THE PLANETARY NEBULA SYSTEM OF M33

ROBIN CIARDULLO, 1 PATRICK R. DURRELL, MARY BETH LAYCHAK, 1 KIMBERLY A. HERRMANN, AND KENNETH MOODY

Department of Astronomy and Astrophysics, Pennsylvania State University, 525 Davey Laboratory, University Park, PA 16802;

rbc@astro.psu.edu, pdurrell@astro.psu.edu, mary@cfht.hawaii.edu, herrmann@astro.psu.edu, moody@astro.psu.edu

GEORGE H. JACOBY

WIYN Observatory, P.O. Box 26732, Tucson, AZ 85726; jacob@wiyn.org

AND

JOHN J. FELDMEIER 2

Department of Astronomy, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106-1712;

johnf@bottom.astr.cwru.edu

Received 2003 November 19; accepted 2004 June 17

ABSTRACT

We report the results of a photometric and spectroscopic survey for planetary nebulae (PNs) over the entire body of the Local Group spiral galaxy M33. We use our sample of 152 PNs to show that the bright end of the galaxy’s [O III] 5007 planetary nebula luminosity function (PNLF) has the same sharp cutoff seen in other galaxies. The apparent magnitude of this cutoff, along with the IRAS DIRBE foreground extinction estimate of $E(B-V) = 0.041$, implies a distance modulus for the galaxy of $(m-M)_0 = 24.86 \pm 0.07$ (0.94 $\pm$ 0.05 Mpc). Although this value is $\sim$15% larger than the galaxy’s Cepheid distance, the discrepancy likely arises from differing assumptions about the system’s internal extinction. Our photometry, which extends more than 3 mag down the PNLF, also reveals that the faint end of M33’s PNLF is nonmonotonic, with an inflection point $\sim$2 mag below the PNLF’s bright limit. We argue that this feature is due to the galaxy’s large population of high core mass planetaries and that its amplitude may eventually be a useful diagnostic for studies of stellar populations. Fiber-coupled spectroscopy of 140 of the PN candidates confirms that M33’s PN population rotates along with the old disk, with a small asymmetric drift of $\sim$10 km s$^{-1}$. Remarkably, the population’s line-of-sight velocity dispersion varies little over $\sim$4 optical disk scale lengths, with $\sigma_{\text{rad}} \sim 20$ km s$^{-1}$. We show that this is due to a combination of factors, including a decline in the radial component of the velocity ellipsoid at small galactocentric radii and a gradient in the ratio of the vertical to radial velocity dispersion. We use our data to derive the dynamical scale length of M33’s disk and the disk’s mass-to-light ratio. Our most likely solution suggests that the surface mass density of M33’s disk decreases exponentially, but with a scale length that is $\sim$2.3 times larger than that of the system’s IR luminosity. The large scale length also implies that the disk’s $V$-band mass-to-light ratio changes from $M/L_V \sim 0.3$ in the galaxy’s inner regions to $M/L_V \sim 2.0$ at $\sim$9 kpc. Models in which the dark matter is distributed in the plane of the galaxy are excluded by our data.

Subject headings: dark matter — distance scale — galaxies: distances and redshifts — galaxies: individual (M33) — galaxies: kinematics and dynamics — planetary nebulae: general

Online material: machine-readable table

1. INTRODUCTION

For many years, H I and optical rotation curves have been used to establish the presence of dark matter halos around spiral galaxies (e.g., Faber & Gallagher 1979; Ashman 1992; Combes 2002). However, the mass profiles of these halos are still in doubt: from the rotation curves alone, it is impossible to decouple the gravitational contribution of the galaxy’s visible disk mass from that of its dark halo (see, e.g., Rubin 1997 and references therein). Most investigations begin with the assumption that a disk’s mass-to-light ratio is constant with radius (e.g., Kent 1986; Palunas & Williams 2000; Sofue et al. 2003), but evidence in support of this hypothesis is limited. In fact, independent estimates of disk mass exist for only a handful of spiral galaxies (e.g., Bottema 1993; Bottema et al. 1987; van der Kruit et al. 2001; Gerssen et al. 1997, 2000), and in these systems, the data are restricted to the galaxies’ inner $\sim$2 scale lengths. Thus, while the information to date is consistent with the constant mass-to-light ratio hypothesis, the applicability of these studies to the dark matter-dominated outer regions of spiral galaxies is limited.

The best way to independently determine the mass within a spiral’s disk is through the motions of its old disk stars. However, the traditional approach to doing this via absorption-line spectroscopy is extremely challenging. Because the surface brightness of spiral galaxies declines exponentially with radius, observations at large radii are quite demanding and require intricate procedures for data reduction. This limits the effectiveness of the technique to the inner 1 or 2 disk scale lengths. Fortunately, there is an alternative. Instead of measuring the Doppler widths of integrated starlight, one can derive velocity

1 Visiting Astronomer, Kitt Peak National Observatory, National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA), Inc., under cooperative agreement with the National Science Foundation.

2 NSF Astronomy and Astrophysics Postdoctoral Fellow.
dispersions directly from the motions of individual stars. Planetary nebulae (PNs) are the ideal targets for this purpose. PNs are excellent tracers of the galactic light (Ciardullo et al. 1989), easily identifiable via their strong \([\text{O} \text{ iii}] \lambda 5007\) emission and distinctive \(I(5007)/I(\text{H} \alpha)\) ratio (Ciardullo et al. 2002), and measurable to a precision of a few kilometers per second via medium-dispersion spectrographs. Moreover, because the progenitors of PNs are intermediate- and low-mass stars, their velocity dispersion is representative of the old components of a galactic disk.

Here we describe a photometric and spectroscopic survey of the PN system of the Local Group spiral galaxy M33 (NGC 598). M33 is an excellent target for such an investigation. First, the galaxy is nearby, making large numbers of planetaries available for study. Second, since M33’s photometric structure is well known at both optical (de Vaucouleurs 1959; Guidoni et al. 1981; Kent 1987) and infrared (Regan & Vogel 1994; Hippelein et al. 2003) wavelengths, information is available on the distribution of both the young and old stars of the system. Finally, M33’s H \(_{\alpha}\) rotation curve is well measured out to \(\sim 20\) kpc (i.e., \(\gtrsim 10\) disk scale lengths). Consequently, detailed models exist for the orientation of the galactic disk and the system’s total potential (Rogstad et al. 1976; Corbelli & Schneider 1997; Corbelli & Salucci 2000).

We begin our study by describing our photometric survey of the galaxy and comparing our \([\text{O} \text{ iii}] \lambda 5007\) and \(\text{H} \alpha\) PN measurements to those of Magrini et al. (2000, 2001). We demonstrate that there is generally good agreement between the samples, although our \([\text{O} \text{ iii}] \lambda 5007\) magnitudes are systematically brighter than the Magrini et al. (2000) values by \(\sim 0.12\) mag. In \(\S 3\) we present M33’s \([\text{O} \text{ iii}] \lambda 5007\) planetary nebula luminosity function (PNLF) and show that at faint magnitudes, the function has the same nonmonotonic behavior as the PNLF of the Small Magellanic Cloud. We interpret this turnover as a population effect and speculate on how the shape of the PNLF can be used to probe the mix of stellar populations within a galaxy. In \(\S 4\) we use our data to derive a new PNLF-based distance to the galaxy and discuss a possible reason for the \(2\sigma\) difference between our measurement and that inferred from the galaxy’s Cepheid variables.
After analyzing the photometric data, we describe our spectroscopic survey of the galaxy’s PN system. In § 5 we detail our observing and reduction procedures and quantify the errors in our velocity measurements. In §§ 6 and 7 we use these data to examine the stellar kinematics of the galaxy. We present the system’s line-of-sight velocity dispersion, estimate the stellar asymmetric drift, and use the epicyclic approximation to model the galaxy’s velocity ellipsoid. We show that the line-of-sight velocity dispersion of M33’s planetaries varies very little with galactocentric radius and that this is due to a combination of factors, including a decline in the system’s radial velocity dispersion at small galactocentric distances and a gradient in the ratio of the system’s vertical to radial velocity dispersion. We also show that M33’s vertical velocity dispersion decreases with a scale length that is greater than that of the dispersion. We also show that M33’s vertical velocity dispersion varies very little with galactic radius and that this is due to a combination of factors, including a decline in the system’s radial velocity dispersion at small galactocentric distances and a gradient in the ratio of the system’s vertical to radial velocity dispersion. We also show that M33’s vertical velocity dispersion decreases with a scale length that is greater than that of the galaxy’s infrared light. In § 8 we discuss possible systematic errors that may affect our result, including the variation of the PN scale height with galactocentric distance. Finally, we use our results to estimate the surface mass density of M33’s disk and constrain the radial profile of the galaxy’s dark halo.

2. IDENTIFICATION OF PLANETARY NEBULAE

M33’s PN candidates were identified on Kitt Peak 4-m telescope MOSAIC CCD images taken as part of the NOAO Local Group Galaxies survey program (Massey et al. 2002). This data set consists of deep broadband (UBVRI) and narrowband ([O iii], [S ii], and Hα + [N ii]) images of 10 star-forming galaxies of the Local Group. Our program made use of the survey’s M33 frames in [O iii] 5007, Hα, B, and V. These data cover a 72′ × 36′ region of the galaxy in three overlapping 36′ × 36′ fields and extend out to a radius of ~10 kpc, or 4 times the galaxy’s disk scale length in blue light (de Vaucouleurs 1959). Our survey region is shown in Figure 1; the properties of the survey frames are summarized in Table 1.

To place the Massey et al. (2002) survey frames on a standard system, additional exposures in [O iii] 5007 (900 s) and Hα (600 s) were taken with the same telescope/instrument setup on the UT nights of 2001 December 22 and 24. The seeing during these observations was moderately poor, 1.7′′ in [O iii] and 1.2′′ in Hα, but the sky was photometric, and images of the Stone (1977) spectrophotometric standards BD +28° 4211 and Feige 110 were acquired immediately before and after the exposures. These supplemental images were bias-subtracted, flat-fielded, and remapped onto the tangent plane using the MSCRED routines within IRAF (Valdes 1998). The latter two steps were performed after the iterative subtraction of the telescope’s ghost pupil from the nights’ dome flats.

PN candidates were identified by blinking the on-band [O iii] 5007 and Hα survey images against their V-band counterparts. All spatially unresolved emission-line objects with little or no continuum flux were classified as PN candidates. Because medium bandwidth continuum images were not part of the Massey et al. (2002) survey, this condition effectively meant that all PN candidates had to be either invisible in V or have an instrumental $V - 5007$ color less than ~3 (the ratio of the filters’ bandpasses, corrected for transmission differences and the inclusion of [O iii] 4959 in the V filter). Since the seeing on the survey frames was generally very good, ≤1.0′′ (~4 pc) in [O iii] 5007 and 0.8′′ (~3 pc) in Hα, this simple classification scheme was generally unambiguous. To confirm the nature of the few sources with instrumentally colors uncomfortably close to the discrimination threshold, i.e., with $2 < V - 5007 < 3$, we also examined their appearance on the survey’s B frames. Those objects with B-band fluxes close to that expected from a B main-sequence star were classified as H ii regions and eliminated from the analysis. We note that under these criteria, the two objects closest to our B magnitude threshold are ~1.5 mag brighter in B than any other PN in our sample.

Once found, the PN candidates on each [O iii] frame were measured astrometrically using a grid of ~200 reference stars from the USNO-A2.0 catalog (Monet et al. 1998). The formal internal errors of our plate solutions were ~0.5″ in both right ascension and declination. However, since M33’s nuclear field overlapped the survey’s northern and southern fields by 50%, almost two-thirds of our PN candidates appeared in two (or three) frames and thus could be measured multiple times. The standard astrometric error between these measurements, ~0.15″, is slightly better than the internal errors of the individual plate solutions. This suggests that part of ~0.3 scatter in the solutions comes from the proper motions of the reference stars and that our relative astrometry is probably good to better than ~0.2″.

Relative [O iii] 5007 and Hα photometry of the PN candidates and several bright field stars was accomplished using the DAOPHOT point-spread function (PSF) fitting routines within IRAF (Stetson 1987, 1992; Stetson et al. 1990). Our photometry was then checked by using DAOPHOT’s SUBSTAR
option to subtract off a scaled-PSF representation of each PN from its position on the frame. An examination of the residuals revealed that, while the vast majority of the subtractions were excellent, a few objects in or near the galaxy’s spiral arms were slightly under- or oversubtracted, typically by ~0.1 mag. The cause of this error was an incorrect estimate of the galactic background, which (particularly in Hα) often included irregular H II regions, supernova remnants, and diffuse knots of emission. When this occurred, the PN magnitudes were manually adjusted until the residuals of the subtractions appeared reasonable.

After determining the raw PN instrumental magnitudes on each frame, the data were merged onto a common system by solving the least-squares condition required to match the magnitudes of stars in the regions of field overlap (Ciardullo et al. 1987). The instrumental magnitudes were then placed on the standard AB system by using the 2001 December data to compare large-aperture measurements of field stars to similar standards. We estimate the error associated with this step to be ~0.03 mag. Finally, to transform the measured AB magnitudes to monochromatic fluxes, we blueshifted the transmission curves of the [O III] 5007 and Hα filters to correct for their temperature at the telescope and applied the photometric procedures for emission-line objects described in Jacoby et al. (1989). In doing so, we assumed that M33’s PNs have a mean heliocentric radial velocity of ~180 km s⁻¹ and a dispersion of ~85 km s⁻¹ about that mean. Of course, the latter value is not strictly applicable to a rotating disk galaxy, but the error introduced by the peculiar velocity of an individual PN is never more than 0.01 mag.

Table 2 lists the positions and emission-line strengths of our 152 PN candidates. The [O III] 5007 magnitudes in the table are related to monochromatic flux by

$$m_{5007} = -2.5 \log F_{5007} - 13.74,$$

where the flux is given in ergs cm⁻² s⁻¹ (Jacoby 1989). Our typical photometric errors, as derived internally from DAOPHOT and externally from objects with multiple measurements, are given in Table 3.

We note that a catalog of M33 PNs has previously been published (Magrini et al. 2000, 2001). The two data sets agree reasonably well. Except for a few anomalous outliers, the positions of PNs common to the two surveys are virtually identical: in right ascension and declination, our values are systematically larger than the Magrini et al. (2001) measurements by 0.07 ± 0.06 and 0.04 ± 0.04, respectively, and the overall dispersion between individual coordinates is 0.7. Similarly, in Hα, there is little systematic difference between the photometry of the two surveys. On average, our Hα + [N II] log fluxes are 0.06 ± 0.05 mag fainter than those of Magrini et al. (2001), and, while the dispersion between the measurements is rather large (~0.25 mag for objects with log fluxes greater than ~15), this may partly be due to the differing contributions of [N II] within the filters’ bandpasses. In fact, the only significant difference between the two surveys is in the [O III] 5007 zero point. Although our [O III] 5007 magnitudes agree with those of Magrini et al. (2001) in the relative sense (the dispersion between the two measurements is ~0.1 mag for objects brighter than $m_{5007} \sim 23$), the zero point of our system is systematically brighter by 0.12 ± 0.05 mag. This offset and the dispersion between the measurements are shown in Figure 2.

Table 3 lists the photometric differences of the 152 PN candidates that were not recovered in our survey are listed in Table 4, along with the probable reason for the discrepancy. The implied fraction of

### Table 2: M33 Planetary Nebula Candidates

| ID | $\alpha$ (J2000.0) | $\delta$ (J2000.0) | $m_{5007}$ | $n_c$ | log $F_{H\alpha}$ | F([O III])/F(Hα) | Error | $n_c$ | $v_0$ | $\sigma_c$ | M01d | Notes |
|----|-------------------|-------------------|-----------|------|-----------------|-----------------|-------|------|------|----------|------|-------|
| 1  | 01 32 09.04       | 30 22 05.7        | 23.51     | 1    | -15.321         | 2.628           | 0.176 | 2    | -139.2| 3.6      | ...  | O     |
| 2  | 01 32 26.54       | 30 25 49.8        | 23.53     | 2    | -15.390         | 3.042           | 0.204 | 2    | -127.7| 1.9      | 1     | O     |
| 3  | 01 32 38.03       | 30 24 00.6        | 21.82     | 2    | -14.243         | 1.046           | 0.019 | 2    | -130.3| 1.6      | 2     | O     |
| 4  | 01 32 39.80       | 30 37 41.0        | 22.07     | 2    | -14.602         | 1.891           | 0.047 | 3    | -208.4| 1.7      | 3     |       |
| 5  | 01 32 40.08       | 30 35 15.3        | 23.38     | 2    | -15.343         | 3.125           | 0.154 | 2    | -160.4| 2.7      | ......|       |
| 6  | 01 32 42.72       | 30 12 25.5        | 22.63     | 1    | -14.895         | 2.218           | 0.085 | 2    | -106.7| 2.0      | 4     |       |
| 7  | 01 32 44.29       | 30 43 25.9        | 24.74     | 2    | -15.655         | 1.837           | 0.153 | 3    | -185.9| 3.0      | 6     |       |
| 8  | 01 32 49.16       | 30 51 09.1        | 21.45     | 2    | -14.526         | 2.826           | 0.042 | 3    | -192.2| 2.0      | 7     |       |
| 9  | 01 32 52.96       | 30 41 57.7        | 22.49     | 2    | -14.778         | 1.930           | 0.041 | 2    | -167.3| 3.0      | 9     |       |
| 10 | 01 32 53.72       | 30 32 25.8        | 23.73     | 2    | -14.720         | 0.541           | 0.023 | 3    | -136.3| 5.4      | 12    | O     |

Notes.—O: object also observed in 2003 January observing run; NT: object not targeted spectroscopically; H: possible halo PN. Units of right ascension are hours, minutes, and seconds, and units of declination are degrees, arcminutes, and arcseconds. Table 2 is published in its entirety in the electronic edition of the Astronomical Journal. A portion is shown here for guidance regarding its form and content.

a Number of photometric measurements of the PN.
b Hα + [N II] flux in units of ergs cm⁻² s⁻¹.
c Number of HYDRA setups in which PN was targeted (2002 October).
d ID number from Magrini et al. (2001).
contaminants (~22%) is consistent with the 28% contamination rate estimated by Magrini et al. (2003) from follow-up spectrophotometry of the Hα, [N II], and [S II] lines of 36 PN candidates. However, we note that there are some discrepancies for individual objects. Specifically, Magrini et al. (2001) object 8, which is classified by Magrini et al. (2003) as a PN, is excluded from our sample on the basis of its detectable continuum emission and [O III] λ5007–to–Hα + [N II] line ratio (see below). Conversely, Magrini et al. (2001) objects 41 and 42 are included in our study, even though they are classified by Magrini et al. (2003) as probable supernova remnants.

Are all the PN candidates listed in Table 2 genuine? At M33’s Cepheid distance of ~820 kpc (Freedman et al. 2001), our typical seeing of 1.0” in [O III] λ5007 and 0.8” in Hα + [N II] implies that all H ii regions or supernova remnants greater than ~3 pc in size are resolved. This is significantly better than the limit of ~7 pc associated with the poorer seeing exposures of the Magrini et al. (2000) survey, and as the catalog of Hodge et al. (1999) demonstrates, the density of H ii regions in M33 decreases rapidly over this range. Moreover, since the Massey et al. (2002) survey frames reach a limiting magnitude of V ~ 23.2, all of M33’s stars brighter than M_V ~ -1.3 are visible via their continuum. Consequently, unless an H ii region is excited by a lone main-sequence star with spectral type B3 or later, it will not make it into our sample. Finally, at the bright end of the [O III] λ5007 luminosity function, the excitation properties of our objects demonstrate that the fraction of contaminants is negligible. As pointed out by Ciardullo et al. (2002), PNs inhabit a distinctive cone in [O III] λ5007–Hα + [N II] emission line space. While faint PNs (those more than ~3 mag down the [O III] λ5007 luminosity function) can have [O III] λ5007–to–Hα + [N II] line ratios anywhere between 0.3 < R < 3, PNs in the top ~1 mag of the PNLF all have R ≥ 2. This contrasts with the great majority of compact H ii regions and supernova remnants, which typically have R < 1 (Shaver et al. 1983). As Figure 3 illustrates, all of our [O III]–bright PN candidates are high-excitation objects. Late-type B stars cannot create nebulae with such high [O III]–to–Hα ratios: the central stars of all our candidates must be extremely hot and optically faint. Taken together, the above arguments strongly suggest that the contamination fraction in our sample is extremely low.

3. THE PLANETARY NEBULA LUMINOSITY FUNCTION

The top panel of Figure 4 displays M33’s [O III] λ5007 PNLF. The sharp cutoff at the bright end of the distribution is obvious, as is the function’s slow decline at fainter magnitudes. Remarkably, both features are consistent throughout the galaxy. M33 has a sizable metallicity gradient (~0.11 ± 0.02 dex kpc^-1; Garnett et al. 1997), and the mean oxygen abundance of PNs in the galaxy’s inner disk (R < 15’) is likely to be ~0.5 dex larger than that of PNs outside this radius. Yet, as the middle and bottom panels of Figure 4 show, the [O III] λ5007 luminosity functions of the two samples are statistically identical. This constancy supports the conclusion of Ciardullo et al. (2002) that the PNLF cutoff is independent of metallicity for all but the most metal-poor systems.

A second interesting feature exhibited in Figure 4 is the decline in the PNLF at magnitudes fainter than m_{5007} ~ 22.5. This rollover is partially real and partially due to photometric incompleteness. To determine the effect of incompleteness, we began by excluding PNs located within 210” of the galaxy’s nucleus; object detections in this region are difficult because of the bright, irregular features of the galactic background. We then measured the local sky associated with each remaining object, identified the worst (most uncertain) background in the sample, and computed the signal-to-noise ratio each PN would have if it were projected on that background. Artificial star experiments have shown that PN identifications become incomplete when the signal-to-noise ratio of a detection drops below ~10 (Ciardullo et al. 1987; Hui et al. 1993). Our putative signal-to-noise ratio measurements therefore translate directly into a completeness limit. Our completeness limit, m_{5007} ~ 23.75, is more than a magnitude fainter than the point where the observed PNLF begins to decline.

We can confirm the reality of the dip at m_{5007} ~ 23 by comparing the radial distribution of “bright” PNs (m_{5007} ≤ 22.5) to a similar distribution for PNs with 23.0 ≤ m_{5007} ≤ 23.75. If incompleteness were important, then we might expect the latter sample to be missing PNs at small radii, where detections are more difficult because of the brighter surface brightness of the underlying galaxy. As Figure 5 and the Kolmogorov-Smirnov statistic show, this is not the case: outside the inner ~210”, the cumulative radial distributions for “bright” and “faint” PNs are identical. The agreement between the two distributions demonstrates that our estimate of the detection limit is reasonable and the decline in the luminosity function at m_{5007} ~ 23 is real.

The decrease in the luminosity function at fainter magnitudes, although unexpected, is not difficult to explain. In fact, there are at least two possible mechanisms that, in theory, can produce a nonmonotonic PNLF. The first is internal extinction in the host galaxy. The central extinction in M33 is A_V ~ 0.9 mag, and there is a significant amount of dust distributed throughout the galaxy’s disk (Regan & Vogel 1994; Hippelein et al. 2003). Because the scale height of PNs should
be larger than that of the interstellar medium (Mihalas & Binney 1981), one would expect the bright end of the PNLF to be dominated by objects foreground to the dust layer. The observation of bright PNs above the dust lanes of the edge-on spiral galaxies NGC 891 (Ciardullo et al. 1991) and NGC 4565 (Jacoby et al. 1996) supports this idea. However, at fainter luminosities, the contribution of extincted PNs may conceivably be dominated by objects foreground to the dust layer. The observation of bright PNs above the dust lanes of the edge-on spiral galaxies NGC 891 (Ciardullo et al. 1991) and NGC 4565 (Jacoby et al. 1996) supports this idea. However, at fainter luminosities, the contribution of extincted PNs may conceivably be enough and if the total amount of extinction is large, then both the number of PNs and the apparent magnitude of the PNLF cutoff may be affected by extinction. However, our analysis suggests that internal extinction, by itself, cannot turn the exponentially increasing PNLF of equation (2) into the peaked luminosity function displayed in Figure 4.

An alternative explanation for M33's PNLF comes from stellar evolution. According to the initial mass–final mass relation (Weidemann 2000), old stellar populations produce low-mass central stars with evolutionary timescales much longer than the timescale for nebular expansion (Vassiliadis & Wood 1994). To a first approximation, such objects can be modeled as nonevolving central stars surrounded by freely expanding gaseous envelopes. The $[O\text{ III}]\lambda 5007$ emission of these PNe will decrease with time, $t$, following the relation

$$F_{5007} \propto N_0 N_e V \propto N_e \propto \frac{1}{R^2} \propto \frac{1}{t^2},$$

where $N_0$ is the number density of $O^{++}$ ions, $N_e$ is the electron density, $V$ is the volume of the nebula, and $R$ is the nebular radius. For an ensemble of objects, the number of PNs with $[O\text{ III}]\lambda 5007$ magnitudes between $M$ and $M + dM$ will then be

$$N(M_{5007}) \propto \frac{dM}{d ln M} \propto e^{0.307M}.$$

This is the exponential law first proposed by Henize & Westerlund (1963) and used by Ciardullo et al. (1989) in their PNLF calculations. However, it may not be applicable for younger stellar populations. In these systems, the typical PN will have a higher mass core and a much shorter timescale for stellar evolution. If this timescale is comparable to, or shorter than, the timescale for envelope expansion, then a PN's $[O\text{ III}]$...
Evolution will be driven by the luminosity evolution of its central star, rather than the expansion of its nebula. Since post-AGB stars spend a relatively long time at high luminosity (as they cross the H-R diagram at the end of shell burning), a short time at intermediate luminosity (when nuclear reactions stop), and a long time at low luminosity (as their cooling rate slows), the PNLF of high-mass objects will be double peaked and similar to that calculated for post-AGB stars by Vassiliadis & Wood (1994).

Of course, in order to model real systems, one must include central star evolution, nebular expansion, and gasdynamics in the analysis, as well as the mix of stellar populations within a galaxy (see Marigo et al. 2004). Nevertheless, these simple arguments do explain the different PNLFs found in Local Group galaxies. As Figure 6 demonstrates, the low-mass stars of M31’s bulge have a PNLF that is well represented at the faint end by an exponential (Ciardullo et al. 2002). Conversely, M33 and the SMC are actively forming stars and their PNLFs are consistent with the bimodal distribution of Vassiliadis & Wood (1994). Given the distinctive nature of the two limiting cases, it is possible that the strength of the “dip” in the luminosity function may someday be a useful probe of stellar populations that are otherwise difficult to observe.

### 4. THE DISTANCE TO M33

Although the shape of the PNLF at fainter magnitudes may change with time, the absolute magnitude of the function’s bright-end cutoff stays amazingly constant (Jacoby et al. 1992; Ciardullo 2003). Whether this is due to convergence in the initial mass–final mass relation (Jacoby 1997), a fortuitous correlation between central star UV flux and circumstellar extinction (Ciardullo & Jacoby 1999; Ciardullo 2003), or the contribution of some other PN-like object (Marigo et al. 2004) is unclear. Nevertheless, the insensitivity of the PNLF cutoff to population age and metallicity makes the feature extremely useful as an extragalactic standard candle.

To derive M33’s PNLF distance and its formal uncertainty, we followed the procedures of Ciardullo et al. (1989). We took the analytic form of the PNLF given in equation (2), convolved it with the photometric error function and shifted to the most likely distance of the galaxy. The agreement confirms that our PN sample is complete to $m = 23.75$ to within 3/4 of the nucleus.
Fig. 6.—[O III] $\lambda$5007 PNLFs of three Local Group stellar populations. The data for M31’s bulge are taken from Ciardullo et al. (2002); the SMC luminosity function comes from Jacoby & De Marco (2002). The absolute magnitudes are derived using the Cepheid distances to the galaxies (Freedman et al. 2001; Udalski et al. 1999) and the Galactic extinctions given by Schlegel et al. (1998). While the PNLF of M31’s bulge is a monotonically increasing function, the luminosity functions of the two star-forming systems decline at fainter magnitudes.

As pointed out by Ciardullo et al. (2002), a more likely explanation for M33’s distance discrepancy involves divergent assumptions about the galaxy’s internal reddening. M33’s Cepheid distance uses an extinction estimate that is derived from multicolor photometry of the Cepheids themselves (Freedman et al. 1991, 2001), while the PNLF method assumes the IRAS DIRBE Galactic value. The difference between these values, $E(B-V) = 0.17$, is one of the largest observed for any Cepheid galaxy. If we were simply to adopt the Cepheid reddening estimate of $E(B-V) = 0.21$ (as did Magrini et al. 2000), or if the internal reddening were reduced to one typical of other Cepheid galaxies, the PNLF and Cepheid distance indicators would be in much better agreement. Additional support for this hypothesis comes from the galaxy’s red giant stars. The distance moduli derived by Kim et al. (2002) using the tip of the red giant branch $[24.81 \pm 0.04 \text{ (random)} \pm 0.13 \text{ (systematic)}]$ and the location of the red clump $[24.80 \pm 0.04 \text{ (random)} \pm 0.05 \text{ (systematic)}]$ are both consistent with the PNLF value.

5. PLANETARY NEBULA SPECTROSCOPY

The radial velocities of M33’s PNs were measured with the WIYN telescope, the Hydra bench spectrophotograph, and a 600 line mm$^{-1}$ grating blazed at 10.1 in first order. Most of the PNs were observed on the photometric nights of 2002 October 3–5, with the 2" red fiber cable and a telescope-instrument combination that produced spectra with 2.8 Å (168 km s$^{-1}$) resolution and 1.4 Å pixel$^{-1}$ dispersion over the wavelength range between 4500 and 7000 Å. The data were taken using six fiber setups, each of which targeted ~50 PNs with approximately five fibers devoted to the sky. The total exposure time for each setup was ~3.5 hr, and, in order to allow for the removal of cosmic rays and other instrumental artifacts, the integrations were subdivided into a series of 30 minute exposures. To provide a check on the repeatability of the measurements, ~75% of the PNs were observed multiple times, and seven PNs, ranging in brightness between 21.0 < $m_{5007}$ < 25.0, were targeted in four or more setups.

To supplement these data, additional spectra for 24 PN candidates were obtained by R. Chandar and H. Ford on the nights of 2003 January 3–5 using HYDRA’s 3” blue fiber cable and a 400 line mm$^{-1}$ grating blazed in first order at 4.2. These data, which covered the wavelength range between 3660

---

we limited the range of the fit to $m_{5007} < 22.5$. In addition, in order to correct for foreground Galactic extinction, we adopted the IRAS DIRBE-based reddening value of $E(B-V) = 0.041$ (Schlegel et al. 1998) and used the relation of Cardelli et al. (1989) with $R_V = 3.1$ to go from differential to total $\lambda$5007 extinction. Assuming an absolute magnitude for the PNLF cutoff of $M^* = -4.47$ (Ciardullo et al. 2002), the most likely distance to M33 is $(m - M)_0 = 24.86$ (0.94 Mpc) with a formal fitting error of $+0.05/-0.10$ mag. When we include the systematic errors associated with the observation’s photometric zero point (0.03 mag), the filter response curve (0.03 mag), and the Galactic foreground extinction (0.03 mag), the most likely PNLF distance modulus becomes $(m - M)_0 = 24.86^{+0.07/-0.11}$, or $0.94^{+0.03/-0.05}$ Mpc.

Our PNLF distance modulus of $(m - M)_0 = 24.86^{+0.07/-0.11}$ is 0.30 mag larger than M33’s Cepheid distance modulus of 24.56 ± 0.10. If these uncertainties are accurate, then a simple propagation of errors implies that the two measurements are discrepant at the ~2σ level. There are two possible explanations for this discrepancy.

The first involves a metallicity dependence in either (or both) of the distance indicators. Thirteen galaxies have both PNLF and Cepheid distance measurements, and a detailed analysis of these data has been performed by Ciardullo et al. (2002). Their results are striking: except for the most metal-poor systems, the distance residuals are perfectly consistent with the internal uncertainties of the methods. Moreover, although a systematic shift is seen in low-metallicity ([O/H] ≤ −0.5) systems, the offset is exactly that predicted for the PNLF by Dopita et al. (1992). This agreement leaves little room for further metallicity corrections in either the PNLF or Cepheid methods.

Nevertheless, since our distance estimate to M33 assumes $M^* = -4.47$, it is possible that metallicity does play a role in the distance discrepancy. However, in order to do so, the mean oxygen abundance of the PNs in our sample must be ~0.2 dex below that of the Large Magellanic Cloud. The data of Garnett et al. (1997) show that this is unlikely: only in the outermost regions of the galaxy do the H$\alpha$ region abundances drop to this level. Moreover, if $M^*$ were being affected by metallicity, the strong abundance gradient across the galaxy would cause our measurement of $m^*$ in M33’s outer disk to be fainter than that of the inner disk. As Figure 4 demonstrates, this is not observed. In fact, the PNLF cutoff for the sample of PNs with $R > 15'$ is marginally brighter than that for the PNs inside this radius (by 0.15 ± 0.17 mag). This constancy strongly suggests that the low metallicity of M33 is not causing us to overestimate $M^*$ and the distance to the galaxy.

As pointed out by Ciardullo et al. (2002), a more likely explanation for the distance discrepancy involves divergent assumptions about the galaxy’s internal reddening. M33’s Cepheid distance uses an extinction estimate that is derived from multicolor photometry of the Cepheids themselves (Freedman et al. 1991, 2001), while the PNLF method assumes the IRAS DIRBE Galactic value. The difference between these values, $E(B-V) = 0.17$, is one of the largest observed for any Cepheid galaxy. If we were simply to adopt the Cepheid reddening estimate of $E(B-V) = 0.21$ (as did Magrini et al. 2000), or if the internal reddening were reduced to one typical of other Cepheid galaxies, the PNLF and Cepheid distance indicators would be in much better agreement. Additional support for this hypothesis comes from the galaxy’s red giant stars. The distance moduli derived by Kim et al. (2002) using the tip of the red giant branch $[24.81 \pm 0.04 \text{ (random)} \pm 0.13 \text{ (systematic)}]$ and the location of the red clump $[24.80 \pm 0.04 \text{ (random)} \pm 0.05 \text{ (systematic)}]$ are both consistent with the PNLF value.
and 6860 Å, had a slightly lower dispersion than the 2002 October data (~6 Å resolution at 1.56 Å pixel⁻¹), but longer (~5 hr) exposure times partially compensated for this drawback. These 2003 January observations yielded three additional PN candidates and reduced the measurement errors of 21 other objects.

Data reduction was accomplished with the DOHYDRA task within IRAF. Flat-fielding was performed using dome flat exposures obtained at the beginning of each night, and wavelength calibrations were found using a series of CuAr comparison arcs taken before, during, and after the science exposures of each setup. The solutions derived from these comparison arcs had an rms dispersion of ~0.03 Å (~2 km s⁻¹) and were stable throughout each setup. Next, the spectra were linearized to a common wavelength scale and sky subtracted using an average sky spectrum determined from the setup’s sky fibers. Because our observations focused on regions of the spectrum well away from any airglow emission and were performed during dark time at moderately high (R ~ 2000) dispersion, the details of this subtraction make very little difference to the final result. After sky subtraction, the individual exposures of each setup were co-added and shifted into the barycentric rest frame to produce a single summed spectrum for each object.

PN radial velocities were derived in a two-step process. First, we obtained an interim velocity for each PN using the centroids of the bright emission lines of [O III] λλ4959, 5007, Hα, Hβ, and (when possible) [N II] λλ6548, 6584. We then used these velocities to create an ultrahigh signal-to-noise ratio PN template by shifting all of the spectra into the rest frame and co-adding the data. Using this template, we then determined our final PN velocities by cross-correlating the individual spectra against the template spectrum with the XCSAO task of the RVSAO package (Kurtz & Mink 1998) of IRAF. Generally speaking, for faint emission line objects, velocities derived in this way are slightly more accurate than velocities found from the weighted average of individual emission-line measurements. In the case of our M33 survey, where most of the PNs are bright, the improvement to our measurements was marginal. The mean difference between velocities determined via cross-correlation and those found by the line-centroiding (EMSAO) task within RVSAO was only 0.15 ± 0.34 km s⁻¹.

At this stage, we verified the internal consistency of our velocity measurements by cross-correlating the spectra of each individual setup against the template spectrum and intercomparing the results for the 113 PNs with multiple measurements. Table 5 summarizes the results of this analysis by listing the mean velocity offset of each setup with respect to the others and the number of velocity pairs used to derive these values. From the table, it is clear that systematic setup-to-setup velocity differences are minimal: the only setup whose offset is not consistent with zero is that of setup 4, and its velocities are low by only ~2.9 km s⁻¹. Table 5 also lists the standard error between the individual pairs of PN measurements. This error provides an upper limit to the internal accuracy of our measurements.

The tenth column of Table 2 lists our velocity measurements for M33’s PNs. The ninth column of the table gives the number of fiber setups employed for each object, excluding the observations of 2003 January. Those PNs targeted in January are identified in the thirteenth column, and their velocities are the weighted mean of the results from the two observing runs. In total, of the 152 PN candidates identified on our survey frames, 151 were targeted by WIYN, and 140 were detected via their emission lines. (Of the 11 objects that were not detected, 7 are extremely faint.) Note that the errors quoted in the eleventh column are the internal uncertainties of the cross-correlation technique and do not include any setup-to-setup systematic uncertainty. As demonstrated in Table 5, such errors are small, ≤3 km s⁻¹.

### 6. THE M33 PLANETARY NEBULA VELOCITY FIELD

The large-scale velocity distribution of M33’s PNs is illustrated in Figure 1. The most obvious feature displayed in the figure is the galaxy’s rotation. M33 is inclined ~56° to the line of sight (see Zaritsky et al. 1989 and references therein), and this inclination is reflected in the radial velocities of the PNs. In fact, since M33’s spheroid-to-disk ratio is extremely small, S/D ~ 0.02 (Bothun 1992; Regan & Vogel 1994), all but ~2 PNs should be rotating along with the galactic disk.

To explore the dynamics of M33’s PN population, we began by assuming that, in the mean, the system’s PNs are confined to the disk and move in circular orbits about the galactic nucleus, which is at α = 1h33m50s, δ = +30°39′36″79″ (J2000; Massey et al. 1996). We then defined the inclination and orientation of this disk. On average, M33 is inclined i ~ 56° to the line of sight and has a major-axis position angle of θ = 23° (Kent 1987; Zaritsky et al. 1989); these values are consistent with the isophotal contours of the inner parts of the galaxy (Kent 1987), as well as the kinematics of the system’s H II regions (Zaritsky et al. 1989). However, such a model is probably too simplistic. The kinematics of M33’s H I gas demonstrates the presence of a significant warp in the galaxy’s outer regions (Rogstad et al. 1976; Corbelli & Schneider 1997; Corbelli & Salucci 2000). This warping is also confirmed by the asymmetrical shape of the galaxy’s spiral arms (Sandage & Humphreys 1980). To include this effect, we therefore adopted the model of Corbelli & Salucci (2000), in which the inclination and position angle of M33’s disk vary with galactocentric distance. Fortunately, our analysis depends very little on the details of this assumption: over the inner ~30′ of the galaxy (i.e., over most of our survey region), the difference between the Corbelli & Salucci (2000) model and the simple i = 56°, θ = 23° geometry is slight (no more than 4° in inclination and 5° in position angle). Nevertheless, since M33’s warp does become important outside this range, and since a few of our PNs do have large isophotal radii, we prefer this approach over the less sophisticated flat-disk assumption.

### Table 5

| Setups | Δv | σ_{pair} | N |
|--------|----|----------|---|
| 1–2   | −0.9 | 1.0 | 5 |
| 1–3   | −1.5 | 2.7 | 8 |
| 1–4   | −3.7 | 4.9 | 22 |
| 1–5   | +0.7 | 3.3 | 22 |
| 1–6   | −0.3 | 7.2 | 17 |
| 2–3   | +0.8 | 8.6 | 18 |
| 2–4   | −7.9 | 6.4 | 6 |
| 2–5   | −3.2 | 6.4 | 19 |
| 3–4   | −0.4 | 5.3 | 12 |
| 3–5   | −6.0 | 4.9 | 13 |
| 3–6   | −5.6 | 5.8 | 9 |
| 4–5   | −2.6 | 4.8 | 16 |
| 4–6   | +2.6 | 3.3 | 10 |
| 5–6   | +1.0 | 6.3 | 21 |
| 5–6   | +1.6 | 7.7 | 11 |
After defining M33’s inclination and position angle, we deprojected the position of each PN to determine its galactocentric radius \((R)\) and position angle with respect to the galaxy’s major axis \((\phi)\). The rotation speeds of the PNs, \(V_*\), were then inferred from the line-of-sight radial velocities, \(V_{\text{rad}}\) via
\[
V_* = \frac{V_{\text{rad}} - V_{\text{sys}}}{\sin i \cos \phi}. \tag{5}
\]
For the systematic velocity, we used \(V_{\text{sys}} = -180\ km\ s^{-1}\); this value is based on both radio (de Vaucouleurs et al. 1991) and optical (Huchra et al. 1999) observations of the galaxy and is comfortably close to the mean radial velocity of \(\langle V_{\text{rad}} \rangle = -176\ km\ s^{-1}\) found from the PNs of this study.

Figure 7a compares our PN rotational velocities with the rotation curve of M33’s atomic \((\text{H} I)\) and molecular \((\text{CO})\) gas (Corbelli 2003). All objects except 14 PNs within 10’ of the minor axis have been plotted. Overall, the agreement is quite good. The data display some amount of scatter, but this is due to the peculiar radial, tangential, and vertical velocities of the PNs. When these noncircular motions are amplified by the \(\cos \phi\) division, the result is the dispersion seen in the figure. Also apparent is the small, but significant, offset between the mean circular velocity of the PNs and the rotation speed of the gas. Figure 7b shows this offset more clearly by binning the data by galactocentric radius with 18 objects per bin. The amplitude of the velocity lag, or asymmetric drift, is \(\sim 10\ km\ s^{-1}\), less than half the drift velocity observed for old-disk stars in the vicinity of the Sun (Ratnatunga & Upgren 1997; Dehnen & Binney 1998). However, since the ratio of M33’s drift speed to rotation velocity is approximately the same as that for the solar neighborhood, our measurement suggests that the dynamical state of M33’s disk is not too dissimilar to that of the Milky Way. We return to the issue of asymmetric drift in \(\S\ 7\), where we use it to help constrain the shape of the system’s velocity ellipsoid.

Another way to examine the motion of M33’s PNs is to plot their line-of-sight velocities after removing the contribution of the rotating stellar disk (defined as the \(\text{H} i + \text{CO}\) gas velocity minus the asymmetric drift). This is done in Figure 8; Figure 8a shows the individual PN velocities, while Figure 8b bins the velocities to show the line-of-sight velocity dispersion. The figure again demonstrates that the noncircular motions of the PNs are significant, with \(\sigma_{\text{rad}} \sim 25\ km\ s^{-1}\). This is more than twice the \(9 \pm 4\ km\ s^{-1}\) dispersion found for the galaxy’s \(\text{H} II\) regions (Brandt 1965; Zaritsky et al. 1989). In addition, our innermost measurement of \(\sigma_{\text{rad}} = 22 \pm 3\ km\ s^{-1}\) is very close to the galaxy’s central value of \(24.0 \pm 1.2\ km\ s^{-1}\) found from absorption-line spectroscopy (Gebhardt et al. 2001). Since M33 shows no dynamical evidence for a central black hole or mass concentration, this agreement confirms that we are, indeed, measuring a velocity dispersion that is representative of old disk stars.

Figure 8 also supports our contention that there are very few nondisk PNs in our sample. Of the 140 PNs surveyed, only
for \( \sigma_R \) was more difficult. In the vicinity of the Sun, the ratio of the vertical to radial velocity dispersion for old disk stars is \( \sigma_z/\sigma_R \approx 0.5 \) (Bienaymé 1999), and surveys of other spiral galaxies consistently find ratios in the range \( 0.3 < \sigma_z/\sigma_R < 0.9 \) (e.g., Gerssen et al. 1997, 2000; Bottema 1993; van der Kruit & de Grijs 1999). Unfortunately, almost all these data are for stars between \( \sim 1 \) and \( \sim 2 \) disk scale lengths from the galactic center: there are no measurements of \( \sigma_R \) or \( \sigma_z \) over the entire range of radii considered in this paper. Thus, our only guidance on the shape of M33’s velocity ellipsoid comes from numerical simulations. According to the models of Villumsen (1985) and Jenkins & Binney (1990), disk heating by molecular clouds and spiral structure should drive an initially isotropic velocity dispersion toward values of \( \sigma_z/\sigma_R \) as low as \( \sim 0.4 \). For a lower limit on \( \sigma_R \), we can therefore demand that \( \sigma_R > \sigma_z \). For the upper limit on this quantity, we can use the requirement that M33’s disk be stable against buckling instabilities, which occur when \( \sigma_z/\sigma_R \gtrsim 0.3 \) (Toomre 1966; Araki 1985; Merritt & Sellwood 1994). To be slightly conservative, we therefore limited our analysis to \( 0.25 < \sigma_z/\sigma_R < 1.0 \).

Figure 9 displays the likelihood contours produced by our data. As is obvious from the figure, \( \sigma_R \) is better constrained than \( \sigma_z \). This is understandable: not only did our boundary conditions ensure that \( \sigma_R > \sigma_z \), but the former variable also receives the contribution of \( \sigma_p \), via the epicyclic approximation. As a result, the line-of-sight velocity dispersion largely reflects the behavior of \( \sigma_R \); our limits on \( \sigma_z \) come primarily from the constraints placed on the shape of the velocity ellipsoid and secondarily from the fits.

The more important property shown in Figure 9 is the absence of an obvious radial gradient. There is very little difference between the fits for PNs near the center of M33 and PNs at \( \gtrsim 4 \) disk scale lengths. If the shape of the galaxy’s velocity ellipsoid is roughly constant with radius, then the data imply that \( \sigma_z \) must also be very nearly constant. On the other hand, if the vertical velocity dispersion truly declines with radius (as it must in a constant mass-to-light ratio disk), the dispersion ratio \( \sigma_z/\sigma_R \) must also decline.

To make further progress, we can apply two additional constraints to our data. The first is the requirement that M33’s disk be stable. According to Toomre (1964), in order for a thin stellar disk to be stable against axisymmetric perturbations, its radial velocity dispersion must satisfy the condition

\[
\sigma_R > \frac{3.36 G \Sigma}{\kappa},
\]  

where the epicyclic frequency, \( \kappa \), is obtainable from the rotation curve via

\[
\kappa = \frac{V_c}{R} \left( 2 + 2 \frac{\partial \ln V_c}{\partial \ln R} \right)^{1/2}
\]  

(Binney & Tremaine 1987). In addition, for stability against nonaxisymmetric perturbations (i.e., bar formation), the disk must also satisfy the (slightly more stringent) condition

\[
\sigma_R > \frac{3.36 G \Sigma 2V_c}{R \kappa}
\]  

(Morozov 1980, 1981a, 1981b). If we combine this equation with the isothermal disk approximation given by equation (6),
we obtain a constraining relationship between the vertical and radial velocity dispersions

\[ \sigma_z < \kappa \left( \frac{\pi z_0 R \sigma_R}{6.72V_c} \right)^{1/2} \]  

(12)

To use this relation, we need to know \( z_0 \), the scale height of PNs in M33’s disk. Our probability contours of Figure 9, in combination with the buckling instability condition, allow us to exclude values of \( z_0 \leq 90 \) pc in M33’s outer regions (\( R > 6 \) kpc). To improve on this number, however, we must turn to other galaxies. In the Milky Way, measurements of \( z_0 \) are notoriously uncertain, with published values ranging all the way from 90 to 300 pc (e.g., Pottasch 1984, p. 119; Maciel 1989; Zijlstra & Pottasch 1991; Corradi & Schwarz 1995; Phillips 2003). Unfortunately, in other galaxies the situation is worse: internal extinction in the planes of edge-on galaxies makes reliable measurements of the PN distribution nearly impossible (see Ciardullo et al. 1991; Jacoby et al. 1996). For lack of a better alternative, we were therefore forced to adopt an intermediate value of \( z_0 \sim 175 \) pc as the most likely scale height of our test particles. This value has two justifications. First, a Milky Way scale height of \( z_0 \sim 175 \) pc implies a vertical PN distribution similar to that of Galactic main-sequence F stars (Mihalas & Binney 1981) and a total PN population that is consistent with PN observations in other galaxies (Peimbert 1993). Second, Bizyaev & Mitronova (2002) have found a correlation between the infrared \( (K_s) \) central surface brightness of a galaxy and the ratio between the galaxy’s vertical and radial scale lengths. If one applies this relation to M33 (Regan & Vogel 1994) and normalizes the result using the ratio of our Galaxy’s PN scale height to the total “old thin disk” scale height (Chen et al. 2001), then \( z_0 \) again becomes \( \sim 175 \) pc. Obviously, this value carries a substantial uncertainty of at least \( \sim 25\% \).

We note that, in theory, the scale height of a galactic disk may change with radius. If it does, then the effect of the stability
criterion will also have a radial dependence. In practice, however, this is not of great concern. Numerous studies of late-type edge-on spiral galaxies in the optical and infrared (e.g., van der Kruit & Searle 1982; Shaw & Gilmore 1990; de Grijs & van der Kruit 1996; de Grijs & Peletier 1997; Fry et al. 1999; Bizyaev & Mitronova 2002) have found that radial variations in the vertical scale heights of stars are always small, ≤10%. This is true even when the galaxies are tidally interacting: a study by Schwarzkopf & Dettmar (2001) shows that the fractional variation of $z_0$ in disturbed edge-on systems is typically less than 14%. Thus, even if M33’s disk is warped at large radii, our assumption of a constant value for the scale height should be acceptable.

The solid lines of Figure 9 show the effect the stability requirement has on our probability contours. In the inner regions of the galaxy, the requirement is satisfied almost everywhere: virtually all the solutions lie below the stability cutoff. However, at large radii, the criterion severely limits the allowable values for the vertical velocity dispersion. Such an effect is unavoidable: if M33’s disk is stable and isothermal, then the vertical velocity dispersion at large distances from the nucleus must be small.

Figure 10 applies the stability criterion and marginalizes Figure 9’s remaining probability contours over the parameters $\sigma_g$ (middle panel) and $\sigma_z$ (bottom panel). The top panel of the figure shows the behavior of $\sigma_z$ if one asserts that M33’s stellar disk is at the limit of stability; this solution also represents the largest values of $\sigma_z$ possible under the isothermal disk approximation. In the figure, we have illustrated the shapes of the non-Gaussian errors via the widths of the shaded areas. The figure contains several features of note.

First, as anticipated, the requirement of stability produces a radial gradient in the vertical velocity dispersion. Remarkably, this gradient is much shallower than that expected from a constant mass-to-light ratio disk. In the $V$ band, M33’s disk scale length is $\sim$9.1 (Guidoni et al. 1981); in the infrared, where the effects of extinction gradients are less, the scale length shrinks to $\sim$5.8 (Regan & Vogel 1994). Our kinematic measurements imply a physical scale length for M33’s disk that is $\sim$13′ or $\sim$3.5 kpc. This scale length does not change much, even if one requires that the disk be pressing the limits of stability.

The other notable feature of Figure 10 is the behavior of the radial component of M33’s velocity ellipsoid. Outside $\sim$2.5 kpc, $\sigma_R$ declines with a scale length that is more than 3 times greater than that of the galaxy’s optical light; this is slightly larger than the scale length one would expect from the survey data of Bottema (1993), but not exceptionally so. However, inside $\sim$2.5 kpc, $\sigma_R$ levels off and may even decline at small galactocentric radii. Although such behavior is not common, neither is it unique: the line-of-sight velocity dispersions of NGC 3198 and NGC 6503 suggest that these galaxies also have low central values of $\sigma_R$ (Bottema 1993).

The turnover of M33’s radial velocity dispersion is most likely a manifestation of the requirement that $\sigma_R$ be much less than the galaxy’s rotation speed. M33’s rotation velocity increases from zero near the galactic center to $\sim$90 km s$^{-1}$ at $\sim$3 kpc. If $\sigma_R$ did not turn over, it would dominate galactic rotation within $\sim$1.5 kpc of the nucleus, and the result would be a “hot” stellar population in the galaxy’s central regions. If such a population exists in M33, it is extremely weak (Bothun 1992; Regan & Vogel 1994).

The turnover of the radial velocity dispersion explains why the system’s line-of-sight dispersion is so flat: the increase in $\sigma_z$ near the center of the galaxy is offset by the decline in $\sigma_R$. This flattening, however, does not affect the large-scale trend in the dispersion ratio. In the interior of the galaxy, $\sigma_z/\sigma_R \sim 0.6$; by $\sim$9 kpc, the ratio has dropped to $\leq$0.4. This range of values is consistent with the numerical models of disk heating computed by Villumsen (1985) and Jenkins & Binney (1990). More importantly, our inferred $\sigma_z/\sigma_R$ gradient agrees with the analytic results of Carlberg (1987), who showed that the velocity ellipsoid of a disk must become rounder as the stellar velocity dispersion increases. Our recovery of this result supports the validity of our $\sigma_z/\sigma_R$ decomposition.

There is one more constraint that can be applied to M33’s stellar kinematics. From the Jeans equation, the rotational velocity of stars in a stellar disk is related to the circular velocity of the system’s gas by

$$V_c^2 - \langle V_c \rangle^2 = \sigma_R^2 \left( \frac{\sigma_R}{\sigma_z} - 2 \frac{\partial \ln \sigma_R}{\partial \ln R} - \frac{\partial \ln \nu}{\partial \ln R} + 1 \right)$$

- \frac{R}{\nu} \frac{\partial \nu}{\partial z}$$

where $\nu$ is the stellar density (Binney & Tremaine 1987). If we use the epicyclic approximation and assume that M33’s PNs

![Figure 10](image-url)
are scattered throughout a constant scale height isothermal disk, then this equation for asymmetric drift simplifies to

$$V_c^2 - \langle V_z \rangle^2 \simeq \sigma_z^2 \left( \frac{1}{2} \frac{\partial \ln V_c}{\partial \ln R} - 2 \frac{\partial \ln \sigma_R}{\partial \ln R} - 2 \frac{\partial \ln \sigma_z}{\partial \ln R} \right).$$  

(14)

This equation, when combined with a model for the radial dependence of $\sigma_z$ and $\sigma_R$, places a further constraint on M33’s galactic kinematics.

Based on the galaxy’s luminosity profile and the data displayed in Figure 10, it is reasonable to parameterize the radial dependence of M33’s vertical velocity dispersion with a simple exponential of scale length $R_c$. An exponential function with scale length $R_c$ can also be used to fit the galaxy’s radial velocity dispersion, as long as the measurements are restricted to galactocentric distances greater than $\sim 2.5$ kpc. Inside this radius, however, the exponential law breaks down, and the form of the $\sigma_R$ versus $R$ relation is unknown.

Rather than guess the radial dependence of $\sigma_R$ in M33’s inner regions, we elected instead to exclude from the analysis all measurements of $\sigma_R$ and asymmetric drift within 2.5 kpc of the nucleus. This allowed us to parameterize the radial dependence of $\sigma_z$ and $\sigma_R$ with exponentials and rewrite equation (14) as

$$V_c^2 - \langle V_z \rangle^2 \simeq \sigma_z^2 \left( \frac{2R}{R_c} + \frac{2R}{R_c} \frac{1}{2} + \frac{1}{2} \frac{\partial \ln V_c}{\partial \ln R} \right).$$  

(15)

Using this equation, we simultaneously solved for the values of $R_c$, $R_R$, $\sigma_R(R = 0)$, and $\sigma_z(R = 0)$ that best fitted both the velocity dispersions of Figure 10 and the asymmetric drift data of Figure 7. To perform this fit, we minimized the $\chi^2$-like statistic

$$\chi^2 = -2 \sum \ln \left[ \frac{P(y_i)}{P(y_m)} \right],$$  

(16)

where $P(y_i)$ is the probability of observing value $y_i$ from a probability distribution $P$ and $P(y_m)$ is the probability of observing the most likely value. This statistic, which reduces to the standard $\chi^2$ value for normally distributed errors, enabled us to consider the non-Gaussian uncertainties associated with $\sigma_R$ and $\sigma_z$ and the asymmetric drift errors in a consistent manner. It also permitted us to estimate the uncertainties in our scale lengths via a modified jackknife analysis. In this procedure, we repeatedly removed one PN at random from each of the bins and refitted the data. When analyzed in this way, the best-fit scale length for $\sigma_z^2$ became 13.5 (with 90% of the simulations lying between 12.2 and 14.9), and the most likely value for the scale length of $\sigma_R$ was 33.5 (with 90% of the simulations between 30.5 and 38.6). The results do not change much if one forces the disk to marginal stability: under this condition, the derived scale length of $\sigma_z^2$ declines by less than $\sim 5\%$. Once again, these scale lengths are significantly larger than the 5.8 value found from K-band photometry. In fact, our 3 $\sigma$ limits on M33’s dynamical scale length lie between 10'8 and 16'0, and none of our 20,000 simulations produced a scale length as low as the galaxy’s optical scale length of 9'. Our best-fit solution is illustrated in Figures 7b and 10 by dotted lines.

If the disk of M33 is indeed isothermal, then we can substitute the vertical velocity dispersions of Figure 10 into equation (6) to produce estimates for the galaxy’s disk mass surface density. These densities can then be combined with optical (Guidoni et al. 1981) and infrared (Regan & Vogel 1994) surface photometry to generate values for the total disk mass-to-light ratio. Finally, we can remove the contribution of M33’s interstellar medium from the mass-to-light values by subtracting the surface mass density of galactic H i and H_2 gas (Heyer et al. 2004) from the total mass density and applying a first-order correction for internal extinction, using the central extinction value ($A_V = 0.9$ mag) and extinction scale length ($\sim 10'2$) derived by Regan & Vogel (1994). These mass estimates and mass-to-light ratios are presented in Table 6.

The errors quoted in Table 6 reflect only the formal uncertainties associated with our measurement of $\sigma_z$; not included are the systematic errors associated with our adopted value of $\sigma_R$ and the use of the isothermal disk approximation. The former uncertainty is small: a 25% increase in the PN scale height (from 175 to 220 pc) results in only a $\sim 5\%$ decrease in the derived values for M33’s disk mass surface density and scale length.

The error associated with the isothermal disk approximation is slightly more formidable. Overall, the isothermal approximation provides an adequate fit to the vertical structure of edge-on spiral galaxies (van der Kruit & Searle 1982) and the stellar kinematics of the solar neighborhood (e.g., Wielen 1977). However, near the galactic plane, the approximation breaks down: the stellar density distribution is more sharply peaked than the isothermal model would predict (Fuchs & Wielen 1987; de Grijs & van der Kruit 1996). This has led van der Kruit (1988) to consider a family of alternative models, which have isothermal [sech² ($z/z_0$)] and exponential [exp ($-z/z_0$)] scale distributions as their limiting cases. According to these models, M33’s surface mass can be up to 33% larger than that derived using the simple isothermal assumption.

The final systematic uncertainty that is not included in Table 6 is one that affects our mass-to-light ratio measurements
in M33’s outer regions. Reliable optical and infrared surface photometry exists only for M33’s inner ~20′; outside this region, one must extrapolate the galaxy’s surface brightness from its measured scale length. The length of this extrapolation is significant, so a small error in scale length propagates into a large error in $M/L$. Moreover, if M33’s disk is truncated (e.g., Kregel et al. 2002), or if a small amount of excess luminosity is left over from a tidal interaction, then the derived mass-to-light ratios will be seriously in error. In fact, the extrapolated $V - K$ color for the outer disk of M33 is not consistent with any population synthesis model (e.g., Maraston 1998). Given the uncertainties associated with infrared photometry, our anomalously large values for $M/L_K$ in the outer disk of the galaxy are almost certainly due to errors in the extrapolated $K$-band surface photometry.

8. THE SCALE LENGTH OF M33’S DISK

Our value for M33’s disk scale length is more than twice that of the galaxy’s infrared light and 1.5 times that of the galaxy’s $V$-light. This discrepancy implies a gradient in the galaxy’s disk mass-to-light ratio. Specifically, our observations imply that M33’s disk $M/L$ slowly increases with radius. Part of this increase is due to the growing contribution of M33’s interstellar medium, which has a density distribution that is much flatter than that of the stars (Heyer et al. 2004). However, even when this component is taken into account, a positive gradient in the disk mass-to-light ratio remains. The sign of this gradient runs contrary to expectations: in most models of galaxy formation, it is the inner regions of a galaxy that form first and contain the older, higher mass-to-light ratio stars (e.g., Eggen et al. 1962; Searle & Zinn 1978; Fall & Efstathiou 1980; Steinmetz & Müller 1994). Moreover, multicolor imaging (Bell & de Jong 2000) and $H_\alpha$ surface photometry (Ryder & Dopita 1994) tend to support this picture: most disks appear to possess significant age gradients that are consistent with the inside-out scenario. Thus, we must consider the possibility that our mass measurements are biased. In fact, there are several effects that could conceivably introduce a systematic error into our estimates of $\sigma_z$. If any of these exist, then our value for the galaxy’s kinematic scale length could be in error.

The first possible bias comes from contamination by emission-line objects that are not PNs. In M33, $H_\alpha$ regions and supernova remnants far outnumber PNs, and, as Population I objects, their velocity dispersions are significantly less than those of the PNs (e.g., Zaritsky et al. 1989). If the fraction of contaminants is higher near the center of M33 than it is at large radii, then a systematic error in the scale length will be the result. Some support for this hypothesis comes from the image quality of the [O $\ii$] $\lambda 5007$ frame of M33’s central field: the seeing on that image (1′05) is the poorest of any frame in the Massey et al. (2002) survey.

Still, for all the reasons detailed in § 2, contamination is unlikely to be a significant problem in our survey. All of our PN candidates were detected in both [O $\ii$] $\lambda 5007$ and $H_\alpha$, and the excellent seeing on the latter image (and on the broadband $V$ frame) should have excluded virtually all $H_\alpha$ regions from our sample. Similarly, given the survey’s image quality and the intrinsic ratio of PNs to supernovae, it is difficult to envision how the fraction of faint compact supernova remnants could be large enough to cause a systematic error in $R_z$. At worst, contamination should cause only a slight underestimate in $\sigma_z$ throughout the galaxy.

A second possible explanation for the large kinematic scale length involves a selection effect. The inner regions of M33 are quite dusty, with internal extinction values of $A_V \approx 1$ mag (Israel & Kennicutt 1980; Berkhuijsen 1983; Regan & Vogel 1994; Petersen & Gammelgaard 1997). Consequently, PN surveys in this area are biased against objects in and on the far side of the galactic plane. Under the isothermal disk approximation, this bias is irrelevant, since the velocity dispersion of such a disk is independent of galactic latitude. However, high above the plane, where the stellar orbits reach turnaround, the approximation must break down. If most of the PNs in our sample are high-latitude objects, then it is possible that our PN velocities are not representative of the true kinematic structure of the disk.

Since M33’s internal extinction declines rapidly with radius (Israel & Kennicutt 1980; Berkhuijsen 1983; Regan & Vogel 1994; Petersen & Gammelgaard 1997), the above selection effect has the potential for creating an artificial gradient in $\sigma_z$. However, the effect goes in the wrong direction. If extinction is to produce an overestimate of the disk’s kinematic scale length, the PNs at high galactic latitude must have a lower dispersion than those near the plane. However, models developed to fit the luminosity profiles of edge-on galaxies all have vertical velocity dispersions that increase with galactic latitude (van der Kruit 1988). Thus, while internal extinction may indeed be censoring our data, the effect cannot produce the observed discrepancy.

A third way of explaining our large value of $R_z$ involves our basic assumptions about the kinematics of M33’s disk and the stability criteria. As Figure 9 demonstrates, the PN velocities, by themselves, do not fix $\sigma_z$. Instead, it is the velocity dispersion measurements, in combination with the requirements of dynamic stability, that produce the galaxy’s dispersion gradient. If our warped-disk model for M33 is incorrect, or if the stability criteria have an additional radial dependence, or if the isothermal approximation breaks down at large radii, then an error in the scale length could be the result.

The first of these possibilities can quickly be excluded: in the area covered by our survey, the effects of M33’s warp are minimal. In fact, if we were to model the galaxy using a simple flat-disk geometry, the derived kinematic scale length of the system would increase by ~10%.

The latter two uncertainties are more difficult to assess. Certainly, the stability criterion given by equation (11) is only an approximation, as it assumes spiral arm pitch angles that are close to zero and neglects (among other things) the thickness of the disk and the existence of gaseous subsystems. Nevertheless, numerical experiments suggest that the formulation is approximately correct (e.g., Griv et al. 1999; Khoperskov et al. 2003). Similarly, although the limit on $\sigma_z/\sigma_R$ from the buckling (or fire hose) instability is not precisely known, the numerical experiments to date all indicate that a dispersion ratio of ~0.3 is reasonable (Araki 1985; Merritt & Sellwood 1994; Sellwood 1996). Finally, although one expects the isothermal approximation to become less accurate at large radii (as a result of the longer timescales for stellar interactions), surface photometry of edge-on spiral galaxies suggests otherwise (e.g., van der Kruit & Searle 1982; Fry et al. 1999; Bizyaev & Mitronova 2002).

The final way of creating an artificial gradient in M33’s disk mass-to-light ratio is through the introduction of a radial gradient in the PN scale height. Our measurement of the disk scale length assumes that $z_0$ is constant throughout the galaxy, and, for disks in general, this is a good assumption (e.g., Shaw & Gilmore 1990; de Grijs & van der Kruit 1996; de Grijs & Peletier 1997). However, our analysis also assumes that PNs
belong to the thin disk system and that the PN scale height does not change with radius.

The former assumption is certainly valid. Although deep observations of edge-on galaxies often reveal the existence of red “thick-disk” envelopes (e.g., Dalcanton & Bernstein 2002), the amount of luminosity in this component is small enough so that thick-disk PNs should be exceedingly rare. PN surveys of the edge-on spiral galaxies NGC 891 (Ciardullo et al. 1991) and NGC 4565 (Jacoby et al. 1996) confirm this fact: very few objects are found in the thick-disk region. Still, within the thin disk, the scale height of PNs could still change. For example, in the solar neighborhood, the vertical velocity dispersion of a stellar population is roughly proportional to the square root of the population’s age (e.g., Wielen 1977; Jahreiss & Wielen 1983). Since PNs are derived from a mix of populations with turnover masses between 1 \( M_\odot \lesssim M \lesssim 5 \ M_\odot \), their dispersion is a weighted average over the age range that corresponds to these masses, i.e., from \( t_1 \sim 10^{10} \) yr to \( t_2 \sim 10^8 \) yr. In other words,

\[
\langle \sigma_z^2 \rangle = \int_{t_1}^{t_2} \Phi(t) \frac{dN}{dm_m} \frac{dm_m}{dt} \sigma_z^2(t) \, dt,
\]

where \( \Phi(t) \) is the region’s star formation rate history and \( m_m \) is the main-sequence turnover mass of a population with age \( t \). Since velocity dispersion is related to scale height via equation (6), this means that a gradient in a galaxy’s star formation rate history can produce an artificial gradient in the PN scale height.

Figure 11 quantifies this effect for the case in which \( \Phi \) declines exponentially with scale time \( \tau \). In late-type spiral galaxies such as M33, the evidence suggests that the rate of star formation has been roughly constant over a Hubble time (Kennicutt et al. 1994). If this is the case, then \( \tau \gg 1 \), and the dependence of \( \sigma_z \) on \( \Phi \) is weak. Nevertheless, if the star formation rate history of M33’s inner disk is substantially different from that of its outer disk, a systematic error in our mass determinations may be the result.

Unfortunately, the effect shown in Figure 11 is much too small to explain M33’s large kinematic scale length: in order to be consistent with a constant \( M/L_\odot \) disk, the scale height of M33’s PNs would have to change by almost an order of magnitude, from \( \gtrsim 500 \) pc near the galactic center to \( \lesssim 100 \) pc at \( 8 \) kpc. Not only is this many times greater than what can be achieved by changing the galaxy’s star formation rate history, but the direction of the effect is wrong. If the stars of M33’s interior are systematically older than the stars of the outer disk (as might be expected from an inside-out galaxy formation scenario), then our assumption of a constant \( z_0 \) would produce an overestimate of the surface mass in the inner regions of the galaxy and an underestimate of surface mass in the outer disk. This would cause us to underestimate the scale length of the system. Moreover, the discrepancy between the galaxy’s photometric and dynamical scale lengths would be exacerbated by the reaction of the system’s infrared light to such a gradient. While the \( K \) band is far superior to the optical when it comes to viewing stellar luminosity through the veil of interstellar extinction, it is not immune to population effects. Just as the OB stars of present-day star formation distort surface brightness measurements made in \( B \), red supergiants left over from star formation of the recent past contaminate data taken in \( K \). Although photometric enhancements in the infrared are less than those in the optical, they are still substantial (Regan & Vogel 1994), and if the inner regions of M33 are older than the stars of the outer disk, the galaxy’s IR luminosity profile will appear more diffuse than the mass. This is the opposite of what is seen.

We note that, in principle, the question of M33’s star formation history can be answered via PN spectrophotometry. Kaler & Jacoby (1990) have shown that the ratio of nitrogen to oxygen in a PN correlates with its core mass: objects with \( M_c < 0.65 \ M_\odot \) have \( N/O \sim 0.2 \), while higher mass objects have \( 1 \lesssim N/O \lesssim 2 \). Consequently, this discontinuity, which is most likely caused by dredge-up in asymptotic giant branch (AGB) stars with \( M \gtrsim 3 \ M_\odot \) (Iben & Renzini 1983), can be used to identify the PNs of high-mass progenitors. Furthermore, Dopita et al. (1997) have shown that one can estimate PN central star masses directly by using deep spectrophotometry to locate the stars on the H-R diagram and then comparing their locations to post-AGB evolutionary models. This technique, when combined with an initial mass–final mass relation (i.e., Weidemann 2000), can allow us to trace the star formation history of different parts of M33’s disk and test for the presence of age gradients in the PN sample.

9. M33’s Disk and Halo

There have been numerous efforts to probe the radial distribution of M33’s dark halo via the shape of the galaxy’s rotation curve (e.g., Giraud 2000; Corbelli & Salucci 2000; Corbelli 2003). However, these efforts have all been limited by the unknown mass-to-light ratio of the galaxy’s disk. The disk mass estimates of Table 6 provide a new constraint on M33’s mass distribution and help break the disk-halo degeneracy.

To demonstrate this constraint, we began by fitting a smooth Brandt (1960) curve to the HI and CO rotation measurements of Corbelli (2003). We then used our mass model to infer a rotation curve of the disk alone. The radial profile of the dark halo was then estimated by simply differencing the two components. Figure 12 illustrates this disk-halo decomposition.

From the rotation curves of Figure 12 and the mass-to-light ratios of Table 6, it is clear that there is very little dark matter in M33’s central regions. Based on our PN velocity dispersion measurements, the disk mass-to-light ratios for M33’s central \( \sim 2 \) kpc are \( M/L_K \sim 0.2 \) and \( M/L_V \sim 0.3 \). Both these values are close to those predicted for stellar populations dominated by young (\( \tau \sim 150 \) Myr old) stars (Maraston 1998; Bell & de Jong 2001). More importantly, these numbers are statistically
Whatever the reason for M33’s increasing disk mass-to-light ratio, the fact that the galaxy’s disk mass-to-light ratios are consistent with those expected from young stellar populations strongly suggests that dark matter is not an important constituent of the disk. Models, such as the molecular gas hypothesis of Pfenniger et al. (1994), are certainly ruled out.

Our analysis is based on the velocities of only 140 PNs, so the spatial resolution of our measurements is poor. However, if we assume that M33’s disk has an exponential profile and that our surface mass densities are accurate, then we can use our disk/halo decomposition to place limits on the radial profile of the galaxy’s dark matter. According to cold dark matter models of galaxy formation, systems such as M33 should contain “cusps” of dark matter at their center, with power-law density profiles in the range $-1.5 < \alpha < -1.5$, where $\rho \propto r^\alpha$ (e.g., Navarro et al. 1996; Moore et al. 1998; Davé et al. 2001; Klypin et al. 2001; Power et al. 2003). As is illustrated in the bottom panel of Figure 12, our central slope is much shallower than this, with $\alpha \approx -0.5$. This result is robust: if we alter the PN scale height by 50%, or change the disk scale length by 25%, or increase the total disk mass by 33% (by either adopting an exponential law for the disk’s vertical profile or changing the central value of $\sigma_z$), the slope of M33’s dark matter core remains low. The result also agrees with the conclusions of Corbelli (2003), who ruled out the presence of a steep dark matter core in M33 on the basis of the galaxy’s rotation curve and the assumption of a constant (K band) mass-to-light ratio disk. Our analysis removes the latter assumption but is otherwise similar.

If our analysis is correct, then M33 is another example of a galaxy whose interior dark matter profile is flat. High-resolution rotation curves now exist for a large number of low-luminosity and low surface brightness spiral galaxies and dwarfs, and every one is consistent with a value of $\alpha > -1$ (e.g., Borriello & Salucci 2001; Swaters et al. 2003). It is still unclear whether this shallow slope is a critical problem for the cold dark matter paradigm (de Blok et al. 2003), or whether the discrepancy can be explained by the limitations in the measurements or the models (Swaters et al. 2003; Spekkens & Giovanelli 2004). Nevertheless, M33 appears to be another example of this problem.

Outside $\sim 1$ kpc, our analysis suggests that M33’s dark matter halo obeys the inverse square law density dependence that is expected from isothermal and Navarro et al. (1996) distributions. Because the dynamical scale length of the disk is larger than the photometric scale length, the system’s total disk mass is larger than that expected from the maximal disk hypothesis. Within a radius of 10 kpc, our velocity dispersion measurements yield a disk mass $\sim 9.1 \times 10^9 M_\odot$. This is $\sim 2.5$ times larger than one would obtain using $V$-band surface photometry and the maximum disk hypothesis and $\sim 6$ times larger than one would derive from a similar analysis in $K$. Still, within this radius, M33’s disk contains only $\sim 30\%$ of the total mass of the galaxy.

Of course, M33’s disk does not end at 10 kpc. Although our PN detections end near this 10 kpc limit, observations of M33’s HI gas extend to nearly twice that distance (Corbelli 2003). The interpretation of the extended HI gas data is made slightly more difficult by the increasing importance of the galactic warp at large radii. Nevertheless, if the models of Corbelli & Schneider (1997) and Corbelli & Salucci (2000) are accurate, then M33’s rotational velocity increases from $\sim 110$ to $\sim 135$ km s$^{-1}$ between 10 and 20 kpc. This suggests that the bulk of M33’s dark matter halo still resides outside the region of our survey.

Fig. 12.—Top: Comparison of M33’s HI and CO rotation curve with the rotation curve that would be produced from the disk alone. The solid line is a Brandt (1960) law fit to the HI and CO velocity measurements; the shaded regions indicate the 1, 2, and 3 $\sigma$ uncertainties about our best-fit mass model (as determined from a jackknife analysis). Bottom: Density law inferred for the galaxy’s dark halo (under the assumption of spherical symmetry), again with 1, 2, and 3 $\sigma$ uncertainties.
10. CONCLUSIONS

Until now, our knowledge of the stellar kinematics of spiral galaxies has come principally from observations of stars in the solar neighborhood and secondarily from absorption measurements in the inner regions of face-on and edge-on spiral galaxies. Our PN observations have improved upon this situation. For the first time, we have been able to measure the shape of the stellar velocity ellipsoid throughout a galactic disk and directly measure the variation in a galaxy’s disk mass surface density. Our data suggest that, in the inner regions of galaxies, maximum disk models for the rotation curve are appropriate and the contribution of dark matter is negligible. However, the data also show that the mass-to-light ratios of galactic disks do change with radius, and the \( M/L \) of the inner disk may be a factor of \( \geq 4 \) smaller than that of the outer disk.

Still, the existence of this gradient does not change the fact that most of a galaxy’s material lies well away from the disk in a dark halo that roughly follows an inverse square density law. Models in which the dark matter is in the form of molecular gas are excluded by our observations.

We would like to thank E. Corbelli for providing M33’s rotation curve in digital form, D. Willmarth for assisting in the setup for our PN spectroscopy, and R. Chandar for providing additional PN spectra for our analysis. In addition, we would like to thank M. Bershady, J. Sellwood, S. Sigurdsson, and S. McGaugh for useful comments during the preparation of this paper. This work was supported by NSF grant AST 00-71238.
