The Optical Light Curve of 0957+561A,B: Time Delay, Hubble Constant, Microlensing

J. Wambsganss, T. Kundic, E. L. Turner,
W. N. Colley, J. R. Gott, III, J. E. Rhoads,
Y. Wang, L. E. Bergeron, K. A. Gloria,
D. C. Long, S. Malhotra, R. Schmidt

Abstract

Continued photometric monitoring of the gravitational lens system 0957+561A,B in the g and r bands with the Apache Point Observatory 3.5 m telescope during 1996 shows a sharp g band event in the trailing (B) image light curve at the time predicted in an earlier paper for the so called "short value" of the differential time delay. This success confirms the "short delay", and the absence of any such feature at a delay near 540 days rejects the "long delay" for this system, thus resolving a long standing controversy. A series of statistical analyses of our light curve data yield a quite robust best fit delay of 417 ± 3 days (2σ).

Recent improvements in the modeling of the lens system allow us to derive a global value (at z = 0.36) of Hubble's constant H_0 using Refsdal's method, a simple and direct distance determination based on experimentally verified and securely understood physics and geometry. The result is $H_0 = 63 \pm 12$ km/s/Mpc (for $\Omega = 1$) where this 95% confidence interval is dominated by the remaining lens model uncertainties. This determination of the extragalactic distance scale (10% accurate at 1-σ) is now of comparable quality, in terms of both statistical and systematic uncertainties, to those based on more conventional techniques.

We use the difference light curve $m_A(t) - m_B(t + \Delta t)$ to derive limits for the occurence of microlensing. We can exclude halo models for the lensing galaxy in which objects in the mass range $10^{-3} M_\odot$ to $10^{-4} M_\odot$ make up the dark matter. Further monitoring and comparison between the lightcurves will be very useful for more stringent limits on or determination of a dark matter component consisting of compact objects.

Introduction: Refsdal's Method

More than 30 years ago, Refsdal (4) pointed out that the differential light propagation time delay between two or more gravitationally lensed images of a background object establishes an absolute physical distance scale ($c\Delta t$) in the system. Thus, the distance to a high–redshift object is directly measured, yielding a value of Hubble's constant, H_0. We briefly review the main strengths of the lensing method in determination of the extragalactic distance scale:

1. It is a geometrical method based on the well understood and experimentally verified physics of General Relativity in the weak–field limit. By contrast, most conventional astronomical techniques for measuring extragalactic distances rely either on empirical relationships or on our understanding of complex astrophysical processes, or both.

2. It provides a direct, single step measurement of H_0 for each system and thus avoids the propagation of errors along the "distance ladder" which is no more secure than its weakest rung.

3. It measures distances to cosmologically distant objects, thus precluding the possibility of confusing a local with a global expansion rate. Note that both observed CMB fluctuations and COBE normalized numerical simulations of large–scale structure formation suggest the possibility of 10-20% *rms* expansion rate fluctuations even on scales of order 10,000 km/s for some cosmological models (Turner, Cen & Ostriker 1992).

4. Independent measurement of H_0 in two or more lens systems with different source and lens redshifts allows a powerful internal consistency check on the answer. Although an inaccurate model of the lens mass distribution or other systematic problems could yield an incorrect distance in any particular system, no known or imagined problem will consistently give *the same* wrong answer when applied to different lenses. Thus, if a small number of time delay measurements all give the same H_0, this value can be regarded as correct with considerable confidence.

Despite these potent virtues, the practical realization of Refsdal's method for measuring H_0 has proven quite challenging and has been long delayed. For the lens system 0957+561A,B discovered by Walsh et al. (7), by far the best studied case, there are two basic reasons. First, there has been sufficient ambiguity in detailed models of the mass distribution in the lensing galaxy and associated cluster to allow values of H_0 different by a factor of two or more to be consistent with the same measured time delay; fortunately, this problem has been much alleviated by recent theoretical and observational work. Second, despite extensive optical and radio monitoring programs extending over a period of more than 15 years, values of the differential time delay discrepant by more than 30% have continued to be debated in the literature. In particular, most studies have given a delay either in the ranges 400–420 or 530–540 days. These two rough values, the "short delay" and the "long delay" have been obtained *both* by applying the same statistical techniques to different data sets *and* by applying different statistical techniques to the same data. Moreover, even application of a single technique to a single (radio) data set has produced best estimate delays that move far outside the nominal formal high confidence interval as additional points in the light curve accumulate. The history of the 0957+561A,B time delay, which can certainly be described as confusing and controversial, is reviewed by (5).

We report a robust determination of the time delay which we believe effectively resolves the controversy in favor of the short delay. In addition, we use this delay and the results of recent theoretical and observational studies of the lens mass distribution to calculate a global measure of Hubble's constant of accuracy comparable to that of the best conventional techniques, both in terms of statistical and systematic errors.

Our time delay determination differs from all previous ones in that the appearance of a sharp, large amplitude feature in the g band light curve of the trailing image B during 1996 was *predicted in advance* based on observations of the leading image A. With the 1996 data we show that the short time delay prediction was quantitatively correct while the long time delay prediction is excluded. Recent observations by other groups (11) support our result.

Data and Confirmed Prediction

The photometric monitoring program for Q0957+561A, B at the Apache Point Observatory (APO[1]) 3.5 meter telescope started in late 1994 and lasted to mid–1995 (1995 season) and was continued in the late 1995 to mid–1996 season (1996 season). The 1996 light curves consist of 51 g band and 54 r band brightness measurements made from 25 November 1995 to 6 July 1996, as compared to the 46 g and 46 r band data points obtained from 2 December 1994 to 27 May 1995 in the 1995 season. In 1996, the mean photometric error of a nightly measurement was 10 millimag in g band and 9 millimag in r band, compared to respective values of 12 and 12 in the previous season. The light curves for the two seasons are thus of quite similar quality. Figure 1 displays our g and r data for images A and B in the 1995 and 1996 seasons, in other words our complete photometric data set. This data is available in the `elt/:0957` subdirectory of the anonymous ftp area at `astro.princeton.edu`.

The predictions of the short time delay set forth in (2) can be confirmed by inspection of Fig. 1 without need for the detailed statistical analysis presented in the next section. The sharp event which (2) noted in the g band light curve of image A in early 1995 is marked in Fig. 1, as are the times of its predicted appearance in the 1996 image B light curve. An event of just the expected amplitude and duration is seen in image B at the time corresponding to the short delay, and no such feature is seen at a time corresponding to the long delay. Although we will proceed to a quantitative analysis below, the qualitative success of the short delay prediction is compelling.

Note that Fig. 1 also displays the rather quiescent behavior of image A in 1996, which corresponds to the predicted 1997 behavior of image B, modulo the now determined delay and magnitude offset. It also emphasizes the good fortune of the occurrence of the sharp g band event in image A shortly after the beginning of our monitoring program; no comparable feature has been seen since!

[1]APO is privately owned and operated by the Astrophysical Research Consortium (ARC), consisting of the University of Chicago, Institute for Advanced Study, Johns Hopkins University, New Mexico State University, Princeton University, University of Washington, and Washington State University.

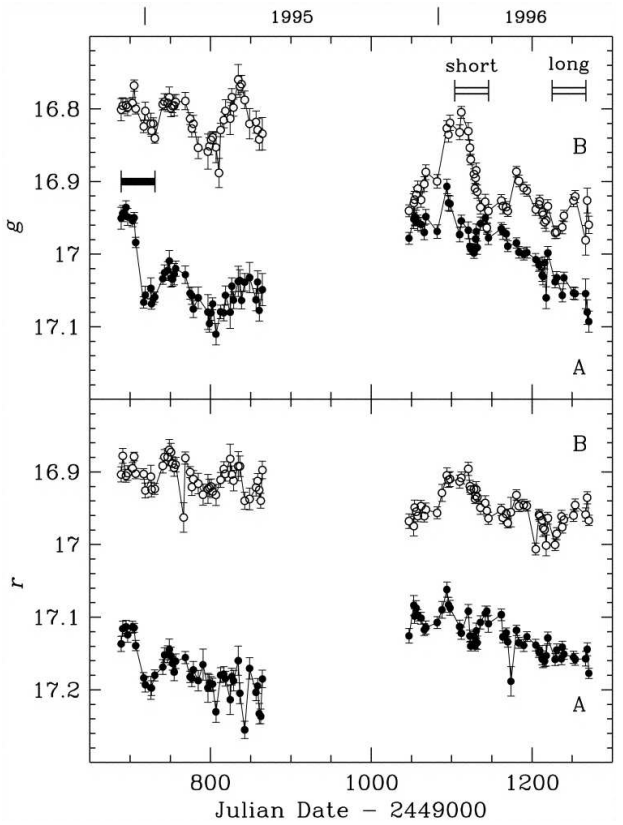

Figure 1: The 1995 and 1996 g- and r-band light curves (with 1 σ error bars for image A (*filled circles*) and B (*open circles*) of the gravitationally lensed quasar 0957+561. The sharp event observed at the beginning of the 1995 season (1994 December - 1995 January) in the g light curve of image A is marked with a horizontal bar. Two predictions for the 1996 light curve of image B were published in (2). They are also marked here: 1996 February and 1996 June. The data support the shorter time delay [from (1)].

Also note that photometric events in the r band are consistently smaller in amplitude than those in the g band. Since most previous monitoring programs have concentrated their attention in red bands (probably to minimize problems with moonlight), this may explain some of the difficulty in obtaining a definitive time delay measurement.

The Best Fit Delay and Its Uncertainty

We used a whole series of statistical comparisons of our 1995 image A light curve with the 1996 image B light curve. The goals of these analyses are 1) to determine the best estimate differential time delay for 0957+561A,B; 2) to quantitatively assess its error (95% confidence interval); and 3) to demonstrate that the data provide a *robust* delay measurement. This last point deserves brief elaboration: Previous attempts to measure the 0957+561A,B time delay from radio and optical data have often depended as much on the statistical techniques employed as upon the data, *i.e.*, different methods gave different answers when applied to the same data Moreover, even the best delay as determined by a single method was unstable against the inclusion or exclusion of a few data points in some cases (PHKRS). The data shown in Fig. 1 produce no such difficulties or ambiguities.

We employ four different statistical methods on four subsets of the complete data set and four parameterized models of how the image A and B light curves are related, practically evaluating 64 possible combinations. Table 1 in (1) and Figure 2 here summarize the results of our statistical comparisons of the 1995 image A light curve with the

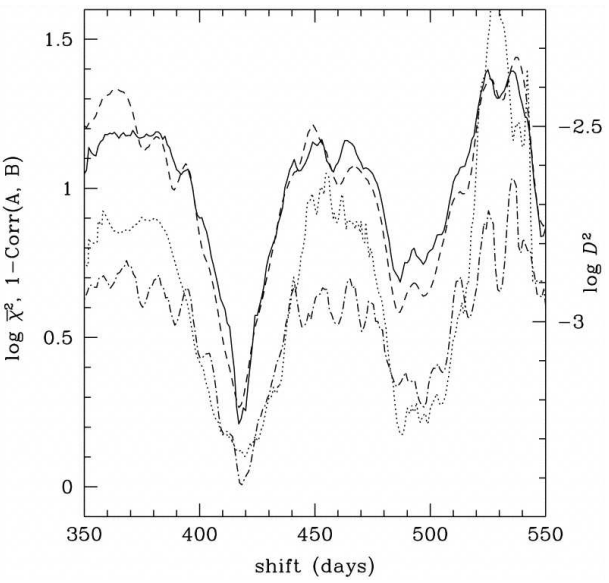

Figure 2: Figure of merit for various statistical methods as a function of time delay Δt based on g-band light curves. The minima correspond to the best fit delays. All methods give minima at $\approx$ 417 - 420 days and strongly reject a delay of $\approx$ 540 days [from (1)].

Of the at least 11 different statistical analyses previously applied to 0957+561A,B optical and radio light curves by various authors (tabulated by (5)), we have selected four representative ones intended to span the range of reasonable approaches for analysis of our data.

1. Linear method: First and simplest, each light

curve and its errors are linearly interpolated, and the data points from the other light curve are shifted so as to minimize $\bar{\chi}^2$. The number of degrees of freedom is equal to the number of overlapping data points in the shifted light curves minus the number of fitted parameters. Time delays with a small overlap of data are thus penalized with respect to delays where the overlap is significant.

2. PRH method: As emphasized by (3), interpolation of the light curves is a crucial ingredient in any method, and we adopt their rather elegant, though somewhat assumption laden, interpolation scheme plus χ^2 minimization as our second method. The PRH method relies heavily on a model of the structure function of the source's intrinsic variability.

3. PHKRS method: As an alternative to χ^2 minimization and as a more assumption–free approach, we adopt a non–parametric technique suggested by PHKRS as a third method.

4. Cross–correlation method: Finally, we have employed a conventional cross-correlation analysis as our fourth method; this technique also requires interpolation but seems to be rather insensitive to its details. Here we report cross–correlation results using only the simplest linear interpolation scheme.

The statistical methods described above have been applied to four subsets of our light curve data. First, we have considered only the g band data; we believe that this gives the best and most reliable delay due to the high signal–to–noise provided by the sharp event. Second, we have treated the r band data separately; it provides a partially independent, though lower signal–to–noise, delay determination. Third, we have combined the g and r data to find overall best fit parameter values. Fourth, we have used only the JD 2449689–2449731 interval of the 1995 image A light curve, *i.e.*, the sharp g band event alone, to search the 1996 image B light curve for a best fit. This last subsampling of the data corresponds most closely to testing the prediction in (2) and is not an *a posteriori* (an hence invalid) editing of the data only because of that prediction.

The simplest model of the mapping of the A image light curve into that of the B image is just a fixed offset in time, Δt, and a fixed magnitude offset, Δm. These two parameters correspond to the macrolensing differential delays and image magnifications. We also consider models in which the magnitude offset is allowed to be a linear function of time at a rate μ, motivated by the possibility that one or both images are being affected by microlensing events with characteristic time scales long compared to the delay and the extent of the observed light curves Finally, we know that the image B photometry is contaminated by the light of the lensing galaxy G1, especially in the r band due to the galaxy's red color relative to the quasar source. Thus, we consider a parameter $m_{\rm gal}$ which represents a constant flux added to that of the source in image B. We have experimented with models in which up to three parameters are allowed to vary as well as ones in which the less interesting ones are neglected.

Table 1 in (1) shows the results of a selected set of combinations of our statistical methods, light curve mapping models and versions of the data. All 64 possible combinations are not shown for the sake of clarity and because some of the statistical methods and some versions of the data do not lend themselves well to the determination of some parameters. However, we emphasize that we are aware of no combination of method, data subset and model which gives a Δt value significantly different from those shown. Figure 2 displays the variation of the minimized statistical parameter with respect to Δt for the g band light curves with each of our four methods. In this Figure microlensing and galaxy terms are not included in the fits.

The most important conclusion to be drawn from Table 1 in (1) and Fig. 2 here is that the best fit value of Δt is *extremely robust*, changing by less than 1% whatever method, data version or parameterized model is used. The delay error estimates produced by the various statistical analyses are somewhat more varied, with those methods that assume greater knowledge of the statistical character of the light curves naturally producing smaller error estimates, but in no case is the estimated error greater than 3%. As we shall see in the next section, the error in the measured delay now contributes negligibly to the error in the deduced global value of H_0.

Because it is simple and seems to work as well

as any other approach for this data set, we adopt the results of the linear method applied to the g band light curves with just a $\Delta t, \Delta m$ model as our final numbers. This gives a delay of 417 ± 3 days and a g band magnitude offset of -118^{+6}_{-8} millimag (B is brighter than A). Figure 3 displays this fit and the corresponding independent fit (420 days, -215 millimag) for the r band data; it also shows the cleanly rejected fit which corresponds to the long delay of about 540 days in the figure insets.

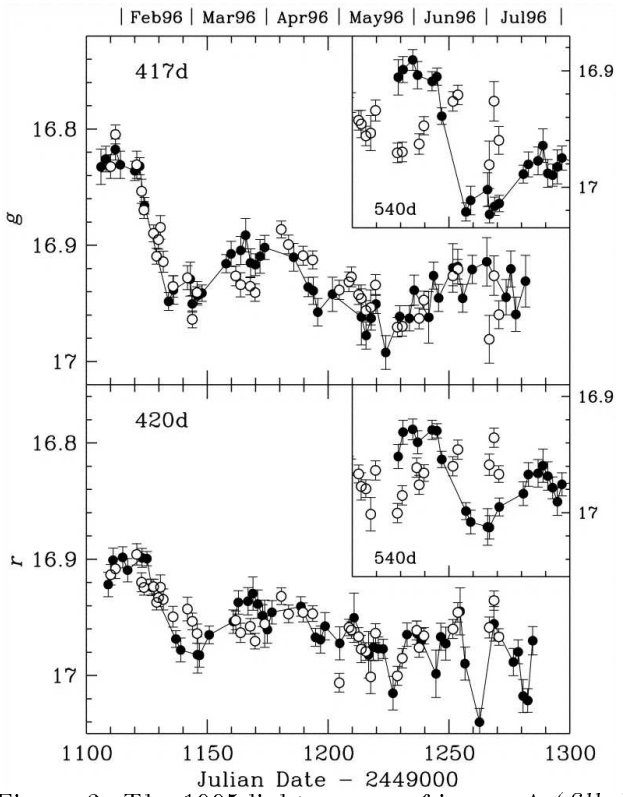

Figure 3: The 1995 light curves of image A (*filled circles*) shifted by the optimal values of the time delay Δt and the magnitude offset Δm are superposed on the 1996 image B data (*open circles*). Insets show the overlapping regions of the A and B light curves assuming the long delay of 540 days [from (1)].

In addition to providing best fits and error estimates, the PRH method produces an optimal (under the assumed statistical properties of the source light curve) reconstruction of the underlying single light curve sampled in the two images, and a "snake" corresponding to its uncertainty at each point. This reconstruction and its comparison to the data are shown in Figure 4.

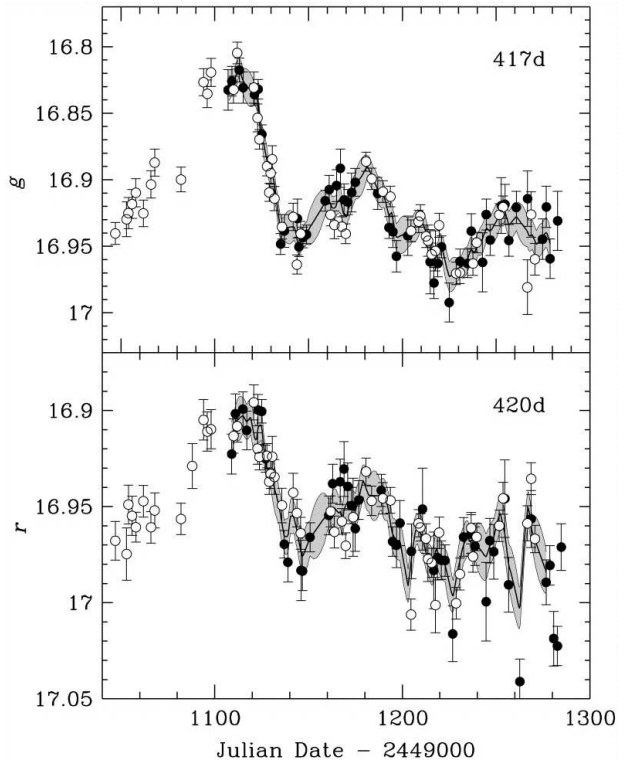

Figure 4: Optimal reconstruction of the shifted and combined 0957+561 image A and image B lightcurves with the technique described in (3). The shaded region ("snake") corresponds to the 1σ condfidence interval of the reconstruction (error bars are 1σ) [from (1)].

Measuring the Global Value of H_0

For a complex gravitational lens system such as 0957+561, reliable measurement of the differential time delay may still leave one far from the goal of a good measurement of H_0 (*e.g.*, Schechter et al.1996), simply because of uncertainties in the lens mass distribution and its implied conversion factor between $c\Delta t$ and the source angular diameter distance (10) This was once a major obstacle, but fortunately an extensive recent theoretical modeling study of 0957+561 by Grogin & Narayan (6, hereafter GN) appears to have resolved the major ambiguities with the exception of one effect noted by (9) In particular, GN use 15 known observables to constrain a set of lens mass models having 9 free parameters. The relation between time delay and distance is controlled primarily by the total projected mass of the lens within the circle defined by the diameter connecting the two source images $M(< r)$ and by its derivative

dM/dr. GN find that the observed image separation $\Delta\theta = 6.1 arcsec$ and the observed mapping of the complex 5-component structure in the VLBI images of the image A and B nearly fix $M(< r)$ and dM/dr at the image positions, respectively, independent of other features of the lens model (8).

Moreover, changing Ω_0 from 1.0 to 0.1 increases H_0 by only 7%. Introducing the cosmological constant while keeping the universe flat results in an increase of just 4% in the $\Omega_0 = 0.25$, $\Omega_\Lambda = 0.75$ model.

There is one remaining degeneracy which cannot be removed by the observed properties of the lensing event itself (5), namely how much of the lensing is contributed by mass associated with the galaxy G1 versus how much is supplied by the associated and superimposed galaxy cluster. Since these two must sum to a known value (in order to produce the observed splitting), the degeneracy may be parameterized by either the one or the other. GN adopt $\sigma_{\rm obs}$, the central line–of–sight velocity dispersion of G1, to measure its contribution and κ, the dimensionless lensing convergence contributed by the cluster, to parameterize its effect. In order to derive an H_0 value, we need to independently measure one or the other parameter. Measuring both provides an internal consistency check.

One can still make an indirect estimate of κ from the total mass of the cluster, its core radius and the location of G1 with respect to the cluster center.

We adopt $\sigma_{\rm cl} = 720 \pm 250$ km/s as our 95% confidence region for the cluster velocity dispersion; using $r = 22 arcsec$ for the distance of G1 from the center of the cluster and $r_c = 5 arcsec$ for the cluster core radius our estimate for the cluster surface density at the position of the lensing event is $\kappa = 0.22 \pm 0.14$ (2σ), more details in (1). This result is insensitive to the precise value of r_c, but it could increase considerably if G1 were much closer to the center of the cluster.

Thus, including the quoted errors in the GN model and those of our time delay and then adding the dominant contribution from the κ uncertainty, we find

$$H_0 = 64^{+4.5}_{-5.3} \left(\frac{1-\kappa}{0.78}\right) \ {\rm km\,s^{-1}\,Mpc^{-1}}$$
$$= 64^{+12}_{-13} \ {\rm km\,s^{-1}\,Mpc^{-1}}$$

where the errors reflect the total 95% confidence interval. The primary remaining systematic uncertainty in this calculation is associated with the mean redshift of the background galaxy population used to map the cluster mass distribution. If it were allowed to vary between 0.75 and 3, the derived value of H_0 would range between 58 and 68 km/s/Mpc, respectively. Considering very extreme possibilities, H_0 would vary from 37 to 70 km/s/Mpc for mean background galaxy redshifts between 0.5 and infinity.

Recently Falco et al. (1997, private communication) obtained a new measurement of the velocity dispersion in G1 from a high signal–to–noise Keck LRIS spectrum. Although there are some puzzling features of their data, a surprisingly rapid variation in $\sigma_{\rm obs}$ along the slit in particular, they obtain a value roughly in the range $\sigma_{\rm obs} = 275 \pm 30$ km/s (2σ). This is much more accurate than—but consistent with—the only previously reported value of 303 ± 50 km/s (Rhee 1991), which was based on a significantly lower signal–to–noise spectrum. Inserting this new value of $\sigma_{\rm obs}$ and again propagating all the relevant errors, we find

$$H_0 = 64^{+7.5}_{-9.8} \left(\frac{\sigma_{\rm obs}}{275 \ {\rm km\,s^{-1}}}\right)^2 \ {\rm km\,s^{-1}\,Mpc^{-1}}$$

$$= 64^{+13}_{-14} \ {\rm km\,s^{-1}\,Mpc^{-1}} \quad .$$

where we again quote 2σ error intervals. The GN model error, included in our H_0 uncertainty, allows for finite aperture effects and possible anisotropy of stellar orbits.

It is important and reassuring that these two entirely independent methods of resolving the galaxy–cluster degeneracy in 0957+561 give very similar and entirely consistent results. For example, if one assumes an extremely low value of 0.5 for the mean redshift of background galaxies in the κ method, this would yield a very small value of H_0 (see above), but would also mildly contradict the measurement mentioned above by implying a G1 velocity dispersion of only 210 km/s.

We cannot properly average the two values since they are based on the same lens mass model (*i.e.*, are not entirely independent), but they can be approximately summarized by $H_0 = 64 \pm 13$ km/s/Mpc at 95% confidence, or $\pm$ 10% at 1σ.

Limits on "Machos" in the lens galaxy

The difference lightcurve between image A and the time shifted image B as shown in Figure 5 shows very little variations. The points with their errorbars are consistent with identical light curves. In particular, there is no gradient visible, which could be the indication of a long-term microlensing event (but see (13)). Conservatively, one can state that here there is no observable difference between the two light curves larger than $\Delta m = 0.05$ mag.

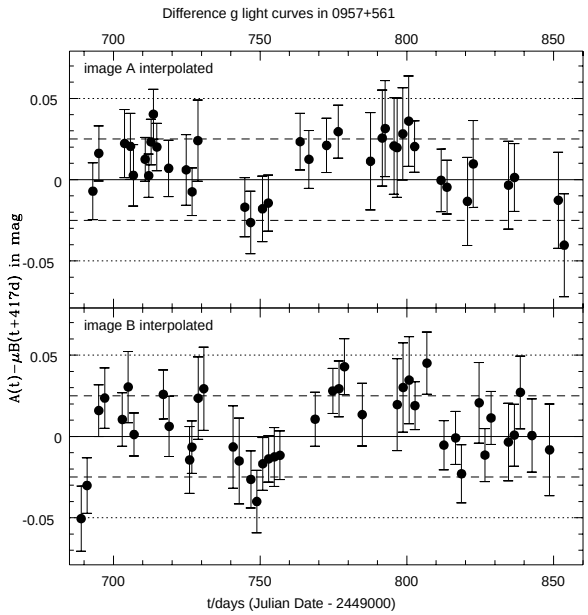

Figure 5: Difference g-band light curves between images A and B (time-shifted and scaled). Top panel: image A light curve linearly interpolated for the epochs of image B observations; bottom panel: vice versa.

We performed ray shooting simulations for the values of surface mass density κ and external shear γ corresponding to the images A ($\kappa = 0.32$, $\gamma = 0.18$) and B ($\kappa = 1.17$, $\gamma = 0.83$). We assumed that the matter consists of compact objects of the same mass, and varied this mass in factors of ten from $m = 10^{-1} to 10^{-4} M_{\odot}$. Evaluated of order 100 000 light curves with a length corresponding to the observational interval of about 6 months, we determined the difference between the maximum and the minimum Δm in these microlensed lightcurves. In Figure 6 the fraction of all lightcurves is displayed which show variations *more* than $\Delta m = 0.05$mag. The four groups of histograms correspond to the four test masses indicated below. The three histograms in each group indicate the lightcurves for image A alone (leftmost), image B alone (middle), and the combined probability that in none of the two lightcurves microlensing can produce a large change.

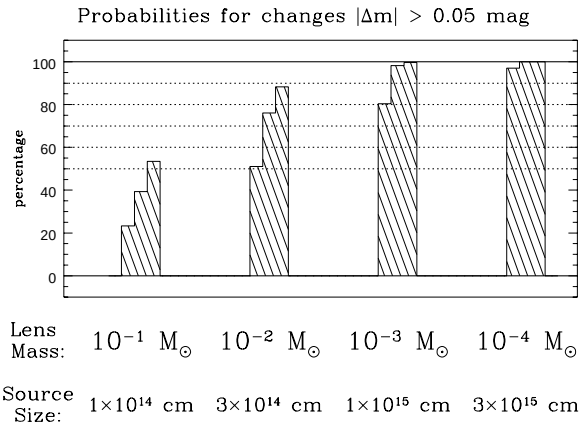

Figure 6: Expected percentages of lightcurves which show deviations of more than $\Delta m = 0.05 mag$ due to microlensing for different masses of compact objects in the halo of the lensing galaxy.

For the high masses ($m = 10^{-1} M_{\odot}$) no conclusive statements can be drawn: in about 50% of the cases one expects variations of $\Delta m \geq 0.05$mag. For the smaller masses of $m = 10^{-2} M_{\odot}$ the probability of no change is only 10%, one more observing period indicating no difference could rule out this mass range. The two lower masses are already practically ruled out as making the dark matter halo in the lensing galaxy of 0957+561.

Summary and Conclusions

Our main results may be summarized as follows: The sharp photometric event predicted in the 0957+561B light curve for 1996 in (2) based on the observed 1995 image A light curve, was observed at the time corresponding to the short (about 420 day) delay, thus confirming its validity. No such event was observed at a time corresponding to the long (540 day) delay, thus rejecting it. The time delay determined by the data presented here and in (1,2) is quite robust and has a best fit value of 417 ± 3 days. Combining this value with the latest theoretical models of the lens mass distribution plus two independent measurements of the relative contributions of the galaxy G1 and the associated galaxy cluster indicates that the global value of H_0 lies in the interval 51 to 77 km/s/Mpc with a most probable value in the range 58–70 km/s/Mpc. Although the possibility of significant systematic error remains, especially in the lens mass distribution model, *this result is likely as accurate and reliable as the best conventional measures of* H_0.

From the fact that the two light curves of images A and B are consistent with being identical (modulo time delay and different magnifications), we conclude that no significant microlensing was going on for the observing periods. By comparison with numerical simulations, we thus exclude massive compact objects of masses $m = 10^{-3} - 10^{-4} M_\odot$ making up the dark halo or the intra-cluster matter in the lensing galaxy.

Turning to a more general perspective, 0957+561A,B appears to have provided a first and so far best success for Refsdal's method of measuring H_0, but it is clear that there are good prospects of both improving its accuracy and extending its applications to other gravitational lens systems. Further theoretical exploration of the lens mass model space is desirable to verify the uniqueness of the GN model. Rapid progress can also be expected using other lens systems. Recently a convincing, though only moderately accurate, pair of time delays in the 1115+080A,B,C system was reported (14). The accuracy can almost certainly be improved. It remains to be seen whether or not sufficiently unique and well constrained models of this system can be obtained to allow useful H_0 measurements.

No known or imagined effect or systematic error would cause different lens systems to give *the same* wrong answer consistently. Thus, if good time delay measurements and well constrained lens models yield the same value of H_0 for a few different lens systems (with different angular sizes, lens and source redshift), the problem of measuring the global value of Hubble's constant will have been effectively solved.

In addition, this lens system may also help to solve one of the other big puzzles, namely what is the dark matter in the halos of galaxies. As suggested by (12), compact objects in halos of galaxies that produce multiple images of quasars must affect the individual light curves by their microlensing effects. The occurence or absence of any such differential changes in the light curves of many multiple quasars is a tool as powerful as the very massive searches for Machos in our galaxy.

The rate of discovery of new lens systems is likely to increase, so we may expect yet more suitable "Golden Lens"-systems to be found in the future.

References

(1) Kundic, T. et al. ApJ 482, 75 (1997)

(2) Kundic, T. et al. ApJ 455, L5 (1995)

(3) Press, W. H. et al. ApJ 385, 404 (1992) and Press, W. H. et al. ApJ 385, 416 (1992)

(4) Refsdal, S. MNRAS 128, 307 (1964) and Refsdal, S. MNRAS 132, 101 (1966)

(5) Haarsma, D. B. et al. ApJ 479, 102 (1997)

(6) Grogin, N. A., & Narayan, R.: ApJ 464, 92 & 473, 570

(7) Walsh, D. et al. Nature 279, 381 (1979)

(8) Garrett, M. A. et al. MNRAS 270, 457 (1994)

(9) Falco, E. E. et al. ApJ 289, L1 (1985) and Gorenstein M.V. et al. (1988).

(10) Falco, E. E. et al. ApJ 372, 364 (1991) and Kochanek, C. S. ApJ 382, 58 (1991)

(11) Oscoz, A. et al. ApJ 470, L19 (1996) and Oscoz, A. et al. ApJ 479, L89 (1997)

(12) Gott, J. R. ApJ 243, 140 (1981)

(13) Schild, R. ApJ 464, 125 (1996)

(14) Schechter et al. ApJ 475, L85 (1997)