PRODUCTION OF MILKY WAY STRUCTURE BY THE MAGELLANIC CLOUDS

MARTIN D. WEINBERG
Department of Physics and Astronomy, University of Massachusetts, Amherst, MA 01003-4525; weinberg@phast.umass.edu

Received 1995 July 17; accepted 1995 September 21

ABSTRACT

Previous attempts at disturbing the Galactic disk by the Magellanic Clouds relied on direct tidal forcing. However, by allowing the halo to respond actively rather than remain a rigid contributor to the rotation curve, the Clouds may produce a wake in the halo which then distorts the disk. Recent work reported here suggests that the Magellanic Clouds use this mechanism to produce disk distortions sufficient to account for the location, position angle, and sign of the H I warp and the observed anomalies in stellar kinematics toward the Galactic anticenter and LSR motion. More generally, the observed response depends on the gravitational potential and therefore provides a diagnostic for the structure of the halo and its extent.

Subject headings: galaxies: halos — galaxies: kinematics and dynamics — Galaxy: halo — Galaxy: structure — Magellanic Clouds

1. INTRODUCTION

All components of spiral galaxies are coupled by their mutual gravitational field. The fact that we only directly observe the luminous components focuses attention on the disk alone in attempts to understand their striking structure. For instance, “grand design” is often correlated with a companion or interloping galaxy and understood to be a result of differential or tidal acceleration of the disk. Similarly, researchers have tried to implicate our own fairly massive pair of companions, the Magellanic Clouds, in producing the observed warp in the Milky Way’s disk. However, the tidal force, which scales as the inverse distance cubed, is nearly an order of magnitude too small to do the job (see Binney 1992 for a review).

The dark halo was postulated more than 20 yr ago to account for extended flat rotation curves, and data continue to imply the existence of this unseen mass. True to the original intent, most dynamical theories regard the halo as inert, providing gravitational support only for the luminous components. However, the estimated mass of the halo is comparable to that of the disk inside of the Sun’s Galactocentric radius and dominates at larger radii. Therefore, any structure in the halo must surely affect the disk, and in this Letter, I implicate the halo as a coconspirator for exciting disk structure.

In short, the long-range gravitational response of the halo can carry the external disturbance by the Magellanic Clouds to sufficiently small radii that the luminous disk can be affected (§ 2), and the disk response is predicted using perturbation theory (§ 3). Previous work has demonstrated that interactions between galaxies and between components of galaxies can be studied perturbatively with good results. Here the theory is extended to include the effects of a disk embedded in a combined luminous and dark matter halo. The implications and observational comparisons are discussed in § 4.

2. EFFECT OF THE CLOUDS ON THE HALO/SPHEROID COMPONENT

Mass estimates for the Large Magellanic Cloud (LMC) range from $6 \times 10^8$ (Meatheringham et al. 1988) to $1.5 \times 10^{10}$ $M_\odot$ (Schommer et al. 1992), which is from 10% to 30% of the stellar mass of the Milky Way’s disk (Binney & Tremaine 1987). Because the Clouds are distant, $R_\odot \approx 50$ kpc, their trajectory traces the extent of the dark halo, and considerable modeling and observational effort have been put to the task. Recently, the space velocity of the LMC has been redetermined from radial velocity and proper motions (Jones, Klemola, & Lin 1994). Assuming a spherical halo and Galactocentric coordinate system with the LSR along the x-axis ($x = -8.5$ kpc), and moving toward positive y, one may straightforwardly derive the following instantaneous orbit: $-76^\circ \pm 13^\circ$ inclination, $-2^\circ \pm 10^\circ$ longitude of ascending node, and $-36^\circ \pm 3^\circ$ argument of perigalacticon.$^2$

The dark halo causes the Clouds’ orbit to decay by dynamical friction: our Galaxy will eventually absorb the Clouds (Murai & Fujimoto 1980; Lin & Lynden-Bell 1982; Lin, Jones, & Klemola 1995). The response of the halo to the interloping satellite can be thought of as a gravitational wake (e.g., Mulder 1983); since the wake trails and has mass, it exerts a backward pull on the satellite. This view of dynamical friction reproduces the standard results, but can include self-gravity (Weinberg 1989), and, most important for our purposes here, it gives us a way to estimate the distortion to the Galaxy itself.

Specifically, the response can be calculated by solving the collisionless Boltzmann equation (CBE) after all initial transients have decayed (time asymptotic limit). Since the satellite is small compared to the main galaxy, a perturbation expansion is a fair predictor of the true disturbance. The general approach has been tested and compared with n-body simulations in a variety of contexts with good agreement (e.g., Hernquist & Weinberg 1989; Weinberg 1993) and will be used here to consider the effect of the satellite on both the halo and the luminous disk.

For an example, let us consider a hypothetical dwarf companion in an extended halo,$^3$ and let the companion be a mass point on an eccentric orbit ($e = 0.4$) similar to the LMC and

$^2$ The quoted errors are propagated by Monte Carlo analysis from the quoted standard errors in distance and proper motion, but are dominated by the latter.

$^3$ I will refer to the total nondisk distribution of mass as the halo; the halo, then, contains the luminous spheroid as well as extended dark matter components.

1 Alfred P. Sloan Foundation Fellow.
currently at pericenter with coordinates \((X, Y) = (50 \text{ kpc}, 0 \text{ kpc})\). Consistent with current estimates, the halo is truncated at 100 kpc with a mass of \(6 \times 10^{11} \, M_\odot\) (following Lin et al. 1995). The concentration of a King profile for the halo, \(c = 0.67\), was chosen to produce an overall flat rotation curve for the combined halo and exponential disk with scale length \(a = 3.5 \text{ kpc}\). For simplicity, and consistent with our emphasis on the outer Galaxy, this model does not include a bulge.

The \(m = 1\) and \(m = 2\) density components of the resulting wake are shown in Figures 1 and 2 on an absolute scale. Although the LMC does excite a local wake in the outer halo, the eccentric LMC orbit has two orbital frequencies whose harmonics determine regions in phase space which couple strongly and nonlocally. The strongest response results from the confluence of two conditions: (1) the existence of low-order, and therefore strongly responding, commensurabilities between the LMC and halo orbit frequencies, and (2) significant mass in phase space near these resonances. In the present case of a flat outer rotation curve, low-order resonances occur at radii roughly between 2 and \(n\) times smaller than the satellite orbital radius for a \(n:1\) resonance. In addition, resonances will cluster where frequencies are changing quickly and commensurability is easier to achieve. These regions occur near the characteristic radii, where the profile is in transition to its asymptotic form. Figures 1 and 2 illustrate these effects, showing the dominant response deep within the halo and near the halo core radius. Most important for this inquiry, the orbiting satellite can influence the inner Galaxy, including the luminous disk, via the global gravitational response of the halo!

\(^4\) The rotation curve drops only 30 km \(s^{-1}\) between 8.5 and 50 kpc.

3. COMBINED EFFECT OF THE CLOUDS AND THE HALO ON THE GALACTIC DISK

To estimate the amplitude of the disk response to the halo and LMC together, I extended the perturbative solution of the CBE to include a flat disk embedded in the halo. For now, the disk disturbance is confined to the plane (see § 4.2 for a discussion of warping). Both the disk and the halo react gravitationally to each other, to themselves, and to the external satellite. The initial disk has the exponential density profile typical of spirals; the distribution function is constructed self-consistently by a quadratic programming technique (Dejonghe 1989) penalized to prefer a cold tangential distribution with an appropriate velocity dispersion. The gravitational field of each component is represented by a truncated series of orthogonal functions (e.g., Clutton-Brock 1972, 1973; Hernquist & Ostriker 1992). With this discretization, the response is a solution to a matrix equation. The individual galactic components are gravitationally coupled by matrices which transform one set of orthogonal functions to all others. Since an orthogonal polar disk component, proportional to \(\exp(\text{i} m \phi)\), couples to all spherical components \(Y_{lm}\) with \(l \geq m\), all desired values \(l\) for a single \(m\) must be considered together. Once the perturbed distribution function is known, all observable quantities such as density and line-of-sight velocity are immediate consequences.

Here we consider the lowest order nonaxisymmetric component, \(m = 1\). As in Figures 1 and 2 for the halo, the \(m = 2\) disk response is smaller by an order of magnitude near the solar circle, and this trend continues to higher order. A detailed discussion will be presented in a later paper. The \(l = m = 1\) disturbance shifts the center of mass. A significant fraction of the halo response will be a center-of-mass or barycentric shift (Weinberg 1989), since the LMC is in the outer halo. A pure barycentric profile would be proportional to \(dpdr\), and deviations from this pattern are evident in Figure
1. The embedded disk responds more strongly to the halo and its wake than the direct LMC tide, and is distorted more than shifted. Figure 3 depicts the distortion of the disk by the halo and the LMC, with the orbit plane tilted and rotated as stated in §2. Both absolute and relative density scales are shown. The dominant response is near the solar circle and has significant amplitude: several percent at the solar circle, reaching 20% at $R_g \approx 12.5$ kpc and 30% near $R_g \approx 3.5$ kpc. This should lead to a detectable offset between the inner and outer Galaxy.

Finally, note that the LMC has a retrograde orbit relative to the disk rotation, and the response has a negative pattern speed relative to the disk rotation. Therefore, the slow retrograde patterns in Figures 1 and 2 cannot be directly responsible for classic prograde disk features, but could be the triggering source.5

4. IMPLICATIONS AND SUMMARY

4.1. Comparison with Inferred Milky Way Asymmetries

Stellar spectroscopy of K giants and carbon stars (Lewis & Freeman 1989; Metzger & Schechter 1994) indicate a net streaming motion relative to the LSR toward the Galactic anticenter of approximately 10 km s$^{-1}$. An axisymmetric galaxy would have no streaming. Most explanations for the implied nonaxisymmetries are based on weak oval or quadrupole distortions. It is clear from Figure 3 that the disk response to the LMC-engendered halo wake may bear on this interpretation. Figure 4 shows the predicted kinematic streaming toward the anticenter as observed from the Galactic center. The streaming is close to zero until $R_g \approx 6.5$ kpc, where it increases monotonically to a peak at $R_g \approx 12.5$ kpc and decreases monotonically to a low value at $R_g \approx 20$ kpc. Viewed from the LSR, the outer Galaxy would appear to be receding at roughly 5 km s$^{-1}$. This is consistent with the observed stellar kinematics, especially in view of the uncertainties in the halo profile and the LMC mass, which can change both the scale and the amplitude.

Blitz & Spergel (1991) use a quadrupole distortion to deduce an outgoing LSR and a circularly symmetric outer Galaxy based on HI kinematics. While somewhat inconsistent with observed and predicted stellar kinematics, the time-dependent LMC wake might easily cause a transient gas response which differs from the stellar response but has similar amplitude. More generally, these estimates suggest observable velocity signatures near and outside the solar circle caused by the Magellanic Clouds, but more detailed modeling will be required for precise predictions.

4.2. Implications for the Galactic Warp

Although the inferred LMC orbital plane intersects the Galactic disk near the observed warp maxima, quantitative analysis of the tidal distortion predicts amplitudes almost an order of magnitude smaller than observed (Hunter & Toomre 1969). Unfortunately, the nonlocal resonant coupling described in §3 is very unlikely to be important here, because the vertical stellar motions have high frequencies compared to LMC orbital frequencies. Nevertheless, the halo response does reinforce the vertical force on the disk in two ways: (1) the wake increases the total mass in the perturbation by roughly a factor of 2, and (2) the response tends to be at smaller radii, which further increases its effect. Overall, we may expect greater than a factor of 2 increase in bending amplitude near the edge of the stellar disk.

To estimate the forced vertical response of the Galactic disk, I coupled the Hunter-Toomre machinery to the com-
combined vertical force from the LMC tide and halo wake. To demonstrate the largest effect, the Schommer et al. (1992) mass estimate is taken, rather than the smaller Meatherrington et al. (1988) value used for all of the previously quoted results. The peak height of the induced warp occurs near \( R_g \approx 25 \) kpc and varies from 400 to 600 pc with the azimuth of the LMC orbit (Fig. 5). The wake in the outer halo lags the satellite by \( 160^\circ \) and dominates the combined vertical force because of its proximity. The current height of the warp is predicted to be 450 pc, with position consistent with its observed location roughly toward Galactic longitudes 90° and 270°. Similar to the conclusions of Hunter & Toomre (1969), one still has to stretch the parameters to account for the observed amplitude. However, the warp calculation does not include self-gravity or the possible resonant interaction with the global bending modes, both of which could sustain or amplify the warp. It is encouraging, therefore, that the predicted position and amplitude at maximum are comparable to observations. To summarize, this work suggests that an active halo distorted by the Clouds can drive the warp, given the uncertainty in the LMC parameters and the limitations of the estimate.

4.3. Future Work

The details of the predictions made here depend on the properties of both the LMC and the potential model for the Milky Way, especially the halo. For example, increasing the core radius, with fixed mass and tidal radius, or increasing the tidal radius, with fixed mass and core radius, leads to weaker disk responses, probably because of the smaller halo mass inside the solar circle. Conversely, the dark matter halo is invoked to account for flat rotation curves. The possibility that the halo has dynamically observable consequences, in addition to producing the rotation curve, provides an independent way to investigate the properties of dark halos in general. Clearly, much remains to be done before the full implications of this mechanism are understood. In addition to the shortcomings already discussed, the following three broad topics are ripe for rapid progress:

1. Effect of the three-dimensional structure of the disk.—This is a straightforward application for \( n \)-body simulation, but it will require very large numbers of particles. Alternatively, one could use a hybrid approach, with an analytic model for the halo distortion together with an \( n \)-body disk.

2. Gas response and comparison with observed HI.—It is likely that the time-dependent effects are crucial to understanding the stellar and gas kinematics together. Both periodic orbit analyses and direct \( n \)-body simulations with gas could be performed and compared.

3. Nonlinear development.—The results presented here describe a linear response only. The slow retrograde distortions might be amplified by the standard mechanisms to drive spiral structure and perhaps inner bars.

I thank Leo Blitz, Jerry Sellwood, and Scott Tremaine for discussion, comments, and suggestions, and Klaus Schittkowski for use of his quadratic programming code. This work was supported in part by NASA grant NAG 5-2873.

REFERENCES

Binney, J. 1992, ARA&A, 30, 51

Binney, J., Tremaine, S. 1987, Galactic Dynamics (Princeton: Princeton Univ. Press)

Blitz, L., & Spergel, D. N. 1991, ApJ, 370, 205

Clutton-Brock, M. 1972, ApSS, 16, 101

———. 1973, ApSS, 23, 55

Dejonghe, H. 1989, ApJ, 343, 113

Hernquist, L., & Ostriker, J. P. 1992, ApJ, 386, 375

Hernquist, L., & Weinberg, M. D. 1989, MNRAS, 238, 407

Hunter, C., & Toomre, A. 1969, ApJ, 155, 747

Jones, B. F., Klemola, A. R., & Lin, D. N. C. 1994, AJ, 107, 1333

Lewis, J. R., & Freeman, K. C. 1989, AJ, 97, 139

Lin, D. N. C., Jones, B. F., & Klemola, A. R. 1995, ApJ, 439, 652

Lin, D. N. C., & Lynden-Bell, D. 1982, MNRAS, 189, 707

Meatherringham, S. J., Dopita, M. A., Ford, H. C., & Webster, B. L. 1988, ApJ, 327, 651

Metzger, M. R., & Schechter, P. L. 1994, ApJ, 420, 177

Mulder, W. A. 1983, A&A, 117, 9

Murai, T., & Fujimoto, M. 1980, PASJ, 32, 581

Schommer, R. A., Olszewski, E. W., Suntzeff, N. B., & Harris, H. C. 1992, AJ, 103, 447

Weinberg, M. D. 1989, MNRAS, 239, 549

———. 1993, ApJ, 410, 543