



**HAL**  
open science

## Revisiting relaxation in globular clusters

Chris Hamilton, Jean-Baptiste Fouvry, James Binney, Christophe Pichon

► **To cite this version:**

Chris Hamilton, Jean-Baptiste Fouvry, James Binney, Christophe Pichon. Revisiting relaxation in globular clusters. *Monthly Notices of the Royal Astronomical Society*, 2018, 481 (2), pp.2041-2061. 10.1093/mnras/sty2295 . hal-02055691

**HAL Id: hal-02055691**

**<https://hal.sorbonne-universite.fr/hal-02055691>**

Submitted on 4 Mar 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Revisiting relaxation in globular clusters

Chris Hamilton<sup>1,2\*</sup>, Jean-Baptiste Fouvry<sup>3†</sup>, James Binney<sup>1</sup>, Christophe Pichon<sup>4,5,6</sup>

<sup>1</sup>*Rudolf Peierls Centre for Theoretical Physics, 1 Keble Road, Oxford, OX1 3NP, UK*

<sup>2</sup>*Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK*

<sup>3</sup>*Institute for Advanced Study, Einstein Drive, Princeton NJ 08540, USA*

<sup>4</sup>*Institut d’Astrophysique de Paris and UPMC, CNRS (UMR 7095), 98 bis Boulevard Arago, 75014, Paris, France*

<sup>5</sup>*Korea Institute of Advanced studies (KIAS), 85 Hoegiro, Dongdaemun-gu, Seoul, 02455, Republic of Korea*

<sup>6</sup>*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK*

3 September 2018

## ABSTRACT

The classical theory of cluster relaxation is unsatisfactory because it involves the Coulomb logarithm. The Balescu–Lenard (BL) equation provides a rigorous alternative that has no ill-defined parameter. Moreover, the BL equation, unlike classical theory, includes the cluster’s self-gravity. A heuristic argument is given that indicates that relaxation does not occur predominantly through two-particle scattering and is enhanced by self-gravity. The BL equation is adapted to a spherical system and used to estimate the flux through the action space of isochrone clusters with different velocity anisotropies. A range of fairly different secular behaviours is found depending on the fraction of radial orbits. Classical theory is also used to compute the corresponding classical fluxes. The BL and classical fluxes are very different because (a) the classical theory materially under-estimates the impact of large-scale collectively amplified fluctuations and (b) only the leading terms in an infinite sum for the BL flux are computed. A complete theory of cluster relaxation likely requires that the sum in the BL equation be decomposed into a sum over a finite number of small wavenumbers complemented by an integral over large wavenumbers analogous to classical theory.

**Key words:** Galaxy: kinematics and dynamics – galaxies: kinematics and dynamics – galaxies: star clusters: general – methods: numerical

## 1 INTRODUCTION

As systems of almost coeval stars, globular clusters have been vital to the development and testing of the theory of stellar structure and evolution. As wonderfully clean dynamical systems that have enough stars to allow precise statistical characterisation but few enough stars to evolve significantly by stellar-dynamical relaxation within a Hubble time, their study has been important for the development of stellar dynamics. Yet notwithstanding intensive study for more than a century, aspects of these system remain mysterious. In particular we don’t understand how or where they formed. Moreover, the efficiency with which gas must have been converted into stars when they formed becomes more surprising as our knowledge of the general star-formation process grows.

In this paper we argue that the dynamical evolution of globular clusters is also less well understood than it should be with a critique of the standard theory of two-body re-

laxation. Then we re-formulate, in the context of a model globular cluster, the theory of the long-time (secular) evolution of a stellar system using a sophisticated approach that has recently explained the secular evolution of collisionless stellar discs.

The conceptual foundation of the theory of secular evolution is that the system evolves through a series of steady-state solutions of the collisionless Boltzmann equation. By Jeans’ theorem, any such solution can be described by a distribution function (DF) that depends on position and velocity ( $\mathbf{x}, \mathbf{v}$ ) only through constants of stellar motion. Since we focus on spherical systems, we know that the system’s mean-field Hamiltonian  $H = \frac{1}{2}v^2 + \Phi(\mathbf{x})$  admits three action integrals  $J_i$ , and we may assume that at each time the DF takes the form  $f(\mathbf{J})$ . On a timescale much longer than the system’s crossing time, small fluctuations of the potential  $\Phi$  around the slowly evolving mean spherical form for which the  $J_i$  are the action integrals cause the actions of stars to change, and thus the functional form of  $f$  to evolve.

The most obvious difference between the real potential  $\Phi_N(\mathbf{x})$  and the spherical approximation for which the  $J_i$  are computed, is the local singularities  $\Phi_N(\mathbf{x}) \sim -Gm/|\mathbf{x} - \mathbf{x}_i|$

\* E-mail: ch783@cam.ac.uk

† Hubble Fellow

near the location  $\mathbf{x}_i$  of each star. The classical theory of relaxation (e.g. Binney & Tremaine 2008, chapter 7) is based on the idea that as a star moves through the cluster it is repeatedly scattered by the singularities that lie along its path. Following Chandrasekhar (1943) one calculates as a function of impact parameter  $b$  and relative velocity the velocity change  $\delta\mathbf{v}$  that an encounter causes, and then adds these changes by integrating over impact parameters and relative velocities to obtain diffusion coefficients for the Fokker-Planck equation that describes the diffusion of stars through phase space (Chandrasekhar 1949):

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial \mathbf{v}} \cdot \left( \mathbf{\Delta}_1 f - \frac{1}{2} \frac{\partial}{\partial \mathbf{v}} \cdot \left[ \mathbf{\Delta}_2 f \right] \right). \quad (1)$$

Here the first-order diffusion coefficient  $\mathbf{\Delta}_1$  is the average  $\langle \delta\mathbf{v} \rangle$  of the  $\delta\mathbf{v}$ , while the second-order diffusion coefficient  $\mathbf{\Delta}_2$  is obtained by averaging the  $\delta\mathbf{v}$  in quadrature.

The calculation is made tractable by neglecting the acceleration caused by the mean-field potential  $\Phi$ . That is, one assumes that in the absence of a nearby singularity, the star's trajectory would be a straight line, and in its presence it is a Keplerian hyperbola. This approximation is only valid when the impact parameter  $b$  is much less than the distance to the cluster's centre. The inconvenient truth is that the integral for  $\langle \delta\mathbf{v} \rangle$  over impact parameters diverges at large  $b$ , signalling that encounters with distant stars contribute significantly to  $\langle \delta\mathbf{v} \rangle$ .

Fortunately, the divergence is only logarithmic and the text-book work-around for this conceptual difficulty is to cut the integral off at a distance  $b_{\max}$  comparable to the system's half-mass radius, and report  $b_{\max}$  through the value of the ‘‘Coulomb logarithm’’, which is the natural logarithm of the ratio of  $b_{\max}$  to the impact parameter that gives rise to scattering by 90 degrees at the mean relative velocity. The venerable age of this work-around should not blind us to its unsatisfactory nature: the divergence of the integral for  $\langle \delta\mathbf{v} \rangle$  is no mathematical nicety but reflects a real physical issue – the significant contribution of fluctuations in the potential generated by distant rather than nearby stars (Weinberg 1993). Clearly these stars do not cause Rutherford-like scattering, and they do not act in isolation. In so far as long-range encounters contribute significantly to the secular evolution of the system, the classical formulation of the problem is flawed.

The purpose of this paper is first to argue that the theory of the secular evolution of globular clusters should be reformulated, second to explain what the reformulated theory looks like, and finally to present a worked example of the reformulated theory in action. In Section 2 we give an order-of-magnitude calculation that implies that relaxation is driven not by individual encounters but by fluctuations in the number of particles in different sub-regions of the system. This conclusion suggests that the self-gravity of the system, which is completely neglected in the traditional approach, can play a significant role. In Section 3 we give the governing equations and explain their physical content. Section 4 deals with the application of these equations to spherical systems. Sections 3 and 4 are rather technical in nature and readers may like to skip straight from Section 2 to Section 5, which examines the diffusive fluxes in a family of clusters that have the spatial structure of Henon's isochrone, but have differing degrees of velocity anisotropy. We show

that the flux computed from the BL equation is larger when self-gravity is included than it is otherwise. We explore the extent to which the BL flux grows as the velocity anisotropy increases towards the value at which the cluster is prone to the radial-orbit instability. In Section 6, we compare the BL flux to the classical theory. We discuss these results in Section 7, and conclude in Section 8.

## 2 WHAT DRIVES SECULAR EVOLUTION

A simple back-of-the-envelope calculation confirms the importance of large-scale fluctuations that fall outside the scope of a Rutherford scattering calculation. We consider a system of mass  $M$  with  $N$  stars and characteristic scale  $R$ , in which the characteristic internal speed is  $\sigma = \sqrt{GM/R}$ . A subregion of size  $r = xR$  contains mass  $M_r \simeq x^3 M$  and  $n \simeq x^3 N$  stars, so on account of Poisson noise  $M_r$  fluctuates by  $\delta M_r = M_r/\sqrt{n} = x^{3/2} M/\sqrt{N}$  during times  $\delta t = r/\sigma$ . Consider a point that is distance  $yR$  from our subregion. At this point a single fluctuation in the subregion's gravitational attraction will change the velocity of a test star by

$$\delta v = \frac{G\delta M_r}{(yR)^2} \delta t = \frac{GMx^{3/2}}{(yR)^2 \sqrt{N}} \frac{xR}{\sigma} = \frac{\sigma x^{5/2}}{y^2 \sqrt{N}}. \quad (2)$$

This formula states that for given  $y$ , large volumes  $x \simeq 1$  perturb  $v$  very much more strongly than small volumes  $x \ll 1$ . Against this trend we must bear in mind that (a)  $y \geq x$ , (b) as  $x$  decreases the number of subregions perturbing increases as  $x^{-3}$ , and (c) the time within which the contribution (2) comes about decreases with  $x$ , so in a given time each small subregion makes many more contributions to  $v$  than does a large subregion.

For the moment we assume that the contributions to  $v$  from different subregions are statistically independent, so it's appropriate to add the  $\delta v$  in quadrature. There are  $\sim 4\pi(y/x)^2$  subregions of scale  $x$  that are distance  $yR$  from our point, and in a global crossing time  $t_{\text{cross}} = R/\sigma$  each such subregion contributes  $x^{-1}$  times. So in a crossing time all these subregions change  $v^2$  by

$$(\Delta v)^2 = 4\pi \frac{y^2}{x^3} (\delta v)^2 = 4\pi \frac{\sigma^2 x^2}{y^2 N}. \quad (3)$$

Now we have to sum over  $y = x, 2x, 3x, \dots, 1$ . We convert the sum to an integral using  $dy = x$  and have

$$\sum \frac{1}{y^2} \simeq \frac{1}{x} \int_x^1 \frac{dy}{y^2} = \frac{1}{x} \left( \frac{1}{x} - 1 \right) \simeq \frac{1}{x^2}. \quad (4)$$

Hence in a crossing time the subregions of scale  $x$  change  $v^2$  by

$$(\Delta v)^2 \simeq 4\pi\sigma^2/N. \quad (5)$$

Remarkably, this is independent of  $x$ , so by this reckoning regions of each scale  $xR$  contribute equally to changing  $v^2$ .

At this point it is important to consider correlations between the fluctuations in subregions. On first examining a cluster, the fluctuations between the masses in different subregions characterised by a given value of  $x$  will be negligible on all but the largest scales: the basis of Poisson statistics is the selection of items from an essentially infinite pool of objects and this model will begin to fail only when the number

of items selected,  $x^3N$ , becomes comparable to  $N$ . When  $x \sim 1$  and the assumption of statistical independence fails, it does so because fluctuations in one half of the cluster are *anti*-correlated with fluctuations in the other half. Consequently, as  $x \rightarrow 1$  we should be adding values of  $\delta v$  *linearly* rather than quadratically because the downward fluctuation in one hemisphere pushes a test star in the same direction as the upward correlation in the other hemisphere. Hence, by neglecting correlations between subregions we are merely under-estimating the importance of large scales relative to small scales.

We must also consider temporal fluctuations: since an over-density in one subregion must run out through its neighbours, the overdensity in adjacent cells of a given scale are temporally correlated. Our calculation allows for this correlation by taking the persistence time of an overdensity to be  $\delta t = xR/\sigma$ , but to incorporate this insight into our final value for  $(\Delta v)^2$  we must add the contributions of all relevant scales  $x$ : the contribution from the smallest values of  $x$  are based on a picture in which after a time  $\delta t = xR/\sigma$  mass is reassigned to cells from some fictional outside reservoir rather than by passing mass between adjacent cells. Only by including independently the contributions to  $(\Delta v)^2$  from larger values of  $x$  do we arrive at a value that reflects temporal correlations between over-densities. Since the number of scales between the smallest one,  $x_{\min}$  and the whole cluster is  $-\ln x_{\min}$ , we multiply equation (5) by  $-\ln x_{\min}$ . The smallest subregion it's sensible to consider cannot be smaller than a decent multiple of the inter-particle distance  $\sim R/N^{1/3}$ . When it is equal to this distance, we have

$$(\Delta v)_{t_{\text{cross}}}^2 \simeq \frac{4\pi\sigma^2 \ln N}{3N}. \quad (6)$$

The relaxation time is the time required for fluctuations to change any velocity by order of itself, thus for  $(\Delta v)^2$  to accumulate to  $\sigma^2$ . From equation (6) it follows that

$$t_{\text{relax}} \simeq \frac{N}{4 \ln N} t_{\text{cross}}. \quad (7)$$

This result is essentially identical with equation (7.108) for the half-mass relaxation time in Binney & Tremaine (2008).

In an ideal gas the number of molecules in a given volume is given by Poisson statistics as was assumed above, and the time evolution of fluctuations can be computed by considering them to arise from thermally excited sound waves. The self-gravity of a stellar system makes the system more compressible on large scales than on small scales, where self-gravity is unimportant and an ideal gas provides a valid model. Gravity reduces the energy density associated with a sound wave of given amplitude – as the wave's wavelength rises towards the Jeans length, the energy density falls to zero. Hence, when thermally excited, waves with longer wavelengths will have larger amplitudes than Poisson statistics would imply. It follows that in a real system fluctuations near the size of the system (the Jeans length) dominate, contrary to our finding above of equal contributions from all scales (Weinberg 1993).

In summary, we can recover the traditional formula for the “two-body” relaxation time by computing the effects of Poisson fluctuations in the mass density within the system. Given that self-gravity makes a gas more compressible than a classical ideal gas, on the largest spatial scales, we expect

the actual density fluctuations in a cluster to have larger amplitudes than simple shot noise predicts. It follows that the classical relaxation time is too long by a factor that depends on the amount by which self-gravity enhances fluctuations.

Weinberg (1993) showed that popular models of clusters have a weakly damped mode in which the core oscillates along some line in antiphase to the envelope that surrounds it. The physics of this dipole mode, in which the core exchanges momentum with the halo, is heavily dependent on the fact that the cluster generates its own gravitational field rather than being confined with an external field from which it could draw momentum at will. In Section 5 we will show that in the self-gravitating case a cluster's relaxation is significantly accelerated as a result of this mode being stimulated by Poisson noise.

Clusters with strongly radially biased velocity dispersions are subject to the radial-orbit instability (Fridman & Poliachenko 1984; Palmer & Papaloizou 1987; Binney & Tremaine 2008, §5.5.2). In such a cluster a quadrupolar distortion of the cluster's initial spherical shape grows exponentially. In essence, when pressure in the tangential directions falls below a certain threshold, the cluster becomes Jeans unstable in these directions. A cluster with a less radially biased velocity distribution may not be Jeans unstable but the frequencies of quadrupolar modes will be low because the system is highly compressible in the tangential directions (May & Binney 1986). In such a case we have to expect enhanced stellar diffusion in angular momentum and a shorter relaxation time than (7). In Section 5 we quantify this prediction.

### 3 THE BAILESCU–LENARD EQUATION

Over the last 20 years an approach to the secular evolution of stellar systems that differs radically from that of Spitzer–Chandrasekhar has emerged. In an important series of papers Weinberg demonstrated the significance of self-gravity in the responses of stellar systems to perturbation. Weinberg (1989) showed that the time required for dynamical friction to drag a satellite to the centre of its host galaxy is 2–3 times longer when self-gravity is included in the calculation than when it is not. This conclusion was obtained by extending to spherical systems the approach to stellar dynamics based on angle-action coordinates and orthonormal potential-density pairs that Kalnajs (1976) introduced in the context of discs. This apparatus enabled for the first time a satisfactory treatment of the large-scale distortions of a stellar system. Weinberg (1991) went on to determine the frequencies of the fundamental modes of anisotropic stellar systems and thus to determine the level of radial anisotropy at which the radial-orbit instability sets in. Using the same apparatus Weinberg (1994) showed that popular models of globular clusters have very weakly damped modes, a result that will prove crucial below.

Whereas in his previous papers the focus had been on the impact of an externally applied perturbation, Weinberg (1993) for the first time discussed the impact of Poisson noise. He did so in the context of the periodic cube of stars, and for this idealised system he was able to derive a well defined collision integral for the Boltzmann equation. By varying the ratio  $r$  of the Jeans length to the size of the

cube, he could show that self-gravity, by amplifying the excitation of large-scale fluctuations by Poisson noise, dramatically accelerates the system's relaxation when  $r$  is close to unity. Weinberg (1998) extended this analysis from the periodic cube to real systems. Using angle-action coordinates and orthonormal potential-density pairs he computed the dressed response of the system to a single particle, and the mean energy invested in thermally excited low-order modes.

Heyvaerts (2010), in the spirit of Weinberg's analysis of the periodic cube, approached the problem of secular evolution by returning to the BBGKY hierarchy of equations for the  $n$ -particle distribution functions  $f^{(n)}$  that one obtains by integrating the  $6N$  dimensional Liouville equation for the dynamics of an  $N$ -particle system over the coordinates of all but the first  $n$  particles. Working in angle-action coordinates to order  $1/N$ , and taking full account of the system's self-gravity, he obtained an expression for the 2-particle distribution function  $f^{(2)}$ . Inserting this into the BBGKY equation for the 1-particle DF, he obtained a Fokker-Planck equation together with expressions for its diffusion coefficients. This equation resembles an equation derived by Balescu (1960) and Lenard (1960) for the secular evolution of an electrostatic plasma, and in their honour we shall refer to it as the 'BL equation'. In Heyvaerts' derivation of the BL equation there is no suggestion of particles scattering one another. Instead, particles 1 and 2 interact when their frequency vectors  $\Omega_i$  satisfy a resonance condition  $\mathbf{n}_1 \cdot \Omega_1 + \mathbf{n}_2 \cdot \Omega_2 = 0$ , where the  $\mathbf{n}_i$  are vectors with integer components.

Rauch & Tremaine (1996) introduced the concept of 'resonant relaxation' in the context of star clusters that are dominated by a central massive black hole. In these systems the mean potential is nearly Keplerian, so each star can be replaced by the elliptical, Gaussian wire one obtains by averaging its motion around a Kepler orbit. The wires apply torques to each other, with the consequence that they precess. Unless the precession frequencies are closely matched, differential precession causes the sign of each angular momentum exchange to change quite rapidly, and hence average to zero. Consequently, the important interactions are those between stars that have commensurable precession frequencies – hence the term *resonant* relaxation. Heyvaerts' work implies that similar physics applies to any stellar system.

Chavanis (2012) derived the same equation as Heyvaerts by a different, perhaps more physically intuitive, route that does not involve the BBGKY hierarchy. He expressed the 1-particle DF as a sum of a part  $f_0$  that evolves only on a secular timescale, and a fluctuating part  $f_1$  that averages to zero on an orbital timescale. The potential was similarly decomposed into mean and fluctuating parts,  $\Phi_0$  and  $\Phi_1$ . Then he could easily show that  $f_0$  satisfies

$$\frac{\partial f_0}{\partial t} = -\langle [f_1, \Phi_1] \rangle, \quad (8)$$

where  $[.,.]$  is a Poisson bracket and  $\langle . \rangle$  is an ensemble average. Equations for  $f_1$  and the potential  $\Phi_1$  to which it gives rise are readily found, and Chavanis solved them for initial conditions characteristic of Poisson statistics. When the solutions are inserted into equation (8), the BL equation is recovered.

The BL equation can be written

$$\frac{\partial f_0(\mathbf{J}, t)}{\partial t} = -\frac{\partial}{\partial \mathbf{J}} \cdot \mathbf{F}, \quad (9)$$

where  $\mathbf{F}(\mathbf{J})$  is the flux of stars through action space. This flux falls naturally into two parts:

$$\mathbf{F}(\mathbf{J}) = -\mathbf{D}_1(\mathbf{J})f_0(\mathbf{J}) - \mathbf{D}_2(\mathbf{J}) \cdot \frac{\partial f_0}{\partial \mathbf{J}}, \quad (10)$$

where for brevity we have suppressed the slow time dependence of all quantities. The term proportional to the first-order diffusion coefficient  $\mathbf{D}_1$  describes dynamical friction: the constant drift back to lower actions and lower energy that in thermal equilibrium perfectly balances the stochastic drive away from the origin of action space. The second-order diffusion coefficient  $\mathbf{D}_2$  is analogous to an anisotropic thermal conductivity, which permits heat to diffuse from hotter (large  $f_0$ ) to colder regions at a particular speed in each direction. If  $\mathbf{D}_1$  vanished, the only steady-state solution would be  $f = 0$ , which would be established after the term proportional to  $\mathbf{D}_2$  had driven stars out into the infinite phase-space volume that exists at large  $|\mathbf{J}|$ .

Heyvaerts' derivation of the BL equation explicitly focuses on two-particle effects in as much as he starts by solving for the function that describes two-particle correlations. Chavanis' derivation arrives at a picture in which secular evolution occurs through interactions between pairs of resonantly coupled particles as a consequence of the mathematical accident that the collision term on the r.h.s. of equation (8) is non-vanishing only because the fluctuations in the DF and in the potential are correlated. In Chavanis' derivation the conceptual focus is on the impact of fluctuations in density and potential, and such fluctuations are by no means confined to pairwise interactions. The system's self-gravity plays a role because the fluctuating part of the DF,  $f_1$  evolves in a potential to which it contributes through Poisson's equation.

In detail, with  $f_0$  normalised such that

$$(2\pi)^3 \int d^3\mathbf{J} f_0 = N\mu, \quad (11)$$

with  $\mu$  the mass of a single particle and  $N$  the number of particles, we have the following expressions for the diffusion coefficients

$$\begin{aligned} \mathbf{D}_1(\mathbf{J}) &= -\frac{1}{2}(2\pi)^4 \mu \sum_{\mathbf{nn}'} \int d^3\mathbf{J}' |E_{\mathbf{nn}'}(\mathbf{J}, \mathbf{J}', \mathbf{n} \cdot \Omega)|^2 \\ &\quad \times \mathbf{n}' \cdot \frac{\partial f_0}{\partial \mathbf{J}'} \delta(\mathbf{n}' \cdot \Omega' - \mathbf{n} \cdot \Omega) \mathbf{n}, \\ \mathbf{D}_2(\mathbf{J}) &= \frac{1}{2}(2\pi)^4 \mu \sum_{\mathbf{nn}'} \int d^3\mathbf{J}' |E_{\mathbf{nn}'}(\mathbf{J}, \mathbf{J}', \mathbf{n} \cdot \Omega)|^2 \\ &\quad \times f_0(\mathbf{J}') \delta(\mathbf{n}' \cdot \Omega' - \mathbf{n} \cdot \Omega) \mathbf{n} \otimes \mathbf{n}. \end{aligned} \quad (12)$$

In these formulae  $\mathbf{n}, \mathbf{n}'$  are 3-vectors with integer components, and  $\Omega \equiv \Omega(\mathbf{J})$  is the vector formed by the frequencies of the star with actions  $\mathbf{J}$ . Analogously  $\Omega' \equiv \Omega(\mathbf{J}')$ . The occurrence of Dirac delta functions in equations (12) tells us that diffusion is possible only to the extent that stars have resonating frequencies: the rate at which stars with actions  $\mathbf{J}$  diffuse is proportional to a sum over all resonating stars, regardless of their location within the system. The impact of self gravity is encoded in the complex *susceptibility coefficients*  $E_{\mathbf{nn}'}(\mathbf{J}, \mathbf{J}', \omega)$ . To explain the content of the suscepti-

bility coefficients we have to introduce additional apparatus in the next section.

On account of the extensive similarities between the expressions for  $\mathbf{D}_1$  and  $\mathbf{D}_2$ , the action-space flux of stars (10) can be written in a relatively compact form

$$\mathbf{F}(\mathbf{J}) = \frac{1}{2}(2\pi)^4 \mu \sum_{\mathbf{nn}'} \mathbf{n} \int d^3 \mathbf{J}' |E_{\mathbf{nn}'}(\mathbf{J}, \mathbf{J}', \mathbf{n} \cdot \boldsymbol{\Omega})|^2 \times \left( \mathbf{n}' \cdot \frac{\partial}{\partial \mathbf{J}'} - \mathbf{n} \cdot \frac{\partial}{\partial \mathbf{J}} \right) f_0(\mathbf{J}) f_0(\mathbf{J}') \delta(\mathbf{n}' \cdot \boldsymbol{\Omega}' - \mathbf{n} \cdot \boldsymbol{\Omega}). \quad (13)$$

The factor  $\mu$  before the summation sign above signals that the flux is proportional to  $1/N$ .

### 3.1 Potential-density pairs

To solve Poisson's equation for the fluctuating potential  $\Phi_1$  given the fluctuating DF  $f_1$  one makes use of orthonormal potential-density pairs (Kahnajns 1976). Let  $(\Phi^{(p)}, \rho^{(p)})$  be such a pair. Then

$$\nabla^2 \Phi^{(p)}(\mathbf{x}) = 4\pi G \rho^{(p)}(\mathbf{x}), \quad (14)$$

and

$$\int d^3 \mathbf{x} (\Phi^{(a)})^* \rho^{(p)} = -\mathcal{E} \delta_{pq}, \quad (15)$$

where  $(\Phi^{(a)})^*$  is the complex conjugate of  $\Phi^{(a)}$  and  $\mathcal{E}$  is a convenient constant with the dimensions of energy. These pairs make it straightforward to compute the potential generated by any given density distribution:

$$\rho_1(\mathbf{x}, t) = \sum_p A_p(t) \rho^{(p)}(\mathbf{x}) \quad (16)$$

implies that

$$\Phi_1(\mathbf{x}, t) = \sum_p A_p(t) \Phi^{(p)}(\mathbf{x}), \quad (17)$$

with

$$A_p(t) = -\frac{1}{\mathcal{E}} \int d^3 \mathbf{x} \Phi(\mathbf{x})^* \rho_1(\mathbf{x}, t). \quad (18)$$

When expressed in action-angle variables, the basis functions depend on all six coordinates. We work with their Fourier transforms with respect to the angle variables, which are defined by

$$f_1(\boldsymbol{\theta}, \mathbf{J}, t) = \sum_{\mathbf{n}} \hat{f}_1(\mathbf{n}, \mathbf{J}, t) e^{i\mathbf{n} \cdot \boldsymbol{\theta}}, \quad \Phi_1(\boldsymbol{\theta}, \mathbf{J}, t) = \sum_{\mathbf{n}} \hat{\Phi}_1(\mathbf{n}, \mathbf{J}, t) e^{i\mathbf{n} \cdot \boldsymbol{\theta}}. \quad (19)$$

A straightforward calculation starting from the linearised collisionless Boltzmann equation shows that the Laplace transform

$$\tilde{f}_1(\mathbf{n}, \mathbf{J}, \omega) = \int_0^\infty dt \hat{f}_1(\mathbf{n}, \mathbf{J}, t) e^{i\omega t} \quad (20)$$

of  $\hat{f}_1(\mathbf{n}, \mathbf{J}, t)$  is given by

$$\tilde{f}_1(\mathbf{n}, \mathbf{J}, \omega) = \frac{\mathbf{n} \cdot \frac{\partial f_0}{\partial \mathbf{J}} \tilde{\Phi}_1(\mathbf{n}, \mathbf{J}, \omega) - i \hat{f}_1(\mathbf{n}, \mathbf{J}, 0)}{\mathbf{n} \cdot \boldsymbol{\Omega} - \omega}. \quad (21)$$

This equation gives the dynamical relationship between an

initial condition  $\hat{f}_1(\mathbf{n}, \mathbf{J}, 0)$  and both the disturbances to the DF and the potential that it provokes.

The disturbed DF and potential are also linked by Poisson's equation  $\nabla^2 \Phi_1 = 4\pi G \int d^3 \mathbf{v} f_1$ . We multiply equation (21) by  $\Phi^{(p)*} \sum_{\mathbf{n}} e^{i\mathbf{n} \cdot \boldsymbol{\theta}}$  and integrate over phase space. Then  $\tilde{A}_p(\omega)$ , the Laplace transform that describes  $\Phi_1$ , appears on both sides. Gathering on the left the terms proportional to  $\tilde{A}_p(\omega)$ , we obtain the matrix equation

$$\sum_{p'} \epsilon_{pp'}(\omega) \tilde{A}_{p'}(\omega) = i \frac{(2\pi)^3}{\mathcal{E}} \int d^3 \mathbf{J} \sum_{\mathbf{n}} \frac{\hat{f}_1(\mathbf{n}, \mathbf{J}, 0) [\hat{\Phi}^{(p)}(\mathbf{n}, \mathbf{J})]^*}{\mathbf{n} \cdot \boldsymbol{\Omega} - \omega}, \quad (22)$$

where

$$\epsilon_{pp'}(\omega) \equiv \delta_{pp'} - M_{pp'} \quad (23)$$

with

$$M_{pp'} \equiv \frac{(2\pi)^3}{\mathcal{E}} \int d^3 \mathbf{J} \sum_{\mathbf{n}} \frac{\mathbf{n} \cdot \frac{\partial f_0}{\partial \mathbf{J}}}{\mathbf{n} \cdot \boldsymbol{\Omega} - \omega} [\hat{\Phi}^{(p)}(\mathbf{n}, \mathbf{J})]^* \hat{\Phi}^{(p')}(\mathbf{n}, \mathbf{J}). \quad (24)$$

The matrix of susceptibility coefficients  $E_{\mathbf{nn}'}(\mathbf{J}, \mathbf{J}', \omega)$  that appears in equations (12) for the diffusion coefficients is the inverse of the matrix  $\boldsymbol{\epsilon}(\omega)$  when the latter is written in the  $(\mathbf{n}, \mathbf{J})$  coordinate system:

$$E_{\mathbf{nn}'}(\mathbf{J}, \mathbf{J}', \omega) \equiv \frac{1}{\mathcal{E}} \sum_{pp'} \hat{\Phi}^{(p)}(\mathbf{n}, \mathbf{J}) \epsilon_{pp'}^{-1}(\omega) [\hat{\Phi}^{(p')}(\mathbf{n}', \mathbf{J}')]^*, \quad (25)$$

It allows us to obtain from equation (22) the key relation

$$\tilde{\Phi}_1(\mathbf{n}, \mathbf{J}, \omega) = (2\pi)^3 i \int d^3 \mathbf{J}' \sum_{\mathbf{n}'} E_{\mathbf{nn}'}(\mathbf{J}, \mathbf{J}', \omega) \frac{\hat{f}_1(\mathbf{n}', \mathbf{J}', 0)}{\mathbf{n}' \cdot \boldsymbol{\Omega}' - \omega}. \quad (26)$$

### 3.2 Role of self-gravity

There are two terms in the numerator of equation (21) for the disturbance to the DF. The first describes the impact of self-gravity, and disappears in the limit  $G \rightarrow 0$ , while the second describes the evolution of the initial disturbance in the mean-field potential. The former term gives rise to the 'response matrix' integral  $M_{pp'}$  in the definition (23) of  $\boldsymbol{\epsilon}$ . In the limit  $G \rightarrow 0$ , the integral's prefactor  $\mathcal{E}^{-1} \sim G^{-1}$  diverges, but each of the potential basis functions in the integrand vanishes like  $G$ , so the self-gravitating response vanishes with  $G$  as we expect. We can switch off self-gravity in the BL equation by simply putting  $M_{pp'} = 0$ , thereby recovering the inhomogeneous Landau equation (Chavanis 2013b).

According to equation (26),  $E_{\mathbf{nn}'}(\mathbf{J}, \mathbf{J}', \omega)$  gives the change in the potential at frequency  $\omega$  that is provoked by a change in the initial conditions. Schematically we have an equation for the potential response of the form  $\mathcal{L}\tilde{\Phi}_1 = \mathcal{S}$ , with  $\mathcal{L}(\omega)$  a linear operator and  $\mathcal{S}$  a source term. If the system has a normal mode at frequency  $\omega$  (that is, a solution to  $\mathcal{L}(\omega)\tilde{\Phi}_1 = 0$ ) there isn't a unique relationship between the response of the system and the driving perturbation  $\mathcal{S}$ , because to any solution of the governing equation we can add a multiple of the normal mode. It follows that when  $\omega$  is an eigenfrequency of the system,  $\boldsymbol{\epsilon}$  is a singular matrix. Its inverse  $\mathbf{E}$  therefore diverges as  $\omega$  tends to the frequency of the normal mode.

The sums over  $\mathbf{n}$  and  $\mathbf{n}'$  in equation (13) suggest that diffusion is simply driven by pairs of resonating particles. This is a serious over-simplification, however, because the contribution of each pair is proportional to  $|E_{\mathbf{nn}'}|^2$ . This factor is large if the frequency at which the pair is communicating lies near an eigenfrequency of the whole system. One star perturbs a star at some distance from it by exciting oscillations in the entire cluster. If the cluster is responsive at the frequency in question, the two resonating stars can communicate effectively and rather quickly exchange significant energy and angular momentum, so we have rapid diffusion through phase space. In a tepid stellar disc the impact of the  $|E_{\mathbf{nn}'}|^2$  factor can be enormous, because the first star may launch a leading spiral wave, which is swing amplified as it is reflected off the forbidden zone around corotation before it is absorbed by the second star near the wave's inner Lindblad resonance. Fouvry et al. (2015) showed that on account of swing amplification the relaxation time in a realistically cool disc can be  $\sim 1000$  times shorter than one predicts when swing amplification is neglected. Below we show that self-gravity can significantly shorten the relaxation time of a non-rotating spherical system.

#### 4 APPLICATION TO SPHERICAL SYSTEMS

In the case of a spherical system, convenient action-space coordinates are the radial action  $J_r$ , the total angular momentum  $L$  and the component  $L_z$  of angular momentum parallel to the  $z$  axis. We confine ourselves to systems with no special axis; that is, we consider systems for which the DF depends only on  $J_r$  and  $L$ , so  $f_0$  is independent of  $L_z$ , which at given  $L$  merely encodes the inclination of the angular momentum vector with respect to the  $z$  axis.

Before the foregoing apparatus can be applied to any spherical system, a conceptual difficulty has to be finessed. The problem is that in a spherical system stars have only two independent frequencies. That is, there is a one-parameter family of vectors  $\mathbf{n}$  such that  $\mathbf{n} \cdot \boldsymbol{\Omega} = 0 \forall \mathbf{J}$ .<sup>1</sup> This is a significant issue for the derivation of the BL equation since the derivations of Heyvaerts (2010) and Chavanis (2012) assume that in each patch of action space the frequencies  $\Omega_i$  can be used as coordinates instead of the actions  $J_i$ , and when the frequencies are systematically degenerate this is not the case. It is also a problem as regards equations (12) for the diffusion coefficients. Indeed, in the  $(J_r, L, L_z)$  coordinate system, any vector  $\mathbf{n} = (0, 0, n)$  yields  $\mathbf{n} \cdot \boldsymbol{\Omega} = 0$  because  $\Omega_3 = \partial H / \partial L_z = 0 \forall \mathbf{J}$ . Physically, the vanishing of  $\Omega_3$  expresses the fact that in a spherical potential orbital planes do not precess. Mathematically vectors of the form  $\mathbf{n} = (0, 0, n)$  are problematic because they make the arguments of the Dirac delta functions vanish throughout action space, so the corresponding integral over actions becomes undefined.

To resolve the above problem, we define the 2-vector  $\tilde{\mathbf{J}} \equiv (J_r, L)$ .

We similarly define integer 2-vectors  $\tilde{\mathbf{n}} = (n_1, n_2)$  and frequency 2-vectors  $\tilde{\boldsymbol{\Omega}} = \partial H / \partial \tilde{\mathbf{J}}$  and observe that

$$\mathbf{n} \cdot \boldsymbol{\Omega} = \tilde{\mathbf{n}} \cdot \tilde{\boldsymbol{\Omega}} \quad (28)$$

because  $\Omega_3 = 0$ . It will be convenient to quantify  $J_3 = L_z$  through the inclination  $\beta$  defined by

$$\cos \beta \equiv \frac{J_3}{J_2} \quad (0 \leq \beta \leq \pi). \quad (29)$$

In terms of  $\beta$  the action-space volume element is

$$d^3 \mathbf{J} = dJ_r L dL d(\cos \beta) = d^2 \tilde{\mathbf{J}} J_2 d(\cos \beta). \quad (30)$$

In the spherical case, it is natural to take the potential-density pairs to be proportional to spherical harmonics  $Y_\ell^m$ . Thus we write<sup>2</sup>

$$\begin{aligned} \Phi^{(p)}(\mathbf{x}) &\equiv \Phi_{\ell mn}(r, \vartheta, \phi) = Y_\ell^m(\vartheta, \phi) U_n^\ell(r), \\ \rho^{(p)}(\mathbf{x}) &\equiv \rho_{\ell mn}(r, \vartheta, \phi) = Y_\ell^m(\vartheta, \phi) D_n^\ell(r), \end{aligned} \quad (31)$$

where the  $(U_n^\ell, D_n^\ell)$  are real radial functions. Several authors (e.g. Clutton-Brock 1973; Weinberg 1989; Hernquist & Ostriker 1992; Rahmati & Jalali 2009) have proposed choices for  $U_n^\ell(r)$ . In Section 5, we give an explicit example of such basis functions. Then one can show (see Tremaine & Weinberg 1984) that

$$\hat{\Phi}^{(p)}(\mathbf{n}, \mathbf{J}) = \delta_{m^p}^{\delta_{n_3}^p} i^{m^p - n_2} Y_{\ell^p}^{n_2}(\pi/2, 0) R_{n_2 m^p}^{\ell^p}(\beta) W_{\ell^p n^p}^{\tilde{\mathbf{n}}}(\tilde{\mathbf{J}}), \quad (32)$$

where<sup>3</sup>

$$\begin{aligned} R_{n_2 m}^\ell(\beta) &= \sum_t (-1)^t \frac{\sqrt{(\ell + n_2)! (\ell - n_2)! (\ell + m)! (\ell - m)!}}{(\ell - m - t)! (\ell + n_2 - t)! t! (t + m - n_2)!} \\ &\times [\cos(\beta/2)]^{2\ell + n_2 - m - 2t} [\sin(\beta/2)]^{2t + m - n_2} \end{aligned} \quad (33)$$

is the spin- $\ell$  Wigner rotation matrix, and

$$W_{\ell n}^{\tilde{\mathbf{n}}}(\tilde{\mathbf{J}}) = \frac{1}{\pi} \int_0^\pi d\theta_1 U_n^\ell(r(\theta_1)) \cos[n_1 \theta_1 + n_2(\theta_2 - \psi)]. \quad (34)$$

Here  $\psi$  is the angular coordinate within the orbital plane. It depends on both  $\theta_1$  and  $\theta_2$ , but the difference  $\theta_2 - \psi$  is a function of  $\theta_1$  alone, so the integral is well defined.

When  $f_0$  is independent of  $L_z$ , the only  $\beta$  dependence in the response matrix from equation (24) comes from the rotation matrices  $R$ . Given the orthogonality of these matrices when integrated over  $\cos \beta$  (e.g. Edmonds 1996)

$$\int_{-1}^1 d(\cos \beta') R_{n_2 n_3}^{\ell^p}(\beta') R_{n_2 n_3}^{\ell^q}(\beta') = \frac{2\delta_{\ell^p}^{\ell^q}}{2\ell^p + 1}, \quad (35)$$

we can write the response matrix as

$$M_{pq}(\omega) = \delta_{\ell^p}^{\ell^q} \delta_{m^p}^{m^q} \xi_{\ell^p n^p n^q}(\omega), \quad (36)$$

where

$$\begin{aligned} \xi_{\ell^p n^p n^q}(\omega) &\equiv \frac{(2\pi)^3}{\mathcal{E}} \sum_{\substack{n_1 \\ |n_2| \leq \ell^p \\ (\ell^p - n_2) \text{ even}}} \frac{2}{2\ell^p + 1} |Y_{\ell^p}^{n_2}(\pi/2, 0)|^2 \\ &\times \int d^2 \tilde{\mathbf{J}} L \frac{\tilde{\mathbf{n}} \cdot \partial f_0 / \partial \tilde{\mathbf{J}}}{\omega - \tilde{\mathbf{n}} \cdot \tilde{\boldsymbol{\Omega}}} W_{\ell^p n^p}^{\tilde{\mathbf{n}}}(\tilde{\mathbf{J}}) W_{\ell^p n^q}^{\tilde{\mathbf{n}}}(\tilde{\mathbf{J}}). \end{aligned} \quad (37)$$

<sup>1</sup> In the Kepler potential of a massive black hole there is only one independent frequency and this problem is still more acute (Fouvry et al. 2017).

<sup>2</sup> The indices on the spherical harmonics run over  $\ell = 0, 1, 2, \dots$  and  $m = -\ell, -\ell + 1, \dots, \ell - 1, \ell$ .

<sup>3</sup> In equation (33) the sum over  $t$  is restricted such that all arguments of the factorial operators are  $\geq 0$ .

It follows that the matrix  $\epsilon = \mathbf{I} - \mathbf{M}$  defined by equation (23) is

$$\epsilon_{pq} = \delta_p^q - \delta_{\ell^q}^{\ell^q} \delta_{m^p}^{m^q} \xi_{\ell^p n^p n^q}(\omega). \quad (38)$$

Then putting  $\delta_p^q = \delta_{\ell^q}^{\ell^q} \delta_{m^p}^{m^q} \delta_{n^p}^{n^q}$  we arrive at

$$\epsilon_{pq}^{-1} = \delta_{\ell^q}^{\ell^q} \delta_{m^p}^{m^q} N_{\ell^p n^p n^q}(\omega), \quad (39)$$

where  $N_{\ell^p n^p n^q}(\omega)$  is the inverse of  $[\delta_{n^p}^{n^q} - \xi_{\ell^p n^p n^q}(\omega)]$ .

We insert this expression for  $\epsilon^{-1}$  into equation (25) for the susceptibility coefficients and use the reality of  $R$ ,  $W$  and  $Y_\ell^m(\pi/2, 0)$ , to obtain

$$\begin{aligned} E_{\mathbf{nn}'}(\mathbf{J}, \mathbf{J}', \omega) &= \frac{1}{\mathcal{E}} \sum_{\ell^p n^p} \sum_{\substack{\ell^q n^q \\ |m^p| \leq \ell^p \\ |m^q| \leq \ell^q}} \delta_{m^p}^{n^3} \delta_{m^q}^{n_3'} \delta_{\ell^p}^{\ell^q} \delta_{m^p}^{m^q} \\ &\times \left[ i^{(n_2 - n_2')} i^{(m^p - m^q)} \right] \left[ Y_{\ell^p}^{n_2}(\pi/2, 0) Y_{\ell^q}^{n_2'}(\pi/2, 0) \right] \\ &\times \left[ R_{n_2 m^p}^{\ell^p}(\beta) R_{n_2' m^q}^{\ell^q}(\beta') \right] \\ &\times \left[ W_{\ell^p n^p}^{\tilde{\mathbf{n}}}(\tilde{\mathbf{J}}) W_{\ell^q n^q}^{\tilde{\mathbf{n}}'}(\tilde{\mathbf{J}}') \right] N_{\ell^p n^p n^q}(\omega). \end{aligned} \quad (40)$$

The sums on  $\ell^q$ ,  $m^p$  and  $m^q$  are now trivially executed. The non-trivial part of  $E_{\mathbf{nn}'}$  that is independent of  $\beta$  is

$$\begin{aligned} \Lambda_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}'}^{\ell}(\tilde{\mathbf{J}}, \tilde{\mathbf{J}}', \omega) &\equiv Y_{\ell}^{n_2}(\pi/2, 0) Y_{\ell}^{n_2'}(\pi/2, 0) \\ &\times \sum_{n^p n^q} W_{\ell n^p}^{\tilde{\mathbf{n}}}(\tilde{\mathbf{J}}) W_{\ell n^q}^{\tilde{\mathbf{n}}'}(\tilde{\mathbf{J}}') N_{\ell n^p n^q}(\omega), \end{aligned} \quad (41)$$

and with this definition  $|E_{\mathbf{nn}'}|^2$  can be written

$$\begin{aligned} |E_{\mathbf{nn}'(\mathbf{J}, \mathbf{J}', \omega)}|^2 &= \frac{1}{\mathcal{E}^2} \delta_{n_3}^{n_3'} \\ &\times \sum_{\ell^p} \sum_{\ell^q} \Lambda_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}'}^{\ell^p}(\tilde{\mathbf{J}}, \tilde{\mathbf{J}}', \omega) \Lambda_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}'}^{\ell^q}(\tilde{\mathbf{J}}, \tilde{\mathbf{J}}', \omega) \\ &\times R_{n_2 n_3}^{\ell^p}(\beta) R_{n_2 n_3}^{\ell^q}(\beta) R_{n_2' n_3'}^{\ell^p}(\beta') R_{n_2' n_3'}^{\ell^q}(\beta'). \end{aligned} \quad (42)$$

The diffusion coefficients (12), for which we require  $|E_{\mathbf{nn}'}|^2$ , involve integrals over the actions, whose volume element is given by (30). Given that the DF  $f_0$  does not depend on  $J_3$ , we can execute the integral over  $\cos \beta$  up front. In view of the orthogonality relation (35) we have

$$\begin{aligned} \int dJ_3' |E_{\mathbf{nn}'(\mathbf{J}, \mathbf{J}', \mathbf{n} \cdot \tilde{\mathbf{\Omega}})}|^2 &= \frac{1}{\mathcal{E}^2} \delta_{n_3}^{n_3'} J_2' \\ &\times \sum_{\ell} \frac{2}{2\ell+1} \left| \Lambda_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}'}^{\ell}(\tilde{\mathbf{J}}, \tilde{\mathbf{J}}', \tilde{\mathbf{n}} \cdot \tilde{\mathbf{\Omega}}) \right|^2 \left| R_{n_2 n_3}^{\ell}(\beta) \right|^2. \end{aligned} \quad (43)$$

We now turn to evaluating the flux  $\mathbf{F}$  from equation (13). The assumed form  $f_0(\mathbf{J})$  of the DF together with equation (28) enables us to simplify equation (13) to

$$\begin{aligned} \mathbf{F}(\mathbf{J}) &= \frac{1}{2} (2\pi)^4 \mu \sum_{\mathbf{nn}'} \mathbf{n} \int d^2 \tilde{\mathbf{J}}' \int dJ_3' |E_{\mathbf{nn}'(\mathbf{J}, \mathbf{J}', \mathbf{n} \cdot \tilde{\mathbf{\Omega}})}|^2 \\ &\times \left( \tilde{\mathbf{n}}' \cdot \frac{\partial}{\partial \tilde{\mathbf{J}}'} - \tilde{\mathbf{n}} \cdot \frac{\partial}{\partial \tilde{\mathbf{J}}} \right) f_0(\tilde{\mathbf{J}}) f_0(\tilde{\mathbf{J}}') \delta(\tilde{\mathbf{n}}' \cdot \tilde{\mathbf{\Omega}}' - \tilde{\mathbf{n}} \cdot \tilde{\mathbf{\Omega}}). \end{aligned} \quad (44)$$

Consequently the only dependence of  $\mathbf{F}$  on the third components of  $\mathbf{n}$  and  $\mathbf{n}'$  is given by equation (43), which states that non-zero contributions to  $\mathbf{F}$  arise only when  $n_3 = n_3'$ . For given values of  $\tilde{\mathbf{n}}$  and  $\tilde{\mathbf{n}}'$  we can sum over all values of

$n_3 = n_3'$  and take advantage of the identities

$$\begin{aligned} \sum_{n_3} \left| R_{n_2 n_3}^{\ell}(\beta) \right|^2 &= 1, \\ \sum_{n_3} n_3 \left| R_{n_2 n_3}^{\ell}(\beta) \right|^2 &= n_2 \cos \beta = n_2 L_z / L, \end{aligned} \quad (45)$$

to find

$$\mathbf{F} = \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix} = \sum_{\tilde{\mathbf{n}}} \begin{pmatrix} n_1 \\ n_2 \\ n_2 J_3 / J_2 \end{pmatrix} \mathcal{F}_{\tilde{\mathbf{n}}}(\tilde{\mathbf{J}}), \quad (46)$$

where

$$\begin{aligned} \mathcal{F}_{\tilde{\mathbf{n}}}(\tilde{\mathbf{J}}) &= (2\pi)^4 \frac{\mu}{\mathcal{E}^2} \sum_{\tilde{\mathbf{n}}'} \int d^2 \tilde{\mathbf{J}}' J_2' \sum_{\ell} \frac{|\Lambda_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}'}^{\ell}(\tilde{\mathbf{J}}, \tilde{\mathbf{J}}', \tilde{\mathbf{n}} \cdot \tilde{\mathbf{\Omega}})|^2}{2\ell+1} \\ &\times \left( \tilde{\mathbf{n}}' \cdot \frac{\partial}{\partial \tilde{\mathbf{J}}'} - \tilde{\mathbf{n}} \cdot \frac{\partial}{\partial \tilde{\mathbf{J}}} \right) f_0(\tilde{\mathbf{J}}) f_0(\tilde{\mathbf{J}}') \delta(\tilde{\mathbf{n}}' \cdot \tilde{\mathbf{\Omega}}' - \tilde{\mathbf{n}} \cdot \tilde{\mathbf{\Omega}}). \end{aligned} \quad (47)$$

Computing the divergence of  $\mathbf{F}$  from equation (46) we find

$$\text{div } \mathbf{F} = \frac{\partial F_1}{\partial J_1} + \frac{\partial F_2}{\partial J_2} + \frac{F_2}{J_2}. \quad (48)$$

Since  $F_1$  and  $F_2$  only depend on  $\tilde{\mathbf{J}}$ , it follows that  $\text{div } \mathbf{F}$  does not depend on  $J_3$ . Hence, a DF that initially is independent of  $L_z = J_3$  will remain so as diffusion proceeds. This result ensures that an initially stable spherical cluster remains spherical.

We will work with the integral with respect to  $J_3$  of the BL equation. We define

$$\bar{f}(\tilde{\mathbf{J}}) \equiv \int_{-J_2}^{J_2} dJ_3 f_0(\tilde{\mathbf{J}}) = 2J_2 f_0(\tilde{\mathbf{J}}), \quad (49)$$

and note that since  $\text{div } \mathbf{F}$  is independent of  $J_3$ ,

$$\begin{aligned} \int_{-J_2}^{J_2} dJ_3 \text{div } \mathbf{F} &= 2J_2 \text{div } \mathbf{F} \\ &= 2 \left( \frac{\partial (J_2 F_1)}{\partial J_1} + \frac{\partial (J_2 F_2)}{\partial J_2} \right). \end{aligned} \quad (50)$$

Hence the integral of the BL equation with respect to  $J_3$  is

$$\frac{\partial \bar{f}}{\partial t} = - \frac{\partial}{\partial \tilde{\mathbf{J}}} \cdot \bar{\mathbf{F}}(\tilde{\mathbf{J}}), \quad (51)$$

where

$$\bar{\mathbf{F}} \equiv \sum_{\tilde{\mathbf{n}}} \tilde{\mathbf{n}} \bar{\mathcal{F}}_{\tilde{\mathbf{n}}}(\tilde{\mathbf{J}}) \quad (52)$$

with

$$\begin{aligned} \bar{\mathcal{F}}_{\tilde{\mathbf{n}}}(\tilde{\mathbf{J}}) &\equiv \frac{1}{2} (2\pi)^4 \frac{\mu}{\mathcal{E}^2} \sum_{\tilde{\mathbf{n}}'} \int d^2 \tilde{\mathbf{J}}' J_2 J_2' \sum_{\ell} \frac{|\Lambda_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}'}^{\ell}(\tilde{\mathbf{J}}, \tilde{\mathbf{J}}', \tilde{\mathbf{n}} \cdot \tilde{\mathbf{\Omega}})|^2}{2\ell+1} \\ &\times \left( \tilde{\mathbf{n}}' \cdot \frac{\partial}{\partial \tilde{\mathbf{J}}'} - \tilde{\mathbf{n}} \cdot \frac{\partial}{\partial \tilde{\mathbf{J}}} \right) \frac{\bar{f}(\tilde{\mathbf{J}})}{J_2} \frac{\bar{f}(\tilde{\mathbf{J}}')}{J_2'} \delta(\tilde{\mathbf{n}}' \cdot \tilde{\mathbf{\Omega}}' - \tilde{\mathbf{n}} \cdot \tilde{\mathbf{\Omega}}). \end{aligned} \quad (53)$$

Equations (51) to (53) describe diffusion of stars in the  $J_r L$  plane; no references to the  $L_z$  coordinate remain. The equations are formally similar to those that govern the relaxation of a two-dimensional stellar disc (Fouvry et al. 2015). In equation (53) the dressed friction was already computed by Weinberg (1989) in his equation (53). Self-gravity can be switched off by setting  $M_{pq} = 0$  when computing the coefficients  $\Lambda_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}}'^{\ell}$ . Fluxes computed with self-gravity switched

off are referred to as ‘bare’, whereas fluxes which include self-gravity are ‘dressed’.

Since  $L$  and  $J_r$  are inherently positive, it is a logical necessity that the component of  $\bar{\mathbf{F}}$  perpendicular to the  $L$  and  $J_r$  axes should vanish at those axes, so stars cannot diffuse to negative values of  $L$  or  $J_r$ . The factor  $J_2 = L$  after the infinitesimal in the definition (53) of  $\bar{\mathcal{F}}_{\bar{\mathbf{n}}}(\bar{\mathbf{J}})$  guarantees that  $\bar{\mathbf{F}}$  vanishes as  $L \rightarrow 0$ . For the case  $J_r \rightarrow 0$ , we can change integration variable in the coefficients  $W_{\ell n}^{\bar{\mathbf{n}}}(\bar{\mathbf{J}})$  defined by equation (34), so that they are integrals over  $r$  from  $r_p$  to  $r_a$ . They will vanish as  $J_r \rightarrow 0$  because then  $r_p \rightarrow r_a$  so the limits of integration become identical. By equation (41)  $\Lambda_{\bar{\mathbf{n}}\bar{\mathbf{n}}'}^{\ell}$  vanishes with  $W_{\ell n}^{\bar{\mathbf{n}}}$  causing  $\bar{\mathcal{F}}_{\bar{\mathbf{n}}}$  to vanish also. Hence there can be no flux across the  $L$  axis either.

## 5 APPLICATION TO THE ISOCHRONE

The natural testbed for the formalism presented above is the isochrone model since for it alone we have analytic expressions  $\theta(\mathbf{x}, \mathbf{v})$  and  $\mathbf{J}(\mathbf{x}, \mathbf{v})$  for the angle-action variables.

Henon (1959) derived the isochrone potential

$$\Phi_0(r) \equiv \Phi_{\text{iso}}(r) = -\frac{GM}{b + \sqrt{b^2 + r^2}}, \quad (54)$$

where  $b$  sets the model’s linear scale and  $M$  is the model’s mass, as the only potential in which  $\Omega_r$  is independent of  $L$  at given energy  $E$ . The Hamiltonian for motion in this potential is

$$H_0(\bar{\mathbf{J}}) = -\frac{(GM)^2}{2 \left[ J_r + \frac{1}{2} (L + \sqrt{L^2 + 4GMb}) \right]^2}, \quad (55)$$

Hénon (1960) gave the ergodic DF  $f(E)$  that self-consistently generates the isochrone potential. Since we wish to probe the impact of velocity anisotropy on the rate of relaxation, we take the DF  $f_0$  to be a member of the Osipkov-Merritt family of DFs (May & Binney 1986; Binney & Tremaine 2008). Specifically,

$$f_0(Q) = \frac{(1-Q)^{-4} M}{128\sqrt{2}\pi^3 (GMb)^{3/2}} \left( \sqrt{Q} [27 + 77\gamma - (66 + 286\gamma)Q] + (320 + 136\gamma)Q^2 - (240 + 32\gamma)Q^3 + 64Q^4 \right) + \frac{3 \arcsin \sqrt{Q}}{\sqrt{1-Q}} \left[ (17\gamma - 9) + (28 - 44\gamma)Q + (16 - 8\gamma)Q^2 \right], \quad (56)$$

where

$$Q \equiv -\frac{b}{GM} \left( E + \frac{L^2}{2R_a^2} \right), \quad \text{and} \quad \gamma \equiv (b/R_a)^2. \quad (57)$$

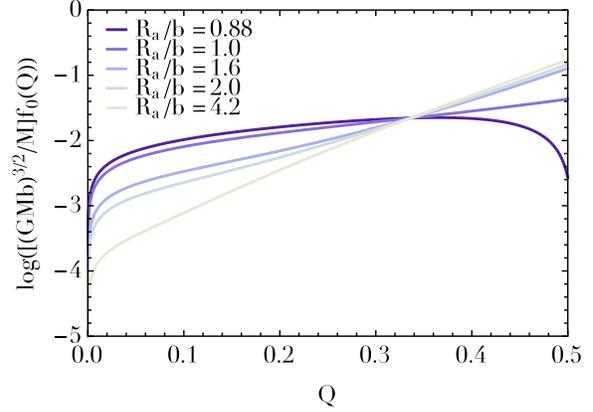
With

$$\beta = 1 - \frac{\langle v_t^2 \rangle}{2\langle v_r^2 \rangle}, \quad (58)$$

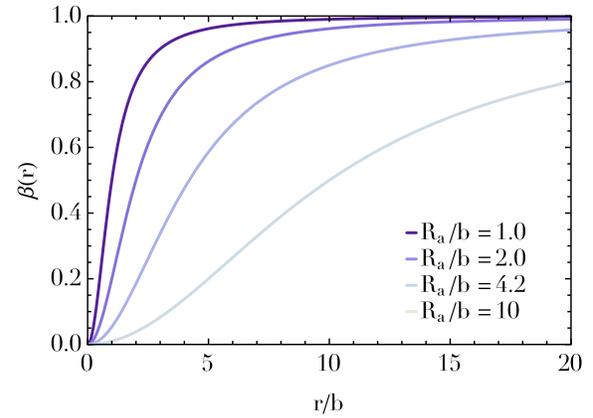
where  $v_t = \sqrt{v_\phi^2 + v_\theta^2}$  is the tangential speed, for any Osipkov-Merritt model we have (Binney & Tremaine 2008)

$$\beta(r) = \frac{1}{1 + R_a^2/r^2}. \quad (59)$$

Hence  $R_a$  is called the ‘anisotropy radius’ and is the radius near which the velocity distribution transitions from isotropic at  $r \ll R_a$  to radially biased at  $r \gg R_a$ . Fig. 1



**Figure 1.** The Osipkov-Merritt DF  $f_0(Q)$  (equation (56)) as a function of  $Q$ . Note that for  $f_0$  to be well defined we require  $R_a > 0.874b$ .



**Figure 2.** Anisotropy  $\beta \equiv 1 - \langle v_t^2 \rangle / \langle v_r^2 \rangle$  as a function of radius  $r$  for various anisotropy radii  $R_a$ .

shows the function  $f_0(Q)$  for several values of  $R_a$ , while Fig. 2 shows the corresponding runs of  $\beta(r)$ .

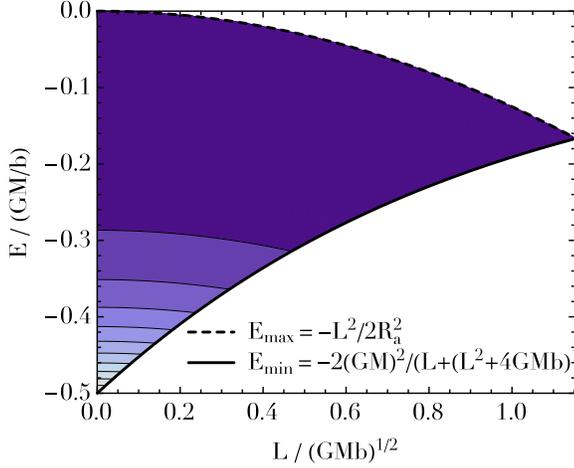
The model’s stars are confined to the part of the  $(E, L)$  plane that is shaded in Figs. 3 and 4. The lower boundary of this region is set by the requirement that  $E$  be not smaller than the energy of a circular orbit at the given value of  $L$ , which is

$$E_{\text{min}} = -\frac{2(GM)^2}{(L + \sqrt{L^2 + 4GMb})^2}. \quad (60)$$

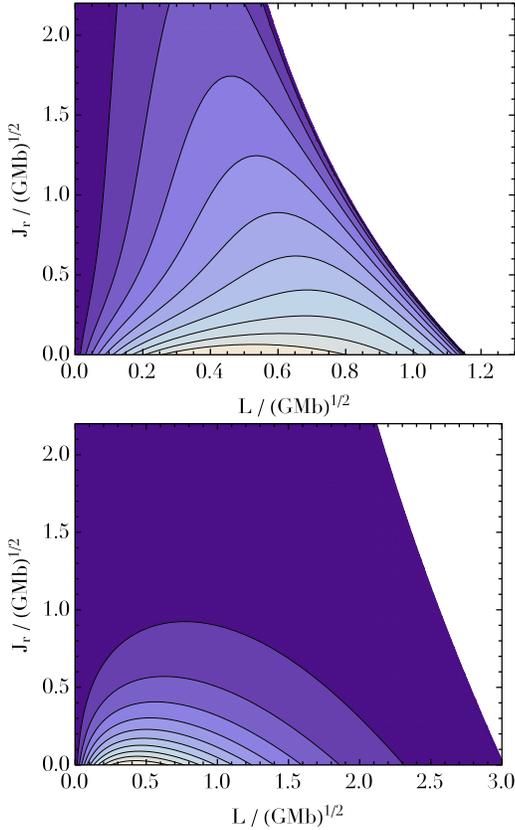
The upper boundary is simply the curve  $Q = 0$ . In addition, to have  $f_0 \geq 0$  everywhere, we require  $R_a > 0.874b$ . By the Doremus-Feix-Baumann theorem (Binney & Tremaine 2008, §5.5) the radial modes of a spherical model are all stable if  $\partial f_0 / \partial E < 0$  everywhere, and models with  $\partial f_0 / \partial L > 0$  tend to be unstable. We consider only models with  $R_a > 0.97b$  to ensure that  $\partial f_0 / \partial E < 0$  and  $\partial f_0 / \partial L < 0$  everywhere.

In significantly anisotropic models the boundary at  $Q = 0$  generates large gradients in  $f$  that are problematic numerically. Consequently, for models with  $R_a < 10b$  we have smoothed  $f$  by multiplying it by  $e^{-0.01/Q}$ .

We use the radial basis functions of Weinberg (1989)



**Figure 3.** Contours of  $f_0(Q(E, L))$  in the allowed part of  $(E, L)$  space, for  $R_a = 2b$ . Contours are spaced linearly from the minimum (dark) to the maximum (light) values taken by  $f_0(Q(E, L))$ .



**Figure 4.** Contour plots of the two-dimensional DF  $\bar{f} = 2Lf_0(Q)$  when  $R_a = 2b$  (above) and  $10b$  (below).

rescaled by a factor  $M/\sqrt{R}$ , so they become

$$U_n^\ell(r) = -\frac{GM}{R} \frac{4\pi\sqrt{2}}{\alpha_{\ell n} |j_\ell(\alpha_{\ell n})|} j_\ell(\alpha_{\ell n} r/R) \quad (61)$$

$$D_n^\ell(r) = \frac{M}{R^3} \frac{\alpha_{\ell n} \sqrt{2}}{|j_\ell(\alpha_{\ell n})|} j_\ell(\alpha_{\ell n} r/R), \quad (62)$$

where the density is assumed to vanish beyond the truncation radius  $R$ ,  $j_\ell$  is a spherical Bessel function of the first

kind and  $\alpha_{\ell n}$  is the  $n$ th zero of  $j_\ell$ . For these basis functions the normalising constant in equation (15) is

$$\mathcal{E} = 4\pi \frac{GM^2}{R}. \quad (63)$$

### 5.1 Calculation of the response matrix

To ensure that our code is applicable to any spherical model and not just isochrone models, we use as coordinates for orbit space the radii  $r_p$  and  $r_a$  of peri- and apo-centre rather than  $J_r$  and  $L$ . These radii are the roots of the equation

$$E - \Phi(r) - \frac{L^2}{2r^2} = 0. \quad (64)$$

Hence, given  $r_p$  and  $r_a$  it follows easily that

$$E = \frac{r_a^2 \Phi(r_a) - r_p^2 \Phi(r_p)}{r_a^2 - r_p^2}; \quad L = \sqrt{\frac{2[\Phi(r_a) - \Phi(r_p)]}{r_p^{-2} - r_a^{-2}}}. \quad (65)$$

Putting  $H = E$  in the Hamiltonian (55) we recover  $J_r(r_p, r_a)$ . Integrals over  $d^2\tilde{\mathbf{J}}$  can be expressed as integrals over  $(r_p, r_a)$  using the Jacobian

$$\frac{\partial(J_r, L)}{\partial(r_p, r_a)} = \frac{\partial(J_r, L)}{\partial(E, L)} \frac{\partial(E, L)}{\partial(r_p, r_a)} = \frac{1}{\Omega_r} \frac{\partial(E, L)}{\partial(r_p, r_a)}. \quad (66)$$

Since  $f_0 = f_0(Q(E, L))$ , we have

$$\tilde{\mathbf{n}} \cdot \frac{\partial f_0}{\partial \tilde{\mathbf{J}}} = -\frac{b}{GM} \frac{df_0}{dQ} \left[ n_1 \Omega_r + n_2 \left( \Omega_\vartheta + \frac{L}{R_a^2} \right) \right]. \quad (67)$$

To perform the integration numerically over the  $(r_p, r_a)$  plane implied by equation (37), we divide the plane into regions labelled by  $i$ . The  $i$ th region is centred at  $(r_p^i, r_a^i)$  and covers the square  $r_p \in [r_p^i - \Delta r/2, r_p^i + \Delta r/2]$ ,  $r_a \in [r_a^i - \Delta r/2, r_a^i + \Delta r/2]$ . Then in each region we separately Taylor expand the frequency  $\tilde{\mathbf{n}} \cdot \tilde{\boldsymbol{\Omega}} - \omega$  that appears as the denominator in equation (37) and its numerator. The ratio of Taylor expansions is then analytically integrated through the region. The numerical steps involved are exactly those that have been detailed extensively in Fouvry et al. (2015) so we will not repeat the recipe here.

The most computationally expensive part of this procedure is the calculation from equation (34) of the matrix elements  $W$  at each point on the  $(r_p, r_a)$  grid. It is important that we use a large, dense grid, as the biggest contributions to  $\tilde{\mathbf{n}} \cdot \partial f_0 / \partial \tilde{\mathbf{J}}$  come from the edge of the model, where  $|df_0/dQ|$  is largest. However, for small values of  $R_a$  the requirement  $Q > 0$  usefully restricts the portion of the  $(r_p, r_a)$  grid within which the matrix elements of  $W$  need to be computed (Fig. 3). For this reason fluxes are more easily computed for models with small values of  $R_a$  than large ones.

### 5.2 Recovering unstable modes

To demonstrate that our evaluation of the response matrix  $\mathbf{M}(\omega)$  is correct, we recover the known unstable modes of our system. For a spherical system the perturbations can be conveniently decomposed into spherical harmonics  $Y_\ell^m$ , and we can consider each harmonic  $\ell$  separately. Modes correspond to frequencies  $\omega = \omega_0 + i\eta$  with  $\eta > 0$  for which the

matrix  $\epsilon = \mathbf{I} - \mathbf{M}(\omega)$  has a zero eigenvalue; in other words, we seek  $\omega$  such that

$$\det \epsilon = 0. \quad (68)$$

The mode has frequency  $\omega_0$  and growth rate  $\eta$ . Since from equation (38)

$$\epsilon_{pq} = \delta_\ell^q \delta_{m^q}^{m^q} [\delta_{n^q}^{n^q} - \xi_{\ell n^q}(\omega)]. \quad (69)$$

for each fixed  $\ell$  we need to compute the determinant of the matrix  $\delta_{n^q}^{n^q} - (\xi_\ell)_{n^q n^q}$ .

Saha (1991) used the Osipkov-Merritt DF (56) to study the radial-orbit instability in anisotropic isochrone models. He found that a mode with  $\ell = 2$  becomes unstable when  $R_a$  falls below  $\sim 4b$ . In terms of the natural unit of time  $T_1 = \sqrt{b^3/GM}$ , growth rates  $\eta$  are very small ( $\eta \ll 0.001/T_1$ ) when the mode first becomes unstable but grow towards  $\eta \approx 0.025/T_1$  as  $R_a \rightarrow b$ .<sup>4</sup>

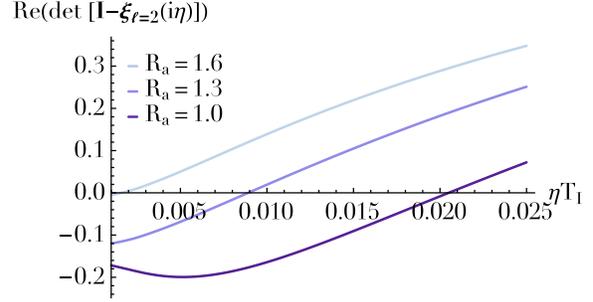
The calculations in Saha (1991) used 10 radial basis functions  $U_n^\ell$  from a different family from that used here, and focused on  $\ell = 2$ . Since in a spherical system the growth rate must be independent of the azimuthal part of the spherical harmonics, Saha set  $m = 0$ . The unstable mode has vanishing pattern speed, so in searching for it Saha set  $\omega_0 = 0$ . We follow all of these conventions. In our standard computation the truncation radius of the basis functions is  $R = 20b$ , and the grid in  $\{(r_p^i, r_a^i)\}$  covers the range  $r^i \in [0.15b, 19.85b]$  with grid spacing  $\Delta r = 0.1b$ . Finally, the maximum index  $n_1$  in the sum of equation (37) is  $n_1 = n^{\max} = 4$ . The justification for this choice is that the isochrone's frequencies satisfy

$$\frac{\Omega_2}{\Omega_1} = \frac{1}{2} \left( 1 + \frac{J_2}{\sqrt{J_2^2 + 4GMb}} \right), \quad (70)$$

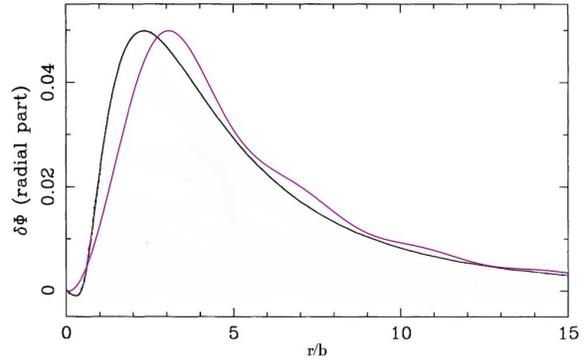
so orbits with  $J_2 \ll 2\sqrt{GMb}$  satisfy  $2\Omega_2 - \Omega_1 \approx 0$ . Consequently, we expect equation (37) to be dominated by the term with  $(n_1, n_2) = \pm(1, -2)$ .

Since  $\eta$  is very small and all other contributions to the response matrix are real, the imaginary part of  $\det[\mathbf{I} - \xi_{\ell=2}(i\eta)]$  is negligible. Hence to locate the unstable modes, we need only plot  $\text{Re}(\det[\mathbf{I} - \xi_{\ell=2}(i\eta)])$  for various values of the growth rate  $\eta$ , and pick out the  $\eta$  for which  $\text{Re}(\det[\mathbf{I} - \xi_{\ell=2}(i\eta)]) = 0$ . Fig. 5 shows  $\text{Re}(\det[\mathbf{I} - \xi_{\ell=2}(i\eta)])$  as a function of  $\eta$  for three values of  $R_a$ . Each curve crosses the  $x$  axis, so all these models are subject to the radial-orbit instability. However, the model with  $R_a = 1.6b$  only just manages to cross the  $x$  axis, suggesting that the limiting value for instability,  $R_{a,\max}$  lies close to  $1.6b$ . This conclusion is consistent with the conclusion of May & Binney (1986) that  $R_{a,\max} \simeq 1.67b$ . The top row of Table 1 gives our growth rates for these four models plus indications that models with  $R_a/b \geq 2$  are stable. The second and third rows of Table 1 give corresponding results from Saha (1991). Our growth rates tend to be slightly smaller than the rates Saha obtained from linear theory, and significantly smaller than the rates Saha inferred from N-body simulations, especially for the marginally unstable model.

<sup>4</sup> Saha's Figure 2(b) shows the growth rates  $\eta$  for various anisotropies  $\beta$ . To convert between Saha's  $\beta$  and  $R_a$ , use his Figure 1.



**Figure 5.** Plot of  $\text{Re}(\det[\mathbf{I} - \xi_{\ell=2}(0 + i\eta)])$  in units of  $T_1$  as a function of the growth rate  $\eta$  for  $R_a/b = 1.0, 1.3, 1.6$  (from bottom to top). Crossing the  $\eta$  axis indicates the existence of the radial-orbit instability.



**Figure 6.** The radial part of our  $\Phi_1$  (what Saha calls  $\delta\Phi$ ) in purple, superimposed on Saha's version of the same mode (black), for the unstable  $\ell = 2$  mode in the model with  $R_a = b$ . The normalisation of the vertical axis is arbitrary. (Graphics edited from Saha (1991)).

Linear theory enables us to predict the shape (but not the amplitude) of the perturbation to the potential associated with an unstable mode. Indeed, with  $\eta$  the growth rate of the unstable mode,  $\xi_{\ell=2}(i\eta)$  is a Hermitian matrix (to see this, take the Hermitian conjugate of equation (37) and use  $W_{\ell n}^{\bar{n}} = W_{\ell n}^{-\bar{n}}$ ). It has an eigenvector  $\mathbf{X}$  with eigenvalue unity, and the perturbation to the potential is

$$\Phi_1(\mathbf{x}) \propto \text{Re} \left[ \sum_{n=1}^{n^{\max}} X^n \Phi^{(n,2,0)}(\mathbf{x}) \right] \quad (71)$$

$$= \left[ \sum_{n=1}^{n^{\max}} X^n \frac{j_2(\alpha_{2n}r/R)}{\alpha_{2n} |j_2(\alpha_{2n})|} \right] P_2(\cos \theta). \quad (72)$$

The purple line in Fig. 6 shows  $\Phi_1$  when  $R_a = b$ , while the black curve reproduces the corresponding plot from Saha (1991). Overall the agreement is good.

### 5.2.1 Convergence study

Table 2 shows how the recovered growth rates of the radial-orbit instability ( $\ell = 2$ ) are affected by changing the four key parameters of the computation. These are (i) the spacing  $\Delta r$  of the grid in  $r_p$  and  $r_a$ ; (ii) the truncation radius  $R$  of the basis functions (which also sets the upper limit on values of  $r_p$  and  $r_a$ ); (iii) the maximum included value of

**Table 1.** Growth rates for the radial-orbit instability in the isochrone model. We compare values from our standard computation with values from Saha (1991). The notation  $\times_y$  means that we could not recover an unstable mode, and that the value of  $\text{Re}(\det[\mathbf{I} - \xi_{\ell=2}(i\eta)])$  at  $\eta = 0$  was  $y$ .

$R_a/b$	1.0	1.3	1.6	2.0	3.0	4.2
Our $\eta T_1$	0.021	0.009	0.002	$\times_{0.20}$	$\times_{0.43}$	$\times_{0.53}$
Saha linear $\eta T_1$	0.024	0.009	0.004	$< 0.001$	$\ll 0.001$	$\lll 0.001$
Saha N-body $\eta T_1$	0.029	0.015	0.010	–	–	–

**Table 2.** Growth rates for the radial-orbit instability in the isochrone model computed for various values of the parameters that control the precision of the calculations. The tabulated values are of  $\eta T_1$  and the values in bold are from our standard computation. The notation  $\times_y$  means that we could not recover an unstable mode, and that the value of  $\text{Re}(\det[\mathbf{I} - \xi_{\ell=2}(i\eta)])$  at  $\eta = 0$  was  $y$ .

Anisotropy radius $R_a/b$	1.0	1.3	1.6	2.0	3.0	4.2
$\Delta r/b = 0.05$	0.0210	0.0111	$\times_{0.01}$	$\times_{0.19}$	$\times_{0.43}$	$\times_{0.53}$
$\Delta r/b = 0.07$	0.0220	0.0100	$\times_{0.02}$	$\times_{0.19}$	$\times_{0.43}$	$\times_{0.53}$
$\Delta r/b = \mathbf{0.1}$	<b>0.0205</b>	<b>0.0089</b>	<b>0.0016</b>	$\times_{0.20}$	$\times_{0.43}$	$\times_{0.53}$
$\Delta r/b = 0.15$	0.0200	0.0081	0.0068	$\times_{0.17}$	$\times_{0.44}$	$\times_{0.54}$
$\Delta r/b = 0.2$	0.0245	0.0097	0.0057	$\times_{0.22}$	$\times_{0.44}$	$\times_{0.54}$
$\Delta r/b = 0.3$	0.0222	0.0085	$\times_{0.03}$	$\times_{0.21}$	$\times_{0.45}$	$\times_{0.55}$
$R/b = 10$	0.0203	$\times_{0.01}$	$\times_{0.15}$	$\times_{0.27}$	$\times_{0.39}$	$\times_{0.45}$
$R/b = 16$	0.0216	0.0084	0.0016	$\times_{0.22}$	$\times_{0.41}$	$\times_{0.50}$
$R/b = \mathbf{20}$	<b>0.0205</b>	<b>0.0089</b>	<b>0.0016</b>	$\times_{0.20}$	$\times_{0.43}$	$\times_{0.53}$
$R/b = 25$	0.0180	0.0081	0.0026	$\times_{0.17}$	$\times_{0.43}$	$\times_{0.56}$
$n_1^{\max} = 1$	0.0208	0.0089	0.0014	$\times_{0.19}$	$\times_{0.43}$	$\times_{0.54}$
$n_1^{\max} = 2$	0.0204	0.0088	0.0015	$\times_{0.19}$	$\times_{0.43}$	$\times_{0.54}$
$n_1^{\max} = \mathbf{4}$	<b>0.0205</b>	<b>0.0089</b>	<b>0.0016</b>	$\times_{0.20}$	$\times_{0.43}$	$\times_{0.53}$
$n_1^{\max} = 6$	0.0205	0.0089	0.0017	$\times_{0.19}$	$\times_{0.42}$	$\times_{0.53}$
$n_1^{\max} = 8$	0.0206	0.0090	0.0017	$\times_{0.19}$	$\times_{0.42}$	$\times_{0.52}$
$n^{\max} = 8$	0.0182	0.0079	0.0001	$\times_{0.22}$	$\times_{0.47}$	$\times_{0.58}$
$n^{\max} = \mathbf{10}$	<b>0.0205</b>	<b>0.0089</b>	<b>0.0016</b>	$\times_{0.20}$	$\times_{0.43}$	$\times_{0.53}$
$n^{\max} = 13$	0.0220	0.0096	0.0020	$\times_{0.17}$	$\times_{0.38}$	$\times_{0.48}$
$n^{\max} = 16$	0.0224	0.0098	0.0023	$\times_{0.15}$	$\times_{0.34}$	$\times_{0.43}$

the quantum number  $n_1$  associated with  $J_r$ ; (iv) the number  $n^{\max}$  of basis functions summed over (equation (71)). The results for the significantly unstable models,  $R_a = 1$  and 1.3, show gratifyingly little sensitivity to these parameters. The most significant fact is that dropping  $R$  to  $10b$  killed the instability in the model with  $R_a = 1.3b$ . The results for the marginally unstable model  $R_a = 1.6b$  do change significantly with the values of parameters, but clear trends are not evident. It is notable that experimenting with different parameter values does not change the conclusion that the model with  $R_a = 2b$  is stable. No modification produced a curve of  $\text{Re}(\det[\mathbf{I} - \xi_{\ell=2}(i\eta)])$  that differed significantly from those plotted in Fig. 5.

Tables 1 and 2 in conjunction with Fig. 6 give us confidence that the response matrix is being calculated correctly.

### 5.3 Calculation of the Balescu–Lenaard flux

The magnitude of the diffusive flux  $\bar{\mathbf{F}}$  is a direct measure of a system’s relaxation rate: as the relaxation timescale tends to infinity, the diffusive flux tends to zero. Hence we now compute from equations (52) and (53) the diffusive flux in the  $LJ_r$  plane. In Section 6 we will compare this flux with that predicted by classical theory.

The sum over  $\ell$  in equation (53) can be pulled out front, so we obtain the flux

$$\bar{\mathbf{F}} = \sum_{\ell} \bar{\mathbf{F}}_{\ell}, \quad (73)$$

as a sum of contributions from each multipole. For given  $\tilde{\mathbf{n}}, \tilde{\mathbf{n}}'$ , the integrand on the right of equation (53) can be straightforwardly computed from equations (41) and (39), so the only challenge is the evaluation of the integral over  $\tilde{\mathbf{J}}'$ , which has the general form

$$\int d^2\tilde{\mathbf{J}}' g(\tilde{\mathbf{J}}') \delta(h(\tilde{\mathbf{J}}')). \quad (74)$$

Following Fouvry et al. (2015), this may be written

$$\int_C d\sigma(\tilde{\mathbf{J}}') \frac{g(\tilde{\mathbf{J}}')}{|\nabla h(\tilde{\mathbf{J}}')|}, \quad (75)$$

where  $C$  is the curve along which the resonant condition  $0 = h(\tilde{\mathbf{J}}') = \tilde{\mathbf{n}}' \cdot \tilde{\mathbf{\Omega}}' - \tilde{\mathbf{n}} \cdot \tilde{\mathbf{\Omega}}$  is satisfied, and  $d\sigma(\tilde{\mathbf{J}}')$  is the line element on  $C$ . Details of the evaluation of the resulting line integral can be found in Fouvry et al. (2015).

### 5.3.1 Wavevectors to consider

To evaluate the BL flux we have to perform a sum over all possible ‘pairs’  $(\tilde{\mathbf{n}}, \tilde{\mathbf{n}}')$  of two-dimensional vectors with integer components. By a slight abuse of language, we shall refer to objects like  $\tilde{\mathbf{n}}$  as ‘wavevectors’. In principle there are infinitely many of these wavevector pairs, but some of them do not contribute to the flux. To eliminate pairs which do not contribute we first note that for each value of  $\ell$  we can restrict ourselves to pairs for which the quantity  $\Lambda_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}'}^\ell$  defined by equation (41) is non-zero. From the appearance in this equation of  $\tilde{n}_2$  and  $\tilde{n}'_2$  as superscripts of spherical harmonics it is clear that the sum can be restricted to  $|\tilde{n}_2| \leq \ell$  and  $|\tilde{n}'_2| \leq \ell$ . Moreover, the spherical harmonics vanish unless  $\ell - \tilde{n}_2$  and  $\ell - \tilde{n}'_2$  are even. Hence the values to be included are

$$\begin{aligned} \ell = 0 : \quad \tilde{n}_2 &= 0 & \ell = 1 : \quad \tilde{n}_2 &= \pm 1 \\ \ell = 2 : \quad \tilde{n}_2 &= \pm 2, 0 & \ell = 3 : \quad \tilde{n}_2 &= \pm 3, \pm 1 \\ & & \ell = 4 : \quad \tilde{n}_2 &= 0, \pm 2, \pm 4, \end{aligned} \quad (76)$$

and similarly for  $\tilde{n}'_2$ . Unfortunately,  $\tilde{n}_1$  and  $\tilde{n}'_1$  are unrestricted. However one can easily show from the definitions in Section 4 that the pair  $(\tilde{\mathbf{n}}, \tilde{\mathbf{n}}')$  makes exactly the same contribution to the flux as the pair  $(-\tilde{\mathbf{n}}, -\tilde{\mathbf{n}}')$ , halving the computation time. In addition one can show that for  $\ell \leq 2$  (and respecting the rules above) the quantity  $\tilde{\mathbf{n}} \cdot \tilde{\boldsymbol{\Omega}}$  has the same sign throughout action space (equation (70)). Since the resonant condition requires  $\tilde{\mathbf{n}} \cdot \tilde{\boldsymbol{\Omega}} = \tilde{\mathbf{n}}' \cdot \tilde{\boldsymbol{\Omega}}'$ , once  $\tilde{\mathbf{n}}$  has been chosen, only one sign of  $\tilde{n}'_1$  can be of interest, restricting the pair count further. A corresponding statement does not hold for  $\ell \geq 3$ .

If for the sake of definiteness we consider only terms with  $|\tilde{n}_1|, |\tilde{n}'_1| \leq 2$ , then when  $\ell = 0$  and thus  $\tilde{n}_2 = \tilde{n}'_2 = 0$ , we require  $|\tilde{n}_1|, |\tilde{n}'_1| > 0$  and have to include the four relevant pairings from  $(\tilde{n}_1, \tilde{n}'_1) = (1, \pm 1), (1, \pm 2), (2, \pm 1), (2, \pm 2)$ . However, the corresponding pair count rises to 25 for  $\ell = 1$ , to 49 for  $\ell = 2$ , to well over 100 for  $\ell = 3$ , etc. Clearly, increasing the upper bound on values of  $|\tilde{n}_1|$  to be considered leads to still more intimidating pair counts and we must seek to identify the wavevector pairs that make the largest contributions to  $\bar{\mathbf{F}}$ . Since computing time scales as  $(\Delta r)^{-4}$ , we do this by performing a preliminary computation on a coarse grid  $\Delta r \sim 0.25$ .

The  $\ell = 1$  fluxes for models with  $10^5$  stars and  $R_a = 2b$  (left),  $R_a = 4.2b$  (centre) and for the isotropic model (right) are shown in the upper row of Fig. 7. Red arrows show the dressed (self-gravitating) flux, and black arrows show the corresponding bare flux computed by setting the response matrix  $M_{pq} = 0$  (i.e. with self-gravity switched off). These fluxes are obtained by summing over all pairs  $(\tilde{\mathbf{n}}, \tilde{\mathbf{n}}')$  with  $\tilde{n}_1, \tilde{n}'_1 \in [-2, 2]$ . In all three models,  $\bar{\mathbf{F}}_1$  generally shifts stars to lower  $J_r$  but with a characteristic swirling action that causes  $L$  to sometimes decrease and sometimes increase. In anisotropic models this swirling pattern is broken at the right-hand edge of the populated part of the  $LJ_r$  plane, where the restriction  $Q > 0$  gives rise to a cliff-edge in  $\bar{f}(\tilde{\mathbf{J}})$ : stars naturally diffuse over this edge notwithstanding the fact that a small distance to the left of this edge stars are diffusing in almost the opposite direction. Since  $\tilde{\mathbf{n}}$  sets the direction of its contribution to  $\bar{\mathbf{F}}$  (equation (46)), the swirling pattern in the upper panels of Fig. 7 indicates

that the dominant value of  $\tilde{\mathbf{n}}$  changes as one moves through action space.

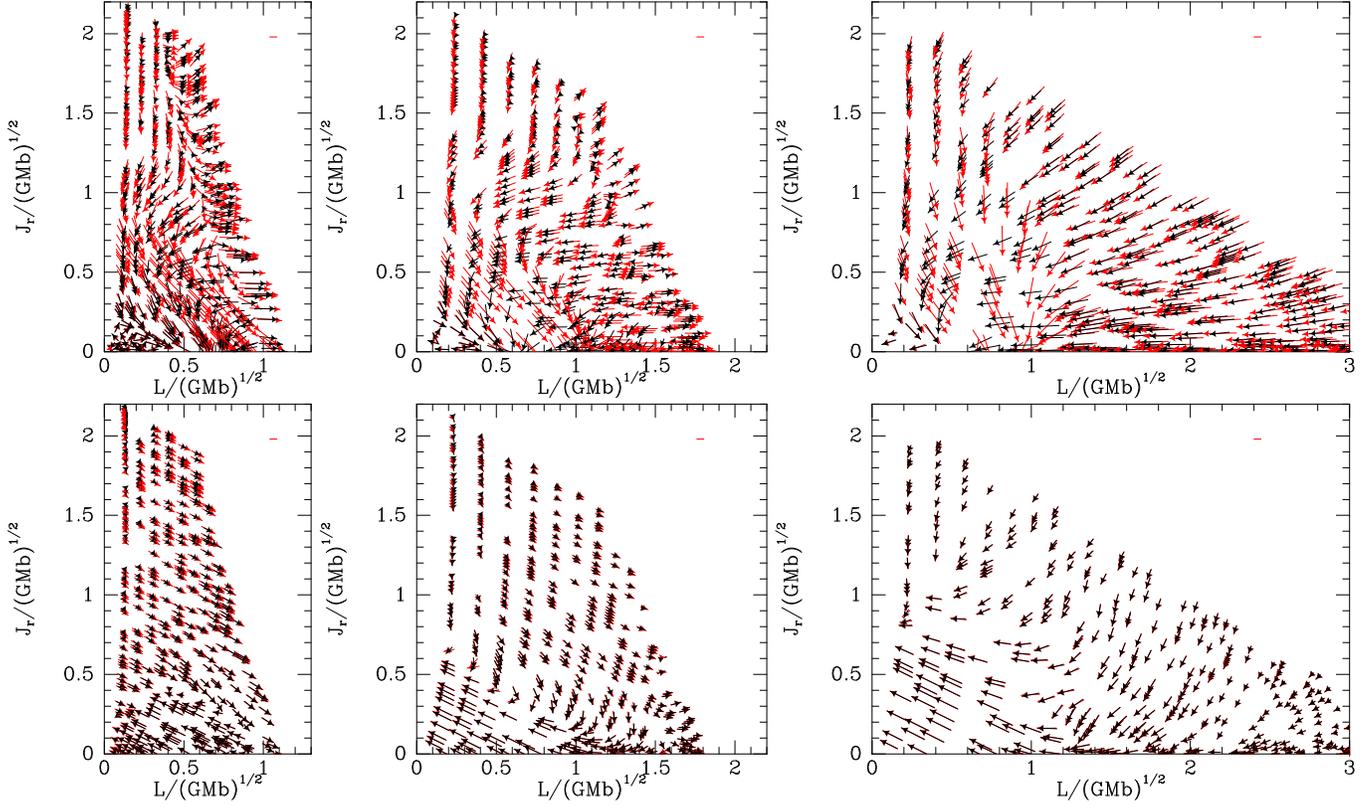
In Fig. 7 bare and dressed fluxes are plotted in black and red arrows, respectively, with black over-plotting red where they coincide. The length of an arrow is proportional to  $\log_{10}[\bar{\mathbf{F}}/(10^{-12} Mb^{-2})]$ . If  $|\bar{\mathbf{F}}| < 10^{-12} Mb^{-2}$  no arrow is plotted, hence the blank regions in the panels. A red line at the upper right of each panel indicates how the length of an arrow increases when  $|\bar{\mathbf{F}}|$  is multiplied by 100. The shortness of this line makes clear the large dynamical range covered by  $\bar{\mathbf{F}}$ , largely as a consequence of the large dynamic range in each model’s phase-space density  $f_0$ . In much of the upper three panels of Fig. 7 the red and black arrows are aligned but the red arrow is significantly longer. In the lower left corner arrows of each colour point in significantly different directions, indicating that the contributions from some vectors  $\tilde{\mathbf{n}}$  are amplified by self-gravity much more than others.

The strong enhancement of  $\bar{\mathbf{F}}_1$  by self-gravity is a natural consequence of the finding of Weinberg (1991) and Weinberg (1994) that the  $\ell = 1$  modes of clusters can be very weakly damped. These modes involve the cluster’s core and halo moving in antiphase along a line, so the linear momenta of the moving parts cancel. The impact of these modes on  $\bar{\mathbf{F}}_1$  is much enhanced by including self-gravity. Displacing the core with respect to the halo obviously becomes harder as the extent of radial anisotropy increases, and careful examination of the arrows in the upper row of Fig. 7 confirms that including self-gravity enhances  $\bar{\mathbf{F}}_1$  less when  $R_a = 2b$  than in the isotropic model. We note that for the secular evolution of discs the tapering of the inner and outer regions that was implemented in Fouvry et al. (2015) prevented any contributions from these  $\ell = 1$  modes.

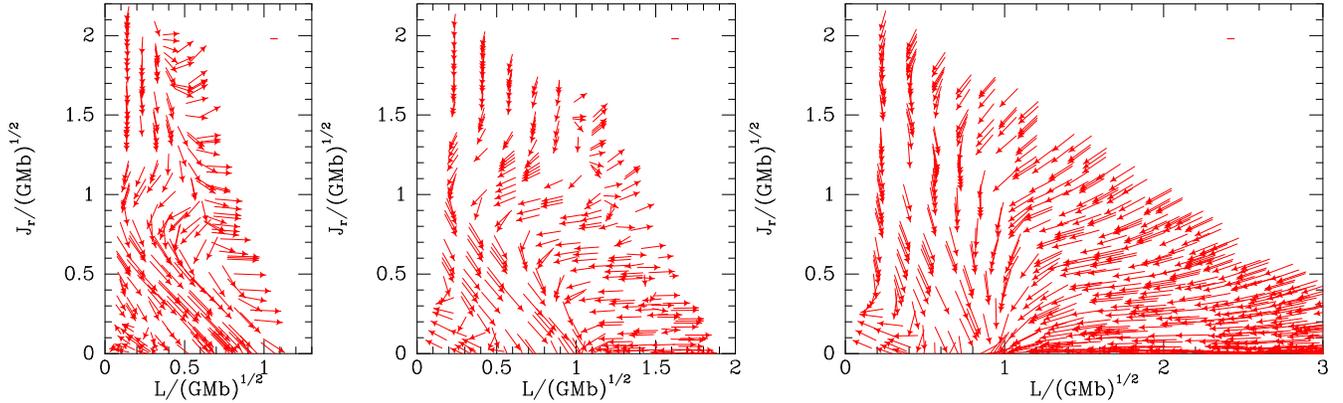
The lower three panels of Fig. 7 show the fluxes  $\bar{\mathbf{F}}_2$ , obtained by adding the contributions of the 49 pairs  $(\tilde{\mathbf{n}}, \tilde{\mathbf{n}}')$  with  $\tilde{n}_1 \in [-2, 2]$  and  $\tilde{n}_2 \in [-2, 2]$ . In these panels red is much less in evidence than in the upper panels, indicating that couplings associated with  $\ell = 2$  are less strongly affected by self-gravity than those associated with  $\ell = 1$ . Red is most evident in the panel for  $R_a = 2b$ , as is to be expected given the vulnerability of this model to quadrupole distortions (May & Binney 1986; Saha 1991). In all three lower panels by far the largest  $\ell = 2$  fluxes occur in the bottom left corner, so physically in the cluster’s core. There the flux shifts stars to larger  $J_r$  and smaller  $L$ . In the nearly unstable model  $R_a = 2b$ , the flux outside this core region moves stars to smaller  $J_r$  and larger  $L$ , and thus tends to reduce the anisotropy that inclines the model to the radial-orbit instability.

Fig. 8 shows sums of the  $\ell = 1$  and  $\ell = 2$  fluxes. In the region of action space that is populated in all three models, the structure of the flow does not differ greatly between models. As  $R_a$  increases, the populated part of phase space extends into a region in which the flow is towards smaller  $L$ .

Away from the cluster cores, the total fluxes are very similar to the  $\ell = 1$  fluxes because the latter are significantly larger than the  $\ell = 2$  fluxes. In the core  $\bar{\mathbf{F}}_2$  dominates and pushes stars to smaller  $L$  and slightly larger  $J_r$ . Around the core  $\bar{\mathbf{F}}_1$  and  $\bar{\mathbf{F}}_2$  are in opposite directions but the net flow reduces  $J_r$ . In the model with  $R_a = 2b$  the flow generally increases  $L$ , thus diminishing the model’s radial anisotropy.



**Figure 7.** Flux vectors  $\bar{\mathbf{F}}_\ell$  in clusters with  $N = 10^5$  equal-mass particles with  $R_a/b = 2$  (left) and 4.2 (centre), and in the isotropic model (right). The fluxes  $\bar{\mathbf{F}}_1$  are shown in the upper row, with the fluxes  $\bar{\mathbf{F}}_2$  shown below. The dressed flow is shown by red arrows on top of which are printed black arrows describing the bare flow. The length of an arrow is proportional to  $\log_{10}[|\bar{\mathbf{F}}|/(10^{-12} Mb^{-2})]$ , so that an arrow of zero length corresponds to  $|\bar{\mathbf{F}}_\ell| = 10^{-12} Mb^{-2}$ . The bar at top right corresponds to a hundred-fold increase in  $|\bar{\mathbf{F}}_\ell|$ .

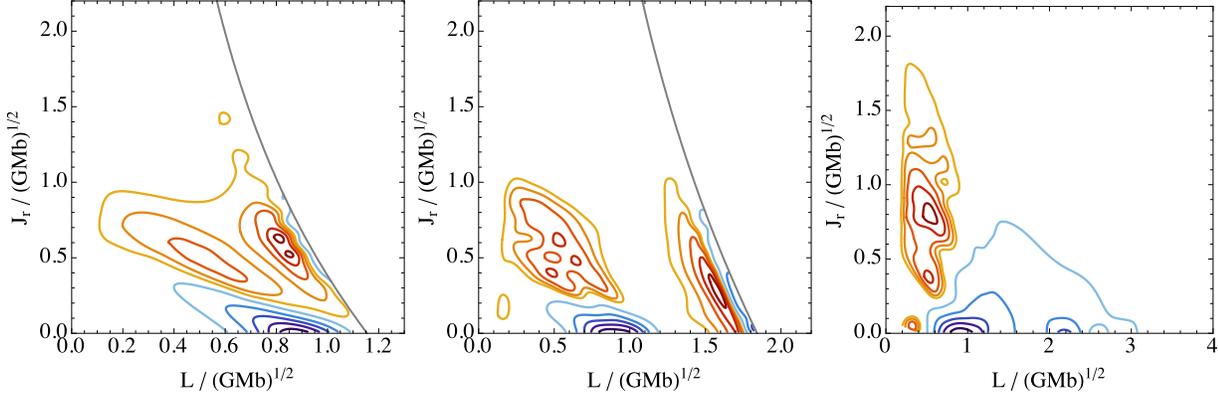


**Figure 8.** The result of adding the  $\ell = 1$  and 2 contributions to  $\bar{\mathbf{F}}$ . As in Fig. 7 panels are for models with  $R_a = 2b$  (left),  $R_a = 4.2b$  (centre) and the isotropic model (right) and the arrows are scaled in the same way.

We find the largest of  $\bar{\mathbf{F}}_0$ ,  $\bar{\mathbf{F}}_3$  and  $\bar{\mathbf{F}}_4$  is always smaller than  $\bar{\mathbf{F}}_2$ , so we do not present our values for them.

Fig. 9 plots  $\partial \bar{f} / \partial t$  for the models with  $R_a/b = 2$  and 4.2 (left and centre) and the isotropic model (right) computed from the divergence of the dressed BL fluxes plotted in Fig. 8. The red contours indicate regions in which  $\bar{f}$  is decreasing, while blue contours mark regions in which it is increasing. The data are quite noisy because they involve differentiation of the numerically computed fluxes. However, in all three models  $\bar{f}$  increases at  $J_r \simeq 0$  and  $L \simeq \sqrt{GMb}$  and

decreases at high  $J_r$  and small  $L$ . In order to illustrate the variety of diffusion features that can generically arise from the BL equation, in Appendix A we briefly consider the case of a tangentially anisotropic DF and illustrate again how self-gravity may impact the properties of the orbital diffusion.



**Figure 9.** The rate of change of the DF,  $\partial\bar{f}/\partial t = -\partial/\partial\tilde{\mathbf{J}} \cdot \bar{\mathbf{F}}$  for the models with  $R_a = 2b$  (left),  $R_a = 4.2b$  (centre) and the isotropic model (right). These rates are minus the divergence of the dressed BL fluxes. Red contours indicate that  $\bar{f}$  is decreasing, blue that it is increasing. The boundary of the allowed action space domain is represented by the heavy black line.

## 6 COMPARISON WITH CLASSICAL THEORY

Let us now compare the results of the BL formalism to those of the classical (Spitzer–Chandrasekhar) theory of local scattering. In Appendix B we show that classical theory predicts

$$\frac{\partial\bar{f}}{\partial t} = -\sum_{i=1}^2 \frac{\partial}{\partial J_i} (\bar{F}_1^i + \bar{F}_2^i), \quad (77)$$

where the fluxes  $\bar{\mathbf{F}}_i$  are given by

$$\begin{aligned} \bar{F}_1^i &= (2\pi)^{-2} \bar{f} \int d\theta_r d\psi \left[ D^i - \frac{1}{2} \Gamma_{jn}^i D^{jn} \right] \\ \bar{F}_2^i &= -\frac{1}{2} (2\pi)^{-2} \frac{\partial}{\partial J_j} \left( \bar{f} \int d\theta_r d\psi D^{ij} \right), \end{aligned} \quad (78)$$

(here we are using the summation convention over repeated indices). Here  $D^i$  and  $D^{jn}$  are rates of change of expectation values that can be obtained from the Rosenbluth potentials (Rosenbluth et al. 1957),  $\Gamma_{jn}^i$  is a Christoffel symbol that emerges from the functional dependence of the actions on velocity, and  $\theta_r$  and  $\psi$  are angles appearing in this dependence.

Evaluation of the Rosenbluth potentials for an anisotropic velocity distribution is costly, so, as is standard practice in studies of globular-cluster evolution (e.g. Cohn 1979; Drukier et al. 1999), we evaluate the Rosenbluth potentials using an equivalent ergodic DF

$$f_{\text{iso}}(E, r) = \int_0^1 dc f \left( E, \frac{r^2 v_E^2}{2R_a^2} (1 - c^2) \right), \quad (79)$$

where  $v_E(r)$  is the speed at  $r$  of a particle of energy  $E$ . After computing the Rosenbluth potentials with  $f_{\text{iso}}$ , the appropriate anisotropic DF  $\bar{f}$  is used in equations (78) for the flux.

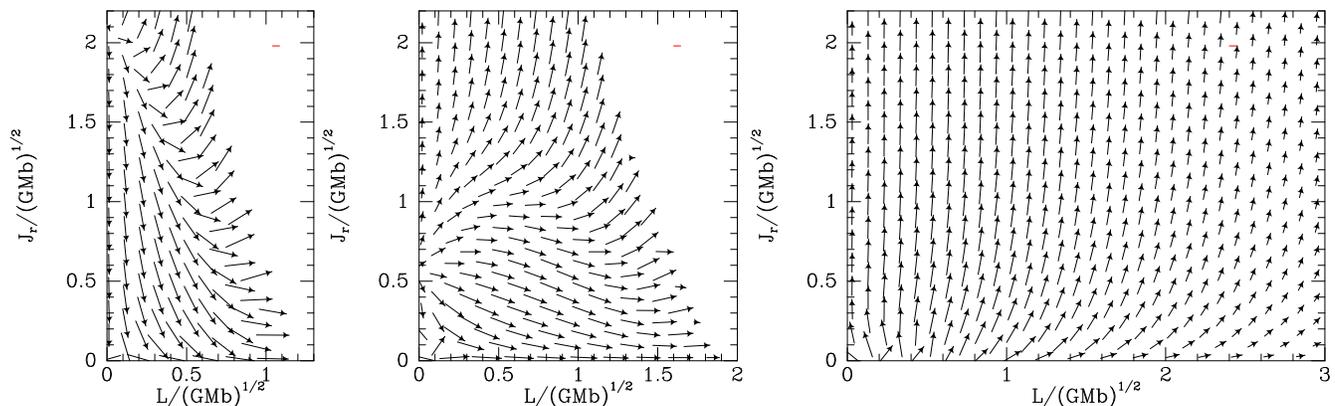
Fig. 10 shows for two finite values of  $R_a$ , namely  $R_a = 2b$  and  $4.2b$ , and for the isotropic model, the resulting classical flux vectors scaled as in Figs. 7 and 8 for a cluster of  $N = 10^5$  equal-mass stars under the assumption that  $\ln \Lambda = 8.9$  (Section B1). The flow is structured by stagnation points. For  $R_a = 2b$  there is a single stagnation point that lies on the  $J_r$  axis at  $J_r \sim 2.0\sqrt{GMb}$ . This point moves down the  $J_r$  axis as  $R_a$  increases, so that by  $R_a = 4.2b$  it is at  $J_r \sim 0.6\sqrt{GMb}$ . The  $R_a = 4.2b$  plot also shows signs of another

stagnation point on the  $L$  axis, roughly at  $L \sim 0.1(GMb)^{1/2}$  (see Appendix B2). For larger  $R_a$  we have a single stagnation point on the  $L$  axis, where it sticks as  $R_a$  tends to infinity. Above the stagnation point and near the  $J_r$  axis the flow is vertically upwards, while below the stagnation point the flow is to lower  $J_r$  and higher  $L$ .

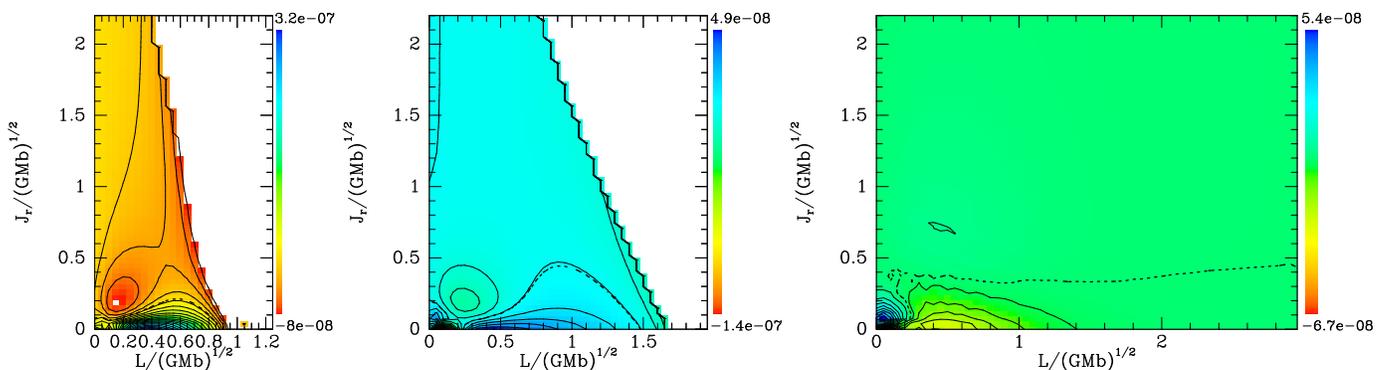
The classical flow pattern for the model with  $R_a = 2b$  is not unlike that obtained from the BL equation (left panel of Fig. 8). Given that in computing the BL fluxes we have only included the  $\ell = 1, 2$  terms and confined ourselves to wave vectors  $\tilde{\mathbf{n}}$  with  $\tilde{n}_1 \in [-2, 2]$ , the extent of the agreement between the two panels on the left of Fig. 8 and the corresponding panels of Fig. 10 is remarkable. As we proceed to larger  $R_a$ , the differences between the classical and the BL flow becomes pronounced. In the isotropic model the BL flow is everywhere towards smaller  $J_r$  whereas the classical flow is towards larger  $J_r$ . The tendency of the classical flow to increase  $J_r$  is consistent with the well known tendency of the halos of clusters to become radially biased.

Fig. 11 shows the rate of change of the DF  $\partial\bar{f}/\partial t$  (in units of  $\sqrt{M/Gb^5}$ ) for the models with  $R_a/b = 2$  and  $4.2$  (left and centre) and the isotropic model (right), all computed from the divergence of the classical flux vectors shown in Fig. 10. In each panel the dashed contour divides the regions in which the DF is increasing from those at which it is decreasing. Whereas in the isotropic model  $\bar{\mathbf{F}}$  tends to shift stars from nearly circular orbits to more eccentric ones, in the highly anisotropic model  $R_a = 2b$  the portion of action space associated with eccentric orbits is depopulated to feed growth in the region associated with nearly circular orbits.

Fig. 12 compares the magnitudes of the BL fluxes in each panel of Fig. 8 to the classical fluxes plotted in Fig. 10. The upper row is for the bare fluxes while the lower row is for dressed fluxes. The colour scales indicate logarithms to base 10 of the ratio of the BL flux to the classical flux. Whereas the bare BL fluxes are almost everywhere significantly smaller than the classical fluxes, outside the core the dressed fluxes are as large as or larger than the classical fluxes. The BL fluxes, both bare and dressed, tend to increase relative to the classical fluxes as  $R_a$  increases.



**Figure 10.** Flux through  $\tilde{\mathbf{J}}$  space,  $\overline{\mathbf{F}}(\tilde{\mathbf{J}})$ , computed using the local approximation for models with  $R_a = 2b$  (left),  $4.2b$  (centre) and the isotropic model (right). The length of an arrow is proportional to  $\log_{10}[|\overline{\mathbf{F}}|/(10^{-12}Mb^{-2})]$ , so that an arrow of zero length corresponds to  $|\overline{\mathbf{F}}| = 10^{-12}Mb^{-2}$  as in Fig. 7. The red bar at top right corresponds to a hundred-fold increase in  $|\overline{\mathbf{F}}|$ .



**Figure 11.** The rate of change of the DF,  $\partial\tilde{f}/\partial t = -\partial/\partial\tilde{\mathbf{J}} \cdot \overline{\mathbf{F}}$ , computed in the local approximation for the model with  $R_a = 2b$  (left),  $R_a = 4.2b$  (centre), and for the isotropic model (right). In each case we plot 20 solid contours, spaced linearly from the minimum to maximum value of  $(\partial\tilde{f}/\partial t)/\sqrt{M/Gb^5}$ . The dashed contours divide regions of increasing and decreasing phase-space density.

## 7 DISCUSSION

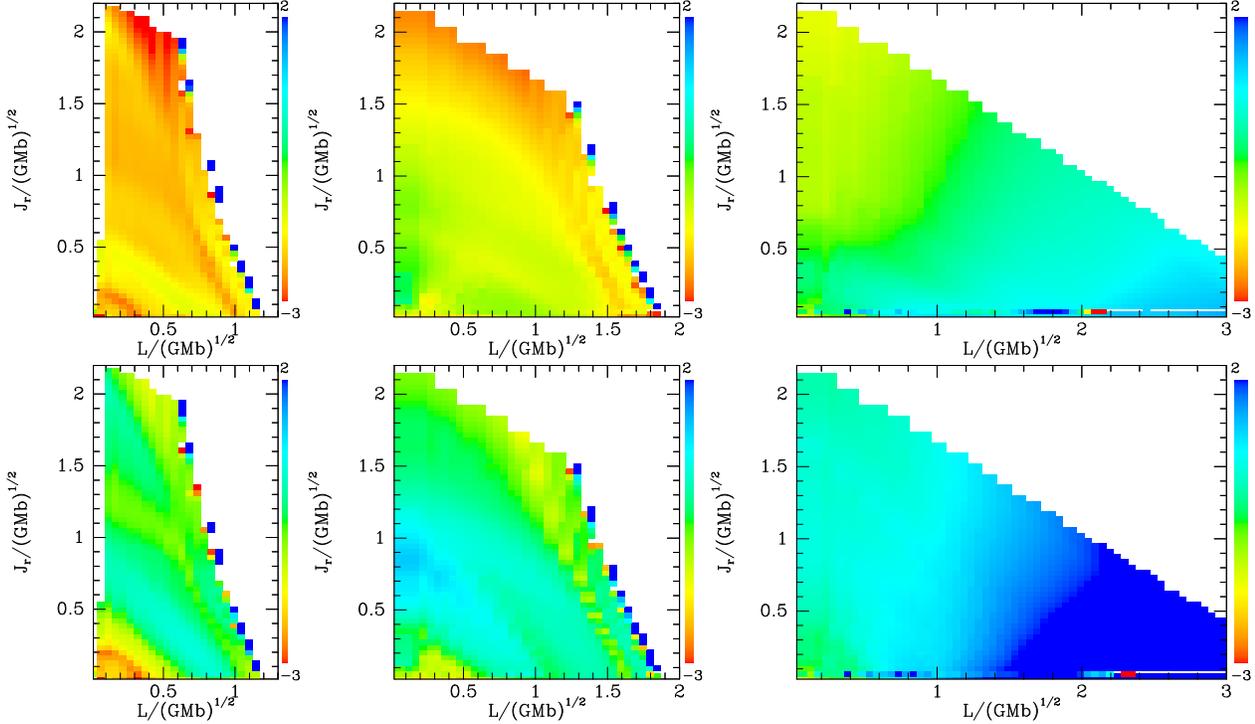
Chavanis (2013a) (see also Chavanis 2013b) has shown that the BL equation yields the classical theory of relaxation when one neglects inhomogeneity and collective effects. Indeed, in the limit of a homogeneous system, the three components of velocity take on the role of actions, the Cartesian coordinates  $x_i$ , which increase uniformly along unperturbed trajectories, play the role of angles, and the sum over discrete vectors  $\mathbf{n}$  becomes an integral over wavevectors. Neglect of collective effects enables one to execute this integral and recover the classical theory. Thus there is little doubt that if we could sum over all resonating pairs  $(\tilde{\mathbf{n}}, \tilde{\mathbf{n}}')$  and all  $\ell$ , we would obtain a flux that embraced the classical flux but went beyond it in that (i) our flux would include the impact of large-scale amplified fluctuations that are not properly handled in the classical theory, and (ii) the flux would not depend on an ill-determined parameter such as  $\log \Lambda$ .

Since collective effects should be unimportant on small scales, it follows from Chavanis' work that for some choice of  $\log \Lambda$ , classical theory should yield a good approximation to the portion of the diffusive flux that arises from encounters at small impact parameters – we shall refer to this as the ‘small-scale’ BL flux. With our current computational machinery it is not feasible to compute the small-scale BL

flux by summing to high  $\ell$  and/or high  $|\tilde{n}_1|$  or  $|\tilde{n}_2|$ . Hence our BL fluxes arise exclusively from the part of the BL sum that returns the flux driven by interactions between stars for which the local approximation of classical theory is certainly invalid.

For our comparison to the classical flux we have used  $N = 10^5$ . The BL flux is proportional to the mass  $\mu$  of a star and therefore scales with the number of particles like  $1/N$  (Section 3). The classical flux is proportional to  $\mu \log \Lambda$  (Appendix B) and so scales like  $\ln N/N$ . Therefore the ratio of BL to classical flux scales with particle number like  $1/\ln N$ .

Given the ill-defined value of the Coulomb logarithm, the magnitude of the classical flux is uncertain. We have adopted  $\log \Lambda = 8.9$ , to get an idea of the magnitude of the small-scale BL flux. Our key finding is that, when the cluster's self-gravity is taken into account, the large-scale BL flux is as large as or even larger than the small-scale flux. Including self-gravity dramatically increases the diffusion generated by thermal excitation of the weakly damped, low frequency  $\ell = 1$  mode of clusters. This finding makes perfect sense physically, because this mode involves the cluster's core and envelope oscillating in antiphase and periodically exchanging momentum. If the stars move in a fixed gravitational field, this mode of communication is impractical



**Figure 12.** Ratios of the magnitudes of BL fluxes to classical fluxes. Upper row: bare fluxes; lower row: dressed fluxes. The left and central panels are for the models with  $R_a = 2b$  and  $4.2b$ , respectively, while the right panel is for the isotropic model. The colour scales shows  $\log_{10}(|\mathbf{F}_{\text{BL}}|/|\mathbf{F}_{\text{Class}}|)$ .

because the core then oscillates at a much higher frequency than the halo, and each section will exchange momentum with the source of the fixed gravitational field rather than the other half of the cluster.

When there are many orbits that move between the core and the halo, it will be harder for the core and the halo to oscillate in antiphase, so the  $\ell = 1$  mode will be harder to excite. Hence for such modes it is natural that the magnitude of the large-scale BL flux relative to that of the small-scale flux is decreased by making the cluster radially anisotropic.

It is well known that as radial anisotropy increases the energy required to make a cluster prolate decreases, and at a critical level of anisotropy it vanishes so the cluster is liable to spontaneous bar formation. Our results confirm the view of May & Binney (1986) that the critical level of anisotropy is reached near  $R_a = 1.7b$ . We find, as expected, that the  $\ell = 2$  contribution to the dressed BL flux associated with quadrupole symmetry increases as  $R_a$  decreases through thermal excitation of the quadrupole mode. We have not, however, computed a model that is sufficiently close to the critical anisotropy for the enhancement by self-gravity to be comparable to that found for the  $\ell = 1$  contribution to the flux in every model.

We have found that contributions to  $\overline{\mathbf{F}}$  from  $\ell = 0, 3, 4$  are much smaller than contributions from  $\ell = 1, 2$ . From these conclusions it does not follow that it was legitimate to neglect these and other  $\overline{\mathbf{F}}_\ell$ , nor does the large magnitude of our dressed BL fluxes justify limiting the sums over wavevectors to small  $\tilde{\mathbf{n}}$ . Indeed, classical theory, which consists of an approximate summation over the neglected terms, strongly suggests that the terms we have dropped are collectively im-

portant. Each wavevector  $\tilde{\mathbf{n}}$  in the sum contributes a flux in its own unique direction. The direction of our BL flux is set by the vectors  $\tilde{\mathbf{n}}$  we have chosen to include, while the classical flux is the result of summing over an infinite number of vectors, and in consequence it can point in a radically different direction, as comparison of the panels on the right of Figs. 8 and 10 shows to be the case in the isotropic model.

This discussion suggests that the partial BL summation on the one hand and the classical theory on the other hand provide distinct approximations to the flux. A summation over the leading discrete  $\tilde{\mathbf{n}}$ , combined with summation over a myriad of other  $\tilde{\mathbf{n}}$  that make small individual contributions to the flux may eventually prove sufficient (Weinberg 1986). Most of these small contributions are expected to come from small-scale fluctuations in the potential that are probably adequately modelled by the orbit-averaged classical theory because neither inhomogeneity nor collective effects are important on small scales. Consequently, a proper procedure may involve computing the dominant terms in the dressed BL expressions for the diffusion coefficients and approximating the remaining infinite sum by the classical diffusion coefficients using a smaller value of  $\log \Lambda$ . The remaining problems are (i) expediting the computation of the BL terms, which is unreasonably costly with our current code and (ii) estimating the correct value of  $\log \Lambda$ .

## 7.1 N-body simulations

If classical theory is so misleading, why have N-body simulations of cluster relaxation not exposed this fact? A likely answer has two parts. First, classical theory has the free pa-

parameter  $\log \Lambda$ , and second analysis of N-body simulations is hard and the action-space flux  $\bar{\mathbf{F}}$  is rarely computed. Rather the rate at which the model's radial density profile evolves is computed, and  $\log \Lambda$  is adjusted to make the classically predicted rate agree with the measured rate. Theuns (1996) did compute the rate at which particles in N-body clusters (in this case isotropic King models) diffused in energy. He concluded that on average they diffused about twice as fast as classical theory predicted, and that outside the core classical theory underestimated the size of the dynamical friction coefficient by up to a factor of 10. These conclusions are roughly in agreement with the lower-right panel of our Fig. 12. Additionally, the excess power due to large scale modes has been seen in N-body simulations (for example those of Weinberg (1998); see also Weinberg (2001)). The patterns of the weakly damped modes themselves might simply be too hard to see in real space because they have such low frequencies and are soon washed out by noise (Weinberg 2001). Sellwood (2015) performed N-body simulations of homogeneous spheres and found a scaling of the energy relaxation with  $N$  which strongly suggested the dominance of large-scale collective modes driven by Poisson noise. On the other hand, Sellwood (2015) also simulated Hernquist and Plummer spheres and found fairly good agreement with the classical picture. The changes to the relaxation rate that can result from tuning  $\log \Lambda$  to any sensible value are typically only  $\sim 10 - 20\%$  and so this alone cannot account for the consistent agreement which has been found between N-body and Fokker-Planck models. Finally we note that Kim et al. (2008) measured a significantly smaller discrepancy in energy diffusion rate from the classical expectation than Theuns (1996).

The validity of our conclusions should be tested by direct N-body simulations of clusters. A suitable programme would comprise an ensemble of short simulations of isochrone clusters. Individual simulations could be integrated for times significantly shorter than the half-mass relaxation time, so their mean-field potentials would remain close to that of the isochrone, and the actions of particles could be readily computed. By comparing actions computed at successive timesteps, fluxes across a grid of lines in  $LJ_r$  space could be determined. Moreover, the difference between the actual potential and the isochrone potential, projected onto a set of basis functions  $\Phi_{n\ell m}$  could be monitored so one could estimate how much of the flux is generated by large-scale fluctuations that are significantly enhanced by self-gravity and cannot be well described by classical theory. Good statistics would be accumulated by stacking results from simulations that differed only in their initial conditions. We plan to present the results of such a study in the near future.

## 8 CONCLUSIONS

The classical theory of relaxation in star clusters has for decades formed one of the pillars of stellar dynamics. Yet it is well known to rest on shaky foundations in that its predictions all involve integrals that diverge logarithmically at large impact parameters. Conventionally this divergence is mastered by imposing an upper limit on impact param-

eters. Predictions then become proportional to the logarithm of this cutoff, so the theory has a free parameter.

A further weakness of the classical theory is that it neglects the cluster's self-gravity. In Section 2 we gave a back-of-the-envelope argument that by accounting for Poisson fluctuations alone one can recover the classical expression for the relaxation time. Since self-gravity makes a star cluster more compressible than an ideal gas it follows that slow, large-scale fluctuations in density are liable to drive a cluster to relax faster than classical theory predicts.

In two important papers Weinberg (1993, 1998) presented arguments that the self-gravity significantly speeds the relaxation of stellar systems through excitation of large-scale modes. However, he did not provide an equation for the secular evolution of the DF when large-scale modes and self-gravity are properly handled. Such an equation, derived from a rigorous, self-consistent theory of relaxation that contains no arbitrary constants, appeared first in Heyvaerts (2010). The structure of this 'BL equation' suggests a physical picture of how systems relax that is radically different from the classical picture of successive stellar encounters. In Section 3 we summarised this theory, and in Section 4 we modified it so it could be applied to spherical systems, which have a degeneracy that precludes application of the unmodified equation.

According to the new theory, relaxation does not occur through localised encounters between stars but through pairs of stars communicating with each other by shaking the entire system at a specific frequency  $\omega$ . The stars do not need to be near each other in real space, although proximity will make communication easier. The effectiveness of communication between stars depends also on whether they can communicate at a frequency that lies near a weakly damped mode of the entire system. Self-gravity is important because it determines the modal structure. Fouvry et al. (2015) applied the new theory to stellar discs and showed that it could for the first time explain the emergence of strong spiral structure in N-body simulations of linearly stable discs. In particular, they showed that when self-gravity is properly included, realistic stellar discs can relax  $\sim 1000$  times faster than classical theory predicts. Moreover, through a mechanism proposed by Sellwood & Carlberg (2014), the resonant nature of the disc's relaxation steers the disc to a configuration in which it is unstable as a collisionless system, so it eventually develops a strong bar on a dynamical timescale.

In Section 5 we put forward the first implementation of the BL theory to three spherical clusters that all have the spatial structure of Hénon's isochrone but distribution functions with different degrees of radial velocity anisotropy. The problem one encounters is the need to sum over infinitely many resonant interactions, each defined by a pair of two-dimensional vectors  $\tilde{\mathbf{n}}, \tilde{\mathbf{n}}'$  with integer components. We tried to identify, and then sum, the interactions that make the largest contributions. The sum can be partitioned into contributions from each angular momentum quantum number  $\ell$ , and we considered  $\ell = 0, \dots, 4$ . We found that the sums for  $\ell = 1, 2$  dominate strongly. However, the number of vector pairs that could contribute significantly grows rapidly with  $\ell$  and the size of the vectors, and it is possible for some configurations that large numbers of interactions that individually contribute little together contribute as much as the most important interactions (Weinberg 1986). This be-

ing so, our partial sum may under-estimate the magnitude of the BL flux through phase space. In all three clusters, the partial BL fluxes are much larger when self-gravity is included than otherwise. This is to be expected because the terms we have included arise from large-scale distortions of the cluster, which are precisely the fluctuations enhanced by self-gravity.

In Appendix B we recovered the fluxes predicted by classical theory. In Section 6 we compared these fluxes with our partial BL fluxes. In the isotropic cluster, the fluxes have quite different directions, and the classical flux is often larger than the bare partial BL flux, consistent with our having missed the contributions of myriads of interactions that individually contribute little. The dressed partial BL fluxes are, by contrast, often larger than the classical flux, suggesting that the classical flux significantly under-estimates the rate of relaxation. The dominance of the dressed BL flux over the classical flux is most pronounced in the isotropic cluster and at large  $L$ .

We conclude that classical theory is liable to under-estimate the rate of relaxation through neglect of self-gravity and non-local resonances. Whereas this can lead to the relaxation rate of a stellar disc being under-estimated by a factor  $\sim 1000$ , for a spherical system the magnitude of the relaxation rate predicted by (large scale) BL theory is typically  $\sim$  a few times that of classical theory (but is often in a completely different direction and can be significantly larger depending on the anisotropy of the cluster). However, brute-force application of the BL equation may not yet provide a fully viable alternative to classical theory because of the evaluation of the infinite sum over interaction vectors. A way must be found to combine a sum over the most important resonance vectors, which are associated with large-scale fluctuations and the important effects of self-gravity, with an integral like that executed in classical theory to include the effects of small-scale fluctuations, to which self-gravity contributes little. However, we should first confirm our finding that classical theory misses out some important physics through analysis of tailored N-body simulations.

## ACKNOWLEDGEMENTS

JJB is supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 321067. JBF acknowledges support from Program number HST-HF2-51374 which was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy Incorporated, under NASA contract NAS5-26555. This research was in part carried out within the framework of the Spin(e) collaboration (ANR-13-BS05-0005 <http://cosmicorigin.org>)

## REFERENCES

- An J. H., Evans N. W., 2006, *AJ*, 131, 782  
 Balescu R., 1960, *Physics of Fluids*, 3, 52  
 Binney J., Tremaine S., 2008, *Galactic Dynamics: Second Edition*. Princeton University Press

- Carroll S. M., 2004, *Spacetime and geometry. An introduction to general relativity*  
 Chandrasekhar S., 1943, *ApJ*, 97, 255  
 Chandrasekhar S., 1949, *Reviews of Modern Physics*, 21, 383  
 Chavanis P.-H., 2012, *Physica A*, 391, 3680  
 Chavanis P.-H., 2013a, *European Physical Journal Plus*, 128, 126  
 Chavanis P.-H., 2013b, *A&A*, 556, A93  
 Clutton-Brock M., 1973, *ApSpSc*, 23, 55  
 Cohn H., 1979, *ApJ*, 234, 1036  
 Drukier G. A., Cohn H. N., Lugger P. M., Yong H., 1999, *The Astrophysical Journal*, 518, 233  
 Edmonds A., 1996, *Angular Momentum in Quantum Mechanics*. Princeton University Press  
 Fouvry J.-B., Pichon C., Magorrian J., 2017, *A&A*, 598, A71  
 Fouvry J. B., Pichon C., Magorrian J., Chavanis P. H., 2015, *A&A*, 584, A129  
 Fridman A. M., Poliachenko V. L., 1984, *Physics of gravitating systems. II*. Springer-Verlag Berlin Heidelberg  
 Henon M., 1959, *Annales d'Astrophysique*, 22, 126  
 Hénon M., 1960, *Annales d'Astrophysique*, 23, 474  
 Hernquist L., Ostriker J. P., 1992, *ApJ*, 386, 375  
 Heyvaerts J., 2010, *MNRAS*, 407, 355  
 Kalnajs A. J., 1976, *ApJ*, 205, 745  
 Kim E., Yoon I., Lee H. M., Spurzem R., 2008, *Monthly Notices of the Royal Astronomical Society*, 383, 2  
 Lenard A., 1960, *Annals of Physics*, 10, 390  
 May A., Binney J., 1986, *MNRAS*, 221, 13P  
 Palmer P. L., Papaloizou J., 1987, *MNRAS*, 224, 1043  
 Rahmati A., Jalali M. A., 2009, *MNRAS*, 393, 1459  
 Rauch K. P., Tremaine S., 1996, *NAstr*, 1, 149  
 Rosenbluth M. N., MacDonald W. M., Judd D. L., 1957, *Physical Review*, 107, 1  
 Saha P., 1991, *MNRAS*, 248, 494  
 Sellwood J. A., 2015, *MNRAS*, 453, 2919  
 Sellwood J. A., Carlberg R. G., 2014, *ApJ*, 785, 137  
 Theuns T., 1996, *MNRAS*, 279, 827  
 Tremaine S., Weinberg M. D., 1984, *MNRAS*, 209, 729  
 Weinberg M. D., 1986, *ApJ*, 300, 93  
 Weinberg M. D., 1989, *MNRAS*, 239, 549  
 Weinberg M. D., 1991, *ApJ*, 368, 66  
 Weinberg M. D., 1993, *ApJ*, 410, 543  
 Weinberg M. D., 1994, *ApJ*, 421, 481  
 Weinberg M. D., 1998, *MNRAS*, 297, 101  
 Weinberg M. D., 2001, in Woodward C. E., Bica M. D., Shull J. M., eds, *Tetons 4: Galactic Structure, Stars and the Interstellar Medium Vol. 231 of Astronomical Society of the Pacific Conference Series, The Milky Way as a Key to Structural Evolution in Galaxies*. p. 53

## APPENDIX A: TANGENTIALLY BIASED DISTRIBUTION FUNCTIONS

The self-consistent isochrone DF introduced in equation (56) is radially anisotropic. Following An & Evans (2006), a tangentially biased DF for the isochrone potential is given by

$$f_0(E, L) = p(E) + q(E)L^2, \quad (\text{A1})$$

where

$$p(E) = \frac{3M(1-E')^{-4}}{2^{15/2}\pi^3(GMb)^{3/2}} \left( (16E'^2 - 12E' + 3) \frac{15 \sin^{-1} \sqrt{E'}}{\sqrt{1-E'}} \right. \\ \left. - (45 - 150E' + 144E'^2 - 208E'^3 + 64E'^4) \sqrt{E'} \right),$$

and

$$q(E) = \frac{3M(1-E')^{-5}}{2^{15/2}\pi^3(GMb)^{3/2}} \left( (40E'^2 - 24E' + 5) \frac{15 \sin^{-1} \sqrt{E'}}{\sqrt{1-E'}} \right. \\ \left. - (75 - 310E' + 400E'^2 - 928E'^3 + 576E'^4 - 128E'^5) \sqrt{E'} \right),$$

and  $E' \equiv -bE/GM$ . As emphasised in the BL equation (53), we may consider the averaged DF,  $\bar{f}(J_r, L) = 2Lf_0$ . In Fig. A1, we illustrate the respective contributions of the  $\ell = 1$  and  $\ell = 2$  to the diffusion flux, with and without collective effects.

In the first row of Fig. A1, we note the importance of collective effects in amplifying the flux and altering its direction for the  $\ell = 1$  component, as induced by the existence of a weakly damped  $\ell = 1$  mode in these systems (Weinberg 1994). In particular, in the top-right panel, the diffusion takes place predominantly along the direction  $(-1, 1)$ . In the tangentially anisotropic model, stars are being picked up and moved along the  $(-1, 1)$  line. Indeed, for perfectly circular orbits, one has  $(-1, 1) \cdot \tilde{\boldsymbol{\Omega}} = 0$ . As there are numerous almost-circular orbits in this model, many stars will have a characteristic frequency  $\tilde{\mathbf{n}} \cdot \tilde{\boldsymbol{\Omega}} \simeq 0$ . Hence, many pairs of stars will have matching frequencies, allowing them to interact resonantly. The direction of the flux is then along the the most strongly-amplified resonance vector  $\tilde{\mathbf{n}}$ , which in this case is  $(-1, 1)$ . On the contrary, in the bottom row of Fig. A1, we note that collective effects do not boost the amplitude of the  $\bar{\mathbf{F}}_2$  flux, because the tangential model has so few radial orbits, i.e. so few orbits prone to undergo a strong self-gravitating amplification via the  $\ell = 2$  response.

## APPENDIX B: CLASSICAL DIFFUSION COEFFICIENTS

We use classical Spitzer–Chandrasekhar theory to compute the diffusive action-space flux. Our strategy is as follows. At each phase of an orbit of given  $(J_r, L)$  we compute the probability flux and then average this flux over orbital phase. We obtain the flux at a given location by transforming the classical velocity-space Fokker-Planck equation to curvilinear coordinates for velocity-space that include  $J_r$  and  $L$ . The requisite tensor calculus is found in e.g. Carroll (2004).

We use as our velocity-space coordinates  $\mathbf{P} = (J_r, L, \psi)$ , where  $\psi$  is the angle between the tangential component of the velocity and a fixed direction  $\mathbf{e}_x$  in the plane perpendicular to the current position vector  $\mathbf{r}$ . In terms of the new velocity coordinates the Cartesian components of velocity are

$$(v_r, v_x, v_y) = \left( \sqrt{2(E - \Phi) - L^2/r^2}, \frac{L}{r} \cos \psi, \frac{L}{r} \sin \psi \right),$$

where  $E = E(J_r, L)$ , so

$$\mathbf{L} = (0, -L \sin \psi, L \cos \psi). \quad (\text{B1})$$

Given that  $dE = \Omega_r dJ_r + \Omega_\vartheta dL$  it follows that

$$dv_r = \frac{\Omega_r dJ_r + (\Omega_\vartheta - L/r^2) dL}{v_r}. \quad (\text{B2})$$

Hence,

$$\frac{\partial \mathbf{v}}{\partial \mathbf{P}} = \begin{pmatrix} \frac{\Omega_r}{v_r} & \frac{\Omega_\vartheta - L/r^2}{v_r} & 0 \\ 0 & \frac{\cos \psi}{r} & -v_t \sin \psi \\ 0 & \frac{\sin \psi}{r} & v_t \cos \psi \end{pmatrix}. \quad (\text{B3})$$

To compute the inverse matrix we note that in any potential we have

$$\frac{\partial L}{\partial \mathbf{v}} = \frac{\mathbf{L} \times \mathbf{r}}{L} \quad \text{and} \quad \frac{\partial \psi}{\partial \mathbf{v}} = \frac{r}{L^2} \mathbf{L}, \quad (\text{B4})$$

and in the case of the isochrone there is a simple expression for  $\partial J_r / \partial \mathbf{v}$ . By applying the operator  $\partial / \partial \mathbf{v}$  to

$$J_r = \frac{GM}{\sqrt{-2H}} - \frac{1}{2} (L + \sqrt{L^2 + 4GMb}), \quad (\text{B5})$$

and using  $\partial H / \partial \mathbf{v} = \mathbf{v}$  and equation (B4) one easily shows that

$$\frac{\partial J_r}{\partial \mathbf{v}} = \frac{\mathbf{v} - \Omega_\vartheta \mathbf{L} \times \mathbf{r} / L}{\Omega_r}. \quad (\text{B6})$$

In matrix notation

$$\frac{\partial P^i}{\partial v^j} = \begin{pmatrix} \frac{v_r}{\Omega_r} & \frac{1}{\Omega_r} \left( v_x - \frac{\Omega_\vartheta L_y r}{L} \right) & \frac{1}{\Omega_r} \left( v_y + \frac{\Omega_\vartheta L_x r}{L} \right) \\ 0 & L_y r / L & -L_x r / L \\ 0 & L_x r / L^2 & L_y r / L^2 \end{pmatrix} \\ = \begin{pmatrix} \frac{v_r}{\Omega_r} & \frac{1}{\Omega_r} \left( v_x - \frac{\Omega_\vartheta L_y r}{L} \right) & \frac{1}{\Omega_r} \left( v_y + \frac{\Omega_\vartheta L_x r}{L} \right) \\ 0 & r \cos \psi & r \sin \psi \\ 0 & -(r/L) \sin \psi & (r/L) \cos \psi \end{pmatrix}.$$

The invariant distance in velocity space is

$$ds^2 = dv_r^2 + dv_x^2 + dv_y^2 \\ = \frac{\Omega_r^2}{v_r^2} dJ_r^2 + \left[ \frac{(\Omega_\vartheta - L/r^2)^2}{v_r^2} + \frac{1}{r^2} \right] dL^2 \\ + \frac{2\Omega_r}{v_r^2} (\Omega_\vartheta - L/r^2) dJ_r dL + \frac{L^2}{r^2} d\psi^2 \quad (\text{B7})$$

so the non-vanishing components of the metric are

$$g_{JJ} = \frac{\Omega_r^2}{v_r^2}, \quad g_{LL} = \left[ \frac{(\Omega_\vartheta - L/r^2)^2}{v_r^2} + \frac{1}{r^2} \right], \\ g_{LJ} = \frac{\Omega_r}{v_r^2} (\Omega_\vartheta - L/r^2), \quad g_{\psi\psi} = \frac{L^2}{r^2}. \quad (\text{B8})$$

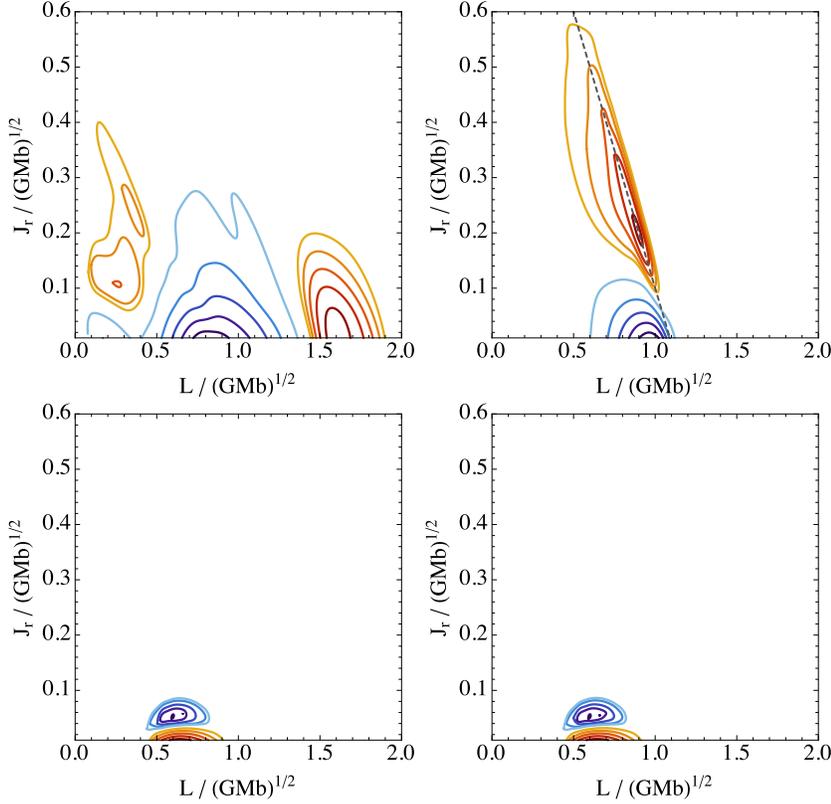
We will need

$$g \equiv \det g_{ij} \\ = \frac{L^2}{r^2} \left\{ \frac{\Omega_r^2}{v_r^2} \left[ \frac{(\Omega_\vartheta - L/r^2)^2}{v_r^2} + \frac{1}{r^2} \right] - \left[ \frac{\Omega_r}{v_r^2} (\Omega_\vartheta - L/r^2) \right]^2 \right\} \\ = \frac{L^2 \Omega_r^2}{r^4 v_r^2}. \quad (\text{B9})$$

In Cartesian velocity coordinates the Fokker-Planck equation is

$$\frac{\partial f}{\partial t} = -\partial_i \left[ \langle \delta v^i \rangle_\mathbf{f} f - \frac{1}{2} \partial_j \langle \delta v^i \delta v^j \rangle_\mathbf{f} f \right], \quad (\text{B10})$$

where the Einstein summation convention is used,  $\partial_i \equiv \partial / \partial v^i$ , and we use superscripts for the indices of velocity components in preparation for the introduction of a non-trivial metric. To obtain the corresponding equation in



**Figure A1.** The rate of change of the DF  $\partial \bar{f} / \partial t = -\partial \bar{\mathbf{F}}_\ell / \partial \tilde{\mathbf{J}}$  for the tangentially anisotropic DF from equation (A1), for  $\ell = 1$  (top row) and  $\ell = 2$  (bottom row), without collective effects (left column) and with collective effects (right column). The convention used is the same as in Fig. 9. In the top-right panel, the dashed gray line is aligned with direction  $(-1, 1)$ .

curvilinear coordinates  $P^i$ , we have to transform the vector  $\langle \delta v^i \rangle_f$  and the tensor  $\langle \delta v^i \delta v^j \rangle_f$  according to

$$\begin{aligned} D^i &\equiv \frac{\partial P^i}{\partial v^j} \langle \delta v^j \rangle_f \\ D^{ij} &\equiv \frac{\partial P^i}{\partial v^m} \frac{\partial P^j}{\partial v^n} \langle \delta v^m \delta v^n \rangle_f \end{aligned} \quad (\text{B11})$$

and to convert the partial derivatives in equation (B10) into covariant derivatives. We exploit the standard results

$$\begin{aligned} \nabla_k A^k &= \frac{1}{\sqrt{g}} \partial_k (\sqrt{g} A^k) \\ \nabla_k B^{kn} &= \frac{1}{\sqrt{g}} \partial_k (\sqrt{g} B^{kn}) + \Gamma_{kj}^n B^{kj}, \end{aligned} \quad (\text{B12})$$

to deduce that in a general coordinate system  $f$  obeys (Rosenbluth et al. 1957)

$$\begin{aligned} \frac{\partial f}{\partial t} &= -\frac{1}{\sqrt{g}} \partial_i \left\{ \sqrt{g} \left[ D^i f - \frac{1}{2\sqrt{g}} \partial_j (\sqrt{g} D^{ij} f) - \frac{1}{2} \Gamma_{jn}^i D^{jn} f \right] \right\} \\ &= -\frac{1}{\sqrt{g}} \partial_i \left\{ \sqrt{g} \left[ D^i - \frac{1}{2} \Gamma_{jn}^i D^{jn} \right] f - \frac{1}{2} \partial_j (\sqrt{g} D^{ij} f) \right\}. \end{aligned} \quad (\text{B13})$$

When this equation is multiplied by the infinitesimal phase-space volume  $d^3 \mathbf{x} d^3 \mathbf{v} = \sqrt{g} d^3 \mathbf{x} d^3 \mathbf{P}$ , the left side becomes the rate of change of the stellar mass in that element. Integrating this over all variables except  $J_r$  and  $L$  we obtain the rate of change of stellar mass with actions within  $(dJ_r dL)$

of  $(J_r, L)$ :

$$\begin{aligned} \frac{\partial N(J_r, L)}{\partial t} &= \int d^3 \mathbf{x} d\psi \sqrt{g} \frac{\partial f}{\partial t} \\ &= - \int d^3 \mathbf{x} d\psi \partial_i \left\{ \sqrt{g} \left[ D^i - \frac{1}{2} \Gamma_{jn}^i D^{jn} \right] f \right. \\ &\quad \left. - \frac{1}{2} \partial_j (\sqrt{g} D^{ij} f) \right\}. \end{aligned} \quad (\text{B14})$$

Since  $\partial_3 = \partial_\psi$  and we are integrating  $\psi$  from 0 to  $2\pi$ , the sum over  $i$  can be restricted to  $i = 1, 2$ . So, taking the remaining differentials out of the integral, we can rewrite equation (B14) in the form

$$\frac{\partial N}{\partial t} = -(2\pi)^3 \sum_{i=1}^2 \partial_i (\bar{F}_1^i + \bar{F}_2^i), \quad (\text{B15})$$

where we have broken the flux in  $(J_r, L)$  space into two parts

$$\begin{aligned} \bar{F}_1^i &\equiv (2\pi)^{-3} f \int d^3 \mathbf{x} d\psi \sqrt{g} \left[ D^i - \frac{1}{2} \Gamma_{jn}^i D^{jn} \right] \\ \bar{F}_2^i &\equiv -\frac{1}{2} (2\pi)^{-3} \partial_j \left( f \int d^3 \mathbf{x} d\psi \sqrt{g} D^{ij} \right). \end{aligned} \quad (\text{B16})$$

and the sum over  $j$  can be restricted to  $j = 1, 2$ .

We obtain an alternative expression for  $N(\tilde{\mathbf{J}})$  by integrating  $f$  over  $\boldsymbol{\theta}$  and  $L_z$  and recalling the definition (49):

$$N(\tilde{\mathbf{J}}) = (2\pi)^3 2L f(\tilde{\mathbf{J}}) = (2\pi)^3 \bar{f}(\tilde{\mathbf{J}}). \quad (\text{B17})$$

Eliminating  $N$  between equations (B14) and (B17) we have

$$\frac{\partial \bar{f}}{\partial t} = - \sum_{i=1}^2 \partial_i (\bar{F}_1^i + \bar{F}_2^i). \quad (\text{B18})$$

We use spherical polar coordinates to execute the spatial integrals in the definition of  $\bar{\mathbf{F}}_\alpha$ . Since the system is spherically symmetric, the angular integrals produce a factor  $4\pi$ . The radial integrals run between the peri- and apo-centre implied by  $(J_r, L)$  and are conveniently turned into integrals over radial phase using  $dr = v_r d\theta_r / \Omega_r$ . Using  $\sqrt{g} = L\Omega_r / (r^2 v_r)$  from equation (B9), we then find

$$\begin{aligned} \bar{F}_1^i &= (2\pi)^{-2} \bar{f} \int d\theta_r d\psi \left[ D^i - \frac{1}{2} \Gamma_{jn}^i D^{jn} \right] \\ \bar{F}_2^i &= -\frac{1}{2} (2\pi)^{-2} \partial_j \left( \bar{f} \int d\theta_r d\psi D^{ij} \right). \end{aligned} \quad (\text{B19})$$

From Binney & Tremaine (2008) §7.4.4 we have

$$\begin{aligned} \langle \delta \mathbf{v} \rangle_{\mathbf{f}} &= 4\pi G(m + m') \log \Lambda \frac{\partial h}{\partial \mathbf{v}} \\ \langle \delta v_i \delta v_j \rangle_{\mathbf{f}} &= 4\pi G m' \log \Lambda \frac{\partial^2 g}{\partial v_i \partial v_j}, \end{aligned} \quad (\text{B20})$$

where  $\log \Lambda$  is the Coulomb logarithm (Section B1) and the Rosenbluth potentials are

$$\begin{aligned} h(\mathbf{v}) &\equiv Gm \int d^3 \mathbf{v}' \frac{n(\mathbf{v}')}{|\mathbf{v} - \mathbf{v}'|} \\ g(\mathbf{v}) &\equiv Gm \int d^3 \mathbf{v}' n(\mathbf{v}') |\mathbf{v} - \mathbf{v}'|. \end{aligned} \quad (\text{B21})$$

Here  $\int d^3 \mathbf{v} n(\mathbf{v})$  is the number density of stars in real space. If we restrict ourselves to an ergodic DF, so

$$n(\mathbf{v}') = n(|\mathbf{v}'|) = f(H)/m \quad (\text{B22})$$

the Rosenbluth potentials only depend on the magnitude of  $\mathbf{v}$ . Integrating over all directions of  $\mathbf{v}'$  we obtain

$$\begin{aligned} h(v) &= 4\pi G \left\{ \frac{1}{v} \int_0^v dv' v'^2 f(v') + \int_v^\infty dv' v' f(v') \right\} \\ g(v) &= \frac{4\pi G}{3} \left\{ \int_0^v dv' f(v') \frac{v'^2}{v} (v'^2 + 3v^2) \right. \\ &\quad \left. + \int_v^\infty dv' f(v') v' (v^2 + 3v'^2) \right\}. \end{aligned} \quad (\text{B23})$$

Hence the averages over field stars are

$$\begin{aligned} \langle \delta \mathbf{v} \rangle_{\mathbf{f}} &= 4\pi G(m + m') \log \Lambda \frac{dh}{dv} \frac{\mathbf{v}}{v} \\ &\equiv h_1(v) \mathbf{v} \\ \langle \delta v_i \delta v_j \rangle_{\mathbf{f}} &= 4\pi G m' \log \Lambda \left\{ \frac{d^2 g}{dv^2} \frac{v_i v_j}{v^2} + \frac{1}{v} \frac{dg}{dv} \left( \delta_{ij} - \frac{v_i v_j}{v^2} \right) \right\} \\ &\equiv [g_2(v) - g_1(v)] \frac{v^i v^j}{v^2} + g_1(v) \delta_{ij}, \end{aligned} \quad (\text{B24})$$

where

$$\begin{aligned} h_1(v) &\equiv 4\pi G(m + m') \frac{\log \Lambda}{v} \frac{dh}{dv} \\ g_1(v) &\equiv 4\pi G m' \log \Lambda \frac{1}{v} \frac{dg}{dv} \\ g_2(v) &\equiv 4\pi G m' \log \Lambda \frac{d^2 g}{dv^2}. \end{aligned} \quad (\text{B25})$$

and the required derivatives are

$$\begin{aligned} \frac{dh}{dv} &= -\frac{4\pi G}{v^2} \int_0^v dv' v'^2 f(v') \\ \frac{dg}{dv} &= \frac{4\pi G}{3} \left\{ \int_0^v dv' v'^2 f(v') \left( 3 - \frac{v'^2}{v^2} \right) \right. \\ &\quad \left. + 2v \int_v^\infty dv' v' f(v') \right\} \\ \frac{d^2 g}{dv^2} &= \frac{8\pi G}{3} \left\{ \frac{1}{v^3} \int_0^v dv' v'^4 f(v') + \int_v^\infty dv' v' f(v') \right\}. \end{aligned} \quad (\text{B26})$$

When we insert equations (B24) into equations (B11) we encounter several instances of

$$\mathbf{v} \cdot \frac{\partial P^i}{\partial \mathbf{v}} = \left( \frac{v^2 - \Omega_\vartheta L}{\Omega_r}, L, 0 \right) \quad (\text{B27})$$

from which we construct

$$\begin{aligned} \mathbf{M} &\equiv \mathbf{v} \cdot \frac{\partial P^i}{\partial \mathbf{v}} \cdot \frac{\partial P^j}{\partial \mathbf{v}} \\ &= \begin{pmatrix} (v^2 - \Omega_\vartheta L)^2 & v^2 - \Omega_\vartheta L & L & 0 \\ \frac{\Omega_r^2}{v^2 - \Omega_\vartheta L} & \Omega_r & L^2 & 0 \\ \frac{\Omega_r}{0} & 0 & 0 & 0 \end{pmatrix}. \end{aligned} \quad (\text{B28})$$

We also encounter

$$\begin{aligned} \mathbf{N} &\equiv \frac{\partial P^i}{\partial \mathbf{v}} \cdot \frac{\partial P^j}{\partial \mathbf{v}} \\ &= \begin{pmatrix} v^2 - 2\Omega_\vartheta L + \Omega_\vartheta^2 r^2 & L - \Omega_\vartheta r^2 & 0 \\ \frac{\Omega_r^2}{L - \Omega_\vartheta r^2} & \Omega_r & r^2 \\ \frac{\Omega_r}{0} & 0 & \frac{r^2}{L^2} \end{pmatrix}. \end{aligned} \quad (\text{B29})$$

With these definitions the  $(J_r, L, \psi)$  diffusion coefficients are

$$\begin{aligned} D^i &= h_1(v) \left( \frac{v^2 - \Omega_\vartheta L}{\Omega_r}, L, 0 \right) \\ D^{ij} &= \left[ \frac{g_2(v) - g_1(v)}{v^2} \mathbf{M} + g_1(v) \mathbf{N} \right]. \end{aligned} \quad (\text{B30})$$

Before the right-hand side of equation (B13) can be evaluated, it remains only to compute the Christoffel symbols that appear in it. This is most easily done by exploiting the fact that they vanish in the Cartesian system, so in the  $(J_r, L, \psi)$  system they are given by

$$\Gamma_{jn}^i = \frac{\partial P^i}{\partial v^k} \frac{\partial^2 v^k}{\partial P^j \partial P^n} \quad (\text{B31})$$

The matrix of first derivatives is given by equations (B4) to (B6). Differentiating the matrix (B3), we find the second derivatives are

$$\frac{\partial^2 \mathbf{v}}{\partial J_r \partial \mathbf{P}} = \begin{pmatrix} \frac{1}{v_r} \frac{\partial \Omega_r}{\partial J_r} - \frac{\Omega_r^2}{v_r^3} & 0 & 0 \\ \frac{1}{v_r} \frac{\partial \Omega_\vartheta}{\partial J_r} - (\Omega_\vartheta - L/r^2) \frac{\Omega_r}{v_r^3} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}^T \quad (\text{B32})$$

$$\frac{\partial^2 \mathbf{v}}{\partial L \partial \mathbf{P}} = \begin{pmatrix} \frac{1}{v_r} \frac{\partial \Omega_r}{\partial L} - (\Omega_\vartheta - L/r^2) \frac{\Omega_r}{v_r^3} & 0 & 0 \\ \frac{1}{v_r} \left( \frac{\partial \Omega_\vartheta}{\partial L} - \frac{1}{r^2} \right) - \frac{(\Omega_\vartheta - L/r^2)^2}{v_r^3} & 0 & 0 \\ 0 & -\frac{\sin \psi}{r} & \frac{\cos \psi}{r} \end{pmatrix}^T$$

$$\frac{\partial^2 \mathbf{v}}{\partial \psi \partial \mathbf{P}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\frac{\sin \psi}{r} & -v_t \cos \psi \\ 0 & \frac{\cos \psi}{r} & -v_t \sin \psi \end{pmatrix}. \quad (\text{B33})$$

The matrix formed by a Christoffel with a given lower left index is obtained by premultiplying each of these matrices by  $\partial \mathbf{P} / \partial \mathbf{v}$ .

### B1 The Coulomb logarithm

The isochrone sphere provides natural units  $M$  and  $b$  of mass and length, and from them we derive the natural unit of time

$$T_1 = \sqrt{\frac{b^3}{GM}}. \quad (\text{B34})$$

The Coulomb logarithm is traditionally defined to be

$$\log \Lambda = \ln \left( \frac{b_{\max} \sigma^2}{2G\mu} \right), \quad (\text{B35})$$

where  $b_{\max}$  is the largest impact parameter we should consider when one star encounters another and  $\sigma$  is the typical speed of stars. For the isochrone the virial theorem yields

$$\sigma^2 = \frac{GM}{b} \left( \frac{1}{4}\pi - \frac{2}{3} \right). \quad (\text{B36})$$

Hence

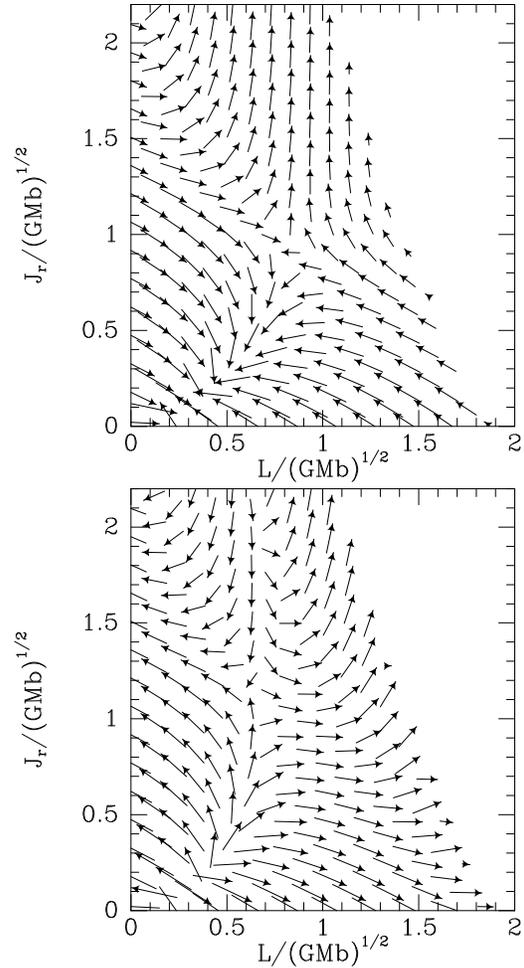
$$\log \Lambda = \ln \left( N \left( \frac{1}{4}\pi - \frac{2}{3} \right) \frac{b_{\max}}{2b} \right). \quad (\text{B37})$$

In the core one might reasonably argue that  $b_{\max}$  is of order  $b$ , while at some radius  $r$  outside the core  $b_{\max}$  could be a significant fraction of  $r$ . We will assume  $b_{\max} \sim b$  so  $\log \Lambda = \ln(0.059N) \simeq 8.7$  for  $N = 10^5$ .

### B2 Numerical details

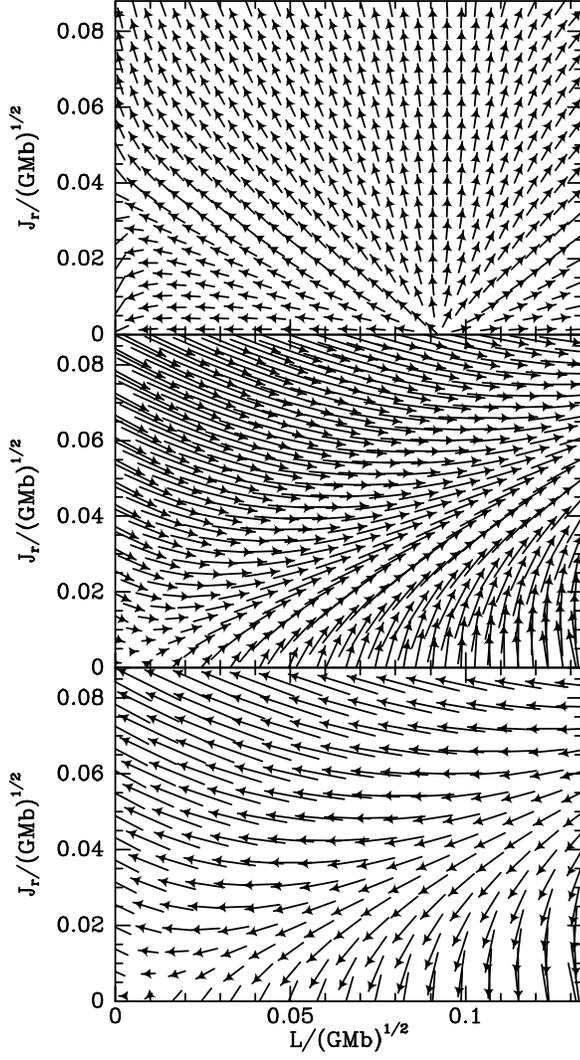
Numerical evaluation of the diffusive flux is quite delicate because components of the Christoffel symbols diverge in the limits  $L \rightarrow 0$  and  $v_r \rightarrow 0$ . The components that diverge as  $v_r \rightarrow 0$  cancel when  $\Gamma$  is multiplied into  $D$ . The divergence at  $L \rightarrow 0$  permits  $\bar{\mathbf{F}}_1$ , the definition of which (equation (B19)) has a leading factor  $L$ , to remain finite as  $L \rightarrow 0$ . Fig. B1 shows why it is important that  $\bar{\mathbf{F}}_1 \neq 0$  at  $L = 0$  by plotting  $\bar{\mathbf{F}}_1$  and  $\bar{\mathbf{F}}_2$  separately. Both  $\bar{\mathbf{F}}_1$  and  $\bar{\mathbf{F}}_2$  indicate unphysical fluxes of stars across the boundaries  $L = 0$  and  $J_r = 0$ . Physically the total flux,  $\bar{\mathbf{F}}_1 + \bar{\mathbf{F}}_2$ , across these boundaries has to vanish, and it does so because the components of  $\bar{\mathbf{F}}_1$  and  $\bar{\mathbf{F}}_2$  perpendicular to the boundaries cancel to a good approximation. Any error in the numerical scheme is liable to spoil this cancellation, so the general structure of the flow shown in Fig. 10 gives confidence in the soundness of the numerics.

However, along the bottom of Fig. 10 the arrows suggest that stars are flowing into  $LJ_r$  space across the  $L$  axis. Fig. B2 shows this region for the case  $R_a = 4.2b$  at higher resolution, for  $\bar{\mathbf{F}}_1 + \bar{\mathbf{F}}_2$  (top) and for  $\bar{\mathbf{F}}_1$  and  $\bar{\mathbf{F}}_2$  below. We see that the net flow is away from a singular point at  $(L, J_r) \simeq (0.095, 0)\sqrt{GMb}$  while neither  $\bar{\mathbf{F}}_1$  nor  $\bar{\mathbf{F}}_2$  has such a singular point and that away from the singular point the



**Figure B1.** Lower panel: the contribution from the divergence  $\bar{\mathbf{F}}_2 = -\frac{1}{2}\partial_j(Lf \int d\theta_r d\psi D^{ij})$  to the flow in  $LJ_r$  space. Upper panel: the complementing contribution to the flow,  $\bar{\mathbf{F}}_1$ .

net flow is parallel to the  $L$  axis because the vertical components of  $\bar{\mathbf{F}}_1$  and  $\bar{\mathbf{F}}_2$  precisely cancel. The singular point in the net flow is almost certainly a stagnation point, so by virtue of remarkable cancellations no stars pass over either axis.



**Figure B2.** Top: overall flow  $\bar{F}_1 + \bar{F}_2$  in the region of the singular point of the model with  $R = 4.2b$ . Centre: the flow  $\bar{F}_1$  driven by dynamical friction. Bottom: the flow  $\bar{F}_2$  that depends on  $\text{grad } f$ .