cumulative shear from the observer to the source. An important question is whether the effect of LSS may be large enough to influence observed lenses, and in particular their use to measure the Hubble constant.

In chapter 3 we present a detailed calculation of the two primary effects of LSS on an observed lens system. Previously, Seljak (1994) had estimated roughly the magnitude of the effects on multiple images. One effect on lenses is a shear, or asymmetry in the deflecting gravitational potential. Formally, we can understand shear as resulting from an expansion of the deflection angle due to LSS. Since LSS varies on a large scale compared with typical image separations, it is meaningful to

expand $\vec{\alpha}(\vec{\theta})$ in powers of $\vec{\theta}$. The first term is a constant, unobservable deflection. The next term is linear in $\vec{\theta}$, $\alpha_i = M_{ij}\theta^j$, with a sum over j. The traceless part of the 2×2 matrix M_{ij} is the shear.

The most important effect of shear is in producing four-image systems (quads). Many confirmed lens systems are quads, since they are easy to identify and tend to be highly magnified (Kochanek 1991b, 1995; Wallington & Narayan 1993). These systems are inconsistent with an axi-symmetric lens, for which all the images would have to be colinear. Lens models of quads typically find a shear of order $7\% - 11\%$ (e.g. Kochanek 1991a). If due to the lensing galaxy itself, this would imply a projected ellipticity (= 1 minus the ratio of minor to major axis) for the mass of $\approx 35\% - 50\%$. By contrast, the typical value observed for ellipticals is $\approx 20\%$ (e.g. Ryden 1992; Schechter 1987). Since the cross section for producing quads increases with shear, observed quads should be biased towards high shear, whatever its origin (Kassiola & Kovner 1993). Recent observations of a four-image "Einstein cross" with the Hubble Space Telescope WFPC2 (Ratnatunga et al. 1995) found an ellipticity in the potential of 26%, which implies a mass ellipticity of 60%. The light distribution was found to have an ellipticity of only 32%. One possible explanation is that the dark matter halo is highly flattened compared to the light distribution. Other galaxies or clusters near the line of sight to the source may contribute to this shear, as can also LSS.

The other effect, corresponding to the trace of M_{ij}, is a convergence or divergence of image rays, which leaves the entire lens geometry unaffected except for an overall change of scale. This can be seen from the time delay corresponding to a pure trace, which is proportional to θ^2. Since the geometrical time delay is also quadratic in θ, any trace term can be absorbed by a change in the overall coefficient, i.e. a change in distance. Since observed time delays between images are used to infer the overall distance scale and thus the Hubble constant, a trace term affects the inferred Hubble constant. This effect has been analyzed for a single deflector (Falco et al. 1985; Narayan 1991). Adding up statistically the effect of LSS, we find an induced uncertainty in determinations of the Hubble constant of order $5\% - 10\%$ rms, depending on the lens and source redshifts, and on the adopted LSS power spectrum.

While the rms effect of LSS can be derived analytically, the probability distribution along different lines of sight is non-Gaussian, and can be investigated numerically by shooting rays through N-body simulations of LSS (Wambsganss et al. 1997 is a recent example). In such investigations it is common not to compute the path of each light ray continuously, but instead to use the multiple lensing approximation of a series of individual deflections, each calculated with the thin-lens equations. Using the exact equations of motion, in chapter 2 we find conditions under which multiple lensing is valid. If the correlation between rays separated by an angle θ is desired for sources at a comoving distance r, with LSS on a scale d dominating the deflection, then a number of lens planes $N \gg \theta r/d$ is typically required, and this is satisfied in most applications with $N \sim 100$.

1.3 Gravitational waves (GW's)

Another case where continuous deflections occur is if the universe is filled with a stochastic background of gravitational waves. Gravitational waves are an elusive prediction of general relativity, and their existence has only been indirectly confirmed. Since gravity wave detectors such as LIGO are being built, any astrophysical phenomenon which can put a limit on the amplitude of existing gravity waves is of great interest.

Gravity waves are described by the same weak field limit of general relativity noted above. They can be derived by searching for wave solutions of the linearized field equations. Gravitational waves can be expressed as a transverse, traceless tensor. The waves propagate at the speed of light, and the non-zero components are transverse to the direction of propagation. There are two possible polarizations, denoted + and $\times$. For example, for a wave propagating in the z-direction, we have

$$h_{\mu\nu} = A_{+} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + A_{\times} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

where the row order is x, y, z, and t. The physical effects of gravity waves can be found using the geodesic equations. The separation of two nearby test particles, as measured in the reference frame of one of them, oscillates when a gravity wave travels by them. Only the component of the separation transverse to the wave direction is affected.

Gravitational waves are important because they carry energy and momentum, and they are generated by any matter distribution which has a changing quadrupole moment. Generating strong gravity waves requires very compact bodies moving relativistically, and some possible sources are binary systems which include a neutron star or black hole, or a period of inflation in the early universe. Gravity waves have never been detected directly, only indirectly from binary pulsar timing. Measurements on the famous Hulse-Taylor pulsar PSR 1913+16 confirmed a number of relativistic effects (e.g. Weisberg & Taylor 1984). Since the force law is not exactly $1/r^2$, there is a precession of the perihelion at a rate of 4.2 degrees per year. Also as the pulsar orbits around its companion, gravitational waves are radiated, the orbit loses energy, and the orbital period shortens in accordance with general relativity, for the measured orbital parameters. A number of projects currently under way will attempt to pioneer the field of gravity wave astronomy. Figure 1-2, taken from Bar-Kana (1994), shows a comparison of experimental limits with several theoretical spectra, in terms of the dimensionless strain which is closely related to the gravity wave amplitude. The curve marked $n_T = 0$ is a scale-invariant spectrum normalized to the large-scale Cosmic Microwave Background anisotropy as measured by the COBE satellite. This curve is an upper limit to most models of inflation. The other curves cannot be produced by inflation, but could still exist given the present limits from pulsar timing (upper left corner). Also shown are several projected experimental limits. LIGO is

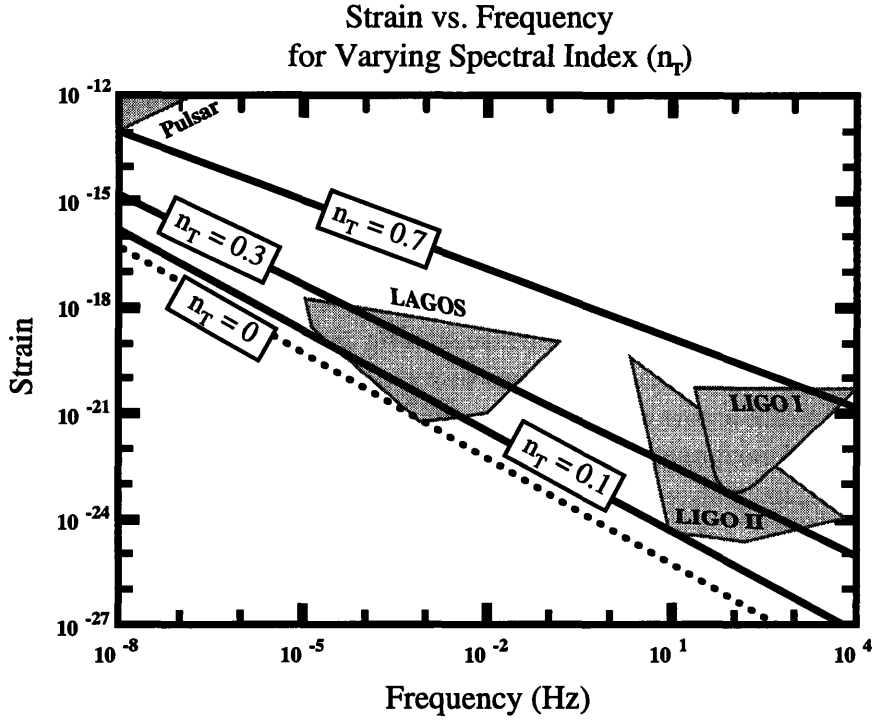

Figure 1-2: Comparison of current and projected limits on gravity waves with theoretical spectra.

the Laser Interferometer Gravitational Wave Observatory, currently being built by an MIT-Caltech collaboration. LIGO I refers to the interferometer at initial operation, and LIGO II is the expected sensitivity after several years of operation and improvement. LAGOS (also called LISA) is a laser interferometer in space which is planned for the future.

Since gravity waves are a perturbation of the metric, they can produce lensing: Light rays are bent in the perturbed metric, and this could have several observable effects. One of these is a shear, analogous to LSS. Another is proper motion, where a fixed source appears to move, since its light rays are bent as gravity waves go by. We show in chapter 4 that the limit on gravity waves from inflation, as derived from their lensing effects, is weaker than the limit from their contribution to the anisotropy of the Cosmic Microwave Background, by several orders of magnitude. A quantitative estimate of these effects due to more general gravity wave spectra (i.e. not necessarily from inflation) also fails to yield an interesting constraint, compared to previous limits. This physically interesting phenomenon turns out not to be observationally accessible at the present.

1.4 Lens observations and the Hubble constant

Lenses can be used to measure the size of the universe. As outlined above, the idea is to observe the angular positions of images and their relative magnifications, and then

to trace the full trajectories of the light rays by fitting a model for the lens to the data. Since the delays in the arrival times among the images depend on the distances to the lens and source, and so on the Hubble constant (for given redshifts), measuring the delays allows a determination of the Hubble constant.

Since each lens can potentially yield an independent, single step H_0 measurement, obtaining values in a number of lenses with a variety of morphologies and constraints on models may help eliminate systematic errors. Observational constraints on the lens model may include precise image positions and flux ratios as well as data on the objects responsible for the lensing. Resolved structure in images as revealed by radio interferometry provides many constraints, since it is essentially the same as observing multiple sources with the same lens. Whether lensing involves a galaxy, group, or cluster, the position, velocity dispersion, and other observational probes of the mass distribution of the lensing objects all yield constraints on lens models. Measuring time delays requires observing time variations in the image intensities with sufficient accuracy and time resolution. As the number of lenses being carefully monitored increases, more time delays are being determined, such as the preliminary measurement in PKS 1830-211 (Van Ommen et al. 1995). Flux measurements are not the only possibility, as shown by the promising measurements of variations in polarization fraction in the images of B 0218+357 (Corbett et al. 1996).

The quadruply imaged quasar PG1115+080 was the second lens to be discovered (Weymann et al. 1980). It is radio quiet, but optical Hubble Space Telescope images (Kristian et al. 1993) were recently analyzed by Schechter et al. (1997) and by Keeton & Kochanek (1997). They found that lens models which include the effect of the lens galaxy and that of the nearby group of galaxies discovered by Young et al. (1981) can fit the image positions well. They however still found great freedom in the H_0 values predicted by these lens models for a given time delay. This freedom results in part from a degeneracy between the radial profile of the lens mass and the inferred H_0, which is typical of symmetric four-image configurations (Kochanek 1991a; Wambsganss & Paczyński 1994) In PG1115+080 future observations of the lensing galaxy light profile and, ultimately, a direct measurement of its central velocity dispersion may constrain or eliminate this degeneracy.

Schechter et al. observed intensity variations in the images of PG1115, and used the method of Press et al. (1992) to determine the best-fitting time delays among the three light curves A, B, and C (the two A images are close together and were combined). They found a Hubble constant of ~ 40 km/s/Mpc from the largest (B-C) delay, and a ratio of $|t_{AC}/t_{AB}| \sim 0.7$ for the two smaller delays. This ratio is independent of overall distance, and so the measured ratio can be compared directly with that predicted by lens models. In chapter 5 of this thesis, I present a new method for determining time delays among the light curves of various images in a gravitational lens. The method is based on constructing a simple model for the source variation and forming a χ^2 measure of the agreement of this same variation with all of the lightcurves. This method provides a simpler and more assumption-free alternative to Press et al., but as a method based on χ^2 minimization it has similar advantages: first, it yields an approximate reconstruction of the source variation and of other parameters such as relative time delays; second, it easily incorporates different assumptions about the

relation between the light curves and about the data measurement errors; finally, it can be applied to more than two light curves by addition of χ^2. We apply this method to the light curves of Schechter et al., including correlated measurement errors in the analysis, as well as the possibility that microlensing, the extra magnification of images by individual stars in the lens galaxy, may cause the different images to vary by different factors in flux. We find a value of $25.0^{+3.3}_{-3.8}$ days (95% confidence) for the delay between components B and C (close to the 24 day value of Schechter et al., and so leading to a similar value of the Hubble constant for a given lens model). However, the ratio t_{AC}/t_{BA} of the two intermediate delays is poorly determined at $1.13^{+.18}_{-.17}$ (68% confidence), close to the value predicted by lens models (~ 1.4) unlike the Schechter et al. value (~ 0.7). The variation ratios of C with respect to A and of A with respect to B are both different from 1, $1.39^{+.16}_{-.20}$ and $.79^{+.10}_{-.12}$ (95% confidence), respectively. This is an indication of a microlensing gradient, and this type of microlensing may allow us to set both an upper and a lower limit on the size of the quasar optical emission region.

At present the most promising gravitational lens for measuring the Hubble constant is the double 0957+561, the first lens ever discovered (Walsh et al. 1979). Despite a history of systematic difficulties, recent measurements indicate that a robust measurement of the Hubble constant (H_0) with an accuracy comparable to that of more conventional techniques may be within reach. These measurements include a precise optical time delay (Kundić et al. 1997) which is consistent with the latest results from radio monitoring (Haarsma et al. 1997), and a measurement of the velocity dispersion of the lens galaxy (Falco et al. 1997). However, the poor reduced χ^2 of the best-fitting lens models of Grogin & Narayan (1996a, 1996b, hereafter GN) may indicate that large systematic uncertainties remain in the modeling results. The models are strongly constrained by VLBI data which resolves jets in the two images into several components, which we require a lens model to put in a one-to-one correspondence. These jet components were observed by Garrett et al. (1994), who fit their positions in order to determine derivatives of several components of the relative magnification matrix of the two images. These derivatives were in turn used by GN to constrain lens models. In chapter 6 of this thesis, we reconsider the lens models by comparing them directly to the positions of the jet components, a procedure which has several advantages. It removes a layer of complexity involved in the fitting performed by Garrett et al. in which several degrees of freedom were lost. But perhaps the most important advantage is that the errors on the VLBI positions may reflect the actual measurement errors more accurately than the errors derived on the derivatives of the relative magnification matrix. The assumption in the lens modeling of Gaussian errors may be particularly invalid for these derivatives, since the errors are of order 100% on some of them.

The basic model of GN represents the lens galaxy as a softened power-law sphere, a density profile which allows for both a core radius and an arbitrary radial index. It also approximates the effect of the surrounding cluster as a constant shear term. For this model, if we use the values of 417 ± 3 days for the time delay (Kundić et al. 1997) and 266 ± 12 km s^{-1} for the line-of-site velocity dispersion (Falco et al. 1997),

we obtain for the best-fit model of GN (for which they found a reduced χ^2 of 6.9),

$$H_0 = 59 \pm 14 \text{ km sec}^{-1} \text{Mpc}^{-1} , \tag{1.14}$$

where the errors are 95% confidence limits. For our best fit model, although we use a different lens position with smaller errors, we obtain a more satisfactory reduced χ^2 of 3.4, and a slightly lower

$$H_0 = 55^{+13}_{-12} \text{ km sec}^{-1} \text{Mpc}^{-1} . \tag{1.15}$$

Bibliography

[1] Bar-Kana, R. 1994, Phys. Rev. D, 50, 1157

[2] Bardeen, J. M. 1980, Phys. Rev. D, 22, 1882

[3] Blandford, R. D., & Kochanek, C. S. 1989, in Dark Matter in the Universe, ed. J. Bahcall, T. Piran, & S. Weinberg (Singapore: World Scientific)

[4] Blandford, R. D., Saust, A. B., Brainerd, T. G., & Villumsen, J. V. 1991, MNRAS, 251, 600

[5] Corbett, E. A., Browne, I. W. A., & Wilkinson, P. N. 1996, in Astrophysical Applications of Gravitational Lensing, Proceedings IAU Symposium 173, ed. C. S. Kochanek & J. N. Hewitt (Dordrecht: Kluwer), 37

[6] Falco, E. E., Gorenstein, M.V., & Shapiro, I. I. 1985, ApJ, 289, L1

[7] Falco, E. E., Shapiro, I. I., Moustakas, L. A., & Davis, M. 1997, preprint astro-ph/9702152

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

[9] Grogin, N. A., & Narayan, R. 1996a, ApJ, 464, 92 (GN)

[10] Grogin, N. A., & Narayan, R. 1996b, erratum, ApJ, 473, 570 (GN)

[11] Haarsma, D. B., Hewitt, J. N., Lehár, J., & Burke, B. F. 1997, preprint astro-ph/9607080

[12] Jaroszynski, M., Park, C., Paczynski, B., & Gott, J. R. 1991, ApJ, 365, 22

[13] Kaiser, N. 1992, ApJ, 388, 272

[14] Kassiola, A., & Kovner, I. 1993, ApJ, 417, 450

[15] Keeton, C. R., & Kochanek, C. S. 1997, preprint astro-ph/9611216

[16] Kochanek, C. S. 1991a, ApJ, 373, 354

[17] Kochanek, C. S. 1991b, ApJ, 379, 517

[18] Kochanek, C. S. 1995, ApJ, 453, 545

[19] Kristian, J. et al. 1993, AJ, 106, 1330

[20] Kundić, T., Turner, E. L., Colley, W. N., Gott, J. R., Rhoads, J. E., Wang, Y., Bergeron, L. E., Gloria, K. A., Long, D. C., Malhorta, S., & Wambsganss, J. 1997, preprint astro-ph/9610162

[21] Miralda-Escudé, J. 1991, ApJ, 380, 1

[22] Mould, J., Blandford, R., Villumsen, J., Brainerd, T., Smail, I., Small, T., & Kells, W. 1994, MNRAS, 271, 31

[23] Narayan, R. 1991, ApJ, 378, L5

[24] Press, W. H., Rybicki, G. B., & Hewitt, J. N. 1992, ApJ, 385, 404

[25] Ratnatunga, K. U., Ostrander, E. J., Griffiths, R. E. & Im, M. 1995, ApJ, 453, L5

[26] Refsdal, S. 1964, MNRAS, 128, 307

[27] Refsdal, S. 1966, MNRAS, 132, 101

[28] Ryden, B. S. 1992, ApJ, 396, 445

[29] Schechter, P. L., 1987, IAU Symposium 127, "Elliptical Galaxies," 217

[30] Schechter, P. L., et al. 1997, preprint astro-ph/9611051

[31] Schneider, P., Ehlers, J., & Falco, E.E. 1992, Gravitational Lenses (New York: Springer)

[32] Seljak, U. 1994, ApJ, 436, 509

[33] Weisberg, J. M., & Taylor, J.H. 1984, Phys. Rev. Lett. 52, 1348

[34] Van Ommen, T. D., Jones, D. L., Preston, R. A., & Jauncey, D. L. 1995, ApJ, 444, 561

[35] Villumsen, J. V. 1995, preprint astro-ph/9507007

[36] Wallington, S., & Narayan, R. 1993, ApJ, 403, 517

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

[38] Wambsganss, J., Cen, R., Xu, G., & Ostriker, J. P. 1997, ApJ, 475, 81

[39] Wambsganss, J., & Paczyński, B. 1994, AJ, 108, 1156

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

[41] Young, P., Deverill, R. S., Gunn, J. E., Westphal, J. A., & Kristian, J. 1981, ApJ, 244, 723

Chapter 2

Gravitational Lensing from Hamiltonian Dynamics

2.1 Introduction

The deflection of light by massive bodies is an old problem having few pedagogical treatments. The full machinery of general relativity seems like a sledge hammer when applied to weak gravitational fields. On the other hand, photons are relativistic particles and their propagation over cosmological distances demands more than Newtonian dynamics. In fact, for weak gravitational fields or for small perturbations of a simple cosmological model, it is possible to discuss gravitational lensing in a weak-field limit similar to Newtonian dynamics, albeit with light being deflected twice as much by gravity as a nonrelativistic particle.

The most common formalism used for deriving the equations of gravitational lensing is based on Fermat's principle: light follows paths that minimize the time of arrival. For weak static gravitational fields gravity deflects light as though it travels in a medium with variable index of refraction $n = 1 - 2\phi$ where ϕ is the dimensionless gravitational potential. While the proof of these statements requires general relativity, the application does not. See Schneider, Ehlers & Falco (1992) for a full exposition; see Blandford & Kochanek (1987) for a nice condensed version of gravitational lensing theory.

In this chapter we present an alternative formulation of the theory of gravitational lensing based on Hamiltonian dynamics. Why bother with another approach? Our answer is two-fold. First, students are less familiar with Fermat's principle than they are with Hamilton's principle. It is not trivial to explain why and under what conditions gravity behaves like a variable index of refraction, and from this to justify the least time principle (or vice versa). Second, the extension of the usual lensing formalism to cosmology is somewhat tricky, especially with inhomogeneity all along the line of sight. In the past this has led to rather extreme models of cosmic structure such as the swiss-cheese model in which the spacetime is unperturbed aside from spheres in which the mass is gathered into point masses in the centers (Kantowski 1969). Our approach to gravitational lensing can take advantage of the powerful

formalism for describing perturbed cosmological models that has been in place for more than a decade (Bardeen 1980; Kodama & Sasaki 1984; Bertschinger 1996).

Some work has already been done along at least part of the lines we develop in these notes. This includes papers of Linder (1990), Seljak (1994), and chapter 3 of this thesis.

In §2.2 we introduce our formulation of Hamilton's principle, and demonstrate its equivalence to standard formulations of general relativity. In §2.3 we apply our formulation to derive, in an Einstein-DeSitter universe, the lens equation and time delay in general and for a thin lens, in agreement with, e.g. Schneider et al. In §2.4 we extend these results to the case of an open or closed universe. The results for a thin lens agree with Schneider et al., and those for continuous deflection agree with Seljak (1995). As another example of the flexibility of the Hamiltonian approach, we derive the equations for lensing in a flat universe due to gravitational waves (GW's) in §2.5, in agreement with the equations derived by Linder (1988) from the geodesic equation. We then find the time delay and extend the results to a curved background. Finally, in §2.6 we consider the accuracy of multiple lensing when it is used for ray tracing in N-body simulations, as an approximation to the actual continuous deflection of rays. We include the propagation of various errors as we attempt to find conditions which ensure good accuracy.

Our notations and conventions for general relativity are drawn from Misner, Thorne, & Wheeler (1973). In particular, the metric signature is $(-,+,+,+)$ and Greek indices are four-dimensional.

2.2 Hamilton's principle for particles in general relativity

The trajectories of all particles including photons can be characterized by 4-position as a function of an affine parameter, $x^\mu(\lambda)$. Freely-falling particles follow geodesics of the spacetime metric. The usual derivation of the geodesic equations of motion is based on a Lagrangian prescription: the trajectory is required to make the total proper time (the action) stationary under small variations. Specifically, starting with the action

$$S_1\left[x^\mu(\lambda)\right] = \int \left[g_{\mu\nu}(x)\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\right]^{1/2} d\lambda = \int L_1(x, dx/d\lambda)\, d\lambda \ , \tag{2.1}$$

variation of the trajectory leads to the Euler-Lagrange equations

$$\frac{d}{d\lambda}\left[\frac{\partial L}{\partial\left(dx^\mu/d\lambda\right)}\right] - \frac{\partial L}{\partial x^\mu} = 0 \ , \tag{2.2}$$

from which one obtains the geodesic equation

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu{}_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda} = 0 \ . \tag{2.3}$$

The full derivation can be found in any general relativity text (e.g. Misner, Thorne & Wheeler 1973).

The Lagrangian of equation (2.1) is not unique. Any Lagrangian may be used that yields the same equations of motion. For example, equation (2.3) also follows if one squares the Lagrangian appearing in equation (2.1). This means that we may use a different action functional (Schneider et al. 1992, §3.3)

$$S_2\left[x^\mu(\lambda)\right] = \int \frac{1}{2} g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}\, d\lambda = \int L_2(x, dx/d\lambda)\, d\lambda \ . \tag{2.4}$$

Note that with this Lagrangian the canonical momentum conjugate to x^μ equals the momentum one-form of the particle (up to a constant factor depending on the choice of affine parameter λ):

$$p_\mu \equiv \frac{\partial L}{\partial x^\mu} = g_{\mu\nu} \frac{dx^\nu}{d\lambda} \ . \tag{2.5}$$

An alternative formulation of particle dynamics is based on treating the position and conjugate momentum as independent quantities during the extremization of the action. This Hamiltonian formulation follows from performing a Legendre transformation from $L(x, dx/d\lambda, \lambda)$ to $H(x, p, \lambda) \equiv p_\mu dx^\mu/d\lambda - L$ with $dx^\mu/d\lambda$ written in terms of x^μ, p_μ and λ by inverting equation (2.5). Using the Lagrangian of equation (2.4) yields the Hamiltonian

$$H_2(x^\mu, p_\nu, \lambda) = \frac{1}{2} g^{\mu\nu}(x) p_\mu p_\nu \ . \tag{2.6}$$

Although the value of this expression exactly equals L_2, it is written in terms of the variables appropriate for Hamiltonian dynamics. Extremization of the action now yields Hamilton's equations in four-dimensional covariant form:

$$\frac{dx^\mu}{d\lambda} = \frac{\partial H_2}{\partial p_\mu} = g^{\mu\nu} p_\nu \ , \quad \frac{dp_\mu}{d\lambda} = -\frac{\partial H_2}{\partial x^\mu} = g^{\beta\gamma} \Gamma^\alpha{}_{\beta\mu} p_\alpha p_\gamma \ . \tag{2.7}$$

These can be combined to give equation (2.3).

Although the covariant Hamiltonian formulation is elegant, it is inconvenient for practical use. One does not want to parametrize every photon's path by a different parameter λ; rather, one would like to specify the spatial position and direction of motion of photons at a time given by an observer's clock. Thus, it is desirable to abandon manifest covariance to use instead a description of trajectories in space as a function of time, $x^i(\tau)$ with $\tau = x^0$, just as is done in nonrelativistic mechanics. However, the equations of motion must be derived in a way that properly accounts for the nonflat spacetime and coordinate invariance of general relativity.

This wish is fulfilled with a non-covariant Hamiltonian prescription starting with the action of equation (2.4), multiplied by a factor of 2 for convenience (Bertschinger 1996). In fact, the Legendre transformation follows automatically on expanding a

sum over repeated indices:

$$S_3[x^i(\tau), p_j(\tau)] = 2S_2 = \int p_\mu dx^\mu = \int \left(p_0 + p_i \frac{dx^i}{d\tau} \right) d\tau \ . \tag{2.8}$$

Note that the position of the subscripts and superscripts is significant; p_0 and p_i are the time and space components of the momentum one-form p_μ and they must not be confused with the components of the momentum 4-vector $p^\mu = g^{\mu\nu} p_\nu$. From equation (2.8) we can read off the Hamiltonian and the conjugate momenta using the fact that $S = \int L d\tau$ where $L(x^i, \dot{x}^j, \tau)$ is the Lagrangian and it is related to the Hamiltonian by $L = p_i \dot{x}^i - H$. The Hamiltonian therefore is $H = -p_0$ — despite appearances, we shall see that this is *not* in general the *proper* energy — and the momentum conjugate to x^i is $p_i = g_{i\mu} dx^\mu / d\lambda$.

Applying Hamilton's principle of least action to equation (2.8) requires that we write $H = -p_0$ in terms of the phase space coordinates (x^i, p_j, τ). Ordinarily this would be done starting with the Lagrangian through $p_i = \partial L / \partial \dot{x}^i$. However, this is not possible here because equations (2.5) and (2.8) show that the Lagrangian is proportional to the unknown function $d\tau/d\lambda$. The main point of our approach is to eliminate the affine parameter altogether. The solution is to normalize the momentum using the mass-shell constraint:

$$g^{\mu\nu} p_\mu p_\nu = -m^2 \tag{2.9}$$

where m is the particle mass (zero for a photon). The constancy of $g^{\mu\nu} p_\mu p_\nu$ follows from equations (2.6) and (2.7) which imply $dH_2/d\lambda = \partial H_2 / \partial \lambda = 0$. Solving equation (2.9) for p_0 yields

$$H_3(x^i, p_j, \tau) = -p_0 = \frac{g^{0i} p_i}{g^{00}} + \left[\frac{(g^{ij} p_i p_j + m^2)}{-g^{00}} + \left(\frac{g^{0i} p_i}{g^{00}} \right)^2 \right]^{1/2} \ . \tag{2.10}$$

By building into our variational principle the constraint of equation (2.9) and by inverting $x^0(\lambda)$ to reparametrize the 8-dimensional trajectory $[x^\mu(\lambda), p_\mu(\lambda)]$, we have reduced the number of functions to vary from 8 to 6. Otherwise the variational principle based on equation (2.8) is fully equivalent to that based on equation (2.4).

Requiring the action of equation (2.8) to be stationary under independent variations of $x^i(\tau)$ and $p_i(\tau)$ (with $p_i \delta x^i = 0$ at the endpoints) yields the familiar form of Hamilton's equations:

$$\frac{dx^i}{d\tau} = \frac{\partial H_3}{\partial p_i} \ , \quad \frac{dp_i}{d\tau} = -\frac{\partial H_3}{\partial x^i} \ . \tag{2.11}$$

Although these equations distinguish time and space, when supplemented by equation (2.9) they are equivalent to equations (2.7) and hence to equations (2.3). As a bonus, we have learned from general relativity how to obtain the standard nonrelativistic Hamilton's equations and how to interpret the conjugate momentum for any coordinate system.

Our final step in this section is to determine the Hamiltonian for the problem

at hand. We shall assume that the spacetime is described by a linearly perturbed Friedmann-Robertson-Walker (FRW) spacetime with line element (Bardeen 1980)

$$ds^2 = a^2(\tau)\left[-(1+2\phi)d\tau^2 + (1-2\phi)\gamma_{ij}dx^i dx^j\right] , \tag{2.12}$$

where we set c = 1. Here τ is the conformal time, $a(\tau)$ the expansion factor (related to redshift z by $a(\tau) = 1/(1+z)$), and we are using comoving coordinates $\vec{x}$. Also ϕ is the scalar, Newtonian potential obeying the cosmological Poisson equation

$$\nabla^2\phi = 4\pi G a^2 \delta\rho , \tag{2.13}$$

where $\bar{\rho}$ is the mean mass density of the universe and $\delta\rho = \rho - \bar{\rho}$ is the local density perturbation.

Substituting the metric implied by equation (2.12) into equation (2.10) with $m = 0$ yields the Hamiltonian for a photon:

$$H(x^i, p_j, \tau) = p(1+2\phi) , \quad p \equiv \left(\gamma^{ij}p_i p_j\right)^{1/2} . \tag{2.14}$$

We have neglected all terms of higher order than linear in ϕ. Not surprisingly, in a perturbed FRW spacetime the Hamiltonian equals the momentum plus a small correction for gravity. However, it differs from the proper energy measured by a stationary observer, $E = -V^\mu p_\mu$, because the 4-velocity of such an observer is $V^\mu = (a^{-1}(1-\phi), 0, 0, 0)$ (since $g_{\mu\nu}V^\mu V^\nu = -1$) so that $E = a^{-1}(1+\phi)p$. The latter expression is easy to understand because a^{-1} converts comoving to proper energy, and ϕ is the gravitational energy per unit mass (energy). Why is the Hamiltonian different? The answer is that it is conjugate to the time coordinate τ which does not measure proper time. The job of the Hamiltonian is to provide the equations of motion and not to equal the energy. The factor of 2 in equation (2.14) is important — it is responsible for the fact that light is deflected twice as much as nonrelativistic particles in a gravitational field.

Hamilton's equations applied to equation (2.14) yield

$$\frac{dx^i}{d\tau} = n^i(1+2\phi) , \quad \frac{dp_i}{d\tau} = -2p\nabla_i\phi + \gamma^k{}_{ij}p_k n^j(1+2\phi) , \quad n^i \equiv \frac{\gamma^{ij}p_j}{p} . \tag{2.15}$$

We have defined a unit vector n^i in the photon's direction of motion (normalized so that $\gamma_{ij}n^i n^j = 1$). The symbol $\gamma^k{}_{ij} = \frac{1}{2}\gamma^{kl}(\partial_i\gamma_{jl} + \partial_j\gamma_{il} - \partial_l\gamma_{ij})$ is a connection coefficient for the spatial metric that vanishes if we are in flat space and use Cartesian coordinates.

In the following sections we shall represent three-vectors (and two-vectors) using arrows above the symbol.

2.3 Lensing in a flat universe

2.3.1 The lens equation

With $\gamma_{ij} = \delta_{ij}$ and $\gamma^k{}_{ij} = 0$, the equations of motion (2.15) for $\vec{x}$ and $\vec{n}$ become

$$\frac{d\vec{x}}{d\tau} = \vec{n}(1+2\phi) \ , \quad \frac{d\vec{n}}{d\tau} = -2\left[\vec{\nabla}\phi - \vec{n}(\vec{n}\cdot\vec{\nabla}\phi)\right] \ , \tag{2.16}$$

where $\phi = \phi(\vec{x},\tau)$ is evaluated on the photon path.

We choose the z-axis in the direction from the observer to the lens, with the observer at the origin. It is not important where exactly the z-axis crosses the lens plane, only that this is a fixed fiducial axis, and that actual image rays are at small angles to the z-axis. We separate the normal (x and y) components and the z component, i.e. we write

$$\vec{x} = (\vec{\xi},\chi) \ , \quad \vec{n} = (-\vec{\theta},-(1-\theta^2/2)) \ , \tag{2.17}$$

where $\vec{\xi}$ and $\vec{\theta}$ are two-dimensional vectors. We use χ and r interchangeably to denote values of the z-coordinate, since some of the results keep the same form in the non-flat case, when χ and r are no longer equal (§2.4). The minus signs are introduced in $\vec{n}$ because it points toward the observer. The z-component of $\vec{n}$ is determined by the normalization $\delta^{ij}n_i n_j = 1$, and we have kept the second order term in θ. With the definitions of equation (2.17), the observed image angle is $\vec{\theta}_0 = \vec{\theta}(\chi = 0)$.

Typically both θ and ϕ are $\sim 10^{-5}$, so we work to first order in both variables, and neglect cross-terms. It follows that $\chi(\tau)$ is single-valued, so that we may reparametrize the trajectories as functions of χ rather than of τ. The equations of motion now become

$$\frac{d\vec{\xi}}{d\chi} = \vec{\theta} \ , \quad \frac{d\vec{\theta}}{d\chi} = -2\vec{\nabla}_\perp\phi \ , \quad \frac{d\tau}{d\chi} = -(1-2\phi) \ , \tag{2.18}$$

where $\vec{\nabla}_\perp\phi = \partial\phi/\partial\vec{\xi}$ plus a correction that is $O(\theta)$ times this which we drop. We can solve these equations to get

$$\begin{aligned}
\vec{\theta}(\chi) &= \vec{\theta}_0 - 2\int_0^\chi \vec{\nabla}_\perp\phi(\vec{\xi}(\chi'),\chi',\tau(\chi'))\,d\chi' \ , \\
\vec{\xi}(\chi) &= r\vec{\theta}_0 - 2\int_0^\chi (r-r')\vec{\nabla}_\perp\phi(\vec{\xi}(\chi'),\chi',\tau(\chi'))\,d\chi' \ , \\
\tau(\chi) &= \tau_0 - \chi + 2\int_0^\chi \phi(\vec{\xi}(\chi'),\chi',\tau(\chi'))\,d\chi' \ .
\end{aligned} \tag{2.19}$$

These are only formal solutions, since ϕ is evaluated on the unknown path. The lens equation for a source at $\chi = \chi_S$ is $\vec{\xi}(\chi_S) = \vec{\xi}_S$, and we define a source angle $\vec{\beta} = \vec{\xi}_S/r_S$.

2.3.2 The thin lens equation

We now consider a "thin" lens at $\chi = \chi_L$, i.e. a lens which exerts a significant force over a z extent $\Delta\chi_{\rm lens}$ which is $\ll r_S$ (and also $\ll r_L$ and $\ll r_{LS} \equiv r_S - r_L$). The lens

equation is then

$$r_S\vec{\beta} = r_S\vec{\theta}_0 - 2r_{LS}\int_0^{\chi_S} \vec{\nabla}_\perp\phi(\vec{\xi}(\chi'),\chi',\tau(\chi'))\,d\chi' \ . \tag{2.20}$$

Since $\vec{\nabla}_\perp\phi$ is non-zero only for $|\chi-\chi_L| \lesssim \Delta\chi_{\rm lens}$, we have set $r-r' = r_S - r' = r_S - r_L = r_{LS}$ in the integrand, neglecting $O(\Delta\chi_{\rm lens}/r_{LS})$ corrections. Thus

$$\vec{\beta} = \vec{\theta}_0 - \vec{\alpha} \ , \tag{2.21}$$

where $\vec{\alpha} = (r_{LS}/r_S)\vec{\gamma}$ is the scaled deflection angle, and the (actual) deflection angle is $\vec{\gamma} = 2\int_0^{\chi_S}\vec{\nabla}_\perp\phi(\vec{\xi}(\chi),\chi,\tau(\chi))\,d\chi$. Since we have a thin lens, we may extend the integral to $\int_{-\infty}^{\infty}d\chi$. We also may evaluate ϕ in the integrand not at $(\vec{\xi}(\chi),\chi,\tau(\chi))$, but at $(\vec{\xi}(\chi_L),\chi,\tau(\chi_L))$,with $\vec{\xi}(\chi_L) = r_L\vec{\theta}_0$. This is justified since over the range of z where the integrand is non-zero, $\vec{\xi}(\chi)$ changes very little. Specifically, $\vec{\xi}(\chi_L)$ is $O(\theta_0 r_L)$, but $d\vec{\xi}/d\chi = \vec{\theta}$. The photon direction $\vec{\theta}$ changes rapidly near the lens, but the angle is always small, $\theta \sim \theta_0$. Thus $\Delta\vec{\xi}$ across the lens is $O(\theta_0\Delta\chi_{\rm lens})$, which is small compared to $\xi(\chi_L)$ and negligible if ϕ changes significantly only on a much larger scale than $\Delta\vec{\xi}$. In using the time $\tau(\chi_L)$, we are also assuming that ϕ does not change significantly with time during the lensing, which is true if the lens and its components have velocities $v \ll c$. Thus we have

$$\vec{\gamma}(\vec{\xi}) = 2\int_{-\infty}^{\infty}\vec{\nabla}_\perp\phi(\vec{\xi},\chi)\,d\chi \ , \tag{2.22}$$

where here and henceforth in this section, $\vec{\xi}$ stands for $\vec{\xi}(\chi_L) = r_L\theta_0$ and ϕ (and ρ) are evaluated at $\tau_L = \tau(\chi_L) = \tau_0 - \chi_L$ (We ignore the (second order: see §2.3.3) corrections to $\tau(\chi)$ assuming that the lens potential is not changing very rapidly during the lensing).

We want to relate $\vec{\gamma}$ to the lensing mass distribution. The (proper) mass density perturbation $\delta\rho$ is related to ϕ through equation (2.13). We have (in 2-D) that $\vec{\nabla}\cdot\vec{\gamma} = 2\int_{-\infty}^{\infty}\nabla_\perp^2\phi(\vec{\xi},\chi)\,d\chi$. We let $\vec{\eta} = a_L\vec{\xi}$ measure proper distance in the lens plane (as implied by the metric (2.12)). Now, the proper surface mass density projected onto the lens plane (i.e. the projected mass per unit area $d^2\eta$) is defined as $\Sigma(\vec{\xi}) = \int_{-\infty}^{\infty}\delta\rho(\vec{\xi},\chi)\,a_L d\chi = a_L^{-1}\int_{-\infty}^{\infty}\nabla^2\phi(\vec{\xi},\chi)/(4\pi G)\,d\chi$ (We have evaluated the scale factor at τ_L assuming that a does not change much over the short time when the deflection is occuring). One term we can integrate by parts: $\int_{-\infty}^{\infty}\partial_\chi^2\phi(\vec{\xi},\chi)\,d\chi = \partial_\chi\phi(\vec{\xi},\chi)|_{-\infty}^{\infty} = 0$. Thus $\Sigma(\vec{\xi}) = a_L^{-1}\int_{-\infty}^{\infty}\nabla_\perp^2\phi(\vec{\xi},\chi)/(4\pi G)\ d\chi$, so

$$\vec{\nabla}\cdot\vec{\gamma} = 8\pi G a_L^{-1}\Sigma(\vec{\xi}) \ . \tag{2.23}$$

From equation (2.22) it also immediately follows that $\partial_x\gamma_y - \partial_y\gamma_x = 0$. Thus $\vec{\gamma}$ is mathematically equivalent to a 2-D conservative force, determined by Σ just as the gravitational force is determined by ρ in 3-D. The corresponding 2-D harmonic

potential Ψ, for which $\vec{\gamma} = \partial_{\vec{\eta}}\Psi$, is given from equation (2.22) by

$$\Psi(\vec{\xi},\tau) = 2a_L \int_{-\infty}^{\infty} \phi(\vec{\xi},\chi,\tau)\, d\chi \ . \tag{2.24}$$

We can now apply the 2-D version of Gauss' Law. For an axially symmetric $\Sigma(\xi)$, we choose the Gaussian curve to be a circle C of radius ξ, and then

$$\gamma(\xi) \times 2\pi\xi = \oint_C \vec{\gamma}\cdot d\vec{\xi}\,' = \int\int \vec{\nabla}\cdot\vec{\gamma}\ dA = 8\pi G a_L^{-1} M_{\rm enc}(\xi) \ , \tag{2.25}$$

where $d\vec{\xi}\,'$ points radially outward along C, dA is an area element of the interior of C, and $M_{\rm enc}(\xi)$ is the total projected mass enclosed by C. The direction of $\vec{\gamma}$ is given by symmetry, so for any axisymmetric surface mass density,

$$\vec{\gamma}(\eta) = \frac{4GM_{\rm enc}(\eta)}{\eta}\ \hat{\eta} \ , \tag{2.26}$$

expressed in terms of proper distances. In particular, this gives the usual 2-D Green's function (i.e. deflection due to a point mass), with the corresponding potential $\Psi = 4GM \ln|\eta/(a_L r_L)|$. By linearity, we can also write down the expressions for a general Σ:

$$\begin{aligned} \vec{\gamma}(\vec{\eta}) &= 4G\int d^2\vec{\eta}\,'\ \Sigma(\vec{\eta}\,')\ \frac{\vec{\eta}-\vec{\eta}\,'}{|\vec{\eta}-\vec{\eta}\,'|^2} \ , \\ \Psi(\vec{\eta}) &= 4G\int d^2\vec{\eta}\,'\ \Sigma(\vec{\eta}\,')\ \ln\left(\frac{|\vec{\eta}-\vec{\eta}\,'|}{a_L r_L}\right) \ . \end{aligned} \tag{2.27}$$

2.3.3 The time delay

The equation of motion for $\chi(\tau)$ is $d\chi/d\tau = \vec{n}\cdot\hat{z}\,(1+2\phi)$. We have approximated this above as $d\chi/d\tau = -(1+2\phi)$. This was sufficient to derive the lens equation (indeed, for this purpose $d\chi/d\tau = -1$ suffices), but in order to find the difference in time taken along different image rays, we need the next order term in the angles, which for a thin lens makes a contribution of the same order as the 2ϕ term. There is no first order correction, so we go to second order: $d\chi/d\tau \simeq -(1-\theta^2/2)(1+2\phi) \simeq -(1-\theta^2/2+2\phi)$. The variable τ measures time at the observer, as we see from the metric (2.12), since $a(\tau_0) = 1$ and $\phi \to 0$ far from the lens. Thus the time for an image ray relative to that for the trivial ray along the z axis is

$$\Delta\tau = \int_0^{\chi_S} \left|\frac{d\tau}{d\chi}\right| d\chi - \chi_S = \int_0^{\chi_S} \left(\frac{1}{2}\theta^2 - 2\phi\right) d\chi \ . \tag{2.28}$$

The first term on the RHS is usually called the geometrical time delay, since it arises from the increased coordinate distance that the ray travels. The second term is the potential, gravitational, or Shapiro time delay, and it accounts for both the slowing down of clocks in a gravitational field and the extra proper distance caused by the

gravitational distortion of space. Like the first term, the second term is also a second order correction to r_S, since ϕ is non-zero only within $\sim \Delta\chi_{\rm lens}$ of the lens plane. Written in terms of the trajectory $\vec{\xi}(\chi)$, we have

$$\Delta\tau = \int_0^{\chi_S} \left[\frac{1}{2} \left(\frac{d\vec{\xi}}{d\chi} \right)^2 - 2\phi \right] d\chi \ . \tag{2.29}$$

In this form, the time delay is a functional of $\vec{\xi}(\chi)$. (We assume that we can evaluate ϕ in this equation at $\tau(\chi) = \tau_0 - \chi$, which is accurate as long as ϕ does not change significantly over a time scale $\Delta\tau$). Since the integration limits are fixed, we can easily check consistency with Fermat's principle, which states that the actual path of a photon is the path of least (or stationary) arrival time. Using the integrand of equation (2.29) as the Lagrangian, the Euler-Lagrange equations (with χ as the independent variable) immediately yield

$$\frac{d^2\vec{\xi}}{d\chi^2} = -2\vec{\nabla}_\perp \phi \ , \tag{2.30}$$

in agreement with the equations of motion (2.18), to first order.

2.3.4 The time delay in the thin lens case

We want to evaluate $\Delta\tau = \int_0^{\chi_S} \left(\frac{1}{2}\theta^2 - 2\phi \right) d\chi$ under the thin lens approximation of §2.3.2. In this case, $\vec{\theta}(\chi) = \vec{\theta}_0$ between the observer and the lens, and $\vec{\theta}(\chi) = \vec{\theta}_0 - \vec{\gamma}$ between the lens and the source, so the geometrical time delay is simple. We also see from equation (2.24) that the potential time delay is $-a_L^{-1}\Psi$. Thus

$$\Delta\tau = \frac{1}{2}\theta_0^2 r_L + \frac{1}{2}|\vec{\theta}_0 - \vec{\gamma}|^2 r_{LS} - a_L^{-1}\Psi \ . \tag{2.31}$$

Once again we can show consistency with Fermat's principle, which here takes the form $\partial\Delta\tau/\partial\vec{\theta}_0 = 0$ at fixed $\vec{\beta}$ (i.e. for a fixed source). Before applying this we must first write the geometrical time delay in terms of $\vec{\theta}_0$ and $\vec{\beta}$, using the lens equation $\vec{\beta} = \vec{\theta}_0 - \vec{\alpha}$. This yields

$$\Delta\tau = \frac{1}{2}\frac{r_L r_S}{r_{LS}}|\vec{\theta}_0 - \vec{\beta}|^2 - a_L^{-1}\Psi \ , \tag{2.32}$$

except for an unimportant $\vec{\theta}_0$-independent constant. Fermat's principle again yields the lens equation (2.21). In terms of angular diameter distances (e.g. $D_{LS} = a_S r_{LS}$), the time delay is

$$\Delta\tau = a_L^{-1}\left[\frac{1}{2}\frac{D_{OL}D_{OS}}{D_{LS}}|\vec{\theta}_0 - \vec{\beta}|^2 - \Psi \right] . \tag{2.33}$$

We prefer to use comoving distances since they are more convenient in cases of continuous deflection, e.g. for the lensing effect of Large-Scale Structure (see chapter

3).

2.4 Lensing in an open or closed universe

In general, the equations of motion (2.15) yield for $\vec{n}$:

$$\frac{dn_i}{d\tau} = -2\nabla^i_\perp \phi - \gamma^i{}_{jk} n^j n^k (1+2\phi) \ , \tag{2.34}$$

where $\nabla^i_\perp = (\gamma^{ij} - n^i n^j)\nabla_j$. We begin with the perturbed FRW line element in the standard form

$$ds^2 = a^2(\tau)\Big[-(1+2\phi)d\tau^2 + (1-2\phi)[d\chi^2 + r^2(\chi)d\Omega^2]\Big] \ , \tag{2.35}$$

in terms of spherical comoving coordinates, where $d\Omega^2 = d\theta_\chi^2 + \sin^2\theta_\chi d\phi_\chi^2$, and

$$r(\chi) = \begin{cases} K^{-1/2}\sin K^{1/2}\chi & \text{if } K > 0. \\ \chi & \text{if } K = 0. \\ (-K)^{-1/2}\sinh(-K)^{1/2}\chi & \text{if } K < 0. \end{cases} \tag{2.36}$$

The curvature constant is $K = (\Omega_0 - 1)H_0^2$. The function $a(\tau)$ is given by the Friedmann equation,

$$\left(\frac{da}{d\tau}\right)^2 = a^2\left(\frac{8\pi G\bar{\rho}a^2}{3} - K\right) \ . \tag{2.37}$$

We replace the angular coordinates by coordinates on the celestial sphere:

$$\begin{aligned} \theta_1 &= 2\tan\frac{\theta_\chi}{2}\cos\phi_\chi \ , \\ \theta_2 &= 2\tan\frac{\theta_\chi}{2}\sin\phi_\chi \ , \\ d\Omega^2 &= \left(d\theta_1^2 + d\theta_2^2\right)\Big/\left[1+\frac{1}{4}\theta^2\right]^2 \ , \end{aligned} \tag{2.38}$$

where $\theta^2 = \theta_1^2 + \theta_2^2$. We write

$$\vec{x} = (\theta_1, \theta_2, \chi) \ , \quad \vec{n} = \left(-\frac{\omega_1}{r}, -\frac{\omega_2}{r}, -(1-\omega^2/2)\right) \ , \tag{2.39}$$

where $\omega^2 = \omega_1^2 + \omega_2^2$. In the small angle approximation θ^2, $\omega^2 \ll 1$ and the weak field approximation $\phi^2 \ll 1$, we obtain

$$\frac{d\theta^i}{d\chi} = \frac{\omega^i}{r(\chi)} \ , \quad \frac{d\omega^i}{d\chi} = -\frac{2}{r(\chi)}\frac{\partial\phi}{\partial\theta_i} - \frac{r'(\chi)}{r(\chi)}\omega_i \ , \quad \frac{d\tau}{d\chi} = -\left(1 - 2\phi + \frac{1}{2}\omega^2\right) \ . \tag{2.40}$$

We define $\xi_i = r\theta_i$, which measures comoving distance (to order θ_χ^2). The solutions

of equations (2.40) are

$$\begin{aligned}\omega_i(\chi) &= -\frac{2}{r(\chi)}\int_0^\chi r(\chi')\partial_{\xi_i}\phi(\vec{\xi}(\chi'),\chi',\tau(\chi'))\,d\chi' \ , \\ \xi_i(\chi) &= r(\chi)\theta_i(0) - 2\int_0^\chi r(\chi-\chi')\partial_{\xi_i}\phi(\vec{\xi}(\chi'),\chi',\tau(\chi'))\,d\chi' \ , \end{aligned} \tag{2.41}$$

and to lowest order $\tau(\chi) = \tau_0 - \chi$. Once again, the lens equation is $\vec{\xi}(\chi_S) = \vec{\xi}_S$, and we define a source angle $\vec{\beta} = \vec{\xi}_S/r_S$ (where $r_S = r(\chi_S)$, and the angular diameter distance is $D_{OS} = a_S r_S$).

For a thin lens, all the expressions in §2.3.2 remain unchanged (where $\vec{\nabla}_\perp\phi = \partial\phi/\partial\vec{\xi}$, and now $r_{LS} = r(\chi_S - \chi_L)$). The time delay, however, has a different form in general:

$$\Delta\tau = \int_0^{\chi_S}\left|\frac{d\tau}{d\chi}\right| d\chi - \chi_S = \int_0^{\chi_S}\left(\frac{1}{2}\omega^2 - 2\phi\right) d\chi \ . \tag{2.42}$$

The time delay is here measured with respect to the undeflected path to the source (i.e. the path $\theta_i(\chi) = \beta_i$). We can write $\Delta\tau$ as a functional of $\vec{\xi}(\chi)$:

$$\Delta\tau = \int_0^{\chi_S}\left(\frac{1}{2}\left|\frac{d\vec{\xi}}{d\chi} - \frac{r'(\chi)}{r(\chi)}\vec{\xi}\right|^2 - 2\phi\right) dr \ . \tag{2.43}$$

The Euler-Lagrange equations with respect to χ yield

$$\frac{d^2\vec{\xi}}{d\chi^2} + K\vec{\xi} = -2\partial_{\xi_i}\phi \ , \tag{2.44}$$

which can also be derived from equations (2.40).

For a thin lens, $\omega_i = 0$ between the observer and the lens, and $\omega_i = -r_L\gamma_i/r(\chi)$ between the lens and the source. Using the identity

$$\frac{d}{d\chi}\left[\frac{r(\chi-\chi')}{r(\chi)r(\chi')}\right] = \frac{1}{r^2(\chi)} \ , \tag{2.45}$$

we reobtain equation (2.32) identically.

2.5 Lensing due to GW's

We now apply the Hamiltonian approach to lensing in the presence of GW's. The line element is

$$ds^2 = a^2(\tau)\left[-d\tau^2 + (\gamma_{ij} + h_{ij})dx^i dx^j\right] \ , \tag{2.46}$$

which yields the following Hamiltonian for photons:

$$H(x^i, p_j, \tau) = p\left(1 - \frac{1}{2}h_{nn}\right) \ , \quad h_{nn} \equiv h_{ij}n^i n^j \ . \tag{2.47}$$

Hamilton's equations are then

$$\frac{dx^i}{d\tau} = n^i\left(1+\frac{1}{2}h_{nn}\right) - h^{ji}n_j \ , \quad \frac{dp_i}{d\tau} = \frac{1}{2}p_j n_k \nabla_i h^{jk} + \gamma^k_{ij}p_k n^j\left(1+\frac{1}{2}h_{nn}\right) \ . \quad (2.48)$$

These equations in turn yield

$$\frac{dn^i}{d\tau} = \frac{1}{2}(\gamma^{ij}-n^in^j)n_l n_k \nabla_j h^{lk} - \gamma^i_{jk}n^j n^k\left(1+\frac{1}{2}h_{nn}\right) + (\gamma^{ij}-\frac{1}{2}n^in^j)n^k n_l h^{lm}\nabla_m \gamma_{jk} \ . \quad (2.49)$$

2.5.1 GW's in a flat universe

In this section we adopt the flat coordinates and approximations used in §2.3.1, and work to first order in h_{ij}. The equations of motion are

$$\frac{d\xi_i}{d\chi} = \theta_i - h_{\chi i} \ , \quad \frac{d\theta_i}{d\chi} = \frac{1}{2}\nabla_i h_{\chi\chi} \ , \quad (2.50)$$

with solutions

$$\begin{aligned} \theta_i(\chi) &= \theta_i(0) + \frac{1}{2}\int_0^\chi \nabla_i h_{\chi\chi}(\vec{\xi}(\chi'),\chi',\tau(\chi'))\,d\chi' \ , \\ \xi_i(\chi) &= r\delta_i(0) + \int_0^\chi \left[\frac{1}{2}(r-r')\nabla_i h_{\chi\chi}(\chi') + h_{\chi i}(0) - h_{\chi i}(\chi')\right] d\chi' \ , \end{aligned} \quad (2.51)$$

where (for comparison with Linder (1988)) we define $\delta_i(0) \equiv d\xi_i/d\chi\,(0) = \theta_i(0) - h_{\chi i}(0)$.

The time delay equals

$$\Delta\tau = \frac{1}{2}\int_0^{\chi_S}(\theta^2 + h_{\chi\chi})d\chi \ , \quad (2.52)$$

which is consistent with Fermat's principle and equations (2.51).

2.5.2 GW's in an open or closed universe

In the notation of §2.4, the equations of motion are

$$\frac{d\theta_i}{d\chi} = \frac{\omega_i}{r} - \frac{h_{\chi i}}{r^2} \ , \quad \frac{d}{d\chi}(r\omega_i) = \frac{1}{2}\nabla_i h_{\chi\chi} \ , \quad (2.53)$$

with solutions

$$\begin{aligned} \omega_i(\chi) &= \frac{1}{2\,r(\chi)}\int_0^\chi \nabla_i h_{\chi\chi}(\chi')\,d\chi' \ , \\ \xi_i(\chi) &= r(\chi)\theta_i(0) + \int_0^\chi \left[\frac{r(\chi-\chi')}{2\,r(\chi')}\nabla_i h_{\chi\chi}(\chi') - \frac{r(\chi)}{r^2(\chi')}h_{\chi i}(\chi')\right] d\chi' \ . \end{aligned} \quad (2.54)$$

In these expressions, the components h_{ij} (and also ∇_i) are expressed in coordinates $x_i = (\theta_1, \theta_2, \chi)$. For comparison with the flat case, we convert to coordinates $\tilde{x}_i = (\xi_1, \xi_2, \chi)$. To the order we need, the conversions are

$$h_{\chi\chi} = \tilde{h}_{\chi\chi} + 2\,\tilde{h}_{\chi i}\,\theta_i\,r'(\chi)\ , \quad h_{\chi i} = \tilde{h}_{\chi i}\,r(\chi)\ . \tag{2.55}$$

We then have

$$\begin{aligned}
\omega_i(\chi) &= \frac{1}{2r(\chi)}\int_0^\chi \left[r(\chi')\partial_{\xi_i}\tilde{h}_{\chi\chi}(\chi') + 2\,r'(\chi')\tilde{h}_{\chi i}\right]\,d\chi'\ , \\
\xi_i(\chi) &= r(\chi)\theta_i(0) + \int_0^\chi \Big\{\frac{1}{2}r(\chi-\chi')\partial_{\xi_i}\tilde{h}_{\chi\chi}(\chi') \\
&\qquad + \frac{1}{r(\chi')}\left[r(\chi-\chi')r'(\chi') - r(\chi)\right]\tilde{h}_{\chi i}\Big\}\,d\chi'\ .
\end{aligned} \tag{2.56}$$

The time delay is

$$\Delta\tau = \frac{1}{2}\int_0^{\chi_S}(\omega^2 + h_{\chi\chi})d\chi\ , \tag{2.57}$$

once again consistent with Fermat's principle.

2.6 Multiple Lensing

When there are several distinct lenses, each satisfying the approximations needed for a thin lens (§2.3.2), the overall path can be described simply by iterating the thin lens equation (2.21). For this to be accurate, the lenses must be sufficiently separated in χ so that as light passes each lens plane the influence on it of mass in the other lens planes can be neglected. Our derivation of the multiple lens equations in this case would be very similar to the standard one (e.g. Schneider et al. 1992). We therefore consider instead the accuracy of multiple lensing as an approximation to continuous deflection. Multiple lensing has been used in this way for studies of lensing due to large-scale structure in N-body simulations, for average galaxy distortions from weak shear (Jaroszynski et al. 1991; Blandford et al. 1991), for weak magnifications of sources (Wambsganss et al. 1996), and for strong lensing statistics (Wambsganss et al. 1995). We determine the conditions needed for multiple lensing to accurately reproduce the true photon trajectories given by integrating the 3-dimensional deflection. In particular, we look for conditions on the number N of lens planes (or equivalently on the distance between successive lens planes) that ensure satisfactory accuracy.

For generality we consider the non-flat case, and use the coordinates (ξ_1, ξ_2, χ) from §2.4. We consider the multiple lensing approximation, in the following form: We divide the z-axis into N sections, with boundaries at $\chi = \Delta^i$, $i = 1, ..., N$ (and $\Delta^0 = 0$). The deflection from each layer is computed at its midpoint $\chi = \zeta^i = (\Delta^{i-1} + \Delta^i)/2$, $i = 1, ..., N$, using the thin lens approximation with the mass between Δ^{i-1} and Δ^i projected onto the lens plane at ζ^i. Superscripts denote lens plane number, so e.g. $\vec{\xi}^i \equiv \vec{\xi}(\zeta^i)$. Also, $r^i \equiv r(\zeta^i)$ and $r^{ij} \equiv r(\xi^j - \xi^i)$. Then with $\vec{\theta}_0$ the

initial direction as in §2.4, and $\vec{\gamma}^i$ the deflection at lens plane i computed from Σ as in §2.3.2, the multiple lensing equations are

$$\vec{\xi}^j = r^j \vec{\theta}_0 - \sum_{i=1}^{j-1} r^{ij} \vec{\gamma}^i(\vec{\xi}^i) \tag{2.58}$$

(with no sum for $j = 1$).

The exact solution, however, is the second equation in (2.41), which can be written in the form

$$\vec{\xi}^j = r^j \vec{\theta}_0 - \left[\sum_{i=1}^{j-1} \Gamma^{ij} + \Pi^j \right] , \tag{2.59}$$

where

$$\Gamma^{ij} = \int_{\Delta^{i-1}}^{\Delta^i} 2\, r(\chi^j - \chi')\, \partial_{\vec{\xi}} \phi(\vec{\xi}(\chi'), \chi', \tau(\chi')) d\chi' ,$$
$$\Pi^j = \int_{\Delta^{j-1}}^{\zeta^j} 2\, r(\chi^j - \chi')\, \partial_{\vec{\xi}} \phi(\vec{\xi}(\chi'), \chi', \tau(\chi')) d\chi' .$$

Starting with this exact solution, we show the steps needed to reach equation (2.58), and derive the conditions needed for the validity of each step, up to factors of order unity.

In order to make a semi-quantitative estimate of errors, we make some simplifying assumptions about the statistics of the deflections. We let k be the typical wavenumber which dominates the gradient of the deflection. Whether ϕ is modelled as a random field or by a superposition of individual deflectors with some characteristic scale, we define k so that $\sum_{i,j} \left| \partial_{\xi_i} \partial_{\xi_j} \phi(\vec{\xi}, \chi, \tau(\chi)) \right|^2 \simeq k^2 \left| \partial_{\vec{\xi}} \phi \right|^2$. This means that the error in $\vec{\xi}$ that can be tolerated in evaluating $\partial_{\vec{\xi}} \phi(\vec{\xi}, \chi, \tau(\chi))$ is of order k^{-1}. We also define

$$\vec{\gamma}_*^i = \int_{\Delta^{i-1}}^{\Delta^i} 2\, \partial_{\vec{\xi}} \phi(\vec{\xi}^i, \chi', \tau^i) d\chi' . \tag{2.60}$$

We assume that the Δ^i for various i are roughly equally spaced, so that all the $\vec{\gamma}_*^i$ are of the same order, γ_*, and that the total deflection $\gamma_{\rm tot} = \left| \sum_{i=1}^N \vec{\gamma}_*^i \right|$ is a random walk of roughly kr_S steps, so that $\gamma_{\rm tot} \propto \sqrt{kr_S}$ (as is true for LSS: Seljak 1994), and also $\propto \sqrt{N}$ (which is true assuming $k^{-1} \lesssim r_S/N$, or $N \lesssim kr_S$). We also assume $N \gg 1$.

2.6.1 Minimal Errors

First we assume that when we calculate $\vec{\xi}^j$, all the $\vec{\xi}^i$ $(i < j)$ are known exactly, and we consider the error in $\vec{\xi}^j$ resulting from other approximations. In the next subsection we consider the effect of errors in the $\vec{\xi}^i$, as well as the errors in average statistical properties which may be small even when the errors for individual light rays are large.

The first step of the approximation is to evaluate ϕ in the integrand of Γ^{ij} at $(\vec{\xi}^i, \chi', \tau^i)$ instead of $(\vec{\xi}(\chi'), \chi', \tau(\chi'))$. Since we are taking $\vec{\xi}$ to be independent of χ',

we require that across every layer $|\Delta\vec{\xi}| \ll k^{-1}$. Since $|\Delta\vec{\xi}|$ is of order $(\theta^0 + \gamma_{\rm tot})r_S/N$, this yields the condition $N/(kr_S\gamma_{\rm tot}) \gg 1$ as well as $N/(kr_S\theta^0) \gg 1$. Evaluating $\vec{\xi}$ at a single time also assumes that the timescale $\delta\tau$ over which $\partial_{\vec{\xi}}\phi$ changes is large compared to the time for light to traverse the $i-$th layer. The former is $\sim 1/(kv)$ with v the typical velocity of a lens or matter clump of size k^{-1}, and the latter is $\sim r_S/(cN)$. We thus find $N \gg r_S/(c\,\delta\tau)$ or $N \gg kr_S v/c$.

We next need to set $\Gamma^{ij} = r^{ij}\vec{\gamma}^i_*$, i.e. set $r(\chi^j - \chi') = r^{ij}$ in the integrand of Γ^{ij}. The contribution from each i to ξ^j is $\sim r^{ij}\gamma_* \sim r_S(j-i)\gamma_*/N$, while the error we are proposing to make is $\sim r_S\gamma_*/N$. Both the total and the error random-walk as we add up all $i < j$, and the total error in ξ^j is of order $r_S\gamma_*\sqrt{j}/N$ (a relative error of $\sim 1/j$). Note that including the error in ξ^j caused by neglecting Π^j (equation (2.59)) does not significantly change this estimate.

The next step is to set $\vec{\gamma}^i_* = \gamma^i$, both considered as 2-dimensional functions of ξ^i. Following §2.3.2, we have the 2-D equations

$$\begin{aligned}
\vec{\nabla}_\perp \times \vec{\gamma}^i_* &= 0\,, \\
\vec{\nabla}_\perp \cdot \vec{\gamma}^i_* &= 2\int_{\Delta^{i-1}}^{\Delta^i} (\nabla^2\phi - \partial^2_\chi\phi)d\chi' = 8\pi G\Sigma^i/a^i - 2\partial_\chi\phi|^{\Delta^i}_{\Delta^{i-1}}\,, \qquad (2.61)
\end{aligned}$$

with Σ^i the mass density in layer i projected onto the lens plane at ζ^i. We can write $\vec{\gamma}^i_* = \vec{\gamma}^i + \vec{\epsilon}^i(\vec{\xi}^i) - \vec{\epsilon}^{i-1}(\vec{\xi}^i)$, where $\vec{\epsilon}^i(\vec{\xi})$ is a deflection field determined by $\vec{\nabla}_\perp \times \vec{\epsilon}^i(\vec{\xi}) = 0$, $\vec{\nabla}_\perp \cdot \vec{\epsilon}^i(\vec{\xi}) = -2\,\partial_\chi\phi(\vec{\xi}, \Delta^i, \tau(\Delta^i))$. We assumed $N \gg r_S/(c\,\delta\tau)$, the same condition required above, in order to replace $\tau(\zeta^i)$ with $\tau(\Delta^i)$. Assuming also that in this layer $|\Delta\vec{\xi}| \ll k^{-1}$ (as above), we can write $\vec{\gamma}^i_* \simeq \vec{\gamma}^i + \vec{\epsilon}^i(\vec{\xi}(\Delta^i)) - \vec{\epsilon}^{i-1}(\vec{\xi}(\Delta^{i-1}))$.

Physically, the $\vec{\epsilon}$ terms are made up of two contributions: the force of mass outside layer i on the photon when it passes through layer i, and also the error in the deflection from layer i due to the fact that the full deflection does not occur (i.e. the light ray does not come in and out on straight lines from an infinite distance). It is the fact that these two effects can be combined in the form of a boundary term which makes the resulting error small, as we now show. The contribution of all layers to ξ^j is now $\sum_{i=1}^{j-1}\Gamma^{ij} = \sum_{i=1}^{j-1} r^{ij}\left[\vec{\gamma}^i + \vec{\epsilon}^i(\vec{\xi}(\Delta^i)) - \vec{\epsilon}^{i-1}(\vec{\xi}(\Delta^{i-1}))\right] = \sum_{i=1}^{j-1}\left[r^{ij}\vec{\gamma}^i + (r^{ij} - r^{i+1\,j})\vec{\epsilon}^i(\vec{\xi}(\Delta^i))\right] - r^{1j}\vec{\epsilon}^0(\vec{\xi} = \vec{0})$. Assuming homogeneity, $\vec{\epsilon}^i$ is of the same order as $\vec{\gamma}^i$, but because of the suppression with $r^{ij} - r^{i+1\,j}$ compared to r^{ij}, the total error is again of order $r_S\gamma_*\sqrt{j}/N$. Interestingly, there is an extra boundary term at the observer, $r^{1j}\vec{\epsilon}^0$, which makes a $\sim 1/\sqrt{j}$ relative error. This term is, however, a $\vec{\theta}_0-$independent constant and is therefore statistically irrelevant unless various observer positions are compared.

2.6.2 Final errors

In the preceding paragraph, we assumed that we knew exactly the $\vec{\xi}^i$ on lens planes $i < j$, and found that we still make an error in $\vec{\xi}^j$ of order $r_S\gamma_*\sqrt{j}/N$, where we must also assume $N/(kr_S\gamma_{\rm tot}) \gg 1$, $N/(kr_S\theta^0) \gg 1$, and $N \gg r_S/(c\,\delta\tau)$. We still have two important complications left to add. First, the errors that we have derived in

each $\vec{\xi}^i$ (replacing j by i) may in turn cause a larger error in $\vec{\xi}^j$, when these errors are included for $i < j$. This process of error propagation can be repeated recursively, and it must converge for the total error to be small. Second, our conditions may be too restrictive, since we have not fully used the fact that we require accuracy only for statistical properties, when many different pairs of rays are averaged over.

We consider first the effect of propagating the error of order $r_S\gamma_*\sqrt{i}/N$ which we have derived in $\vec{\xi}^i$ when $\vec{\xi}^i$ is used in the calculation of $\vec{\xi}^j$, for $j > i$. This results in a relative error in $\vec{\gamma}^i$ of order $kr_S\gamma_*\sqrt{i}/N$. We now use induction in order to estimate the effect of recursively propagating these errors. Thus, we call this the $n = 1$ step, and work out the general case of going from step n to $n + 1$. At the n-th step, for every i the relative error in $\vec{\gamma}^i$ is of order

$$\delta_n(i) = \frac{1}{N} C_n (kr_S\gamma_{\rm tot})^{Q_n} \left(\frac{i}{N}\right)^{P_n} .$$

The induction which follows shows that this is the correct form, with $Q_1 = 1$, $P_1 = 1/2$, and $C_1 = 1$, ignoring n-independent factors.

The induced actual (not relative) error in $r^{ij}\vec{\gamma}^i$ is of order $\delta_n(i)\gamma_* r_S(j-i)/N$. To sum up the total error in $\vec{\xi}^j$, we use the approximation of a random walk: For a random walk with $j \gg 1$ independent steps of size $f(i)$, the mean square sum is

$$\left\langle \left(\sum_{i=1}^{j} f(i) \right)^2 \right\rangle = \sum_{i=1}^{j} \left\langle f^2(i) \right\rangle \simeq \int_0^j f^2(x)\, dx \ .$$

Once we have the rms induced error in $\vec{\xi}^j$, we multiply it by k to get $\delta_{n+1}(j)$. As a result we derive

$$\delta_n(i) = \frac{1}{N} \sqrt{\frac{2^{n-1}}{(3n-2)!}}\, (kr_S\gamma_{\rm tot})^n \left(\frac{i}{N}\right)^{\frac{3}{2}n-1} . \qquad (2.62)$$

We want to have $\delta_n(i) \ll 1$ for $i = N$, and for all $n \geq 1$ (roughly up to $n = N$). For $kr_S\gamma_{\rm tot} \lesssim 3.5$, the most restrictive condition is $N/(kr_S\gamma_{\rm tot}) \gg 1$. For larger values up to $kr_S\gamma_{\rm tot} \lesssim 10$, we add the condition $3.5N/(kr_S\gamma_{\rm tot})^2 \gg 1$. For $kr_S\gamma_{\rm tot} \lesssim 19$, we also have $36N/(kr_S\gamma_{\rm tot})^3 \gg 1$, etc. These conditions quickly begin to exclude $N \sim 100$ as $kr_S\gamma_{\rm tot}$ becomes large.

We now address the second issue noted above. In using N-body simulations we are often only interested in statistical properties of lensing. Consider first the conditions derived in the previous subsection, before propagating the errors. The statistical error in integrating ϕ along a straight $\vec{\xi}$ =constant path in each layer instead of the actual path (which is at a tiny angle to the $z-$axis) is very small. The statistics of shear, magnification, and multiple imaging depend on the correlation between nearby ray directions, say one with initial direction $\vec{\theta}_a = \vec{0}$ and the other with $\vec{\theta}_b = \vec{\theta}_0$ (cf. Seljak 1994, who notes that absolute deflections are not observable while relative deflections are). So for example, instead of needing to replace $\vec{\xi}(\chi')$ by $\vec{\xi}^i$ in layer i, we require only that $\vec{\xi}_a(\chi') - \vec{\xi}_b(\chi')$ equal $\vec{\xi}_a^i - \vec{\xi}_b^i$ to within about k^{-1}, for $\Delta^{i-1} < \chi' < \Delta^i$. This

yields approximately $(\theta^0 + |\gamma_{\rm tot}(\vec{\theta}_0) - \gamma_{\rm tot}(\vec{\theta}_0 = 0)|)r_S/N \ll k^{-1}$.

If we compare directions in a small angular range so that $kr_S\theta^0 \ll 1$, then Taylor-expanding $\gamma_{\rm tot}$ (and also every $\vec{\gamma}_*^i$) as a function of $\vec{\theta}_0$ (for a given field $\phi(\vec{\xi}, \chi, \tau)$) yields $|\gamma_{\rm tot}(\vec{\theta}_0) - \gamma_{\rm tot}(\vec{\theta}_0 = 0)| \sim kr_S\theta^0\gamma_{\rm tot}$. Reasoning in this way we find that the conditions in the previous subsection are weakened to $N/(kr_S\gamma_{\rm tot}kr_S\theta^0) \gg 1$, and still also $N/(kr_S\theta^0) \gg 1$. When we propagate the errors, an error of $\Delta\vec{\xi}^i$ again leads to a relative error in $\vec{\gamma}^i$ of order $k|\Delta\vec{\xi}^i|$. Thus errors propagate the same, only with one overall factor of $kr_S\theta^0$. On the other hand, if we compare directions at a larger separation $kr_S\theta^0 \gtrsim 1$, the two directions become only weakly correlated and $|\gamma_{\rm tot}(\vec{\theta}_0) - \gamma_{\rm tot}(\vec{\theta}_0 = 0)| \sim \gamma_{\rm tot}$. However, we still may not require the full conditions derived above. We might expect that if we average over many directions then the uncorrelated terms will average out, while the correlated parts - which we are interested in, and which may be small for each individual pair of rays - will dominate the average. In addition, if we have power on a range of scales, small scale modes ($kr_S\theta^0 \gg 1$) are more weakly correlated between the two rays and tend to cancel out in averaging, which lowers the dominant k and makes all the above conditions easier to satisfy with a smaller N. The time condition $N \gg r_S/(c\,\delta\tau)$ is also weaker for statistics since even though individual deflectors may evolve on a timescale of $1/(kv)$ the average statistics of LSS may change on a much longer timescale.

To illustrate the typical numbers, we take $\gamma_{\rm tot} \sim 1'$ (Seljak 1994) and $r_S \sim 1$ Gpc. Then $kr_S\gamma_{\rm tot} \sim 0.3(1\ {\rm Mpc}/k^{-1})$. For weak lensing on arcminute scales (where $k^{-1} \sim 10$ Mpc), N of $50 - 100$ is probably safe. For lensing which involves scales of $k^{-1} \sim 100$ kpc, multiple lensing is still accurate if we consider only $kr_S\theta^0 \ll 1$, which includes e.g. the statistics of point sources which are not multiply imaged, where the shear and magnification depend on comparing infinitesimally close pairs of rays. If we consider angular correlations of shear, or multiply imaged sources, the question of accuracy is less clear. The conditions we have derived are only rough estimates. There are various small factors which we have not included. In any realistic case, there will always be power on a range of scales, and as indicated above this leads to a dependence between the angular scale θ^0 and the dominant k^{-1}. Thus numerical tests with N-body simulations are needed to test whether multiple lensing works even just statistically, when both $kr_S\gamma_{\rm tot} \gtrsim 1$ and $kr_S\theta^0 \gtrsim 1$.

Bibliography

[1] Bardeen, J. M. 1980, Phys. Rev. D, 22, 1882

[2] Bertschinger, E., in Cosmology and Large Scale Structure, proc. Les Houches Summer School, Session LX, ed. R. Schaeffer, J. Silk, M. Spiro, and J. Zinn-Justin (Amsterdam: Elsevier Science) (preprint astro-ph/9503125)

[3] Blandford, R. D., & Kochanek, C. S. 1989, in Dark Matter in the Universe, ed. J. Bahcall, T. Piran, & S. Weinberg (Singapore: World Scientific)

[4] Blandford, R. D., Saust, A. B., Brainerd, T. G., & Villumsen, J. V. 1991, MNRAS, 251, 600

[5] Jaroszynski, M., Park, C., Paczynski, B., & Gott, J. R. 1991, ApJ, 365, 22

[6] Kantowski, R. 1969, ApJ, 155, 89

[7] Kodama, H., & Sasaki, M. 1984, Prog. Theor. Phys. Suppl. 78, 1

[8] Linder, E. 1988, ApJ, 328, 77

[9] Linder, E. 1990, MNRAS, 243, 362

[10] Misner, C. W., Thorne, K. S., & Wheeler, J. A. 1973, Gravitation (San Franciso: Freeman)

[11] Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses (Berlin: Springer-Verlag)

[12] Seljak, U. 1994, ApJ, 436, 509

[13] Seljak, U. 1995, Ph.D. Thesis, Massachusetts Institute of Technology

[14] Wambsganss, J., Cen, R., Ostriker, J. P., & Turner, E. L. 1995, Science, 268, 274

[15] Wambsganss, J., Cen, R., Xu G., & Ostriker, J. P. 1997, ApJ, 475, 81

Chapter 3

Effect of Large-Scale Structure on Multiply Imaged Sources

This chapter[1] deals with the observable effects of light deflection due to density fluctuations in the universe.

3.1 Introduction

Since gravitational lenses and sources typically lie at significant redshifts, light rays are deflected by large-scale structure (LSS) as they traverse the enormous distance from source to observer. These deflections are not large enough to produce multiple images, but they do distort the shapes of sources. Such weak lensing has been investigated both analytically (e.g. Miralda-Escudé 1991; Kaiser 1992) and in N-body simulations (Jaroszynski et al. 1991; Blandford et al. 1991). These studies find a shear of order 1%, coherent over a scale of $\approx 1°$, in a flat cold dark matter (CDM) model. This shear may in principle be detected observationally as a coherent distortion of background galaxies, when averaged over a sufficiently large angular field in order to be separated from the random scatter of intrinsic ellipticities (e.g. Mould et al. 1994; Villumsen 1995).

The shear due to LSS can also affect strong lensing, when it acts in addition to a strong primary lens, a galaxy or cluster near the line of sight to the source. This effect is enhanced compared to weak lensing, because of the small angular separations between multiple images. Also, the higher redshift of quasars compared to faint galaxies increases the cumulative shear from the observer to the source. Seljak (1994) estimated the dependence of the rms value of this shear on the power spectrum of density fluctuations, and found it to be of order 10% for a source at redshift 3. Seljak also considered the effect of LSS on the time delay, and showed that the lowest order terms cancel out in the total time delay. However, since these canceling terms are separately much larger than the time delay from the primary lens, even higher order

[1]Based on the publication Bar-Kana, R. 1996, ApJ, 468, 17

terms might still dominate the time delay and threaten the effort to determine the Hubble constant from lensing.

In order to find precisely how LSS affects the observables of a lens system, Surpi et al. (1996) set up the lens equation in the presence of a lens plus LSS. They made an expansion for the position of a light ray in powers of its deflection from the unperturbed straight path, and kept only the lowest order term. This term is equivalent to a constant angular deflection at the lens. They thus concluded that LSS leaves all observables (such as relative image positions) unchanged to lowest order. Indeed, since the actual source position is unobservable, the effect of this lowest order term can be removed from the lens equation by subtracting the constant angle out of the source angle. This approximation of keeping the lowest order term is not a good one, however, since the shear due to LSS arises from *relative* deflections between different light rays, which involve higher order terms in the expansion. We follow a similar approach but include these higher order terms in order to study the observable effects of LSS.

In this chapter we analyze the effect of LSS on the lens equation and time delay. We derive the lens equation in §§3.2 and 3.3, and find it to have a form similar to the generalized quadrupole lens of Kovner 1987 (§3.4). We express the perturbed lens equation in terms of integrals along the line of sight of the scalar, Newtonian potential. These integrals are random variables of zero mean, whose variances and covariances can be evaluated in terms of the power spectrum of density perturbations (§3.5). We find that the effective shear in the lens equation is not simply the integrated shear from the observer to the source, but is reduced by 40% or more, depending on the lens redshift. For realistic power spectra that include modeling of non-linear effects, the effective shear is of order 6% for a source at redshift 3. In addition, the accumulated shear from the observer to the lens can significantly affect the observables as well as the appearance of the lens itself, if the lens is at a relatively high redshift. In §3.3 we also discuss how our results determine the effect of LSS on angular diameter distances.

The most important effect of shear is in producing four-image systems (quads). Many confirmed lens systems are quads, since they are easy to identify and tend to be highly magnified (Kochanek 1991b, 1995; Wallington & Narayan 1993). These systems are inconsistent with an axi-symmetric lens, for which all the images would have to be colinear. Lens models of quads typically find a shear of order $7\% - 11\%$ (e.g. Kochanek 1991a). If due to the lensing galaxy itself, this would imply a projected ellipticity (= 1 minus the ratio of minor to major axis) for the mass of $\approx 35\% - 50\%$. By contrast, the typical value observed for ellipticals is $\approx$ 20% (e.g. Ryden 1992; Schechter 1987). Since the cross section for producing quads increases with shear, observed quads should be biased towards high shear, whatever its origin (Kassiola & Kovner 1993). In particular this includes a bias toward an alignment between the shear caused by the galaxy and the external shear. High resolution observations of lensing galaxies can determine the degree of agreement or inconsistency between the observed ellipticity and the inferred shear in specific cases. Recent observations of a four-image "Einstein cross" with the Hubble Space Telescope WFPC2 (Ratnatunga et al. 1995) found an ellipticity in the potential of 26%, which implies a mass ellipticity

of 60%. The light distribution was found to have an ellipticity of only 32%. One possible explanation is that the dark matter halo is highly flattened compared to the light distribution, but other observations may not support the existence of such large differences in typical galaxies (for a recent review see Sackett 1996). Another possibility is that a LSS shear of order 8% has been added on to the 7% shear of the lens. In fact, the directions of the total shear and that due to the light distributions are different by about 13°, so the disagreement is larger. A recent HST observation of a lensed arc (Eisenhardt et al. 1996) similarly found an observed ellipticity of about half that implied by the best fit lens model. Note, however that other possible sources of external shear, namely additional galaxies or clusters near the line of sight to the source, must be properly accounted for before the contribution of LSS can be determined.

In §3.4 we also consider the effect of LSS on relative time delays of images. The related phenomenon of amplification of sources due to large-scale structure has been studied by Babul & Lee (1991), but not in the presence of a primary lens. We show that the effect on time delays is enhanced through a combination of two separate effects. LSS thus limits our ability to derive accurate values of the Hubble constant from lensing. The induced uncertainty depends on the lens and source redshifts and on the LSS power spectrum, but in §3.5 we find it to be of order $5\% - 10\%$ at 1σ. This uncertainty may have either sign since LSS may effectively produce a negative mass density (negative is measured w.r.t. the mean density of the universe, not w.r.t. zero).

In §3.6 we choose a simple lens distribution, a singular isothermal sphere, and illustrate the effect of LSS on relative image positions and time delays, as well as the caustics and critical curves of the lens system. In §3.7 we apply our formalism to the transition from strong to weak lensing and demonstrate its agreement with previous studies of weak lensing (a detailed derivation is given in Appendix B, §3.10). Finally, in §3.8 we give our conclusions and comment on possible applications of our results.

We assume a flat universe throughout, in the absence of an accurate fitting formula for the time dependent, non-linear power spectrum in a curved background. Our formalism is, however, easily generalized to a closed or open universe, as we show in Appendix A (§3.9).

3.2 Formalism

We work in the framework of a flat Robertson-Walker metric with small-amplitude scalar metric fluctuations. In the longitudinal gauge (Bardeen 1980) we can write the line element as

$$ds^2 = a^2(\tau)[-(1+2\phi)d\tau^2 + (1-2\phi)d\vec{x}\cdot d\vec{x}\ ]\ , \tag{3.1}$$

where we set c = 1. Here τ is the conformal time, $a(\tau)$ the expansion factor, and we are using comoving coordinates $\vec{x}$. Redshift in a flat, matter-dominated ($\Omega_m = 1$) universe is given by $1/(1+z) = a(\tau) = (H_0\tau/2)^2$, with $H_0 = 100h\,\mathrm{km\,sec^{-1}\,Mpc^{-1}}$

the present Hubble constant. Also ϕ is the scalar, Newtonian potential obeying the cosmological Poisson equation

$$\nabla^2\phi = 4\pi G a^2 \bar{\rho}\ \delta\ , \tag{3.2}$$

where $\bar{\rho}$ is the mean density of the universe and $\delta = (\rho - \bar{\rho})/\bar{\rho}$ is the local density perturbation. We describe statistical properties of ϕ in terms of its Fourier transform $\phi(\vec{k},\tau)$, where $\phi(\vec{x},\tau) = \int d^3k \phi(\vec{k},\tau) e^{i\vec{k}\cdot\vec{x}}$. Its ensemble mean and variance are $\langle\phi(\vec{k},\tau)\rangle = 0$ and $\langle\phi(\vec{k},\tau)\phi^*(\vec{k}',\tau)\rangle = P_\phi(k,\tau)\delta^3(\vec{k}-\vec{k}')$, where $P_\phi(k,\tau)$ is the power spectrum of the potential at time τ, simply related to the density power spectrum by $P_\phi(k,\tau) = \left(4\pi G a^2(\tau)\bar{\rho}(\tau)\right)^2 k^{-4} P_\rho(k,\tau)$.

We place the observer at the origin of coordinates and the primary lens [2] on the z-axis. We use r to denote values of the z-coordinate, with z_L and z_S reserved for lens and source redshift, respectively. Note that the z-axis is only a coordinate axis used for reference and not the actual path of any light ray. We let $\vec{n}$ denote a unit vector in the photon's direction of motion and $\vec{x}$ denote its position. To first order in ϕ, in the metric (3.1) they obey

$$\frac{d\vec{n}}{d\tau} = -2\left[\vec{\nabla}\phi - \vec{n}(\vec{n}\cdot\vec{\nabla}\phi)\right]\ , \qquad \frac{d\vec{x}}{d\tau} = \vec{n}(1+2\phi)\ . \tag{3.3}$$

In this and similar expressions in this section, ϕ is to be evaluated on the actual photon path, not on the z-axis.

We now assume that the angle between $\vec{n}$ and the z-axis is small (e.g. Seljak 1994), and consider the components of $\vec{n}$ and $\vec{x}$ perpendicular to the z-axis. They obey

$$\frac{d\vec{n}_\perp}{d\tau} = -2\vec{\nabla}_\perp\phi\ , \qquad \frac{d\vec{x}_\perp}{d\tau} = \vec{n}_\perp\ , \tag{3.4}$$

where $\vec{\nabla}_\perp\phi$ denotes the derivative of the potential transverse to the z-axis. In the approximation of small angles, these equations are the same as the Newtonian equations of motion for a particle moving in a gravitational field, except for the factor of 2 from General Relativity. The absolute mean of ϕ is not observable, since the perturbations in the metric are defined about the large-scale mean. Indeed, we may choose our space and time units so that the large-scale value of ϕ is zero at the Local Group. Then ϕ is a random variable with rms value of order 10^{-4} for the observed LSS power spectrum. Equation (3.3) implies that the photon path obeys $r(\tau) = \tau_0 - \tau$ with $O(\phi)$ corrections, where τ_0 is the present value of τ. The relation of comoving distance to redshift is, e.g. in an Einstein-de Sitter universe, $r(z) = 2H_0^{-1}[1-(1+z)^{-1/2}]$. Thus comoving distances are simply related to the measured redshifts (to $O(\phi)$), and so we use comoving distances rather than angular diameter distances (r_L and r_S refer below to the lens and source, respectively). In a homogeneous universe with no LSS, angular diameter distances are given by $D = r/(1+z)$. In general, if an object at

[2] "The lens" refers to some reference point in the lens plane, such as the center if the lens is axi-symmetric.

comoving distance r has a proper diameter R and is observed to subtend an angle θ, then the angular diameter distance is defined to be $D = R/\theta$. This differs from D in the homogeneous case by terms larger than $O(\phi)$. As explained in §3.1, this is precisely the effect which we calculate below, and so we discuss angular diameter distances further in §3.3.

We can trace the photon trajectory backwards in time using equations (3.4), with the final conditions $\vec{x}_\perp = 0$ and $\vec{n}_\perp = \vec{n}^0_\perp$ at the observer $r = 0$. Between the observer and the lens, we find that

$$\begin{aligned} \vec{n}_\perp(\tau) &= \vec{n}^0_\perp + 2\int_\tau^{\tau_0} \vec{\nabla}_\perp \phi(\tau')d\tau' \ , \\ \vec{x}_\perp(\tau) &= -(\tau_0 - \tau)\vec{n}^0_\perp - 2\int_\tau^{\tau_0} (\tau' - \tau)\vec{\nabla}_\perp \phi(\tau')d\tau' \ . \end{aligned} \tag{3.5}$$

When the photon is at the lens, its direction of motion is $\vec{n}_\perp(\tau_L)$. It is then deflected so that at the source side of the lens its direction of motion is $\vec{p}_\perp = \vec{n}_\perp(\tau_L) + \vec{\gamma}$. The deflection angle $\vec{\gamma}$ is evaluated at $\vec{x}^L_\perp \equiv \vec{x}_\perp(\tau_L)$, and is determined by the mass distribution of the primary lens. Equations (3.4) then imply that, between the lens and the source,

$$\begin{aligned} \vec{n}_\perp(\tau) &= \vec{p}_\perp + 2\int_\tau^{\tau_L} \vec{\nabla}_\perp \phi(\tau')d\tau' \ , \\ \vec{x}_\perp(\tau) &= \vec{x}_\perp(\tau_L) - (\tau_L - \tau)\vec{p}_\perp - 2\int_\tau^{\tau_L} (\tau' - \tau)\vec{\nabla}_\perp \phi(\tau')d\tau' \ . \end{aligned} \tag{3.6}$$

For a given source position $\vec{x}^S_\perp$, the lens equation is then $\vec{x}_\perp(\tau_S) = \vec{x}^S_\perp$.

The total proper time delay, relative to the $\phi = 0$ path along the z-axis, is given by (e.g. Schneider et al. 1992)

$$\Delta t = \int_{\tau_S}^{\tau_0} \left[\frac{1}{2}\left(\frac{d\vec{x}_\perp}{d\tau}\right)^2 - 2\phi \right] d\tau \ - (1 + z_L)\psi(\vec{x}^L_\perp) \ . \tag{3.7}$$

The first term is the geometrical time delay, the second is the potential time delay due to LSS, and the last is the potential time delay due to the lens, given by

$$\psi(\vec{x}_\perp) = 4G \int d^2\xi' \Sigma(\vec{\xi}') \ \log \left| \frac{\vec{\xi} - \vec{\xi}'}{(1 + z_L)^{-1} r_L} \right| \ , \tag{3.8}$$

where $\Sigma(\vec{\xi})$ is the projected mass density of the lens, and $\vec{\xi} = (1 + z_L)^{-1}\vec{x}_\perp$ measures *proper* distance in the lens plane. We let $r_{LS} = r_S - r_L$, and then the scaled deflection angle is given by

$$\vec{\alpha} = \frac{r_{LS}}{r_S}\vec{\gamma} = \frac{r_{LS}}{r_S}\frac{\partial \psi}{\partial \vec{\xi}} \ . \tag{3.9}$$

3.3 Lensing in the presence of LSS

Equations (3.5) and (3.6) cannot in general be solved analytically, since they involve integrals over the potential ϕ evaluated on the (unknown) photon path. We therefore expand ϕ about its value on the z-axis (as in Surpi et al. 1996):

$$\phi(r\hat{z} + \vec{x}_\perp) \approx \phi + \vec{x}_\perp \cdot \vec{\nabla}_\perp \phi \ + \frac{1}{2}(\vec{x}_\perp \cdot \vec{\nabla}_\perp)^2 \phi \ , \tag{3.10}$$

where the right-hand side (RHS) is evaluated on the z-axis. The second term on the RHS leads to an unobservable constant deflection, and the third to a relative deflection between light rays. Unlike Surpi et al., we include the third term. To lowest order, in the resulting expansion for $\vec{\nabla}_\perp \phi$ we substitute for $\vec{x}_\perp$ the expressions given in equations (3.5) and (3.6) with ϕ evaluated on the z-axis. Hereafter, all expressions involving ϕ are evaluated on the z-axis.

The expansion (3.10) should be valid as long as the LSS power on scales smaller than the deflection $\vec{x}_\perp$ is negligible. We find below, however, that the shear is produced by modes over a broad range of wavelengths. Moreover, the higher order terms in the expansion formally diverge at small wavelengths in an rms sense, e.g. for a scale-invariant spectrum, at fixed $\vec{x}_\perp$. In reality, $\vec{x}_\perp$ depends on the initial direction and on ϕ. This worry is resolved by using a different expansion, equivalent to summing this entire series (see §3.7). This alternate expansion is convergent, and shows that the contribution of small wavelength modes is cut off. For strong lensing we find that the terms in equation (3.10) suffice for an accurate analysis. Note that we have not assumed at any point that $\delta < 1$ for the density. Our expansions remain valid even when we include non-linear modes for which $\delta \gg 1$.

We are not interested in any deflection which is common to all light rays, since such a constant angle only affects the unobservable absolute position of the source. We can subtract out such terms to all orders simply by measuring displacements relative to some light ray instead of the z-axis. We define this fiducial ray as the light ray (null geodesic) passing through the observer and through the lens, and extended out to r_S (see figure 3-1). This ray is deflected by LSS throughout its path, but is not deflected by the primary lens. The quantities $\vec{x}_\perp$ and $\vec{n}_\perp$ measured relative to the corresponding quantities for this fiducial ray we denote by $\vec{d}_\perp$ and $\vec{l}_\perp$, respectively. Then $\vec{\theta} = -\vec{l}_\perp^{\,O}$ is the observed angle of a light ray relative to the observed lens position. Note that $\vec{d}_\perp^{\,L} = \vec{x}_\perp^{\,L} \equiv r_L \vec{X}$.

We define dimensionless 2×2 symmetric tensors,

$$\begin{aligned} \mathrm{F}_{ij}(\tau_1, \tau_2) &= -\frac{2}{\tau_1 - \tau_2} \int_{\tau_2}^{\tau_1} (\tau - \tau_2)(\tau_1 - \tau) \partial_i \partial_j \phi(\tau) d\tau \ , \\ \mathrm{G}_{ij}(\tau_1, \tau_2) &= -2 \int_{\tau_2}^{\tau_1} (\tau_1 - \tau) \partial_i \partial_j \phi(\tau) d\tau \ . \end{aligned} \tag{3.11}$$

The traceless part of F_{ij} is the shear tensor of weak lensing (e.g. Kaiser 1992).

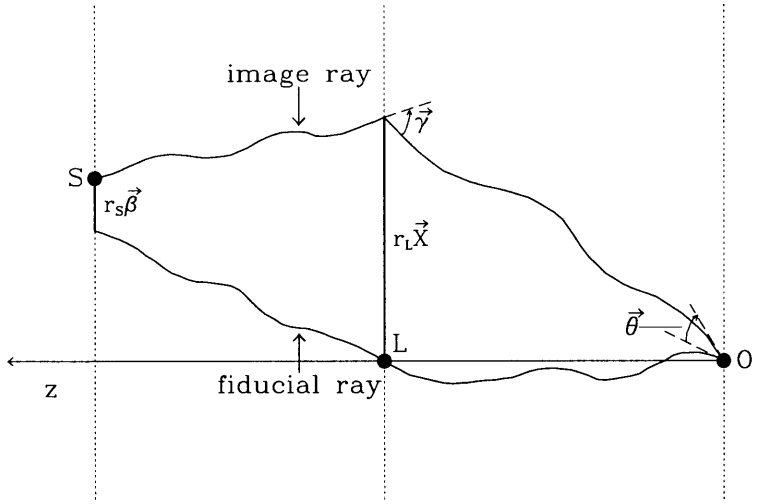

Figure 3-1: Sketch showing the fiducial ray and an image ray, with distances in comoving coordinates.

Between the observer and the lens,

$$\begin{aligned} l^i_\perp(\tau) &= -\theta^i - \theta_j \mathrm{G}^{ij}(\tau_0,\tau) \ , \\ d^i_\perp(\tau) &= (\tau_0-\tau)\left[\theta^i + \theta_j \mathrm{F}^{ij}(\tau_0,\tau)\right] \ . \end{aligned} \tag{3.12}$$

Equations (3.12) suggest a simple physical interpretation for the two tensors, in our approximation. For two rays that end up at the origin at τ_0 with a small angular separation $\vec{\theta}$, $(\tau_0-\tau)\mathrm{F}_{ij}\theta^j$ measures the change in their relative separation at τ, compared to having no LSS. [3] $\mathrm{G}_{ij}\theta^j$ similarly measures the induced change in their relative directions at time τ.

If we let $\vec{p}_\perp = \vec{l}_\perp(\tau_L) + \vec{\gamma}\left(\vec{d}^L_\perp\right)$ then, between the lens and the source,

$$\begin{aligned} l^i_\perp(\tau) &= p^i_\perp + p^j_\perp \mathrm{G}^i_j(\tau_L,\tau) - d^j_\perp(\tau_L)\left[\mathrm{G}^i_j(\tau_L,\tau) + \mathrm{G}^i_j(\tau,\tau_L)\right]/(\tau_L-\tau) \ , \\ d^i_\perp(\tau) &= d^i_\perp(\tau_L) - (\tau_L-\tau)\left[p^i_\perp + p^j_\perp \mathrm{F}^i_j(\tau_L,\tau)\right] + d^j_\perp(\tau_L)\mathrm{G}^i_j(\tau,\tau_L) \ . \end{aligned} \tag{3.13}$$

The lens equation is $\vec{d}_\perp(\tau_S) = \vec{d}^S_\perp$.

Defining $\vec{\beta} = \vec{d}^S_\perp/r_S$, and denoting e.g. $\mathrm{F}^{ij}(\tau_0,\tau_L)$ by F^{ij}_{OL}, the lens equation becomes

$$\beta^i = \theta^i + \theta_j \mathrm{F}^{ij}_{OS} - (\alpha^i + \alpha_j \mathrm{F}^{ij}_{LS}) \ , \tag{3.14}$$

where $\vec{\alpha}$ is evaluated at

$$d^i_\perp(\tau_L) = r_L(\theta^i + \theta_j \mathrm{F}^{ij}_{OL}) \ . \tag{3.15}$$

We thus conclude that to our order of approximation, LSS affects the lens equation through three terms, which are easily understood. Two rays separated by an angle θ^i at the observer would, in the absence of lensing or LSS, be separated by a (comoving)

[3] Repeated indices are summed over the x and y directions. There is no distinction between upper and lower indices.

distance $r_L\theta^i$ at the lens and $r_S\theta^i$ at the source. The LSS shear changes these distances to $r_L(\theta^i + \theta_j \mathrm{F}^{ij}_{OL})$ and $r_S(\theta^i + \theta_j \mathrm{F}^{ij}_{OS})$ respectively. The deflection between the two rays by an angle $-\gamma^i$ at the lens leads to an additional separation of $-r_{LS}\gamma^i = -r_S\alpha^i$ at the source, or $-r_S(\alpha^i + \alpha_j \mathrm{F}^{ij}_{LS})$ when we include the effect of LSS shear. There are no cross-terms between these three effects in our approximation, where only terms linear in $\vec{\theta}$ and $\vec{\alpha}$ appear in $\vec{d}_\perp$.

The magnification matrix is

$$\frac{\partial\beta^i}{\partial\theta_j} = \delta^{ij} - (-\mathrm{F}^{ij}_{OS} + \Psi^{ij} + \Psi^i_k \mathrm{F}^{kj}_{OL} + \Psi^j_k \mathrm{F}^{ki}_{LS}) \,, \tag{3.16}$$

where $\Psi^{ij} = \partial_{\xi_i}\partial_{\xi_j}\psi/(4\pi G\Sigma_{crit})$ is the shear matrix of the primary lens, in units of $\Sigma_{crit} = r_S(1+z_L)/(4\pi G r_L r_{LS})$, and Ψ^{ij} in equation (3.16) is evaluated at $\vec{d}^L_\perp$. With the usual sign conventions in lensing, the constant LSS shear is $-\mathrm{F}^{ij}_{OS}$. This term would still be present even in the absence of the lens (see §3.7 below). Note that $\partial\beta^i/\partial\theta_j$ is in general *not* symmetric, which it would be in the absence of LSS. In other words, LSS can rotate images.

As noted in §3.1, we have also calculated the effect of LSS on angular diameter distances. Indeed, an object which subtends an angle $\vec{\theta}$ at the observer measures a comoving distance $\vec{d}^L_\perp$ on the lens plane, given by equation (3.15). The same object measures a proper distance $\vec{R} = \vec{d}^L_\perp/(1+z_L)$, which follows (to $O(\phi)$) from the line element (3.1) taken at constant τ. Then $R^i = D^{ij}_{OL}\theta_j$, where the angular diameter "distance" $D^{ij}_{OL} = (\delta^{ij} + F^{ij}_{OL})r_L/(1+z_L)$ is a tensor when LSS is present. Thus R at a given θ depends on orientation, and also $\vec{R}$ may have a different direction than $\vec{\theta}$, so when giving "distances" to lenses it is preferable to use the comoving distance r_L which is well defined (up to corrections of $O(\phi)$, i.e. 0.01%) in terms of z_L. As we show in §3.5, the components of F_{OL} are of order a few percent, much larger than $O(\phi)$ corrections. Some other authors (e.g. Ehlers & Schneider 1986; Watanabe 1992; Sasaki 1993) have also considered the effect of LSS on angular diameter distances, but they used an oversimplified model in which some fraction of the mass density in the universe is distributed in clumps. Theory and observation of LSS indicate that a description in terms of a random field with positive and negative fluctuations over a range of scales is more realistic (see also the related discussion in Seljak 1994).

3.4 The lens equation and time delay

The lens equation (3.14) is similar in form to the generalized quadrupole lens of Kovner (1987). Kovner cosidered multiple lensing in which there is one primary lens and additional lenses with linear deflection laws. In that case Kovner showed how to write the lens equation in the form of a thin-lens equation, which simplifies the analysis of properties of the lens mapping, such as image multiplicities and the time delay between images. LSS is different in that the deflection is accumulated

continuously, but the final result can be similarly simplified. Letting

$$X^i = \theta^i + \theta_j \mathrm{F}^{ij}_{OL} \,, \tag{3.17}$$

$$Y^i = \beta^i - \beta_j \mathrm{F}^{ij}_{LS} \,, \tag{3.18}$$

the lens equation becomes

$$Y^i = X^i - \left[X_j \mathrm{F}^{ij}_{\mathrm{eff}} + \alpha^i(\vec{X}) \right] \,, \tag{3.19}$$

$$\mathrm{F}^{ij}_{\mathrm{eff}} = -\mathrm{F}^{ij}_{OS} + \mathrm{F}^{ij}_{LS} + \mathrm{F}^{ij}_{OL} \,, \tag{3.20}$$

where we write $\vec{\alpha}$ as a function of $\vec{X}$ rather than $r_L\vec{X}$. We find that $\delta^{ij} - \mathrm{F}^{ij}_{\mathrm{eff}}$ plays the same role as the "telescope matrix" of Kovner, which in his case is in general symmetric. We find a symmetric $\mathrm{F}_{\mathrm{eff}}$ only because we are working to first order in the LSS shear. The effective shear $\mathrm{F}^{ij}_{\mathrm{eff}}$ is in general significantly weaker than F^{ij}_{OS}, as we show in §3.5. Still, this shear should be of order 6% r.m.s., compared e.g. with a galaxy of ellipticity 20%, which produces a shear of $\approx 4\%$.

We thus have a simple description of the lens mapping: The source plane is slightly distorted to become the $\vec{Y}$ plane, as given by equation (3.18), so e.g. a circular source appears elliptical in the $\vec{Y}$ plane. Equation (3.19) then gives the lens mapping from the $\vec{Y}$ plane to the $\vec{X}$ plane. Finally, the (observed) image plane is also a slightly distorted picture of the $\vec{X}$ plane, as in equation (3.17). Only the $\vec{Y} \mapsto \vec{X}$ map is non-linear, and it determines the geometry of the lens mapping. Thus e.g. the probability of producing quads depends on the sum of $\mathrm{F}_{\mathrm{eff}}$ and any intrinsic asymmetric shear from the ellipticity of the lens galaxy. The shear $\mathrm{F}_{\mathrm{eff}}$ should tend to make the observed galaxy light distribution inconsistent with the observed lensing. If the lens is at high redshift, however, then the distortion in equation (3.17) is also important, since $\vec{\theta}$ is observed and not $\vec{X}$. In this case, the lens itself is distorted by LSS, if it is observed. Because this induced ellipticity is likely to be wrongly interpreted as intrinsic to the galaxy, it tends to confuse observers as to the actual direction of the galaxy's internal shear, but the effect is important only if the intrinsic ellipticity itself is not too large. Since the source plane is not directly observable, the distortion in equation (3.18) does not affect lens reconstruction, but it is important for absolute magnifications (given by equation (3.16)), and for measuring shape distortions (§3.7).

We can calculate the time delay explicitly using equation (3.7). However, it is easier to use Fermat's principle, which implies that (for a given $\phi(\vec{x}, \tau)$) the lens equation must be equivalent to $\partial \Delta t / \partial \vec{\theta} = 0$ at fixed $\vec{\beta}$ (e.g. Schneider et al. 1992). Thus the time delay is the same as that corresponding to the thin-lens equation (3.19) which, up to $\vec{X}$-independent terms, equals [18]

$$\Delta t = \frac{1}{2} \frac{r_L r_S}{r_{LS}} \left[(\vec{X} - \vec{Y})^2 - \mathrm{F}^{ij}_{\mathrm{eff}} X_i X_j \right] - (1 + z_L) \psi(r_L \vec{X}) \,. \tag{3.21}$$

We might have expected linear terms of the form $\theta^i \mathrm{C}_i$ to make large contributions to Δt, where C_i is independent of $\vec{\theta}$ and $\vec{\beta}$, e.g. $\mathrm{C}_i = r_{LS} \int_{\tau_S}^{\tau_L} \partial_i \phi(\tau) d\tau$ or $\mathrm{C}^i =$

$F^{ij}_{LS} \int_{\tau_S}^{\tau_L} (\tau_L - \tau) \partial_j \phi(\tau) d\tau$. Such terms do appear in the geometric and potential time delays, but Fermat's principle shows that they must drop out in the total time delay. These cancellations can also be demonstrated explicitly with equation (3.7).

In addition to shear effects, the lens geometry is also affected by the trace parts of $F_{\rm eff}$ and F_{OL}. These traces cannot be determined through lens reconstruction, since they only affect the unobservable overall scales of the lens size and distance. However, they do affect the determination of the Hubble constant from lensing. To show the various effects, we first set $F^{ij}_{OL} = 0$, and consider just equation (3.19) and the effect of the trace part of $F_{\rm eff}$. We also allow for a focusing term κ_{cl} due to a cluster surrounding the lens galaxy. Equation (5) of Narayan (1991) applies here:

$$\Delta t \propto \sigma^2 r_L \ , \tag{3.22}$$

where σ represents some characteristic velocity of the lens system. The proportionality constant in this equation depends on a number of parameters which can in principle be determined for each pair of images from lens modeling. One of these parameters involves σ and $\kappa_{\rm cl}$, in the combination

$$\zeta = \frac{r_S}{\sigma^2 r_{LS}} \left[1 - \left(\frac{1}{2} {\rm Tr}\ F_{\rm eff} + \kappa_{\rm cl} \right) \right] \tag{3.23}$$

(as follows from equation (6) of Narayan). Thus if Δt is measured (and σ is not) then from the product $\zeta \Delta t$, we can try to deduce H_0 given z_L, z_S, and an assumed deceleration parameter q_0. The real H_0 is different from the H_0 deduced assuming $\kappa_{\rm cl} = {\rm Tr}\ F_{\rm eff} = 0$ by a factor of $[1 - (\frac{1}{2}{\rm Tr}\ F_{\rm eff} + \kappa_{\rm cl})]$. If both Δt and σ can be measured independently, then as noted by Narayan we can (if $F^{ij}_{OL} = 0$) circumvent the unknowns in equation (3.23) and use equation (3.22) to obtain r_L directly, and thus H_0 for an assumed q_0.

We now add in also the effect of the trace part of F_{OL} in equation (3.17), which is an unobservable magnification of the lens plane produced by foreground structure. Since the time delay is proportional to the square of the angular scale (e.g. Schneider et al. 1992), equation (3.22) is replaced by

$$\Delta t \propto \sigma^2 r_L \left(1 + \frac{1}{2} {\rm Tr}\ F_{OL} \right)^2 \ . \tag{3.24}$$

Reasoning as above, we see that if only Δt is measured then (from equations (3.23) and (3.24)) the real global H_0 is different from the deduced H_0 (for an assumed q_0) by a factor of $[1 + \frac{1}{2}{\rm Tr}(2F_{OL} - F_{\rm eff}) - \kappa_{\rm cl}]$, to linear order. LSS thus produces a 1σ uncertainty in determinations of H_0 of

$$\sigma_{H_0,1} = \ {\rm rms\ of}\ \frac{1}{2}\ {\rm Tr}\ (2F_{OL} - F_{\rm eff}) \ . \tag{3.25}$$

Contrary to Falco et al. (1985), we cannot derive an upper bound on H_0 from the Δt measurement since while $\kappa_{\rm cl} \geq 0$, the LSS term may be negative or positive. Even with measurements of both Δt and σ^2, we cannot measure H_0 exactly, since when we

use equation (3.24) we are subject (for a given q_0) to a 1σ uncertainty of

$$\sigma_{H_0,2} = \text{ rms of Tr } \mathrm{F}_{OL} \ . \tag{3.26}$$

Thus, LSS creates uncertainties in determinations of the Hubble constant from lensing which apply even to perfect lens models determined by an arbitrarily large number of observables. If a precise measurement of H_0 is sought from lens time delays, then at least for some redshifts and LSS power spectra these uncertainties may not be very small, as we show in §3.5.

3.5 LSS effects in realistic models

We have shown above that the effects of LSS on lensing enter through the symmetric tensors F_{OL}, F_{LS} and F_{OS}. For a given lens and source, these tensors will affect the lens mapping as we showed in §3.4, possibly with observable effects which we illustrate in §3.6. In this section we estimate the typical magnitude of these tensors that is expected based on theory and observation of LSS, and its dependence on the redshifts of the lens and source.

The tensor components are random variables of zero mean, with variances and covariances given in terms of the power spectrum of ϕ. For example, if $\tau_1 \geq \tau_2 \geq \tau_3$, then following the method of Kaiser (1992) we find that

$$\begin{aligned} \langle \mathrm{F}_{ij}(\tau_1,\tau_2)\, \mathrm{F}_{kl}(\tau_1,\tau_3)\rangle \;&=\; 2\pi^2 \mathrm{Q}_{ijkl} \int_0^\infty k^5 dk \\ &\int_{\tau_2}^{\tau_1} \frac{(\tau-\tau_2)(\tau_1-\tau)}{\tau_1-\tau_2} \frac{(\tau-\tau_3)(\tau_1-\tau)}{\tau_1-\tau_3} P_\phi(k,\tau) d\tau \ , \end{aligned} \tag{3.27}$$

$$\text{where } \mathrm{Q}_{ijkl} = \begin{cases} 3 & \text{if } ijkl \text{ are all equal.} \\ 1 & \text{if of } ijkl \text{ two} = x \text{ and two} = y. \\ 0 & \text{otherwise.} \end{cases}$$

This assumes that the dominant contribution comes from modes with wavelengths that are much smaller than the distance $\tau_1 - \tau_2$. This is satisfied for standard forms of the power spectrum and relevant distances.

We follow the approach of Seljak (1996) for calculating rms shear. For the linear power spectrum we take a scale-invariant spectrum with a CDM type transfer function (Bardeen et al. 1986), which is normalized by σ_8, the mass fluctuation averaged over a sphere of radius $8h^{-1}$ Mpc, and whose peak is determined by $\Omega_{m0}h$. Galaxy and cluster surveys are consistent with $\sigma_8 \approx 0.8$ and $\Omega_{m0}h \approx 0.25$ (e.g. Peacock & Dodds 1994). We then find the non-linear power spectrum using the mapping proposed by Hamilton et al. (1991) and extended by Peacock & Dodds, in the improved form of Jain et al. (1995), which they show agrees with N-body simulations at the relevant scales, for an $\Omega_m = 1$ universe. We find that the dominant contribution in equation (3.27) comes from wavenumbers $k \approx 3$ h Mpc^{-1}, with a broad range of two decades on each side contributing significantly. We therefore require a power spectrum that is accurate well into the non-linear regime.

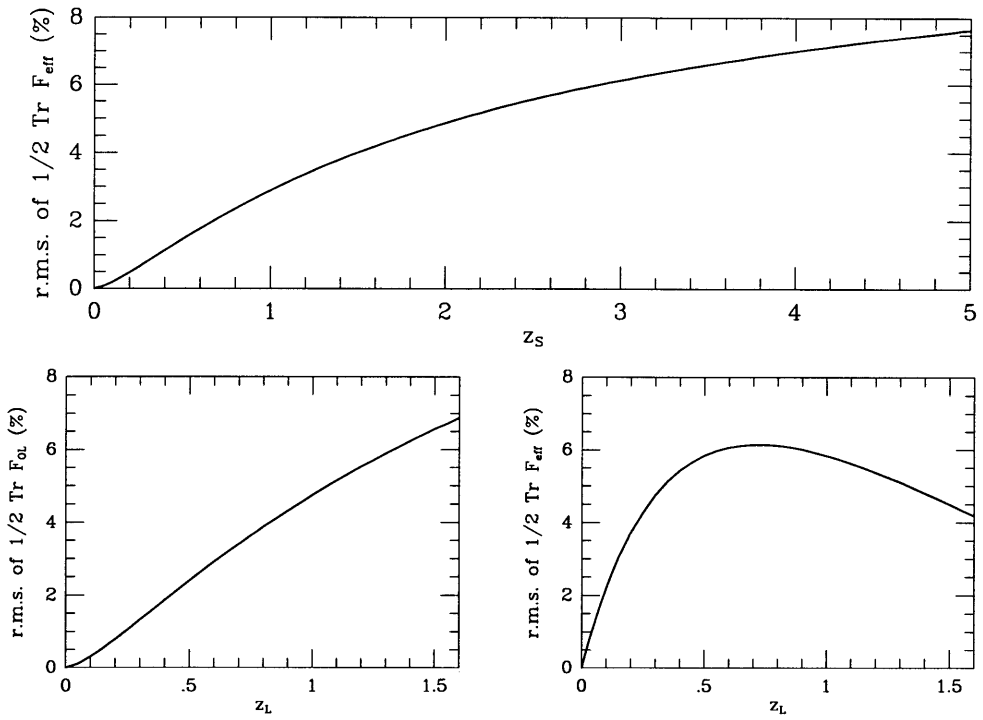

Figure 3-2: The top plot shows the rms value of $\frac{1}{2}$ Tr $\mathrm{F_{eff}}$, as a function of z_S, with z_L set so that $r_L = \frac{1}{2} r_S$. The bottom right plot shows the same quantity, but as a function of z_L, for fixed $z_S = 3$. The bottom left plot shows the rms value of $\frac{1}{2}$ Tr F_{OL} as a function of z_L. All curves use the non-linear power spectrum, with $\Omega_m = 1$, h= 0.25, and $\sigma_8 = 0.8$.

The shear due to F_{OS} is defined as $\Gamma = \sqrt{(\Gamma_1)^2 + (\Gamma_2)^2}$, where $\Gamma_1 = \frac{1}{2}(\mathrm{F}^{11}_{OS} - \mathrm{F}^{22}_{OS})$ and $\Gamma_2 = \mathrm{F}^{12}_{OS}$. Equation (3.27) shows that Γ has the same rms value as $\frac{1}{2}$ Tr F_{OS}, which is the convergence or surface mass density κ due to F_{OS}. The same is true for $\mathrm{F_{eff}}$. Because each of the tensors F_{OL} and F_{LS} is correlated (and so tends to be aligned) with F_{OS}, $\mathrm{F_{eff}}$ tends to have smaller components than F_{OS}. For a given z_S, the rms shear of $\mathrm{F_{eff}}$ is maximized at $\approx 60\%$ of that of F_{OS}, approximately at z_L for which $r_L = \frac{1}{2} r_S$. In the top plot of figure 3-2 we show the rms value of $\frac{1}{2}$ Tr $\mathrm{F_{eff}}$ as a function of z_S, at this maximizing z_L. In the bottom right plot we show the same quantity, but as a function of z_L, for a fixed source at redshift 3. This quantity can be estimated for other redshifts through its scaling $\propto r_L r_{LS}/\sqrt{r_S}$, which is accurate to better than 10%. The bottom left plot shows the rms value of $\frac{1}{2}$ Tr F_{OL} as a function of z_L. This quantity scales approximately as $\propto r_L^{3/2}$, and also equals $\frac{1}{2}\sigma_{H_0,2}$ as shown in §3.4. All curves use the non-linear power spectrum, which gives rms shear larger than the linear spectrum by a factor of ≈ 2.5. Equation (3.27) gives a statistical tendency for perpendicular alignment between F_{OL} and $\mathrm{F_{eff}}$, which increases $\sigma_{H_0,1}$ relative to $\frac{1}{2}$ Tr $\mathrm{F_{eff}}$. At $z_S = 3$ and $r_L = \frac{1}{2} r_S$, $\sigma_{H_0,1} = 11.7\%$, and it scales approximately as $\propto r_L \sqrt{r_S - r_L/3}$. Since the effect of LSS accumulates over distance, we find that the induced shears and time delay uncertainties are all smaller at lower redshifts. For the 0957+561 redshifts ($z_L = .36, z_S = 1.41$), the r.m.s. $\frac{1}{2}$ Tr $\mathrm{F_{eff}}$ is 3.7%, $\sigma_{H_0,1} = 5.9\%$, and $\sigma_{H_0,2} = 3.3\%$. In addition, note that the effects of LSS on lens reconstruction

disappear as $z_L \to 0$, even if z_S is large.

The rms shear increases with σ_8, in exact proportion for the linear power spectrum, faster for the non-linear spectrum. The rms shear also grows with h (at fixed σ_8), by $\approx 35\%$ for $h = 0.5$ compared to $h = 0.25$. As an additional example, tilted CDM (e.g. Cen et al. 1992) with $h = 0.5$, $\sigma_8 = .6$, and power spectrum index $n = 0.8$ lowers the shear by $\approx 30\%$ compared to Figure 2. The rms shear can also be calculated for models with $\Omega_{m0} \neq 1$ with modified formulas (see Appendix A, §3.9).

3.6 Illustration of the effect of LSS

Kovner (1987) analyzed in some generality the properties of the lens mapping for an axi-symmetric lens perturbed by a weak shear. We simply wish to illustrate the possible observable effects of a shear of the magnitude that we obtained in §3.5. We choose a particular symmetric lens distribution, a singular isothermal sphere, with deflection law $\vec{\alpha}(r_L\vec{\theta}) = \vec{\theta}/\theta$. We use equations (3.14) – (3.16) to find the caustics and critical curves of the lens system. The critical curves are the points in the image plane for which the magnification $\det^{-1}(\partial\beta^i/\partial\theta_j)$ is infinite, and the caustics are the corresponding points in the source plane. The caustics also determine image multiplicities, in that a source located outside all the caustics has a single image, and each time a source moves inside a caustic two images are added (except that for a singular surface density, one image is captured in the core when multiple images are produced). For a given source position, we can thus find all image positions, magnifications, and also time delays with equation (3.21).

The components F^{ij}_{OL}, etc. are random variables, with covariances obtained as in §3.5 above. We choose $z_L = 0.78$ and $z_S = 3.0$, and take a particular example:

$$\mathrm{F}_{OL} = \begin{pmatrix} -3.87 & 0.50 \\ 0.50 & -2.04 \end{pmatrix} \%\,, \quad \mathrm{F}_{LS} = \begin{pmatrix} -0.70 & 3.68 \\ 3.68 & 2.20 \end{pmatrix} \%\,, \quad \mathrm{F}_{\mathrm{eff}} = \begin{pmatrix} 6.65 & -6.56 \\ -6.56 & -0.31 \end{pmatrix} \%\,.$$

Figure 3-3 shows the caustics in the source plane (upper panels) and critical curves in the image plane (lower panels), for the lens alone and for the lens plus LSS. For the latter case, the $\vec{Y}$ (distorted source) plane and $\vec{X}$ (lens) plane are also shown. Also plotted are two source positions and the corresponding image positions. Table 3.1 lists the image positions, magnifications, and relative time delays. LSS changes the image configurations significantly. It displaces images from the line to the lens, in the two-image configuration, and also produces four-image systems when $|\vec{\beta}|$ is small.

3.7 Weak lensing and strong lensing

The approximation of equation (3.10) suffices for consideration of strong lensing, where $|\vec{\theta}|$ is very small ($\approx$ a few arcseconds). In weak lensing, however, the shear is observed at larger angles (arcminutes), and the variation of LSS shear with angle is important. Moreover, as noted above, there are potential convergence problems with our expansion of ϕ, even in the strong lensing case. Our formalism allows us to make

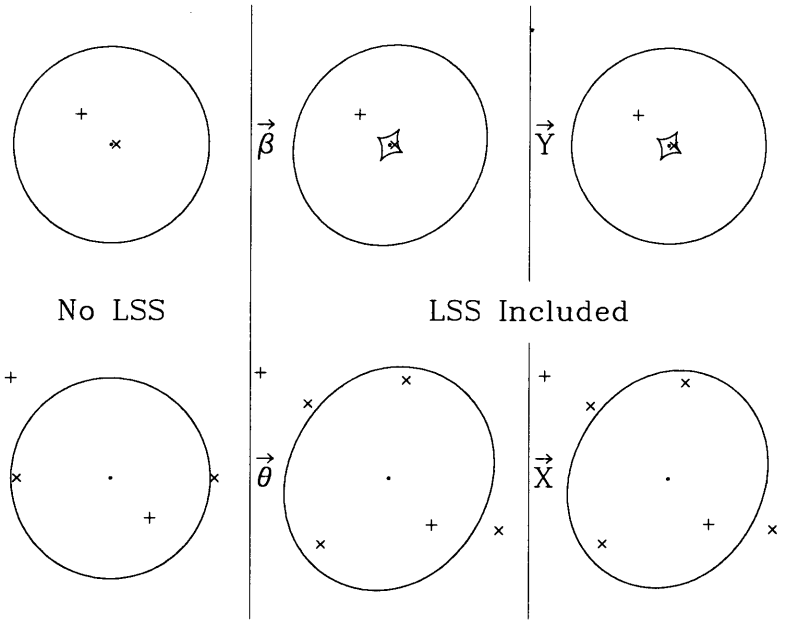

Figure 3-3: Caustics in the source plane (upper panels) and critical curves in the image plane (lower panels) for a singular isothermal sphere, with no LSS, and with LSS. For the latter case, the $\vec{Y}$ (distorted source) plane and $\vec{X}$ (lens) plane are also shown. Also plotted are two source positions (marked + and ×) and the corresponding images for each. A dot shows the $\vec{\theta} = \vec{\beta} = 0$ position.

a more rigorous and powerful expansion, and to study the transition from strong to weak lensing. Note that for shape distortions we must use $\vec{\beta}$ and $\vec{\theta}$ rather than $\vec{Y}$ or $\vec{X}$, since we are interested in the observed compared to the intrinsic ellipticity of background galaxies.

We replace equation (3.10) with

$$\phi(r\hat{z} + \vec{x}_\perp) = \sum_{n=0}^{\infty} \frac{1}{n!} (\vec{x}_\perp \cdot \vec{\nabla}_\perp)^n \phi(r\hat{z}) \ , \tag{3.28}$$

where this expression must be consistent with equations (3.5) and (3.6) for $\vec{x}_\perp$. We now include derivatives to all orders in equation (3.28), but only keep terms linear in ϕ. This means that in the following expansions in powers of $\vec{\theta}$ and $\vec{\gamma}$, we are finding each coefficient up to relative corrections of the same order as the rms LSS shear.

Between the observer and the lens, equation (3.5) requires

$$\vec{x}_\perp(\tau) = -(\tau_0 - \tau)\vec{n}^0_\perp - 2\int_\tau^{\tau_0} (\tau' - \tau) e^{\vec{x}_\perp(\tau')\cdot\vec{\nabla}_\perp} \vec{\nabla}_\perp \phi(\tau') d\tau' \ , \tag{3.29}$$

where the exponential denotes the corresponding Taylor series expansion (and ϕ on the RHS is again evaluated on the z-axis). We can find all terms linear in ϕ in the solution by substituting for $\vec{x}_\perp(\tau')$ in the RHS the ϕ-independent term $-(\tau_0 - \tau')\vec{n}^0_\perp$.

Table 3.1: Positions of the images shown in figure 3-3. Also listed are the absolute magnifications (with a sign giving the image parity), and the time delay in units of $r_L r_S / r_{LS}$ relative to the earliest image to arrive at the observer.

Source at $\vec{\beta} = (-0.30, 0.30)$, $\vec{Y} = (-0.31, 0.30)$				
Model	Image Plane ($\vec{\theta}$)	Lens Plane ($\vec{X}$)	Magnification	Relative Δt
No LSS	$(-1.00, 1.00)$	—	3.36	—
	$(0.39, -0.39)$	—	-1.36	0.84
With LSS	$(-1.30, 1.05)$	$(-1.24, 1.03)$	2.82	—
	$(0.43, -0.46)$	$(0.41, -0.45)$	-2.05	0.98
Source at $\vec{\beta} = \vec{Y} = (0.05, 0)$				
No LSS	$(1.04, 0)$	—	21.0	—
	$(-0.94, 0)$	—	-19.0	0.099
With LSS	$(1.10, -0.52)$	$(1.06, -0.50)$	6.18	—
	$(-0.82, 0.75)$	$(-0.78, 0.73)$	11.2	0.098
	$(0.18, 0.98)$	$(0.17, 0.96)$	-9.22	0.121
	$(-0.68, -0.66)$	$(-0.66, -0.65)$	-5.85	0.165

We now find that

$$\begin{aligned} l^i_\perp(\tau) &= -\theta^i + 2\int_\tau^{\tau_0} \left[e^{(\tau_0-\tau')\vec{\theta}\cdot\vec{\nabla}_\perp} - 1\right] \partial^i\phi(\tau')d\tau' \,, \\ d^i_\perp(\tau) &= (\tau_0-\tau)\theta^i - 2\int_\tau^{\tau_0} (\tau'-\tau)\left[e^{(\tau_0-\tau')\vec{\theta}\cdot\vec{\nabla}_\perp} - 1\right] \partial^i\phi(\tau')d\tau' \,. \end{aligned} \tag{3.30}$$

If we let $\vec{p}_\perp = \vec{l}_\perp(\tau_L) + \vec{\gamma}(\vec{d^L_\perp})$ then between the lens and the source we similarly find that

$$\begin{aligned} l^i_\perp(\tau) &= p^i_\perp + 2\int_\tau^{\tau_L} \left[e^{(\vec{d^L_\perp}-(\tau_L-\tau')\vec{p}_\perp)\cdot\vec{\nabla}_\perp} - 1\right] \partial^i\phi(\tau')d\tau' \,, \\ d^i_\perp(\tau) &= d^i_\perp(\tau_L) - (\tau_L-\tau)p^i_\perp - 2\int_\tau^{\tau_L} (\tau'-\tau)\left[e^{(\vec{d^L_\perp}-(\tau_L-\tau')\vec{p}_\perp)\cdot\vec{\nabla}_\perp} - 1\right] \partial^i\phi(\tau')d\tau' \,. \end{aligned} \tag{3.31}$$

The lens equation is

$$\begin{aligned} \beta^i &= \theta^i - \alpha^i - \frac{2}{r_S}\int_{\tau_L}^{\tau_0} (\tau-\tau_S)\left[e^{(\tau_0-\tau)\vec{\theta}\cdot\vec{\nabla}_\perp} - 1\right] \partial^i\phi(\tau)d\tau \\ &\quad - \frac{2}{r_S}\int_{\tau_S}^{\tau_L} (\tau-\tau_S)\left[e^{((\tau_0-\tau)\vec{\theta}-(\tau_L-\tau)\vec{\gamma})\cdot\vec{\nabla}_\perp} - 1\right] \partial^i\phi(\tau)d\tau \,, \end{aligned} \tag{3.32}$$

with $\vec{\alpha}$ and $\vec{\gamma}$ evaluated at $\vec{d^L_\perp}$ calculated from equation (3.30). The magnification matrix is

$$\frac{\partial\beta^i}{\partial\theta_j} = \delta^{ij} + \mathrm{A}^{ij} - \Psi^{ij} - \mathrm{B}^{ij} \,, \tag{3.33}$$

where

$$\begin{aligned}
\mathrm{A}_{ij} &= -\frac{2}{r_S}\int_{\tau_L}^{\tau_0} e^{(\tau_0-\tau)\vec{\theta}\cdot\vec{\nabla}_\perp}(\tau-\tau_S)(\tau_0-\tau)\partial_i\partial_j\phi(\tau)d\tau \\
&\quad -\frac{2}{r_S}\int_{\tau_S}^{\tau_L} e^{((\tau_0-\tau)\vec{\theta}-(\tau_L-\tau)\vec{\gamma})\cdot\vec{\nabla}_\perp}(\tau-\tau_S)(\tau_0-\tau)\partial_i\partial_j\phi(\tau)d\tau \ , \\
\mathrm{B}_{ij} &= -\frac{2}{r_L}\Psi_i^k\int_{\tau_L}^{\tau_0} e^{(\tau_0-\tau)\vec{\theta}\cdot\vec{\nabla}_\perp}(\tau-\tau_L)(\tau_0-\tau)\partial_k\partial_j\phi(\tau)d\tau \\
&\quad -\frac{2}{r_{LS}}\Psi_j^k\int_{\tau_S}^{\tau_L} e^{((\tau_0-\tau)\vec{\theta}-(\tau_L-\tau)\vec{\gamma})\cdot\vec{\nabla}_\perp}(\tau-\tau_S)(\tau_L-\tau)\partial_k\partial_i\phi(\tau)d\tau \ .
\end{aligned}$$

If we expand the exponentials to first order in equation (3.32) and zeroth order in equation (3.33), we recover the results of §3.3. If lensing is not strong, B_{ij} is small, and in the external shear A_{ij} we can set $\vec{\gamma} = (\vec{\theta}-\vec{\beta})r_S/r_{LS}$. In the limit of weak lensing, we can set $\vec{\gamma}=0$ to get

$$\mathrm{A}_{ij} = -\frac{2}{r_S}\int_{\tau_S}^{\tau_0} e^{(\tau_0-\tau)\vec{\theta}\cdot\vec{\nabla}_\perp}(\tau-\tau_S)(\tau_0-\tau)\partial_i\partial_j\phi(\tau)d\tau \ . \tag{3.34}$$

This expression can be used to calculate two point correlation functions of ellipticity. E.g., we can write down $\langle \mathrm{TrA}(\theta=0)\mathrm{TrA}(\theta)\rangle$ and evaluate this expectation value in Fourier space. The exponential of $ir\vec{\theta}\cdot\vec{k}$ (in Fourier space) oscillates rapidly at high k, which cuts off small wavelengths and prevents any divergence. The result, which is derived fully in Appendix B (§3.10), agrees with previous analyses of weak lensing in the absence of a primary lens (e.g. Blandford et al. 1991; Miralda-Escudé 1991; Kaiser 1992). These analyses have found that the relative change in the angular correlation of ellipticity is smaller than 10% (in an rms sense) for θ less than about an arcminute. For the non-linear spectrum, we find this to be true below about $10''$ (see also Seljak 1996), thus justifying our keeping only linear terms in $\vec{\theta}$ for strong lensing. Our result (3.33) is more general than weak lensing, as it includes a primary lens (Ψ^{ij}) and cross-terms (B^{ij}).

We can also get quadratic and higher-order terms in the gradients of ϕ by iterating this procedure. Given a solution $\vec{x}_\perp^{(j)}$ we substitute it in the RHS of equation (3.29) and find the next order solution $\vec{x}_\perp^{(j+1)}$. The exponential of $i\vec{k}\cdot\vec{x}_\perp^{(j)}$ ensures that small wavelength modes are cut off in the calculation of $\vec{x}_\perp^{(j+1)}$. This corresponds to the physical intuition that on average $\phi(r\hat{z}+\vec{x}_\perp^{(j)})-\phi(r\hat{z})$ is determined by power on scales of order $|\vec{x}_\perp^{(j)}|$. If we calculate the rms shear at a point (i.e. for $\theta=\gamma=0$) corresponding to $\vec{x}_\perp^{(j+1)}$, for a given $\vec{x}_\perp^{(j)}(\tau)$, the answer is the same as for $\vec{x}_\perp^{(j)}(\tau)=0$ as long as the angular deflection is small and LSS power on the scale of r_S is negligible. There is, of course a statistical correlation between $\vec{x}_\perp^{(j)}(\tau)$ and $\phi(\tau)$, but it is typically weak since $\vec{x}_\perp^{(j)}(\tau)$ is determined by the accumulated deflection from τ to τ_0, a distance many times larger than the coherence length of LSS. The first correction to F_{OS} is a relative correction of order 1%, if ϕ is Gaussian, and the corrections to F_{OS} are expected to be small also for a non-Gaussian ϕ produced through hierarchical clustering.

3.8 Conclusions

We have shown that LSS can have significant effects on strong gravitational lensing. This suggests that lens reconstruction done without including LSS might reach incorrect conclusions about the distribution of matter in the lensing galaxy or cluster. It also raises the possibility of constraining the amplitude of the power spectrum directly, if lensing observations can be used to detect the effect of LSS.

The effect of LSS is simply described by two symmetric tensors. Including only the effect of $\mathrm{F_{eff}}$, we find that the observed power spectrum of LSS requires the presence of an external shear of order 6% if $z_S = 3$. This can significantly affect the cross sections for image multiplicities in lens systems. In particular, it can produce more images than would be created in the absence of LSS. This implies that in addition to the usual magnification bias, which increases the observed number of quads relative to doubles, there is a bias in quads toward lines of sight with relatively large effective shear from LSS.

The second effect, given by F_{OL}, produces a magnification and shear between the observer and the lens. This term enters the lens equation differently from the effective shear and should be included in lens modeling. It also distorts the lens plane, which contributes an ellipticity to the observed lens galaxy, and converts the angular diameter "distance" into a tensor, though the comoving distance is still simply defined in terms of the observed redshift. Even if lens reconstruction can model the lens potential exactly, we find that LSS creates an absolute uncertainty ($\approx 5 - 10\%$ at 1σ) in deductions of the Hubble constant from time delays. Among lens systems, the uncertainty is smaller for those with lower lens and source redshifts.

Models of quadruple lenses typically find a shear of order 10% in addition to an axi-symmetric mass distribution. If this is due to the ellipticity of the lens galaxy, it may imply a larger ellipticity than that observed in the galaxy light distribution, as confirmed in a number of cases by recent observations. Since quads tend to be produced more easily when the shear due to the galaxy and the effective shear due to LSS are aligned, it is important to compare the magnitudes of the observed and modeled shears for consistency, and not only their directions. If the shear is due instead to other galaxies or clusters near the line of sight to the source, these additional lenses may not be found where expected. Only high-resolution observations and careful modeling of particular lens systems will determine if the shear may in part be due to LSS. When the parameters of many lenses are confidently known, it may become possible to study the redshift dependence of the shear. E.g., LSS does not affect lens reconstruction if the lens is at very low redshift. The original Einstein Cross 2237+0305 has $z_L = 0.04$, $z_S = 1.7$, and the lens light distribution seems to be consistent with the lensing mass (Rix et al. 1992). Other methods must be used to investigate independently whether the mass in galaxies is more flattened than the light distribution or not. E.g., an affirmative answer is suggested by an optical plus X-ray study of the elliptical galaxy NGC 720 (Buote & Canizares 1994).

Constraining the effects of LSS on strong lensing should complement observations of weak lensing due to LSS. For measurements of weak lensing the sources are background galaxies, and the interpretation is complicated by the unknown source

redshift distribution, while for some strong lenses the redshifts of the lens and source are known. If the characteristic source redshift for weak lensing is $\approx 0.7-1$ then the shear due to LSS is significantly smaller than for strong lensing (e.g. figure 3-2). In addition, since measurements of weak lensing with high signal to noise require relativly large angular fields, the rms shear is further reduced. On the other hand, weak lensing due to LSS can in principle be distinguished from other effects by averaging over a wide field, an option not available in strong lensing. Demanding consistency between determinations of the effects in these two regimes should allow us to learn more about the distribution of matter in the universe.

3.9 Appendix A

To calculate LSS shear in a curved background requires slight modifications of our formulas (e.g. Miralda-Escudé 1991; Seljak 1996). In a general Robertson-Walker model, the line element is

$$ds^2 = a^2(\tau)\Big[-(1+2\phi)d\tau^2 + (1-2\phi)[d\chi^2 + \sin_K^2\chi(d\theta^2+\sin^2\theta d\phi^2)]\Big] , \qquad (3.35)$$

in terms of spherical comoving coordinates, where we are now using the variable $\chi = \tau_0 - \tau$. We have defined

$$\sin_K\chi \equiv \begin{cases} K^{-1/2}\sin K^{1/2}\chi & \text{if } K>0. \\ \chi & \text{if } K=0. \\ (-K)^{-1/2}\sinh(-K)^{1/2}\chi & \text{if } K<0. \end{cases} \qquad (3.36)$$

The curvature is $K = (\Omega_0 - 1)H_0^2$. The relation between redshift and τ is given by the Friedmann equation.

In a curved geometry, a deflection by angle $\delta\vec{\theta}$ at χ' leads to a perpendicular displacement at χ of $\delta\vec{x}_\perp = \delta\vec{\theta}\sin_K(\chi-\chi')$. In our approximation of §3.3, these deflections simply add linearly. Thus, our expressions for $\vec{x}_\perp$ or $\vec{d}_\perp$ remain valid if we replace any expression of the form $\tau_1-\tau_2$ with $\sin_K(\tau_1-\tau_2)$, so e.g. $r_{LS} = \sin_K(\tau_L-\tau_S)$. Thus the lens equation (3.14), the magnification matrix (3.16), and (again by Fermat's principle) the time delay (3.21) all have the same form except that now

$$\mathrm{F}_{ij}(\tau_1,\tau_2) = -\frac{2}{\sin_K(\tau_1-\tau_2)}\int_{\tau_2}^{\tau_1}\sin_K(\tau-\tau_2)\sin_K(\tau_1-\tau)\partial_i\partial_j\phi(\tau)d\tau . \qquad (3.37)$$

3.10 Appendix B

In the limit of weak lensing with no primary lens, our formalism reproduces previously derived results. From equation (3.34), we find an angular correlation function

$$\langle \mathrm{TrA}(0)\mathrm{TrA}(\theta)\rangle = \frac{4}{r_S^2}\int_0^{r_S} dr_1 \int_0^{r_S} dr_2\, e^{r_1\vec{\theta}\cdot\vec{\nabla}_\perp} r_1(r_S-r_1)r_2(r_S-r_2)\left\langle \vec{\nabla}_\perp^2\phi(\tau_1)\vec{\nabla}_\perp^2\phi(\tau_2)\right\rangle ,$$

where we have used $r_1 = \tau_0 - \tau_1$, etc. We convert to Fourier space, and use spherical coordinates $\{k,\ \theta_k,\ \phi_k\}$. We use the approximation that only k values for which $kr_S \gg 1$ (i.e., wavelengths much smaller than the source distance) make an important contribution. This implies that $\int_0^{r_S} dr_1 \int_0^{r_S} dr_2 \approx \int_0^{r_S} dr_1 \int_{-r_S}^{r_S} du$, with $u = r_2 - r_1$, and also that we can set $r_2 = r_1$ in the distance terms in the integrand. Letting $\omega = ku$ and denoting r_1 now by r, in Fourier space our expression becomes

$$4 \int_0^{r_S} dr \int_{-kr_S}^{kr_S} d\omega \int d^3k\ e^{ikr\theta \sin\theta_k \cos\phi_k} e^{i\omega \cos\theta_k} \sin^4\theta_k\ r^2 \left(1 - \frac{r}{r_S}\right)^2 k^3 P_\phi(k, \tau = \tau_0 - r)\ ,$$

where in the $\vec{k}$ integration we chose the x-axis in the direction of $\vec{\theta}$. Under the approximation of $kr_S \gg 1$,

$$\int_{-kr_S}^{kr_S} d\omega e^{i\omega \cos\theta_k} = 2\pi\delta(\cos\theta_k)\ .$$

Our expression thus equals

$$8\pi \int_0^{r_S} dr\ r^2 \left(1 - \frac{r}{r_S}\right)^2 \int_0^\infty k^5 dk P_\phi(k, \tau = \tau_0 - r) \int_0^{2\pi} d\phi_k\ e^{ikr\theta \cos\phi_k}\ ,$$

or, finally,

$$\langle \mathrm{TrA}(0)\mathrm{TrA}(\theta)\rangle = 16\pi^2 \int_0^{r_S} dr\ r^2 \left(1 - \frac{r}{r_S}\right)^2 \int_0^\infty k^5 dk P_\phi(k, \tau = \tau_0 - r) J_0(kr\theta)\ . \tag{3.38}$$

This correlation function of Tr A equals that of twice the shear of A, which has also been derived previously through other methods (e.g. Blandford et al. 1991; Miralda-Escudé 1991; Kaiser 1992).

Bibliography

[1] Babul, A., & Lee, M. H. 1991, MNRAS, 250, 407

[2] Bardeen, J. M. 1980, Phys. Rev. D, 22, 1882

[3] Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15

[4] Blandford, R. D., Saust, A. B., Brainerd, T. G., & Villumsen, J. V. 1991, MNRAS, 251, 600

[5] Buote, D. A., & Canizares, C. R. 1994, ApJ, 427, 86

[6] Cen, R., Gnedin, N. Y., Kofman, L. A., & Ostriker, J. P. 1992, ApJ, 399, L11

[7] Ehlers, J., & Schneider, P. 1986, A&A, 168, 57

[8] Eisenhardt, P. R., Armus, L., Hogg, D. W., Soifer, B. T., & Neugebauer, G., & Werner, M. W. 1996, ApJ, 461, 72

[9] Falco, E. E., Gorenstein, M.V., & Shapiro, I. I. 1985, ApJ, 289, L1

[10] Hamilton, A. J. S., Kumar, P., Lu, E., & Matthews, A. 1991, ApJ, 374, L1

[11] Jain, B., Mo, H. J., & White, S. D. M. 1995, MNRAS, 276, L25

[12] Jaroszynski, M., Park, C., Paczynski, B., & Gott, J. R. 1991, ApJ, 365, 22

[13] Kaiser, N. 1992, ApJ, 388, 272

[14] Kassiola, A., & Kovner, I. 1993, ApJ, 417, 450

[15] Kochanek, C. S. 1991a, ApJ, 373, 354

[16] Kochanek, C. S. 1991b, ApJ, 379, 517

[17] Kochanek, C. S. 1995, ApJ, 453, 545

[18] Kovner, I. 1987, ApJ, 316, 52

[19] Miralda-Escudé, J. 1991, ApJ, 380, 1

[20] Mould, J., Blandford, R., Villumsen, J., Brainerd, T., Smail, I., Small, T., & Kells, W. 1994, MNRAS, 271, 31

[21] Narayan, R. 1991, ApJ, 378, L5

[22] Peacock, J. A., & Dodds, S. J. 1994, MNRAS, 267, 1020

[23] Ratnatunga, K. U., Ostrander, E. J., Griffiths, R. E. & Im, M. 1995, ApJ, 453, L5

[24] Rix, H. W., Schneider, D. P., Bahcall, J. N. 1992, AJ, 104, 959

[25] Ryden, B. S. 1992, ApJ, 396, 445

[26] Sackett, P. D. 1996, in Proc. IAU Symposium 173, "Astophysical applications of gravitational lensing", C. Kochanek and J. Hewitt, eds., Kluwer Academic, Dordrecht

[27] Sasaki, M. 1993, Prog. Theor. Phys., 90, 753

[28] Schechter, P. L., 1987, IAU Symposium 127, "Elliptical Galaxies," 217

[29] Schneider, P., Ehlers, J., & Falco, E.E. 1992, Gravitational Lenses (New York: Springer)

[30] Seljak, U. 1994, ApJ, 436, 509

[31] Seljak, U. 1996, ApJ, 463, 1

[32] Surpi, G. C., Harari, D. D., & Frieman, J. A. 1996, ApJ, 464, 54

[33] Villumsen, J. V. 1995, preprint astro-ph/9507007

[34] Wallington, S., & Narayan, R. 1993, ApJ, 403, 517

[35] Watanabe, K., Sasaki, M., & Tomita, K. 1992, ApJ, 394, 38

Chapter 4

Limits on a Stochastic Background of Gravitational Waves from Gravitational Lensing

This chapter[1] deals with the observable effects of light deflection due to a stochastic background of gravitational waves.

4.1 Introduction

Events in the early Universe may have left a stochastic background of gravitational waves (GW's). In particular, a generic prediction of inflation is a relic spectrum of GW's (Starobinskii 1979; Rubakov et al. 1982; Starobinskii 1985; Abbott & Wise 1984). Detecting these elusive remnants would not only establish this prediction of general relativity, but also serve as a critical test for inflation. While the predicted background may be too weak for direct detection (Bar-Kana 1994; Liddle 1994; Turner et al. 1993; Sahni 1990; Sahni & Souradeep 1992), it could be detected indirectly through its effect on light propagation in the Universe. Even if the effects of GW's cannot be distinguished observationally from other effects, observers who assume no GW's might reach incorrect conclusions about the distribution of matter in the Universe.

GW's may be produced by many sources. Astrophysical sources, such as close binary systems which include a neutron star or black hole, radiate GW's, and numerous individual sources may superpose to create a stochastic background. At the Planck time, quantum fluctuations in the metric are significant and may produce gravitons. Phase transitions in the universe may lead to topological defects such as cosmic strings, which generate GW's. A period of inflation may leave behind a significant amount of GW's. Whatever the source, any spectrum which extends over wavelengths comparable to the present horizon would contribute to the quadrupole

[1]Based on the publication Bar-Kana, R. 1996, Phys. Rev. D, 54, 7138

anisotropy of the cosmic microwave background (CMB) (Krauss & White 1992; Davis et al. 1992; Turner 1993). Such a spectrum is therefore limited by the anisotropy measured by the COBE DMR experiment (Smoot et al. 1992; Gorski et al. 1994). For our calculations we adopt a scale-invariant primordial spectrum, i.e., one which has constant energy density per logarithmic frequency, which we assume produces the entire measured quadrupole anisotropy. Inflationary models predict slightly tilted spectra which are responsible only for some fraction of the anisotropy (e.g. Davis et al. 1992), and so are generally weaker than our adopted case.

In inflation, GW's are produced in conjunction with density fluctuations. The initial nearly scale-invariant power spectrum of density fluctuations evolves as modes reenter the horizon after inflation, and as structure later forms in a universe dominated by dark matter. The present spectrum is strongly constrained by galaxy and cluster surveys, and can be used to study the effects of LSS on lensing. The induced effects are small but potentially observable. In weak lensing, the effect is a coherent distortion of background galaxies by an ellipticity of the order of a few per cent (e.g. Miralda-Escudé 1991; Mould et al. 1994; Kaiser 1992). In strong lensing, the primary effect is an external shear which may be significant for observed four-image systems (chapter 3 of this thesis; Seljak 1994).

In general, the influence at a given time of a weak metric perturbation on light propagation is simply described by two effects. Their magnitudes were estimated for LSS by Seljak (1994), as we summarize here. The first effect is a constant deflection, the same for all nearby light rays. This deflection simply displaces the "true" angular position of an observed lens or source, and is not directly observable. In the case of LSS, deflections from coherent structures of size ~ 1 Mpc add up in a random walk, giving an overall deflection of order a few arcminutes at redshift 1, which scales as the square root of comoving distance r. The second effect is a *relative* deflection between nearby light rays, which produces a focusing and shear with observable effects on weak and strong lensing. For two rays at initial angle θ, each coherent structure at a distance r causes a relative deflection proportional to their separation of $\approx r\theta$. The additional random walk gives a relative angular fluctuation of $\approx 0.07\ \theta$ at redshift 1, which scales as $r^{3/2}$.

It was suggested by Allen (1989, 1990) that gravity waves could significantly affect the time delays in a multiply imaged system. It was later pointed out (Surpi et al. 1996; Frieman et al. 1994) that a correct analysis must include the lensing constraint, i.e., the fact that image rays in the presence of GW's follow different paths than for no GW's, so that all rays go from the source to a common destination, the observer. These later authors also showed that both LSS and GW's have no observable effects on lensing, to lowest order. However, in their lowest order expansion they assumed that two image rays that are observed at an angular separation θ are separated by a distance of exactly $r\theta$ on the lens plane at a distance r. In other words, they neglected the relative deflection between light rays, and therefore only included an overall, constant deflection due to LSS or GW's.

We can easily see why this assumption leads to no observable effects. In the absence of metric perturbations, we can write the lens equation for a thin lens as (e.g., Schneider et al. 1992) $\vec{\beta} = \vec{\theta} - \vec{\alpha}_{\rm lens}(\vec{\theta})$, where $\vec{\theta}$ and $\vec{\beta}$ are the image and source

angles, respectively, and $\vec{\alpha}_{\rm lens}$ is the scaled deflection angle, which is determined by the mass distribution of the lens. If we neglect relative deflections, then LSS or GW's can only cause an angular shift $\vec{\alpha}_L$ between the observer and the lens, and a shift $\vec{\alpha}_{LS}$ between the lens and the source. Then the lens equation becomes $\vec{\beta} = \vec{\theta} - \vec{\alpha}_{\rm lens}(\vec{\theta}) + \vec{\alpha}_S$, where $\vec{\theta}$ is now measured relative to the *observed* (and shifted) lens position, and $\vec{\alpha}_S$ involves $\vec{\alpha}_L$ and $\vec{\alpha}_{LS}$ (see §4.3 for the full details). The constant (i.e., $\vec{\theta}$-independent) deflection $\vec{\alpha}_S$ has no effect on any observables of the lens system (e.g., Schneider et al. 1992), since $\vec{\beta}$ is not directly observable . Fermat's principle then implies that the lens equation must be equivalent to $\partial \Delta t / \partial \vec{\theta} = 0$ at fixed $\vec{\beta}$, where Δt is the relative time delay. There is thus no observable effect on the time delay, either, since it can be derived from the lens equation, up to (unobservable) $\vec{\theta}$-independent terms.

This approximation of neglecting the relative deflection may not be a good one. Indeed, such deflection can have observational consequences, which may be sufficiently large to detect in the case of LSS (chapter 3; Seljak 1994). In this chapter, we compute the rms total and relative deflections between light rays induced by a scale-invariant stochastic background of GW's. Unlike LSS, GW's oscillate with time, and so the effect of short wavelength modes does not amplify, as light rays deflect one way in crests and the opposite way in troughs. In addition, the energy density and thus also the amplitude of sub-horizon GW's redshift away as the universe expands. The lensing effect is thus dominated by wavelengths on the scale of the distance to the source. Each such mode acts as a single coherent structure, and so both the total and relative deflections due to GW's scale approximately linearly with distance. The effect of different modes must be convolved with a particular power spectrum and include the above-mentioned decay of each mode as the universe expands. We find simple integral expressions for the scale-invariant spectrum. The total and relative deflections are smaller than those caused by large-scale structure by factors of the order of 10^2 and 10^4, respectively. We do not need to explicitly set up the lens equation, since the rms shear in the lens equation is directly related to the rms relative deflection of light rays, which we calculate. This fact is demonstrated for LSS in chapter 3, and we give a general proof in §4.3 below. Our results imply that the static effects of the GW spectrum on lensing are negligible compared to those of LSS, and cannot be detected in practice.

In addition to the static effects of LSS and GW's on lensing, it is possible that the fluctuation in the induced deflection with time would be directly manifested as an observed proper motion of images. In other words, the sources do not really move but the light rays from the sources are deflected and so the sources appear to move. We find that even LSS can only produce motions of order 10^{-8} arcsec per year from this effect. This corresponds to ~ 50 km/s at a distance of a Gpc, and the effect of GW's is smaller still by a factor of $\sim 10^2$. Since typical peculiar velocities are much larger, the proper motion induced by deflection of light due to LSS is unobservable, and the same is true for the COBE-normalized scale-invariant spectrum of GW's.

However, we may try to use shear or proper motions of imaged sources to improve existing limits on stochastic GW's at a range of astrophysical wavelengths. There are only a few such limits known: Single-pulsar timing yields $\Omega_\lambda < 1 \times 10^{-8}$ at $\lambda \approx 2$ pc

(Thorsett & Dewey 1996; Kaspi et al. 1994), binary pulsar timing implies $\Omega_\lambda < 0.04$ over $\lambda \approx 2$ pc to 1 kpc and $\Omega_\lambda < 0.5$ up to 10 kpc (Thorsett & Dewey; Bertotti et al. 1983), and the observed angular correlation function of galaxies sets a limit of $\Omega_\lambda < 10^{-3}$ over $\lambda \approx 100$ kpc-100 Mpc (Linder 1988b). These limits apply to any stochastic background of GW's, whether cosmological in origin or generated at low redshift as a superposition of many discrete sources. For a cosmological spectrum that existed at early times, there are also big bang nucleosynthesis constraints of $\Omega_\lambda < 10^{-4}$ for $\lambda < 100$ pc (Carr 1980) and CMB limits of $\Omega_\lambda < 10^{-12}$ at Horizon wavelengths (from COBE) and $\Omega_\lambda < 10^{-8}$ for $\lambda > 1$ Mpc from small-scale anisotropy (Linder 1988a).

Linder (1988b) suggested that highly magnified lensed sources could increase the sensitivity to detecting proper motions due to GW's. The angular deviations induced by GW's produced by an individual source were discussed by Fakir (1994). Pyne et al. (1996) considered detecting proper motions (of unlensed sources) due to GW's through VLBI measurements, but our approach is simpler than theirs. For an image of a lensed source, only an angular deflection of the source *relative* to the lens is easily observed, and we find that this relative motion is small when we assume an isotropic GW background. Thus we do not find an interesting limit on the energy density.

4.2 Formalism

In this section we review the formalism describing gravity waves, their cosmological evolution, and their effect on lensing, as well as the usual formalism of gravitational lensing. We work in the framework of a flat Robertson-Walker metric with small-amplitude tensor metric fluctuations. For weak perturbations, we can consider the effect of GW's without including LSS, since the cross terms between them would be of higher order. In comoving coordinates we can write the line element as

$$ds^2 = a^2(\tau)[-d\tau^2 + (\delta_{ij} + h_{ij})dx^i dx^j] \; . \tag{4.1}$$

Here τ is the conformal time, $a(\tau)$ the expansion factor, and we have set $c = 1$. We expand the metric perturbation in plane waves ($k = 2\pi/\lambda$),

$$h_{lm}(\vec{x},\tau) = \int d^3k \; h^n(\vec{k},\tau)\epsilon^n_{lm}(\widehat{k})\, e^{-i\vec{k}\cdot\vec{x}}, \tag{4.2}$$

where ϵ^n_{lm} is the polarization tensor which depends on the direction $\widehat{k}$ (l and m are spatial indices ranging from 1 to 3, while n ranges over the polarization components $+, \times$). For a wave propagating in the z-direction, the nonvanishing components are in the x-and y- rows and columns:

$$\epsilon^+_{lm} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; \epsilon^\times_{lm} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \; .$$

For other propagation directions $\vec{k}$, we rotate $\epsilon \to R\epsilon R^{\rm T}$, with R the standard 3×3 rotation matrix.

GW's with a given wavevector $\vec{k}$ are produced during inflation and then stretched outside the horizon. The amplitude is constant outside the horizon, but once a mode reenters its energy redshifts as a^{-4}. For the inflationary spectrum the effect of very short wavelength modes is negligible, and so we can assume that all modes enter during the matter-dominated era, for which the exact time evolution is given (e.g. Abbott & Wise 1984) by a spherical Bessel function, $3j_1(k\tau)/(k\tau)$. This time evolution is also correct for all modes long after matter-radiation equality. Inflation produces Gaussian, stochastic perturbations. The Fourier components have zero ensemble mean and a covariance

$$< h^i(\vec{k},\tau_1)h^j(\vec{q},\tau_2) >= A_{\rm T}k^{-3}\left[\frac{3j_1(k\tau_1)}{k\tau_1}\right]\left[\frac{3j_1(k\tau_2)}{k\tau_2}\right]\delta^3(\vec{k}+\vec{q})\delta_{ij}, \qquad (4.3)$$

for the scale-invariant k^{-3} spectrum. Note that we do not assume the short-wavelength approximation $h^i(\vec{k},\tau) \propto a^{-1}(\tau)\, e^{ik\tau}$. The contribution to Ω at the present (averaged over several periods) is

$$\Omega_\lambda = \frac{d\Omega_{\rm GW}}{d\ln k} = \frac{3\pi}{2}A_{\rm T}(k\tau_0)^{-2}\ , \qquad (4.4)$$

where $\tau_0 = 2H_0^{-1}$ is the present value of τ, and throughout we set $H_0 = 75$ km sec^{-1} Mpc^{-1}. Normalization to the full CMB quadrupole anisotropy gives $A_{\rm T} = 6 \times 10^{-11}$.

Consider a photon emitted from a source toward an observer at the origin, with the photon's final direction defined as (minus) the z-axis. We use r to denote values of the z-coordinate (with z_S denoting the source redshift, not its z-coordinate). GW's affect the distance-redshift relation, but this effect is separate from that of the angular deflections which we are interested in, and it introduces only small additional corrections in these quantities (Linder 1988b). We can thus neglect this effect, and assume that the photon path obeys $r(\tau) = \tau_0 - \tau$. In a flat, matter-dominated universe, $r_S = 2H_0^{-1}[1-(1+z_S)^{-1/2}]$. The components perpendicular to the z-axis of the photon direction obey (Linder 1988b)

$$\frac{dx^i}{d\tau}(\tau) - \frac{dx^i}{d\tau}(\tau_0) = h_{zi}(\tau) - h_{zi}(\tau_0) - \frac{1}{2}\int_\tau^{\tau_0}\nabla_i h_{zz}(\tau')\, d\tau'\ . \qquad (4.5)$$

Integrating this we find, for the perpendicular components of the position (with respect to $\vec{x}(\tau_0) = 0$),

$$x^i(\tau) = \int_\tau^{\tau_0}\left[\frac{1}{2}(\tau'-\tau)\nabla_i h_{zz}(\tau') + h_{zi}(\tau_0) - h_{zi}(\tau')\right]d\tau'\ . \qquad (4.6)$$

We define a (two-component) angle $\beta^i = x^i(\tau)/r(\tau)$.

In gravitational lensing with a primary thin lens at a distance r_L (but no LSS or

GW's) the lens equation is (e.g., Schneider et al. 1992)

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}_{\rm lens}(\vec{\theta}) \ , \tag{4.7}$$

where $\vec{\theta}$ is the observed image angle, $\vec{\beta}$ is the source angle (defined as $\vec{x}_S/r_S$, in terms of the perpendicular position of the source), and $\vec{\alpha}_{\rm lens}$ is the deflection angle scaled by r_{LS}/r_S (we define $r_{LS} = r_S - r_L$). In this case, the fiducial z-axis is defined to be in the observed direction of the lens. The distortion of the image of a small source is given by the inverse of the Jacobian matrix

$$\frac{\partial \beta^i}{\partial \theta^j} = \delta^{ij} - \Psi^{ij} \ , \tag{4.8}$$

where Ψ^{ij} is also termed the shear tensor of the lens.

4.3 Shear induced by GW's on lensing

In this section we follow the approach used for LSS by Seljak (1994), i.e., we compute some of the same quantities for GW's and compare the results. As stated in §4.1, we do not need to include a lens explicitly, as we now justify. In the presence of a metric perturbation, but without a primary lens, the lens equation has the form $\vec{\beta} = \vec{\theta} - \vec{\alpha}_{OS}(\vec{\theta})$, where $\vec{\alpha}_{OS}$ results from the accumulated deflection between the observer and the source. As defined in §4.2, the shear tensor for an image at $\vec{\theta}$ due to the perturbation is $F^{ij} = \partial \alpha^i_{OS}(\vec{\theta})/\partial \theta^j$. On the other hand, the relative deflection at $\vec{\theta}$ between two rays separated by a tiny angle $\vec{\gamma}$ is $\vec{\alpha}_{OS}(\vec{\theta} + \vec{\gamma}) - \vec{\alpha}_{OS}(\vec{\theta})$. We denote the rms of this quantity by $\sigma_{\Delta\beta}$. We average over directions of $\vec{\gamma}$ (which for this calculation is equivalent to fixing $\vec{\gamma}$ and assuming that F^{ij} is isotropic) and take $\gamma \to 0$, obtaining the relation [2]

$$\left(\frac{\sigma_{\Delta\beta}}{\gamma}\right)^2 = \frac{1}{2}\left\langle F^{ij}F_{ij}\right\rangle \ , \tag{4.9}$$

all evaluated at position $\vec{\theta}$. Thus, $\sigma_{\Delta\beta}/\gamma$ yields an estimate of the magnitude of the shear tensor. Indeed, it fully characterises rms values of F^{ij}, since for an isotropic field

$$\langle F_{ij}F_{kl}\rangle = \frac{1}{8}\left[\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right]\langle F^{mn}F_{mn}\rangle \ . \tag{4.10}$$

If we also include a primary lens, in the lens equation we simply add up all deflections linearly, assuming all deflections are small. For the primary lens alone, we have Eq. (4.7). Figure 4-1 shows this setup schematically. In the presence of a metric perturbation, we trace a light ray that is observed at $r = 0$ to come from the

[2] Repeated indices are summed over the x and y directions. There is no distinction between upper and lower indices.

direction $\vec{\theta}$, back to the source. We find a different form for the lens equation:

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}_{OS}^{(2,3)} - \vec{\alpha}_{\rm lens}(\vec{\theta} - \vec{\alpha}_{OL}^{(2)}) \ . \tag{4.11}$$

Here $\vec{\alpha}_{OS}^{(2,3)}$ refers to the integrated deflection caused by LSS or GW's along the paths labeled 2 and 3 in Figure 4-1, defined so that the total induced change in $\vec{x}(\tau_S)$ equals $r_S\vec{\alpha}_{OS}^{(2,3)}$. Similarly, $r_L\vec{\alpha}_{OL}^{(2)}$ is the induced change in $\vec{x}(\tau_L)$. In integrating the deflections along the unperturbed paths 2 and 3, we are assuming that the relative deflections due to LSS or GW's are small compared to θ and $\alpha_{\rm lens}$, which is true for the cases which we consider below.

When the perturbation is included, $\vec{\theta}$ is no longer an observable, since it is measured with respect to the unperturbed position of the lens. The observed position of the lens (whose actual position has not changed) is now $\vec{\theta}_{\rm lens} = \vec{\alpha}_{OL}^{(1)}$, so the lens equation in terms of the observable $\vec{\theta}' = \vec{\theta} - \vec{\theta}_{\rm lens}$ is

$$\vec{\beta} = \vec{\theta}' + \vec{\alpha}_{OL}^{(1)} - \vec{\alpha}_{OS}^{(2,3)} - \vec{\alpha}_{\rm lens}(\vec{\theta}' + \vec{\alpha}_{OL}^{(1)} - \vec{\alpha}_{OL}^{(2)}) \ . \tag{4.12}$$

If we now calculate the shear tensor resulting from equation (4.12), it will contain the shear of the primary lens, shear terms from the perturbation, and also cross terms. For simplicity, in the case of GW's we only estimate one characteristic magnitude, that of the shear resulting from $\vec{\alpha}_{OS}^{(2,3)}$, by evaluating the corresponding $\sigma_{\Delta\beta}/\gamma$. In chapter 3, all the different shear terms are studied for LSS, showing that $\sigma_{\Delta\beta}/\gamma$ indeed estimates the relative magnitude of the various corrections due to LSS. Since we find that $\sigma_{\Delta\beta}/\gamma$ is much smaller for GW's, we do not have any motive to explore equation (4.12) further in this case. Instead of the path $(2,3)$, we may use a straight path from $r = 0$ to $r = r_S$ to evaluate the rms of various quantities in this section, since $\alpha_{\rm lens} \ll 1$ and so the components of vectors and tensors as well as the relative distances of points along the path (both of which enter into the rms calculations) are unchanged (except for $O(\alpha_{\rm lens})$ corrections). Thus we only need to consider the effect of GW's in the absence of a primary lens.

Consider first a single light ray with $\vec{\theta} = 0$. In the absence of GW's (or LSS) it would follow the straight line $x^i(\tau) = 0$ for all τ. We now include the effect of GW's, and compute the rms fluctuation in the photon's perpendicular displacement at the source, $\sigma_\beta = \left\langle \vec{\beta}(\tau_S) \cdot \vec{\beta}(\tau_S) \right\rangle^{1/2}$. This is a measure of the common deflection of all image rays, and is therefore not observable, but it is useful for the calculations that follow. We use equation (4.6) and convert the expression to Fourier space. Consider first only the h_{zz} term, whose contribution to σ_β^2 we denote $\sigma_{\beta,a}^2$. The polarization gives $(\epsilon_{zz}^+)^2 + (\epsilon_{zz}^\times)^2 = \sin^4\theta_k$, where $\vec{k} = (k, \theta_k, \phi_k)$ in spherical coordinates. Performing the angular $\vec{k}$ integrations then yields

$$\sigma_{\beta,a}^2 = \frac{432\pi}{r_S^2} A_{\rm T} \int_{\tau_S}^{\tau_0} d\tau_1 \int_{\tau_S}^{\tau_0} d\tau_2 \int_0^\infty dk\ k(\tau_1 - \tau_S)(\tau_2 - \tau_S) \frac{j_1(k\tau_1)}{k\tau_1} \frac{j_1(k\tau_2)}{k\tau_2} \frac{j_3[k(\tau_1 - \tau_2)]}{[k(\tau_1 - \tau_2)]^3} \ .$$

The $j_3[k(\tau_1 - \tau_2)]/[k(\tau_1 - \tau_2)]^3$ term represents a further suppression of short wavelength modes due to phase cancellations among different waves in the assumed isotropic

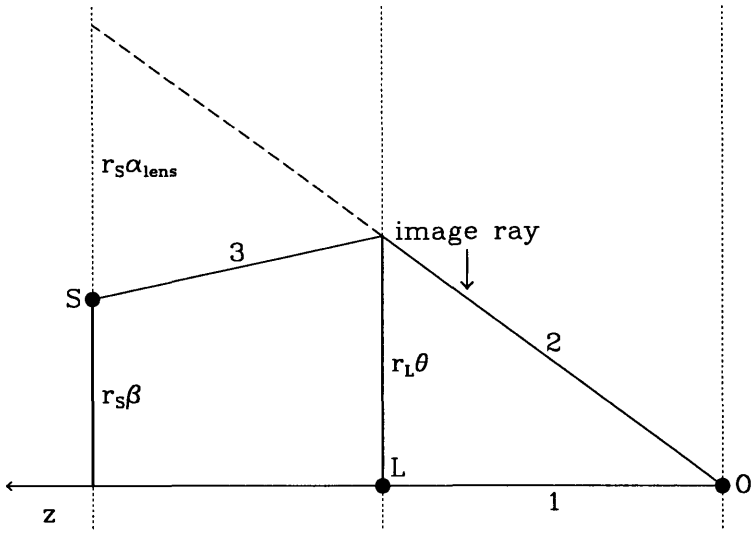

Figure 4-1: Sketch showing positions of the observer, lens and source, as well as an image ray and several comoving distances.

stochastic background. Letting $s = k\tau_1$ and $q = \tau_2/\tau_1$, we can simplify this expression to a double integral,

$$\sigma^2_{\beta,a} = \frac{864\pi}{r_S^2} A_T \int_{qs}^1 \left\{ q\tau_0 \left(\frac{1}{2}\tau_0 - \tau_S\right) - \tau_S r_S + \tau_S^2 \left[\frac{1}{2q} + \ln(q/q_S)\right] \right\} W(q) dq \ , \quad (4.13)$$

where $q_S = \tau_S/\tau_0$ and with $u \equiv (1-q)s$ we define

$$W(q) = \int_0^\infty \frac{j_1(s)}{s} \frac{j_1(qs)}{qs} \frac{j_3(u)}{u^3} \, s \, ds \ . \quad (4.14)$$

Similarly, the contribution of the h_{zi} terms of Eq. (4.6) is

$$\sigma^2_{\beta,b} = \frac{144\pi}{r_S^2} A_T \int_{qs}^1 \left[\frac{\tau_S^2}{2q^2} + \tau_0 \left(\frac{1}{2}\tau_0 - \tau_S\right) \right] G(q) dq \ , \quad (4.15)$$

where

$$G(q) = \int_0^\infty \left\{ \frac{1}{5}\left[\frac{j_1(s)}{s}\right]^2 + \frac{1}{5}\left[\frac{j_1(qs)}{qs}\right]^2 - 2\frac{j_1(s)}{s}\frac{j_1(qs)}{qs}\left[\frac{j_1(u)}{u} - 2\frac{j_2(u)}{u^2}\right] \right\} \frac{ds}{s} \ . \quad (4.16)$$

Integrating over angles gives a zero cross term, and so $\sigma^2_\beta = \sigma^2_{\beta,a} + \sigma^2_{\beta,b}$. Numerically, we find that $\sigma_\beta = 5 \times 10^{-6}$ ($z_S = 1$), 9×10^{-6} ($z_S = 3$). This is much smaller than the estimates for LSS (Seljak 1994), 6×10^{-4} ($z_S = 1$), 7×10^{-4} ($z_S = 3$).

To estimate the relative deflection between rays at $\vec{\theta} = 0$, we choose two directions (labeled A and B) separated at the observer by an infinitesimal angle γ, and find the rms difference between the deflections due to GW's in these two directions, $\sigma_{\Delta\beta} = \left\langle \left[\vec{\beta}_A(\tau_S) - \vec{\beta}_B(\tau_S)\right]^2 \right\rangle^{1/2}$. We cannot evaluate this with the method used for LSS, which assumes that horizon size modes are negligible (Kaiser 1992). Instead we must

calculate $\sigma^2_{\Delta\beta}$ explicitly and keep all the terms to lowest order in γ, i.e., quadratic order. These include terms which come from multiplying polarization components for the different directions A and B. The final result is

$$(\sigma_{\Delta\beta}/\gamma)^2 = 4\sigma^2_{\beta,a} + \sigma^2_{\Delta\beta,a} + \sigma^2_{\Delta\beta,b} + \sigma^2_{\Delta\beta,c} \ , \tag{4.17}$$

where

$$\begin{aligned}
\sigma^2_{\Delta\beta,a} &= \frac{576\pi}{r_S^2} A_{\rm T} \int_{qs}^1 \left[\frac{\tau_0 \tau_S}{q}(1+q)^2 - \left(\tau_0^2 + \frac{\tau_S^2}{q}\right)(1+q) \right. \\
&\qquad \left. + (\tau_0 + \tau_S) r_S \ln(q/q_S) \right] I_1(q) dq \ , \\
\sigma^2_{\Delta\beta,b} &= \frac{1728\pi}{r_S^2} A_{\rm T} \int_{qs}^1 \Big\{ (\tau_0+\tau_S)^2 - \tau_0^2(1+q)^2 + 2q\tau_S(\tau_S - q\tau_0) \\
&\qquad + \left[\tau_S\tau_0(1+q^2) + q(\tau_0+\tau_S)^2 \right] \ln(q/q_S) \Big\} I_2(q) dq \ , \\
\sigma^2_{\Delta\beta,c} &= \frac{288\pi}{r_S^2} A_{\rm T} \int_{qs}^1 \left[\frac{\tau_S^2}{2q^2} + \tau_0 \left(\frac{1}{2}\tau_0 - \tau_S \right) \right] I_3(q) dq \ , \\
I_1(q) &= \int_0^\infty \frac{j_1(s)}{s} \frac{j_1(qs)}{qs} \left[\frac{j_2(u)}{u^2} - 3\frac{j_3(u)}{u^3} \right] s \, ds \ , \\
I_2(q) &= \int_0^\infty \frac{j_1(s)}{s} \frac{j_1(qs)}{qs} \frac{j_4(u)}{u^4} s^3 \, ds \ , \\
I_3(q) &= \int_0^\infty \left\{ \frac{2}{15} \left[\frac{j_1(s)}{s} \right]^2 + \frac{2}{15} \left[\frac{j_1(qs)}{qs} \right]^2 - 2 \frac{j_1(s)}{s} \frac{j_1(qs)}{qs} \left[\frac{j_1(u)}{u} - 3\frac{j_2(u)}{u^2} \right] \right\} \frac{ds}{s} \ .
\end{aligned} \tag{4.18}$$

Then $\sigma_{\Delta\beta}/\gamma = 7 \times 10^{-6}$ ($z_S = 1$), 1.3×10^{-5} ($z_S = 3$). By contrast, LSS gives a $\sigma_{\Delta\beta}/\gamma = 0.07$ ($z_S = 1$), 0.14 ($z_S = 3$). For LSS, the relative deflection is greatly increased by coherent deflections for short wavelength modes, but for GW's the effect of these modes is cut off by the redshifting as well as the temporal oscillations. We also used the relation $kr_S\gamma \ll 1$ in the calculation of $\sigma_{\Delta\beta}$. The reason we find a $\sigma_{\Delta\beta}$ of order $\gamma\sigma_\beta$ is that long wavelength modes overlap over the two light rays, and the relative deflection is small compared to the total deflection. Indeed, a Taylor expansion suggests that in general $\sigma_{\Delta\beta}/\gamma \sim kr_S\sigma_\beta$, and $kr_S \sim 1$ is dominant for this GW spectrum. As shown above, the shear tensor (which is also used in weak lensing) is closely related to $\sigma_{\Delta\beta}/\gamma$, and so the mean square ellipticity at a point induced by GW's is of order 10^{-5}, again negligible compared to the few per cent expected from LSS (e.g., Kaiser 1992).

We can also try to derive general limits on GW's at astrophysical wavelengths from the induced shear. To obtain a limit on Ω_λ, we compute the $\sigma_{\Delta\beta}/\gamma$ produced by an isotropic background of GW's at a single wavenumber k. Note that for modes at a given k, we can use equations (4.3) and (4.4) even for short wavelengths (with $A_{\rm T}$ a normalization factor, separate for each k), for times τ long after matter-radiation equality. Since GW's at Horizon wavelengths are already strongly constrained by the CMB as noted above, we restrict our calculation to the case $kr_S \gg 1$, in which case

the h_{zi} terms in Eq. (4.6) can be neglected. We can estimate from Eq. (4.6) that in order of magnitude σ_β^2 should equal $A_T/(k\tau_0)^4$, and thus that $\sigma_{\Delta\beta}^2/\gamma^2 \sim A_T/(k\tau_0)^2$. However, we find from the exact calculation that there is no term this large, only higher order terms in $1/(k\tau_0)$. We outline in the Appendix (§4.6) a mathematical argument showing this cancellation at small wavelengths. This result requires both the phase cancellations that come in averaging over an isotropic background, and also the oscillation with time of the GW's. With different assumptions, e.g., if we analyzed GW's from a particular source, which are then not isotropic, stronger limits may be possible.

4.4 Proper motions induced by LSS and GW's

We now consider the fluctuation of the angular deflection of image rays with time, and the resulting proper motion. If the deflection of image rays induced by LSS or GW's changes significantly during an observation of a lens system, then the slow shift in alignment between the lens and the source will change the impact parameter at the lens of a given ray from the source. The images will therefore move, and even tiny motions may be detected since the source motion is magnified if it is lensed by a primary lens. We first show that this effect is still expected to be too small to measure for LSS and for the GW power spectrum that we have considered above. However, given the weakness of existing limits on GW's at astrophysical wavelengths (§4.1), we also consider possible limits on a general GW spectrum.

Again we consider a single light ray from the observer out to some distance r_S, in the absence of a primary lens (we consider the effect of a lens below). Given a ray with a fixed direction at the observer, its position $x^i(r_S)$ at r_S moves with time, and it is this motion which we evaluate. In practice, we are interested in a fixed source at r_S, in which case its *apparent* position will drift with the same speed but in the opposite direction. For LSS we have (e.g., Seljak 1994)

$$x^i(r_S) = -2\int_{r=0}^{r_S}(r_S - r)\nabla_i\phi(\tau = \tau_0 - r)dr, \tag{4.19}$$

in terms of the Newtonian potential (or scalar metric perturbation) ϕ. We are now using the parameter r rather than τ, since as time changes all comoving distances remain fixed. The only change is the time of evaluation of ϕ, and so to find $dx^i(r_S)/d\tau_0$ from $x^i(r_S)$ we simply replace $\phi(\tau = \tau_0 - r)$ by $\dot{\phi}(\tau = \tau_0 - r)$, with the partial time derivative in $\dot{\phi}$ taken at a fixed position. The rms value of $dx^i(r_S)/d\tau_0$ depends on the power spectrum of $\dot{\phi}$, a quantity which has been estimated by various authors in connection with the Rees-Sciama effect on the CMB (e.g., Rees & Sciama 1968; Martinez-Gonzalez et al. 1992, 1994; Kofman & Starobinsky 1985; Kamionkowski & Spergel 1994; Seljak 1996). While the integrated deflection is dominated by short (~ 1 Mpc) wavelengths, the LSS potential only evolves on a cosmological time scale. In an Einstein-deSitter universe, ϕ is time independent in the linear regime of small density perturbations, but in this case too $\dot{\phi}$ becomes nonzero when nonlinear structure forms. In general, therefore, the proper motion induced by LSS is of order $\sigma_\beta/\tau_0 \simeq 10^{-8}$

arcsec per year. For the gravity wave spectrum considered above, horizon size modes are dominant, and so here too the induced proper motion is of order σ_β/τ_0, with a σ_β smaller by $\sim 10^2$ than for LSS. Any observed proper motion will thus be dominated by peculiar velocities of hundreds of km/s generated, e.g., by the velocity dispersion of stars in a galaxy or galaxies in a galaxy group or cluster.

We now estimate the lensing limit on stochastic GW's in general, at any wave number k. VLBI observations can directly measure or limit proper motions, and this then implies a limit on GW's. Again we restrict ourselves to wavelengths with $kr_S \gg 1$, and consider first the apparent motion of a source that is not lensed by a primary lens. The apparent motion due to GW's of a fixed object at distance r_S is $-d\vec{\beta}(r_S)/d\tau_0$. Up to corrections of order $1/kr_S$, the mean square of this motion is

$$\left\langle \left[\frac{d}{d\tau_0}\beta(r_S)\right]^2 \right\rangle = \frac{18\pi A_T}{5k^2\tau_0^4}\left[1+\frac{2}{3}\cos(2k\tau_0)\right] \ . \tag{4.20}$$

However, when there are both a lens and a source, a GW background will produce correlated proper motions in both, and the relative motion may be small. Limits from VLBI on proper motions in gravitational lenses were recently considered by Kochanek et al. (1996), and we proceed similarly. We may hope for strong limits because, in the presence of lensing, a proper motion of the source relative to the lens is magnified into a larger proper motion of the images. Furthermore, only a *relative* motion between images needs to be detected, as opposed to a more difficult measurement of motion with respect to an external reference frame, since if the source moves (relative to the lens), the different images do *not* all move together. In general, different values of the magnification matrix at the different image positions will produce relative motions between images of the same order of magnitude as the absolute motions. Moreover, pairs of highly magnified images generally have antiparallel motions (Kochanek et al.)

To analyze how proper motion due to GW's may be magnified, we start from Eq. (4.12), and consider the same equation a time Δt later, when the deflections from GW's have changed. E.g., $\vec{\alpha}_{OL}^{(1)}$ has changed to $\vec{\alpha}_{OL}^{(1)} + \Delta^{(1)}$, and a total change $\vec{\Delta}$ in the observed $\vec{\theta}'$ has been induced. Expanding the lens equation to first order in the small changes and solving for $\vec{\Delta}$, we obtain

$$\Delta_i = \Delta_i^{(2)} - \Delta_i^{(1)} + M_i^j\left[\Delta_j^{(2,3)} - \Delta_j^{(2)}\right] \ , \tag{4.21}$$

where the magnification matrix M_i^j equals the inverse of $\delta_i^j - \partial_i\alpha_{\rm lens}^j$ (and is evaluated at $\vec{\theta}' + \vec{\alpha}_{OL}^{(1)} - \vec{\alpha}_{OL}^{(2)}$). Consider first the magnified term, $\Delta_j^{(2,3)} - \Delta_j^{(2)}$. Averaging over directions of $\vec{\Delta}^{(2,3)} - \vec{\Delta}^{(2)}$ we obtain a result analogous to Eq. (4.9) for the mean square. Since M^{ij} is symmetric for a thin lens (Schneider et al.), it has two real eigenvalues m_a and m_b (where the magnification $M = |m_a m_b|$). Letting

$$\widetilde{M} = \left[\frac{1}{2}(m_a^2 + m_b^2)\right]^{1/2} \ , \tag{4.22}$$

we find that

$$\text{rms } \left| M_i^j \left[\Delta_j^{(2,3)} - \Delta_j^{(2)} \right] \right| = \widetilde{M} \times \text{rms } \left| \vec{\Delta}^{(2,3)} - \vec{\Delta}^{(2)} \right| . \tag{4.23}$$

In Eq. (4.23) we may evaluate the rms on the right-hand side using a straight path (as in §4.3). Letting

$$\frac{d}{d\tau_0}\vec{\beta}_{LS} = -\frac{d}{d\tau_0}\vec{\beta}(r_S) + \frac{d}{d\tau_0}\vec{\beta}(r_L) , \tag{4.24}$$

we find that in $\left\langle \dot{\beta}_{LS}^2 \right\rangle$ there is no term of order $A_{\rm T}/(k^2\tau_0^4)$ [as in Eq. (4.20)], but only higher order terms in $1/(k\tau_0)$. Once again this small wavelength cutoff results from combining the time oscillation of GW's and the phase cancellations in averaging over an isotropic background (see the Appendix, §4.6), and as a result there is only a very weak limit on Ω_λ.

4.5 Conclusions

Gravitational lensing is affected by perturbations to the homogeneous and isotropic background metric. Such perturbations, whether they are caused by LSS or GW's, may produce a number of effects on light propagation. One such effect is an overall shift in the angular positions of nearby objects, which is not observable. Another is a relative difference between the induced shifts in nearby light rays. This relative deflection manifests itself as a shear which may cause weak lensing and also affect strong lensing. A third effect is a fluctuation of the angular position of distant objects with time, leading to a directly observable proper motion.

The actual amplitude of long wavelength modes of LSS and GW's is limited by the quadrupole anisotropy of the CMB. Even if both make comparable contributions to the anisotropy, LSS is dominant in its effects on lensing. This results from cancellations due to the time oscillation of short wavelength gravity waves, as well as the redshifting of their amplitude. For LSS, on the other hand, the effect of small coherent structures is amplified as the deflection executes a random walk. We find that the relative deflection due to GW's is four orders of magnitude smaller than that of LSS, and is therefore not observable.

The induced proper motions expected for LSS or for GW's generated in inflation are small compared to typical peculiar velocities, and thus are not observable. The motions are also too small to yield interesting limits on the energy density of GW's at shorter wavelengths.

After this paper was submitted for publication, the bending of light by gravity waves was analyzed differently by Kaiser & Jaffe (1996), for the case of short (sub-horizon) wavelengths, in a nonexpanding flat space (i.e., neglecting the redshifting of the amplitude of GW's). That simplified analysis shows that the relative proper motion between two sources is small not only if they are at different redshifts along the same line of sight (in agreement with our calculation of $\left\langle \dot{\beta}_{LS}^2 \right\rangle$ in §4.4), but also if they are separated on the sky by a small angle. The treatment presented in Kaiser & Jaffe changes quantitatively if expansion is included, but not qualitatively for GW's

with a period short compared to the redshifting time scale (i.e., a Hubble time).

4.6 Appendix

In trying to set limits on GW's at short wavelengths $kr_S \gg 1$, we twice encountered a weaker limit than simple dimensional analysis would suggest: Once in calculating $\sigma_{\Delta\beta}/\gamma$ or shear in §4.3, and then in estimating the magnified proper motion in §4.4. In this Appendix we outline the first of these calculations and show how this result emerges. The second calculation can be done similarly.

From Eq. (4.6) in the limit of short wavelengths (compared with the present Horizon), we derive

$$\begin{aligned}\sigma^2_{\Delta\beta}/\gamma^2 &= \frac{1728\pi}{r_S^2} A_{\rm T} k^4 \int_{\tau_S}^{\tau_0} d\tau_1 \int_{\tau_S}^{\tau_0} d\tau_2 (\tau_1-\tau_S)(\tau_2-\tau_S) \\ &\quad \times (\tau_0-\tau_1)(\tau_0-\tau_2)\frac{\cos(k\tau_1)}{(k\tau_1)^2}\frac{\cos(k\tau_2)}{(k\tau_2)^2}\frac{j_4[k(\tau_1-\tau_2)]}{[k(\tau_1-\tau_2)]^4} \ .\end{aligned}$$

The $j_4[k(\tau_1-\tau_2)]/[k(\tau_1-\tau_2)]^4$ term comes from the angular $\vec{k}$ integrations, including the angular dependence of the polarizations and assuming an isotropic background. Letting $x = k\tau_1$ and $u = k(\tau_1-\tau_2)$ (also $x_0 = k\tau_0$, etc.) leads to

$$\frac{864\pi}{k^2 r_S^2} A_{\rm T} \int_{x_S}^{x_0} dx \frac{(x-x_S)(x_0-x)}{x^2} \int_{x_S-x}^{x_0-x} du \frac{(x+u-x_S)(x_0-x-u)}{(x+u)^2}\frac{j_4(u)}{u^4} \times [\cos u(1+\cos 2x) - \sin u \sin 2x] \ .$$

We evaluate only the first $\cos u$ term here, since the other terms in the square brackets can be evaluated similarly. Note that dimensional analysis suggests that the x and u integrals should give a term of order 1 (not larger, because of the oscillating integrand).

To do the u integral we separate the smooth and oscillating parts and then repeatedly integrate by parts: We let $w^{[n]}(u)$ be the n-th indefinite integral of $[\cos u]\, j_4(u)/u^4$ with respect to u, and $v^{[n]}(u)$ the n-th derivative of $(x+u-x_S)(x_0-x-u)/(x+u)^2$ with respect to u. For each n such that $w^{[n]}(u)$ converges for $u \to \pm\infty$, we fix the arbitrary constant by $w^{[n]}(\infty) + w^{[n]}(-\infty) = 0$ (Any constant will do for other n). Then the u integral equals a series of terms evaluated at the two limits of integration,

$$\sum_{n=0}^{\infty}(-1)^n \left\{ w^{[n+1]} v^{[n]}|_{u=x_0-x} - w^{[n+1]} v^{[n]}|_{u=x_S-x} \right\} \ .$$

Since the two series of terms can be handled similarly, we evaluate here only the $u = x_0 - x$ terms. We do the x integration in the same way as the u integration. Thus we continue to integrate $w^{[n+1]}(x)$ with respect to x, and let $v^{[n,m]}(x)$ be the m-th derivative of $(x-x_S)(x_0-x)v^{[n]}(x_0-x)/x^2$ with respect to x. The contribution

to $\sigma^2_{\Delta\beta}/\gamma^2$ from the terms we have kept is

$$\frac{864\pi}{x_S^2} A_{\mathrm{T}} \sum_{n,m=0}^{\infty} (-1)^{n+1} \left\{ w^{[n+m+2]}(0) v^{[n,m]}(x_0) - w^{[n+m+2]}(x_0 - x_S) v^{[n,m]}(x_S) \right\} .$$

Now, $v^{[n,m]}(x)$ at $x \gg 1$ is of order $x^{-(n+m)}$, $w^{[n]}(0)$ is 0 or a constant, and we find that $w^{[n]}(x)$ at $|x| \gg 1$ is of order $|x|^{n-5}$. This last fact, that $w^{[n]}(\pm\infty)$ converges for the first few n, depends on the specific function $w^{[0]}(x)$ which in turn is determined by the two physical assumptions of time oscillation and angular averaging. The only term from the final sum that could give a contribution of order A_{T}/x_S^2 is the $n = m = 0$ term. We find that $w^{[2]}(0)$ is a nonzero constant, but since $v^{[0,0]}(x) = 0$ identically, there is no term of this lowest order.

Bibliography

[1] Abbott, L. F. and Wise, M. B. 1984, Nucl. Phys. B, 244, 541

[2] Allen, B. 1989, Phys. Rev. Lett., 63, 2017

[3] Allen, B. 1990, Gen. Relativ. Grav., 22, 1447

[4] Bar-Kana, R. 1994, Phys. Rev. D, 50, 1157

[5] Bertotti, B., Carr, B. J., & Rees, M. J. 1983, MNRAS, 203, 945

[6] Carr, B. J. 1980, A&A, 89, 6

[7] Davis, R. L., Hodges, H. M., Smoot, G. F., Steinhardt, P. J., & Turner, M. S. 1992, Phys. Rev. Lett., 69, 1856

[8] Fakir, R. 1994, ApJ 426, 74

[9] Frieman, J. A., Harari, D. D., & Surpi, G. C. 1994, Phys. Rev. D, 50, 4895

[10] K. M. Gorski et al. 1994, ApJ, 430, L89

[11] Kaiser, N. 1992, ApJ, 388, 272

[12] Kaiser, N., & Jaffe, A. 1996, submitted to ApJ, astro-ph/9609043 (1997).

[13] Kamionkowski, M., & Spergel, D. N. 1994, ApJ, 432, 7

[14] Kaspi, V. M., Taylor, J. H., & Ryba, M. F. 1994, ApJ, 428, 713

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

[16] Kofman, L., & Starobinsky, A. 1985, Sov. Astron. Lett., 11, 271

[17] Krauss, L. M., & White, M. 1992, Phys. Rev. Lett., 69, 869

[18] Liddle, A. R. 1994, Phys. Rev. D, 49, 3805

[19] Linder, E. V. 1988a, ApJ, 326, 517

[20] Linder, E. V. 1988b, ApJ, 328, 77

[21] Martinez-Gonzalez, E., Sanz, J. L., & Silk, J. 1992, Phys. Rev. D., 46, 4196

[22] Martinez-Gonzalez, E., Sanz, J. L., & Silk, J. 1994, ApJ, 436, 1

[23] Miralda-Escudé 1991, J., ApJ, 380, 1

[24] Mould, J., Blandford, R., Villumsen, J., Brainerd, T., Smail, I., Small, T., & Kells, W. 1994, MNRAS, 271, 31

[25] Pyne, T., Gwinn, C. R., Birkinshaw, M., Eubanks, T. M., & Matsakis, D. N. 1996, ApJ, 465, 566

[26] Rees, M. J., & Sciama, D. W. 1968, Nature, 517, 611

[27] Rubakov, V. A., Sazhin, M. V., & Veryaskin, A. V. 1982, Phys. Lett., 115B, 189

[28] Sahni, V. 1990, Phys. Rev. D, 42, 453

[29] Sahni, V., & Souradeep, T. 1992, Mod. Phys. Lett. A, 7, 3541

[30] Schneider, P., Ehlers, J., & Falco, E.E. 1992, Gravitational Lenses (New York: Springer)

[31] Seljak, U. 1994, ApJ, 436, 509

[32] Seljak, U. 1996, ApJ, 460, 549

[33] Smoot, G. F. et al. 1992, ApJ, 396, L1

[34] Starobinskii, A. A. 1979, Pis'ma Zh. Eksp. Teor. Fiz., 30, 719 [JETP Lett., 30, 682]

[35] Starobinskii, A. A. 1985, Pis'ma Astron. Zh., 11, 323 [Sov. Astron. Lett., 11, 133]

[36] Surpi, G. C., Harari, D. D., & Frieman, J. A. 1996, ApJ, 464, 54

[37] Thorsett, S. E., & Dewey, R. J. 1996, Phys. Rev. D, 53, 3468

[38] Turner, M. S. 1993, Phys. Rev. D, 48, 5539

[39] Turner, M. S., White, M., & Lidsey, J. E. 1993, Phys. Rev. D, 48, 4613

Chapter 5

Analysis of Time Delays in the Gravitational Lens PG1115+080

This chapter[1] presents an analysis of time delays in a gravitational lens, which can be used to estimate the Hubble constant as well as to constrain models of the lens mass distribution.

5.1 Introduction

The quadruply imaged quasar PG1115+080 was the second lens to be discovered (Weymann et al. 1980). It is radio quiet, but optical Hubble Space Telescope images (Kristian et al. 1993) were recently analyzed by Schechter et al. (1997, hereafter SCH) and by Keeton & Kochanek (1997). They found that lens models which include the effect of the lens galaxy and that of the nearby group of galaxies discovered by Young et al. (1981) can fit the image positions well. They however still found great freedom in the Hubble constant (H_0) values predicted by these lens models for a given time delay.

In four-image configurations where the images lie at roughly the same distance from the lens, there is a well known degeneracy between the radial profile of the lens mass and the inferred H_0 (Kochanek 1991; Wambsganss & Paczyński 1994). In this situation the relative image positions do determine the total enclosed mass within the ring of images, but they are not very sensitive to the radial profile of the mass. Changing the radial profile affects the convergence at the images, and this changes the inferred H_0 value in a similar way to the effect of a constant mass sheet (Falco et al. 1985; Narayan 1991). In PG1115+080 future observations of the lensing galaxy light profile and, ultimately, a direct measurement of its central velocity dispersion may constrain or eliminate this degeneracy.

Recently SCH measured light curves for the A1, A2, B and C images of PG1115+080, and used them to determine multiple delays. The bright A1 and A2 images are close

[1]Submitted to ApJ, also preprint astro-ph/9701068

together and should have a very small time delay ($\sim$ a few hours), so they were combined into a single A=A1+A2 curve. SCH used the Press et al. (1992a, 1992b) method and found that C leads A and A leads B, with $t_{BC} = 23.7 \pm 3.4$ days and $t_{AC} = 9.4 \pm 3.4$ days which yields a ratio $r_{ABC} \equiv t_{AC}/t_{BA}$ of 0.7 ± 0.3. It is useful to express the two independent quantities as t_{BC} and r_{ABC}, since t_{BC} can be taken to fix the Hubble constant for a given lens model while r_{ABC} is independent of overall distance, and can be compared directly with the ratio predicted by lens models. The models mentioned above are consistent in predicting $r_{ABC} = 1.4$ to within about 0.1, and SCH noted the 2σ discrepancy with their fitted value. In their analysis SCH assumed that the measurement errors in the light curves are uncorrelated, and also that the fractional flux variations are the same for each component.

In this chapter we present a more detailed analysis of the light curves in PG1115+080. We first present a new method based on χ^2 minimization, which has many of the advantages of Press et al. but is simpler and allows for a more conservative assessment of errors in the reconstructed parameters. We then apply this method to PG1115+080, and include correlated measurement errors in the analysis. In addition to relative time delays, we also allow for different factors of variation in flux, which may arise from microlensing of the images. In §5.2 we present our χ^2 method, and discuss its distinct features and free parameters. We then discuss the physical meaning of the various parameters that the method can account for and attempt to extract from data. In §5.3 we apply our method to fitting the PG1115+080 light curves, singly, in pairs, and all together, and discuss the results and implications. Finally in §5.4 we summarize our results and point out some of the significant returns possible from further monitoring.

5.2 Method and physical parameters

Suppose we have a light curve of N points, (t_i, v_i) $i = 1, \ldots, N$, where t_i is time (in days) and v_i is the intensity (in magnitudes) at time t_i, and we construct a model for $v(t)$ with M adjustable parameters a_j, $j = 1, \ldots, M$. Then if the v_i measurements are independent, with normally distributed errors of standard deviation σ_i for each i, the χ^2 measure of goodness-of-fit is

$$\chi^2 = \sum_{i=1}^{N} \left[\frac{v_i - v(t_i; a_1 \ldots a_M)}{\sigma_i} \right]^2 . \tag{5.1}$$

When the measurement errors are correlated, with a covariance matrix $\mathbf{Q}$, then

$$\chi^2 = \sum_{i,j=1}^{N} W_{ij} \left[v_i - v(t_i; a_1 \ldots a_M) \right] \left[v_j - v(t_j; a_1 \ldots a_M) \right] , \tag{5.2}$$

where the weight matrix $\mathbf{W}$ is the matrix inverse of $\mathbf{Q}$.

For the model we take L points (t_k, v_k) with the t_k just covering the range of the data and equally spaced. The v_k are then the parameters, and the model for

$v(t)$ is a curve going through all the (t_k, v_k) with interpolation used for other values of t. Note that we interpolate between the v_k in the model, we do not interpolate between the noisy data points directly. When we fit multiple light curves, a single such model is used for all of them, or equivalently the light curves are all combined into a single light curve which is then compared to the model. Additional parameters can be added to account for relative shifts of an entire light curve relative to the others, before comparing to the model. A relative magnification corresponds (in magnitudes) to adding a constant to all the measured v_i of an image. A relative time delay corresponds to adding a constant to all the measured t_i of an image. For given time delays, all the other parameters are linear (i.e. the χ^2 is a quadratic form in these parameters) if we use only interpolation which is linear in the v_k (but arbitrarily non-linear in the t_k). Minimizing the χ^2 then corresponds to solving a linear system, and we solve it using Cholesky decomposition (e.g. Press et al. 1992c, §2.9).

Our model is very flexible. There is freedom to choose both the number of interpolation points L and the interpolation method. We use several methods of interpolation. Given some time t between (t_k, v_k) and (t_{k+1}, v_{k+1}), in linear interpolation we take for $v(t)$ the linearly interpolated value between v_k and v_{k+1}. In cubic interpolation we interpolate using the cubic polynomial through the four points at $k-1$, k, $k+1$ and $k+2$ (with linear interpolation used for $k = 1$ or $k = L-1$). In cubic spline interpolation (e.g. Press et al. 1992c, §3.3), we draw a cubic polynomial between every pair of points k and $k+1$ so that the entire curve is continuous in the second derivative. We find in §5.3 below that changing the interpolation method has a rather minor effect. Note that with any of these interpolation methods there is some inter-dependence among all of the v_k, unless there is some k_0 with no data points between t_{k_0} and t_{k_0+1}. In this case the parameters v_k for $k > k_0$ would decouple from those with $k < k_0$, at least with linear interpolation. There are 37 points in each light curve in the PG1115+080 data, and the final data point follows a gap of 38 days. Hereafter we drop this last point, and then we are never in a situation where such a k_0 exists.

The number of points in each light curve is thus 36, and we use an L of 10, 20 and 30. Below we find that the differences in the derived parameters using the various L are fairly large, so this freedom provides a more conservative and robust estimate of parameter uncertainties. Using small L corresponds essentially to a larger-scale smoothing of the data, though once again the smoothing is not directly of the raw data. If we use a very small L, we are over-smoothing the signal, and we may expect an unreliable reconstruction if the actual signal is not smooth enough on timescales $\leq \Delta T/L$, where ΔT is the total length of the data. If on the other hand we use a large L, there are fewer constraints on the v_k, and we may get large oscillations between nearby v_k as the fitting procedure tries to compensate for individual measurement errors. A range of L thus allows an important check on robustness.

When we use a method based on χ^2 minimization for finding time delays among a number of curves, we may worry about being biased towards large time delays. This is because large time delays imply little overlap among the light curves, and so smaller χ^2 since there are fewer constraints coming from requiring smoothness when the light curves are combined into a single curve. In our method, with a fixed L, this

bias is partially offset since larger time delays imply larger effective smoothing as the total ΔT for the combined light curves increases. In practice we find negligible bias in the Monte Carlo tests described in §5.3 below.

We wish to also account for the effect of microlensing in the fitting. Microlensing is the additional magnification of each image due to lensing by stars or dark matter clumps in the lensing galaxy. It has been detected in 2237+0305 (e.g. Wambsganss et al. 1990) as well as in 0957+561 (e.g. Schild 1996). Such microlensing may cause an amplification of order 0.6 magnitudes, which in PG1115+080 is expected to vary over a time scale of $\sim$ 10 years (Witt et al. 1995). However, even if there is no significant variation over the span of the measured light curves, there may be a 'static' microlensing effect which causes different images to vary by different factors in flux, as suggested by SCH. Specifically, this arises if the quasar emission region which produces the mean (i.e. time independent part of the) flux is not identical to the region responsible for the small intensity variations with time. A possible physical picture is that the variation is caused by a small jet emanating from the core or a sudden localized surge of mass accretion rather than by a coherent change in the entire emission region.

In such a situation, the overall magnifications M_i of the mean flux F and m_i of the variation f may be unequal for a given image i, and their ratio m_i/M_i may be different for different images. In magnitudes, we would then measure (except for an overall factor of $-5/2$)

$$\log_{10}(M_iF + m_if) = \log_{10}(M_iF) + \log_{10}\left(1 + \frac{m_if}{M_iF}\right) \simeq \log_{10}(M_iF) + \frac{m_i}{M_i}\frac{f}{F\ \log 10}\ , \tag{5.3}$$

where we assume $f/F \ll 1$. The variation in magnitudes may therefore have different amplitudes in different images if the m_i/M_i ratio is not the same for all i.

The spatial scale for microlensing is set by the Einstein radius of a typical deflector at the lens plane, projected onto the source plane. In PG1115+080 this radius is

$$\zeta_E = 5 \times 10^{-3}\ [\langle M\rangle /0.1M_\odot]^{1/2}\ \mathrm{pc} = 1000\ [\langle M\rangle /0.1M_\odot]^{1/2}\ \mathrm{AU}\ , \tag{5.4}$$

where $\langle M\rangle$ is the average stellar mass (Witt et al. 1995). In order to have m_i and M_i not be equal, one or both of the emission regions must have structure on scales of order ζ_E or smaller: if they were both smooth then the microlensing would be washed out, and we would have $m_i = M_i$ for all i. On the other hand, if the two emission regions overlapped and their union had a very small extent compared to ζ_E, then the magnification due to microlensing would not vary over this scale, and again we would have $m_i = M_i$. Thus we reach the interesting conclusion that the indication of different m_i/M_i for different i (see §5.3 below) implies that the size of the quasar optical emission region *must be of order* ζ_E, say to within a factor of 10.

There are, however, a number of reasons to be cautious about this statement. The two emission regions may have different sizes, and in this case we might expect the mean region to contain the variation region, and then the different m_i/M_i would imply the lower bound only on the larger emission region. On the other hand, the variation

may involve relativistic motions, and such high velocities can cause microlensing variations on small time scales that we might confuse as differences in m_i/M_i. More generally, the stochastic nature of optically-thick microlensing (e.g. Witt et al. 1995) implies the possibility of a rare statistical fluctuation over the relatively short time span of the current data. In addition, ζ_E depends on what we assume for $\langle M \rangle$, and there could be a large range of masses contributing. Note also that the magnification due to the macrolens itself varies with position, but it should only vary over much larger spatial scales. There could be a different amount of contamination by light from the lens galaxy in each image, and this would lead to apparent changes in the M_i with no corresponding changes in the m_i. However, in PG1115+080 this effect seems to be too small to affect the m_i/M_i ratios significantly. Finally, there may still be unmodelled systematic errors in the measured light curves which could give the appearance of unequal m_i/M_i ratios. With a longer time span of measurements it may be possible to eliminate some or all of these uncertainties. See also Gould & Miralda-Escudé (1997) for an independent, illuminating discussion of the possible observational consequences of differential microlensing of accretion disks in quasars.

In measuring the quasar image light curves, SCH used two nearby stars, *B and *C, as photometric references. All intensities (i.e. for the images as well as for *B) were measured with respect to *C, and SCH then subtracted one half of the magnitude of *B from the quasar components, for each observation. The correlation between the quasar and *B may result from systematic errors such as the adopted point spread function, which varies across the field and so may not be identical in *B and in the quasar components. In our analysis we treat the fluctuations of *B as stochastic noise, with a component which is common to *B and to all the quasar components, as well as an independent source of error. We then let the χ^2 minimization decide the size of the error common to *B and all the quasar images, separately in each observation, but guided by our best estimate of the typical size of the correlation. Thus, we add *B to our method effectively as a fourth light curve, but for the underlying model of *B we take just a constant intensity in time.

5.3 Results for PG1115+080

Our model for component A consists of L points as described in §5.2. For B we use the same model, but with an overall magnitude shift m_{BA} of A with respect to B (accounting for mean magnification), a time shift t_{BA} due to the time delay, and an overall variation ratio α_{BA} multiplying the variation (accounting for differential microlensing as explained in §5.2). Similarly, we have m_{AC}, t_{AC}, and α_{AC}, as well as m_{BC}, t_{BC}, and α_{BC}. When we fit all the curves simultaneously, we have the constraints $m_{BC} = m_{BA} + m_{AC}$, $t_{BC} = t_{BA} + t_{AC}$, and $\alpha_{BC} = \alpha_{BA}\alpha_{AC}$. The only parameters which require non-linear χ^2 minimization are t_{BA}, t_{AC}, α_{BA}, and α_{AC}. We minimize with respect to these parameters using direction set methods (Press et al. 1992c, §10.5).

We vary a number of parameters in the model. We try $L = 10$, 20 and 30 interpolation points. We try linear, cubic and cubic spline interpolation. For the

measurement errors, we begin with the uncorrelated errors for A, B and C of 2.3, 7.5 and 4.0 millimagnitudes (mmag) at 1σ, respectively (SCH). We also take an uncorrelated error of 1.2 mmag for the bright star *B. We then add an error of standard deviation σ_n common to *B and all QSO components, and also separate errors of $f_1\sigma_n$ in *B and $f_2\sigma_n$ in all the QSO components. If we assume $f_1 = f_2 = 1$, then since *B has a measured variability of 8 mmag at 1σ, we take $\sigma_n \simeq 6$ mmag. As we show below, reasonable variations in σ_n, f_1 and f_2 affect the minimum χ^2 value but not appreciably values of the derived parameters such as the time delays. On the other hand, having no correlated errors at all ($\sigma_n = 0$) does change the derived parameters. All the errors are naturally assumed to be Gaussian in the χ^2 method.

We can also take Gaussian errors for Monte Carlo trials. On the other hand, since the method reconstructs a model for the input signal, it also reconstructs an estimate of the measurement errors in all the light curves in each observation. We take these actual errors instead of Gaussian errors for a second, bootstrap-like Monte Carlo analysis. Thus we have a set of 36 observations, and in the Monte Carlo trials for each day we pick one of the 36 at random (with replacement) and add the reconstructed errors on that day (in A, B, C, and *B) to the matching simulated light curves (i.e. A, B, C, and *B, respectively). Note that this is not a rigorous bootstrap, since after the χ^2 minimization the errors are no longer independent. However, this procedure should give us an idea of the actual size and mutual correlations of the measurement errors, and should be more robust and conservative than the Gaussian Monte Carlo trials. For the input signal in the Monte Carlo trials we take a smooth curve (see Figure 1) with roughly the same shape as the variation in the actual data. This way, the input is identical to the real signal on large time scales, and we assume that any apparent small scale variability is dominated by measurement noise, which seems likely.

We define our 'standard' model as having $L = 20$, cubic spline interpolation, $\sigma_n = 6$ mmag and $f_1 = f_2 = 1$. Thus the assumed covariance matrix of errors for this model in mmag squared is

$$\begin{pmatrix} 77.3 & 72 & 72 & 36 \\ 72 & 128.3 & 72 & 36 \\ 72 & 72 & 88.3 & 36 \\ 36 & 36 & 36 & 73.3 \end{pmatrix} \tag{5.5}$$

with rows from the top in the order A, B, C, and *B. Hereafter we use these standard settings except where otherwise noted.

We begin by fitting each quasar component separately, always along with *B. Table 5.1 shows the resulting χ^2 values for various L. The number of data points is 72 (36 each for a quasar component and for *B), and the number of parameters is $L + 1$ (1 for the mean value of *B), which yields $71 - L$ for the number of degrees of freedom (ndof). The χ^2 is in most cases less than the ndof, which is reasonable since we are assuming that a significant portion of the error is correlated among all the quasar components, and if we take each of them separately then we are not including this very strong constraint. Table 5.1 also suggests that if our error estimates are

roughly correct for A and C, they may be somewhat underestimated for component B. Note that the χ^2 value is higher than the ndof at $L = 10$ only for B, which is the faintest component and so has the largest errors. This may indicate higher non-Gaussianity for the errors in this component, since with lower L the fewer parameters cannot effectively compensate for outlier points.

Next we fit *B and two quasar components at a time, with the results displayed in Table 5.2. There are $36*3 = 108$ data points, and $L+4$ parameters (e.g. *B mean, m_{BC}, t_{BC} and α_{BC}). Also shown are the one sided 68% confidence limits derived from 300 bootstrap Monte Carlo trials. Both the bootstrap errors and the parameters (t_{BC} etc.) that we input into the Monte Carlo trials are in each case (i.e. for each pair of components and for each L) set according to the values fitted from the data in that case. So e.g., for the BC pair with $L = 20$, a value of $t_{BC} = 25.3$ days is used as input for the Monte Carlo trials which in turn determine the error of $^{+2.3}_{-2.0}$ days on t_{BC}. We only show the results of varying L since this leads to the largest variations (see below). Note that the BA time delay has the largest fractional uncertainty, as we expect: B has the largest errors, so the uncertainty in t_{BA} is greater than in t_{AC}, and while C has larger errors than A, the BC delay is larger and much better resolved than is t_{BA}. For the variation ratios, there is some variation with L, but only α_{BC} is consistent with 1, while the others disagree with 1 at the $4-5\sigma$ level for each L.

Finally we combine all three quasar components, with results shown in Table 5.3. Uncertainties, where shown, are again one sided 68% confidence limits derived from bootstrap Monte Carlo trials. All the models are based on our standard inputs, except for the changes shown in the first column. In rows 4 and 5, σ_n is in mmag. In rows 6 and 7, f_1 and f_2 refer to assumptions about the correlated errors (see the beginning of this section). In rows 8 and 9, we use cubic and linear interpolation, respectively (see §5.2 above). As above, $r_{ABC} \equiv t_{AC}/t_{BA}$. We try a large range of model assumptions, most of which show very little variation in parameter values relative to varying L. The exception is the last row, marked 'SCH-like'. Here we use our method, but with assumptions that correspond closely to those of SCH: we assume no correlated errors, the uncorrelated errors are doubled, *B is not included in the fitting but rather we subtract half of *B from the quasar light curves, and we set the variation ratios to 1. The 'SCH-like' parameter values are compatible with those of SCH to about 1σ.

For the standard case ($L = 20$), there are $4 * 36 = 144$ data points, and $L + 7$ parameters. In each case, we can calculate $\alpha_{BC} = \alpha_{BA}\alpha_{AC}$. E.g. for $L = 20$ the result is $\alpha_{BC} = 1.10^{+.08}_{-.08}$. With $L = 20$, the 1σ uncertainties from Gaussian Monte Carlo trials are, for comparison with Table 5.3, $^{+0.9}_{-0.9}$ days for t_{AC}, $^{+1.1}_{-1.1}$ days for t_{BA}, $^{+1.4}_{-1.3}$ days for t_{BC}, $^{+.13}_{-.14}$ for r_{ABC}, $^{+.07}_{-.06}$ for α_{AC}, $^{+.04}_{-.04}$ for α_{BA}, and $^{+.07}_{-.07}$ for α_{BC}. The 95% confidence limits from bootstrap Monte Carlo trials are, also for the $L = 20$ case, $^{+2.0}_{-2.2}$ days for t_{AC}, $^{+2.9}_{-3.2}$ days for t_{BA}, $^{+3.3}_{-3.8}$ days for t_{BC}, $^{+.31}_{-.41}$ for r_{ABC}, $^{+.16}_{-.20}$ for α_{AC}, $^{+.10}_{-.12}$ for α_{BA}, and $^{+.15}_{-.19}$ for α_{BC}. For this standard case, Figure 1 shows the quasar components A ($\star$), B ($\bullet$), and C (o) (including the final point in each light curve, which was excluded from the fitting). All have been corrected in each observation only by the error component which is common to all the quasar components and to *B, as reconstructed by the fitting. Other than this A is shown as observed (except for a vertical shift to have zero mean), but B and C are shifted and scaled according to the values of m_{AC}, m_{BA},

t_{AC}, t_{BA}, α_{AC}, and α_{BA}. Also shown are the reconstructed signal (solid line) and the input used in the Monte Carlo trials (dotted line). Figures 2 and 3 show one dimensional χ^2 plots, as a function of t_{BC} and r_{ABC}, respectively. In each plot, at every point the parameter shown is fixed and χ^2 is minimized with respect to all other parameters. The plots do not show any strong local minima that could be confused with the global minimum. The formal 1σ uncertainties derived from these curves are 1.0 days for t_{BC} and .12 for r_{ABC}. Figure 4 shows a two dimensional χ^2 plot as a function of t_{AC} and t_{BA}, around the minimum (marked $\times$). Note that the result of SCH (marked o) is outside even the outermost contour, which delineates the formal 99.99% confidence level.

We adopt the $L = 20$ results and bootstrap uncertainties, since they are compatible with the values for the other L to about 1σ. Thus $t_{BC} = 25.0^{+1.5}_{-1.7}$ days, and Tables 5.2 and 5.3 indicate that this 6% uncertainty (at 68% confidence) is a reasonable estimate. Our value used with lens models reduces the induced H_0 only by 5% relative to the SCH value. On the other hand, r_{ABC} varies fairly strongly with different assumptions, and is only weakly constrained at $r_{ABC} = 1.13^{+.18}_{-.17}$. This result is close to the values around 1.4 predicted by lens models, but the uncertainty is too large to be able to decide among different types of models (Keeton & Kochanek 1997). The fitting also recovers $m_{AC} = 2.033 \pm .005$, $m_{BA} = -2.534 \pm .007$, and $m_{BC} = -.501 \pm .008$, which yield variation-subtracted magnification ratios for the mean flux. These agree with SCH and also with the flux ratios of Kristian et al. (1993), since the quasar does not vary much over the time scale of the time delays. Our extremely accurate magnification ratios are not useful for lens modelling, since they are still likely to be greatly altered by microlensing, as suggested by the variation ratios.

Regarding the variation ratios, $\alpha_{BC} = 1.10^{+.08}_{-.08}$ is consistent with 1, while $\alpha_{AC} = 1.39^{+.07}_{-.08}$ is greater than 1 and $\alpha_{BA} = .79^{+.05}_{-.06}$ is less than 1, both at $4 - 5\sigma$. These values are also roughly consistent with Table 5.2. To make perhaps a more direct test of the significance of this result, we also perform Monte Carlo tests on input which has variation ratios equal to 1, but they are allowed to vary in the fitting. The result indicates that if the variation ratios were really 1, we would measure $\alpha_{AC} = 1.00^{+.07}_{-.07}$ ($^{+.14}_{-.17}$ at 95% confidence) and $\alpha_{BA} = 1.00^{+.10}_{-.10}$ ($^{+.19}_{-.23}$ at 95% confidence), so in this sense the $L = 20$ results for α_{AC} and α_{BA} are significantly different from 1 at 5σ and 2σ, respectively. As discussed in §5.2, this may indicate that the quasar optical emission region must be of order 1000 AU in linear dimension.

5.4 Conclusions

We have developed a method for determining time delays among light curves of multiple images of a gravitational lens. The method constructs a simple model for the actual source variation, using interpolation between a number of equally spaced values. It then performs a combined χ^2 minimization by fitting all of the light curves to this model simultaneously, which is similar to the method of Press et al. (1992a, b). The ability to vary the basic parameters of the model over a large range lends robustness to our method. Most of the parameters are linear and so the χ^2 minimiza-

tion is easily done. In addition to the time delays, the other non-linear parameters are relative variation ratios, which account for different fractional variation in flux for different images. We interpret this physically as evidence for differential microlensing, i.e. a different magnification due to microlensing for the varying region from the region giving rise to the mean flux.

Applying our method to the light curves of PG1115+080 observed by SCH, we find a value of $25.0^{+3.3}_{-3.8}$ days (95% confidence) for the delay between components B and C, and a ratio t_{AC}/t_{BA} for the two smaller delays of $1.13^{+.18}_{-.17}$ (68% confidence). Unlike SCH, we include correlated measurement errors as well as the above mentioned variation ratios in the analysis. Our result for t_{BC} agrees with SCH, but the ratio r_{ABC} does not. Our result for r_{ABC} does agree with lens models, but we find that with the present data it cannot be derived accurately enough to help in fitting lens models. For the variation ratios, we find $\alpha_{AC} = 1.39^{+.16}_{-.20}$ and $\alpha_{BA} = .79^{+.10}_{-.12}$ (95% confidence), each indicating differential microlensing at a significance of $4-5$ times our estimated 1σ uncertainties. If confirmed as the data accumulates, this would imply that the size of the quasar optical emission region is of order 1000 AU, for microlenses of $\langle M \rangle = 0.1 M_{\odot}$. Further data may also allow a determination of the time variation of the two microlensing magnifications.

Bibliography

[1] Falco, E. E., Gorenstein, M.V., & Shapiro, I. I. 1985, ApJ, 289, L1

[2] Gould, A., & Miralda-Escudé, J. 1997, preprint astro-ph/9612144

[3] Keeton, C. R., & Kochanek, C. S. 1997, preprint astro-ph/9611216

[4] Kochanek, C. S. 1991, ApJ, 373, 354

[5] Kristian, J. et al. 1993, AJ, 106, 1330

[6] Narayan, R. 1991, ApJ, 378, L5

[7] Press, W. H., Rybicki, G. B., & Hewitt, J. N. 1992a, ApJ, 385, 404

[8] Press, W. H., Rybicki, G. B., & Hewitt, J. N. 1992b, ApJ, 385, 416

[9] Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992c, Numerical Recipes in Fortran: The Art of Scientific Computing, Sec. Edition (New York: Cambridge Univ. Press)

[10] Schechter, P. L., et al. 1997, preprint astro-ph/9611051 (SCH)

[11] Schild, R. E. 1996, ApJ, 464, 125

[12] Wambsganss, J., & Paczyński, B. 1994, AJ, 108, 1156

[13] Wambsganss, J., Schneider, P., & Paczynski, B. 1990, ApJ, 358, 33L

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

[15] Witt, H. J., Mao, S., & Schechter, P. L. 1995, ApJ, 443, 18

[16] Young, P., Deverill, R. S., Gunn, J. E., Westphal, J. A., & Kristian, J. 1981, ApJ, 244, 723

Table 5.1: Single quasar component fitting

L	ndof	A, χ^2	B, χ^2	C, χ^2
10	61	40.4	70.1	40.2
20	51	29.7	39.9	29.8
30	41	27.4	30.3	27.6

Table 5.2: Fitting of quasar components in pairs

L	ndof	χ^2	time delay [days]	var ratio
BC components				
10	94	166	$25.0^{+1.7}_{-1.8}$	$1.07^{+.09}_{-.09}$
20	84	129	$25.3^{+2.3}_{-2.0}$	$1.09^{+.09}_{-.09}$
30	74	76.4	$26.7^{+1.8}_{-1.8}$	$1.03^{+.06}_{-.07}$
AC components				
10	94	119	$11.0^{+0.9}_{-0.9}$	$1.35^{+.06}_{-.06}$
20	84	88.7	$11.0^{+0.9}_{-0.9}$	$1.52^{+.09}_{-.08}$
30	74	72.7	$12.3^{+1.3}_{-1.0}$	$1.40^{+.09}_{-.10}$
BA components				
10	94	224	$11.7^{+2.0}_{-2.0}$	$0.74^{+.05}_{-.05}$
20	84	124	$6.75^{+1.7}_{-1.5}$	$0.61^{+.05}_{-.04}$
30	74	78.3	$7.22^{+1.2}_{-1.2}$	$0.46^{+.03}_{-.03}$

Table 5.3: Fitting of the three quasar components

Input	ndof	χ^2	t_{AC} [days]	t_{BA} [days]	t_{BC} [days]	r_{ABC}	α_{AC}	α_{BA}
$L = 10$	127	232	$12.5^{+0.9}_{-0.9}$	$14.0^{+1.6}_{-1.6}$	$26.5^{+1.7}_{-1.7}$	$0.89^{+.12}_{-.13}$	$1.45^{+.06}_{-.06}$	$.80^{+.05}_{-.05}$
$L = 20$	117	187	$13.3^{+0.9}_{-1.0}$	$11.7^{+1.5}_{-1.6}$	$25.0^{+1.5}_{-1.7}$	$1.13^{+.18}_{-.17}$	$1.39^{+.07}_{-.08}$	$.79^{+.05}_{-.06}$
$L = 30$	107	148	$14.9^{+1.4}_{-1.3}$	$10.6^{+1.3}_{-1.3}$	$25.5^{+1.8}_{-1.6}$	$1.41^{+.27}_{-.25}$	$1.43^{+.08}_{-.08}$	$.72^{+.04}_{-.05}$
$\sigma_n = 4$	117	266	13.0	11.8	24.8	1.10	1.38	.80
$\sigma_n = 8$	117	157	13.4	11.7	25.1	1.15	1.39	.79
$f_1 = 0$	117	235	12.8	11.4	24.2	1.12	1.34	.80
$f_2 = 0$	117	249	12.8	11.8	24.6	1.08	1.38	.81
cubic	117	183	13.4	11.9	25.2	1.13	1.39	.80
linear	117	181	13.5	11.7	25.1	1.15	1.40	.79
SCH-like	84	135	$8.9^{+1.7}_{-1.6}$	$11.9^{+2.8}_{-2.5}$	$20.9^{+2.9}_{-2.6}$	$0.75^{+.26}_{-.29}$	1	1

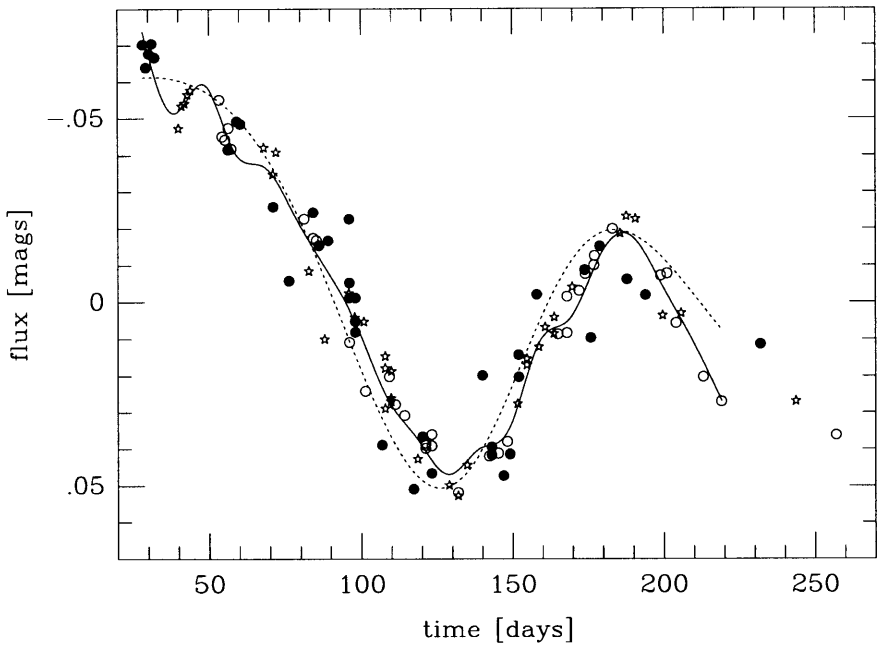

Figure 5-1: Light curves for components A (⋆), B (•), and C (◦). B and C have been shifted to match A, and all have been partly corrected for errors (see text). Also shown are the reconstructed signal (solid line), the input used in Monte Carlo trials (dotted line), and the final point in each light curve, which was excluded from the fitting.

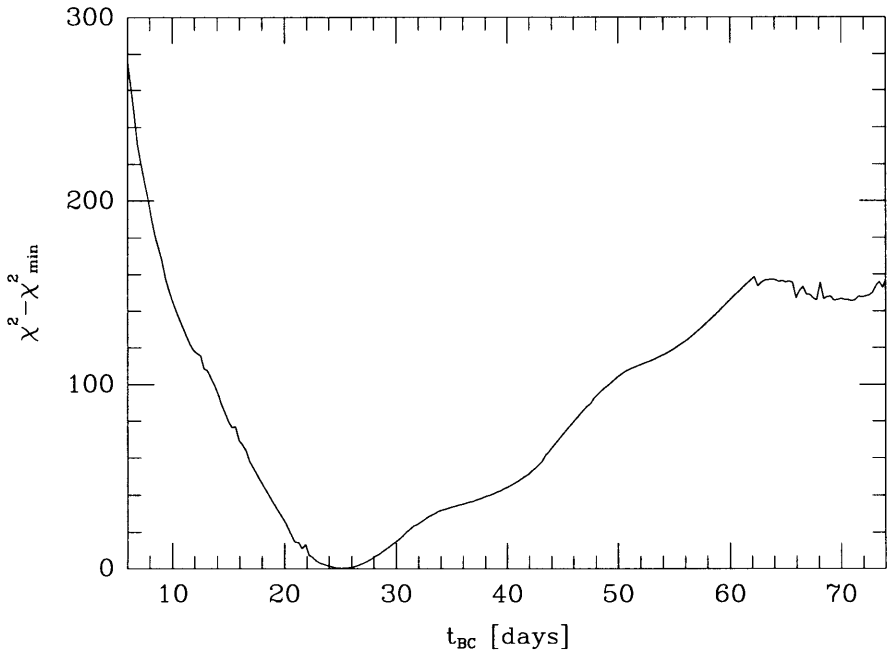

Figure 5-2: Value of χ^2 versus assumed delay t_{BC}, relative to the χ^2 value at its global minimum.

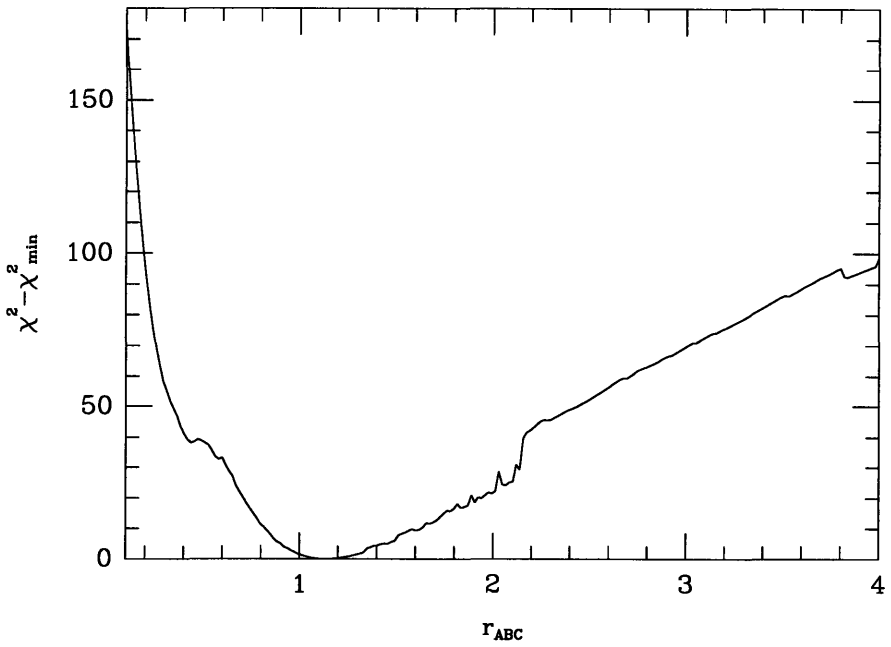

Figure 5-3: Value of χ^2 versus assumed ratio r_{ABC}, relative to the χ^2 value at its global minimum.

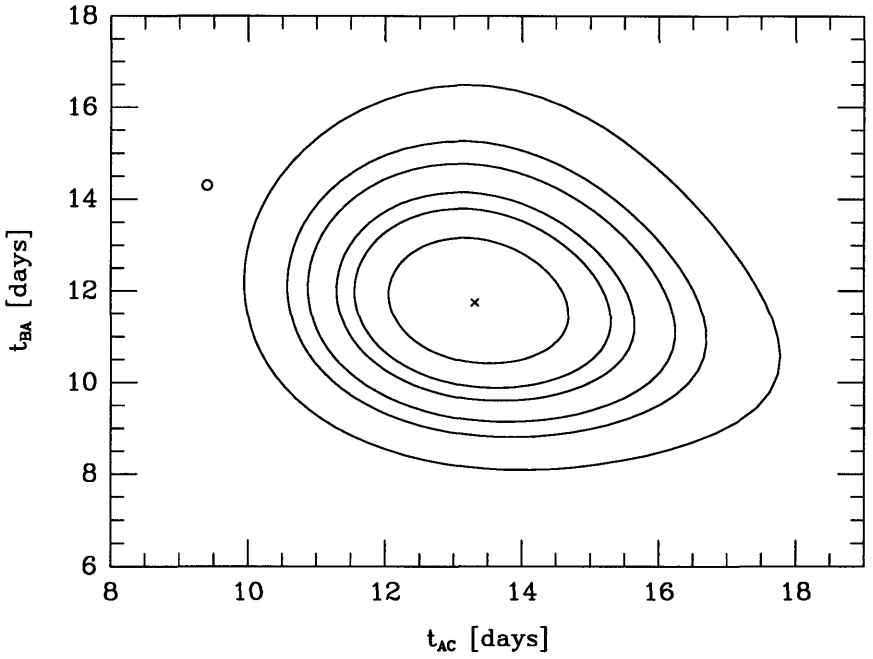

Figure 5-4: χ^2 contours in the t_{AC}, t_{BA} plane. The point marked $\times$ is $\chi^2 = 187$, the global mimimum. The contours are drawn at $\Delta\chi^2$ =2.30, 4.61, 6.17, 9.21, 11.8 and 18.4, the 1σ, 90%, 2σ, 99%, 3σ and 99.99% confidence levels for two parameters. The point marked o is the SCH result.

Chapter 6

Lens Models of 0957+561 and the Hubble Constant

6.1 Introduction

The observational field of gravitational lensing began with the discovery of the double quasar 0957+561 (Walsh et al. 1979). Monitoring was initiated almost immediately, with the goal of measuring a time delay between the arrival times of light via the two image paths. However, it was found that achieving the measurement accuracy required for a precise determination was very challenging. Optical (Lloyd 1981; Keel 1982; Florentin-Nielsen 1984; Schild & Cholfin 1986; Vanderriest et al. 1989; Schild & Thomson 1995) and radio (Lehár et al. 1992) monitoring programs produced extensive data, but analyses by a host of sophisticated techniques (e.g. Press, Rybicki & Hewitt 1992a, 1992b; Pelt et al. 1994, 1996) could not resolve the conflict between groups obtaining delays near 400 days and those finding delays close to 540 days. Only recently an optical detection of a sharp event in both images has resulted in a confirmation of the short delay with a 1% measurement of the time delay (Kundić et al. 1997). Some additional confidence in this measurement comes from the consistency with the latest results from radio monitoring (Haarsma et al. 1997). With the time delay measured, observations have begun to attempt to resolve the cluster degeneracy in 0957+561. This degeneracy in the Hubble constant determination is caused by the convergence due to the cluster, which can be traded back and forth with the convergence due to the lensing galaxy without affecting the image configuration. Thus a direct measurement of the mass of either the galaxy or the cluster is required in order to fix the cluster convergence and remove the degeneracy.

The cluster mass distribution can be determined from weak lensing, the shape distortions of background galaxies due to lensing by the cluster. Such a measurement is being attempted for the 0957+561 cluster (Fischer et al. 1997). However, the 0957+561 cluster is relatively not very massive, which makes the measurement difficult. Furthermore, what is needed to resolve the Hubble constant degeneracy is the value of the cluster convergence at the positions of the images. The contribution

of the lens galaxy itself to the weak lensing must be accounted for if the weak lensing measurements are to be useful. As for the lens galaxy, its velocity dispersion can be measured by taking its spectrum and measuring the width of spectral lines. This yields the velocity dispersion of the stars and not the dark matter, so the interpretation is difficult. For a given dark matter distribution, as specified by a lens model, it is possible to convert between the two but only under certain assumptions (see below). Until recently, the only velocity dispersion measurement was that of Rhee (1991), who found a value of 303 ± 50 km s^{-1}. Just this measurement error translates to a 33% uncertainty in H_0. A newer measurement by Falco et al. (1997, hereafter FSMD) with increased precision found 279 ± 12 km s^{-1}, but it also found a difference of 50 km s^{-1} between the velocity dispersion measured within or outside $0\farcs2$ of the galaxy center. Such a gradient may be due to a few billion $M_\odot$ black hole at the center, or it may indicate a systematic error. In the former case, FSMD suggest using the outside value not dominated by the black hole, i.e. 266 ± 12 km s^{-1}

The other component essential to translating a time delay measurement to a Hubble constant value is a well-constrained lens model. The most recent effort to explore models of 0957+561 was by Grogin & Narayan (1996a, 1996b, hereafter GN). They considered two types of models for the lens, and approximated the effect of the surrounding cluster as an additional constant shear term. The basic model of GN represents the lens galaxy as a softened power-law sphere (SPLS), a density profile which allows for both a core radius and an arbitrary radial index. The other type of model was adopted from earlier work by Falco et al. (1991, hereafter FGS). This model gives the galaxy a King profile, a generalization of the singular isothermal sphere which includes a core radius, and goes from constant density at the center to an isothermal profile at large radii. The models are strongly constrained by VLBI data which resolves jets in the two images into several components. These jet components were observed by Garrett et al. (1994, hereafter G94), who fit their positions to compute derivatives of several components of the relative magnification matrix of the two images. This matrix and its derivatives were in turn used by GN to constrain lens models. In this chapter we reconsider the lens models by comparing them directly to the positions of the jet components, a procedure which has several advantages over that of GN. It removes a layer of complexity involved in the fitting performed by G94, in which several degrees of freedom were lost. Also, GN used the positions of several components as constraints, in addition to using the magnification matrix and its derivatives, but these different constraints are not independent. More troubling may be the fact that the errors are of order 100% on some of the derivatives of the relative magnification matrix, and so treating the errors as Gaussian may not be reasonable. The errors given on the jet component positions may more accurately reflect the actual measurement errors. These positions have correlated errors, which we include.

One worry in using the results of GN is the poor reduced χ^2 of even their best-fitting lens models. Their SPLS gave a χ^2 per degree of freedom $\bar{\chi}^2 = 6.9$, and their FGS model reached $\bar{\chi}^2 = 3.8$. When we use the VLBI components directly we find below significantly better reduced χ^2 values for similar types of models, though the models still do not fit very well.

We also explore other possible models for the cluster. Since the cluster produces the external shear required in lens models, the magnitude and direction of this shear has implications for the position of the cluster center and its mass. Furthermore, for a given cluster mass profile the lens model can determine the cluster position and mass, which in turn determines the cluster convergence and resolves the time delay degeneracy. We assume that the cluster can be described as a Singular Isothermal Sphere (SIS). This is likely to be approximately true, but in the future the weak lensing measurements discussed above may determine the true profile more precisely.

6.2 Lens models

Using the notation of Chapter 1, we write the lens equation as

$$\vec{\beta} = \vec{\theta} - \vec{\alpha} \, . \tag{6.1}$$

The SPLS profile (GN) is characterized by a spherically symmetric volume density profile,

$$\rho(r) = \rho_0 \left(1 + \frac{r^2}{r_c^2}\right)^{(\eta-3)/2} , \tag{6.2}$$

with a corresponding projected surface density

$$\Sigma(\xi) = \Sigma_0 \left(1 + \frac{\xi^2}{r_c^2}\right)^{(\eta-2)/2} , \tag{6.3}$$

where $\Sigma_0 = \rho_0 r_C B(1/2, 1-\eta/2)$ and B is the Euler beta function. The deflection law is

$$\vec{\alpha}(\vec{\theta}) = \left(\frac{\alpha_E^2}{\theta^2}\right) \left[\frac{(\theta^2 + \theta_c^2)^{\eta/2} - \theta_c^\eta}{\alpha_E^\eta}\right] \vec{\theta} \, , \tag{6.4}$$

where $\alpha_E = \alpha_0^{2/(2-\eta)} \theta_C^{-\eta/(2-\eta)}$ and in radians

$$\alpha_0 = \left(\frac{8\pi G \Sigma_0 r_c^2}{c^2 D \eta}\right)^{1/2} , \tag{6.5}$$

with $D = D_{OL} D_{OS} / D_{LS}$. The parameters are thus a normalization α_E, core radius θ_c, and power η.

The FGS galaxy is described by a King profile. FGS adopted an analytic approximation introduced by Young et al. (1981) for the deflection law:

$$\begin{aligned} \vec{\alpha}(\vec{\theta}) \text{ [radians]} &= \left(\frac{D_L}{D}\right) \left(\frac{\sigma_v^2}{c^2}\right) \alpha_*(\theta) \, \hat{\theta} \, , \\ \alpha_*(\theta) &= 53.2468 \, f\left(1.155 \frac{\theta}{\theta_c}\right) - 44.0415 \, f\left(0.579 \frac{\theta}{\theta_c}\right) , \end{aligned} \tag{6.6}$$

$$f(x) \;=\; \frac{\sqrt{1+x^2}-1}{x}\,. \tag{6.7}$$

The parameters are a velocity dispersion σ_v and a core radius θ_c. In order to fit the data, FGS also included a black hole of mass $M_{\rm bh}$ at the center of the galaxy, which produces an

$$\vec{\alpha}(\vec{\theta}) = \left(\frac{\alpha_{\rm bh}^2}{\theta^2}\right)\vec{\theta}\,, \tag{6.8}$$

where the Einstein radius is

$$\alpha_{\rm bh} = \left(\frac{4GM_{\rm bh}}{c^2 D}\right)^{1/2} = 0\farcs91\left(\frac{M_{\rm bh}}{10^{11}\,h^{-1}\,M_\odot}\right)^{1/2}\,. \tag{6.9}$$

In order to add some ellipticity to the lens galaxy, we use the tilted Plummer potential from GN: $\vec{\alpha}(\vec{\theta})$ is given by the gradient with respect to $\vec{\theta}$ of the potential

$$\Psi = \frac{\alpha_E^2}{\eta}\left[\frac{u^2+\theta_c^2}{\alpha_E^2}\right]^{\eta/2}\,, \tag{6.10}$$

where $u^2 = r^2[1-\epsilon_p\cos 2(\theta-\theta_\epsilon)]$. The corresponding density profile is ellongated in the direction of θ_ϵ (measured from North through East), with an ellipticity e of (from GN):

$$1-e = \left(\frac{1-\epsilon_p}{1+\epsilon_p}\right)^{1/2}\left[\frac{1-\epsilon_p(2/\eta-1)}{1+\epsilon_p(2/\eta-1)}\right]^{\frac{1}{2-\eta}}\,. \tag{6.11}$$

In fact, for large ellipticities this elliptical potential model can have unphysical corresponding density contours (Blandford & Kochanek 1987). Because of this, GN restricted their model to the small ellipticity of $e = 0.3$ measured for the lens galaxy light profile by Bernstein et al. (1993). However, there is no guarantee that the dark matter has the same shape as the light profile, so it is interesting to test the ability of the lensing data to constrain the dark matter ellipticity. We thus let the ellipticity vary freely, and even at large ellipticities the elliptical potential model should give us an idea of the freedom that ellipticity adds to the modelling.

The lensing galaxy in 0957+561 is a massive galaxy near the center of a galaxy cluster. Following FGS, we assume that the cluster deflection varies on a large scale compared to the image separation, so we expand the cluster deflection about the lens galaxy and take it to have a linear deflection law, $\alpha_i = M_{ij}\theta^j$. The traceless part of the matrix M_{ij} is a shear γ with direction ϕ, where

$$M = \gamma\begin{pmatrix}\cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi\end{pmatrix}\,. \tag{6.12}$$

The trace part is a convergence κ, which corresponds to the degeneracy identified by Falco et al. (1985): Given any lens model, if we multiply the deflection $\alpha(\theta)$ by the factor $(1-\kappa)$ and at the same time include a convergence κ in the model, the relative image positions and magnifications remain unchanged. The time delay

changes, however, by the factor $(1-\kappa)$, inducing an uncertainty in the derived Hubble constant unless κ can be determined. GN note that because of this, models really only determine the scaled shear $\gamma' = \gamma/(1-\kappa)$, and (for a given measured time delay) a scaled value of h which we denote h', where

$$H_0 = 100\,h\,\mathrm{km\,s^{-1}\,Mpc^{-1}} \tag{6.13}$$

is standard notation and we also have

$$H_0 = 100\,h'\,(1-\kappa)\,\mathrm{km\,s^{-1}\,Mpc^{-1}}\ . \tag{6.14}$$

As noted above, a direct measurement of the mass of the lens galaxy or the cluster can determine κ.

As an independent attempt to determine κ, we also model the cluster as an SIS with a free position, letting the fit determine the position as well as the velocity dispersion. For this model, $\rho(r) \propto 1/r^2$, $\Sigma(\xi) \propto 1/\xi$, and

$$\vec{\alpha}(\vec{\theta}\,) = b_{\rm cl}\hat{\theta}'\ , \qquad b_{\rm cl} = 4\pi\left(\frac{\sigma_{\rm cl}}{c}\right)^2\frac{D_{OL}}{D} = 17\farcs3\left(\frac{\sigma_{\rm cl}}{1000\,\mathrm{km\,s^{-1}}}\right)^2\ , \tag{6.15}$$

where $\sigma_{\rm cl}$ is the velocity dispersion of the cluster and $\vec{\theta}' = \vec{\theta} - \vec{\theta}_{\rm cl}$. The cluster parameters in this case are thus $\sigma_{\rm cl}$ and the coordinates $(x_{\rm cl}, y_{\rm cl})$ of the cluster center $\vec{\theta}_{\rm cl}$ with respect to the lens galaxy position. GN considered this type of profile for the cluster but did not use it as part of their lens model. Bernstein et al. (1993) included an isothermal cluster in some of their models.

6.3 Observational constraints

While optical observations of 0957+561 can yield only the two image positions, radio images have resolved the source and revealed internal structures. Early VLBI observations (Porcas et al. 1981) found that both components have a core-jet radio structure. Improved maps (Gorenstein et al. 1988) resolved the A and B jets into three components each, enabling a determination of the relative magnification matrix. The maps of G94 further resolved the jets into five components each, denoted $A_{2\ldots6}$ and $B_{2\ldots6}$ (A_1, B_1 denote the core), and provided evidence for a magnification gradient. As explained above in the introduction, we use the data of G94 by fitting directly to the jet component positions, thus bypassing the complicated non-linear fitting of the magnification matrix and its derivatives to these positions. G94 give the jet component positions in radial coordinates, with respect to the cores. While the errors are given for each radius and position angle, there are also correlated errors among the radius and position angle of each component, as well as correlations among the positions of all the A components and separately among all the B components. These correlations are significant, since the positions of the Gaussian components are the result of a combined overall fit to the VLBI map of each image. We obtained these correlations from Garrett (1997), and include them in the model fitting. Following

G94, we also include two magnification ratios as constraints, a B/A ratio of $0.75 \pm .02$ for the core, and $0.63 \pm .03$ for the jet, the latter compared with the model at the most luminous jet component (A_5, B_5). For the $A_1 - B_1$ separation we adopt the value of $(-1\farcs25271, 6\farcs04662)$ with $0\farcs00004$ uncertainty from Gorenstein et al. (1988), as do GN. This yields a total of 26 constraints, 4 for each of 6 pairs of positions (relative to the lens galaxy position, discussed below) plus the two magnification ratios.

Models having a smooth surface mass density produce a third image of 0957+561, typically demagnified and near the center of the lens galaxy. No such image has been seen down to a 5σ limit of 1/30 the flux of image B (Gorenstein et al. 1984). We follow the approach of GN of penalizing models only once they begin to exceed this 5σ limit, which GN achieve by adding to the χ^2 a term

$$\chi^2_{C/B} = \begin{cases} 0 & C/B < 1/30 \\ \frac{(C/B - 1/30)^2}{(1/150)^2} & C/B > 1/30 \end{cases} \tag{6.16}$$

In the SPLS model, the core radius determines the degree of central mass concentration and is the parameter most sensitive to the third-image flux limit.

The final pair of data are the lens galaxy coordinates, whose values are currently disputed. GN took the optical center of brightness (Stockton 1980) $G1$ at $(0\farcs19, 1\farcs00)$ from image B, with an uncertainty in each component of 30 milliarcseconds (mas). However, two radio observations have found nearby sources that disagree with each other. A point-source G at $(0\farcs151, 1\farcs051)$ was found in VLA observations (Roberts et al. 1985), while a point source G' at $(0\farcs181, 1\farcs029)$ was found by VLBI (Gorenstein et al. 1983), with both measurements claiming 1 mas accuracy. New HST observations yield a lens center close to G' (Bernstein 1997), and so we adopt the G' position but with 10 mas accuracy. Such an accuracy is a reasonable guess for HST, and we choose not to adopt the VLBI uncertainty, since even if G' is indeed associated with the lens galaxy we probably cannot trust the position of the radio source to coincide with the center of the lens potential to an accuracy as high as 1 mas.

In summary, we have 26 constraints, plus one more for non-singular models which produce a third image. The SPLS and FGS models have 17 parameters : 3 for the lens galaxy profile , 2 for the cluster shear, and 12 for the 6 source positions. The number of degrees of freedom (dof) is thus 10 for the SPLS model, 9 for the FGS model, and 8 for the Plummer model with ellipticity. An additional degree of freedom is lost if the cluster is modelled as an SIS instead of an external shear.

6.4 Stellar velocity dispersion

As noted above, the recent measurement of the velocity dispersion by FSMD offers the possibility of resolving the cluster degeneracy in 0957+561, but its interpretation requires translating the stellar velocity dispersion into a constraint on the dark matter distribution of the lens galaxy. GN follow the approach of Kochanek (1993), who assumes that the stars are in equilibrium and then uses the Jeans equation, which describes this equilibrium, to find the line-of-sight velocity dispersion for a given

dark matter distribution. Kochanek, however, only gives the formula for a singular isothermal sphere dark matter distribution. In what follows we derive a more general formula, carefully noting the various assumptions necessary for this approach to be valid.

We must first assume spherical symmetry, for the three-dimensional profiles $\nu(r)$ of the luminous matter density and $M(r)$ of the total (dark) mass enclosed within radius r. We assume a constant anisotropy factor q for the radial and tangential stellar velocity dispersions, $\sigma_\theta^2(r) = \sigma_\phi^2(r) = (1-q)\sigma_r^2(r)$. Then Binney & Tremaine (1987) give the Jeans equation in this case:

$$\frac{GM(r)}{r} = -\sigma_r^2 \left(\frac{d\ln\nu}{d\ln r} + \frac{d\ln\sigma_r^2}{d\ln r} + 2q \right) . \tag{6.17}$$

This equation is a first-order linear differential equation for σ_r^2 which is solved by

$$\sigma_r^2(r) = G \int_r^\infty \frac{\nu(s)}{\nu(r)} \frac{M(s)}{s^2} \left(\frac{s}{r} \right)^{2q} ds . \tag{6.18}$$

By projection, the line-of-sight velocity component at a projected distance R is given by (Binney & Tremaine 1987)

$$I(R) v_{\rm los}^2(R) = 2 \int_R^\infty \left(1 - q \frac{R^2}{r^2} \right) \frac{\nu \sigma_r^2 r\, dr}{\sqrt{r^2 - R^2}} , \tag{6.19}$$

where $I(R)$ is the surface brightness profile. Substituting equation 6.18 into equation 6.19 and reversing the order of integration we finally derive

$$I(R) v_{\rm los}^2(R) = 2G \int_R^\infty dr\, \nu(r) M(r) r^{(2q-2)} \int_0^{\sqrt{r^2-R^2}} \frac{du}{(u^2+R^2)^q} \left(1 - q \frac{R^2}{u^2+R^2} \right) . \tag{6.20}$$

For 0957+561 we follow GN by taking the observed de Vaucouleurs profile

$$I(R) = I_e \exp\left\{ -7.67 \left[\left(\frac{R}{R_e} \right)^{1/4} - 1 \right] \right\} , \tag{6.21}$$

with $R_e = 4\farcs5 \pm 0\farcs64$ (Bernstein et al. 1993), and using Abel inversion to compute

$$\nu(r) = -\frac{1}{\pi} \int_r^\infty \frac{dI(R)}{dR} \frac{dR}{\sqrt{R^2 - r^2}} . \tag{6.22}$$

In using these various formulas, the assumption of spherical symmetry is key. In practice, the galaxy is observed to have an ellipticity of 0.30 (Bernstein et al. 1993), and lens models allow a substantial ellipticity for the dark matter (see below). There can be other projection effects if the galaxy is triaxial, and the induced uncertainty in $v_{\rm los}$ from these various possibilities is unclear. Even in the case of spherical symmetry we must assume that the velocity anisotropy is small, i.e. that q is not far from zero. If an accurate determination of κ is ever made from a direct measurement of the

Table 6.1: Fitted parameters and χ^2 values, SPLS + shear (10 dof)

Parameter	Value	χ^2	Value
α_E	$2\farcs507$	$\chi^2_{\rm gal}$	11.4
θ_c	$0''$	$\chi^2_{\rm pos}$	16.9
η	1.159	$\chi^2_{\rm mag}$	5.8
γ'	0.245	$\chi^2_{C/B}$	0
ϕ	$-65\fdg6$	$\chi^2_{\rm tot}$	34.1

cluster mass distribution, its interpretation will probably be less problematic than that of $v_{\rm los}$.

6.5 Results

To evaluate the χ^2 of the models we compare them to the data directly in the image plane. This inversion of the lens mapping makes all the parameters non-linear, and we search for the χ^2 minimum using direction set methods (Press et al. 1992, §10.5). For the SPLS model, our best-fit model has a χ^2 of 34.1, and thus a reduced χ^2 of $\bar{\chi}^2 = 34.1/10 = 3.41$. The parameter values are given in Table 6.1, and are all within the 95% confidence limits of the best-fit model of GN. Also given are contributions to the total chisq $\chi^2_{\rm tot}$ from the galaxy position ($\chi^2_{\rm gal}$), from the image positions after fitting for the galaxy ($\chi^2_{\rm pos}$), from the two magnification ratios described above ($\chi^2_{\rm mag}$), and from the third-image flux limit ($\chi^2_{C/B}$). Estimates of the observables from this model are compared with their measured values in Table 6.2. The largest disagreement is for the x component of the galaxy position, which the model prefers to be off by $0\farcs034$. Figure 6-1 illustrates the great precision demanded from the model by the G94 data. It shows the reconstructed positions of the six source components and the corresponding A and B image component positions as given by the model (o). The G94 positions (×) are also shown for comparison. The differences between the observed and reconstructed positions correspond to $\chi^2_{\rm pos}$.

With the value of 417 ± 3 days (at 95% confidence) for the time delay (Kundić et al. 1997), our model yields an

$$H_0 = 79^{+6}_{-2}\,(1-\kappa)\,{\rm km\,s^{-1}\,Mpc^{-1}}\ . \tag{6.23}$$

This is close to the $h' = 0.82$ value of the corresponding model of GN. The uncertainties are 95% bounds derived (as in GN) from the condition $\Delta\chi^2 = 4\bar{\chi}^2$ (and we also added the small uncertainty in the time delay measurement). Figure 6-2 shows a plot of the minimum χ^2 as a function of h'. We find this curve using the method of Lagrange multipliers as suggested by GN; for each value of the Lagrange multiplier λ, minimizing the quantity $\chi^2 + \lambda h'$ yields a point on the $\chi^2(h')$ curve. The minimum is at $h' = 0.789$ and the 95% confidence interval is $0.771 < h' < 0.853$.

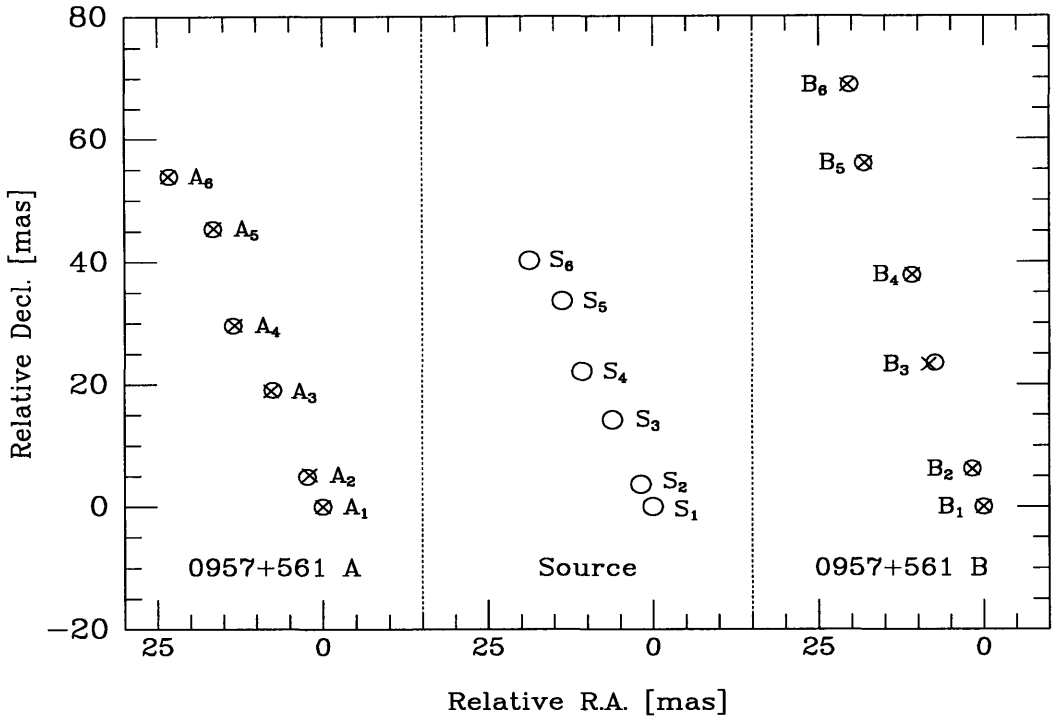

Figure 6-1: Comparison of SPLS model to G94 VLBI data. The reconstructed sources are shown along with corresponding model positions for the A and B image components (o). The G94 positions (×) are also shown. The three panels are at arbitrary offsets but have the same scale.

Calculating $v_{\rm los}$ for our model, with isotropic orbits $q = 0$, we obtain 320 km s^{-1} which implies an

$$H_0 = 79^{+10}_{-7} \left(\frac{\sigma_{\rm obs}}{320\,{\rm km\,s^{-1}}}\right)^2 {\rm km\,s^{-1}\,Mpc^{-1}}\ . \tag{6.24}$$

Here we have included uncertainties in converting to a stellar velocity dispersion (§6.4): a 4% uncertainty from a 2σ variation in the measured effective radius of the light, and an 8% uncertainty from stellar anisotropy allowing $|q| \leq 0.2$. Note that the latter is an assumption, i.e. if we allow a larger range $|q| \leq 0.5$ the induced error in H_0 can be as large as 28%.

So far we have not included the uncertainty due to large-scale structure, but we wish to include it in the final values of the Hubble constant. We analyze the effect of large-scale structure on the Hubble constant in Chapter 3. There we show that the large-scale structure uncertainty depends on how the lens model is normalized. Large-scale structure normally causes an uncertainty Δ_1, but for models which we normalize to the observed velocity dispersion, the velocity dispersion effectively constrains part of the effect of large-scale structure, and a smaller uncertainty Δ_2 is left over. Given the source and lens redshifts for 0957+561, Chapter 3 suggests 2σ uncertainties of $\Delta_1 = 11.8\%$ and $\Delta_2 = 6.6\%$, but the large-scale structure power spectrum adopted there tends to be at the high end of typical models (see e.g. Figure 2 of Keeton et al.

Table 6.2: Estimates of observables, SPLS + shear

Observable	Model estimate	$\frac{\text{Estd.} - \text{Obsvd.}}{\text{Obsvd. Err.}}$
$A_1 - B_1$	$(-1\farcs25271, 6\farcs04662)$	(-.003,.02)
$A_2 - A_1$	$(0\farcs00234, 0\farcs00487)$	(.14,-.38)
$A_3 - A_1$	$(0\farcs00750, 0\farcs01903)$	(-.32,.10)
$A_4 - A_1$	$(0\farcs01352, 0\farcs02956)$	(.47,-.17)
$A_5 - A_1$	$(0\farcs01652, 0\farcs04537)$	(.92,-.42)
$A_6 - A_1$	$(0\farcs02318, 0\farcs05388)$	(-.70,.82)
$B_2 - B_1$	$(0\farcs00175, 0\farcs00624)$	(.25,.39)
$B_3 - B_1$	$(0\farcs00746, 0\farcs02350)$	(-1.6,.44)
$B_4 - B_1$	$(0\farcs01089, 0\farcs03783)$	(.08,-.03)
$B_5 - B_1$	$(0\farcs01818, 0\farcs05593)$	(2.0,-.54)
$B_6 - B_1$	$(0\farcs02033, 0\farcs06885)$	(-1.4,.14)
B/A, core	0.715	-1.7
B/A, jet	0.680	1.7
$G' - B_1$	$(0\farcs215, 1\farcs025)$	(3.4,-.38)

1997). We adopt $\Delta_1 = 9.8\%$ and $\Delta_2 = 5.5\%$, and hereafter we add the appropriate uncertainty to final values of h, but not to intermediate values of h'.

Thus if we use 266 ± 12 km s^{-1} for the line-of-site velocity dispersion (FSMD), then this measurement error dominates the final uncertainty, and we get at 2σ

$$H_0 = 55^{+13}_{-12}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}\ , \tag{6.25}$$

while 279 ± 12 km s^{-1} yields $h = 0.60 \pm .13$ (We follow FSMD in quoting 1σ errors on their velocity dispersion measurements, but we give 2σ errors on H_0).

As noted by GN, there is some degeneracy in the fitting between the two radial parameters, η and θ_c. However, this degeneracy begins to break down when the core radius is large enough that the third image is not sufficiently demagnified. Indeed, if we fit an isothermal SPLS model, forcing $\eta = 1$, the model requires a core radius of $\theta_c = 0\farcs127$ and yields a $\bar{\chi}^2 = 6.16$, with about half the χ^2 coming from the constraint on the third-image flux.

The FGS model has a $\bar{\chi}^2 = 3.04$, somewhat lower than the SPLS. Two distinct models are in fact found with nearly identical $\bar{\chi}^2$ values, with the parameter sets given in Table 6.3. One solution yields an $h' = 0.79$ and the other an $h' = 0.86$, compared to $h' = 0.94$ for the FGS model of GN. In both of our models the errors are dominated by the image positions, with the galaxy position contributing only about 0.06 to $\bar{\chi}^2$. To normalize these models using $\sigma_{\rm obs}$, we can assume that $\sigma_{\rm obs}$ can be substituted for σ_v. In this case we obtain $h = 0.42$ or $h = 0.50$ for the two models, using the higher value of $\sigma_{\rm obs}$. However, this normalization is not self-consistent. Indeed, FSMD note that their measurement of a velocity gradient suggests the existence of a black hole of mass $4 \times 10^9 h^{-1} M_\odot$. A $100 \times 10^9 h^{-1} M_\odot$ point mass, as required by the models,

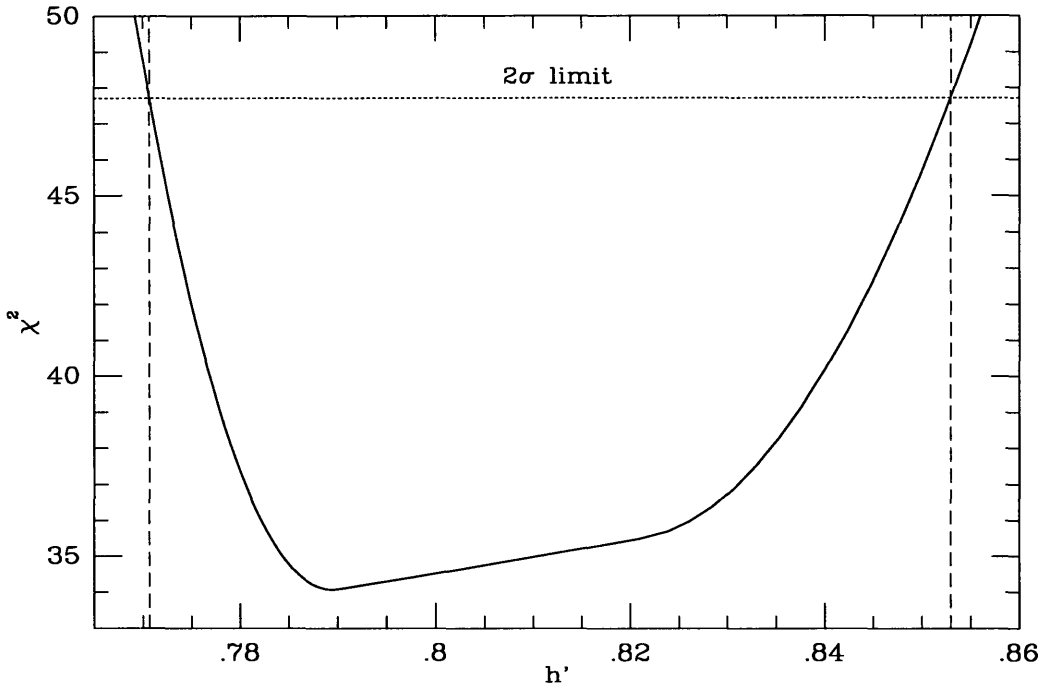

Figure 6-2: Plot of minimum χ^2 for SPLS models producing a given h', modelled after Figure 3 of GN. Dashed vertical lines show the 95% confidence limits. The steep rise for $h' > .83$ results from the third-image flux limit, as noted by GN.

would be expected to produce a much larger gradient, and thus appears unphysical.

For comparison with GN, we fit the SPLS model as above but using the optical G1 position and uncertainty. The result is $\bar{\chi}^2 = 2.27$, with parameter values given in Table 6.4 in the column labeled Value(1). This model yields an $h' = 0.79$. On the other hand, we may attempt to fit to the VLBI lens position, along with its 1 mas uncertainty. The resulting parameter values are also given in Table 6.4, in the column labeled Value(2). While the Hubble constant remains about the same, $h' = 0.77$, the model cannot fit the lens galaxy to this accuracy and yields $\bar{\chi}^2 = 70.4$. As a different check on robustness, we fit the SPLS model to the lower-resolution VLBI data of Gorenstein et al. (1988). For this we return to the VLBI lens galaxy position with

Table 6.3: Fitted parameters, FGS + shear (9 dof)

Parameter	Value(1)	Value(2)
θ_c	$2\farcs66$	$2\farcs24$
σ_v	384 km s^{-1}	367 km s^{-1}
$M_{\rm bh}$	$102 \times 10^9 h^{-1} M_\odot$	$106 \times 10^9 h^{-1} M_\odot$
γ'	0.192	0.220
ϕ	$-64\fdg1$	$-64\fdg4$

Table 6.4: Fitted SPLS + shear parameters, other data (10 dof)

Parameter	Value(1)	Value(2)	Value(3)	Value(4)
α_E	$2\farcs455$	$2\farcs450$	$2\farcs480$	$2\farcs506$
θ_c	$0''$	$0''$	$0\farcs0004$	$0''$
η	1.152	1.160	1.155	1.159
γ'	0.259	0.254	0.251	0.248
ϕ	$-66\fdg3$	$-66\fdg8$	$-66\fdg0$	$-65\fdg7$

10 mas error, and just as we modelled the six component positions of G94 we now model the four components of Gorenstein et al. (but assuming no correlated errors among jet components in this case). We also use the same two magnification ratios we used above, with the jet value of $B/A = 0.63 \pm .03$ compared to the models at the brightest jet component, the middle one of the three. We list the parameters in the column labeled Value(3) in Table 6.4. They are not very different from our standard SPLS model of Table 6.1, and the Hubble constant is also similar at $h' = 0.79$. With 6 degrees of freedom, the agreement is slightly worse at $\bar{\chi}^2 = 3.58$. Finally, we return to the model and data of Table 6.1, but include only the data as given in G94, i.e. without any correlated errors. With the parameters shown in the column labeled Value(4) in Table 6.4, we find $h' = 0.79$ and $\bar{\chi}^2 = 4.19$.

Returning to our standard set of data, we model the galaxy once again as an SPLS but consider an SIS cluster with variable position. There are now 9 degrees of freedom, and the result is a $\bar{\chi}^2 = 3.55$, slightly worse than the external shear model of Table 6.1. The parameter values are given in Table 6.5. Our input assumption is an SIS profile for the cluster, but as output we get estimates of the cluster center and velocity dispersion as well as the velocity dispersion of the lens galaxy and the Hubble constant, all with no remaining degeneracy. Regarding the lens, this model predicts a velocity dispersion of 270 km s^{-1}, in agreement with the observed values given above. The Hubble constant is

$$H_0 = 59^{+7}_{-6}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}\ . \tag{6.26}$$

In this 2σ error we include uncertainties from the model and from large-scale structure, but note that the value relies on the assumption of an SIS cluster profile.

Regarding properties of the cluster, Figure 6-3 compares observations and lens model predictions for the cluster center and velocity dispersion. The top plot shows position in terms of right ascension and declination, both measured with respect to the VLBI galaxy position G'. The image positions A and B are also shown. The bottom plot has the same x-axis, and plots values of the cluster velocity dispersion. The SPLS + shear model yields an external shear magnitude and direction. If produced by a spherical cluster of arbitrary radial profile, the shear direction confines the cluster center to lie on the straight line shown going through G'. If the cluster is an SIS, the shear magnitude determines the velocity dispersion as a function of distance from G',

Table 6.5: Fitted parameters, SPLS + SIS (9 dof)

Parameter	Value
α_E	$2\farcs050$
θ_c	$0''$
η	1.249
$x_{\rm cl}$	$18\farcs4$
$y_{\rm cl}$	$11\farcs1$
$\sigma_{\rm cl}$	652 km s^{-1}

and this dispersion is the curve shown in the bottom plot. The SPLS + SIS model determines the cluster position $C_{\rm SIS}$ with a velocity dispersion $\sigma_{\rm SIS} = 652$ km s^{-1}. Fischer et al. (1997) found cluster centroids of $C_{\rm lens} = (18'', 13'')$ from weak lensing and $C_{\rm gal} = (26'', 19'')$ from galaxy number contours, although each of these positions has a large 1σ uncertainty of order $15''$. Still, they are both offset from G' in the same direction, and also agree approximately with the positions from the two lens models. We note that external shear is symmetric about G', i.e. for a given external shear the cluster can be anywhere on the line shown through G', on *either side* of G'. However, if we try to find an SIS cluster on the other side, no minimum is found and the model pushes the SIS out to infinite distance, recovering the external shear solution. Thus, while the cluster is far enough away so that its effect is mostly an external shear, the model can also detect the effects of higher order terms, but our best-fit SIS position is better than infinity with a formal significance of only about 1σ. Weak lensing determines a velocity dispersion of $\sigma_{\rm lens} = 698 \pm 210$ km s^{-1}. Garrett et al. (1992) and Angonin-Willaime et al. (1994) measured redshifts for 21 probable cluster members, obtaining $\sigma_{\rm gal} = 715 \pm 130$ km s^{-1}. For the two measured values, these 1σ uncertainties are shown in the figure. The close agreement with the SIS lens model is encouraging. Figure 6-3 suggests that the lens galaxy may indeed *not* be right at the center of the cluster.

Returning to the external shear model of the cluster, we consider the effect of including a galaxy ellipticity using the tilted Plummer potential model. If we try the elliptical galaxy without an external shear, the model is driven to an ellipticity of 1 but still cannot fit, with a $\bar{\chi}^2 = 182$. With external shear, we find two minima with similar χ^2 values. The parameter sets are given in Table 6.6. The first model has a $\bar{\chi}^2 = 3.06$, somewhat lower than the SPLS model of Table 6.1, and an $h' = 0.96$. The second model yields $\bar{\chi}^2 = 3.07$ and $h' = 0.86$. Thus the addition of ellipticity allows a value of h' that is 20% higher than that of the spherical SPLS model. Also, once we allow the underlying mass distribution to be elliptical, it becomes very hard to use the observed stellar velocity dispersion. On the other hand, these models may be unrealistic. They correspond to a density ellipticity of 0.71 and 0.61, respectively. These are much larger than the observed $e = 0.3$ (Bernstein et al. 1993), and in very different directions from the observed θ_ϵ of about 55°. Since the images are located near the effective light radius of $4\farcs5$, where the stellar mass may contribute

Table 6.6: Fitted parameters, elliptical Plummer (8 dof)

Parameter	Value(1)	Value(2)
α_E	$2\farcs886$	$2\farcs719$
θ_c	$0''$	$0''$
η	1.132	1.142
γ'	0.414	0.374
ϕ	$-35\fdg7$	$-43\fdg0$
ϵ_p	0.422	0.333
θ_ϵ	$-24\fdg3$	$-28\fdg8$

substantially, the total mass probably cannot be so distorted compared to the light shape. New HST images of the lens galaxy should allow a more precise determination of the light profile and shape. Adding ellipticity also allows larger external shear and at a different direction from that found by the SPLS model. Future weak lensing measurements of the cluster mass center and profile may limit the external shear that can be included in this "two-shear" model.

6.6 Conclusions

We have modelled 0957+561 using the high-resolution VLBI data of G94, using as direct constraints their measurements of the positions of six radio components in each image. Compared to GN, even though we use a different lens position with smaller errors of 10 mas, we find a substantially improved reduced χ^2 of 3.41 for the SPLS + shear model. The change in the Hubble constant is, however, small. We find

$$H_0 = 79^{+6}_{-2}\,(1-\kappa)\,\mathrm{km\,s^{-1}\,Mpc^{-1}}\ , \tag{6.27}$$

and if we use the two values of the line-of-site velocity dispersion, 266 ± 12 km s^{-1} and 279 ± 12 km s^{-1}, we obtain $h = 0.55^{+.13}_{-.12}$ and $h = 0.60 \pm .13$ at 2σ, respectively. Normalizing lens models using $\sigma_{\rm obs}$ requires a number of simplifying assumptions, so as an independent check we also try an SPLS + SIS model. This model gives a Hubble constant of

$$H_0 = 59^{+7}_{-6}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}\ , \tag{6.28}$$

and predicts a cluster center and velocity dispersion as well as a lens galaxy velocity dispersion, all of which agree with observations. The result, however, relies on the assumption of an SIS profile. The true profile may be determined in the future from weak lensing measurements of the cluster mass distribution.

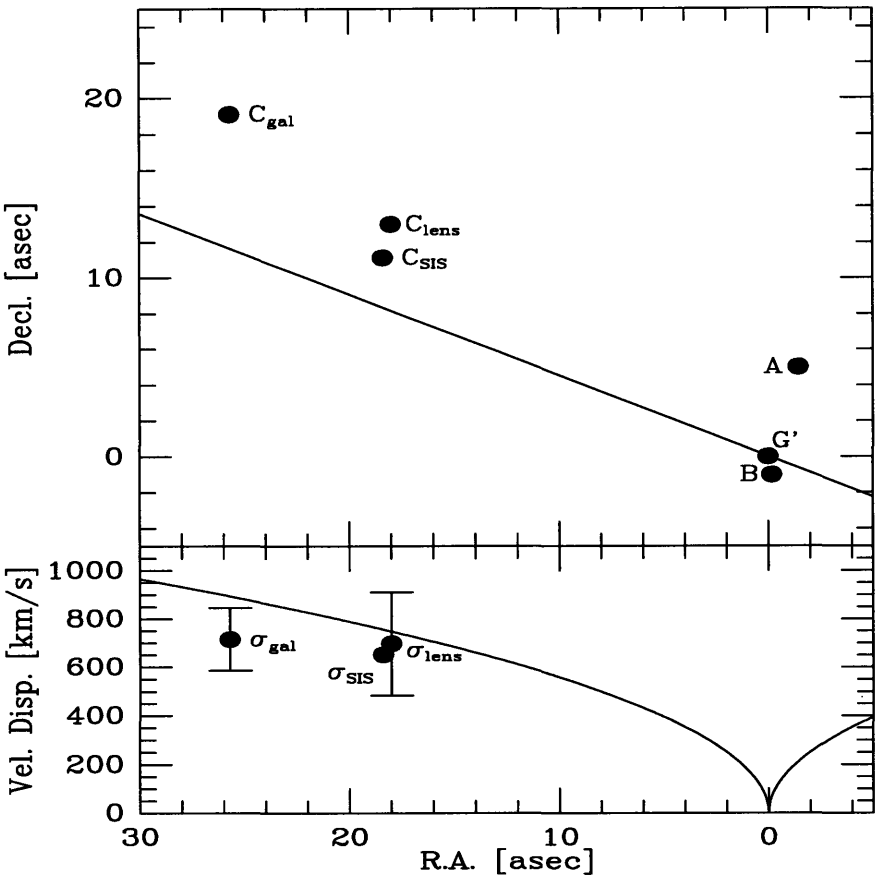

Figure 6-3: Comparison of observations to lens model predictions for the cluster center and velocity dispersion. Both plots have right ascension on the x-axis, and the top plot also shows the declination while the bottom plot shows corresponding values for the cluster velocity dispersion. Positions are measured with respect to the lens galaxy G'. Also shown in the top plot are the image positions A and B, the line of possible cluster centers from the SPLS + shear model, the cluster position ($C_{\rm SIS}$) from the SPLS + SIS model, and observed centers from weak lensing ($C_{\rm lens}$) and from galaxy number contours ($C_{\rm gal}$). In the bottom plot, the solid curve is the velocity dispersion from the SPLS + shear model. Also shown are the velocity dispersion ($\sigma_{\rm SIS}$) from the SPLS + SIS model, and observed dispersions from weak lensing ($\sigma_{\rm lens}$) and from redshift measurements of cluster members ($\sigma_{\rm gal}$).

Bibliography

[1] Angonin-Willaime, M. C., Soucail, G., & Vanderriest, C. 1994, A& A, 291, 411

[2] Bernstein, G. M. 1997, personal communication

[3] Bernstein, G. M., Tyson, J. A., & Kochanek, C. S. 1993, AJ, 105, 816

[4] Binney, J., & Tremaine, S. D. 1987, Galactic Dynamics (Princeton: Princeton University Press)

[5] Blandford, R. D., & Kochanek, C. S. 1987, ApJ, 321, 658

[6] Falco, E. E., Gorenstein, M.V., & Shapiro, I. I. 1985, ApJ, 289, L1

[7] Falco, E. E., Shapiro, I. I., Moustakas, L. A., & Davis, M. 1997, preprint astro-ph/9702152

[8] Falco, E. E., Gorenstein, M. V., & Shapiro, I. I. 1991, ApJ, 372, 364

[9] Fischer, P., Bernstein, G., Rhee, G., & Tyson, J. A. 1997, A&A, 113, 521

[10] Florentin-Nielsen, R. 1984, A&A, 138, L19

[11] Garrett, M. A. 1997, personal communication

[12] Garrett, M. A., Calder, R. J., Porcas, R. W., King, L. J., Walsh, D., & Wilkinson, P. N. 1994, MNRAS, 270, 457 (G94)

[13] Garrett, M. A., Walsh, D., & Carswell, R. F. 1992, MNRAS, 254, 27p

[14] Gorenstein, M. V., Cohen, N. L., Shapiro, I. I. Rogers, A. E. E., Bonometti, R. J., Falco, E. E., Bartel, N, & Marcaide, J. M. 1988, ApJ, 334, 42

[15] Gorenstein, M. V., Shapiro, I. I. Rogers, A. E. E., Cohen, N. L., Corey, B. E., Porcas, R. W., Falco, E. E., Bonometti, R. J., Preston, R. A., Rius, A., & Whitney, A. R. 1984, ApJ, 287, 538

[16] Gorenstein, M. V., et al. 1983, Science, 219, 54

[17] Grogin, N. A., & Narayan, R. 1996a, ApJ, 464, 92 (GN)

[18] Grogin, N. A., & Narayan, R. 1996b, erratum, ApJ, 473, 570 (GN)

[19] Haarsma, D. B., Hewitt, J. N., Lehár, J., & Burke, B. F. 1997, preprint astro-ph/9607080

[20] Keel, W. C. 1982, ApJ, 255, 20

[21] Keeton, C. R., Kochanek, C. S., & Seljak, U. 1997, preprint astro-ph/9610163

[22] Kochanek, C. S. 1993, ApJ, 419, 12

[23] Kundić, T., Turner, E. L., Colley, W. N., Gott, J. R., Rhoads, J. E., Wang, Y., Bergeron, L. E., Gloria, K. A., Long, D. C., Malhorta, S., & Wambsganss, J. 1997, preprint astro-ph/9610162

[24] Lehár, J., Hewitt, J. N., Roberts, D. H., & Burke, B. F. 1992, ApJ, 384, 453

[25] Lloyd, C. 1981, Nature, 294, 727

[26] Pelt, J., Hoff, W., Kayser, R., Refsdal, S., & Schramm, T. 1994, A&A, 286, 775

[27] Pelt, J., Kayser, R., Refsdal, S., & Schramm, T. 1996, A&A, 305, 97

[28] Porcas, R. W., Booth, R. S., Browne, I. W. A., Walsh, D., & Wilkinson, P. N. 1981, Nature, 289, 758

[29] Press, W. H., Rybicki, G. B., & Hewitt, J. N. 1992a, ApJ, 385, 404

[30] Press, W. H., Rybicki, G. B., & Hewitt, J. N. 1992b, ApJ, 385, 416

[31] Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical Recipes in Fortran: The Art of Scientific Computing, Sec. Edition (New York: Cambridge Univ. Press)

[32] Rhee, G. 1991, Nature, 350, 211

[33] Roberts, D. H., Greenfield, P. E., Hewitt, J. N., Burke, B. F., & Dupree, A. K. 1985, ApJ, 293, 356

[34] Schild, R. E., & Cholfin, B. 1986, ApJ, 300, 209

[35] Schild, R., & Thomson, D. J. 1995, AJ, 109, 1970

[36] Stockton, A. 1980, ApJ, 242, L141

[37] Vanderriest, C., Schneider, J., Herpe, G., Chevreton, M., Moles, M., & Wlérick, G. 1989, A&A, 215, 1

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

[39] Young, P., Gunn, J. E., Kristian, J., Oke, J. B., & Westphal, J. A. 1981, ApJ, 244, 736