NEW MODELING OF THE LENSING GALAXY AND CLUSTER OF Q0957+561: IMPLICATIONS FOR THE GLOBAL VALUE OF THE HUBBLE CONSTANT

KYU-HYUN CHAE

Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260

Received 1999 February 5; accepted 1999 June 9

ABSTRACT

The gravitational lens 0957+561 is modeled utilizing recent observations of the galaxy and the cluster as well as previous VLBI radio data that have been reanalyzed recently. The galaxy is modeled by a power-law elliptical mass density with a small core, while the cluster is modeled by a nonsingular power-law sphere, as indicated by recent observations. Using all of the currently available data, the best-fit model has \( \chi^2_{\text{min}} / N_{\text{dof}} \approx 6 \), where the \( \chi^2 \) value is dominated by a small portion of the observational constraints used; this value of the reduced \( \chi^2 \) is similar to that of the recent FGSE (Falco, Gorenstein, & Shapiro elliptical) best-fit model of Barkana et al. However, the derived value of the Hubble constant is significantly different from the value derived from the FGSE model. We find that the value of the Hubble constant is given by \( H_0 = 69^{+13}_{-12}(1 - \kappa) \) and \( 74^{+15}_{-18}(1 - \kappa) \) km s\(^{-1}\) Mpc\(^{-1}\) with and without a constraint on the cluster’s mass, respectively, where \( \kappa \) is the convergence of the cluster at the position of the galaxy and the range for each value is defined by \( \Delta \chi^2 = \chi^2_{\text{min}} / N_{\text{dof}} \). At present, the best achievable fit for this system is not as good as for PG 1115+080, which has also recently been used to constrain the Hubble constant, and the degeneracy is large. Possibilities for improving the fit and reducing the degeneracy are discussed.

Subject headings: distance scale — gravitational lensing — quasars: individual (Q0957+561)

1. INTRODUCTION

The gravitationally lensed “double quasar” Q0957+561 is the first lensed system for which the time delay between image components has been measured with considerable confidence and accuracy (Haarasma et al. 1999; Pelt et al. 1998; Kundic et al. 1997; Pipers 1997; Oscoz et al. 1997; Schild & Thomson 1997). In principle, the measured time delay (417 ± 3 days; Kundic et al. 1997) can be turned into an accurate determination of the Hubble constant, \( H_0 \), using the method outlined byRefsdal (1964, 1966; see Schneider, Ehlers, & Falco 1992 for a pedagogical introduction). For this method, the potential of the lens (i.e., its mass distribution) must be well determined to be consistent with observational constraints. Many important observational constraints on Q0957+561 have come from VLBI observations of the radio jets and cores of the two image components (Garrett et al. 1994; Gorenstein et al. 1988, 1984). The Garrett et al. (1994) data have been reanalyzed just recently, and improved constraints are given by Barkana et al. (1999, hereafter Ba99).

The lensing of Q0957+561 is due to the combined effect of a massive elliptical galaxy and a cluster at redshift \( z = 0.36 \), of which the elliptical galaxy is the brightest member. The cluster has been studied using optical observations (Angonin-Willame, Soucail, & Vanderriest 1994; Garrett, Walsh, & Carswell 1992), X-ray observations (Chartas et al. 1998, 1995), and weak-lensing effects (Fischer et al. 1997; Dahle, Maddox, & Lilje 1994). These studies indicate that the lensing galaxy is positioned close to the center of the cluster, and the cluster’s convergence in the region of the image is significant. However, the cluster’s mass distribution could not be accurately determined from those studies. The lensing galaxy has been detected both optically (Bernstein et al. 1997, hereafter Be97; Stockton 1980) and in the radio (Roberts et al. 1985; Gorenstein et al. 1983). The more recent Hubble Space Telescope (HST) observations by Be97 appear to have significantly reduced the uncertainty in the position of the galaxy. The HST optical position (G1) is consistent with the VLBI position (G), but is inconsistent with the VLA position (G). The Be97 observations also reveal new possible lensed features of two blobs and two knots, along with an arc, in the system.

Present observational constraints (the VLBI and HST constraints) now clearly exclude both the Grogin & Narayan (1996) softened power-law sphere (SPLS) model and the Falco, Gorenstein, & Shapiro (1991) model (FGS) of a King-profile sphere with a central black hole, which were previously the best-fit models. Ba99 extended the SPLS and FGS models to include an arbitrary ellipticity to the galaxy, and the new versions of models (softened power-law elliptical mass distribution [SPEMD] and FGS elliptical [FGSE]) fit the observations better. According to their study, the FGSE model gives a somewhat better fit than the SPEMD model \( [\chi^2_{\text{min}} / N_{\text{dof}} = 6 \text{ versus } 10] \). Notably, the value of \( H_0 \) derived from the FGSE model is \( \geq 100 \) km s\(^{-1}\) Mpc\(^{-1}\). In previous models of Q0957+561, the lensing cluster has been modeled by a quadrupole term (e.g., Ba99; Grogin & Narayan 1996; Falco et al. 1991) because of its simplicity, and assuming that higher-order terms can be neglected in the expansion of the cluster potential. In addition, Ba99 considered a singular isothermal sphere (SIS) to model the cluster for the FGSE model of the galaxy, while Bernstein & Fischer (1999) considered higher-order terms in the cluster expansion. Ba99 found that for the FGSE model of the galaxy, an SIS model of the cluster did not lead to either an improved \( \chi^2_{\text{min}} \) or a significantly different value of \( H_0 \) compared to the case using a quadrupole term. Thus, for the FGSE model of the galaxy, it appears that consideration of a realistic cluster mass model is not crucial. However, the quadrupole term alone may not accurately account for the cluster contribution, especially since recent observations indicate that the cluster’s mass center can be very close to the galaxy (Fischer et al. 1997; Angonin-
willaime et al. 1994), and there is no a priori reason that we
need only consider the sis model for the cluster. Bernstein
& Fischer's (1999) consideration of higher-order terms can
be a better approximation of the cluster potential than the
single quadrupole term; however, the low \( \chi^2 \) values of their
best-fit models depended, in large part, on their selective use
of the current observational constraints (see § 5).

In this study we consider a more general model of the
cluster, i.e., a power-law sphere, while modeling the galaxy
with a power-law elliptical mass density with a small core
(i.e., the SPEMD model with a fixed small core; see §§ 2 and
3). We also consider the case in which the cluster is elliptical.
The reduced \( \chi^2 \) of the best-fit model of this type is similar
to that of the best-fit FGSE model. However, the derived
value of \( H_0 \) is significantly different from the value
derived from the FGSE model. It turns out that neither the
core radius nor the power-law radial index of the cluster can
be determined from fitting the model to the current obser-
vation constraints. This is because varying either the radial
index or the core radius (within a wide range) does not
significantly alter the fit, while the convergence of the cluster
(\( \kappa \)) at the position of the galaxy (and thus the derived
value of \( H_0 \)) varies. Thus, we can only derive a scaled
by a factor of \( (1 - \kappa) \) value of \( H_0 \). The outline of this paper is
as follows. We first review the adopted observational con-
straints (§ 2). In § 3, we present the lens model and review its
parameters. In § 4, we present the results of fitting the model
to the observational constraints, and derive a value for \( H_0 \).
In § 5, we discuss possible future improvements to the deter-
mination of \( H_0 \) using the Q0957 + 561 lens.

2. SUMMARY OF CURRENT OBSERVATIONAL
CONSTRAINTS

The optical image of Q0957 + 561 consists of two point-
like quasars, A and B, \( \approx 6' \) apart on the sky. This is only a
weak constraint on a lens model. However, the radio image
of Q0957 + 561 provides many strong constraints. In partic-
ular, the extended radio jets along with the radio cores can
be used to derive both accurate relative positions of the jets
with respect to their cores and between the cores, and the
relative magnification matrix between the corresponding
components of A and B. The separation between the cores
was measured with an uncertainty of 0·00004 using VLBI
observations (Gorenstein et al. 1984; Falco et al. 1991).
Further VLBI observations at \( \lambda = 13 \) cm by Gorenstein et
al. (1988) identified three jets in each quasar, which were
used to derive a relative magnification matrix \( [M_{\theta x}] =
[\partial x_x/\partial x_x] \) between quasars A and B. More recently, VLBI
observations at \( \lambda = 18 \) cm by Garrett et al. (1994) separated
the jets into five components for each quasar, from which
they derived an improved relative magnification matrix
along its gradient. However, Ba99 employed improved
analysis procedures to reanalyze the raw data of Garrett et
al. (1994); in particular, they corrected some significant
errors in the VLBI data reduction packages adopted by
Garrett et al. (1994) and fitted the jet component param-
eters (i.e., positions, fluxes, and Gaussian model parameters)
and the magnification transformation parameters simulta-
nously in a single step, and so we will use the constraints
given by Ba99 in this study.

Radio observations of the system have resulted in the
detection of radio sources near the optical center of the
lensing galaxy (Roberts et al. 1985; Gorenstein et al. 1983).
While both the VLA position G (Roberts et al. 1985) and
the VLBI position G' (Gorenstein et al. 1983) were reported
to have an uncertainty of 1 mas, they differed by 30 \( \sigma \) from
each other. The optical position (G1) of the lensing galaxy
as measured by Stockton (1980) is more consistent with the
VLBI position; however, his measurement uncertainty is
large (30 mas). Thus, in previous studies that modeled the
lens, it was not clear which observed position to use as a
constraint. However, HST observations of the system by
Be97 were used to determine the optical position of the
lensing galaxy with an uncertainty of 3.5 mas. The HST
position is consistent with the VLBI position within 10 mas,
but it is clearly inconsistent with the VLA position.
Whether the HST position or the VLBI position are used to
constrain the lens model (§ 4) makes little difference;
however, we will use the VLBI position. Another constraint
to be used is the core flux ratio, B/A, derived from VLBI/
VLA observations and observations of the optical emission
lines (Conner, Lehar, & Burke 1992; Schild & Smith 1991).
The above constraints, derived mostly from VLBI obser-
vation, are used as our "primary constraints" (by this we
mean that these constraints have been derived from con-
firmed lensed features, and that they are used in the first
step of the fitting procedure; see § 3) on the lens model; they
are summarized in Table 1.

The Be97 HST observations detected possible lensed
images of other background sources. Blobs 2 and 3 are
probably images of a common background galaxy. Knots 1
and 2, which form an arc, are also probably images of a
common source. While a lens model based on the primary
constraints easily satisfies the constraints from the knots,
the constraints from the blobs turn out to be useful for
distinguishing between models. The constraints from the
blobs and knots are used as "secondary constraints," which
are summarized in Table 2. The total number of primary
and secondary constraints is 25. All of these constraints are
used to define the goodness of fit, \( \chi^2 \), taking into account
correlation coefficients for some constraints (Tables 3 and
4).

| TABLE 1 |
| PRIMARY LENSING CONSTRAINTS |
|-----------------------------|
| Observational Constraint   | Value (uncertainty) | Reference |
| \( \Delta \sigma (A_1 - A_2) \) (mas) | 16.6 (0.1) | 1 |
| \( \Delta \sigma (A_2 - A_3) \) (mas) | 45.6 (0.1) | 1 |
| \( \Delta \sigma (B_1 - B_2) \) (mas) | 18.32 (0.07) | 1 |
| \( \Delta \sigma (B_2 - B_3) \) (mas) | 55.8 (0.2) | 1 |
| \( \Delta \sigma (A_1^b - A_2^b) \) | 1.15 (0.03) | 1 |
| \( \Delta \sigma (A_2^b - A_3^b) \) | -0.56 (0.03) | 1 |
| \( \phi_1 (\deg) \) at \( A_1^b \) | 18.76 (0.04) | 1 |
| \( \phi_2 (\deg) \) at \( A_3^b \) | 107 (7) | 1 |
| \( \Delta \sigma (A_1^b - A_2^b) \) | 1.27 (0.03) | 1 |
| \( \Delta \sigma (A_2^b - A_3^b) \) | -0.58 (0.04) | 1 |
| \( \Delta \sigma (A_1 - B_1^s) \) (mas) | -1.25254 (0.00004) | 1, 2 |
| \( \Delta \sigma (A_1^b - B_1^s) \) (mas) | 6.04662 (0.00004) | 1, 2 |
| Relative magnitude \( (B_1/A_1) \) | 0.747 (0.015) | 3 |
| \( \Delta \sigma (G' - B_1^s) \) (mas) | 0.181 (0.001) | 4 |
| \( \Delta \sigma (G' - B_1^s) \) (mas) | 1.029 (0.001) | 4 |

Note.—All positions are for J1950.0.

\(^a\) See Table 3 for the correlation coefficients.
\(^b\) See Table 4 for the correlation coefficients.
\(^c\) Since the VLBI position of \( G' \) is consistent with the HST position
from Bernstein et al. (1997) at a 10 mas level, we have chosen to use this
position. However, the model fit is unaffected if the HST position is used.

References.—(1) Barkana et al. 1999; (2) Falco et al. 1991; (3) Conner
et al. 1992; (4) Gorenstein et al. 1983.
Another observational constraint is the upper limit on the relative brightness of any third image (image C) predicted by a lens model that has a smooth mass distribution at the central region of the galaxy. The VLBI observations by Gorenstein et al. (1984) set a 5σ upper limit of C/B = 3.3%, while the optical observations by Stockton (1980) found C/B < 2%. The relative brightness of the third image is mostly related to the mass distribution at the central region of the lens, which is effectively described by a core-radius parameter. The observed low values of the upper limit on the relative brightness of a third image implies that the core radius of the lens model is very small (see § 3). If the core radius is sufficiently small, then the lensing properties of the model become insensitive to the value of the core radius, since mass distributions with different core radii (and different central densities) are essentially the same outside the small central regions. Thus, we can fix the core radius at a small value that ensures that the relative magnification of the third image predicted by the model is always less than the observational upper limit. This approach is employed in this study.

3. AN OBSERVATIONALLY MOTIVATED MODEL OF THE LENS

We consider the following forms for the mass distributions of the lensing galaxy,

\[ \Sigma_{\text{gal}}(r, \theta) = \frac{\Sigma_1}{[1 + (r/r_1)^2][1 + e_1 \cos 2(\theta - \theta_1)][(v_1 - 1)/2]^{1/2}}, \]

and cluster,

\[ \Sigma_{\text{cl}}(r') = \frac{\Sigma_2}{[1 + (r'/r_2)^2][v_2 - 1/2]^{1/2}}. \]

Here \( \Sigma_i \) (i = 1, 2) are the central surface mass densities, \( r_i \) are the core radii, and \( v_i \) are the radial indices. Radial indices \( v = 1, 2, \) and 3 correspond to radial profiles of constant surface density, isothermal, and modified Hubble law, respectively, so it is related to the parameter \( \eta \) used by Grogin & Narayan (1996) and Ba99 via \( v = 3 - \eta \). The parameter \( e_1 \) (greater than 0) is related to the galaxy’s ellipticity \( e_1 = (1 - b/a) \), where \( a \) and \( b \) are the major and minor axes, respectively) via \( e_1 = 1 - [(1 - e_1)/(1 + e_1)]^{1/2} \). The parameter \( \theta_1 \) is the position angle (north through east) of the galaxy. Primed coordinates are used for the cluster’s mass distribution, since the cluster’s center is offset from the galaxy’s center. The relative position of the cluster with respect to the galaxy is represented by the distance between them, \( d_{12} \), and the position angle of the cluster’s center as viewed from the galaxy, P.A.12. The above model will be referred to as the “PEM + PS” (power-law elliptical mass + power-law sphere) model.

Throughout, we assume a standard cosmology with a cosmological matter-energy parameter \( \Omega_{\text{matter}} = 1 \) and a cosmological vacuum-energy parameter \( \Omega_{\Lambda} = 0 \), and we use the usual definition \( h = H_0/100 \) km s\(^{-1}\) Mpc\(^{-1}\). For this choice of cosmology, we have \( 1' \approx 3.0 \) h\(^{-1}\) kpc at the redshift of the lens. For an open universe with \( \Omega_{\text{matter}} = 0.3 \) (and \( \Omega_{\Lambda} = 0 \)), the estimate of \( h \) increases by \( \approx 7\% \). As explained in § 2, the present observational constraint on the relative brightness of a third image implies that the core radius of the galaxy, \( r_1 \), should be very small. For example,
the upper limit on the core radius of a galaxy with radial index \( v_1 = 1.72 \) in equation (1) is \( r_1 \approx 10^{-3} \) pc for \( C/B \leq 3.3\% \). We fix the core radius at \( r_1 = 0.1 \) h\(^{-1}\) pc, which is small enough to satisfy the observational constraint for any value of \( v_1 \), but is an arbitrary choice otherwise (see § 2). The other parameters of the galaxy (i.e., central density, ellipticity, and position angle) are free parameters for the model. Thus, our model of the lensing galaxy can be considered as a dark matter–dominated mass model. It will be interesting to compare the determined mass distribution of the lensing galaxy with its observed light distribution (Be97).

Present observational studies of the lensing cluster indicate that the cluster’s mass distribution is more consistent with an extended core (Chartas et al. 1998; Fischer et al. 1997), although the uncertainties in the derived values are too large for them to be directly used to constrain the lens model. Thus, we fix the core radius at an arbitrary large value (e.g., several arcseconds) and derive a value of \( h \) with the fixed core radius. We then study how the derived value of \( h \) is affected if we vary the core radius from the chosen value. There is no observational constraint on the radial index of the cluster. Thus, we fix the radial index at an arbitrary value, e.g., \( v_2 = 2 \) (isothermal profile) and then study how the derived value of \( h \) is affected when \( v_2 \) is varied.

To calculate the lensing properties (i.e., deflection, magnification, and relative light travel time) of the mass distribution implied by equation (1), we use the series method by Chae, Khersonsky, & Turnshek (1998); for details, the reader is referred to the paper and references therein. In fitting the observational constraints of Q0957+561, we must use an accurate and robust method of calculation, since the fractional errors of the VLBI data are as small as \( \sim 10^{-5} \). The calculational accuracy of the series method can be made orders of magnitude smaller than the observational errors by controlling the truncation of the series. This was tested for the isothermal case of \( v = 2 \) using the analytic solution of Kormann, Schneider, & Bartelmann (1994). Since the series converge well for an arbitrary value of \( v \), we can be confident that the series calculation gives sufficiently accurate results. Given that the total number of constraints is 25 and the total number of free parameters (including the eight coordinates of the four sources, i.e., core, jet, blob and knot, and two redshifts of the blob and knot) is 17, the minimization of the \( \chi^2 \) is not a simple numerical task. For this reason, we first fit the model to the primary constraints only, so that we can work with smaller numbers of constraints and free parameters. Once best-fit values based on the primary constraints are determined, the secondary constraints and the remaining parameters are considered, and then the entire set of parameters is adjusted to minimize the \( \chi^2 \).

4. RESULTS OF FITTING THE MODEL TO THE OBSERVATIONAL CONSTRAINTS: BOUNDS ON THE HUBBLE CONSTANT

Because of the degeneracies in the cluster’s core radius (\( r_2 \)) and radial index (\( v_2 \)), we first consider fixed values of \( r_2 \) and \( v_2 \), and then study how varying them affects the derivation of \( h \). We set \( v_2 = 2 \) and \( r_2 = 15 \) h\(^{-1}\) kpc (\( \sim 5' \)). We find that the model has significant degeneracies in parameters \( v_1 \) (the galaxy’s radial index) and \( d_{12} \) (the separation between the cluster and the galaxy). Thus, we fix the parameters \( v_1 \) and \( d_{12} \) for each model and, by incrementing them in steps of 0.02 and 1°, respectively, obtain a grid of models in the two-dimensional parameter space spanned by \( v_1 \) and \( d_{12} \). A model with \( v_1 = 1.72 \) and \( d_{12} = 9'' \) has the lowest value of \( \chi^2 \) of \( \chi^2_{\text{min}} \approx 49.6 \) with 8 degrees of freedom (\( \chi^2_{\text{min}} = 6.2 \)). As the parameters \( v_1 \) and \( d_{12} \) are varied from the best-fit values, \( \chi^2 \) increases moderately. A confidence limit (CL) ellipse defined by \( \Delta \chi^2 = \chi^2_{\text{min}} \) is shown in Figure 1 (solid line). Although this results in a larger region of parameter space than the conventional 1σ (i.e., 68% CL) region, we adopt it as in Ba99 and Grogin & Narayan (1996), since the best-fit model has a poor overall \( \chi^2 \).

For the best-fit model (\( v_1 = 1.72, d_{12} = 9'' \)), we find \( h = 0.54 \), which is obtained from \( h = \Delta \chi^2_{\text{model}} / \Delta \chi^2_{\text{observed}} = (225 \text{ day})/(417 \text{ day}) = 0.54 \), where \( \Delta \chi^2_{\text{model}} \) is calculated for \( h = 100 \) km s\(^{-1}\) Mpc\(^{-1}\). In the range, \( 0.39 < h < 0.70 \). However, this value of \( h \) was obtained for the fixed cluster parameters of \( v_2 = 2 \) and \( r_2 = 15 \) h\(^{-1}\) kpc. Although these choices are the best values obtained by Fischer et al. (1997), their allowed range of \( r_2 \) is large, and there is no a priori reason that the cluster’s mass distribution should have an exact isothermal profile. Thus, we must study how the above derived value of \( h \) is affected as the cluster parameters are varied. The effect of varying the cluster’s radial index turns out to be trivial; for a profile shallower (\( v_2 < 2 \)) or steeper (\( v_2 > 2 \)) than the isothermal,
the fit results are virtually the same except that the cluster's convergence at the position of the galaxy ($\kappa$), i.e., at $r = 0$, is altered and the value of $h$ is scaled by a factor of the ratio of the values of $(1 - \kappa)$. Figure 2 shows the values of $\chi^2$, $h / (1 - \kappa)$, and $\kappa$ for various example cluster parameter sets. We see that for a wide range of $1.5 \leq v_2 \leq 2.5$ around the isothermal value ($v_2 = 2$) at core radius $r_2 = 15 \ h^{-1} \text{ kpc}$ ($\approx 5''$), $\Delta \chi^2 < 0.12 \chi^2_{\text{min}}$ and $h/(1 - \kappa)$ changes by no more than 1.4% while $\Delta \kappa < 0.22$. The effect of varying the cluster's core radius can be more complicated. Nevertheless, we find that for the observational range of $r_2$ from Fischer et al. (1997), the effect is simple (see Fig. 2). Namely, when the core radius is increased or decreased from $r_2 = 5''$ for the isothermal profile, the $\chi^2$ varies little ($\Delta \chi^2 < 0.18 \chi^2_{\text{min}}$), the value of $\kappa$ is altered, and the value of $h$ scales, to a good approximation (up to an error of 2.8%), according to $(1 - \kappa)$ for $0' \leq r_2 \leq 10'$. In summary, for a range of the cluster's radial index and core radius ($1.5 \leq v_2 \leq 2.5, 0' \leq r_2 \leq 10'$) that we investigated with the current observational constraints, we find that there is no significant change of $\chi^2$ ($\Delta \chi^2 < 0.26 \chi^2_{\text{min}}$), and that the change of $h$ is dictated by $(1 - \kappa)$ to an error no larger than $\approx 4\%$. Therefore, the mass-sheet degeneracy (Falco, Gorenstein, & Shapiro 1985) appears to persist even when we use a full potential model of the cluster. (We remind ourselves that when the quadrupole approximation was adopted in previous lens models of Q0957+561, this degeneracy was exact.) However, for each radial profile of the cluster, the $\chi^2$ has a “minimum” value at a certain finite core radius (e.g., at $r_2 \approx 2''$ for the isothermal case), and the $\chi^2$ increases consistently as the core radius is varied, although we concluded above that the variation of the $\chi^2$ was insignificant with the current observational constraints. There is a subtlety here. Even though the total $\chi^2$ remains virtually invariant as the core radius is varied for a fixed radial index, the $\chi^2$ contributions due to the primary constraints and those due to the secondary constraints vary significantly. For instance, for the isothermal profile ($v_2 = 2$), the model with $r_2 \approx 0''$ has $\chi^2_p \approx 35.6$ and $\chi^2_s \approx 14.0$, while the model with $r_2 \approx 10''$ has $\chi^2_p \approx 40.2$ and $\chi^2_s \approx 10.2$, where $\chi^2_p$ and $\chi^2_s$ denote the $\chi^2$ contributions due to the primary and secondary constraints, respectively. As will be seen in § 5, the radio jet positions have the largest $\chi^2$ contribution out of the primary constraints, as does the blob magnification ratio out of the secondary constraints. This suggests that better observational knowledge of the radio jets and the blob magnification ratio could break the present “superficial” degeneracy of the core radius for a given value of the radial index. When the radial index is varied for a fixed value of the core radius, the $\chi^2_p$ and $\chi^2_s$ each change by only $\approx 1$ for $\Delta \nu = 1$. This means that the degeneracy of the cluster's radial index is much more significant with the present observational constraints. Unless new independent observational constraints are revealed, the degeneracy of the cluster's radial index cannot be broken.

As seen in the above, the degeneracy of the cluster's core radius is weaker than that of the cluster's radial index. However, the determination of $h/(1 - \kappa)$ is not significantly affected by varying the core radius. To illustrate this, CL ellipses are drawn in Figure 1 for two other values of $r_2$, i.e., $r_2 \approx 3''$ (dashed line) and $r_2 \approx 7''$ (dotted line). As seen in the figure, the confidence region of the parameter space is altered only slightly whether the core radius is increased or decreased. The $\Delta \chi^2 = \chi^2_{\text{min}}$ range of $h/(1 - \kappa)$ is nearly unchanged for $0' \leq r_2 \leq 10'$. From the grid of models with $v_2 = 2$ and $r_2 = 15h^{-1} \text{ kpc} (\approx 5'')$, we find $h = 0.74^{+0.18}_{-0.17} (1 - \kappa)$.

As we considered a CL by $\Delta \chi^2 = \chi^2_{\text{min}}$ in the above discussion, we could consider a CL by $\Delta \chi^2 = 4\chi^2_{\text{min}}$. However, the confidence region within so defined a CL is too large to be useful to constrain $h$. If the cluster's mass is constrained, the ranges of the parameter space can be reduced. Unfortunately, the cluster's mass is not well determined from present observations. Figure 3 shows confidence regions of the parameter space of the model with $v_2 = 2$ and $r_2 = 15h^{-1} \text{ kpc}$, including the cluster's mass as a constraint determined directly from weak lensing effects by Fischer et al. (1997; $\Sigma_0 = 0.36 \pm 0.11 h \times 10^{10} M_\odot \text{ kpc}^{-2}$ for $v = 2$ and $r_2 \approx 5''$). The parameter space is slightly better constrained compared to the case in which the cluster's mass is not constrained (Fig. 1). We find $\chi^2_{\text{min}} \approx 5.6$ for $v_1 = 1.70$ and $d_{12} = 10''$. The derived value of the Hubble constant is $h = 0.69^{+0.15}_{-0.13}(1 - \kappa)$ with the cluster's mass constrained. The measured velocity dispersions of the lensing galaxy (Tonry & Franx 1999; Falco et al. 1997; Rhe 1991) can in principle be used to constrain the mass of the galaxy, which can be useful for reducing the degeneracy in the parameter space of the PEM + PS model. Stellar-orbit modeling methods have been applied to the best-fit SPLS mass profile of Grogin & Narayan (1996) to infer the mass of the galaxy (Romanowsky & Kochanek 1999). Similar studies in the future for the PEM galaxy model could be useful.

---

2 Here we do not constrain the cluster's position, and thus we do not constrain the cluster's convergence at $r = 0$, although we do constrain its mass.
Fig. 3.—Confidence regions, in the same parameter space as in Fig. 1, of the model with the cluster core radius of 5'. Here the cluster's mass is constrained using the observationally derived value from Fischer et al. (1997). The parameter space is slightly better constrained compared with the case (Fig. 1) in which the cluster's mass is not used as an observational constraint. In fact, if the cluster's mass were determined more accurately by observations, the confidence region of parameter space could be reduced significantly, thereby reducing the allowed range of $H_0$.

Table 5 summarizes the values of the parameters for the model with $v_2 = 2$ and $r_2 = 15 \, h^{-1} \, \text{kpc}$, with the cluster's mass unconstrained. The theoretical images of the radio core and the brightest radio jet for the best-fit model are shown in Figure 4. Figure 5 shows theoretical images of the optical blob and knot along with an arc. The model galaxy's radial profile is somewhat shallower than the isothermal profile $v = 2$, i.e., the mass within radius $r$ increases faster than the isothermal case as $r$ increases. The model galaxy's ellipticity ($v_{\text{model}} = 0.01 - 0.38$) is similar to the ellipticity of the observed light $v_{\text{light}} = 0.05 - 0.49$; position angle of the light distribution is P.A. $= 32\degree - 68\degree$ (Be97). The position angle of the best-fit model (P.A. $= 64\degree$) is in agreement with the light distribution. However, in the $\Delta \chi^2 = \chi^2_{\text{min}}$ range, $6\degree \leq \text{P.A.} \leq 166\degree$. Here it is interesting to note that for the majority of 17 gravitational lens galaxies studied by Keeton, Kochanek, & Falco (1998), the position angles of the light and the mass (i.e., lens model) are the same to $\leq 10\degree$. In other words, not all of the models within the $\Delta \chi^2 = \chi^2_{\text{min}}$ range around the best-fit model are consistent with the general result obtained by Keeton et al.
Recent observations of the lensing galaxy and cluster of Q0957 + 561 (Be97; Fischer et al. 1997) motivated modeling the galaxy by an elliptical mass distribution and modeling the cluster by a power-law sphere with an extended core. The PEM + PS model (§ 3) was fitted to both the VLBI constraints (Ba99) and the HST constraints (Be97) of the system to derive a value for $H_0$. The best-fit model has $\chi^2_{\text{min}} \approx 6$, which is better than the best fits achievable by previous models. Here it should be emphasized that the relative improvement of the fit was achieved by using an elliptical mass distribution for the galaxy and a mass sphere for the cluster, replacing the quadrupole term. The only other model that gives a fit of comparable quality is the FGSE model of Ba99, in which the galaxy is modeled by an elliptical King-profile mass distribution with an extended core along with a large effective black hole (BH) mass at the galaxy's center, while the cluster is modeled using a quadrupole term or a singular isothermal sphere (SIS). The FGSE model suggests that the Hubble constant is $h = 1.23^{+0.23}_{-0.33}$ and $1.13^{+0.32}_{-0.32}$ ($\chi^2 = 4\chi^2_{\text{min}} \text{ CL}$) using the quadrupole term and the SIS for the cluster, respectively. A few comments appear to be relevant regarding the FGSE model. First, the model requires a very large point mass of $\sim 10^{11} M_{\odot}$, which does not seem observationally motivated or supported. Second, the value of $H_0$ derived from the FGSE model is inconsistent with other recent independent determinations of $H_0$, especially the determination from the lens PG 1115 + 080 by Impey et al. (1998), who found $h \gtrsim 0.70$ regardless of the choice of lens models. On the other hand, the PEM + PS model predicts that $h = 0.69^{+0.19}_{-0.15} (1 - \kappa)$ and $0.74^{+0.18}_{-0.17} (1 - \kappa)$ ($\chi^2 = \chi^2_{\text{min}} \text{ CL}$) with the cluster's mass constrained and unconstrained, respectively. Since $\kappa > 0$, the PEM + PS model predicts a significantly lower value for the Hubble constant than the FGSE model. Just recently, Bernstein & Fischer (1999) have considered breaking power-law index of the galaxy to model its potential by up to three subregions of independent elliptical mass distribution, and have included higher-than-quadrupole order terms in the multipole expansion of the cluster potential. Although the models of Bernstein & Fischer (1999) are more sophisticated than previous (published) models and give $\chi^2 \sim N_{\text{deg}}$, direct comparison with the above models (i.e., PEM + PS, FGSE) is not possible because they did not use the full set of magnification constraints (but used only the flux ratios) and deemphasized the observed jet positions by using $\approx 5 \sigma$ in the definition of their $\chi^2$, and because of other differences in the fit procedure, e.g., their fixing the redshifts of the blobs and knots at $z = 1.41$, calculating the positional $\chi^2$ on the source plane by transforming the image position variances to it. They have found $h = 1.04^{+0.21}_{-0.22} (1 - \kappa) (95\% \text{ CL})$.

The mere values of $\chi^2_{\text{min}}$ of the FGSE model and the PEM + PS model cannot distinguish between them. Even worse, neither of the models can be considered to be acceptable if the adopted observational errors are true and assumed to be normally distributed, even though they give

3 Using the same constraints as used by Bernstein & Fischer (1999), but calculating the positional $\chi^2$ on the image plane (thus with no approximations) and allowing the redshifts of the blobs and knots each to be free parameters, we find, for example, that a PEM + PS model with $v_1 = 1.86$ and $d_{1,2} = 4'$ has $\chi^2 \approx 10$ with $N_{\text{deg}} = 4$, which is comparable, in fit quality, to the best-fit models of Bernstein & Fischer (1999).
current “best” fits. In other words, we are at present faced with two difficulties in determining $H_0$ from the Q0957 + 561 lens. One is the poor fit of the present best models to the present observational constraints. This might be, from a modeler’s point of view, partly because the reported measurement errors were underestimated and/or because possible systematic errors (see below) could not be taken into account. In fact, Ba99 point out that the radio jet positions and the magnification constraints derived from the VLBI data using even the improved analysis procedure should be treated with some caution. For example, superluminal motion over the period of the time delay or a milliarcsecond scale deflection by a globular cluster could have influenced the observed jet positions by a scale much larger than their formal uncertainties of $\sim 0.1$ mas. For this reason, it is useful to identify the individual contributions of the observational constraints to the total $\chi^2$. Table 6 shows the $\chi^2$ contributions of the observational constraints for the PEM + PS model. Remarkably, the jet positions contribute most to the $\chi^2$, and they have the second largest $\chi^2/N_{\text{constraint}}$ after the blob magnification ratio. This could be an indication that the above-mentioned effects are present. Nonetheless, the poor overall fit is an indication that more sophisticated and realistic models need to be considered in the future to better describe the mass distribution of the lens. As seen in § 4, the introduction of an ellipticity to the cluster does not improve $\chi^2_{\text{min}}$ at least with the present observational constraints.

Up to now, in lens research an elliptical mass density of a constant ellipticity (e.g., the functional form of eq. [1]) has been used, and regarded as relatively realistic, to model a lensing object. While the mass distribution of equation (1) is very flexible in that all of its parameters are unrestricted, it has several limitations in describing the true mass distribution of a lensing object, e.g., the elliptical lensing galaxy of Q0957 + 561. First, a truncation radius is not included in its functional form. For this reason, its total mass diverges for a radial profile not steeper than the modified Hubble profile (i.e., for $\nu \leq 3$). Second, the ellipticity of equation (1) is constant over the entire range of $r$. The observed light distribution of the lensing galaxy in Q0957 + 561 (Be97) shows a varying ellipticity as a function of $r$ in the inner region of the galaxy. Even when dark mass is the dominant mass component, it would be surprising if the ellipticity of the mass distribution was constant over the entire range of $r$. Third, the position angle of equation (1) is constant over the entire range of $r$. We note that the observed position angle of the lensing galaxy in Q0957 + 561 (Be97) is scattered for a range of $r$. Finally, equation (1) neglects small-scale perturbations or substructure (Mao & Schneider 1998) but assumes a smooth distribution of mass. Bernstein & Fischer’s (1999) models partly account for the second and third points; however, the evaluation of their models depends on interpretation of the current observational constraints. Realistic incorporation of the above-mentioned deviations from equation (1) may improve the fit unambiguously in the future.

In this and previous studies (Ba99; Grogin & Narayan 1996), attempts have been made to put constraints on $H_0$ using lens models despite the poor fits. We are faced with another difficulty in those attempts; two classes of models with similar quality fits can give disagreeing values of $H_0$, and the range of $H_0$ for each class of model is large. This “degeneracy problem” was also encountered in the recent efforts by Impey et al. (1998) to determine $H_0$ using the lens PG 1115 +080. Their dark mass–dominated models and constant mass-to-light ratio model give similar quality fits ($\chi^2/\nu \approx 3$–4 with $N_{\text{deg}} = 1$), but imply $h = 0.44 \pm 0.04$ and $h = 0.65 \pm 0.05$ (for $\Omega_0 = 1$), respectively. Here we focus on the degeneracies in the models of Q0957 +561. Since the FGSE model requires a point mass of $\sim 10^{11} M_\odot$ at the center of the galaxy, which appears to be rather unnatural (even if we interpret it as an effective mass rather than a physical BH mass), we would favor the PEM + PS model over the FGSE model unless the FGSE model was found to give a much better fit than the PEM + PS model. At the same time, the large range of $h$ in the PEM + PS model can be significantly reduced if the cluster’s mass distribution is more accurately determined from observations in the future (e.g., using Chandra X-ray observations, or improved data on weak lensing effects), and/or the galaxy’s mass is securely inferred from the measured velocity dispersions of the galaxy using, e.g., realistic stellar dynamics models for the PEM galaxy model. Observational studies of the two optical blobs and two knots (e.g., measurements of their redshifts and better astrometric and photometric data) will also be useful.

In conclusion, the PEM + PS model of Q0957 +561 shows appreciable ($\approx 2$–3 $\sigma$) discrepancies with some of the present observational constraints (in particular, the radio jet positions and the optical blob magnification ratio), while it is consistent with many others. For this reason, the determined value of $H_0$ should be taken with some caution until the discrepancies are resolved. Nevertheless, the modeled mass distributions of the galaxy and the cluster are consistent with the observations of the galaxy and the cluster (in fact, the PEM + PS model was motivated by recent observations of the galaxy and the cluster). This increases the likelihood that this lens can be used to determine $H_0$ accurately in the future with the aid of better data and more realistic lens models. It is worth mentioning that our present determined value of $H_0$ is consistent with the independent determination by Impey et al. (1998) using the lens PG 1115 +080 and recent determinations from Type Ia supernovae observations, provided that the cluster’s con-

| Constraints          | $N_{\text{constraint}}$ | $\chi^2$ | $\chi^2/N_{\text{constraint}}$ |
|----------------------|------------------------|----------|--------------------------------|
| Jet positions        | 4                      | 19.00    | 4.75                           |
| Core positions       | 2                      | 0.00     | 0.00                           |
| G1 positions         | 2                      | 0.05     | 0.03                           |
| Magnification constraints | 6              | 18.65    | 3.11                           |
| Core magnification ratio | 1               | 0.08     | 0.08                           |
| Blob 2 positions     | 2                      | 3.24     | 1.62                           |
| Blob 3 positions     | 2                      | 0.18     | 0.09                           |
| Blob magnification ratio | 1              | 8.26     | 8.26                           |
| Knot 1 positions     | 2                      | 0.00     | 0.00                           |
| Knot 2 positions     | 2                      | 0.00     | 0.00                           |
| Knot magnification ratio | 1              | 0.19     | 0.19                           |
| Total                | 25                     | 49.6     |                                |

$^4$ The large $\chi^2$ contribution of the blob magnification ratio may not be a surprising result, since current data on the blobs’ surface brightnesses do not permit a very reliable estimate of the centroid magnification ratio (G. Bernstein 1998, private communication).
vergence in the region of the galaxy of Q0957+561 is \( \kappa \sim 0.1-0.2 \), as estimated from recent observations (Chartas et al. 1998; Fischer et al. 1997). For example, if we adopt \( \kappa \approx 0.26 \pm 0.08 \) (1 \( \sigma \)), derived by Bernstein & Fischer (1999) from the Fischer et al. (1997) observation, we find

\[
H_0 = 51^{+14}_{-10} \text{ and } 55^{+14}_{-14} \text{ km s}^{-1} \text{ Mpc}^{-1}
\]

with and without a constraint on the cluster's mass, respectively. These values of \( H_0 \) are in agreement with the values derived from Type Ia supernovae observations (Branch 1998, \( H_0 = 60 \pm 10 \) km s\(^{-1}\) Mpc\(^{-1}\); Schaefer 1998, \( H_0 = 55 \pm 8 \) km s\(^{-1}\) Mpc\(^{-1}\)) as well as the values obtained by Impey et al. (1998). This could be an indication that the results from both Q0957+561 and PG 1115+080 are on the right track. The coming years will be an exciting period for our observational and theoretical efforts to determine the Hubble constant more accurately directly from the time delays of gravitationally lensed systems.

I would like to thank G. Bernstein for comments on the HST blobs and providing the revised blob magnification ratio prior to publication and R. Barkana for sending a revised version of the article cited in this paper. My special thanks are due to D. A. Turnshek for helpful discussions, careful reading of the manuscript, and numerous comments, and his encouragement and support. Financial support from the Andrew Mellon fellowship is gratefully acknowledged. The anonymous referee provided a very useful review of the paper, which was helpful in clarifying and improving the analysis and presentation.

Note added in proof.—With regard to the data presented here in Table 6, G. Bernstein has recently informed us that their new HST/STIS resolved images of blobs 2 and 3 show a magnification ratio of blob 2/blob 3 \( \approx 0.8 \), which is in good agreement with the model predicted value of \( \approx 0.81 \). Thus, our best fit now improves to \( \chi^2_{\text{min}} \approx 5 \).