

# Quasar Lensing

Frederic Courbin<sup>1,2</sup>, Prasenjit Saha<sup>3</sup>, and Paul L. Schechter<sup>4</sup>

<sup>1</sup> Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860,  
Departamento de Astronomía y Astrofísica, Casilla 306, Santiago 22, Chile

<sup>2</sup> Université de Liège, Institut d'Astrophysique et de Géophysique,  
Allée du 6 août, Bat. B5C, Liège 1, Belgium

<sup>3</sup> Astronomy Unit, School of Mathematical Sciences  
Queen Mary and Westfield College, London E1 4NS, UK

<sup>4</sup> Massachusetts Institute of Technology,  
70 Vassar Street, Cambridge, MA 02139, USA

**Abstract.** Massive structures, such as galaxies, act as strong gravitational lenses on background sources. When the background source is a quasar, several lensed images are seen, as magnified or de-magnified versions of the same object. The detailed study of the image configuration and the measurement of “time-delays” between the images yield estimates of the Hubble parameter  $H_0$ . We describe in a simple way the phenomenon of strong lensing and review recent progress made in the field, including microlensing by stars in the main lensing galaxy.

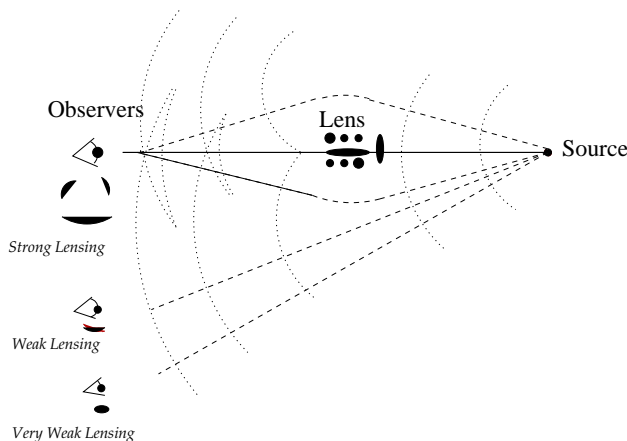
## 1 Concepts

### 1.1 The formation of multiple images

There are several ways of understanding the effect of gravity on light in the context of lensing. We start with an approach which lends itself particularly well to pictorial representation.

**Wavefronts** A schematic wavefront is illustrated in Figure 1. Spreading outwards from a point source, the wavefront is initially spherical. But as it passes through the gravitational field of the lens the wavefront gets delayed and bent; we can interpret this effect as a slowing-down of light by a gravitational field, usually called the Shapiro time delay [123]. Where the wavefront crosses an observer, they see an image in the direction normal to the wavefront, and images will be (de)magnified and/or distorted according to how curved the wavefront is as it crosses the observer. If the lens is strong enough, the wavefront can fold in on itself, producing multiple images. If moreover the source is variable, different images will show that variability with time delays proportional to the spacing between these folds, i.e., the cosmological distance scale.

It is possible to make the above explanation quantitative within the wavefront picture [103,55], but for calculations that is usually not the most convenient route. Notice that the wavefront picture has a single source and multiple



**Fig. 1.** Schematic illustration [1] of the wavefront and the different regimes of lensing. Lensed quasars fall in the strong lensing regime; the other regimes are important in lensing by clusters of galaxies.

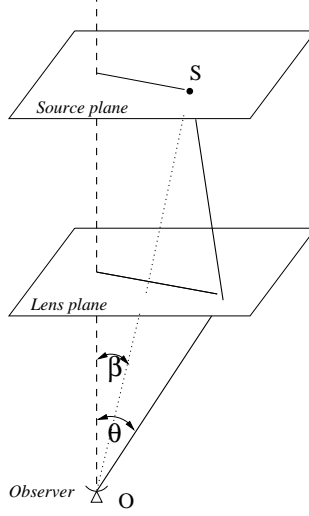
observers, whereas astrophysical problems generally involve multiple sources and a single observer. So calculations are easier if we use a relative of the wavefront called the arrival-time surface [16,90].

**Arrival times** Consider Figure 2: in the usual astrophysical approximation of small angles and thin lenses, this figure shows a virtual light ray getting deflected by the lens and reaching the observer from direction  $\vec{\theta}$ , the source being at angular position  $\vec{\beta}$ . (Vector signs denote 2D angles on the sky.) The arrival time is the light travel time—with irrelevant constants discarded—of such a virtual ray as a function of  $\vec{\theta}$ , with  $\vec{\beta}$  held fixed. It has two contributions: a ‘geometrical’ part and a ‘gravitational’ part [16]. The geometrical part is simply the difference between the continuous and dotted paths in Figure 2, and is given by

$$t_{geom}(\vec{\theta}) = \frac{1}{2}(1 + z_L) \frac{D_L D_S}{c D_{LS}} (\vec{\theta} - \vec{\beta})^2, \quad (1)$$

where  $z_L$  is the lens redshift and the  $D$  factors are angular diameter distances as shown in Figure 2. The gravitational part is the Shapiro time delay in a gravitational field from general relativity, and depends on the surface density  $\Sigma(\vec{\theta})$  of the lens. A concise way of writing the Shapiro delay is

$$t_{Shapiro}(\vec{\theta}) = (1 + z_L) \frac{8\pi G}{c^3} \nabla^{-2} \Sigma(\vec{\theta}). \quad (2)$$



**Fig. 2.** Illustration of a virtual light ray:  $\vec{\beta}$  is the unlensed sky position of the source, and  $\vec{\theta}$  is its apparent position. In the text, we use  $D_L$ ,  $D_S$ , and  $D_{LS}$  for angular diameter distances from observer to lens, observer to source, and lens to source.

Here  $\nabla^{-2}$  denotes the inverse of a 2D Laplacian with respect to  $\vec{\theta}$ ,<sup>1</sup> and  $\nabla^{-2}\Sigma(\vec{\theta})$  is some sort of 2D potential.

Putting equations (1) and (2) together we have the arrival time  $t(\vec{\theta})$  in full:

$$t(\vec{\theta}) = \frac{1}{2}(1+z_L)\frac{D_LD_S}{cD_{LS}}(\vec{\theta}-\vec{\beta})^2 - (1+z_L)\frac{8\pi G}{c^3}\nabla^{-2}\Sigma(\vec{\theta}). \quad (3)$$

From Fermat's principle, real light rays take paths that make the arrival time stationary. Thus the condition for images is

$$\vec{\nabla}t(\vec{\theta}) = 0. \quad (4)$$

Equation (3) looks formidable, but it will become much less so once we introduce some scales.

<sup>1</sup> By this we mean an operator that solves Poisson's equation in 2D. Thus, if  $\nabla^2 f(\vec{\theta}) = g(\vec{\theta})$ , we write  $f(\vec{\theta}) = \nabla^{-2}g(\vec{\theta})$ . The explicit form of the inverse Laplacian is as an integral

$$f(\vec{\theta}) = \int \ln|\vec{\theta}-\vec{\theta}'|g(\vec{\theta}')d^2\vec{\theta}'$$

but we will not need it in this article.

**Some scales** Consider a point-mass lens and a point source along the same line of sight, i.e.,  $\vec{\beta} = 0$  and  $\Sigma(\vec{\theta}) = M\delta(\vec{\theta})$ . The arrival time then becomes

$$t(\vec{\theta}) = \frac{1}{2}(1 + z_L) \frac{D_L D_S}{c D_{LS}} \theta^2 - (1 + z_L) \frac{4G}{c^3} \ln \theta, \quad (5)$$

since  $\nabla^{-2}\delta(\vec{\theta}) = \ln \theta / (2\pi)$ , and there is a minimum at  $\theta = \theta_E$  where

$$\theta_E^2 = \frac{4GM}{c^2} \frac{D_{LS}}{D_L D_S}. \quad (6)$$

This corresponds to a ring image, called an Einstein ring, and  $\theta_E$  is called the Einstein radius. If the source is much further than the lens

$$\theta_E \simeq 0.1 \text{ arcsec} \times \left[ \frac{M \text{ in } M_\odot}{D_L \text{ in pc}} \right]^{\frac{1}{2}}. \quad (7)$$

The combination of a point lens and colinear source is very improbable, but the Einstein radius is a very useful concept, for two reasons. First, even if there is no Einstein ring in a multiple-image system, the image separation still tend to be of order  $\theta_E$ . Secondly, the Einstein radius also supplies a scale for  $\Sigma$ , by the following argument.

From the two-dimensional analog of Gauss's flux law, for any circular mass distribution  $\Sigma(\theta)$ ,  $\nabla^2 t(\vec{\theta})$  will depend only on the enclosed mass. So not just a point mass, but any circular distribution of the mass  $M$ , will produce an Einstein ring from a colinear source, *provided it fits within  $\theta_E$* . The condition of a mass fitting into its own Einstein radius is known as 'compactness'. And because from (6) the area within an Einstein radius is itself proportional to the mass, compactness is equivalent to the density exceeding some critical density. Working out the algebra we easily get this critical density<sup>2</sup> to be

$$\Sigma_{\text{crit}} = \frac{c^2}{4\pi G} \frac{D_L D_S}{D_{LS}}. \quad (8)$$

From equation 3 we can also define a time scale

$$T_0 = (1 + z_L) \frac{D_L D_S}{c D_{LS}}, \quad (9)$$

which is of order the light travel time, or a Hubble time in cosmological situations. The interesting time scale in lensing, however, is not  $T_0$ , but

$$T_0 \times \langle \text{image separations} \rangle^2, \quad (10)$$

being the scale of arrival-time *differences* between images. We will meet the latter presently, in the approximate equation (14).

<sup>2</sup> In this article the units of  $\Sigma$  are  $M_\odot \text{ arcsec}^{-2}$ . Some authors prefer  $M_\odot \text{ kpc}^{-2}$ . This difference of convention means that different authors' equations may differ by factors of  $D_L$  or  $D_L^2$ .

**The arrival-time surface** Using the scales introduced above, we can render dimensionless the arrival time (3),

$$\tau(\vec{\theta}) = \frac{1}{2}(\vec{\theta} - \vec{\beta})^2 - 2\nabla^{-2}\kappa(\vec{\theta}). \quad (11)$$

Here the scaled arrival time  $\tau$  and the scaled surface density  $\kappa$  (also called convergence) are both dimensionless. The last term in equation (11) is called the lens or projected potential

$$\psi(\vec{\theta}) \equiv 2\nabla^{-2}\kappa(\vec{\theta}). \quad (12)$$

The physical arrival time and density are

$$t(\vec{\theta}) = \tau(\vec{\theta}) \times T_0, \quad \Sigma(\vec{\theta}) = \kappa(\vec{\theta}) \times \Sigma_{\text{crit}}, \quad (13)$$

and the scales are approximately

$$T_0 \simeq h^{-1} z_L (1 + z_L) \times 80 \text{ days arcsec}^{-2}. \quad (14)$$

and

$$\Sigma_{\text{crit}} \simeq h^{-1} z_L \times 1.2 \cdot 10^{11} M_{\odot} \text{ arcsec}^{-2}, \quad (15)$$

where  $h$  is the Hubble constant in units of 100 km/s/Mpc.

The scaled arrival time  $\tau(\vec{\theta})$  in (11), visualized as a surface, is called the arrival-time surface. Much of lensing theory is effectively the study of the arrival-time surface and its derivatives, as we see below.

Note that although the wavefront and the arrival time surface look similar and indeed are closely related [90], they are not quite the same thing. The wavefront is a surface in real space whereas the arrival time surface is in  $(\vec{\theta}, \tau)$  space and thus a little more abstract.

**Images and magnification** The condition for images, from Fermat's principle and following (4) is

$$\vec{\nabla}\tau(\vec{\theta}) = 0, \quad \text{or} \quad \vec{\beta} = \vec{\theta} - \vec{\nabla}\psi. \quad (16)$$

The latter form is called the lens equation. Its interpretation is that the observer sees an image wherever the arrival-time surface has a minimum, maximum, or saddle point. Then consider the second derivative of  $\tau(\vec{\theta})$ , or curvature of the arrival-time surface. We have

$$\vec{\nabla}\vec{\nabla}\tau(\vec{\theta}) = \mathbf{1} - \vec{\nabla}\vec{\nabla}\psi(\vec{\theta}), \quad (17)$$

a 2D tensor. (The bold-face  $\mathbf{1}$  denotes an isotropic tensor—identity matrix in component notation.) Meanwhile, taking the gradient of the lens equation (16) gives

$$\vec{\nabla}\vec{\beta} = \mathbf{1} - \vec{\nabla}\vec{\nabla}\psi(\vec{\theta}). \quad (18)$$

The curious term  $\vec{\nabla}\vec{\beta}$  expresses how much source-plane displacement is needed to produce a given small image displacement; i.e., the inverse of magnification<sup>3</sup>. Equation (18) tells us that magnification is a 2D tensor, and depends on  $\vec{\theta}$  but not  $\vec{\beta}$ ; let us write magnification as  $\mathbf{M}$ . Comparing the last two equations we have

$$\mathbf{M}^{-1} = \vec{\nabla}\vec{\nabla}\tau(\vec{\theta}). \quad (19)$$

Equation (19) means that the curvature of the arrival-time surface is the inverse of the magnification. Thus, broad low hills and shallow valleys in the arrival-time surface correspond to highly magnified images; needle-sharp peaks or troughs correspond to images demagnified into unobservability.

By curvature, we mean a tensor curvature, which depends on directions:  $\mathbf{M}$  and  $\mathbf{M}^{-1}$  are symmetric 2D tensors, so their components form  $2 \times 2$  matrices. In particular, we have

$$\mathbf{M}^{-1} = \begin{pmatrix} 1 - \partial^2\psi/\partial\theta_x^2 & -\partial^2\psi/\partial\theta_x\partial\theta_y \\ -\partial^2\psi/\partial\theta_x\partial\theta_y & 1 - \partial^2\psi/\partial\theta_y^2 \end{pmatrix}. \quad (20)$$

Comparing equations (11) and (19 or 20) we see that the trace of  $\mathbf{M}^{-1}$  must be  $2(1 - \kappa)$ . Thus  $\kappa$ , originally defined as the surface density in suitable units, also has the interpretation of an isotropic magnification. Accordingly,  $\kappa$  is known as the convergence. The traceless part of  $\mathbf{M}^{-1}$  is called the shear and its magnitude is denoted by  $\gamma$ ; it changes the shape of an image but not its size. In full, we have

$$\mathbf{M}^{-1} = (1 - \kappa) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \gamma \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix} \quad (21)$$

where  $\phi$  denotes the direction of the shear. Note that any symmetric  $2 \times 2$  matrix can be written in the form (21). All we have done here is interpret  $\kappa$  and  $\gamma$ .

The determinant

$$|\mathbf{M}| = [(1 - \kappa)^2 - \gamma^2]^{-1} \quad (22)$$

defines a scalar magnification, or ratio of image area to source area for an infinitesimal source.

Surface brightness is conserved by lensing. Although we will not prove it here, this is a consequence of the fact that the lens equation is a gradient map. Magnification changes only angular sizes and shapes on the sky. Thus a constant surface brightness sheet stays a constant brightness sheet when lensed. (Were this not the case, the microwave background would get wildly lensed by large scale structure.) However, an unresolved source will have its brightness amplified according to (22).

---

<sup>3</sup> An alternative notation,  $\partial\vec{\beta}/\partial\vec{\theta}$ , reminds one of this physical interpretation.

**Saddle-point contours, critical curves, caustics** The equations (11) for the arrival-time surface, (16) for the image positions, and (19) for the magnification are elegant, but they do not give us much intuition for the shape of the arrival-time surface, the possible locations of images, and the likely magnification in real systems that we might observe. To gain some intuition, it is very useful to introduce [16] three special curves in the image and source planes.

Consider the arrival-time surface and contours of constant  $\tau$ . In the absence of lensing  $\tau(\vec{\theta})$  is a parabola, and the image is at its minimum, or  $\vec{\theta} = \vec{\beta}$ . For a small lensing mass, the shape changes slightly from being a parabola and the minimum moves a little. But for large enough mass, a qualitative change occurs, in that a contour becomes self-crossing. There are two ways in which a self-crossing can develop: as a kink on the outside of a contour line, or a kink on the inside. These are illustrated in Figure 3. The outer-kink type is a lemniscate and the inner-kink type a limaçon. With the original contour having enclosed a minimum, a lemniscate produces another minimum, plus a saddle-point at the self-crossing, while a limaçon produces a new maximum plus a saddle point. (The previous sentence remains valid if we interchange the words ‘maximum’ and ‘minimum’.) The process of contour self-crossing can then repeat around any of the new maxima and minima, producing more and more new images, but always satisfying

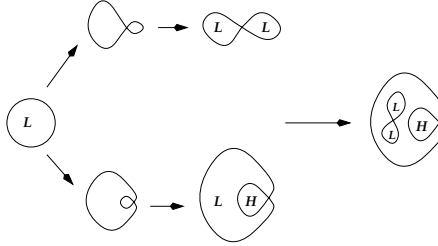
$$\text{maxima} + \text{minima} = \text{saddle points} + 1. \quad (23)$$

The self-crossing or saddle-point contours form a sort of skeleton for the multiple-image system. Lensed quasars characteristically have one of two configurations: double quasars have a limaçon, while quads have a lemniscate inside a limaçon, as in the rightmost part of Figure 3. Both cases have one maximum (marked ‘H’ in the figure), which will be located at the center of the lensing galaxy. Since galaxies tend to have sharply-peaked central densities, the arrival-time surface at the maximum will be sharply peaked as well; the corresponding image is highly demagnified and (almost) always unobservable. Thus lensed quasars are doubles or quads: an incipient third or fifth image hides at the center of the lensing galaxy.

Critical curves are curves on the image plane where the magnification is infinite. More formally, they are curves where  $\mathbf{M}^{-1}$  has a zero eigenvalue. From the definition (19) it is clear that at minima of  $\tau$  both eigenvalues of  $\mathbf{M}^{-1}$  will be positive, at maxima both eigenvalues will be negative, and at saddle points one eigenvalue will be positive and one negative. Thus critical curves separate regions of the image plane that allow minima, saddle points, and maxima.

If we map critical curves to the source plane via the lens equation (16) we get caustic curves. Caustics separate regions on the source plane that give rise to different numbers of images.

We discuss examples of saddle-point contours, critical curves, and caustics in the next section.



**Fig. 3.** Multiple images via saddle-point contours in the arrival-time surface. Here  $L$  marks minima and  $H$  marks maxima.

## 1.2 An illustrative macro-model

We have already met the point lens, which in dimensionless form has lens potential

$$\psi(\vec{\theta}) = \theta_E \ln \theta \quad (24)$$

where  $\theta_E$  is effectively a parameter expressing the total mass. Solving the lens equation, we see that images are at

$$\vec{\theta} = \frac{1}{2} \left( \beta \pm \sqrt{\beta^2 + 4\theta_E^2} \right) \hat{\beta}, \quad (25)$$

where  $\hat{\beta}$  denotes a unit vector in the direction of  $\vec{\beta}$ . The scalar magnification is given by

$$|\mathbf{M}|^{-1} = 1 - \frac{\theta_E^4}{\theta^4}. \quad (26)$$

Another commonly used lens is the isothermal lens (so called because of its relation to isothermal spheres in stellar dynamics, and a good zeroth order model for disk-galaxy halos and giant ellipticals — more on this subject in the modeling section); it has  $\kappa(\vec{\theta}) = \frac{1}{2}\theta_E/\theta$  and lens potential

$$\psi(\vec{\theta}) = \theta_E \theta, \quad (27)$$

For  $\beta < \theta$  there are two images at

$$\vec{\theta} = \vec{\beta} + \theta_E \hat{\beta}, \quad \vec{\theta} = \vec{\beta} - \theta_E \hat{\beta} \quad (28)$$

and for  $\beta > \theta$  the second of these disappears. The constant image-separation in equation (28) is a peculiar feature of the isothermal. The scalar magnification is given by

$$|\mathbf{M}|^{-1} = 1 - \frac{\theta_E}{\theta}. \quad (29)$$

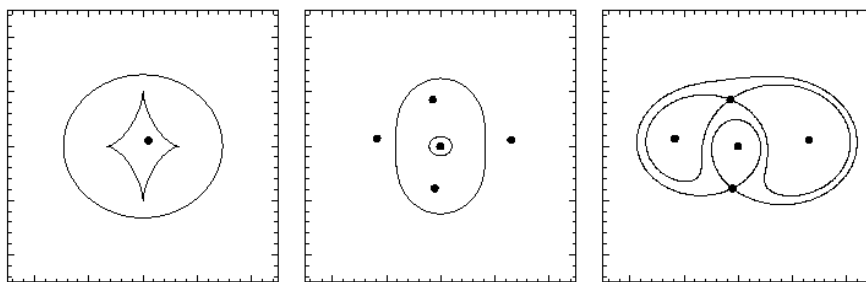
Lacking any ellipticity, these lenses by themselves cannot produce quads. But with some added ellipticity, quads and indeed all the main qualitative features of quasar lenses can be reproduced, as we now show.



As an example, consider the potential

$$\psi(\vec{\theta}) = (\theta^2 + \epsilon^2)^{\frac{1}{2}} + \frac{1}{2}\gamma\theta^2 \cos(2\phi) \quad (30)$$

where  $\phi$  is the polar angle and  $\epsilon$  and  $\gamma$  are adjustable parameters;  $\epsilon$  gives the isothermal a non-singular core, and  $\gamma > 0$  contributes ‘external shear’ which in this case amounts to extra lensing mass outside the lens in the  $y$  direction. We take  $\epsilon = 0.1$  and  $\gamma = 0.2$ , and then examine what happens for different source positions, through caustics, critical curves, and saddle-point contours. A similar potential, but with the scale and shear orientation adjustable, will be used later (cf. equation 44) to fit data on observed systems.



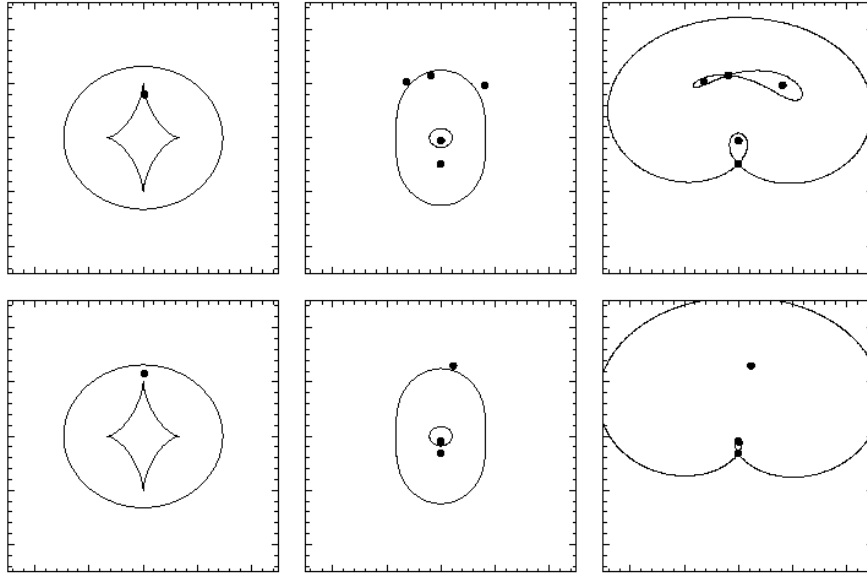
**Fig. 4.** A central quad: one with source near the center. **Left panel:** source positions and caustics; **middle panel:** image positions and critical curves; **right panel:** image positions and saddle-point contours. In this figure, and in Figures 5 to 7, the left hand panels (showing the source plane) have a scale half that of the other panels (showing the image plane).

Figure 4 shows the situation with the source close to the center. The left panel shows what is happening in the source plane, while the middle and right panel show what is happening in the image plane. Several interesting things may be seen.

- The two caustic curves in the source plane (left panel) demarcate regions from where a source produces 1, 3, and 5 images. In this case the source is well within the inner caustic, and that results in five images. The other panels shows these five images, along with the critical curves (middle panel) or the saddle-point contours (right panel). But the image near the center is highly demagnified, and observationally such a system would be a quad. Let us call it a ‘central quad’, to distinguish it from other quads we will see below.
- The two critical lines are maps of the caustics to the image plane, but the *inner* caustic maps to the *outer* critical line. (Also, the long axes of both of these are aligned with the potential.) Recall that critical lines

are where an eigenvalue of  $\mathbf{M}$  changes sign. The consequence for this lens is that any image outside both critical curves is a minimum, any image between the critical curves is a saddle point, and any image inside both critical curves is a maximum, all irrespective of the source position. For the current source position we can verify these statements by comparing the middle and right panels.

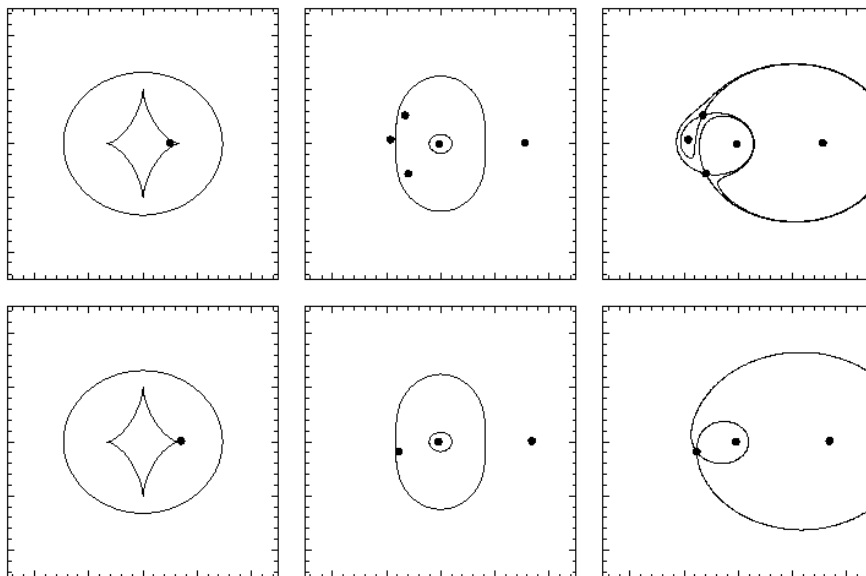
- The time-ordering of the quad's images is evident from the saddle-point contours—compare with Figure 8.
- The arrival-time contours and the arrangement of the images appear to be squeezed in the  $y$  direction. Such squeezing is characteristic along the long axis of the potential, and the images appear pop out along the short axis of the potential.



**Fig. 5.** A long axis quad and double. Note how, as the source crosses the diamond caustic, two images merge on the tangential critical line and then disappear.

Figure 5 shows the situation with the source is displaced along the long axis of the potential. As the source nears the inner caustic curve, two of the images approach the outer critical curve. We call this configuration a long-axis quad. Two minima and a saddle point are fairly close together, displaced in the same direction as the source, while another saddle point is on the opposite side of the lens center. This 3+1 image arrangement reveals the direction of the source displacement. Meanwhile, as with the core quad,

the arrangement of the images is squeezed along the direction of the long axis of the potential. As the source crosses the inner caustic curve, a minimum and a saddle point merge on the outer critical curve, and then disappear. The system is now a double, which we may call a long-axis double.

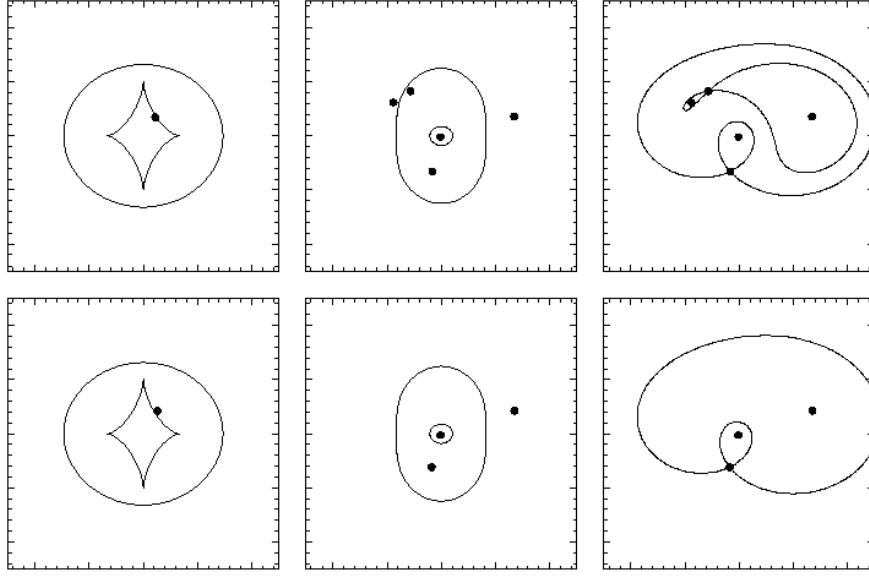


**Fig. 6.** A short axis quad and double.

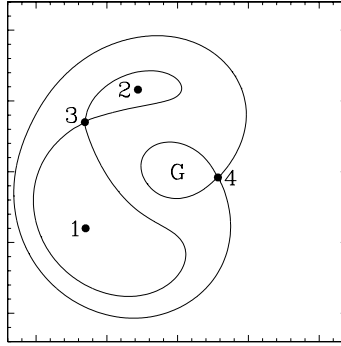
Figure 6 has the source displaced along the short axis of the potential, producing configurations we call a short-axis quad and a short-axis double. The morphology of a short-axis quad resembles that of the long axis quad, but one can tell them apart. First, one of the four images is far from the others, but this time it is a minimum, not a saddle point. Secondly, the 1+3 image arrangement indicates direction of the source displacement, and it is perpendicular to the long axis of the potential which can be inferred from the squeezing of the image arrangement. Moving the source outside the inner caustic again causes two images to merge, leaving a short-axis double. The morphology of a short axis-double is the same as that of a long-axis double.

Figure 7 has the source displaced obliquely to the potential, producing what we call an inclined quad and an inclined double. These are more common than the long and short-axis types, and easily distinguished because of their asymmetry.

Examining the saddle-point contours in Figures 4 to 7, the order of arrival times of the images is nearly always evident. We can summarize image-ordering in quads in the following simple rules, illustrated in Figure 8: (i) Im-



**Fig. 7.** An inclined quad and double.



**Fig. 8.** Saddle point contours in a generic quad. Images 1 and 2 are minima, 3 and 4 are saddle points; the fifth image would be a maximum at the galaxy's centre G.

ages 1 and 2 are opposite in Position Angle (PA), (ii) 3 and 4 are opposite in PA, (iii) 1 is the furthest or nearly the furthest from the lens centre, (iv) 4 is the furthest or nearly the furthest from the lens centre, (v) if there are a nearly merging pair, they are 2 and 3. For some cases it is not possible to decide between 1 and 2, but otherwise there is never an ambiguity. For doubles, time ordering is trivial: the image further from the galaxy arrives first.

With a little practice, it is easy to sketch the saddle-point contours (including image ordering), and from there the critical curves and caustics, of any quasar lens just from the morphology.

We may summarize the conclusions of this section as follows:

- From the morphology of a quad, it may be immediately recognized as one of (i) central, (ii) long axis, (iii) short axis, or (iv) inclined; doubles may be recognized as (i) axis, or (ii) inclined, but long and short axis doubles need more information to distinguish. The ‘axis’ in each case is of course the axis of the potential, including any external shear; so morphology already gives some idea mass distribution.
- Morphology of quads or doubles also reveals the time-ordering of images.

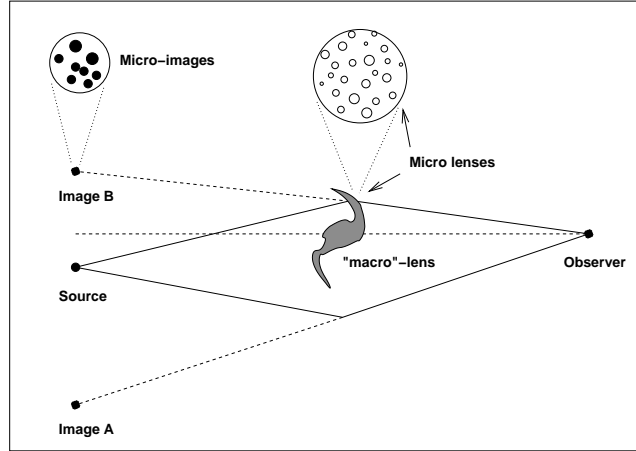
### 1.3 Lenses within lenses: microlensing

Stars comprise an appreciable fraction of the mass in lensing galaxies. These stars produce small scale fluctuations in lens potentials which will be seen to have substantial effects on the magnifications of quasar images.

**Random star fields** Suppose that light from a source passes through a screen of  $N$  equal point masses with random positions, and that the Einstein rings of individual masses are very small compared to the mean spacing between them. The delay equation (3) consists of a single geometric term and a great many Shapiro terms. Each Shapiro term produces three stationary points: a singular and completely demagnified maximum at the angular position of the point mass, a significantly demagnified saddle point close to the point mass and on the side opposite the source position, and a minimum not far from the unperturbed source position. The minimum will be the only non-negligible image.

Were the same stars to increase in mass without changing position, the saddle points would move further away from the stars, increasing in brightness. As their masses continue to increase, close pairs will create new saddles between them. For each new ridge, a new valley will form on the side furthest from the source. Just after the formation of this new pair of images, the curvature along the line connecting them is very small and they are very highly magnified. As the masses continue to increase, the images separate and grow fainter, though the new minimum will never be fainter than the unmagnified source.

Thus the number of images increases from  $N + 1$  to  $N + 3$  to  $N + 5$  and so forth. If we lacked the resolution to see the individual images, but only the combined light, we would find that for the most part the combined brightness increases steadily, but with bright flashes as new pairs of images are created. At any time our star field would have some average surface density and an associated dimensionless convergence,  $\kappa$ . For an ensemble of such sources placed randomly behind such a screen, we would expect an average scalar



**Fig. 9.** Schematic representation of microlensing by stars in a doubly imaged system. In this example, the unresolved “sea” of stars in the main lensing galaxy is responsible for “microlensing” of one of the quasar images.

magnification of  $(1 - \kappa)^{-2}$  (see equation 22), but there would be fluctuations depending upon the accidents of source position. Additional images begin to appear (in the absence of external shear) when  $\kappa$  approaches unity.

The general phenomenon of the amplification of unresolved images by stars (or other point masses) in intervening galaxies is termed microlensing. The situation is illustrated schematically in Figure (9). The large numbers of highly demagnified saddle points are not shown.

**Mandatory microlensing** In the thought experiment of the preceding subsection, additional positive parity images (minima) and their accompanying saddle points formed when  $\kappa$  approached unity. The average density interior to the Einstein ring of an isolated microlens is just the critical density, with  $\kappa \equiv 1$ . The criterion for substantial microlensing is therefore  $\kappa \sim 1$ .

Now let us suppose that the galaxy lensing a multiply imaged quasar is comprised entirely of point masses. The average surface density interior to the galaxy’s Einstein ring is exactly the critical density. Unless the galaxy is very highly concentrated, the surface density at the Einstein ring must be a substantial fraction of the critical density – one half in the case of an isothermal lens. The covering factor of the microlenses’ Einstein rings must therefore be a substantial fraction of unity. Thus microlensing must be important, *if* the galaxy is comprised entirely of point-like objects.

Microlensing will be important only if the Einstein rings of the particles comprising the galaxy are larger than the projection of the source onto the sky. There are two ways in which this might fail to occur for lensed quasars. First, the source might be large compared to the the Einstein rings of the

galaxy's stars. Second, most of the mass in the galaxy might be in particles with masses very much smaller than that of a star, as we suspect would be the case for dark matter. Our understanding of quasar sources and the distribution of dark matter within galaxies is as yet so limited that we cannot say with certainty whether microlensing should or should not be expected. As we shall see later there is considerable observational evidence that the conditions for microlensing are met in at least some lensed quasars.

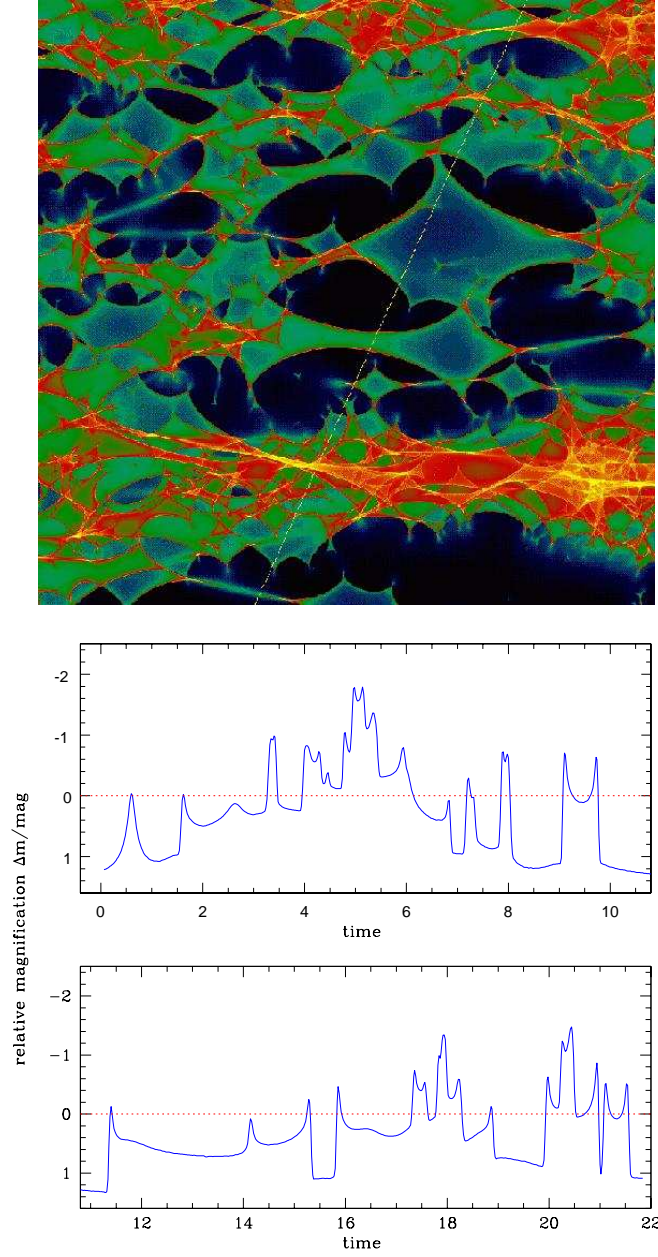
It should be noted, by contrast, microlensing of sources in the Magellanic Clouds by stars (and dark objects) in the halo of our Milky Way is an exceedingly rare event. The covering factor for halo object Einstein rings is at most  $10^{-6}$ . The largest source of this difference is the small distance to the Clouds and the correspondingly large value of  $\Sigma_{\text{crit}}$ .

**Static and kinetic microlensing** In the above *Gedanken* experiment neither the source nor stars were moving. Imagine a symmetric lens which forms two quasar images exactly opposite each other. The images pass through regions of identical surface density and shear, and would, in the absence of microlensing, undergo the same magnification. But since they pass through different random star fields, they suffer different amounts of microlensing. The magnifications predicted from the global galaxy potential would be only approximate – one would have to take into account the local fluctuations. Static microlensing produces “errors” in the predicted fluxes.

Imagine further that the quasar consists of two components, one smaller than the typical size of stellar Einstein rings and the other larger. The smaller component would be microlensed but the larger component would not.

The motions of the source and the microlensing pattern add an additional complication. Taking the microlens positions to be fixed, as the source moves the microlensing will change. To order of magnitude, the source must move an amount equal to the Einstein radius of the microlens to produce a substantial change. If the stars are moving, they must move an amount comparable to the sizes of their Einstein radii to produce substantial changes. The temporal changes in the brightness of an unresolved source are the result of such kinetic microlensing.

**Microlensing caustics** As described in Section 1, critical curves are the locii in the image plane along which pairs of images merge or are created as one varies the position of a background source. The scalar magnification is infinite along the critical curves. This property suggests a relatively straightforward computational scheme for identifying caustics, which are the locii in the source plane which produce images on the critical curves. Given a set of (random) microlens positions, one projects rays back from the observer uniformly in solid angle. These land with high (low) surface density in regions of high (low) magnification and the caustics readily emerge when one plots a “spot diagram” for these rays. Such a plot also allows rapid computation of kinetic microlensing light curves for a moving source – one simply takes



**Fig. 10. Top:** network of micro-caustics in a lensing galaxy. The local convergence  $\kappa$  is 0.5 and the shear  $\gamma$  is 0.6. The bright regions correspond to high magnification while the darker ones show de-magnification. **Bottom:** predicted light curve when a source crosses the caustics along the straight line in the top panel. The time scale is arbitrary (Figure courtesy Joachim Wambsganss).



linear cuts through the source plane spot diagram. The magnification is proportional to the local density of spots. Figure 10 shows such a plot, with a predicted light curve when a source crosses a network of caustics.

**Quantitative microlensing** Microlensing is fundamentally statistical in nature. It has been surprisingly resistant to analytic techniques, and most quantitative work has been carried out via simulations. These have shown [134,73] that fluctuations of a magnitude or more are possible for highly magnified images. Moreover saddlepoints behave differently from minima, with larger fluctuations for the former than for the latter [146,117]. Among the few interesting analytic results are an exact expression for the magnification probability distribution at high magnification [121], and an expression for the mean number of positive parity microimages (minima) as a function of  $\kappa$ .

#### 1.4 The effect of cosmology

The main observables in lensing, image positions and magnifications, are all dimensionless; only time delays are dimensional. The effect of cosmology is to set the scale of time delays, and we can think of it as setting  $T_0$ , the time scale in equation (9). Cosmology really enters only through the angular-diameter distances, so fixing  $T_0$  also fixes the other important scale,  $\Sigma_{\text{crit}}$ .

The time scale has a dependence of the form

$$T_0 = h^{-1} z_L (1 + z_L) \times \langle \text{weak function of } z_L, z_S, \Omega_0, \Omega_\Lambda \rangle \quad (31)$$

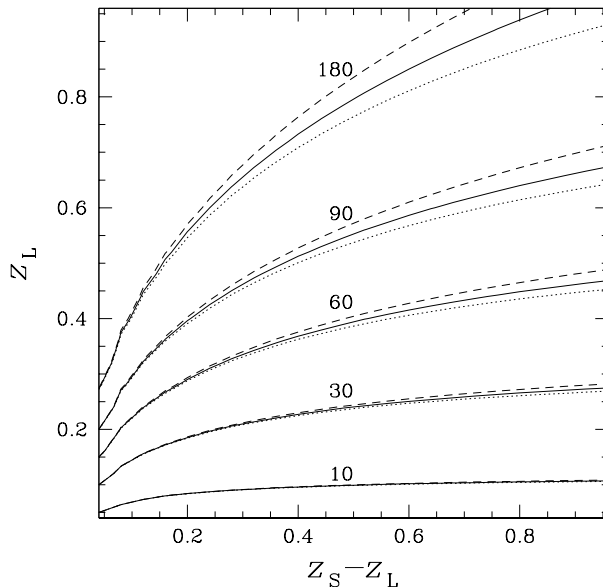
and is analytic [37] but messy, so we do not reproduce it here. Instead we illustrate it in Figure 11 for some cosmologies. It is worth remarking that

- $T_0 \propto h^{-1}$  exactly;
- for  $z_S \gg z_L$  the approximation (14) applies.
- for the same  $h$ , an Einstein-de-Sitter cosmology gives large  $T_0$ , an open cosmology gives small  $T_0$ , with the currently favored flat  $\Lambda$ -cosmology being intermediate; but the differences are small.

The simple dependence on  $h$  make it attractive to use time delays to try and measure  $h$ . One can even imagine putting in several time delays on a sort of Hubble diagram to try and constrain  $\Omega_0, \Omega_\Lambda$ . Both these ideas are due to Refsdal [104,105].

#### 1.5 Degeneracies

Lensed images correspond to minima, saddle points, and maxima of the arrival-time surface; the rest of the arrival-time surface is unobservable. Thus, lensing observables do not uniquely specify a lens; another lens that preserves  $\tau(\vec{\theta})$  and its derivatives at image positions but changes them elsewhere will



**Fig. 11.** Contour plots of  $T_0$  as a function of  $z_S - z_L$  and  $z_L$ . The labels are in units of  $h^{-1} \text{days arcsec}^{-2}$ . The dashed curves are for  $\Omega_0 = 1, \Omega_A = 0$ , the solid curves are for  $\Omega_0 = 0.3, \Omega_A = 0.7$ , and the dotted curves are for  $\Omega_0 = 0.1, \Omega_A = 0$ .

produce exactly the same lensing data. In this sense, lenses are subject to degeneracies.

An example, which we have already used when deriving the critical density, is the monopole degeneracy: any circularly symmetric redistribution of mass inwards of all observed images, and any circularly symmetric change in mass outside all observed images will change  $\tau$  by at most an irrelevant constant in the image region, and hence have no effect on lensing observables. This means in particular that doubles and quads contain no information about the monopole part of the interior mass distribution, though they constrain the total mass enclosed. So in the example in Section 1.2 our choice of core radius was irrelevant; it specified the location of the inner critical curve and the outer caustic, but those played no part since images and sources never went near them.

In addition to degeneracies of the above type, which all involve localized changes in the arrival-time surface, there is one special degeneracy which is particularly serious: the mass disk degeneracy [32,93,122,112]. In this the  $\tau$  scale of the whole arrival-time surface gets stretched or shrunk. To derive it we rewrite equation (11) first discarding a  $\frac{1}{2}\bar{\beta}^2$  term since it is constant over

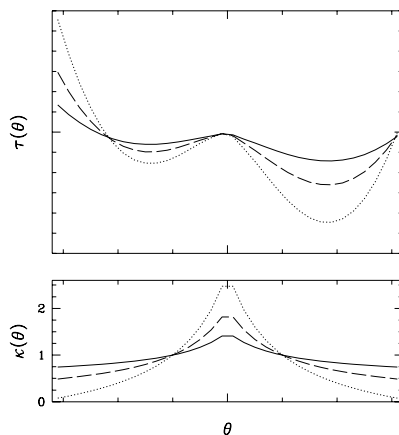
the arrival-time surface, and then using  $\nabla^2 \vec{\theta}^2 = 4$ , to get

$$\tau(\vec{\theta}) = 2\nabla^{-2}(1 - \kappa) - \vec{\theta} \cdot \vec{\beta}. \quad (32)$$

Now the transformation

$$1 - \kappa \rightarrow s(1 - \kappa), \quad \vec{\beta} \rightarrow s\vec{\beta}. \quad (33)$$

where  $s$  is a constant which just rescales time delays while keeping the image structure the same; but since the source plane is rescaled by  $s$  all magnifications are scaled by  $1/s$ , leaving relative magnifications unchanged. The effect on the lens is to make it more like or less like a disk with  $\kappa = 1$ . Figure 12 illustrates. Note that in (33)  $s$  can become arbitrarily small; it can not become arbitrarily large because otherwise  $\kappa$  will become negative somewhere in the image region. (Negative  $\kappa$  *outside* the image region can always be avoided by adding an external monopole).



**Fig. 12.** Illustration of the mass disk degeneracy, showing the surface density (lower panel) and the arrival time (upper panel) for three circular lenses. The units, except for  $\kappa$ , are arbitrary. The arrival time indicates a saddle point (which looks like a local minimum in this cut), a maximum, and a minimum. The dashed curves correspond to a non-singular isothermal lens. Stretching the time scale amounts to making lens profile steeper (dotted curves) and shrinking the time scale amounts to making the lens profile shallower (solid curves).

From the modelers point of view, the mass disk degeneracy is a degeneracy in the central concentration of the lens, or the steepness of the radial profile, and we will meet this many-headed monster again in the modeling section. An easy way of remembering its effect is “the lens gets steeper as the universe gets smaller”. The lensing data stay exactly the same, and the mass inside an

Einstein radius is unaffected [112], but the sources before lensing get larger and brighter, and  $h$  gets bigger. Which reminds us that this degeneracy is particularly inimical to measuring  $h$  from lensed quasars, where it dominates the uncertainty. In principle it could be broken in various ways: if the intrinsic brightness of sources were known, or if sources at very different redshifts were lensed by the same lens [2], or indeed if  $h$  were known from some other method. But there seem no immediate prospects for any of these.

Another kind of degeneracy is associated with a non-lensing observable that is often observed in connection with lensing, velocity dispersion. Lenses follow an approximate relation between Einstein radius (or some surrogate for it in non-circular lenses such as the size of the outer critical curve) and the line-of-sight velocity dispersion:

$$\theta_E \simeq 2'' \times \frac{\langle v_{\text{los}}^2 \rangle}{(300 \text{ km s}^{-1})^2}. \quad (34)$$

To see why there should be such a relation, we rewrite the expression (6) for the Einstein radius of a circular mass distribution as

$$\frac{GM}{\theta_E D_L} = \frac{c^2}{4} \frac{D_S}{D_{LS}} \theta_E. \quad (35)$$

Now, the left hand side in (35) will be of order  $\langle v_{\text{los}}^2 \rangle$  because of the virial theorem, leading to (34). The trouble is that the relation (34) cannot be made more precise, because the exact coefficient that would go into it depends on the mass distribution in a very complicated way. In general, more centrally concentrated mass distributions would give larger velocity dispersions. On the other hand, an isothermal sphere in stellar dynamics gives  $(3\pi/2)\langle v_{\text{los}}^2 \rangle$  for the left hand side in (35) while a barely compact homogeneous sphere gives  $5\langle v_{\text{los}}^2 \rangle$  — almost the same number despite the very different mass profile.

## 2 Observations

### 2.1 Historical background

While the concept of light deflection by massive bodies was already proposed by Isaac Newton in the 18th century [89], the astrophysical and cosmological potential of the phenomenon was, with notable exceptions, taken seriously only after discovery in of the first multiply imaged quasar by Walsh, Carswell & Weymann [132]. The observation of two well separated images of the same source at  $z = 1.41$  not only confirmed the existence of what had previously been seen largely as a theoretical curiosity, but also established gravitational lensing as a new field of astrophysics. Indeed, the existence of even a single lensed quasar, lent considerable hope to the application of Refsdal's method [103,104] for determining the Hubble parameter  $H_0$ . Proposed in 1964, the method is based on the measurement of the light variations in the lensed

images of a distant source. The time lag, or so-called “time-delay” between the arrival times of the luminous signal from each image of the source to the observer, is directly related to  $H_0$  and to the mass distribution in the lensing object. Measuring the time-delay therefore provides us, via a mass model for the lensing galaxy, with an estimate of  $H_0$ . Refsdal originally proposed to apply his method to distant supernovae. The discovery of quasars by Schmidt [120] offered new prospects in using even more distant light sources.

Measuring time-delays is far from trivial: the angular separations between the lensed images are usually small, typically 1-2 arcsec, and not all quasars are willing to show measurable photometric variations. In addition, characterizing the mass distribution responsible for the lensing effect, assuming the lensing galaxy is detected at all, was very challenging at the time of the first discoveries. CCD detectors, were only just coming into use. They were hard to obtain and had small formats and high read noise. The uncontrolled thermal environments of telescopes produced mediocre seeing, typically larger than the angular separations observed in most presently known objects (see Tables at the end of this chapter). The Hubble Space Telescope (HST) was more than a decade off in the future. Despite these difficulties, searching for new systems suitable for cosmological investigations became a major activity in the early eighties. Based on the argument that some of the brightest quasars might be magnified versions of a lower luminosity object (e.g., Sanitt [114]), systematic searches for new multiply imaged sources were undertaken among the apparently brightest quasars. These searches yielded the discovery of more doubles, like UM 673 [125], but also new image configurations. PG 1115+080 [137] was thought to be triple but turned out to be an off-axis quadruple with higher resolution observations [44]. More symmetric quadruples, such as the “cloverleaf” [77], were also found. Almost simultaneously, radio searches yielded their first results. As the radio emitting regions of quasars are larger than the optical ones, lensed radio loud quasars were often found to be complete Einstein rings: MG 1131+0416, MG 1654+1346, PKS 1830-211 [45,66,124]. The observation of complete or partial rings offers more constraints than 2 or even 4 point source images and led to the development of more accurate models [59].

During this same period, systematic campaigns were initiated to measure time delays, much of it concentrated on the first lens discovered: Q 0957+561. Early reports [118], [38] gave contradictory results. Vanderriest et al. [131] and Schild et al. [119] derived a value of 415 days, from ground based optical observations. Press et al. [100] reanalyzed Vanderriest’s data and published a very different time delay: 536 days, a value supported by the radio monitoring results obtained at the Very Large Array (VLA) [107]. The dictum attributed to Rutherford [8], “If your experiment needs statistics, you ought to have done a better experiment,” appears to have been borne out. Improved optical [65,91] and radio monitorings [42] have finally settled the issue. They reconcile the optical and radio time delays and lead to the value of  $\Delta t = 417 \pm 3$  days.

The controversy over Q 0957+561 reflects the difficulty of measuring time delays. Quasars do not commonly show very sharp light variations, and their light curves are often corrupted by the erratic photometric variations induced by microlenses (stars) in the lensing galaxy (see Chapter 1 and Section 2.3 of the present Chapter). Photometric monitoring over a period considerably longer than the time delay is therefore necessary. Temporal sampling must also be sufficiently frequent to average out short timescale microlensing variations. Microlensing may corrupt quasar light curves but it is of considerable interest for constraining the statistical mass of MACHOs (see Chapter 1) in the lens [135] and the size of the lensed source [133]. With particularly good data obtained over a wide wavelength range, it might even be possible to reconstruct some of the quasar's accretion disk parameters, such as, size, inclination and details of the spectral energy distribution of the accretion disk as a function of distance from the AGN's center [3,149].

Progress with CCD detectors, with radio interferometers and with image processing techniques has made it possible to overcome at least some of the observational limitations on time delay measurements. The list of systems with known time delays is rapidly growing, with optical and radio time delays both available in some cases. Schechter et al. [115] obtained optical light curves for PG 1115+080 and two time delays between two images and the group of blended bright images A1+A2. Three time delays have been measured from radio VLA observation, in the quadruply imaged quasar CLASS B1608+656 [34]. The two bright radio doubles PKS 1830-211 and B 0218+357 are two other cases with known time delays (e.g., [76,14]). Note also the lucky case of B 1600+434 which has both optical [20] and radio [64] time delays and even overlapping light curves. Many more time delays have recently been obtained at the Nordic Optical Telescope or at ESO [22,23,48].

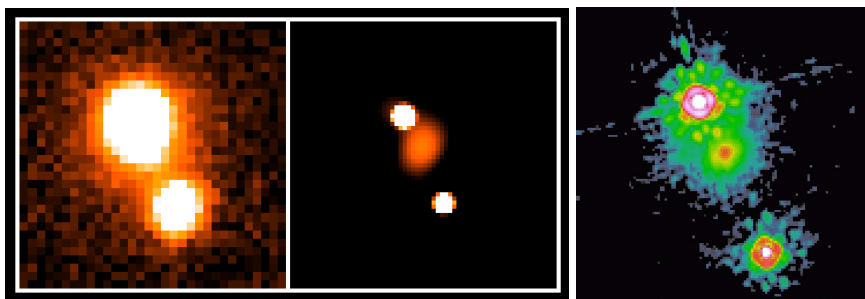
The level of interest in lensed quasars has followed a more or less predictable course. The considerable excitement following what was effectively the birth of the field in 1979 was followed by extraordinary growth, as measured by the number of papers published [127] and number of lenses known [61]. The phenomenon is no longer so novel, it is entering a more mature, and astrophysically and cosmologically, more productive stage. Observational and theoretical advances have proceeded in parallel, with considerable improvement in "best" estimates of  $H_0$ , in weighing distant galaxies, and in probing their stellar and dust content. And as with other areas of astrophysics, there is an increasing tendency toward large, international teams, marking the substantial demands of the enterprise.

## 2.2 Observational constraints in quasar lensing

Given the small deflection angles involved in multiply imaged quasars, high resolution observations are required to measure accurately the main observables: the position of the quasar images and of the lens, the time-delay, and the magnification at the position of the images.

**The image configuration and the time delay:** For most known lens systems it has become relatively straightforward nowadays to obtain astrometry of the requisite precision, especially when HST images are available. However, adequate temporal sampling is also required as soon as the goal is to measure the time delay. Photometric monitoring with  $0.1''$  resolution would be possible with HST but one could not realistically expect the large numbers of orbits necessary. Until recently, such work was restricted to radio wavelengths, less affected by weather conditions than optical ones and providing data with higher resolution on a more regular basis. Scheduling is also facilitated as one can observe in day time. But as only 10% of quasars are radio loud, this restricts the available sample of lenses.

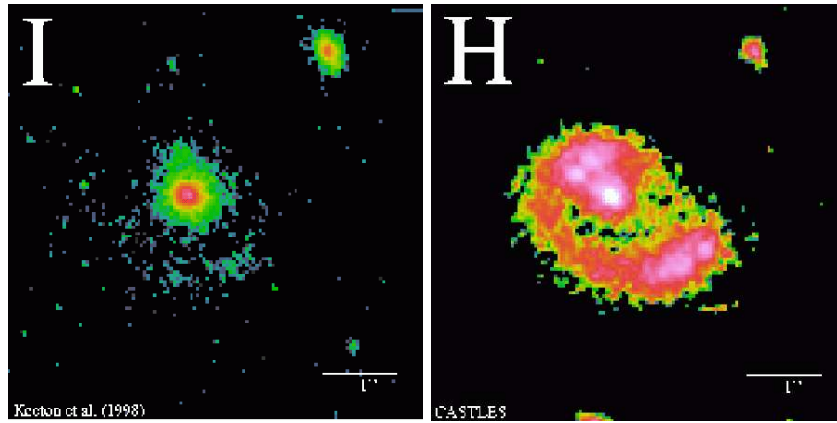
Recent advances in image processing techniques have extended the range of ground-based imaging into the subarcsecond regime. Typically reconstruction and deconvolution techniques can only be used with relatively high signal-to-noise observations, but in such cases improvements of a factor of two in resolution have been possible [79].



**Fig. 13.** Two ways of obtaining high resolution images of lensed quasars. From left to right: near-IR ground based image of HE 1104-1805. It has been obtained with the ESO/MPI 2.2m telescope in the  $J$ -band under average seeing conditions ( $0.7''$ ). Its resolution has been improved down to  $0.27''$  on the deconvolved version of the data displayed in the middle panel. The lensing galaxy is obvious, between the QSO images. Its position and elongated shape oriented with a PA of about 30 degree are confirmed by the  $H$ -band HST/NICMOS image shown on the right (HST image from the CASTLEs survey).

Figure 13 shows an example of high resolution data of the doubly imaged quasar HE 1104-1805 [139], as might be obtained either with the HST (here, in the near-IR) or from post-processed (deconvolved) ground based images. The data presented in this figure are sufficient to infer the image and lens positions with an accuracy from a few milli-arcsec (quasar images) to a few tens of milli-arcsec (lensing galaxy). In fact, the combination of space observatory data [72] and post-processed ground based data now allow for accurate photometric monitoring in the optical and for detailed modeling.

The range in properties of lens systems is such that there is no single factor which consistently limits one's ability to carry out a determination of  $H_0$  through time delay measurement. It seems that there is no “golden lens,” no ideal case that will give a “best” measurement of  $H_0$ . In some cases, the error on the time delay dominates (for example HE 2149-2745 [22]), while in other systems more symmetric about the center of the lensing galaxy, the errors introduced by the astrometry of the quasar images will dominate [115,56,28,50]. In still other cases the erratic variations of the light curves introduced by microlensing events in the lensing galaxy are the main source of error [141,20]. It therefore seems more reasonable to monitor as many systems as possible rather than trying to concentrate on a particular one which might have its own unknown sources of systematic errors.

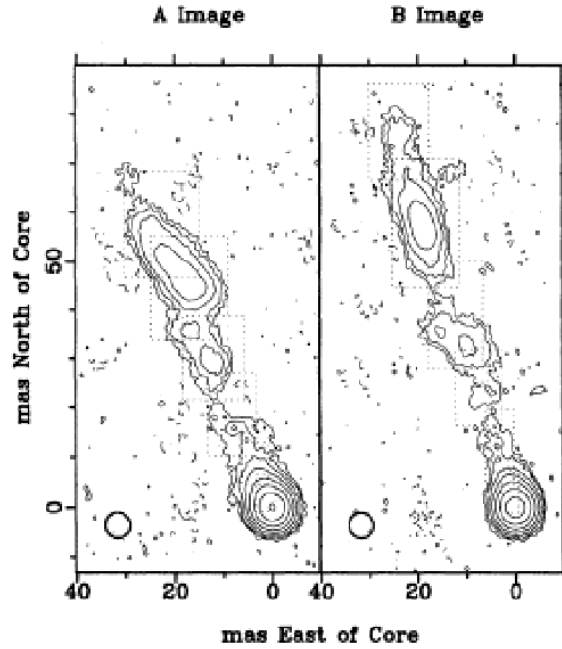


**Fig. 14.** The lensed radio source MG 1131+0456 is a nice example of system with the source only visible in the near-IR and longer wavelengths. On these HST images obtained by the CASTLES group, only the lens is visible (left) in the optical *I*-band. On the *H*-band image (right), the source is seen as a almost full Einstein ring.

**Distances to the source and lens** As seen from equation (3), modeling lensed quasars requires knowledge of the distance  $D_L$  to the lens, and of the distance  $D_S$  to the source. While the lensing galaxies are not especially faint by current standards, measuring their redshifts is non-trivial. In the optical, the background quasar is often bright and hides the much fainter lensing galaxy. In some lucky cases, the lensing galaxy shows emission lines in superposition on the quasar spectrum [129], but this is not the rule. To date, no HST spectrum has been taken of a lensing galaxy, but application of deconvolution techniques [29] to spectra obtained on ground based 10m class telescopes have proved useful and have yielded the measurement of several lens redshifts [75,22].

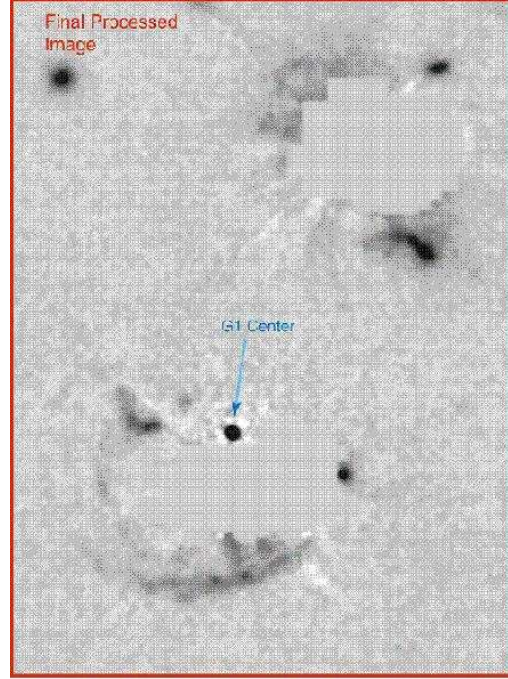


There are a number of cases where the lens and source have such different spectral energy distributions that they must be observed at very different wavelengths. MG 1131+0456 (see Figure 14) and PKS 1830-211 are examples of systems where the source can be seen only in the near-IR. In the case of PKS 1830-211, the source's redshift could be determined only from IR spectroscopy [74]. Other more extreme cases like MG 1549+3047 show the lens only in the optical/near-IR and the source only in the radio. As such systems show no light contamination by the background source, they allow for a detailed study of the lens. In the case of MG 1549+3047, the velocity dispersion of lensing galaxy could even be measured [69,70]. A major drawback however, is that the redshift of the source remains unknown.



**Fig. 15.** The double quasar Q 0957+0561 observed in VLBI [24] at 6 cm with a resolution of 6 milli-arcsec. The two quasar images show a very detailed radio jet that is used to place constraints on the lens model.

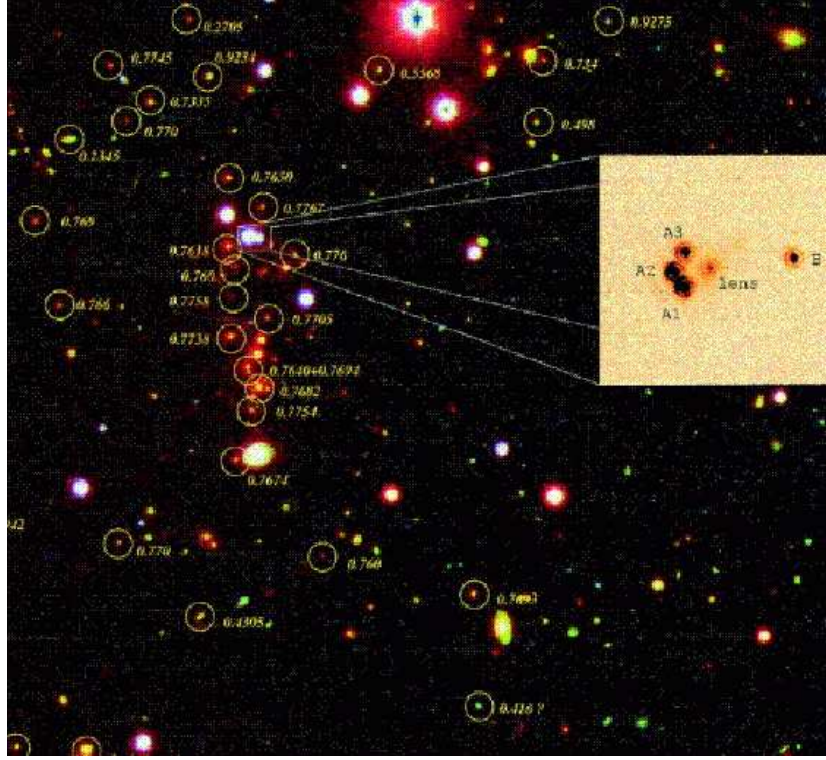
**The quasar host galaxy and background objects** At very high angular resolution, it becomes possible, beyond measuring the position and brightness of the quasar, to resolve details in the distorted and amplified quasar host. Observing the distorted quasar host galaxy brings extra constraints on the



**Fig. 16.** Using the HST, arcs and arclets are discovered in the field of Q 0957+0561 [12] and help to determine the mass profile of the lensing galaxy. The quasar components have been subtracted on this STIS image, provided to us by Gary Bernstein and Phil Fischer, prior to publication. Several arcs are visible as well as members of a foreground galaxy cluster at  $z \sim 0.35$  [5]. G1 is the center of the main lensing galaxy which has also been removed from the image.

lensing potential, and helps to see distant quasar host galaxies (up to redshift 4.5) that would have been missed without the lensing magnification [72,57]. Figure 14 shows an example of a red quasar host galaxy where small details are unveiled, at the resolution of the HST (about  $0.15''$  in the  $H$ -band). Such information is of importance as each detail might be identified in the counter-image and used to place additional constraint on the reconstruction of the lensing galaxy's mass profile. In the radio, using Very Long Baseline Interferometry (VLBI) with resolution on the order of the milli-arcsec, “blobs” can be seen in the lensed images of the radio jet in the source. Such observations, producing the spectacular maps shown in Figure 15 are restricted to very few objects with such high spatial resolution [24,109,130].

Arcs and arclets are sometimes seen in the immediate vicinity of the quasar components. These objects, as shown in Figure 16, might be companions to the quasar or simply unrelated background sources [12]. As in the case of the quasar host, they probe the lensing potential, with the further ad-



**Fig. 17.** The quadruply imaged quasar RX J0911+0551 [7] at  $z = 2.8$  and the intervening cluster at  $z = 0.77$  [58] which significantly modifies the overall gravitational potential responsible for the lensing effect. The field of view is  $3.5' \times 3.5'$ .

vantage that they do not lie at the same position as the quasar and therefore probe the lensing potential in a location otherwise inaccessible.

**Intervening clusters/groups** Isolated lenses may be the exception rather than the rule. Multiply imaged quasars often lie close to the line of sight to foreground groups and even clusters of galaxies. A massive galaxy cluster ( $\sigma \sim 600 - 1000 \text{ km s}^{-1}$ ), even situated several tens of arcsecs away from a system, will modify the expected image position and the time delay, hence also modifying the inferred value for  $H_0$ . Therefore one has to set constraints not only on the astrometry and shape of the main lens, but also on additional objects that may modify the total gravitational potential responsible for a given image configuration.

In the case of B 1600+43 [54,20] the lens is an edge-on spiral at  $z = 0.41$  [33] with a lower redshift spiral a few arcsec South-East [54]. The quadruply imaged quasar PG 1115+080 can be modeled only by taking into account a small group of 4-5 galaxies at  $z = 0.31$  (the same as the lens) about  $20''$

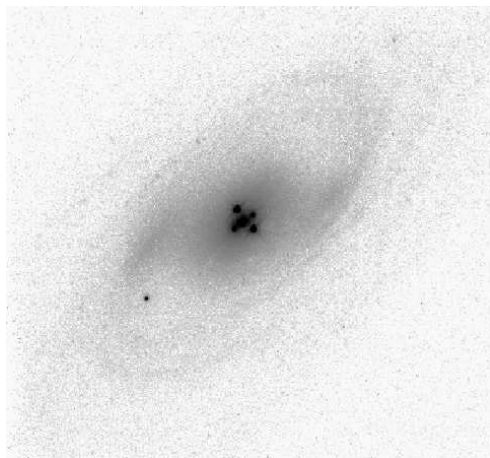
away from the line of sight [56]. Two spectacular examples of intervening clusters are RX J0921+4529 which is situated in an X-ray cluster at  $z = 0.32$  [86] and RX J0911+0551 (see Figure 17). The later is lensed by a galaxy at  $z = 0.77$ , a member of an X-ray cluster centered about  $30''$  to the south [19,85]. The cluster’s velocity dispersion has been measured from the redshifts of 24 members [58] as  $836^{+180}_{-200} \text{ km s}^{-1}$ .

### 2.3 Microlensing of the quasar images

As was shown in section 1.3 that microlensing of individual quasar images is not unexpected, although it depends upon the size of the source and the constituents of the lensing galaxy. The angular scale associated with such microlenses is smaller than that of the macro lens by  $\sqrt{M_{\text{micro}}/M_{\text{macro}}}$ . Stellar mass microlenses therefore produce microarcsecond splittings, not accessible to present-day instrumentation. But the image magnification by microlenses is nonetheless observable – flickering of the combined flux from the unresolved microimages can be detected as the stars are move randomly in the lensing galaxy. While the observational evidence for microlensing is fragmentary, there is enough to indicate that this phenomenon, predicted immediately after the discovery of Q 0957+561 [25], plays an important role in the lensing of quasars.

The first hints of microlensing were found in the doubly imaged quasar Q 0957+561 [131]. The difference light curve between the two components (once corrected for the time delay) showed slow variations unrelated to the intrinsic variability of the quasar. These additional variations are thought to be the explanation of different time delays measured by different investigators. They have also been identified as a potentially interesting tool to set constraints on the stellar content of the lens and on the internal structure of the lensed quasar on parsec scales (see for example [102,95,133]).

Q 2237+0305, also known as the “Einstein cross” (see Figure 18), was quickly recognized after its discovery [49] to be particularly susceptible to microlensing. The redshift of the lens,  $z_L = 0.04$ , is so low that the apparent angular velocity of the microlenses, in projection on the plane of the sky, is much higher than in other systems. Moreover the Einstein rings of these microlenses have a larger angular diameter, making it more likely that they are larger than the source. The quasar is therefore expected to show frequent and rapid variations, the mean time separating each microlensing event being approximately the time required for the microlenses to run across a distance equal to the diameter of their Einstein ring (see equation 7). Note however that time scales involved can be significantly different from those calculated with this naive approach. As can be seen from Figure 10 there are regions of high magnification, in particular close to the cusps of caustics, which are exceedingly narrow, much smaller than the projection back onto the source plane of the Einstein ring [133]. There are also plateaus, larger than the

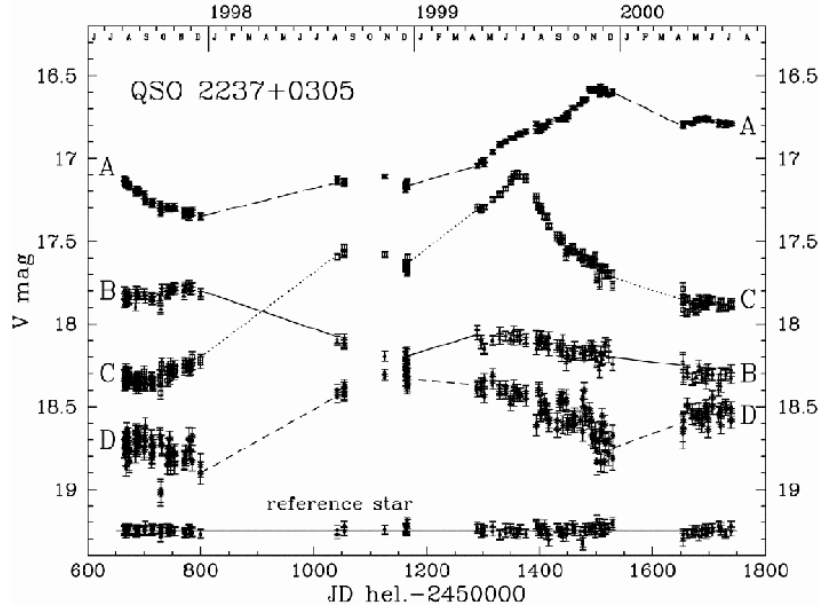


**Fig. 18.** HST V-band image of Q 2237+0305. Four quasar images at  $z = 1.69$  are seen about  $1''$  away from the nucleus of a much lower redshift lensing galaxy ( $z = 0.04$ ). The high density of stars in the lens' nucleus and their high projected angular velocity make of Q 2237+0305 a privileged object for the study of microlensing.

Einstein ring, over which the magnification is relatively constant, and usually less than unity.

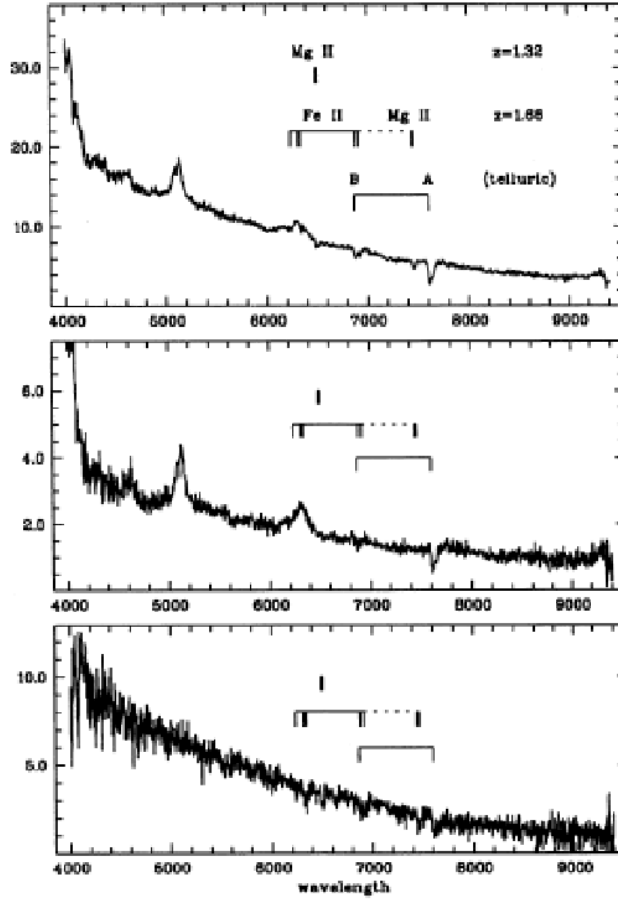
As the optical path to each quasar image intersects the lensing galaxy at very different locations, microlensing-induced variations in the light curves of the quasar images are uncorrelated. Intrinsic variations of the quasar would be seen identical in each image, separated by the time delay. Time scales involved for microlensing events in the Einstein cross were predicted to be of the order of a few months [133], spectacularly confirmed by optical monitoring [92,148]. This is much longer than the time delay of the system (about 1 day), making it easy to discriminate between intrinsic variations of the source and microlensing events. Figure 19 illustrates this: erratic variations of the 4 light curves (especially component C) are seen, with a typical time-scale of a few months. At the scale of the plot, intrinsic variations of the quasar would be seen simultaneously in all light curves.

Flickering of quasar light curves is not the only signature of microlensing. As noted in section 1.3 the Einstein radius of microlenses is small and may be comparable in size to the inner regions of quasars. One may therefore observe differential magnification of regions of different sizes. As different regions of quasars are thought to have different colors, this implies *chromatic* magnification. There are many instances where the flux ratios for quasar images are quite different at different wavelengths [84]. Static microlensing is often invoked as a possible explanation. If one region varies and another does not, static microlensing might also produce chromatic differences in the quasar light curves.



**Fig. 19.** Optical V-band light curves for the four quasar images in the Einstein cross, Q 2237+0305. The time-delay in this system is of the order of a day. The very different behaviour of the four light curves, with slow variations of the order of a month, strongly support the idea of microlensing induced variability [148].

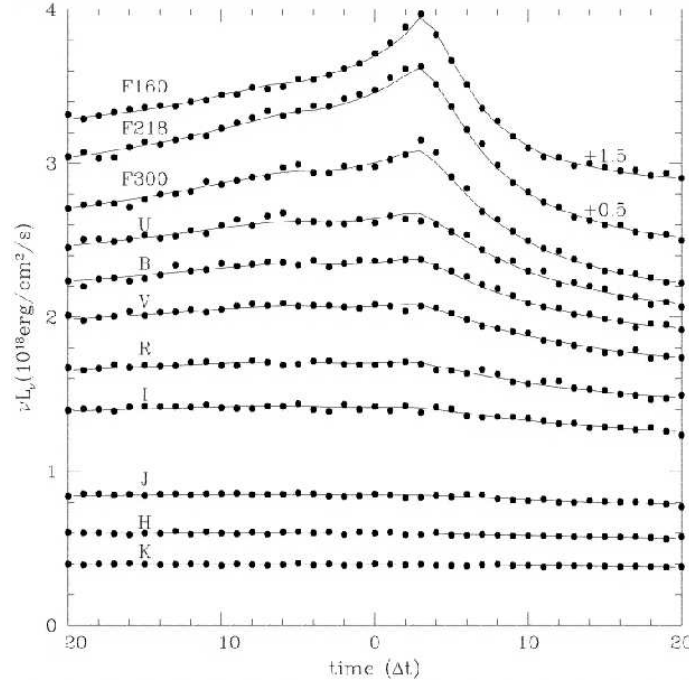
More generally, the spectral differences among the regions of a quasar will involve the presence or absence of emission lines. One might therefore expect differential magnification of the emission lines and continuum. The deblending of closely separated quasar spectra is not trivial. Fortunately, the relatively wide angular separation system, HE 1104-1805, appears to show the phenomenon [139,142]. In this double system (see Figure 13), the spectra of the two components are identical in the emission lines but show a different continuum, suggestive of microlensing. Figure 20 shows the difference spectrum (bottom panel) between the two quasar images. In order to subtract properly the emission lines from the spectrum of component A (top panel), one has to subtract a scaled version of the spectrum of component B (middle panel). The scaling factor of 2.8 is found to be stable with time and wavelength, even in the near-IR [30]. This suggests that emission lines are unaffected by microlensing. The difference continuum is blue and shows photometric variations. Part of these variations are intrinsic to the quasar, and are used to infer the time delay of the system, but additional flickering can be attributed to microlensing [142]. With higher signal-to-noise spectra of HE 1104-1805, it has been found that some emission lines might be affected by microlensing as well [75]. For example the red side of the CIII emission



**Fig. 20.** Spectra of the two quasar images in HE 1104-1805 [139]. The two first panels starting from the top show the quasar spectra. The bottom panel shows the difference between the spectrum of the brightest component of the system and a scaled version of the spectrum of the faint component. A scaling factor of 2.8 is necessary to subtract the emission lines from the spectrum. The labels in the figure are related to absorption lines by the lensing galaxy.

line does not subtract perfectly after subtraction of the B spectrum. This is also true for HE 2149-2745 [22], but there are no similar observations so far for other systems.

In principle, this effect can be used to study the detailed structure of active galactic nuclei. For the sake of argument, one adopts a standard model for of active galactic nuclei [17] where, for example, the continuum region is much smaller than the broad line region. The continuum region itself is composed of an accretion disk which radiation field is less and less energetic when



**Fig. 21.** Expected light curves for the microlensing events in a quasar at  $z = 1.695$  [3]. Microlensing events have larger amplitudes (larger magnification) at short wavelengths. The time axis is in arbitrary units which depends on the velocity and redshift of the microlenses.

going from the center of the accretion disk to its outer parts. As wavelength decreases with increasing energy, the accretion disk radiates bluer photons in the center than in the outer parts. Let us now make the realistic assumption [133] that the mean size of the micro-caustics corresponds about to the angular size of a quasar accretion disk. Simple geometric considerations shows that the smaller regions (compared with the caustics) are entirely magnified, while only a fraction of the larger regions are amplified. That is, the emission line region will be less magnified by microlensing than the continuum, and the outer parts of the accretion disk will be less magnified than the inner parts. In other words, the flickering of the light curves expected from microlensing should be stronger in the blue than in the red, and should even be invisible in the emission lines.

This very simple scheme is a lot more complicated in practice, simply because one can not map the actual network of micro-caustics present in a given lensed system: this would require a map of the mass distribution in the lensing galaxy ! Still, one can propose a quasar model and predict the statistical behaviour of the light curves, as a function of wavelength. Such



a theoretical work has been investigated [3,149,150], with the goal to derive the relative sizes of quasar emission regions. Figure 21 illustrates how a given distribution of microlenses preferentially magnifies the innermost (blue) regions of quasar accretion disks, hence producing light curves with luminosity peaks progressively increasing while observing from the near-IR  $K$ -band to the ultraviolet  $U$ -band [3].

Unfortunately, the amount of data available so far is too small to implement any of the proposed method to probe quasar structure from microlensing.

### 3 Models

The small number of observables in lensing means that the observational data, no matter how accurate, can be fit by a huge variety of lens models. The space of allowed models must be narrowed by the adoption of priors which reflect our understanding of the relative astrophysical plausibility of different mass models. There are two strategies for doing this. One is to adopt a parameterized mass distribution, where the parameters are chosen to include the reasonable and important variations expected among lensing galaxies. The other strategy is to keep the mass map free-form, but impose astrophysical priors as constraints on it.

We now discuss both these modeling strategies.

#### 3.1 Parameterized models

When building parametrized models, the small number of observables then demands a small number of parameters. Eliminating a parameter (e.g. the octupole moment of the gravitational potential) means that some aspect of the lensing galaxy is not being modeled. A wise choice of parameters models those aspects which are important for the task at hand.

Fortunately there is a vast literature on the mass distributions and gravitational potentials of galaxies. For example (and quite importantly for the interpreting time delays) we know that galaxies have mass density profiles which vary roughly as  $1/r^2$ , giving galaxies flat rotation curves and flat velocity dispersion profiles. For the sake of discussion we put forward here a “standard” model which incorporates much of what we know about nearby galaxies.

**Some simple models** We start with the simple monopole potential and describe a number of additional terms which correspond, at least roughly, to what one might expect for galaxies in a variety of contexts, adding degrees of freedom which we have reason to believe nature exploits.

*Singular isothermal sphere* The singular isothermal sphere is a cornerstone of galaxy dynamics [15]. It gives the flat rotation curves and constant velocity dispersion profiles characteristic (to a first approximation) of spiral and elliptical galaxies respectively. It has a three dimensional potential  $\Phi = v_c^2 \ln r$ , where the circular velocity  $v_c = \sqrt{2}\sigma$ , with  $\sigma$  the one dimensional velocity dispersion. Integrating this along the line of sight and multiplying by  $2D_{LS}/(D_L D_{SC}^2)$  gives the lens potential

$$\psi(\vec{\theta}) = \theta_E \theta, \quad (36)$$

which is the same as equation (27) but now with  $\theta_E = 4\pi\sigma^2 D_{LS}/D_{SC}^2$  (measured in radians) giving the lens strength. We recall (cf. equation 28) that such a lens produces two colinear images, one on each side of the lens, with magnifications given by equation (29). The infinite second derivative of the potential at the origin gives infinite demagnification of a central third image.

*Power-law monopole* The singular isothermal sphere is a special case of the power-law monopole

$$\psi(\vec{\theta}) = \frac{\theta_E^2}{(1+\alpha)} \left( \frac{\theta}{\theta_E} \right)^{1+\alpha}, \quad (37)$$

where the central concentration index,  $\alpha$ , measures the deviation from isothermality. The normalization has been chosen so that the strength  $\theta_E$  is again the radius of the Einstein ring. As the exponent  $\alpha$  approaches  $-1$ , the potential approaches that of a point mass.

*Self-similar power law quadrupole* Since galaxies are not circularly symmetric, there is no reason why their effective potentials should be. The flattening of a potential is dominated by a quadrupole term, which if we use polar coordinates with  $\theta = (\theta, \phi)$ , varies as  $\cos 2\phi$ . A simple model for the effective potential which incorporates the non-negligible quadrupole of galaxies incorporates quadrupole term with the same radial dependence of the monopole, giving equipotentials which similar scaled versions of each other,

$$\psi(\vec{\theta}) = \frac{\theta_E^2}{(1+\alpha)} \left( \frac{\theta}{\theta_E} \right)^{1+\alpha} [1 + \gamma \cos 2(\phi - \phi_\gamma)]. \quad (38)$$

An on-axis source gives 4 images whose distance from the lens center is approximately equal to the lens strength  $\theta_E$ . Dimensionless  $\gamma$  gives the flattening of that quadrupole and  $\phi_\gamma$  gives its orientation.<sup>4</sup> The special case  $\alpha = 0$  gives a flattened system with the flat rotation curve and constant velocity dispersion profile characteristic of isothermals. While the equipotentials are self similar for all  $\alpha$ , the equipotentials and the equidensity contours are both self-similar only for the  $\alpha = 0$ , isothermal case.

<sup>4</sup> Following Kochanek (1992) we (somewhat confusingly) use the same symbol,  $\gamma$ , for the flattening as is used for the shear. The flattening and the shear are equal at  $\theta = \theta_E$ , but not elsewhere.

*Tidal quadrupole (first order tide)* Equipotentials which have the same shape are esthetically appealing but highly idealized. In particular galaxy equipotentials will deviate strongly from self-similarity if the quadrupole is due to a tide from a neighboring galaxy or cluster of galaxies. In that case the quadrupole term shows a  $\theta^2$  dependence on distance from the center of the lens, as was seen in equation (30). Among others, [65] have noted that the quadrupoles of many lensed systems appear to be due to tides rather than to the flattening of the lensing galaxies. A simple potential incorporating these features is

$$\psi(\vec{\theta}) = \frac{\theta_E^2}{(1+\alpha)} \left( \frac{\theta}{\theta_E} \right)^{1+\alpha} + \frac{1}{2} \gamma \theta^2 \cos 2(\phi - \phi_\gamma). \quad (39)$$

This is much like the self-similar power law quadrupole of equation (38). While the monopole term is, as in the previous cases, a power law, the quadrupole term has the  $\theta^2$  dependence characteristic of a tide. In general we expect a lens to have both a tidal quadrupole, from neighboring galaxies, and something like the self-similar quadrupole due to the flattening of the lensing galaxy itself.

*Clusters as mass sheets* Galaxies typically reside in groups and clusters, with considerably more mass (dark matter) associated with the cluster than with the individual galaxies. One must therefore take the gravitational potential of the associated cluster into account. The scale of a cluster is much larger than that of a galaxy, so its surface density projected onto the galaxy is to first order constant. A mass sheet of uniform density produces an effective potential

$$\psi_s(\vec{\theta}) = \frac{1}{2} \kappa_s \theta^2 \quad (40)$$

where  $\kappa_s$  is the dimensionless convergence associated with the mass sheet.

As was shown in section 1, differentiating twice one finds that a superposed mass sheet stretches an image configuration by a constant factor  $1/(1 - \kappa_s)$  without changing any of the dimensionless ratios associated with the image configuration. A model that failed to take account of such a mass sheet would predict too long a differential time delay by the just this same stretch factor. But there is no way of knowing from image positions or relative magnifications whether or not such a mass sheet is present. This formal degeneracy *demands* that one bring to bear “external” information regarding the projected density of any such mass sheet.

*Clusters and higher order tidal terms* In the above paragraphs we have identified two distinct effects of clusters of galaxies: they introduce tidal and mass sheet terms into the effective lensing potential. There are many lenses for which the first order tidal terms are so strong (e.g. [65]) that higher order terms are likely to be important. The simplest way to do this is to drop the tidal term above and to model the cluster as an isothermal at position  $\vec{\Theta}$  with effective potential

$$\psi_c(\vec{\theta}) = \Theta_E |\vec{\Theta} - \vec{\theta}|. \quad (41)$$

This model has three free parameters (replacing the two tidal parameters,  $\gamma$  and  $\phi_\gamma$ ), the lens strength  $\Theta_E = 4\pi\sigma^2 D_{LS}/D_{SC}^2$  where  $\sigma$  is the the velocity dispersion of the cluster, and the polar coordinates  $(\Theta, \Phi)$ .

While the cluster potential can be written quite compactly in this form, it obscures the connection between the cluster properties and the tidal and mass sheet terms described above. Taking the lensing galaxy to be at the origin of our coordinate system, we can expand the cluster potential in powers of  $\theta/\Theta$ , where  $\Theta$  is the distance of the cluster from the origin. Dropping constant terms, we find

$$\begin{aligned} \psi_c(\vec{\theta}) = & -\Theta_E \vec{\Theta} \cdot \left( \frac{\vec{\theta}}{\Theta} \right) \\ & + \frac{1}{4} \Theta_E \Theta \left( \frac{\theta}{\Theta} \right)^2 - \frac{1}{4} \Theta_E \Theta \left( \frac{\theta}{\Theta} \right)^2 \cos 2(\phi - \Phi) \\ & + \text{terms of order } \left( \frac{\theta}{\Theta} \right)^3 \text{ and higher.} \end{aligned} \quad (42)$$

The first term gives a constant deflection  $\Theta_E$  away from the cluster, showing that the source position  $\vec{\beta}$  may be rather far from the origin and the lensing galaxy. The second term is just that of a mass sheet with  $\kappa_s = \frac{1}{2}(\Theta_E/\Theta)$  while the third is a tidal term with shear  $\gamma = \kappa$ . Noting that the coefficient of the shear term is negative, we find that the position angle of the shear,  $\phi_\gamma$ , as defined in equations (38) and (39), is at right angles to the position angle of the cluster,  $\Phi$ .

The equality of the shear and convergence suggests a possible resolution of the mass sheet degeneracy: measure the shear and infer the convergence. We adopt this approach with an obvious caveat. To the extent that clusters and groups are *not* isothermal, such a “shear inferred” mass sheet correction will introduce a systematic error in a derived Hubble constant.

The terms of order  $(\theta/\Theta)^3$  are useful because they break the classical tidal degeneracy. Since  $\gamma = \frac{1}{2}(\Theta_E/\Theta)$ , we might produce an equally strong first order tide by putting an isothermal cluster with twice the Einstein radius at twice the distance. Alternatively, we might put the cluster at position  $-\vec{\Theta}$  without changing the first order tide. Keeping the higher order terms resolves these ambiguities. But rather than add many coefficients, it is conceptually simpler and more economical to replace the two parameters of a first order tide with the three parameters of a circularly symmetric cluster. There are several lenses (e.g. RX J0911+0551, PG 1115+080 and B1422+231) for which higher order tidal effects have been used to determine the position and lensing strength of the associated cluster.

*Yet more degrees of freedom* Even in the absence of tides, there is no reason to insist that the monopole and quadrupole terms of a galaxy potential have

the same dependence on  $\theta$ , *i.e.* that the potential be self-similar. The self-similar model presented above can readily be extended to allow for separate  $\theta$  exponents, permitting the potential to get rounder or flatter with increasing  $\theta$ . Nor is there any reason, in principle, why we should limit ourselves to monopole and quadrupole terms. Purely elliptical density profiles produce potentials with higher order multipoles. Some ellipticals are “boxy” while others are “disky” [11], and these too should have higher order multipoles. Power laws like equations (37) and (38) give unphysical mass and density divergences, and should in principle be cut off at small or large  $\theta$  or broken somewhere in between. Finally, we might argue that it is naive to assume that the dark matter in a galaxy is centered on its starlight, and that we should take the central coordinates of the lensing potential to be free parameters.

With all these possibilities, it is no surprise that different investigators modeling the same system come up with different potentials and derive different values of  $H_0$  from the same time delay. The number of measurements which constrain the potential is small, so one cannot allow oneself the luxury of adding extra parameters just for the sake of insurance. In introducing new parameters the two questions to be kept in mind are the degree to which they degree to which they affect the deflections and distortions, which constrain the potential, and the degree to which they affect the delays, which give the Hubble constant.

**Useful approximations and rules of thumb** For the sake of simplicity, suppose that a lens has the power-law monopole potential of equation (37). Using the lens equation, we substitute the gradient of the effective potential,  $\partial\psi/\partial\vec{\theta}$ , for the deflection,  $\vec{\theta} - \vec{\beta}$ , in the time delay equation. Under the assumption that two quasar images,  $A$  and  $B$  are roughly equidistant from the center of the potential, we predict a differential time delay

$$\tau_B - \tau_A \approx \frac{1}{2}T_0 (\theta_A^2 - \theta_B^2) (1 - \alpha), \quad (43)$$

where  $T_0$  is the time scale defined in equation (9) and  $\theta$  is measured in radians. Had we not assumed circular symmetry,  $\theta_A^2$  and  $\theta_B^2$  would each have a coefficient which differed from unity by a factor of order gamma, usually less than 10%. Equation (43) has the important and useful property that it depends only upon observable quantities, assuming that the position of the lensing galaxy can be measured.

Several useful lessons can be drawn from equation (43). First, the more distant image leads the closer image (cf. Figure 8). Second, if  $\theta_A \approx \theta_B$  high astrometric accuracy is needed in measuring the position of the lensing galaxy for high precision in the predicted time delay. Third, the predicted delay scales as the square of the separation. The differential time delay of Q 0957+561, 172, is therefore atypically long, resulting from its large (6.1'') separation and relative asymmetry. Fourth, if the lens potential is more sharply peaked than a singular isothermal sphere, the predicted time delay is longer. In particular, a point mass model, with  $\alpha = -1$ , predicts a time delay

twice as large (yielding a Hubble constant twice as large for a given observed delay) as the corresponding singular isothermal,  $\alpha = 0$  model. Either  $\alpha$  must be measured with high accuracy from the observed image configuration or we must bring external considerations to bear upon our models. In comparing models by different investigators for the same system, one must pay particular attention to the way in which the degree of central concentration has been treated.

**Fitting models** *How and what to fit ?* On first thought it seems straightforward to adopt a lens potential and a source, find the predicted images, compare those with the observed images and adjust the parameters associated with the potential and the source so as to get better agreement. On closer examination one discovers that the lens equation can only rarely be solved in closed form for image positions. Worse yet, one finds that small changes in parameters can cause pairs of predicted images to merge and disappear. What does one do in a gradient search when a small trial step causes an image to cease to exist ? Fortunately robust methods for fitting data have been developed, some of which are publicly available [57].

The fluxes of images can readily be measured to 1% accuracy, but the differences between optical and radio flux ratios are of order 10-30%. Given the very much greater accuracy of positions, one might be tempted to dispense with magnifications entirely. But for double lenses, even the simplest non-circular models have one too many parameters to permit fitting using positions alone. Moreover, fitting fluxes can help avoid aforementioned disappearing image problem. It is therefore helpful to use fluxes, but with full awareness of the associated pitfalls.

Image positions constrain the first derivative of the effective potential. Magnifications constrain the second derivative. In some systems more than one time delay can be measured. The first measured delay goes to solving for the dimensioned combination of angular diameter distances in equation (43) but the ratios of the second and subsequent delays to the first give dimensionless constraints on the effective potential itself. Though not yet incorporated in most parametrized models, such constraints are in principle quite powerful. But a disadvantage so far is that in practice the uncertainties in all measured delays for a given system are roughly the same, as measured in days; so while the fractional uncertainties in the longest delay is typically better than 10%, the fractional uncertainties in the shorter delays are correspondingly greater.

**What constitutes “good enough”?** There is little difficulty in finding models for lens systems which fit the observed data perfectly. The number of constraints is small, and the number of free parameters is large, and so it should be possible to find an  $N$  parameter model which fits the  $N$  available constraints perfectly. But that leaves no room for reality testing. Ideally one hopes to find a model with  $< N$  parameters for which the predicted images

agree with the observed images within the measured uncertainties, giving an acceptable fit to the data.

The words “unacceptable fit” have a damning ring which tends to end discussion. Were we able to measure the relative positions of the lensed images to one part in a million, the deflections due to individual stars within the lensing galaxy become important. At that point we would be unlikely to ever get an acceptable fit from a macromodel. But the differences in the time delays induced by such microlensing are small.

For the purpose of interpreting time delay measurements a less stringent definition of acceptable may be in order. Consider the case of Q 0957+561. Errors in the positions of 100 milliarcseconds introduce negligible changes in the time delay predicted by equation (43). While one can concoct a parameterized model for which small differences in the positions produce large changes in the predicted time delays, these are, with the exception of the central concentration degeneracy, somewhat artificial.

**The central concentration degeneracy** The central concentration degeneracy has already surfaced in our discussion, first theoretically as the mass disk degeneracy, and then in the approximate rule for computing time delays, equation (43), and it appears yet again in connection with free-form models. It has also surfaced many times in the literature. A particularly thorough treatment can be found in [130], though it is evident as well in other works [13,115]. Briefly, it has proven exceedingly difficult to constrain the (radial) second derivative of the monopole term of the effective potential. Several factors contribute. In double systems the associated parameters are coupled to the quadrupole amplitude. In quadruple systems the images all tend to lie at roughly the same distance from the center of the lens – otherwise the system wouldn’t be quadruple. The radial displacements of these images depend not only on the concentration parameter but also on higher order multipoles. Einstein rings may be less susceptible to this degeneracy because the rings are resolved in the radial direction, though this is controversial [62,113].

What makes this degeneracy pernicious is its strong influence on the predicted time delay, increasing them by a factor  $(1 - \alpha)$  in the parameterization of our power-law models (equations (37) and (38)). In the face of this, one has two choices: to search for a “golden” lens which doesn’t suffer from it or to bring external constraints to bear.

Golden lenses, at least 24 carat golden lenses, are rare. MG J0414+0534 would at first sight seem as good a candidate any, with a core and 3 VLBI features, each quadruply imaged. But [130] conclude “It is clear that useful information on the radial profile of MG J0414+0534 is unavailable from this data.” Alas even if it were, the object has shown little sign of variability [82,6].

Measurements of velocity dispersion gradient [129] have been made of the lensing galaxies in Q 0957+561 and PG 1115+080, which in principle constrain the degree of central concentration of the potential. This is a par-

ticularly difficult measurement because of the competition from the optical images of the lensed quasar. Moreover the effective radius of the lensing galaxy tends to be considerably smaller than the Einstein ring, making it difficult to obtain measurements out to the region of interest.

An alternative approach [108] is to use what one knows about the potentials of nearby elliptical galaxies. They compiled data on the potentials of nearby elliptical galaxies for which not only velocity dispersions but higher order moments of the line of sight velocity distribution had been measured. Their data show a mean power-law index  $\langle\alpha\rangle = -0.2$ , with a scatter of roughly 0.2 about that value. This is somewhat more centrally concentrated than for the isothermal index,  $\alpha = 0$ , but not nearly so concentrated as the point mass index,  $\alpha = -1$ . If the data fail to constrain the power-law exponent, fixing it at its mean value would introduce errors in the predicted time delays of roughly 20%. But since the observed power-law index is so close to the isothermal value,  $\alpha = 0$ , and since the power law index makes so little difference in the quality of the fit (otherwise it would be well constrained), one does little harm in fixing the power-law index at its isothermal value and making a post-hoc correction to the predicted time delay.

**A proposed “standard” model for lenses** As the preceding sections make only too clear, predicted time delays and derived Hubble constants depend sensitively upon how lens potentials are modeled. In particular, they are sensitive to the degree of central concentration of the lens model, which is especially difficult to constrain using lensed images alone. Such model differences have led to widely divergent predicted delays and derived Hubble constants for what are essentially the same data.

Absent the discovery of a 24 carat lens, one can still make progress measuring the Hubble constant by accepting that most lenses are underconstrained and adopting a “standard” model for which the associated systematic errors are well understood and which is sufficiently simple that it can be applied to a large fraction of the known lensed systems.

*The proposed standard* In the belief that it will take us within striking distance of  $H_0$ , we propose the following “standard” effective potential,

$$\psi(\vec{\theta}) = \theta_E \theta + \frac{1}{2} \gamma \theta^2 \cos 2(\phi - \phi_\gamma), \quad (44)$$

which is the isothermal variant of the tidal power-law plus quadrupole of equation (39). To the extent that they are understood, the systematic and random errors associated with this model are as follows.

As noted above the assumption of isothermality,  $\alpha = 0$ , introduces a systematic error, but this can readily be corrected by multiplying the predicted time delay by the factor  $1 - \langle\alpha\rangle$ . We choose to fit  $\alpha = 0$  because in most cases the available data fail to constrain  $\alpha$  any better than this external constraint and to avoid fussing about second and third generation standards as the appropriate mean value of  $\alpha$  is further refined.



In double systems there are too few constraints to permit discrimination between the tidal isothermal as in our proposed standard model and a self-similar isothermal. Among quadruple systems tides appear to be more important than the flattening of the lenses [56], but then tides may explain the relatively large number of quadruple systems [65].

For our proposed standard the differential time delay is given by

$$\tau_B - \tau_A \approx T_0 \times \{ \theta_A^2 [1 + \gamma \cos 2(\phi_A - \phi_\gamma)] - \theta_B^2 [1 + \gamma \cos 2(\phi_B - \phi_\gamma)] \}. \quad (45)$$

Had we instead adopted the isothermal variant of the self-similar power-law potential of equation (38), the square bracketed terms would have reduced to unity as in (43). If we have made the wrong choice, and if the orientation of the shear,  $\phi_\gamma$ , is random, our choice of a tidal quadrupole introduces a random error but not a systematic one. If  $\gamma$  is small, the effect is not large. If  $\gamma > 0.1$ , the quadrupole term is so large that an external tide seems the more likely possibility. So either we make a small random error or we make the right choice.

If we believe that the shear is largely tidal, it seems reasonable to assume that the tide is due to an isothermal potential, and that there is an associated convergence  $\kappa = \gamma$ . The predicted time delays of equation (45) should therefore be multiplied by a factor  $(1 - \gamma)$  to account for the “mass sheet” associated with the tidal perturber. We cannot avoid making a systematic error here, but we make a larger error in failing to correct for the projected surface densities associated with tides than in making the correction. Our doubly corrected prediction is therefore

$$(\tau_B - \tau_A)_c = (1 - \langle \alpha \rangle)(1 - \gamma)(\tau_B - \tau_A). \quad (46)$$

In summary, our standard model is a tidal singular isothermal. We fit the model to the available constraints and use equation (45) to compute the time delay. We apply a correction factors of  $1 - \langle \alpha \rangle$  and  $1/(1 - \gamma)$  to the predicted time delay to account for, respectively, the mean power-law index observed in ellipticals and the projected surface density associated with tides.

*Application of the proposed standard* Our standard model is by no means new. It is one of two models used by the CASTLES group to analyze the lens data they have assembled. We note, however, that they do not apply the corrections for central concentration and convergence that we adopt in the previous section.

The CASTLES model for PG 1115+080 has a shear of 0.12 with a predicted C-B time delay of 12<sup>d</sup>8 for an  $h = 1$  EdS universe. Applying the corrections of equation (46) using a mean concentration,  $\langle \alpha \rangle = -0.2$ , gives a predicted time delay of 13<sup>d</sup>5. Using Barkana’s value of 25<sup>d</sup>0 for the observed delay gives  $h = 0.53$ .

The CASTLES group has not yet posted a SIS+shear model for RX J0911+0551, but Schechter gives a shear of 0.307 and a predicted time delay of 120<sup>d</sup>.5 between B and (A1+A2+A3). Again applying equation (46) we get a corrected prediction of 100<sup>d</sup>.2. [47] report preliminary delay measurement of 200<sup>d</sup>, giving  $h = 0.50$  for an EdS universe.

### 3.2 Free-form models

Free-form models build a lens as a superposition of a large number of small components, with minimal assumptions about the form of the full lens. They are motivated by three considerations.

1. The fewness of observables in quasar lensing, and the presence of degeneracies, means that any one lens reconstructed from observations is highly non-unique. One needs some systematic way of searching through possible lens reconstructions.
2. The high accuracy of observations, despite their fewness, means that data always show deviations from the parametrized models discussed above. Models with more parameters can fit the data to observational accuracy, but it is not known what all the essential parameters are. Are twisting isodensity contours important? Does ellipticity vary significantly with radius?
3. The most important observational constraints from lensing (being image positions, tensor magnifications, and time delays) are linear, which makes it straightforward to fit lenses by superposition.

We will refer to the small components as pixels, but in fact they can be any kind of components and not necessarily small. For example, they may be Fourier or harmonic terms in the mass profile [130]. But here we will discuss in detail the case where the pixels are mass tiles with uniform but adjustable surface density [111,138].

Consider a lens made up of  $N$  pixels each with mass profile  $\kappa_n f_n(\vec{\theta})$ . Here  $\kappa_n$  is an adjustable parameter.<sup>5</sup> Let  $Q_n(\vec{\theta})$  be the integral of  $\nabla^{-2} f_n(\vec{\theta})$  over the  $n$ -th pixel. In other words, let  $\kappa_n Q_n(\vec{\theta})$  be the  $n$ -th pixel's contribution to the lens potential at  $\vec{\theta}$ . For square tiles or Gaussian tents  $Q_n$  is known but messy [111,2]. For harmonic components or other eigenfunctions of  $\nabla^2$ ,  $Q_n(\vec{\theta})$  is simply proportional to  $f_n(\vec{\theta})$ . For a pixelated lens the arrival time surface (11) becomes

$$\tau(\vec{\theta}) = \frac{1}{2}\vec{\theta}^2 - \vec{\theta} \cdot \vec{\beta} - \sum_n \kappa_n Q_n(\vec{\theta}) \quad (47)$$

where we have discarded a  $\frac{1}{2}\vec{\beta}^2$  term from (11) since it is constant over the surface.

---

<sup>5</sup> Hence we deprecate the alternative name ‘non-parametric’ for this method, favoring ‘free-form’ or ‘pixelated’.

We may now implement three kinds of observational constraints.

1. Image positions: an image observed at  $\vec{\theta}_i$  implies

$$\vec{\nabla}\tau(\vec{\theta}_i) = 0. \quad (48)$$

(We can safely neglect the uncertainty in  $\vec{\theta}_i$ , since image astrometry is typically at the mas level). A multiple-image system derives from the same  $\vec{\beta}$ , but that  $\vec{\beta}$  is unknown. So each such system introduces  $2(\langle \text{images} \rangle - \langle \text{sources} \rangle)$  constraints.

2. A time-delay measurement between images at  $\vec{\theta}_i$  and  $\vec{\theta}_j$  implies

$$\tau(\vec{\theta}_i) - \tau(\vec{\theta}_j) = h T_0^{-1} \times \langle \text{obs delay} \rangle. \quad (49)$$

In a quad there may be two or three independent time delays.

3. Tensor magnifications are measured from images of a multiple-component source. The implied constraints can be included simply by treating the images of separate components as independent image systems. A scalar magnification, or simple flux ratio, cannot be included in this way; however, flux ratios are sensitive to microlensing and thus less suitable for constraining macro-models.

All these constraints are linear in the unknowns  $\kappa_n(\vec{\theta})$  and  $\vec{\beta}$ . Schematically, we may write

$$\begin{pmatrix} \text{Lensing} \\ \text{data} \end{pmatrix} = \begin{pmatrix} \text{A} & \text{messy} & \text{but} & \text{linear} & \text{operator} \\ \text{also} & \text{involving} & \text{the} & \text{same} & \text{data} \end{pmatrix} \begin{pmatrix} \text{The} \\ \text{lens's} \\ \text{projected} \\ \text{mass} \\ \text{distribution} \end{pmatrix}$$

Note the un-square matrix: there are many more pixels than data, i.e., the reconstruction problem is highly underdetermined.

It is easy to find lens profiles formally consistent with observational constraints as above, but most of them will not look anything like galaxies. We now try to exclude the latter with additional constraints — in Bayesian terminology we apply a prior. A reasonable prior is the following.

- $\kappa_n \geq 0$ ,
- $180^\circ$  rotation symmetry (optional),
- density gradient  $\leq 45^\circ$ ,
- $\kappa_n \leq 2\langle \text{average of neighbors} \rangle$ , except for the central pixel,
- $\kappa$  steeper than  $|\vec{\theta}|^{-1/2}$ , based on stellar dynamics evidence that the 3D density in galaxies is steeper than  $r^{-1.5}$ .

The observational and prior constraints confine allowed lenses to a convex polyhedron in the space of  $(\kappa_1, \kappa_2, \dots, \kappa_N)$ . This can now be searched by a random-walk technique, yielding an ensemble of models. And then one can estimate  $h$  or whatever from that ensemble [138].

There are three caveats associated with this technique.

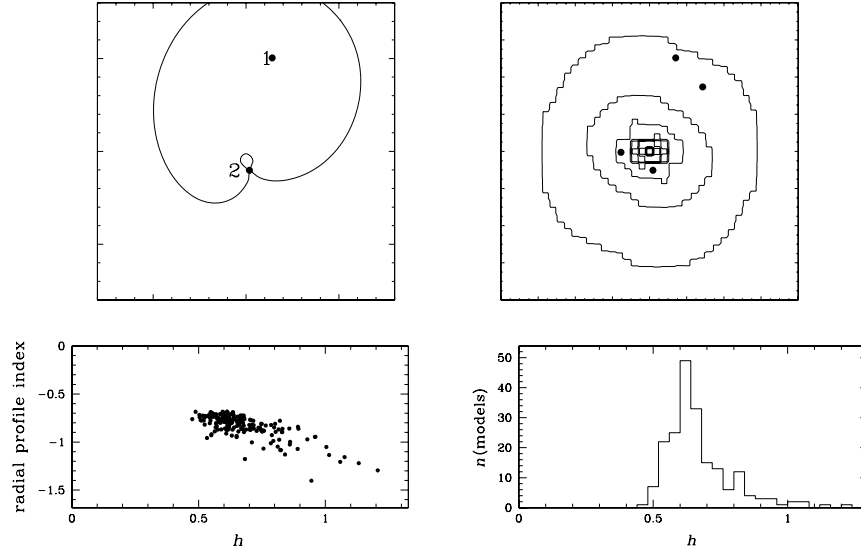
- The results depend on the prior, and the above prior certainly has too little information. But it has at least the advantage that uncertainty estimates will be conservative.
- Having the mass on tiles means that the models cannot have central density cusps, contrary to what galaxies are thought to have. But far from the center, this is not an issue because of the monopole degeneracy.
- There is too much pixel-scale structure. This is not an issue for  $h$ , but if one wanted input for microlensing computations then the pixel-scale structure needs to be smoothed out.

**Four well-known systems** We now describe some new results obtained by one of us (PS) with L.L.R. Williams, on four lenses. For each lens, there was an ensemble of 200 models. The  $\kappa_n$  controlled  $\sim 500$  mass tiles, plus a parametrized external shear.

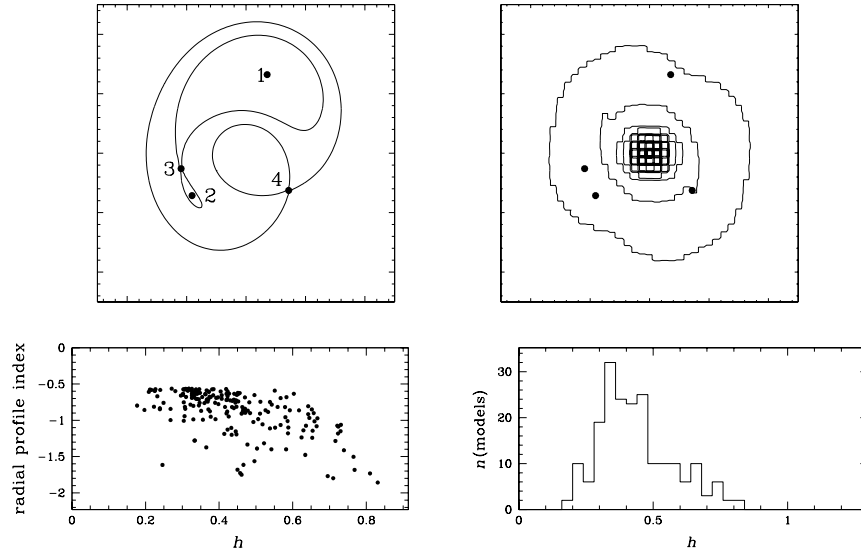
*Q 0957+561* The reconstructions use the positions, tensor magnifications [39], and time delays [65] of the quasar, and another double formed by a knot in the quasar’s host galaxy [13]. Figure 22 shows (i) the image configuration and schematic saddle-point contours for the quasar, (ii) the ensemble-average mass map, (iii) the  $h$  values from each model in the ensemble plotted against the radial-profile index of that model between the innermost and outermost images, and (iv) a histogram of the  $h$  values from the ensemble. The radial profile index corresponds roughly to  $\alpha - 1$  for small values of  $\alpha$  as defined in equation (37).

Two things are very noticeable in Figure 22. The first is the largeness of the uncertainty in  $h$ ; even in this lens with VLBI structure giving tensor magnifications and a time delay accurate to 1%,  $h$  values between 0.5 and 1 are all admissible. The second noticeable thing is the near-proportionality of  $h$  and the radial index, and it points us to the dominant source of the uncertainty: changing the radial index is almost equivalent to applying the mass disk degeneracy transformation, which rescales the time delays, and hence  $h$ , while having no effect on image positions or tensor magnifications.

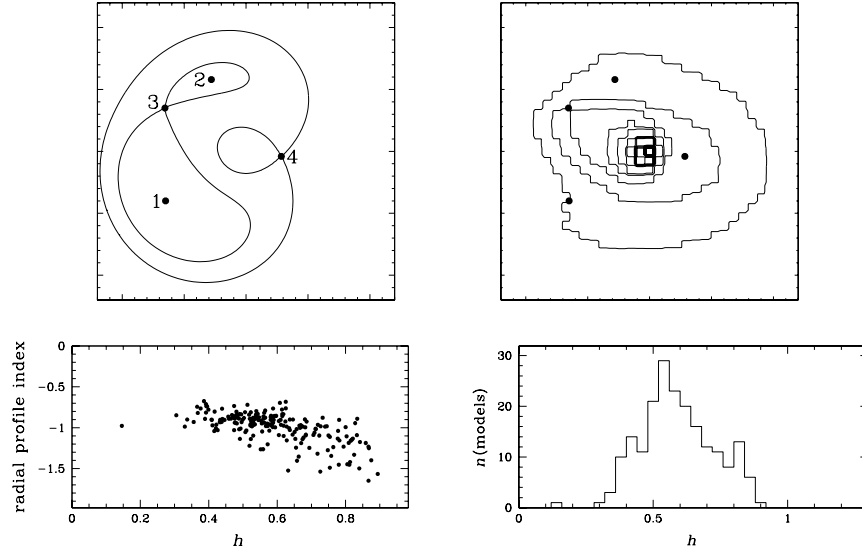
*PG 1115+080* Here the reconstructions use only image positions of the quasar (an inclined quad) and time delays [115,9]. Figure 23 shows the results, following the same plan as before. Again  $h$  has a large uncertainty, but is strongly correlated with the radial profile. But the distribution of  $h$  values is on average lower than for 0957+561. This promises improved results if results for several lenses are combined.



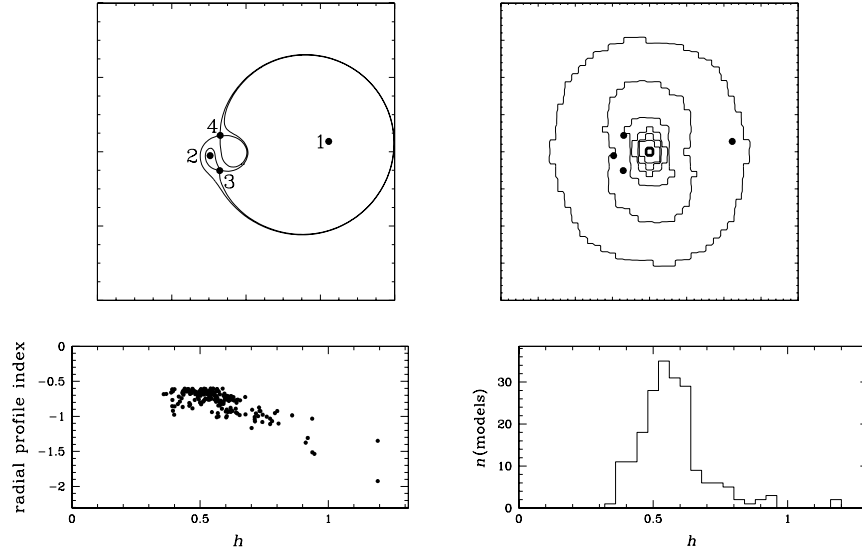
**Fig. 22.** Models of Q 0957+561. **Upper left:** schematic image configuration and saddle-point contours for the quasar. **Upper right:** ensemble-average reconstructed mass map; contours are  $\kappa = \frac{1}{3}, \frac{2}{3}, \dots$ . The four images marked are the quasar double and another double from a knot in the host galaxy. **Lower left:**  $h$  against radial index for all 200 models in the ensemble. **Lower right:** Histogram of  $h$  from the ensemble.



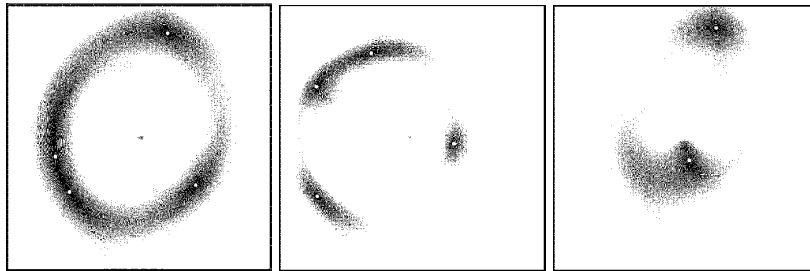
**Fig. 23.** Models of PG 1115+080. Panels arranged as in Figure 22.



**Fig. 24.** Models of B 1608+656. Panels arranged as in Figures 22 and 23.



**Fig. 25.** Models of RX J0911+055. Panels arranged as in Figures 22, 23, and 24.



**Fig. 26.** Ring and arc models resulting from plotting arrival-time surfaces with dense contours. **Left panel:** PG 1115+080 with contours 80 min apart; **middle panel:** B 1608+656 with contours 2 hr apart; **right panel:** 0957+561 with contours 1 day apart.

*B 1608+656* The reconstructions from this inclined quad use image positions and time delays [34]. The lensing galaxy in this system appears to be a binary [18], so the  $180^\circ$ -rotation symmetry is not imposed. Figure 24 shows the results. It is interesting that the mass profile comes out elongated towards the visible second galaxy, even though the reconstructions had no information about the light from the lensing galaxies.

*RX J0911+055* This is a short-axis quad with a preliminary time delay [47], and Figure 25 shows the results.

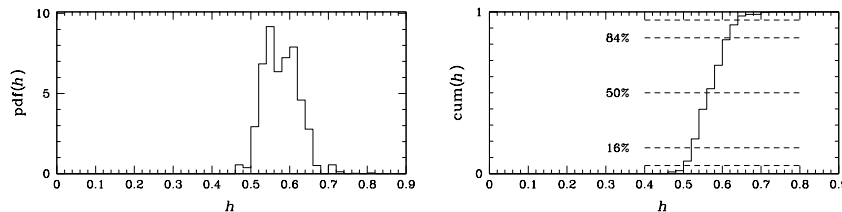
**Ring and arcs** The models described above are designed to fit images of (one or more) point sources. But having produced a model, one can check what sort of image it produces for extended sources. For a source with a conical or tent profile for brightness, the image is particularly easy to produce. We just have to make a dense contour map of the arrival-time surface for the center of the source and then view this map from a distance so that the contour lines blur into a grayscale [138,113]; the ratio

$$\frac{\tau\text{-spacing between contours}}{\text{thickness of contour lines}}$$

is proportional to the source size.

Figure 26 shows ring and arc images generated in this way from the ensemble-average models of PG 1115+080, B 1608+656, and Q 0957+561. These may be compared with published images of observed rings [50,18,57]. For PG 1115+080 and B 1608+656 the model and observed rings overlay extremely well. (Recall that the modeling procedure used no ring/arc data.) For Q 0957+561 the agreement is not so good: this may indicate simply that the models are less good, or it may indicate that the observed arc is the image not of the quasar host galaxy but another galaxy, possibly at different redshift.

**Combined  $h$  results** Returning to estimates of  $h$ , in Figure 27 we show the result of combining the  $h$  distributions from all four systems above.



**Fig. 27.** Combined  $h$  results from Q 0957+561, PG 1115+080, B 1608+656, and RX J0911+055: histogram on the left and cumulative plot on the right.

The combined result is

$$\begin{aligned} H_0 &= 56^{+6}_{-4} \quad (68\%) \\ &= 56^{+8}_{-6} \quad (90\%) \end{aligned}$$

The reference cosmology is Einstein de Sitter; for other cosmologies the numbers would be 5–10% higher.

## 4 Summary and future prospects

The present article concentrates on some selected aspects of quasar lensing, and in particular on their use for determining  $H_0$ . Lensed quasars have the advantage over other methods that they do not rely on the knowledge of any standard candle. The disadvantage is that precise modeling of the potential well responsible for the lensing effect is required.

It has been shown that, with present day instrumentation and efficient post-processing techniques, “mass production” of time-delays is possible, even using 2m class telescopes under average seeing conditions [21]. A typical precision on a time-delay determination is of the order of 10%, sometimes better. However, in many cases, most of the error on  $H_0$  comes from the lens models used to convert the time-delay into  $H_0$ . The problem can be overcome in two ways: (1) by using any prior knowledge available on (lens) galaxies and, (2) by improving the observations to constrain better the gravitational potential (main lens and any intervening cluster/group) in each individual lensed quasar.

The effectiveness of quasar lensing in producing a competitive value for  $H_0$  therefore depends on our knowledge of the physics of galaxies in general. Gravitational lensing itself should be able to set suitable constraints on galaxy mass profiles, for example through the statistical study of galaxy halos using



galaxy-galaxy lensing (see chapter on Dark Matter Halos). The development of two-dimensional spectrographs used to derive the full velocity field for many galaxies of all types will also yield important clues to the detailed mass distribution in galaxies. Both methods, direct or indirect, should constrain the degree of concentration of the mass in galaxies and the extent and shape of dark matter halos, two quantities which are often unconstrained in present days lens models and which imply the exploration of huge parameter spaces followed by a choice of a “best” model or a best family of models.

Improving the observations of individual lensed systems is also important. Lensing galaxies have an effect not only on quasars but also on galaxies in the vicinity of quasars. They should be seen under the form of arcs or arclets, as long as the angular resolution and depth are sufficient. The Advanced Camera for Surveys (ACS), on board of the HST, shall provide us at least with depth, hence with more background sources susceptible to be lensed, just as is the quasar. Since we usually observe only 2 or 4 quasar images, observing even a few arclets is a significant constraint for the lens model. In addition, with the depth of the ACS, most lensed quasars should show their distorted host galaxy, and bring even more constraints on the models. Constraining lens models using many arclets will probably become an efficient method with the launch of the Next Generation Space Telescope.

Measuring  $H_0$  is not the only application of quasar lensing. Once adequate observational constraints are available, or even assuming  $H_0$  can be measured independently by other methods, one shall use lensing to map the mass distribution in lensing galaxies and to infer basic parameters on the structure of quasars, using the chromatic variations due to microlensing events. Spectrophotometric monitoring is the next obvious observational step in the field, in order to enable such applications.

Whether we will learn about the mass distribution in galaxies once  $H_0$  is measured by other means, or the opposite, will depend on the speed of the progress made in the fields of the physics of galaxies, galaxy-galaxy lensing and on the possible discoveries of new methods to infer  $H_0$  with a high precision.

## References

1. H.M. AbdelSalam: D.Phil. thesis, Oxford (1998)
2. H.M. AbdelSalam, P. Saha, L.L.R. Williams: *NewAR* **42**, 157 (1998)
3. E. Agol, J. Krolik: *ApJ* **524**, 49 (1999)
4. C. Alard, R. Lupton: *ApJ* **503**, 325 (1998)
5. M.-C. Angonin-Willaime, G. Soucail, C. Vanderriest: *A&A* **291**, 411 (1994)
6. M.-C. Angonin-Willaime, C. Vanderriest, F. Courbin: *A&A* **347**, 434 (1999)
7. N. Bade, J. Siebert, S. Lopez, et al.: *A&A* **317**, L13 (1997)
8. N.T.J Bailey: *The Mathematical Approach to Biology and Medicine*, New York, Wiley (1967)
9. R. Barkana: *ApJ* **489**, 21 (1997)

10. R. Barkana, R.D. Blandford, D.W. Hogg: ApJ **513**, L91 (1999)
11. R. Bender, C. Moellenhoff: A&A **177**, 71 (1987)
12. G. Bernstein, P. Fischer, T.J. Tyson, et al.: ApJ **483**, L79 (1997)
13. G. Bernstein, P. Fischer: AJ **118**, 14 (1999)
14. A.D. Biggs, I.W.A. Browne, P. Helbig, et al.: MNRAS **304**, 349 (1999)
15. J. Binney, S. Tremaine: *Galactic Dynamics*, Princeton University Press (1987)
16. R.D. Blandford, R. Narayan: ApJ **310**, 568 (1986)
17. R.D. Blandford, H. Netzer, L. Woltjer, et al.: *Active galactic nuclei*, Springer-Verlag (1990)S
18. R.D. Blandford, G. Surpi: *Gravitational Lensing: Recent Progress and Future Goals*, ASP Conf. Series, Vol 237 (eds. Brainerd & Kochanek), p103 (2001)
19. I. Burud, F. Courbin, C. Lidman, et al.: ApJ **501**, L5 (1998)
20. I. Burud, J. Hjorth, A.O. Jaunsen, et al.: ApJ **544**, 117 (2000)
21. I. Burud: D. Phil. thesis, Université de Liège, Belgium (2001)
22. I. Burud, F. Courbin, P. Magain, et al.: A&A **383**, 71 (2002)
23. I. Burud, J. Hjorth, F. Courbin, et al.: A&A, in press (2002)
24. R.M. Campbell, J. Lehar, B.E. Corey, et al.: AJ **110**, 2566 (1995)
25. K. Chang, S. Refsdal: Nature **282**, 561 (1979)
26. V.H. Chavushyan, V.V. Vlasjuk, J.A. Stepanian, et al.: A&A **318**, L67 (1997)
27. J.-F. Claeskens, J. Surdej, M. Remy: A&A **305**, L9 (1996)
28. F. Courbin, P. Magain, C.R. Keeton, et al.: A&A **324**, L1 (1997)
29. F. Courbin, P. Magain, M. Kirkove, et al.: ApJ **529**, 1136 (2000)
30. F. Courbin, C. Lidman, G. Meylan, et al.: A&A **360**, 853 (2000)
31. C.P. O'Dea, S.A. Baum, C. Stanghellini, et al.: AJ **104**, 1320 (1992)
32. E.E. Falco, M.V. Gorenstein, I.I. Shapiro: ApJ **289**, L1 (1985)
33. C.D. Fassnacht, J.G. Cohen: AJ **115**, 377 (1998)
34. C.D. Fassnacht, T.J. Pearson, A. Readhead, et al.: ApJ **527**, 498 (1999)
35. C.D. Fassnacht, R.D. Blandford, J.G. Cohen, et al.: AJ **117**, 658 (1999)
36. P. Fischer, D. Schade, F. Barrientos: ApJ **503**, L127 (1998)
37. M. Fukugita, T. Futamase, M. Kasai, et al.: ApJ **393**, 3 (1992)
38. R. Florentin-Nielsen, K. Augustesen: IAUC **3945**, 2 (1984)
39. M.A. Garrett, R.J. Calder, R.W. Porcas, et al.: MNRAS **270**, 457 (1994)
40. J.R. Graham, M.C. Liu: ApJ **449**, L29 (1995)
41. M. D. Gregg, L. Wisotzki, R.H. Becker, et al.: AJ **119**, 2535 (2000)
42. D.B. Haarsma, J.N. Hewitt, J. Lehar, et al.: ApJ **510**, 64 (1999)
43. H.-J. Hagen, D. Reimers: A&A **357**, L29 (2000)
44. E.K. Hege, E.N. Hubbard, P.A. Strittmatter, et al.: ApJ **248**, L1 (1981)
45. J.N. Hewitt, E.L. Turner, D.P. Schneider, et al.: Nature **333**, 573 (1988)
46. J.N. Hewitt, E.L. Turner, C.R. Lawrence, et al.: AJ **104**, 968 (1992)
47. J. Hjorth, I. Burud, A.O. Jaunsen, et al.: *Gravitational Lensing: Recent Progress and Future Goals*, ASP Conf. Series, Vol 237 (eds. Brainerd & Kochanek), p125 (2001)
48. J. Hjorth, I. Burud, A.O. Jaunsen, et al.: astro-ph/0205124 (2002)
49. J. Huchra, M. Gorenstein, S. Kent, et al.: AJ **90**, 691 (1985)
50. C.D. Impey, E.E. Falco, C.S. Kochanek, et al.: ApJ **509**, 551 (1998)
51. M.J. Irwin, R.A. Ibata, G.F. Lewis, et al.: ApJ **505**, 529 (1998)
52. N. Jackson, A.G. De Bruyn, S. Myers, et al.: MNRAS **274**, L25 (1995)
53. N. Jackson, S. Nair, I.W.A. Browne, et al.: MNRAS **296**, 483 (1998)
54. A.O. Jaunsen, J. Hjorth: A&A **317**, L39 (1997)

55. R. Kayser, S. Refsdal: *A&A* **128**, 156 (1983)
56. C. R. Keeton, C. S. Kochanek: *ApJ* **487**, 42 (1997)
57. C.R. Keeton, E.E. Falco, C.D. Impey, et al.: *ApJ* **542**, 74 (2000)
58. J.-P. Kneib, J.G. Cohen, J. Hjorth: *ApJ* **544**, 35 (2000)
59. C.S. Kochanek, R. Narayan: *ApJ* **401**, 461 (1992)
60. C.S. Kochanek, E.E. Falco, R. Schild, et al.: *ApJ* **479**, 678 (1997)
61. C.S. Kochanek, E.E. Falco, C.D. Impey, et al.: *The CASTLE Survey*,  
<http://cfa-www.harvard.edu/glensdata/>
62. C.S. Kochanek, C.R. Keeton, B.A. McLeod: *ApJ*, **547**, 50 (2001)
63. L. V. E. Koopmans, A.G. Bruyn, D.R. De Marlow, et al.: *MNRAS* **303**, 727 (1999)
64. L.V.E. Koopmans, A.G. de Bruyn, E. Xanthopoulos, et al.: *A&A*, **356**, 391 (2000)
65. T. Kundic, E.L. Turner, W.N. Colley, et al.: *ApJ* **482**, 75 (1997)
66. G.I. Langston, D.P. Schneider, S. Conner, et al.: *AJ* **97**, 1283 (1989)
67. C.R. Lawrence, D.P. Schneider, D.P. Schmidt, et al.: *Science* **223**, 46 (1984)
68. C. R. Lawrence, D.P. Schneider, M. Schmidet, et al.: *Science* **223**, 46 (1984)
69. J. Lehar, G.I. Langston, A.D. Silber, et al.: *AJ* **105**, 847 (1993)
70. J. Lehar, A.J. Cooke, C.R. Lawrence, et al.: *AJ* **111**, 1812 (1996)
71. J. Lehar, B.F. Burke, S.R. Conner, et al.: *AJ* **114**, 58 (1997)
72. J. Lehar, E.E. Falco, C.S. Kochanek, et al.: *ApJ* **536**, 584 (2000)
73. G.F. Lewis, M.J. Irwin: *MNRAS* **276**, 103 (1996)
74. C. Lidman, F. Courbin, G. Meylan, et al.: *ApJ* **514**, L57 (1999)
75. C. Lidman, F. Courbin, J.-P. Kneib, et al.: *A&A* **364**, L62 (2000)
76. J.E.J. Lovell, D.L. Jauncey, J.E. Reynolds, et al.: *ApJ* **508**, L51 (1998)
77. P. Magain, J. Surdej, J.-P. Swings, et al.: *Nature* **334**, 327 (1988)
78. P. Magain, J. Surdej, C. Vanderriest: *A&A* **253**, L13 (1992)
79. P. Magain, F. Courbin, S. Sohy: *ApJ* **494**, 472 (1998)
80. R. McMahon, M. Irwin: *Gemini* **36**, 1 (1992)
81. D. R. Marlow, S.T. Myers, D. Rusin, et al.: *AJ* **118**, 654 (1999)
82. B. Moore, N. Katz, G. Lake: *IAUS* **171**, 203 (1996)
83. N.D. Morgan, A. Dressler, J. Maza, et al.: *AJ* **118**, 1444 (1999)
84. N.D. Morgan, R.H. Becker, M.D. Gregg, et al.: *AJ* **121**, 611 (2001)
85. N.D. Morgan, G. Chartas, M. Malm, et al.: *ApJ* **555**, 1 (2001)
86. J. A. Munoz, E.E. Falco, C.S. Kochanek, et al.: *ApJ* **546**, 769 (2001)
87. S.T. Myers, C.D. Fassnacht, S.G. Djorgovski, et al.: *ApJ* **447**, L5 (1995)
88. S.T. Myers, D. Rusin, C.D. Fassnacht, et al.: *AJ* **117**, 2565 (1999)
89. I. Newton: *Optics 2nd Ed., query 1* (1704)
90. R. Nityananda: *Current Science* **59**, 1044 (1990)
91. A. Oscoz, E. Mediavilla, L.J. Goicoechea, et al.: *ApJ* **479**, L89 (1997)
92. R. Ostensen, S. Refsdal, R. Stabell, et al.: *A&A* **309**, 59 (1996)
93. B. Paczyński: *ApJ* **301**, 503 (1986)
94. A.R. Patnaik, I.W.A. Browne, D. Walsh, et al.: *MNRAS* **259**, 1 (1992)
95. J. Pelt, R. Schild, S. Refsdal: *A&A* **336**, 829 (1998)
96. J. Pelt, R. Kayser, S. Refsdal, T. Schramm: *A&A* **305**, 97 (1996)
97. J. Pelt, W. Hoff, R. Kayser, et al.: *A&A* **286**, 775 (1994)
98. P.M. Phillips, M.A. Norbury, L. V. E. Koopmans, et al.: *MNRAS* **319**, L7 (2000)
99. F.P. Pijpers: *MNRAS* **289**, 933 (1997)

100. W.H. Press, G.B. Rybicki, J.N.W. Hewitt: ApJ **385**, 416 (1992)
101. K.U. Ratnatunga, E.J. Ostrander, R.E. Griffiths, et al.: ApJ **453**, L5 (1995)
102. S. Refsdal, R. Stabell, J. Pelt, et al.: A&A **360**, 10 (2000)
103. S. Refsdal: MNRAS **128**, 295 (1964)
104. S. Refsdal: MNRAS **128**, 307 (1964)
105. S. Refsdal: MNRAS **132**, 101 (1966)
106. J. Rhoads, S. Malhotra, T. Kundic: AJ **111**, 642 (1996)
107. D.H. Roberts, J. Lehar, J.N.W. Hewitt, et al.: Nature **352**, 43 (1991)
108. A.J. Romanowsky, C.S. Kochanek: ApJ **516**, 18 (1999)
109. E. Ros, J.C. Guirado, J.M. Marcaide, et al.: A&A **362**, 845 (2000)
110. D. Rusin, D.R. Marlow, M. Norbury, et al.: AJ **122**, 591 (2001)
111. P. Saha, L.L.R. Williams: MNRAS **292**, 148 (1997)
112. P. Saha: AJ **120**, 1654 (2000)
113. P. Saha, L.L.R. Williams: AJ **122**, 585 (2001)
114. N. Sanitt: Nature **234**, 199 (1971)
115. P.L. Schechter, C.D. Bailyn, B. Robert et al.: ApJ **475**, L85 (1997)
116. P. L. Schechter, M.D. Gregg, R.H. Becker, et al.: AJ **115**, 1371 (1998)
117. P. L. Schechter, J. Wambsganss: astro-ph/0204425 (2002)
118. R.E. Schild, B. Cholfín: ApJ **300**, 209 (1986)
119. R.E. Schild: AJ **100**, 1771 (1990)
120. M. Schmidt: Nature **197**, 1040 (1963)
121. P. Schneider: ApJ **319**, 9 (1987)
122. P. Schneider, C. Seitz: A&A **294**, 411 (1985)
123. I.I. Shapiro: Phys. Rev. Lett. **13**, 789 (1964)
124. R. Subrahmanyan, D. Narashima, A. Pramesh-Rao, et al.: MNRAS **246**, 263 (1990)
125. J. Surdej, P. Magain, J.-P. Swings, et al.: Nature **329**, 695 (1987)
126. J. Surdej, M. Remy, A. Smette et al.: *Proc. 31st Liege Int. Astroph. Coll. 'Gravitational Lenses in the Universe'*, p. 153 (1993)
127. A. Surdej: *Lensing bibliographie*: [http://vela.astro.ulg.ac.be/themes/extragal/-gravlens/bibdat/engl/glb\\_homepage.html](http://vela.astro.ulg.ac.be/themes/extragal/-gravlens/bibdat/engl/glb_homepage.html)
128. C.M. Sykes, I.W.A. Browne, N.J. Jackson, et al.: MNRAS **301**, 310 (1998)
129. J.L. Tonry, C.S. Kochanek: AJ **117**, 2034 (1999)
130. C.S. Trotter, J.N. Winn, J.N. Hewitt: ApJ **535**, 671 (2000)
131. C. Vanderriest, J. Schneider, G. Herpe, et al.: A&A **215**, 1 (1989)
132. D. Walsh, R.F. Carwell, R.J. Weymann: Nature **279**, 381 (1979)
133. J. Wambsganss, P. Schneider, B. Paczynski: ApJ **358**, L33 (1990)
134. J. Wambsganss: ApJ **386**, 19 (1992)
135. J. Wambsganss, R. Schmidt: NewAR **42**, 101 (2000)
136. S.J. Warren, P.C. Hewett, G.F. Lewis, et al.: MNRAS **278**, 139 (1996)
137. R. J. Weymann, D. Latham, J. Roger: Nature **285** 641 (1980)
138. L.L.R. Williams, P. Saha: AJ **119**, 439 (2000)
139. L. Wisotzki, T. Koehler, R. Kayser, et al.: A&A **278**, L15 (1993)
140. L. Wisotzki, T. Koehler, S. Lopez, et al.: A&A **315**, L405 (1996)
141. L. Wisotzki, O. Wucknitz, S. Lopez, et al.: A&A **339**, L73 (1998)
142. L. Wisotzki, N. Christlieb, M.C. Liu, et al.: A&A **348**, L41 (1999)
143. J. N. Winn, J.N. Hewitt, P.L. Schechter, et al.: AJ **120**, 286 (2000)
144. J.N. Winn, J.N. Hewitt, A.R. Patnaik, et al.: AJ **121**, 1223 (2001)
145. J.N. Winn, N.D. Morgan, J.N. Hewitt, et al.: AJ **123**, 10 (2002)

- 146. H.J. Witt: ApJ **403**, 530 (1993)
- 147. H.J. Witt, S. Mao, P.L. Schechter: ApJ **443**, 18 (1995)
- 148. P. R. Wozniak, A. Udalski, M. Szymanski, et al.: ApJ **540**, L65 (2000)
- 149. J.S.B. Wyithe, R.L. Webster, E.L. Turner, et al.: MNRAS **315**, 62 (2000)
- 150. J.S.B. Wyithe, R.L. Webster, E.L. Turner: MNRAS **318**, 1120 (2000)
- 151. E. Xanthopoulos, I.W.A. Browne, L.J. King: MNRAS **300**, 649 (1998)

## 5 Inventory of known systems

The numerous quasar surveys carried out to date and others still under way have led to the discovery of many lensed systems. Consequently, it is becoming increasingly difficult to keep track of all new lenses discovered. We try here to provide the reader with a list of all known cases; we apologize in advance to those who will not see their favorite system, probably because it has been too recently discovered. Note also that we list only lensed quasars. There are other cases of multiply imaged distant galaxies, discovered for example in HST deep fields (see for example [10]). Basic information such as coordinates, source and lens redshifts are given, together with the reference of the discovery paper. When several references are listed, the first ones corresponds to the discovery paper, and the others to the time delay measurement, when available. Time delays are given relative to the leading image. For example,  $\Delta t(\text{BA})$  means that image B is the leading image. Note finally that most objects have been observed or will be observed with the HST, either in the context or individual observing programs or through the CASTLE Survey whose main results are summarized at <http://cfa-www.harvard.edu/glensdata/>.

**Table 1.** List of confirmed doubles.

Object	Coords (2000)	Redshifts	Notes
Q 0142-100	$\alpha$ : 01h 45m 16.50s	$z_s=2.72$	
Surdej et al. [125]	$\delta$ : -09d 45m 17.00s	$z_l=0.49$	
CTQ 414	$\alpha$ : 01h 58m 41.44s	$z_s=1.29$	
Morgan et al. [83]	$\delta$ : -43d 25m 04.20s	$z_l=?$	
B 0218+357	$\alpha$ : 02h 21m 05.48s	$z_s=0.96$	$\Delta t(\text{BA}) = 10.5$
O’dea et al. [31]	$\delta$ : +35d 56m 13.78s	$z_l=0.68$	$\pm 0.4$ days
Biggs et al. [14]			
HE 0512-3329	$\alpha$ : 05h 14m 10.78s	$z_s=1.57$	
Gregg et al. [41]	$\delta$ : -33d 26m 22.50s	$z_l=0.93(?)$	
CLASS B0739+366	$\alpha$ : 07h 42m 51.20s	$z_s=?$	
Rusin et al. [?]rusin01)	$\delta$ : +36d 34m 43.70s	$z_l=?$	
MG 0751+2716	$\alpha$ : 07h 51m 41.46s	$z_s=3.20$	Ring
Lehar et al. [71]	$\delta$ : +27d 16m 31.35s	$z_l=0.35$	
HS 0818+1227	$\alpha$ : 08h 21m 39.10s	$z_s=3.12$	
Hagen & Reimers [43]	$\delta$ : +12d 17m 29.00s	$z_l=0.39$	
APM 08279+5255	$\alpha$ : 08h 31m 44.94s	$z_s=3.87$	
Irwin et al. [51]	$\delta$ : +52d 45m 17.70s	$z_l=?$	
SBS 0909+532	$\alpha$ : 09h 13m 01.05s	$z_s=1.38$	
Kochanek et al. [60]	$\delta$ : +52d 59m 28.83s	$z_l=0.83$	
RXJ 0921+4528	$\alpha$ : 09h 21m 12.81s	$z_s=1.66$	
	$\delta$ : +45d 29m 04.40s	$z_l=0.31$	
FBQ 0951+2635	$\alpha$ : 09h 51m 22.57s	$z_s=1.24$	
Schechter et al. [116]	$\delta$ : +26d 35m 14.10s	$z_l=?$	
BRI 0952-0115	$\alpha$ : 09h 55m 00.01s	$z_s=4.5$	
McMahon & Irwin [80]	$\delta$ : -01d 30m 05.00s	$z_l=?$	
Q 0957+561	$\alpha$ : 10h 01m 20.78s	$z_s=1.41$	$\Delta t(\text{BA}) = 417$
Walsh et al. [132]	$\delta$ : +55d 53m 49.40s	$z_l=0.36$	$\pm 3$ days
Kundić et al. [65]			
LBQS 1009-0252	$\alpha$ : 10h 12m 15.71s	$z_s=2.74$	
Surdej et al. [126]	$\delta$ : -03d 07m 02.00s	$z_l=?$	
Q 1017-207	$\alpha$ : 10h 17m 24.13s	$z_s=2.55$	
Claeskens et al. [27]	$\delta$ : -20d 47m 00.40s	$z_l=?$	
FSC 10214+4724	$\alpha$ : 10h 24m 37.58s	$z_s=2.29$	Ring
Graham & Liu [40]	$\delta$ : +47d 09m 07.20s	$z_l=?$	
B 1030+074	$\alpha$ : 10h 33m 34.08s	$z_s=1.54$	
Xanthopoulos et al. [151]	$\delta$ : +07d 11m 25.50s	$z_l=0.60$	
HE 1104-1805	$\alpha$ : 11h 06m 33.45s	$z_s=2.32$	$\Delta t(\text{AB}) = 260$
Wisotzki et al. [139]	$\delta$ : -18d 21m 24.20s	$z_l=0.73$	$\pm 90$ days
Wisotzki et al. [142]			

**Table 2.** List of confirmed doubles (continued)

Object	Coords (2000)	Redshifts	Notes
B 1127+385	$\alpha$ : 11h 30m 00.13s	$z_s=?$	
Koopmans et al. [63]	$\delta$ : +38d 12m 03.10s	$z_l=?$	
MG 1131+0456	$\alpha$ : 11h 31m 56.48s	$z_s=?$	
Hewitt et al. [45]	$\delta$ : +04d 55m 49.80s	$z_l=0.84$	
B 1152+199	$\alpha$ : 11h 55m 18.37s	$z_s=1.02$	
Myers et al. [88]	$\delta$ : +19d 39m 40.39s	$z_l=0.44$	
Q 1208+1011	$\alpha$ : 12h 10m 57.16s	$z_s=3.80$	
Magain et al. [78]	$\delta$ : +09d 54m 25.60s	$z_l=?$	
SBS 1520+530	$\alpha$ : 15h 21m 44.83s	$z_s=1.86$	$\Delta t(\text{BA}) = 130$
Chavushyan et al. [26]	$\delta$ : +52d 54m 48.60s	$z_l=0.71$	$\pm 6$ days
Burud et al. [23]			
MG 1549+3047	$\alpha$ : 15h 49m 12.37s	$z_s=?$	Ring
Lehar et al. [69]	$\delta$ : +30d 47m 16.60s	$z_l=0.11$	
B 1600+434	$\alpha$ : 16h 01m 40.45s	$z_s=1.59$	$\Delta t(\text{BA}) = 51$
Jackson et al. [52]	$\delta$ : +43d 16m 47.80s	$z_l=0.42$	$\pm 4$ days (radio)
Burud et al. [20]			$\Delta t(\text{BA}) = 47$
Koopmans et al. [64]			$\pm 11$ days (radio)
PMN J1632-0033	$\alpha$ : 16h 32m 55.98s	$z_s=3.42$	
Winn et al. [145]	$\delta$ : -00d 33m 04.50s	$z_l=?$	
FBS 1633+3134	$\alpha$ : 16h 33m 48.99s	$z_s=1.52$	
Morgan et al. [?]	$\delta$ : +31d 34m 11.90s	$z_l=?$	
MG 1654+1346	$\alpha$ : 16h 54m 41.83s	$z_s=1.74$	Ring
Langston et al. [66]	$\delta$ : +13d 46m 22.00s	$z_l=0.25$	
PKS 1830-211	$\alpha$ : 18h 33m 39.94s	$z_s=2.51$	Ring
Subrahmanyan et al. [124]	$\delta$ : -21d 03m 39.70s	$z_l=0.89$	$\Delta t(\text{BA}) = 26$
Lovell et al. [76]			$\pm 8$ days
PMN J1838-3427	$\alpha$ : 18h 38m 28.50s	$z_s=2.78$	
Winn et al. [143]	$\delta$ : -34d 27m 41.60s	$z_l=?$	
B 1938+666	$\alpha$ : 19h 38m 25.19s	$z_s=?$	Full ring
Rhoads et al. [106]	$\delta$ : +66d 48m 52.20s	$z_l=0.88$	
PMN J2004-1349	$\alpha$ : 20h 04m 07.02s	$z_s=?$	
Winn et al. [144]	$\delta$ : -13d 49m 31.65s	$z_l=?$	
B 2114+022	$\alpha$ : 21h 16m 50.75s	$z_s=?$	
	$\delta$ : +02d 25m 46.90s	$z_l=0.32/0.59$	
HE 2149-2745	$\alpha$ : 21h 52m 07.44s	$z_s=2.03$	$\Delta t(\text{BA}) = 103$
Wisotski et al. [140]	$\delta$ : -27d 31m 50.20s	$z_l=0.49$	$\pm 12$ days
Burud et al. [22]			
B 2319+051	$\alpha$ : 23h 21m 40.80s	$z_s=?$	
Wisotski et al. [140]	$\delta$ : +05d 27m 36.40s	$z_l=0.62$	

**Table 3.** List of central quads.

Object	Coords (2000)	Redshifts	Notes
CLASS B0128+437	$\alpha$ : 01h 31m 16.26s	$z_s=?$	
Phillips et al. [98]	$\delta$ : +43d 58m 18.00s	$z_l=?$	
HST 1411+5211	$\alpha$ : 14h 11m 19.60s	$z_s=2.81$	
Fischer et al. [36]	$\delta$ : +52d 11m 29.00s	$z_l=0.46$	
H 1413+117	$\alpha$ : 14h 15m 46.40s	$z_s=2.55$	
Magain et al. [77]	$\delta$ : +11d 29m 41.40s	$z_l=?$	
HST 14176+5226	$\alpha$ : 14h 17m 36.51s	$z_s=3.4$	
Ratnatunga et al. [101]	$\delta$ : +52d 26m 40.00s	$z_l=0.81$	
B 1555+375	$\alpha$ : 15h 57m 11.93s	$z_s=?$	
Marlow al. [81]	$\delta$ : +37d 21m 35.90s	$z_l=?$	
Q 2237+0305	$\alpha$ : 22h 40m 30.34s	$z_s=1.69$	
Huchra et al. [49]	$\delta$ : +03d 21m 28.80s	$z_l=0.04$	

**Table 4.** List of short axis quads.

Object	Coords (2000)	Redshifts	Notes
B 1422+231	$\alpha$ : 14h 24m 38.09s	$z_s=3.62$	
Patnaik et al. [94]	$\delta$ : +22d 56m 00.60s	$z_l=0.34$	

**Table 5.** List of long axis quads.

Object	Coords (2000)	Redshifts	Notes
RXJ 0911.4+0551	$\alpha$ : 09h 11m 27.50s	$z_s=2.8$	$\Delta t(\text{BA}) = 150$
Bade et al. [7]	$\delta$ : +05d 50m 52.00s	$z_l=0.77?$	$\pm 12$ days
Hjorth et al. [48]			
HST 12531-2914	$\alpha$ : 12h 53m 06.70s	$z_s=?$	
Ratnatunga et al. [101]	$\delta$ : -29d 14m 30.00s	$z_l=?$	
B 2045+265	$\alpha$ : 20h 47m 20.35s	$z_s=1.28$	
Fassnacht et al. [35]	$\delta$ : +26d 44m 01.20s	$z_l=0.87$	



**Table 6.** List of inclined quads.

Object	Coords (2000)	Redshifts	Notes
0047-2808	$\alpha$ : 00h 49m 41.89s	$z_s=3.60$	
Warren et al. [136]	$\delta$ : -27d 52m 25.70s	$z_l=0.49$	
HE 0230-2130	$\alpha$ : 02h 32m 33.10s	$z_s=2.16$	
Wisotzki et al. [142]	$\delta$ : -21d 17m 26.00s	$z_l=?$	
MG 0414+0534	$\alpha$ : 04h 14m 37.73s	$z_s=2.64$	
Hewitt et al. [46]	$\delta$ : +05d 34m 44.30s	$z_l=0.96$	
B 0712+472	$\alpha$ : 07h 16m 03.58s	$z_s=1.34$	
Jackson et al. [53]	$\delta$ : +47d 08m 50.00s	$z_l=0.41$	
PG 1115+080	$\alpha$ : 11h 18m 17.00s	$z_s=1.72$	$\Delta t(\text{AB}) = 11.7$
Weymann et al. [137]	$\delta$ : +07d 45m 57.70s	$z_l=0.31$	$\pm 1.2$ days
Schechter et al [116]			$\Delta t(\text{CB}) = 25.0$ $\pm 1.6$ days
B 1608+656	$\alpha$ : 16h 09m 13.96s	$z_s=1.39$	$\Delta t(\text{BA}) = 31$
Myers et al. [87]	$\delta$ : +65d 32m 29.00s	$z_l=0.63$	$\pm 7$ days
Fassnacht et al. [34]			$\Delta t(\text{BC}) = 36$ $\pm 7$ days $\Delta t(\text{BD}) = 76$ $\pm 10$ days
MG 2016+112	$\alpha$ : 20h 19m 18.15s	$z_s=3.27$	
Lawrence et al. [68]	$\delta$ : +11d 27m 08.30s	$z_l=1.01$	

**Table 7.** List of systems with more than four images.

Object	Coords (2000)	Redshifts	Notes
B 1359+154	$\alpha$ : 14h 01m 35.55s	$z_s=3.24$	6 images
Myers et al [88]	$\delta$ : +15d 13m 25.60s	$z_l=?$	
B 1933+507	$\alpha$ : 19h 34m 30.95s	$z_s=2.63$	10 images
Sykes et al. [128]	$\delta$ : +50d 25m 23.60s	$z_l=0.76$	