THE ASTROPHYSICAL JOURNAL, 378:L5–L9, 1991 September 1

GRAVITATIONAL LENSING, TIME DELAY, AND ANGULAR DIAMETER DISTANCE

RAMESH NARAYAN
Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138
Received 1991 May 28; accepted 1991 June 7

ABSTRACT

Combining a model of the mass distribution in the gravitational lens Q0957+561 with measurements of the velocity dispersion of the primary lensing galaxy and the time delay between the two images, the Hubble constant has been recently estimated to be $H_0 = 37 \pm 14$ km s^{-1} Mpc^{-1}. The result depends weakly on the assumed value of q_0 and may have systematic errors due to nonuniqueness of the model. It is shown here that the quantity that is most directly measured using a gravitational lens system is not H_0 but rather the angular diameter distance $d_{\rm OL}$ from the observer to the lens. The determination of $d_{\rm OL}$ requires neither a knowledge of the distance to the source nor any cosmological assumption other than local isotropy and homogeneity transverse to the line of sight, again with possible systematic effects. In the case of Q0957+561, where the lens is at a redshift of $z_L = 0.36$, the result is $d_{\rm OL} = 1700 \pm 600$ Mpc. If the mass distribution of the cluster surrounding the primary lens in Q0957+561 can be determined through independent observations, then it may be possible to estimate both H_0 and q_0.

Subject headings: cosmology — dark matter — galaxies: distances — gravitational lenses

1. INTRODUCTION

Long before the discovery of the first gravitational lens, Refsdal (1964, 1966) proposed that a measurement of the time delay between the multiple images produced by such a lens could be used to measure the Hubble constant H_0 and possibly the deceleration parameter q_0. With the discovery of Q0957+561 (Walsh, Carswell, & Weymann 1979), and subsequently several other examples of lensing, this idea has now become a practical possibility. Most of the attention has been focused on Q0957+561, in part because this has been the most promising source for the measurement of time delay but also because VLBI observations (e.g., Gorenstein et al. 1988) have provided reasonably strong constraints on the mass model of the lens.

Using a simple but realistic model of Q0957+561 that fits all the observations, Falco, Gorenstein, & Shapiro (1991, hereafter FGS) provided the following formula for the Hubble constant, where H_0 is expressed in terms of observables,

$$H_0 = (90 \pm 10)\left(\frac{\sigma}{390 \text{ km s}^{-1}}\right)^2 \left(\frac{\Delta t}{1 \text{ yr}}\right)^{-1} \text{ km s}^{-1} \text{ Mpc}^{-1} . \quad (1)$$

Here, σ is the velocity dispersion of the lensing galaxy (which is at a redshift of $z_L = 0.36$) and Δt is the time delay between the two lensed images of the background quasar (at redshift $z_S = 1.41$). The error estimate includes uncertainties in the mass distribution of the lens model (however, see Kochanek 1991). Recent measurements indicate that $\sigma = 303 \pm 50$ km s^{-1} (Rhee 1991) and $\Delta t = 536 \pm 12$ (95% confidence) days (Lehar et al. 1991; Press, Rybicki, & Hewitt 1991). These measurements give $H_0 = 37 \pm 14$ km s^{-1} Mpc^{-1}, though Kochanek (1991) and Roberts et al. (1991) argue that the velocity dispersion of the halo of the lensing galaxy may be larger than that of the stars by a factor of up to $(1.5)^{1/2}$ (cf. Turner, Ostriker, & Gott 1984), in which case $H_0 = 56 \pm 20$ km s^{-1} Mpc^{-1}. In any case it seems that Refsdal's original proposal has finally become a reality.

Although it is usual in this subject to express the results in terms of H_0, it is well known that what is in effect measured is an angular diameter distance, which is then converted to H_0 through a cosmographic model. The obvious question then is the following. Exactly which angular diameter distance, or combination of distances, does one measure through gravitational lensing? There are at least four candidates: the angular diameter distance from the observer to the lens $d_{\rm OL}$, the observer-source distance $d_{\rm OS}$, the lens-source distance $d_{\rm LS}$, and the following effective distance which appears frequently in the theory of gravitational lensing,

$$d_{\rm eff} \equiv d_{\rm OL}\, d_{\rm OS}/d_{\rm LS} . \quad (2)$$

It is shown in this *Letter* that gravitational lensing uniquely determines the observer-lens distance $d_{\rm OL}$ and that the result requires neither a knowledge of the distance to the source nor any assumption regarding the cosmological model of the universe. The proof of this result is given in § 2, and some consequences are explored in § 3.

2. PROOF

Consider a reference null geodesic that reaches the observer after passing through the center of the lensing galaxy (defined to be a projected point with no gravitational deflection). Erect a "source plane" perpendicular to this geodesic, and represent the position of the source by the transverse vector $\boldsymbol{r}_S$ on this plane. In the absence of lensing, the angular position of the source as seen by the observer will be $\boldsymbol{\theta}_S = \boldsymbol{r}_S/d_{\rm OS}$. Consider now a light ray that travels from the source to a point at transverse position $\boldsymbol{r}_L$ on the "lens plane" and then travels from there to the observer. The observer sees the ray arriving along the direction $\boldsymbol{\theta}_I = \boldsymbol{r}_L/d_{\rm OL}$ (we use I for image). The excess "geometrical" time delay along this ray, compared to the reference ray, may be written under quite general conditions of isotropy in the form (see the Appendix)

$$t_{\rm geom} = \frac{(1 + z_L)}{2}\, d_{\rm eff}(\boldsymbol{\theta}_I - \boldsymbol{\theta}_S)^2 . \quad (3)$$

Note that here and below we use units such that $c = 1$.

Consider next the "gravitational" time delay $t_{\rm grav}$ introduced by the lens. Following FGS let us model the lens as consisting of a lensing galaxy plus dark matter associated with a surrounding cluster. We may then write the gravitational time delay measured by a clock in the lens frame as

$$t_{\rm grav} = -\sigma^2 d_{\rm OL}[\chi(\theta_I) + \tfrac{1}{2}q_S(\boldsymbol{\theta}_I - \boldsymbol{\theta}_S)^2 + \tfrac{1}{2}q_A(\boldsymbol{\theta}_I - \boldsymbol{\theta}_S)^2] . \quad (4)$$

The first term on the right-hand side of equation (4) is the time delay due to the galaxy written in a general form that encompasses most cases of lens modeling. The function $\chi(\theta_I)$, which represents the angular "shape" of the line-of-sight–integrated two-dimensional potential of the galaxy, will normally be known to within a few adjustable parameters, either from an assumed model (e.g., a King profile with an additional central point mass as in the FGS model of Q0957+561) or from an observed surface brightness distribution coupled with the assumption of a constant mass-to-light ratio (e.g., the model by Schneider et al. 1988 for Q2237+031). To convert the dimensionless $\chi(\theta_I)$ to a two-dimensional potential, one needs to normalize first by $d_{\rm OL}$ to go from an angular scale to a physical linear scale and second, by another factor that characterizes the dynamical mass of the galaxy. The latter factor is written as σ^2, where σ represents an observable velocity. If the galaxy is an isothermal system, then σ is simply the velocity dispersion of the stars and dark matter. In more complicated models, σ will represent some characteristic velocity of the system, for instance the rotation velocity or velocity dispersion at some fiducial angular position relative to the center. Although the details will depend on the particular model used to describe the mass distribution of the galaxy, nevertheless, the form written in equation (4) must almost always be valid, as can be seen purely from a dimensional argument.

The rest of equation (4) corresponds to a simple model of the dark matter associated with the cluster (FGS). The second term on the right-hand side describes an isotropic focusing term, while the third term represents shear introduced by the dark matter. Both terms are scaled by $\sigma^2 d_{\rm OL}$ for later convenience. This model of the dark matter is not as general as the galaxy model considered above but is arguably adequate so long as the cluster is smooth on the scale of the image separation. Note that in equation (4) q_S is a scalar parameter while q_A is a traceless symmetric 2 × 2 matrix described by two parameters which may be taken to be the magnitude and orientation of the shear. All three parameters are assumed to be essentially unconstrained by the observations, except for the requirement that $q_S \geq 0$ since dark matter cannot have negative mass density. The term involving q_A in equation (4) is a quadratic form that should technically be written as $(\boldsymbol{\theta}_I - \boldsymbol{\theta}_S)_R\, q_A(\boldsymbol{\theta}_I - \boldsymbol{\theta}_S)_c$ distinguishing row and column vectors of $(\boldsymbol{\theta}_I - \boldsymbol{\theta}_S)$; an abbreviated notation has been used for simplicity. Another minor point is that FGS express the dark matter time delay in terms of $\boldsymbol{\theta}_I^2$ rather than $(\boldsymbol{\theta}_I - \boldsymbol{\theta}_S)^2$ as written here. It is easily shown that this difference is irrelevant since it merely redefines the unobservable "true" position of the source. (Gorenstein, Falco, & Shapiro 1988 call this a "prismatic" transformation.)

The total time delay at the observer is given by

$$\begin{aligned} t_{\rm tot} &= t_{\rm geom} + (1 + z_L)t_{\rm grav} \\ &= (1 + z_L)\sigma^2 d_{\rm OL}[\tfrac{1}{2}\zeta(\boldsymbol{\theta}_I - \boldsymbol{\theta}_S)^2 - \chi(\boldsymbol{\theta}_I) - \tfrac{1}{2}q_A(\boldsymbol{\theta}_I - \boldsymbol{\theta}_S)^2] , \quad (5) \end{aligned}$$

where the scalar parameter ζ is defined by

$$\zeta \equiv \frac{d_{\rm OS}}{\sigma^2 d_{\rm LS}} - q_S . \quad (6)$$

Fermat's principle states that the images are located at the positions $\boldsymbol{\theta}_I$ at which $t_{\rm tot}$ is an extremum for fixed $\boldsymbol{\theta}_S$ (Schneider 1985; Blandford & Narayan 1986). Setting $\partial t_{\rm tot}/\partial\boldsymbol{\theta}_I = 0$, we thus obtain the lens equation that must be satisfied by each of the multiple images,

$$(\zeta - q_A)(\boldsymbol{\theta}_I - \boldsymbol{\theta}_S) = d\chi/d\boldsymbol{\theta}_I . \quad (7)$$

Note that $d\chi/d\boldsymbol{\theta}_I \equiv \boldsymbol{\alpha}(\boldsymbol{\theta}_I)/\sigma^2$, where $\boldsymbol{\alpha}(\boldsymbol{\theta}_I)$ is the usual deflection angle due to the primary lensing galaxy at impact parameter $d_{\rm OL}\,\boldsymbol{\theta}_I$. Differentiating equation (7) once more with respect to $\boldsymbol{\theta}_S$ gives the magnification matrix at each image,

$$M \equiv \frac{\partial\boldsymbol{\theta}_I}{\partial\boldsymbol{\theta}_S} = \frac{\zeta - q_A}{\zeta - q_A - d^2\chi/d\boldsymbol{\theta}_I^2} . \quad (8)$$

Equations (5)–(8) are the basic relations needed for the modeling and interpretation of multiply imaged quasars. Except for equation (5) which has the dimensions of time, the others are dimensionless. The data to be fitted are the positions θ_I of the n images of the source, and $n - 1$ relative magnifications. In favorable cases such as Q0957+561 observations provide the complete 2 × 2 relative magnification matrix, but more often one merely obtains the determinants of the matrices. The parameters at the disposal of the modeler are the scalar ζ (one parameter), the matrix q_A (two parameters), the source position $\boldsymbol{\theta}_S$ (two parameters), and the internal degrees of freedom of $\chi(\theta_I)$ which, in the case of the FGS model of Q0957+561, consist of two dimensionless parameters, viz., an angular core radius and the relative strength of the central point mass. Having solved for the parameters, if one also has a measurement of the differential time delay between images and/or the dynamical velocity σ, then various deductions can be made as discussed below. Points (1)–(4) are already known, but points (5)–(7) seem to be new.

1. Following Refsdal, suppose we do not include any dark matter in the model. Setting $q_S = 0$, $q_A = 0$, equation (6) shows that the model-fitted value of ζ directly gives σ^2, provided that $d_{\rm OS}/d_{\rm LS}$ is known (from z_S and an assumed q_0). Substituting this into equation (5), one can then use a measured time delay to obtain H_0 (corresponding to the assumed q_0). Alternatively, if Δt is not available but σ is measured, then equation (6) gives $d_{\rm OS}/d_{\rm LS}$, which may be used to deduce q_0. Finally, if Δt and σ are both measured, then H_0 as well as q_0 may be obtained.

2. If a more realistic model including dark matter is considered, then equation (6) reveals the fundamental degeneracy identified by Falco, Gorenstein, & Shapiro (1985). The expression for ζ involves two parameters, σ and q_S. This means that the solution for ζ obtained from the lens model corresponds not to a unique model but rather to a one-parameter *family* of models in which the mass in the galaxy (described by σ) and the mass in the cluster (given by q_S) are mutually adjusted. Consequently, since the time delay depends on σ, a measurement of Δt does not provide a unique solution for H_0. Nevertheless, since q_S cannot be negative (dark matter cannot have negative density), equation (6) may still be used to estimate the maximum value of σ and hence the maximum value of H_0 (Borgeest & Refsdal 1984). For the FGS model of Q0957+561, the results are $\sigma_{\rm max} = 390$ km s^{-1} and $(H_0)_{\rm max} = (90 \pm 10)(\Delta t/1\ {\rm yr})^{-1}$ km s^{-1} Mpc^{-1} for $q_0 = \frac{1}{2}$. Using the measured Δt, this gives $(H_0)_{\rm max} = 61 \pm 7$ km s^{-1} Mpc^{-1}.

3. Equation (5) shows that if the time delay is measured, then one obtains an estimate of $\sigma^2 d_{\rm OL}$. It is easy to verify that this quantity determines the mass distribution of the lens expressed in angular coordinates. Thus, one can obtain the surface mass density per unit solid angle across the lens or the

mass enclosed within any angular contour in the lens. These estimates will depend of course on the assumed "shape" of the mass distribution, which goes into the assumed $\chi(\theta_I)$, but will not require any knowledge of H_0, q_0, or the mass of the cluster. This interesting result was discovered by Borgeest (1986).

4. If both Δt and σ are measured, then the model degeneracy discussed in point (2) above is broken and a unique mass model is obtained. This leads to a specific solution for H_0 corresponding to an assumed q_0, as discussed in § 1.

5. If σ is measured and ζ, q_A and the internal parameters in $\chi(\theta_I)$ are all obtained through the lens modeling, then the only unknown in the right-hand side of equation (5) is $d_{\rm OL}$. Therefore, a measurement of the differential time delay between a pair of images provides a unique solution for $d_{\rm OL}$. This proves the main result claimed in this paper. The key point is that the other angular diameter distances, $d_{\rm OS}$ and $d_{\rm LS}$, enter the relevant equations only in the particular dimensionless combination ζ. Therefore, once ζ is fitted in the model, the actual value of these distances is unimportant. Equation (1) from FGS may now be recast into the following result for the angular diameter distance to the lensing galaxy in Q0957+561,

$$d_{\rm OL} = (700 \pm 80)\left(\frac{\sigma}{390 \text{ km s}^{-1}}\right)^{-2}\left(\frac{\Delta t}{1 \text{ yr}}\right) \text{ Mpc} . \qquad (9)$$

For a measured $\sigma = 303 \pm 50$ km s^{-1} and $\Delta t = 536 \pm 12$ days, this gives $d_{\rm OL} = 1700 \pm 600$ Mpc, or $d_{\rm OL} = 1100 \pm 400$ Mpc if one includes the correction factor of 1.5 (Kochanek 1991; Roberts et al. 1991). It should be emphasized that equation (9) is independent of q_0. It is even independent of the quasar redshift!

6. Given σ and ζ, equation (6) provides a lower bound on $d_{\rm OS}/d_{\rm LS}$, viz., $d_{\rm OS}/d_{\rm LS} > \sigma^2\zeta$. In the FGS model, $\zeta/c^2 = 1.10 \times 10^{-5}$ (km s^{-1})$^{-2}$, hence

$$\frac{d_{\rm OS}}{d_{\rm LS}} > 1.67\left(\frac{\sigma}{390 \text{ km s}^{-1}}\right)^2 . \qquad (10)$$

In favorable cases this might be translated into a bound on q_0. For $\sigma \sim 300$ km s^{-1} as measured by Rhee (1991), there is no useful bound. However, if the true σ of the halo of the lensing galaxy is $300(1.5)^{1/2}$ km s^{-1} (Kochanek 1991; Roberts et al. 1991), then $d_{\rm OS}/d_{\rm LS} > 1.50$, which correponds to $q_0 > 0.1$.

7. If the dark matter density parameter q_S is measured, for example by mapping the spatial distribution of the galaxies in the cluster and measuring their velocity dispersion, then equation (6) would provide a unique solution for $d_{\rm OS}/d_{\rm LS}$ rather than merely a bound. One would then have a solution for both H_0 and q_0.

3. DISCUSSION

The main result of this *Letter* is that a gravitational lens model coupled with a measured time delay gives an estimate of the angular diameter distance $d_{\rm OL}$ from the observer to the lens. The estimate is independent of cosmological parameters such as H_0, q_0, or λ and is valid so long as the geometrical and gravitational time delays may be written in the forms assumed in equations (3) and (4). Equation (3) is applicable, provided the universe is isotropic and locally homogeneous transverse to the line of sight (see the Appendix), while equation (4) requires that the cluster dark matter be sufficiently smooth on the scale of image separation (§ 2). In the case of Q0957+561, using the FGS model and current observations, it is estimated that $d_{\rm OL} = 1700 \pm 600$ Mpc provided one assumes that the measured velocity dispersion reflects the mass distribution of the whole galaxy including its halo. If one introduces a correction factor of $(1.5)^{1/2}$ (Kochanek 1991; Roberts et al. 1991) between the velocity dispersion of the stars and the halo, then $d_{\rm OL} = 1100 \pm 400$ Mpc. The quoted error estimates are dominated by the uncertainty in the measured value of σ. Improved observations would be most useful.

Interestingly, the estimation of $d_{\rm OL}$ from observations of a gravitational lens requires no knowledge of the distance to the source. This is somewhat obvious in retrospect. The point is that the component of the dark matter that is described by the parameter q_S behaves exactly like a perfect converging lens. Since q_S is considered a free parameter, this means that the effective focal length of the dark matter lens is undetermined. Consequently, the distance to the source too is unconstrained. Instead, it is the particular combination of parameters making up ζ that turns out to be important, but this is determined directly in the model-fitting process.

The fact that one can measure the angular diameter distance to the lens without knowing the distance to the source means of course that it is not necessary to know the redshift of the source. This may turn out to be a useful feature of the method. The most promising lens candidates for cosmography are likely to be radio-loud sources because (1) they tend to have extended structure such as rings (e.g., MG 1131+0456; Hewitt et al. 1988), and (2) VLBI observations, where possible, can potentially provide much more detailed information on the magnification matrix than one normally obtains with nonradio sources. This *Letter* shows that a radio source with no optical identification could still be a powerful tool for cosmography provided the redshift and velocity dispersion of the lens are measured, and a time delay is obtained.

The insensitivity to source distance also means that any deviation of the universe from homogeneity can be tolerated so long as it occurs on the far side of the lens and the deviations are on a sufficiently large angular scale that the concept of angular diameter distance survives over the scale of the image separations ($\sim 10''$, ~ 100 kpc). In the case of Q0957+561, there are unconfirmed reports of a second cluster of galaxies behind the primary lens. According to the result proved in this *Letter*, the presence of such a cluster makes no difference to the determination of $d_{\rm OL}$ so long as its effect is merely to introduce an additional convergence, i.e., another contribution to the unknown q_S. In fact, a stronger statement can be made. Even if the second cluster contributes both convergence and shear—indeed, even if there are several clusters at different redshifts behind the lens contributing variable amounts of convergence and shear—these will be absorbed into q_S and q_A, and the determination of $d_{\rm OL}$ is unaffected. Only if the universe has inhomogeneities between the lens and the observer is there a significant effect. In this case, if the additional contribution is a pure convergence, then the present analysis continues to be valid, but gravitational lensing will now measure the *local* angular diameter distance along the line of sight to the lens, which will in general deviate from the global average for the universe as a whole. However, if there is also shear, then even the concept of a scalar $d_{\rm OL}$ breaks down and one must replace it by a 2×2 matrix. This will introduce two additional parameters in the model. (It is sufficient to restrict attention to a symmetric matrix for $d_{\rm OL}$, since the antisymmetric part corresponds to an unmeasurable and irrelevant rotation). The necessary extension of the modeling procedure is straightforward, but the additional parameters will require more observational constraints for a reliable solution.

An interesting sidelight of the present analysis is that Q0957+561 may already be on the verge of setting weak limits on q_0 (see § 2, point [6]). One may in fact be able to do better. If the mass distribution of the surrounding cluster in Q0957+561 can be observationally determined, then one may be able actually to determine both H_0 and q_0. Detailed observations of the spatial and velocity structure of the cluster would be very useful.

Another point worth highlighting is that the surface mass density of the lens per unit solid angle can be estimated from a measured time delay without knowing the distance to the lens, the velocity dispersion of the galaxy, or the mass of the cluster (Borgeest 1986, § 2, point [3]). This means that one can obtain the absolute mass of the galaxy out to any isophote, the reliability of the estimate being limited only by the degree of uniqueness of the lens model. Unfortunately, this cannot be translated to a mass-to-light ratio for the lens because one cannot estimate the luminosity of the galaxy—the "light"—without knowing the distance to the galaxy.

The author gratefully acknowledges useful discussions with R. D. Blandford, E. E. Falco, J. N. Hewitt, and W. H. Press. This work was supported in part by the National Science Foundation under a Presidential Young Investigator award (grant AST 8957107).

APPENDIX

Consider a ray traveling from the point $\boldsymbol{r}_L$ on the lens plane to the point $\boldsymbol{r}_O$ on the "observer plane." The excess geometrical time of this ray as measured by the observer, compared to a reference ray that travels from $\boldsymbol{r}_L = 0$ to $\boldsymbol{r}_O = 0$, may be written rather generally in the form (note $c = 1$)

$$t_{\rm OL}(\boldsymbol{r}_O, \boldsymbol{r}_L) = \frac{1}{2}\left(\frac{r_O^2}{d_1} - \frac{2\boldsymbol{r}_O \cdot \boldsymbol{r}_L}{d_2} + \frac{r_L^2}{d_3}\right). \tag{A1}$$

In the lens plane, the same ray has a travel time equal to $t_{\rm OL}/(1 + z_L)$. In writing equation (A1), it is assumed that the distances d_1, d_2, d_3 are simple scalars (their physical meaning is given below), which requires that the medium be isotropic. Further, by employing a quadratic dependence of the time delay on $\boldsymbol{r}_O$, $\boldsymbol{r}_L$, it is being assumed that the medium is homogeneous transverse to the ray over the region of interest ($\sim 10''$, ~ 100 kpc). However, no assumption of homogeneity is made along the direction of the ray. In the language of geometrical optics, $t_{\rm OL}(\boldsymbol{r}_O, \boldsymbol{r}_L)$ is the so-called point characteristic corresponding to the two end-points of the ray (cf. Born & Wolf 1980).

In the observer frame, the angle between the above ray and the reference ray is

$$\boldsymbol{\theta}_{\rm OL} \equiv \boldsymbol{\theta}_I = -\frac{\partial t_{\rm OL}}{\partial \boldsymbol{r}_O} = -\frac{\boldsymbol{r}_O}{d_1} + \frac{\boldsymbol{r}_L}{d_2}. \tag{A2}$$

This shows that d_1 is the "parallax distance" from O to L and that $d_2 \equiv d_{\rm OL}$ is the observer-lens angular diameter distance. The angle made by the ray in the lens plane is

$$\boldsymbol{\theta}_{\rm LO} = \frac{1}{(1 + z_L)}\frac{\partial t_{OL}}{\partial \boldsymbol{r}_L} = -\frac{\boldsymbol{r}_O}{(1 + z_L)d_2} + \frac{\boldsymbol{r}_L}{(1 + z_L)d_3}. \tag{A3}$$

Thus, $(1 + z_L)d_3$ is the parallax distance from the lens to the observer.

We may similarly write the geometrical time as measured in the lens frame for a ray traveling from $\boldsymbol{r}_S$ (in the source plane) to $\boldsymbol{r}_L$ as

$$t_{\rm LS}(\boldsymbol{r}_L, \boldsymbol{r}_S) = \frac{1}{2}\left(\frac{r_L^2}{d_4} - \frac{2\boldsymbol{r}_L \cdot \boldsymbol{r}_S}{d_5} + \frac{r_S^2}{d_6}\right). \tag{A4}$$

As before we find that $d_5 \equiv d_{\rm LS}$ is the lens-source angular diameter distance.

The total geometrical time as measured by the observer is $t_{\rm geom} = t_{\rm OL} + (1 + z_L)t_{\rm LS}$. In the absence of any deflection by the lens, the trajectory of the ray that goes from $\boldsymbol{r}_S$ to $\boldsymbol{r}_O = 0$ may be found by minimizing $t_{\rm geom}$ with respect to $\boldsymbol{r}_L$ (Fermat's principle). If we define the angle made by this ray at the observer as $\boldsymbol{\theta}_S$, we find

$$\boldsymbol{r}_S = d_2 d_5\left[\frac{1}{d_4} + \frac{1}{(1 + z_L)d_3}\right]\boldsymbol{\theta}_S \equiv d_{\rm OS}\boldsymbol{\theta}_S. \tag{A5}$$

This expresses the observer-source angular diameter distance $d_{\rm OS}$ in terms of the other distances defined above.

It is now a simple matter to compute the excess geometrical time for a ray to travel from $\boldsymbol{r}_S = d_{\rm OS}\boldsymbol{\theta}_S$ to $\boldsymbol{r}_L = d_{\rm OL}\boldsymbol{\theta}_I$ and then to the observer at $\boldsymbol{r}_O = 0$. The result, as measured by the observer, is

$$t_{\rm geom}(\boldsymbol{\theta}_I, \boldsymbol{\theta}_S) = \frac{(1 + z_L)}{2}\frac{d_{\rm OL}d_{\rm OS}}{d_{\rm LS}}(\boldsymbol{\theta}_I - \boldsymbol{\theta}_S)^2, \tag{A6}$$

which is the formula used in the text.

To interpret equation (A6) physically, note that the deflection angle $\boldsymbol{\alpha}$ of the ray in the lens plane and its lateral offset $\Delta\boldsymbol{r}_L$ are given by

$$\boldsymbol{\alpha} = \frac{d_{\rm OS}}{d_{\rm LS}}(\boldsymbol{\theta}_I - \boldsymbol{\theta}_S), \qquad \Delta\boldsymbol{r}_L = d_{\rm OL}(\boldsymbol{\theta}_I - \boldsymbol{\theta}_S). \tag{A7}$$

Thus equation (A6) becomes

$$t_{\rm geom} = \frac{(1+z_L)}{2}\,\Delta \boldsymbol{r}_L \cdot \boldsymbol{\alpha}\,. \tag{A8}$$

This result is self-evident when one thinks in terms of wavefronts (cf. Kayser & Refsdal 1983) and remembers that these are assumed to be quadratic. Apart from the redshift factor, equation (A8) is no more than a restatement of the trivial result that, if $f(x) = ax^2/2$, then $f(x) = xf'(x)/2$.

If there is anisotropy, then the inverse distances d_1^{-1}–d_6^{-1} in equations (A1) and (A4) have to be replaced by (2×2) matrices. The present analysis continues to be valid provided the equations are appropriately generalized to allow for the matrices (see Kovner 1987).

REFERENCES

Blandford, R., & Narayan, R. 1986, ApJ, 310, 568
Borgeest, U. 1986, ApJ, 309, 467
Borgeest, U., & Refsdal, S. 1984, A&A, 141, 318
Born, M., & Wolf, E. 1980, Principles of Optics (Oxford: Pergamon)
Falco, E. E., Gorenstein, M. V., & Shapiro, I. I. 1985, ApJ, 289, L1
———. 1991, ApJ, 372, 364 (FGS)
Gorenstein, M. V., Cohen, N. L., Shapiro, I. I., Rogers, A. E. E., Bonometti, R. J., Falco, E. E., Bartel, N., & Marcaide, J. M. 1988, ApJ, 334, 42
Gorenstein, M. V., Falco, E. E., & Shapiro, I. I. 1988, ApJ, 327, 693
Hewitt, J. N., Turner, E. L., Schneider, D. P., Burke, B. F., Langston, G. I., & Lawrence, C. R. 1988, Nature, 333, 537
Kayser, R., & Refsdal, S. 1983, A&A, 128, 156
Kochanek, C. S. 1991, ApJ, in press
Kovner, I. 1987, ApJ, 318, L1
Lehar, J., Hewitt, J. N., Roberts, D. H., & Burke, B. F. 1991, ApJ, submitted
Press, W. H., Rybicki, G. B., & Hewitt, J. N. 1991, ApJ, submitted
Refsdal, S. 1964, MNRS, 128, 295
———. 1966, MNRAS, 132, 101
Rhee, G. 1991, Nature, 350, 211
Roberts, D. H., Lehar, J., Hewitt, J. N., & Burke, B. F. 1991, Nature, in press
Schneider, D. P., Turner, E. L., Gunn, J. E., Hewitt, J. N., Schmidt, M., & Lawrence, C. R. 1988, AJ, 95, 1619
Schneider, P. 1985, A&A, 143, 413
Turner, E. L., Ostriker, J. P., & Gott, J. R., III. 1984, ApJ, 284, 1
Walsh, D., Carswell, R. F., & Weymann, R. J. 1979, Nature, 279, 381