

UNSTABLE BAR AND SPIRAL MODES OF DISK GALAXIES

MIR ABBAS JALALI¹ AND C. HUNTER²

Department of Mathematics, Florida State University, Tallahassee, FL 32306-4510
submitted to The Astrophysical Journal

ABSTRACT

We present new analytical and numerical results for modes of flat stellar disks which lie in potentials with soft centers. Stars primarily circulate in one direction. We identify two modes of angular wavenumber 2: a more central fundamental mode and a more extensive and more spiral (trailing) secondary mode. The fundamental mode is particularly sensitive to the population of stars of low angular momentum. Depending on that population, a small fraction of the whole, the fundamental mode varies from a small compact bar to a trailing spiral that is almost as wound as the secondary mode. Modes transfer angular momentum from the central to the outer regions of the disk. Most of them release gravitational energy and convert it to kinetic energy, which also flows outwards through the disk. Few Fourier components contribute significantly to this transfer. All modes rotate too rapidly to have an inner Lindblad resonance, and are unstable unless there is a sufficiently large external halo or bulge.

Subject headings: stellar dynamics — galaxies: kinematics and dynamics — galaxies: spiral — galaxies: structure

1. INTRODUCTION

The publication by Lin & Shu (1964) of a density wave theory of spiral structure has led to a broad literature. One strand of that literature consists of stellar dynamic studies of flat stellar systems. Lin & Shu assumed tight-winding of waves, and near-circularity of the stellar orbits of the unperturbed disk. Using these same approximations, Toomre (1964) showed that the stability of stellar disks to axisymmetric waves requires that the parameter $Q = \kappa \sigma_R / 3.36 G \Sigma_D > 1$. Here κ is the epicyclic frequency, σ_R is the radial velocity dispersion, Σ_D is the disk density, and G is the gravitational constant. This result has proved to be remarkably robust and widely applicable, despite the simplifying approximations used in deriving it. Yet it was soon recognized that Toomre's criterion does not provide a complete stability criterion for stellar disks when N -body experiments (Miller, Prendergast & Quirk 1970, Hohl 1971, Ostriker & Peebles 1973) revealed that disks with Q safely greater than 1 are prone to large-scale bar-like instabilities.

Kalnajs (1971, 1977) developed the matrix method for using linear perturbation theory to study the stability of stellar systems. His method imposes no restrictions on the forms of either the instability, or the unperturbed stellar orbits. The matrix method is sufficiently complex that it has not been widely used. Its first comprehensive use was by Zang (1976) to study scale-free singular isothermal disks. Some of his results are given in papers by his PhD supervisor Toomre (1977, 1981). Zang's methods were re-used and extended by Evans & Read (1998a,b) to scale-free disks with other power-laws. Orbits of any energy in a scale-free disk are simply scaled versions of those at any other energy. Though this is indeed a simplification, applications of the matrix method even to scale-free disks are by no means simple.

We study stellar disks whose potentials are locally har-

monic at their centers. We adopt Toomre's (1981) terminology, and label these potentials as soft-centered, in contrast to the sharp centers of the singular potentials of scale-free disks. §2 contains new developments of the matrix method. We show that this method generally requires extra boundary integral terms when applied to disks whose stars all circulate in the same direction. The boundary integrals arise from near-radial orbits. We extend the linear perturbation theory to second order in §2.3. Those results allow us later to study how modes re-arrange the angular momentum and energy.

To understand the influence of orbital populations on the responses of the disks, we study a variety of disks. That variety allow us also to study the effects of the mass and extent of the disk. We use three axisymmetric potentials for our disks; those of Kuzmin's disk, the isochrone disk, and the soft-centered logarithmic potential. We use known distribution functions (DFs) for the first two, but none are currently available for cored exponential disks embedded in the soft-centered logarithmic potential. We construct some in Appendix C.

We give numerical results for modes of angular wavenumber $m = 2$ in §3. Some of our models are the same as those of others, and we compare results. Prior work, some of which was done using N -body simulations, includes that of Kalnajs (1978), Earn & Sellwood (1995) and Pichon & Cannon (1997) for the isochrone disk, and that of Athanassoula & Sellwood (1986, hereafter AS), Sellwood & Athanassoula (1986, hereafter SA), Hunter (1992), Pichon & Cannon (1997), and Polyachenko (2004, 2005) for Kuzmin's disk. Vauterin & Dejonghe (1996) studied modes of a cored exponential disk, though in a different potential, while Sellwood (1989) studied the uncored and mildly singular exponential disk in the same soft-centered logarithmic potential as ours. Our numerical methods are described in Jalali & Hunter (2005, hereafter JH).

We analyze our results in §4, and relate them to Polyachenko's (2004, 2005) unified theory of bar-like and spi-

¹ E-mail: mjalali@math.fsu.edu

² E-mail: hunter@math.fsu.edu

ral modes. We summarize our findings in §5. Appendix A shows how to pass from the usual Eulerian form of the perturbation theory to its Lagrangian form, and how to interpret the boundary integrals as boundary fluxes.

2. PERTURBATION THEORY

This section presents the dynamical theory on which our work is based. §2.1 carries out an Eulerian linear perturbation analysis of the collisionless Boltzmann equation in action-angle variables, following Kalnajs (1971, 1977), to derive a matrix eigenvalue problem for finding modes. We apply this analysis in §2.2 to unidirectional disks in which all stars circulate in the same direction. We show that the matrix formulation may then need additional boundary integral terms which have been unjustifiably ignored in some earlier work. We extend the perturbation theory to second order in §2.3 to the extent necessary to discuss the transfer of angular momentum, and kinetic and potential energy. We discuss orbits and Polyachenko's (2004, 2005) unified theory in §2.4. In §2.5 we discuss the cutout functions which we have used in some cases to generate stellar disks with central holes.

2.1. Linear Perturbation Theory

We study the stability of a collisionless stellar disk composed of a distribution of stars moving in orbits in a circularly symmetric potential $V_0(R)$. The unperturbed system is described by a DF $f_0(E, L)$ (Binney & Tremaine 1987) where

$$E = \frac{1}{2} (v_R^2 + v_\phi^2) + V_0(R), \quad L = Rv_\phi, \quad (1)$$

are the energy and angular momentum which remain constant along an orbit. The density corresponding to the unperturbed DF is

$$\Sigma_D = \int f_0 d\mathbf{v}, \quad (2)$$

where $d\mathbf{v}$ denotes an element of the two-dimensional velocity space. This density is the one that produces the potential $V_0(R)$ only in the fully self-consistent case; in our work it may provide only a part of that potential. The development of the DF for the perturbed state is described by the collisionless Boltzmann equation

$$\partial f / \partial t + [f, \mathcal{H}] = 0, \quad (3)$$

where $[,]$ denotes a Poisson bracket and $\mathcal{H}$ is the Hamiltonian

$$\mathcal{H} = \frac{1}{2} (v_R^2 + v_\phi^2) + V(R, \phi, t). \quad (4)$$

We expand the DF as $f = f_0 + f_1 + f_2 + \dots$ and the Hamiltonian as $\mathcal{H} = \mathcal{H}_0 + V_1 + V_2 + \dots$, where $V = V_0 + V_1 + V_2 + \dots$ is the corresponding expansion of the potential. Collecting terms of first and second orders then gives the equations

$$\frac{\partial f_1}{\partial t} + [f_1, \mathcal{H}_0] = -[f_0, V_1], \quad (5)$$

$$\frac{\partial f_2}{\partial t} + [f_2, \mathcal{H}_0] = -[f_0, V_2] - [f_1, V_1]. \quad (6)$$

It is necessary that the perturbed densities due to the changes from the unperturbed DF are precisely those

needed to produce the corresponding components of the perturbed density, so that

$$V_j(\mathbf{x}, t) = -G \int \frac{\Sigma_j(\mathbf{x}', t) d\mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|}, \quad j \geq 0, \quad (7)$$

$$\Sigma_j(\mathbf{x}, t) = \int f_j(\mathbf{x}, \mathbf{v}, t) d\mathbf{v}, \quad j > 0. \quad (8)$$

Here $d\mathbf{x}$ and $d\mathbf{v}$ denote elements of the two-dimensional position and velocity spaces. Equation (7) is true when $j = 0$ because we shall use Σ_0 to denote the self-consistent density for the potential V_0 . Equation (2) replaces the $j = 0$ case of equation (8).

It is convenient to work with the actions

$$J_R = \frac{1}{2\pi} \oint v_R dR, \quad J_\phi = L. \quad (9)$$

of the unperturbed motion, and to express the unperturbed DF and Hamiltonian as $f_0(J_R, J_\phi)$ and $\mathcal{H}_0(J_R, J_\phi)$ respectively. The advantage of using action variables is that their conjugate angle variables (θ_R, θ_ϕ) increase uniformly with time

$$\frac{d\theta_R}{dt} = \frac{\partial \mathcal{H}_0}{\partial J_R} = \Omega_R(J_R, J_\phi), \quad \frac{d\theta_\phi}{dt} = \frac{\partial \mathcal{H}_0}{\partial J_\phi} = \Omega_\phi(J_R, J_\phi). \quad (10)$$

Perturbations must be periodic in the angles, and hence have Fourier series in these angles (Kalnajs 1971). We write them in the complex form

$$f_1 = e^{i(m\theta_\phi - \omega t)} \sum_{l=-\infty}^{\infty} \tilde{f}_l(J_R, J_\phi) e^{il\theta_R}, \quad (11)$$

$$V_1 = e^{i(m\theta_\phi - \omega t)} \sum_{l=-\infty}^{\infty} \tilde{V}_l(J_R, J_\phi) e^{il\theta_R}, \quad (12)$$

with the understanding that it is their real parts which give the physical solution. Following Landau (1946), we suppose that the frequency ω is complex with real and imaginary parts

$$\omega = m\Omega_p + is, \quad (13)$$

with Ω_p representing an angular pattern speed and s a growth rate. We suppose initially that $s > 0$ so that we have a growing disturbance which was infinitesimally small infinitely long ago. The possibility of stationary oscillations and real values of ω has to be considered via analytical continuation to $s = 0$ from $s > 0$.

The solutions (11) and (12) are those for a single angular wavenumber m . Perturbations of all angular wavenumbers are possible, but do not interact at the first order when, as here, the unperturbed state is axisymmetric. In fact we shall be concerned almost entirely with the case of $m = 2$.

Substituting expansions (11) and (12) into equation (5) and matching Fourier coefficients yields the relation

$$\tilde{f}_l(J_R, J_\phi) = \frac{\tilde{V}_l(J_R, J_\phi)}{l\Omega_R + m\Omega_\phi - \omega} \left(l \frac{\partial f_0}{\partial J_R} + m \frac{\partial f_0}{\partial J_\phi} \right). \quad (14)$$

The potential V_1 , and the density Σ_1 which causes it and is obtained from integrating f_1 as in equation (8), can also be expanded in position space as

$$V_1 = e^{i(m\phi - \omega t)} \sum_{j=0}^{\infty} c_j \psi_j^m(R), \quad (15)$$

$$\Sigma_1 = e^{i(m\phi - \omega t)} \sum_{j=0}^{\infty} c_j \sigma_j^m(R), \quad (16)$$

where $\psi_j^m(R)$ and $\sigma_j^m(R)$ are a complete set of real basis functions, and c_j are constant coefficients. We multiply equation (16) by $e^{i(\omega t - m\phi)}\psi_j^m(R)$ and integrate over position space to get $\sum_{k=0}^{\infty} D_{jk} c_k$, where

$$D_{jk}(m) = 2\pi \int_0^{\infty} \psi_j^m(R) \sigma_k^m(R) R dR, \quad (17)$$

are the components of a constant matrix $\mathbf{D}(m)$. It is diagonal if ψ_j^m and σ_k^m form a biorthogonal set. Alternatively, we can rewrite Σ_1 using its integral (8), and then carry out the integration over phase space using action and angle variables. This requires that we calculate Fourier coefficients of the basis potential functions

$$\Psi_{l,j}^m(J_R, J_\phi) = \frac{1}{\pi} \int_0^\pi \psi_j^m(R) \times \cos[l\theta_R + m(\theta_\phi - \phi)] d\theta_R, \quad (18)$$

$$\tilde{V}_l = \sum_{j=0}^{\infty} c_j \Psi_{l,j}^m, \quad (19)$$

(Kalnajs 1977; Tremaine & Weinberg 1984). Using equation (14) to relate the Fourier coefficients of f_1 to those of V_1 yields

$$[\mathbf{M}(m, \omega) - \mathbf{D}(m)]\mathbf{c} = \mathbf{0}, \quad (20)$$

where the components of the matrix $\mathbf{M}(m, \omega)$ are defined as

$$\frac{M_{jk}}{4\pi^2} = \sum_{l=-\infty}^{\infty} \int_0^{\infty} dJ_R \int_{-\infty}^{\infty} \frac{\left(l \frac{\partial f_0}{\partial J_R} + m \frac{\partial f_0}{\partial J_\phi} \right)}{l\Omega_R + m\Omega_\phi - \omega} \Psi_{l,j}^m \Psi_{l,k}^m dJ_\phi, \quad (21)$$

with integration over the whole of action space. We suppose that the integrand decays sufficiently rapidly as $J_R \rightarrow \infty$ and $J_\phi \rightarrow \pm\infty$ for convergence. The linear equations (20) have a non-trivial solution for the coefficient vector $\mathbf{c}$ only if

$$\mathcal{M}(m, \omega) \equiv |\mathbf{M}(m, \omega) - \mathbf{D}(m)| = 0. \quad (22)$$

This determinantal equation defines a nonlinear eigenvalue problem for ω . Once ω is found, its eigenvector $\mathbf{c}$ gives the physical shape of the perturbation.

2.2. Boundary Integrals

The DF of a unidirectional disk for which all the stars rotate in the prograde direction has the form

$$f_0(J_R, J_\phi) = f_0^P(J_R, J_\phi) H(J_\phi), \quad (23)$$

where H is the Heaviside function. The derivative of this DF with respect to J_ϕ is

$$\frac{\partial f_0}{\partial J_\phi} = \frac{\partial f_0^P}{\partial J_\phi} H(J_\phi) + f_0^P(J_R, 0) \delta(J_\phi). \quad (24)$$

The matrix $\mathbf{M}(m, \omega)$ then has two components

$$\mathbf{M}(m, \omega) = \mathbf{M}^A(m, \omega) + \mathbf{M}^B(m, \omega), \quad (25)$$

whose elements are

$$\begin{aligned} \frac{M_{jk}^A}{4\pi^2} &= \sum_{l=-\infty}^{\infty} \int_0^{\infty} dJ_R \int_0^{\infty} \frac{\left(l \frac{\partial f_0^P}{\partial J_R} + m \frac{\partial f_0^P}{\partial J_\phi} \right)}{l\Omega_R + m\Omega_\phi - \omega} \Psi_{l,j}^m \Psi_{l,k}^m dJ_\phi, \\ \frac{M_{jk}^B}{4\pi^2} &= \sum_{l=-\infty}^{\infty} \int_0^{\infty} dJ_R \left[\frac{m f_0^P(J_R, 0) \Psi_{l,j}^m \Psi_{l,k}^m}{l\Omega_R + m\Omega_\phi - \omega} \right]_{J_\phi=0}. \end{aligned} \quad (27)$$

Because $\Omega_\phi = \Omega_R/2$ for radial orbits, the denominator of the boundary integral (27) reduces to $-\omega$ for $l = -m/2$ when m is even. Modes with $\omega = 0$ are therefore precluded when $f_0^P(J_R, 0) \neq 0$.

The DF (23) also drops abruptly to zero at the circular orbit limit $J_R = 0$ so that one should also include an extra $H(J_R)$ factor in the DF. Differentiation of f_0 with respect to $J_R = 0$ then gives a $\delta(J_R)$. However it gives no boundary integral, regardless of the value of f_0 at $J_R = 0$. The reason is that its integrand, for which the $m f_0^P(J_R, 0)$ of equation (27) is replaced by $l f_0^P(0, J_\phi)$, vanishes at $J_R = 0$. That is because the Fourier coefficients $\Psi_{l,j}^m(0, J_\phi)$ vanish for all $l \neq 0$ because orbits with $J_R = 0$ are circular, and the non-zero $l = 0$ Fourier coefficient is annulled by its factor l . Hence the simpler form (23) of the DF suffices.

The boundary integral (27) does not arise if the unperturbed DF contains no radial orbits so that $f_0^P(J_R, 0) = 0$. This is the case for the unidirectional DFs used by Zang (1976) and Evans & Read (1998a,b) because their DFs contain positive powers of $L = J_\phi$ as factors, and hence contain no radial orbits. As Gerhard (1991) discusses for the analogous problem of spherical systems, most DFs for soft-centered potentials which produce finite densities in their centers tend to isotropy there and so have radial orbits. His reasoning applies to thin disks too. Only orbits with low angular momenta penetrate near the center, and the only alternative to radial orbits is a singular distribution of non-radial orbits, as when all orbits are circular. The need for radial orbits to provide a non-zero central density disappears if the density of the stellar disk drops to zero in the center, as it does with the cutouts which we discuss in §2.5.

The omission of the boundary integral terms (27) by Hunter (1992) invalidates his results. As we shall see in §3.1, the effect is substantial. Pichon & Cannon (1997) confirmed Hunter's results, but they repeated his error. We show in Appendix A that the omission means neglecting the contributions to potential energy and angular momentum due to the perturbation of radial orbits. Boundary integrals are avoided if one uses a Lagrangian instead of an Eulerian perturbation theory (Kalnajs 1977). The two forms of the theory are complementary and Appendix A also shows how relevant Lagrangian results can be derived simply from the Eulerian theory. That analysis does not apply to the quite different version of Lagrangian theory used by Vandervoort (1999).

2.3. Angular Momentum and Energy

The perturbations to angular momentum and energy are of second order, and so their calculation requires second order terms from (6). Summing the contributions from all elements in phase space gives

$$\mathcal{L} = \iint J_\phi f d\mathbf{J} d\Theta, \quad (28)$$

for the total angular momentum, and

$$\mathcal{K} = \frac{1}{2} \iint (v_R^2 + v_\phi^2) f d\mathbf{J} d\Theta, \quad (29)$$

for the total kinetic energy. To compute the gravitational energy, we must distinguish between the part V_0^D of the unperturbed gravitational potential V_0 which is due to

the stars of the disk, and the remainder V_0^{ext} which is provided by some other and external source. The perturbational terms V_j , $j > 0$, of the potential all arise from the perturbed DF, and so all belong to the self-gravitational potential. The double contribution of the external potential to the gravitational energy is taken care of by writing it as

$$\mathcal{W} = \frac{1}{2} \iint (V + V_0^{\text{ext}}) f d\mathbf{J} d\Theta. \quad (30)$$

The leading corrections to these quantities due to the perturbations are of second order because all the first order terms vanish when integrated over θ_ϕ . They are

$$\mathcal{L}_2 = \iint J_\phi f_2 d\mathbf{J} d\Theta, \quad (31)$$

$$\mathcal{K}_2 = \iint (\mathcal{H}_0 - V_0) f_2 d\mathbf{J} d\Theta = \mathcal{K}_{2,1} + \mathcal{K}_{2,2}, \quad (32)$$

after using the zeroth order part of equation (4) to substitute for the velocities, and, because $V_0^{\text{ext}} = V_0 - V_0^{\text{D}}$,

$$\begin{aligned} \mathcal{W}_2 &= \frac{1}{2} \iint (V_2 f_0 + V_1 f_1 - V_0^{\text{D}} f_2 + 2V_0 f_2) d\mathbf{J} d\Theta, \\ &= \frac{1}{2} \iint f_1 V_1 d\mathbf{J} d\Theta + \iint V_0 f_2 d\mathbf{J} d\Theta, \\ &= \mathcal{W}_{2,1} + \mathcal{W}_{2,2}. \end{aligned} \quad (33)$$

The component $\mathcal{W}_{2,2} = -\mathcal{K}_{2,2}$ because their defining integrals match. The step to the second line of equation (33) uses the fact that the first and third terms on the first line cancel. This is seen by transforming the integrations to $(\mathbf{x}, \mathbf{v})$ space, and then using equations (2), (8) for $j = 2$, and Poisson integrals like (7) to relate potentials to densities, to express them as identical integrals of the product of the densities Σ_{D} and Σ_2 .

The simplest integral to calculate is that for the first component of $\mathcal{W}_{2,1}$ in equation (33). As we noted earlier, the physical parts of our solutions are given by the real parts of our complex solutions. The real parts of f_1 and V_1 are given by the sums $\frac{1}{2}(f_1 + \bar{f}_1)$ and $\frac{1}{2}(V_1 + \bar{V}_1)$, where a bar denotes a complex conjugate. Therefore

$$\begin{aligned} \mathcal{W}_{2,1} &= \frac{1}{8} \iint (f_1 + \bar{f}_1)(V_1 + \bar{V}_1) d\mathbf{J} d\Theta, \\ &= \frac{1}{8} \iint (f_1 \bar{V}_1 + V_1 \bar{f}_1) d\mathbf{J} d\Theta, \\ &= e^{2st} \sum_{l=-\infty}^{\infty} W_{2,1}^l, \end{aligned} \quad (34)$$

where the components $W_{2,1}^l$ are defined by

$$\begin{aligned} W_{2,1}^l &= \pi^2 \int d\mathbf{J} \left(l \frac{\partial f_0}{\partial J_R} + m \frac{\partial f_0}{\partial J_\phi} \right) \\ &\quad \times \frac{[l\Omega_R + m(\Omega_\phi - \Omega_p)] |\tilde{V}_l|^2}{|l\Omega_R + m\Omega_\phi - \omega|^2}. \end{aligned} \quad (35)$$

We have used here, and will again, the fact that the only products which do not vanish on integration over θ_ϕ are those which pair a conjugate with a non-conjugate quantity.

We show in appendix B that

$$\mathcal{L}_2(t) = e^{2st} \sum_{l=-\infty}^{\infty} L_2^l, \quad \mathcal{K}_{2,1}(t) = e^{2st} \sum_{l=-\infty}^{\infty} K_{2,1}^l, \quad (36)$$

where the components of these sums are defined by the integrals

$$\begin{aligned} L_2^l &= -m\pi^2 \int d\mathbf{J} \left(l \frac{\partial f_0}{\partial J_R} + m \frac{\partial f_0}{\partial J_\phi} \right) \\ &\quad \times \frac{|\tilde{V}_l|^2}{|l\Omega_R + m\Omega_\phi - \omega|^2}, \end{aligned} \quad (37)$$

$$\begin{aligned} K_{2,1}^l &= -\pi^2 \int d\mathbf{J} \left(l \frac{\partial f_0}{\partial J_R} + m \frac{\partial f_0}{\partial J_\phi} \right) \\ &\quad \times \frac{(l\Omega_R + m\Omega_\phi) |\tilde{V}_l|^2}{|l\Omega_R + m\Omega_\phi - \omega|^2}. \end{aligned} \quad (38)$$

To each of the area integrals (35), (37) and (38) must be added the boundary integrals given by the delta function term of equation (24) for the prograde DF (23). The integrals can be combined to give the simple relation

$$K_{2,1}^l + W_{2,1}^l = \Omega_p L_2^l, \quad (39)$$

between the separate components. Although each of these components can be found directly from the first order solution, computing $\mathcal{W}_{2,2} = -\mathcal{K}_{2,2}$ requires more, as we show in Appendix B.

The second order corrections $\mathcal{L}_2$ and $\mathcal{W}_{2,1}$ to the angular momentum and gravitational energy have simple representations in terms of the real and imaginary parts of the matrix $\mathbf{M} = \mathbf{M}_R + i\mathbf{M}_I$. Combining equations (37) and (35) with expansion (19), we get the quadratic forms

$$\mathcal{L}_2(t) = -\frac{m}{4s} e^{2st} \bar{\mathbf{c}}^T \mathbf{M}_I \mathbf{c}, \quad (40)$$

$$\mathcal{W}_{2,1}(t) = \frac{1}{4} e^{2st} \bar{\mathbf{c}}^T \mathbf{M}_R \mathbf{c}, \quad (41)$$

where the superscript T denotes transposition of the column vector $\mathbf{c}$ to generate a row vector. These expressions are real because the matrix $\mathbf{M}$ is symmetric. Moreover multiplying equation (20) by $\bar{\mathbf{c}}^T$ and separating real and imaginary parts ($\mathbf{D}$ is also symmetric) gives

$$\bar{\mathbf{c}}^T \mathbf{M}_R \mathbf{c} = \bar{\mathbf{c}}^T \mathbf{D} \mathbf{c}, \quad \bar{\mathbf{c}}^T \mathbf{M}_I \mathbf{c} = 0. \quad (42)$$

The second relation shows that $\mathcal{L}_2(t) = 0$, which it must be because the disk is not subject to any external torques, and hence its angular momentum is conserved. There is no such restriction on sizes of the separate Fourier components represented by the terms for different l in the sum (36), other than that they must sum to zero. Similarly the sizes of the different components of the potential and kinetic energy vary, because only the total energy $\mathcal{E}$ is constrained to be zero, with $\mathcal{E}_2(t) = \mathcal{K}_2(t) + \mathcal{W}_2(t) = \Omega_p \mathcal{L}_2(t) = 0$. Hence the two components of the kinetic and potential energy are related in the same way:

$$\mathcal{K}_{2,1}(t) = -\mathcal{W}_{2,1}(t), \quad \mathcal{K}_{2,2}(t) = -\mathcal{W}_{2,2}(t). \quad (43)$$

The fact that the second order components grow twice as fast as those of first order is not paradoxical. It reflects the fact that our analysis can describe only the early stages of the growth of an instability. If we introduce a small ordering parameter ε into our expansion $f = f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + \dots$ of the DF to measure the size of the perturbation relative to that of the unperturbed state, then we see that the linearization breaks down,

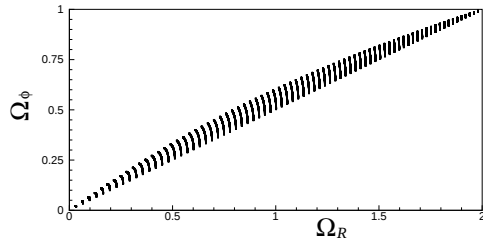


FIG. 1.— The orbital frequency space of the soft-centered logarithmic potential (53) in units of v_0/R_C . The curves are contours of constant E in equal steps in e^{-E} . They show that Ω_R depends only weakly on the angular momentum L , which varies from zero on the radial orbit boundary $\Omega_\phi = 0.5\Omega_R$, to $L_c(E)$ on the curved upper circular orbit boundary. Similarly narrow lens-shaped plots are obtained for other soft-centered potentials. When normalized to the same frequency range as here, the largest value of $\Omega_i = \Omega_\phi - \Omega_R/2$, which is 0.106 here, is 0.130 for Kuzmin's disk and 0.119 for the isochrone. Ω_R depends only on E for the isochrone, so that curves of constant E are then exactly straight.

and our analysis is unreliable, after a time t such that $\varepsilon e^{st} = O(1)$. The second order components are then of magnitude $\varepsilon^2 e^{2st}$, which is also $O(1)$. Our perturbation theory has then ceased to be valid. It is useful only when our expansion is well-ordered, that is for times for which εe^{st} is small.

2.4. Orbits and Polyachenko's Unified Theory

Figure 1 shows an orbital frequency space for prograde orbits. Its shape is characteristic of those of other soft-centered potentials. The upper curved boundary is formed by circular orbits for which $\Omega_R = \kappa(R)$ is the epicyclic frequency. The straight lower boundary is formed by radial orbits for which $\Omega_\phi = \Omega_R/2$. The slimness of the region in between these boundaries, where all the intermediate orbits lie, shows how limited is the range of

$$\Omega_i = \Omega_\phi - \Omega_R/2, \quad (44)$$

for orbits of *all* geometrical shapes.

Polyachenko's (2004, 2005) unified theory of bar-like and spiral modes supposes that global galactic structure can be understood in terms of low frequency modes for which $|\Omega_p - \Omega_i|$ is small. Lynden-Bell (1979) showed that orbits which he classified as abnormal then align with, and hence support, bar-like perturbations. Polyachenko (2004) shows further that the Fourier series (11) for the perturbed DF is then given to leading order by the single $l = -1$ component. This approximation is equivalent to dropping all but the $l = -1$ term from the matrix $\mathbf{M}(2, \omega)$. We test its accuracy by comparing the pattern speeds using it with those obtained from the full matrix; the results are listed in the rightmost columns of our tables. Polyachenko finds eigensolutions by a different method, which has the advantage that he is able to derive a linear eigenvalue problem for ω , albeit one which must be solved for an unknown function of the two action variables. The calculations in Polyachenko (2005) add in additional l values, most importantly $l = 0$ and $l = 1$.

2.5. Disks with Central Holes

Kormendy (1977) proposed modeling stellar disks with an inner-truncated exponential so that they have holes in their centers. We construct such disks by applying a

cutout function $H_{\text{cut}}(L)$ to change an unperturbed DF $f_0(E, L)$ to $H_{\text{cut}}(L)f_0(E, L)$. In so doing, we assume that the part of the gravitational potential which the cutout stars had previously provided, is instead provided by a central bulge which is so hot that it does not respond to disturbances in the disk.

Cutouts were introduced first in Zang's (1976) pioneering study of the singular scalefree logarithmic potential. His purpose was to eliminate stars with short dynamical time scales and high dynamical frequencies. Dynamical frequencies are bounded in soft-centered potentials are bounded (cf Figure 1) and so avoid that problem. With some models, we apply a cutout function

$$H_{\text{cut}}(L) = 1 - e^{-(L/L_0)^2}, \quad (45)$$

where L_0 is some angular momentum scale. Like Zang's, it removes stars for which L is significantly less than L_0 , but has no effect on stars with $L \gg L_0$. It removes stars on radial and near-radial orbits, as well as stars of low energy because they too have low angular momenta. The result of the cutout (45) is to give an active surface density

$$\Sigma_{\text{act}} = \int H_{\text{cut}}(L) f_0(E, L) d\mathbf{v}, \quad (46)$$

which tends to zero at the center, and hence models a central hole. No boundary integrals (27) arise for such cutout unidirectional disks because their unperturbed DF vanishes at $L = 0$. The active densities corresponding to the cutout (45) is shown in several Figures. They show that it truncates far less sharply than Kormendy's exponential.

3. M=2 MODES OF DISKS

This section presents our numerical results. We have found two modes for most models. Classifying them is less straightforward than it is for simpler physical systems, for which the fundamental mode has the lowest frequency and simplest structure, and modes of successively higher order have higher frequencies and more complex structure. The most important instability is that with the highest growth rate. Often it also has the simplest structure. However, we find that small changes in the orbital population have a much greater effect on relative growth rates than they do on radial structure, with the result that the mode with the simplest structure is not always the fastest growing. For that reason we choose structure as our criterion for determining which modes are fundamental and which secondary. We label them 1 and 2 respectively.

We display contour plots of the perturbed density

$$\Sigma_1 = P(R) \cos[2\phi - \omega t + \vartheta(R)], \quad (47)$$

obtained from the real part of equation (16). The phase $\vartheta(R)$ is arbitrary to within an additive constant because eigenvectors are arbitrary to within a complex constant multiple, and so modes are oriented arbitrarily. We draw only the contours for positive perturbed density; those for negative levels have the same pattern rotated by 90° , and occupy the blank sectors. The levels of the contours are in steps of 10% of the maximum of Σ_1 from 10% to 90%. The length scale of all plots is that of the core radius R_C of the potential.

Solid and dotted circles on the contour plots mark the radii R_{CR} and R_{OLR} of circular orbits in co-rotation resonance (CR) and outer Lindblad resonance (OLR) respectively with a neutral $s = 0$ mode with the pattern speed Ω_p of that mode. All pattern speeds are too large for there to be an inner Lindblad resonance (ILR). Orbits of any shape, not just circular ones, may be resonant. For example, the orbits which are in a CR with a pattern speed Ω_p are those which lie on the horizontal line $\Omega_\phi = \Omega_p$ which cuts through the lens-shaped region of Figure 1. They are spread out in space and not confined to one specific circle. They are more concentrated near that circle when most orbits are near-circular, but not otherwise. Similarly the orbits in an OLR are those for which $\Omega_\phi = \Omega_p - \Omega_R/2$, and lie on a line through Figure 1 with slope $-1/2$. They too range from circular to radial. They have lower orbital frequencies than those in the CR because they lie further out in the disk. An ILR can occur only for the limited range of Ω_p values for which the line $\Omega_\phi = \Omega_p + \Omega_R/2$ of slope $1/2$ intersects the lens-shaped region of Figure 1. Our unstable modes with growth rates $s > 0$ have no resonances, only near-resonances at which the denominators of the matrix components (21) are small when s is small.

We next plot the radial variations of the amplitude $P(R)$ of the perturbed density (full curve) and the unperturbed density (dashed curve). Below this is a bar chart which displays the Fourier components L_2^l , $K_{2,1}^l$ and $W_{2,1}^l$. They are evaluated after normalizing the eigenvector $\mathbf{c}$ so as to make the positive and negative components of $e^{-2st}\Omega_p\mathcal{L}_2$ sum to ± 1 respectively. We find that $K_{2,1}$ is always positive. This does not necessarily imply that all modes release gravitational energy and convert it to kinetic energy. Such a release occurs only if the sum of the two components $K_{2,1} + K_{2,2}$ is positive. Its sign follows from the tabulated values of the ratio $K_{2,2}/K_{2,1}$. A mode converts gravitational energy to kinetic energy if this ratio exceeds -1, and vice versa if the ratio is less than -1. There is no conversion if the ratio is exactly -1.

3.1. Kuzmin Disks

The potential of Kuzmin's disk is

$$V_0(R) = -\frac{GM}{\sqrt{R^2 + R_C^2}}, \quad (48)$$

where R_C is the core radius (Kuzmin 1956, Binney & Tremaine 1987). We describe modes of two classes of self-consistent DFs, in units in which $G = M = R_C = 1$.

Miyamoto Models

Miyamoto (1971) models have a single parameter n_M . The orbital population becomes more nearly circular with increasing n_M , and ultimately cold in the limit $n_M \rightarrow \infty$. We use models for which all orbits circulate in the same direction (Hunter 1992).

The frequencies of the fundamental modes for the self-consistent ($L_0 = 0$) disks in Table 1 differ substantially from those given by Hunter (1992) and Pichon & Cannon (1997). Those results omit the boundary integral $\mathbf{M}^B$ and are incorrect. Neglecting $\mathbf{M}^B$ gives a pattern speed of $\Omega_p = 0.357$ with $R_{\text{CR}} = 1.717$ and a growth rate of

$s = 0.295$ for $n_M = 3$. The true fundamental mode of the $n_M = 3$ model is the compact and rapidly growing central bar shown in Figure 2a, not the larger bar shown in Figure 11 of Pichon & Cannon. The amplitude $P(R)$ of its perturbed density in Figure 2c has a single peak. The secondary mode is slower growing, slower propagating, and more extensive. It has a double-peaked amplitude, and a more spiral structure which is also largely confined within the CR circle. It is plotted in JH and resembles the secondary mode shown in the right panels of Figure 3. Growth rates and pattern speeds increase with increasing n_M , modes become increasingly centrally concentrated, and the $W_{2,2}$ terms reinforce the transfer from gravitational to kinetic energy.

The bar chart in Figure 2e displays a standard pattern which is common to all but two of those we show. We show in §4.1 how this pattern can be understood from the formulae of §2.3. Angular momentum is lost by the $l < 0$ Fourier components, primarily $l = -1$, and gained by the $l \geq 0$ components. All Fourier components lose $W_{2,1}$ gravitational energy, much from the $l = -1$ component, while all components gain $K_{2,1}$ kinetic energy except for $l = -1$. Every bar chart shows that a few Fourier components are significant for most of the transfer of angular momentum and energy.

The second block of results in Table 1 shows the effect of applying the cutout function (45) with $L_0 = 0.2$ to the same models. The removal of many low angular momentum orbits reduces the active mass of the disk by a little more than 10%. Its density now drops gradually to zero in the center, but there is no sharp barrier. The effect on the fundamental mode is large, as the right panels of Figure 2 show. Its amplitude is still single peaked, though that peak has moved out to $R \approx 0.4$. The mode now rotates so rapidly that no orbits are in CR, and it extends out to the OLR circle. Its bar chart is quite different, but is similar to that of Figure 11e for another fundamental mode of a cutout disk. The $l = -1$ components are small, and the flow of angular momentum and $K_{2,1}$ is from the $l < 1$ components, primarily $l = 0$, and to the $l \geq 1$ components, primarily $l = 1$.

The cutout reduces the growth rate of the fundamental mode so much that it no longer grows fastest. The now faster growing secondary mode is changed much less. The cutout smoothens its spiral, and moves its two humps outwards. The inner hump in the region where there is large mass reduction, is diminished relative to the outer hump. The cutout affects the fundamental mode so much more than the secondary mode because of its large effect on the more central orbits which are the major participants in the fundamental mode but not in the secondary mode. Both modes of the cutout $n_M = 3$ model atypically induce a transfer from kinetic to gravitational energy.

The results for higher n_M are similar. None of the fundamental modes have CRs. The pattern speeds of secondary modes decrease, though not enough for there to be ILRs. All growth rates are decreased, though the fundamental mode is still, though barely, the faster growing for $n_M = 7$. As n_M increases and orbits become more circular, the relative proportion of angular momentum absorbed by the $l = 1$ component of the secondary modes increases.

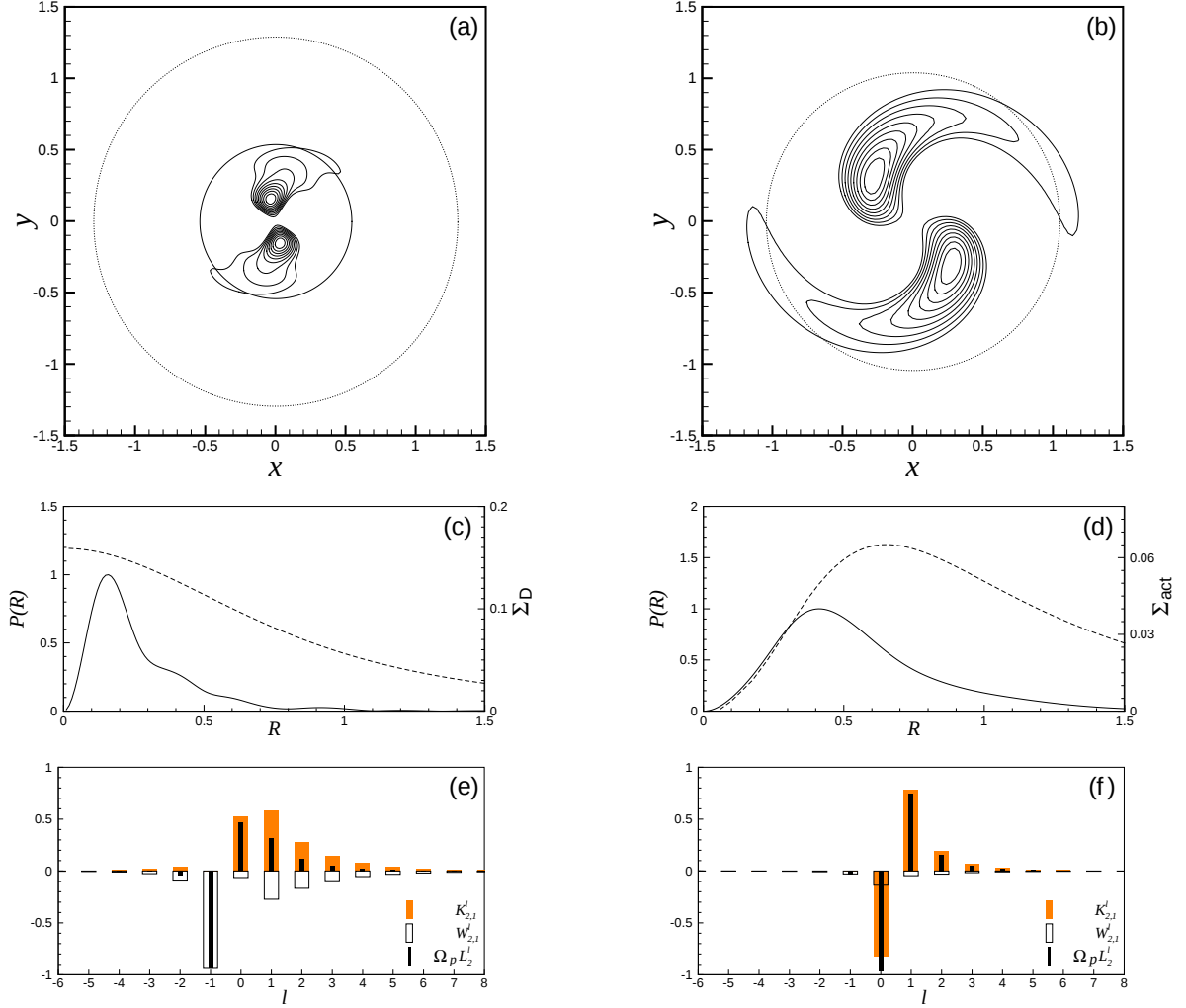


FIG. 2.— The left panels show the fundamental mode of the self-consistent $n_M = 3$ Miyamoto model for Kuzmin's disk. The right panels show how an $L_0 = 0.2$ cutout changes that mode. Here and below, the top panels show positive contours of the perturbed density Σ_1 , in steps of $0.1\Sigma_1$. The solid and dotted circles mark the co-rotation and outer Lindblad resonance circles. The middle panels show the wave amplitude (solid line and left scale) and unperturbed density (dashed line and right scale). The bottom panels show Fourier components of kinetic energy (grey bars), gravitational energy (white bars), and angular momentum (thin bars).

TABLE 1. EIGENVALUES FOR $m = 2$ MODES OF UNIDIRECTIONAL MIYAMOTO MODELS FOR KUZMIN'S DISK.

n_M	L_0	M_{act}	mode	Full Model						$l = -1$ only
				Ω_p	s	$\mathcal{K}_{2,2}/\mathcal{K}_{2,1}$	R_{CR}	R_{OLR}	Ω_p	
3	0	1.000	1	0.825	0.939	1.58	0.541	1.296	0.649	
3	0	1.000	2	0.418	0.265	1.88	1.483	2.246	0.270	
5	0	1.000	1	0.913	1.216	1.61	0.359	1.176	0.738	
5	0	1.000	2	0.530	0.409	1.86	1.154	1.878	0.323	
7	0	1.000	1	0.963	1.465	2.58	0.227	1.115	0.805	
7	0	1.000	2	0.643	0.588	2.64	0.895	1.609	0.372	
3	0.2	0.868	1	1.023	0.114	-3.26	...	1.045	...	
3	0.2	0.868	2	0.336	0.203	-1.67	1.811	2.632	0.259	
5	0.2	0.882	1	1.049	0.222	0.53	...	1.017	...	
5	0.2	0.882	2	0.384	0.259	-0.64	1.607	2.390	0.294	
7	0.2	0.892	1	1.067	0.311	1.27	...	0.997	...	
7	0.2	0.892	2	0.430	0.302	-0.17	1.443	2.200	0.321	

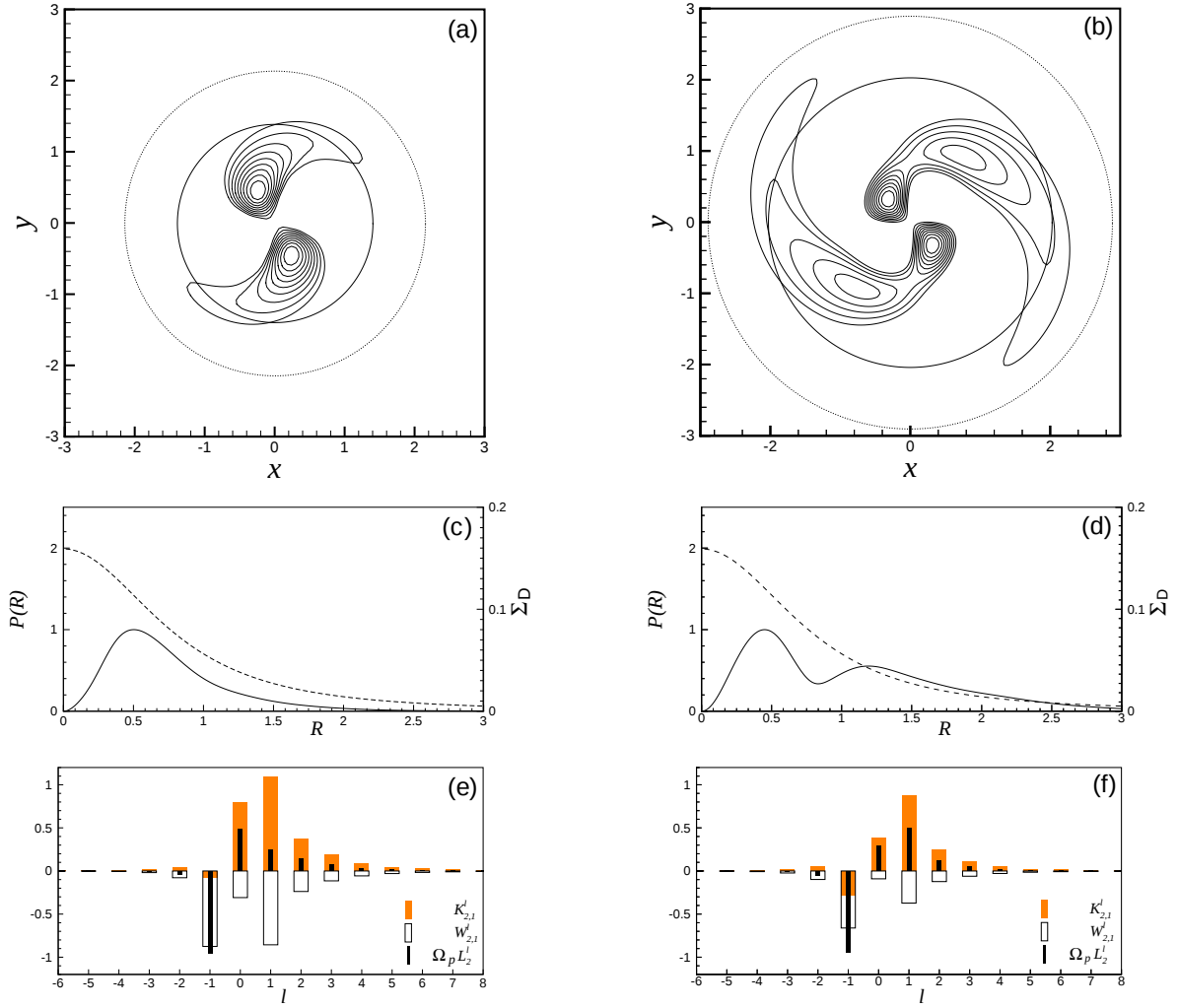


FIG. 3.— The fundamental mode (left panels) and secondary mode (right panels) of a tapered $m_K = 6$, $J_c = 0.25$, AS model with 4.6% of its orbits retrograde.

TABLE 2. EIGENVALUES FOR $m = 2$ MODES OF AS MODIFIED KALNAJS MODELS FOR KUZMIN'S DISK

m_K	β	J_c	M_{retro}	mode	Full Model			AS		Pol ^a		$l = -1$ only	
					Ω_p	s	$\mathcal{K}_{2,2}/\mathcal{K}_{2,1}$	Ω_p	s	Ω_p	s	Ω_p	
4	0	0.40	0.084	1	0.335	0.174	1.83	0.168	0.020	...	...	0.193	
4	0	0.40	0.084	2	0.175	0.016	-1.05	...	...	...	...	0.144	
6	0	0	0	1	0.746	0.711	1.42	...	...	...	...	0.569	
6	0	0	0	2	0.358	0.161	0.60	...	...	...	...	0.240	
6	0	0.25	0.046	1	0.445	0.308	0.52	0.233	0.066	...	...	0.264	
6	0	0.25	0.046	2	0.294	0.109	1.85	0.165	0.058	0.24	0.058	0.207	
6	3	0.60	0.154	2	0.158	0.027	-4.00	0.145	0.014	0.14	0.02	0.144	
8	4	0.90	0.160	1	0.199	0.064	-1.47	0.173	0.035	...	...	0.174	

^aPolyachenko (2004, 2005)

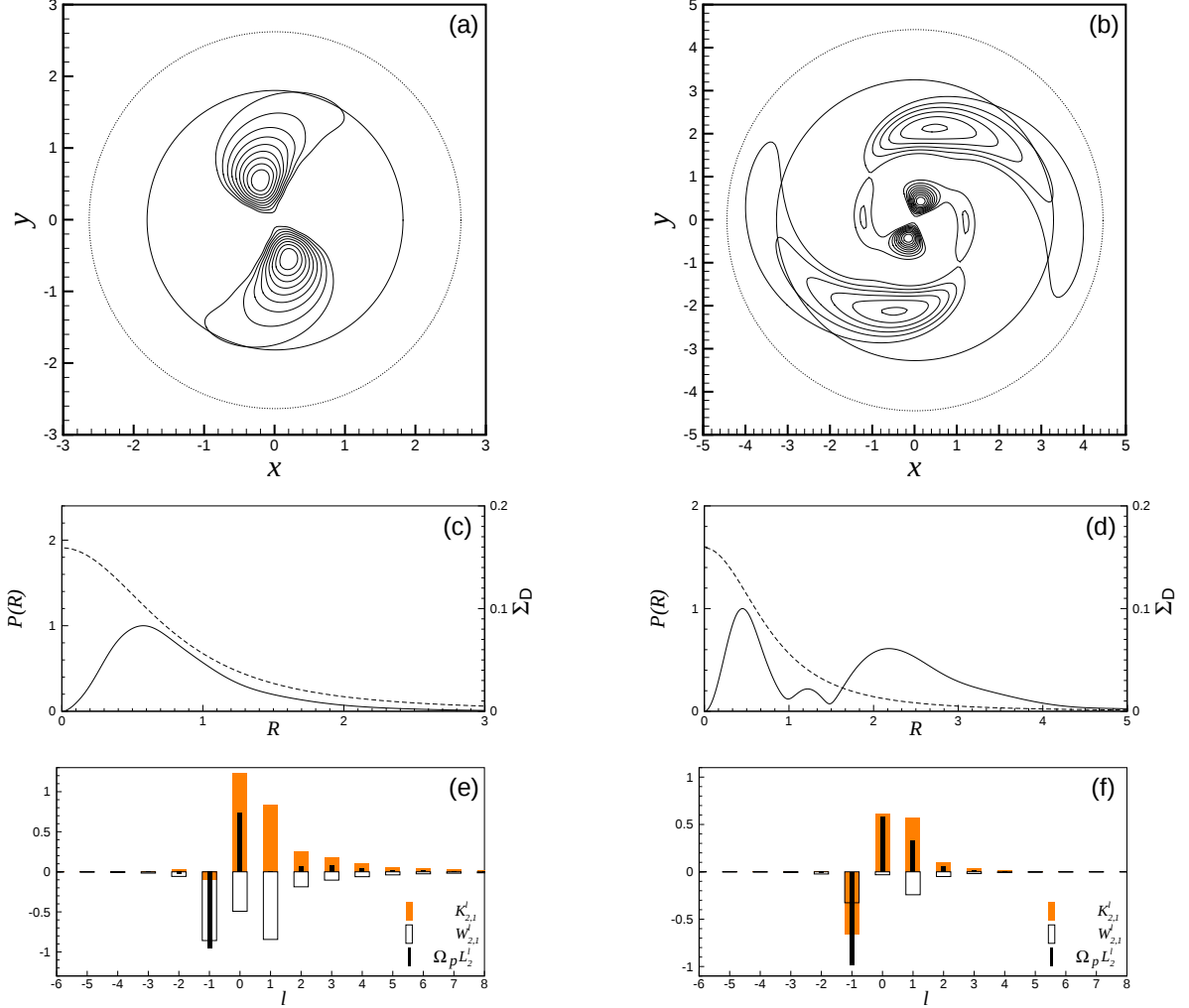


FIG. 4.— The left panels show the fundamental mode of the $m_K = 4$ AS model, tapered with $J_c = 0.4$ and with 8.4% of its orbits retrograde. The right panels show our computed mode for their $\beta = 3$, $m_K = 6$, $J_c = 0.6$ model with 15.4% retrograde orbits.

Kalnajs Models

Toomre's stability parameter Q increases monotonically outwards for Miyamoto models. Kalnajs (1976) models, which also depend on an integer parameter m_K which increases as the models cool, keep Q almost constant. SA modified these models with the addition of two extra parameters: an angular momentum J_c and β . A fractional mass M_{retro} with angular momenta in the range $(0, J_c)$ have their rotations reversed, following equation (5) of Zang & Hohl (1978). This gives a smoothly tapered DF, and eliminates the discontinuity of the unidirectional model. Table 2 gives our results for several AS models, and also a unidirectional $m_K = 6$, $J_c = 0$ model which is new. The fundamental mode of the latter has not been plotted because it is a compact and rapidly rotating bar, a little larger, but otherwise just like that shown in the left panels of Figure 2. The left panels of Figure 3 show how much that compact bar is modified by a $J_c = 0.25$ tapering which reverses less than 5% of the orbits. Like the cutout, tapering makes the fundamental mode smoother, more spiral, more extensive, and reduces its growth rate. But the taper di-

minishes the pattern speed, whereas the cutout increases it. Both tapering and cutout cause a large change in the low angular momentum orbits which dominate in the center of the disk where the fundamental mode is concentrated. Whereas the cutout removes many of them, the tapering merely reverses some rotations, and makes the DF isotropic for $|L| \ll J_c$. The secondary mode in the right panels of Figure 3 has the usual double peak, and more extensive and spiral form. It resembles that for the same model shown in Figure 1 of Polyachenko (2005).

The fundamental mode for the tapered $m_K = 4$, $J_c = 0.4$ model is one of the few for which SA plot a mode shape. The major difference between their Figure 5b and our Figure 4a is that theirs is less spiral and has some central structure which our Figure 4a lacks. This mode is less spiral than that of the cooler and more sharply tapered disk shown in Figure 3a, and there is a remarkable dearth of angular momentum transferred to its $l = 1$ component.

A positive value of β removes orbits with low energies and angular momentum high relative to $L_c(E)$, and adds to those with higher energies and lower angular momen-

tum. It is easy to show from their Appendix A that

$$f_0(E, 0) = \frac{(m_K - 2\beta E_s^2)}{2\pi^2} \exp[(\beta(1 - E_s^2))], \quad E_s = \frac{R_C E}{GM}. \quad (49)$$

The value of β can not exceed $m_K/2$ because f_0 is then negative at the center of the disk where the scaled energy $E_s = -1$. Many of AS's models have marginal values $\beta = -m_K/2$. Their f_0 values peak at intermediate values of E_s . The $(m_K, \beta, J_c) = (6, 3, 0.6)$ model is one to which Polyachenko (2004) applied his simplified theory. Our Figure 4b shows the same concentrated central structure as in Polyachenko's Figure 6, but his plot has a stronger second hump around $R = 1$ than our weaker one, and lacks the extended outer spiral that we find. The structure of this mode suggests that it is a secondary one, though it is the only one we have found. It is closer to having an ILR than any other mode in our survey. Figure 4f is noteworthy for its large negative $l = -1$ component of $\mathcal{K}_{2,1}$; these components are usually small. This mode is one which converts kinetic to gravitational energy.

Table 2 shows discrepancies between our results and those of AS and Polyachenko. Some discrepancies between matrix theory and simulations with finite numbers of particles and gravity softening are to be expected. Further comparisons showed that discrepancies increase as J_c decreases. This leads us to suspect that they are due in large part to the marked sensitivity to the population of near-radial orbits; note how much the $m_K = 6$ mode changes between $J_c = 0$ and $J_c = 0.25$. We found also that eigenvalues are sensitive to the accuracy with which the central regions are handled. Results for isochrone disks in Table 3 below are in excellent agreement with those of others. These models do not have tapers. SA found the mode of a cold disk accurately, but its orbits are all circular and not at all radial.

3.2. Isochrone Disks

The isochrone disk has the potential

$$V_0(R) = -\frac{GM}{R_C + \sqrt{R^2 + R_C^2}}, \quad (50)$$

where R_C is again its core radius. We again use units in which $G = M = R_C = 1$. Kalnajs's (1976) self-consistent models for this disk contain an integer parameter m_K which increases as the models cool. They have fairly uniform values of the Toomre's Q . Some stars are reversed to become retrograde according to the prescription given in equation (13) of Earn & Sellwood (1995). Combining that with the $x \rightarrow 0$ limit of equation (26) of Kalnajs (1976) and using Gradshteyn & Ryzhik (1980) formula (7.126.1) gives

$$f_0 = \frac{m_K}{6\pi^2 [1 + J_R + |J_\phi|]^{2m_K-2}}, \quad J_\phi < 0, \quad (51)$$

for the DF of retrograde stars. There are now retrograde stars of all angular momenta, not the limited ranges of the models of §3.1. Integration over phase space gives

$$M_{\text{retro}} = \frac{m_K}{3(m_K - 2)(2m_K - 3)}, \quad (52)$$

for their fractional mass.

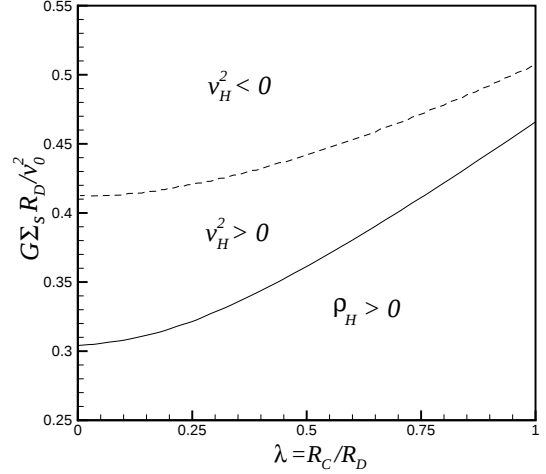


FIG. 5.— The parameter space for cored exponential disks (54) in the soft-centered logarithmic potential (53). Only models on or below the solid line have physically possible spherical halos with density $\rho_H \geq 0$ at all radii.

Table 3 compares results; the agreement is now gratifyingly close. Both pattern speeds, growth rates, and spirality increase with increasing m_K , as they do with Miyamoto models with increasing n_M . Again there are no ILRs. ILRs occur only when $\Omega_p < 0.0593$ in our units, for which the ranges of both Ω_R and Ω_ϕ are only a half of what they are in Figure 1.

The wave patterns for the two modes for $m_K = 12$ are displayed in Earn & Sellwood's Figure 1 and also in JH. As usual, the secondary mode is more extended than the fundamental, and the amplitude of its spiral arm is two-peaked, versus the single peak of the fundamental mode. Our bar charts, which are given in JH, show that the $l = 1$ component absorbs much more angular momentum and kinetic energy than the $l = 0$. This is different from Figure 3, but is in line with the trend we noted with Miyamoto models with increasing n_M . The secondary mode of the warmest $m_K = 6$ model is the only one for which kinetic is converted to gravitational energy. The mode in Figure 15 of Vauterin & Dejonghe (1996), for a model which also has some retrograde orbits though a different DF and potential, is well matched by the fundamental mode of the $m_K = 9$ isochrone.

3.3. Exponential Disks

The soft-centered logarithmic potential, widely used in galactic studies because of its near-flat rotation curve, is

$$V_0(R) = v_0^2 \ln \sqrt{1 + R^2/R_C^2}, \quad (53)$$

where v_0 is the flat rotation velocity, and R_C is the core radius. Densities of the stellar disks of spiral galaxies are found to decay exponentially with distance (Freeman 1970), and are quite unlike the self-consistent density for the potential (53), given in equation (51) of JH, which decays only as R^{-1} at large distances like the singular Mestel (1963) disk. We study disks with densities

$$\Sigma_D(R) = \Sigma_s \exp \left[-\lambda \sqrt{1 + R^2/R_C^2} \right], \quad \lambda = R_C/R_D, \quad (54)$$

TABLE 3. EIGENVALUES OF KALNAIS MODELS FOR THE ISOCHRONE DISK.

m_K	M_{retro}	mode	Full Model		K78 ^a		PC97 ^b		$l = -1$ only	
			Ω_p	s	Ω_p	s	Ω_p	s	Ω_p	
6	0.056	1	0.169	0.080	-0.11	0.170	0.075	0.170	0.075	0.094
6	0.056	2	0.121	0.035	-1.58	...	...	...	...	0.079
9	0.029	1	0.235	0.149	-0.05	0.235	0.145	0.235	0.145	0.127
9	0.029	2	0.183	0.089	-0.92	...	...	...	...	0.109
12	0.019	1	0.292	0.217	0.07	0.295	0.210	0.295	0.210	0.150
12	0.019	2	0.234	0.148	-0.60	...	...	0.230	0.145	0.133

^aKalnajs (1978).^bPichon & Cannon (1997).TABLE 4. EIGENVALUES FOR $m = 2$ MODES OF EXPONENTIAL DISKS WITH $N = 6$.

mode	R_D	$\Sigma_s R_D$	L_0	M_{act}	Full Model				$l = -1$ only	
					Ω_p	s	R_{CR}	R_{OUT}	Ω_p	
1	1	0.42	0	1.000	0.914	1.151	0.444	1.688	0.698	
2	1	0.42	0	1.000	0.693	0.324	1.040	2.334	0.364	
1	1.2	0.38	0	1.000	0.840	0.935	0.646	1.871	0.591	
2	1.2	0.38	0	1.000	0.607	0.059	1.309	2.701	0.330	
1	1.4	0.36	0	1.000	0.805	0.794	0.737	1.967	0.540	
2	1.4	0.36	0	1.000	0.467	0.179	1.893	3.572	0.294	
1	1.6	0.34	0	1.000	0.768	0.642	0.834	2.077	0.488	
2	1.6	0.34	0	1.000	0.443	0.119	2.024	3.775	0.277	
1	1	0.42	0.1	0.967	1.170	0.259	...	1.212	...	
2	1	0.42	0.1	0.967	0.537	0.288	1.571	3.082	0.331	
1	1	0.42	0.3	0.913	1.044	0.228	...	1.423	...	
2	1	0.42	0.3	0.913	0.439	0.274	2.047	3.811	0.302	

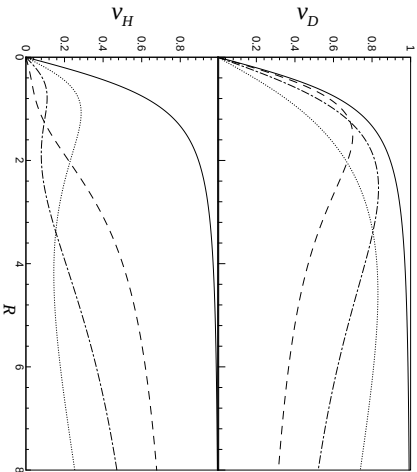


FIG. 6.— The rotational velocities due to the cored exponential disk (54) (upper panel) and its halo/bulge (lower panel), in units in which $R_C = v_0 = 1$. The dashed, dot-dashed, and dotted lines correspond to $(R_D, G\Sigma_s R_D) = (0.5, 0.6)$, $(1, 0.42)$ and $(2, 0.32)$, respectively, while the solid line shows the circular velocity of the soft-centered logarithmic potential (53).

where R_D is the length-scale of the exponential decay, and Σ_s is a density scale. The central density of the disk is $\Sigma_s e^{-\lambda}$ and its total mass $M_D = 2\pi\Sigma_s R_D(R_C + R_D)e^{-\lambda}$. Like Vauterin & Dejonghe (1996), we have included the core radius R_C of the potential in the density (54) so as to avoid the mildly singular behavior of the uncored exponential disk $\Sigma_D = \Sigma_s \exp(-R/R_D)$. The difference by which the potential (53) exceeds that due to the disk (54) must be provided by some external source, halo or bulge. Figure 5 shows how parameter space is restricted by the requirement that the external source be spherical.

Figure 6 plots rotational velocities for some feasible disks. DFs for exponential disks (54) which contain an integer parameter N are given in equation (C3). Figure 7 plot some of their basic properties. These models are our most varied. After normalizing to units in which $G = v_0 = R_C = 1$, three parameters remain: the length scale R_D and density scale Σ_s of the disk, and N . A cutoff introduces a further parameter L_0 . N measures the tendency of the orbits to circularity because, as Figure 7a shows, disks become increasingly centrifugally supported with increasing N . Hence N plays a role similar to those of the parameters m_M and m_K of previous

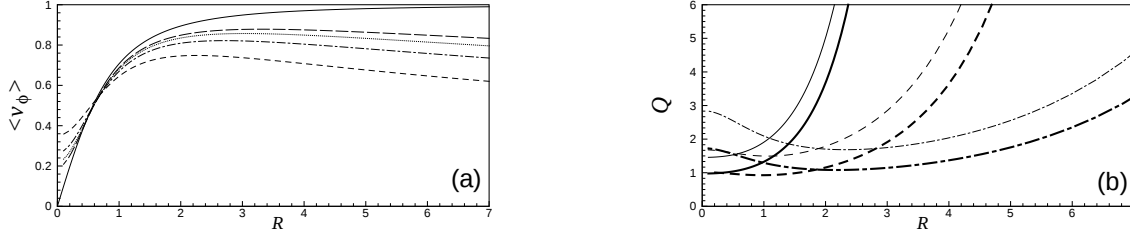


FIG. 7.— (a) Circular velocity (solid line) and mean rotation velocities of the exponential disk. Dashed, dot-dashed, dotted and long-dashed lines respectively show the mean rotation velocity $\langle v_\phi \rangle$ of exponential disks with $N = 2, 4, 6$ and 8 . All curves are for $R_C = R_D = v_0 = 1$. (b) Toomre's Q for the exponential disk for $R_C = 1$ with $N = 2$ (thin lines) and $N = 6$ (thick lines). Solid, dashed and dot-dashed lines here correspond to $(R_D, G\Sigma_s R_D) = (0.5, 0.6), (1, 0.42)$ and $(2, 0.32)$, respectively as in Figure 6.

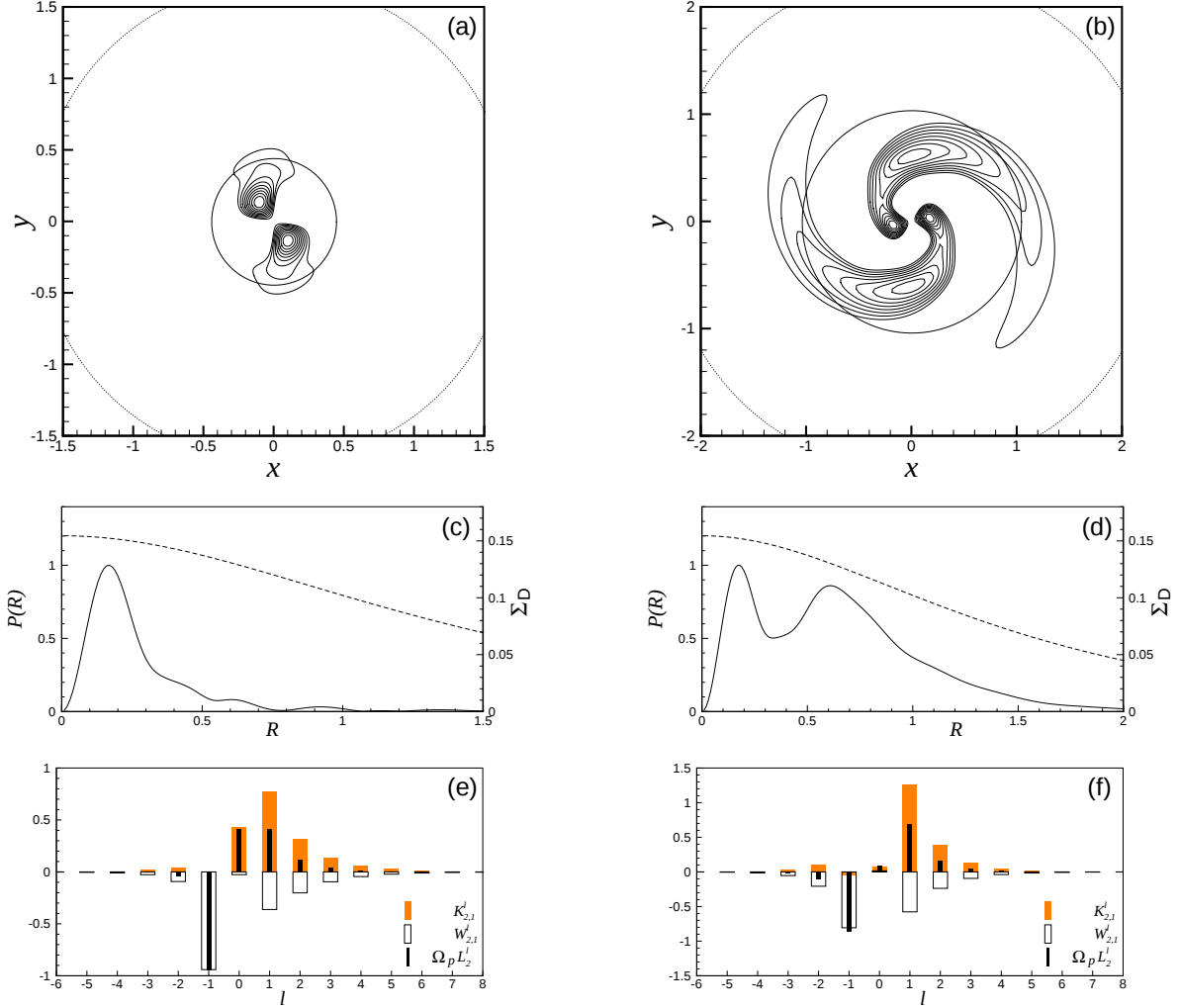


FIG. 8.— Modes of the exponential disk for $N = 6$, $R_D = 1$, and $\Sigma_s R_D = 0.42$. The left panels are for the fundamental mode, and the right panels are for the secondary mode.

models. We varied them, and so now restrict exponential disk models to the single case of $N = 6$ so as to study the consequences of varying other parameters. As Figure 7b shows, Toomre's Q for $N = 6$ is close to unity over a substantial central part of the disk, whose size increases with R_D when $R_D \geq 1$.

Table 4 lists two modes for six different $N = 6$ models. It lacks values of the ratio $\mathcal{K}_{2,2}/\mathcal{K}_{2,1}$ because we were un-

able to compute accurate values of $\mathcal{W}_{2,2}$ with our chosen basis functions, which are those of Clutton-Brock (1972), and Aoki & Iye (1978). They find modes well, but the R^{-3} decay of their σ_j^0 coupled with the growth of the logarithmic potential V_0 causes the integrals in the sum (B17) to grow slowly with increasing j .

Figure 8 shows two modes of the first $R_D = 1$, $\Sigma_s R_D = 0.42$, model. The fundamental mode is a rapidly rotating

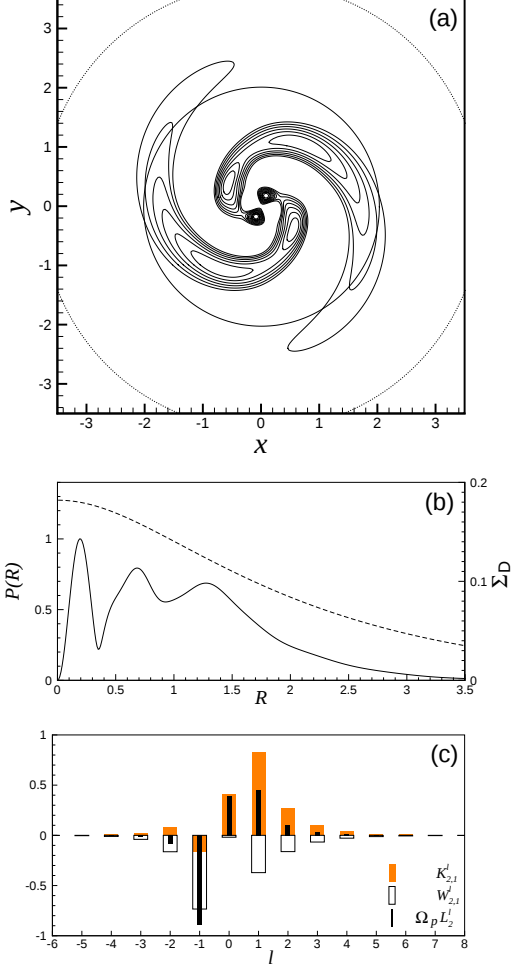


FIG. 9.— The secondary mode of an $N = 6$ exponential disk of larger extent with $R_D = 1.6$ and $\Sigma_s R_D = 0.34$.

bar confined within the CR circle, and resembles that of the Kuzmin disk in Figure 2a. The secondary mode has a lower pattern speed and growth rate, a spiral form, and its amplitude has the usual two peaks within the CR circle. Its bar chart is remarkable for its small $l = 0$ components.

The first four models of Table 4 follow the effect of varying R_D . The parameter $\Sigma_s R_D$ changes too so as to remain within the physically allowed region of Figure 5. Its values are near marginal in that they are approximately 90% of their allowed maximum. Despite the decrease of $\Sigma_s R_D$, the total disk mass grows as its scale length R_D increases, though, as Figure 6 shows, the halo/bulge grows in importance as R_D decreases, and the disk becomes progressively less maximal. The structure of the fundamental bar mode does not change along this sequence; it remains a compact and rapidly rotating central bar. Figure 9 for the larger $R_D = 1.6$ shows that the amplitude of the secondary mode has developed a third hump. This development occurs around $R_D \approx 1.25$, and seems to be related to the anomalously low growth rate at $R_D = 1.2$ (See Table 4). The bar chart of Figure 9c for $R_D = 1.6$ shows much larger $l = 0$ components than Figure 8f.

Figure 10 illustrates how decreasing Σ_s and hence the mass of the disk, whilst keeping its length scale R_D fixed, stabilizes the disk. The transition of the fundamental mode from stability to instability appears to take place through a pitchfork bifurcation; that of the secondary mode through a tangent bifurcation. The order in which the two modes are stabilized is different for the two different values of R_D . When, as here, $\Sigma_s R_D$ is decreased at a fixed R_D , then the part of the rotational velocity due to the halo/bulge components increase from those shown in Figure 6. The stabilization shown in figure 10 is therefore similar to that which is achieved by sufficiently massive halos (Kalnajs 1972, Ostriker & Peebles 1973, Hohl 1975). Decreasing Σ_s boosts the X of Toomre's (1981) swing-amplifier theory to about 1.6 and 2.3 respectively at the CR circle when the fundamental and $R_D = 1$ secondary modes are stabilized. The stabilizing value of X varies with the shear Γ , which is 0.6 and 0.8 at the CR circle in the two cases.

The stabilization of both modes in the neighborhood of $\Sigma_s R_D \approx 0.3$ for both $R_D = 1$ and $R_D = 1.6$ suggests that the stability boundary approaches the boundary of the physically feasible region plotted in Figure 5 as R_D increases, i.e. as λ decreases. It raises the possibility that the two boundaries intersect before the $\lambda \rightarrow 0$ limit of an exponential disk in a singular logarithmic potential, for which $\Sigma_s R_D = 0.304$ is reached. Our current computer algorithms are not capable of approaching that limit, but they do show the $R_D = 2$ disk to be quite stable. This suggests that the classical exponential disk ($\lambda = 0$) with a completely flat rotation curve ($R_C = 0$) may be stable against bisymmetric excitations. This disk is much less than maximal because the everywhere flat rotation curve requires a substantial contribution from a spherical central bulge/halo.

The final Figure 11 and the last two models of Table 4 show again the large effects of cutting out a small fractional mass of low angular momentum orbits. The effects are similar those we found in §3.1 for Miyamoto disks. Comparison with Figure 8 shows that the fundamental mode is changed much more than the secondary one. Its growth rate is diminished substantially, and its pattern speed is increased so much that there is no longer a CR circle. It is more spiral and extensive and its peak amplitude has moved out from the region of diminished density. Its bar chart has undergone a large change and now resembles that of Figure 2f with small $l = -1$ components and the flow of both angular momentum and $K_{2,1}$ is primarily from $l = 0$ to $l = 1$. The more extensive secondary mode is changed much less by the cutout, though the inner hump of its amplitude has almost disappeared. The $l = 0$ components are again significant in the bar chart of Figure 11f, unlike Figure 8f for $R_D = 1$, but like Figure 9c for $R_D = 1.6$. The lesser changes between $L_0 = 0.1$ and $L_0 = 0.3$ show the importance of the transition between $L_0 = 0$ and $L_0 = 0.1$. The singular behavior of e^{-L^2/L_0^2} makes the study of small values of L_0 and the approach to the limit $L_0 \rightarrow 0$ computationally difficult.

We can not compare our results directly with those of Sellwood (1989). Though for the same potential (53), his disks are different, and have the larger range $2 \leq R_D/R_C \leq 8.33$ of radii. His $m = 2$ modes are also

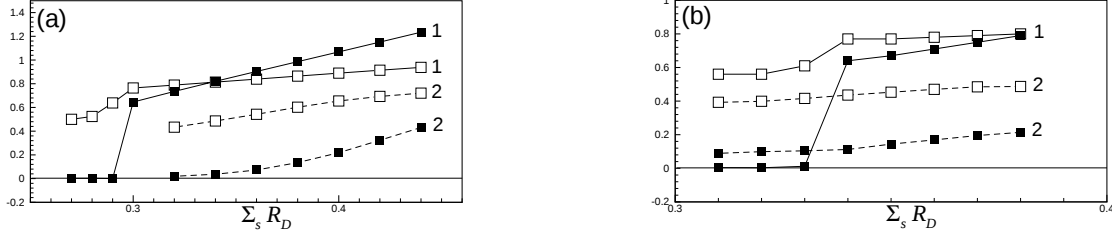


FIG. 10.— (a) The variation of Ω_p (open squares) and s (filled squares) of the $N = 6$, $R_D = 1$ exponential disk as the mass of the disk is varied. Data for the fundamental modes (1) and secondary modes (2) are connected by solid and dashed lines, respectively. The secondary modes are double peaked. (b) Same as panel (a) but for $N = 6$, $R_D = 1.6$. The secondary modes now have triple peaks.

too fast for ILRs. The spiral mode for $R_D/R_C = 5$ shown in his Figure 3 extends out to the OLR circle, and so is more extensive than any of ours. Sellwood finds 0.6 to be the critical value of the parameter $v_0(R_D/GM_D)^{1/2}$. That parameter becomes $v_0[\lambda e^\lambda/2\pi\Sigma_s G R_C(1+\lambda)]^{1/2}$ with our disk. It is 0.86 when stability is achieved for $\Sigma_s R_D = 0.29$ in Figure 10a for $\lambda = 1$, and 0.59 for $\Sigma_s R_D = 0.33$ when only the fundamental mode has stabilized in Figure 10b for $\lambda = 1.6$. Even so, our values are significantly less than the 1.1 which Efsthathiou, Lake & Negroponte (1982) find to be necessary for their N -body experiments, though that larger value may well be needed for nonlinear stability.

4. DISCUSSION

We analyze and explain the properties of the bar charts in §4.1, and discuss the implications of our findings for Polyachenko's (2004) simplified theory in §4.2.

4.1. Transfer of Angular Momentum and Energy

The rate of change of the perturbed angular momentum (B6) contains a factor $s/|\Omega_R + m\Omega_\phi - \omega|^2$, which tends to $\pi\delta[m(\Omega_\phi - \Omega_p) + l\Omega_R]$ as $s \rightarrow 0$. As Lynden-Bell & Kalnajs (1972, hereafter LBK) note, this implies that a neutrally stable wave emits and absorbs angular momentum only at resonances. For the unstable modes we discuss, the factor $s/|\Omega_R + m\Omega_\phi - \omega|^2$ peaks at resonances, sharply so when s is small. This peaking implies that emission and absorption occurs mainly near resonances. Figure 1 shows that, for given pattern speed $\Omega_p > 0$, the orbits associated with $l \geq 0$ resonances lie at successively smaller values of the orbital frequencies Ω_R and Ω_ϕ , and hence successively further out in the disk. That is the reason for interpreting the flows of angular momentum, and also kinetic energy, which are proportional to the quantities shown in the bar charts, as outward flows.

It is evident from equations (35), (37), and (38) that the partial derivatives $\partial f_0/\partial J_R$ and $\partial f_0/\partial J_\phi$ play a major role in determining the signs of the Fourier components shown in the bar charts. All the other terms in L_2^l are either magnitudes or constants. LBK argued that both derivatives are negative for a physically reasonable DF. If this is so, then equation (37) shows that $L_2^l > 0$ for $l \geq 0$. The compensating negative values of L_2^l , necessary because there is no net change of angular momentum, can and must occur for $l < 0$. The fact that L_2^{-1} is negative, as we find, can be accounted for by $|\partial f_0/\partial J_R|$ being much larger than $|\partial f_0/\partial J_\phi|$. The prediction matches the

standard pattern we find in §3. The assumptions that $\partial f_0/\partial J_R < 0$ and $|\partial f_0/\partial J_R| \gg |\partial f_0/\partial J_\phi|$ are generally valid. Figure 6 of Kalnajs (1976) gives examples; the contours of the two unidirectional DFs shown there decrease much more rapidly with increasing J_R than with J_ϕ . However the modified $\beta = m_K/2$ models of §3.1 are exceptional because they have regions in which $\partial f_0(E, L)/\partial E$, and hence $\partial f_0(J_R, J_\phi)/\partial J_R$, are positive.

Though positive values of $\partial f_0/\partial J_R$ are unusual, positive values of $\partial f_0/\partial J_\phi$ are not. They are unavoidable with disks which are totally, or mostly, unidirectional. If $f_0^P(J_R, 0) = 0$, as it is for the unidirectional models of Zang (1976) and Evans & Read (1998a) and our cutout disks, then $\partial f_0/\partial J_\phi$ must be positive for some positive values of J_ϕ because otherwise $f_0(J_R, J_\phi)$ could never become positive for $J_\phi > 0$. Similarly $\partial f_0/\partial J_\phi$ must be positive for some $J_\phi > -J_c$ for the tapered models of Table 2. Furthermore the explicit formula (51) shows that $\partial f_0/\partial J_\phi > 0$ for $J_\phi < 0$ for the isochrone models of §3.2. Although no positive values of $\partial f_0/\partial J_\phi$ are visible in Figure 6 of Kalnajs (1976), these are disks for which the boundary integral

$$-m^2\pi^2 \int dJ_R \left[\frac{f_0^P(J_R, 0)|\tilde{V}_l|^2}{|l\Omega_R + m\Omega_\phi - \omega|^2} \right]_{J_\phi=0}, \quad (55)$$

must then be included in the expression for L_2^l . It is negative for every l , as are the contributions to the integral (37) from positive values of $\partial f_0/\partial J_\phi$. Unidirectional disks for which $f_0^P(J_R, 0) > 0$ can be regarded as $J_c \rightarrow 0$ limits of tapered disks; the boundary integral (55) then accounts for the effect of the large positive values of $\partial f_0/\partial J_\phi$ which occur in the taper.

Despite the occurrence of regions of positive values of $\partial f_0/\partial J_\phi$, they are generally either too small, or confined to too limited regions, to modify the standard bar chart pattern. However, cutout disks have central regions in which $\partial f_0/\partial J_\phi$ is large and positive. That is the reason for the negative values of L_2^0 in the bar charts of Figures 2f and 11e for fundamental modes of cutout disks. The bar charts for the secondary modes of these cutout disks have the standard form, as is seen in Figure 11f. This is because $\partial f_0/\partial J_\phi < 0$ for the more distant orbits which are the more important for the secondary mode.

Integral (38) for $K_{2,1}^l$ differs from integral (37) for L_2^l by a factor $(\Omega_\phi + l\Omega_R/2)$. This is positive for $l \geq -1$ for all direct orbits, but negative for $l \leq -2$ (See Figure 1), and is the reason why the signs of $K_{2,1}^l$ match those of L_2^l for $l \geq -1$, but are their opposites for $l \leq -2$.

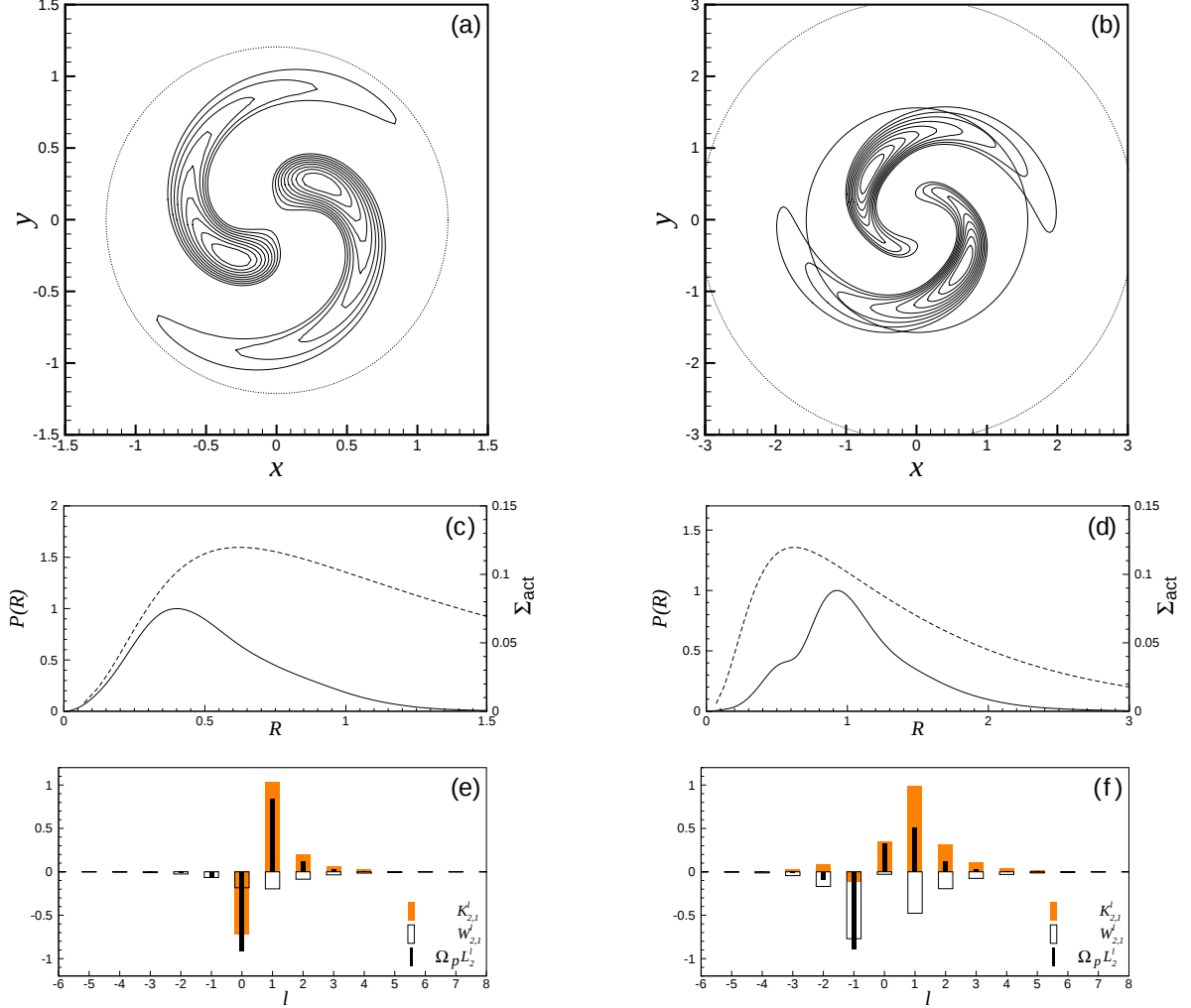


FIG. 11.— Modes of the $N = 6$, $R_D = 1$, and $\Sigma_s R_D = 0.42$ exponential disk with an $L_0 = 0.1$ cutout. The active surface density is shown by the dashed lines in the central panels. The left panels show the fundamental mode and the right panels the secondary mode.

This pattern holds even for the exceptional cases, with the result that $K_{2,1}^l$ is positive for all l , except for $l = -1$ in standard bar charts and $l = 0$ in our exceptional ones.

Integral (35) for $W_{2,1}^l$ differs from that for L_2^l by a factor $(\Omega_p - \Omega_\phi - l\Omega_R/2)$. This factor is positive for all orbits for $l = 0$ modes when the pattern speed $\Omega_p > 1$, which is why $W_{2,1}^0$ is negative like L_2^0 for the exceptional fundamental modes. $W_{2,1}^l$ and L_2^0 otherwise have opposite signs for other modes with slower pattern speeds and $L_2^0 > 0$, and which are dominated by orbits within CR. In fact $W_{2,1}^l$ is negative for all l in all bar charts. This is explained by noting that $(\Omega_p - \Omega_\phi - l\Omega_R/2)$ is negative for modes which lie principally within OLR for $l \geq 1$, and positive for $l \leq -1$, so that the signs of $W_{2,1}^l$ are respectively the opposite and the same as those of L_2^l .

A striking feature of the bar charts is the rapidity with which the quantities displayed there decrease with increasing $|l|$. There are two reasons for this. One is the increase in the denominators of their integrands in the regions where most orbits lie. The $l = -1$ terms are generally important, even though there is no ILR, because the denominator term $2(\Omega_\phi - \Omega_p) - \Omega_R = 2(\Omega_i - \Omega_p)$ is

never large for any orbit. Another reason is the decay of the Fourier coefficients $\tilde{V}_l$ with increasing $|l|$.

4.2. Polyachenko's Theory and Abnormal Orbits

Tables 1 through 4 show consistently that eigenvalues obtained from $l = -1$ components only, as in Polyachenko's (2004) simplified theory, are consistently too low. This is so in Table 3 for isochrone disks for there is little disagreement as to the correct frequencies. Most pattern speeds Ω_p well exceed the narrow range of Ω_i values which the cored potentials allow, and so the assumption of small $|\Omega_p - \Omega_i|$, on which the idea of ignoring all but $l = -1$ components is based, is not then valid. Our bar charts show clearly that a few Fourier components other than $l = -1$ are always significant. Polyachenko (2005) adds $l = 0$ and $l = 1$ components to a Kalnajs model of Table 2 and finds that the $l = 1$ terms have the greater effect on the frequency. Some of our bar charts show $l = 1$ terms which are larger than the $l = 0$ ones, but not all do. We caution against equating $l = -1$, $l = 0$, and $l = 1$ components with ILR, CR, and OLR respectively. Not only are resonances more than circles as we argued in the introductory remarks to §3,

but Fourier components in angle space, which is what the different l components represent, are not the same as resonances. There is indeed an inter-relation, but they are conceptually distinct entities. The $l = -1$ components are almost always highly significant, even though no orbits of any shape or size are in ILR in any of our modes. Also the amplitudes of most modes become small beyond the CR circle despite the size of their $l = 1$ bars and non-negligible $l = 2$ bars. Remarkably just the $l = 0$ and $l = 1$ Fourier components dominate for the rapidly rotating fundamental modes we find for unidirectional disks, from which a small relative mass of low angular momentum stars have been removed and replaced by an inert central bulge.

The orbits which participate in our unstable modes are predominantly abnormal as defined by Lynden-Bell (1979). All sufficiently central, as well as more radial, orbits are abnormal. In fact even circular orbits are normal only for $R/R_C > 2.44$ for Kuzmin's disk, $R/R_C > 3.73$ for the isochrone, and $R/R_C > 25.3$ for the soft-centered logarithmic potential. Our modes lie primarily within these limits. A quite different situation arises with the scale-free potentials $V_0(R) = \text{sgn}(\alpha)R^\alpha$ for $\alpha \neq 0$, and $V_0(R) = \ln R$ for $\alpha = 0$, studied by Evans & Read (1998a,b). Even central orbits can then be normal because the normal/abnormal classification then depends solely on the ratio $y = L/L_c(E)$, and is independent of energy. Specifically, an orbit is abnormal if

$$- \left[f(y) + \frac{\pi y}{2} \right]^2 f''(y) - \left[\frac{2 - \alpha}{2 + \alpha} \right] \left[f'(y) + \frac{\pi}{2} \right]^2 [f(y) - y f'(y)] > 0, \quad (56)$$

where the function $f(y) = \pi J_R/L_c(E)$ and depends also on α . Its derivative $f'(y) = -\pi \Omega_\phi/\Omega_R = \frac{1}{2}|g(\alpha, y)|$, where $g(\alpha, y)$ is the function used by Touma & Tremaine (1997). The criterion (56) gives $y < 0.723$ for the scale-free logarithmic potential, so that all the more circular orbits with large $L/L_c(E)$ are normal. For $\alpha = -0.25$ and a falling rotation curve, there is a wide range of normal orbits, and only orbits with $y < 0.496$ are abnormal. For $\alpha = 0.25$ and a rising rotation curve, only nearly circular orbits are normal, and all those with $y < 0.973$ are abnormal. All orbits are abnormal for $\alpha > 0.275$. Interestingly Evans & Read, who looked specifically at the cases of $\alpha = \pm 0.25$, found that unstable modes grow more vigorously for the rising rotation curve case of $\alpha = 0.25$ with many abnormal orbits. Their modes have ILRs because the range of Ω_i is unbounded.

5. SUMMARY

This paper contains new analytical results in the theory of modes of stellar disks. The most significant are the extensions of the linear perturbation theory to second order, and the need for boundary integrals. The boundary integrals led us to realize how important near-radial orbits are. Kalnajs (1971) and LBK gave formulas for angular momentum and its Ω_p multiple the total energy, but our expressions for the two separate energy components are new. The new theory is used in our calculations.

We have computed $m = 2$ modes for a variety of stellar disks, some old and some new, some self-consistent

and some not. All have soft-centered potentials. Apart from the two marginal $\beta = -m_K/2$ cases of Table 2, we have consistently found two modes. One is a centrally concentrated fundamental mode with a single peak in amplitude, and the other a more extensive, more spiral, and usually two-peaked, secondary mode. There may be others. We find that the shape, pattern speed, and growth rate of the fundamental mode are especially sensitive to the population of low angular momentum orbits. The fundamental modes of self-consistent unidirectional disks are compact and rapidly growing central bars. The growth rate of this mode is reduced substantially, and its shape becomes more extensive and spiral, if low angular momentum orbits are either removed or reversed. Removing them increases the pattern speed, while reversing them decreases it. Removing them also makes the more extensive and more spiral secondary mode the faster growing in all but one of the cases in Tables 1 and 4. We believe that this sensitivity is the reason for differences between our pattern speeds and those of AS. These occur for tapered disks, but not for the non-tapered isochrone disks.

The sensitivity of the fundamental mode to the population of low angular momentum orbits, many of which are near-radial, suggests a connection with the phenomenon of *radial orbit instability*. That phenomenon, reviewed recently in Merritt (1999), is found to occur when radial orbits are sufficiently predominant in hot anisotropic and non-rotating spherical stellar systems. The sensitivity we find arises in cool and rotating systems with few near-radial orbits.

Our bar charts are the first to show in detail how modes transfer angular momentum and energy between Fourier components in angle space. Their overall patterns are constrained, as we discuss in §4.1, but variations within the pattern are evident. Importantly, the bar charts show that only a handful of Fourier components account for almost all of the transfer of angular momentum and energy. Hence accurate calculations of modes should be achievable with relatively few Fourier components, even if more than the three Polyachenko (2005) uses in his §5.1. This result is important for the more realistic case in which the halo is responsive, rather than inert as in our non-self-consistent cases. Athanassoula (2002) finds that halo/disk interactions are significant for the transfer of angular momentum. It is reasonable to hope that they too can be well accounted for with few Fourier components. Because our analysis of the transfer is derived from a second order extension of a linear theory, it can describe only the early stages of an instability, and not its later fully nonlinear development.

With two exceptions, modes are largely confined within the CR circle, but are too fast for there to be any orbits in ILR. The exceptional cases, which occur only when low angular momentum stars are removed, are too fast for any orbits to be in CR. Those modes lie within the OLR circle. The lack on an ILR means that modes can propagate into, and be reflected from, the soft centers of our disks, even in cases in which our densities drop to zero there. The modes are unstable, sometimes rapidly so, as swing-amplifier theory (Toomre 1981) predicts, and we achieve stability in §3.3 only by decreasing the active mass of the disk. We have also found that orbital population, which plays little role in swing-amplifier theory,

does matter.

Alar Toomre for their helpful comments. This work has been supported in part by the National Science Foundation through grant DMS-0104751.

We thank Evgenii Polyachenko, Jerry Sellwood, and

APPENDIX

LAGRANGIAN DESCRIPTION, AND BOUNDARY INTEGRALS AS FLUXES

The matrix analysis of §2 uses an Eulerian description of phase space. Kalnajs (1977) gives an alternative Lagrangian description. Its advantage is that it automatically includes any contributions from the motion of boundaries in phase space. Kalnajs's Lagrangian analysis, with our definition (12) of Fourier coefficients, yields the formula

$$M_{jk}(m, \omega) = -4\pi^2 \sum_{l=-\infty}^{\infty} \int d\mathbf{J} f_0(J_R, J_\phi) \left(l \frac{\partial}{\partial J_R} + m \frac{\partial}{\partial J_\phi} \right) \left[\frac{\Psi_{l,j}^m \Psi_{l,k}^m}{l\Omega_R + m\Omega_\phi - \omega} \right], \quad (\text{A1})$$

with no derivatives of f_0 . Our equation (21) follows after integrating (A1) by parts with respect to J_R and J_ϕ . Those integrations introduce boundary terms at any boundary of the region of integration in action space unless one or other of f_0 and the term in square brackets vanish at that boundary. In the case of the wholly prograde DF of equation (23), integration by parts with respect to J_ϕ gives precisely the boundary integral $\mathbf{M}^B$ as defined in (27), plus the area integral (26) after also integrating with respect to J_R . This shows that the boundary integrals, which arise from the step function in the Eulerian description of §2.2, can be explained as due to the motion of that boundary. In particular, the boundary integral $\mathbf{M}^B$ arises from the perturbation of a non-zero population of radial orbits.

LBK uses a Lagrangian analysis to calculate the contribution to $d\mathcal{L}_2/dt$ given by stars with angular momenta in the range (h_1, h_2) . The result, their equation (28), contains two boundary integral terms which match those one obtains from applying the Eulerian equation (B6) to the DF $f_0(J_R, J_\phi)H(J_\phi - h_1)H(h_2 - J_\phi)$ which represents the stars with angular momenta in the range (h_1, h_2) . LBK interpret the boundary integrals as representing fluxes through the two boundaries. That interpretation, together with equations (40) and (41) which give a physical significance to the real and imaginary parts of the matrix $\mathbf{M}$, shows that neglecting the boundary integral terms (27) for a prograde disk of stars means neglecting the contributions to the total potential energy and angular momentum which arise from the perturbation of radial orbits. The numerical results reported in §3.1 shows that this neglect can cause large errors.

ANGULAR MOMENTUM AND ENERGY

Both $\mathcal{L}_2$ and $\mathcal{K}_{2,1}$ contain integrals of the form

$$\mathcal{I}_2(t) = \iint S(J_R, J_\phi) f_2 d\mathbf{J} d\Theta, \quad (\text{B1})$$

for different functions S . We evaluate such integrals using

$$\frac{d\mathcal{I}_2(t)}{dt} = \iint S(J_R, J_\phi) \frac{\partial f_2}{\partial t} d\mathbf{J} d\Theta = \iint S(J_R, J_\phi) \left(\frac{\partial f_2}{\partial t} + [f_2, \mathcal{H}_0] \right) d\mathbf{J} d\Theta, \quad (\text{B2})$$

because the added terms have angle derivatives of periodic functions and so integrate to zero over the angles. Equation (6) gives the last term in parentheses in (B2) as the sum of two terms. The first contributes

$$\iint -S[f_0, V_2] d\mathbf{J} d\Theta = \iint S \frac{\partial f_0}{\partial J_\nu} \frac{\partial V_2}{\partial \theta_\nu} d\mathbf{J} d\Theta = \iint \frac{\partial}{\partial \theta_\nu} \left(S V_2 \frac{\partial f_0}{\partial J_\nu} \right) d\mathbf{J} d\Theta = 0. \quad (\text{B3})$$

Here the subscript ν represents the pair of subscripts (R, ϕ) , and we suppose the summation convention to apply to it. The last step again uses the fact that any integral of an integrand which is a derivative with respect to an angle vanishes. The second term contributes

$$\begin{aligned} \iint -S[f_1, V_1] d\mathbf{J} d\Theta &= \iint S \left[\frac{\partial f_1}{\partial J_\nu} \frac{\partial V_1}{\partial \theta_\nu} - \frac{\partial f_1}{\partial \theta_\nu} \frac{\partial V_1}{\partial J_\nu} \right] d\mathbf{J} d\Theta = \\ &= \iint \left[-f_1 \frac{\partial V_1}{\partial \theta_\nu} \frac{\partial S}{\partial J_\nu} + \frac{\partial V_1}{\partial \theta_\nu} \frac{\partial (S f_1)}{\partial J_\nu} - \frac{\partial (S f_1)}{\partial \theta_\nu} \frac{\partial V_1}{\partial J_\nu} \right] d\mathbf{J} d\Theta. \end{aligned} \quad (\text{B4})$$

The combination of the second and third components vanishes because it can be combined to an integral of a divergence:

$$\iint \left[\frac{\partial}{\partial J_\nu} \left(S f_1 \frac{\partial V_1}{\partial \theta_\nu} \right) - \frac{\partial}{\partial \theta_\nu} \left(S f_1 \frac{\partial V_1}{\partial J_\nu} \right) \right] d\mathbf{J} d\Theta. \quad (\text{B5})$$

The angle derivatives integrate to zero, but so too do the derivatives with respect to the actions. That is because the differentiated quantities vanish at the limits in action space, as $J_\phi \rightarrow \pm\infty$ and $J_R \rightarrow \infty$ where the perturbation tends to zero, and at $J_R = 0$ where V_1 is independent of θ_R because $\Psi_{l,j}^m(0, J_\phi) = 0$ for $l \neq 0$ (cf §2.2). The remaining first component of (B4) can be evaluated for $S = J_\phi$ as in (34), to give

$$\frac{d\mathcal{L}_2(t)}{dt} = -2m\pi^2 e^{2st} \sum_{l=-\infty}^{\infty} \int d\mathbf{J} \left(l \frac{\partial f_0}{\partial J_R} + m \frac{\partial f_0}{\partial J_\phi} \right) \frac{|\tilde{V}_l|^2}{|l\Omega_R + m\Omega_\phi - \omega|^2}. \quad (\text{B6})$$

This result agrees with that of LBK whose Fourier coefficients are larger than ours by a factor of $4\pi^2$, and whose ω has the opposite sign. For $S = \mathcal{H}_0$, we get

$$\frac{d}{dt} \iint \mathcal{H}_0 f_2 d\mathbf{J} d\Theta = -2s\pi^2 e^{2st} \sum_{l=-\infty}^{\infty} \int d\mathbf{J} \left(l \frac{\partial f_0}{\partial J_R} + m \frac{\partial f_0}{\partial J_\phi} \right) \frac{(l\Omega_R + m\Omega_\phi) |\tilde{V}_l|^2}{|l\Omega_R + m\Omega_\phi - \omega|^2}, \quad (\text{B7})$$

A simple consequence is that

$$\frac{d\mathcal{E}_2}{dt} = \frac{d(\mathcal{K}_2 + \mathcal{W}_2)}{dt} = \Omega_p \frac{d\mathcal{L}_2}{dt}, \quad (\text{B8})$$

where $\mathcal{E}$ is the total energy. The undifferentiated values which are quoted in §2.3 follow because of the simple time dependence on e^{2st} .

A deeper analysis of equation (6), though not its full solution, is needed to evaluate the integral

$$\mathcal{W}_{2,2} = -\mathcal{K}_{2,2} = \iint V_0 f_2 d\mathbf{J} d\Theta. \quad (\text{B9})$$

We rewrite equation (6) as

$$\frac{\partial f_2}{\partial t} + [f_2, \mathcal{H}_0] + [f_0, V_2] = -[f_1, V_1]. \quad (\text{B10})$$

The left hand side of (B10), which is homogeneous in subscript 2 quantities, has the same form as the first order problem for which we derived the homogeneous linear equations (22). Equation (B10) leads in a similar way to inhomogeneous linear equations. Its right hand side contains both axisymmetric terms and non-axisymmetric ones with wavenumber $2m$. We need consider only the axisymmetric terms and the part of the solution for f_2 which they cause, because only they will contribute to the integral (B9) for $\mathcal{W}_{2,2}$. They have a Fourier expansion

$$e^{2st} \sum_{l=-\infty}^{\infty} \tilde{N}_l(J_R, J_\phi) e^{il\theta_R}, \quad \text{where} \quad e^{2st} \tilde{N}_l = \frac{1}{(2\pi)^2} \int d\Theta e^{-il\theta_R} \left(-\frac{1}{4}[f_1, \bar{V}_1] - \frac{1}{4}[\bar{f}_1, V_1] \right). \quad (\text{B11})$$

We represent the potential and density of the axisymmetric part of f_2 by series

$$V_2 = e^{2st} \sum_{j=0}^{\infty} a_j \psi_j^0(R), \quad \Sigma_2 = e^{2st} \sum_{j=0}^{\infty} a_j \sigma_j^0(R), \quad (\text{B12})$$

like those of equations (15) and (16) but now with axisymmetric basis functions. We use Fourier expansions

$$f_2 = e^{2st} \sum_{l=-\infty}^{\infty} \tilde{g}_l(J_R, J_\phi) e^{il\theta_R}, \quad V_2 = e^{2st} \sum_{l=-\infty}^{\infty} \tilde{U}_l(J_R, J_\phi) e^{il\theta_R}, \quad \tilde{U}_l = \sum_{j=0}^{\infty} a_j \Psi_{l,j}^0, \quad (\text{B13})$$

like those of equations (11), (12), and (18), with Fourier coefficients $\Psi_{l,j}^0$ defined as in equation (18) for $m = 0$. Matching Fourier coefficients in equation (B10) gives

$$(2s + il\Omega_R) \tilde{g}_l - il \frac{\partial f_0}{\partial J_R} \tilde{U}_l = \tilde{N}_l. \quad (\text{B14})$$

Then, following the same procedure as used in §2.1, we obtain the matrix equation

$$[\mathbf{M}(0, 2is) - \mathbf{D}(0)] \mathbf{a} = \mathbf{h}, \quad (\text{B15})$$

where the components of the column vector $\mathbf{h}$ are given by

$$h_j = 4\pi^2 \sum_{l=-\infty}^{\infty} \int d\mathbf{J} \frac{i\tilde{N}_l \Psi_{l,j}^0}{(l\Omega_R - 2is)}. \quad (\text{B16})$$

The matrix $\mathbf{M}(0, 2is)$ is real because each $\pm l$ pair in the sum (21) combines two complex conjugate quantities, because $\Psi_{l,j}^0$ is even in l [cf eq. (18)]. The right hand side $\mathbf{h}$ of equation (B15) is real because the $\pm l$ pairs in the sum (B16) also combine two complex conjugate quantities, due also to the fact that $\tilde{N}_l = \tilde{N}_{-l}$ because the $\tilde{N}_l$ are the Fourier coefficients of a real function. Hence equation (B15) is a real matrix equation, and its solution for the unknown vector $\mathbf{a}$ is real. Knowing $\mathbf{a}$, we can evaluate

$$\mathcal{W}_{2,2} = \int V_0 \Sigma_2 d\mathbf{x} = 2\pi e^{2st} \sum_{j=0}^{\infty} a_j \int_0^\infty V_0(R) \sigma_j^0(R) R dR. \quad (\text{B17})$$

The reason why it is so much easier to compute $\mathcal{W}_{2,1}$, $\mathcal{K}_{2,1}$, and $\mathcal{L}_2$ is that they need only the single Fourier coefficient $\tilde{g}_0$. Equation (B14) gives $\tilde{g}_0$ simply as $\tilde{N}_0/2s$ and no matrix equation is needed.

The computation of $\mathbf{a}$ can be checked by verifying that the total mass due to the axisymmetric density Σ_2 vanishes. This is

$$\mathcal{M}_2(t) = 2\pi e^{2st} \sum_{j=0}^{\infty} a_j \int_0^{\infty} \sigma_j^0(R) R dR = 0. \quad (\text{B18})$$

Formally, the constancy of $\mathcal{M}_2$ follows from the analysis of Appendix B; it is the $S = 1$ case of integral (B1).

The double sums of integrals needed for calculating the elements of the column vector $\mathbf{h}$, found after substituting $\tilde{N}_l$ as given by equation (B11) in equation (B16), are given in Appendix A of JH. Their most significant feature is that they, and hence the solution for $\mathcal{W}_{2,2} = -\mathcal{K}_{2,2}$ intermingle different Fourier components, unlike $\mathcal{L}_2$, $\mathcal{W}_{2,1}$, and $\mathcal{K}_{2,1}$ for which the Fourier components can be separated as in equations (35) and (36).

MODELS FOR EXPONENTIAL DISKS

The difference $V_0^{\text{ext}}(R) = V_0(R) - V_0^{\text{D}}(R)$ by which the total unperturbed potential (53) exceeds that due to the disk (54) must be provided by some halo or bulge component of the galaxy, whose density must be everywhere positive. Figure 5 illustrates the consequences of requiring that that external density be spherical. That density is given by Poisson's equation as

$$\rho_{\text{H}}(r) = \frac{1}{4\pi G} \frac{1}{r^2} \frac{d}{dr} [r v_{\text{H}}^2(r)], \quad v_{\text{H}}^2(R) = R \frac{dV_0^{\text{ext}}(R)}{dR} \geq 0, \quad (\text{C1})$$

(Zang 1976) where $r = \sqrt{R^2 + z^2}$. Here $v_{\text{H}}^2(R)$ gives the amount by which the square of the circular velocity for the galaxy exceeds that of the disk alone. The physically necessary condition $\rho_{\text{H}}(r) \geq 0$ requires that $v_{\text{H}}^2(R) \geq 0$, but is more restrictive. The region below the solid curve in Figure 5 gives the range of the dimensionless combinations of parameters $G\Sigma_s R_{\text{D}}/v_0^2$ and $R_{\text{C}}/R_{\text{D}}$ which it allows. The boundary value $G\Sigma_s R_{\text{D}}/v_0^2 = 0.304$ of the solid curve at $\lambda = 0$ applies to the limit of the uncored exponential disk in a singular logarithmic potential. An alternative statement of the condition $\rho_{\text{H}}(r) \geq 0$ in this limit is $v_0(R_{\text{D}}/GM_{\text{D}})^{1/2} \geq 0.723$, where M_{D} is the mass of the uncored exponential disk.

Sellwood (1989), and unpublished work by Toomre described there, studied modes of the uncored exponential disk in the soft-centered logarithmic potential (53). The critical value of $v_0(R_{\text{D}}/GM_{\text{D}})^{1/2}$ needed to ensure that $\rho_{\text{H}} > 0$ for this case is not greatly changed when the logarithmic potential is soft-centered. It is reduced only to 0.691 when $\lambda = R_{\text{C}}/R_{\text{D}} = 0.5$. For $\lambda = 0.2$ and $v_0(R_{\text{D}}/GM_{\text{D}})^{1/2} = 0.6$ as in the model displayed in Sellwood's Figures 1 and 2, not only does ρ_{H} become negative, but $v_{\text{H}}^2(R) < 0$ for $1.4 < R/R_{\text{D}} < 3.0$. The fixed halo therefore exerts an outward force in this range. (The singular nature of the potential of the uncored exponential disk also causes $\rho_{\text{H}} < 0$, but only for $R/R_{\text{D}} \ll 1$ which is well below the softening length used in computations.)

Giovanelli & Haynes (2002) analyzed a large number of rotation curves and have found that the ratio of the scale length of the steep inner rise of the rotation curve to the scale length of the exponential varies from 0.63 for the most luminous galaxies, to 1.28 for their least luminous. They fit a parametric model in which the rotation curve decays exponentially towards its form in the outer regions, and hence their ratios provide only the approximate guidance that it is reasonable to use values of order unity for our ratio $R_{\text{C}}/R_{\text{D}}$. Figure 6 plots total circular velocity, and the parts provided by the disk and the halo/bulge for three different relative sizes of the exponential disk which cover more than their range. The value of $G\Sigma_s R_{\text{D}}/v_0^2$ is close to 90% of the maximum allowable for that $R_{\text{C}}/R_{\text{D}}$ in each case. The least extensive disk makes the largest contribution to the rotation curve at the center, but its relative contribution then declines rapidly. The disk's contribution tracks the rotation curve considerably further in the intermediate case with $R_{\text{D}} = R_{\text{C}}$, and requires only a relatively small contribution from a central halo/bulge to make up the deficit. As R_{D} increases, the disk becomes less maximal and an increasingly large spherical central halo/bulge is needed.

We construct DFs by using the identity $\exp(-V_0/v_0^2) \sqrt{1 + R^2/R_{\text{C}}^2} = 1$ to partition the density (54) in the form

$$\Sigma_{\text{D}}(R) = \Sigma_s e^{-\lambda e^{\Phi}} e^{-2N\Phi} \left(1 + \frac{R^2}{R_{\text{C}}^2}\right)^N, \quad \Phi = \frac{V_0}{v_0^2}, \quad \lambda = \frac{R_{\text{C}}}{R_{\text{D}}}. \quad (\text{C2})$$

Here N is an integer parameter which allows us to generate a family of models. Binomial expansion of $(1 + R^2/R_{\text{C}}^2)^N$ gives Σ_{D} as a sum of powers of R^2 multiplied by functions of the potential Φ . We then use Sawamura's (1988) method, following Evans & Collett (1993) §3.1, to find the DF as

$$f_0^P(E, L) = \Sigma_s \sum_{n=0}^N \binom{N}{n} \left(\frac{L}{R_{\text{C}}}\right)^{2n} g_n(E), \quad g_n(E) = \frac{(-1)^{n+1}}{2^n \sqrt{\pi} \Gamma(n+1/2)} \frac{d^{n+1}}{dE^{n+1}} \left(e^{-2NE/v_0^2} e^{-\lambda e^{E/v_0^2}} \right). \quad (\text{C3})$$

Figure 7a plots the mean rotation velocity $\langle v_{\phi} \rangle$ for four different values of N , and shows that the disks become increasingly cool and centrifugally supported with increasing N . The mean rotation velocity exceeds the circular velocity near the center, due to the $n = 0$ isotropic term in the DF (C3). Its density, which is

$$\Sigma_{\text{iso}} = \Sigma_s \left(1 + \frac{R^2}{R_{\text{C}}^2}\right)^{-N} \exp \left[-\frac{\sqrt{R^2 + R_{\text{C}}^2}}{R_{\text{D}}} \right], \quad (\text{C4})$$

is confined to the central region. It is compressed as N increases, and disappears as $N \rightarrow \infty$ along with all radial orbits. The other interesting feature of Figure 7a is that $\langle v_\phi \rangle$ declines away from v_0 at large distances. Evans & Collett's (1993) models for a simple uncored exponential disk in a singular logarithmic potential have the same property - see their Figure 2a - and for the same reason. Although $\langle v_\phi \rangle \rightarrow v_0$ as $N \rightarrow \infty$, their equation (3.17) also shows that $\langle v_\phi \rangle$ decreases for increasing R for fixed N , as in their Figure 2a and our Figure 7a.

Figure 7b plots the Toomre stability parameter Q for two different N values and for the same three sizes of the exponential disk as in Figure 6. The cooler $N = 6$ models have regions of varying extent in which Q is close to its marginal value of 1. The growth of Q with increasing R is due primarily to decreasing Σ_D because the radial velocity dispersion $\sigma_R = \sqrt{\langle v_R^2 \rangle}$ falls off only mildly.

REFERENCES

- Aoki, S., & Iye, M. 1978, PASJ, 30, 519
 Athanassoula, E. 2002, ApJ, 569, L83
 Athanassoula, E., & Sellwood, J. A. 1986, MNRAS, 221, 213 (AS)
 Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton: Princeton Univ. Press)
 Clutton-Brock, M. 1972, Ap&SS, 16, 101
 Earn, D. J. D., & Sellwood, J. A. 1995, ApJ, 451, 533
 Efsthathiou, G., Lake, G., & Negroponte, J. 1982, MNRAS, 199, 1069
 Evans, N. W., & Collett, J. L. 1993, MNRAS, 264, 353
 Evans, N. W., & Read, J. C. A. 1998a, MNRAS, 300, 83
 Evans, N. W., & Read, J. C. A. 1998b, MNRAS, 300, 106
 Freeman, K. C. 1970, ApJ, 160, 811
 Gerhard, O. E. 1991, MNRAS, 250, 812
 Giovanelli, R., & Haynes, M. P. 2002, ApJ, 571, L107
 Gradshteyn, I. S., & Ryzhik, I. M., 1980, Table of Integrals, Series and Products (New York: Academic Press)
 Hohl, F. 1971, ApJ, 168, 343
 Hohl, F. 1975, in IAU Symp. 69, Dynamics of Stellar Systems, ed. A. Hayli (Dordrecht: Reidel) 349
 Hunter, C. 1992, in Astrophysical Disks, ed S. F. Dermott, J. H. Hunter Jr., & R. E. Wilson (New York: Annals NY Acad. Sci. 675), 22
 Jalali, M. A., & Hunter, C. 2005, preprint (astro-ph/0503255) (JH)
 Kalnajs, A. J. 1971, ApJ, 166, 275
 Kalnajs, A. J. 1972, ApJ, 175, 63
 Kalnajs, A. J. 1976, ApJ, 205, 751
 Kalnajs, A. J. 1977, ApJ, 212, 637
 Kalnajs, A. J. 1978, in IAU Symp. 77, Structure and Properties of Nearby Galaxies, ed. E. M. Berhuijsen & R. Wielebinski (Dordrecht: Reidel) 113
 Kormendy, J. 1977, ApJ, 217, 406
 Kuzmin, G. G. 1956, AZh, 33, 27
 Landau, L. D. 1946, J. Phys. USSR, 10, 25
 Lin, C. C., & Shu, F. H. 1964, ApJ, 140, 646
 Lindblad, B. 1959, in Handbuch der Physik, 53, ed S. Flügge (Berlin: Springer), 21
 Lynden-Bell, D. 1979, MNRAS, 187, 101
 Lynden-Bell, D., & Kalnajs, A. J. 1972, MNRAS, 157, 1 (LBK)
 Merritt, D. 1999, PASP, 111, 129
 Mestel, L. 1963, MNRAS, 126, 553
 Miller, R. H., Prendergast, K. H., & Quirk, W. J. 1970, ApJ, 161, 903
 Miyamoto, M. 1971, PASJ, 23, 21
 Ostriker, J. P., & Peebles, P. J. E. 1973, ApJ, 186, 467
 Pichon, C., & Cannon, R. C. 1997, MNRAS, 291, 616
 Polyachenko, E. V. 2004, MNRAS, 348, 345
 Polyachenko, E. V. 2005, MNRAS, 357, 559
 Sawamura, M. 1988, PASJ, 40, 279
 Sellwood, J. A. 1989, MNRAS, 238, 115
 Sellwood, J. A., & Athanassoula, E. 1986, MNRAS, 221, 195 (SA)
 Toomre, A. 1964, ApJ, 139, 1217
 Toomre, A. 1977, ARA&A, 15, 437
 Toomre, A. 1981, in Structure and Evolution of Normal Galaxies, ed S. M. Fall & D. Lynden-Bell (Cambridge: Cambridge Univ. Press), 111
 Touna, J., & Tremaine, S. 1997, MNRAS, 292, 905
 Tremaine, S., & Weinberg, M. D. 1984, MNRAS, 209, 729
 Vandervoort, P. O. 1999, MNRAS, 303, 393
 Vauterin, P., & Dejonghe, H. 1996, A&A, 313, 465
 Zang, T. A. 1976, PhD Thesis, Massachusetts Institute of Technology, Cambridge, MA
 Zang, T. A., & Hohl, F. 1978, ApJ, 226, 521