THE ASTROPHYSICAL JOURNAL, 376:467–479, 1991 August 1

GALACTIC DISK WARPS

KONRAD KUIJKEN
Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, Ontario, M5S 1A1, Canada
Received 1990 November 21; accepted 1991 January 31

ABSTRACT

Warped equilibrium configurations of galactic disks embedded in flattened haloes are considered. The relevant torques are calculated in their full nonlinearity, extending the results of linear-mode calculations. It is shown that for two specific examples, NGC 2841 and M33, it is possible to reproduce closely the observed warp shapes with these models. It is further demonstrated that the observed shapes of a galaxy's warp and rotation curve do not suffice to determine the disk mass uniquely, but can only provide upper limits on the disk-to-halo mass ratio. In the case of M33, this limit is about one-third of a maximal disk decomposition of the rotation curve.

Qualitative differences between the linear theory and the present work are discussed.

Subject headings: galaxies: individual (M33, NGC 2841) — galaxies: internal motions — galaxies: structure

1. INTRODUCTION

The origin of galactic disk warps, a common phenomenon in the universe, is uncertain. Mechanisms involving tidal disturbances from nearby galaxies, intergalactic magnetic fields, or steady state oscillations of the disk have all been put forward as explanations, but unfortunately at present there is no clear consensus on which of them to choose. A promising model is one in which the warp represents an equilibrium configuration of a disk "misaligned" inside a flattened halo, with the disk precessing about the halo axis of symmetry. Models of this kind have been calculated with a linearized theory, in which it is possible to view a warped equilibrium as a standing, neutrally stable bending wave in the disk (Sparke & Casertano 1988).

The detailed investigation by these authors is built on a chain of work spanning more than 20 years, starting from Lynden-Bell's (1965) suggestion that a stellar disk might have a bending mode analogous to the wobble of a spinning penny that is thrown into the air. Hunter & Toomre (1969) showed with a WKB analysis that most isolated stellar disk models do not in fact admit such a normal bending mode, however: unless the stellar density drops quite fast at the outer edge of the disk, bending waves traveling outwards will not be reflected back towards the centre of the disk, and hence cannot form a standing wave. Later, when the evidence for massive halos surrounding galactic disks grew, it was realized that these might hold the key to long-lasting warps. Thus, Tubbs & Sanders (1980) showed that a massive spherical halo can slow down the differential precession (which tends to wind up and destroy warps) down to Gyr time scales, and Petrou (1980) showed that suitably adjusting the halo flattening as a function of radius can remove the winding altogether. In work presented at IAU 100, Toomre (1983) and Dekel & Shlosman (1983) considered self-gravitating disks embedded in a flattened potential (as provided by a massive distant ring, or by a flattened massive halo), and demonstrated that they *can* have a warped mode, without the strong restrictions on the sharpness of the outer edges of isolated disks found by Hunter & Toomre (1969). Sparke (1984) and Sparke & Casertano (1988) then investigated such models in detail, showing that they can look quite realistic, and discussed the criteria for existence of bending modes.

Maintaining a warp in a galactic disk requires the disk to experience some external torque: without it, the warp would wind up into a spiral bending disturbance and disappear, Kahn & Woltjer (1959). Because the disk is spinning, this torque will cause a rigid precession of the whole warped disk. Since warps appear to be very common (in a sample of 20 well-studied, nearby edge-on disk galaxies Bosma 1991 finds the H I disks of four galaxies unwarped, and 12 others "clearly" warped; Sánchez-Saavedra, Battener, & Florido 1990 see warps in optical photographs of 49% of a sample of 86 disk galaxies), they have to be excited regularly if they are not long-lived features.

Many warped disks have no visible nearby galaxy that may be accused of having disturbed it in the recent past. In these galaxies, at least, there may not have been any interaction that could have set up a warp in the last several Gyr (assuming that there are no large "dark" galaxies roaming through space). Moreover, in most warped galaxies the line of nodes is quite straight, at least out to a certain radius (Briggs 1990), indicating that differential precession has not been severe in the inner parts. This implies a long lifetime for the warp. If the only forces of relevance to the disk are gravitational, this then requires torques acting on the disk to halt, or at least slow down, the differential precession.

Recently an intriguing correlation between the direction of the warps of disk galaxies has been found in a sample of galaxies covering a large part of the sky (Battaner, Florida, & Sánchez-Saavedra 1990). Such large coherence in warp directions cannot be explained in terms of the internal structure of galaxies, and, if true, argues for a larger-scale cause such as intergalactic magnetic fields. In this paper we will not investigate this nevertheless very interesting suggestion further.

Another possible model for warps is discussed by Ostriker & Binney (1989). They argue that realignment of the angular momentum of a galaxy on a Hubble time scale is plausible in current models of secondary infall onto extant galaxies. While the galaxy realigns, it warps, providing a nice explanation for the ubiquity of warps which does not require flattening of dark halos.

In this paper, we concentrate on the possibility that the torques responsible for the warp are provided by a flattened dark halo whose plane of symmetry is misaligned with the disk plane. In particular, we extend the work of Sparke & Casertano (1988, hereafter SC) by performing calculations which are fully (nonlinearly) developed in the misalignment angles between the various parts of the disk and the halo equator plane.

There is some evidence that such misalignments are readily generated when disk galaxies form. Simulations by Katz (1990, private communication) show that inclusion of gasdynamical processes in calculations of primary collapse of density perturbations naturally leads to flattened halos with misaligned disks, since the gas forms quite a lumpy structure which can exchange angular momentum with the halo before finally settling into a disk. He reports typical misalignments of 30° and halo flattenings of 0.1 in the potential. Observations of the polar ring in NGC 4650A (Sackett & Sparke 1990) strongly suggest that this galaxy is embedded in a halo which is flattened like an E3–E7 galaxy, providing some confirmation for these models.

Dubinski & Carlberg (1990) simulated the dissipationless collapse of cold dark matter halos with a high-resolution N-body simulation, in the absence of any gas. They find that such halos are strongly flattened, quite triaxial, and (within the resolution of the calculation, which was 1.4 kpc) essentially coreless. It is not clear how the process of disk formation would affect the halo core, although it seems unlikely that, if the disks are very massive, the resulting potential would still have a core radius small compared to the disk scale length. If the bulk of the disk gas is accreted after the halo has formed, then strong misalignments might be expected to occur only in halos with quite spherical cores, otherwise the gas would quickly find the halo plane and only small misalignments could result. In any case, these cosmological simulations suggest that models such as will be considered here may be relevant to the structure of warped disks.

Section 2 of this paper covers the basic equations, including some discussion of how far the linear approximations can be considered valid; § 3 contains discussion of some warp models of truncated exponential disks in flattened halos, and § 4 presents comparison of models with Begeman's (1987) observations of the disk galaxy NGC 2841, and the data of Reakes & Newton (1978) on M33. Section 5 contains the conclusions.

2. DETAILS OF THE CALCULATION

The galactic disk will be taken to be completely cold, i.e., all stars will be assumed to move along closed orbits. It is expected that small radial excursions of stars would not grossly affect the structure of the disk. Numerical calculations of closed orbits in the potential of a uniformly precessing disk have shown that in the precessing frame the closed orbits remain remarkably round, as is the case for the so-called "anomolous" orbit families (which do not lie in any of the planes of symmetry of the potential) in tumbling elliptical galaxies investigated by Heisler, Merritt, & Schwartzschild (1982). The same conclusion was reached by Zachilas & Petrou (1988, private communication), who studied orbits in the potential of a flattened halo and a tilted disk fixed inside it. It is therefore justifiable to follow the methods of Toomre (1983) and Sparke (1984), and calculate the dynamics of the disk in terms of a rigid-ring model, each ring representing a set of stars uniformly spread around their orbit. Further justification is provided by the work of Ostriker & Binney (1989), who demonstrate that to leading order in the precession rate of the potential rigid rings and stellar orbits respond identically.

The potential Φ_H of the flattened halo embedding the disk will be taken to be of the form

$$\Phi_H(R, z) = \tfrac{1}{2}V_c^2 \ln(R_H^2 + R^2 + z^2/q^2) \tag{1}$$

in cylindrical polar coordinates (R, ϕ, z). The flattening ϵ of the potential is given by $q = 1 - \epsilon$, V_c is the asymptotic circular speed on the equator of the halo, and R_H is the halo core radius. The potential energy $V_{H,i}$ of a massive ring of radius R_i and mass m_i, inclined at angle θ_i to the halo plane is given by

$$V_{H,i} = \int_0^{2\pi} \frac{m_i\, d\psi}{2\pi} \Phi_H[(R_i^2 - R_i^2 \sin^2\theta_i \sin^2\psi)^{1/2}, R_i \sin\theta_i \sin\psi]$$

$$= m_i V_c^2 \ln(1 + \sqrt{1 + a_i \sin^2\theta_i}) + \text{constant}\,, \tag{2}$$

where

$$a_i = \frac{R_i^2(q^{-2} - 1)}{R_H^2 + R_i^2}\,. \tag{3}$$

The torque $L_{H,i}$ exerted by the halo on the ring is therefore

$$L_{H,i} = -\frac{\partial V_{H,i}}{\partial \theta_i}$$

$$= -\frac{1}{2} m_i V_c^2 \sin 2\theta_i \frac{a_i}{\sqrt{1 + a_i \sin^2\theta_i}\,(1 + \sqrt{1 + a_i \sin^2\theta_i})}\,. \tag{4}$$

Note that in such a halo the torque felt by a ring at a given inclination approaches a constant value beyond a few core radii. Also, the flattening and the radial mass profile are described by the same core radius, something which need not apply to real galactic halos. This last point will be discussed further below. The halo potentials used in the calculations of SC have a similar coupling.

In an equilibrium situation, the disk uniformly precesses about the halo axis—there is therefore some advantage in moving to the reference frame which rotates about the halo symmetry axis at the disk's precession frequency Ω. In this frame the figures of the disk and halo appear stationary, although the different rings that make up the disk still spin about their individual axes. If the ring i spins at a speed V_i (in the positive sense as viewed down the halo's symmetry axis), the coriolis force associated with the frame rotation subjects it to a fictitious torque $L_{C,i}$ given by

$$L_{C,i} = -m_i R_i V_i \Omega \sin\theta_i\,. \tag{5}$$

As will be seen below, and is also apparent from the calculations of SC, the typical pattern speeds are small compared to the angular frequencies even in the outer parts of the disk, so that centrifugal terms can be ignored.

If the disk is to be stationary (in the rotating frame) then the total torque it experiences from halo and Coriolis forces must be zero. It then follows from equations (4) and (5) that, provided $|\theta_i| < \pi/2$, Ω is negative in an oblate ($\epsilon > 0$) halo (i.e., the disk precession is retrograde), while the opposite is true in a prolate halo. It is also important to note that the coriolis and halo torques have different angular dependences: the approximate $\sin 2\theta$ dependence of the halo torque combined with the fact that $L_C \propto \sin\theta$ mean that, to keep these two in balance,

approximately $\Omega \propto \cos\theta$, where θ here is some mean misalignment of the (warped) disk plane with the halo. Thus, increasing the tilt of the midplane of the disk reduces the precession frequency of the entire pattern.

Finally, to treat the self-interaction of the disk, it is necessary to calculate the torque between two concentric, mutually inclined rings. The mutual potential energy V_{ij} between two rings i and j inclined to each other by an angle δ is (see e.g., Ostriker & Binney 1989)

$$V_{ij} = -\frac{Gm_i m_j}{2\pi^2\sqrt{R_i R_j}} \int_0^{2\pi} d\psi\, kK(k)\,, \tag{6}$$

with

$$k^2 = \frac{4R_i R_j\sqrt{1-\sin^2\psi\,\sin^2\delta}}{R_i^2 + R_j^2 + 2R_i R_j\sqrt{1-\sin^2\psi\,\sin^2\delta}}\,, \tag{7}$$

and $K(k)$ is the complete elliptical integral of the first kind. If the line of nodes in the halo equatorial plane is the same for both rings, then $\delta = \theta_i - \theta_j$. In this case, the torque L_{ij} on ring i from ring j is

$$L_{ij} = -\frac{\partial V_{ij}}{\partial\theta_i} = -\frac{Gm_i m_j}{16\pi^2\sqrt{R_i R_j}} \sin 2\delta \int_0^{2\pi} d\psi\, \frac{k(2-k^2)E(k)\sin^2\psi}{(1-k^2)(1-\sin^2\delta\sin^2\psi)}\,. \tag{8}$$

$E(k)$ is the complete elliptical integral of the second kind, which satisfies $d[kK(k)]/dk = E(k)/(1-k^2)$. In general, this integral cannot be simplified further, so that numerical approximations to it were needed for the calculations described below.

Figure 1 shows the torque L_{ij}, in arbitrary units, as a function of mutual inclination for two rings with radii in the ratio r:1. Evidently, the torque between two rings of similar radii peaks at very small angles—approximately, for r close to 1, the maximum occurs at the angle $\theta_{\max} \simeq 1.2|r-1|$ radians. So in a disk in which the warp angle θ increases at a rate more steeply than $|d\theta/dR| = 1.2/R$, the self-gravity of the disk (at least the short-range portion thereof) is stretched to the limit. Therefore, a simple scaling up of $\theta(R)$ by a constant factor, which in the linear theory yields a new solution, produces a configuration in which the torques required for equilibrium cannot be provided by the local disk gravity, and at the very least the shape of the equilibrium warp of this amplitude must change significantly from the linear mode shape. In disks in which the linear mode has a very steeply changing $\theta(R)$ (mostly extreme fast warps, see below), this maximum slope can be reached at small amplitudes, where the halo and coriolis torques are still well-described by linear approximations. In some cases, therefore, it is the slope of the warp which is the limiting factor for the linear calculations, rather than the amplitude of the misalignment between the disk and the halo planes. This is also evident from the calculations for disks with massless outer regions which will be considered in § 3 along with full models of truncated exponential disk galaxies.

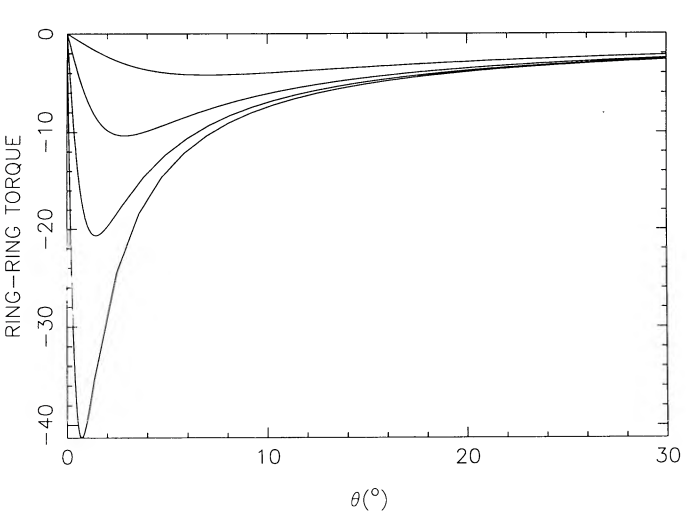

FIG. 1.—Torques between two nearby concentric rings of radius ratio r:1, as a function of their mutual tilt angle θ. Curves for (*top to bottom*) $r = 1.1$, 1.04, 1.02, and 1.01 are shown. Note the limited θ-extent of the linear regime as $r \to 1$.

In their calculations, SC identified two kinds of warp. The first, so-called type-I warps, occur in halos with small core radii. In these disks the massive inner part, hardly influenced by the weak torques exerted by the tenuous outer gas, feels a strong halo torque. This sets a high precession speed for the warp mode. The outer parts are forced to precess at the same speed if the galaxy is in a warp mode, so, because of their larger radius, they feel a larger Coriolis torque (assuming that the circular speed is approximately constant). The specific halo torque, on the other hand, is basically the same for inner and outer rings if the halo core radius is small. Thus in the outer regions the Coriolis term will dominate, lifting the rings there further out of the halo plane. Thus type-I warps (which I will call "fast") bend steadily away from the halo plane.

The type-II, or "slow" warps occur in disks embedded in halos with large core radii. In such disks, the halo torque rises through most of the massive regions of the disk, rather than quickly reaching a maximum as in the fast warps. Consequently, the inner disk does not feel as strong a torque, and the galaxy precesses more slowly. In the outer disk, the halo torque can therefore dominate the Coriolis term, forcing the disk to bend down towards the halo plane. As R increases, though, eventually the Coriolis term will be largest once again and, if the disk is sufficiently large, its edge will turn up, away from the halo equator.

Toomre (1990, private communication) has pointed out that halos are unlikely to remain inert in the presence of a precessing stellar disk, but that instead dynamical friction will try to align these two components. This effect is more serious for more massive stellar disks and for higher precession speeds (fast warps).

3. WARPING OF A TRUNCATED EXPONENTIAL DISK

SC considered the linear bending modes of a truncated exponential disk in a flattened halo. The calculations presented in this section extend their results into the nonlinear regime, by requiring equilibrium of all the disk's rings under the combination of the torques exerted by the other rings, the halo and the Coriolis force. As in the linear theory, in equilibrium all rings are required to have the same line of nodes in the halo equatorial plane. If the disk is modeled with N rings, there are N equations (the torque on each ring must be zero) to be satisfied, with $N+1$ unknowns, namely the angles of tip θ_i of all rings, and the overall precession speed Ω. After fixing the tilt $\theta_{\rm in}$ of the innermost ring, the remainder of the unknowns can

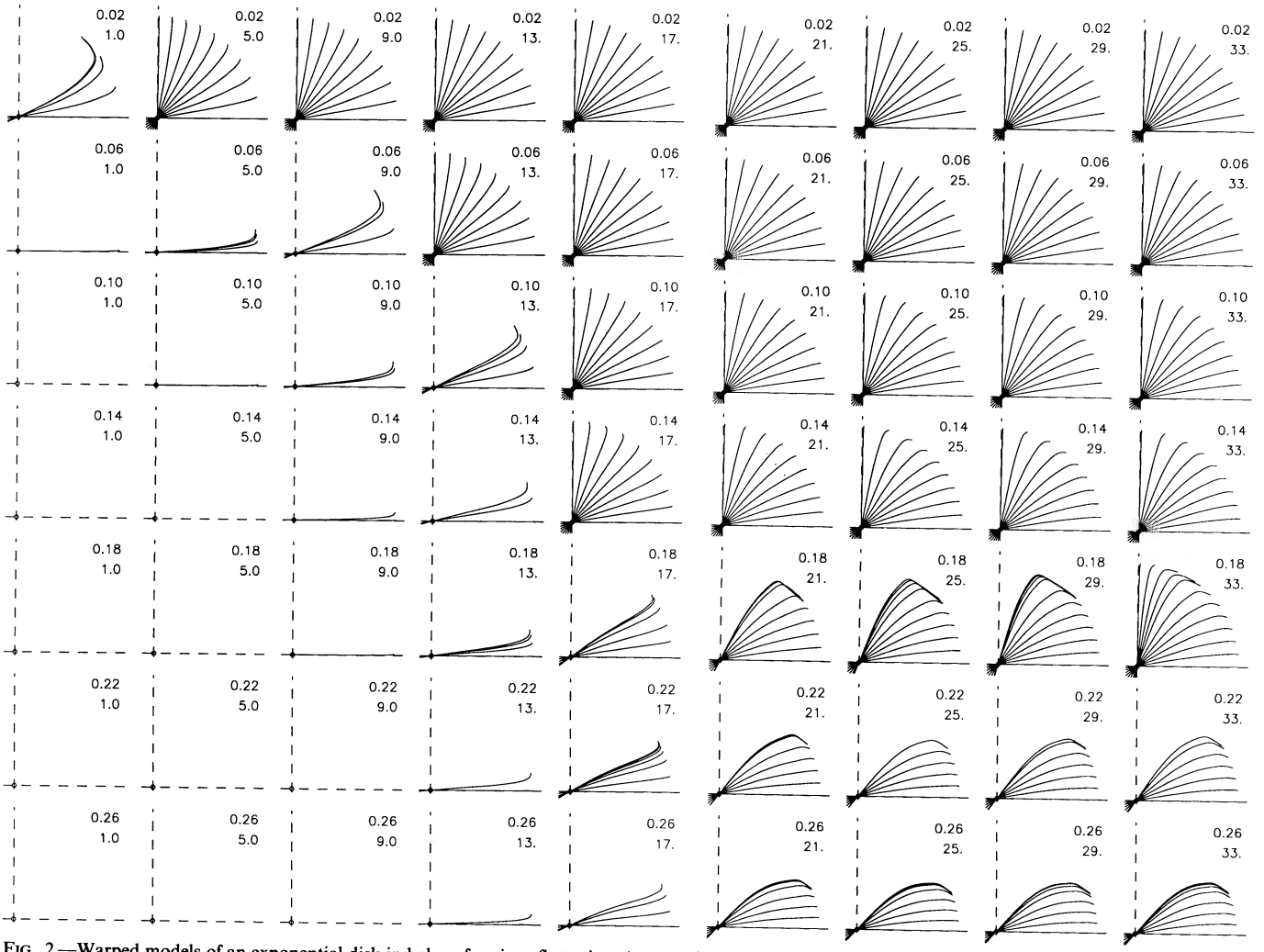

FIG. 2.—Warped models of an exponential disk in halos of various flattenings (see text for details). Halo flattening increases from top to bottom (ϵ = 0.02, 0.06, 0.10, 0.14, 0.18, 0.22, and 0.26 are shown), and the core radius increases from left to right (R_H = 1, 5, 9, 13, 17, 21, 25, 29, and 33 kpc).

be solved for, yielding equilibrium models at a range of inclinations of the inner ring. (Note that these models are still not fully self-consistent, since they replace the closed orbits with nearby circular paths. For very inclined rings in a very flattened halo, this assumption becomes less appropriate, although it is probably not a serious concern in practice.) All calculations shown in this paper have $N = 50$–60 rings, concentrated towards the disk's outer edge where warping was taking place. Experiments with models having 100 rings gave essentially the same answers.

In Figure 2, the resulting equilibria of one particular truncated exponential disk in a variety of dark halo potentials are shown. The disk has a mass of $7 \times 10^{10}\ M_\odot$, with a radially exponential surface density profile of scale length 4 kpc out to the truncation radius of 20 kpc, beyond which it is smoothly tapered to zero at a radius of 24 kpc following the prescription of SC. The dark halo has an asymptotic circular speed V_c of 200 km s^{-1}. The different panels of Figure 2 show the possible equilibrium configurations for the disk, in halos of different flattening ϵ and core radius R_H. In all cases the galaxy is viewed from the halo's equatorial plane, edge-on to all the rings. The halo plane is horizontal. As already noted by SC, misalignment does not imply warping: quite a few of these models show little or no warping, even though they do not lie in the symmetry plane and the halo is quite flattened.

In any given halo model, equilibrium solutions were sought for successively increased values of $\theta_{\rm in}$, using multidimensional Newton-Raphson iteration. Each solution found was rescaled (linearly with $\theta_{\rm in}$ for the individual tilt angles θ_i, and with $\cos\theta_{\rm in}$ for the precession frequency Ω) for use as a starting approximation to the equilibrium with the next $\theta_{\rm in}$. Usually, $\theta_{\rm in}$ could be increased in steps of 10° and a new solution obtained; however, near regions of parameter space where the equilibrium solutions change rapidly or disappear, smaller steps needed to be taken. In practice, when the Newton-Raphson algorithm failed to converge the increment in $\theta_{\rm in}$ was halved and a smaller step taken. The first solution was obtained starting from the linear mode solution described by SC (which in turn was calculated by starting with a constant tilt solution and the precession frequency $\Omega_{\rm lin}$ given by SC in the low-ϵ limit). The search ended when the step size dropped below 1°, or 90° was reached. All solutions thus calculated are displayed in the figure.

Empty panels in Figure 2 represent those halos in which there are no isolated linear modes for this particular disk (see SC): in these cases the solution to the linearized equations does

not converge as the number of rings is increased, but remains irregular and highly oscillatory near the disk edge. At fixed ellipticity, there may be a minimum core radius for which a linear mode can exist; this arises because halos with smaller core radii cause faster precession rates, which can eventually be so fast that bending waves of this frequency propagate to the disk edge rather than reflect and form a standing wave. With halo core radii slightly above the minimum (e.g., the $\epsilon = 0.10$, $R_H = 9$ kpc model) a (highly warped) linear mode does indeed exist. However, once the nonlinear terms are included in the torque calculations, it is found that the maximum amplitude of the tilt, set by θ_{in}, is quite small. For larger amplitudes, no equilibrium exists, as the disk's self-gravity is unable to maintain a uniform precession of the entire disk.

3.1. *A Simple Example: Test Particles Around a Massive, Precessing Disk*

Many of the features of these results can be demonstrated qualitatively with a very simple model. Consider a disk galaxy consisting of two regions: a massive flat inner disk, and a series of massless "test rings" in the outer regions. The whole system is embedded in a flattened halo as above, with core radius well inside the inner disk's edge. If ϵ is not too large, then $L_H(R, \theta) \simeq -\frac{1}{2}\epsilon V_c^2 \sin 2\theta$ in the region of the test rings. If the plane of the massive disk is now tilted at an angle θ_d, it will precess about the halo axis at a frequency of $\Omega_0 \cos \theta_d$, with the constant Ω_0 determined by the rotation speed of the disk and by the halo potential. Such a galaxy can have a linear mode of sorts, although the absence of self-gravity in the outer parts precludes propagation of energy by bending waves, so that these waves are kinematic only.

This model has the advantage that the outer rings only interact with the inner disk and the halo, and their equilibrium positions do not depend on each other, making the system much easier to solve. We now study such kinematic modes at radii large compared to that of the massive inner disk. This should at least be a reasonable approximation to galaxies, because typically the inner (optical) disk is virtually unwarped, while the gas disk extending to larger radii is essentially massless for present purposes.

The linearized specific torques due to disk, halo, and Coriolis force on a ring of radius R and tilt θ are $-f^2(R)(\theta - \theta_d)$, $-\epsilon V_c^2 \theta$ and $-RV\Omega_0 \theta$, which sum to zero provided that

$$\theta(R) = \frac{f^2(R)\theta_d}{V\Omega_0 R + \epsilon V_c^2 + f^2(R)} . \quad (9)$$

Here $f^2(R)$, which governs the torque due to the inner disk, is a positive, decreasing function which tends to 0 as $R \to \infty$, and ϵ and $\Omega_0 V$ have opposite signs. $\theta(R)$ is continuous and the solution stable provided that the denominator is positive, which it is for all R less than some radius R_{crit}. At this radius, there is a resonance between free vertical oscillations about a circular orbit in the halo plane (i.e., precession of the ring) and the periodic forcing due to the precessing massive disk. Thus, in spite of the lack of self-gravity which could propagate bending waves out through the distant gas, energy is being deposited at a resonance of the disk's precession. This is analogous to the distinction in the theory of spiral structure between kinematic response to a potential perturbation in a planar disk, which requires a forcing to be put in by hand, and tightly wound density waves which can reflect and form standing waves. The latter can be described by the WKBJ approximation (Hunter 1969), being driven by local gravity; the former cannot.

In the notation of SC, the outer gas can join in the disk's precession out to the radius R_{crit} at which its free precession frequency $-(\mu_{tot}^2 - \Omega^2)/2\Omega$ becomes equal to the disk's precession speed. Here μ_{tot} is the vertical oscillation frequency of the gas in the plane of the halo, and Ω the circular frequency. As long as the galaxy has radius below R_{crit}, it has a continuous linear bending mode. Although self-gravity will undoubtedly affect the detailed structure of a warped disk, in a fast warp the strongest bending does occur in the outer regions where the density tapers to zero, and a large part of the torques influencing the outer rings is nonlocal in that case. Thus it appears that the outer vertical resonance is at least as important an ingredient as the local self-gravity for understanding the structure of warped disks. This picture is supported by the calculations set out above for the truncated exponential disk.

To investigate the nonlinear equilibria, consider the quadrupole approximation, in which the disk exerts a torque on a ring at radius R of $-\frac{1}{2}f^2(R) \sin 2(\theta - \theta_d)$; $f^2 \propto R^{-3}$. Then putting the total specific torque on such a ring to zero yields

$$L(R, \theta) = -\tfrac{1}{2}f^2(R) \sin 2(\theta - \theta_d) - \tfrac{1}{2}\epsilon V_c^2 \sin 2\theta - RV\Omega_0 \sin \theta \cos \theta_d = 0 . \quad (10)$$

Figure 3 shows, for different values of θ_d, the curves $L(R, \theta) = 0$, for the specific case $\Omega_0 = -0.1$, $\epsilon = 0.1$, $f^2 = R^{-3}$, and $V_c = V(R) = 1$. This galaxy displays a fast warp. At a given value of θ_d, there is a maximum radius $R_{crit}(\theta_d)$ for which stable solutions exist. At that point, a stable and an unstable (not shown) root of equation (10) merge. The dotted curves show the result of the linearized equation (9). There clearly are marked differences between the linearized and full results, especially in the regions where the disk is most strongly warped, which is where observations would carry most weight. The maximum radius of a warped configuration is largest at small amplitudes (which the linear theory applies to), and decreases as the inner disk's inclination grows. (In this example, for an inner disk tilt of only 10° the maximum radius is $0.9R_{crit}$, shrinking further to $0.8R_{crit}$ at a tilt of 30°. Conversely, a galaxy with a radius 90% of R_{crit} can only be in equilibrium if its inner disk is tilted by less then 10°.) Thus, in galaxies with very steep warps in the outer parts, the linear theory should be used with caution, particularly in the region of steepest warping. This is true even at modest angles of inclination.

If the precession frequency is reduced, slow warps result: the outer regions of the galaxy lie closer to the halo plane than the inner disk. Figure 4 illustrates one such case, for $\Omega_0 = -0.03$. There still is an upper limit on the size of the galaxy, and rings just inside this resonance radius bend up away from the plane again. Note that once the tilt of the inner disk exceeds 50°, the maximum outer radius starts to increase again; this is due to the reduction in precession speed of the galaxy with θ_d, and is a general feature of this model. For slow warps, these "kinematic" models are likely to be a worse approximation to reality than for the fast warp shown above, since the steepest portion of the warp now occurs further in towards the centre of the galaxy, and hence self-gravity cannot be ignored safely. Nonetheless, they reproduce the general features of the exponential disk calculations. As discussed in § 2, very steep warps cannot be supported by disk self-gravity, but rather are a balance primarily between coriolis and halo torques.

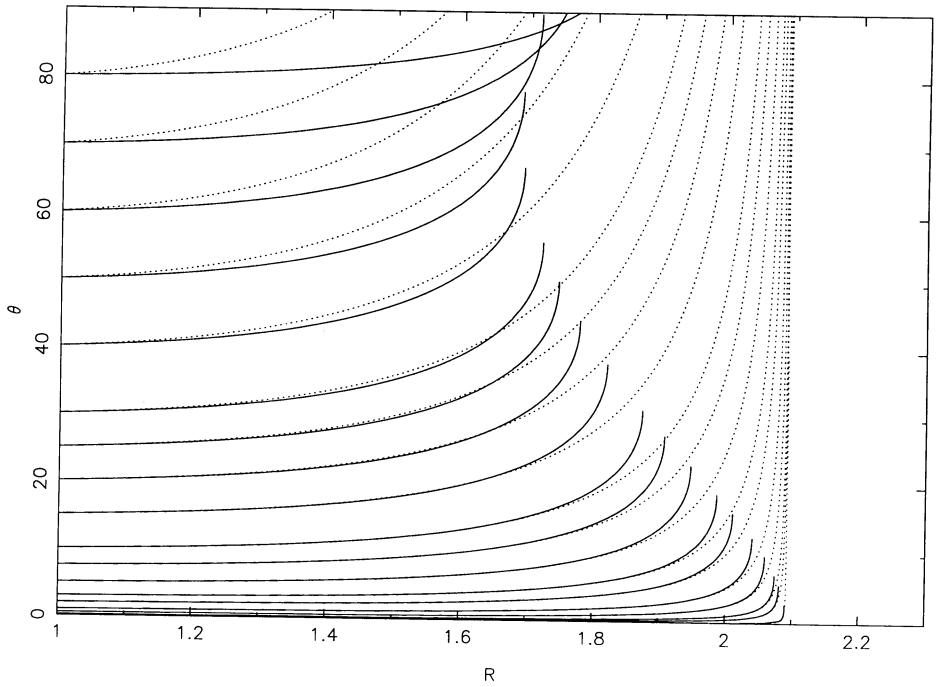

FIG. 3.—Approximate warp models of a solid disk surrounded by massless test-rings. Solid lines show results of the nonlinear calculation, dotted lines show results of the linear one. Models have high precession speed, giving a fast warp.

Inclusion of the 4th-order harmonics in the potential of the disk does not change these results qualitatively, although in detail the equilibrium solutions are different. The effect of the extra terms is to confine the warp more strongly to the disk plane just outside the edge of the disk, resulting in a steepening of the warp shape further out. However, this simple model is still useful as a description of what happens in the full problem.

[For an exponential disk of scale length h and cutoff at $4h$ it can be shown that at small amplitudes the precession frequency Ω_0 in a coreless halo is $\simeq -\epsilon V_c/(1.5h)$, which places the resonance radius $R_{\rm crit}$ calculated with eq. (9) very close to, or even inside, the edge of the disk; in these regions the quadrupole approximation is quite insufficient anyway.]

Similar calculations to these, but ignoring the effects of disk

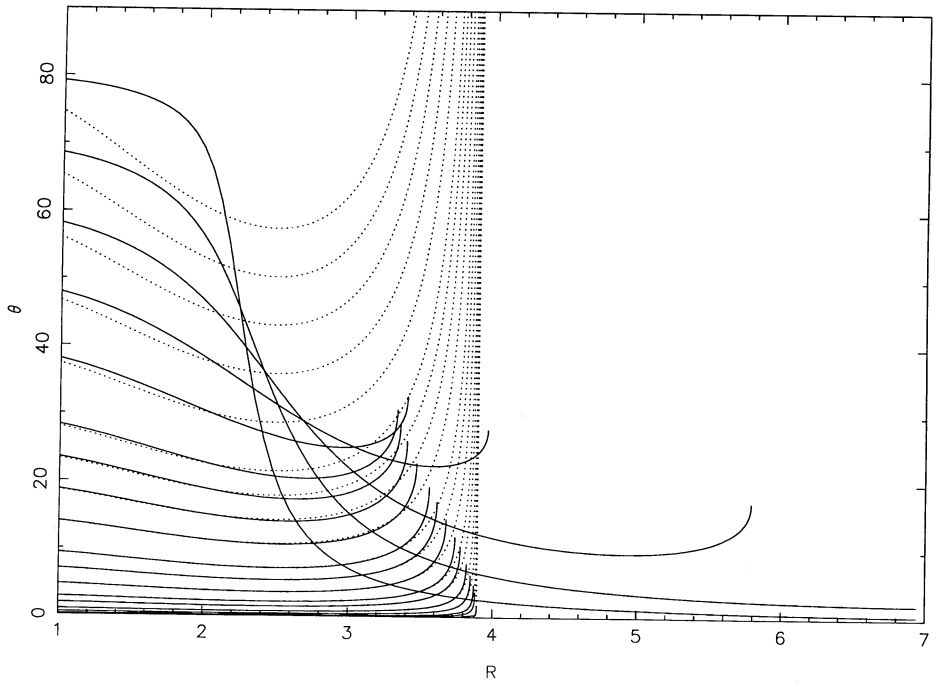

FIG. 4.—Same as Fig. 3, but with a reduced precession speed to yield a slow warp

precession, were carried out by Zachilas & Petrou (1988, private communication).

3.2. *A Model-Independent Approach*

It is possible to treat the observations in a more model-independent way than the method that has been used in this paper so far. Rather than fitting the observed warp parameters of a disk with model grids such as the ones constructed and shown here, and hence deducing the likely parameters of flattened halos surrounding these disks, instead one can start from the observed radial dependence of the disk surface density and inclination to calculate the self-torque of the disk, as a function of radius. Requiring equilibrium (apart from a steady precession of the pattern about some axis perpendicular to the line of nodes) then allows the radial run of the external (halo) torque on the disk to be derived as a function of the precession frequency and the angle $\theta_{\rm in}$ between the inner disk plane and the precession axis.

Writing the *specific* torque L_i/m_i as $\mathscr{L}_i$, and keeping the same subscripts as used in § 2, the equilibrium condition on ring i can be written as

$$\mathscr{L}_{H,i} + \mathscr{L}_{C,i} + \sum_{j \neq i} \mathscr{L}_{ij} = 0 \,. \tag{11}$$

The last term can be calculated from the observations after assuming a mass-to-light ratio, and the second depends on the inclination ($\theta_{\rm in} + \theta_i$) and circular rotation speed of ring i and on the global precession frequency Ω. Thus, assumption of $\theta_{\rm in}$ and Ω allows the halo torques experienced by the disk to be calculated. This shows that in principle there are equilibrium solutions at any precession rate and inclination, although in such cases the torques required of the halo may well be impossible to produce with a realistic (or even realizable) mass distribution. For example, at zero precession, the specific halo torque on each ring is simply equal and opposite to the specific disk torque: since the latter must change sign at some radius within the disk, (otherwise the disk exerts a net torque on itself), so must the former. Then choosing $\theta_{\rm in}$ such that the disk does not intersect the halo plane or meridian results in a halo potential whose flattening changes sign. At the other extreme, it is possible to pick such a large precession speed that the counterbalancing halo torques cannot be realized.

Incidentally, the fact that the total self-torque of the disk $\sum_i \sum_{j \neq i} m_i \mathscr{L}_{ij}$ is zero leads to a simple derivation of the SC formula for the precession frequency at small amplitudes, since it implies that

$$\sum_i m_i (\mathscr{L}_{H,i} + \mathscr{L}_{C,i}) = 0 \;; \tag{12}$$

hence the precession frequency is given by

$$\Omega_0 = \left(\sum_i m_i \mathscr{L}_{H,i} \right) \Big/ \left(\sum_i m_i R_i V_i \sin \theta_i \right) . \tag{13}$$

Writing Ω_H and μ_H for the circular and vertical oscillation frequencies in the halo potential at $z = 0$, and Ω for the circular frequency of the total potential, it is easy to show that $\mathscr{L}_{H,i}(\theta) = (\Omega_H^2 - \mu_H^2) R_i^2 \theta / 2$ to $O(\theta)$. Thus, linearizing the torques in equation (13) in θ, substituting the leading-order (in ϵ) solution $\theta_i = \theta_{\rm in}$ and proceeding to the continuum limit $m_i = 2\pi R \Sigma(R) dR$, for disk surface density Σ, then yields

$$\Omega_0 = \left(\int_0^\infty dR R^3 \Sigma (\Omega_H^2 - \mu_H^2)/2 \right) \Big/ \left(\int_0^\infty dR R^3 \Sigma \Omega \right) , \tag{14}$$

which is identical to the SC formula (21).

It is unnecessary to know the exact expression of the torque between two massive rings for this calculation, only Newton's third law (which translates into the self-adjointedness of the SC operator $\mathscr{J}$, their eq. [16]) which implies that the disk cannot exert a total torque on itself.

4. SPECIFIC GALAXIES

4.1. *The Warped Disk Galaxy NGC 2841*

Assuming that warps are indeed caused by misaligned, flattened halos, to what extent can the observed warp of a galaxy constrain the shape and mass of the halo surrounding it? In this section, that question is investigated with the aid of the galaxy NGC 2841. This galaxy has optical photometry by Kent (1986) and Wevers (1984), and has been observed extensively in the 21 cm H I line by Bosma (1978) and Begeman (1987). Assuming a distance of 9 Mpc, the gas disk extends to at least 43 kpc from the centre, as compared to a luminous disk scale length of between 2 and 4 kpc. Begeman fitted a tilted-ring model to the observations, from which he derived the rotation curve of the gas, and deduced that the disk of this galaxy is warped by up to 16°. This is in good agreement with Bosma's (1978) earlier results. Apart from this distortion, the velocity field of the galaxy is symmetric, and in the plane of the inner disk the lines of nodes of all the rings are within 30° of each other (Briggs 1990). Thus this galaxy comes quite close to fitting the requirements for being in a steady state warp. Even if it is not quite in a steady state, the fact that the line of nodes is quite straight suggests that differential precession of the outer rings is at least substantially slowed down, which may imply that the galaxy is close to a steady state solution. The warp of NGC 2841 is displayed in Figure 5, both as $\theta(R)$ and as an edge-on projection of the disk plane.

As is well-known, the rotation curves of typical disk galaxies do not suffice to derive the mass-to-light ratio Υ_D of the disk if the core radius of the halo is unknown (van Albada & Sancisi 1986; Lake & Feinswog 1989). A maximum value can be derived, though, by requiring that there is no dark matter in the inner parts of the galaxy. For NGC 2841, this "maximum disk" value is about $\Upsilon_D = 11$, depending slightly on the bulge mass fitted, which would imply a halo core radius of 25 kpc. Lower mass to light ratios of the disk require lower halo core radii.

In the model for warped galaxies presented here, the observed warp depends on two parameters which do not affect the rotation curve of the disk (which is a measure of the monopole component of the potential): the halo flattening and the misalignment between the disk and halo planes. Thus, unless the warp shape and amplitude depend quite strongly on these quantities, the halo properties (assuming that the halo is of the form of eq. [1]) will not be well-constrained by the warp shapes, and conversely the observed shape of the warp will not be a very discriminatory test of the model.

A grid of warped equilibria for the disk of NGC 2841 is shown in Figure 6. Each panel corresponds to a different halo, and is analogous to a panel of Figure 2. The halo core radius was derived by fixing Υ_D, calculating the radial disk mass profile from the observed light and H I distributions, and fitting the rotation curve. This halo was then flattened to varying degrees. As can be seen, small core radii (which correspond to small disk masses in these models) do not allow an equilibrium warp, even at very moderate halo flattenings. The disk is forced to precess fastest in these halos, and its self-gravity is weakest; both these facts reduce its ability to maintain a coherent warp.

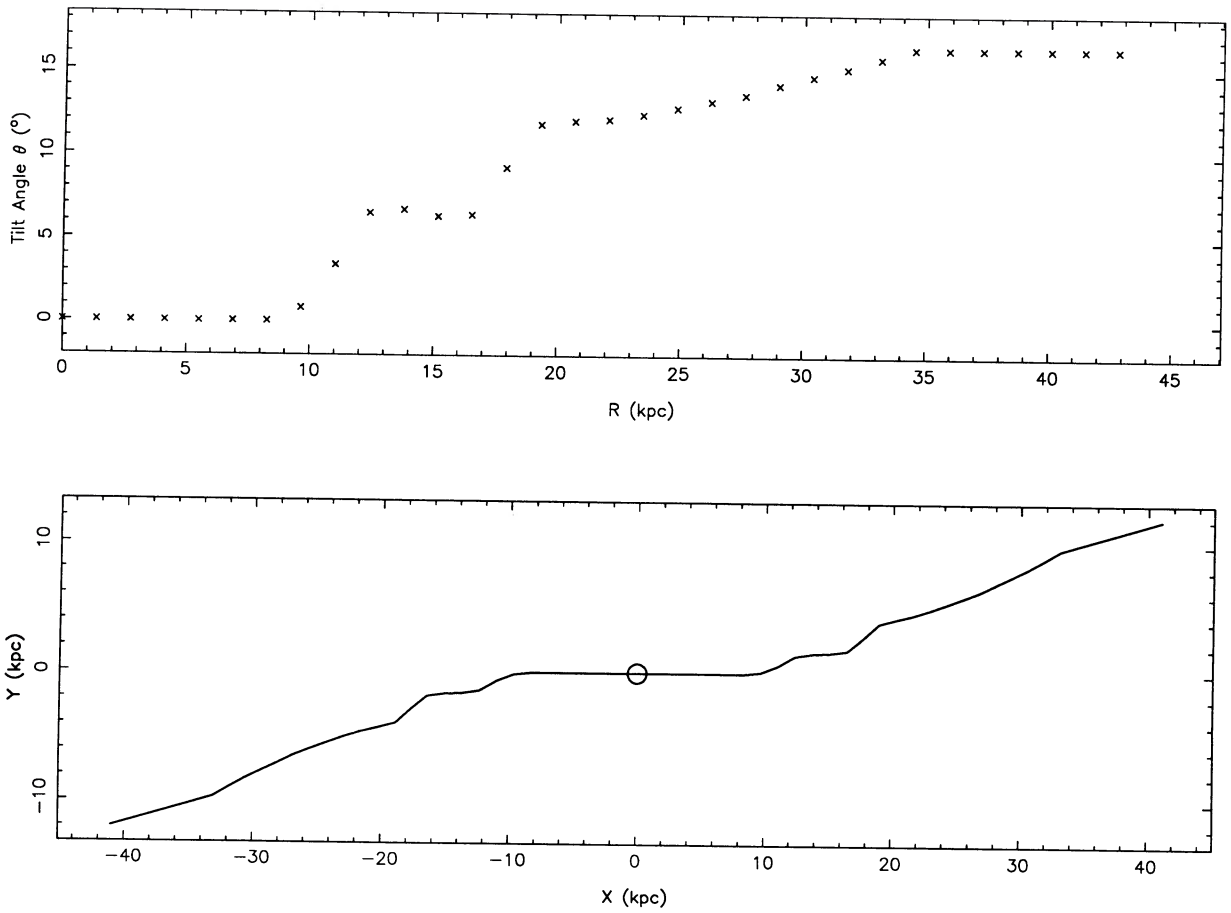

FIG. 5.—Observed warp of NGC 2841, from Begeman (1987). Angles are referred to the plane of the inner disk.

After interpolating these solutions over the inner disk tilt angle, the best-fit to the observed warp can be found for each panel. [Note that the orientation of the halo plane is not observed directly, so the fit has to be made to $\theta(R) - \theta(0)$.] The best overall fit is for an ($\epsilon = 0.22$, $R_H = 20.4$ kpc) halo, for $\Upsilon_D = 10$, which reproduces the observations with an rms scatter of 1°.5. Halos with $\epsilon > 0.13$ and $R_H > 20$ kpc ($\Upsilon_D > 10$) give reasonably good fits, with rms scatters below 3°. The flattening ϵ refers to the iso*potential* surfaces; the flattening of the density is about 3 times this value. Therefore within this grid of models the warp of NGC 2841 can be reproduced only in halos which are flatter than E4 or so, and whose core radius exceeds about 20 kpc. The halo and disk planes are inclined to each other by 18°.

The above models suggest, among other things, that the mass of the disk of NGC 2841 is close to its "maximum disk" value. This is an artifact of the halo parameterization that was chosen: since NGC 2841 is clearly a slow warp (the steepest portion of the $\theta(R)$ curve occurs far from the disk edge), this means that the specific halo torque needs to be rising throughout most of the disk (see discussion at the end of § 2), which in the halo models used here implies that the circular velocity due to the halo also rises slowly, which in turn necessitates a large disk mass to make up the rest of the radial acceleration in the inner regions of the galaxy. If instead this extra radial force is made up with spherically distributed matter, leaving the halo torques unaffected, a lower disk mass is allowed. Thus, removing the coupling between the core radius of the quadrupole and monopole components of the halo allows the warps to be fitted for lower disk masses too. This is shown in Figure 7, which is a grid of models of the NGC 2841 disk, with $\Upsilon_D = 3$. In these calculations a halo potential of the form

$$\Phi_H(R, z) = \frac{1}{2} V_c^2 \ln(R_H^2 + R^2 + z^2/q^2) + \frac{1}{2} V_c^2 \ln\left(\frac{R_c^2 + R^2 + z^2}{R_H^2 + R^2 + z^2}\right) \quad (15)$$

was used, with $R_c < R_H$. Effectively, this gives one halo core radius (R_c) for the rotation curve fit, and a different one (R_H) for the halo torques. It is readily verified with Poisson's equation that the density of such a halo is everywhere positive provided that $q > 2^{-1/2}$.

The best fit that can be obtained in this way occurs at $\epsilon = 0.22$, $R_H = 22$ kpc, misalignment angle 18°. It has a standard deviation of 0.9°, slightly better than the $\Upsilon_D = 10$ solution shown above. The region of acceptable solutions (standard deviation < 2°) is shown in Figure 8, and extends down to $\epsilon = 0.08$. Once we admit these more general halo models, the disk mass is very poorly constrained by the warp shape. In fact, not even a minimum can be set on the disk mass without assuming some form of halo potential, since from any equilibrium it is possible to construct a series of different ones with

the same rotation curve but arbitrarily small disk mass: to do this one reduces the potential of the flattened halo proportionally to M_D, and again makes up the extra circular speed required to fit the rotation curve with spherically distributed halo material. This reduces the specific torques due to halo and disk by the same amounts, and hence also the precession frequency of the entire pattern. Thus, the shape of a warped disk and its rotation curve cannot provide lower limits on the disk mass. Upper limits may arise, though, since if the above scaling is performed to a higher disk mass, spherically distributed halo matter needs to be removed which may lead to negative space densities. In other words, the required flattening of the halo potential may be unrealizable with a positive mass distribution. The galaxy M33, discussed below, may be an example of this.

Repeat calculations for the (unrealistically low) $\Upsilon_D = 1$ (not shown) do indeed show that the warp and rotation curve of NGC 2841 can be modeled well with a halo of the form of equation (15). The observations have an rms scatter of 0°.8 about the best fit in this grid of models, which occurs at an inclination of the inner disk plane of 20°. Since the disk is less coherent in these models, it requires less flattening in the halo to produce the observed warp: the minimum value of ϵ which gives a scatter below 3° is 0.05. If $\Upsilon_D = 1$ the core radius of the flattening must exceed 15 kpc to produce a slow warp which resembles the observations.

Irrespective of the disk mass, the misalignment between the halo and disk planes is about 20°–30° in the best-fit models. This is caused by the fact that the outer disk of NGC 2841 is quite flat, which can only be reproduced by making these orbits lie close to the halo equatorial plane. Thus this plane is almost fixed by the observations. In galaxies which do not have such flat outer regions, this will not apply in general. In the best-fit $\Upsilon_D = 3$ model, the precession period of the disk of NGC 2841 disk is 4.5 Gyr (this scales approximately inversely with the disk mass). The largest radius at which there are stable test-ring orbits is about 79 kpc, independent of Υ_D. This outermost ring is inclined 9°.5 to the halo equator. If more sensitive

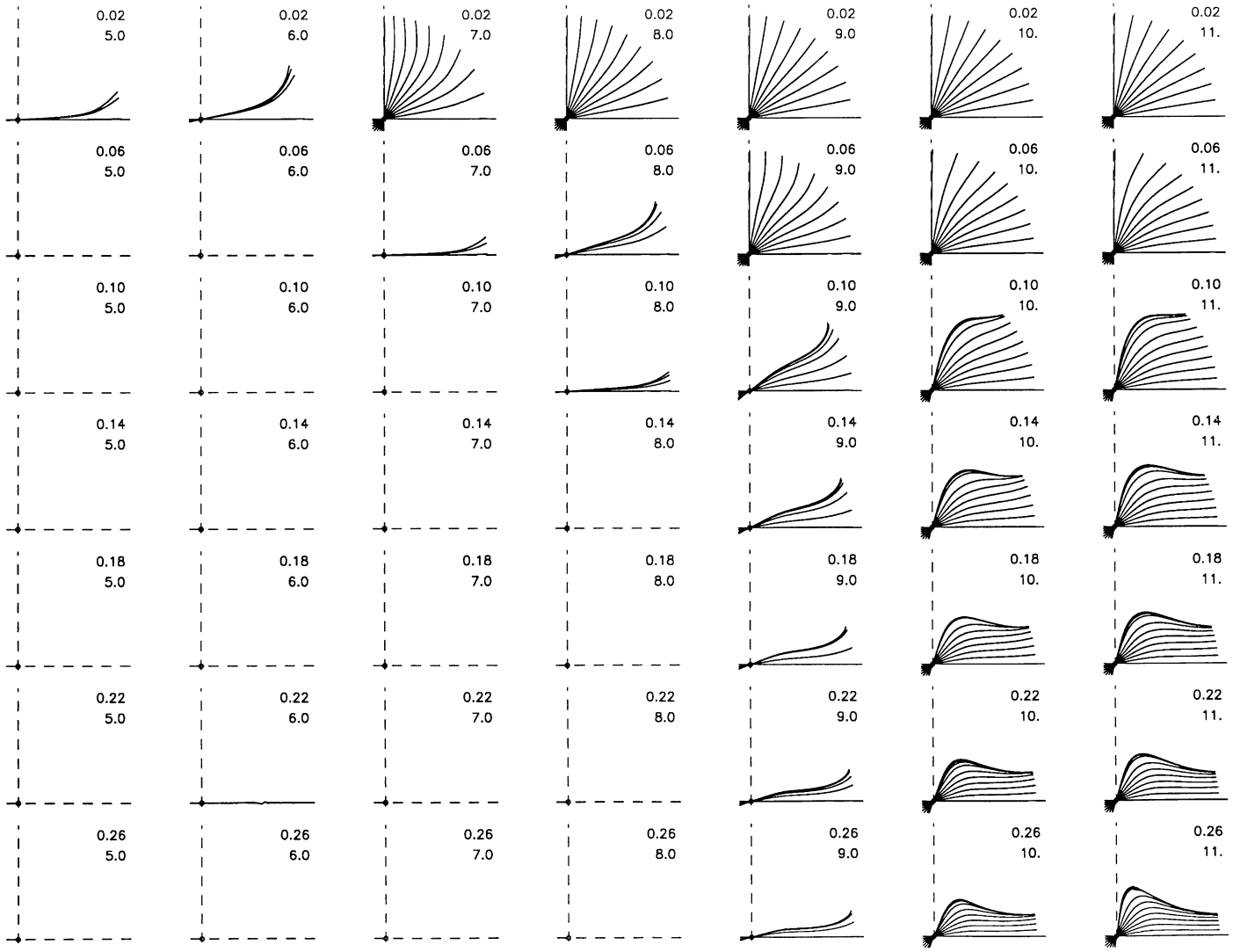

FIG. 6.—Equilibrium solutions for the disk of NGC 2841 in a variety of halos. The core radius is derived by fitting the rotation curve after assuming a value for the disk M/L ratio Υ_D. *Left to right*: $\Upsilon_D = 5, 6, 7, 8, 9, 10, 11$. The halo flattening increases downwards ($\epsilon = 0.02, 0.06, 0.10, 0.14, 0.18, 0.22$, and 0.26).

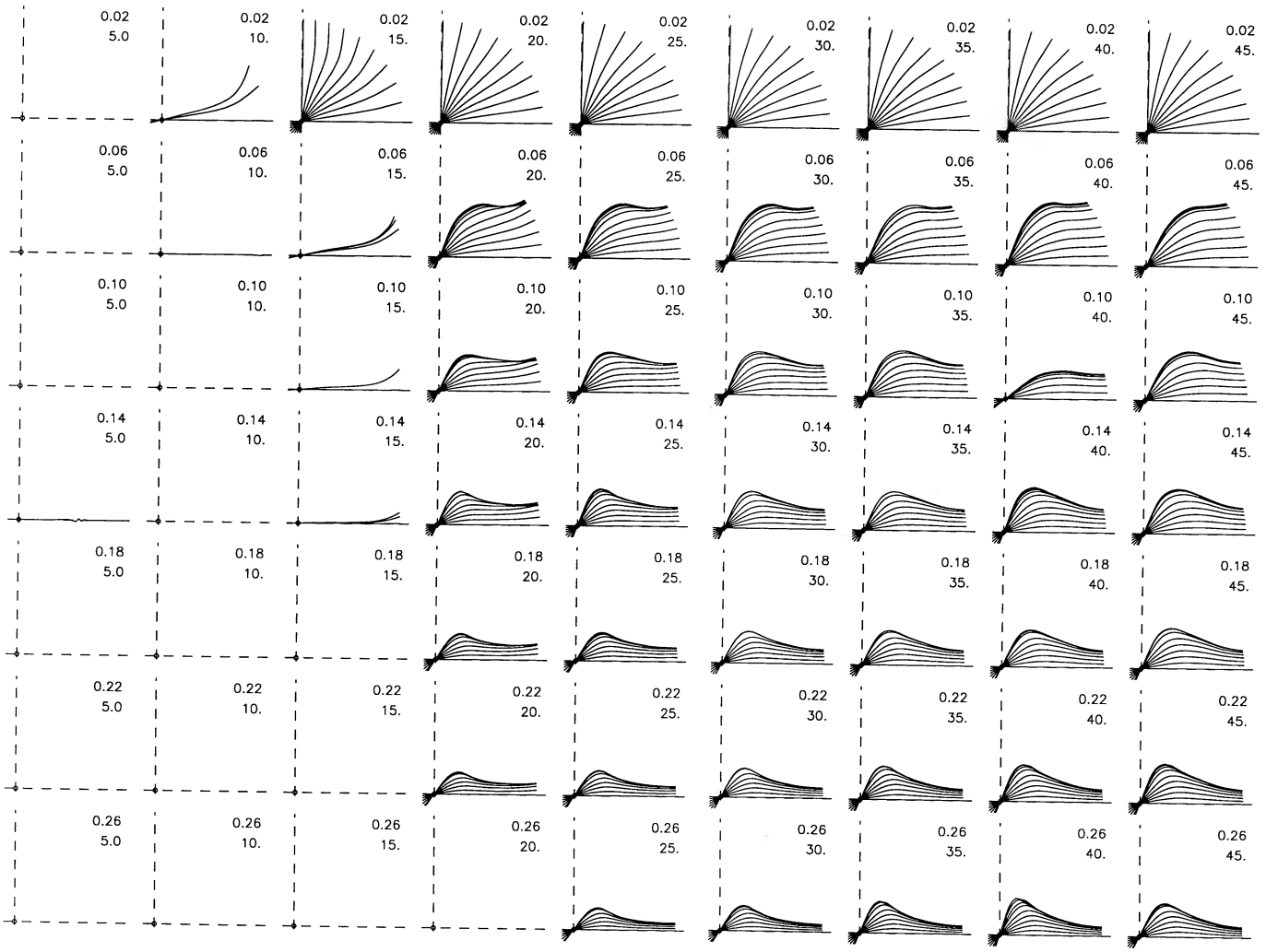

FIG. 7.—Same as Fig. 6, but with a fixed value $\Upsilon_D = 3$. The halo monopole core radius is fitted to the rotation curve, while the core radius of the flattening is varied (*left to right*: $R_H = 5, 10, \ldots, 45$ kpc). ϵ is as in Fig. 6.

observations show that a coherent gas disk extends to radii beyond this point, then this model for the warp will have been disproved.

4.2. *Messier 33*

It has long been known that the local group galaxy M33 has a very pronounced warp in its outer gas disk. This is already evident from the "hat-brims" seen in spatial H I maps of the galaxy. Kinematic data confirm this picture (Rogstad, Wright, & Lockhart 1976; Reakes & Newton 1978): the H I line profiles across most of the face of M33 are split, which is naturally explained if the disk is warped and so oriented that it intersects our line of sight several times. Based on their H I data, Rogstad et al. (1976) constructed a model in which the outer edge of the disk is edge-on to us, and about 45° out of the central disk plane of M33. The change in warp angle is very steep here, occurring in the radius range $R = 5$–9 kpc (assumed distance 690 Mpc). Subsequently, Reakes & Newton (1978) obtained more sensitive observations, which revealed that the gas distribution is more extended than found by Rogstad et al. (1976); they constructed a new tilted-ring model, in which the warp shape is basically the same, but occurs between $R = 7.5$–14 instead. This last model will be compared to the calculations here.

M33 has a nice exponential luminosity profile of scale length 1.6 kpc (de Vaucouleurs 1959), probably with a cutoff at about 4–5 scale lengths (which is where the warp sets in—the Rogstad et al. 1976 model would require the warp to start at a radius of three disk scale lengths, which would be unusually small). The line of nodes in the plane of the inner disk is quite straight: for the Rogstad et al. (1976) model it deviates by about 50° from straightness, and for the Reakes & Newton (1978) model by 30°. Therefore M33 also is not severely affected by differential precession.

Following the same method as applied to NGC 2841 above, the equilibrium warps of M33 were calculated in a variety of model halos. The disk mass-to-light ratio was taken as 1 in the *B* band, one third of its maximum disk value. This will turn out to be close to the allowed maximum, otherwise too great a flattening of the halo potential will be required.

Does the warp of M33 fit the misaligned halo model? First, it is clear that if M33 is a slow-warp case, the halo and disk planes must be misaligned by at least 50°. If M33 has a fast warp, then it has an unusually large amplitude which is diffi-

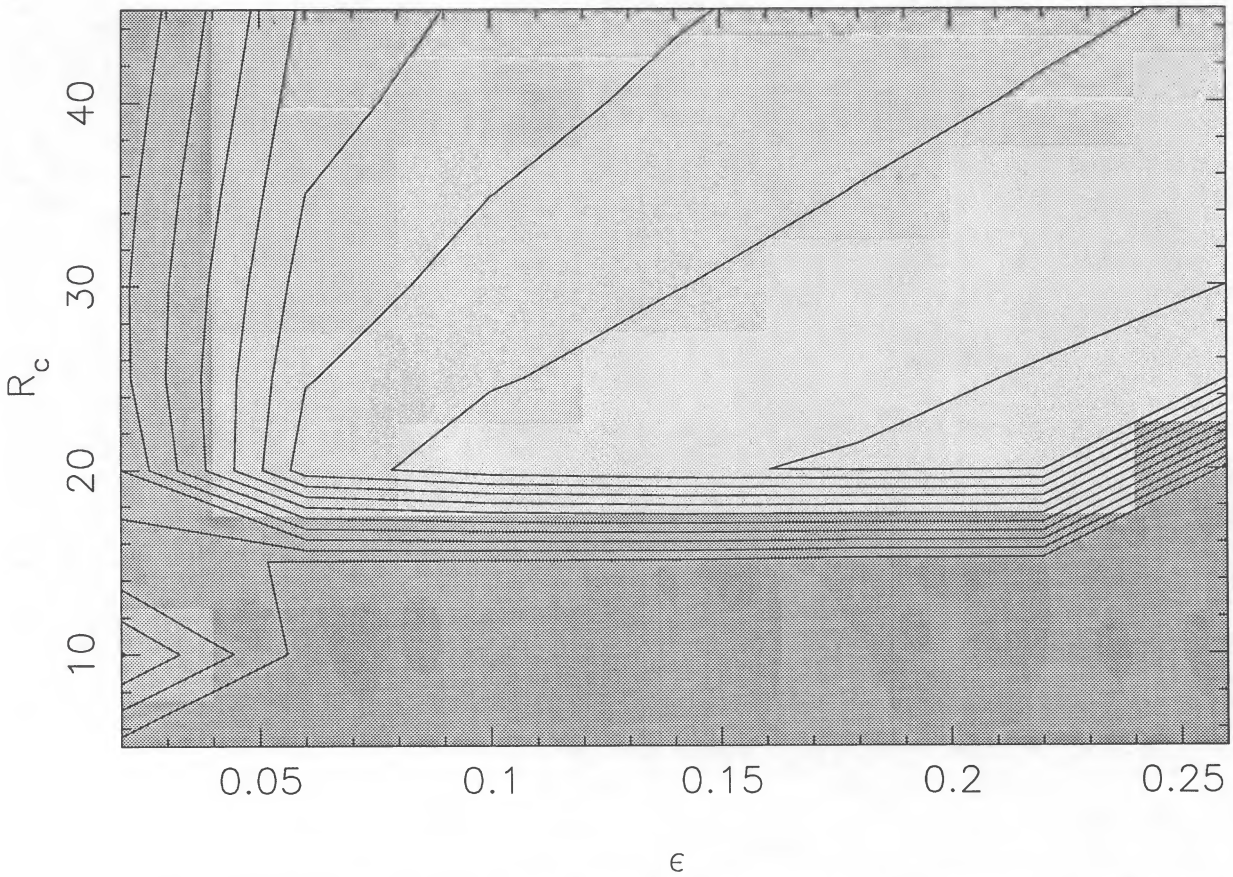

FIG. 8.—Contours of constant rms dispersion about the observations of NGC 2841 for the best-fit model in each panel of Fig. 7 ($\Upsilon_D = 3$). The gray scale is lightest at good fits, and the contours shown are at 1°, 2°, ... rms dispersion.

cult to reproduce in the model calculations. Since at large tilt the torques from halo and Coriolis force act in the same direction, the equilibrium of the outer ring is precarious; nonetheless, if the halo flattening is kept sufficiently small it is possible to construct large-amplitude fast warps. The best-fit fast warp occurs for a halo flattening of 0.01, core radius 1.2 kpc, and inner disk misalignment of 37°. However, the fit is not great: the warp is both steeper and sets in at greater radius than the observations of Reakes & Newton (1978) show, and its amplitude is not quite as large. The comparison is shown in Figure 9.

Slow warps provide a better description, since they are not destroyed at large amplitudes. The rms scatter of the best-fits in various halos is contoured in Figure 10, again for a disk with mass-to-light ratio 1.0. The best-fits, at $\epsilon > 0.20$, $R_H > 10$ kpc require that the inner disk is tilted 61° out of the halo plane. As discussed above for NGC 2841, the required halo flattening is reduced roughly in proportion to Υ_D if the disk mass is decreased. The value used here is already about one-third of the maximum disk value, and cannot be increased significantly without spoiling the positivity of the halo density; on the other hand, it is unlikely that Υ_D is much lower than the value adopted, from the nature of stellar populations. (We note that a maximum-disk model for M33 allows, but does not imply, formation of $m = 2$ structures, whereas a disk mass less than two-thirds of the maximum inhibits such structures (Athanassoula, Bosma, & Papaioannou 1987). No strong two-armed spiral pattern is seen in M33.) This model can reproduce

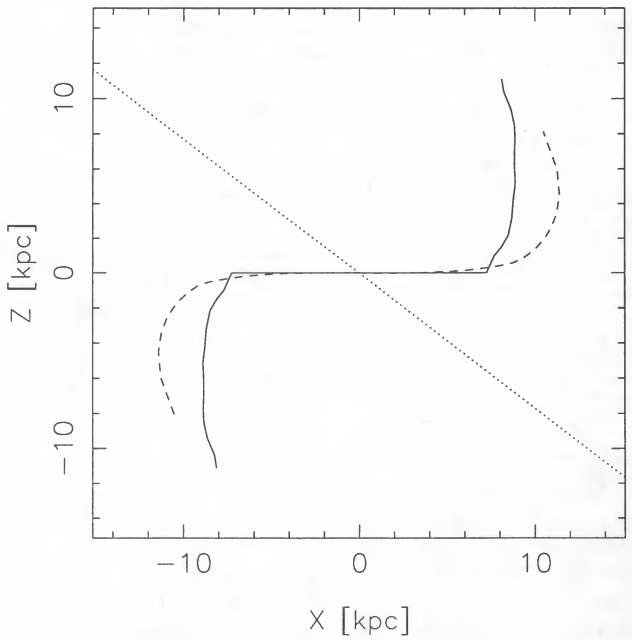

FIG. 9.—Warp of M33 (*solid line*), and the best-fit fast warp to it (*dashed line*). The observed warp extends over a larger range in radius than fast models.

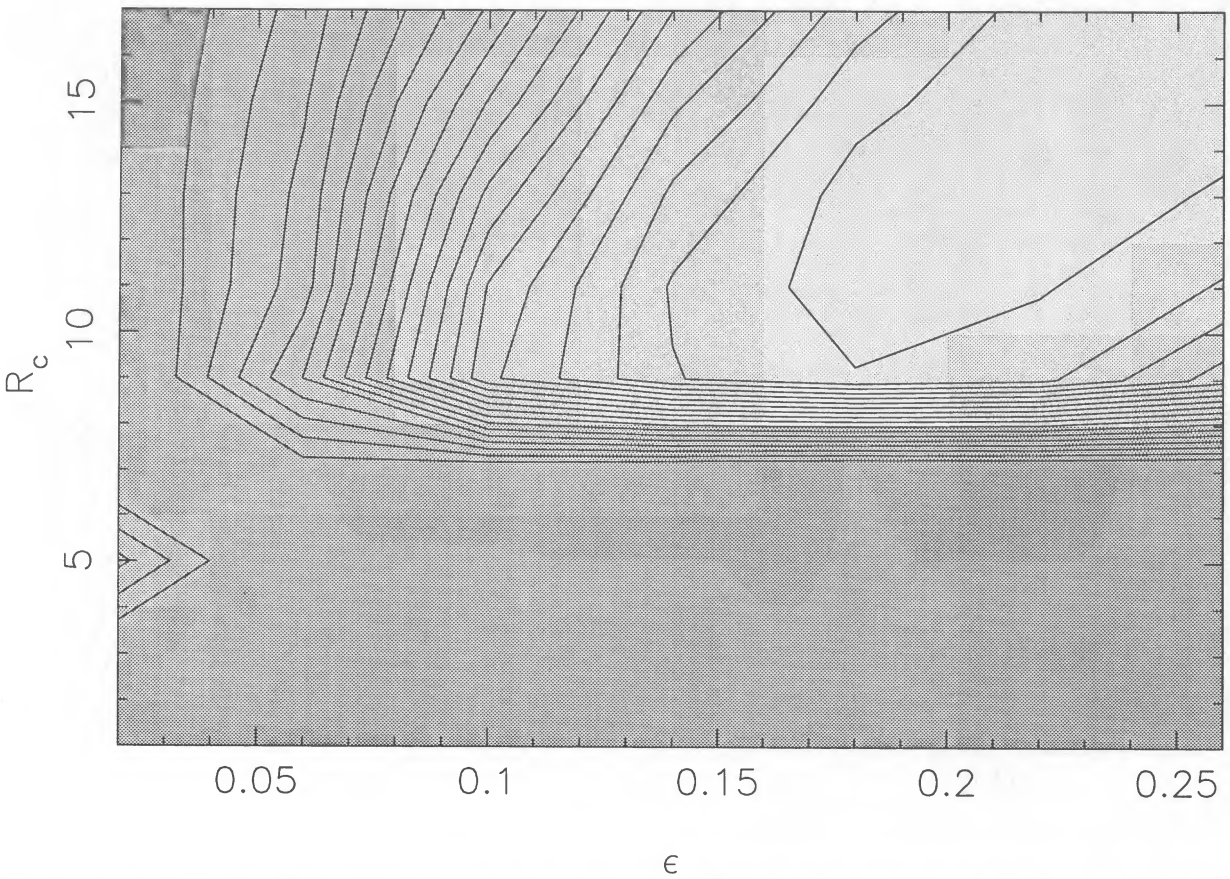

FIG. 10.—Contours of constant rms dispersion between models ($\Upsilon_D = 1$) and observations of the warp of M33. The gray scale is lightest at good fits, and the contours shown are at 2°, 3°, ... rms dispersion.

the observations with an rms scatter of only 2°, as shown in Figure 11. Test rings out to a radius of 50 kpc could take part in the warp; once again, if the disk of M33 were found to extend much further than this while taking part in the global warp, this would be a serious problem for this model.

The large amplitude of M33 makes it a good test case for a comparison between linear and nonlinear models. The best-fits between the observations and freely scaled linear modes (constrained only to have the tilts of the inner and outer rings less than 90° in magnitude) have larger tilts of the inner disk than the full calculations, and are rather poor. The main effect responsible is the reduction of precession speed with increased inner disk tilt, which the linear models do not include. Also, the linear approximation to the disk self-torque is always stronger than the true value (§ 2), making the disk stiffer; thus it cannot reproduce the sharp change in warp angle exhibited in the outer regions of the stellar disk of M33.

5. CONCLUSION

The calculations presented here demonstrate that (1) equilibrium models of uniformly precessing, differentially rotating galactic disks tipped with respect to the equator in flattened halos can resemble observed warped galaxies, specifically NGC 2841 and M33; (2) compared to the linear calculations (Sparke & Casertano 1988) of these equilibria, there are some qualitative differences, notably the disappearance of sharply bent "fast" warp solutions at quite small disk/halo misalign-

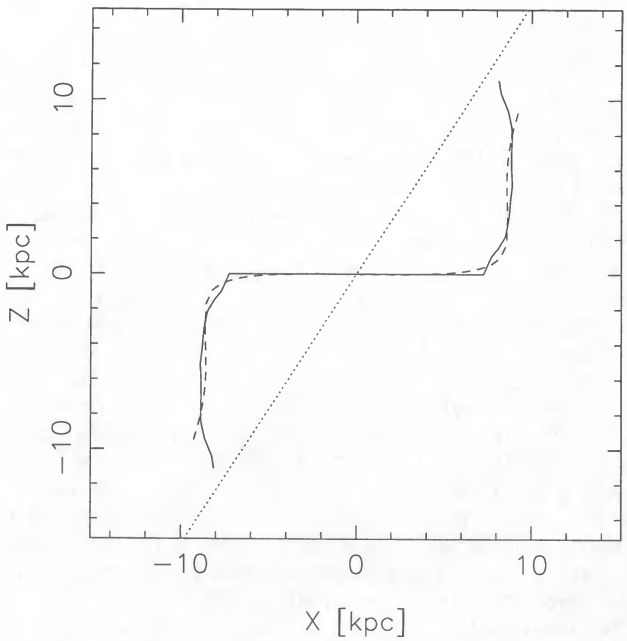

FIG. 11.—Best-fit slow warp (*dashes*) to the M33 data (*solid line*)

ment angles; (3) there is a relation between the core radius *of the quadrupole component* of the halo and the existence of warped equilibrium solutions, which need not, however, translate into a relation involving the halo core radius derived from rotation curves; (4) simple models in which the optical, massive disk is not allowed to warp, and in which the surrounding gas is taken as massless test particles, qualitatively reproduce the important physics in these models; and (5) the warp inferred for M33 from the data of Reakes & Newton (1978) can only be sustained if the disk mass-to-light ratio is at most one-third of the maximum disk value.

It is a pleasure to acknowledge helpful conversations with S. Tremaine, L. Sparke, and A. Toomre. This work was partly supported with a Jeffrey L. Bishop Fellowship.

REFERENCES

Athanassoula, A., Bosma, A., & Papaioannou, S. 1987, A&A, 179, 23
Battaner, E., Florido, E., & Sánchez-Saavedra, M. L. 1990, A&A, 236, 1
Begeman, K. 1987, Ph.D. thesis, Groningen
Bosma, A. 1978, Ph.D. thesis, Groningen
———. 1991, in Warped Disks and Inclined Rings Around Galaxies, ed. F. Briggs, S. Casertano, & P. Sackett (Cambridge: Cambridge Univ. Press), in press
Briggs, F. 1990, ApJ, 352, 15
Dekel, A., & Shlosman, I. 1983, in IAU Symposium 100, Internal Kinematics and Dynamics of Galaxies, ed. E. Athanassoula (Dordrecht: Reidel), p. 187
de Vaucouleurs, G. 1959, ApJ, 130, 728
Dubinski, J., & Carlberg, R. G. 1991, ApJ, in press
Heisler, J., Merritt, D., & Schwarzschild, M. 1982, ApJ, 258, 490
Hunter, C. 1969, ApJ, 157, 183
Hunter, C., & Toomre, A. 1969, ApJ, 155, 747
Kahn, F. D., & Woltjer, L. 1959, ApJ, 130, 705
Kent, S. M. 1986, AJ, 93, 816
Lake, G., & Feinswog, L. 1989, AJ, 98, 166
Lynden-Bell, D. 1965, MNRAS, 129, 299
Ostriker, E. C., & Binney, J. J. 1989, MNRAS, 237, 785
Petrou, M. 1980, MNRAS, 191, 767
Reakes, M. L., & Newton, K. 1978, MNRAS, 185, 277
Rogstad, D. H., Wright, M. C. H., & Lockhart, I. A. 1976, ApJ, 204, 703
Sackett, P. D., & Sparke, L. S. 1990, ApJ, 361, 408
Sánchez-Saavedra, M. L., Battaner, E., & Florido, E. 1990, MNRAS, 246, 458
Sparke, L. S. 1984, ApJ, 280, 117
Sparke, L. S., & Casertano, S. 1988, MNRAS, 234, 873 (SC)
Toomre, A. 1983, in IAU Symposium 100, Internal Kinematics and Dynamics of Galaxies, ed. E. Athanassoula (Dordrecht: Reidel), p. 177
Tubbs, A. D., & Sanders, R. H. 1979, ApJ, 230, 736
van Albada, T. S., & Sancisi, R. 1986, Phil. Trans. Roy. Soc. London, A320, 305
Wevers, B. M. H. 1984, Ph.D. thesis, Groningen