

Noise-driven evolution in stellar systems – I. Theory

Martin D. Weinberg[★]

Department of Astronomy, University of Massachusetts, Amherst, MA 01003-4525, USA

Accepted 2001 July 30. Received 2001 June 28; in original form 2000 July 24

ABSTRACT

We present a theory for describing the evolution of a galaxy caused by stochastic events such as weak mergers, transient spiral structure, orbiting blobs, etc. This noise excites large-scale patterns that drive the evolution of the galactic density profile. In a dark matter halo, the repeated stochastic perturbations preferentially ring the lowest-order modes with only a very weak dependence on the details of their source. The subsequent redistribution of halo mass is determined only by the mechanics of these modes. The halo profile then evolves toward a universal asymptotic form for a wide variety of noise sources. Such a convergence may help explain the similarity of normal galaxy morphology in diverse environments. A variety of other applications are discussed.

Key words: galaxies: evolution – galaxies: haloes – galaxies: kinematics and dynamics – cosmology: theory – dark matter.

1 INTRODUCTION

Galaxies never reach a true equilibrium state. The outer halo in large galaxies is less than 10 dynamical times old, so primordial inhomogeneities will not have had time to phase mix and continued disturbances from mergers will not have relaxed (see Tremaine 1992 for additional discussion). These features, as well as those from other intrinsic sources of noise, such as a population of 10^6 - M_{\odot} massive black holes, gas accretion, orbiting dwarf galaxies, debris streams (Johnston 1998; Morrison et al. 2000) or dark clusters, are amplified by the self-gravity of the halo. These distortions create potentially *observable* asymmetries in the stellar and gaseous Galactic disc at the current epoch. In addition, this noise may be sufficient to drive the evolution of the halo in a variety of environments (Weinberg 2001, hereafter Paper II). This paper provides a theoretical framework for studying the evolution of a near-equilibrium galaxy caused by scattering of orbits by fluctuations in the gravitational potential.

The motivation for this theory is as follows. Previous work has shown that fluctuations in stellar systems on the largest scales can be strongly amplified by their own self-gravity. For example, in the case of particle noise, this means that large-scale fluctuations will greatly exceed their Poisson amplitudes. Weinberg (1993) explored the idealized case of periodic cube; the fluctuations become very large as the system size approaches its Jeans' length. This does not apply directly to an equilibrium galaxy but does suggest that the fluctuations in bound systems will be largest for the discrete modes with the largest spatial scales. For example, Weinberg (1994) argues that galaxies will often have very weakly damped $m = 1$ (sloshing or seiche) modes and these result in large excitations

when excited (see Vesperini & Weinberg 2000). These weakly damped modes are similar to those which cause Landau damping in a plasma (e.g. Binney & Tremaine 1987). Putting this together, one might ask: if noise preferentially excites particular modes with little dependence on the details of the noise source, is it possible that the repetitive stochastic response of the galaxy will drive the equilibrium toward some characteristic profile, independent of its initial conditions?

In order to answer this question and address related applications, this paper concentrates on a theoretical framework for describing the evolution of a galaxy by stochastic fluctuations. Beginning with a description of the linear response of a galaxy to a perturbation, and assuming that the process is Markovian, one may expand the Boltzmann collision term in a series. Analogous to two-body collisions, only the first two terms contribute and the resulting evolution equation has a Fokker–Planck form. This equation is very far from being analytically tractable because the Fokker–Planck coefficients depend on integrals over all phase space under the stochastic perturbation. It is straightforwardly solved numerically, however. This approach is perturbative and assumes that the stellar system remains near an equilibrium and therefore applies to galaxies after formation. Weinberg (Paper II) will demonstrate that the repetitive self-gravitating response to noise does indeed evolve a halo toward an asymptotically universal profile. Damped modes are key to understanding this result. These modes slowly redistribute mass to the outer halo in a way that depends on the mode alone, and repetitive excitation leads to self-similarity.

Given these analytic and numerical complications of this approach, it might appear easier to treat these problems by n -body simulation. However, the interpretation of such simulations are complicated by the intrinsic fluctuations of the gravitational

[★]E-mail: weinberg@astro.umass.edu

potential as a result of the finite number of particles. In particular, the graininess intrinsic to n -body simulations causes the orbital conserved quantities to diffuse in phase space even though the system is in equilibrium. Therefore, resonances whose libration periods are larger than the diffusion time can not exist in an n -body simulation.¹ On the other hand, the actions are conserved for all time in a true collisionless system. This difference-limits the time interval over which an n -body simulation approximates a collisionless stellar system and limits the ability of a simulation to resolve the collective dynamics which drive large-scale patterns and depend on coherence in phase space. In addition, n -body fluctuations provide a background source that may be larger than natural noise sources. Even if the astronomical noise source is strong overall, the true noise is likely to have large spatial scale and low frequency. The response to low-frequency noise may be wiped out by the diffusion caused by the n -body fluctuations. For this reason, the often-used test of doubling or halving the particle number will be specious, because particle noise likely will continue to dominate the desired collisionless effects unless one is unwittingly fortunate. In summary, one needs to understand the details underlying dynamical mechanisms in order to interpret n -body results. Simulations, therefore, are likely to be inefficient and possibly deceptive on their own, but coupled with some analytic underpinning they should be a useful tool for studying noise phenomena in galaxies.

The plan for this paper is as follows. The overall analytic approach is outlined in Section 2, with an explicit derivation for spherical haloes given in Appendix A. This could be extended straightforwardly to most often-used geometries (e.g. elliptical, disc or disc and halo together). We characterize the noise for two general cases, transient and quasi-periodic perturbers (a halo of black holes, for example) in Section 3. These two cases result in qualitatively different behaviour and represent the most plausible astronomical scenarios. We will see that transient noise is probably the most relevant and important, but an understanding of the qualitative difference between the two cases is insightful. Finally, the main findings are summarized and discussed in Section 4. Derivation details appear in the Appendices. Paper II will review the basic physics and apply these methods to understanding the evolution of halo in a noisy environment and may be a better place to begin for those interested in astronomical consequences rather than the kinetic theory.

2 DEVELOPMENT OF THE EVOLUTION EQUATION

The general problem of deriving a phase-space evolution equation for stochastic events will be familiar from the collisional dynamics literature and could be solved following the standard two-body approach: beginning with the collisional Boltzmann equation, one writes the collision term in Master equation form and expands in a Taylor series to derive Fokker–Planck equation (e.g. Binney & Tremaine 1987; Spitzer 1987). For a spherically symmetric system, the perturbed, out-of-equilibrium phase-space distribution is a function of two action and two angle variables. Averaging over times which are short compared to the evolution time but long

¹ Mathematically, resonance with a bar with pattern speed Ω_b in an axisymmetric disc has the form: $n\Omega_r + m\Omega_\phi = m\Omega_b$. This corresponds to an orbit that makes m radial oscillations for every n rotations and therefore has a closed-orbit period $P = 2\pi m/\Omega_r$.

compared to the dynamical time (orbit-averaging), one obtains a Fokker–Planck equation.

Alternatively, recent work in statistical mechanics and noise theory has developed a body of methods for treating stochastic differential equations based directly on transition probabilities. If $P(x', t + \tau|x, t)$ is the transition probability to some new state x' at time $t + \tau$ from the initial state x at time t , then the following integral equation determines all subsequent evolution of the distribution $f(x)$:

$$f(x, t + \tau) = \int dx' P(x, t + \tau|x', t) f(x', t).$$

By expanding the transition probability in moments of $x - x'$ for small τ , one can derive a differential equation of the form:

$$\frac{\partial f(x, t)}{\partial t} = \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial x} \right)^n D^{(n)}(x, t) f(x, t), \quad (1)$$

where the coefficients $D^{(n)}$ are proportional to the time-derivative of the moments in transition probability:²

$$D^{(n)}(x, t) = \frac{1}{n} \lim_{\tau \rightarrow 0} \frac{1}{\tau} \int dx' P(x', t + \tau|x, t) (x' - x)^n. \quad (2)$$

This is known as the *Kramers–Moyal expansion*. The main advantage of the Kramers–Moyal expansion is that the noise process appears explicitly as the solution to an initial-value problem. Because the response of the galaxy to a stochastic event can be computed straightforwardly in perturbation theory or by n -body simulation (described below), we can compute the noise-driven evolution for a wide variety of astronomical scenarios. For this reason, the Kramers–Moyal expansion is better suited to treating general stochastic noise than the Master equation, even though the two approaches are formally equivalent.

For galaxy evolution, the state variable x in the equations above is replaced with the six-dimensional phase-space vector. We are interested in long-term behaviour and may simplify the problem by orbit-averaging the evolution equation. Writing the phase space in action-angle variables, we may use the averaging theorem (Arnold 1978) to turn the orbit average into an angle average. This gives us an evolution equation that is a function of actions alone. We can solve for the change in actions of any orbit in the galaxy using Hamiltonian perturbation theory for any given noise process and similarly derive the change in the phase-space distribution function (Weinberg 1998, hereafter W98). Evaluation of these quantities lead directly to the moments needed for the Fokker–Planck-type evolution equation. The most subtle aspect of the development below is enforcing a consistent time-ordering. The angle averaging defines two time-scales: a short dynamical time-scale and a long evolutionary time-scale. The stochastic perturbations may last for several dynamical time-scales but they are instantaneous from the evolutionary point of view. Although this approximation may seem restrictive and only marginally true for some scenarios of interest, it yields results that are in good agreement with n -body simulation in the case of globular cluster evolution and the closely related case of self-gravitating fluctuations explored W98.

We will use the method described in W98 to derive the coefficients $D^{(1)}$ and $D^{(2)}$. Briefly summarized, we represent distortions in the structure of halo in a biorthogonal basis whose member pairs, d_j^{lm} and u_j^{lm} , satisfy the Poisson equation,

² The factorial divisors from the Taylor series expansion are absorbed in to the coefficients $D^{(n)}$ following Risken (1989).

$\nabla^2 u_j^{lm} = d_j^{lm}$. Any distortion then can be summarized by a set of coefficients with three indices. We can truncate this expansion and still recover most of the power, because large spatial scales are most important in understanding global evolution. Moreover, we can analytically compute the self-gravitating response of the halo to some arbitrary disturbance using Hamiltonian perturbation theory as previously described. This development gives us the perturbed quantity x' for all phase-space variables (see equations 2 and A3) and the appropriate ensemble averages give us the required diffusion coefficients. See Appendix A for additional detail.

By using the same biorthogonal basis as the potential solver for an n -body simulation of a desired transient process (e.g. Clutton-Brock 1972, 1973, Kalnajs 1976, Hernquist & Ostriker 1992, Weinberg 1999), the time-series of coefficients for describing the noise source can be used directly to derive the coefficients $D^{(1)}$ and $D^{(2)}$. The coefficients for other n -body Poisson solvers can be easily computed from phase-space slices at discrete times. The equilibrium distribution corresponds to the time-invariant ('DC') component and does not contribute. We will consider point mass perturbers in Section 3.1 and transient perturbers.

We work out the details for two examples in the following section: quasi-periodic distortions such as orbiting super-massive black holes (Section 3.1) and transient perturbers such as dwarf galaxy mergers, decaying substructure and spreading debris trails (Section 3.2). The consequences of both of these will be explored in Paper II.

3 NOISE MODELS

The noise response can be divided into two classes depending on its frequency spectrum: quasi-periodic or transient. Typical quasi-periodic sources are orbiting black holes, low-mass dwarfs and star clusters. A time-series analysis of expansion coefficients from a quasi-periodic noise source will contribute power at a discrete set of frequencies; a *line* spectrum. In practice, the linewidths are inversely proportional to the evolutionary time-scale for the near-equilibrium gravitational potential. This time is the age of the galaxy at most. This width may be relatively large for an orbiting satellite with mean radius of roughly 100 kpc and periods of several Gyr.

Transient sources have aperiodic time variation. The power from these time-series are characterized by a significant continuum component. The perturbation can couple to modes over a continuous range of frequencies, and therefore any weakly damped modes in the range will be excited automatically. In contrast, the line frequency from a line spectrum needs to nearly coincide with the modal frequency to produce a large excitation. Although commensurabilities in a line spectrum densely cover real frequencies in general, those with large integer multiples are likely to have very small amplitudes and require unrealistically long time-scales to be of practical interest. Typical astronomical examples with continuous spectra include bodies on decaying orbits such as large satellites, disrupting dwarfs and fly-by encounters. The case of orbital decay leads to a continuous spectrum with broadened line-like features.

Evaluation of the response is similar in both cases. For the line spectrum, one sums over discrete responses at each line. For the continuous spectrum, one integrates over the frequencies, which is also a sum over frequencies in practice. In both cases therefore, the response to a single frequency is the building block and we begin with a development common to both cases.

The goal is the calculation of the coefficients needed to solve the evolution equation (1) and defined in detail by equations (A3) and (A4). Any perturbed quantity can be represented as a Fourier series in angles with coefficients depending on actions because orbits in the equilibrium phase space are quasi-periodic and representable as fixed actions and constantly advancing angles. Following W98, the perturbed Hamiltonian is

$$H(\mathbf{I}, \mathbf{w}) = H_o(\mathbf{I}) + H_1(\mathbf{I}, \mathbf{w}) \quad (3)$$

$$= H_o(\mathbf{I}) + \sum_{\mathbf{l}} H_{1\mathbf{l}}(\mathbf{I}) e^{i\mathbf{l} \cdot \mathbf{w}}, \quad (4)$$

where

$$\begin{aligned} H_{1\mathbf{l}}(\mathbf{I}) &= \sum_{l=0}^{\infty} \sum_{m=-l}^l \sum_j Y_{ll_2}(\pi/2, 0) r_{l_2 m}^l(\beta) \mathbf{W}_{ll_2 m}^{l_1 j}(\mathbf{I}) a_j^{lm}(t) \quad (5) \\ &= \sum_{l=0}^{\infty} \sum_{m=-l}^l \sum_j Y_{ll_2}(\pi/2, 0) r_{l_2 m}^l(\beta) \mathbf{W}_{ll_2 m}^{l_1 j}(\mathbf{I}) \int_{-\infty}^{\infty} d\omega e^{i\omega t} \\ &\quad \times \sum_k [\mathbf{M}_{jk}^{lm}(\omega) + \delta_{jk}] b_k^{lm}(\omega), \quad (6) \end{aligned}$$

where $\mathbf{l} = l_1, l_2, l_3$ is a triple of integers, $r_{ij}^l(\beta)$ and $\mathbf{W}_{ll_2 l_3}^{l_1 j}(\mathbf{I})$ are the rotation matrices and action-angle transforms of the gravitational potential basis, u_j^{lm} . For a given spherical harmonic Y_{lm} , only terms with $l_3 = m$ contribute (Tremaine & Weinberg 1984).³ Therefore, features with $m = 1$ symmetry first occur at spherical harmonic order $l = 1$ and those with $m = 2$ symmetry first occur at $l = 2$. We will emphasize these two lowest-order symmetries $l = m = 1$ and $l = m = 2$ below. The time-dependence of the coefficients describing the response in equation (5), $a_j^{lm}(t)$, may be represented as its Fourier transform in time and this allows each frequency to be treated separately. The response matrix \mathbf{M} describes the reaction of the galaxy to the perturbation and depends on the equilibrium profile only. The sum $\mathbf{M}_{jk}^{lm} + \delta_{jk}$ therefore represents both the response and direct forcing. This development gives us equation (6) from equation (5). The interested reader can find the details in W98; the practical importance of this result is the convenient factorization of the response from the perturbation. We may integrate the equations of motion directly to evaluate $\mathbf{I}(\tau + t)$. Hamilton's equations yield

$$\dot{l}_j = -\frac{\partial H}{\partial w_j} = -i \sum_{\mathbf{l}} l_j H_{1\mathbf{l}}(\mathbf{I}) e^{i\mathbf{l} \cdot \mathbf{w}}, \quad (7)$$

and therefore we have

$$\Delta I_j(t + \tau) \equiv I_j(t + \tau) - I_j(t) = \int_t^{t+\tau} dt \dot{I}_j(t). \quad (8)$$

The evolution of the perturbed distribution function in time follows from the linearized, collisionless Boltzmann equation and the total time-derivative for a Hamiltonian system:

$$\dot{f}_1 \equiv \frac{\partial f_1}{\partial t} + [f_1, H] = \frac{\partial f_1}{\partial t} + \frac{\partial H_0}{\partial \mathbf{I}} \cdot \frac{\partial f_1}{\partial \mathbf{w}} = \frac{\partial H_1}{\partial \mathbf{w}} \cdot \frac{\partial f_0}{\partial \mathbf{I}}, \quad (9)$$

³ A regular system is described by an action-angle pair in each dimension. For a sphere, the Fourier analysis on the torus will have three indices. The symmetry of the sphere leads to an identically zero frequency corresponding to the restriction of orbits to a plane. The choice of a particular spherical harmonic, Y_{lm} , implicitly specifies the arbitrary axis and requires $l_3 = m$.

where only the first-order terms are retained and the final equality follows from Liouville's equation. A Fourier series expansion in action–angle variables analogous to the development above for $H_1(t)$ yields

$$\dot{f}_1(\mathbf{I}, \mathbf{w}, t) = \sum_{\mathbf{I}'} e^{i\mathbf{l} \cdot \mathbf{w}_i \mathbf{I}'} \cdot \frac{\partial f_0}{\partial \mathbf{I}} H_{1l}(\mathbf{I}, t), \quad (10)$$

and therefore

$$f_1(\mathbf{I}, \mathbf{w}, t + \tau) = \sum_{\mathbf{I}'} e^{i\mathbf{l} \cdot \mathbf{w}_i \mathbf{I}'} \cdot \frac{\partial f_0}{\partial \mathbf{I}} \int_t^{t+\tau} dt' H_{1l}(\mathbf{I}, t'). \quad (11)$$

To evaluate the evolution equation (equations 1 and A6), we need the first- and second-order action moments defined by equations (2) and (A5). We assume that τ is larger than intrinsic dynamical times consistent with the ordering of our slow and fast time-scales and use the time-asymptotic form for the response matrix \mathbf{M}^{lm} (see W98). Previous work, including detailed comparison to n -body simulations, suggests that this is a very good approximation for time-scales longer than several crossing times. The response of the galaxy to the distortion induces a shift in the actions \mathbf{I} and the overall response causes a change in the distribution function. This is represented in the matrix equation defined by equations (5) and (6) for an external perturbation described by the coefficients $b_j^{lm}(t)$. The stochastic variables are the coefficients $b_j^{lm}(t)$ themselves.

We will denote the ensemble average of the fluctuating coefficients with angle brackets, e.g. $\langle b_j^{lm}(t_1) b_j^{lm}(t_2) \dots b_j^{lm}(t_n) \rangle$. Note that $\langle b_k^{lm}(t_1) b_k^{*lm}(t_2) \rangle$ is related to the density correlation function:

$$\begin{aligned} \langle b_k^{lm}(t_1) b_k^{*lm}(t_2) \rangle &= \int d^3 r_1 \int d^3 r_2 Y_{lm}^*(\theta_1, \phi_1) Y_{lm}(\theta_2, \phi_2) u_j^{*lm} \\ &\quad \times (r_1) u_k^{lm}(r_2) \langle \rho(\mathbf{r}_1, t_1) \rho(\mathbf{r}_2, t_2) \rangle. \end{aligned} \quad (12)$$

Assuming that the process causing fluctuations is independent of time (e.g. a *stationary* process) we can write the correlation in terms of the time difference: $\langle \rho(\mathbf{r}_1, t_1) \rho(\mathbf{r}_2, t_2) \rangle = C(\mathbf{r}_1, \mathbf{r}_2, t_1 - t_2)$. The quantity $\langle b_k^{lm}(t_1) b_k^{*lm}(t_2) \rangle = \langle b_k^{lm}(0) b_k^{*lm}(t_1 - t_2) \rangle$ describes the correlation of random variables b_j^{lm} as a function of time. The limit $t_1 \rightarrow t_2$ gives the mean-squared fluctuation amplitude and was explored and compared to n -body simulations in W98.

The overall conditional probability required for evaluating the Kramers–Moyal moments (see equations 1, 2 and A1) has two contributions. First, the response of the galaxy changes the underlying distribution and consequently the probability of obtaining a given final state. Secondly, the resonant coupling changes the action of an orbit at a particular point in phase space. Altogether we have

$$\begin{aligned} P(\mathbf{I}', t + \tau | \mathbf{I}, t) f(\mathbf{I}, t) &= [f_0(\mathbf{I}) + f_1(\mathbf{I}, t) + \dots] \\ &\quad \times \delta \left[\mathbf{I}' - \mathbf{I} - \int_t^{t+\tau} dt' \dot{\mathbf{I}}(t') \right]. \end{aligned} \quad (13)$$

The first and second moments of $\Delta \mathbf{I} \equiv \mathbf{I}' - \mathbf{I}$ over this distribution are proportional to the square of the perturbation coefficients b^2 and are therefore second-order in the perturbation amplitude. With additional work, one can show that the next contribution is proportional to b^4 and relatively negligible. One now factors out $f(\mathbf{I}, t)$ to get the moments over the conditional probability. The f_1 contribution oscillates rapidly on a time-scale of $\mathcal{O}(\tau)$ and has zero mean. The f_2 term has a non-oscillatory term of $\mathcal{O}(b^2)$. Its contribution to the two moments is then $\mathcal{O}(b^4)$ so we ignore this

contribution as well. Therefore, to $\mathcal{O}(b^2)$ accuracy, our factorization is equivalent to division by $f_0(\mathbf{I})$.

We can now use equations (6), (8), (11) and (13) to evaluate the moment average in equation (A5). After explicit substitution and averaging over angles and factorisation of $f_0(\mathbf{I})$, we have

$$\begin{aligned} \langle \Delta I_j(t + \tau) \rangle &= -l_j \mathbf{I} \cdot \frac{\partial \ln f_0}{\partial \mathbf{I}} |Y_{l_2}(\pi/2, 0) r_{l_2 m}^l(\beta)|^2 \left(\frac{1}{2\pi} \right)^2 \\ &\quad \times \int_{-\infty}^{\infty} d\omega_1 \int_{-\infty}^{\infty} d\omega_2 \sum_{\mu\nu} \sum_{rs} \mathbf{W}_{l_2 m}^{l_1 \mu}(\mathbf{I}) \mathbf{W}_{l_2 m}^{*l_1 \nu}(\mathbf{I}) \\ &\quad \times \mathbf{M}_{\mu r}^{lm}(\omega_1) \mathbf{M}_{rs}^{*lm}(\omega_2) \int_t^{t+\tau} dt_1 \int_t^{t+\tau} dt_2 e^{i(\omega_1 t_1 + \omega_2 t_2)} \\ &\quad \times \langle b_r^{lm}(\omega_1) b_s^{*lm}(\omega_2) \rangle. \end{aligned} \quad (14)$$

The expression for $\langle \Delta I_j(t + \tau) \Delta I_k(t + \tau) \rangle$ is nearly the same, with $\mathbf{l} \cdot (d \ln f_0 / d \mathbf{I})$ in equation (14) replaced with $-l_k$. Although cumbersome in appearance, all quantities in this expression are straightforwardly computed. In particular, the rotation matrices, $r_{l_2 m}^l(\beta)$, have closed-form analytic expressions and the response matrices, $\mathbf{M}_{\mu r}^{lm}(\omega)$, have elements that can be evaluated by numerical quadrature. Everything but the term $\langle b_r^{lm}(\omega_1) b_s^{*lm}(\omega_2) \rangle$ depends only on the equilibrium galaxy profile and describes the mean response to an arbitrary perturbation. For a specific stochastic process, all that remains is to evaluate the Fourier transform of the density correlation function which contains all of the information about the perturbations. We will do this below for quasi-periodic and transient noise sources.

3.1 Orbiting point-mass perturbers

For a halo of black holes, the space density ρ is a sum of delta functions. Expanding the distribution for a single black hole in an action–angle series gives

$$b_j^{lm}(t) = \sum_{\mathbf{I}} Y_{l_2}(\pi/2, 0) r_{l_2 m}^l(\beta) \mathbf{W}_{l_2 m}^{*l_1 \nu}(\mathbf{I}) e^{i\mathbf{l} \cdot \mathbf{w}(t)}, \quad (15)$$

where $\mathbf{w}(t) = \mathbf{w}_o + \Omega t$ and after Fourier transforming in time, we find

$$b_j^{lm}(\omega) = \sum_{\mathbf{I}} Y_{l_2}(\pi/2, 0) r_{l_2 m}^l(\beta) \mathbf{W}_{l_2 m}^{*l_1 \nu}(\mathbf{I}) e^{i\mathbf{l} \cdot \mathbf{w}_o} 2\pi \delta(\omega - \mathbf{l} \cdot \Omega), \quad (16)$$

As in W98, we assume that individual particles are uncorrelated. The wakes from the black holes do give rise to correlations but this is of higher order in $1/N$ in the Bogolyubov, Born, Green, Kirkwood & Yvon (BBGKY) expansion than the lowest-order effect we will consider here (cf. Gilbert 1969). Assuming that individual black-hole orbits are uncorrelated, the number density of particles at $\mathbf{I}_1, \mathbf{w}_1$ at time t_1 and at $\mathbf{I}_2, \mathbf{w}_2$ at time t_2 is

$$\begin{aligned} \mathcal{P}(\mathbf{I}_1, \mathbf{w}_1, t_1; \mathbf{I}_2, \mathbf{w}_2, t_2) &= \mathcal{P}(\mathbf{I}_1, \mathbf{w}_1) \delta(\mathbf{I}_1 - \mathbf{I}_2) \\ &\quad \times \delta[\mathbf{w}_1 - \mathbf{w}_2 + \Omega(\mathbf{I}_1)(t_2 - t_1)], \end{aligned} \quad (17)$$

where $\mathcal{P}(\mathbf{I}, \mathbf{w})$ is the equilibrium particle distribution with

$$N = \int d^3 I d^3 w \mathcal{P}(\mathbf{I}, \mathbf{w}). \quad (18)$$

Direct substitution demonstrates that equation (17) solves the Liouville equation with the initial condition $\mathbf{I}_2 = \mathbf{I}_1$ and $\mathbf{w}_2 = \mathbf{w}_1$ at $t = t_1$. Similarly, integrating equation (17) over all coordinates gives N . The ensemble average $\langle b_r^{lm}(t_1) b_s^{*lm}(t_2) \rangle$ is then the average of $b_r^{lm}(t_1) b_s^{*lm}(t_2)$ over the distribution given by equation (17).

Using these definitions in equation (14) yields (see Appendix C1 for details)

$$\begin{aligned}
 \langle \Delta I_j(t + \tau) \rangle = & -\frac{1}{f_0(\mathbf{I})} l_j \cdot \frac{\partial f_o}{\partial \mathbf{I}} |Y_{ll_2}(\pi/2, 0) r_{l_2 m}^l(\beta)|^2 \\
 & \times \sum_{\mu \nu} \sum_{rs} r_{l_2 m}^l(\beta) r_{l_2 m}^{*l}(\beta) \mathbf{W}_{ll_2 m}^{l_1 \mu}(\mathbf{I}) \mathbf{W}_{ll_2 m}^{*l_1 \nu}(\mathbf{I}) \\
 & \times \left\{ (2\pi)^3 \sum_1 \int d^3 I f_o(\mathbf{I}) |Y_{ll_2}(\pi/2, 0)|^2 \mathbf{W}_{ll_2 m}^{l_1 r}(\mathbf{I}) \right. \\
 & \times \mathbf{W}_{ll_2 m}^{*l_1 s}(\mathbf{I}) \mathbf{M}_{\mu\nu}^{lm}[\mathbf{I} \cdot \boldsymbol{\Omega}(\mathbf{I})] \mathbf{M}_{rs}^{*lm}[\mathbf{I} \cdot \boldsymbol{\Omega}(\mathbf{I})] \\
 & \left. \times \int_t^{t+\tau} dt_1 \int_t^{t_1+\tau} dt_2 e^{i\mathbf{l} \cdot \boldsymbol{\Omega}(t_1 - t_2)} \right\}. \quad (19)
 \end{aligned}$$

The term in $\{\dots\}$ is the temporal correlation function for the response to the point-mass fluctuations and only depends on the time difference $t_1 - t_2$. We may change the double time integration from variables t_1, t_2 to $T = (t_1 + t_2)/2, \tau = t_1 - t_2$. Our diffusion calculation is in the regime $\tau \gg 1/\Omega$. The integral over τ in this limit gives a delta function $\delta[\mathbf{l} \cdot \boldsymbol{\Omega}(\mathbf{I})]$ and the integral over T gives τ . This approach to a delta function has a physical interpretation. Consider placing a black hole in some halo orbit. Initially, a large volume of phase space in the vicinity of the perturber's orbit is disturbed. In time, the phase of the disturbance for orbits more distant from the actions of the perturber drifts out of phase and the initial perturbation mixes away. Conversely, perturbations for nearby orbits will continue to grow in amplitude. The trade-off between phase-space volume and amplitude leads to an overall linear change in action with time. The delta function reflects that only orbits with nearly commensurate frequencies will give rise to secular changes. As a smaller and smaller phase-space volume will have larger and larger amplitude, this approximation will break down for a fixed potential. However, the equilibrium gravitation potential evolves as a result of the perturbation and this prevents breakdown and allows the secular evolution to continue.

We may also gain insight from the morphology of orbits responsible for the secular change. The disturbance must present an asymmetric force distribution on average in order to cause a secular change in the actions. If the orbital frequencies are not commensurate, the long-term average of the perturbing force will be axisymmetric. In this particular case, the ratio of radial to azimuthal frequencies ranges from 1 in the point-mass limit to 2 near the centre of a homogeneous core. Most haloes have ratios $1 < \Omega_1/\Omega_2 < 2$ between and excluding these limits. The commensurability or *closed orbit* condition is $l_1 \Omega_1 - l_2 \Omega_2 = 0$ where $l_1 \in (-\infty, \infty)$ and $l_2 \in [-l, -l + 2, \dots, l - 2, l]$. For $l = 1$, commensurability requires $1/l_1 = \Omega_1/\Omega_2$ and obviously there is no solution between the limiting cases. Similarly, for $l = 2$, we need $2/l_1 = \Omega_1/\Omega_2$ and this also admits no solution. Therefore for most systems, no commensurabilities are available for harmonic orders $l = 1, 2$. A closed $l = 1$ orbit will look like a Keplerian orbit, executing one radial oscillation for every one azimuthal and a closed $l = 2$ orbit will look like a bisymmetric oval (e.g. stationary bar orbit), executing two radial oscillations for every azimuthal oscillation. Neither exist in any significant measure in most extended stellar systems. Because of these symmetries, orbiting point-mass perturbers can not excite the most weakly damped $m = 1$ and $m = 2$ discrete modes.

We will concentrate on the evolution of haloes modelled as a

collisionless spherical distribution for application in Paper II and adopt the traditional E, J, J_z or $E, \kappa \equiv J/J_{\max}(E), \cos \beta \equiv J_z/J$ phase-space variables for computational purposes. In addition, we will not consider any processes with a preferred axis so we can average the evolution equation (A6) over β with no loss of information. If we further restrict ourselves to an isotropic distribution, we can average over κ to yield a 1+1 dimensional Fokker–Planck equation in time and energy, E ; this is described in Appendix B. The Fokker–Planck equation (equation A6) is non-linear, just as for globular cluster evolution, because the diffusion coefficients depend on the distribution function. However, it is straightforward to solve this numerically by iteration.

This application is a generalization of two-body relaxation in a stellar system, such as a globular cluster or galactic nucleus. There are two significant differences with the standard approach. First, the development here and in W98 describes the large-scale low-order structure rather than the small-scale structure included in the homogeneous local approximation of Rosenbluth, MacDonald & Judd (1957; see Binney & Tremaine 1987 for a summary). Secondly, the current approach includes self-gravity of the response. Self-gravity is important only for the features at the largest scales. In the development here, the standard results would be recovered by eliminating self-gravity and extending the series in large harmonic order l . One finds that the contribution diverges as $\ln l$ (Weinberg 1986). This is analogous to the standard logarithmic divergence parametrized by the $\ln \Lambda \propto \log(b_{\max}/b_{\min})$ term. Here, the large impact parameter cut-off is intrinsic to the equilibrium model. The small impact parameter cut-off is now a cut-off in the maximum of l included in equation (11). Computational expense makes a direct numerical check of this limit prohibitive.

Fig. 1(a) shows the Fokker–Planck coefficients and flux for a $W_0 = 3$ King model halo with total mass $M = 1$ and total potential energy $W = -1/4$. In these units the core radius is $r_c = 1.33$ and the circular orbit with r_c has energy $E = -0.44$. The three panels show the run of coefficients in energy D_E and D_{EE} and the particle flux defined as the argument of the partial differential in E in equation (B3). The coefficients scale inversely as the number of particles N and are shown here with $N = 1$. This will be applied to study a halo consisting of $10^6 M_\odot$ black holes in Paper II. The dominant resonance is 2:3 (three radial for every two azimuthal oscillations). These orbits only occur at energies $E = -0.32$ or larger, in other words in the outer halo. The next contributing resonance is 3:4 at energies about $E = -0.22$. The advection and diffusion is predominantly in the outer halo, near the resonant perturbers. The inner halo does respond to these zero-frequency perturbations but more weakly (see Murali & Tremaine 1998 for additional discussion of the zero-frequency response). The outward decreasing flux in the inner halo causes an advection of particles inward for $E \lesssim -0.47$ and the outward increasing flux beyond $E \approx -0.47$ causes an advection of particles outward.

3.2 Transient processes

Here, we describe the results of bombarding the galaxy isotropically with bits of mass on linear trajectories. This is an idealization of interactions during the epoch of galaxy evolution. The perturbations are *shots*, not quasi-periodic, and therefore this process has a continuous spectrum of perturbation frequencies. This case and the previously considered point-mass case represent two extremes.

The expansion coefficients are straightforward for the perturbers

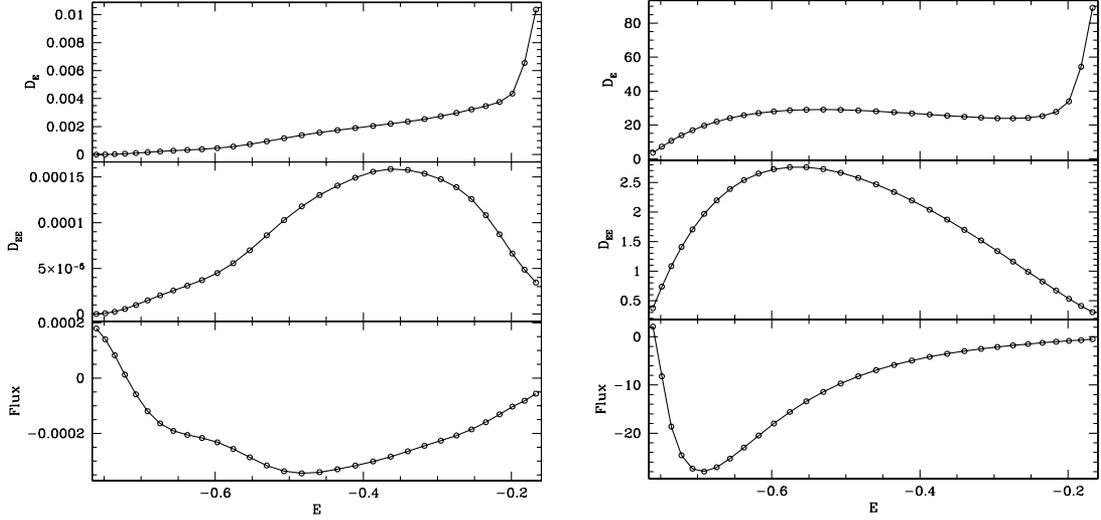


Figure 1. Isotropized Fokker–Planck coefficients and flux for (a; left-hand panel) orbiting point masses and (b; right-hand panel) perturbers on linear trajectories or ‘shrapnel’. The three subpanels show the advection and diffusion coefficients, D_E and D_{EE} , and the energy flux.

inside of the galaxy:

$$\begin{aligned}
 b_i^{lm}(t) &= \int d^3r Y_{lm}^*(\theta, \phi) u_i^{*lm}(r) \delta(\mathbf{r} - \mathbf{b} - \mathbf{v}t) \\
 &= Y_{lm}^*[\theta(t), \phi(t)] u_i^{*lm}[\mathbf{r}(t)].
 \end{aligned} \quad (20)$$

The argument of the delta function in this expression describes the linear trajectory. An arbitrary trajectory can be treated similarly. Any perturbation can always be constructed from mass elements on arbitrary trajectories and therefore it is sufficient to consider the response to a point mass. If the perturber is outside the galaxy, we use the multipole expansion with the density rather than the potential member of the biorthogonal pair to evaluate the coefficients:

$$\begin{aligned}
 b_i^{lm}(t) &= -\frac{1}{4\pi} \int d^3r Y_{lm}^*(\theta, \phi) d_i^{*lm}(r) \frac{1}{|\mathbf{r} - \mathbf{r}(t)|} \\
 &= -\frac{1}{4\pi} \int d^3r Y_{lm}^*(\theta, \phi) d_i^{*lm}(r) \sum_l \sum_{m=-l}^l \\
 &\quad \times \frac{4\pi}{2l+1} \frac{r_{<}^l}{r_{>}^{l+1}} Y_{lm}^*[\theta(t), \phi(t)] Y_{lm}^*(\theta, \phi) \\
 &= -\frac{1}{2l+1} Y_{lm}^*[\theta(t), \phi(t)] \int dr r^2 d_i^{*lm}(r) \frac{r^l}{r(t)^{l+1}}.
 \end{aligned} \quad (21)$$

For both cases, substituting into equation (14) and rearranging yields

$$\begin{aligned}
 \langle \Delta I_j(t + \tau) \rangle &= -\frac{1}{f_0(\mathbf{I})} I_j \cdot \frac{\partial f_0}{\partial \mathbf{I}} |Y_{l_2 m}(\pi/2, 0)|^2 \sum_{\mu\nu} \sum_{rs} \\
 &\quad \times r_{l_2 m}^l(\beta) r_{l_2 m}^{*l}(\beta) \mathbf{W}_{l_2 m}^{l_1 \mu}(\mathbf{I}) \mathbf{W}_{l_2 m}^{*l_1 \nu}(\mathbf{I}) \\
 &\quad \times 4\pi \int_{-\infty}^{\infty} d\omega \mathbf{M}_{\mu r}^{lm}(\omega) \mathbf{M}_{\nu s}^{*lm}(\omega) \langle b_r^{lm}(\omega_1) b_s^{*lm}(\omega_2) \rangle.
 \end{aligned} \quad (22)$$

See Appendix C2 for details. The expression for $\langle \Delta I_j(t + \tau) \Delta I_k(t + \tau) \rangle$ follows by analogy with the equations (14) and (C7) in Section 3.1. These moments describe the change in action for a

single encounter. The rate of change in actions follows from the convolution with the ensemble flux of all encounters.

Fig. 1(b) shows the diffusion coefficients and flux for a dwarfs passing through the halo on linear trajectories (‘shrapnel model’). Each encounter is a single shot and independent of the others. The trajectories have constant velocity chosen to be $\sqrt{2}$ times the peak halo circular velocity for the same King model described in Section 3.1. Here (and in Paper II) we assume that the flux of shrapnel is uniform and average over impact parameters. The diffusion coefficients scale with the impact rate of shrapnel and the square of the perturber mass. Both quantities are chosen to be unity in Fig. 1(b). Scalings and results for realistic scenarios are described in Paper II.

For transient perturbations, the continuous spectra will have some power near the frequencies of weakly damped modes and these dominate the response. The amplitude of the weakly damped $m = 1$ mode is an order of magnitude larger than the remainder of the response in all cases explored. The subsequent evolution of the halo mass distribution, then, is determined only the mechanics of these modes. Therefore, no matter what shape the frequency spectrum takes, the coefficients and flux in Fig. 1(b) will have similar shapes, depending only on the equilibrium model. In particular, the results are nearly unchanged if the incoming velocity is increased or decreased by a factor of 2, or the trajectory is parabolic or hyperbolic rather than linear. Similarly, an orbiting satellite sinking as a result of dynamical friction has a continuous spectrum with features of a line spectrum at higher frequencies because of the orbital motion and a broad low-frequency component because of the decay. The decay broadens the line-like part of the spectrum, of course, but primary resonances are easily distinguished. Direct computation shows that the coefficients and flux are nearly identical to the shrapnel model in Fig. 1(b) up to an overall multiplicative factor, because the coupling is dominated by the weakly-damped modes. This also suggests that the aperiodic noise sources will dominate the response and therefore the response will depend only weakly on the details of the source.

The effect of the damped mode on the halo is described graphically in Fig. 2. The pattern speed of the damped mode, Ω_p is smaller than the any orbital frequency in the halo. However, there is a commensurate combination of frequencies $2\Omega_\phi - \Omega_r = \Omega_p$.

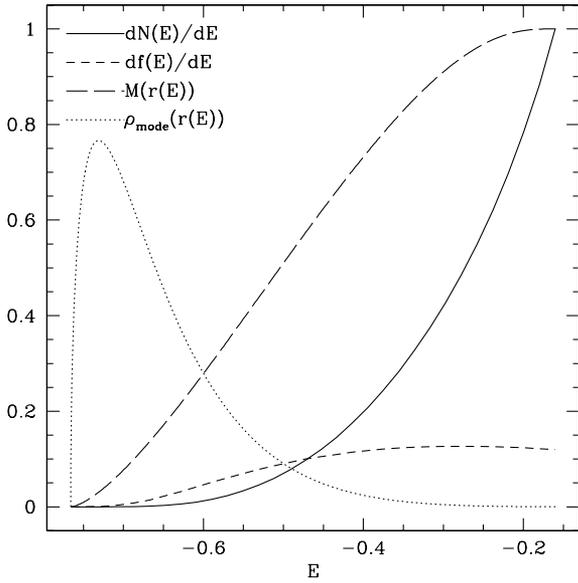


Figure 2. Features of the weakly damped $m = 1$ mode in the $W_0 = 3$ King model halo. The excited damped mode will deposit angular momentum and energy at resonances, changing the phase-space distribution. The profile of this change is shown in two ways: (i) the change in particle number in energy space, $dN(E)/dE$ and (ii) the change in phase-space distribution with energy, $df(E)/df$. The enclosed mass $M(r)$ and the profile of the mode in physical space $\rho_{\text{mode}}(r)$ is also shown for comparison. The normalizations for phase-space and density distributions are arbitrary and chosen for convenience.

Because the frequency of a weakly damped mode has a very small imaginary component, there will be halo orbits very close to resonance and these will dominate the response. The resonance exists from energies at roughly core radius outward. The near resonant orbits become increasingly radial at high energies. The modal profile may be found by solving for the null space of the matrix $\delta_{ij} - M_{ij}$ whose determinant is the dispersion relation (see Weinberg 1994). The change in phase-space density follows by using Liouville’s equation to express the second-order rate of change of the distribution function, \dot{f}_2 (analogous to equation 9), in terms of the first-order solutions for the distribution function f_1 and Hamiltonian H_1 . The change in the number of orbits at a given energy, $dN(E)/dE$, (solid curve in Fig. 2) and the change in the phase-space density $df(E)/df$ (short-dashed curve) shows that the mode deposits most of its energy and angular momentum in the outer halo. The density profile of the mode, ρ_{mode} (dotted curve) is plotted in the same figure using change of variables $E = \Phi(r)$ where $\Phi(r)$ is the halo gravitational potential. The enclosed mass for the halo model M (long-dashed curve) is shown for comparison. The density profile of the mode peaks inside of the core and is small at the half-mass radius. This is the reason for weak damping: the coupling is strongest in the outer halo where its amplitude is very small. The character of weakly damped modes in halo models without cores is similar: angular momentum and energy is deposited in the outer halo and the relative density of the mode peaks inside the scalelength.

4 SUMMARY

Recent results of large-scale structure simulations and observations of substructure in the Milky Way halo (stellar streams and disrupting dwarf galaxies) provides ample evidence that galaxies

are not in equilibrium and motivates the question: what is the long-term response of a stellar system to a fluctuating potential? This paper presents a general method for computing noise-driven evolution that incorporates the full self-gravitating response of a stochastic process. By working in the perturbation limit, linearity guarantees that the process is Markovian. The response of the stellar system to repeated stochastic events is naturally treated using the matrix method and the standard statistical techniques for stochastic differential equations.

Expansions of integral equations for stochastic processes yield an infinite number of terms. However, a general theorem (Pawula 1967, see Appendix for additional discussion) demonstrates that the truncation at quadratic order is consistent for a Markovian process and the resulting evolution equation takes the Fokker–Planck form.⁴ Evaluation of the Fokker–Planck coefficients requires the specifying the response of stellar system individual events in the stochastic process.

We explicitly develop techniques for two general situations: quasi-periodic perturbers and transient perturbers. The former case is motivated by halo of super-massive black holes (e.g. Lacey & Ostriker 1985). The latter case includes almost everything else. For example, unbound dwarf encounters (fly-by), orbiting substructure decaying as a result of dynamical friction, disrupting dwarfs and mixing tidal debris are transient perturbations. There is no practical constraint on deriving the Fokker–Planck coefficients for transient noise as long as the process can be represented as an expansion in some biorthogonal basis. For dwarfs on decaying or unbound orbits, this is particularly easy and can be done by quadrature. Alternatively, one may construct an n -body simulation using the expansion method and let the simulation produce the time-series of coefficients directly.

The theory described here is more difficult than an n -body simulation but is necessary. The nature of noise-dominated dynamics described here shows that exploration of noise-driven evolution by n -body simulation *alone* will be difficult at best for two reasons: (i) Poisson fluctuations from the distribution of bodies (and possibly numerical noise from the Poisson solver) results in background noise that may swamp the investigation even for 10^6 particles (W98); and (ii) the Poisson noise at small scales causes diffusion of an orbit’s conserved quantities and this may eliminate the otherwise important resonance structure. Therefore, the traditional test of doubling or halving the number of bodies in the simulation is not sufficient. For example, a simulation strongly dominated by diffusion is likely to remain so when the particle number is doubled. Similarity of the evolution in this case would not imply convergence to the collisionless limit.

A key finding is the importance of weakly damped modes to the evolution of haloes and therefore to galaxy evolution on the whole. The long-lived weakly damped modes depend only on the equilibrium halo profile. If they can be excited by noise, their dynamics will dominate the redistribution of angular momentum, energy and mass. The companion paper (Paper II) applies the method described here to investigate the evolution of haloes in noisy environments. Such environments may be found in the epoch just after galaxy formation or in clusters and groups at present times. We find that both unbound encounters and decaying substructure excites these modes and drive the halo profile to a universal time-asymptotic form.

There are a number of other possible implications and

⁴This is also true for two-body collisions *if* one eliminates strong encounters.

applications ripe for study. The spectrum at all harmonic orders l have discrete damped modes; the $l = m = 1$ mode is emphasized here because it is the most weakly damped and dominates the fluctuation spectrum. However, an excited, weakly damped $l = m = 2$ mode in a dark halo may interact with and trigger disc features, such as transient bar formation and spiral arms. The long-term influence of halo features on galactic structure in the 10×10^9 yr since formation may be significant. Conversely, the response of the halo caused by transient structure in ‘maximal’ discs may effect the evolution of the disc itself, either reinforcing the disc features or hastening their evolution. An interesting possibility is a excitation of disc bars by the halo or, perhaps, by the disc and halo together. In the latter case, the dark halo may help reinforce, rather than only dissipate, bars through dynamical friction (e.g. Weinberg 1985, Debattista & Sellwood 2000). For non-galactic application, consider the distribution of binary semi-major axes a in the field star populations. This observed to be proportional to a^{-1} over several orders of magnitude. Preliminary work suggests that this may be explained as the result of fluctuations from the noisy environment found in proto-stellar clusters and molecular clouds.

ACKNOWLEDGMENTS

I thank Enrico Vesperini for comments and discussion. It is also a pleasure to thank the anonymous referee for many probing questions and insightful comments. This work was support in part by NSF AST-9529328 and AST-9988146.

REFERENCES

- Arnold V. I., 1978, *Mathematical Methods of Classical Mechanics*. Springer-Verlag, New York
- Binney J., Tremaine S., 1987, *Galactic Dynamics*. Princeton University Press, Princeton, NJ
- Clutton-Brock M., 1972, *A&SS*, 16, 101
- Clutton-Brock M., 1973, *A&SS*, 23, 55
- Debattista V. P., Sellwood J. A., 2000, *ApJ*, 543, 704
- Edmonds A. R., 1960, *Angular Momentum in Quantum Mechanics*. Princeton University Press, Princeton, NJ
- Gilbert I. H., 1968, *ApJ*, 152, 1043
- Hernquist L., Ostriker J. P., 1992, *ApJ*, 386, 375
- Johnston K. V., 1998, *ApJ*, 495, 297
- Kalnajs A. J., 1976, *ApJ*, 205, 745
- Lacey C. G., Ostriker J. P., 1985, *ApJ*, 299, 633
- Morrison H. L., Mateo M., Olszewski E. W., Harding P., Dohm-Palmer R. C., Freeman K. C., Norris J. E., Morita M., 2000, *AJ*, 119, 2254
- Murali C., Tremaine S., 1998, *MNRAS*, 296, 749
- Pawula R. F., 1967, *Phys. Rev.*, 162, 186
- Risken H., 1989, *The Fokker–Planck Equation*. Springer-Verlag, Berlin
- Rosenbluth M. N., MacDonald W. M., Judd D. L., 1957, *Phys. Rep.*, 107, 1
- Spitzer L., 1987, *Dynamical Evolution of Globular Clusters*. Princeton University Press, Princeton, NJ
- Tremaine S., 1992, in Holt S. S., Verter F., eds, *AIP Conf. Proc. Vol. 278, Back to the Galaxy*. Am. Inst. Phys., New York, p. 599
- Tremaine S., Weinberg M. D., 1984, *MNRAS*, 209, 729
- Vesperini E., Weinberg M. D., 2000, *ApJ*, 534, 598
- Weinberg M. D., 1985, *MNRAS*, 213, 451
- Weinberg M. D., 1986, *ApJ*, 300, 93
- Weinberg M. D., 1993, *ApJ*, 410, 543
- Weinberg M. D., 1994, *ApJ*, 421, 481
- Weinberg M. D., 1999, *AJ*, 117, 629
- Weinberg M. D., 2001, *MNRAS*, 328, 321 (Paper II, this issue)

APPENDIX A: DERIVATION OF A FOKKER–PLANCK EQUATION

The natural coordinates for the Boltzmann equation are action–angle variables. For a collisionless equilibrium, the actions are constant and the angles advance at constant rate. The initial distribution function then depends on actions alone, $f = f(\mathbf{I})$. As outlined in Section 2, one begins the derivation with the definition of transition probability:

$$f(\mathbf{I}, t + \tau) = \int d\mathbf{I}' P(\mathbf{I}, t + \tau | \mathbf{I}', t) f(\mathbf{I}', t), \quad (\text{A1})$$

where an average over the rapidly oscillating angles is implied and P is the conditional probability that a state has action \mathbf{I} at time $t + \tau$ if it had \mathbf{I}' at time t initially. A Taylor series expansion of the integrand of equation (A1) in the action perturbation $\Delta \equiv \mathbf{I}' - \mathbf{I}$ is followed by a change of variables to and integration over Δ . In the limit $\tau \rightarrow 0$ this expansion leads directly to

$$\frac{\partial f(\mathbf{I}, t + \tau)}{\partial t} = \sum_{n=1}^{\infty} \left(-\frac{\partial}{\partial \mathbf{I}} \right)^n D^{(n)}(\mathbf{I}, t) f(\mathbf{I}, t), \quad (\text{A2})$$

where the moments $D^{[n]}$ will be defined explicitly below. Note that P is the probability of a change in \mathbf{I} caused by stochastic events despite the appearance of continuity. Therefore, the formal time-derivatives in the expansion and phase-space integral should be considered as the limit for small τ (but for τ greater than a dynamical time) of the ensemble average of stochastic events. If we let ξ be the stochastic value of \mathbf{I} at some later time, the expansion coefficients describing the stochastic variables ξ is:

$$D^{(n)}(x, t) = \frac{1}{n!} \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle [\xi(t + \tau) - \mathbf{I}]^n \rangle_{\xi(t)=\mathbf{I}}. \quad (\text{A3})$$

In N dimensions, this becomes:

$$D_{j_1 j_2 \dots j_n}^{(n)}(\mathbf{I}, t) = \frac{1}{n!} \lim_{\tau \rightarrow 0} \frac{1}{\tau} M_{j_1 j_2 \dots j_n}^{(n)}(\mathbf{I}, t, \tau), \quad (\text{A4})$$

where the moments $M^{(n)}$ are

$$M_{j_1 j_2 \dots j_n}^{(n)}(\mathbf{x}, t, \tau) = \langle (\bar{I}_{j_1} - I_{j_1})(\bar{I}_{j_2} - I_{j_2}) \dots (\bar{I}_{j_n} - I_{j_n}) \rangle. \quad (\text{A5})$$

The angle brackets denote the expectation over the conditional probability of obtaining the state variable \bar{I}_j at time $t + \tau$ given I_j at time t . The approach sketched here is described by Risken (1989, chapter 4).

Although the Kramers–Moyal expansion has an infinite number of terms in general (cf. equation 1), the Pawula Theorem (Pawula 1967) shows that consistency demands that the expansion either stops after two terms and takes the standard Fokker–Planck form or must have an infinite number of terms. The response function guarantees that the Kramers–Moyal expansion terminates after two terms in the limit of weak perturbations because higher order terms are vanishingly small in comparison. This is consistent with our intuition that repetitive, weak, stochastic excitation in a galaxy is a Markov process and the transition probability should take a Gaussian form following the Central Limit Theorem. The evolution equation is then a Fokker–Planck equation:

$$\frac{\partial f(\mathbf{I}, t)}{\partial t} = \mathbf{L}_{\text{FP}}(\mathbf{I}, t) f(\mathbf{I}, t), \quad (\text{A6})$$

where

$$\mathbf{L}_{\text{FP}}(\mathbf{I}, t) = -\frac{\partial}{\partial I_i} D_i^{(1)}(\mathbf{I}, t) + \frac{\partial^2}{\partial I_i \partial I_j} D_{ij}^{(2)}(\mathbf{I}, t). \quad (\text{A7})$$

APPENDIX B: AVERAGED FOKKER–PLANCK EQUATION

A number of authors have described coordinate transformations for the multivariate Fokker–Planck equation (Rosenbluth et al. 1957, Risken 1989). The approach is the familiar one: write the equation in terms of scalars, covariant and contravariant vectors and tensors and covariant derivatives only and apply the standard tools of differential geometry. Rosenbluth et al. use the Jacobian of the coordinate transformation as a metric and Risken uses the diffusion matrix. We will use the former approach here. Denote the Jacobian of the coordinate transformation as J . Under a change of coordinates, one can show after a fair bit of algebra that the advection and diffusion coefficients transform as

$$D'_k = \frac{\partial x'_k}{\partial x_i} D_i + \frac{\partial^2 x'_k}{\partial x_i \partial x_j} D_{ij}, \quad (\text{B1})$$

$$D'_{kl} = \frac{\partial x'_k}{\partial x_i} \frac{\partial x'_l}{\partial x_j} D_{ij}. \quad (\text{B2})$$

The phase-space distribution function transforms as $f'(\mathbf{I}) = Jf(\mathbf{I})$ (cf. Risken 1989) and in the new variables, the equation takes the standard Fokker–Planck form:

$$\frac{\partial f'(\mathbf{I}, t)}{\partial t} = \left[-\frac{\partial}{\partial I'_k} D'_k + \frac{\partial^2}{\partial I'_k \partial I'_l} D'_{kl} \right] f'(\mathbf{I}, t). \quad (\text{B3})$$

Now let $\mathbf{I}' = (E, \kappa, \cos \beta)$. Assuming that the distribution function f is time-independent and non-zero, we may integrate equation (B3) over κ and $\cos \beta$. As both of these variables have a bounded domain, the flux through their boundaries must vanish, leaving a single flux term:

$$\frac{\partial \langle f' \rangle}{\partial t} = \frac{\partial}{\partial E} \left\langle \left[-D_{Ej} f' + D_{Ej} \frac{\partial f'}{\partial x^j} \right] \right\rangle_{\text{iso}}, \quad (\text{B4})$$

where the angle brackets denote integration over κ and $\cos \beta$ and implied sum on j is over all three variables. Expanding the contravariant vectors in terms of the original advection and diffusion coefficients one finds that

$$\frac{\partial \langle f' \rangle}{\partial t} = \frac{\partial}{\partial E} \left\langle -D_{Ej} f + \frac{\partial}{\partial x^j} (D_{Ej} f) \right\rangle_{\text{iso}}. \quad (\text{B5})$$

The isotropically averaged Fokker–Planck equation is then

$$\frac{\partial \bar{f}(E)}{\partial t} = \frac{\partial}{\partial E} \left\{ -\langle D_E \rangle_{\text{iso}} \bar{f}(E) + \frac{\partial}{\partial E} [\langle D_{EE} \rangle_{\text{iso}} \bar{f}(E)] \right\}, \quad (\text{B6})$$

where $\langle D_E \rangle_{\text{iso}}$ and $\langle D_{EE} \rangle_{\text{iso}}$ are the isotropically averaged diffusion coefficients:

$$\left\{ \begin{array}{l} \langle D_E \rangle \\ \langle D_{EE} \rangle \end{array} \right\} = \frac{J_{\text{max}}^2(E)}{f(E)} \int d\kappa d(\cos \beta) \kappa \left\{ \begin{array}{l} D_E \\ D_{EE} \end{array} \right\} f(E, \kappa, \beta), \quad (\text{B7})$$

where

$$\bar{f}(E) = J_{\text{max}}^2(E) \int d\kappa d(\cos \beta) \kappa f(E, \kappa, \beta), \quad (\text{B8})$$

and the phase-space volume is

$$P(E) \equiv J_{\text{max}}^2(E) \int d\kappa d(\cos \beta) \kappa. \quad (\text{B9})$$

Note that the standard notation in the globular cluster literature is $f(E) = \bar{f}(E)/P(E)$.

APPENDIX C: DERIVATION OF ACTION MOMENTS AND COEFFICIENTS

C1 Orbiting point-mass perturbers

We apply the two-particle distribution function from Section 3.1, equation (17) to explicitly evaluate the moments in actions. The ensemble average in equation (14) implies an average of possible distributions of point masses consistent with some underlying distribution. Therefore

$$\begin{aligned} \langle b_r^{lm}(\omega_1) b_s^{*l'm'}(\omega_2) \rangle &= \left\langle \sum_{\mathbf{I}} \delta_{mm'} Y_{l_2}(\pi/2, 0) Y_{l'_2}^*(\pi/2, 0) r_{l_2m}^l(\beta) \right. \\ &\quad \times r_{l'_2m'}^{*l'}(\beta) \times \mathbf{W}_{l_2m}^{l_1\nu}(\mathbf{I}) \mathbf{W}_{l'_2m'}^{*l_1\mu}(\mathbf{I}) (2\pi)^2 \\ &\quad \left. \times \delta(\omega_1 + \omega_2) \delta(\omega_1 - \mathbf{I} \cdot \boldsymbol{\Omega}) \right\rangle, \quad (\text{C1}) \end{aligned}$$

and

$$\begin{aligned} \langle b_r^{lm}(t_1) b_s^{*l'm}(t_2) \rangle &= \left(\frac{1}{2\pi} \right)^2 \int_{-\infty}^{\infty} d\omega_1 \int_{-\infty}^{\infty} d\omega_2 e^{i(\omega_1 t_1 - \omega_2 t_2)} \\ &\quad \times \langle b_r^{lm}(\omega_1) b_s^{*l'm}(\omega_2) \rangle. \quad (\text{C2}) \end{aligned}$$

Integration over angles identifies \mathbf{l} with \mathbf{l}' and therefore $m = l_3 = l'_3 = m'$. Now using equations (15), (16), (17) and the orthogonality of rotation matrices,

$$\int d\beta \sin(\beta) r_{\mu\nu}^l(\beta) r_{\mu\nu}^{l'}(\beta) = \frac{2}{2l+1} \delta_{ll'}, \quad (\text{C3})$$

(Edmonds 1960) to evaluate the average, we find that the final part of expression (14) becomes

$$\begin{aligned} &\left(\frac{1}{2\pi} \right)^2 \left\langle \int_{-\infty}^{\infty} d\omega_1 \int_{-\infty}^{\infty} d\omega_2 e^{i(\omega_1 t_1 + \omega_2 t_2)} b_r^{lm}(\omega_1) b_s^{*lm}(\omega_2) \right\rangle \\ &= (2\pi)^3 \sum_{\mathbf{I}} \int d^3 I f_o(\mathbf{I}) e^{i\mathbf{l} \cdot \boldsymbol{\Omega}(\mathbf{I})(t_1 - t_2)} |Y_{l_2}(\pi/2, 0)|^2 \\ &\quad \times |r_{l_2m}^l(\beta)|^2 \mathbf{W}_{l_2m}^{l_1r}(\mathbf{I}) \mathbf{W}_{l_2m}^{*l_1s}(\mathbf{I}), \quad (\text{C4}) \end{aligned}$$

where the integration over ω_1 and ω_2 has become a sum over discrete frequencies denoted by \mathbf{I} . We now substitute this development into equation (14) and find

$$\begin{aligned} \langle \Delta I_j(t + \tau) \rangle &= -\frac{1}{f_o(\mathbf{I})} l_j \mathbf{l} \cdot \frac{\partial f_o}{\partial \mathbf{I}} |Y_{l_2}(\pi/2, 0)|^2 \sum_{\mu\nu} \sum_{rs} \\ &\quad \times r_{l_2m}^l(\beta) r_{l_2m}^{*l}(\beta) \mathbf{W}_{l_2m}^{l_1\mu}(\mathbf{I}) \mathbf{W}_{l_2m}^{*l_1\nu}(\mathbf{I}) \\ &\quad \times \left\{ (2\pi)^3 \sum_{\mathbf{I}} \int d^3 I f_o(\mathbf{I}) |Y_{l_2}(\pi/2, 0)|^2 r_{l_2m}^l(\beta) r_{l_2m}^l(\beta) \right. \\ &\quad \times \mathbf{W}_{l_2m}^{l_1r}(\mathbf{I}) \mathbf{W}_{l_2m}^{*l_1s}(\mathbf{I}) \mathbf{M}_{\mu r}^{lm}[\mathbf{l} \cdot \boldsymbol{\Omega}(\mathbf{I})] \mathbf{M}_{s\nu}^{*lm}[\mathbf{l} \cdot \boldsymbol{\Omega}(\mathbf{I})] \\ &\quad \left. \times \int_t^{t+\tau} dt_1 \int_t^{t+\tau} dt_2 e^{i\mathbf{l} \cdot \boldsymbol{\Omega}(\mathbf{I})(t_1 - t_2)} \right\}. \quad (\text{C5}) \end{aligned}$$

Returning to equation (A3), we can now read off our diffusion

coefficients in action variables:

$$\begin{aligned}
D_j^{(1)}(\mathbf{I}, t) &= \lim_{\tau \rightarrow 0} \frac{\langle \Delta I_j(t + \tau) \rangle}{\tau} = -\frac{1}{f_0(\mathbf{I})} l_j \mathbf{I} \cdot \frac{\partial f_o}{\partial \mathbf{I}} |Y_{l_2}(\pi/2, 0)|^2 \\
&\times \sum_{\mu\nu} \sum_{rs} r_{l_2m}^l(\beta) r_{l_2m}^{*l}(\beta) \mathbf{W}_{l_2m}^{l_1\mu}(\mathbf{I}) \mathbf{W}_{l_2m}^{*l_1\nu}(\mathbf{I}) \\
&\times \left\{ (2\pi)^3 \sum_{\mathbf{I}} \int d^3 I f_o(\mathbf{I}) |Y_{l_2}(\pi/2, 0)|^2 r_{l_2m}^l(\beta) r_{l_2m}^l(\beta) \right. \\
&\times \mathbf{W}_{l_2m}^{l_1r}(\mathbf{I}) \mathbf{W}_{l_2m}^{*l_1s}(\mathbf{I}) \mathbf{M}_{\mu\nu}^{lm}[\mathbf{I} \cdot \boldsymbol{\Omega}(\mathbf{I})] \mathbf{M}_{rs}^{*lm}[\mathbf{I} \cdot \boldsymbol{\Omega}(\mathbf{I})] 2\pi \delta \\
&\left. \times [\mathbf{I} \cdot \boldsymbol{\Omega}(\mathbf{I})] \right\}. \tag{C6}
\end{aligned}$$

$$\begin{aligned}
D_{jk}^{(2)}(\mathbf{I}, t) &= \lim_{\tau \rightarrow 0} \frac{\langle \Delta I_j(t + \tau) \Delta I_k(t + \tau) \rangle}{2\tau} \\
&= \frac{l_j l_k}{2} |Y_{l_2}(\pi/2, 0)|^2 \sum_{\mu\nu} \sum_{rs} r_{l_2m}^l(\beta) r_{l_2m}^{*l}(\beta) \mathbf{W}_{l_2m}^{l_1\mu}(\mathbf{I}) \mathbf{W}_{l_2m}^{*l_1\nu}(\mathbf{I}) \\
&\times \left\{ (2\pi)^3 \sum_{\mathbf{I}} \int d^3 I f_o(\mathbf{I}) |Y_{l_2}(\pi/2, 0)|^2 r_{l_2m}^l(\beta) r_{l_2m}^l(\beta) \right. \\
&\times \mathbf{W}_{l_2m}^{l_1r}(\mathbf{I}) \mathbf{W}_{l_2m}^{*l_1s}(\mathbf{I}) \mathbf{M}_{\mu\nu}^{lm}[\mathbf{I} \cdot \boldsymbol{\Omega}(\mathbf{I})] \mathbf{M}_{rs}^{*lm}[\mathbf{I} \cdot \boldsymbol{\Omega}(\mathbf{I})] 2\pi \delta \\
&\left. \times [\mathbf{I} \cdot \boldsymbol{\Omega}(\mathbf{I})] \right\}. \tag{C7}
\end{aligned}$$

The limit $\tau \rightarrow 0$ must be taken in the sense that τ is small compared to the evolutionary time-scale as a result of the fluctuations but large compared to the dynamical time. The time-dependence of the diffusion coefficients reminds us that the underlying equilibrium distribution $f_o(\mathbf{I})$ changes on an evolutionary time-scale but, for the purposes of computation, is held fixed on a dynamical time-scale. The integrals may be simplified by noting that $d^3 I = dE dJ d(\cos \beta)/\Omega_1(E, J)$. We can do the integral in β using the orthogonality of the rotation matrices as previously described. For a given equilibrium distribution function $f_o(\mathbf{I})$, the term in curly brackets in equations (C6) and (C7) can be computed once and for all as they are independent of the local value of the actions.

C2 Transient processes

The development for transient ‘shot’ noise is similar to that above and this section will emphasize the differences. The Fourier transform needed for equation (14) is carried out most easily assuming the perturber is in the equatorial plane of the coordinate system. We denote the Fourier transform of $b_j^{lm}(t)$ as $\hat{b}_j^{lm}(\omega)$. In practice, the transform is easily computed numerically by fast Fourier transform (FFT). Equations (20) and (21) in Section 3.2 describe the time-dependence of coefficients for any trajectory $r(t), \theta(t), \phi(t)$. A cloud of points and more generally any phase-space distortion yielding $b_i^{lm}(t)$ or coefficients at discrete times from an n -body simulation are possible input to the FFT. Now, we can evaluate the final line in equation (14) by changing coordinates from t_1, t_2 to $T = (t_1 + t_2)/2, \tau = t_1 - t_2$. We have

$$e^{i(\omega_1 t_1 + \omega_2 t_2)} = e^{i(\omega_1 [T + \tau/2] + \omega_2 [T - \tau/2])} = e^{i(\omega_1 + \omega_2)T} e^{i(\omega_1 - \omega_2)\tau/2}.$$

Using this in the last line of equation (14) we find

$$\begin{aligned}
&\int_t^{t+\tau} dt_1 \int_t^{t+\tau} dt_2 e^{i(\omega_1 t_1 + \omega_2 t_2)} \langle b_r^{lm}(\omega_1) b_s^{*lm}(\omega_2) \rangle \\
&= \int_t^{t+\tau/2} dT 4\pi \delta(\omega_1 - \omega_2) e^{i2\omega T} \langle b_r^{lm}(\omega_1) b_s^{*lm}(\omega_2) \rangle \\
&= 4\pi \delta(\omega_1 - \omega_2) e^{i\omega \tau} \frac{\sin \omega \tau}{\omega} \langle \bar{b}_r^{lm}(\omega_1) \bar{b}_s^{*lm}(\omega_2) \rangle. \tag{C8}
\end{aligned}$$

where $\omega \equiv \omega_1 = \omega_2$. In deriving the second equality, we note that the bombardment must occur between t and $t + \tau$ and use the shift properties of the Fourier transform to write $\langle b_r^{lm}(\omega_1) b_s^{*lm}(\omega_2) \rangle = e^{-2i\omega t} \langle \bar{b}_r^{lm}(\omega_1) \bar{b}_s^{*lm}(\omega_2) \rangle$ where the transform $\bar{b}(\omega)$ denotes the transform of an event centred about the temporal origin. In the limit $\tau \rightarrow 0$, this expression becomes $4\pi \delta(\omega_1 - \omega_2) \tau$. Substituting this back into equation (14), we can perform one of the ω integrals straight away. After rearranging we have

$$\begin{aligned}
\langle \Delta I_j(t + \tau) \rangle &= -\frac{1}{f_0(\mathbf{I})} l_j \mathbf{I} \cdot \frac{\partial f_o}{\partial \mathbf{I}} |Y_{l_2}(\pi/2, 0)|^2 \sum_{\mu\nu} \sum_{rs} r_{l_2m}^l(\beta) r_{l_2m}^{*l}(\beta) \\
&\times \mathbf{W}_{l_2m}^{l_1\mu}(\mathbf{I}) \mathbf{W}_{l_2m}^{*l_1\nu}(\mathbf{I}) 4\pi \int_{-\infty}^{\infty} d\omega \mathbf{M}_{\mu\nu}^{lm}(\omega) \mathbf{M}_{rs}^{*lm}(\omega) \\
&\times \langle b_r^{lm}(\omega_1) b_s^{*lm}(\omega_2) \rangle. \tag{C9}
\end{aligned}$$

The expression for $\langle \Delta I_j(t + \tau) \Delta I_k(t + \tau) \rangle$ follows by analogy with the equations (C6) and (C7).

Symmetry suggests the choice of perturbing orbits on the equatorial plane should not effect the final results. This can be demonstrated explicitly using the rotational properties of the spherical harmonics. Let $\mathbf{R}_{mm'}^l(\alpha, \beta, \gamma)$ be the rotation matrix with the Euler angles α, β, γ . Then

$$Y_{lm}(\theta, \phi) = \sum_{m'=-l}^l Y_{lm'}(\theta', \phi') \mathbf{R}_{mm'}^l(\alpha, \beta, \gamma),$$

where the primed coordinates refer to the rotated coordinate system, and therefore we have

$$b_j^{lm}(t) = \sum_{m'=-l}^l b_j^{lm'}(t) \mathbf{R}_{mm'}^l(\alpha, \beta, \gamma), \tag{C10}$$

For an isotropic spherical system the response operator $\mathbf{M}_{\mu\nu}^{lm}(\omega)$ is independent of m . We can exploit this and the rotational properties of $b_r^{lm}(\omega)$ to simplify the computation of $\langle b_r^{lm}(\omega_1) b_s^{*lm}(\omega_2) \rangle$. First, express $b_r^{lm}(\omega)$ in any convenient coordinate system and then use the rotation matrices to rotate this to any orientation:

$$\begin{aligned}
&\langle b_r^{lm}(\omega_1) b_s^{*lm}(\omega_2) \rangle \\
&= \left\langle \sum_{m'=-l}^l \mathbf{R}_{mm'}^l(\alpha, \beta, \gamma) b_r^{lm'}(\omega_1) \sum_{m''=-l}^l \mathbf{R}_{mm''}^l(\alpha, \beta, \gamma) b_s^{*lm''}(\omega_2) \right\rangle. \tag{C11}
\end{aligned}$$

Summing over all values m for a given l we have

$$\left\langle \sum_{m=-l}^l b_r^{lm}(\omega_1) b_s^{*lm}(\omega_2) \right\rangle = \left\langle \sum_{m'=-l}^l b_r^{lm'}(\omega_1) b_s^{*lm'}(\omega_2) \right\rangle, \tag{C12}$$

having used the orthogonality of rotation matrices. We assume that the ensemble average includes random events from all directions and therefore only the same event will be correlated in the computation of $\langle b_r^{lm}(\omega_1) b_s^{*lm}(\omega_2) \rangle$.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.