Mon. Not. R. astr. Soc. (1989) **239,** 549–569

Self-gravitating response of a spherical galaxy to sinking satellites

Martin D. Weinberg *Institute for Advanced Study, Princeton, NJ 08540, USA*

Accepted 1988 October 25. Received 1988 October 25; in original form 1988 May 3

Summary. I present a method for computing the response of a spherical stellar system to a periodic perturbation. The perturbation can be a normal mode or an externally applied force. Any physically realistic model [general distribution function $f(E, J)$ with $\rho > 0$] can be investigated. The technique is an adaptation of a method used by Kalnajs for studying discs.

As an application, I derive an analytic solution for the orbital decay of a satellite in a self-gravitating galaxy. This problem is central to our understanding of galactic cannibalism. Strictly, the scenario is applicable to the merger of a dwarf and, for example, a cD galaxy. However, the analysis presented here should help provide a physical framework for understanding more complicated realistic systems.

The central result is that the orbital decay time in the self-gravitating case is increased by a factor of 2–3 over the non-self-gravitating case. The response of the galaxy to the satellite is dominated by the dipole contribution; the self-gravity causes a phase shift which may decrease the orbital torque by an order of magnitude depending on galactocentric position. I present and contrast both the self-gravitating and non-self-gravitating wakes and clarify the importance of local and global effects. The results suggest that Chandrasekhar's formula will not adequately represent the orbital decay of an extended satellite in general.

1 Introduction

The 'sinking satellite' is an important merger scenario. It appears in such diverse endeavours as in calibrating the galaxian luminosity function, in the study of formation and evolution of bright cluster ellipticals (e.g. Tremaine 1981), and in understanding the morphological evolution of disc galaxies (Quinn & Goodman 1986). For the most part, researchers have investigated these problems using one of two models: an application of Chandrasekhar's (1943) dynamical friction formula or direct numerical simulation (e.g. Lin & Tremaine 1983; White 1983; Bontekoe & van Albada 1987, hereafter BvA; Zaritsky & White 1987, hereafter ZW). Neither approach is ideal; as I discuss below, the former ignores the fact that orbital structure is generally important in the system, while the latter suffers from discreteness effects which limit phase space resolution.

Chandrasekhar's formula describes the drag on a body moving through an infinite homogeneous background of particles. It may be derived by (i) linearizing the equations of motion, (ii) computing the net momentum change between a background particle and the body, and then (iii) determining the rate of momentum exchange between the body and all particles in the system. Summing over the change in momentum for all particle trajectories gives a logarithmic divergence at large and small impact parameters. However, the contribution at impact parameter smaller than that needed for a 90° scattering, $b_{90°}$, gives a subdominant contribution and may be omitted. In addition, there is no contribution for impact parameters larger than the size of the system, R. The divergence, then, has a natural truncation which is traditionally parametrized by the quantity $\ln \Lambda \approx \ln(R/b_{90°})$. The drag force on a body of mass M moving at speed v through the infinite homogeneous medium is usually written

$$M\frac{dv}{dt} = -\frac{4\pi G^2 M^2}{v^2}\rho(<v)\ln\Lambda, \tag{1}$$

where $\rho(<v)$ is the density of stars with speeds less than v. This approach may be extended to other geometries by considering the momentum exchange between the moving body and the trajectories or orbits of the background system (Mulder 1983).

The theory of angular momentum exchange in spherical systems was discussed by Tremaine & Weinberg (1984, Paper I). The spherical system is inhomogeneous (in r) and is bounded in two dimensions (θ, ϕ). The qualitative difference in the trajectories or orbits changes the physics of the drag. We found that angular momentum was only exchanged at commensurabilities or *resonances* between the oscillation frequency of the perturbation and the orbital frequencies of the stars. I used this theory to compute the drag on a satellite orbiting in an isothermal sphere (Weinberg 1986, Paper II). Palmer & Papaloizou (1985) also considered this problem using a different formalism and derived an equivalent expression for the torque. An advantage of this treatment is that there is no undetermined logarithmic factor, $\ln \Lambda$. Conversely, because the singular isothermal sphere is scale-free, the spherical theory and Chandrasekhar's formula have identical dependence on physical quantities allowing a direct determination of an effective $\ln \Lambda$. This use of Chandrasekhar's formula will only be strictly valid for scale-free models.

A shortcoming of the analytic theories described above is that they ignore the response of the galaxy to the wake raised by the satellite. In other words, the self-gravity of the response is not taken into account. An exception is Palmer & Papaloizou (1985) who were able to find upper limits to the self-consistent energy exchange rate between the satellite and primary, although they did not produce a self-gravitating decay curve *per se*. Owing to the difficulty in extending the analytic theories, most previous studies of self-gravitating orbital decay have been simulations. White (1983) investigated the lowest order self-gravitating corrections using N-body simulations. He found that the orbital decay times were significantly longer than those in the non-self-gravitating simulation by Lin & Tremaine (1983). The assumption that the centre of the primary galaxy remains at rest in the inertial frame in Lin & Tremaine's calculation was assigned the blame. White referred to this assumption as a 'pinned' centre. White, by allowing some degree of barycentric motion, assumed an 'unpinned' centre. More recently, BvA and ZW performed detailed simulations of the satellite–primary interaction. The orbital decay times from both groups are in good agreement. However, BvA conclude that the decay is a local process and find no significant difference between simulations with pinned and unpinned centres, whereas ZW assert the importance of global effects but find that the local description gives a good estimate of the decay rate.

Although the simulations illustrate the overall features of the process, the literature attests to

many unanswered questions. This confusion underlines the difficulty in investigating a process with an N-body simulation that depends on detailed structure of a phase space distribution. For this reason, I take a quasi-analytic approach to determine the response of a self-gravitating galaxy. I say 'quasi-analytic' because of the great numerical effort required to solve the formal analytic solutions. In this paper, I present a general technique for studying the response of a spherical system to a perturbation. Both forced problems such as satellite orbital decay and self-excited problems such as the stability and modal structure of the system can be studied. As a first application, I will attempt to shed some light on the theoretical aspects of self-gravitating orbital decay.

I devote Section 2 to a description of the technique. The reader who is not interested in the technical details may wish to skim Section 2 and proceed to Section 3 where the results of the sinking satellite investigation are presented. In Section 3, I first contrast the response of the spherical galaxy in both the self-gravitating and non-self-gravitating cases and identify the dominant contribution. An examination of the wake illustrates the global nature of the primary galaxy's response and elucidates the importance of self-gravity. Finally, the torque and the orbital decay profiles are presented. I defer a detailed comparison between these theoretical results and the numerical simulations to a companion paper (Hernquist & Weinberg 1989). I conclude with Section 4.

2 Derivation of self-gravitating response

In this section, I derive an expression for the response of a self-gravitating spherical galaxy to a satellite of fixed potential in a circular orbit. Although the discussion is directed toward this specific scenario, the formalism described here has several other applications so I will describe it in detail. For example, this technique could be generalized to compute the drag on non-circular satellite orbits. A more exciting application, though, is in investigating dynamical stability. Stability criteria based on a variational principle for phase space distributions which are functions of energy alone have been discussed by Antonov (1962) and Doremus, Feix & Baumann (1971). The situation for anisotropic spherical models is less established although unstable models have been found. In particular, Barnes (1986), Barnes, Goodman & Hut (1986) and Palmer & Papaloizou (1987) have demonstrated the existence of a *radial orbit instability* which distorts the sphere into a triaxial figure. Using the formalism presented in this section, the dispersion relation can be calculated and the stability of a model with any distribution function $f(E, J)$ can be ascertained. The dispersion relation for the stellar sphere has also been discussed by Fridman & Polyachenko (1984).

2.1 COMPUTING THE RESPONSE

The self-consistent response of the galaxy to a perturbation is described by the simultaneous solution of the collisionless Boltzmann equation and Poisson's equation. However, because of the spherical geometry, the system of equations is difficult to solve, even in the linearized regime. The natural coordinates for solving each of the two equations are different which results in the somewhat cumbersome formalism described in Section 2.1.2. The simultaneous solution of the two equations leads to a coupled set of integrodifferential equations. Rather than solving this system directly, it is first transformed to a matrix equation in Section 2.1.3. The solution to the original set of equations, then, is given by an expression similar to an eigenequation, $c_j = M_{ji} d_i$; the vector on the left-hand-side, c_j, is the response of the system to the imposed perturbation on the right-hand-side, d_i (Section 2.1.4). The specialization of the formalism to the case of satellite decay is derived in Sections 2.1.5 and 2.1.6.

2.1.1 Statement of the problem

The collisionless Boltzmann equation and the Poisson equation may be written

$$\frac{\partial f}{\partial t}+[f, H]=0, \tag{2}$$

$$\nabla^2\Phi=4\pi G\rho, \tag{3}$$

where f is the distribution function for the galaxy, H is the single-particle Hamiltonian, Φ is the total gravitational potential and ρ is the total density. The Poisson bracket (e.g. Goldstein 1950) is a convenient coordinate-free notation for

$$[f, g]=\sum_{i=1}^{3}\left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i}-\frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}\right), \tag{4}$$

where (q_i, p_i) are canonical variables, for example, the Cartesian positions and momenta. The Hamiltonian, H, for a particle of unit mass may be written

$$H=\sum_{i=1}^{3}\frac{p_i^2}{2}+\Phi(\mathbf{r}). \tag{5}$$

Let us assume the existence of an equilibrium (time-independent) solution of equations (2) and (3): f_0, H_0, Φ_0 and ρ_0. I now want to compute the response of the system to the existence of the satellite. To make the calculation tractable, I assume that the satellite mass is small and linearize equation (2). The perturbed potential is the sum of two terms: $\Phi_1=\Phi_1^{\rm ext}+\Phi_1^{\rm resp}$. The first term describes the external perturbation itself which is specified *a priori* and the second term is the response of the system to the external force which is the desired unknown quantity. Similarly, the perturbed density may be written $\rho_1=\rho_1^{\rm ext}+\rho_1^{\rm resp}$. The linearized forms of equations (1) and (2) become

$$\frac{\partial f_1}{\partial t}+[f_1, H_0]+[f_0, \Phi_1]=0, \tag{6}$$

$$\nabla^2\Phi_1=4\pi G\rho_1, \tag{7}$$

where the subscripts 0 and 1 indicate unperturbed and first-order quantities as usual.

2.1.2 Formal solution

If we can find a single basis for which equations (6) and (7) are separable, the first-order quantities can be expanded in complete sets of functions and the solution can be reduced to algebra and quadrature. Unfortunately, the natural basis for each equation is different. Let us first consider equation (6). In a spherical system, the orbital motion is confined to a plane and an orbit has two characteristic frequencies: a radial frequency and an azimuthal frequency. The radial frequency and radial period (time from periapse to periapse) is given by

$$t_{\rm r}=\frac{2\pi}{\Omega_1}=2\int_{r_{\rm p}}^{r_{\rm a}}\frac{dr}{[2(E-\Phi_0)-J^2/r^2]^{1/2}}, \tag{8}$$

where $r_{\rm p}$ and $r_{\rm a}$ are the periapse and apoapse, and E and J are the energy and angular

momentum of the orbit, respectively. In one radial period the star advances by an angle in the plane of its orbit $\Delta\psi$, where $\Delta\psi = \Omega_2 t_r$ and

$$\frac{\Omega_2}{\Omega_1} = \frac{J}{\pi}\int_{r_p}^{r_a} \frac{dr}{r^2[2(E-\Phi_0)-J^2/r^2]^{1/2}}. \tag{9}$$

The radial action is

$$I_r = \frac{1}{2\pi}\oint dr p_r = \frac{1}{\pi}\int_{r_p}^{r_a} dr[2(E-\Phi_0)-J^2/r^2]^{1/2}. \tag{10}$$

Following the method developed in Paper I, we may express equation (6) in the canonical variables $(I_1,I_2,I_3)=(I_r,J,J_z)$ which follow from Hamilton–Jacobi theory. The conjugate coordinates are the angles (w_1,w_2,w_3). The quantities I_j are constants of motion and the w_j mod 2π are periodic functions of time; these variables are related to the action-angle variables by a trivial transformation. The angle w_1 describes the phase of a star's orbit in its radial oscillation. It is defined to be zero at periapse and increases by 2π in one radial period. The angle w_1 is given by

$$w_1 = \Omega_1 \int_{C_1} \frac{|dr|}{[2(E-\Phi_0)-J^2/r^2]^{1/2}}, \tag{11}$$

where the integration contour C_1 goes from periapse to the current position. The integrals are line integrals which increase monotonically along the orbit. The angle w_2 describes the *mean* angular phase of the orbit in the orbital plane. It is given by

$$w_2 = \psi + \int_{C_1} |dr|\,[2(E-\Phi_0)-J^2/r^2)]^{-1/2}(\Omega_2 - J/r^2), \tag{12}$$

where

$$\psi = J\int_{C_2} |d\theta|\,[J^2 - J_z^2/\sin^2\theta]^{-1/2}. \tag{13}$$

The integration contour C_2 starts at an ascending node (that is $\theta=\pi/2$, $\dot\theta<0$). The quantity ψ is the angle from the ascending node to (r,θ,ϕ), measured in the orbit plane along the direction of orbital motion. Because of the radial motion, the true phase is only identical to w_2 for circular orbits. The third angle variable w_3 is the azimuth of the ascending node which is constant since E is independent of J_z. From Hamilton's equations and the above expressions, we may write

$$\dot w_j = \Omega_j = \frac{\partial H}{\partial I_j} = \frac{\partial E}{\partial I_j}, \tag{14}$$

and find $\Omega_3 = 0$. Since the trajectories in a spherical system are conditionally periodic and the quantities I_j are invariant by construction, it is natural to expand a particular quantity χ in a Fourier series in the variables I_j, w_j,

$$\chi(\mathbf{r},\mathbf{v},t) = \sum_{l_1,l_2,l_3} \chi_{l_1 l_2 l_3}(\mathbf{I},t)\, e^{i\mathbf{n}\cdot\mathbf{w}}, \tag{15}$$

where $\mathbf{n}=(l_1,l_2,l_3)$ is a triple of integers and $\mathbf{w}=(w_1,w_2,w_3)$ is the angles. Similarly, the obvious choice of basis for equation (7) is spherical harmonics. A particular harmonic (l,m) of the response potential may be written

$$\Phi_1^{\rm resp}(r,\theta,\phi)=Y_{lm}(\theta,\phi)U_{lm}(r). \tag{16}$$

Using results from Paper I (equations 50–54), a perturbation in the form of equation (16) may be written in the form of equation (15):

$$\Phi_1^{\rm resp}(r,\theta,\phi)=\sum_{l_1=-\infty}^{\infty}\sum_{l_2=-l}^{l}\Psi_{l_1l_2l_3}(I_1,I_2,I_3)\,e^{i\mathbf{n}\cdot\mathbf{w}}, \tag{17}$$

where $l_3=m$ for a particular harmonic (l,m),

$$\Psi_{l_1l_2l_3}(I_1,I_2,I_3)=\sum_{l=1}^{\infty}\left(\frac{2}{1+\partial_{l_30}}\right)V_{ll_2l_3}(\beta)\,W^{l_1}_{ll_2l_3}(I_1,I_2), \tag{18}$$

$$V_{ll_2l_3}(\beta)=r^l_{l_2l_3}(\beta)\,Y_{ll_2}(\pi/2,0)i^{l_3-l_2}, \tag{19}$$

and

$$W^{l_1}_{ll_2l_3}(I_1,I_2)=\frac{1}{\pi}\int_0^{\pi}dw_1\,\cos[l_1w_1-l_2(\psi-w_2)]U_{ll_3}(r). \tag{20}$$

The angle β is the inclination of the orbital plane to the equatorial plane. I have defined $V_{lnm}=0$ for $|n|>l$ or $|m|>l$. The $r^l_{mn}(\beta)$ are components of the rotation matrices (e.g. Edmonds 1960). The external potential may be expanded similarly (see Paper II).

With this machinery, equations (6) and (7) may now be solved as an initial value problem. I begin by rewriting equation (6) explicitly in the variables I_j, w_j,

$$\frac{\partial f_1}{\partial t}+\frac{\partial f_1}{\partial\mathbf{w}}\cdot\frac{\partial H_0}{\partial\mathbf{I}}-\frac{\partial f_0}{\partial\mathbf{I}}\cdot\frac{\partial\Phi_1}{\partial\mathbf{w}}=0, \tag{21}$$

and perform a Fourier–Laplace transform. I define the transforms in the following way:

$$f_{l_1l_2l_3}=\mathscr{F}[f]\equiv\frac{1}{(2\pi)^3}\int_0^{2\pi}dw_1\int_0^{2\pi}dw_2\int_0^{2\pi}dw_3\,e^{-i\mathbf{n}\cdot\mathbf{w}}f, \tag{22}$$

$$\tilde g=\mathscr{L}[g]\equiv\frac{1}{2\pi}\int_0^{\infty}dt\,e^{i\omega t}g, \tag{23}$$

where $c\equiv Im(\omega)>0$ and is chosen sufficiently large that the integral in equation (23) converges. The relations $\mathscr{F}[\partial f/\partial w_j]=il_j\mathscr{F}[f]$ and $\mathscr{L}[\partial g/\partial t]=-i\omega\mathscr{L}[g]-g(0)/2\pi$ together with equations (22) and (23) are then used to transform equation (21) which yields

$$i\omega\tilde f_{l_1l_2l_3}-i\mathbf{n}\cdot\mathbf{\Omega}\,\tilde f_{l_1l_2l_3}+i\mathbf{n}\cdot\frac{\partial f_0}{\partial\mathbf{I}}\tilde\Psi_{l_1l_2l_3}=0. \tag{24}$$

I have assumed that the perturbed distribution function vanishes at $t=0$. In the case of a satellite on a circular orbit with angular frequency Ω_s, the external force has a periodic time

dependence

$$\Phi_1^{\rm ext}(\mathbf{r},t)=U^{\rm ext}(r,\theta,\phi-\Omega_s t). \tag{25}$$

It follows from equations (16), (17) and (23) that

$$\tilde{\Psi}_{l_1 l_2 l_3}=\tilde{\Psi}^{\rm resp}_{l_1 l_2 l_3}-\frac{1}{2\pi i}\frac{1}{\omega-l_3\Omega_s}\tilde{\Psi}^{\rm ext}_{l_1 l_2 l_3}. \tag{26}$$

The perturbed density then follows directly from equation (25):

$$\tilde{\rho}_1^{\rm resp}=\int d^3v\tilde{f}_1=\int d^3v\sum_{l_1,l_2,l_3}\tilde{f}_{l_1 l_2 l_3}e^{i\mathbf{n}\cdot\mathbf{w}}.$$

$$=-\int d^3v\sum_{l_1,l_2,l_3}\mathbf{n}\cdot\frac{\partial f_0}{\partial\mathbf{I}}\tilde{\Psi}_{l_1 l_2 l_3}(\mathbf{I})\frac{1}{\omega-\mathbf{n}\cdot\boldsymbol{\Omega}}e^{i\mathbf{n}\cdot\mathbf{w}} \tag{27}$$

For self-consistency, $\tilde{\rho}_1^{\rm resp}$ given in equation (27) must be equal to the density implied by $\Phi_1^{\rm resp}$ through Poisson's equation (7), $\tilde{\rho}_1^{\rm P}$. This yields a messy set of integrodifferential equations for the $\tilde{\Psi}_{l_1 l_2 l_3}$:

$$\tilde{\rho}_1^{\rm P}=\tilde{\rho}_1^{\rm resp}. \tag{28}$$

The right-hand-side of equation (28) is complicated by the need to transform between two sets of canonical variables, (**w**, **I**) to (**r**, **v**), in order to perform the triple integral in equation (27).

2.1.3 Solution in terms of a matrix equation

An alternative approach to solving equation (28) has been discussed by Kalnajs (1977). By introducing a space of perturbations where the interaction energies between two perturbations define a scalar product, equation (28) may be converted to a matrix equation which may be solved by standard methods. In addition, this technique circumvents the inconvenient transformation between the two sets of canonical variables required to solve equation (28). These simplifications are gained at the expense of the introduction of a function space as I will now describe.

Let us assume the existence of a complete biorthonormal set of potential-density pairs (u_i^{lm}, d_i^{lm}) with the following scalar product

$$-\frac{1}{4\pi G}\int drr^2 u_i^{lm*}d_j^{lm}=\begin{cases}1 & \text{if } i=j\\ 0 & \text{otherwise.}\end{cases} \tag{29}$$

The pair (u_i^{lm}, d_i^{lm}) is constructed to satisfy Poisson's equation, $\nabla^2 u_i^{lm}=4\pi G d_i^{lm}$. Completeness ensures that the expansions

$$\Phi_1^{\rm resp}=\sum_{lm}Y_{lm}(\theta,\phi)\sum_j a_j^{lm}(t)u_j^{lm}(r), \tag{30}$$

$$\Phi_1^{\rm ext}=\sum_{lm}Y_{lm}(\theta,\phi)\,e^{-im\Omega_s t}\sum_j b_j^{lm}u_j^{lm}(r), \tag{31}$$

$$\rho_1^{\rm resp}=\sum_{lm}Y_{lm}(\theta,\phi)\sum_j a_j^{lm}(t)\,d_j^{lm}(r), \tag{32}$$

exist. I will discuss a particular choice of such a set in Section 2.2. The vector $\mathbf{b}^{lm}$ in equation (31) contains all information about the structure of the satellite potential. Since the satellite orbit is circular, $\mathbf{b}^{lm}$ is independent of time. The vector $\mathbf{a}^{lm}$ in equations (30) and (32) describes the response which, in general, has a complicated time dependence. Equation (28) may be replaced by the equivalent condition

$$\int d^3r Y_{lm} u_i^{lm*} \tilde{\rho}_1^{\rm P} = \int d^3r Y_{lm} u_i^{lm*} \tilde{\rho}_1^{\rm resp}, \tag{33}$$

owing to the completeness of the $Y_{lm}u_i^{lm}$. Using equations (26)–(33) we solve for $\mathbf{a}^{lm}$ to obtain

$$\tilde{a}_i^{lm} = \frac{1}{4\pi G} \int d^3r d^3v\, Y_{lm}^*(\theta,\phi) u_i^{lm*}(r) \sum_{l_1,l_2,l_3} \mathbf{n}\cdot\frac{\partial f_0}{\partial \mathbf{I}} \tilde{\Psi}_{l_1 l_2 l_3}(\mathbf{I}) \frac{1}{\omega - \mathbf{n}\cdot\boldsymbol{\Omega}} e^{i\mathbf{n}\cdot\mathbf{w}} \tag{34}$$

$$= \frac{(2\pi)^3}{4\pi G} \int d^3I \sum_{l_1,l_2,l_3} \mathbf{n}\cdot\frac{\partial f_0}{\partial \mathbf{I}} \frac{1}{\omega - \mathbf{n}\cdot\boldsymbol{\Omega}} V^*_{l,-l_2,-l_3}(\beta) W^{-l_1,i*}_{l,-l_2,-l_3}(\mathbf{I})$$

$$\times V_{l,l_2,l_3}(\beta) W^{l_1,j}_{l,l_2,l_3}(\mathbf{I}) \left[\tilde{a}_j^{lm} - \frac{1}{2\pi i}\frac{1}{l_3\Omega_{\rm s}} \tilde{b}_j^{lm}\right], \tag{35}$$

where I have used equations (18)–(20) to go from equation (34) to (35), employed the relation $\int d\Omega\, Y^*_{l'm'}(\theta,\phi) Y_{lm}(\theta,\phi) = \delta_{l'l}\delta_{m'm}$, and defined

$$W^{l_1,j}_{l,l_2,l_3} \equiv \frac{1}{\pi}\int_0^\pi dw1\, \cos[l_1 w_1 - l_2(\psi - w_2)] u_j^{ll_3}(r). \tag{36}$$

Equation (35) may be further simplified. First, since distribution functions are often described in terms of (E,J), it is convenient to couch equation (35) in these variables; one finds $d^3I = dE\, dJ\, d\beta\, J \sin\beta/\Omega_1(E,J)$ where $\cos\beta \equiv I_3/I_2$ and β is the inclination of an orbit. In addition, it follows from the symmetries of the spherical harmonics and the rotation matrices that $V_{l,n,m} = (-1)^m V_{l,-n,-m}$ and $W^{k,j*}_{l,n,m} = (-1)^m W^{-k,j}_{l,-n,-m}$. Using the relation $\int_0^\pi d\beta \sin\beta r^l_{l_2,l_3}(\beta) r^{l'}_{l_2,l_3}(\beta) = 2\delta_{l'l}/(2l+1)$ and the relations above, equation (35) becomes

$$\tilde{a}_i^{lm} = M_{ij}(\omega)\left[\tilde{a}_j^{lm} - \frac{1}{2\pi i}\frac{1}{(\omega - l_3\Omega_{\rm s})}\tilde{b}_j^{lm}\right], \tag{37}$$

where summations over like indices are implied and

$$M_{ij}(\omega) = \frac{(2\pi)^3}{4\pi G}\int\int \frac{dE\, dJ\, J}{\Omega_1(E,J)} \sum_{l_1,l_2,l_3} \frac{2}{2l+1} \mathbf{n}\cdot\frac{\partial f_0}{\partial \mathbf{I}} \frac{1}{\omega - \mathbf{n}\cdot\boldsymbol{\Omega}} \times |Y_{l,l_2}(\pi/2,0)|^2 W^{l_1,i*}_{l,l_2,l_3}(\mathbf{I}) W^{l_1,j}_{l,l_2,l_3}(\mathbf{I}). \tag{38}$$

The above equation describes the response coefficients $\tilde{\mathbf{a}}^{lm}$ in the ω-domain. It follows that $M^*_{ij}(\omega) = M_{ij}(-\omega^*)$ and that $M_{ij} = A_{ij} + i{\rm Im}(\omega)B_{ij}$ where A and B are Hermitian. Thus M_{ij} (and $[\mathbf{1} - \mathbf{M}]_{ij}$) is Hermitian only if Im $\omega = 0$.

2.1.4 *Physical interpretation*

The physical content of the response described by equation (37) may be better discussed by first writing equation (37) in symbolic form,

$$\tilde{\mathbf{a}}^{lm} = \mathbf{M}(\omega)\cdot(\tilde{\mathbf{a}}^{lm} + \tilde{\mathbf{b}}'^{lm}), \tag{39}$$

where the vector is $\tilde{\mathbf{b}}'^{lm}$ by

$$\tilde{\mathbf{b}}'^{lm} = -\frac{1}{2\pi i}\frac{1}{(\omega - l_3\Omega_s)}\tilde{\mathbf{b}}'^{lm}. \tag{40}$$

The vector on the left-hand-side of equation (39) represents the physical response of the spherical galaxy to the perturbation represented by the vector on the right-hand-side. The matrix $\mathbf{M}(\omega)$, then, is the response operator. In the case of orbital decay, the perturbation is the galaxy's own response plus the external perturbation. If there is no external perturbation, $\tilde{\mathbf{b}}'^{lm} = 0$ and equation (39) defines a non-linear eigenvalue problem:

$$\tilde{\mathbf{a}}^{lm} = \mathbf{M}(\omega)\cdot\tilde{\mathbf{a}}^{lm}. \tag{41}$$

An eigenvalue with Im $\omega > 0$ indicates a growing mode (dynamical instability) in the background model. The forced (non-self-gravitating) response is obtained by eliminating $\mathbf{a}^{lm}$ on the right-hand-side of equation (39),

$$\tilde{\mathbf{a}}^{lm} = M(\omega)\cdot\tilde{\mathbf{b}}'^{lm}. \tag{42}$$

Since a self-gravitating response is the consistent response of the system to its own response and the applied perturbation, one might imagine that we can solve equation (39) by an iterative application of equation (42), that is $\tilde{\mathbf{a}}^{lm}|_{i+1} = \mathbf{M}(\omega)\cdot\tilde{\mathbf{a}}^{lm}|_i$ where $\tilde{\mathbf{a}}^{lm}|_0 = \tilde{\mathbf{b}}'^{lm}$. If the process converges, we find

$$\mathbf{a}^{lm} = \tilde{\mathbf{a}}^{lm}|_\infty = [\mathbf{1} + \mathbf{M} + \mathbf{M}^2 + \ldots]\cdot\tilde{\mathbf{b}}'^{lm} \tag{43}$$

$$= [\mathbf{1} - \mathbf{M}]^{-1}\mathbf{M}\cdot\tilde{\mathbf{b}}'^{lm}. \tag{44}$$

Equation (44) is equivalent to equation (37).★

2.1.5 Derivaion of the forced response

In order to obtain the response $\mathbf{a}^{lm}$ in the time domain, we must take the inverse Laplace transform of equation (37) to obtain: $\mathscr{L}^{-1}[\tilde{g}] = \int_{-\infty+ic}^{\infty+ic} d\omega e^{-i\omega t}\tilde{g}$ for some $c > 0$ as discussed following equation (23). Since $\tilde{\mathbf{a}}$ has no singularities except for simple poles in the upper-half plane Im $\omega \geq c$ by construction, the contour may be deformed as in the derivation of Landau damping (Landau 1946; Krall & Trivelpiece 1973). At an arbitrary time, the inverse Laplace transform is difficult to perform. However, I am not interested in transients but in the long-term behaviour of the response. For the satellite decay problem, I assume a dynamically stable background model (e.g. Antonov 1962) with a well-behaved distribution function and an external potential of the form of equation (25). Thus, after large times ($t \to \infty$) all transients will have decayed, leaving the forced purely oscillatory response.† Taking the inverse Laplace transform of equation (37), we obtain

$$\mathbf{a}^{lm} = \mathscr{R}^{lm}(l_3\Omega_s + i\varepsilon)\cdot\mathbf{b}^{lm}, \tag{45}$$

where

$$\mathscr{R}^{lm}(\omega) \equiv [\mathbf{1} - \mathbf{M}(\omega)]^{-1}\mathbf{M}(\omega). \tag{46}$$

★ The condition for convergence of equation (43) are much more restrictive than that for the validity of equation (44).

†We have no guarantee that there are no other oscillatory responses *a priori*. However, this can be easily checked for any given model and is the case for the $n = 3$ polytrope.

The quantity $\varepsilon > 0$ reminds us of the assumption Im $\omega > 0$ and may be thought of as a causality condition. The operator $\mathscr{R}$ relates the external perturbation $\mathbf{b}^{lm}$ to the response $\mathbf{a}^{lm}$. The non-self-gravitating limit is then recovered by taking $\mathscr{R} \to \mathbf{M}$.

Taking the limit $\varepsilon \to 0^+$ in equation (45), we may express $\mathbf{M}$ in terms of the Cauchy principal value,

$$\lim_{\varepsilon \to 0^+} M_{ij}(l_3\Omega_s + i\varepsilon) = M^R_{ij}(l_3\Omega_s) + iM^I_{ij}(l_3\Omega_s), \tag{47}$$

where the real and imaginary parts of $\mathbf{M}$ are

$$M^R_{ij}(l_3\Omega_s) = \frac{8\pi^3}{4\pi G}\int\int \frac{dE\,dJ\,J}{\Omega_1(E,J)} \sum_{l_1,l_2,l_3} \frac{2}{2l+1}\,\mathbf{n}\cdot\frac{\partial f_0}{\partial \mathbf{I}}\,P\frac{1}{l_3\Omega_s - \mathbf{n}\cdot\boldsymbol{\Omega}}$$

$$\times |Y_{l,l_2}(\pi/2,0)|^2\, W^{l_1,i}_{l,l_2,l_3}(\mathbf{I})\, W^{l,j}_{l,l_2,l_3}(\mathbf{I}), \tag{48}$$

and

$$M^I_{ij}(l_3\Omega) = -\frac{8\pi^4}{4\pi G}\int\int \frac{dE\,dJ\,J}{\Omega_1(E,J)} \sum_{l_1,l_2,l_3} \frac{2}{2l+1}\,\mathbf{n}\cdot\frac{\partial f_0}{\partial \mathbf{I}}\,\delta(l_3\Omega_s - \mathbf{n}\cdot\boldsymbol{\Omega})$$

$$\times |Y_{l,l_2}(\pi/2,0)|^2\, W^{l_1,i}_{l,l_2,l_3}(\mathbf{I})\, W^{l_1,j}_{l,l_2,l_3}(\mathbf{I}). \tag{49}$$

Equations (45)–(49) describe the response of the stable collisionless sphere to a fixed perturbing potential with harmonic time variation. The imaginary part of $\mathbf{M}$ describes the resonant response and the real part of $\mathbf{M}$ describes the non-resonant response of the sphere to the perturbation. Equation (45) replaces the integrodifferential equation given by equation (28) by an infinite-order matrix equation. Of course, the expansion in equations (30)–(32) may be truncated; in practice, only ten terms are needed for reasonable accuracy. Given $\mathbf{M}(l_3\Omega_s)$, the numerical solution of equation (45) is then straightforward.

2.1.6 Determination of the wake and orbital decay

Using equations (30) and (32), we can obtain the response potential and density. Since the potential perturbation is real, the potential and density response must be real and may be written

$$\Phi_1^{\rm resp} = \frac{1}{2}\sum_{l,m}\left[Y_{lm}(\theta,\phi)\sum_j a_j^{lm} u_j^{lm}(r) + Y^*_{lm}(\theta,\phi)\sum_j a_j^{lm*} u_j^{lm}(r)\right], \tag{50}$$

$$\rho_1^{\rm resp} = \frac{1}{2}\sum_{l,m}\left[Y_{lm}(\theta,\phi)\sum_j a_j^{lm} d_j^{lm}(r) + Y^*_{lm}(\theta,\phi)\sum_j a_j^{lm*} d_j^{lm}(r)\right]. \tag{51}$$

This expression of the perturbed quantities facilitates the consideration of a particular (l,m) component. Alternatively, it is straightforward to show that equations (50) and (51) are equivalent to the average of the (l,m) and $(l,-m)$ terms of the expansion as they must be. In the case of satellite decay, I refer to a response density given by equation (51) as a *wake*.

To compute the orbital decay, we need an expression for the torque on the orbiting satellite due to the wake. Since the satellite moves in the x–y plane (*cf.* equation 25), we need the torque along the z-axis which is given by

$$-(\mathbf{r}\times\nabla\Phi_1^{\rm resp})\cdot\hat{\mathbf{z}} = -\frac{\partial\Phi_1^{\rm resp}}{\partial\phi}. \tag{52}$$

The total torque felt by the perturber is

$$\tau_z = \int d^3r \rho^{\text{ext}} \left(-\frac{\partial \Phi_1^{\text{resp}}}{\partial \phi} \right)$$
$$= -\frac{1}{2} 4\pi G \sum_{l,m} mIm[\mathbf{b}^{lm} \cdot \mathscr{R}(l_3 \Omega_s) \cdot \mathbf{b}^{lm}]. \tag{53}$$

The second of equation (53) is obtained by substituting equation (50) in the first of equation (53) and using the orthogonality of the Y_{lm} and equation (29). Note that if we eliminate self-gravity by taking $\mathscr{R} \to M$, the expression for the torque given in equation (53) reduces to the expression previously given in Paper II (equation 25). The torque τ_z is an implicit function of the satellite radius through Ω_s. For a given model, it is then easy to integrate equation (53) to obtain the orbital decay curve, $r = r(t)$.

2.2 BIORTHONORMAL FUNCTIONS

Biorthonormal functions suitable for disc systems have been discussed by Kalnajs (1971) and Clutton-Brock (1972). Fridman & Polyachenko (1984) give an appropriate set for finite-radius spherical systems which are described below.

In the spherical system, we know that the radial wave functions must satisfy

$$\frac{d^2 U_{lm}}{dr^2} + \frac{2}{r}\frac{dU_{lm}}{dr} - \frac{l(l+1)}{r^2} U_{lm} = 4\pi G \rho_{lm}, \tag{54}$$

for harmonic term (l,m). On the boundary, we must match the solution of Laplace's equation which yields the condition

$$\left[\frac{dU_{lm}}{dr} + \frac{(l+1)}{r} U_{lm} \right] \Bigg|_{r=R} = 0. \tag{55}$$

For the rest of this section, I will drop the subscripts lm to simplify the notation. Letting $y = U(r) r^{1/2}$, one finds that equation (54) becomes

$$\frac{d^2 y}{dr^2} + \frac{1}{r}\frac{dy}{dr} - \frac{(l+1/2)^2}{r^2} y = 4\pi G \rho r^{1/2}. \tag{56}$$

Since the left-hand-side is Bessel's equation minus y, it is natural to try to find solutions of the form

$$y_n^{lm} = J_{l+1/2}(\alpha_n r/R), \tag{57}$$

and consequently

$$\rho_n^{lm} = -\frac{1}{4\pi G} \left(\frac{\alpha_n}{R} \right)^2 r^{-1/2} J_{l+1/2}(\alpha_n r/R), \tag{58}$$

where $n = 1, 2, \ldots$ and R is the radius of the sphere and the α_n are determined through equation (55). This pair is complete but it remains to show that the pair is orthogonal. From the

orthogonality relations for Bessel functions, it is straightforward to show that

$$\int_0^R drrU_j^{lm}(r)\rho_k^{lm}(r)=\begin{cases}-\dfrac{1}{4\pi G}\dfrac{\alpha_n^2}{2}[J_{l+1/2}^2(\alpha_n)-J_{l-1/2}(\alpha_n)J_{l+3/2}(\alpha_n)] & \text{if } j=k\\ 0 & \text{otherwise.}\end{cases} \tag{59}$$

The solutions must also satisfy the boundary condition, equation (56), which yields the condition $\alpha_n J_{l-1/2}(\alpha_n)=0$. In terms of spherical Bessel functions, the orthogonal pair may be written

$$u_n^{lm}(r)=-\frac{4\pi G\sqrt{2}}{\alpha_n|j_l(\alpha_n)|}R^{-1/2}j_l\left(\alpha_n\frac{r}{R}\right), \tag{60}$$

$$d_n^{lm}(r)=\frac{\alpha_n\sqrt{2}}{|j_l(\alpha_n)|}R^{-5/2}j_l\left(\alpha_n\frac{r}{R}\right), \tag{61}$$

with the scalar product given by

$$-\frac{1}{4\pi G}\int_0^R drr^2u_i^{lm}(r)\,d_j^{lm}(r)=\delta_{ij}. \tag{62}$$

We are not by any means limited to finite radius spherical systems. For example, Clutton-Brock (1973) has also presented a biorthonormal set defined over the half-infinite interval which is in principle well-suited to studying spherical systems of infinite extent. Preliminary investigations suggest that calculation with this set will be costly owing to the large values of n that may be necessary to resolve the satellite potential at radii of interest.

2.3 NUMERICAL IMPLEMENTATION

To compute the orbit of the satellite as a function of time, $r(t)$, we need $\mathbf{M}(l_3\Omega_s)$ for a run of $\Omega_s(r)$. In the case of the forced response, I use $\mathbf{M}$ in the form of equations (48) and (49). For the quadratures, I use the variables E, $\kappa\equiv J/J_{\max}(E)$ rather than E,J directly. Note that equation (48) involves a two-dimensional integration and may involve a principal part, while equation (49) only involves a one-dimensional integration over κ owing to the delta function. The principal parts are evaluated by transforming to logarithmic coordinates and changing ranges of integration to symmetrize the integrand. The accuracy of this prescription is limited by the size of the floating-point word. The performance bottleneck is in the evaluation of the Fourier transforms of the perturbed potential, $W_{l,l_2,l_3}^{l_1 j}(\mathbf{I})$, which require the computation of individual orbits through the quantities $r(w_1)$ and $(\psi-w_2)(w_1)$. I tabulate these functions and determine the W's for all j for each component (l_1,l_2,l_3) at a given κ. The cost can be greatly reduced at the expense of storage overhead by maintaining tables of $r(w_1)$ and $(\psi-w_2)(w_1)$ for each κ and interpolating for desired values of E.

In practice, I use a Simpson's rule integrator with Romberg extrapolation on a (E,κ) grid of $17\,(2^4+1)$ square to compute the integrals in equations (48) and (49). I compute all resonances with $|l_1|\le\max(8,2l)$ and expand the potential in the biorthonormal set (u_j,d_j) up to $j=10$. By varying the parameters, this choice is seen to give the response to an accuracy of 1 per cent. Computing a particular $\mathbf{M}(l_3\Omega_s)$ for harmonic l takes roughly $(l/2+1)\times 5$ CPU hr on a SUN 3/50 and depends, of course, on the galaxy model and the value of $l_3\Omega_s$. Once $\mathbf{M}$ is computed for all desired values of (l,m), the torque, density and potential response, are quickly obtained.

3 Results and discussion

In this section, I apply the method presented in Section 2 to investigate orbital decay. For historical reasons, I take the primary galaxy to be a $n=3$ polytrope with a Plummer law satellite in circular orbit. This case has been investigated by several workers in the field (White 1983, BvA, ZW).

I choose the gravitational constant G, the total mass of the primary, M, and the radius of the primary, R, to be $G=M=1$, $R=3$ (Hénon 1973) unless otherwise stated. This implies that the total energy of the primary is $W=-1/4$. The satellite at radius r_s with core radius r_c has the potential of the form (*cf.* equation 25)

$$U_1^{\text{ext}}(\mathbf{r})=-\frac{m_s}{\sqrt{(\mathbf{r}-\mathbf{r}_s)^2+r_c^2}}. \tag{63}$$

The value of m_s will be specified as needed.

3.1 BARYCENTRIC MOTION

We begin by discussing the response of the galaxy to a uniform acceleration. This idealized situation has two important applications. First, it provides a means of checking the apparatus. Given the numerical complexity in computing the response matrices (equations 48 and 49), an independent check is imperative. Second, it describes the response of the galaxy to a satellite at large galactocentric radii. This response results in barycentric motion.

Recall that a body moving in a spatially constant force field experiences no relative acceleration. Thus, applying a perturbation of the form $\Phi^{\text{ext}}=-\varepsilon' \text{Re}\, Y_{11}(\theta,\phi)r\propto -x$ to the system should cause a uniform translation along x-axis. Since the theory in Section 2 is only correct in the perturbation limit, we demand that the net displacement of the system still be small as $t\to\infty$. This is satisfied by a perturbation which is harmonic in time: $\Phi^{\text{ext}}=-\varepsilon' \text{Re}\, Y_{11}(\theta,\phi-\omega t)r=-\varepsilon\cos\theta r e^{i(\phi-\omega t)}$. The motion of the centre of the sphere is then $\mathbf{r}=(x,y,z)=-\varepsilon\omega^{-2}(\cos\omega t, \sin\omega t, 0)$ and the expected potential response is

$$\Phi^{\text{resp}}(\mathbf{r})=\Phi_0\left(r-\frac{\varepsilon\hat{\mathbf{r}}\cdot\hat{\mathbf{x}}}{\omega^2}\right)-\Phi_0(r)=-\frac{\varepsilon\hat{\mathbf{r}}\cdot\hat{\mathbf{x}}}{\omega^2}\frac{d\Phi_0}{dr}. \tag{64}$$

Note that the response is in phase (strictly, 180° out of phase) with the applied perturbation. The numerical solution for the response, equation (50), agrees with the expected result, equation (64), to the accuracy of the calculation for $\omega\in[0.01, 100.0]$. This is a significant test of the implementation since reproducing the proper self-gravitating response requires a non-trivial combination of the resonant and non-resonant interactions.

A satellite in circular orbit at r_s outside the sphere of the central galaxy (radius R) produces perturbation proportional to $(r/r_s)^l$. Thus the lowest order harmonics give the strongest response and as $r_s/R\to\infty$, the dipole ($l=1$) contribution will dominate. This is identical to the situation described in the paragraph above and the response is consequently given by equation (64). In fact, the translation causes the primary to orbit about the common primary-satellite centre of mass; this is easily shown by analogy with the Kepler two-body problem for circular orbits.

The density response ρ_1^{resp} of the galaxy to a satellite orbiting outside ($r_s=3.5$, $r_c=0.01$) is shown in Fig. 1(a). This response is nearly 180° out of phase and thus exerts *no* torque on a satellite. Although multipole components $l=1,\ldots,5$ have been included in this figure,

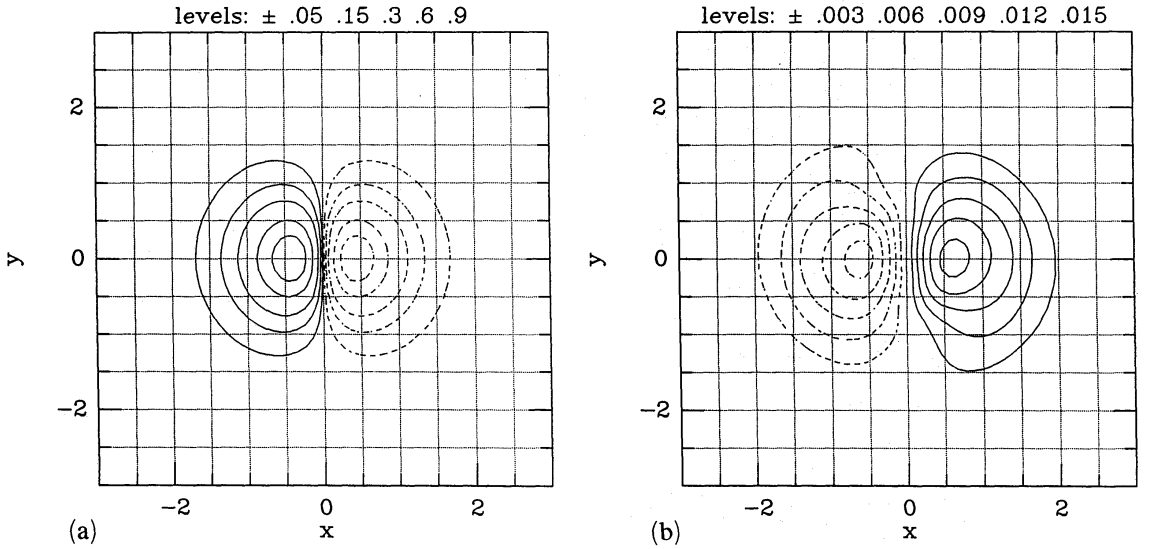

Figure 1. Density response ρ_1 of the sphere to a satellite ($r_c = 0.01$, $m_s = 1$, $l = 1, \ldots, 5$) orbiting at $(x, y) = (3.5, 0)$ for with (a) and without (b) self-gravity. The solid (dashed) lines show overdensity (underdensity). The contour levels for $|\rho_1|$ are stated above the figure. A very small core radius was chosen for this example to insure that the satellite is completely outside the primary.

components with $l > 1$ make negligible contribution. If the self-gravity is *not* included, the $l = 1$ response is no longer symmetric about the x axis and does not represent the barycentric motion. The response ρ_1^{resp} in the non-self-gravitating case is shown in Fig. 1(b). The centre of mass of the primary is shifted into the 4th quadrant resulting in a torque on the satellite orbit. This should be no surprise since barycentric motion is an aspect of the system's self-gravity. For a satellite orbiting inside the primary, the self-gravitating response still involves a centre of mass motion but this will not be a uniform translation in general.

The dramatic qualitative difference between the self-gravitating and non-self-gravitating dipole response may help explain the difference in satellite decay times for the restricted three-body simulations of Lin & Tremaine (1983) and White (1983). In the former calculation, the unperturbed potential of the primary galaxy was fixed in inertial space, just as the case shown in Fig. 1(b); this wake results in a strong dipolar torque. White, in one of his experiments, allowed a more massive body representing the centre of the galaxy to respond consistently to the forcing. It is now clear that including the barycentric motion by 'unpinning' the primary centre in this way is equivalent to admitting some part of the $l = 1$ self-gravitating response. Consequently, the dipolar torque may have been much reduced leading to a longer orbital decay time. Conversely, restricted N-body simulations (e.g. Lin & Tremaine 1983) are non-self-gravitating and must have a 'pinned' centre for consistency. However, the N-body simulator must be careful in applying this approximation: simply unpinning the centre is an incomplete representation of even the dominant $l = 1$ response especially for $r_s < R$!

3.2 DENSITY AND POTENTIAL RESPONSE

In this section, I illustrate the density response by presenting several sets of examples of the perturbed density ρ_1 in the orbital plane of the satellite. The first set, Fig. 2, shows the self-gravitating and non-self-gravitating response for a satellite with $r_c = 0.3$ in circular orbit at $r = 1$. Approximately 63 per cent of the mass of the primary is contained within $r = 1$. The satellite's location is $x = r$, $y = 0$ and is indicated by the symbol ☆. The non-self-gravitating response is obtained by setting $\mathscr{R} = \mathbf{M}$ in equation (45). To illustrate the contribution from the

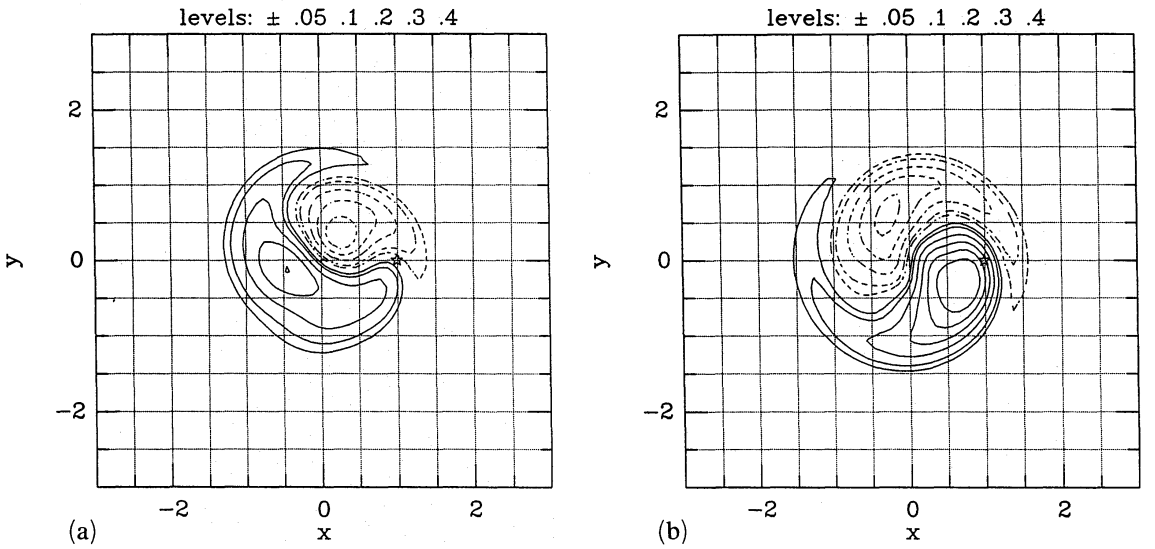

Figure 2. Density response ρ_1 of the sphere to a satellite ($r_c = 0.3$, $m_s = 1$, $l = 1, \ldots, 5$) orbiting at $(x, y) = (1, 0)$ for with (a) and without (b) self-gravity. The position of the satellite is indicated by ★.

various harmonics to the density ρ_1 shown in Fig. 2, the contribution of the harmonic $l = 1$ only shown in Fig. 3 and the contribution of harmonics $l = 2$–5 is shown in Fig. 4. Clearly, the $l = 1$ contribution dominates the response. We note that the shapes of the self-gravitating and non-self-gravitating $l = 1$ wakes are similar but the self-gravitating wake lags by roughly 90°. The self-gravitating response is slightly stronger. For the higher order harmonics shown in Fig. 4, the strength of the self-gravitating wake is stronger by roughly 20 per cent with only a small phase lag relative to the non-self-gravitating wake.

Since the differences between the self-gravitating and non-self-gravitating responses are dominated by the $l = 1$ term, let us concentrate on understanding Fig. 3. Again, it is helpful to first consider the case of a satellite orbiting outside the primary. In Section 3.1, we saw that a satellite fully outside the primary galaxy excites a body translation, the barycentric response. This response has the form $-Y_{11}(\theta, \phi)\varepsilon\hat{\mathbf{r}}\cdot\hat{\mathbf{x}}\omega^{-2}d\Phi_0/dr$ and is illustrated in Fig. 1(a) for a satellite orbiting at $r = 3.5$. The response is nearly in phase with the perturbation and exerts

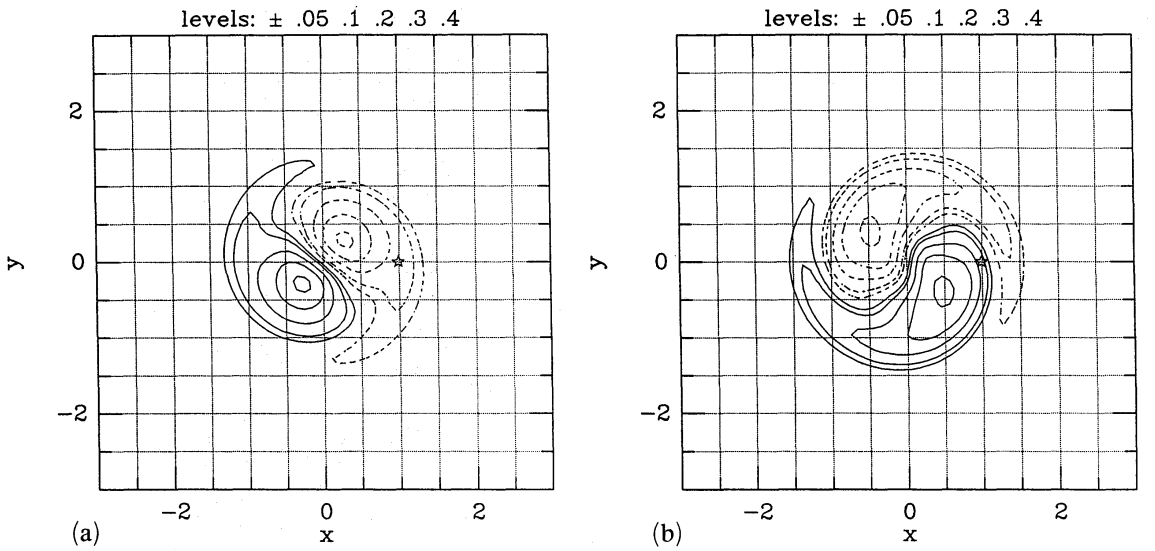

Figure 3. Density response as in Fig. 2 but for $l = 1$ contribution only.

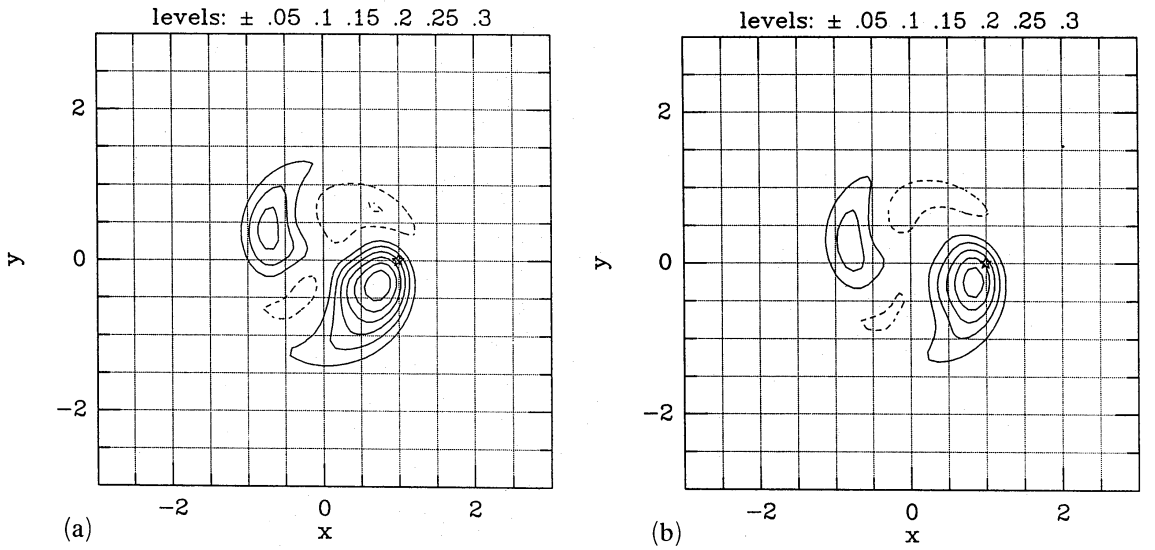

Figure 4. Density response as in Fig. 2 but for $l=2, \ldots, 5$ contributions only.

little torque. The present situation for a satellite within the primary is more complicated; the response of the primary not simply a body translation. Nonetheless, for a particle in the inner galaxy, there is not much difference between a satellite orbiting in the outer parts of the halo or outside the galaxy entirely and we expect similar behaviour. Indeed, the phase angle and morphology of the wake seems midway between the barycentric response (Fig. 1a) and the non-self-gravitating response (Fig. 2b). Of course, this explanation is not rigorous since the self-consistent response must satisfy equation (45). Nevertheless, the shift in the phase of the self-gravitating wake to the 3rd quadrant implies a lower torque since the moment arm of the force in the direction of the satellite's motion from is smaller.

At small r, the barycentric motion is much less important. This is well-illustrated by the wake for a satellite with $r_c=0.1$ orbiting at $r=0.12$. The wakes are qualitatively similar in both the self-gravitating (Fig. 5a) and non-self-gravitating (Fig. 5b) cases. Both have similar phase lags and are dominated by the dipole ($l=1$) component, although the self-gravitating wake has higher amplitude. The effect of the barycentric response is small, consistent with the small

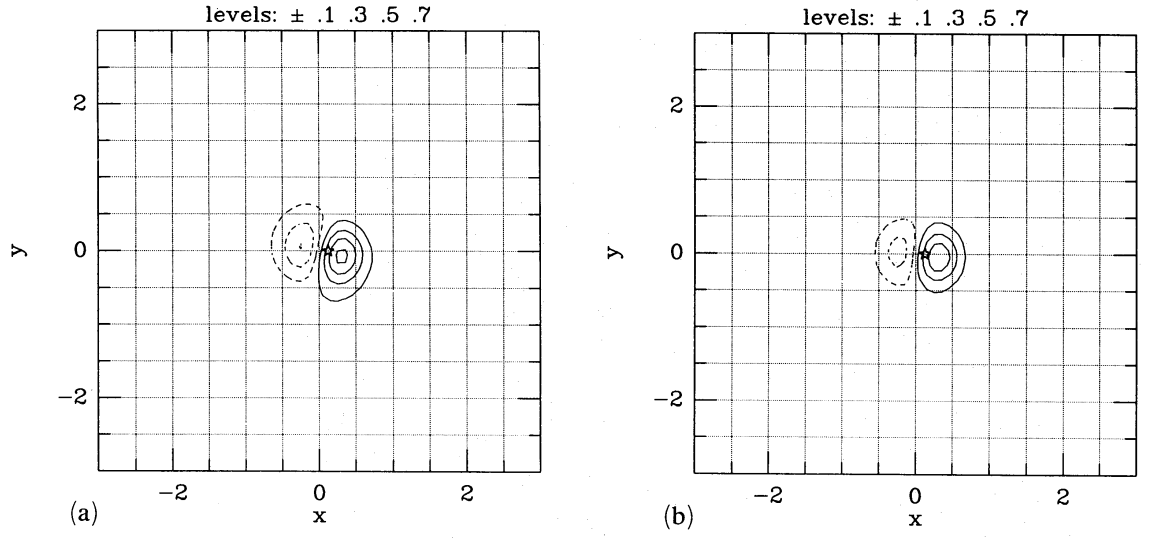

Figure 5. Density response for a satellite ($r_c=0.1$, $m_s=1$, $l=1, \ldots, 5$) orbiting at $(x, y)=(0.12, 0)$ for with (a) and without (b) self-gravity. The position of the satellite is indicated by ☆.

fraction of mass enclosed by the satellite's orbit. The sequence of Figs 1, 2 and 5 illustrates the dominance of the barycentric motion at large radii, its importance at intermediate radii, and its relative insignificance to the response deep within the primary.

Examining Figs 1–5 together, we clearly see that both the self-gravitating and non-self-gravitating responses are global; the wake involves a significant fraction of the galaxy. Based on the success of Chandrasekhar's formula in reproducing the results of the full self-gravitating simulation in the case of the $n=3$ polytrope (BvA), one might infer that the decay is a local process, which the results above demonstrate is not so. The reason the local theory appears to work well in this case stems from the way in which Chandrasekhar's theory is applied. To compute the orbital decay using Chandrasekhar's formula, one represents a particular location in the galaxy by an infinite homogeneous medium with the local density and velocity structure. If r_c is the size of the perturber, then we know that contributions to the drag in an infinite homogeneous medium come roughly equally from logarithmic scales larger than b. To the extent that $\ln R/r_c \gg 1$, the orbits in the spherical system will appear rectilinear to the satellite over many decades in scale and the application Chandrasekhar's formula will accurately represent the physical situation. In the present case, $\ln 10 \not\gg 1$ and Chandrasekhar's formula will not give a good description of the underlying dynamics. The fact that the resultant torque can be parametrized by Chandrasekhar's formula for a range of r_s attests to the self-similarity of the orbital structure in the primary, not the locality of the response. For example, in Paper II, I considered the non-self-gravitating response of a Plummer law satellite orbiting in a singular isothermal sphere. Since the isothermal sphere is self-similar in radius, the response will also be self-similar. For this reason, the spherical torque formula and the application of Chandrasekhar's formula have the same scaling. Thus, the spherical theory may be used to fit for $\ln \Lambda$ which, in this case, embodies the orbital geometry. In general, we should not expect that Chandrasekhar's formula will successfully describe the orbital decay of an extended satellite. For example, a fit of Chandrasekhar's formula to a N-body simulation of orbital decay in a King model leads to $\ln \Lambda \propto r$ (Bontekoe 1988).

3.3 ORBITAL DECAY

The torque on the satellite by the response is computed using equation (53). The contribution to the torque for each harmonic l for the satellite orbit (with $r=1$) discussed in Section 3.2 is shown in Table 1. The self-gravitating contribution of the $l=1$ term is smaller than the non-

Table 1. Torque ($r=1$, $r_c=0.3$).

l	m	τ_z (Self)	τ_z (Non-self)	τ_z (Self)/τ_z (Non-self)
1	1	-1.57×10^{-1}	-8.97×10^{-1}	0.18
2	2	-2.67×10^{-1}	-1.98×10^{-1}	1.35
3	1	-2.95×10^{-2}	-1.90×10^{-2}	1.55
3	3	-5.47×10^{-2}	-4.90×10^{-2}	1.12
3	1 + 3	-8.42×10^{-2}	-6.80×10^{-2}	1.24
4	2	-7.72×10^{-3}	-6.21×10^{-3}	1.24
4	4	-2.00×10^{-2}	-1.97×10^{-2}	1.05
4	2 + 4	-2.77×10^{-2}	-2.59×10^{-2}	1.07
5	1	-1.42×10^{-3}	-1.12×10^{-3}	1.27
5	3	-3.71×10^{-3}	-3.21×10^{-3}	1.16
5	5	-9.36×10^{-3}	-9.00×10^{-3}	1.04
5	1 + 3 + 5	-1.45×10^{-2}	-1.33×10^{-2}	1.09

self-gravitating contribution by a factor of ≈ 5. For all other harmonics, self-gravity strengthens the torque by as much as 35 per cent. This, of course, is consistent with the appearance of the wake considered in Section 3.2.

The number of harmonics needed to reproduce the torque accurately depends on r_c. For $r_c \to 0$, the contribution from each harmonic is proportional to $1/l$ and the sum diverges logarithmically. For a finite r_c, the contributions drop off quickly above a critical l. This critical $l \propto r_c^{-1}$ as discussed in Paper II. As one might expect, self-gravity becomes less important for larger l (see Table 1). Thus, for computational economy, I replace all terms $l > 4$ with the non-self-gravitating results which may be computed cheaply using the technique from Paper II. I verify that all sums have converged by $l = 8$ to better than 2 per cent for $r_c/R = 0.1$ by estimating the convergence from extrapolation to large l.

The effects of self-gravity on the decay are nicely summarized by examining $\tau_z(r)$ which is shown in Fig. 6. The upper (lower) set of values shows τ_z for the self-gravitating (non-self-gravitating) case. The open circles show the points at which the τ_z is evaluated. The decay curve $r(t)$ is computed by integrating $\tau_z(r) = dJ/dt$ for a circular orbit. The resulting orbital decay curves are given in Fig. 7. For this figure, I have used the units $G = 1$, $M = 2$, $R = 1$ for consistency with ZW. The decay time for the non-self-gravitating case is shorter by a factor of 2–3 than the self-gravitating case.

The self-gravitating result can be compared with the N-body simulations for the identical model by BvA (their Model A and fig. 3) and ZW (their fig. 1). Both groups find that the satellite decays from $r = 1$ to $r = 0.2$ in approximately 20 (ZW) time units whereas the analytic theory predicts the decay should occur in 30 time units. This difference is due to non-linear effects. For a satellite of mass $m_s = 0.03M$ rather than $m_s = 0.1M$, we find that simulation and analytic results are in good agreement for both the self-gravitating and non-self-gravitating cases (Hernquist & Weinberg 1989). Unfortunately, I cannot compare my non-self-gravitating case with BvA's Model C. Although the particles move in a rigid potential field, this field is tied to the centre of the particle distribution which moves about the combined primary-satellite centre of gravity. This case is neither purely self-gravitating or non-self-gravitating. Nonethe-

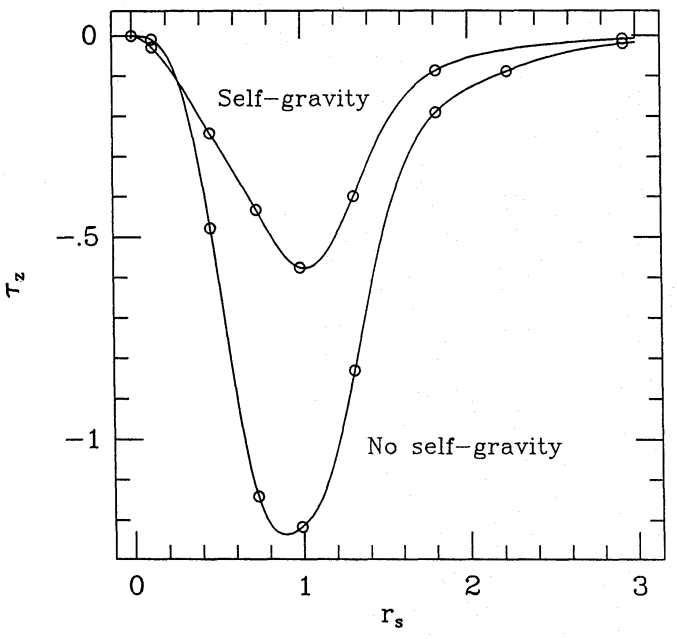

Figure 6. The self-gravitating and non-self-gravitating torque shown as a function of radius for the circular orbits. The calculated points (open circles) are fit by a cubic spline (solid line).

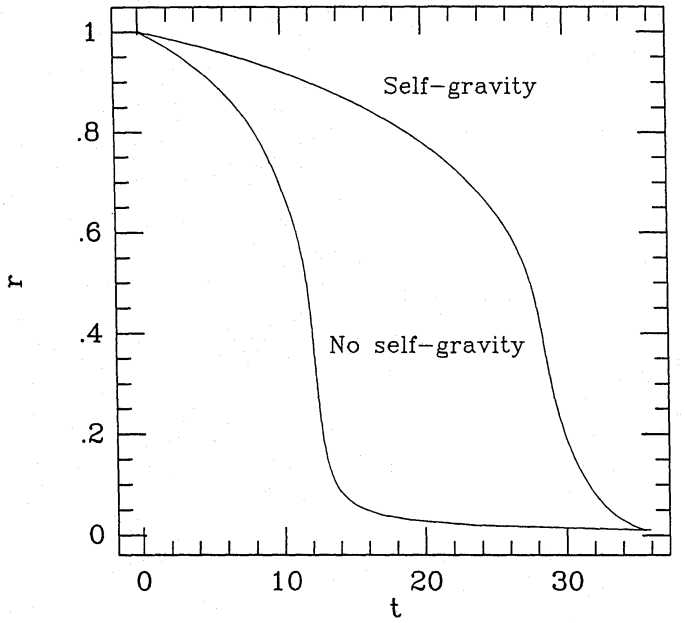

Figure 7. Radius versus time for the orbital decay with units $G=1$, $M=2$, $R=1$. Both the self-gravitating and non-self-gravitating cases are shown. The time required to sink from $r=1.0$ to $r=0.2$ differs by ~ 2.4.

less, the fact that Model A and C give nearly the same decay profile illustrates that the admission of barycentric motion is the most important difference between the self-gravitating and non-self-gravitating dynamical friction for an $n=3$ polytrope. I again emphasize that this result will not be true in general.

4 Summary

In this paper, I report on theoretical progress made on the sinking satellite problem. Using the method of Section 2, I have computed the self-gravitating dynamical friction on a Plummer law satellite orbiting in an $n=3$ polytrope. The ratio of satellite to core radius size of the system is $r_c/R=0.1$. The basic results are:

(i) The self-gravitating response of the galaxy significantly effects the orbital decay. The decay time is ~ 2.5 times *longer* with self-gravity.

(ii) Result (i) can be understood in terms of the barycentric motion. For a satellite orbiting outside the primary, the dipole ($l=1$) perturbation by the satellite gives rise to a uniform force field. The self-gravitating dipole response of the galaxy corresponds to a uniform translation directed opposite to the position vector of the satellite and the magnitude of the translation is the radius of its barycentric orbit. The associated wake is symmetric and in phase with the perturbation, hence it exerts no torque on the satellite. In contrast, the non-self-gravitating torque is dominated the dipole contribution. For a satellite inside the primary, the dipole response is more complicated, but is qualitatively similar for the fraction of the galaxy well inside the satellite's orbit. This qualitative difference in self-gravitating and non-self-gravitating dipole responses is responsible for the difference in decay times. For higher-order multipoles ($l>1$), the self-gravitating response is larger but qualitatively similar to the non-self-gravitating response.

(iii) For the soft satellite ($r_c/R=0.1$) considered here, by BvA, and by ZW, the response of the primary is *global* in both the self-gravitating and non-self-gravitating cases. The success of

Chandrasekhar's formula in fitting the results obtained by BvA is probably due to a similarity in scaling and does not by itself indicate a purely local response. However, for cases with $\ln R/r_c \gg 1$, the orbits in the spherical system will appear rectilinear to the satellite over many decades in scale and the application Chandrasekhar's formula will accurately represent the physical situation.

These conclusions should not be sensitive to the choice of (non-pathological) model. However, the relative importance of various multiple contributions and the ratio of times for self-gravitating and non-self-gravitating orbital decay will be model dependent. The model considered here illustrates the physics of satellite decay and is probably too unrealistic for direct comparison with observations. In particular, the rotation curve of an $n=3$ polyrope is not a good representation of galaxian rotation curves and a Plummer model most likely gives the satellite too much mass at large radii. Nonetheless as an estimate, for a galaxy with $M=2\times10^{11}M_\odot$ and $R=40$ kpc, we find that a satellite with $m_s=2\times10^{10}M_\odot$ decays from the edge to 5 kpc in 10^{10} yr.

The model applies to the orbital decay of a single extended satellite. We have seen that the self-gravitating global response is crucial to understanding the decay process. If there is more than one such satellite orbiting the primary, their responses would interfere and decrease the drag in general. For example, two diametrically oriented satellites in an $n=3$ polytrope each of mass m_s will take nearly twice the time to decay as a single satellite of mass m_s. For very compact satellites whose size is very much smaller than the scale length of the primary, the present analysis is less relevant. For example, I would not expect that the global effects discussed in this paper to influence our understanding of globular cluster orbits near the Galactic Centre.

The calculation presented here is the first self-consistent analytic determination of the response and subsequent decay of two interacting galaxies and certainly demonstrates the power of the approach. The technique described in Section 2 for computing the self-gravitating response of a spherical galaxy is quite general and can be used to investigate, for example, the modes of the system as well as a forced response. It may even be possible to extend the technique to spheroidal and triaxial galaxies. Work on these problems is in progress.

Acknowledgments

It is a pleasure to thank James Binney, Jeremy Goodman, Herb Rood, Martin Schwarzchild, Scott Tremaine and Tim de Zeeuw for stimulating discussions and comments on the manuscript. I thank Lars Hernquist for the same and for aiding in interpreting the *N*-body simulations. I also acknowledge the encouragement of John Bahcall. Supercomputing time was supplied through a grant provided by the Pittsburgh Supercomputing Centre. This research was supported in part by a fellowship from the Siemens Corporation.

References

Antonov, V. A., 1962. Vestnik Leningrad Univ., **19,** 96.
Barnes, J., 1986. In: *Dynamics of Star Clusters, IAU Symp. No. 69,* p. 297, ed. Hayli, A., Reidel, Dordrecht.
Barnes, J., Goodman, J. & Hut, P., 1986. *Astrophys. J.,* **300,** 112.
Bontekoe, Tj. R., 1988. *PhD thesis,* University of Groningen.
Bontekoe, Tj. R. & van Albada, T. S., 1987 (BvA). *Mon. Not. R. astr. Soc.,* **224,** 349.
Chandrasekhar, S., 1943. *Astrophys. J.,* **97,** 251.
Clutton-Brock, M., 1972. *Astrophys. Space Sci.,* **16,** 101.
Clutton-Brock, M., 1973. *Astrophys. Space Sci.,* **23,** 55.

Doremus, J. P., Feix, M. R. & Baumann, G., 1971. *Phys. Rev. Lett.*, **26,** 725.
Edmonds, A. R., 1960. *Angular Momentum in Quantum Mechanics,* Princeton University Press, Princeton.
Fridman, A. M. & Polyachenko, V. L., 1984. *Physics of Gravitating Systems,* Vol. II, p. 282, Springer-Verlag, New York.
Goldstein, H., 1950. *Classical Mechanics,* Addison-Wesley, Reading.
Hénon, M., 1973. *Astr. Astrophys.*, **24,** 229.
Hernquist, L. & Weinberg, M. D., 1989. *Mon. Not. R. astr. Soc.,* **238,** 407.
Kalnajs, A. J., 1977. *Astrophys. J.*, **212,** 637.
Kalnajs, A. J., 1971. *Astrophys. J.*, **212,** 637.
Krall, N. A. & Trivelpiece, A. W., 1973. *Principles of Plasma Physics,* McGraw-Hill, New York.
Landau, L. D., 1946. *J. Phys. USSR,* **10,** 25.
Lin, D. N. C. & Tremaine, S., 1983. *Astrophys. J.*, **264,** 364.
Mulder, W. A., 1983. *Astr. Astrophys.*, **117,** 9.
Palmer, P. L. & Papaloizou, J., 1985. *Mon. Not. R. astr. Soc.*, **215,** 691.
Palmer, P. L. & Papaloizou, J., 1987. *Mon. Not. R. astr. Soc.*, **224,** 1043.
Quinn, P. J. & Goodman, J., 1986. *Astrophys. J.*, **309,** 472.
Tremaine, S. 1981. In: *The Structure and Evolution of Normal Galaxies,* p. 67, eds Fall, S. M. & Lynden-Bell, D. Cambridge University Press.
Tremaine, S. & Weinberg, M. D., 1984. *Mon. Not. R. astr. Soc.*, **209,** 729 (Paper I).
Weinberg, M. D., 1986. *Astrophys. J.*, **300,** 93 (Paper II).
White, S. D. M., 1983. *Astrophys. J.*, **274,** 53.
Zaritsky & White, S. D. M., 1987 (ZW). Preprint.