THE ASTROPHYSICAL JOURNAL, 599:70–85, 2003 December 10

THE HUBBLE CONSTANT FROM THE GRAVITATIONAL LENS B1608+656[1]

L. V. E. KOOPMANS,[2,3] T. TREU,[4,5,6] C. D. FASSNACHT,[7] R. D. BLANDFORD,[2,8,9] AND G. SURPI[2,10]

Received 2003 June 10; accepted 2003 August 19

ABSTRACT

We present a refined gravitational lens model of the four-image lens system B1608+656 based on new and improved observational constraints: (1) the three independent time delays and flux ratios from Very Large Array observations, (2) the radio-image positions from Very Large Baseline Array observations, (3) the shape of the deconvolved Einstein ring from optical and infrared *Hubble Space Telescope* images, (4) the extinction-corrected lens-galaxy centroids and structural parameters, and (5) a stellar velocity dispersion, $\sigma_{\rm ap} = 247 \pm 35$ km s^{-1}, of the primary lens galaxy (G1), obtained from an echelle spectrum taken with the Keck II Telescope. The lens-mass model consists of two elliptical mass distributions with power-law density profiles and an external shear, totaling 22 free parameters, including the density slopes that are the key parameters for determining the value of H_0 from lens time delays. This has required the development of a new lens code that is highly optimized for speed. The minimum-χ^2 model reproduces all observations very well, including the stellar velocity dispersion and the shape of the Einstein ring. A combined gravitational lens and stellar dynamical analysis leads to a value of the Hubble constant of $H_0 = 75^{+7}_{-6}$ km s^{-1} Mpc^{-1} (68% CL; $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$). The nonlinear error analysis includes correlations between all free parameters, in particular the density slopes of G1 and G2, yielding an accurate determination of the random error on H_0. The lens galaxy G1 is $\sim$5 times more massive than the secondary lens galaxy (G2) and has a mass density slope of $\gamma'_{\rm G1} = 2.03^{+0.14}_{-0.14} \pm 0.03$ (68% CL) for $\rho \propto r^{-\gamma'}$, very close to isothermal ($\gamma' = 2$). After extinction correction, G1 exhibits a smooth surface brightness distribution with an $R^{1/4}$ profile and no apparent evidence for tidal disruption by interactions with G2. Given the scope of the observational constraints and the gravitational lens models, as well as the careful corrections to the data, we believe this value of H_0 to be little affected by known systematic errors ($\lesssim$5%).

Subject headings: distance scale — galaxies: elliptical and lenticular, cD — galaxies: kinematics and dynamics — galaxies: structure — gravitational lensing

On-line material: color figures

1. INTRODUCTION

The physics behind gravitational lensing is well understood and primarily based on gravity, i.e., general relativity (e.g., Schneider, Ehlers, & Falco 1992). Since Refsdal (1964), it has been known that arcsecond-scale strong gravitational lenses—those with multiply imaged sources—provide a tool to measure the expansion rate of the universe, i.e., the Hubble constant (H_0), if the mass distribution of the lens(es) and the time delay(s) between the lensed images are known. This provides an elegant one-step and global method to measure H_0, independent of local distance ladder techniques (e.g., Parodi et al. 2000; Saha et al. 2001; Freedman et al. 2001), which could be prone to the accumulation of unknown systematics and seems difficult to improve beyond the current 10% precision on the value of H_0.

More recently, cosmic microwave background (CMB) observations with *Wilkinson Microwave Anisotropy Probe* (*WMAP*) have been combined with different cosmological data sets (i.e., Lyα and large-scale structure observations) to indirectly infer a value of $H_0 = 71^{+4}_{-3}$ km s^{-1} Mpc^{-1} (Bennett et al. 2003; Spergel et al. 2003) similar to that from the *Hubble Space Telescope* (*HST*) Key Project, which finds $H_0 = 72^{+8}_{-8}$ km s^{-1} Mpc^{-1} (Freedman et al. 2001). Because H_0 cannot be measured directly from the CMB data alone, a combination with other data sets is necessary, and thus it is prone to its own systematic errors (e.g., Seljak, McDonald, & Makarov 2003; Bridle et al. 2003). Even so, the agreement between the different values of H_0 from different techniques, including those from Sunyaev-Zeldovich observations (e.g., Mason, Myers, & Readhead 2001), is striking. Since H_0 is one of the most fundamental cosmological parameters, however, its measurement should be subject to multiple cross-examinations, with the ultimate goal of breaking the 10% precision limit.

Although gravitational lensing can provide a straightforward and potentially very precise measurement of H_0, its major problems are that it requires (1) accurately measured time delays obtained from long monitoring campaigns and (2) an accurate determination of the mass distribution in the lenses and the field (i.e., the "lens potential"). At present, about a dozen lens systems have measured time delays with different degrees of precision (see Courbin, Saha, &

[1] Based on observations collected at W. M. Keck Observatory, which is operated jointly by the California Institute of Technology and the University of California, and with the NASA/ESA *Hubble Space Telescope*, obtained at STScI, which is operated by AURA, under NASA contract NAS5-26555.

[2] California Institute of Technology, Theoretical Astrophysics, Mail Code 130-33, Pasadena, CA 91125.

[3] Current address: STScI, 3700 San Martin Drive, Baltimore MD 21218.

[4] California Institute of Technology, Astronomy, Mail Code 105-24, Pasadena, CA 91125.

[5] Current address: University of California, Los Angeles, 8308 Mathematical Sciences Building, Los Angeles, CA 90095-1547.

[6] Hubble Fellow.

[7] Department of Physics, University of California, 1 Shields Avenue, Davis, CA 95616.

[8] Kavli Institute for Particle Physics.

[9] Stanford University.

[10] Current address: FairIsaac, 5935 Cornerstone Court West, San Diego, CA 92121-3711.

Schechter 2002 for a recent summary). This has shifted attention to understanding the often underestimated complexity of the mass distribution of the lens. For example, the first discovered gravitational lens system, Q0957+561 (Walsh, Carswell, & Weymann 1979), has an exquisitely measured time delay with an error $\lesssim 1\%$ (e.g., Kundić et al. 1997; Oscoz et al. 1997, 2001; Haarsma et al. 1999; Colley & Schild 2000; Slavcheva-Mihova, Oknyanskij, & Mihov 2001; Goicoechea 2002; Ovaldsen et al. 2003) but a complex lens potential that includes a nearby cluster. Even though much effort has been put into constraining this particular lens system, the lens potential is still not known to adequate precision to allow a satisfactory determination of H_0 (Keeton et al. 2000 and references therein).

One of the dominant degeneracies in gravitational lens models—even for systems that are relatively isolated and have no major external perturbers (e.g., groups or clusters)—is that between the radial mass profile and the inferred value of H_0. Often point-image constraints (i.e., positions and flux ratios) can be satisfied equally well by mass distributions that have very different slopes of their radial mass profile (e.g., Wambsganss & Paczyński 1994; Refsdal & Surdej 1994; Witt, Mao, & Schechter 1995; Witt, Mao, & Keeton 2000; Wucknitz & Refsdal 2001; Surpi & Blandford 2001; Wucknitz 2002; Williams & Saha 2000). This is similar to the well-known mass-sheet degeneracy (Gorenstein, Shapiro, & Falco 1988). In general, steeper (shallower) mass profiles lead to higher (lower) inferred values of H_0. To break this degeneracy and reliably measure H_0, one has to determine the mass profile of the lens galaxies. Often the most effective way is to use additional external constraints, for example from stellar kinematics of the lens galaxy (e.g., Treu & Koopmans 2002a; Koopmans & Treu 2002).

An often cited result is that of the gravitational lens system PG 1115+080, which gives a value of $H_0 = 44 \pm 4$ km s^{-1} Mpc^{-1} (Impey et al. 1998) for an isothermal lens-mass distribution ($\rho \propto r^{-2}$). Not only is this value low compared with most other lens and nonlens measurements, but the quoted error does not include one of the major sources of uncertainty—the unknown radial mass profile. In fact, early-type galaxies can have a considerable scatter in their radial mass profiles (e.g., Gerhard et al. 2001) around the effective radius, and steeper mass density profiles can increase H_0 well above 60 km s^{-1} Mpc^{-1} (Impey et al. 1998).

It was recently shown that if one includes the observed stellar velocity dispersion (Tonry 1998) of the lens galaxy in PG 1115+080 as additional constraint, the effective slope of the mass density profile can be measured directly and also properly included in the error estimate on H_0. The effective slope of the density profile for this particular system is found to be steeper than isothermal, leading to a higher value of $H_0 = 59^{+12}_{-7} \pm 3$ km s^{-1} Mpc^{-1} (Treu & Koopmans 2002b). This measurement is in better agreement with most other methods, and its larger error bars include the contribution of the residual uncertainty on the mass distribution of the lens. This analysis has led us to initiate a new program with Keck and the Very Large Telescope (VLT) to measure the stellar velocity dispersion of a number of additional systems with known time delays, in order to determine a more precise value of H_0 from gravitational lensing, less affected by the radial mass profile degeneracy.

In this paper we focus on the gravitational lens system B1608+656 (source redshift $z_s = 1.39$, lens redshift $z_l = 0.63$) (Myers et al. 1995; Snellen et al. 1995; Fassnacht et al. 1996). The system is unique in that all three time delays are known between its four lensed images (Fassnacht et al. 1999, 2002b), with accuracies of a few percent. In addition, multicolor images are available in the *HST* archive that show the lensed arcs of the radio-source host galaxy connecting to a full Einstein ring (e.g., Blandford, Surpi, & Kundić 2001; Kochanek, Keeton, & McCloud 2001; Surpi & Blandford 2003). The multicolor images allow us to correct the positions and surface brightness distributions of the lens galaxies for extinction, removing a major source of systematic error common to previous modeling efforts. Finally, we have measured the stellar velocity dispersion of the dominant lens galaxy from a spectrum obtained with the Echelle Spectrograph and Imager (ESI; Sheinis et al. 2002) on the Keck II Telescope. With this additional information and the improved constraints on the time delays, we are well suited to refine the lens models of this system and remove many of the major degeneracies and systematic biases in the determination of H_0.

This paper is organized as follows. In §§ 2 and 3 we discuss the observations and data reduction of the *HST* and Keck data, respectively. In § 4 we describe the observational constraints obtained from Very Large Array (VLA) and Very Large Baseline Array (VLBA) radio observations,[11] *HST* images, and the Keck spectrum. In § 5 we briefly discuss a new lensing code and the χ^2-minimization procedure. In §§ 6 and 7 the modeling results and inferred value of H_0 are discussed. In § 8 we summarize and discuss our results. Throughout this paper, we assume $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$.

2. IMAGING

The main objectives of the analysis of the multicolor *HST* imaging are to obtain (1) accurate extinction-corrected images of the two main lens galaxies, from which their true centroids and structural parameters (i.e., position angle, ellipticity, and effective radius) can be determined and (2) a deconvolved Einstein ring after subtraction of a model of the lens galaxies. The shape of the Einstein ring places additional constraints on the azimuthal structure of the lens model (Kochanek et al. 2001). We note that the analysis of the *HST* images presented here is independent of that in Surpi & Blandford (2003, hereafter SB03) but produces nearly identical color maps. Finally, we bring all astrometric data obtained from the *HST* images to a common "reference frame," i.e., that of the high-resolution VLBA observations. This allows us to combine the radio and optical/infrared observations in a self-consistent way, taking proper care of potential differences (i.e., offsets, rotations, and scalings) between the different data sets.

2.1. *Observations and Data Reduction*

The *HST* images of the gravitational lens system B1608+656 are available from the *HST* archive. Since the primary concern here is obtaining high spatial resolution and an accurate model of the point-spread function (PSF),

[11] The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

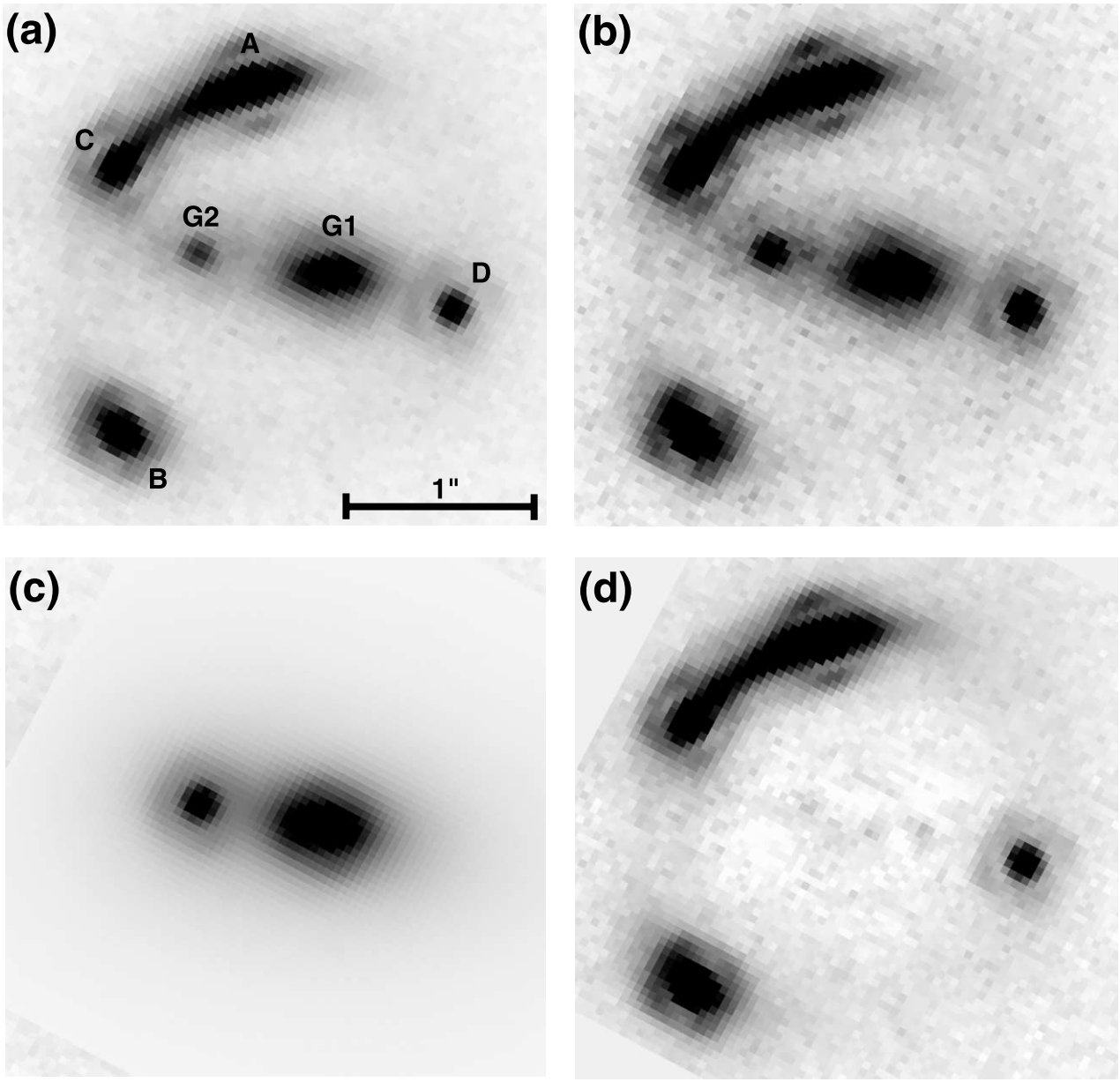

FIG. 1.—Compilation of the *HST*/F160W images of B1608+656: (*a*) The original reduced image of B1608+656. (*b*) Extinction-corrected image (see text). (*c*) Model surface brightness distributions of G1 and G2, fitted to the extinction-corrected brightness distributions. (*d*) Original image of B1608+656 after subtraction of the lens galaxy models extincted by the inferred dust screen.

we select the images with the best resolution and sampling in each of the available bands (see also the discussion in SB03): 41,344 s in seven exposures with the Near Infrared Camera and Multiobject Spectrograph (NICMOS) Camera 1 (NIC1) through filter F160W (GO-7422; PI: Readhead) and four exposures on the Planetary Camera (PC) of the Wide Field Planetary Camera 2 (WFPC2) through two filters, for total exposure times of 11,500 and 11,600 s in the F814W and F606W bands, respectively (*HST* GO-6555; PI: Schechter).

The images are first reduced using a series of IRAF[12] scripts based on the DRIZZLE package (Fruchter & Hook 2002). The scripts clean and combine the images by iteratively refining the measurement of relative offsets and the identification of cosmic rays and defects (similar to the scripts used in Koopmans & Treu 2002 and Treu et al. 2003). The final F160W "drizzled" image is shown in the top left panel of Figure 1. Images of the system in the other filters are similar to the ones shown in SB03 and are not repeated here.

[12] IRAF (Image Reduction and Analysis Facility) is distributed by the National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy under cooperative agreement with the National Science Foundation.

2.2. *Dust Extinction Correction*

The B1608+656 system is characterized by two lens galaxies, G1 and G2, with a strong color gradient in the region connecting the two galaxies. The positions of the centroids of the two galaxies with respect to the multiple images were observed to vary by as much as 86 mas between the F160W

and F606W images (SB03; Koopmans & Fassnacht 1999; Fassnacht et al. 2002b).

To remove this source of systematic uncertainty, we assume that the observed color gradients arise from dust extinction (see SB03) and that the two lens galaxies have the same intrinsic colors. We use the available colors to measure the color excesses and correct for dust extinction based on an extinction law (see below for a discussion). To check for systematics, we perform the correction for each pair of filters, in the reference frame of each one of the filters. The results from each pair of filters are very similar and affect the results only at a very small level, which will be taken into account in the final error analysis on H_0. In the interest of space we will describe as an example only the procedure for the F814W-F160W pair in the F160W reference frame, mentioning the differences in the results for the other frames where relevant. First, we deconvolve the F814W image using a Lucy-Richardson algorithm and a PSF generated using Tiny TIM (Krist 1993). Second, we align the deconvolved F814W image to the F160W image, assuming that the positions of the point sources are the same in each filter and not affected by dust[13] and that the transformation is a rotation and translation with arbitrary axis scaling. Third, the deconvolved and transformed F814W image is convolved with a NICMOS PSF generated by Tiny Tim, to match the resolution of the F160W image. Fourth, the ratio of the matched images in the two filters is used to produce a color map, which is found to be very similar to the one shown in SB03. Fifth, it is assumed that the intrinsic colors of G1 and G2 are spatially uniform and equal to the minimum of the color measured in the region including G1 and G2 (see also Surpi & Blandford 2003). Finally, the color excess is converted into extinction in the individual filters using the following relations, computed using the Galactic extinction law by Cardelli, Clayton, & Mathis (1989) for $R_V = 3.1$: $\mathrm{A(F160W)}/A_V = 0.415$, $\mathrm{A(F814W)}/A_V = 1.091$, and $\mathrm{A(F606W)}/A_V = 1.561$. The extinction-corrected image[14] in F160W is shown in the top right panel of Figure 1.

This procedure is repeated for each pair of filters, in the reference frame of each filter. We emphasize that this procedure does not assume a particular smooth surface brightness profile for the lens galaxies, nor does it force the centroids of the lenses to coincide in the different filters. However, the production of a smooth extinction-corrected surface brightness distribution for the lens galaxies strongly suggests, although does not guarantee, that the extinction correction has been done properly. Similarly, the consistency of the astrometry between various filters provides an additional test of the uncertainties related to the assumption of a specific dust extinction law, the assumption of coincident centroids of the lensed images, and the PSF modeling used for convolution and deconvolution purposes. These tests are discussed in the next two subsections, together with surface photometry and astrometry measurements.

2.3. *Surface Photometry*

Surface photometry is performed on the extinction-corrected F814W and F160W images, as described in Treu et al. (1999, 2001). The F606W image is not used for this task, given that it has the largest dust extinction and the lowest signal-to-noise ratio (S/N), especially on G2. The galaxies are modeled as two $R^{1/4}$ profiles (see Blandford et al. 2001) that are fitted simultaneously to obtain effective radii (R_e) and total magnitudes (Table 1). Only the region well inside the Einstein ring is used in the fit, to minimize contamination by the extended ring structure. Very similar photometric parameters are obtained in the two filters, demonstrating that uncertainties in the alignment of the images and in the PSF modeling do not affect significantly the derived surface photometry. The best-fitting model in filter F160W is shown in the lower left panel of Figure 1.

TABLE 1
OBSERVED PHOTOMETRIC QUANTITIES OF THE LENS GALAXIES G1 AND G2

Parameter	G1	G2
F160W (mag)	16.41 ± 0.08	18.18 ± 0.05
F814W (mag)	18.23 ± 0.12	19.99 ± 0.10
$R_{e,\mathrm{F160W}}$ (arcsec)	0.62 ± 0.05	0.27 ± 0.03
$R_{e,\mathrm{F814W}}$ (arcsec)	0.58 ± 0.06	0.26 ± 0.04
$b/a = (1-e)$	0.5 ± 0.1	$\equiv 1$
Major axis P.A. (deg)	79 ± 2	N/A

NOTES.—All magnitudes are in the Vega system. The systematic uncertainty on the NICMOS zero point (see the NICMOS instrument handbook; Mobasher 2002) is not included. Magnitudes are the total magnitudes obtained from fitting the galaxy surface brightness profiles with an $R^{1/4}$ law. Magnitudes have been corrected for internal extinction, as discussed in the text, but are not corrected for Galactic extinction. The axial ratio of G2 is set to unity during the fit.

To inspect the residuals from the fit, we apply the dust extinction map to the best-fitting $R^{1/4}$ model and subtract the reddened model from the original image. The brightness contribution of the galaxies at the ring radius is found to be negligible. The residual image is shown in the lower right panel of Figure 1, demonstrating that indeed the data are consistent with two relatively smooth spheroids, reddened by a dust lane (Surpi & Blandford 2003), without strong evidence that the primary lens galaxy G1 is much affected by tidal interactions with G2. The ring structure is very clear in the residual image (e.g., Kochanek et al. 2001; Surpi & Blandford 2003), and it is even more prominent in the deconvolved residual image (Fig. 5).

2.4. *Astrometry*

The extinction-corrected images are used to measure the locations (in pixels) of the centroids of G1 and G2, using the task *imcentroid* in IRAF. The deconvolved residual images are used to measure the centroids of the multiple images of the lensed source, as well as the shape of the ring (see § 4.2 and right panels in Fig. 5).

Since we perform all measurements from the *HST* images in their original pixel frames, we need to align the optical and infrared frames to the much more accurate VLBA radio frame. We determine the transformation (i.e., rotation, translation, and scaling) that minimizes the positional differences between the lensed image centroids measured from the *HST* images and their radio counterparts, weighting properly by the measurement errors. We allow for different pixel scales in the x- and y-directions on the NIC1 and PC chips. We find pixel scales of $0\farcs0457/0\farcs0456$ for the PC and

[13] Since the centroids of the lensed images are very sharp, they change very little owing to extinction (see also § 2.4).

[14] Note that this correction is made only to the lens galaxies and not to the Einstein ring.

TABLE 2
ASTROMETRY OF THE B1608+656 SYSTEM

	VLBA/8.4 GHz		*HST*/F814W		*HST*/F160W	
IMAGE	Δx (arcsec)	Δy (arcsec)	Δx (arcsec)	Δy (arcsec)	Δx (arcsec)	Δy (arcsec)
A.....	$\equiv$0.0000	$\equiv$0.0000	−0.0003	+0.0009	−0.0029	+0.0042
B.....	−0.7380	−1.9612	−0.7376	−1.9588	−0.7342	−1.9498
C.....	−0.7446	−0.4537	−0.7455	−0.4528	−0.7469	−0.4446
D	+1.1284	−1.2565	+1.1298	−1.2563	+1.1365	−1.2546
G1.....	N/A	N/A	+0.4124	−1.0592	+0.4261	−1.0581
G2.....	N/A	N/A	−0.3091	−0.9306	−0.2897	−0.9243

0.″0434/0.″0431 for NIC1, in excellent agreement with the values in the *HST* instrument handbook (Mobasher 2002).

The positional differences between the optical/infrared and radio image positions are $\lesssim$10 mas in right ascension and declination for each of the individual lensed images. Since the transformation is determined from the full set of four images, we estimate an error of ∼4 mas in x and y on the transformation between the radio and NICMOS frames. Using this transformation, we map the trace of the brightness peak of the Einstein ring to the radio frame (see § 4.2). Since the transformation errors are negligible compared with the positional determination of the brightness peak along spokes, one can assume for all practical purposes that the radio data and the Einstein rings are perfectly aligned.

We use the same transformation to bring the galaxy centroids in F160W to the radio frame. The combined centroid errors and transformation errors quadratically add to an error of 6 mas in x and y. We adopt this value as the centroiding errors of the galaxies with respect to the lensed radio images. The measurement errors on the radio positions are negligible ($\ll$1 mas; Koopmans & Fassnacht 1999). We repeat the same procedure for the F814W image.

After transformation to the radio frame, the positions of the lens galaxies G1 and G2 agree to within 14 and 20 mas between the F160W and F814W filters, respectively, as opposed to 72 and 38 mas in the uncorrected frame. Slightly larger differences are found using the F606W filter, as expected, because this is the filter where uncertainties in the dust corrections are the most severe.[15] All astrometric measurements from the F160W and F814W filters are listed in Table 2. The deconvolved Einstein rings in the F160W and F814W filters are shown in the left panels of Figure 5.

3. SPECTROSCOPY

In this section we describe the spectroscopic observations of B1608+656 with ESI. The main objective of the observations was to obtain the stellar velocity dispersion of the primary lens galaxy (G1), the measurement of which allows us to determine the power-law slope of its radial mass profile and break the degeneracy with the inferred value of the Hubble constant. Special care is required in order to determine an unbiased stellar velocity dispersion and a correct uncertainty, since G1 is a poststarburst galaxy. This is accomplished through extensive Monte Carlo simulations of artificial spectra with different fractions of "old" and "young" stellar populations.

3.1. *Observations and Data Reduction*

We observed B1608+656 using ESI (Sheinis et al. 2002) on the Keck II Telescope on 2000 July 3, for a total integration time of 5400 s (3 × 1800 s). The seeing was good, with FWHM ≈ 0.″85, and the night was clear. Between each exposure, we dithered along the slit to allow for a better removal of sky residuals at the red end of the spectrum. The slit (20″ in length) was positioned at P.A. = 83°, i.e., aligned to within 4° of the major axis of G1. The slit width of 0.″75 yields an instrumental resolution of $\sigma \sim 20$ km s^{-1}, which is adequate for measuring the stellar velocity dispersion and removing narrow sky emission lines. The centering of the galaxy in the slit was constantly monitored by means of the ESI viewing camera—the galaxy was bright enough to be visible in exposures of a few seconds duration—and we estimate the centering perpendicular to the slit to be accurate to better than 0.″1.

Data reduction was performed using the IRAF package EASI2D,[16] as discussed in Koopmans & Treu (2002) and Sand, Treu, & Ellis (2002, 2003). A one-dimensional spectrum was extracted (Fig. 2) by summing the signal within an aperture corresponding to 1.″7 × 0.″75. The S/N of the spectrum in the vicinity of the G band (∼4304 Å) is 20 Å^{-1}. Note the very prominent Balmer absorption features Hγ and Hδ, typical of a K+A post starburst population (Dressler & Gunn 1983; Myers et al. 1995; Surpi & Blandford 2003).

3.2. *The Stellar Velocity Dispersion of G1*

The luminosity-weighted velocity dispersion $\sigma_{\rm ap}$ of G1 was measured with the Gauss-Hermite pixel-fitting software (van der Marel 1994) on the spectral region covering the observed wavelength ∼6750–7200 (Fig. 2), excluding Hδ and including the G band, Hγ, and Fe λ4383 (Trager et al. 1998). As kinematic templates we used spectra of G–K giants observed at twilight with a 0.″3 slit width, appropriately smoothed to match the instrumental resolution of the 0.″75 slit. The regions around 6870 Å, affected by atmospheric absorption, and the Hγ feature were masked during the fit. The result is robust with respect to changes in the spectral region used in the fit, the fit parameters (e.g., the

[15] The maximum difference in the centroid coordinates between F606W and F160W is now 21 mas, with an rms scatter of 14 mas, as opposed to 86 and 45 mas, respectively, in the uncorrected images of SB03.

[16] Developed by D. Sand and T. Treu (Sand et al. 2003).

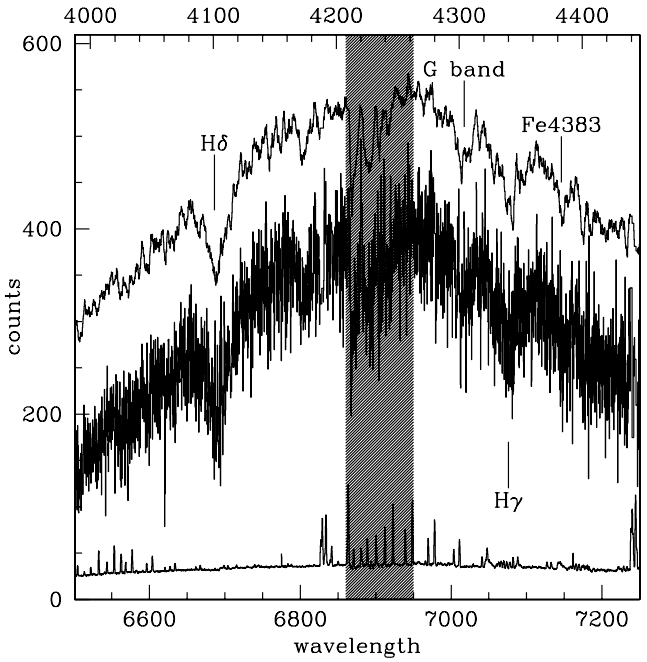

FIG. 2.—Spectrum of B1608+656-G1 in one of the echelle orders around the *G* band (in counts, not flux-calibrated). The lower histogram shows the noise per pixel. A boxcar smoothed version of the spectrum is shown above the original spectrum to guide the eye. The wavelength range used for the kinematic fit was 6750–7200 Å (observed), which includes the *G*-band, Hγ, and Fe λ4384 features, but excludes Hδ. The atmospheric *B* band (*hatched band*) and Hγ were masked out in the kinematic fit.

order of polynomial used to reproduce the continuum level), and the adopted template. The best fit gives $\sigma_{\rm ap} = 247 \pm 35$ km s^{-1} in an aperture of $1\farcs7 \times 0\farcs75$ centered on G1.

To minimize the amount of corrections to the data, we will use only this measurement of the luminosity-weighted velocity dispersion (as opposed to the central velocity dispersion, defined in a standard aperture) in the kinematic analysis (§ 4), properly taking the seeing and aperture effects into account by applying those corrections to the models. However, to facilitate comparison with earlier work and other applications of this measurement, we also give the central velocity dispersion,[17] $\sigma = 269 \pm 40$ km s^{-1}, within a standard circular aperture of radius $R_e/8$.

We note here that if one adopts the measured central velocity dispersion and the structural parameters measured from the extinction-corrected images (Table 1), to determine the offset of G1 from the local fundamental plane, one finds the galaxy to be significantly brighter than quiescent early-type galaxies of the same mass and at similar redshifts. This is consistent with what is typically found for K+A galaxies (e.g., van Dokkum & Stanford 2003). In contrast, this result differs from the findings of Rusin, Kochanek, & Keeton (2003b), who estimated a larger velocity dispersion from the image separation and measured a lower luminosity for G1, resulting in an offset from the local fundamental plane consistent with quiescent evolution.

The rest of this section describes Monte Carlo simulations aimed at ensuring that our measurement provides not only

[17] The correction from luminosity-weighted velocity dispersion to central velocity dispersion is computed on the basis of the range of observed velocity dispersion profiles of local E/S0 galaxies and taking into account seeing effects as described in Treu et al. (2001).

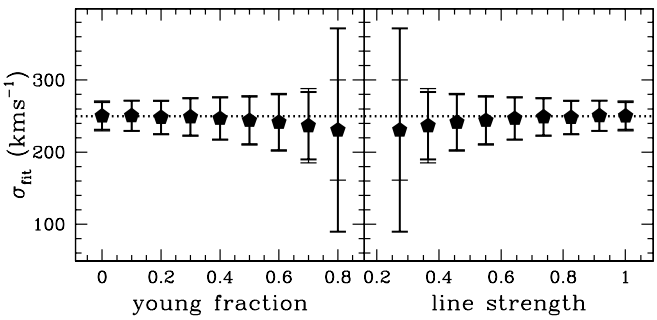

FIG. 3.—Monte Carlo simulations of the stellar velocity dispersion recovery from artificial spectra of B1608+656-G1. The dotted horizontal line shows the input value, and the pentagons show the average value measured from 1000 realizations. The heavy error bars show the average uncertainty estimated by the code, while the light error bars show the rms scatter of the recovered values

an unbiased value of the stellar velocity dispersion of G1 but also a realistic estimate of the uncertainty. In fact, although the S/N and resolution are more than adequate to measure the velocity dispersion of a uniformly old stellar population (e.g., Treu et al. 2001), the spectrum of G1 is more challenging, because of the contamination by young stars. Our Monte Carlo approach works as follows. First, we model the spectrum as a combination of a G8 III star, to represent the old population, and of an A5 V star, to represent the young population (Dressler & Gunn 1983; Franx 1993). The stellar spectra are smoothed to a velocity dispersion of 250 km s^{-1}, taking into account the initial spectral resolution of the stellar spectra,[18] and then combined with variable weights. Second, for each set of weights (hereafter "old" fraction and "young" fraction), 1000 realizations are constructed, adding noise to reproduce the S/N of our data. Third, we run the Gauss-Hermite pixel-fitting software (van der Marel 1994) on each realization to simulate the stellar velocity dispersion measurement.

The results of the Monte Carlo simulations are shown in Figure 3. In the left panel we plot the average and rms scatter (*solid pentagons with thin error bars*) of the recovered velocity dispersion $\sigma_{\rm fit}$ as a function of the fraction of A5 V light (young fraction). For young fractions up to $\sim$0.6, the procedure recovers the input velocity dispersion to within 4%. Only for a young fraction larger than 0.7 does the procedure become biased. This happens because, as the young fraction increases, the information content from the metal absorption features gets washed away by the contribution of the young population. Similarly, for young fractions less than 0.6, the rms scatter (*thin error bars*) agrees remarkably well with the average uncertainty returned by the code (*thick error bars*). We can estimate the fraction of young stars in the spectrum of G1 by comparing the observed depth of the *G* band and of Hγ with those in the model spectra. In Figure 4 we show that a composite spectrum with $\sim$40% of young population matches the observed spectrum, while fractions $\sim$70% and higher produce much deeper Hγ absorption lines than observed. We therefore conclude that the fraction of young stellar light in G1 must be smaller than $\sim$70% and that the value of $\sigma_{\rm ap}$ is the correct and unbiased

[18] The G8 III star is from the ESI data, and the A5 V star is taken from the library of Jacoby, Hunter, & Christen (1984); see Treu et al. (2001) for a discussion of the resolution of that library.

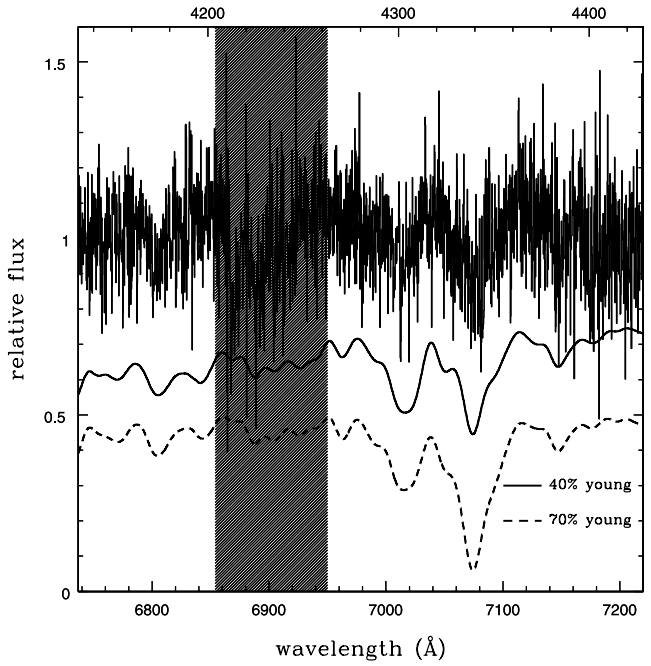

FIG. 4.—Fit of σ_{ap} for G1 in B1608+656 for two different fractions of the young stellar population (shifted for clarity). A young fraction of 70% clearly overpredicts the depth of the Hγ feature.

measurement of the stellar velocity dispersion of G1. Similarly the error estimate is unbiased.

A more quantitative confirmation of this result comes from the inspection of the line strength, which is used by the Gauss-Hermite pixel-fitting software (see definition in van der Marel 1994) to measure the depth of the absorption features: the larger the line strength, the deeper the absorption features. In the models, by construction, the line strength of the old stellar population anticorrelates very well with the young fraction. In fact, a very similar conclusion is drawn if line strength is used as an independent variable to show the result of the Monte Carlo simulations (Fig. 3, *right*). When the line strength of the old stellar population is larger than 0.4 (corresponding approximately to a young fraction smaller than 0.7), the S/N and resolution of our spectrum are sufficient to obtain an unbiased measurement of the stellar velocity dispersion. For a real composite population the interpretation of this parameter is more complex, since it depends on the chemical abundances and possible additional contributions to the continuum. Nevertheless, it provides a useful quantitative measure of the depth of the metal absorption features, which is the main parameter that controls the accuracy of the fit. The best-fitting template yields line strength of 0.5 ± 0.05, also confirming that the measurement of velocity dispersion is unbiased.

4. MASS MODEL AND OBSERVATIONAL CONSTRAINTS

We model the lensing and kinematic properties of both lens galaxies, G1 and G2, assuming that their luminous plus dark matter mass distributions are elliptical (i.e., oblate in three dimensions[19]) and have a radial power-law dependence, (i.e., $\rho \propto r^{-\gamma'}$). In particular, we assume for the lens models a dimensionless surface mass density of

$$\kappa(x, y) = b\left[x^2 + \left(\frac{y}{q_l}\right)^2\right]^{(1-\gamma')/2}, \quad (1)$$

where q_l is the axial ratio of the surface mass distribution and γ' the slope of the corresponding density distribution ($\gamma' = 2$ is designated as "isothermal" henceforth). The mass distribution has its centroid at (x_l, y_l) and a major-axis position angle of θ_l (measured north through east). In addition, we allow for a single external shear for the system with strength γ_{ext} and position angle θ_{ext}.

We choose to use power-law mass models, since they have been successful in reproducing many of the detailed lensing observations (e.g., Kochanek 1995; Cohn et al. 2001; Muñoz, Kochanek, & Keeton 2001; Rusin et al. 2002; Winn, Rusin, & Kochanek 2003) and, more recently, also kinematic observations (e.g., Treu & Koopmans 2002a, 2003) of lens systems and lens ensembles (e.g., Rusin et al. 2003a), even though we acknowledge that this choice might not be unique. We feel at this point, however, that there is no strong incentive to go beyond these models. From a practical standpoint, including the constraints from the Einstein ring becomes computationally very expensive for more complex lens models (see § 5 for a discussion).

For the dynamical modeling (i.e., determining the stellar velocity dispersion), we solve the spherical Jeans equations, assuming a spherical mass distribution with the same power-law index γ' as used for the lens models (i.e., eq. [1]). We model the luminous mass distribution as a trace population—embedded in the luminous plus dark matter potential—with a Hernquist luminosity-density profile (Hernquist 1990). We have also examined Jaffe (1983) models and find differences typically $\lesssim 1\%$ for the models that we explore. Henceforth, we only use the Hernquist profile, which closely follows a $R^{1/4}$ brightness profile over the radial range of interest and is analytically tractable. Since the break radius of the dark matter halo is well beyond the Einstein radius, we assume it to be much larger than the latter, with negligible effects on the calculation of kinematic quantities (see Treu & Koopmans 2002a and Koopmans & Treu 2003 for a full discussion). Finally, we allow for anisotropy of the stellar velocity ellipsoid, modeling the radial anisotropy parameter β (Binney & Tremaine 1987) with a Osipkov-Merritt (Osipkov 1979; Merritt 1985a, 1985b) model with anisotropy radius r_i. For $r_i = \infty$, the stellar velocity dispersion is fully isotropic. We also take the seeing and aperture into account. We refer to Koopmans & Treu (2003) for a further discussion of our kinematic model.

We now continue with a more detailed description of the precise set of constraints that we use in the lensing and kinematic models that are presented in § 6.

4.1. *The Radio and Optical Image Constraints*

The positions of the four lensed radio images can most accurately been determined from high-resolution radio observations. We use the VLBA positions listed in Koopmans & Fassnacht (1999), which are consistent with VLA observations to within ~1 mas (Fassnacht et al. 2002b). We choose positional errors on the VLBA image positions of 1 mas, which are considerably larger than the formal measurement errors (tens of microarcseconds). This error is

[19] Some early-type galaxies show evidence for iso–surface-density twists that could be due to triaxial mass distributions.

TABLE 3
GALAXY CENTROID POSITIONS AND POINT-IMAGE CONSTRAINTS ON THE B1608+656 SYSTEM

Image	Δx (arcsec)	Δy (arcsec)	$S_{\rm norm}$	Δt (days)
A.....	$\equiv 0.0000 \pm 0.001$	$\equiv 0.0000 \pm 0.001$	2.020 ± 0.404	31.5 ± 1.5
B.....	-0.7380 ± 0.001	-1.9612 ± 0.001	1.000 ± 0.200	$\equiv 0.0$
C.....	-0.7446 ± 0.001	-0.4537 ± 0.001	1.034 ± 0.207	36.0 ± 1.5
D	$+1.1284 \pm 0.001$	-1.2565 ± 0.001	$0.347 \pm \infty$	77.0 ± 1.5
G1.....	$+0.4261 \pm 0.006$	-1.0581 ± 0.006	N/A	N/A
G2.....	-0.2897 ± 0.006	-0.9243 ± 0.006	N/A	N/A

NOTES.—The positions for components A–D are from Koopmans & Fassnacht 1999, the time delays and flux ratios are from Fassnacht et al. 2002b, and the G1 and G2 positions are from this paper.

approximately that by which the most massive mass substructures in the lens (expected to be on the order of $\sim 10^9$ $M_\odot$) could potentially change the image positions. Since mass substructures are "randomly" distributed, such positional shifts can practically be regarded as random errors.

Radio flux ratios might not be reliable constraints on macro lens models, either because of mass substructure (e.g., Mao & Schneider 1998; Metcalf & Madau 2001; Keeton 2001b; Chiba 2002; Metcalf & Zhao 2002; Dalal & Kochanek 2002; Bradač et al. 2002, 2003; Keeton, Gaudi, & Petters 2003), because of radio microlensing (e.g., Koopmans & de Bruyn 2000; Schechter & Wambsganss 2002), or because of interstellar medium (ISM) propagation effects (Koopmans et al. 2003). Hence, we conservatively adopt a large error of 20% on the individual radio fluxes (i.e., $\sim$28% on the flux ratios). This is supported by the fact that the formal errors on the flux ratios within a single season are much smaller than the differences between seasons (Fassnacht et al. 2002b). For the flux ratios we adopt the approximate median value from three independent monitoring seasons of B1608+656 (Fassnacht et al. 2002b). Since previous models had difficulties reproducing the flux ratio of image D (Koopmans & Fassnacht 1999, 2002b), it is not used as a constraint. Hence, if the refined lens models can recover the flux of image D, we would have additional confidence in the lens models.

Finally, we use the time delays between the VLA images as constraints (Fassnacht et al. 1999; Fassnacht et al. 2002b). Time-delay ratios are particularly strong constraints on the mass profile, since they are almost a direct measure of the depth of the potential at the lensed-image positions. Since there are three independent time delays and one Hubble constant to solve for, they add two constraints on the mass model (we emphasize here that time delays are affected by substructure only at a negligible level, $\ll 1\%$). All point-image constraints are listed in Table 3.

4.2. *The Einstein-Ring Constraints*

The F160W image of B1608+656 exhibits a fully connected Einstein ring around the two lens galaxies (Kochanek et al. 2001; Surpi & Blandford 2003). The structure of Einstein rings can be used as additional constraint on the gravitational lens models (Blandford et al. 2001; Kochanek et al. 2001). In particular, the unit vector along a radial spoke at the position of the brightness peak (on the spoke), when projected on the source plane, is perpendicular to the gradient of the surface brightness distribution of the unlensed source (Kochanek et al. 2001). When the isophotes of the source are assumed to be elliptical—often a reasonable assumption—the trace of brightness peaks (along radial spokes) places extra constraints on the lens potential at the minimal cost of adding four additional free parameters (ring-source position, ellipticity, and position angle). Moreover, for an elliptical source, the trace of the ring is independent of assumptions about the radial behavior of the surface brightness distribution of the source.

Even though the S/N decreases with increasing distance from the central nucleus of the lensed source along the ring, it is still sufficient to fully trace the ring (see Fig. 1; Kochanek et al. 2001). This can be done even more accurately on the deconvolved image, where the Einstein ring is "sharper" in the radial direction (Fig. 5). We note that color gradients over the image have very little effect on the brightness peaks of the sharp ring (see also Kochanek et al. 2001). To trace the deconvolved Einstein ring, we define 90 independent spokes, separated by 4° in angle (consistent with the improved image resolution after deconvolution), radiating from a point roughly in the middle of the ring. The origin of the spokes is arbitrary. The brightness distribution of the deconvolved ring along each spoke is subsequently fitted with a Gaussian. From this fit, the position of the brightness peak and the FWHM of the ring are obtained. The resulting fitted ring is shown in Figure 5. To ensure that the ring is not significantly affected by uncertainties produced by the deconvolution, we repeat the same procedure for the F814W image. We assume that the error on the ring peak is proportional to the ring width, since it is hard to estimate an error from a deconvolved image. We adopt $0.42 \times \sigma_r$ as error for F160W (with $\sigma_r = {\rm FWHM}/2.35$ from the Gaussian fit), which gives $\chi^2/{\rm dof} \sim 1$, only weakly dependent on the particular lens model. This particular choice is not crucial, however, and somewhat higher or lower scale factors can be chosen without affecting the modeling results. Hence, we put more weight on those parts of the ring that are sharpest (near the radio images) and less weight on the less well-defined parts of the ring. Notice the remarkable agreement between the traces of the ring in F160W and F814W (Fig. 5), lending credit to the trace (compare also with Kochanek et al. 2001, who find a very similar trace from the original, not deconvolved, ring).

4.3. *Constraints from the Lens Galaxies*

As for the Einstein ring, we use the transformation discussed in § 2.4 to bring the positions of the two lens galaxies in the F160W and F814W images to that of the VLBA

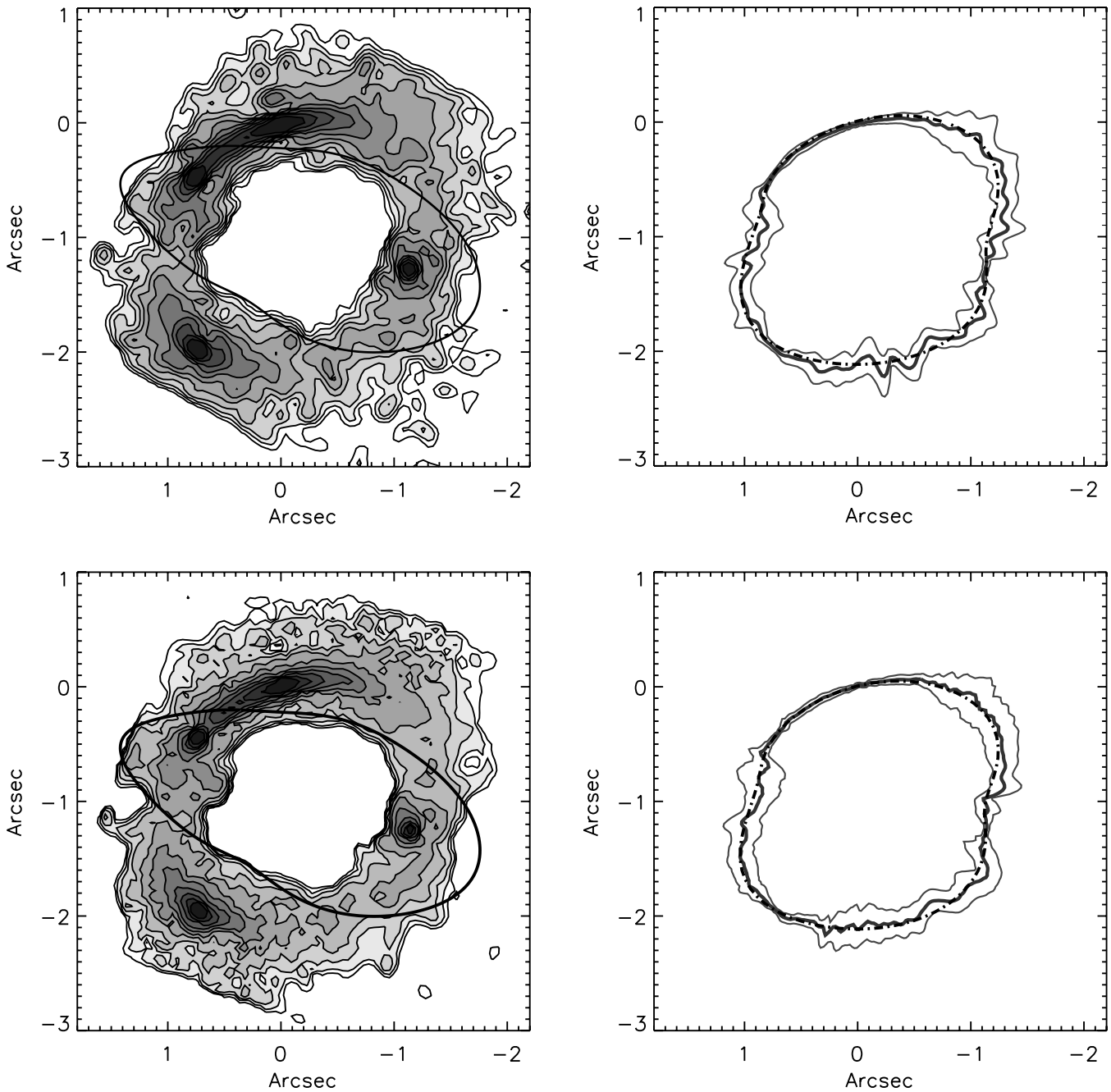

FIG. 5.—Deconvolved Einstein rings of B1608+656 in the F160W (*upper left*) and F814W (*lower left*) bands, after extinction correction and subtraction of a model of the two lens galaxies. The contour levels increase by factors of 2 (arbitrary units). The solid curve indicates the outer critical curve (see text) of our best lens model (§ 6.4). The right panels indicate the trace of the Einstein ring (*thick solid line*) and the $\pm 1 - \sigma_r$ width of the ring (*two thin solid lines*). To smooth the trace, we plot it for 360 spokes (only 90 are used in the modeling; see text). The dot-dashed curves show the minimum-χ^2 SPLE1+D models. [*See the electronic edition of the Journal for a color version of this figure*.]

frame with positional errors of 6 mas with respect to the radio images (Table 3). The good agreement of the galaxy centroids from different filters—after extinction correction—seem to indicate that the extinction corrections applied to the F160W image have been done properly. Confidence in the extinction correction is further supported by the fact that, after correction, G1 is well described by a $R^{1/4}$ brightness profile (see also Blandford et al. 2001). It seems highly unlikely that an improper extinction correction would produce a well-behaved smooth surface brightness distribution from an originally patchy image. In the lens models, therefore, we adopt the light centroids from the extinction-corrected F160W images, transformed to the radio frame, as the final values for the mass centers of G1 and G2.

In addition, modeling of the extinction-corrected brightness distribution gives a position angle of $79^\circ \pm 2^\circ$ (north to east) of G1. Since the extinction-corrected galaxies appear relatively smooth and unperturbed, it is safe to assume that G1 is most likely not much affected by tidal interactions

with G2. The position angle of G1 is used as a prior on the position angle of its mass distribution.

As a final model constraint, we use the stellar velocity dispersion of G1, $\sigma_{ap} = 247 \pm 35$ km s^{-1} (§§ 6 and 7), which provides information on the slope of its radial mass profile.

5. LENSING CODE AND χ^2-MINIMIZATION

To model the B1608+656 system, we have developed a new lens code, highly optimized for speed and accuracy. This is necessary because most of the models require χ^2-minimization in a 22-dimensional parameter space (see § 6.1 for a listing of all free model parameters). Since the structure of the Einstein ring is included as a constraint, each iteration (i.e., χ^2 calculation) is computationally expensive, especially if there are no analytic solutions for the lens potential. Furthermore, a full nonlinear error analysis of the parameters of the lens model requires a repetitive minimization of the χ^2 function for different sets of fixed lens parameters.

To maximize speed, the code is first optimized in its use of memory, which can be critical when copying large data structures (i.e., those containing all constraints and intermediate results). Second, analytic solutions are used wherever possible, as well as publicly available fast algorithms developed for specific power-law mass models (e.g., Barkana 1998; Chae, Khersonsky, & Turnshek 1998; Chae 2002). A significant speedup of the χ^2-minimization and error analysis is gained by using the MINUIT minimization package from CERN (James 1994), which has been optimized for finding function minima in large-dimensional parameter spaces with a minimum number of function evaluations. We have compared the results from the new code, for B1608+656 and various simple test systems, with those from a previously used code (e.g., Koopmans & Fassnacht 1999) and from the publicly available code from Keeton (2001a). We find that they all agree to within the expected numerical errors.

The increase in speed of the new code over the other codes, when we include the Einstein ring data for B1608+656, can be several orders of magnitude, depending on the specific situation. For example, χ^2-minimization with 22 free parameters usually takes minutes, compared to hours for the code by Keeton (2001a), when we use mass models that are not analytically tractable (e.g., nonisothermal power-law density profiles). A full nonlinear error analysis for all parameters—a built-in option with MINUIT—takes several hours with the new code but only several minutes extra for a single parameter (e.g., H_0). A similar analysis with the code by Keeton (2001a) is expected to take several days and is clearly not feasible for systems of this complexity.

To find the global χ^2 minimum, we first optimize all free model parameters using only the radio data as constraints. This rapidly leads to a model that fits all the radio constraints extremely well. In the second step, the Einstein ring data are added to the constraints. Since the mass model parameters are already relatively close to the best solution, it takes only several minutes to converge to the χ^2 minimum. To check that indeed the correct global χ^2 minimum has been found, the models are started from different initial conditions and also using different minimization techniques (i.e., simulated annealing and downhill simplex methods). Monte Carlo minimization is done once a χ^2 minimum is reached. Finally, the minimum in all parameter-space directions is traced while performing a nonlinear error analysis.[20] None of this has led us to suspect that the global χ^2 minimum has not been found. This is also supported by the fact that the observed lens properties are reproduced well by the lens models.

6. RESULTS

In this section we describe the results from the lens models. After listing the free parameters and adopted priors of the general model in § 6.1, we consider in § 6.2 a simplified problem in which the slopes of the galaxies are fixed to isothermal. This is done to facilitate comparison with previous work and to illustrate the effect of correcting the centroid positions for dust extinction. In § 6.3 we explore the full range of parameter space and hence its effects on H_0, allowing the slopes to vary to reproduce the lensing geometry. Finally, in § 6.4 we add the dynamical constraints to reduce the uncertainties and break the mass profile degeneracy of G1. All results have been calculated using the new lens code and are based *only* on constraints from the radio data, the NICMOS F160W images, and the stellar kinematics. Constraints from the WFPC2 F814W images will be used as additional check on systematic effects in § 7.3.

6.1. *Free Parameters and Priors*

To allow maximum freedom in the lens models and a proper calculation of the significance of the derived value of the Hubble constant, we allow nearly all of the model parameters (i.e., position, lens strength, position angle, ellipticity, and density slope) of each lens galaxy to vary, constraining several with Gaussian priors. The source adds three additional free parameters (i.e., position and flux), the Hubble constant adds one, and the external shear adds two. The model of the elliptical source that describes the Einstein ring (§ 4.2) adds an additional four free parameters (position, ellipticity, and position angle). In total there are up to 22 free parameters.

We use Gaussian priors on the galaxy positions, placing them at their observed positions and allowing for 1 σ errors of 6 mas in x and y (Table 3). Hence, χ^2 increases by unity if the galaxy positions deviate by 1 σ from the observed positions. Similarly, we allow the position angle of G1 to vary using a Gaussian prior of $79^\circ \pm 2^\circ$. Since B1608+656 is part of a small group (e.g., Fassnacht et al. 2002a), we allow for an external shear. The group is not massive and has a low velocity dispersion (C. D. Fassnacht et al. 2003, in preparation). We therefore place a conservative 0.00 ± 0.10 (1 σ positive) prior on the strength of the external shear, although this prior is not particularly important.

Finally, we place a Gaussian prior on the density slope of the secondary galaxy (G2), for which we have no kinematic information. We choose $\gamma'_{G2} = 2.00 \pm 0.10$, fully consistent with the spread in values of γ' found from 0047−285 and MG 2016+211 (Treu & Koopmans 2002a; Koopmans & Treu 2003). Since G2 is ~5 times less massive than G1, its contribution to the properties of the lens system is less pronounced. We discuss this prior further in § 7.3.

[20] The nonlinear error analysis consists of reminimizing χ^2 for a range of fixed values for the parameter of interest (James 1994). The 68% confidence limits, quoted in this paper, correspond to the parameter range with $\Delta\chi^2 = 1$ with respect to the minimum-χ^2 solution.

6.2. *Isothermal Lens Model*

To allow a comparison with previous modeling efforts (Koopmans & Fassnacht 1999; Kochanek et al. 2001; Fassnacht et al. 2002b), we first set the density slopes of both G1 and G2 to isothermal (i.e., $\gamma' = 2$). The structure of the ring is included as a constraint, and the priors are as discussed in § 6.1.

We have listed the model parameters of the minimum-χ^2 solution in Table 4, where this model is designated as SIE. Compared with previous isothermal lens models (Koopmans & Fassnacht 1999; Fassnacht et al. 2002b), we find that the inferred value of the Hubble constant, $H_0 = 71^{+5}_{-6}$ km s^{-1} Mpc^{-1} (68% CL), has increased by about 6–8 km s^{-1} Mpc^{-1}. The error includes all random errors and correlations between free parameters. This increase is partly due to the two main lens galaxies being closer to each other after extinction correction. Although each galaxy is isothermal, their combined potential becomes steeper, leading to an increase in H_0 (e.g., Wucknitz 2002; Williams & Saha 2000). Other contributions come from including external shear and widening the errors on the image flux ratios and positions, which change the χ^2 surface.

Besides reproducing all point-image constraints well, an excellent fit (nearly identical to Fig. 5) to the boxy structure of the Einstein ring is also found. A comparison by eye to the fit in Kochanek et al. (2001, their Fig. 5) shows that our fit is significantly better, in particular between the image pairs B-C and D-A. The difference can probably be attributed to the fact that we use a deconvolved Einstein ring shape and extinction-corrected lens galaxy centroids. We find a source axial ratio of $q_s = 0.64 \pm 0.03$ and position angle of $\theta_s = 113 \pm 2$. The axial ratio is similar, but the position angle of the ring source is significantly different from the value in Kochanek et al. (2001). The optical source has a minor offset from the source for the radio images ($\lesssim 0\farcs03$), which could, for example, be attributed to a small asymmetry (e.g., lopsidedness or $m = 1$ mode) of the host galaxy or deviations from ellipticity. Forcing the centroid of the host galaxy to coincide with the radio source increases the χ^2 of the ring fit (by $\sim$30%) but changes the resulting mass model parameters only marginally. Since $m = 1$ modes for the brightness distribution of galaxies are not uncommon and automatically accounted for in models with a free host-galaxy centroid, we choose not to force such a coincidence. We find values identical to those derived from the F160W image when the ring shape from the F814W data is used, and also when using the code from Keeton (2001a).

As a further check, we calculate the stellar velocity dispersion of G1 on the basis of this isothermal lens model. Details of the calculation method are outlined in Koopmans & Treu (2003) and § 4. We also come back to this in more detail in § 6.4. We find $\sigma_{\rm ap} = 230$ km s^{-1} for isotropic models ($r_i = \infty$) and $\sigma_{\rm ap} = 250$ km s^{-1} for anisotropic models with $r_i = R_{\rm eff}$, inside the proper aperture and corrected for seeing effects. These values agree remarkably well with the measurement of $\sigma_{\rm ap} = 247 \pm 35$ km s^{-1}, even though this information has not been used in the isothermal models. This suggests that the simplest isothermal lens model is probably not too far from the final model.

6.3. *Power-Law Lens Models*

The next step is to allow the slope of the dominant lens galaxy (G1) to vary. Since the isothermal models show that G1 is $\sim$5 times more massive than G2, the contribution of G2 to the structure of the Einstein ring and the lensed images is relatively small. In general, we find that the properties of G2 are ill-defined if we leave its density slope (or ellipticity; see § 7.3) entirely unconstrained. We therefore

TABLE 4
LENS-MASS MODEL PARAMETERS FOR B1608+656

Parameter	SIE	SPLE1	SPLE2	SPLE1+D (isotopic)	SPLE1+D (anisotropic)	SPLE2+D (isotropic)	SPLE2+D (anisotrpic)
x_l (arcsec)	0.425	0.426	0.425	0.425	0.425	0.425	0.426
	−0.291	−0.290	−0.290	−0.291	−0.290	−0.291	−0.290
y_l (arcsec)	−1.071	−1.069	−1.069	−1.069	−1.069	−1.069	−1.069
	−0.929	−0.928	−0.928	−0.928	−0.928	−0.928	−0.928
b (arcsec$^{\gamma'-1}$)	0.531	0.553	0.553	0.526	0.555	0.531	0.553
	0.288	0.263	0.263	0.269	0.263	0.268	0.263
q_l	0.606	0.607	0.606	0.604	0.603	0.605	0.606
	0.340	0.308	0.308	0.318	0.307	0.316	0.308
θ_l (deg)	76.8	77.0	77.0	77.0	77.0	77.0	77.0
	69.2	68.5	68.5	68.4	68.4	68.4	68.5
γ'	≡2.00	$1.99^{+0.20}_{-0.20}$	$1.99^{+0.14}_{-0.15}$	$2.05^{+0.14}_{-0.13}$	$2.00^{+0.14}_{-0.15}$	$2.04^{+0.11}_{-0.12}$	$2.00^{+0.12}_{-0.12}$
	≡2.00	2.12	2.12	2.12	2.12	2.12	2.12
$\gamma_{\rm ext}$	0.085	0.081	0.081	0.077	0.081	0.078	0.081
$\theta_{\rm ext}$ (deg)	6.8	10.6	10.6	13.4	10.4	12.9	10.6
x_s (arcsec)	0.095	0.103	0.102	0.088	0.102	0.090	0.102
y_s (arcsec)	−1.070	−1.070	−1.070	−1.069	−1.070	−1.069	−1.070
q_h	0.639	0.640	0.640	0.634	0.640	0.635	0.640
θ_h (deg)	113.2	113.0	113.0	112.1	113.1	112.3	113.0
H_0 (km^{-1} s Mpc^{-1})	71^{+5}_{-6}	74^{+10}_{-11}	74^{+7}_{-6}	76^{+7}_{-6}	74^{+7}_{-6}	75^{+6}_{-5}	74^{+6}_{-5}
χ^2	99.8	98.0	98.0	98.1	98.0	98.1	98.0

NOTES.—The first (second) line for each parameter denotes values for G1 (G2). The 68% confidence limits on H_0 and $\gamma'_{\rm G1}$ have been determined from a full nonlinear error analysis, where all free parameters are varied for a range of fixed values of these parameters until χ^2 increases by unity. We limit ourself to these two most important parameters, since this calculation is computationally very expensive.

place a prior on its slope, $\gamma'_{G2} = 2.00 \pm 0.10$, as suggested from similar mass models of other lens systems (§ 6.1). The slope of G1 is left fully unconstrained. We designate this model as the "singular power-law ellipsoid" (SPLE) mass model. The SPLE1 model is that model with a prior *only* on the power-law slope of G2.

The results are listed in Table 4. We find a slightly better fit to the data (i.e., $\Delta\chi^2 = -1.8$) than for the isothermal models. Most notably, the slope of G1 is found to be $\gamma'_{G1} = 1.99^{+0.20}_{-0.20}$, nearly identical to isothermal. The slope of G2 is somewhat steeper than isothermal, which leads to a slight increase in the Hubble constant, $H_0 = 74^{+10}_{-11}$ km s^{-1} Mpc^{-1} (68% CL) compared with the SIE mass models. Since the density slopes are included as free parameters, the resulting error on H_0 is larger than for the isothermal models that have fixed density slopes. The error on H_0 is less than 15%, even without kinematic constraints. If we include a similar prior on the slope of G1 as on G2 (the SPLE2 model), the results are again nearly identical (Table 4). The error on H_0 decreases considerably, because it is dominated by the uncertainty on the density slopes, as discussed previously. As for the isothermal models, the point-image constraints and the shape of the Einstein ring are reproduced well.

6.4. *Lensing and Dynamical Models*

The final step in the modeling effort is to include the stellar velocity dispersion of G1 as an additional constraint. This is not as simple as for a case with a single lens galaxy. Since we have two lens galaxies, we cannot simply attribute all of the mass within the Einstein radius (or critical curve) to only the lens galaxy G1.

We therefore follow a different approach. First we determine the minimum-χ^2 parameters for a range of models with fixed values of γ'_{G1} ranging from 1.6 to 2.4. We then calculate the mass enclosed within an aperture with a radius of 1″ centered on G1, using the lens model parameters. Using the enclosed mass and the slope of the density profile of G1, we can then determine the stellar velocity dispersion inside the observed aperture (for details, see § 4 and Koopmans & Treu 2003). The model velocity dispersion is compared with the observed value and the resulting $\Delta\chi^2_\sigma$, assuming a Gaussian error, is added to the χ^2_l from the lens constraints alone.

Since b, q, and γ' all enter into the determination of the mass enclosed within 1″, the minimum of $\chi^2_l + \Delta\chi^2_\sigma$ is not necessarily the true minimum. However, since γ'_{G1} is by far the parameter that most dominates the determination of the inferred stellar velocity dispersion,[21] for all practical purposes the difference between this approach and the correct approach (which is to calculate the dispersion at every χ^2 evaluation) is negligible compared with the final error on H_0. We note that calculating the dispersion at every minimization step would slow down the minimization process by several orders of magnitude and is not feasible.

The resulting best-fit models for the lensing and dynamical models, designated as SPLE1+D and SPLE2+D, are listed in Table 4. We do calculations for two different anisotropy radii, for $r_i = \infty$ (isotropic) and $r_i = R_{\rm eff}$ (anisotropic). Since the observed stellar velocity dispersion is so close to that already predicted from the lens models alone, we see that the final results barely differ from the results presented in § 6.3, except that the error on H_0 has decreased. In particular, we find that the SPLE2 models have errors on H_0 nearly identical to those of the SPLE1+D models. Hence, both the lens and lens plus dynamics models clearly prefer an isothermal mass distribution for G1. In Table 5 we have listed the recovered point-image properties from the best SPLE1+D model, assuming $r_i = \infty$. Nearly identical results are found for $r_i = R_{\rm eff}$. The best-fit model of the Einstein ring is shown in Figure 5, along with the observed ring. Figure 6 shows the critical and caustic curves and the time-delay surface. For comparison, the outer critical curve is also overplotted on the data on the left panel of Figure 5. Note that the critical curve is consistent with going through the saddle points of the ring, as expected. This provides a consistency check of the model, independent of the assumption of ellipticity for the source.

TABLE 5

RECOVERED POINT-IMAGE PROPERTIES FROM THE MINIMUM-χ^2 SPLE1+D (ISOTROPIC) MODEL

Image	Δx (arcsec)	Δy (arcsec)	$S_{\rm norm}$	Δt (days)
A.....	+0.00005	+0.00003	2.0408	32.42
B.....	−0.73823	−1.96125	0.9313	0.00
C.....	−0.74461	−0.45379	1.0869	36.20
D	+1.12845	−1.25630	0.4230	76.51

7. THE HUBBLE CONSTANT FROM B1608+656

In this section we discuss the value of H_0 from B1608+656 and its random and systematic errors in more detail, based on the class of models presented in § 6.

7.1. *Models Using Only Lensing Constraints*

The SIE and SPLE1 and SPLE2 models are mass models based solely on the gravitational lensing constraints (i.e., no stellar dynamics) and give a value of H_0 ranging from 71 to 74 km s^{-1} Mpc^{-1} (Table 4). We adopt the value of $H_0 = 74^{+10}_{-11}$ km s^{-1} Mpc^{-1} as the best lensing-only determination, which includes the uncertainty on the density slope of the G1 without any assumptions except for a prior on the slope of G2. We discuss the latter in more detail in § 7.3. We note that including a prior on G1 similar to that on G2 does not change the result, since the slope of G1 is already found to be nearly isothermal, i.e., $\gamma'_{G1} = 1.99^{+0.20}_{-0.20}$ (68% CL) without a prior. The lensing-only models therefore predict an isothermal density profile for G1 to within 10%.

7.2. *Models Using Lensing and Dynamical Constraints*

All other models we consider include the stellar velocity dispersion of G1 as a constraint. In § 6 we showed that the simplest lens models that we consider, with two isothermal galaxies, predict a stellar velocity dispersion close to the observed value (within 1 σ for both the isotropic and strongly anisotropic models). This agreement shows that the slope of the density profile of G1 cannot deviate from isothermal too strongly. Indeed, it is found that the slope of G1 is isothermal to within a few percent if we include the stellar velocity dispersion as an additional constraint. We finally adopt the median value of $H_0 = 75^{+7}_{-6}$ km s^{-1} Mpc^{-1}

[21] We note that the effect of changing the aperture, effective radius, etc., are completely negligible here.

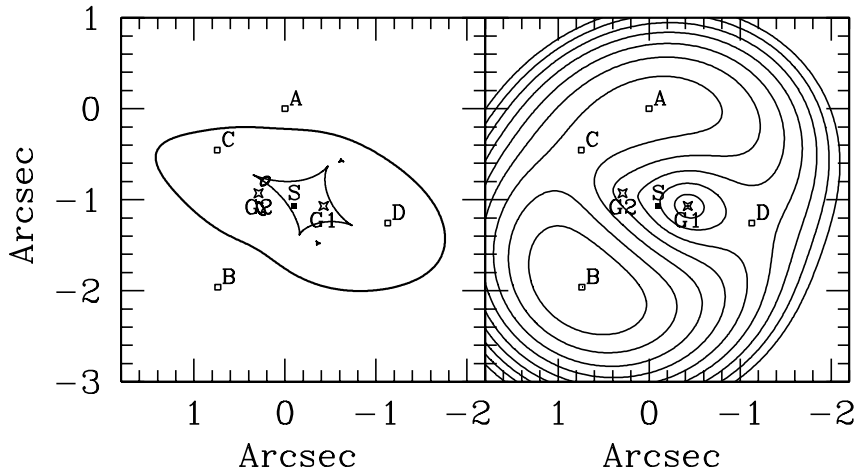

FIG. 6.—*Left:* Critical (*thick*) and caustic (*thin*) curves of the SPLE1+D model. The galaxy positions are indicated by stars, the image positions by open squares, and the source by a filled square. *Right:* Contours indicate constant time delays starting at $\Delta t = 0$ at image B and increasing in steps of 10 h^{-1} days. [*See the electronic edition of the Journal for a color version of this figure.*]

as the best determination. A systematic error of ± 1 km s^{-1} Mpc^{-1} is estimated, due to the unknown anisotropy of G1. We note that including the stellar dynamics has not significantly changed the value of H_0 but reduces its error considerably, to less than 10%. Similarly, the slope of G1 is $\gamma'_{\rm G1} = 2.03^{+0.14}_{-0.14} \pm 0.03$ (68% CL), where we take the average value of the SPLE1+D models with $r_i = R_{\rm eff}$ and $r_i = \infty$. The latter error is the uncertainty due to the unknown anisotropy of the stellar velocity ellipsoid of G1 (Table 4).

7.3. *Systematics*

Finally, we discuss the dominant systematic errors that could potentially be left in the models. The results have not drastically changed since earlier models, but given the wealth of new constraints (e.g., the Einstein ring, improved time-delay measurements, stellar dynamics,), the results have become more robust. The error on H_0 has shrunk even though we have included more free parameters (e.g., external shear, varying slopes, and galaxy positions) in the model. The smaller error on H_0 is due to improvements in the measurement of the time delays and the inclusion of the Einstein ring and stellar dynamics. Support for the refined models is further given by the flux ratio of image D, which is recovered to within 22% (Tables 3 and 5), even though it has not been included as a constraint on the models. All previous models had great difficulty with this. Hence, at first glance, the results appear little affected by remaining systematics.

Even so, we remain careful about assumptions, the most important of which is that for the prior on the density slope of G2 (§ 6.1). If we increase the width of the prior, we find that H_0 increases somewhat (i.e., the density profile of G2 becomes steeper), but also that its axial ratio decreases to very small values $\lesssim 0.2$ (P.A. remains similar to those in Table 4). Since G2 is most likely a disk galaxy (because of the copious amounts of dust associated with it), a small axial ratio is not unlikely, but a value of $\sim$0.1–0.2 appears unrealistic, especially since G2 has no discernible ellipticity in the F160W image (Fig. 1). If we take priors of $q_{\rm G2} = 1.0 \pm 0.2$ and loosen the prior on the density slope to $\gamma'_{\rm G2} = 2.0 \pm 0.3$—both of which are also reasonable, given the observations—we find that the resulting value of H_0 increases by at most a few percent.

Next, we have assumed a spherical mass distribution for G1 in our stellar dynamical models. Since we measure a luminosity-weighted dispersion within a relatively large aperture ($\gtrsim R_{\rm eff}$), many assumptions about the phase-space density of stars have little effect on the final measured stellar velocity dispersion. For example, anisotropy has little effect on H_0. A change of $\pm 5\%$ in the average dispersion would also affect H_0 on a level of $\lesssim 3\%$. We therefore believe the assumption of spherical symmetry to be of little effect, given the current error on the observed stellar velocity dispersion of G1 ($\sim$14%). We also notice that the slope of G1 from lensing alone and from stellar dynamics fully agree, giving no indication of large remaining systematic errors.

As a check on the astrometry, dust correction, and deconvolution procedure, we have used the Einstein ring from the F814W observations as a constraint. We follow exactly the same procedure as for the F160W image. We find that the two lens galaxies lie slightly farther apart than in the NICMOS images, as discussed previously. This suggests that a smaller value of H_0 will be found. However, the best SPLE1+D model produces a somewhat steeper density profile for G1 (by 2%), which compensates for the somewhat larger lens galaxy separation and leads to an identical value of H_0. Hence, it appears that inclusion of the ring "forces" the combined lens potential of G1 and G2 to be slightly steeper than isothermal, leading to a robust value of H_0 independent of which data set is used.

We therefore conclude that given the assumptions and priors—which are all consistent with the observations of B1608+656—remaining systematic errors on H_0 probably add up to not more than about 5%, dominated by those from assumptions in the stellar dynamical model and the priors on G2.

However, an important systematic error that could still affect the value of H_0 is a mass-sheet degeneracy (Gorenstein et al. 1988) due to the small group around the lens system (Fassnacht et al. 2002a). Although our model reproduces the observed stellar velocity dispersion well for an isothermal mass distribution, a strong mass sheet could have led us to overestimate the enclosed

galactic mass and therefore to underestimate the inferred density slope, given the observed velocity dispersion constraint. The observed quasi-isothermality, however, does not necessarily argue against a mass sheet since the mass density profile does not have to be exactly isothermal. We note also that the effect on the slope partially compensates the well-known scaling of H_0 with κ_c for a given mass model. We defer a precise quantification of the convergence (κ_c) and shear of this group, and their effects on the lens system and H_0, to a forthcoming publication with additional data on the group (C. D. Fassnacht et al. 2003, in preparation), but we believe it to be relatively small on the basis of the "poorness" of the group and its small velocity dispersion.

Finally, we note that the effects of the cosmological model are negligible. Changing Ω_Λ from 0.7 to 0.0, with $\Omega_m = 0.3$, increases H_0 by 2%. A drastic change to ($\Omega_m = 1.0$, $\Omega_\Lambda = 0.0$) decreases H_0 by 7%. However, within the current limits on the cosmological model set by *WMAP* (Spergel et al. 2003), changes in H_0 are $\lesssim 1\%$.

8. DISCUSSION AND CONCLUSIONS

We have presented significantly improved and refined mass models of the gravitational lens B1608+656—compared with previous modeling efforts—with the aim of determining a value of the Hubble constant that is less affected by previously known systematics (e.g., radial mass profile, dust extinction).

New constraints on the mass model include the following: (1) the stellar velocity dispersion of the dominant lens galaxy (G1), as measured with ESI; (2) the deconvolved Einstein ring seen in the *HST* F160W and F814W images—the former of which is little affected by dust—corrected for the contribution from the lens galaxies; (3) the extinction-corrected lens-galaxy centroids and structural parameters, the former being one of the major uncertainties in previous lens models; and (4) recent improvements in the determination of the three independent time delays in this four-image lens system (Fassnacht et al. 2002b).

Lens models have also been improved in allowing for many additional free parameters compared with previous modeling efforts (Koopmans & Fassnacht 1999; Fassnacht et al. 2002b), including the galaxy positions, the position angle of G1, an external shear, and the density slopes of G1 and G2. Some of these parameters are constrained with observational priors. The freedom in the lens model (up to 22 free parameters) allows for a proper analysis of the error on the inferred value of H_0, including all observational errors and correlations between free parameters.

The improvements in the observations *and* lens models have led to a Hubble constant from B1608+656 that is less affected by known systematic errors than previously. We obtain $H_0 = 75^{+7}_{-6}$ km s^{-1} Mpc^{-1} (68% CL), with a $\pm 5\%$ error contributed by known systematic errors (for $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$). This value is higher than found from previous models (e.g., Koopmans & Fassnacht 1999; Fassnacht et al. 2002b), predominantly because, after extinction correction, the two lens galaxies are somewhat closer together than previously thought, their density profiles are slightly steeper than isothermal, and external shear is included in the models.

First, we find that the new lens models reproduce all radio-image constraints to within the errors. This agreement might imply that we have been too conservative in the error estimates and might also have slightly overestimated the final error on H_0. The flux ratio of image D is also reproduced to within 22%, a considerable improvement over previous models (e.g., Koopmans & Fassnacht 1999; Fassnacht et al. 2002b). Image D is the closest image to the dominant lens galaxy and might be affected by structure in the stellar (or dark matter) mass distribution of lens galaxy G1, which is poorly represented by the global mass model.

Second, the lens models reproduce the observed stellar velocity dispersion of G1, and its mass profile is found to be nearly isothermal, $\gamma'_{\rm G1} = 2.03^{+0.14}_{-0.14} \pm 0.03$ (68% CL). This result is not as straightforward as one might expect, since the inferred stellar velocity dispersion of G1 depends strongly on the mass that G2 contributes to the lens properties. If the lens model predicted too large a mass for G2, the mass of G1 would decrease (because lensing tightly fixes the mass enclosed by the lensed images), as would the inferred stellar velocity dispersions. Similarly, if the mass is correct, but the density slope of G1 inferred from lensing alone is wrong, the inferred velocity dispersion would be incorrect as well. The fact that the lensing model (without dynamical constraints) already predicts a mass model for G1 that accurately reproduces the observed stellar velocity dispersions makes the models, in our opinion, very believable. The addition of the dynamical constraint therefore does not change the results from lensing alone, but it significantly tightens the error on the value H_0.

Third, the lens models reproduce the boxy shape of the Einstein ring both in the F160W and F814W bands well and give the same value of H_0 irrespective of which of these data sets we use as a constraint. In addition, the critical curves (see Fig. 5) cross the Einstein ring at the saddle points of its brightness distribution, as required (e.g., Blandford et al. 2001).

We note that the resulting value H_0 from B1608+656 seems difficult to reconcile with Kochanek (2002, 2003), who finds that perfectly isothermal lens galaxies should lead on average to $H_0 = 48 \pm 3$ km s^{-1} Mpc^{-1} (see also Kochanek & Schechter 2003). We find a significantly higher value of H_0 even if the dominant lens galaxy (G1) is *assumed to be* isothermal. One possible explanation is that the time-delay systems examined by Kochanek (2002,2003) have lens galaxies with luminous plus dark matter mass profiles steeper than isothermal, enough to yield H_0 in better agreement with that obtained from B1608+656. Although indeed—*on average*—isothermal models appear to provide a good representation of the mass distribution of early-type galaxies (e.g., Treu & Koopmans 2002a; Koopmans & Treu 2003; Rusin et al. 2003a), significant departures in several individual cases is not necessarily inconsistent with the observed spread of mass slopes of local early-type galaxies (e.g., Fig. 1 in Gerhard et al. 2001; Bertin & Stiavelli 1993), especially at small radii where the stellar mass starts to dominate (e.g., Bertola et al. 1993). The lens galaxy in PG 1115+080 (Treu & Koopmans 2002b) provides one such example of a lens-galaxy mass profile steeper than isothermal inside the Einstein radius.

Conversely, a mass-sheet degeneracy could have led us to overestimate H_0 from B1608+656, although the mass-sheet degeneracy is somewhat broken by the kinematic measurement. In this respect, we only note that a convergence of at least $\kappa_c \sim 0.4$ is needed to bring our value of H_0 in agreement with 48 ± 3 km s^{-1} Mpc^{-1}, an unrealistically high

value, and we defer further discussion to a forthcoming paper (C. D. Fassnacht et al. 2003, in preparation).

Another viable alternative would be that the time-delay systems examined by Kochanek (2002, 2003) have not been modeled in enough detail yet and could also be affected by nearby clusters or groups as seen near some of these systems (e.g., RXJ 0911; Kneib, Cohen, & Hjorth 2000; Hjorth et al. 2002). A simple prediction is therefore the following: *If* these lens systems are isothermal and *not* affected by external perturbers, then the stellar velocity dispersion should be that predicted from a simple isothermal lens model (as for B1608+656). *If*, on the other hand, the observed stellar velocity dispersion is much higher that predicted from isothermal models, this would be a clear sign of a steeper mass profile (as for PG 1115+080). Hence, stellar kinematics of lens galaxies is a key measurement to further study this apparent inconsistency.

Even so, the analysis of B1608+656 has demonstrated that when enough information is available, it is possible to reach accuracies of $\lesssim 10\%$ (random) on the value of H_0 (or $\lesssim 15\%$ when including systematic errors) taking the radial mass profile properly into account. A similar uncertainty was obtained by applying a joint lensing and dynamical analysis to PG 1115+080 (Treu & Koopmans 2002b). This level of uncertainty is comparable to the accuracy achieved by the individual methods (SNae–Ia, fundamental plane, etc.) that contributed to the final value of H_0 from the *HST* Key Project (Freedman et al. 2001). Hence, obtaining sufficiently good data—including on the local environments—for a handful of gravitational lens systems seems to be a promising way to break the 10% limit on the uncertainty on the value of H_0.

Finally, the analyses of B1608+656 and PG 1115+080 have demonstrated that there are no simple ways to reach this goal, and it does not suffice to assume a radial mass profile, without measuring it for each individual system. That approach probably does not lead to a reliable value of H_0. High-quality observations and detailed lens-mass models that allow for many degrees of freedom are required to obtain an accurate (and realistic) estimate of the Hubble constant and its uncertainty. We are currently collecting spectroscopic data for additional lens systems with measured time delays, with this aim in mind. In addition, optical and radio monitoring programs that should provide more (and/or improved) time delays in the near future are continuing, and forthcoming observations with *HST* could provide improved constraints on the lens galaxies.

We thank Richard Ellis and Jean-Paul Kneib for useful comments on this manuscript and stimulating conversations. We are grateful to an anonymous referee for comments that helped clarify the manuscript. The use of the Gauss-Hermite pixel fitting software developed by R. P. van der Marel is gratefully acknowledged. The ESI data were reduced using software developed in collaboration with D. Sand. We acknowledge the use of the *HST* data collected by the CASTLES collaboration. L. V. E. K. and T. T. acknowledge support by NASA through grant AR-09960 from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, under NASA contract NAS5-26555. L. V. E. K. also acknowledges an STScI Fellowship grant. T. T. acknowledges support from NASA through Hubble Fellowship grant HST-HF-01167.01-A. L. V. E. K. and R. D. B. acknowledge NSF grants AST 99-00866 and 0206286, and G. S. acknowledges NASA grant NAG5-7007 and CONICET in Argentina. We thank J. Miller, M. Bolte, R. Guhathakurta, D. Zaritsky, and all the people who worked to make ESI such a nice instrument. Finally, the authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

REFERENCES

Barkana, R. 1998, ApJ, 502, 531
Bennett, C., et al. 2003, ApJ, ApJS, 148, 97
Bertin, G., & Stiavelli, M. 1993, Rep. Prog. Phys., 56, 493
Bertola, F., Pizzella, A., Persic, M., & Salucci, P. 1993, ApJ, 416, L45
Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton: Princeton Univ. Press)
Blandford, R., Surpi, G., & Kundić, T. 2001, in ASP Conf. Ser. 237, Gravitational Lensing: Recent Progress and Future Goals, ed. T. G. Brainerd & C. S. Kochanek (San Francisco: ASP), 65
Bradač, M., Schneider, P., Lombardi, M., Steinmetz, M., & Koopmans, L. V. E. 2003, A&A, submitted (astro-ph/0306238)
Bradač, M., Schneider, P., Steinmetz, M., Lombardi, M., King, L. J., & Porcas, R. 2002, A&A, 388, 373
Bridle, S. L., Lahav, O., Ostriker, J. P, & Steinhardt, P. J. 2003, Science, 299, 1532
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Chae, K. 2002, ApJ, 568, 500
Chae, K., Khersonsky, V. K., & Turnshek, D. A. 1998, ApJ, 506, 80
Chiba, M. 2002, ApJ, 565, 17
Cohn, J. D., Kochanek, C. S., McLeod, B. A., & Keeton, C. R. 2001, ApJ, 554, 1216
Colley, W. N., & Schild, R. E. 2000, ApJ, 540, 104
Courbin, F., Saha, P., & Schechter, P. L. 2002, Gravitational Lensing as an Astrophysical Tool, ed. F. Courbin & D. Minniti (Springer: Berlin), 1
Dalal, N., & Kochanek, C. S. 2002, ApJ, 572, 25
Dressler, A., & Gunn, J. E. 1983, ApJ, 270, 7
Falco, E. E., Gorenstein, M. V., & Shapiro, I. I. 1985, ApJ, 289, L1
Fassnacht, C. D., Lubin, L. M., Squires, G. K., & Readhead, A. C. S. 2002a, AAS Meeting, 201, 8009
Fassnacht, C. D., Pearson, T. J., Readhead, A. C. S., Browne, I. W. A., Koopmans, L. V. E., Myers, S. T., & Wilkinson, P. N. 1999, ApJ, 527, 498
Fassnacht, C. D., Womble, D. S., Neugebauer, G., Browne, I. W. A., Readhead, A. C. S., Matthews, K., & Pearson, T. J. 1996, ApJ, 460, L103
Fassnacht, C., Xanthopoulos, E., Koopmans, L. V. E., & Rusin, D. 2002b, ApJ, 581, 823
Franx, M. 1993, ApJ, 407, L5
Freedman, W. L., et al. 2001, ApJ, 553, 47
Fruchter, A. S., & Hook R. N. 2002, PASP, 114, 144
Gerhard, O., Kronawitter, A., Saglia, R. P., & Bender, R. 2001, AJ, 121, 1936
Goicoechea, L. J. 2002, MNRAS, 334, 905
Gorenstein, M. V., Shapiro, I. I., & Falco, E. E. 1988, ApJ, 327, 693
Haarsma, D. B., Hewitt, J. N., Lehar, J., & Burke, B. F. 1999, ApJ, 510, 64
Hernquist, L. 1990, ApJ, 356, 359
Hjorth, J., et al. 2002, ApJ, 572, L11
Impey, C. D., Falco, E. E., Kochanek, C. S., Lehár, J., McLeod, B. A., Rix, H.-W., Peng, C. Y., & Keeton, C. R. 1998, ApJ, 509, 551
Jacoby, G. H., Hunter, D. A., & Christian, C. A. 1984, ApJS, 56, 257
Jaffe, W. 1983, MNRAS, 202, 995
James, F. 1994, MINUIT: Function Minimization and Error Analysis, Reference Manual, ver. 94.1, CERN Program Library Long Write-up D506
Keeton, C. R.. 2001a, ApJ, submitted (astro-ph/0102340)
———. 2001b, ApJ, submitted (astro-ph/0111595)
Keeton, C. R., Gaudi, B. S., & Petters, A. O. 2003, ApJ, in press
Keeton, C. R., et al. 2000, ApJ, 542, 74
Kneib, J., Cohen, J. G., & Hjorth, J. 2000, ApJ, 544, L35
Kochanek, C. S. 1995, ApJ, 445, 559
———. 2002, ApJ, 578, 25
———. 2003, ApJ, 583, 49
Kochanek, C. S., Keeton, C. R., & McLeod, B. M. 2001, ApJ, 547, 50
Kochanek, C. S., & Schechter, P. L. 2003, Measuring and Modeling the Universe, ed. W. L. Freedman (Cambridge: Cambridge Univ. Press), in press (astro-ph/0306040)
Koopmans, L. V. E., & de Bruyn, A. G. 2000, A&A, 358, 793
Koopmans, L. V. E., & Fassnacht, C. D. 1999, ApJ, 527, 513
Koopmans, L. V. E., & Treu, T. 2002, ApJ, 568, L5

Koopmans, L. V. E., & Treu, T. 2003, ApJ, 583, 606
Krist, J. 1993, in ASP Conf. Ser. 52, Astronomical Data Analysis Software and Systems II, ed. R. J. Hanisch, R. J. V. Brissenden, & J. Barnes (San Francisco: ASP), 536
Kundić, T., et al. 1997, ApJ, 482, 75
Mao, S., & Schneider, P. 1998, MNRAS, 295, 587
Mason, B. S., Myers, S. T., & Readhead, A. C. S. 2001, ApJ, 555, L11
Merritt, D. 1985a, AJ, 90, 1027
———. 1985b, MNRAS, 214, 25
Metcalf, R. B., & Madau, P. 2001, ApJ, 563, 9
Metcalf, R. B., & Zhao, H. 2002, ApJ, 567, L5
Mobasher, B. 2002, *Hubble Space Telescope* User's Guide, ed. B. Mobasher (Baltimore: STScI)
Muñoz, J. A., Kochanek, C. S., & Keeton, C. R. 2001, ApJ, 558, 657
Myers, S. T., et al. 1995, ApJ, 447, L5
Oscoz, A., Mediavilla, E., Goicoechea, L. J., Serra-Ricart, M., & Buitrago, J. 1997, ApJ, 479, L89
Oscoz, A., et al. 2001, ApJ, 552, 81
Osipkov, L. P. 1979, AZh Pis'ma, 5, 77
Ovaldsen, J. E., Teuber, J., Schild, R. E., & Stabell, R. 2003, A&A, 402, 891
Parodi, B. R., Saha, A., Sandage, A., & Tammann, G. A. 2000, ApJ, 540, 634
Refsdal, S. 1964, MNRAS, 128, 307
Refsdal, S., & Surdej, J. 1994, Rep. Prog. Phys., 57, 117
Rusin, D., Kochanek, C. S., & Keeton, C. R. 2003a, ApJ, 595, 29
Rusin, D., Norbury, M., Biggs, A. D., Marlow, D. R., Jackson, N. J., Browne, I. W. A., Wilkinson, P. N., & Myers, S. T. 2002, MNRAS, 330, 205
Rusin, D., et al. 2003b, ApJ, 587, 143
Saha, A., Sandage, A., Tammann, G. A., Dolphin, A. E., Christensen, J., Panagia, N., & Macchetto, F. D. 2001, ApJ, 562, 314
Sand, D. J., Treu, T., & Ellis, R. S. 2002, ApJ, 574, L129
Sand, D. J., Treu, T., Smith, G. P., & Ellis, R. S. 2003, ApJ, submitted (astro-ph/0309465)
Schechter, P. L., & Wambsganss, J. 2002, ApJ, 580, 685
Schneider, P., Ehlers, J., & Falco, E. E. 1992, Gravitational Lenses (Berlin: Spinger)
Seljak, U., McDonald, P., & Makarov, A. 2003, MNRAS, 342, L79
Sheinis, A. I., Bolte, M., Epps, H. W., Kibrick, R. I., Miller, J. S., Radovan, M. V., Bigelow, B. C., & Sutin, B. M. 2002, PASP, 114, 851
Slavcheva-Mihova, L. S., Oknyanskij, V. L., & Mihov, B. M. 2001, Ap&SS, 275, 385
Snellen, I. A. G., de Bruyn, A. G., Schilizzi, R. T., Miley, G. K., & Myers, S. T. 1995, ApJ, 447, L9
Spergel, D., et al. 2003, ApJS, 148, 258
Surpi, G., & Blandford, R. 2001, in ASP Conf. Ser. 237, Gravitational Lensing: Recent Progress and Future Goals, ed. T. G. Brainerd & C. S. Kochanek (San Francisco: ASP), 103
———. D. 2003, ApJ, 584, 100 (SB03)
Tonry, J. L. 1998, AJ, 115, 1
Trager, S. C., Worthey, G., Faber, S. M., Burstein, D., & Gonzalex, J. J. 1998, ApJS, 116, 1
Treu, T., Ellis, R. S., Kneib, J.-P., Dressler, A., Smail, I., Czoske, O., Oemler, A., & Natarajan, P. 2003, ApJ, 591, 53
Treu, T., & Koopmans, L. V. E. 2002a, ApJ, 575, 87
———. 2002b, MNRAS, 337, L6
Treu, T., Stiavelli, M., Casertano, C., Møller, P., & Bertin G. 1999, MNRAS, 308, 1037
Treu, T., Stiavelli, M., Møller, P., Casertano, C., & Bertin G. 2001, MNRAS, 326, 221
van der Marel, R. P. 1994, MNRAS, 270, 271
van Dokkum, P. G., & Stanford, S. A. 2003, ApJ, 585, 78
Walsh, D., Carswell, R. F., & Weymann, R. J. 1979, Nature, 279, 381
Wambsganss, J., & Paczyński, B. 1994, AJ, 108, 1156
Williams, L. L. R., & Saha, P. 2000, AJ, 119, 439
Winn, J. N., Rusin, D., & Kochanek, C. S. 2003, ApJ, 587, 80
Witt, H. J., Mao, S., & Keeton, C. R. 2000, ApJ, 544, 98
Witt, H. J., Mao, S., & Schechter, P. L. 1995, ApJ, 443, 18
Wucknitz, O. 2002, MNRAS, 332, 951
Wucknitz, O., & Refsdal, S. 2001, in ASP Conf. Ser. 237, Gravitational Lensing: Recent Progress and Future Goals, ed. T. G. Brainerd & C. S. Kochanek (San Francisco: ASP), 157