COSMOLOGICAL APPLICATIONS OF GRAVITATIONAL LENSING

R. D. Blandford

Theoretical Astrophysics, 130-33 Caltech, Pasadena, California 91125

R. Narayan

Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, Massachusetts 02138

KEY WORDS: microlensing, arcs, quasars, galaxies, clusters, cosmography, Hubble constant, dark matter

1. INTRODUCTION

1.1 History

The discovery by Walsh et al (1979) of the first bona fide gravitational lens, the doubly-imaged quasar, Q0957 + 561, happened at an opportune time, following several prescient theoretical papers, and just preceding the completion of radio and optical quasar surveys that have since yielded over a dozen examples of this phenomenon. Interest in gravitational lenses stretches back over more than seventy years (Eddington 1919, Lodge 1919). Zwicky (1937a,b) appears to have been the first to realize that gravitational lensing ought to have a major impact on cosmology, specifically by “weighing” nebulae and providing crude telescopes to magnify lensed sources. The discovery of quasi-stellar “point” sources added two more possible uses of lenses, for distance measurement (Klimov 1963, Liebes 1964, Refsdal 1964b) and as probes of the stellar composition of lenses (Chang & Refsdal 1979), both of which may be just coming to fruition. These four topics constitute the primary theme of this review.

The observational challenge of gravitational lensing is great and, in fact, many of the difficulties had already been recognized prior to 1979. (a) Multiple-imaging due to lensing is a comparatively rare phenomenon,
affecting only a fraction of a percent of distant sources (Press & Gunn 1973). (b) Large magnification can disguise the nature of the source (Bar- nothy 1966). (c) The lens mass distribution is uncertain and this is reflected in the lensing properties (Bourassa & Kantowski 1976). (d) Regions of different size in the source can be magnified to differing degrees (Sanitt 1971). (e) Perturbations due to inhomogeneities along the line of sight can introduce additional distortions (Zeldovich 1964; Gunn 1967a,b). Despite these many problems, the past decade has seen gravitational lenses produce pleasing corroboration of many basic ideas in extragalactic astronomy and geometrical optics. We expect that the next decade will see lenses playing a more central role in cosmological research.

1.2 This Review

The field of gravitational lensing has expanded so much in recent years that it is not practical to encompass it entirely in a single Annual Review article. We therefore concentrate only on those aspects of the subject that promise to impact cosmology, dealing successively with cosmography (Section 4), the inferences that can be drawn about the perturbing masses (Section 5), deductions that can be made concerning the source structure (Section 6), and future prospects for enlarging the sample of known lenses (Section 7). For completeness, however, we begin with abbreviated discussions of two basic topics: the observational status of individual gravitational lenses (Section 2), and the theoretical concepts used to understand the optics of lensing (Section 3). We give only limited bibliography for these two sections and refer the reader to more detailed reviews by Blandford & Kochanek (1987a), Canizares (1987), Refsdal & Kayser (1988), Fort (1990), Surdej (1990), Narayan & Wallington (1992), the recent conference proceedings edited by Moran et al (1989), Mellier et al (1990), and Kayser & Schramm (1992), and, in particular, the excellent monograph of Schneider et al (1992) (cf also Bloikh & Minakov 1989).

2. SECURE LENSES

To date, three distinct classes of multiple imaging have been identified: multiple quasars (i.e. double, triple, and quadruple images of a single quasar), arcs (images of high redshift galaxies formed by clusters), and radio rings (images of extended radio sources formed by intervening galaxies). Unfortunately, the criteria for class membership remain subjective and no list commands universal acceptance. [For instance, a problem with identifying multiply-imaged quasars is that distinct objects can have quite similar spectra (Phinney & Blandford 1986, Bahcall et al 1986, Shaver 1986); conversely, small spectral differences can be attributed to time
variability (Filippenko 1989, Steidel & Sargent 1991). Adopting the most conservative of conditions, we list in Table 1 the “secure” cases in each of the three classes of lensing.

2.1 Multiple Quasars

The most famous, and also still the best studied gravitational lens, is the “double quasar,” Q0957 + 561 (Walsh et al 1979). Two optical and radio images of a $z = 1.41$ quasar, separated by $\sim 6''$, are formed on opposite sides of a brightest cluster galaxy with redshift $z = 0.36$ (Figure 1). Detailed lens modeling can account for the observed images (e.g. Falco et al 1991a).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Secure cases of multiple-imaging gravitational lenses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>$z_s$</td>
</tr>
<tr>
<td>---------</td>
<td>--------</td>
</tr>
<tr>
<td><strong>Multiple Quasars</strong></td>
<td></td>
</tr>
<tr>
<td>Q0957 + 561AB</td>
<td>1.41</td>
</tr>
<tr>
<td>Q0142 - 100AB</td>
<td>2.72</td>
</tr>
<tr>
<td>Q2016 + 112ABC</td>
<td>3.27</td>
</tr>
<tr>
<td>Q0414 + 053ABCD</td>
<td>2.63</td>
</tr>
<tr>
<td>Q1115 + 080A,A2BC</td>
<td>1.72</td>
</tr>
<tr>
<td>H1413 + 117ABCD</td>
<td>2.55</td>
</tr>
<tr>
<td>Q2237 + 031ABCD</td>
<td>1.69</td>
</tr>
<tr>
<td><strong>Arcs</strong></td>
<td></td>
</tr>
<tr>
<td>Abell 370</td>
<td>0.72</td>
</tr>
<tr>
<td>Abell 963</td>
<td>0.77</td>
</tr>
<tr>
<td>Abell 1352</td>
<td>?</td>
</tr>
<tr>
<td>Abell 1525</td>
<td>?</td>
</tr>
<tr>
<td>Abell 1689</td>
<td>?</td>
</tr>
<tr>
<td>Abell 2163</td>
<td>?</td>
</tr>
<tr>
<td>Abell 2218</td>
<td>?</td>
</tr>
<tr>
<td>Abell 2390</td>
<td>0.92</td>
</tr>
<tr>
<td>Cl 0024 + 17</td>
<td>?</td>
</tr>
<tr>
<td>Cl 0302 + 17</td>
<td>?</td>
</tr>
<tr>
<td>Cl 0500 - 24</td>
<td>0.91</td>
</tr>
<tr>
<td>Cl 1409 + 52</td>
<td>?</td>
</tr>
<tr>
<td>Cl 2244 - 02</td>
<td>2.23</td>
</tr>
<tr>
<td><strong>Radio Rings</strong></td>
<td></td>
</tr>
<tr>
<td>MG1131 + 0456</td>
<td>?</td>
</tr>
<tr>
<td>0218 + 357</td>
<td>?</td>
</tr>
<tr>
<td>MG1549 + 3047</td>
<td>?</td>
</tr>
<tr>
<td>MG1634 + 1346</td>
<td>1.75</td>
</tr>
<tr>
<td>1830 - 211</td>
<td>?</td>
</tr>
</tbody>
</table>
and a 1.48 yr time lag has been measured in the variation of the two images (Lehar et al 1992a; Press et al 1992a,b). There is one other secure double quasar—Q0142−100 (Surdej et al 1987), and five candidates—Q1120+019 (Meylan & Djorgovski 1989), Q1208+101 (Maoz et al 1992, Magain et al 1992), Q1429−008 (Hewett et al 1989), Q1635+267 (Djorgovski & Spinrad 1984, Turner et al 1988), and Q2345+007 (Weedman et al 1982, Tyson et al 1986), where the evidence for multiple imaging is not conclusive and where any lensing galaxies, if present, must be highly subluminous. The radio galaxy 3C324 may also be doubly imaged (Le Fevre et al 1987). The source Q2016+112 (Lawrence et al 1984, Heffin et al 1991) is compellingly argued on the basis of spectroscopic and morphological data.
to be triply-imaged by two observed galaxies, but the image geometry is not well understood (however, see Narasimha et al 1987). There are four convincing examples of quadruple imaging: 0414+053 [which, although a radio source may not be a quasar (Hewitt et al 1989; R. Elston & C. R. Lawrence, personal communication)], Q1115+080 [the “triple quasar” (Weymann et al 1980, Young et al 1981a)], H1413+117 [the “clover leaf” (Magain et al 1988)], and Q2237+031 [the “Einstein cross” (Huchra et al 1985); see Figure 2]. Despite the absence of a detected lensing galaxy in 0414+053 and Q1413+117, all four examples are considered secure because the image arrangement is a pattern quite characteristic of a gravitational lens (see Figure 6).

2.2 Arcs

A rather different type of lens was first identified independently by Lynds & Petrosian (1986) and Soucail et al (1987) who discovered a blue, luminous arc ~20" long in the rich cluster of galaxies, Abell 370 (Figure 3). Similar long features have been reported in Abell 963 (Lavery & Henry 1988, Ellis et al 1991), Abell 1352, 1525, 1689, 2163, 2218 (J. A. Tyson, personal communication), Abell 2390 (Pello et al 1991), Cl 0024+17 (Koo 1988), Cl 0302+17 (Fort 1992), Cl 0500−24 (Giraud 1988), Cl 1409+52 (Tyson et al 1990), and Cl 2244−02 (Lynds & Petrosian 1986) (cf Table 1). Detailed theoretical models (Hammer 1987, Kovner 1988, Grossman & Narayan 1988, Lynds & Petrosian 1989) showed that arcs would be formed quite naturally by a deep cluster potential well producing a large

Figure 2  The “Einstein cross” gravitational lens Q2237+031. (Left) Spiral galaxy lens with redshift z = 0.039. (Right) Four quasar images surrounding the bright central bulge of the lens obtained using the Faint Object Camera on Hubble Space Telescope. (Images supplied courtesy C. Blades, D. Macchetto, and T. Tyson.)
Figure 3 Giant arc formed by the rich cluster of galaxies Abell 370. The source is a blue galaxy ($z = 0.72$) located close to a caustic. Also visible are many smaller arclets, elongated tangentially with respect to the cluster center. (Image supplied courtesy B. Fort and G. Soucail.)

tangential magnification in the image of a background galaxy (Paczyński 1987a).

The majority of the above clusters, plus about six other clusters in which long arcs have not yet been reported, exhibit numerous (up to 60) smaller scale arclets (Tyson et al 1990, Fort 1992, Smail et al 1991, Soucail 1991). These are again background galaxies, but they are much less magnified than the major arcs and are mostly singly imaged.

2.3 Radio Rings

The first radio ring, MG 1131 +0456, was discovered by Hewitt et al (1988) (Figure 4) and is formed when a galaxy lies in front of a compact radio source and distorts the image of the latter into the form of an “Einstein ring.” Four more radio rings have been discovered: 0218 +357 (Patnaik et al 1992), MG1549 +3047 (Lehar et al 1992b), MG1634 +1346 (Lang-
3. GRAVITATIONAL LENSES OPTICS

Although basic geometrical optics is all that is required to describe gravitational lenses, the development of the subject has generated some fresh physical insights. Several complementary formalisms have been derived (cf Nityananda 1990a,b), but we shall confine our attention to those that are most appropriate for application to cosmology. Schneider et al (1992) provide a more complete treatment of the subject.

3.1 Ray Deflections and the Lens Equation

The basic ray geometry is shown in Figure 5. For the moment, consider a homogeneous Friedmann-Robertson-Walker (FRW) universe of cosmological density parameter $\Omega_0$ in which angular diameter distances $D(z_i, z_f)$
Figure 5 Basic ray geometry of gravitational lensing. A light ray from a source S at redshift $z_s$ is incident on a deflector or lens L at redshift $z_d$ with impact parameter $\xi$ relative to some fiducial lens "center." Assuming the lens is thin compared to the total path length, its influence can be described by a deflection angle $\delta(\xi)$ (a two-vector) suffered by the ray on crossing the "lens plane." The deflected ray reaches the observer O, who sees the image of the source apparently at position $\theta$ on the sky. The true direction of the source, i.e. its position on the sky in the absence of the lens, is indicated by $\beta$. Also shown are the angular diameter distances $D_d, D_s, D_{ds}$, separating the source, deflector, and observer.

Proper lengths $\xi$ located at redshift $z_j$ to angles $\theta_i$ subtended by these lengths when observed from redshift $z_i < z_j$:

$$D(z_i, z_j) = \frac{\xi_j}{\theta_i} = \frac{2c}{H_0} \frac{(1 - \Omega_0 - G_i G_j) (G_i - G_j)}{\Omega_0^2 (1 + z_i) (1 + z_j)^2}, \quad G_{i,j} = (1 + \Omega_0 z_{i,j})^{1/2}, \quad 1.$$ 

where $H_0 = 100h$ km s$^{-1}$ Mpc$^{-1}$ is the Hubble constant. For a source at redshift $z_s$ and a single deflector at redshift $z_d$, we follow convention and set $D_d = D(0, z_d)$, $D_s = D(0, z_s)$, and $D_{ds} = D(z_d, z_s)$. Elementary geometry gives the lens equation (e.g.Refsdal 1964a), which connects the source position $\beta$ and the image position $\theta$ (Figure 5),

$$\beta = \theta - \alpha(\theta), \quad 2.$$ 

where the reduced deflection angle $\alpha(\theta)$ and the true deflection angle $\delta(\xi)$ are related by

$$\alpha(\theta) = D_d \delta(D_d \theta)/D_s.$$ 

Deflection angles of interest in astrophysical applications are always small. We are therefore justified in using the weak field approximation of General Relativity. For a thin lens, $\delta = \nabla \psi/c^2$, where $\psi(\xi)$ is twice the 2D Newtonian potential obtained by solving the 2D Poisson equation, $\nabla^2 \psi(\xi) = 8\pi G \Sigma(\xi)$, corresponding to surface mass density $\Sigma(\xi)$. [A weak gravitational lens is, in its essentials, equivalent to a flat space deflector]
with refractive index \(1 - 2\phi/c^2\), where \(\phi\) is the 3D Newtonian potential with respect to infinity (Eddington 1919).]

From the structure of Equation 2 it is clear that, for a given lens, there is a unique source position \(\beta\) for each image position \(\theta\). However, the converse is not true. For a nontrivial deflection law \(\alpha(\theta)\), it is possible to find more than one solution \(\theta\) satisfying Equation 2 for a given \(\beta\), thus giving rise to multiple imaging. (In general, gravitational lenses are not like simple optical lenses which have \(\alpha(\xi)\) varying linearly with \(\xi\).)

The magnification tensor of the \(i\)th image relative to the unlensed source is

\[
[\mu^{(i)}] = \frac{\partial\theta^{(i)}}{\partial \beta} = \left[ J - \frac{\partial \alpha}{\partial \theta^{(i)}} \right]^{-1},
\]

which is a symmetric \(2 \times 2\) matrix. This is not directly measurable, but the relative magnification between two images, described by the transformation matrix \([\mu^{(i)}] = [\mu^{(j)}][\mu^{(i)}]^{-1}\) can be estimated when the images are resolved. This is in general nonsymmetric and provides four independent observables. The flux magnification associated with the \(i\)th image is the absolute value of the determinant of the magnification tensor, i.e. \(\mu^{(i)} = |[\mu^{(i)}]|\). For unresolved images, only the relative flux magnifications, \(\mu^{ij} = |[\mu^{(i)}]|\), are measurable.

### 3.2 Simple Lens Models

The deflection angle for impact parameter \(\xi\) relative to a point mass \(M\) is given by \(\alpha(\xi) = 4GM\xi/c^2\xi^2\). A source on the optic axis will form an Einstein ring (Chwolson 1924, Einstein 1936) of angular radius

\[
\theta_E = \left( \frac{4GM}{c^2D} \right)^{1/2} = 3 \left( \frac{M}{M_\odot} \right)^{1/2} \left( \frac{D}{1\text{Gpc}} \right)^{-1/2} \mu\text{arcsec}, \quad D \equiv \frac{D_d D_s}{D_{ds}},
\]

where we have introduced an effective lens distance \(D\). An off-axis source will produce two images on opposite sides of the lens—a magnified image outside the Einstein ring, and one inside with magnification diminishing as \((\theta_E/\beta)^4\) for \(\beta \gg \theta_E\). For source positions \(\beta \lessgtr \theta_E\) the two images have roughly comparable magnifications. Therefore, the effective cross section for lensing is usually taken to be \(\pi\theta_E^2\). Using this estimate of the cross section, the optical depth to lensing for a point source at high redshift is of order the fractional mass density \(\Omega_0\) of the universe in point mass lenses (Press & Gunn 1973). For extended sources of angular size \(\lesssim \theta_E\), this is also the optical depth corresponding to significant image distortion due to lensing.

The point mass lens is an appropriate model for computing deflections.
by individual stars and black holes. However, even when the lens is not point-like, one can still use Equation 4 for rough estimates. For instance, one could take the characteristic angular separation of the images (say in a multiply-imaged quasar) to be a measure of $2\theta_E$ and thus estimate the mass $M$ of the lens “enclosed” by the images. Further, for non-point-like lenses, the optical depth to lensing for a high redshift source is roughly given by the density parameter $\Omega_E$ corresponding to the fractional lens mass enclosed within the respective Einstein rings of the lenses.

In the opposite limit, when the mass distribution has a length scale that is much larger than the size of the image region, one can make a multipole expansion of the potential $\psi$ to quadratic order and rotate the coordinate system so that

$$(\alpha_1, \alpha_2) = \kappa (\theta_1) + \gamma (\theta_1, -\theta_2),$$

where the subscripts 1 and 2 refer to components along two orthogonal axes. The associated magnification matrix of such a quadratic lens is

$$[\mu] = \begin{pmatrix} (1-\kappa-\gamma)^{-1} & 0 \\ 0 & (1-\kappa+\gamma)^{-1} \end{pmatrix}.$$  

The parameter $\kappa$ is the convergence, which measures the isotropic part of the magnification:

$$\kappa = \frac{\Sigma}{\Sigma_{cr}}, \quad \Sigma_{cr} = \frac{c^2}{4\pi G D'} = 0.35 \left( \frac{D'}{1\text{Gpc}} \right)^{-1} \text{g cm}^{-2}, \quad D' = \frac{D_{ds}D_d}{D_s},$$

where $\Sigma_{cr}$ is known as the critical density and $D' = D(D_{ds}/D_s)^2$ is a second effective lens distance. Normally a lens will have $\Sigma > \Sigma_{cr}$ somewhere in order to produce multiple images, but this is not strictly necessary (Subramanian & Cowling 1986). The parameter $\gamma$ is known as the shear and measures the anisotropic stretching of the image. It arises from matter lying outside the beam. Quadratic lenses are used to describe background galaxies surrounding individual stars and sometimes also for galaxy clusters around galaxies.

The intermediate case, when the image separations are comparable with the size of the potential, is harder to treat. For galaxies and clusters, simple spherical models have been often used (e.g. Clark 1972, Gott & Gunn 1974, Young et al 1980, Yakovlev et al 1983). These models can be parametrized by the 1D velocity dispersion $\sigma = 300\sigma_{300}$ km s$^{-1}$. For a singular isothermal sphere, the deflection angle is constant and given by $\alpha = 4\pi\sigma^2/c^2 = 2.6''\sigma_{300}^2$, so that
\[ \theta_E = \frac{4\pi \sigma^2 D_{ds}}{c^2 D_s} = 2.6' \sigma^2_{100} \frac{D_{ds}}{D_s}. \]

For source positions \( \beta < \theta_E \), there are two images at \( \theta = \beta \pm \theta_E \). Technically, there is also a third image at \( \theta = 0 \), but this has zero magnification because the surface density is singular at the center of the lens. The cross section for multiple-imaging is given by \( \pi \theta_E^2 \).

Two refinements are often introduced when modeling galaxy and cluster lenses. (Bourassa et al. 1973; Bourassa & Kantowski 1975, 1976; Sanitt 1976; Bray 1984; Blandford & Kochanek 1987b; Kovner 1987a). (a) The singularity at the center is removed by introducing a finite core radius \( \xi_c \) such that the deflection angle is diminished for \( \xi < \xi_c \) and vanishes for \( \xi = 0 \). The surface density \( \Sigma_c \) at the center of the lens now is finite, and multiple imaging is possible only if \( \Sigma_c \) exceeds the (distance-dependent) critical density \( \Sigma_{cr} \). When there is multiple imaging, in addition to two images on either side of the lens as in the singular isothermal lens, there is also a third image in the core region which for small \( \xi_c \) is usually weak in comparison to the other two. (b) It is often essential to break the circular symmetry of the lens. This can be done by introducing a quadrupolar component in either the mass distribution or the potential of the lens, or by adding an external shear to a circularly symmetric lens. We refer to such models as elliptical lenses. Figures 6 and 7 show some image configurations that arise with elliptical lenses in front of point and extended sources.

### 3.3 Caustics and Congruences

The angular diameter distance quoted in Equation 1 is only useful in a homogeneous universe. Inhomogeneous universes necessitate drastic approximations. Most treatments of this problem assume that the universe is uniform on the scale of the Hubble radius and that the relationship between redshift and cosmological time is the same as that in a FRW universe of similar mean density.

We are interested in the evolution of the cross section of a null geodesic congruence (a bundle of rays) as it propagates backward in time away from the observer, specifically in the mapping from the observer's angle space to proper distance perpendicular to some central or fiducial ray. Let us generalize Figure 5 and set the vector, \( \xi(\lambda) = (\xi_1, \xi_2) \), equal to this offset as measured all along the bundle, not just in the lens plane. The subscripts 1 and 2 represent components with respect to an orthonormal basis parallel propagated along the fiducial ray, and \( \lambda \) is the affine parameter. Next, define the complex number \( \xi(\lambda) = \xi_1 + i\xi_2 \). The angle between a ray in the congruence and the fiducial ray at the observer can also be represented as
a complex number $\xi_0$, where a dot denotes differentiation with respect to $\lambda$. We are now in a position to generalize the notion of angular diameter distance by defining a two component vector, $\mathcal{D} = (\mathcal{D}_1, \mathcal{D}_2)$, where both $\mathcal{D}_1$ and $\mathcal{D}_2$ are complex, using the general linear relation

$$\xi = \mathcal{D}_1 \xi_0 + \mathcal{D}_2 \dot{\xi}_0^\ast.$$  

The real part of $\mathcal{D}_1$ measures the expansion of the ray while its imaginary
part describes pure rotation. (In practice, rotation is usually small and $D_1$ is approximately real.) $D_2$ measures the shear. All the information about the local image distortion is contained in $D$. The conventional angular diameter distance, whose square is the ratio of the source area to the solid angle it subtends, is given by $|D_1|^2 - |D_2|^2$, and suffices for point sources where only the flux can be measured.

In an inhomogeneous universe containing Newtonian matter, $D$ can be shown to evolve according to

$$D = \begin{pmatrix} \mathcal{R} & \mathcal{F}^* \\ \mathcal{F} & \mathcal{R} \end{pmatrix} D,$$

where the quantity $\mathcal{R} = -(1+z)^2(\phi_{,11} + \phi_{,22}) = -4\pi(1+z)^2G \rho$ describes focusing by matter lying within the congruence with proper density $\rho$, and $\mathcal{F} = -(1+z)^2(\phi_{,11} - \phi_{,22} + 2i\phi_{,12})$ describes the influence of matter external to the congruence (e.g. Penrose 1966, Blandford et al 1991). This formalism immediately gives expressions for the magnification tensors, $[\mu]$ (cf Equation 3), whose definition we can now generalize by identifying $\beta$ with the angle which would be subtended by the proper length $\xi$ in the
source plane in a FRW universe of similar average density to the inhomogeneous universe under consideration. (See Ehlers & Schneider 1986 for an alternative choice of reference universe.)

In the limiting case when all the matter in the universe apart from the lens is isolated from the congruence ($\mathcal{V}_2 = 0$), the lack of focusing by matter in the beam (save for the lens) compared to a FRW universe of the same $\Omega_0$ increases the angular diameter distance of the source (Dashevskii & Zel'dovich 1965, Dyer & Roeder 1972, Nottale 1983, Nottale & Hammer 1984, Kasai et al. 1990). The increase is about 30% for a source with $z_s \sim 2$ in an Einstein–De Sitter universe. However, the cumulative shear caused by external matter usually produces a second-order focusing which leads to a diminished net effect. In general, if multiple imaging is uncommon, the distribution of magnifications due to smoothly distributed matter is dominated by the convergence rather than shear (Lee & Paczyński 1990, Watanabe & Sasaki 1990). The total flux is always conserved when suitably averaged over all directions (Weinberg 1976, Peacock 1986).

If the focusing (say due to a lensing galaxy) is strong enough to make the rays cross along any congruence (Figure 8), then multiple images must form and we have an example of gravitational lensing. At the point where the rays cross, known as a conjugate point to the observer, the conventional angular diameter distance vanishes ($|\mathcal{D}_1| = |\mathcal{D}_2|$) and the formal magnification diverges. The locus of these conjugate points is a two-dimensional surface, a caustic sheet, to which the rays are tangent (see Blandford

![Figure 8](image)

*Figure 8* An infinitesimal conical bundle of rays is shown drawn backwards from an observer, past an elliptical lens, and touching two caustic sheets. The second caustic sheet, on the left, has a cusp line perpendicular to the plane of the diagram, while the first caustic sheet has a cusp in the plane. Representative cross sections of the bundle are indicated at the bottom. Where the bundle touches a caustic, its cross section degenerates to a straight line. Beyond this point, the bundle is "inverted," and a source located here will acquire two additional images. In general there could be many caustic sheets behind a complex lens, but with a single elliptical lens there are only two sheets (which may penetrate each other, cf Blandford & Narayan 1986).
& Narayan 1986 for a schematic diagram showing the caustic sheets associated with an elliptical lens). Equivalently, we can think in terms of wavefronts normal to the rays merging at a caustic (Kayser & Refsdal 1983). For a source at a fixed redshift, the source plane intersects the caustic sheets at caustic lines. The images of these lines are known as critical curves (cf Figures 6, 7).

In the generic situation, the caustic sheet corresponds to a fold caustic. When a source crosses a fold, an extra pair of images will either be created or destroyed. These image pairs will be stretched toward each other along a direction essentially perpendicular to the projection of the caustic on the sky (Blandford & Kovner 1988). Because of the stretching, the images will be bright; an example is the pair of bright images, $A_1A_2$, in Q1115+080. The magnifications of the two images will be inversely proportional to their separations and also inversely proportional to the square root of the distance of the source from the caustic (Benson & Cooke 1979, Ohanian 1983, Blandford & Narayan 1986, Kayser & Witt 1989). Therefore, for a fold caustic, the cross section, $\sigma(>\mu)$, for the magnification to be greater than $\mu$ has a universal scaling, $\sigma \propto \mu^{-2}$, for $\mu \gg 1$. Equivalently, the differential cross section scales as $d\sigma/d\mu \propto \mu^{-3}$.

Every time a ray touches a caustic (grazes it tangentially), the associated image is inverted, i.e. its parity is reversed. (Polarization directions are parallel propagated and unaffected.) In Q0957+561, the A image is believed to be uninverted while the ray associated with the B image has touched one caustic, so the two images are approximate reflections of each other. A faint third image ought to be formed in the galaxy nucleus, inverted twice through roughly orthogonal planes, hence rotated through $\sim 180^\circ$.

Fold surfaces meet at cusp lines, which correspond to a cusp caustic. Sources lying just inside cusps create three bright images (plus any additional images that are not associated with the cusp). Sources lying just outside cusps have one of their images highly brightened. In this region, the cross section for large $\mu$ scales as $\sigma(>\mu) \propto \mu^{-5/2}$, or $d\sigma/d\mu \propto \mu^{-7/2}$. Cusps are believed to play an important role in the luminous arcs. Cusp lines meet at points associated with higher order singularities, but these have not yet been identified in the observations. The closest point of the caustic to the observer is generically a cusp. When a source is located close to this point, the lens is said to be marginal and may produce one or three bright images (Narayan et al 1984, Kovner 1987d,e).

In general, for a nonsingular lens, caustic surfaces separate regions with image multiplicities differing by two. Since far from the lens a source has but one image, therefore the total number of images has to be odd (Burke 1981, McKenzie 1985).
3.4 Time Delay

The propagation time from the source to the observer varies from one image to another, and this difference can be measured when the source is variable. For a single lens in a homogeneous universe, the excess time delay associated with the image at \( \theta^{(i)} \), relative to the direct ray in the absence of the lens, is given by (Refsdal 1964b, Cooke & Kantowski 1975, Kayser & Refsdal 1983, Borgeest 1983, Schneider 1985)

\[
t^{(i)} = (1 + z_d) \left[ \frac{D}{2c} (\theta^{(i)} - \beta)^2 - \frac{\psi(\theta^{(i)})}{c^2} \right].
\]

The quantity \( t^{(i)} \) itself cannot be measured, but the relative time delay between two images, \( t^{(ij)} = t^{(i)} - t^{(j)} \), can be. Since \( D \) (cf Equation 4) is inversely proportional to \( H_0 \), a measurement of \( t^{(ij)} \) provides vital information for cosmography (cf Section 4.1). The term proportional to \((\theta^{(i)} - \beta)^2 \) in Equation 11 is due to the geometrical excess path length and the term proportional to \( \psi \) is the gravitational time delay, known as the “Shapiro effect” in the solar system (Shapiro 1964).

In one formulation of gravitational lens theory, the time delay \( t(\theta) \) is taken to be the fundamental quantity and the lens equation (2) is obtained by invoking Fermat’s principle, which states that images are formed at stationary points of \( t \) (Weyl 1922, Schneider 1985, Blandford & Narayan 1986). Equation 11 can be generalized to multiple lenses (Blandford & Narayan 1986, Kovner 1987b), as well as to nonstationary gravitational fields (Kovner 1990, Nityananda & Samuel 1992).

3.5 Microlensing

For a sufficiently compact source, individual stars in a lensing galaxy can modify the magnification relative to that expected from a smoothly-distributed lens (Chang & Refsdal 1979, Gott 1981, Young 1981, Chang 1984, Vietri & Ostriker 1983, Nityananda & Ostriker 1984, Paczyński 1986a). This phenomenon is termed microlensing, in contrast to the effect of the smooth mass distribution which is referred to as macrolensing.

A convenient measure of the influence of an isolated star is its Einstein radius (Equation 4). The optical depth for a source to lie within an Einstein ring is given by \( \tau = \Sigma^*/\Sigma_{cr} \) where \( \Sigma^* \) is the total stellar surface density. This is independent of the masses of the individual stars. When \( \tau \ll 1 \), the faint microimages associated with individual stars are unimportant except on the rare occasions when a star crosses within a few \( \theta_E \) of the line of sight. When this happens the source will brighten and fade on a timescale given by
\[ t_{\text{var}} \sim \frac{D_0 \theta_E}{V} = 15 \left( \frac{M}{M_\odot} \right)^{1/2} \left( \frac{D'}{1 \text{Gpc}} \right)^{1/2} \left( \frac{V}{10^3 \text{ km s}^{-1}} \right)^{-1} \text{ yr}, \]

where \( V \) is the velocity of the star relative to the source-earth line and \( D' \) is given by Equation 7. In contrast to the optical depth, \( t_{\text{var}} \) does depend on the mass of the microlens (Kayser et al 1986, Kayser & Refsdal 1989, Kayser 1992).

As \( \tau \) increases, the microlenses can no longer be considered in isolation. At moderate \( \tau \), the combined, long range action (through shear) of the background stars can be thought of as creating a quadrupole lens at the site of each star (Chang & Refsdal 1984, Nityananda & Ostriker 1984, Lee & Spergel 1990). This can break the circular symmetry and allow slender caustic surfaces to form behind individual stars, leading to extra image pairs. At still higher \( \tau \), when the Einstein rings start to overlap, a complicated caustic network will develop (Schneider & Weiss 1986, Kayser et al 1989, Wambsganss 1990, Witt 1990). The background shear from the large-scale mass distribution of the lens makes the network anisotropic and the variability will be sensitive to the direction of transverse motion relative to the shear (Wambsganss 1990, Wambsganss et al 1990a). Frequent, large amplitude image variation is possible in this regime (Paczyński 1986a, Nemiroff 1986, Schneider & Weiss 1987, Witt 1990). If there is a smooth, supercritical background density, then dramatic demagnifications are also expected when the ray associated with the brightest microimage intercepts a star. The net effect will be to conserve flux so that the mean magnification is the same as if the mass in the stars had been smoothed out (Peacock 1986). Since fold caustics dominate at the highest magnification, the asymptotic probability for large magnification by more than \( \mu \) scales as \( \mu^{-2} \) (Section 3.3, Vietri & Ostriker 1983, Nityananda & Ostriker 1984, Blandford & Narayan 1986), with a normalization that can be computed exactly (Schneider 1987c). Numerical simulations show a significant excess in the cross section (over and above the analytical normalization of the \( \mu^{-2} \) law) at moderate magnifications (Rauch et al 1991).

When a caustic associated with the macroimage is approached, the number of stars contributing to the macroimage as well as the mean magnification increase, rendering numerical simulation impractical (Deguchi & Watson 1987, 1988). For \( \tau \sim 1 \), the number of microimages becomes so large that the fluctuation level in fact diminishes. When \( \tau \gg 1 \), the mean magnification decreases \( \propto \tau^{-2} \) again, but the relative fluctuations, for fixed source size, are found to increase (Deguchi & Watson 1988). In this limit the angular profile of the macroimage consists of a Gaussian core and a power law tail \( \propto \theta^{-4} \) (Katz et al 1986).
Variability associated with microlensing is strongly attenuated if the angular size of the source becomes comparable to the Einstein radius of the microlenses (e.g. Wambsganss 1991, Refsdal & Stabell 1991). This can be used to set limits on the sizes of emission regions of distant sources (cf Section 6.1).

Irwin et al (1989) and Corrigan et al (1991) have reported variability in the images of Q2237 + 031. Since the optical depth for microlensing in this case is \( \tau \sim 0.5 \), the variations have been plausibly interpreted as being the result of microlensing. Schild & Smith (1991) appear to have measured variability due to microlensing in Q0957 + 561.

### 3.6 Observables and Model Fitting

#### 3.6.1 Compact Sources

It is straightforward to trace rays through gravitational lens models to determine the image locations, magnification tensors, and time delays corresponding to a given source position. It is less easy to infer the lensing potential from the observations. The lens equation (2) provides one vector equation (two components) for each of the \( n \) images. Since the true position of the source is not known, this leads to \( 2(n - 1) \) constraints that may be applied to the lensing potential. The relative magnifications of the images provide \( 4(n - 1) \) constraints when the images are resolved and the transformation matrices can be determined, or \( (n - 1) \) in the more usual unresolved case when only flux ratios are available. Finally, time delays may provide additional constraints.

When there are galaxies in the field it is common to assume that they can be modeled with King, De Vaucouleurs, or similar profiles with a few adjustable parameters, or with standard mass-to-light ratios where photometry of the lens is available [as in Q2237 + 031 (Schneider et al 1988)]. The disposition of dark matter in the outer halos, in clusters and presumably also in groups of galaxies is more problematical (Young et al 1981b). This is usually modeled as either an isothermal sphere (with a finite core) or a quadratic lens. The parameters in the lens model are determined by using the constraints from the data, e.g. the image locations, magnifications, and time delays.

An additional consideration is usually included subjectively: The lensing geometry should not be too improbable. However, this has to be evaluated with considerable care as there are selection effects whereby seemingly unlikely alignments may be highly favorable for discovery.

The most important selection effect is magnification bias (sometimes also referred to as amplification bias) (Turner 1980, Turner et al 1984), which occurs for images that are magnified as a consequence of lensing. Although high magnification configurations may have a small a priori cross section, they may actually dominate flux-limited samples if faint
sources are sufficiently more numerous than bright sources (cf Sections 5.2.2, 6.5). Similar biases are possible with marginal lenses (Kovner 1987d).

In general, the largest lensing cross sections are associated with arrangements in which there are two images located on either side of the galaxy as in Q0957+561 and those where there are four images lying roughly on the Einstein ring as in Q2237+031 and Q1115+080 (Blandford & Kochanek 1987b, Pojmanski & Szymanski 1987, Nemiroff 1989). If the galaxy is part of a cluster, the image separation will be enlarged and the magnification will be enhanced, which may give these cases disproportionate importance because of magnification bias (Turner et al 1984). One complication that is probably relevant to Q2016+112 is that strong lenses may be localized at more than one redshift. This opens up a much larger parameter space (Nottale & Chauvineau 1986, Kochanek & Apostolakis 1988, Jaroszyński 1989).

Simple lens models based on elliptical potentials suffice to account for most of the multiply-imaged quasars (e.g. Narayan & Grossman 1989, Blandford et al 1989, Kayser 1990, Kochanek 1991a). More complex potentials, derived from more realistic galactic mass distributions, are used for detailed modeling (e.g. Dyer & Roeder 1981, Narasimha et al 1982, Kayser & Schramm 1988, Schramm 1990, Schneider & Weiss 1991), but the models are not well constrained at present (Kochanek 1991a).

3.6.2 Extended Sources When the source is extended there are effectively as many images to consider as there are resolution elements covered by the source. Consequently, arcs and rings can potentially provide more information about the lensing potential than do multiply-imaged quasars. With a resolved source, different images of the same source element should have identical surface brightness and this can be used as constraints in the modeling. The radio rings are particularly well-suited for such modeling. Typically, some parts of these sources are singly imaged while other parts are multiply-imaged (with either three or five images, of which one in the center may be highly demagnified). Plausible lens and source models can be derived iteratively by matching the observed intensities over the multiple-image regions. This has been done for MG1131+0456 (Kochanek et al 1989), MG1654+1346 (Kochanek 1990a), and 1830−211 (Kochanek & Narayan 1992). The rings could turn out to be the best sources for accurate determination of the lens potential.

4. Cosmography

The most important, and arguably the most difficult application of gravitational lenses is to cosmography—determining the geometry of the uni-
verse on the largest scales. In this section, we consider critically how the Hubble constant has recently been estimated using Q0957 + 561, how the deceleration parameter of FRW models may possibly be measured, and how alternative cosmologies can currently be constrained by observations.

4.1 Hubble Constant

As described in Section 3.4, the difference in travel time $\Delta t$ from the source to the observer along two distinct rays is inversely proportional to the Hubble constant $H_0$, if all the relevant deflection angles are known. A simple geometrical construction explains this result. Consider two spherical wavefronts, one emanating from a point source, the other converging on the observer, touching each other at the lens plane (Figure 9). Images are located at isolated points on the lens plane where the deflection angle $\hat{\alpha}$ equals the angle between the two wave normals. The geometrical part of the time delay is expressible as $(1 + z_d)\xi/2c$, measuring $\xi'$ from the tangent point of the wavefronts. The gravitational time delay is

![Figure 9](image)

The upper panel uses wavefronts to show that the geometrical time delay at the deflector is $\xi/2c$. The bottom panel shows a transformation where (a) a constant density sheet with convergence $\kappa$ is added, (b) the deflection due to the primary lens is reduced to $(1 - \kappa)\hat{\alpha}$, and (c) all linear dimensions are expanded by the factor $(1 - \kappa)^{-1}$. Since there is no net time delay due to the sheet (the geometrical and gravitational delays cancel), all observables, including the time delay, remain unchanged in this transformation. Thus, in order to determine the scale of the universe (i.e. $D_0$ or $H_0$), either the convergence $\kappa$ must be arbitrarily set to zero (Scenario 1), or it must be independently estimated, e.g. by measuring the velocity dispersion of the lens (Scenario 2).
\[-(1 + z_d)\psi(\xi')/c^2,\] where the potential \(\psi\) is proportional to the mass of the deflector, which in turn scales as \(\sigma^2 \xi'\), where \(\sigma\) is a characteristic velocity in the lens. The total time delay is thus proportional to \(\xi'\) and hence inversely proportional to \(H_0\). Normalizing all angles to the image separation \(\Delta \theta\), we have the scaling \(H_0 = K \Delta \theta^2 / \Delta t\), where the constant of proportionality \(K\) depends upon the lens model (Refsdal 1964b; Dyer & Roeder 1980; Kayser & Refsdal 1983; Borgerist 1983; Gaskell 1985; Falco et al 1985; Kayser 1986, 1990; Narayan 1991). Thus, given a reliable time delay measurement and a well-constrained lens model, one can estimate \(H_0\).

In the case of Q0957+561, the time delay appears to have been convincingly measured at both radio and optical wavelengths (Lehar et al 1992a, Press et al 1992a,b, cf also Florentin-Nielsen 1984, Falco et al 1990, Beskin & Oknyanski 1992), giving \(\Delta t = 1.48 \pm 0.03\) yr (though Vanderriest et al 1989 and Schild 1990 obtain \(\Delta t = 1.1\) yr). A widely used model of this system is that developed by Falco et al (1991a), which employs five parameters to characterize the lens; these consist of the one-dimensional velocity dispersion \(\sigma\) of the lens (modified by an additional convergence \(\kappa\) as described below), two “shape” parameters of the lens, namely an angular core radius and a dimensionless compact core mass, and two parameters describing the shear \(\gamma\) due to the other mass in the lens plane such as the surrounding cluster. We comment below on the reliability of this model, but first we discuss a number of scenarios which incorporate various levels of assumptions about the lensing mass and the geometry of the universe. Although the following discussion is focused on Q0957+561 and the Falco et al (1991a) model of this object, most of the arguments will be valid also for other lensed quasars (such as Q1115+080, Narasimha et al 1992) for which time delays may become available in future.

**SCENARIO 1**  
Assume that, except for the scale \(H_0\), the geometry of the universe and in particular the deceleration parameter \(q_0\) and the distance ratio \(D_s/D_ds\) are exactly known. Assume further that the model includes all relevant mass in the lens plane; in particular, assume that there is no dark matter that might contribute an extra convergence \(\kappa\). [The shear \(\gamma\) due to dark matter could be included in the model (cf Kayser 1990) as Falco et al have done for Q0957+561.] Two deductions can then be made. First, the model will uniquely predict the velocity parameter \(\sigma\) of the lens. In Q0957+561, the Falco et al model gives \(\sigma = 390\) km s\(^{-1}\). Secondly, once \(\Delta t\) is measured, \(H_0\) will be uniquely determined. In Q0957+561, assuming \(\Delta t = 1.48\) yr, this gives \(H_0 = 61 \pm 7\) km s\(^{-1}\) Mpc\(^{-1}\) for \(\Omega_0 = 1\) (\(q_0 = 1/2\)). The result is inversely proportional to \(\Delta t\).
SCENARIO 2  Allow now for an unknown amount of convergence $\kappa$ due to smooth dark matter in the lens plane, which in Q0957+561 would be due to the mass associated with the cluster. This is equivalent to adding a quadratic lens with $\alpha \propto \theta$ (Equation 5) and will effectively reduce the curvature of one of the two wavefronts discussed above, say the source-lens wavefront (Figure 9). One then finds that $H_0 = K(1-\kappa)\Delta \theta^2/\Delta t$. This reveals the following fundamental degeneracy in the model: as long as $\kappa$ is undetermined, $H_0$ has no unique solution (Falco et al 1985, Gorenstein et al 1988). There are two possibilities now. First, since $\kappa \geq 0$ (dark matter has positive density), one can still obtain an upper bound on $H_0$ (Borgeest & Refsdal 1984, Falco et al 1985, Kovner 1987c), which for Q0957+561 is the result $61 \pm 7$ km s$^{-1}$ Mpc$^{-1}$ given under Scenario 1. Secondly, it is easy to verify that $\sigma^2/(1-\kappa)$ is a constant. Therefore, the degeneracy may be broken by obtaining $\sigma$ independently. The velocity dispersion of the stars in the lensing galaxy in Q0957+561 has been measured by Rhee (1991) to be $303 \pm 50$ km s$^{-1}$. If the parameter $\sigma$ in the Falco et al model is set equal to this, then one obtains $H_0 \sim 25$--50 km s$^{-1}$ Mpc$^{-1}$ for $\Delta t = 1.48$ yr. However, if the stars are more centrally concentrated than the dark mass in the galaxy, then $\sigma$ could be larger by a factor of up to $(1.5)^{1/2}$ (Turner et al 1984, Kochanek 1991c, Roberts et al 1991), in which case $H_0 \sim 40$--70 km s$^{-1}$ Mpc$^{-1}$.

SCENARIO 3  Give up next any assumption on the geometry of the universe, particularly knowledge of $q_0$ or $D_s/D_d$. Alternatively, let the source redshift $z_s$ be unavailable. Also, allow for large-scale inhomogeneities in the line-of-sight between the lens and the source; in Q0957+561, for instance, there is evidence of a second cluster at redshift 0.5 (Garrett 1992). However, assume that the additional inhomogeneities are quadratic, i.e. that each is completely described by a convergence and a shear (cf Equation 5). In this case, all the additional uncertainties get absorbed into the parameters $\kappa$ and $\gamma$ that have been introduced to describe the dark matter in the lens plane. As a result it can be shown that one is still able to measure the angular diameter distance to the lens (Narayan 1991). For Q0957+561, one obtains $D_d = 1300$--2400 Mpc without any correction factor applied to the measured $\sigma$, and $D_d = 900$--1700 Mpc including the factor of $(1.5)^{1/2}$. These distances can be converted to estimates of $H_0$ if a value of $q_0$ is assumed.

SCENARIO 4  Now include quadratic inhomogeneities between the observer and the lens. The scalar angular diameter distance $D_d$ then needs to be generalized to the complex $\mathcal{D}$ introduced in Section 3.3. Ignoring rotation which factors out, this means one needs three parameters to describe the mapping between angles $\theta$ at the observer and displacements $\xi$ at the lens. Given a sufficiently well-constrained lens (Q0957+561 is probably
inappropriate), one could in principle fit these parameters while at the same time fitting the lens model (Kovner 1987c, Narayan 1991). One can thus solve for $D_d$. Of course, this characterizes only the local line-of-sight to the particular lens (Alcock & Anderson 1985, 1986), and one needs several good lenses to obtain a global estimate of $D_d$ or $H_0$.

**Scenario 5** Suppose there are significant levels of small-scale inhomogeneity in the universe that are not consistent with a quadratic model. Alternatively, suppose the dark matter in the lens plane is not smooth but is lumpy on scales smaller than the image separations. Even in this pessimistic scenario, the lens modeling is still constrained by the requirement that the mass density be positive and by the need not to create additional images in the field. How extreme a Hubble constant can still be tolerated by the observations? The answer depends to some extent on the subjective question of how much latitude one allows oneself in designing an extreme model. However, it would appear that even with considerable freedom one cannot drastically modify the results. For instance, suppose one considered adding a mass perturbation, say a (dwarf) galaxy of mass $M$ and angular size $\theta_c$, at the position of image B in Q0957+561. The time delay of this image will be increased by an amount $\sim 4GM(1+z_d) \ln (\Delta \theta/\theta_c)/c^2 \sim 0.008(M/10^{10} M_\odot) \ln (\Delta \theta/\theta_c)$ yr, the dominant contribution coming from the gravitational component. However, the mass can be bounded above by the requirement that the B VLBI image not be excessively magnified or distorted. This is measured by the ratio of the perturbing galaxy surface density to the critical density, $\sim 0.2(M/10^{10} M_\odot) (\theta_c)^{-2}$. For a perturbing galaxy mass $M$ that does not create extra images or significantly modify the image positions or transformation matrix, only a small change in the relative time delay will be allowed and hence only a modest increase in the true Hubble constant will result. A further encouraging factor is that the relatively undistorted shapes of the long arcs in Abell 370 and other clusters (Section 2.2) argue against significant levels of small-scale mass fluctuations in these cases. If these clusters and lines-of-sight are typical, then small-scale distortions must be unimportant in the majority of gravitational lenses.

From the above discussion, we see that in general a gravitational lens has the potential to provide important cosmographic information. The technique has the virtue of being insensitive to mass fluctuations on large scales because they can be adequately modeled by quadratic lenses, as well as fluctuations on small scales whose effects are likely to be limited. However, the method clearly is sensitive to the details of the mass distribution on the scale of the image separation, i.e. to the particular parameterized form of the lens model employed, and this is its chief weakness. In Q0957+561, there are effectively five constraints from the obser-
vations—two from the relative image position, and three from the relative magnification matrix (technically four constraints, but one of these is poorly determined because the VLBI images are not well-resolved perpendicular to the jet). Since two parameters are used to describe the shear of the cluster, one is left with only three parameters for the main galaxy, of which one is the parameter combination $\sigma^2/(1-\kappa)$. Although the Falco et al model appears to be quite robust, nevertheless, Kochanek (1991c) has found other models that fit the observations equally well. These give estimates of $H_0$ in the range $15\text{--}80\ \text{km s}^{-1}\text{Mpc}^{-1}$ (for $\Delta t = 1.48$ yr). A more comprehensive investigation of allowed models is needed.

Looking to the future, it is clear that lenses that have no obvious cluster surrounding the primary lens are preferable since one has the option of applying Scenario 1. Also, lensed sources having more than two images are likely to be superior to Q0957+561. If the images have resolved VLBI structure and one is able to measure complete transformation matrices, then up to 18 constraints will be available with a quadruply-imaged quasar, plus up to three relative time delays. Potentially even better are the radio rings, where the number of constraints is comparable to the number of resolution elements in the multiply-imaged zones. However, to do useful cosmography, one requires a time-variable core (to measure time delays), and a well-identified lens with a measured redshift and velocity dispersion (though the source redshift is not necessary, see Scenario 3). None of the known quadruply-imaged quasars or radio rings satisfies all these requirements yet.

Another serious uncertainty relates to how one interprets a measured stellar velocity dispersion in the lens. Apart from the factor of $(1.5)^{1/2}$ mentioned above, the velocity distribution may also be anisotropic (e.g. Binney & Tremaine 1987, Foltz et al 1992), which would further complicate the interpretation. Yet another problem is that some of the relative magnifications observed at optical wavelengths may be affected by microlensing (Kayser 1990), leading to errors in the derived lens model. However, this is probably not a consideration for the radio transformation matrix (Falco et al 1991b). Even at radio wavelengths, the core and the jet regions of Q0957+561 are known to have different brightness ratios. This could be interpreted as evidence for gradients in the lens magnification across the source (Conner et al 1992), which will provide additional constraints, enabling the galaxy model to be improved. However, if the discrepancy is due to some kind of microlensing, say due to an intergalactic compact object, then it increases the uncertainties in the model.

4.2 Deceleration Parameter

If a lensing potential is ever specified to better than a few per cent accuracy, it will be possible to infer the distance ratio $D_{as}/D_s$ observationally, and
consequently to solve for $q_0$ (Refsdal 1966a, Lacroix & Schneider 1982). It is also possible, in principle, to measure $q_0$ by comparing the locations and magnifications of arc images of galaxies at different redshifts. In the simple case of a circularly symmetric mass distribution, the arcs will trace the variation of the Einstein radius with source redshift. Unfortunately, for an optimal lens redshift of $z_d \sim 0.6$, the Einstein radius only varies by about 12% as $q_0$ increases from 0 to 0.5. This is much less than the expected uncertainty in the radial variation of the mass distribution in the cluster even if a large number of galaxy velocities are acquired (Tyson 1990). Under extreme conditions, gravitational lensing may complicate alternative determinations of $q_0$ (e.g. Omote & Yoshida 1990).

Undoubtedly, the best test of determinations of $H_0$ and $q_0$ through gravitational lensing will be to obtain similar estimates of these parameters from observations of different objects.

### 4.3 Alternative Cosmologies

Gravitational lenses have certainly affirmed our belief in the redshift distance relation. High redshift sources do appear to lie beyond lower redshift lenses, even when the a posteriori probability of small angular separations occurring is small (Burbidge 1985). Also, the approximate consistency of the Hubble constant derived from Q0957 + 561 with values from more traditional approaches affirms the FRW model and is incompatible with radical alternatives like the chronometric cosmology (Segal 1982) which predicts $\Delta t \lesssim 1$ month (Blandford & Falco 1989, unpublished).

Even in the context of FRW models, some extreme effects are, in principle, possible and certainly worth seeking. If the mean density of the universe were so large (or the cosmological constant, $\Lambda_0$, so negative) that a typical backward propagating congruence had conjugate points or foci (beyond which all images are inverted, cf Section 3.3) at a modest antipodal redshift, the character of the imaging of higher redshift sources would be radically altered (Petrosian & Salpeter 1968, Lebedev & Lebedeva 1985, Gott et al 1989). For instance, intervening galaxy lenses would typically form dim, single images. The existence of multiple images in Q2016 + 112 may then be used to set a lower bound of $z = 3.27$ on the antipodal redshift (Gott et al 1989). This argument is, however, somewhat problematical in the case of this object. There are at least two lensing galaxies (probably at different redshifts) in the line-of-sight to Q2016 + 112 and there is no robust imaging model of the object within the framework of a conventional cosmology (but see Narasimha et al 1987). An antipodal model that invokes two additional unseen galaxies near images A and B may be able to reproduce the observations. The argument of Gott et al is certainly
valid in the case of Q0142–100 which, at a redshift of \( z = 2.72 \), sets a reliable lower bound on the antipodal redshift.

Gravitational lenses can also be used to limit the value of a cosmological constant (Fukugita et al 1990, Turner 1990, Fukugita & Turner 1991, Kochanek 1992). The choice of \( \lambda_0 = \Lambda_0 / 3H_0^2 = 1 - \Omega_0 \), which allows the universe to remain flat and consistent with inflationary cosmological models, is particularly interesting (Peebles 1984). In this case, the optical depth to lensing increases more dramatically with redshift than in an Einstein–De Sitter universe. We defer to the accompanying article by Carroll et al (1992) for a critical discussion of this point.

5. DARK MATTER

Gravitational lensing provides a powerful diagnostic to probe the distribution of matter in the universe, particularly dark matter, over a wide range of scales. In the following sections we discuss some of the possibilities, starting with dark matter in the Milky Way and working outward to the largest scales in the universe.

5.1 Our Galaxy

It is widely recognized that the known populations of stars account for only a fraction of the mass in the halo of the Galaxy (Gilmore et al 1990). It has been suggested that the rest of the mass may consist of one or more varieties of dark compact objects, e.g. "comets," "asteroids," "Jupiters," brown dwarfs, cool white dwarfs, neutron stars, black holes; in the case of black holes, masses of up to \( 10^6 M_\odot \) are allowed (cf Ostriker 1992). All of the above, except comets, can be potentially detected through their gravitational lensing action on background sources (Liebes 1964, Bontz 1979, and especially Paczyński 1986d).

The Einstein radius of a stellar lens of mass \( M \) in the halo is \( \theta_E \approx (M/M_\odot)^{1/2} (D/10 \text{ kpc})^{-1/2} \) mas (cf Equation 4) and the critical surface density is \( \Sigma_c = 3 \times 10^4 (D'/10 \text{ kpc})^{-1} \) g cm\(^{-2}\) (cf Equation 7). Any compact background source that has an angular size \( < \theta_c \) and whose line-of-sight happens to be within \( \sim \theta_E \) of the lens, will be strongly lensed and will appear to vary with a timescale \( t_{\text{var}} \sim 0.2(M/M_\odot)^{1/2} (D'/10 \text{ kpc})^{1/2} (V/200 \text{ km} \text{ s}^{-1})^{-1} \) yr (cf Equation 12). The source will brighten considerably and then fade to its original flux with a symmetric and achromatic light curve that is precisely calculable in the case of an isolated lens and a point source. Therefore, unless either sources or lenses occur predominantly in compact binaries (Griest 1991, Mao & Paczyński 1991), the signature of microlensing events should be distinguishable from intrinsic variability of the
source. However, reliable discrimination will be possible only with a high signal-to-noise ratio and frequent sampling of the light curve.

The mean surface density of the Galaxy at the solar radius (Gilmore et al 1990) is $\lesssim 0.1$ g cm$^{-2} \sim 10^{-6} \Sigma_{cr}$. Therefore, even if the entire mass of the Galaxy is in compact objects, one will still have to observe $\sim 10^6$ sources in order to find one that is microlensed. This could be achieved by monitoring the stars in the LMC or other nearby galaxies, or in the bulge of the Milky Way, on a nightly basis for a few years (Paczyński 1986d, 1991; Griest et al 1991). With sufficiently closely spaced observations (several measurements per night) it may be possible to distinguish microlensing from intrinsic source variability, and even, in principle, to detect signatures of double stars and planets (Mao & Paczyński 1991, Gould & Loeb 1992). Several attempts are underway to put this idea into practice, using photographic methods and CCD arrays at existing telescopes (Milstajn 1990, Vidal-Madjar et al 1991, Paczyński 1992) and with dedicated telescopes and detectors built specifically for this purpose (Alcock et al 1992). However, even under optimistic scenarios, for example a halo bound by objects with masses $\sim 3 \times 10^{-3} M_\odot$, only $\sim 15$ events are predicted per year.

Compact dark objects in the Milky Way may also be found through lens-induced distortions in the images of background sources. This technique is suitable for objects of mass $10^5 M_\odot \lesssim M \lesssim 10^6 M_\odot$, say primordial black holes, which produce distortions on angular scales $\gtrsim 0.1''$. A background optical continuum source such as the Andromeda Galaxy, the Galactic Center, or an extended radio source like Centaurus A has to be imaged with high dynamic range and subarcsecond resolution, and one then looks for characteristic lensing patterns such as Einstein rings and arcs (Turner et al 1990). At least 1 square degree of sky must be imaged in order to have a reasonable chance of success.

Yet another proposal is that compact lenses in the Galaxy may cause variable time-delays that may be detected, e.g. in the signal from a radio pulsar (Krauss & Small 1991). This could, in principle, be used to measure the mass of the deflecting star.

5.2 External Galaxies

5.2.1 Cores The high surface density associated with the cores of galaxies, coupled with the expectation that most of this density is due to stars, implies that images of a background source that happen to be superimposed on the central region of a foreground galaxy are likely to display microlensing effects. Q2237+031, with four images less than 1 kpc from the center of a $z = 0.039$ galaxy, was recognized soon after its discovery to be an excellent candidate (Kayser et al 1986, Schneider et al 1988,
Kayser & Refsdal 1989). Indeed, an apparent microlensing event was seen in 1988 (Irwin et al 1989), when image A brightened by a fraction of a magnitude within a time of order a month. Other events have since been seen in image A and in the other images (Corrigan et al 1991). Detailed numerical simulations (Wambsganss et al 1990a,b; Witt et al 1992) indicate that the observed variations are similar to those expected from a normal stellar mass distribution in the lens with a mass cutoff $\sim 0.1M_\odot$ and a transverse velocity $< 1000$ km s$^{-1}$. This is a unique instance where the dynamical mass of a cosmologically distant star has been measured.

Gravitational lensing can also provide information on the overall mass distribution of the lens. In Q2237+031, the positions and relative magnifications of the four images are well-reproduced by a mass distribution assumed to follow the light (Schneider et al 1988). The derived mass-to-blue light ratio is $\sim 10$, which is consistent with Population II stars. It appears that the core of this galaxy does not contain a significant amount of dark matter.

Further information on the mass distribution in the cores of lensing galaxies comes from observed upper limits on the strength of the central/odd image. The magnification of this image is expected to be $\sim (\xi_c/D_0\theta_E)^2$, where $\xi_c$ is the core radius of the mass distribution. The fact that no central image has been seen in any example of lensing so far suggests that the core radii of the lenses are no more than a few hundred parsecs (Narayan et al 1984, Narayan & Wallington 1992), or that the centers of the lenses have compact masses of $\sim 10^{10}M_\odot$ (Narasimha et al 1986). This is consistent with directly measured core radii of nearby galaxies (Lauer et al 1991). However, a contrary indication is obtained from a number of BL Lac objects that apparently lie behind the cores of low redshift galaxies (Stickel et al 1988a,b, 1989). None of these sources is multiply-imaged, which sets a lower limit $\gtrsim 2$ kpc to the core radii of the galaxies (Narayan & Schneider 1990).

Discussions of the mass distribution in the core become complicated when the additional magnification due to microlensing is included. This is because the optical depth for microlensing is usually large, and so the expected distribution of image intensities is likely to be strongly modified (Section 3.5). Moreover, if $\Sigma > \Sigma_{cr}$ (which is almost always the case for the central image), the primary image can sometimes be "swallowed" by a star (Section 3.5), leading to a highly demagnified image (Chang & Refsdal 1984, Subramanian et al 1985, Paczyński 1986a).

Microlensing at large optical depth can also occasionally give extremely large magnifications. It has been suggested that several apparent low redshift BL Lacs are in fact faint high redshift OVVs that have been microlensed by stars in a foreground galaxy in such a way that the con-
tinuum is highly brightened while the line-emitting regions are not (Ostriker & Vietri 1985, 1986, 1990). The rapid variability of these objects may also be attributed to microlensing (Nottale 1986, 1988; Schneider & Weiss 1987). Magnification bias is used to explain the intrinsically unlikely alignment of the OVV with the galaxy. However, BL Lacs (including those claimed to be lensed) have distinct radio properties, and the radio emission is unlikely to be microlensed since it is thought to come from a region at least as large as the broad-line emitting region (Gear 1991). Also, if the microlensing is due to a normal star, then the bright phase will only last a couple of years, whereas several of the BL Lacs have been known for nearly 20 years (Kayser 1992).

5.2.2 Halos Direct estimates of the masses of gravitational lenses may be obtained in most cases of multiple lensing provided the redshifts of the lens and the source are measured. At the crudest level, the angular separation of the images gives an estimate of $2\theta_E$, which is then converted to the velocity dispersion $\sigma$ of the lens (Equation 8) or the mass enclosed within the image cylinder (Equation 4, but this is not always reliable, cf. Nemiroff 1988a). Estimated values of $\sigma$ range from $\sim 200–500$ km s$^{-1}$. It is believed that in some cases, especially where the derived $\sigma$ turns out to be large, e.g. Q0957+561, there is a significant contribution to the lensing from a galaxy cluster surrounding the primary lens. An upper limit to the mass of a quasar can be obtained whenever another quasar with a higher redshift happens to lie in nearly the same direction and not show any evidence of multiple-imaging (e.g. the quasar pairs Q1548$+115$, Q1038$+528$, Gott & Gunn 1974, Dyer & Roeder 1982).

Detailed models provide a better estimate of the mass of a lens, and also help determine the "shape" of the projected mass density. Such models are available for the multiply-imaged quasars, Q0957+561 (Falco et al 1991a), Q1115+080 (Narasimha et al 1982), Q2016+112 (Narasimha et al 1987), Q0142−100 (Surdej et al 1988), and Q1413+117 (Kayser et al 1990). It appears that the solutions are not unique since differently parametrized models are comparably successful in fitting the observations (Kochanek 1991a), though the various parametrizations do tend to be qualitatively similar to one another in certain essential features of the mass distribution. The radio ring sources have arguably the best-constrained mass models because of the larger number of observational constraints provided by the resolved images (Kochanek et al 1989). However, there could be an undetermined normalization if the redshift of either the lens or the source is not known (as in MG1131+0456). The mass enclosed within the ring in MG1654+1346, where both the lens and the source have been observed, is estimated to be $9 \pm 0.4 \times 10^{10} h^{-1} M_\odot$ (Langston et al 1990) and the mass-to-blue light ratio is $16 h$ solar units (Burke 1992).
Usually, the derived mass of a gravitational lens is the sum of the mass in the galaxy and that in a surrounding cluster (if any). Interestingly, if the time delay between two of the images is measured, then the mass of the galaxy alone (enclosed between the images) may be estimated, independently of the cluster (Borgeest 1986, Narayan 1991). This is possible whenever the cluster can be approximated as a quadratic lens.

Several attempts have been made to check whether the observed numbers of multiply-imaged quasars, and the distribution of image separations, apparent magnitudes, etc are consistent with the number density and masses/velocity dispersions of galaxy-like mass condensations in the universe (Hacyan 1982, Tyson 1983, Turner et al 1984, Dyer 1984, Hinshaw & Krauss 1987, Kochanek & Blandford 1987, Narayan & White 1988, Wu 1989, Kochanek 1991b, Fukugita & Turner 1991, Mao 1991, Narayan & Wallington 1992). The probability that a given quasar will be multiply-imaged in some particular image configuration can be written as the product of factors describing respectively the assumed mass distribution of the lens, the cosmography of the world model, the effects of magnification bias (which depends upon the quasar luminosity function), an angular resolution factor (to model the difficulty of discovering small angular separation pairs), and some allowance for other selection effects (e.g. large magnitude-difference image pairs are hard to find). In the modeling, it is found that the introduction of a core radius enhances the mean image separation but reduces the overall cross section for lensing. Also, in comparing the relative incidence of doubly- and quadruply-imaged quasars, it is essential to consider nonspherical lenses. Overall, it appears from such studies that the observations are consistent with the hypothesis that most of the lensing of quasars is done by galaxies. Some additional deflection due to galaxy groups, clusters, and other large-scale mass distribution is needed to explain a minority of the cases with large image separations (Turner et al 1984, Anderson & Alcock 1986, Katz & Paczyński 1987, Narayan & White 1988, Jaroszynski 1991).

The existence of lens candidates and no detected lensing galaxy down to deep limits (Section 2.1) does leave open the possibility that the universe may contain galaxy-sized condensations made entirely of dark matter. Since lensing galaxies have been seen in roughly half the known cases of lensing, the number density of such dark mass concentrations is probably no larger than that of real galaxies. However, if the dark lenses have large core radii and subcritical central surface density, then they will produce multiple imaging only when two or more lenses happen to line up toward a distant quasar (Subramanian et al 1987). In that case their number density could be substantially larger.

Multiply-imaged quasars probe only the mass distribution interior to
the images, typically the inner regions of lens galaxies. Although lines-of-sight through outer galaxy halos do not lead to multiple-imaging, there can nevertheless be measurable effects in the images of background optical and radio galaxies (Blandford & Jaroszyński 1981, Noonan 1983, Saslaw et al 1985, Kochanek & Lawrence 1990, Kronberg et al 1991). Tyson et al (1984) have set observational limits on the tangential elongation induced in background galaxies by the lensing action of foreground galaxies. The data are generally consistent with galaxy masses derived from rotation curves assuming extended halos (Kovner & Milgrom 1987).

Several intriguing claims have been made that background sources, particularly quasars, BL Lacs and radio galaxies, preferentially occur behind foreground galaxies (Arp 1981, 1982; Sulentic 1981; Tyson 1986; Stocke et al 1987; Webster et al 1988a; Fugmann 1988, 1989, 1990; Hammer & Le Fèvre 1990). Some of these claims appear to be statistically significant. A plausible explanation of the phenomenon is that the sources are magnified by the lensing action of the galaxies and so, because of magnification bias, these sources are found preferentially in the vicinity of the lenses (Canizares 1981, Hammer & Nottale 1986a). There are two opposing effects at work here. For a magnification $\mu$, the observed number density of a given population of sources actually goes down by a factor $\mu$ because the sky is locally “stretched” by the magnification. However, this effect is compensated by the fact that sources that are fainter by up to a factor of $\mu$ will be brought into a flux-limited sample as a result of the brightening. If the logarithmic cumulative counts have a slope steeper than 0.4 per magnitude, then the net observed number density will increase (Narayan 1989, Kovner 1989b, Schneider 1989). Counts of optical quasars are steeper than this limit at the bright end, but shallower at faint magnitudes, with the break occurring at $B \sim 19$ (Boyle et al 1988). Thus, bright quasars, at least, should display preferential association with foreground galaxies. However, the magnitude of the effect is not expected to be large (Vietri & Ostriker 1983; Zuiderwijk 1985; Narayan 1989; Kovner 1989b; Schneider 1989, 1992; Hogan et al 1989), particularly for quasars close to the break magnitude, and this may be in strong conflict with a substantial overdensity seen in one sample (Magain et al 1992). The recent suggestion of a possible double magnification bias (Borgeest et al 1991) in a sample that is flux-limited in two independent wavelength bands, e.g. optical and radio, merits further investigation.

One question yet to be settled is whether quasar-galaxy associations are primarily due to macrolensing by the overall mass distribution of the galaxy or whether microlensing by individual stars in the galactic halo is also important. An argument for the latter is the fact that the preferential association appears to be strongest for compact sources such as optical
and X-ray quasars and BL Lacs (and possibly compact flat spectrum radio sources, though the case for microlensing here is weaker). Also, theoretical arguments indicate that magnification bias is stronger when macro and microlensing are considered together than with the former alone (Vietri & Ostriker 1983; Nityananda & Ostriker 1984; Ostriker & Vietri 1986; Schneider 1986, 1987a,b; Bartelmann & Schneider 1991; Nemiroff 1991). However, it has been claimed that 3C radio sources, most of which are quite extended, have foreground galaxies in their vicinity unusually often (Hammer & Le Fevre 1990). This cannot be due to microlensing. Also, a natural way to reconcile the rarity of the gravitational lensing phenomenon (the probability of a given source being multiply imaged is \( \sim 10^{-4} \)–\( 10^{-3} \)) with the fact that 1830–211, one of the brightest radio sources known, is lensed into an Einstein ring, is to invoke magnification bias. Again, since the source is well-resolved, microlensing cannot be important.

Microlensing in a galaxy halo would be directly confirmed if the associated time variability in a background quasar is ever detected (Chang & Refsdal 1979, 1984; Gott 1981; Young 1981; Subramanian et al 1985). Such variability would reveal the mass of the microlens and provide information on the stellar content of the lens. Schild & Smith (1991) have obtained fairly convincing evidence for microlensing in Q0957+561. A few other claims of variability have been made (Borgeest et al 1992a,b; Altieri & Giraud 1991), but none are unambiguous.

5.3 Clusters of Galaxies

The possibility that galaxy clusters and superclusters may produce gravitational lensing has been considered by several authors (Noonan 1971, Dyer & Roeder 1976, Narayan et al 1984, Sanders et al 1984, Webster 1985, Hammer & Nottale 1986b, Crawford et al 1986, Blandford et al 1987, Kovner 1987f). The discovery of long arcs in a handful of clusters and the subsequent discovery of smaller arcs in these and several other clusters (Section 2.2) has now opened up this exciting field. Compared to lensing by individual galaxies, clusters offer several advantages. (a) Rich clusters cover a substantially larger angular area of the sky, thus increasing the cross section for lensing. (b) The lensed sources in this case, faint galaxies, are much more numerous than quasars, particularly at the low brightness levels achievable today. (c) Since the sources are resolved, image distortions may be directly measured, and one can study even modest effects due to lensing. (d) Since the mass distribution of clusters is considerably less relaxed and ordered than galaxy halos, there is potentially much more new information to be obtained through gravitational lensing.

Several groups have attempted to model the clusters with the longest arcs using a variety of mass models (Hammer 1987; Fort et al 1988; Kovner
1988, 1989a; Grossman & Narayan 1988, 1989; Narasimha & Chitre 1988; Hammer & Rigaut 1989; Wambsganss et al 1989; Nemiroff & Dekel 1989; Bergmann et al 1990; Mellier et al 1990; Hammer 1991; Kassiola et al 1992). Parameters such as the velocity dispersion $\sigma$ of the cluster, the mass-to-light ratio $M/L$, core radius, and ellipticity have been estimated. Since long arcs arise as a result of extreme tangential elongation of the image, usually a cusp caustic, the radius of curvature of such an arc is roughly of order the Einstein radius [but not always, e.g. the “straight arc” in Abell 2390 (Pello et al 1991)]. This can be used to estimate $\sigma$ and $M/L$ of the lens (cf Section 3.2). Derived values of these parameters are generally quite high ($\sigma \gtrsim 1000$ km s$^{-1}$, $M/L \gtrsim 100$ solar units), but consistent with measured cluster velocity dispersions where available (Mellier et al 1988). Almost all studies indicate that the lensing clusters must have smaller core radii (<100 kpc) than indicated by either their optical or X-ray images (e.g. Bahcall 1977). Abell 370 shows the clearest evidence for noncircularity, with a strong indication that the projected mass distribution is elongated along the axis defined by the two central cD galaxies and not traced by the observed galaxies (e.g. Grossman & Narayan 1989, Soucail 1991). In other cases (e.g. Cl 2244-02) it is possible to associate all the mass with the visible cluster galaxies (e.g. Bergmann et al 1990, Mellier et al 1992).

The longest and most luminous arcs are rare since they are due to high redshift bright galaxies that happen to be close to caustics. Images of the sky down to a level of $B = 29$ per square arcsecond reveal a population of faint blue galaxies with a number density as high as $\sim 100$ per square arcminute (Tyson 1988). It now appears that up to $B = 27$ (corresponding to a surface brightness of $B = 29$ per square arcsecond with 0.8 arcsecond seeing), $\sim 80\%$ of the galaxies have redshifts in the range 0.7–1.4 (Fort 1992), consistent with extrapolations of redshift surveys of 24$^m$ galaxies (e.g. Colless et al 1990, Lilly et al 1991), and with other observations (Grossman 1990, Guhathakurta et al 1990). These galaxies are therefore ideal for studying the gravitational lensing influence of clusters at redshifts $\lesssim 0.5$.

The tangentially elongated arclets observed in the fields of several moderate redshift rich clusters, notably A1689 with 50 arclets (Tyson et al 1990) and Abell 370 with 60 (Fort 1992), give two-dimensional maps of a distortion measure which approximately reflect the cluster mass distribution. It is found that the dark matter is reasonably correlated with the cluster red light. This promising technique to trace the mass in clusters is unfortunately limited by lack of knowledge of the individual redshifts and intrinsic ellipticities of the background sources (Kochanek 1990b, Miralda-Escudé 1991a, Knieb et al 1992).
Individual galaxies in a rich cluster can have an exaggerated influence on images of background sources because of the enhancement of their lensing action by the overall mass distribution of the cluster. Particularly noteworthy are the optical rings seen around a few cluster galaxies. If these are due to gravitational lensing, then their analysis leads to interesting constraints on the mass of the galaxy as well as the surface density and shear of the cluster. However, there are alternative explanations of these features (Kochanek & Blandford 1991, Petrosian 1992).

5.4 Large-Scale Structure and Intergalactic Medium

A great deal of evidence has accumulated in recent years that the universe is inhomogeneous on scales of up to 50–100 Mpc (e.g. Geller & Huchra 1989, Saunders et al 1991). Such inhomogeneity will affect the propagation of light and lead to fluctuations in the convergence and shear of light bundles (Section 3.3). This has been studied in a number of papers which discuss distortions in the images of distant sources and changes in the apparent luminosities of the sources (e.g. Kantowski 1969; Dyer & Roeder 1973; Hammer 1985; Dyer 1986; Dyer & Oattes 1988; Schneider & Weiss 1988a,b; Kasai et al 1990; Jaroszyński et al 1990; Babul & Lee 1991; Bartelmann & Schneider 1991; Tomita 1991).

The discovery of the population of faint blue galaxies now provides a potential tool to study these effects. Specifically, large-scale mass inhomogeneities will induce correlated elliptical distortions in images of background sources, which could be measured by mapping large numbers of sources over areas of the sky \( \sim 1 \) square degree in size (Kristian 1967, Blandford et al 1991, Miralda-Escudé 1991b, Kaiser 1992). The magnitude of the correlated ellipticity should be in the range 1–3% (depending on the source redshift distribution) if the cold dark matter cosmology is correct, with much larger signals, should mass trace light. The same technique will also reveal any coherent structures such as walls or voids. Although an early attempt to measure some of these effects only gave upper limits (Valdes et al 1983), recent improvements in techniques make it worthwhile to repeat the observations.

In addition to relatively smooth large-scale inhomogeneities, a fraction of the mass in the universe may be in the form of compact dark objects on mass scales ranging from stars to galaxy clusters. These objects will cause distortions in the images of cosmologically distant sources on angular scales \( \sim \theta_E \) (Press & Gunn 1973). Roughly, the optical depth to lensing is \( \sim \Omega_\rho \), the density parameter corresponding to mass within the projected Einstein rings of the lenses. For compact objects in the mass range \( 10^{-2} \)–\( 10^5 \) \( M_\odot \), a limit of somewhat less than unity may be set on \( \Omega_\rho \) merely by the fact that the line-to-continuum ratio is relatively constant among
quasars (Canizares 1982). Better limits could be obtained from future
observations of cosmologically distant supernovae (Schneider & Wagoner
1987, Linder et al 1988, Rauch 1991). With masses larger than $10^5 M_\odot$,
one would expect occasionally to see image distortions (Blandford &
Jaroszyński 1981), image doubling (Nemiroff & Bistolas 1990, Nemiroff
1991), rings (Burke 1992), or phantom images in radio maps (Kassiola et
al 1991). Current observations limit the $\Omega_0$ in the form of these objects
to be $\lesssim 0.1$. Future observations may be used to put significantly stronger
limits.

Lensing by cosmic strings has been studied by some authors. A straight
string will produce a characteristic lensing pattern consisting of approxi-
mately equal-separation image pairs (Vilenkin 1984, 1986; Gott 1985) with
similar parity. A promising example of this was discovered (Cowie & Hu
1987), but it may be merely a cluster of binary galaxies (Hewitt et al 1990,
Hu 1990). The image of an extended source, such as a galaxy, can exhibit
a discontinuity if the source happens to lie behind a string (Paczyński
1986c), though this signature may not be obvious if strings exhibit small-
scale kinky structure as now thought (Bennett & Bouchet 1988). Small
string loops are expected to have lensing properties similar to those of
galaxies (Hogan & Narayan 1984).

6. QUASARS AND GALAXIES AT HIGH
REDSHIFT

Gravitational lenses can be used as probes of the sources that they magnify.
In this section, we review what has been learned about quasars and galaxies
in this way, and also describe how lensing can sometimes interfere with
our attempts to study distant regions of the universe.

6.1 Quasar Emission Regions

As discussed in Section 3.5, gravitational microlensing leads to significant
flux variation only if the angular size of the source is smaller than the
Einstein radius of the lens. This allows interesting constraints to be set on
the sizes of quasar emission regions (Sanitt 1971). Based on the variability
observed in Q2237+031, the linear size of the continuum emission region
in this source is estimated to be $\lesssim 10^{15}$ cm—the first direct estimation
of the size of a quasar (Wambsganss et al 1990b, Witt et al 1991, Webster et
al 1991). Interestingly, this size is not consistent with a blackbody spectrum
from a standard accretion disk model of the quasar (Rauch & Blandford
1991) and suggests the presence of a nonthermal component. Future multi-
color observations should lead to additional information on the structure
of this quasar (Wambsganss & Paczyński 1991). Interesting effects are also
possible in the broad line emission, particularly if it arises in compact clouds orbiting the quasar nucleus (Nemiroff 1988b, Schneider & Wambsganss 1990). Differential magnification can cause variation in the line profiles between macroimages and this can, in principle, be used to probe the cloud kinematics.

Since the strongest microlensing events, the so-called high-amplification events (HAE), are due to the source crossing a caustic, either a fold or a cusp, and since the generic structure of caustics is well-understood, it may be possible in favorable cases to invert the light curve to obtain the one-dimensional structure of the source (Kayser et al 1986; Grieger et al 1986, 1988; Refsdal 1990; Grieger 1990; cf also Refsdal 1966b). If observations are done simultaneously from Earth and one or more spacecraft in the solar system, then it will be possible to distinguish between microlensing and intrinsic source variability and also to determine the source profile as well as the lens velocity.

6.2 Magnified Images of Galaxies and Radio Sources

In the really spectacular examples of gravitational lensing of resolved sources, such as the blue luminous arcs and radio rings, some parts of the source are magnified by quite large factors. When a successful model is developed, this automatically provides a reconstruction of the source with particularly high resolution information in those parts of the source that lie close to a caustic. The very fact that the arcs are highly magnified allows spectroscopy to be carried out on 27" galaxies, three magnitudes fainter than normal. Several arc redshifts have been thus determined. An even more impressive example is the arc in Abell 2390, where a velocity variation of 378 km s\(^{-1}\) has been measured along the length of the arc (Pello et al 1991). By combining a knowledge of the velocity width with the central surface brightness, it might be possible to measure the Hubble constant, or more realistically, the evolution of the Tully-Fisher relation (Soucail & Font 1991).

Occasionally, a star in the source galaxy may come close enough to a caustic to be brightened for a few hours by a huge factor of up to 10\(^7\) or more (Miralda-Escudé 1991b). Other examples of gravitational lensing give more modest magnifications but may still be useful. VLBI imaging of the core-jet region of a lensed radio quasar (Q0957+561), HST imaging of the fuzz around a lensed optical quasar or its line emitting regions, and the imaging of Lyman-\(\alpha\) emission regions (Q2016+112) are examples of this. It has been suggested that some examples of superluminal velocities seen in the VLBI jets of extragalactic radio galaxies may be the result of gravitational magnification of more modest velocities (e.g. Chitre & Narlikar 1979), and that rapid variability may be associated with micro-
lensing, if the sources are intrinsically extremely compact (Gopal-Krishna & Subramanian 1991, Subramanian & Gopal-Krishna 1991).

6.3 Exploitation of Time Delays

The occurrence of time delays among multiple images of a lensed source can provide unique opportunities to study transient phenomena. For instance, were a supernova to go off in a multiply-imaged region of an arc-galaxy, the explosion will be seen three or more times with delays (Kovner & Paczyński 1988). If a reasonable estimate of the time delay is available (say from a model), one could wait and observe one of the later incarnations of the event in much greater detail than would be possible without prior warning. (However, this will not provide a measurement of \( H_0 \) as the cluster potential is unlikely to be understood well enough.) The same principle could be used with any unusual event in the multiply-imaged quasars, e.g. transient flaring of the optical or X-ray emission, or ejection of a radio VLBI blob.

The BATSE experiment on the Compton Gamma Ray Observatory (Meegan et al 1992) has shown that most gamma-ray bursts are isotropically distributed on the sky and have non-Euclidean source counts, which suggests that they lie at cosmological distances (e.g. Prilutski & Usov 1975, Mao & Paczyński 1992, Piran 1992). If so, gamma-ray bursts constitute a fairly homogeneous population of near point sources with (model-dependent) angular sizes \( \sim 10^{-20} \) radians. Gamma-ray telescopes have only primitive angular resolving power, but submillisecond temporal resolution, and so this hypothesis can be tested by seeking multiple bursts from the same source separated in time by \( \sim 10^{-4}(M/M_\odot) \) s where \( M \) is the lens mass (Paczyński 1986b, 1987b). Galaxy lenses will create repeating bursts with median time delays \( \sim 1 \) month with a probability \( \sim 10^{-3} \) (Mao 1992), and examples ought to be seen after \( \sim 3 \) yr of full operation of BATSE. Microlensing by individual stars will probably not be resolvable, but a cosmologically significant density of compact masses with \( M \gtrsim 100M_\odot \) could be detected. Using the observed time delays and magnification ratios, the masses of the individual lenses could be determined with good accuracy (Narayan & Wallington 1992). Interestingly, for a lens with mass \( \lesssim 10^{18} \) g, the Fresnel length is comparable to the Einstein radius and energy-dependent diffraction effects are predicted (Bliokh & Minakov 1975, Gould 1992).

6.4 Quasar Absorption Lines

Multiply-imaged quasars provide unique opportunities to probe the intervening intergalactic medium between us and the source. Since the geometry of the two ray-paths is known, the presence or absence of common absorp-
tion lines in the multiple spectra leads to powerful constraints on the transverse sizes of Lyman-\(\alpha\) clouds and metal-line systems (Weymann et al 1979, Krolik & Kwan 1979, Young et al 1981c, Weymann & Foltz 1983, Foltz et al 1984, Thomas & Webster 1990, Steidel & Sargent 1990, Smette et al 1992). Sizes of 5–25 kpc have been deduced from Lyman-\(\alpha\) absorption lines observed in Q0957+561 and Q2345+007. Conversely, the redshift distribution of the absorption systems in the images can be used as an argument for, or against, multiple imaging (Duncan 1991, Steidel & Sargent 1991).

For most models of the clouds responsible for the Lyman-\(\alpha\) forest, no significant gravitational lensing is expected by the clouds themselves (Ikeuchi & Turner 1991).

### 6.5 Quasar Luminosity Function

Even before the discovery of the first gravitational lens it was recognized that lensing has the potential to modify our view of cosmologically distant parts of the universe (Barnothy 1965, 1966; Barnothy & Barnothy 1968, 1986; De Silva 1970). One particular question that has been discussed often is whether the observed quasar luminosity function could be significantly affected by lensing (Turner 1980, Avni 1981, Peacock 1982, Setti & Zamarani 1983, Vietri & Ostriker 1983, Vietri 1985, Ostriker & Vietri 1986, Isaacson & Canizares 1989).

Flux conservation demands that the average magnification of a randomly selected source in an inhomogeneous universe must be unity compared to a smooth universe of the same \(\Omega_0\) (Weinberg 1976, Ehlers & Schneider 1986). However, in the former universe one has a distribution of magnifications, \(P(\mu)\), while in the latter all sources have unit magnification. Consequently, the observed luminosity functions can differ in the two cases. Typically, gravitational macrolensing has a \(P(\mu)\) that is peaked around \(\mu \sim 1\) with a power-law tail extending to large \(\mu\). The exponent in the tail is \(-3\) if large magnifications are dominated by fold caustics and \(-7/2\) if due to the exterior single-image region of cusp caustics. If the intrinsic source differential luminosity function \(\phi(L)\) is flatter than \(L^{-3}\), then lensing has only a minor effect on the observed luminosity function. This is true for optical quasars fainter than \(B \sim 19\) (Boyle et al 1988). However, if there is any range of \(L\) for which the intrinsic counts are steeper than \(-3\), e.g. quasars brighter than \(B \sim 19\), then the observed population could potentially have a large contribution from highly-magnified intrinsically-weak sources (because of magnification bias).

Quantitative estimates based on the known populations of galaxies and clusters in the universe indicate that lensed quasars are unlikely to dominate in any magnitude range where there are substantial source counts. Never-
theless, even a modest influence due to lensing, say at bright magnitudes, is of interest since it implies that multiply-imaged quasars will be particularly common in such a population. Based on this line of thinking, many lens searches have been confined to high redshift, high apparent luminosity quasars (cf Section 7.2). This strategy has had some success (Magain et al 1988), but the lack of a greater success rate does suggest that macrolensing has only a weak effect on the quasar luminosity function even at the brightest end ($B \lesssim 17$).

The additional effect due to microlensing has been considered by some authors (Vietri 1985; Ostriker & Vietri 1986; Schneider 1987a,b; Bartelmann & Schneider 1990). Once again, at very large $\mu$, $P(\mu)$ has a power-law character due to the effect of caustics. However, the power-law tail is cut off above a critical $\mu$ that depends on the size of the source and on the mass distribution of the microlenses. It is our opinion that microlensing has only a marginal influence on the quasar luminosity function even at the brightest quasar magnitudes. However, it is possible that microlensing does play a role in some of the quasar-galaxy associations discussed in Section 5.2.2.

Given a model of $P(\mu)$ and an observed luminosity function $\phi(L)$, it is possible in principle to obtain the true luminosity function $\phi_{\text{true}}(L)$ (Schneider 1992).

### 6.6 Microwave Background

Several authors have investigated whether the observed anisotropy of the cosmic microwave background could be significantly modified by gravitational lensing (Dyer 1976; Mitrofanov 1981; Nottale 1984; Chitre et al 1986; Blanchard & Schneider 1987; Linder 1988, 1990a,b; Cole & Efstathiou 1989; Sasaki 1989; Durrer & Kovner 1990; Watanabe & Tomita 1991). Since a stationary lens does not alter the surface brightness of a source, all that a population of gravitational lenses will do is to distort the brightness fluctuations on the sky. This can modify the angular power spectrum of the microwave anisotropy and shift the scale on which these fluctuations are observed. In principle, for extremely strong lensing, the fluctuations can be so badly scrambled as to be wiped out entirely at the resolution of the observations. However, given the present limits on the number density and masses of lenses, this appears to be very unlikely.

It has been pointed out that surface brightness is not preserved if the lens is moving across the line-of-sight (Mitrofanov 1981, Birkinshaw & Gull 1983, Kaiser & Stebbins 1984, Gurvits & Mitrofanov 1986, Khmil' 1988). The temperature ahead of the lens is greater than that behind by $\sim T \alpha z / c$, where $T$ is the mean temperature, $\alpha$ is the deflection angle at the lens, and $v_\perp$ is the perpendicular velocity. To detect this effect with
present-day techniques, one either needs a large $v_\perp/c$ (e.g. a relativistically moving cosmic string) or a large $\alpha$ (e.g. a supercluster-scale lens).

7. SEARCHES FOR GRAVITATIONAL LENSES

The first examples of multiple quasars, rings, and arcs were identified in observing programs directed toward quite different goals. Since then, roughly half of the known instances of gravitational lensing have arisen serendipitously, which makes it difficult to quantify the incidence of lensing and to draw cosmographic deductions. In this section we briefly review attempts to search for lenses in uniform, complete samples of extragalactic sources.

7.1 Radio Surveys

A significant number of lensed objects have been found from radio surveys. The MIT-Greenbank (MG) survey (Bennett et al. 1986) consists of 6000 sources found at 5 GHz using a transit telescope. These were subsequently mapped at the VLA with 0.3'' resolution and those that exhibited unusual compact structure were imaged optically. Subsequent spectroscopy of the nonstellar objects has yielded two secure cases of multiple imaging (plus four candidates) and three rings (Burke 1990).

Gravitational lenses have also been sought in sources observed with MERLIN (Walsh 1992) which is now an ideal instrument for finding candidates with subarcsecond structure. One ring has been found. Radio ring images of background extended radio sources ought to be more common in existing surveys than has hitherto been reported (Kochanek & Lawrence 1990) and there is optimism that many more of these sources, which are promising for measuring the Hubble constant (cf Section 4.1), will be found.

Q0957 + 561 and 1830 − 211 were also found from radio surveys.

7.2 Optical Surveys

Existing quasar surveys (e.g. Boyle et al. 1988) show that the number of quasars brighter than some luminosity $L$ varies as $L^{-\beta}$, where $\beta \sim 2.7$ (Boyle et al. 1988) down to $\sim 19^{m}$. As the dominant lens cross sections (due to fold caustics, cf Section 3.3) decrease only as the inverse square of the magnification, the majority of lensed sources brighter than $19^{m}$ ought to be highly magnified as a result of magnification bias (Turner et al. 1984, Section 6.5). This has so far proved to be the case, and has motivated optical lens searches among samples of bright quasars.

Several groups (e.g. Djorgovski & Meylan 1989, Surdej 1990, Swings et al. 1990) have combined to search for lenses in an ESO key project. CCD images of approximately 400 bright ($M_V \lesssim -29$), distant ($z \gtrsim 1$) quasars
drawn from highly heterogeneous and ill-defined catalogs have been obtained. Two secure and two good candidate lenses have so far been identified, as well as two quasar pairs (cf Section 2.1). Negative searches in smaller bright quasar samples have also been reported (Crampton 1992, Yee 1992). Selection criteria are somewhat poorly defined, which leads to difficulties in drawing quantitative conclusions about the frequency of multiple imaging (cf Section 5.2.2).

These shortcomings may be partly overcome in the APM survey (Webster et al 1988b, Hewett 1992) which establishes a methodology for creating a complete sample. Quasar candidates are obtained by uniformly scanning photographic plates covering a large fraction of the southern sky. This survey appears to be highly efficient at rediscovering known quasars (Foltz et al 1989) and 2500 quasars brighter than $20''$ are believed to be present in the survey. But the survey may be less efficient at identifying multiple quasars with separations similar to the angular resolution because, by selecting unusually elliptical images, it is prejudiced against small separations and, in particular, five image configurations (Kochanek 1991b). One lens candidate with angular separation $5''$ [Q1429+008 (Hewett et al 1989)], and three quasar pairs have so far been found using this method. Further success is dependent upon obtaining spectra of a large number of candidate objects.

A comparatively new search technique uses the Hubble Space Telescope to make short exposures during gaps in the regular observing program of high luminosity quasars (Bahcall et al 1992). 354 bright quasars will be surveyed. Simulations show that, despite the degraded point spread function, this search ought to be sensitive to separations as small as $\sim 0.1''$ and magnitude differences up to $\sim 2''$. As of this writing, one lens candidate has been found [Q1208+101 (Maoz et al 1992)].

8. OUTLOOK

Over the next few years, we can look forward to several new instruments that will be well suited to studying different aspects of the gravitational lens problem. At radio wavelengths, the VLBA should produce far more detailed maps of the radio rings as well as the radio jets in Q0957+561 and Q2016+112, which will motivate correspondingly detailed modeling. The enlarged MERLIN radio telescope may have just the right angular resolution to find several more examples of galaxy-imaged compact radio sources. In the optical, a repaired and fully operational Hubble Space Telescope will have the angular resolution to locate intervening galaxies, and the Keck Telescope will be particularly important for faint object spectroscopy. Detector advances in CCDs and infrared arrays also promise
to make a substantial difference. Because of these and several other developments, the observational future looks bright.

If we can learn anything from the history of cosmology, it is that it is much easier in theory than in observational practice to make secure, quantitative deductions about the large-scale structure and the evolution of the universe. By contrast, if the last decade’s progress in understanding gravitational lenses can teach us anything, it is that we have usually been too conservative in our anticipation of novel phenomena. It will be fascinating to see which of these two opposing trends dominates over the next decade.

ACKNOWLEDGMENTS

This work was supported in part by NSF grants AST 89-17765 (RB) and AST 91-09525 (RN). We thank George Efstathiou, Emilio Falco, Bernard Fort, Paul Hewett, Chris Kochanek, Charles Lawrence, Jordi Miralda-Escudé, Bohdan Paczyński, Kevin Rauch, Sjur Refsdal, Rudy Schild, Peter Schneider, Jean Surdej, Sylvanie Wallington, and Joachim Wambstgas for comments on the draft manuscript. We are grateful to Helen Knudsen and Rosanne Scholey for bibliographic assistance.

Literature Cited

Barnothy, J. M. 1966. Observatory 86: 115
Bartelmann, M., Schneider, P. 1990. Astron. Astrophys. 239: 113
GRAVITATIONAL LENSES IN COSMOLOGY

Fort, B. 1990. See Mellier et al 1990, p. 221

© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System


Kochanek, C. S. 1990a. See Mellier et al 1990, p. 244


© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System


Pello, R., Le-Borgne, J. F., Soucail, G., Mel-}


Penrose, R. 1966. In Essays in Honor of


Petrosian, V. 1992. See Kayser & Schramm


151: 411


*Nature* 321: 569


Astron.* 38: 1


397


Space Sci.* 34: 395


231: 229


Lett.* 381: L39

Rauch, K., Mao, S., Paczyński, B., Wamb~gs


Refsdal, S., Kayser, R. 1988. In The Post-

Recombination Universe, ed. N. Kaiser, A.

Lasenby, p. 297. Dordrecht: Kluwer


Astrophys.* 250: 62


Roberts, D. H., Lehar, J. N., Hewitt, J. N.,


Sanders, R. H., Van Albada, T. S., Ooster-


Saslaw, W. C., Narasimha, D., Chitre, S. M.


Saunders, W., Frenck, C., Rowan-Robin-

son, M., Efstathiou, G., Lawrence, A., et al.


101: 813

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

413


Schneider, P. 1987b. *Astron. Astrophys.* 183:

189


Schneider, P. 1989. *Astron. Astrophys.* 221:

221


press


314: 154


Astrophys.* 237: 42


Astrophys.* 164: 237


Astrophys.* 171: 49


526


1


Astrophys.* 247: 269

Schneider, D. P., Turner, E. L., Gunn, J. E.,

Hewitt, J. N., Schmidt, M., Lawrence, C. R.


*Gravitational Lensing*. Berlin: Springer-

Verlag. In press


19


Astrophys.* 118: L1


789


Smalı, I., Ellis, R. S., Fitchett, M. J.,

Norgard-Nielsen, H. U., Hansen, L.,


Smette, A., Surdej, J., Shaver, P. A., Foltz,


press

Soucail, G. 1991. In *Clusters and Super-

clusters of Galaxies*, ed. A. C. Fabian.

Dordrecht: Kluwer. In press


Astrophys.* 243: 23

Soucail, G., Fort, B., Mellier, Y., Picat, J. P.


Steidel, C. C., Sargent, W. L. W. 1990.

*Astron. J.* 99: 1693


*Astron. J.* 102: 1610


*Astron. Astrophys.* 198: L13


*Astron. Astrophys.* 206: L30


*Astron. Astrophys.* 224: L27

Stocke, J. T., Schneider, P., Morris, S. L.,

Gioia, I. M., Maceccaro, T., Schild, R. E.


*MNRAS* 219: 333

Subramanian, K., Chitre, S. M., Narasimha,


© Annual Reviews Inc. • Provided by the NASA Astrophysics Data System
358 BLANDFORD & NARAYAN

Surdej, J. 1990. See Mellier et al 1990, p. 57
Tyson, J. A. 1990. See Mellier et al 1990, p. 230