IMPROVED EVIDENCE FOR A BLACK HOLE IN M32 FROM HST/FOS SPECTRA.
II. AXISymmetric DYNAMICAL MODELS

ROELAND P. VAN DER MAREL
Institute for Advanced Study, Olden Lane, Princeton, NJ 08540
N. CRETTON AND P. TIM DE ZEEUW
Sterrewacht Leiden, Postbus 9513, 2300 RA Leiden, The Netherlands
AND
HANS-WALTER RIX
Steward Observatory, University of Arizona, Tucson, AZ 85721
Received 1997 May 12; accepted 1997 September 8

ABSTRACT

Axisymmetric dynamical models are constructed for the E3 galaxy M32 to interpret high spatial resolution stellar kinematical data obtained with the Hubble Space Telescope (HST). Models are studied with two-integral, \( f(E, L_z) \), phase-space distribution functions and with fully general three-integral distribution functions. The latter are built using an extension of Schwarzschild’s approach: individual orbits in the axisymmetric potential are calculated numerically, and populated using nonnegative least-squares fitting so as to reproduce all available kinematical data, including line-of-sight velocity profile shapes. The details of this method are described in companion papers by Rix et al. and Cretton et al.

Models are constructed for inclinations \( i = 90^\circ \) (edge on) and \( i = 55^\circ \). No model without a nuclear dark object can fit the combined ground-based and HST data, independent of the dynamical structure of M32. Models with a nuclear dark object of mass \( M_\bullet = 3.4 \times 10^6 M_\odot \) (with 1 \( \sigma \) and 3 \( \sigma \) error bars of 0.7 \( \times 10^6 M_\odot \) and 1.6 \( \times 10^6 M_\odot \), respectively) do provide an excellent fit. The inclined models provide the best fit, but the inferred \( M_\bullet \) does not depend sensitively on the assumed inclination. The models that best fit the data are not two-integral models, but like two-integral models they are azimuthally anisotropic. Two-integral models therefore provide useful low-order approximations to the dynamical structure of M32. We use them to show that an extended dark object can fit the data only if its half-mass radius is \( r_h \lesssim 0.08 \) (0.26 pc), implying a central dark matter density exceeding \( 1 \times 10^8 M_\odot \) pc\(^{-3}\).

The inferred \( M_\bullet \) is consistent with that suggested previously by ground-based kinematical data. However, radially anisotropic axisymmetric constant mass-to-light ratio models are now ruled out for the first time, and the limit on the dark matter density implied by the HST data is now stringent enough to rule out most plausible alternatives to a massive black hole.

The dynamically inferred \( M_\bullet \) is identical to that suggested by existing models for HST photometry of M32 that assume adiabatic growth (over a timescale exceeding \( 10^6 \) yr) of a black hole into a preexisting core. The low activity of the nucleus of M32 implies either that only a very small fraction of the gas that is shed by evolving stars is accreted onto the black hole or, alternatively, that accretion proceeds at very low efficiency, e.g., in an advection-dominated mode.

Subject headings: black hole physics — galaxies: elliptical and lenticular, cD — galaxies: individual (M32) — galaxies: kinematics and dynamics — galaxies: nuclei — galaxies: structure

1. INTRODUCTION

It is generally believed that active galaxies and quasars are powered by the presence of massive black holes (BHs) in their nuclei and that such BHs are present in many, possibly all, quiescent galaxies as well (see Kormendy & Richstone 1995; Lynden-Bell 1996, and Rees 1996 for reviews of this paradigm and its history). Evidence for this can be derived from studies of the dynamics of stars and gas in the nuclei of individual galaxies. The high spatial resolution data that can now be obtained with the Hubble Space Telescope (HST) allows the existing evidence to be strengthened considerably. The present paper is part of a new HST study of the quiescent galaxy M32, in which the presence of a BH has long been suspected based on the steep central rotation velocity gradient and nuclear peak in the velocity dispersion seen in ground-based data (e.g., Tonry 1984, 1987; Dressler 1984; van der Marel et al. 1994a; Bender, Kormendy, & Dehnen 1996). The main results of our project were summarized and discussed in the context of other recent work, in van der Marel et al. (1997a). The acquisition and reduction of the stellar kinematical HST data were described in van der Marel, de Zeeuw, & Rix (1997b, hereafter Paper I). Here we present new dynamical models that we have used to interpret the combined HST and ground-based data.

1 Based on observations with the NASA/ESA Hubble Space Telescope obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555.

2 Hubble Fellow.

3 Alfred P. Sloan Fellow.
Two types of self-consistent\textsuperscript{4} dynamical models have been constructed previously to interpret the ground-based data for M32. Dressler & Richstone (1988) and Richstone, Bower, & Dressler (1990) used a method based on Schwarzschild’s (1979) technique, in which individual orbits are calculated and superposed, to provide a self-consistent model that fits a given set of data. These “maximum entropy” (Richstone & Tremaine 1988) models could fit the (then available) data only by invoking the presence of a central dark mass of \((0.7-8) \times 10^6 M_\odot\). The models are general in the sense that they make no assumptions about the dynamical structure of the galaxy. However, a drawback was that only spherical geometry was considered. Even though the models can be made to rotate, it remains unclear what systematic errors are introduced when they are applied to a flattened (E3) galaxy like M32. An alternative approach has been to construct axisymmetric models with phase-space distribution functions (DFs) that depend only on the two classical integrals of motion, \(f=f(E, L_z)\), where \(E\) is the binding energy and \(L_z\) is the angular momentum component along the symmetry axis, both per unit mass. These models properly take flattening and rotation into account. To fit the M32 data, they require the presence of a central dark mass between \(1.8 \times 10^6 M_\odot\) (van der Marel et al. 1994b; Qian et al. 1995; Dehnen 1995) and \(3 \times 10^6 M_\odot\) (Bender et al. 1996). The disadvantage of these models is that they have a special dynamical structure. The velocity dispersions in the meridional plane are isotropic, \(\sigma_r=\sigma_\theta\), which might not be the case in M32. However, the models do fit the observed line-of-sight velocity profile (VP) shapes without invoking freely adjustable parameters, which provides some reason to believe that the M32 DF may not be too different from the form \(f(E, L_z)\).

The previous work on M32 has shown that models with a BH can fit the ground-based data, but the modeling has not been general enough to demonstrate that a BH is required.\textsuperscript{5} In fact, the spatial resolution of the ground-based data might have been insufficient for this to be the case (see Appendix A for a discussion of this issue). It has certainly not been sufficient to rule out a cluster of dark objects (as opposed to a central BH) on the basis of theoretical arguments; Goodman & Lee (1989) showed that this requires a resolution of \(\lesssim 0.1\). Our new HST data of the nuclear region of M32 were obtained with the HST Faint Object Spectrograph (FOS) through square apertures of 0.068 and 0.191, respectively, yielding the highest spatial resolution stellar kinematic data for M32 obtained to date. The results (see, e.g., Fig. 2 below) show a steeper rotation curve and higher central velocity dispersion than the best ground-based M32 data. The primary goals of our project are to determine whether these new HST data rigorously rule out models without any dark mass and to what extent they constrain the mass and size of the dark object in M32.

To obtain constraints on the presence of a dark object that are least dependent on a priori assumptions about the DF, we need to compare the HST data not only with the predictions of axisymmetric \(f(E, L_z)\) models but also with the predictions of models with a fully general dynamical structure. Orbit superposition techniques provide the most straightforward approach to construct such models. However, orbit superposition is more difficult to implement for the axisymmetric case than for the spherical case: the orbits are not planar and typically possess an additional integral of motion, so that the orbit library must sample three rather than two integrals of motion. Furthermore, a larger parameter space must be explored because of the unknown inclination angle. This implies that larger amounts of CPU time and computer memory are required. However, other than that, there are no reasons why such models would be infeasible. Motivated by the increased speed and memory capacity of computers, we therefore developed a technique to construct fully general axisymmetric orbit superposition models, which fit any given number of observed photometric and kinematic constraints.

Independent software implementations were written by H.-W. R, N. C., and R. v. d. M. Our technique may be viewed as the axisymmetric generalization of the spherical modeling used by Richstone and collaborators, with the important additional feature that we calculate VP shapes and include an arbitrary number of Gauss-Hermite moments in the fit. We take into account the error on each observational constraint to obtain an objective \(\chi^2\) measure for the quality of fit. Our basic algorithm is described in Rix et al. (1997, hereafter R97) and summarized in de Zeeuw (1997). R97 provide an application to the spherical geometry; Creton et al. (1997, hereafter C97) present the extension to the axisymmetric case. Here we summarize the main steps of the axisymmetric algorithm briefly and focus on the application to M32. The resulting models are the most general yet constructed for M32.

The paper is organized as follows. In § 2 we discuss our parameterizations for the stellar mass density and for the potential of the dark object. In § 3 we describe the construction of models with \(f(E, L_z)\) DFs, and in § 4 we compare the predictions of these models, both with and without BHs, to the kinematical data. In § 5 we outline the orbit-supposition technique for constructing models with a fully general dynamical structure, and in § 6 we compare the predictions of these models to the data. We construct models with an extended dark object in § 7. We summarize and discuss our main conclusions in § 8. Readers interested primarily in the results of our models may wish to skip §§ 2, 3, and 5.

2. MASS DENSITY AND POTENTIAL

We adopt a parametrized form for the axisymmetric mass density of M32:

\[
\rho(R, z) = \rho_0(m/b)^2\left[1 + (m/b)^2\right]\left[1 + (m/c)^2\right]^{3/2},
\]

\[
m^2 \equiv R^2 + (z/q)^2.
\]

(1)

The mass density \(\rho\) is related to the luminosity density \(j\) according to \(\rho = Y j\), where \(Y\) is the average mass-to-light ratio of the stellar population (hereafter given in solar \(V\)-band units). Both \(Y\) and the intrinsic axial ratio \(q\) are assumed to be constant (as a function of radius). The projected axial ratio \(q_p\) is determined by the inclination \(i\) according to \(q_p^2 = \cos^2 i + q^2 \sin^2 i\). The parameters \(Y\) and \(i\)
can be freely specified; all other parameters are determined by fitting to the available M32 surface photometry.

The highest spatial resolution surface photometry available for M32 is that presented by Lauer et al. (1992), based on pre-COSTAR HST/WFPC images. Their measurements extend to ~4" from the nucleus. At larger radii, ground-based data are available from Kent (1987) and Peletier (1993). Figure 1 shows the major axis surface brightness measurements from these sources. The solid curve shows the brightness profile predicted by our model, for the latter to be

$$\Psi_{\text{dark}} = GM_\odot(r^2 + e^2)^{-1/2},$$

which is the potential generated by a cluster with a Plummer model mass density (e.g., Binney & Tremaine 1987). For $$\epsilon = 0$$ one obtains the case of a dark nuclear point mass, i.e., a nuclear BH. We do not include the potential of a possible dark halo around M32. There are no (strong) observational constraints on the possible existence and characteristics of such a dark halo, and even if present, it will not affect the stellar kinematics near the nucleus of M32.

3. CONSTRUCTION OF TWO-INTEGRAL MODELS

The regular orbits in general axisymmetric potentials are characterized by three integrals of motion, the binding energy $$E = \Psi - \frac{1}{2}v^2$$, the component of the angular momentum around the symmetry axis $$L_z = Rv_\phi$$, and a nonclassical, or effective, third integral $$I_3$$ (Ollongren 1962; Richstone 1982; Binney & Tremaine 1987). In any given axisymmetric potential, there is an infinity of DFs $$f(E, L_z, I_3)$$ that generate a given axisymmetric mass density $$\rho(R, z)$$. Such models are difficult to construct, primarily because the third integral cannot generally be expressed explicitly in terms of the phase-space coordinates. However, for any mass density $$\rho(R, z)$$, there is exactly one DF that is even in $$L_z$$ and does not depend on $$I_3$$. This unique even "two-integral" DF, $$f(E, L_z)$$, provides a useful low-order approximation to any axisymmetric model and has the convenient property that many physical quantities, including the DF itself, can be calculated semianalytically. We study models of this type for M32 because they have successfully reproduced ground-based M32 data, and because they provide a useful guide for the interpretation of more general three-integral models, which are discussed in §§ 5 and 6.

To calculate the $$f(E, L_z)$$ DFs for our models, we have used a combination of the techniques described in Qian et al. (1995) and Dehnen (1995), even though either technique by itself could have been used to get the same result (in fact, yet another technique to address this problem is described in C97). Initially, four radial regimes are considered: $$m \ll b$$, $$m \approx b$$, $$m \approx c$$, and $$m \gg c$$, with $$m$$ and $$b$$ similar to $$c$$ as defined in equation (1). In these regimes the mass density is approximately $$\rho \propto m^2$$, $$\rho \propto (m/b)^{1 + (m/c)^2}$$, $$\rho \propto (m/b)^{2 + 2b}$$, and $$\rho \propto (m/c)^2$$, respectively. For each of these mass densities $$f(E, L_z)$$ can be calculated with the technique and software of Qian et al. (1995). The DFs for the four regimes are then smoothly patched together in energy, to yield an approximation to the full DF. This approximation is then used as the starting point for Lucy-Richardson iteration as described in Dehnen (1995). This yields the DF $$f(E, L_z)$$, reproducing the model mass density to ~0.3% RMS.

The total DF is the sum of the part $$f_0$$ that is even in $$L_z$$, and the part $$f_1$$ that is odd in $$L_z$$. In principle, $$f_0$$ is determined completely by the mean streaming velocities $$(v_\phi^0)(R, z)$$, but these are not determined well enough by the data to make an inversion practicable. Instead, therefore, $$f_0$$

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{fig1.png}
\caption{Data points are observations of the M32 major axis V-band surface brightness from Lauer et al. (1992), Peletier (1993), and Kent (1987). The R-band data from Kent and Peletier were transformed to the V band by assuming a constant V − R color. The differences between the data sets at large radii are due to uncertainties in the sky subtraction. Measurements are not plotted at radii where the PSF introduces large uncertainties (~0.1 for the Lauer et al. HST data, ~2" for the ground-based data). The solid curve is the brightness profile for the axisymmetric luminous density model used in the present paper. The dashed curve is the profile for the model used by van der Marel et al. (1994b) and Qian et al. (1995).}
\end{figure}
be calculated for any particular observational setup as in Qian et al. (1995). From the VPs, predictions can be calculated for the observable kinematical quantities.

4. PREDICTIONS OF TWO-INTEGRAL MODELS

4.1. Data-Model Comparison

Figure 2 shows the $HST$/FOS data presented in Paper I, obtained with the apertures 0.1-PAIR (0.068 square) and 0.25-PAIR (0.191 square). The figure also shows the highest available spatial resolution ground-based data, obtained by van der Marel et al. (1994a) with the William Herschel Telescope (WHT), and by Bender et al. (1996) with the Canada-France-Hawaii Telescope (CFHT). The spatial resolution for these ground-based observations is roughly 0.9 and 0.5 FWHM, respectively. The predictions of $f(E, L_z)$ models for the WHT and CFHT data have already been discussed in detail by previous authors, and we therefore focus here on a comparison of $f(E, L_z)$ models to the new $HST$ data.

The curves in Figure 2 are the predictions of edge-on $f(E, L_z)$ models with dark nuclear point masses (i.e., BHs). These models have an intrinsic axial ratio $q = 0.73$. The dark mass only influences the kinematical predictions in the central few arcseconds, and the mass-to-light ratio $\Upsilon$ was therefore chosen to fit the normalization of the kinematical data at larger radii. A good fit to the WHT data between $\sim S'$ and $\sim 12'$ is obtained with $\Upsilon = 2.51$. Predictions were calculated for each individual observation, taking into account the aperture position, aperture size, and point-spread function (PSF) for the $HST$ data as given in Paper I. Connecting curves in the figure were drawn to guide the eye. It proved sufficient to study only models with a very simple odd part $f_o$, namely those that produce a total DF in which at every $(E, |L_z|)$ a fraction $F$ of the stars has $L_z > 0$, and a fraction $(1 - F)$ has $L_z < 0$. For each $M_\bullet$, the fraction $F \leq 1$ of stars with $L_z > 0$ in the model was chosen to optimize the $\chi^2$ of the fit to the rotation curve. The displayed models with a BH fit the rotation curve well, more or less independent of $M_\bullet$; models with higher $M_\bullet$ require smaller $F$. The models differ primarily in their predictions for the velocity dispersions. The observed trend of increasing velocity dispersion toward the nucleus is successfully reproduced by models with a nuclear point mass of $M_\bullet \approx (3 \pm 1) \times 10^6 \, M_\odot$. The model without a BH predicts a roughly constant velocity dispersion with radius and is strongly ruled out. In fact, this model also fails to fit the observed rotation velocity gradient in the central arcsecond.

For a quantitative analysis of the best-fitting $M_\bullet$, we define a $\chi^2$ statistic that measures the quality of the model fit to the observed $HST$ velocity dispersions:

$$\chi^2 \equiv \sum_{i=1}^{N} \left( \frac{\sigma_{\text{model}} - \sigma_{\text{obs}}}{\Delta \sigma_{\text{obs}}} \right)^2, \quad \chi^2_{\text{RMS}} \equiv \left( \frac{\chi^2}{N} \right)^{1/2}. \quad (3)$$

Figure 3 shows the relative RMS residual $\chi^2_{\text{RMS}}$ as function of $M_\bullet$. The best fit to the $HST$ dispersions is obtained for $M_\bullet = (2.7 \pm 0.3) \times 10^6 \, M_\odot$. The quoted error is a formal 1$\sigma$ error based on the assumption of Gaussian statistics (probably an underestimate, as there is some hint for systematic errors in the data in addition to random errors).

An alternative way of estimating $M_\bullet$ is to model the average of the four data points within 0':1 from the center: $\sigma_{r<0.1'} = (126 \pm 10) \text{ km s}^{-1}$, where the 1$\sigma$ error is based on the scatter between the data points. Figure 4 shows the observed $\sigma_{r<0.1'}$ as a hatched region; a solid curve shows the predicted value as a function of $M_\bullet$. The predictions fall in the observed range for $M_\bullet = (3.4 \pm 0.9) \times 10^6 \, M_\odot$.

The inclination of M32 cannot be derived from the observed photometry and is therefore a free parameter in the modeling. However, the predictions of $f(E, L_z)$ models are rather insensitive to the assumed inclination (van der Marel et al. 1994b). This was verified by also calculating the predictions of inclined models with $i = 55^\circ$, which have an intrinsic axial ratio $q = 0.55$. A mass-to-light ratio $\Upsilon = 2.55$ was adopted, so that at large radii one obtains the same RMS projected velocity on the intermediate axis (between the major and minor axes) as for an edge-on model. On the major axis the $i = 55^\circ$ model then predicts a slightly higher...
RMS velocity than the edge-on model, resulting in a $\sim 10\%$ smaller best-fitting $M_\bullet$. Apart from this, the conclusions from the inclined models were found to be identical to those for the edge-on models (cf. Fig. 13 below).

4.2. Is the M32 Distribution Function of the Form $f(E, L_z)$?

The $f(E, L_z)$ models that fit the HST data can only be correct if they also fit the ground-based WHT and CFHT data. One may define similar $\chi^2$ quantities as for the HST data, to determine the best-fitting $M_\bullet$ for either of these data sets. Figures 3 and 4 show the relative RMS residual $\chi_{\text{RMS}}$ of the fit to all the dispersion measurements, and the average dispersion $\sigma_{\leq 0.1}$ of the dispersion measurements centered with $0.1'1$ from the nucleus. For the WHT data, $\chi_{\text{RMS}}$ is minimized for $M_\bullet = (1.46 \pm 0.03) \times 10^6 M_\odot$, whereas for the CFHT data it is minimized for $M_\bullet = (2.8 \pm 0.2) \times 10^6 M_\odot$. The central velocity dispersion measured with the WHT is best fit with $M_\bullet = (1.94 \pm 0.05) \times 10^6 M_\odot$, whereas for the CFHT data it is best fit with $M_\bullet = (2.1 \pm 0.2) \times 10^6 M_\odot$. These results are roughly consistent with those of previous authors. Van der Marel et al. (1994b), Qian et al. (1995), and Dehnen (1995) found $M_\bullet \approx (1.8 \pm 0.3) \times 10^6 M_\odot$ for the best fitting edge-on $f(E, L_z)$ model to the ground-based WHT data, while Bender, Kormendy & Dehnen obtained a best fit to their higher spatial resolution CFHT data with $M_\bullet \approx (3.0 \pm 0.5) \times 10^6 M_\odot$.

The latter value is somewhat higher than the one we find here, because it was chosen only to provide a good fit to the CFHT rotation curve; it does not fit the CFHT velocity dispersions very well.

These results indicate that, under the assumption of an $f(E, L_z)$ DF, the different observations cannot all be fit simultaneously with the same $M_\bullet$, even after accounting for the different observational setups. The lowest spatial resolution WHT data require a significantly lower $M_\bullet$ than the highest spatial resolution HST data. This implies that M32 has a DF that is not of the form $f = f(E, L_z)$.

5. CONSTRUCTION OF THREE-INTEGRAL MODELS

To construct more general three-integral models for M32, we extended Schwarzschild’s orbit superposition algorithm. Its basic structure is to calculate an orbit library that samples integral space in some complete and uniform way, to store the time-averaged intrinsic and projected properties of the orbits, and to search for the weighted superposition of orbits that best fits the observed kinematics, while reproducing the mass density $\rho(R, z)$ for self-consistency. Here we summarize the main steps, with emphasis on those aspects that are unique to the M32 application. Complete descriptions of the technique are given in R97 and C97.

We sample integral space with an $(R_c, \eta, w)$ grid. The quantity $R_c(E)$ is the radius of the circular orbit in the equatorial plane with energy $E$. Its angular momentum, $L_{\text{max}}(E)$, is the maximum angular momentum at the given energy. We define $\eta(E, L_z) \equiv L_z/L_{\text{max}}(E)$. For fixed $(R_c, \eta)$, the position of a star in the meridional $(R, z)$ plane is restricted to the region bounded by the zero-velocity curve (ZVC), defined by the equation $E = \Psi_{\text{eff}}$, where $\Psi_{\text{eff}} = \Psi - \frac{1}{2} L_z^2/R^2$ is the “effective gravitational potential” (Binney & Tremaine 1987). We parametrize the third integral at each $(R_c, \eta)$ using an angle $w$, which fixes the position at which an orbit touches the ZVC (cf. Fig. 5).
identical, except for a reversal of the three-dimensional velocity vector at
Because every orbit with touches the ZVC all
orbits at a given ( (see for examples of actual orbits). Orbits with
are in units of the radius of the circular orbit at the given energy. The
The figure shows the meridional plane at the energy for which
only orbits with 
[w] exceed the Hubble time only at radii The integra-
tion of a large set of three-dimensional mass densities (which all fit the
meridional plane mass constraints. It was found that
models with no kinematical constraints could at best fit the
masses with a fractional error of \( < 5 \times 10^{-2} \). Motivated by
this, fractional "errors" of this size were assigned to all the masses in the self-consistency constraints. As described in R97 and C97, we use the nonnegative least-squares (NNLS) routine of Lawson & Hanson (1974) to determine the combination of nonnegative orbital occupancies (which need not be unique) that minimizes the combined \( \chi^2 \).

The model predictions have a finite numerical accuracy, due to, e.g., gridding and discretization. Tests show that the numerical errors in the predicted kinematics of our method are \( \leq 2 \text{ km s}^{-1} \) for the rotation velocities and velocity dispersions, and \( \leq 0.01 \) in the Gauss-Hermite moments (C97).

- In theory, it is sufficient to fit only the meridional plane masses. Projected masses are then fit automatically. In practice, this is not exactly the case, because of discretization. Projected masses were therefore included as separate constraints.

- In principle, one would like to include the observational surface brightness errors in the analysis. Unfortunately, this requires the exploration of a large set of three-dimensional mass densities (which all fit the surface photometry to within the errors), which is prohibitively time consuming. However, the observational errors in the surface brightness are small enough that they are not believed to influence the conclusions of our paper.

The quantity \( R_c \) was sampled using 20 logarithmically spaced values between \( R_{c, \text{min}} = 6.12 \times 10^{-4} \) arcsec, and \( R_{c, \text{max}} = 7.55 \times 10^{3} \) arcsec. This range of radii contains all but a fraction \( 10^{-4} \) of the stellar mass of M32. The quantity \( \eta \) was sampled using an "open" grid (in the same sense that numerical quadrature formulae can be open or closed, e.g., Press et al. 1992) of \( N_\eta = 14 \) values, spaced linearly between \(-1 \) and \( 1 \), i.e., \( \eta_i = -1 + (2i - 1)N_\eta \) for \( i = 1, \ldots, N_\eta \). The quantity \( w \) was sampled using an "open" grid of 7 values, spaced linearly between \( 0 \) and \( w_\text{th} \). Here \( w_\text{th} \) is the angle \( w \) for the "thin tube" orbit at the given \( (E, L_\eta) \) (see Fig. 5). The special values \( \eta = 0 \) and \( \eta = \pm 1 \) (meridional plane and circular orbits) and the special values \( w = 0 \) and \( w = w_\text{th} \) (equatorial plane and thin tube orbits) are presumed to be represented by their closest neighbors on the grid, but are not included explicitly.

An orbit was integrated for each \( (R_c, \eta, w) \) combination, starting with \( v_x = v_y = 0 \) from the ZVC. The integration time was 200 times the period of the circular orbit at the same energy. This is sufficient to properly sample the phase-space trajectory for the large majority of orbits, although it exceeds the Hubble time only at radii \( \gtrsim 100^\circ \). The integra-
tions yield the "orbital phase-space density" for each orbit, as described in C97. These were binned onto \((r, \theta)\) grid in the meridional plane; \((r', \theta')\) grid on the projected \((x, y)\) plane of the sky; and several Cartesian \((x, y, v)\) cubes, with \( v \) the line-of-sight velocity. The first two grids were chosen logarithmic in \( r, r' \), with identical bins to those used for \( R_c \), and linear in \( \theta, \theta' \) in \([0, \pi/2]\). The \((x, y, v)\) cubes were centered on \((0, 0, 0)\), with \( 211 \times 211 \) square spatial cells, and 91 velocity bins of \( 15 \text{ km s}^{-1} \). Spatial cell sizes were adopted of \( 0.025, 0.08, \) and \( 0.5 \), respectively. The \((x, y, v)\) cubes were used to calculate, for each orbit, the predicted line-of-sight velocity histograms for all positions and setups for which kinematical data are available, taking into account the observational PSFs and aperture positions, orientations and sizes, as described in C97.

Construction of a model consists of finding a weighted superposition of the orbits in the library that reproduces two sets of constraints:

1. Self-consistency constraints. The model should reproduce the masses predicted by the luminous density \( \rho(R, z) \) (§ 2), for each cell of the meridional \((r, \theta)\) grid, for each cell of the projected \((r', \theta')\) grid, and for each aperture on the sky for which there are data.\(^6\)

2. Kinematical constraints. The model should reproduce the observed kinematics of the galaxy, including VP shapes. We chose to express this as a set of linear constraints on the Gauss-Hermite moments \( \hat{h}_l \) of the VPs (R97).

\[ \hat{h}_l = \sum_{i, j} c_{ij} \hat{h}_{ij}, \quad c_{ij} = \int_0^{2\pi} \int_0^\pi R_c^{l+1} P_l(\cos \theta) e^{-i2\pi r/R_c} \rho(R_c, \theta) \sin \theta d\theta dR_c. \]

In practice this is unfeasible, because the projected mass constraints are not independent from each other (aperture positions for different data sets partly overlap) and from the meridional plane mass constraints. It was found that models with no kinematical constraints could at best fit the masses with a fractional error of \( < 5 \times 10^{-2} \). Motivated by this, fractional "errors" of this size were assigned to all the masses in the self-consistency constraints.\(^7\) As described in R97 and C97, we use the nonnegative least-squares (NNLS) routine of Lawson & Hanson (1974) to determine the combination of nonnegative orbital occupancies (which need not be unique) that minimizes the combined \( \chi^2 \).

\[ \sum_{i, j} c_{ij} \hat{h}_{ij} = \sum_{i, j} d_{ij} \hat{h}_{ij}, \quad d_{ij} = \int_0^{2\pi} \int_0^\pi R_c^{l+1} P_l(\cos \theta) e^{-i2\pi r/R_c} \rho(R_c, \theta) \sin \theta d\theta dR_c. \]

\[ \hat{h}_l = \sum_{i, j} c_{ij} \hat{h}_{ij}, \quad c_{ij} = \int_0^{2\pi} \int_0^\pi R_c^{l+1} P_l(\cos \theta) e^{-i2\pi r/R_c} \rho(R_c, \theta) \sin \theta d\theta dR_c. \]

\[ \sum_{i, j} c_{ij} \hat{h}_{ij} = \sum_{i, j} d_{ij} \hat{h}_{ij}, \quad d_{ij} = \int_0^{2\pi} \int_0^\pi R_c^{l+1} P_l(\cos \theta) e^{-i2\pi r/R_c} \rho(R_c, \theta) \sin \theta d\theta dR_c. \]
6. PREDICTIONS OF THREE-INTEGRAL MODELS

6.1. Implementation

We have studied three-integral models with dark central point masses, i.e., $\epsilon = 0$ in equation (2). There are then three free model parameters: the inclination $i$, the mass-to-light ratio $\Upsilon$, and the BH mass $M_\bullet$. The parameter space must be explored through separate sets of orbit libraries. As for the two-integral models, we have studied only two, widely spaced, inclinations: $i = 90^\circ$ and $i = 55^\circ$. For each inclination we have sampled the physically interesting range of $(\Upsilon, M_\bullet)$ combinations. For each $(\Upsilon, M_\bullet)$ combination we determined the orbital weights (and hence the dynamical constraints only; the actual NNLS fits were done to both the kinematical and the self-consistency constraints, but contour plots of the total $\chi^2$ look similar. The overall minimum $\chi^2$ values are obtained for: $\Upsilon = 2.1$ and $M_\bullet = 3.4 \times 10^6 M_\odot$ for $i = 90^\circ$; and $\Upsilon = 2.0$ and $M_\bullet = 3.2 \times 10^6 M_\odot$ for $i = 55^\circ$.

Figure 7 compares the kinematical predictions of the best-fitting edge-on and $i = 55^\circ$ models to the data. Two problems with the data must be taken into account when assessing the quality of the fit. First, the HST velocity dispersions show a scatter between some neighboring points that is much larger than the formal errors, most likely due to some unknown systematic effect. The models cannot be expected to reproduce this. Second, the CFHT rotation velocities at radii $\lesssim 0.5' \times 0.5'$ exceed the WHT measurements by an amount that cannot be attributed to differences in spatial resolution but that must be due to some unknown systematic error in either of the two data sets. The WHT data have the smallest error bars and therefore receive most weight in the NNLS fit. As a result, the models tend to underpredict the CFHT rotation curve.

These systematic problems with the data preclude the use of $\chi^2$ as a meaningful statistic to assess which models provide an acceptable fit: if the observations themselves are not mutually consistent, then clearly no model can be statistically consistent with all of them. Although the use of any statistical test is suspect in the presence of systematic errors, one may still assign confidence regions on the model parameters by using the relative likelihood statistic $\Delta \chi^2 = \chi^2 - \chi^2_{\text{min}}$. This statistic merely measures which parameter combinations provide an equally good (or bad) fit as the one(s) that yield the minimum $\chi^2$. If we assume that the observational errors are normally distributed (which, as mentioned, is likely to be an oversimplification), then $\Delta \chi^2$ follows a $\chi^2$ probability distribution with the number of degrees of freedom equal to the number of model parameters (Press et al. 1992).

The best-fitting edge-on model in Figure 7 has $\chi^2 = 690.0$ while the best-fitting $i = 55^\circ$ model has $\chi^2 = 602.5$, both for $N = 416$ degrees of freedom. The fact that $\chi^2 > N$, even for these optimum fits, is due primarily to the systematic errors in the data. To the eye, the models appear to fit the data as well as could be hoped for. The $\chi^2$ values do indicate that the $i = 55^\circ$ model provides a significantly better fit than the edge-on model, implying that M32 is not seen edge on. However, the results presented here do not allow us to derive the actual inclination of M32. That would require a detailed study of the entire range of possible inclinations, which would be more computer intensive. The important conclusion in the present context is that the topology of the $\chi^2$ contours in Figure 6 is virtually identical for both inclinations: the allowed range for $M_\bullet$ is therefore uninfluenced by our ignorance of the true inclination of M32.

The $\Delta \chi^2$ statistic was used to assign confidence values to the contours in Figure 6. At the 68.3% confidence level (1 $\sigma$ for a Gaussian probability distribution), the allowed $M_\bullet$ fall in the range $(3.2-3.5) \times 10^6 M_\odot$ for $i = 90^\circ$, and in the range $(3.1-3.4) \times 10^6 M_\odot$ for $i = 55^\circ$. At the 99.73% confidence level (3 $\sigma$ for a Gaussian probability distribution), they fall in the ranges $(2.5-3.7) \times 10^6 M_\odot$ and $(2.3-3.9) \times 10^6 M_\odot$, respectively. In reality, small numerical errors in the models might have distorted the $\chi^2$ contours. We address this issue in Appendix A. Any numerical errors are small enough that they have no influence on our conclusion that models without a dark mass are firmly ruled out. However, the possibility of small numerical errors does increase the confidence bands on $M_\bullet$. Based on the analysis in Appendix A, we conclude that $M_\bullet = (3.4 \pm 0.7) \times 10^6 M_\odot$ at 68.3% confidence, and $M_\bullet = (3.4 \pm 1.6) \times 10^6 M_\odot$ at 99.73% confidence. These estimates take into account both the observational errors in the data and possible numerical errors in the models and are valid for both inclinations that were studied.

Eight selected models in Figure 6 are labeled as A–H and are listed in Table 1. Figure 8 compares the predictions of these models to the observed rotation velocities and velocity dispersions. Models C and G are the overall best fits for the two inclinations. Models B and F and D and H are

---

* Only one orbit library needs to be calculated for models with the same $M_\bullet/\Upsilon$. The potentials of such models are identical except for a normalization factor, and the orbits are therefore identical except for a velocity scaling. Each $(\Upsilon, M_\bullet)$ combination does require a separate NNLS fit to the constraints.

* A more robust way to incorporate the effects of random errors in the assignment of confidence bands would be to use “bootstrapping,” in which one directly calculates the statistical distribution of models parameters by finding the best-fit parameter combinations for different “realizations” of the data set. Unfortunately, this is computationally infeasible in the present context: even the analysis of the single (available) data set for M32 already takes weeks of CPU time on a high-end workstation.
Fig. 6.—Contour plots of the $\chi^2$ that measures the quality of the fit to the combined HST, CHFT, and WHT data, for orbit-superposition models with $i = 90^\circ$ and $i = 55^\circ$. The model parameters along the abscissa and ordinate are the BH mass $M_\bullet$ and the $(V$ band) mass-to-light ratio $Y$, respectively. The dots indicate models that were calculated, the contours were obtained through spline interpolation (the first three contours define the formal 68.3%, 95.4%, and [heavy contours] 99.73% confidence regions; subsequent contours are characterized by a factor of 2 increase in $\Delta\chi^2$). The bottom right corner in each panel is a region for which no models were calculated. The labeled positions define models that are discussed in detail in the text, in subsequent figures and in Table 1. Models C and G provide the overall best fits.

(approximately) the best-fitting models for $M_\bullet = 1.9$ and $5.4 \times 10^6 M_\odot$, respectively. The latter models are marginally ruled out at the greater than 99% confidence level (see the above discussion), although to the eye they do appear to reproduce the main features of the data. They differ from the overall best-fitting models primarily in their predictions for the HST velocity dispersions. The differences in the predictions for the ground-based data are smaller (and invisible to the eye in Fig. 8), but nonetheless more statistically significant because of the smaller error bars for these data. Models A and E, the best fits without a central dark mass, are indisputably ruled out. The main problem for these models is to fit the central peak in the velocity dispersion. They come rather close to fitting the WHT observations and predict a central dispersion of $\sim 84$ km s$^{-1}$. However, the models without a dark mass fail to reproduce the higher central dispersion of $\sim 91 \pm 2$ km s$^{-1}$ measured with the CFHT (although still only marginally) and do not even come close to reproducing the HST dispersions, which exceed $100$ km s$^{-1}$ in the central 0$^\circ$1.

6.3. Smooth Solutions

The top row in Figure 9 illustrates the orbital occupancies in integral space for the best-fitting edge-on model (model C). Only a small fraction of the orbits in the library is used to fit the constraints, while the remainder receives zero weight. This yields an equilibrium solution of the collisionless Boltzmann equation but is not physically plausible.

Smoothness of the solutions in integral space can be enforced by adding linear regularization constraints to the problem (Zhao 1996; C97). We have explored this only in an ad hoc way, merely to be able to assess the effect of smoothness constraints on the resulting fit to the data. A

### TABLE 1

**Properties of Selected Models for M32**

| Model | $i$ (deg) | $Y$ | $\langle v^2_\phi \rangle^{1/2}$ (km s$^{-1}$) | $\sigma_\phi$ (km s$^{-1}$) | $\sigma_\theta$ (km s$^{-1}$) | $\sigma_r$ (km s$^{-1}$) |
|-------|-----------|-----|---------------------------------|-----------------|-----------------|-----------------|
| A......| 90        | 0.0 | 3.0                            | 80              | 64              | 81              |
| B......| 90        | 1.9 | 2.4                            | 91              | 74              | 83              |
| C......| 90        | 3.4 | 2.1                            | 98              | 79              | 86              |
| D......| 90        | 5.4 | 1.7                            | 109             | 87              | 94              |
| E......| 55        | 0.0 | 3.1                            | 91              | 65              | 80              |
| F......| 55        | 1.9 | 2.4                            | 98              | 70              | 77              |
| G......| 55        | 3.2 | 2.0                            | 112             | 83              | 82              |
| H......| 55        | 5.4 | 1.7                            | 114             | 86              | 80              |

**Note.**—Col. (1) lists the model label as indicated in Fig. 6. Col. (2) lists the inclination. Col. (3) lists the dark point mass $M_\bullet$. Col. (4) lists the average mass-to-light ratio $Y$ of the stellar population (in solar $V$-band units). These parameters determine the gravitational potential of each model. For each potential, orbit superposition was used to determine the dynamical structure that best fits the combined HST and ground-based data for M32. Cols. (5)–(12) summarize the velocity ellipsoid shapes of the resulting models for two radial ranges. Cols. (5)–(8) list the mass-weighted average values of $\langle v^2_\phi \rangle^{1/2}$, $\sigma_\phi$, $\sigma_\theta$, and $\sigma_r$ in the radial range $0^\circ07 \leq r \leq 0^\circ9$. Cols. (9)–(12) list the same quantities for the radial range $0^\circ9 \leq r \leq 12^\circ$. The azimuthal dispersion is defined as $\sigma_\phi \equiv (\langle v^2_\phi \rangle - \langle v^\theta \rangle^2)^{1/2}$. As discussed in the text, models A and E are ruled out by the M32 data, while models C and G provide the best fits.
model is defined as a set of masses $m(R_c, \eta, w)$ in integral space. For each point that is not on the boundary of the $(R_c, \eta, w)$ grid, we measure the smoothness of the model (Press, et al. 1992; eq. [18.5.10]) through the second-order divided differences (in each of the three variables, assuming for simplicity that the distances between adjacent grid points are equal in all directions) of the function $m(R_c, \eta, w) / m_0(R_c)$. The function $m_0(R_c)$ is a rough approximation to the energy dependence of the model, obtained, e.g., by studying the spherical isotropic limit of the given mass density. The
Fig. 8.—Predicted rotation velocities and velocity dispersions for the orbit-superposition models A–D (top panel) and E–H (bottom panel) defined in Fig. 6 and Table 1. Models C and G are the overall best fits, models A and E are the best fits without a central dark mass, and models B and F and D and H are (approximately) the best-fitting models for $M_r = 1.9$ and $5.4 \times 10^6 M_\odot$, respectively. The data are as in Fig. 7. The models without a BH manage to fit the WHT data reasonably well, but are firmly ruled out by the HST data.

Regularization constraints are then that the divided differences should equal $0 \pm \Delta$, where the "error" $\Delta$ determines the amount of smoothing. Models with $\Delta \to \infty$ have no smoothing, while models with $\Delta \to 0$ force $m(R_c, \eta, w)/m_0(R_c)$ to be a linear function on the $(R_c, \eta, w)$ grid.

The second and third rows in Figure 9 show the integral space for model C with the addition of either a modest ($\Delta = 5$) or a large ($\Delta = 0.2$) amount of regularization in the NNLS fit, respectively. As the bottom panels show, the price paid for the increased smoothness is a somewhat poorer fit to the data. However, the fits are still quite good. This demonstrates that the good fits to the data shown in Figure 7 are not primarily the result of the use of implausible distributions in integral space. These distributions result from the numerical properties of the problem, but there also exist smooth solutions which provide similar fits.

6.4. Dynamical Structure

Figure 10 shows the components of the second velocity moment tensor as function of radius, for the edge-on models A–D and the $i = 55^\circ$ models E–H. Table 1 lists for two radial ranges the average RMS velocities in km s$^{-1}$ in each of the spherical coordinate directions and also includes the azimuthal dispersion $\sigma_\phi = [\langle v_\phi^2 \rangle - \langle v_\phi \rangle^2]^{1/2}$.

By contrast to the models with a BH, models A and E without a BH invoke a large amount of radial motion in the central arcsecond to produce a peak in the observed veloc-
Fig. 9.—The top row shows the \((\eta, w)\) integral space (defined as in Fig. 5) for a selected set of energies, for the best-fitting edge-on orbit-superposition model C defined in Fig. 6 and Table 1. Each square in each panel represents an orbit. The (logarithmic) gray scale shows the fraction of the mass at the given energy that was assigned to each orbit by the NNLS fit. Smoother solutions are obtained by adding regularization constraints to the NNLS fit. The second and third rows show the integral space for the same model with a modest and a large amount of regularization, respectively. Most of the mass resides at \(\eta > 0\) (i.e., \(L_z > 0\)), which is obviously required to fit the observed rotation of M32. The bottom panels show the fits of the models to the observed rotation velocities and velocity dispersions. The model without regularization provides the best fit, but even for the smoothest model the fits are still quite good.
The models B–D and F–H, which represent the best fits for different potentials, all have a similar dynamical structure: the second velocity moment tensor is dominated by azimuthal motion. This is most pronounced for the intrinsically flatter, model G rotates faster than the edge-on model C. The predictions for oblate isotropic rotator models ($\sigma_r = \sigma_\theta = \sigma_z$) are shown for comparison. The edge-on model C rotates faster than an oblate isotropic rotator model for radii $r \geq 0.5$; the inclined model G rotates slower than the oblate isotropic rotator model. However, overall the rotation rate is not too dissimilar from that for an oblate isotropic rotator model.

The combined results that the dynamical structure of M32 is not too far from that of an $f(E, L_z)$ model and that and intrinsically flatter, model G rotates faster than the edge-on model C. The predictions for oblate isotropic rotator models ($\sigma_r = \sigma_\theta = \sigma_z$) are shown for comparison. The edge-on model C rotates faster than an oblate isotropic rotator model for radii $r \geq 0.5$; the inclined model G rotates slower than the oblate isotropic rotator model. However, overall the rotation rate is not too dissimilar from that for an oblate isotropic rotator model.

The combined results that the dynamical structure of M32 is not too far from that of an $f(E, L_z)$ model and that
its rotation rate is not too far from that of an oblate isotropic rotator model imply that the velocity dispersions (but not the RMS velocities) in M32 are not far from isotropic. This is confirmed by the velocity dispersion values listed in Table 1. This explains why the value of $M_*$ determined here is similar to that determined from previous studies: these assumed either isotropic velocity dispersions (e.g., Tonry 1987), a range of possible anisotropies that bracket isotropic models (e.g., Richstone et al. 1990), or an $f(E, L_z)$ DF (e.g., van der Marel 1994b). It is an interesting question what physical process could have caused M32 to have the particular dynamical structure inferred here, in particular, because the DFs of (giant) elliptical galaxies are in general not well described by the $f(E, L_z)$ form (e.g., van der Marel 1991; Bender, Saglia, & Gerhard 1994). The answer to this question is currently unknown.

The second velocity moments $\langle v_x^2 \rangle$ and $\langle v_y^2 \rangle$ can be combined with the mixed moment $\langle v_x v_y \rangle$ to determine the tilt of the velocity ellipsoid in the meridional plane. For all the models A–D and E–H we found $\langle v_x v_y \rangle$ to be small, and the velocity ellipsoids are more closely aligned with spherical coordinate axes than with cylindrical coordinate axes (but they are not perfectly aligned with either). This is not uncommon in three-integral models for axisymmetric systems (e.g., Dejonghe & de Zeeuw 1988; Dehnen & Gerhard 1993; de Zeeuw, Evans, & Schwarzschild 1996).

We have not attempted to derive confidence bands on the dynamical structure of M32. This is a much more difficult problem than the derivation of confidence bands on the model parameters ($M_*, \gamma$) and is beyond the scope of the present paper.

7. MODELS WITH AN EXTENDED DARK NUCLEAR OBJECT

The results in the previous section demonstrate that M32 must have a massive dark object in its nucleus. To obtain a limit on the size of this dark object, we have studied models in which it has a finite size, $\sim \epsilon$ (see eq. [2]). Searching the parameter space of three-integral models with different $\epsilon$ is extremely computer intensive. We have therefore restricted ourselves to $f(E, L_z)$ models with extended dark objects. This is not likely to bias our conclusions, because the best-fitting three-integral models found in § 6 are similar to two-integral models.

Figure 12 shows the predictions of edge-on $f(E, L_z)$ models with extended nuclear dark objects, for two representative values of $M_*$. As in the models of § 4.1, the fraction $F$ of stars with $L_z > 0$ in each model was chosen to best fit the rotation curve. By adjusting $F$, adequate fits to the rotation velocities can be obtained for most relevant models. Hence, only the velocity dispersions are shown in the figure. The predicted dispersions are typically constant or decreasing toward the center within the scale radius $r$. The $HST$ data show a much higher velocity dispersion in the center than farther out. Hence, the extension of any possible dark nuclear cluster cannot be large; Figure 12 suggests $\epsilon \lesssim 0.1$. Figure 13 shows a contour plot of the quantity $\gamma^2_\epsilon$ (eq. [3]), measuring the quality of the model fit to the $HST$ dispersion measurements, both for the edge-on case and for $i = 55^\circ$. The best fit is obtained for $\epsilon = 0$ (the point mass case discussed previously). The formal 99.73% confidence level (assuming Gaussian formal errors) rules out all models with $\epsilon \gtrsim 0.06$, independent of the inclination. The models with the largest $\epsilon$ must have a total mass of at least $M_* \approx 4 \times 10^6 M_\odot$.

At the distance of M32, 1" = 3.39 pc. Hence, the upper limit on the scale radius corresponds to $\epsilon = 0.20$ pc. Combined with a total mass in the cluster of $M_* \approx 4 \times 10^6 M_\odot$, this implies a central mass density of at least $\rho_0 = 1.1 \times 10^8 M_\odot$ pc$^{-3}$. The half-mass radius of a Plummer model is $r_h = 1.30e$. Hence, there must be $\sim 2 \times 10^6 M_\odot$ inside $r \lesssim 0.078$. The total $V$-band luminosity inside this radius is $1 \times 10^5 L_\odot$, implying a luminous mass of $2.5 \times 10^5 M_\odot$. Hence, the ratio of the total mass to luminosity inside this radius must be $\gtrsim 22.5$.

The observed kinematics constrain only the amount of mass in the system, not whether this mass is luminous or dark. One can therefore fit the data equally well with models in which the average mass-to-light ratio $\gamma$ of the stellar population increases toward the nucleus and in which there is no dark mass. We have not explicitly constructed such models, but it is clear from the preceding discussion that in such models $\gamma$ must rise from $\sim 2$ in the main body of the galaxy to $\gtrsim 20$ at $r \lesssim 0.1$. Such a drastic variation in mass-to-light ratio would imply a strong change in the stellar population, accompanied by broadband color gradients. The size of these gradients depends on the actual stellar population mix, which is unknown.

---

10 This quantity does not depend sensitively on the assumed density cusp slope at very small radii. Gebhardt et al. (1996) infer a somewhat steeper slope for M32 than used here, but their model has only 10% more luminous mass inside $r \lesssim 0.078$ than ours.
However, one may use the properties of main-sequence stars as a guideline. For these, a change in $\gamma$ from 2 to 20 implies color changes $\Delta(U-B) \approx 0.9$, $\Delta(B-V) \approx 0.6$, and $\Delta(V-I) \approx 0.8$ (using the tables of stellar properties in Allen 1973). Such variations between $0.1$ and $1^\circ$ should have been obvious in photometric observations. However, neither subarcsecond resolution ground-based imaging (Lugger et al. 1992) nor prerefurbishment $HST$ imaging (Crane et al. 1993) have revealed any significant color gradients in the central arcsecond of M32. Postrefurbishment $HST$ observations (T. R. Lauer et al. 1997, private communication) also do not show strong color gradients. Thus, the nuclear mass concentration in M32 cannot be merely due to a change in the mix of ordinary stars in the nuclear region.

The absence of observed color gradients does not exclude the possibility of nuclear concentrations of brown dwarfs, white dwarfs, neutron stars, or stellar-mass BHs in M32. However, at high densities, such clusters of dark objects are not stable over a Hubble time. This was discussed by Goodman & Lee (1989), and their arguments were updated and extended in van der Marel et al. (1997a). The latter paper shows that the new $HST$ limit on the density of dark material in M32 rules out all but the most implausible clusters, leaving a single massive BH as the most likely interpretation of the data.

The kinematical predictions of our models depend on the assumed Plummer form of the extended dark object. If it is a cluster of collapsed objects, this distributed dark mass may itself be cusped (e.g., Gerhard 1994). However, the limit on $\epsilon$ results from the fact that the dispersion of stars with mass density (1) in a Plummer potential does not have a (strong) central peak. This property is common to many alternative types of models, such as those of King (e.g., Binney & Tremaine 1987), Hernquist (1990), and Jaffe (1983). These all produce constant or decreasing dispersions inside their scale radius, as does the Plummer model. Hence, the upper bound on $\epsilon$ derived here is likely to be generic to most plausible density profiles for the extended dark object. In addition, King, Jaffe, and Hernquist models are more centrally condensed than Plummer models and would therefore require dark clusters of even higher central densities to fit the data.

8. CONCLUSIONS AND DISCUSSION

8.1. Summary of Results

The main bottlenecks in proving the presence of nuclear BHs in quiescent galaxies from stellar kinematical data have long been (1) the restricted spatial resolution of ground-based data and (2) lack of sufficiently general dynamical models to rule out constant mass-to-light ratio models beyond doubt. The $HST$ now provides spectra of superior spatial resolution. To fully exploit the potential of these new data, it is imperative to improve the modeling techniques that have been used in the past decade. The situation is considerably more complicated than for gas disks in (active) galaxies, where the assumption of simple circular orbits is often adequate. Interpretation of stellar kinematical data for flattened elliptical galaxies ideally requires axisymmetric (or even better, triaxial) dynamical models with completely general three-integral distribution functions. Such models have not previously been constructed for any stellar kinematical BH candidate galaxy. We therefore developed a technique for the construction of such axisymmetric models and used it to interpret our $HST$ data for M32.

To guide the construction and interpretation of the three-integral models, we first compared the new $HST$ data with the predictions of $f(E, L)$ models, which have been used extensively to interpret ground-based M32 data. Such models have the advantage that the DF can be calculated semianalytically, but have the disadvantage of having a special dynamical structure, with $\sigma_0$, $\sigma_0$ everywhere. There is no a priori reason why any galaxy should have this property. However, the fact that $f(E, L)$ models fit the observed VP shapes inferred from ground-based data to within $\sim 2\%$ (in terms of deviations from a Gaussian) suggested that the M32 DF might in fact be close to the form $f(E, L)$. We find here that $f(E, L)$ models for M32 can also fit the new $HST$ data and that this requires the presence of a nuclear dark
mass, as was the case for the ground-based data. However, the best-fitting dark mass of \( M_\bullet = (2.5-4.5) \times 10^6 \, M_\odot \) is larger than the \( M_\bullet = (2-3) \times 10^6 \, M_\odot \) that best fits the \( \sim 0.5 \) spatial resolution data from the CFHT and is even more different from the \( M_\bullet = (1.5-2) \times 10^6 \, M_\odot \) that best fits the \( \sim 0.9 \) spatial resolution data from the WHT. Thus, under the assumption of an \( f(E, L_z) \) DF, the different data sets cannot be fit with the same \( M_\bullet \). This indicates that the M32 DF is not of the form \( f(E, L_z) \), although it might be close to it.

To obtain a model-independent estimate of the best-fitting \( M_\bullet \), and to firmly rule out models without any dark mass, it is necessary to study more general three-integral models. We have made such models for M32, both with and without central BHs, and for various possible values of the average mass-to-light ratio \( Y \) of the stellar population. The models were constructed to fit all available kinematical HST, CFHT, and WHT data, and the acceptability of each model was assessed through the \( \chi^2 \) of its fit to the data. The models demonstrate explicitly for the first time that there is no axisymmetric constant mass-to-light ratio model that can fit the kinematical data without invoking the presence of a nuclear dark mass, independent of the dynamical structure of M32. A nuclear dark point mass of \( M_\bullet = (3.4 \pm 1.6) \times 10^6 \, M_\odot \) is required (with 1 \( \sigma \) and 3 \( \sigma \) error bars of 0.7 \( \times 10^6 \, M_\odot \) and 1.6 \( \times 10^6 \, M_\odot \), respectively, which includes the possible effect of small numerical errors in the models). This mass is similar to that quoted by most previous papers, but the confidence on the detection of a nuclear dark mass in M32 is now much higher. Constant mass-to-light ratio models still come very close to fitting the ground-based data, and only the new HST data make the case for a nuclear dark mass clear-cut.

The inclination of M32 cannot be inferred from the available surface photometry and is therefore a free parameter in the modeling. Ideally, one would like to construct dynamical models for all possible inclinations (which would be very computer intensive) and determine the inclination that best fits the kinematical data. Here we have taken the more modest approach of constructing models for only two representative inclinations: \( i = 90^\circ \) (edge on) and \( i = 55^\circ \). The intrinsic axial ratios for these inclinations are \( q = 0.73 \) and \( q = 0.55 \), respectively. The three-integral \( i = 55^\circ \) models provide a better fit than the edge-on models, which suggests that M32 is not seen edge on. However, the allowed range for \( M_\bullet \) does not depend sensitively on the assumed inclination: models with no central dark mass are firmly ruled out for both inclinations. So even though a more detailed study of the full inclination range for M32 would improve our knowledge of the true inclination and intrinsic axial ratio of M32, it would probably not change significantly the constraints on the central dark mass.

The best-fitting three-integral models are similar to \( f(E, L_z) \) models in that they have an excess of azimuthal motion. This is why they have been so successful in fitting ground-based data, including available VP shape parameters, and it confirms that they provide a useful low-order approximation to the dynamical structure of M32. However, \( f(E, L_z) \) models do have \( \sigma_r \equiv \sigma_\theta \), which does not reproduce the inequality between \( \sigma_r \) and \( \sigma_\theta \), nor the modest tilt of the velocity ellipsoid indicated by the small \( \langle v_r v_\theta \rangle \) term, seen in the best-fitting three-integral models. This is why \( f(E, L_z) \) models cannot successfully explain all observed features of the kinematical data.

To constrain the size of the dark object in M32 we have constructed \( f(E, L_z) \) models with an extended dark nuclear object. These show that the HST data put an upper limit of \( 0.08 \leq 0.26 \) pc on the half-mass radius of the nuclear dark object, implying a central density exceeding \( 1 \times 10^9 \, M_\odot \) \( \text{pc}^{-3} \). This limit on the density of dark material in M32 essentially rules out nuclear clusters of planets, brown dwarfs, white dwarfs, neutron stars, or smaller mass BHs (van der Marel 1997a). The absence of color gradients in the central arcsecond of M32 implies that the nuclear mass concentration can also not be attributed to a stellar population gradient. A single massive nuclear BH therefore provides the most plausible interpretation of the data.

### 8.2. Dynamical Stability

Axisymmetric dynamical models with a nuclear BH provide an excellent fit to all available kinematical data for M32. However, to be physically meaningful, the models must also be dynamically stable. In van der Marel, Sigurdsson, & Hernquist (1997c) we presented \( N \)-body simulations of the \( f(E, L_z) \) models for M32. The models were found to be completely stable, both for \( i = 90^\circ \) and for \( i = 55^\circ \). This shows that dynamical stability is not a problem for the models and that the inclination of M32 cannot be meaningfully constrained through stability arguments. We have not evolved \( N \)-body models for the best-fitting three-integral models, but we expect these models to be stable as well, given their similarity to \( f(E, L_z) \) models.

### 8.3. Dynamical Relaxation

The two-body relaxation time in M32 can be estimated as in, e.g., Binney & Tremaine (1987; eq. [8-71]). Using the relevant quantities for our best-fitting dynamical model, we find for solar mass stars in the central cusp (\( r \leq 0.5 \) that \( t_{\text{relax}} \approx 3 \times 10^9(0.001)^{-0.065} \) yr. The timescale for “resonant relaxation” (Rauch & Tremaine 1996) is of the same order. The central cusp must therefore be evolving secularly over a Hubble time. However, the diffusion of stars in phase space is slow enough that one may assume the evolution to be through a sequence of quasi-equilibrium models. This justifies our approach of modeling M32 as a collisionless equilibrium system. Studies of the secular evolution of the M32 cusp will be interesting but will not change the need for a nuclear dark object. In fact, the process of dynamical relaxation supports the presence of a dark object: without a dark object the relaxation would proceed at a much more rapid rate that is difficult to reconcile with observations (Lauer et al. 1992).

### 8.4. Triaxiality

One remaining uncertainty in our dynamical modeling is the possibility of triaxiality. After the step from spherical models to axisymmetric models, triaxial models are the obvious next step. However, there are several reasons to believe that for M32 this additional step will be less important. First, M32 is known not to be spherical, but there is no reason why it cannot be axisymmetric. There is no significant isophote twisting in M32 and no minor axis rotation. This does not mean that M32 cannot be triaxial (we might be observing it from one of the principal planes), but it also does not mean that M32 needs to be triaxial. Second, spherical constant mass-to-light ratio models (without a nuclear dark mass) for ground-based M32 data failed to fit by only a few km s\(^{-1}\), and it was quite conceivable that axisym-
mery could fix this (which it did, see Fig. 15 below). However, axisymmetric constant mass-to-light ratio models for the new HST data fail to fit the nuclear velocity dispersion by greater than 50 km s$^{-1}$, and this cannot likely be fixed through triaxiality. Third, theoretical arguments suggest that strongly triaxial models with density cusps as steep as in M32 may not be stable, owing to the fact that regular box orbits are replaced by boxlets and irregular orbits that may not be able to sustain a triaxial shape (Binney & Gerhard 1985; Merritt & Fridman 1996; Merritt & Valluri 1996; see also the review by de Zeeuw 1996).

Rapidly rotating low-luminosity elliptical galaxies like M32 always have steep power-law cusps (Faber et al. 1997) and may therefore be axisymmetric as a class (de Zeeuw & Carollo 1996). This is consistent with statistical studies of their intrinsic shapes (e.g., Merritt & Tremblay 1996). So, apart from the fact that triaxiality is unlikely to remove the need for a central dark object in M32, it may even be so that M32 cannot be significantly triaxial.

8.5. Adiabatic Black Hole Growth

The growth of a black hole into a stellar system is adiabatic if it occurs over a timescale that is “long” (see Sigurdsson, Quinlan, & Hernquist 1995 for a quantitative discussion) compared with the typical orbital period of the stars. For the case of M32, the black hole formation can be considered adiabatic if it took at least 10$^6$ yr. Young (1980) studied the adiabatic growth of BHs in spherical isothermal models with central density $\rho_0$ and core radius $r_0$. The BH growth leaves the mass density at large radii unchanged, but induces a central cusp $\rho \propto r^{-1.5}$ for $r \to 0$. The form of the density profile at intermediate radii is determined by the dimensionless parameter $M_\bullet = M_\bullet / [(4/3)\pi r_0^3 \rho_0^2]$, which measures the ratio of the BH mass to the initial core mass. Lauer et al. (1992) showed that the shape of the M32 brightness profile measured with HST can be well fit with $M_\bullet = 0.33 \pm 0.11$. The radial and density normalization implied by the data are then $r_0 = 3.0$ pc and $\rho_0 = (4.2 \times 10^4) M_\odot pc^{-3}$. This photometric model therefore implies that $M_\bullet / r_0 = (1.6 \pm 0.5) \times 10^6 M_\odot$. Although this result depends somewhat on the assumed isothermality of the initial distribution (Quinlan, Hernquist, & Sigurdsson 1995), it is quite remarkable that our best-fitting dynamical models have exactly $M_\bullet / r_0 = 1.6 \times 10^6 M_\odot$, for both inclinations that we studied. The M32 data are therefore fully consistent with the presence of a BH that grew adiabatically into a preexisting core. This is similar to the situation for M87 (cf. Young et al. 1978; Harms et al. 1994).

Lee & Goodman (1989) extended Young’s calculations to the case of rotating models. For the value of $M_\bullet$ implied by the photometry, their models predict a profile of $\langle v^2 \rangle / (v^2 r)_{1/2}$ that is approximately flat with radius (with amplitude fixed by the axial ratio of the system). However, this result depends very sensitively on the assumed rotation law of the initial model. The radial variations in $\langle v^2 \rangle / (v^2 r)_{1/2}$ seen in our best-fitting models (Fig. 11) are probably equally consistent with the adiabatic growth hypothesis.

8.6. Tidal Disruption of Stars

A star of mass $m_*$ and radius $r_*$ on a circular orbit of radius $r$ will be tidally disrupted if $r \leq r_\ast \equiv (2M_\bullet/m_\ast)^{1/3} r_\ast$ (e.g., Binney & Pettit 1988). Thus, disruption of a solar type star by the BH in M32 will occur inside $r_\ast = 4.2 \times 10^{-6}$ pc = $1.2 \times 10^{-6}$ arcsec. A disruption event will be highly luminous, but is not predicted to occur more often than once every 10$^4$ yr (Rees 1988). The minimum pericenter distance for a star with given $(E, L_\star)$ in a Kepler potential is $r_{p,min} = R_\ast [1 - (1 - \eta)^{1/2}]$, where as before, $R_\ast$ is the radius of the circular orbit at the given energy and $\eta = L_\star / L_{max}$. The kinematical data for M32 only meaningfully constrain the DF for energies with $R_\ast(E) \geq 0.1$. For $R_\ast(E) = 0.1$, only stars with $|\eta| < 5 \times 10^{-3}$ have $r_{p,min} < r_\ast$. The data do not constrain variations in the DF over such a small range in $|\eta|$, and our dynamical models therefore cannot address the existence and properties of the so-called loss cone (Frank & Rees 1976; Lightman & Shapiro 1977). For the $\eta$-grid that we have employed, all solar-type stars on orbits with $R_\ast(E) > 0.1$ have $r_{p,min} > 2 \times 10^7 r_\ast$. Even giants with $r_\ast = 10^7 r_\ast$ have $r_{p,min} > r_\ast$. This justifies our neglect of tidal disruption in the orbit calculations.

8.7. Accretion onto the Black Hole

An interesting question is why BHs in quiescent galaxies are not more luminous (e.g., Kormendy & Richstone 1995). For M32, the total X-ray luminosity is $L_X \approx 10^{38}$ ergs$^{-1}$ (Eskridge, White, & Davis 1996), the far-infrared luminosity is $L_{100} \approx 3 \times 10^{46}$ ergs$^{-1}$ (Knapp et al. 1989), and for the 6 cm radio emission $vL_v < 3 \times 10^{33}$ ergs$^{-1}$ (Roberts et al. 1991). Part or all of the observed X-ray emission may be due to low-mass X-ray binaries, so the total luminosity due to accretion onto the BH in M32 is $L_{acc} < 10^{38}$ ergs$^{-1}$. By contrast, the Eddington luminosity of the BH is $L_{Edd} = 4.3 \times 10^{44}$ ergs$^{-1}$. For a canonical mass-loss rate of 1.5 $M_\odot (10^{11} L_\odot)^{-1}$ yr$^{-1}$ (Faber & Gallagher 1976), the stars that are bound to the BH in M32 shed $1 \times 10^{-4}$ $M_\odot$ yr$^{-1}$ of gas as a result of normal stellar evolution. If a fraction $f$ of this gas is steadily accreted with efficiency $\epsilon$, it produces a luminosity $L_{acc} = f\epsilon L_{Edd} = (6.7 \times 10^{42})$ ergs$^{-1}$. Thus, either the accretion fraction $f$ or the accretion efficiency $\epsilon$ must be very small in M32. Thin disk accretion with $\epsilon \approx 0.1$ requires $f < 1.5 \times 10^{-4}$, which is possible (the accretion fraction is difficult to predict theoretically, because it depends on the hydrodynamics of the stellar winds that shed the gas), but may be implausibly low. Instead, it appears more likely that $\epsilon$ is small, since there is a family of “advection dominated” accretion solutions that naturally predict such low efficiencies. Models of this type successfully explain the “microactivity” of the BH (Sgr A*) in our own Galaxy (Narayan, Yi, & Mahadevan 1995). In a typical accretion model of this type (Narayan & Yi 1995, their Fig. 11), $f \leq 0.16$ suffices to explain the upper bound on $L_{acc}$ for M32.

8.8. Forthcoming Observations

Future observations of M32 will include spectra with the new long-slit HST spectrograph STIS. These will provide significantly better sky coverage than our FOS data, but the spatial resolution will be similar. The high-resolution HST data can be complemented with that from fully two-dimensional ground-based spectrographs, such as OASIS on the CFHT and SAURON on the WHT. These combined data will yield improved constraints on the BH mass, on the orbital structure and inclination of M32, and on possible deviations from axisymmetry.

We dedicate this paper, and its companions R97 and C97, to the memory of Martin Schwarzschild, who pioneered the modeling technique employed here. Martin’s sense of
purpose, his exceptional clarity of thinking, his transparent personal integrity, and, most of all, his genuine warm interest and support remain a great source of inspiration to us. Support for this work was provided by NASA through grant GO-05847.01-94A and through a Hubble Fellowship HF-1065.01-94A awarded to R. P. v. d. M., both from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. N. C. acknowledges financial support from the Swiss government (Etat du Valais) and NUFFIC and the hospitality of Steward Observatory, the MPA Garching, and Geneva Observatory. T. d. Z. is grateful for the generous hospitality of the Institute for Advanced Study. Both he and N. C. received financial support from the Leids Kerkhoven Bosscha Fonds.

APPENDIX A

χ² TOPOLOGY FOR ORBIT-SUPERPOSITION MODELS

In this Appendix we discuss the topology of the χ²(Υ, M_*) contours for the edge-on orbit-superposition models. The top panels of Figure 14 show the χ² contours when only (subsets of) the ground-based WHT data are included in the fit. These panels can be compared with Figure 14d, which shows the contours for the case in which all WHT, CFHT, and HST data are included.

Figure 14a shows the χ² contours when only the major axis V and σ WHT measurements are fit. Binney & Mamon (1982) showed that a large range of gravitational potentials can fit any given observed velocity dispersion profile. The valley seen in the χ² contours is a consequence of this: it outlines a one-parameter family of models that can fit the data with different velocity dispersion anisotropy. For a nonrotating spherical system, only models that require negative second velocity moments are ruled out. For a rotating system like M32, the observed rotation rate sets additional limits on the allowed radial anisotropy. For the case of the major axis V and σ WHT measurements, a no-BH model is just marginally acceptable at 99.73% confidence, see Figure 14a. For the lower-spatial resolution major axis V and σ measurements of Dressler & Richstone (1988) such a model is entirely acceptable. Figure 15 compares the predictions of the best-fit axisymmetric

![Contour plots](image)

Fig. 14.—Contour plots of χ²(M_*, Υ) for edge-on orbit superposition models constructed to fit: (a) the major axis WHT V and σ measurements; (b) all major axis WHT measurements, including VP shapes; (c) all WHT measurements, including various position angles; (d) all WHT, CFHT and HST data (same as Fig. 6). Contours are defined as in Fig. 6. Heavy contours show the formal 99.73% confidence regions. Panel (e) shows how the contours in panel (d) are modified if random errors are added to the predictions for each data point, to simulate numerical errors in the models. Panel (f) shows for 100 simulations as in panel (e) the position of the χ² minimum (solid symbols), and the lowest and highest M_*) that fall within a 99.73% confidence contour (open symbols). These simulations show that numerical errors cannot be responsible for the fact that models with either M_*) < 1.8 × 10^6 M_0 or M_*) > 5.0 × 10^6 M_0 fail to fit the data at this confidence level.
orbit-superposition model without a BH to their data. Richstone et al. (1990) concluded that these data could not be fit by any spherical model without a BH. This is because spherical models allow less rotation and therefore failed to fit the observed rotation velocities. This underscores the importance of making axisymmetric models for flattened galaxies like M32.

VP shape measurements provide independent constraints on the velocity dispersion anisotropy. Figure 14b shows the \( \chi^2 \) contours for edge-on orbit-superposition models when not only the WHT major axis \( V \) and \( \sigma \) measurements are fit, but also the major axis VP shape measurements. With the inclusion of the VP shapes, models without a BH are ruled out. Figure 14c shows the \( \chi^2 \) contours when also the WHT measurements along other position angles are included, which contracts the allowed \( M_\bullet \) range to \( (1.5 \pm 1.1) \times 10^6 M_\odot \) at the formal 99.73% confidence level. The WHT data by themselves therefore rule out axisymmetric models without a BH. However, the models without a BH still come very close to fitting the data, and, e.g., fail to fit the central velocity dispersion by only 1–2 km s\(^{-1}\) (cf. Fig. 8). So one cannot make a particularly strong claim for a BH on the basis of the WHT data alone, because it is conceivable that the fit could be improved with, e.g., only a minor amount of triaxiality. The same holds for the CFHT data, but the new HST data do make the case for a BH in M32 clear-cut.

The contours for the case in which all the available WHT, CFHT and HST kinematical data are included in the fit (Fig. 14d) show one global \( \chi^2 \) minimum and a second local minimum. The presence of a global minimum does not necessarily imply that the combined data constrain a single best-fit potential. It might be that there is a small range of potentials that all fit equally well, but that such a range of constant \( \chi^2 \) would not be evident due to the finite numerical accuracy of our technique. In Figure 14e we show explicitly how the topology of the \( \chi^2 \) contours might have been influenced by the possibility of small numerical errors in our models. It was obtained from Figure 14d by recalculating the \( \chi^2 \) contours after adding a random error \( \Delta V, \sigma \in [-2, 2] \) km s\(^{-1}\) and \( \Delta h, e \in [-0.01, 0.01] \) (cf. § 5) to the prediction for each data point, for each \( (\gamma, M_\bullet) \) combination. The results show that numerical errors can indeed influence the \( \chi^2 \) contours near the \( \chi^2 \) minimum. Thus, the second minimum in Figure 14d might be the result of numerical inaccuracies in our technique. However, the numerical errors are small enough that they have only a negligible effect on the overall \( \chi^2 \) topology. In particular, models without a dark mass remain firmly ruled out.

To assess the possible effect of numerical errors on the confidence bands for \( M_\bullet \), we constructed 100 figures like Figure 14e using different random realizations. For each we determined the position of the \( \chi^2 \) minimum and the minimum and maximum \( M_\bullet \) for which there is an \( \chi \) such that the model with \( (M_\bullet, \gamma) \) falls within the 99.73% confidence region. The results are plotted in Figure 14f. All allowed \( M_\bullet \) values fall in the range \( M_\bullet = (3.4 \pm 0.7) \times 10^6 M_\odot \) at 68.3% confidence. Similar experiments show that \( M_\bullet = (3.4 \pm 0.7) \times 10^6 M_\odot \) at 68.3% confidence. Experiments for \( i = 55^\circ \) produced similar results, and mass ranges that were either the same or slightly smaller. Thus, we conclude that the 1 \( \sigma \) and 3 \( \sigma \) errors on the estimated \( M_\bullet = 3.4 \times 10^6 M_\odot \), are 0.7 and 1.6 \( \times 10^6 M_\odot \), respectively.

REFERENCES

Allen, C. W. 1973, Astrophysical Quantities (London: Athlone)
Bender, R., Kormendy, J., & Dehnen, W. 1996, ApJ, 464, L123
Bender, R., Saglia, R. P., & Gerhard, O. E. 1994, MNRAS, 269, 785
Binney, J. J., & Gerhard O. E. 1995, MNRAS, 215, 409
Binney, J., & Mamon, G. A. 1982, MNRAS, 200, 361
Binney, J., & Petit, J.-M. 1988, in Dynamics of Dense Stellar Systems, ed. D. R. Merritt (Cambridge: Cambridge Univ. Press), 43
Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton: Princeton Univ. Press)
Crane, P., et al. 1993, ApJ, 106, 1371
Cretton, N., de Zeeuw, P. T., van der Marel, R. P., & Rix H.-W. 1997, ApJ, in preparation (C97)
de Brujinne, J. H. J., van der Marel, R. P., & de Zeeuw P. T. 1996, MNRAS, 282, 909
Dehnen, W. 1995, MNRAS, 274, 919
Dehnen, W., & Gerhard, O. E. 1993, MNRAS, 261, 311
Dejonghe, H., & de Zeeuw, P. T. 1988, ApJ, 333, 90
de Zeeuw, P. T. 1996, in Gravitational Dynamics, ed. O. Lahav, E. Terlevich, & R. J. Terlevich (Cambridge: Cambridge Univ. Press), 1
1997, in ASP Conf. Ser. 116, The Nature of Elliptical Galaxies, Proceedings of the Second Stromlo Symposium, ed. M. Arnaboldi, G. da Costa, & P. Saha (San Francisco: ASP), 44
de Zeeuw, P. T., & Carollo, C. M. 1996, in IAU Symp. 171, New Light on Galaxy Evolution, ed. R. Bender, & R. L. Davies (Dordrecht: Kluwer), 47
de Zeeuw, P. T., Evans, N. W., & Schwarzschild, M. 1996, MNRAS, 280, 903
Dressler, A. 1984, ApJ, 286, 97
Dressler, A., & Richstone, D. O. 1988, ApJ, 324, 701
Eskridge, P. B., White, R. E., & Davis, D. S. 1996, ApJ, 463, L59
Faber, S. M., & Gallagher, J. S. 1976, ApJ, 204, 365
Faber, S. M., et al. 1997, AJ, in press
Frank, J., & Rees, M. J. 1976, MNRAS, 176, 633
Gebhardt, K., et al. 1996, AJ, 112, 105
Gerhard, O. E. 1994, in The Nuclei of Normal Galaxies: Lessons from the Galactic Center, R. ed. Genzel, & A. I. Harris (Dordrecht: Kluwer), 267
Goodman, J., & Lee, H. M. 1989, ApJ, 343, L84
Harms, R. J. et al. 1994, ApJ, 435, L35
Hernquist, L. 1990, ApJ, 356, 359
Jaffe, W. 1983, MNRAS, 202, 985
Kent, S. M. 1987, AJ, 94, 306
Knapp, G. R., Gualachakurta, P., Kim, D.-W., & Jura, M. 1989, ApJS, 70, 329
Kormendy, J., & Richstone D. 1995, ARA&A, 33, 581
Lauer, T. R., et al. 1992, AJ, 104, 552
Lawson, C. L., & Hanson, R. J. 1974, Solving Least Squares Problems (Englewood Cliffs: Prentice-Hall)
Lee, M. H., & Goodman, J. 1989, ApJ, 343, 594
Lightman, A. P., & Shapiro, S. L. 1977, ApJ, 211, 244
Lugger, P. M., Cohn, H. N., Cederbloom, S. E., Lauer, T. R., & McClure, R. D. 1992, AJ, 104, 83
Lynden-Bell, D. 1996, MNRAS, 279, 389
Merritt, D. R., & Fridman, T. 1996, ApJ, 460, 136
Merritt, D. R., & Tremblay, B. 1994, AJ, 108, 514
———. 1996, AJ, 111, 2243
Merritt, D. R., & Valluri, M. 1996, ApJ, 471, 82
Narayan, R., & Yi, I. 1995, ApJ, 452, 710
Narayan, R., Yi, I., & Mahadevan, R. 1995, Nature, 374, 623
Ollongren, A. 1962, Bull. Astr. Inst. Netherlands, 16, 241
Peletier, R. F. 1993, A&A, 271, 51
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical Recipes (Cambridge: Cambridge Univ. Press)
Qian, E. E., de Zeeuw, P. T., van der Marel, R. P., & Hunter, C. 1995, MNRAS, 274, 602
Quinlan, G. D., Hernquist, L., & Sigurdsson, S. 1995, ApJ, 440, 554
Rauch, K. P., & Tremaine, S. 1996, New Astron., 1, 149
Rees, M. 1988, Nature, 333, 523
———. 1996, in Gravitational Dynamics, eds. Lahav, O., Terlevich, E., & Terlevich, R. J. (Cambridge: Cambridge Univ. Press)
Richstone, D. O. 1982, ApJ, 252, 496
Richstone, D. O., Bower, G., & Dressler, A. 1990, ApJ, 353, 118
Richstone, D. O., & Tremaine, S. 1988, ApJ, 327, 82
Rix, H.-W., de Zeeuw, P. T., Cretton, N., van der Marel, R. P., & Carollo, C. M. C. 1997, ApJ, 488, 702 (R97)
Roberts, M. S., Hogg, D. E., Bregman, J. N., Forman, W. R., & Jones, C. 1991, ApJS, 75, 751
Schwarzschild, M. 1979, ApJ, 232, 236
Sigurdsson, S., Hernquist, L., & Quinlan, G. D. 1995, ApJ, 446, 75
Tonry, J. L. 1984, ApJ, 283, L27
———. 1987, ApJ, 322, 632
van den Bosch, F. C. 1997, MNRAS, 287, 543
van der Marel, R. P. 1991, MNRAS, 253, 710
van der Marel, R. P., de Zeeuw, P. T., & Rix, H.-W. 1997b, ApJ, 488, 119 (Paper I)
van der Marel, R. P., de Zeeuw, P. T., Rix H.-W., & Quinlan, G. D. 1997a, Nature, 385, 610
van der Marel, R. P., Evans, N. W., Rix, H.-W., White, S. D. M., & de Zeeuw, P. T. 1994b, MNRAS 271, 99
van der Marel, R. P., Rix, H.-W., Carter, D., Franx, M., White, S. D. M., & de Zeeuw, P. T. 1994a, MNRAS, 268, 521
van der Marel, R. P., Sigurdsson, S., & Hernquist, L. 1997c, ApJ, 487, 153
Young, P. 1980, ApJ, 242, 1232
Young, P., Westphal, J. A., Kristian, J., Wilson, C. P., & Landauer, F. P. 1978, ApJ, 221, 721
Zhao, H. S. 1996, MNRAS, 283, 149