Kinetic Theory of Stellar Systems: A Tutorial

Chris Hamilton, Jean-Baptiste Fouvry

Kinetic Theory of Stellar Systems: A Tutorial

Chris Hamilton[1] and Jean-Baptiste Fouvry[2]

[1)] *Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA*

[2)] *CNRS and Sorbonne Université, UMR 7095, Institut d'Astrophysique de Paris, 98 bis Boulevard Arago, F-75014 Paris, France*

(*Electronic mail: chamilton@ias.edu)

(Dated: 22 February 2024)

Stellar systems — globular and nuclear star clusters, elliptical and spiral galaxies and their surrounding dark matter haloes, and so on — are ubiquitous characters in the evolutionary tale of our Universe. This tutorial article is an introduction to the collective dynamical evolution of the very large numbers of stars and/or other self-gravitating objects that comprise such systems, i.e. their *kinetic theory*.

We begin by introducing the basic phenomenology of stellar systems, and explaining why and when we must develop a kinetic theory that transcends the traditional two-body relaxation picture of Chandrasekhar. We then study the individual orbits that comprise stellar systems, how those orbits are modified by linear and nonlinear perturbations, how a system responds self-consistently to fluctuations in its own gravitational potential, and how one can predict the long term evolutionary fate of a stellar system in both quasilinear and nonlinear regimes. Though our treatment is necessarily mathematical, we develop the formalism only to the extent that it facilitates real calculations. Each section is bolstered with intuitive illustrations and we give many examples throughout the text of the equations being applied to topics of major astrophysical importance, such as radial migration, spiral instabilities, and dynamical friction on galactic bars.

Furthermore, in the 1960s and 1970s the kinetic theory of stellar systems was a fledgling subject which developed in tandem with the kinetic theory of *plasmas*. However, the two fields have long since diverged as their practitioners have focused on ever more specialized and technical issues. This tendency, coupled with the famous obscurity of astronomical jargon, means that today relatively few plasma physicists are aware that their knowledge is directly applicable in the beautiful arena of galaxy evolution, and relatively few galactic astronomers know of the plasma-theoretic foundations upon which a portion of their subject is built. Yet once one has become fluent in both Plasmaish and Galacticese, and has a dictionary relating the two, one can pull ideas directly from one field to solve a problem in the other. Therefore, another aim of this tutorial article is to provide our plasma colleagues with a jargon-light understanding of the key properties of stellar systems, to offer them the theoretical minimum necessary to engage with the modern stellar dynamics literature, to point out the many direct analogies between stellar- and plasma-kinetic calculations, and ultimately to convince them that stellar dynamics and plasma kinetics are, in a deep and beautiful and *useful* sense, the same thing.

CONTENTS

I. INTRODUCTION

The evolution of our Universe is a great unfolding drama in which galaxies play a fundamental role. These galaxies grow and die; they pulse and throb; they swallow gas and they belch it back out again; they merge together and they tear one another apart. The study of the collective dynamical behavior that occurs inside galaxies — and of the violent lives they lead, the tumultuous dark matter environments in which they sit, and the star clusters they cradle — is best termed the *kinetic theory of stellar systems*. This is the subject with which we are concerned in this tutorial article.

The kinetic theory of stellar systems rests upon two main theoretical pillars, and our tutorial article aims to introduce colleagues to both of them. The first pillar is the theory of *orbits*. By this we mean the orbital motion of stars (or other self-gravitating objects, like dark matter) in gravitational fields; or, more broadly, the dynamics of point particles subject to some Hamiltonian flow. The basic language of this topic is angle-action coordinates. These are not only the natural variables with which to describe smooth orbits in quiescent gravitational fields, but also to understand how these orbits are modified by *perturbations* of various kinds. These perturbations tend to have the most dramatic effect on orbits near *resonances*, where the frequency of the perturbation matches the natural frequency of the orbit. The first aim of our tutorial article is to introduce colleagues to these concepts. We employ as little heavy formalism as possible, assuming that readers already have a passing familiarity with Hamiltonian mechanics, our aim being to help them understand how they apply *in the stellar dynamical context*.

The second pillar upon which our subject is built is the collective dynamics of many bodies, and here our theoretical foundation is plasma kinetics. Why is this the case in a treatise on gravity? Well, every high school student knows that electromagnetism and gravity are intellectually adjacent subjects: up to a sign, the Coulombic electrical force and the Newtonian gravitational force are equivalent. What is less well appreciated is that this adjacency extends much further than simple $\pm 1/r^2$ force laws. Indeed, the mathematical methods developed to probe the collective behavior of protons and electrons in astrophysical and fusion plasmas are in many cases *precisely the same* as those needed to describe the dynamics of stars, spiral arms and dark matter in galaxies like the Milky Way. Concepts like two-body relaxation, linear response theory, Landau damping, resonant wave-particle interactions and phase mixing all had origins in plasma theory but are now part of the the standard lexicon of galactic dynamics. But these two fields long ago went their separate ways, and in much of today's literature are separated by a near-impenetrable wall of jargon. In the second part of this tutorial article we develop the tools needed to describe collective behavior in galaxies. We aim to provide colleagues with the theoretical minimum needed to appreciate galactic-dynamical problems. In so doing we aim to convey — to those on both the stellar-dynamical and plasma-theoretic sides of the aisle — just some of the deep connections between the two fields, and to encourage much greater cross-talk between these two estranged, yet intimately related, subjects.

In contrast with most other treatises on stellar (or 'galactic') dynamics and galaxy evolution, we approach our subject from the point of view of theoretical physics rather than practical astronomy. Thus, we spill almost no ink over such crucial concepts as mass-to-light ratios and chemical evolution and dust extinction (though we happily refer the reader to the many excellent texts that do[1–3]). Our models are necessarily rather simple, in some cases even crude, since our goal here is to provide physical insight rather than to accurately reproduce observational data. On the other hand, we do *not* simply develop mathematical formalism for its own sake. Almost every formal calculation we present is accompanied by a concrete example of its application. We hope this approach underscores our key message, which is that the kinetic theory of stellar systems is not merely the mathematical fancy of a few specialists. Instead, it can and should aid our understanding of galaxy evolution; and it passes the basic test of all good physics, by telling us things about the real world that we did not already know.

A. A (very) brief astronomy lesson

To set the context for this tutorial article, and to get an idea of the characteristic scales involved, we invite you to consider Figure 1. Panel (a) shows an image of a *disk galaxy* called the Southern Pinwheel, captured by the la Silla Observatory in the mountains of Chile. We have chosen this image not because it is visually stunning — though who could deny that? — but because it exhibits several of the paradigmatic features of disk galaxies, and is in fact thought to resemble our own Galaxy, the Milky Way. Like the Milky Way, the Southern Pinwheel has an elongated or *barred* structure at its center and exhibits beautiful *spiral arms*. The red knots strung along the arms are hydrogen-alpha emission from regions of recent star formation, meaning stars have been born there in the last few tens of Myr (where 1 Myr $\equiv 10^6$ years), while the bluer light comes from much older stars, up to ages of ~ 10 Gyr (where 1 Gyr $\equiv 10^9$ years; the Big Bang happened 13.7 Gyr ago, and galaxies formed in the first few Gyr that followed[3]). The image is haphazardly obscured by thin, dark *dust lanes*, regions wherein visible starlight has been absorbed. The whole galaxy is a few tens of kiloparsecs (kpc) across, where 1 pc $\sim 3 \times 10^{16}$m. (The distance from the Sun to the next-closest

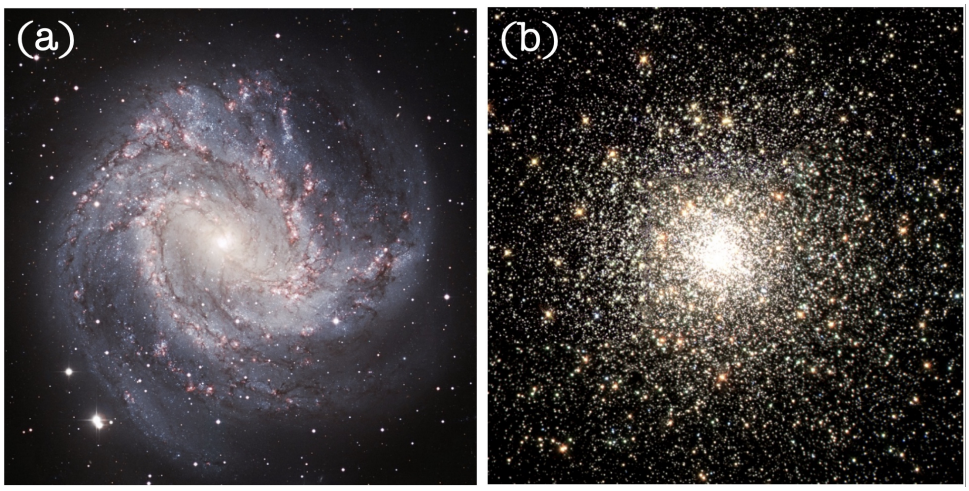

FIG. 1. (a) The spiral galaxy M83 (the Southern Pinwheel). Credit: ESO. (b) The globular cluster M80. Credit: NASA, The Hubble Heritage Team, STScI, AURA.

star, Proxima Centauri, is of order ~ 1 pc). From our perspective, we are seeing the Southern Pinwheel galaxy more or less face-on, but if we were to see it from the side we would find that it was very thin, with a central bulge — like two fried eggs glued back to back. It consists of around $N \sim 10^{11}$ stars, most of which are orbiting in this very thin, nearly two-dimensional disk plane, and they mostly move on (quasi-)*circular orbits* around the center of the galaxy, all rotating in the same direction. This means that the motion of stars in the galaxy is highly ordered, a.k.a. *dynamically cold*. Another way to say this is that the stellar disk, like a plasma beam, has low velocity dispersion. The key *dynamical timescale* in this disk is the azimuthal orbital period of one of the stars, which is typically $t_{\rm dyn} \gtrsim 100$ Myr, and this timescale increases with orbital radius. Since galaxies are ~ 10 Gyr old, we deduce that most stars have completed fewer than 100 orbits in the lifetime of the galaxy.

Now let us contrast this with panel (b) of Figure 1, which is an image of the *globular cluster* M80. This is another optical image, taken with the Hubble Space Telescope. Globular clusters are much smaller than galaxies, with typical radii of several pc — in fact, a galaxy like the Southern Pinwheel typically contains many hundreds or even thousands of globular clusters. Despite their relative smallness, globular clusters have a much higher density of stars than do galaxies. The cluster M80 contains something like $N \sim 10^5$ stars, so the typical distance between stars is $\sim (10^5/\mathrm{pc}^3)^{-1/3} \sim 0.02$pc, of order 50 times smaller than the distance from the Sun to Proxima Centauri. Globular clusters typically contain very little gas, and hence exhibit almost no ongoing star formation. This is why there is a much narrower range of colors in this image compared to the one beside it. Moreover, unlike disk galaxies, globular clusters are (nearly) *spherically symmetric*, so the image of M80 would not look very different if we were able to view it from another angle. Also unlike disk galaxies, they are disordered, *dynamically hot* systems, meaning that the motion of their stars has no strongly preferred direction (like the isotropic Maxwellian in an electrostatic plasma). More mathematically, at any given point $\mathbf{r}$ in the cluster, the local phase space distribution function $f(\mathbf{r}, \mathbf{v})$ depends primarily on the magnitude, rather than the direction, of the velocity $\mathbf{v}$. Though they are of a comparable age to galaxies — having formed ~ 10 Gyr ago — globular clusters are much more 'dynamically old' than galaxies, since the typical orbital period of their stars is only $t_{\rm dyn} \sim 10^5$yr. Thus each star in a globular cluster has undergone $\sim 10^5$ orbits in its lifetime.

To summarize: disk galaxies are *dynamically young and cold*, globular clusters are *dynamically old and hot*.

B. Two mental models

The brief astronomy lesson above motivates us to introduce two simple mental models which we will refer to throughout this tutorial article: the *cold, rotating axisymmetric disk* and the *hot, non-rotating spherical cluster* — see Figure 2 for an illustration. These are really just (over-)simplified versions of disk galaxies and globular clusters respectively, and constitute clean examples of stellar systems that exhibit completely different dynamical behaviors. Some of the characteristic properties of these two prototypical systems are summarized in Table I.

We will find that the cold disk is analogous to a non-thermal plasma (with collective oscillations/instabilities and wave-particle interactions playing a key role), while the hot sphere is more analogous to a thermal plasma or a classical

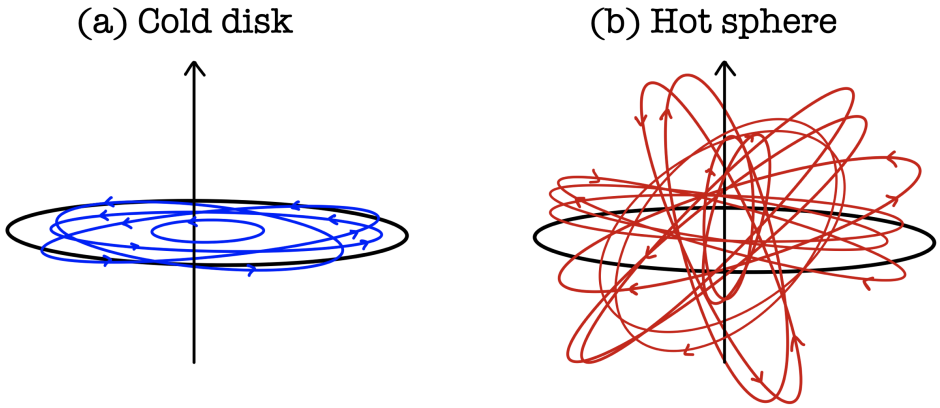

FIG. 2. Schematic illustration of typical orbits in (a) cold disks and (b) hot spheres (idealized versions of the archetypical systems shown in Figure 1). In a cold disk most orbits are nearly circular and confined to low inclinations relative to the disk plane. By contrast, in a hot sphere the orbits are near-isotropically distributed and have a broad range of eccentricities.

TABLE I. Characteristic scales for the two prototypical systems considered in this article, the disk galaxy and the globular cluster.

	Disk galaxy	Globular cluster
Mental model	Cold, rotating axisymmetric disk	Hot, nonrotating sphere
Size	$\sim$ 10s of kpc	$\sim$ several pc
Age	~ 10 Gyr	~ 10 Gyr
Number of stars N	$\sim 10^{11}$	$\sim 10^5$
Orbital period $t_{\rm dyn}$	~ 100 Myr	~ 0.1 Myr

gas (with a kinetic evolution primarily driven by discrete, local two-body encounters). But there are many subtleties and differences to be encountered along the way, and gravitational systems are capable of behaviors that have no precedent in gas or plasma kinetics.

C. Plan for the rest of this article

The rest of this tutorial article is organized as follows. In §II we introduce some of the basic properties of N-body self-gravitating systems, discuss the extent to which self-gravitating systems play by the same thermodynamic rules as gases and plasmas, and give a qualitative idea of how these systems evolve on long timescales. Then we begin with the theory of orbits. In §III we approximate a galaxy's gravitational potential by a smooth 'mean field' potential and study the orbits of individual stars in this mean field, with the help of angle-action variables. In §IV we show how individual orbits respond to both linear and nonlinear perturbations and describe the theory of resonant trapping. Our proper study of kinetic theory begins in §V, where we introduce the Klimontovich equation for an N-body system's *phase space distribution function* (DF) and derive a generic set of equations governing its evolution. In §VI we study the self-consistent linear response of stellar systems to weak perturbations, and then in §VII we use those linear solutions to construct a quasilinear theory of statistically-averaged evolution, culminating in the Balescu–Lenard equation. Our final topic, in §VIII, is nonlinear kinetic theory, particularly in the presence of orbit trapping and with applications to bar-halo friction and the saturation of spiral instabilities.

Throughout these sections we continually draw direct analogies between gravitational and plasma systems, as well as giving many concrete examples of the equations working in practice. Lastly, in §IX we touch on some other areas of galactic dynamics that have plasma-theoretic analogues, and give references for further reading.

We mention here that the standard reference in this subject is the textbook *Galactic Dynamics*[4] by Binney & Tremaine (2008), which we will hereafter refer to as BT08. While our article is far more narrow in scope than BT08, much of what we discuss cannot be found there or in any other textbook or publication. The main exceptions to this rule are the introductory sections §§II-III, where by necessity there is significant overlap with BT08. Where helpful, we point the reader to specific sections of BT08 in which more details can be found.

II. BASIC PROPERTIES OF N-BODY SELF-GRAVITATING SYSTEMS

The simplest theoretical approach to describing a stellar system is to imagine it as a gas of stars, and to apply the traditional methods of equilibrium thermodynamics and two-body collision theory. In this section we will attempt to do just that, and we will find that while these classical notions can be useful in characterizing stellar systems, they also break down completely in some circumstances — and that in the strict sense, in self-gravitating systems there is no such thing as equilibrium. The insights we gain from this failure will point us towards the more complete kinetic description that we will develop in later sections.

A. Virial equilibrium

The first thing to mention is that the speeds of stars in galaxies and star clusters are nearly always of order 10s or 100s of km s^{-1}, much less than the speed of light. This means that relativistic corrections to Newtonian gravity can almost always be ignored. There are exceptions, particularly for stars orbiting around supermassive black holes in galactic centers[5], but for our purposes Newtonian gravity is an excellent approximation. Also, while stars do exhibit a spectrum of different masses, for simplicity we will always assume that our system consists of a single mass species, unless we explicitly state otherwise.

We therefore consider a system of $N \gg 1$ stars of equal mass m, interacting only via Newtonian gravity. The gravitational potential energy due to the interaction between two stars with positions $\mathbf{r}$ and $\mathbf{r}'$ is

$$U(\mathbf{r},\mathbf{r}') = -\frac{Gm^2}{|\mathbf{r}-\mathbf{r}'|}, \tag{1}$$

where $G = 4.3\times10^{-3}$ pc $M_\odot^{-1}$km s^{-1} is Newton's gravitational constant (here $M_\odot = 2.0\times10^{30}$kg is the mass of the Sun). Newton's second law then tells us that the force on particle i is

$$m\frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t} = -\frac{\partial}{\partial\mathbf{r}_i}\sum_{j\neq i}^{N} U(\mathbf{r}_i,\mathbf{r}_j) = -Gm^2\sum_{j\neq i}^{N}\frac{\mathbf{r}_i-\mathbf{r}_j}{|\mathbf{r}_i-\mathbf{r}_j|^3}, \tag{2}$$

where $\mathbf{v}_i \equiv \mathrm{d}\mathbf{r}_i/\mathrm{d}t$ is the ith star's velocity.

Gravity is an exclusively attractive force — two masses always attract and never repel — which means that unlike in plasma, there is no shielding, nor is there any sensible notion of a 'Debye length' over which forces are screened. Since N is large ($\sim 10^{11}$ in the cold disk, $\sim 10^5$ in the hot sphere), and since every star is attracting every other star, what stops the entire system from collapsing under its own weight and forming a huge black hole? The answer depends on the context. In the hot sphere, the random motions of the stars counterbalance the gravitational attraction (and can be thought of as an effective 'pressure'). In the case of a cold disk it is the rotational support (i.e. the fact that the disk is spinning) that counterbalances gravity.

To put these notions on a more mathematical basis, we now derive the *virial theorem*. For an isolated N-body system we can choose choose the origin of our coordinates to be the fixed center of mass. Then the moment of inertia of the N-body system with respect to the origin is

$$I = \sum_{i=1}^{N} m\mathbf{r}_i^2. \tag{3}$$

Let us now take the second derivative of this quantity with respect to time

$$\frac{\mathrm{d}^2 I}{\mathrm{d}t^2} = 2\left(\sum_{i=1}^{N} m\mathbf{v}_i^2 + \sum_{i=1}^{N} m\mathbf{r}_i\cdot\frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t}\right). \tag{4}$$

Using the equation of motion (2), we can write the second term on the right hand side as

$$\begin{aligned}\sum_{i=1}^{N} m\mathbf{r}_i \cdot \frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t} &= -Gm^2 \sum_{j\neq i}^{N} \mathbf{r}_i \cdot \frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3} \\ &= -\tfrac{1}{2} Gm^2 \sum_{i=1}^{N}\sum_{j\neq i}^{N} \frac{(\mathbf{r}_i - \mathbf{r}_j)\cdot(\mathbf{r}_i - \mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|^3} \\ &= \sum_{i=1}^{N}\sum_{j>i}^{N} U(\mathbf{r}_i, \mathbf{r}_j). \end{aligned} \tag{5}$$

For an equilibrium system, the moment of inertia can fluctuate around some average value taken over a few dynamical times, but that average value must not itself change. Denoting such a time average by $\langle . \rangle$, we have that for a system in *virial equilibrium*,

$$2K + W = 0, \tag{6}$$

where

$$K = \left\langle \tfrac{1}{2} m \sum_{i=1}^{N} \mathbf{v}_i^2 \right\rangle, \qquad W = \left\langle \sum_{i<j}^{N} U(\mathbf{r}_i, \mathbf{r}_j) \right\rangle, \tag{7}$$

are the averaged total kinetic and total potential energies of the system respectively. In other words, in virial equilibrium, the kinetic and potential parts of the energy of the system balance each other to within an order-unity factor.

A globular cluster (lower panel of Figure 1) is an excellent example of a system in virial equilibrium. Its smooth, near-spherical shape reflects the fact that it is not, on average, undergoing any violent bulk motions, despite the stars inside it whizzing around like a swarm of bees on the timescale $t_{\rm dyn}$. Note that even though a cluster is in virial equilibrium, that equilibrium *does* still evolve; but the evolution of this 'quasi-stationary state' occurs over a timescale much longer than $t_{\rm dyn}$, as we will see (§II C).

The virial theorem allows us to make quick order-of-magnitude estimates about typical random velocities of stars, σ. For instance, consider that a globular cluster has a mass $M \sim Nm \sim 10^5 M_\odot$ and a radius R of a few parsecs. Using the virial theorem we can calculate the typical velocity dispersion of stars in this cluster. Roughly, we have $K \sim M\sigma^2/2$ and $W \sim -GM^2/R$, so using (6) we get

$$\sigma \sim \sqrt{\frac{GM}{R}} \sim 10\,\mathrm{km\,s^{-1}} \left(\frac{M}{10^5 M_\odot}\right)^{1/2} \left(\frac{R}{4\mathrm{pc}}\right)^{-1/2}, \tag{8}$$

in agreement with what is observed (and much less than c, justifying our use of Newtonian gravity).

Finally, let us contrast globular clusters with an example of a system which is very much *not* in virial equilibrium. Figure 3 shows the Antennae system, which consists of two galaxies that have collided and are in the process of merging. For this system, equation (6) does not hold; instead energy is rapidly being transferred between kinetic and potential forms on the timescale $t_{\rm dyn}$ as the two galaxies perform their cosmic dance. But we know from observations and numerical simulations of such galaxies that they ultimately *do* virialize, and that this virialization is nearly complete after only a few $t_{\rm dyn}$ (which in this case still means a billion years or more). This rather rapid evolution towards a virialized state is called *violent relaxation*[6]. The precise mechanics of violent relaxation, and the question of which particular virial equilibrium state the system will fall into, is a matter of ongoing research not only in stellar dynamics but also in plasma physics[7].

For the rest of this article we will only consider virialized stellar systems whose bulk properties are evolving on timescales $\gg t_{\rm dyn}$. Isolated disk galaxies and globular clusters both fall into this category.

B. Thermal equilibrium?

We mentioned above that virial equilibria are not true equilibria of stellar systems — rather, they are just quasi-equilibria that evolve on timescales much longer than the dynamical time $t_{\rm dyn}$. But why is this the case? After all, this is not what happens in either a gas or an electrostatic plasma: there, the natural thermal equilibrium distribution

FIG. 3. The Antennae galaxies NGC4038 and NGC4039. These two colliding galaxies are in the process of merging into a single structure. It is easy to see by eye that this system is going to undergo significant morphological changes on the typical timescale $\sim t_{\rm dyn}$. Hence, it is not in virial equilibrium. Credit: ESA/Hubble.

of particles' positions and velocities is a homogeneous Maxwellian. Unfortunately, gravity does not play by those thermodynamic rules.

To see why, suppose we wished to describe a globular cluster as a 'stellar gas' in thermal equilibrium. We could make a natural definition of the cluster's temperature T by relating it to the total kinetic energy

$$K = \frac{3}{2} N k_{\rm B} T, \tag{9}$$

where N is the number of stars. Then the total energy of the cluster $E_{\rm tot}$ is the sum of the kinetic and potential energies

$$E_{\rm tot} = K + W. \tag{10}$$

But from the virial theorem (6) we know that $W = -2K$, so $E_{\rm tot} = -K < 0$. The fact that $E_{\rm tot}$ is negative is not unusual: it just means that the system is gravitationally bound. The heat capacity is then

$$C = \frac{\partial E_{\rm tot}}{\partial T} = -\frac{\partial K}{\partial T} = -\frac{3}{2} N k_{\rm B} < 0. \tag{11}$$

Notice that C is *negative*. This *is* strange: if you add heat to the system (increase its total energy $E_{\rm tot}$), it cools down! In other words, the more energy is given to the cluster, the lower the amplitude of the random velocities inside it.

The negativity of the heat capacity implies that it is impossible for a globular cluster ever to reach a thermal equilibrium. For instance, consider a cluster at temperature T immersed in a heat bath of the same temperature. Suppose there is some infinitesimal negative fluctuation in the cluster's energy, $-|\delta E|$. By (11), this will cause the cluster to achieve a higher temperature by $\delta T = C/(-|\delta E|) > 0$, i.e. with higher velocity dispersion. Since the temperature is now higher, the system loses more heat to the bath, which then makes its temperature increase even further, and so on without limit.

Nor can one achieve thermal equilibrium in an isolated stellar system (i.e. in the absence of any heat bath). The reason is that any star with a sufficiently large velocity is unbound from the system and hence will escape to infinity, a process known as *evaporation*. This precludes the system from ever attaining a Maxwellian velocity distribution function (DF), since the cluster would have a finite phase space density out to $v \to \infty$.

One might argue at this point that the Maxwellian is only an idealization, and not being able to *exactly* realize a Maxwellian DF is a rather pedantic argument for the non-existence of thermal equilibrium. (Indeed, all speeds are ultimately bounded by the speed of light, but this does not stop us from employing a Maxwellian DF when describing plasmas!). To justify our claim further, let us calculate the rate at which stars *would* evaporate from a hypothetical Maxwellian stellar system. As we will argue in §III, the motion of individual stars on the timescale $\sim t_{\rm dyn}$ is very accurately described by a *mean field approximation*, in which we ignore the complexities of the N-body system and instead consider the stars to be moving in a *smooth* gravitational potential $\Phi(\mathbf{r})$, in precisely the same way as one

imagines the unperturbed motion in a homogeneous electrostatic plasma to be a straight line. Then the energy per unit mass of any star with velocity v and position $\mathbf{r}$ is $v^2/2+\Phi(\mathbf{r})$, meaning that any star with $v > v_{\rm esc}(\mathbf{r}) \equiv \sqrt{-2\Phi(\mathbf{r})}$ will evaporate. Averaging this over the entire system gives us the typical escape speed

$$\sigma_{\rm esc}^2 \sim \frac{1}{M}\int \mathrm{d}\mathbf{r}\,\rho(\mathbf{r})v_{\rm esc}^2(\mathbf{r}) = -\frac{2}{M}\int \mathrm{d}\mathbf{r}\,\rho(\mathbf{r})\Phi(\mathbf{r}), \tag{12}$$

where $\rho_0(\mathbf{r})$ is the mass density of stars at position $\mathbf{r}$, and M is the total cluster mass. This is obviously related to the total potential energy of the system W. Being careful not to over-count the pairwise interactions, and using the virial theorem (6), we get

$$\sigma_{\rm esc}^2 \sim -\frac{4W}{M} = \frac{8K}{M} = 4\sigma^2, \tag{13}$$

where $\sigma^2 = N^{-1}\langle\sum_{i=1}^N \mathbf{v}_i^2\rangle$ is the cluster's velocity dispersion. Supposing the cluster did have a Maxwellian DF with the same velocity dispersion, we can estimate the fraction of unbound stars by

$$f_{\rm esc} \approx \frac{\int_{2\sigma}^\infty \mathrm{d}v\, v^2 \mathrm{e}^{-3v^2/2\sigma^2}}{\int_0^\infty \mathrm{d}v\, v^2 \mathrm{e}^{-3v^2/2\sigma^2}} \approx \frac{1}{140}. \tag{14}$$

In other words, we would expect about 1% of an isolated stellar system to evaporate away on a *relaxation time* (see the next subsection for the definition of this term). The equivalent situation in a collisional plasma would be that 1% of the electrons disappeared every mean-free time! Thus, even if the cluster 'wanted' to reach a Maxwellian steady-state, it could not do so because stars are constantly evaporating from the high-velocity tail.

Evaporation is a real and important effect in cluster dynamics; indeed there are certain star clusters called *open clusters* with $N \sim 10^3$ (some of which you can see with the naked eye, like the Pleiades) whose lifetime is limited by evaporation to a few hundred Myr. In other words, there were star clusters visible to the dinosaurs that have since evaporated into the vastness of space and are no longer visible to us. But at no point was their DF ever Maxwellian.

With minor modifications, the above arguments also hold for cold, differentially rotating disks, as well as other virialized self-gravitating structures. The bottom line is that *strict thermal equilibrium is impossible in a self-gravitating system*, and there is no universal tendency for stellar systems to 'tend towards the Maxwellian' as gases and plasmas do. In the language of statistical mechanics, in isolated self-gravitating systems there is no microcanonical ensemble, no canonical ensemble, energy is not extensive, and entropy is not bounded (so there exists no maximum entropy state[4,8]). Concepts like entropy are still useful in stellar dynamics, but one has to be very careful about naive application of 'standard' statistical mechanical results.

C. So what does happen to a virialized stellar system?

We have argued that there exists no unevolving thermal equilibrium state for stellar systems. What, then, does happen to stellar systems after they reach virialization? Roughly speaking, the answer is that they move slowly from one virialized quasi-equilibrium state to another on a timescale that we call the *relaxation time*, $t_{\rm relax}$.

How do we calculate the relaxation time? Roughly speaking, it is the timescale over which a typical star 'forgets' its initial orbit. A proper theory of relaxation requires an understanding of stellar orbits (§III), but for now the intuition from gas or plasma kinetics is good enough. Hence we simply ask, supposing an unperturbed star was moving in a straight line, how long would it take for its velocity $\mathbf{v}$ to de-correlate from its initial value, i.e. to accumulate an impulse $|\Delta\mathbf{v}| \sim |\mathbf{v}|$?

In a gas these impulses occur because a given molecule collides with some other molecule roughly once per mean-free time. Hence, after a few such collisions, this molecule has completely forgotten its initial state (see Figure 4a). This is not so in a stellar system. Even in a dense system like a globular cluster, the distance between single stars is such that direct physical collisions are very rare[9]. Instead, in a cluster the evolution of a star's orbit is mostly driven by a very large number of weak nudges from all the other stars it is interacting with simultaneously (Figure 4b). In the next two subsections we will follow Chandrasekhar and estimate the relaxation timescale due to a series of such nudges (§II C 1), a calculation completely analogous to the Spitzer two-body collision calculation from plasma kinetics. Then (§II C 3) we will show that Chandrasekhar's theory is very successful at describing the evolution of hot spheres (in the same way that Spitzer's theory works for describing collisional plasmas) but fails completely to describe the evolution of cold disks (as does Spitzer's theory when dealing with collisionless plasma out of equilibrium). This will be the jumping-off point where stellar dynamics ceases to consist of a series of rough estimates; thereafter, things get quantitative.

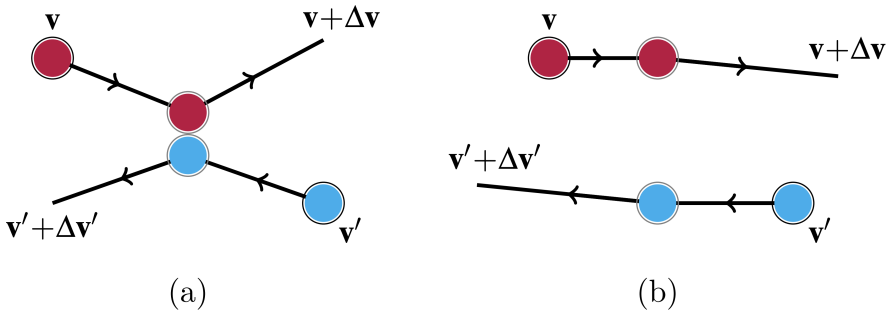

FIG. 4. (a) Relaxation in a gas involves physical contact collisions (or at least, very short-range interactions) between molecules, meaning a molecule's velocity can be changed significantly after a single collision. (b) In a stellar system, by contrast, two-body encounters are mostly weak, long-range nudges. Later in the article we will deal with *collective effects*, whereby one must think beyond bare two-body processes and account for the coherent oscillations of very large numbers of stars at once.

1. Chandrasekhar's theory of two-body relaxation

Here we give the simplest version of Chandrasekhar's calculation of the relaxation time[10], based on his theory of two-body relaxation, and primarily following §1.2.1 of BT08. (For a more complete treatment of Chandrasekhar theory, see Chapter 7 of BT08). What we will present is a gravitational version of Rutherford scattering.

To begin, we will ignore all the complications associated with real stellar systems and instead consider a single 'test' star moving in a statistically homogeneous box of static 'field' stars of mass m. Let the test star travel in a perfectly straight line until it approaches a field star at impact parameter b. The goal now is to calculate the nudge to the test star's velocity, $\delta\mathbf{v}(b)$.

Since we are going to assume that the interaction between the two stars is weak, to lowest order we can approximate the trajectory of the test star as a straight line throughout the encounter (see Figure 5). Suppose without loss of

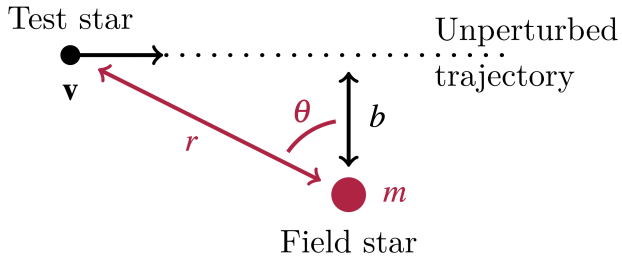

FIG. 5. A test star encounters a field star of mass m at impact parameter b. Assuming the encounter to be weak, the lowest-order nudge to the test star's velocity, $\delta\mathbf{v}$, can be calculated by integrating along the unperturbed trajectory (straight dotted line).

generality that the closest approach of the test star to the field star occurs at $t = 0$. We may then calculate the total impulse it receives by integrating from $t = -\infty$ to $t = \infty$. Then the acceleration the test star feels in the direction parallel to $\mathbf{v}$ while flying towards the field star ($t < 0$) is exactly canceled by the retardation it feels as it flies away ($t > 0$), meaning that there is no overall velocity change in the parallel direction. Hence, we need only to calculate the total *perpendicular* velocity nudge $\delta v_\perp$.

The perpendicular acceleration field felt by the test star is (see Figure 5 for the definition of r and θ):

$$F_\perp = \frac{Gm}{r^2}\cos\theta = \frac{Gm}{b^2}\left[1+\left(\frac{vt}{b}\right)^2\right]^{-3/2}. \tag{15}$$

Integrating over all time, we get

$$\delta v_\perp = \int_{-\infty}^{\infty} \mathrm{d}t\, F_\perp = \frac{Gm}{bv}\int_{-\infty}^{\infty}\frac{\mathrm{d}s}{(1+s^2)^{3/2}} = \frac{2Gm}{bv}. \tag{16}$$

As one might expect, the more massive is the field star, and the smaller is the impact parameter b, the larger is the velocity impulse the test star experiences. Moreover, it is sensible that the impulse decays with v, since the larger

is v the shorter is the duration of the encounter. On the other hand, the formula (16) also predicts its own failure: if $\delta v_\perp$ is sufficiently large that it is comparable to the original speed v, then the assumed straight-line trajectory is obviously a poor approximation. Note that for a fixed m and v, this breakdown occurs at a critical impact parameter $b_{90} = 2Gm/v^2$, so-called because this in fact produces a 90-degree deflection. For impact parameters $b \lesssim b_{90}$ one ought really to solve the two-body encounter problem without making any approximations (which is not difficult — see §3.2 of BT08). Luckily, such strong encounters are rare and make little difference to the bulk evolution of the system. We will ignore them from now on.

Chandrasekhar's next step is to assume that relaxation in a stellar system consists of nothing more than an uncorrelated series of these weak two-body encounters. Under this assumption we can calculate the relaxation time as follows. Let the number density of stars in the box be n. Then in a time T the field star passes a total of $\sim nvT\,2\pi b\,\mathrm{d}b$ field stars at impact parameter $\in [b, b+\mathrm{d}b]$. The mean-square kick to the velocity that the test star receives in time T is therefore

$$\sum \delta v_\perp^2 \approx \int_{b_{\min}}^{b_{\max}} nvT\,2\pi b\,\mathrm{d}b \times \delta v_\perp^2 = \frac{8\pi G^2 m^2 n}{v} T \ln\Lambda, \tag{17}$$

where $\ln\Lambda \equiv \ln(b_{\max}/b_{\min})$ is the *Coulomb logarithm*, which is familiar from the corresponding plasma calculation of charged particle scattering[11].

We seem to have a problem here, since taking either $b_{\min} \to 0$ or $b_{\max} \to \infty$, or both, causes the total velocity impulse to diverge (albeit in the most undramatic way, i.e. logarithmically). So, what values of $b_{\min}$ and $b_{\max}$ should we take? Usually one takes $b_{\max}$ to be comparable to the scale of the system $\sim R$, since there cannot be any encounters more distant than this (this is different from plasma, in which one can safely take $b_{\max}$ of order the Debye screening length[11]). Moreover, one typically takes $b_{\min} \sim b_{90}$ because, as we mentioned above, strong encounters are usually rare and unimportant. In fact, while fixing the Coulomb logarithm accurately is an important problem for those modeling globular cluster evolution[12], from a theorist's standpoint the answer is *it does not matter too much because the model is only heuristic anyway.* For instance, encounters with impact parameters b on the order of the size of the system certainly cannot be thought of as straight lines, since they must take into account the curved geometry of the true orbits of the stars (§II C 3). The standard choice $\Lambda \sim R/b_{90}$ is good enough for obtaining order-of-magnitude estimates, which is all we care about here.

If we now set $\sum \delta v_\perp^2$ equal to the original squared speed v^2, we can identify $T = t_{\mathrm{2BR}}$ as the *two-body relaxation time*

$$t_{\mathrm{2BR}} \sim \frac{v^3}{8\pi G^2 m^2 n \ln\Lambda}. \tag{18}$$

Let us relate this more clearly to the number of stars in the system N. Let the side-length of the box be $\sim R$ — which is a proxy for the radius of the star cluster or galaxy in which we are interested. Then we know from the virial theorem that typically $v^2 \sim GNm/R$ (see equation 8). We may also set the typical crossing time to be $t_{\mathrm{dyn}} \sim R/v$ and the typical number density to be $n \sim N/(4\pi R^3/3)$. Then taking $b_{\max} = R$ and $b_{\min} = b_{90} = 2Gm/v^2$ we get $\Lambda \sim NRv^2/(Gm) \sim N$, so that

$$t_{\mathrm{2BR}} \sim \frac{0.1\,N}{\ln N} t_{\mathrm{dyn}}. \tag{19}$$

To be even more heuristic, we will often drop the factors of 0.1 and $\ln N$ and simply state that the two-body relaxation time is on the order of $t_{\mathrm{2BR}} \sim N t_{\mathrm{dyn}}$.

Plugging in characteristic numbers from Table I, we find that for a globular cluster $t_{\mathrm{2BR}} \sim 10$ Gyr, comparable to the age of the cluster. This suggests that the simple physics of uncorrelated two-body encounters is of major importance for determining the evolution of globular clusters. However, for a disk galaxy this theory suggests evolution should only occur on the excruciatingly long timescale $t_{\mathrm{2BR}} \sim 10^{10}$ Gyr, nine orders of magnitude longer than the age of the Universe[13]. And yet we know from observations and simulations that disk galaxies can evolve significantly on a timescale of $\gtrsim$ a few Gyr. The estimate (19) tells us that star-star encounters are much too inefficient to have produced this evolution, so a more sophisticated kinetic theory is required (§II C 3).

2. 'Collisional' and 'collisionless'

It is worth commenting here on the (often extremely confusing) use of the words *collisional* and *collisionless* when describing stellar systems. These words were partly inherited from plasma theory, where they have a fairly concrete meaning: roughly, interactions that occur on scales smaller than the Debye scale are 'collisions', and if these can be

ignored then the system is 'collisionless'. In a stellar system, however, there is no shielding, and no Debye length, and no natural way to decouple small-scale and large-scale gravitational potential fluctuations — every star is interacting gravitationally with every other star, always. Sometimes, the word 'collisionless' is used as a shorthand for 'a system whose two-body relaxation time (19) is much longer than its age' (e.g., a disk galaxy) while 'collisional' systems are those with $t_{\rm 2BR}$ less than or comparable to their age (basically, star clusters). A closely related definition is to say a system is collisionless if, when solving a particular problem, finite-N effects never need to be considered. We will avoid this terminology a much as we can, although it is so ingrained in the literature that sometimes this is impossible.

3. *Where do we go from here?*

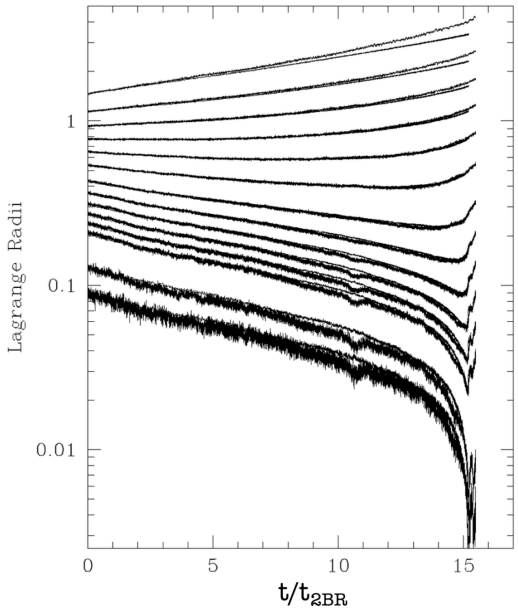

FIG. 6. Long term evolution of a globular cluster (adapted from Joshi *et al.*[14]). The lines show the evolution of the 'Lagrange radii' which enclose 0.35, 1, 3.5, 5, 7, 10, 14, 20, 30, 40, 50, 60, 70, and 80 percent of the cluster mass respectively. In each case, both the results of direct N-body simulation and of a Monte–Carlo calculation[15] based on Chandarasekhar's theory are shown. For most times and Lagrange radii, the two sets of results are indistinguishable.

It turns out that when applied to the evolution of hot spherical clusters, Chandrasekhar's picture does a remarkably good job of estimating the true relaxation time $t_{\rm relax}$ and various other evolutionary properties[12], although the value of the Coulomb logarithm has to be tweaked[17,18]. Moreover, one can use Chandrasekhar's ideas to spin up a full kinetic theory — i.e. a kinetic equation of the form $\partial f/\partial t = C[f]$, with C calculated by assuming relaxation consists of two-body encounters as above — and this works well in hot spheres too[19]. To demonstrate this, we present Figure 6 (adapted from Joshi *et al.*[14]), which shows the evolution of 'Lagrange radii' of a spherical globular cluster in a numerical simulation. These Lagrange radii are the radii enclosing a fixed percentage of the cluster's mass at any given time. First of all, we see that the inner Lagrange radii gradually decrease with time (corresponding to the contraction of the cluster core, leading ultimately to the 'gravothermal catastrophe'[20]) and the outer radii slowly increase (corresponding to the expansion of the cluster's envelope). This evolution happens very slowly (the time unit on the horizontal axis is the two-body relaxation time $t_{\rm 2BR} \gg t_{\rm dyn}$), corresponding to the gradual evolution from one virialized state to another. Looking more closely, one can see that for each Lagrange radius in the panel there are actually two lines plotted. One of these was calculated from a direct N-body simulation of the 'exact' equations of motion of all the stars, whereas the other employed a Monte–Carlo calculation based upon Chandrasekhar's theory of two-body encounters. The fact that across much of the figure these two lines are indistinguishable is testament to the success of Chandrasekhar theory in describing hot spherical clusters.

On the other hand, as mentioned at the end of §II C 1, Chandrasekhar's theory fails completely to describe the evolution of cold disk galaxies. First of all, disk galaxies develop long-range collective structures like the spiral arms in Figure 1a, and these have no place in Chandrasekhar's picture. Also, Chandrasekhar's theory drastically underestimates the timescale on which these systems relax. As an example of this, consider Figure 7 (adapted

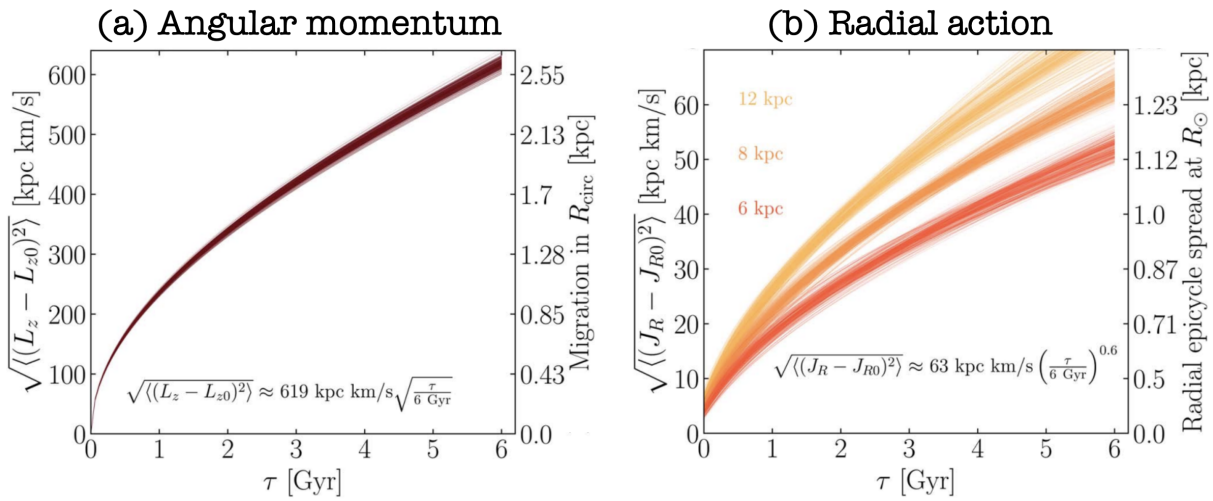

FIG. 7. Adapted from Frankel *et al.*[16]. Each panel shows the evolution of an orbital integral of motion (namely (a) angular momentum and (b) radial action, which will be introduced properly in §III C) of stars in the Milky Way over 6 Gyr, as inferred from GAIA data. In the right panel, different colors correspond to groups of stars with different initial Galactocentric radii.

from Frankel *et al.*[16]). This figure was generated by fitting a simple diffusion model to the orbital actions (to be introduced in §III C) of stars in the Milky Way using data from GAIA, a space observatory which has been mapping the kinematics of stars in our Galaxy since 2013. Panel (a) shows the diffusion of the stellar angular momenta (essentially the guiding radius of near-circular orbits around the galaxy), while panel (b) shows the evolution of those same stars' mean radial actions (essentially their orbital 'eccentricities'). In both panels, we see that significant evolution can occur on the timescale of a few Gyr. To get an estimate for the true relaxation time in this system, note that the orbital angular momenta of stars on circular orbits (speed $v_\phi \sim 250$ km s^{-1}) near the Solar radius in the Milky Way ($R \sim 8$ kpc) is $L = Rv_\varphi \sim 2000$ kpc km s^{-1}. The relaxation time we infer from Figure 7a is therefore of order $t_{\rm relax} \sim (2000/619)^2 \times 6$ Gyr ~ 60 Gyr, many orders of magnitude shorter than would be suggested by the Chandrasekhar estimate (19). And even though this relaxation time is still several times longer than the age of the Galaxy (~ 10 Gyr), it can still have a major effect on galaxy evolution, as we will see in §VIII B 1.

Chandrasekhar's theory fails because it is based on a number of false assumptions, for instance:

- Stellar systems are inherently *inhomogeneous* (see Figure 1!) but the theory assumed homogeneity.
- Because of the inhomogeneity, unperturbed stellar orbits are not endless straight lines, but instead are *quasiperiodic* (§III). This leads to *resonant* interactions.
- Stars do not behave in an independent, uncorrelated way but instead tend to oscillate *collectively*. Thus, one cannot replace all interactions between stars with a superposition of uncorrelated two-body encounters.

Of course these statements apply to both globular clusters and disk galaxies, which begs the question: why does Chandrasekhar's theory work to describe the former but not the latter? The answer lies in the fact that globular clusters are dynamically *hot* whereas disk galaxies are dynamically *cold*. The hotness of clusters means that they act like thermal plasmas: collective effects are weak and successive fluctuations are mostly uncorrelated, as different parts of the system have no way to efficiently communicate with one another. On the other hand, the highly-ordered orbital structure of disk galaxies means that great numbers of stars can have long-lived coherent dynamical interactions. The result is that disk galaxies, like non-equilibrium plasmas, are replete with large-scale collective oscillations and possible instabilities. Such effects are able to redistribute energy and angular momentum much more efficiently — and in a fundamentally different way — than the mechanism of incoherent two-body relaxation.

It follows that to describe the behavior of disk galaxies, we need to leave Chandrasekhar's theory behind and develop a new kinetic theory, one capable of dealing with the complications listed above. That is one of the aims the rest of this tutorial article. But before we can do kinetics, we must learn how to describe the orbits of stars in inhomogeneous, but still fixed and smooth, mean field gravitational potentials (§III), and then figure out what happens to those orbits when they are perturbed (§IV).

III. ORBITS IN MEAN FIELD POTENTIALS

As we mentioned in §II B, on the timescale $\sim t_{\rm dyn}$ the orbits of stars in a virialized N-body system can be very well described by ignoring the discrete nature of system and instead treating individual stars as test particles orbiting in a smooth potential. This is a good approximation for two reasons: (i) the gravitational force between two objects is long range, decreasing as the square of the distance between them, and (ii) the number of stars in the system N is very large. The consequence of (i) is that the force acting on any individual star is typically not dominated by its nearest neighbors, but rather by the collective gravity of simultaneous distant interactions. The consequence of (ii) is that the Poisson fluctuations in the number of stars in a given sub-volume tends to be small, so the collective gravitational force due to stars in this sub-volume fluctuates only weakly on the timescale $t_{\rm dyn}$. The average value of this collective force, combining all the sub-volumes, is then typically very smooth (see also §1.2 of BT08).

Another key feature of virialized stellar systems is that they have rather simple bulk geometry: they are normally approximately spherical (in the case of globular clusters, some dwarf galaxies, etc.), or thin and axisymmetric (like a disk galaxy), or possibly ellipsoidal/triaxial (like some massive elliptical galaxies). This simple geometry leads to (mostly) simple, regular orbits, and not many chaotic orbits. In this section, we will describe how stars orbit within a simple class of smooth mean field potentials.

A. Mean field orbits

In elementary courses on plasma physics there is no great heed paid to 'mean field dynamics' because those dynamics are trivial: in a homogeneous electrostatic plasma, mean field 'orbits' are just straight lines: $\mathbf{r}(t) = \mathbf{v}t + \mathbf{r}(0)$ and $\mathbf{v} = \mathbf{v}(0) = \text{cst}$. On the contrary, stellar systems are inherently inhomogeneous, which means that the unperturbed mean field orbits of stars are not straight lines. This is something plasma physicists also have to deal with when considering orbits in systems with nontrivial geometry, such as in toroidal tokamak fusion plasmas or the Earth's magnetosphere[21–23].

If the gravitational potential is $\Phi(\mathbf{r})$, the equations of motion of a test star are

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{v}, \tag{20a}$$

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\frac{\partial\Phi}{\partial\mathbf{r}}. \tag{20b}$$

In practice, we cannot solve these equations of motion explicitly except in a few very simple potentials (e.g., the Kepler potential $\Phi(\mathbf{r}) \propto -1/|\mathbf{r}|$, and the harmonic potential $\Phi(\mathbf{r}) \propto |\mathbf{r}|^2$). But we can solve the equations *numerically* — this is called 'integrating the orbit'. Modern numerical packages like `Galpy`[24], `Gala`[25] and `AGAMA`[26] allow one to integrate and analyze mean field orbits in pre-defined potentials with very little computational cost.

In general, numerical orbit integration reveals a wide array of possible orbit families — see Chapter 3 of BT08 for an overview. However, for the purposes of this article there is only one simple orbit family we need to visualize, which is the planar rosette orbit shown in Figure 8. That is because this is (nearly) the only type of orbit that is possible in *central* potentials $\Phi(r)$, where $r \equiv |\mathbf{r}|$. This encompasses both our mental models, the hot sphere and the cold axisymmetric disk (Figure 2). By symmetry, motion in central potentials conserves each component of the specific orbital angular momentum vector, $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, and so each such orbit is confined to a two-dimensional plane (which we can take to be the plane $Z = 0$ without loss of generality). From now on we describe orbits in this plane with a standard (ϕ, R) polar coordinate system centered at $r = 0$.

1. *Orbits in central potentials*

We will now show how to describe the rosette-type orbit of Figure 8 quantitatively. We see from Figure 8 that the orbit consists of a combination of an azimuthal circulation in the positive-ϕ direction, and a radial oscillation in R between the fixed *pericentre* $R_{\rm p}$ and *apocentre* $R_{\rm a}$. We can calculate $R_{\rm p}$ and $R_{\rm a}$ in terms of the two conserved quantities of the star's motion, namely its specific energy and specific angular momentum (we will drop the word 'specific' from now on)

$$E = \tfrac{1}{2}\left(\frac{\mathrm{d}R}{\mathrm{d}t}\right)^2 + \tfrac{1}{2}\frac{L^2}{R^2} + \Phi(R), \tag{21a}$$

$$L = R^2\frac{\mathrm{d}\phi}{\mathrm{d}t}. \tag{21b}$$

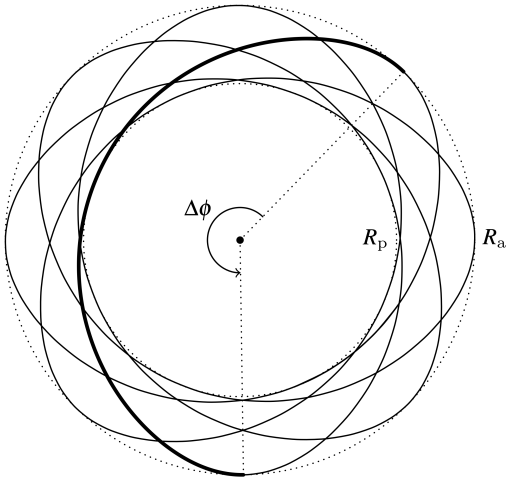

FIG. 8. Nearly all orbits in central potentials $\Phi(r)$ look like this: a two-dimensional planar 'rosette'. Using in-plane polar coordinates (ϕ, R), each orbit is a combination of an azimuthal oscillation in the ϕ direction and a radial oscillation between R_p and R_a. The thick black line shows the trajectory over one radial period, from R_a to R_p and back again.

The peri/apocentre are the radial turning points of the orbit; that is, they are solutions to $v_R(R_\mathrm{p}) = v_R(R_\mathrm{a}) = 0$, where $v_R = \mathrm{d}R/\mathrm{d}t$ is the radial velocity

$$v_R(R) \equiv \sqrt{2[E - \Phi(R)] - L^2/R^2}. \tag{22}$$

In central potentials, there is then a one-to-one correspondence between the integrals of motion (E, L) and the peri/apocentric distances $(R_\mathrm{p}, R_\mathrm{a})$:

$$E(R_\mathrm{p}, R_\mathrm{a}) = \frac{R_\mathrm{a}^2\Phi(R_\mathrm{a}) - R_\mathrm{p}^2\Phi(R_\mathrm{p})}{R_\mathrm{a}^2 - R_\mathrm{p}^2}, \tag{23a}$$

$$L(R_\mathrm{p}, R_\mathrm{a}) = \sqrt{\frac{2[\Phi(R_\mathrm{a}) - \Phi(R_\mathrm{p})]}{R_\mathrm{p}^{-2} - R_\mathrm{a}^{-2}}}. \tag{23b}$$

We can also calculate the frequencies of the radial and azimuthal oscillations as follows. First, the radial frequency is

$$\Omega_R = \frac{2\pi}{T_R}, \tag{24}$$

where T_R is the radial period

$$T_R(R_\mathrm{p}, R_\mathrm{a}) = 2\int_{R_\mathrm{p}}^{R_\mathrm{a}} \frac{\mathrm{d}R}{v_R(R)}. \tag{25}$$

The azimuthal frequency is

$$\Omega_\phi = \Omega_R \frac{|\Delta\phi|}{2\pi}, \tag{26}$$

where $|\Delta\phi|$ is the increment in azimuthal angle over one radial period reading

$$|\Delta\phi(R_\mathrm{p}, R_\mathrm{a})| = 2L\int_{R_\mathrm{p}}^{R_\mathrm{a}} \frac{\mathrm{d}R}{R^2 v_R(R)}. \tag{27}$$

In other words, all the characteristic quantities describing the orbit can be calculated given the two numbers $(R_\mathrm{p}, R_\mathrm{a})$, or equivalently the two constants of motion (E, L). This is in contrast to a homogeneous plasma, where one can label an orbit just by its velocity $\mathbf{v}$.

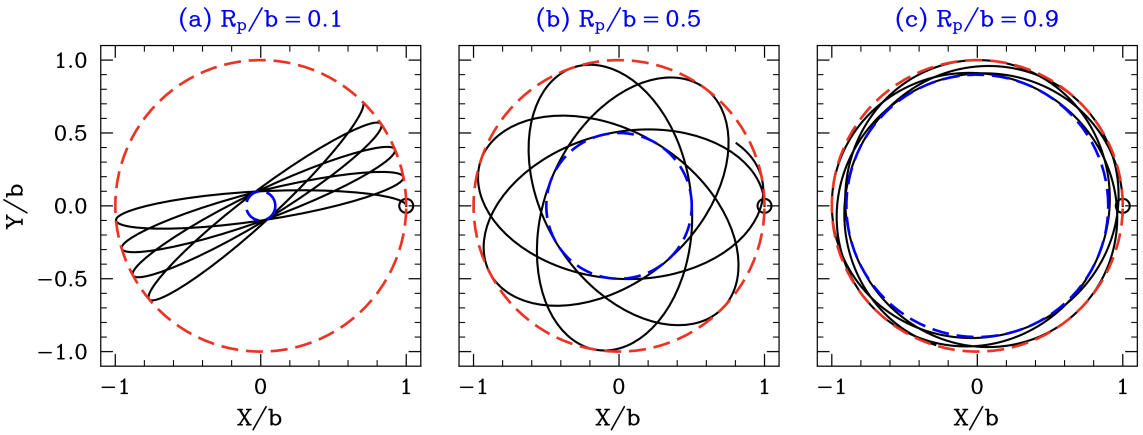

FIG. 9. Three orbits in the Plummer potential (28). In each case the orbit starts at its apocentric distance at the location $(X,Y)=(b,0)$. Each orbit has the same apocenter $R_\mathrm{a}=b$ (red dashed circles) but different pericenters R_p (blue dashed circles), as indicated above each panel. As a result, orbit (a) is almost radial ($J_R \gg L$), orbit (c) is almost circular ($J_R \ll L$), and orbit (b) is somewhere in-between ($J_R \sim L$) — see §III C 3 for the definition of the radial action J_R.

To further illustrate what these rosette orbits look like, consider motion in the spherical *Plummer potential*, which is often used as a simple model for the gravitational potential in globular clusters:

$$\Phi(r) = -\frac{GM}{\sqrt{r^2+b^2}}, \qquad \text{(Plummer potential)}. \tag{28}$$

Here b is the scale radius of the potential (for a globular cluster $b\sim 1\mathrm{pc}$ — see Table I). The Plummer potential has very simple asymptotic forms: for $r\gg b$ it approximates the potential of a point mass, $\Phi\approx -GM/r$, whereas for $r\ll b$ the (non-constant part of the) potential is quadratic $\Phi\propto r^2$, which corresponds to motion in a spherical 'core' of constant density (see BT08, §3.1). In Figure 9 we show three orbits integrated in the potential (28) in the (ϕ,R) plane, each with the same $R_a=b$ but different R_p, as labeled above each panel. Orbit (a) is nearly radial, with $R_\mathrm{p}\ll R_\mathrm{a}$, whereas orbit (c) is nearly circular, with $R_\mathrm{p}\approx R_\mathrm{a}$, and orbit (b) is somewhere in-between.

Of course, the orbit labels $(R_\mathrm{p},R_\mathrm{a})$, or equivalently (E,L), contain no information about the orbit's instantaneous radial or azimuthal *phase*, just like labeling an orbit by $\mathbf{v}$ in a plasma does not tell us its instantaneous position in the box $\mathbf{r}$. But it turns out that the phase of the orbits also exhibit rather simple behavior, as we investigate next.

B. Quasiperiodicity

A crucial feature of orbits in many smooth, regular galactic potentials revealed by numerical orbit integration is that they are *quasiperiodic*. By this we mean that, very often, one can Fourier-analyze a numerically integrated orbit and write it as

$$\mathbf{r}(t) = \sum_{\mathbf{n}} \mathbf{r_n}\, \mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\mathbf{\Omega}t}. \tag{29}$$

Here $\mathbf{n}\in\mathbb{Z}^d$ where d is the number of dimensions, is a dimensionless vector of integers, e.g., $(0,-1,4)$ in the case of $d=3$. The vector $\mathbf{\Omega}$ is made up of d frequencies that are characteristic of the orbit in question. Quasiperiodicity is also a common feature of particle orbits in tokamaks, magnetospheres, and so on.

That the orbits in Figure 9 are quasiperiodic might have been guessed by eye, but this fact also extends to a much broader range of orbits in realistic galactic potentials[27,28]. Mathematically, that an orbit is quasiperiodic can be proven in cases where the number of known independent isolating integrals of motion is (at least) equal to the number of dimensions (which is always true for orbits in spherically symmetric potentials). However, it is a non-trivial empirical fact that quasiperiodicity also extends to orbits in many realistic *axisymmetric* potentials $\Phi(R,Z)$ for which the orbit is fully three-dimensional and there are only two obvious independent constants of motion (E and the component of angular momentum along the symmetry axis, L_Z)[27,28]. Normally in a system with more dimensions than there are independent integrals of motion, one would expect some degree of chaos; instead, it is believed that orbits in axisymmetric potentials almost always possess a third independent integral of motion, whose analytic form we do not know but which makes the orbits regular and hence quasiperiodic (though this is not generically guaranteed[29]).

While triaxial potentials (like ellipsoids) can also harbor families of quasiperiodic orbits, they also tend to have many more chaotic orbits too, which arise in the parts of phase space between the regions of distinct quasiperiodic families. As such, in the rest of this tutorial article we will suppose that the exact potential of the system is close to, or exactly, axisymmetric, and that in the limit of pure axisymmetry there is only one regular (quasiperiodic) family of orbits we need to worry about, and no chaos. The central potentials $\Phi(r)$ are of this nature.

C. Angle-action variables

The upshot of quasiperiodicity is that stellar orbits can be thought of as *a superposition of oscillators* with fixed frequencies $\boldsymbol{\Omega}$. Here we will show how this fact allows us to construct a coordinate system in which this mean field orbital motion is trivial.

Consider a set of canonical coordinates and momenta $(\mathbf{q},\mathbf{p})$, and consider a star moving according to the generic Hamiltonian $H(\mathbf{q},\mathbf{p},t)$. The equations of motion for that star are

$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} = \frac{\partial H}{\partial \mathbf{p}}, \tag{30a}$$

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -\frac{\partial H}{\partial \mathbf{q}}. \tag{30b}$$

We can also use these canonical coordinates to introduce a bilinear antisymmetric functional on phase space called a *Poisson bracket*. For any two functions $g(\mathbf{q},\mathbf{p})$ and $h(\mathbf{q},\mathbf{p})$ we define

$$[g,h] = \frac{\partial g}{\partial \mathbf{q}}\cdot\frac{\partial h}{\partial \mathbf{p}} - \frac{\partial g}{\partial \mathbf{p}}\cdot\frac{\partial h}{\partial \mathbf{q}}. \tag{31}$$

This will be useful later.

A perfectly valid choice for the canonical coordinates $(\mathbf{q},\mathbf{p})$ is the position and velocity $(\mathbf{r},\mathbf{v})$. However, the lack of straight line orbits in stellar systems means that $(\mathbf{r},\mathbf{v})$ are awkward variables with which to formulate a kinetic theory. We would like instead to work as much as possible with quantities that are invariant along orbits. An *integral of motion* $I(\mathbf{r},\mathbf{v})$ is any function of phase space coordinates that is constant along an orbit; as argued in §III B for all the mean field orbits we care about in this tutorial article, there exist at least as many independent integrals of motion as there are degrees of freedom in the system (normally 2 or 3). Of course, one can construct as many mutually *dependent* integrals of motion as one wishes, since any function of integrals of motion is itself an integral of motion.

The obvious next question is: given that these integrals of motion exist, can we use them as coordinates for our phase space? The answer is yes, but if we want to preserve the Hamiltonian structure of our system — i.e. if we want the new coordinates to be *canonical* — then we have to be careful about which integrals to use. It turns out that there is a special set of integrals called *actions* $\mathbf{J}$, which *are* canonical momenta on phase space[30]. Specifically, these actions have the special property that the system's Hamiltonian $H(\mathbf{q},\mathbf{p})$ can be written exclusively in terms of them, $H = H(\mathbf{J})$, and they can be complemented by canonical coordinates or *angles* $\boldsymbol{\theta}$, which have the special property that they are 2π-periodic. In these special angle-action coordinates $(\mathbf{q},\mathbf{p}) = (\boldsymbol{\theta},\mathbf{J})$, Hamilton's equations (30) take the simple form

$$\frac{\mathrm{d}\boldsymbol{\theta}}{\mathrm{d}t} = \frac{\partial H}{\partial \mathbf{J}} = \boldsymbol{\Omega}(\mathbf{J}), \tag{32a}$$

$$\frac{\mathrm{d}\mathbf{J}}{\mathrm{d}t} = -\frac{\partial H}{\partial \boldsymbol{\theta}} = 0, \tag{32b}$$

where we introduced the orbital frequencies $\boldsymbol{\Omega}(\mathbf{J}) = \partial H/\partial \mathbf{J}$. In these coordinates, the orbits are straight lines, as one has

$$\boldsymbol{\theta}(t) = \boldsymbol{\theta}(0) + \boldsymbol{\Omega}(\mathbf{J})\,t \pmod{2\pi}, \tag{33a}$$

$$\mathbf{J}(t) = \mathbf{J}(0) = \mathrm{cst}. \tag{33b}$$

Some further properties of the angle-action coordinates are as follows. Since they are canonical, they leave the Poisson bracket (31) invariant, so that one generically has

$$[g,h] = \frac{\partial g}{\partial \boldsymbol{\theta}}\cdot\frac{\partial h}{\partial \mathbf{J}} - \frac{\partial g}{\partial \mathbf{J}}\cdot\frac{\partial h}{\partial \boldsymbol{\theta}}. \tag{34}$$

As a result, angle-action coordinates also leave the infinitesimal phase space volumes invariant so that $\mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}=\mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{v}=\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}\mathbf{J}$. Furthermore, the periodicity of the angles in d dimensions means that any phase space function can be Fourier expanded as

$$G(\boldsymbol{\theta},\mathbf{J})=\sum_{\mathbf{n}}G_{\mathbf{n}}(\mathbf{J})\,\mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}}, \tag{35}$$

with inverse

$$G_{\mathbf{n}}(\mathbf{J})=\int\frac{\mathrm{d}\boldsymbol{\theta}}{(2\pi)^d}\,G(\boldsymbol{\theta},\mathbf{J})\,\mathrm{e}^{-\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}}. \tag{36}$$

Actions are also *adiabatic invariants*: roughly speaking, if H evolves on a timescale longer than the dynamical time, $t_{\rm dyn}\simeq 1/|\boldsymbol{\Omega}|$, an orbit governed by H will evolve in such a way that $\mathbf{J}=\mathrm{cst}$. (but see Weinberg[31,32,33] for a more precise statement).

Note that we have not *proven* the existence of the angle-action variables, nor have we explained how to generate them from scratch (via the process of writing down a suitable canonical generating function). All this is very standard material which is explained rather neatly (if heuristically) in §3.5 of BT08, or much more rigorously in the books by Lichtenberg and Lieberman[34] and Arnold[35]. Throughout the rest of this tutorial article we will simply assume that the angle-action coordinates can be constructed for the problem at hand. Let us give some simple examples.

1. Example: Angle-action variables for a multiperiodic homogeneous system

As a first example, we consider a fictitious *homogeneous stellar system*, which provides a direct analogy with homogeneous plasmas. To this end, let us assume that our stellar system is placed within a periodic three-dimensional box of side-length L. We also assume that the mean potential vanishes, i.e. $\Phi_0=0$, so that unperturbed trajectories are straight lines. In this case, the system's angle-action coordinates and the associated orbital frequencies are given by

$$\boldsymbol{\theta}=\frac{2\pi}{L}\mathbf{r},\qquad \mathbf{J}=\frac{L}{2\pi}\mathbf{v},\qquad \boldsymbol{\Omega}=\frac{2\pi}{L}\mathbf{v}. \tag{37}$$

This makes clear a crucial correspondence between (fictitious) homogeneous stellar systems and homogeneous electrostatic plasmas. In both systems, mean field orbits are labelled by their velocity $\mathbf{v}$, and, up to a prefactor, these velocities are also the orbit's intrinsic frequencies $\boldsymbol{\Omega}$. This means that (i) the gravitational/electrostatic potential $\Phi(\mathbf{r})$ is a function only of angles $\boldsymbol{\theta}$ and not actions $\mathbf{J}$, and (ii) the orbital frequencies $\boldsymbol{\Omega}(\mathbf{J})$ are trivial (linear) functions of the actions. These two features are special to homogeneous systems, and greatly simplify the kinetic description (see e.g. §VI B 1).

2. Example: Angle-action variables for the 1D harmonic oscillator

Let us now consider the case of a one-dimensional harmonic oscillator whose Hamiltonian is $H=(v^2+\omega^2x^2)/2$. As illustrated in Figure 10, in this case, orbits take the form of concentric circles in the phase space $(x,v/\omega)$.

We define the action via $H=J\omega$, along with the angle θ so that

$$x=\sqrt{\frac{2J}{\omega}}\sin\theta,\qquad v=\sqrt{2J\omega}\cos\theta. \tag{38}$$

Following (34), one can check that $[\theta,J]=1$; thus, (θ,J) are canonical coordinates. In follows that orbits which followed circles in $(x,v/\omega)$ space now traverse straight horizontal lines in (θ,J) space.

3. Example: Angle-action variables for orbits in a central potential

For central potentials, $\Phi(R)$, the two-dimensional rosette motion can be described using action variables $J_1=L$ and $J_2=J_R$, where L is the usual angular momentum (see equation 21) and J_R is the *radial action*

$$J_R=\frac{1}{\pi}\int_{R_{\rm p}}^{R_{\rm a}}\mathrm{d}R\,v_R(R), \tag{39}$$

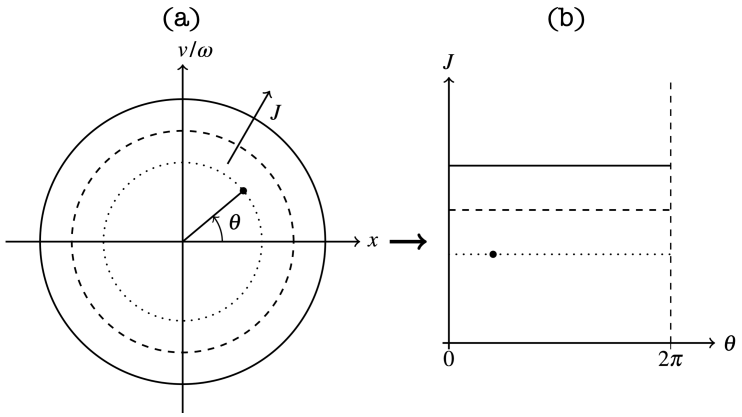

FIG. 10. Phase space diagram of a harmonic oscillator, $H = (v^2 + \omega^2 x^2)/2$. Panel (a) shows particle trajectories in the physical phase space $(x, v/\omega)$, which take the form of concentric circles. Here, the action J should therefore be seen as a label for the circle associated with the orbit, and the angle θ should be seen as the position along the circle. Panel (b) shows the same trajectories in angle-action space (θ, J). In these coordinates, the mean field motions are straight lines. The action J is conserved, while the angle θ evolves linearly in time.

with the radial velocity, v_R, given by equation (22). This is directly analogous to the longitudinal adiabatic invariant $J = m \oint v_\parallel \mathrm{d}s$ for particle motion in a magnetic mirror[36]. The conjugate angles are the azimuthal and radial angles

$$\theta_1 = \theta_\phi = \phi + \int_{R_\mathrm{p}}^{R} \mathrm{d}R' \, \frac{\Omega_\phi - L/R'^2}{v_R(R')}, \tag{40a}$$

$$\theta_2 = \theta_R = \Omega_R \int_{R_\mathrm{p}}^{R} \frac{\mathrm{d}R'}{v_R(R')}. \tag{40b}$$

These angles evolve at the rate of the azimuthal frequency Ω_ϕ and radial frequency Ω_R respectively.

Figure 11 (taken from Hunt *et al.*[37]) presents recent data from the spacecraft GAIA, showing the distribution of millions of stellar orbits in our local patch of the Milky Way disk, plotted in various slices through angle-action space. Each panel is full of structure, some of which is due to selection effects (since we are unable to access data across the whole Galaxy) but much of which is due to dynamical interactions between stars, spiral arms, the Galaxy's central bar, and so on. Though we will not pursue any details here, understanding structures like those in Figure 11 is a major goal of modern Galactic astronomy.

4. *Example: Angle-action variables for near-circular orbits*

We have introduced angle-action coordinates as the natural variables with which to describe quasiperiodic orbits, since those orbits are superpositions of (generally anharmonic) oscillations at discrete frequencies $\boldsymbol{\Omega}(\mathbf{J})$. However, a disadvantage of angle-action coordinates is that we can rarely make explicit the mapping $(\mathbf{r}, \mathbf{v}) \leftrightarrow (\boldsymbol{\theta}, \mathbf{J})$. On the other hand we *can* write down an explicit mapping if we restrict ourselves to near-circular orbits (which is a highly relevant regime for describing cold disks). There, we can employ the *epicyclic approximation*. In this limit, each orbit consists of circular motion at the *guiding center* radius R_g, plus a superposition of *harmonic* oscillations in the radial and azimuthal coordinates. In Appendix A we show how to construct explicit angle-action coordinates in the epicyclic approximation.

D. Summary and outlook

For the remainder of this tutorial article we assume that we can always construct angle-action variables $(\boldsymbol{\theta}, \mathbf{J})$, such that the mean field Hamiltonian depends only on $\mathbf{J}$ (which is itself a nontrivial technical problem[39]). It is then tempting to think of actions as the 'velocity' variables in our new phase space and the angles as our new 'positions'. And indeed, many formulae from plasma kinetics will more-or-less hold in stellar dynamics just by making the substitutions $\mathbf{v} \to \mathbf{J}$ and $\mathbf{r} \to \boldsymbol{\theta}$ (see especially Table II). However, there are two key mathematical differences

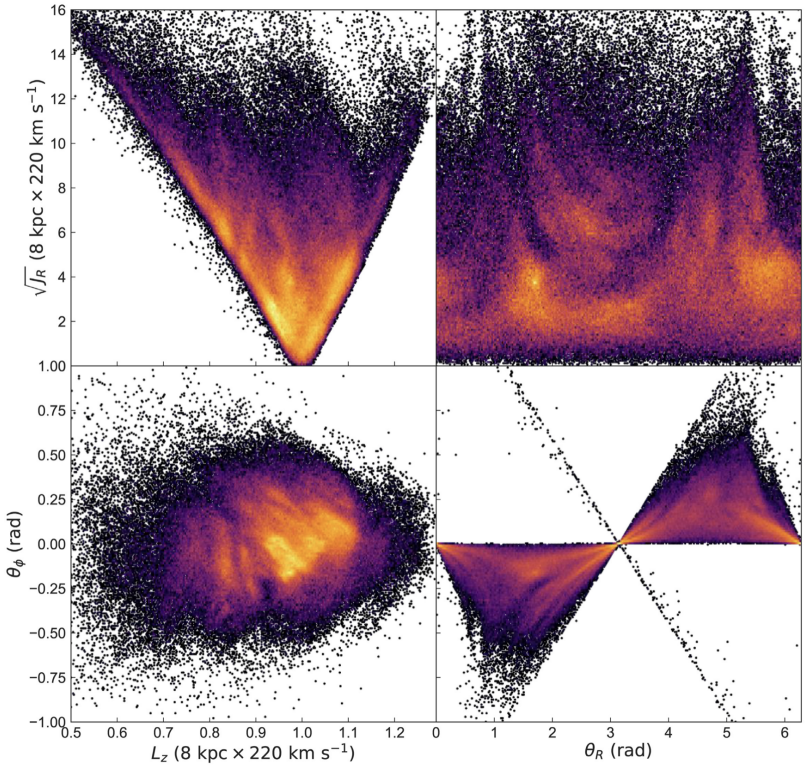

FIG. 11. The distribution of stars' orbits in our local patch of the Milky Way disk (a.k.a. the Solar neighborhood), plotted in angle-action space, using data from the GAIA satellite. This figure was taken from Hunt *et al.*[37]. Another detailed guide to features in the action space (top left panel) can be found in Trick *et al.*[38].

which make stellar kinetics much harder than plasma kinetics. The first is that unlike in a homogeneous plasma, the canonical coordinates change at rates which are in general nonlinear functions of the canonical momenta; i.e. $\boldsymbol{\Omega}(\mathbf{J})$ is typically a complicated (nonlinear) function of $\mathbf{J}$. The other key difference is that the fluctuating gravitational potential in these variables (which we will introduce in §IV) depends on *both* the canonical coordinates and momenta, i.e. $\delta\Phi(\mathbf{r},t) = \delta\Phi(\boldsymbol{\theta},\mathbf{J},t)$. As we will see, these differences make stellar systems in some ways richer dynamically than electrostatic plasmas, but also make explicit calculation more difficult.

IV. PERTURBED ORBITS

Now that we know how to describe stellar orbits in smooth central potentials, we are ready to consider what happens if we perturb those orbits. In particular, in this section we discuss how stellar orbits in an axisymmetric, razor-thin (i.e. two-dimensional) disk are modified when we introduce a rigidly-rotating non-axisymmetric perturbation (e.g. a spiral arm pattern). The discussion will be familiar to plasma physicists who have studied, for example, the interaction of ions with toroidal Alfven waves in tokamaks[21,23].

A. Mean field and perturbations

To begin our construction, at a given instant let the *exact* gravitational potential of the whole two-dimensional system (including external perturbations, if desired) be $\Phi(\mathbf{r}) = \Phi(\phi, R)$. Then, the exact Hamiltonian of a test star is

$$H(\mathbf{r},\mathbf{v}) = \tfrac{1}{2}\mathbf{v}^2 + \Phi(\mathbf{r}). \tag{41}$$

We suppose that Φ is reasonably close to axisymmetry, so it makes sense to introduce the azimuthal average of H as

$$\overline{H}(R,\mathbf{v}) \equiv \tfrac{1}{2}\mathbf{v}^2 + \overline{\Phi}(R), \tag{42}$$

where

$$\overline{\Phi}(R) \equiv \frac{1}{2\pi}\int \mathrm{d}\phi\, \Phi(\mathbf{r}). \tag{43}$$

Now, since $\overline{\Phi}$ is axisymmetric, we know (§III B) that we can construct angle-action variables $(\boldsymbol{\theta},\mathbf{J})$ such that the associated Hamiltonian $\overline{H}$ only depends on $\mathbf{J}$. Having constructed these, we can express any function on phase space in terms of them, via $\mathbf{r}=\mathbf{r}(\boldsymbol{\theta},\mathbf{J})$ and $\mathbf{v}=\mathbf{v}(\boldsymbol{\theta},\mathbf{J})$. In particular, we can return to our original exact Hamiltonian H and consider it to be a function of $\boldsymbol{\theta}$ and $\mathbf{J}$. Moreover, we can split H into a $\boldsymbol{\theta}$-independent 'mean field' part and a $\boldsymbol{\theta}$-dependent 'fluctuation', via

$$H[\mathbf{r}(\boldsymbol{\theta},\mathbf{J}),\mathbf{v}(\boldsymbol{\theta},\mathbf{J})] = H_0(\mathbf{J}) + \delta H(\boldsymbol{\theta},\mathbf{J}), \tag{44}$$

where

$$H_0(\mathbf{J}) \equiv \int \frac{\mathrm{d}\boldsymbol{\theta}}{(2\pi)^2}\, H[\mathbf{r}(\boldsymbol{\theta},\mathbf{J}),\mathbf{v}(\boldsymbol{\theta},\mathbf{J})]. \tag{45}$$

We now consider more closely the mapping from $(\mathbf{r},\mathbf{v})=(\phi,R,v_\phi,v_R)$ to $(\boldsymbol{\theta},\mathbf{J})=(\theta_\phi,\theta_R,J_\phi,J_R)$ that we constructed in §III C 3. We notice that the velocities $\mathbf{v}$ and the radius R do not depend on θ_ϕ, while the azimuthal angle $\phi=\theta_\phi+\lambda(\theta_R,\mathbf{J})$ where λ is a (generally rather complicated, but ϕ-independent) function. It follows that averaging over θ_ϕ is equivalent to averaging over ϕ, i.e.

$$H_0(\mathbf{J}) = \int \frac{\mathrm{d}\theta_\phi \mathrm{d}\theta_R}{(2\pi)^2} H[\theta_\phi+\lambda(\theta_R,\mathbf{J}),R(\theta_R,\mathbf{J}),\mathbf{v}(\theta_R,\mathbf{J})] = \int \frac{\mathrm{d}\theta_R}{2\pi}\,\overline{H}[R(\theta_R,\mathbf{J}),\mathbf{v}(\theta_R,\mathbf{J})] = \overline{H}(\mathbf{J}), \tag{46}$$

where the last equality follows because $\overline{H}$ is independent of $\boldsymbol{\theta}$ by construction. We can therefore speak of orbits in the 'mean field Hamiltonian' H_0 or the 'axisymmetric Hamiltonian' $\overline{H}$, and they mean the same thing. Note, though, that this is *not* true of the potential alone:

$$\Phi_0(\mathbf{J}) \equiv \int \frac{\mathrm{d}\boldsymbol{\theta}}{(2\pi)^2}\Phi(\boldsymbol{\theta},\mathbf{J}) = \int \frac{\mathrm{d}\theta_R}{2\pi}\overline{\Phi}[R(\theta_R,\mathbf{J})] \neq \overline{\Phi}(R), \tag{47}$$

because R depends on θ_R. Thus, although commonly done in the literature, it is misleading to talk of orbits in the 'mean field potential' Φ_0 as if this were the same thing as the axisymmetric potential $\overline{\Phi}$. When one speaks of 'mean field orbits' one is referring to orbits governed by the full mean field Hamiltonian $H_0=\overline{H}$, and not any $\boldsymbol{\theta}$-averaged potential.

Finally, we can combine equations (41), (42) and (44) and the fact that $\overline{H}=H_0$ to deduce that

$$\delta H(\boldsymbol{\theta},\mathbf{J}) = \Phi(\mathbf{r}) - \overline{\Phi}(R) \equiv \delta\Phi(\boldsymbol{\theta},\mathbf{J}) \tag{48}$$

is just the non-axisymmetric part of the potential. Hence, from now on we write

$$H[\mathbf{r}(\boldsymbol{\theta},\mathbf{J}),\mathbf{v}(\boldsymbol{\theta},\mathbf{J})] = H_0(\mathbf{J}) + \delta\Phi(\boldsymbol{\theta},\mathbf{J}), \tag{49}$$

and refer to H_0 and $\delta\Phi$ as the mean field Hamiltonian and the non-axisymmetric potential perturbation respectively.

B. Rotating potential perturbation

We now assume that the non-axisymmetric potential perturbation $\delta\Phi$ rotates in the ϕ direction with fixed pattern speed Ω_p:

$$\delta\Phi = \delta\Phi(\phi-\Omega_\mathrm{p}t,R). \tag{50}$$

Given that the perturbation is periodic in $\phi-\Omega_\mathrm{p}t$, using equation (40a) and the fact that R is a function only of θ_R and $\mathbf{J}$ (through the implicit relation 40b), we can always expand $\delta\Phi$ as a Fourier series (c.f. equation 35)

$$\delta\Phi(\boldsymbol{\theta},\mathbf{J},t) = \sum_{\mathbf{n}\neq\mathbf{0}} \delta\Phi_\mathbf{n}(\mathbf{J})\,\mathrm{e}^{\mathrm{i}(\mathbf{n}\cdot\boldsymbol{\theta}-n_\phi\Omega_\mathrm{p}t)}, \tag{51}$$

where $\mathbf{n}=(n_\phi,n_R)$. For a perturbation with fixed azimuthal harmonic number $m>0$ (e.g. an m-armed spiral), the only terms which survive in the expansion (51) are those with $n_\phi=\pm m$.

C. Resonances

A star's orbit evolves according to Hamilton's equations (32). Thus, using the Fourier expansion (51), its action evolves as $\mathbf{J} = \mathbf{J}(0) + \delta\mathbf{J}(t)$, where

$$\frac{\mathrm{d}\delta\mathbf{J}}{\mathrm{d}t} = -\mathrm{i}\sum_{\mathbf{n}} \mathbf{n}\,\delta\Phi_{\mathbf{n}}(\mathbf{J})\,\mathrm{e}^{\mathrm{i}(\mathbf{n}\cdot\boldsymbol{\theta} - n_\phi\Omega_\mathrm{p}t)}. \tag{52}$$

Given this equation, it will not surprise the reader that the majority of interesting orbital modifications by rigidly rotating potential perturbations occur at *resonances*, i.e. locations $\mathbf{J}$ in phase space where

$$\mathbf{N}\cdot\boldsymbol{\Omega}(\mathbf{J}) = N_\phi\Omega_\mathrm{p}. \tag{53}$$

for some vector of integers $\mathbf{N} = (N_\phi, N_R)$. That is because these are locations where, substituting the unperturbed trajectories (33a)-(33b), the argument of the exponential on the right hand side of (52) is constant.

For concreteness, suppose the perturbation in question has azimuthal harmonic number m. Then the most important values of $\mathbf{N}$ — which we are sometimes referred to as the perturbation's *major resonances* — are

$$\mathbf{N}_\mathrm{CR} = (m, 0) \quad - \quad \text{corotation resonance (CR)}, \tag{54a}$$
$$\mathbf{N}_\mathrm{ILR} = (m, -1) \quad - \quad \text{inner Lindblad resonance (ILR)}, \tag{54b}$$
$$\mathbf{N}_\mathrm{OLR} = (m, 1) \quad - \quad \text{outer Lindblad resonance (OLR)}. \tag{54c}$$

These resonances tend to dominate the in-plane evolution of disk galaxies. They have a simple physical interpretation in the important case of near-circular orbits (see §III C 4 and Appendix A). In that case, the corotation resonance is located at the orbital radius R_CR such that an orbit with guiding radius $R_\mathrm{g} = R_\mathrm{CR}$ advances in azimuth at exactly the same rate as the perturbation, i.e. $\Omega_\mathrm{c}(R_\mathrm{CR}) = \Omega_\mathrm{p}$, with Ω_c the azimuthal frequency for circular orbits. If we now imagine keeping that star's orbit near-circular but increasing R_g slightly, it will begin to circulate at a frequency slightly smaller than Ω_p. As we continue this process, eventually R_g will be sufficiently large that the difference $\Omega_\mathrm{p} - \Omega_\mathrm{c}$ is equal to the frequency of radial epicycles κ (equation A7b); this is the outer Lindblad resonance. Had we performed a similar experiment but instead gradually made the circular orbit *smaller* than R_g, the star would begin to orbit faster than Ω_p, until $\Omega_\mathrm{p} - \Omega_\mathrm{c}$ coincided with $-\kappa$; this is the inner Lindblad resonance. In summary, for epicyclic orbits we can write the resonance condition as

$$m\,(\Omega_\mathrm{p} - \Omega_\mathrm{c}) = n\kappa, \tag{55}$$

with $n = 0, 1, -1$ for CR, OLR and ILR respectively. For illustration of these resonant orbits, see e.g. Figure 6.10 of BT08.

1. Example: orbits perturbed by a spiral potential

To illustrate the morphology of perturbed orbits and the importance of resonances, consider the motion of test stars in the potential $\Phi = \overline{\Phi} + \delta\Phi$, where $\overline{\Phi}$ is the logarithmic halo potential

$$\overline{\Phi}(R) = V_0^2 \ln(R/R_0), \tag{56}$$

and R_0 is an arbitrary scale length. This potential gives rise to a *flat rotation curve*, meaning that all stars on circular orbits ($J_R = 0$) have the same orbital speed: $v_\varphi^2 = R\partial\overline{\Phi}/\partial R = V_0^2 = \mathrm{cst}$. This is in fact a reasonably realistic model for many galactic disks[3]. For $\delta\Phi$, we choose the potential of a steadily-rotating *logarithmic spiral*

$$\delta\Phi(\phi, R, t) = -\varepsilon V_0^2 \cos[\alpha \ln(R/R_0) - m(\phi - \Omega_\mathrm{p}t - \phi_\mathrm{p})], \qquad \alpha \equiv m\cot p. \tag{57}$$

This corresponds to a spiral with m arms, pattern speed Ω_p, and pitch angle p (so $p \to 0$ is a very tightly wrapped trailing spiral, which gets less and less tightly wound as $p \to 90°$ — for much more detail, see §6.1 of BT08). The rotation period of this spiral is simply, $T_\mathrm{p} = 2\pi/\Omega_\mathrm{p}$. In Figure 12 we plot contours of $\delta\Phi/(\varepsilon V_0^2)$ for $m = 2$ and $p = 60°$ (a rather loosely wrapped arm).

We now fix the pattern speed to be $\Omega_\mathrm{p} = 0.154\,V_0/R_0$, so that the corotation radius of circular orbits is $R_\mathrm{CR} = 6.5R_0$. We take the spiral strength to be $\varepsilon = 0.01$, and the initial phase $\phi_\mathrm{p} = 0$. We then initialized the orbits of fifteen test stars in this potential, putting each star on a circular orbit with speed V_0 and azimuthal location $\phi(t = 0) = 0$. The

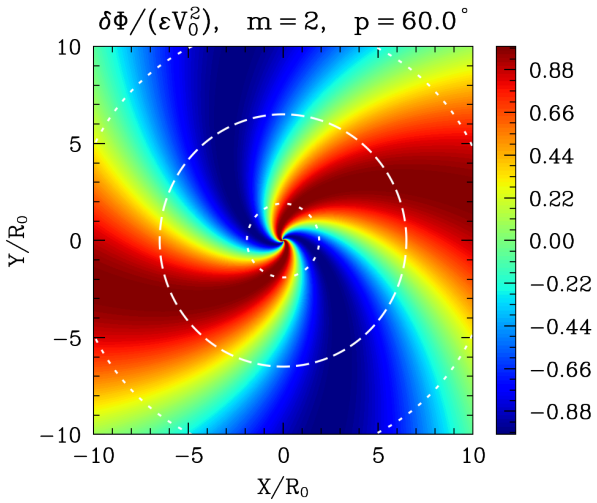

FIG. 12. Contours of the spiral potential, equation (57), for $m = 2$ and $p = 60^\circ$ (the azimuthal phase is arbitrary). The dotted and dashed white curves show the radial locations of the ILR, CR and OLR respectively assuming $\Omega_\mathrm{p} = V_0/(6.5R_0)$, but this is only meant to guide the eye — the logarithmic spiral (57) and the logarithmic halo (56) are both scale-free, so nothing about the figure would change if we modified Ω_p apart from the radii of these white circles.

initial conditions of the stars therefore differed only in their initial orbital radii $R(0)$, which we sampled linearly between R_ILR and R_OLR inclusive. We integrated each orbit for a total time $20T_\mathrm{p}$.

In Figure 13 we show the resulting orbits. In each panel, the black line shows the trajectory of a star whose initial radius $R(0)$ is given at the top of the panel. We plot the trajectory in the (X',Y') frame, which is initially aligned with the inertial $(X,Y) = (R\cos\phi, R\sin\phi)$ frame of the galaxy but rotates relative to that frame in the ϕ direction at rate Ω_p (so the spiral potential (57) is fixed in the (X',Y') frame). In each panel we also show the corotation radius of near-circular orbits R_CR with a dashed red line, and the corresponding Lindblad radii $R_\mathrm{ILR/OLR}$ with dotted red lines.

These panels reveal that initially circular orbits are not dramatically distorted if the star happens to sit far from a resonance. In this case, the orbits can be well described with linear perturbation theory which we introduce in §IV E. On the other hand, the orbits can be changed qualitatively if they are located near a resonance. In particular, panel (a) shows an orbit which is 'trapped' at the ILR, which is the key resonance responsible for e.g. the formation of stellar bars. Meanwhile, panels (g), (h) and (i) show orbits that are trapped at corotation. When viewed in the rotating frame (X',Y') these orbits do not complete any azimuthal loops, but are rather 'stuck' on one side of the spiral, and librate back and forth in the trough of the spiral potential. Trapped orbits at corotation are especially important for determining the slowdown rate of spinning stellar bars (§VIII B 1), the saturation of spiral instabilities (§VIII B 2), and so on. (They are also very familiar to plasma physicists, who are used to particles 'bouncing' in the troughs of waves if the wave's phase velocity matches the particle velocity). Finally, panel (o) shows an orbit trapped at the OLR.

The rest of this section is devoted to a mathematical description of how stellar orbits respond to a perturbation of the form (51). We start by discussing a quantity which is rigorously conserved for such orbits (§IV D). We then describe how one can approximate the evolution of linearly perturbed orbits (§IV E). Finally, we show how the non-linearly distorted orbits that can develop near resonances can be treated analytically in the pendulum approximation (§IV F).

D. Jacobi integral

One thing we can immediately say about test star motion in the Hamiltonian (49) is that it preserves neither angular momentum L or energy E (since $\delta\Phi$ is both non-axisymmetric and time-dependent). Yet, it *does* conserve the Jacobi integral (see §3.3.2 of BT08):

$$E_\mathrm{J} \equiv E - \Omega_\mathrm{p} L. \tag{58}$$

This follows because $H_\mathrm{J} \equiv H - \Omega_\mathrm{p} L$ is just the Hamiltonian in the frame co-rotating with the perturbation, and in that frame the Hamiltonian is time-independent and hence conserved. Since $E_\mathrm{J} = \mathrm{cst.}$, any interaction of a test star

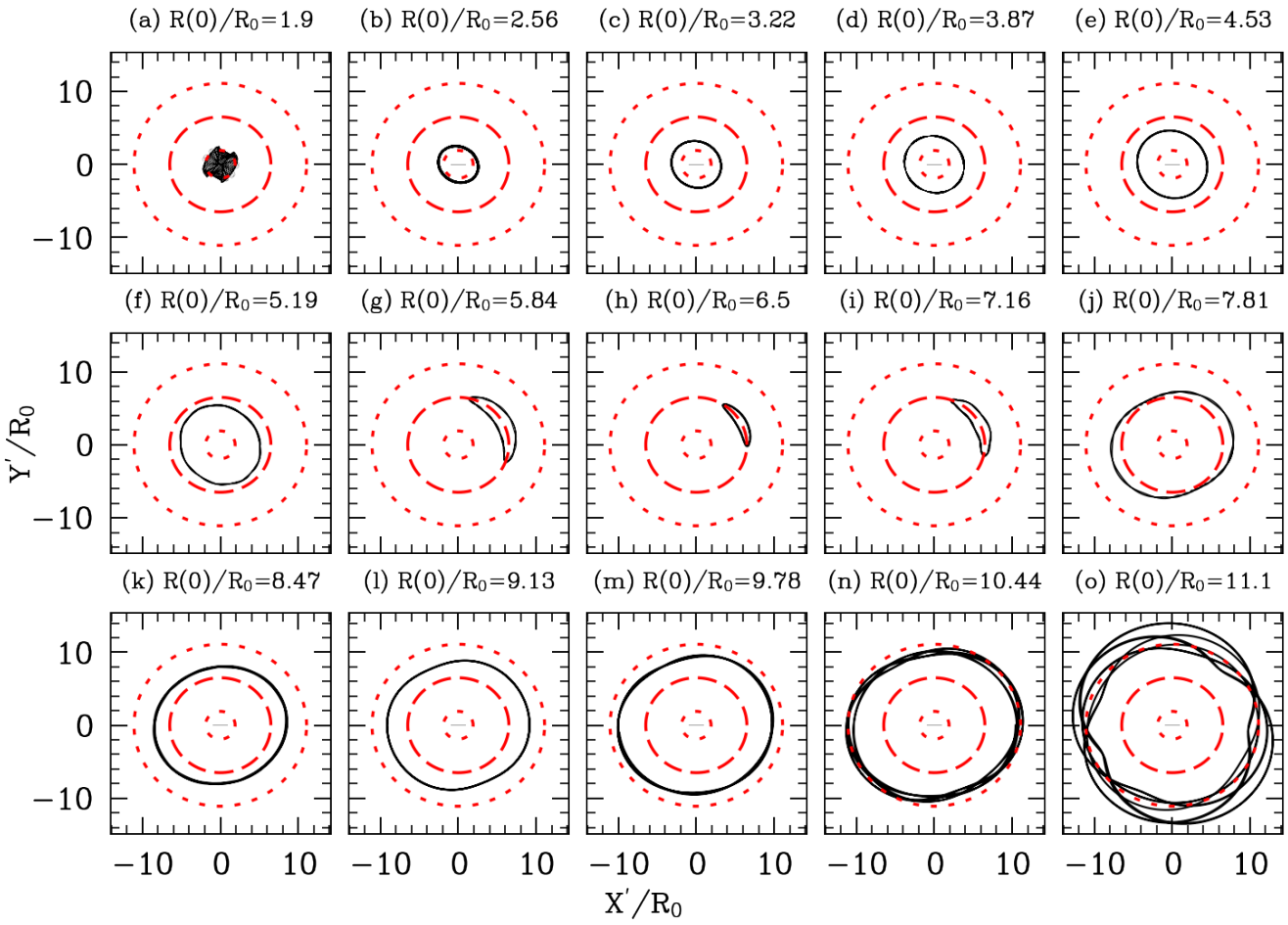

FIG. 13. Orbits of stars in the combined logarithmic halo (56) and spiral (57) potentials (in the frame rotating with the spiral). All stars are initially on circular orbits with speed V_0 but differ in their initial orbital radii $R(0)$. The dotted and dashed lines show the Lindblad and corotation radii, as in Figure 12. In particular, in panel (a) the orbit is trapped at the ILR, while in panels (g)–(i), the orbit is trapped at the CR.

with a constant amplitude, rigidly rotating potential perturbation must produce changes to the star's energy ΔE and angular momentum ΔL that are related by

$$\Delta E = \Omega_{\rm p}\,\Delta L. \tag{59}$$

If ΔE is sufficiently small, we can approximate it by $\Delta E \approx (\partial E/\partial L)\Delta L + (\partial E/\partial J_R)\Delta J_R = \Omega_\phi \Delta L + \Omega_R \Delta J_R$, following the generic definition of orbital frequencies in (32a). This allows us to relate the change in radial action ΔJ_R to the change in angular momentum ΔL:

$$\Delta J_R \approx \frac{\Omega_{\rm p} - \Omega_\phi}{\Omega_R}\Delta L. \tag{60}$$

For instance, stars which corotate with the perturbation, $\Omega_\phi = \Omega_{\rm p}$, do not experience changes to their radial actions to first order.

One can generalize the conservation of the Jacobi integral to potential perturbations of the form $A(t)\delta\Phi(\phi - \Omega_{\rm p} t)$, such as a rigidly-rotating spiral with a time-varying amplitude. In this case, the Hamiltonian is time-dependent even in the rotating frame, so $H_{\rm J}$ is not conserved instantaneously. But the *change* in $H_{\rm J}$ between $t = 0$ and $t = T$ will be $\Delta H_{\rm J}(T) \propto \int_0^T {\rm d}t\, {\rm d}A/{\rm d}t = A(T) - A(0)$. This change is obviously zero if e.g. the perturbation is switched on and off again between times 0 and T, i.e. if $A(0) = A(T) = 0$. We will see a concrete example of this in §IV F 2.

E. Linear perturbations

Suppose the perturbation strength $|\delta\Phi|/|H_0| \sim \varepsilon \ll 1$ (just like in equation (57)). Then to zeroth order in ε we have $\delta\mathbf{J}^{(0)} = \mathbf{0}$, i.e. the star's trajectory is just a straight line through angle-action space, $\mathbf{J} = \mathbf{J}_0$ and $\boldsymbol{\theta} = \boldsymbol{\theta}_0 + \boldsymbol{\Omega} t$. Plugging this back in to the right hand side of (52) and integrating forward in time, we get

$$\delta\mathbf{J}^{(1)}(t) = -\mathrm{i}\sum_{\mathbf{n}} \mathbf{n}\,\delta\Phi_{\mathbf{n}}(\mathbf{J}_0)\,\mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}_0}\,\frac{\mathrm{e}^{\mathrm{i}(\mathbf{n}\cdot\boldsymbol{\Omega}(\mathbf{J}_0) - n_\phi\Omega_\mathrm{p})t} - 1}{\mathrm{i}(\mathbf{n}\cdot\boldsymbol{\Omega}(\mathbf{J}_0) - n_\phi\Omega_\mathrm{p})}, \tag{61}$$

Thus to first order in ε the perturbation (51) nudges the star's orbital action by an amount $\delta\mathbf{J}^{(1)}$ which is proportional to ε in magnitude and oscillatory in time, *except* in the case where the star's mean field orbit is in resonance with the perturber, namely when there exists some pair of integers $\mathbf{n} = \mathbf{N} = (N_\phi, N_R)$ such that (53) holds for $\mathbf{J} = \mathbf{J}_0$. In the resonant case, (61) predicts that $\delta\mathbf{J}$ will grow linearly in time, since the $\mathbf{N}^{\mathrm{th}}$ Fourier component of the forcing (51) is constant in the frame moving with the unperturbed stellar orbit.

In reality, it is precisely near these resonant locations that linear perturbation theory will eventually break down. In that case, one should instead turn to an alternative, *nonlinear* description of the dynamics (§IV F).

F. The pendulum approximation

To describe the near-resonant orbits analytically we turn to the 'pendulum approximation'. The main idea of the pendulum approximation is that near resonances, the linear approximations used in §IV E only truly break down along certain special 'resonant directions' in angle-action space, namely those directions where the argument of the exponential in (51) varies slowly for some resonance $\mathbf{n} = \mathbf{N}$. This allows us to average out the oscillatory motion in the 'nonresonant' directions, and thereby reduce the complicated resonant dynamics to the integrable problem of a one-dimensional pendulum.

For definiteness, let us set the number of dimensions to $d = 2$. Then $\boldsymbol{\theta} = (\theta_\phi, \theta_R)$ and $\mathbf{J} = (J_\phi, J_R)$. We now make a canonical transformation to a new set of coordinates, which are again angle-action coordinates of the unperturbed problem[34,40]. Precisely, we map $(\boldsymbol{\theta}, \mathbf{J}) \to (\boldsymbol{\theta}', \mathbf{J}')$, where $\boldsymbol{\theta}' = (\theta_\mathrm{f}, \theta_\mathrm{s})$ consists of the 'fast' and 'slow' angles

$$\theta_\mathrm{f} \equiv \theta_R, \tag{62a}$$
$$\theta_s \equiv \mathbf{N}\cdot\boldsymbol{\theta} - N_\phi\Omega_\mathrm{p} t, \tag{62b}$$

and $\mathbf{J}' = (J_\mathrm{f}, J_\mathrm{s})$ consists of the corresponding fast and slow actions

$$J_\mathrm{f} \equiv J_R - \frac{N_R}{N_\phi} L, \tag{63a}$$
$$J_\mathrm{s} \equiv \frac{L}{N_\phi}. \tag{63b}$$

(In plasma, the analogue of e.g. equation (62) would be a coordinate transform $x' = x - v_\mathrm{ph} t$, where v_ph is the phase velocity of a wave moving along the x axis). Having made this transformation, we may rewrite the Hamiltonian H (equations (49)-(50)) in terms of the new coordinates:

$$H(\boldsymbol{\theta}', \mathbf{J}') = H_0(\mathbf{J}') - N_\phi\Omega_\mathrm{p} J_\mathrm{s} + \sum_{\mathbf{k}\neq\mathbf{0}} \Psi_{\mathbf{k}}(\mathbf{J}')\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\boldsymbol{\theta}'}, \tag{64}$$

where $\mathbf{k} = (k_\mathrm{f}, k_\mathrm{s})$ is a vector of integers and we have expanded $\delta\Phi$ as a Fourier series in the new angles $\boldsymbol{\theta}'$, i.e. written $\delta\Phi(\mathbf{r}, t) = \sum_{\mathbf{k}} \Psi_{\mathbf{k}}(\mathbf{J}')\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\boldsymbol{\theta}'}$. The coefficients $\Psi_{\mathbf{k}}$ are easily computed for a simple bar or spiral model[41,42]. The special thing about the form (64) of the Hamiltonian is that it has no explicit time dependence (or rather, the time-dependence has been absorbed into the definition of the angle θ_s; see equation 62).

The fast angles θ_f evolve on the orbital timescale whereas θ_s evolves on the much longer timescale $\sim (\mathbf{N}\cdot\boldsymbol{\Omega} - N_\phi\Omega_\mathrm{p})^{-1}$. Thus we choose to average H over the unimportant fast motion, and work instead with the simpler Hamiltonian $h \equiv (2\pi)^{-1}\int \mathrm{d}\theta_\mathrm{f} H$ at fixed J_f. We also expand h around the resonance. To do this, we let the resonant action be $\mathbf{J}_\mathrm{res}$ (meaning that equation (53) is satisfied for $\mathbf{J} = \mathbf{J}_\mathrm{res}$). In the transformed coordinates this is $\mathbf{J}'_\mathrm{res} = (J_\mathrm{f}, J_\mathrm{s,res})$, and so we can define

$$I \equiv J_\mathrm{s} - J_\mathrm{s,res}, \qquad \varphi_k = \theta_\mathrm{s} + \arg(\Psi_k)/k, \tag{65}$$

with $\varphi_k \in [-\pi, \pi]$ (and we used the shorthand $\Psi_{(0,k_s)} = \Psi_k$). Expanding h for small I, and assuming that a single Fourier component k dominates the perturbation, we find that at each fixed J_f the dynamics reduces to motion in the 'slow plane' (θ_s, J_s) dictated by the pendulum Hamiltonian

$$h(\varphi, I) = \frac{1}{2} G I^2 - F \cos k\varphi, \tag{66}$$

where $\varphi = \varphi_k$ and

$$G(J_f) \equiv \left. \frac{\partial^2 H_0}{\partial J_s^2} \right|_{J_{s,res}}, \quad F(J_f) \equiv -2 |\Psi_k|_{J_{s,res}}. \tag{67}$$

Here $J_{s,res}(J_f)$ is the resonant value of the slow action at fixed fast action, $|\Psi_k|_{J_{s,res}}$ is the strength of the bar perturbation on resonance, and $\varphi \in [-\pi, \pi]$ is just a phase-shifted slow angle. Note that in contrast with linear perturbation theory (§IV E), in this case if the perturbation strength is $|\delta\Phi|/|H_0| \sim \varepsilon \ll 1$, then the 'perturbing' part of the Hamiltonian (66) has relative strength $|F/h| \sim \mathcal{O}(\varepsilon^{1/2})$, and the timescale for evolution due to this perturbation is therefore $t_{lib} \sim \varepsilon^{-1/2} t_{dyn}$. Appreciating this difference is crucial when it comes to developing an appropriate kinetic theory (see §V E).

The variables (φ, I) are canonical variables for the pendulum Hamiltonian (66). The pendulum moves at constant 'energy' h, either on an untrapped 'circulating' orbit with $h > F$ or a trapped 'librating' orbit with $h < F$ (the separatrix between these two families is $h = F$). The *libration period* for oscillations around $(\varphi, I) = (0, 0)$ is

$$t_{lib} \equiv 2\pi/\omega_{lib} \quad \text{where} \quad \omega_{lib}(J_f) \equiv \sqrt{kFG}. \tag{68}$$

The maximum width in I of the librating 'island' is at $\varphi = 0$, where it spans $I \in [-I_h, I_h]$ and I_h is the *island half-width*:

$$I_h(J_f) \equiv 2\sqrt{F/G}, \tag{69}$$

which we note is independent of k. The contours of this Hamiltonian are plotted in Figure 14 for the case with $k = 1$. The separatrix between librating and circulating orbit families is shown with a black dashed line.

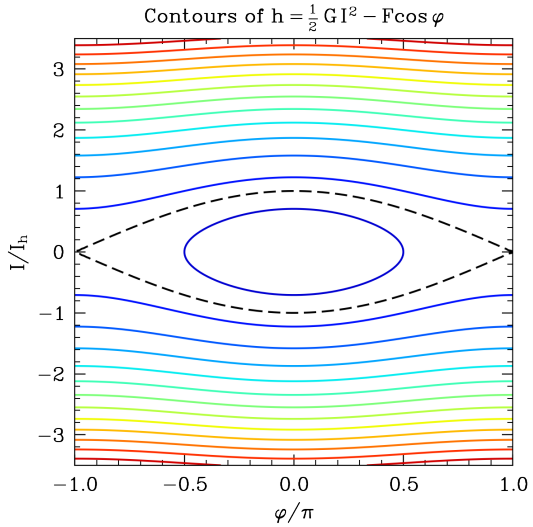

FIG. 14. Pendulum phase space. Colors show contours of constant $h(\varphi, I)$ (see equation 66) for $k = 1$. The dashed black line shows the separatrix between librating (i.e. trapped) and circulating (i.e. untrapped) orbit families.

1. *Example: Corotation resonance, flat rotation curve*

Let us illustrate the pendulum dynamics with a concrete example. This requires stipulation of the resonance **N** under consideration. Let us consider the corotation resonance (CR), as defined in equation (54a), with the resonance

location at the angular momentum value $L = L_{\rm CR}(J_R)$ for a given radial action J_R. Then we have

$$N_\phi = m, \quad N_R = 0, \quad k = 1, \tag{70a}$$
$$J_{\rm s} = L/m, \quad J_{\rm f} = J_R, \quad I = (L - L_{\rm CR})/m, \tag{70b}$$
$$\varphi = m(\theta_\phi - \Omega_{\rm p} t), \quad |\Psi_1| = |\delta\Phi_{(m,0)}|. \tag{70c}$$

We must also stipulate a background potential; the simplest choice is the logarithmic halo potential (56), which gives rise to a flat rotation curve with circular speed V_0. With this choice, for initially circular orbits ($J_{\rm f} = J_R = 0$) we get

$$L = R v_\phi = R V_0, \tag{71a}$$
$$\Omega_{\rm c}(L) = V_0/R = V_0^2/L, \tag{71b}$$
$$G = -N_\phi^2 V_0^2/L_{\rm CR}^2 = -N_\phi^2/R_{\rm CR}^2. \tag{71c}$$

We now combine these two choices, i.e. we consider circular, corotating orbits in a logarithmic halo. In this case the slow action $J_{\rm s}$ is just proportional to the guiding radius R, with the resonant value $J_{\rm s,res}$ occurring at radius at $R_{\rm CR}$, and the slow angle is (up to a phase shift) equivalent to $m[\phi - \int_0^t {\rm d}t' \Omega_{\rm p}(t')]$. Thus, in this case, one can think of the transformation into slow and fast angle-action variables simply as 'moving into the rotating frame' (or in the plasma context, shifting to a frame in which the wave's phase velocity is zero).

With these choices, the libration time (68) can be estimated as

$$t_{\rm lib} \sim \frac{t_{\rm dyn}}{m\varepsilon^{1/2}} = \frac{2\pi}{m\varepsilon^{1/2}\Omega_{\rm p}} \tag{72a}$$
$$\approx 1\,{\rm Gyr} \times \left(\frac{m}{2}\right)^{-1} \left(\frac{\varepsilon}{0.01}\right)^{-1/2} \left(\frac{t_{\rm dyn}}{200\,{\rm Myr}}\right), \tag{72b}$$

where $t_{\rm dyn} = 2\pi R/V_0$ is the circular orbital period at radius R, and all quantities must be evaluated at the resonance location $R = R_{\rm CR}$. Clearly $t_{\rm lib}$ is always several times longer than the orbital period $t_{\rm dyn}$. Moreover, equation (69) lets us estimate the radial width of the resonant island for these orbits:

$$R_{\rm h} = \frac{m I_{\rm h}}{V_0} \sim 2m\varepsilon^{1/2} R_{\rm CR} \tag{73a}$$
$$\approx 1.6\,{\rm kpc} \times \left(\frac{m}{2}\right) \left(\frac{\varepsilon}{0.01}\right)^{-1/2} \left(\frac{R_{\rm CR}}{6.5\,{\rm kpc}}\right), \tag{73b}$$

Thus the resonant island at corotation can span several kiloparsecs in radius. This must be borne in mind whenever one makes a 'narrow resonance' approximation, since $R_{\rm h}$ can also be comparable to the radial scale over which the galaxy's underlying density profile changes significantly (~ 3 kpc for the stellar disk of the Milky Way).

As a more quantitative example, let us apply our pendulum formalism to the fifteen orbits shown in Figure 13. In Figure 15 we show the trajectory of these fifteen stars through the space of azimuthal angle $\phi - \Omega_{\rm p} t + \ln(R_{\rm CR}/R_0)\cot p$ and angular momentum L. Up to constants, these are equivalent to the slow angle-action pair (φ, I) near corotation. With the colored lines we plot contours of the pendulum Hamiltonian (c.f. Figure 14) using the parameters for the corotation resonance (equations 70-71). We see that the pendulum approximation does a good job of predicting the orbit shapes near corotation. The slight mismatch in the orbit shapes is mostly because the pendulum theory ignores all but one Fourier harmonic k (as such, the agreement with theory gets better (worse) if we use more open (tight) spirals, i.e. take $p > 60°$ ($p < 60°$)). However, we see that the pendulum approximation begins to break down close to the ILR and OLR, where resonant effects not associated with corotation start to play an important role in the dynamics.

2. *Example: Radial migration*

Radial migration is the name given to any dynamical process which takes a star on a circular or nearly-circular orbit in a galactic disk and alters its angular momentum (or equivalently its guiding radius) while producing little or no change in its radial action. The ubiquity of radial migration in galactic disks, and the importance of this especially for the evolution of chemical gradients, was first appreciated by Sellwood and Binney [43]. Here we perform two numerical experiments similar to the ones shown in their paper, in order to demonstrate the difference between linear (§IV E) and nonlinear (§IV F) perturbations.

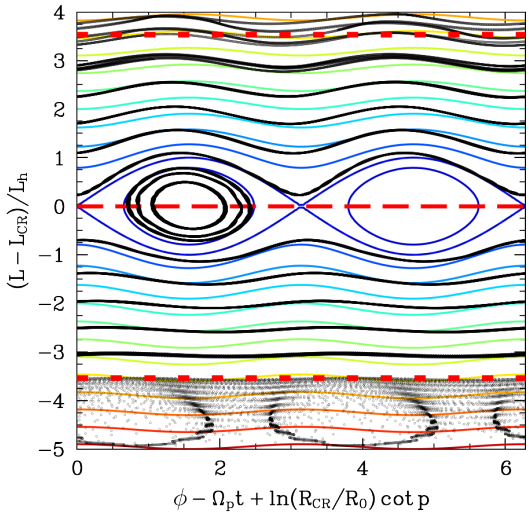

FIG. 15. The same orbits as in Figure 13, but plotted in the space of azimuthal angle (in the rotating frame) and angular momentum (relative to the corotation resonance value). Up to constants, these are equivalent to the slow angle and slow action (equations 62 and 63). The background colored lines are contours of the associated pendulum Hamiltonian (c.f. Figure 14). The dashed and dotted red lines have the same meaning as in Figure 13.

We consider a similar setup to the one we described in §IV C 1 by integrating an array of test particles in the logarithmic halo potential $\overline{\Phi}$ with an accompanying spiral potential perturbation $\delta\Phi$, except this time we modify the potential perturbation (57) by multiplying it by a Gaussian in time:

$$\delta\Phi(\phi, R, t) = -\varepsilon V_0^2 \cos[\alpha \ln(R/R_0) - m(\phi - \Omega_p t - \phi_p)] \times e^{-(t-t_{peak})^2/(2\tau^2)}. \tag{74}$$

This corresponds to a transient spiral which grows and decays on the timescale $\sim \tau$, and whose amplitude reaches its peak at $t = t_{peak}$. All other parameters are set to the same values as in Figure 13. With these parameters we can use equation (72a) to estimate the libration time at peak spiral amplitude to be $t_{lib} \sim 5T_p$. For the initial conditions of the stars we use circular orbits with radii randomly sampled from the range $R(0)/R_0 \in [4, 9]$, which is centered on the corotation radius but does *not* reach as far as the ILR/OLR. We choose the initial azimuthal angles $\phi(0)$ randomly in $[0, 2\pi]$. We integrate for a total time of $10T_p$, and always set the time of peak amplitude to be half-way through the integration, $t_{peak} = 5T_p$.

First, we integrated the equations of motion of 5 000 test stars in this spiral potential, choosing the growth/decay time to be $\tau = 3T_p$. In panels (a) and (b) of Figure 16 we show the resulting evolution of the angular momenta of a randomly selected subset of these stars as a function of azimuthal angle and time. In panel (c) we show a scatter plot of the change in angular momentum, $\Delta L \equiv L(10T_p) - L(0)$, for every star, as well as the mean $\langle \Delta L \rangle$ at each fixed $L(0)$. The red dashed line in these panels shows the corotation resonance while the solid pink lines are displaced from corotation by $\sim L_h$ (equation 69), the approximate resonance width at peak spiral amplitude. We see that close to the resonance, the spiral traps stars and causes them to undergo large shifts in their angular momenta (up to the resonance width $\sim L_h$). As the spiral's amplitude decays it 'drops off' these stars at angular momenta which are often rather different from their initial values. The blue line in panel (c) makes it clear that there is a preference for stars initially inside corotation ($L(0) < L_{CR}$) to receive positive ΔL, and vice versa. To understand this, note that in the limit where the libration time at peak amplitude is very short, a star which was originally located *inside* corotation will librate back and forth across $L = L_{CR}$ several times before being deposited at some new $L = L(0) + \Delta L$. This new L could be either inside or outside L_{CR} with roughly equal probability, so on average the star will have moved *out* ($\Delta L > 0$). Meanwhile, far from the resonance the stars respond roughly adiabatically to the perturbation and so ΔL is clustered around zero. We also mention that in panel (c), many stars gather on the line with slope -2 that passes through $\Delta L = 0$ at corotation. This line corresponds to stars that are initially located at $L(0) = L_{CR} - \ell$ being deposited at $L_{CR} + \ell$[44,45].

Second, in Figure 17 we rerun the same experiment except with a much shorter-lived spiral, $\tau = 0.3T_p$. We see from panel (a) that in this case, no orbits manage to undergo a full libration inside the trapping region. As a result, there is no preferential direction of the transport of stellar angular momenta across the resonance, so the the mean ΔL in panel (c) is approximately zero for all initial $L(0)$. In fact, since the perturbation lasts for so short a time, the resonance location is not so significant, meaning the distribution of ΔL values does not depend particularly strongly

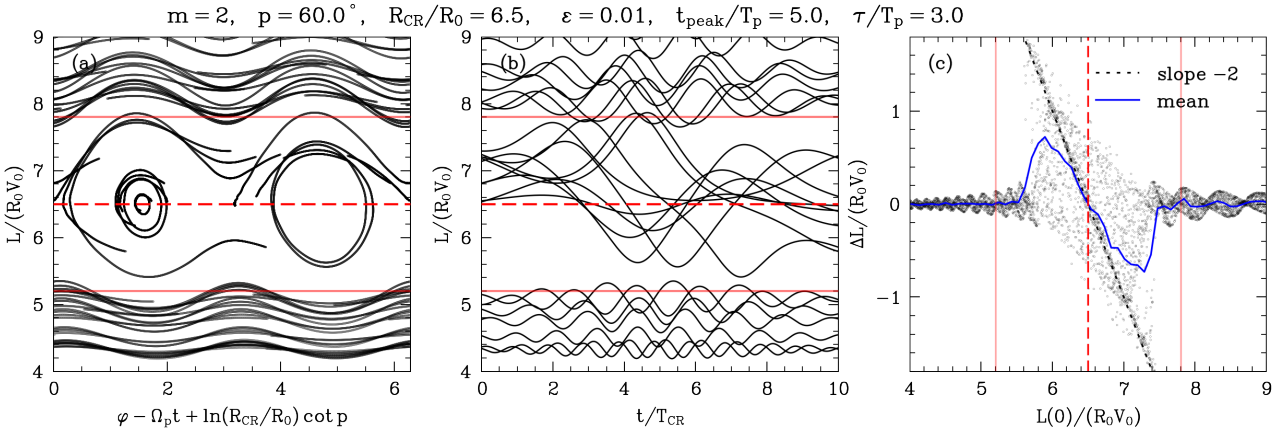

FIG. 16. Radial migration in a galactic disk. We integrate the orbits of 5 000 stars that are initially on circular orbits with random azimuths $\phi(0)$ and radii chosen randomly in the range $R(0)/R_0 \in [4,9]$. The stars are perturbed by a transient spiral potential (74) with lifetime $\sim \tau = 3T_\mathrm{p}$. Panels (a) and (b) show the evolution of the angular momenta of a handful of these stars as functions of azimuthal angle in the rotating frame and time respectively. Panel (c) shows the change in angular momentum experienced by each particle by the end of the simulation as a function of its initial angular momentum $L(0)$. In each panel the red dashed line corresponds to corotation, and the solid pink lines are offset from corotation by approximately the resonance half-width $L_\mathrm{h} = mI_\mathrm{h}$ (equation 69) at peak spiral amplitude.

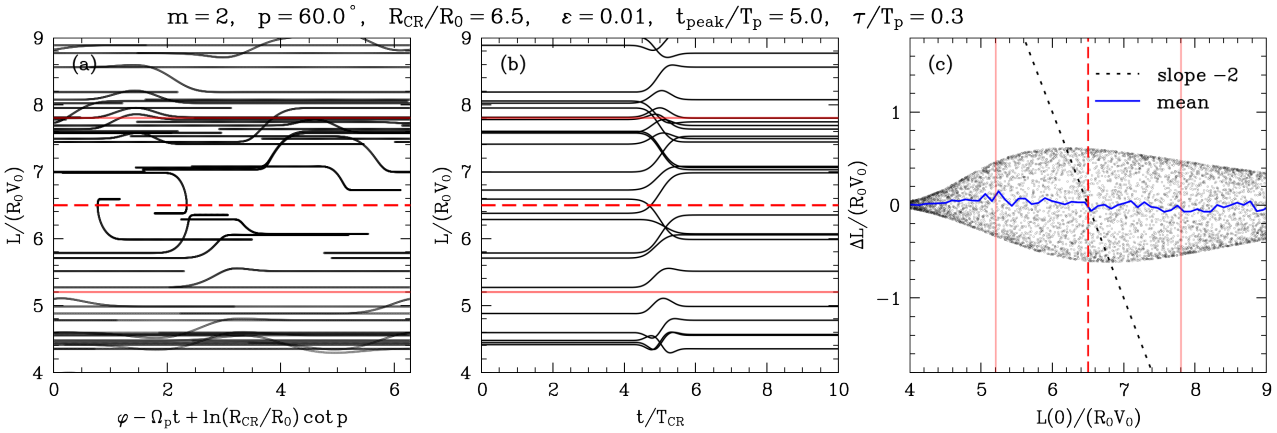

FIG. 17. As in Figure 16, except the spiral grows and decays much more quickly, $\tau = 0.3T_\mathrm{p}$.

on $L(0)$. In fact, the envelope of ΔL values is significantly more broad far away from corotation than it was in Figure 16 (so the variance $\langle(\Delta L)^2\rangle$ is relatively large even though $\langle\Delta L\rangle$ is close to zero).

From the results of these numerical experiments we can extract two key physical principles, which apply ubiquitously in the kinetic theory of both stellar systems and plasmas:

- The efficiency with which different perturbations shuffle the integrals of motion of a distribution of particles depends on the spatio-temporal correlation spectrum of those perturbations (just as in e.g. the classic stochastic acceleration study of Sturrock [46]). It is therefore imperative to develop a theory which can account for the *statistical* evolution of a distribution of particles when exposed to such a perturbation spectrum, be it externally or self-consistently generated. This is what we will do in the rest of this tutorial article, starting in §V.
- If a single coherent perturbation dominates the potential $\delta\Phi$ over a timescale $\gtrsim t_\mathrm{lib}$ then it will trap resonant orbits nonlinearly, and this may drive the system down a substantially different evolutionary path compared to the case with no trapping. But if some process is able to interrupt the trapping process on a timescale $\ll t_\mathrm{lib}$, then nonlinear trapping can be ignored and one can apply linear perturbation theory (§IV E) with confidence. This 'interruption' can be due to the rapid decay of the perturbation as in the example above, or because of a time-dependent Ω_p and the accompanying movement of the resonance location through phase space[41,47], or it can be due to some collisional/diffusive process which kicks particles in and out of the resonance region. We

address the last of these possibilities in §VIII B. We also mention that the interruption of nonlinear trapping by time-dependent wave amplitudes, collisionality, and so on is very commonly studied in plasma theory[48–51].

V. FUNDAMENTALS OF KINETIC THEORY

In §II, we introduced the idea that an arbitrary gravitationally-bound ensemble of stars reaches virial (quasi-)equilibrium after a few dynamical timescales $t_{\rm dyn}$. It then moves through a series of such equilibria on the relaxation timescale $t_{\rm relax} \gg t_{\rm dyn}$. Then, in §§III-IV, we saw how to describe the quasiperiodic orbits of stars in smooth, regular, time-independent potentials $\overline{\Phi}$, became familiar with the machinery of angle-action variables, and discussed the modification of orbits by some perturbation $\delta\Phi$. Here we will set up the remainder of the article by bringing these two sets of ideas together. Our aim is to describe a quasi-equilibrium stellar system as a collection of quasiperiodic orbits, and then to develop a quantitative theory of how fluctuations in the true gravitational potential cause that quasi-equilibrium to change slowly over time.

A. Klimontovich description

We start with the microscopic *Klimontovich* (DF), or empirical DF,

$$f_{\rm M}(\mathbf{r},\mathbf{v},t) = m\sum_{i=1}^{N}\delta\big(\mathbf{r}-\mathbf{r}_i(t)\big)\delta\big(\mathbf{v}-\mathbf{v}_i(t)\big). \tag{75}$$

(Here and in what follows, the subscript 'M' stands for 'microscopic'). This function encodes the exact position $\mathbf{r}_i(t)$ and velocity $\mathbf{v}_i(t)$ of every star i at time t, which are themselves the exact solutions to the Newtonian equations of motion (2). Therefore, for indistinguishable particles, the Klimontovich function (75) contains complete information about the system. Of course, at this stage it is totally useless because if we knew all this information we would not need kinetic theory. Nevertheless, describing the system in this way is what will allow us to most conveniently formulate that kinetic theory.

Note that f is normalized such that $\int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v} f = Nm = M$, the total mass of the stars. Thus, if there is no external forcing, the exact microscopic gravitational potential is

$$\Phi_{\rm M}(\mathbf{r},t) = \int \mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{v}'\,\psi(\mathbf{r},\mathbf{r}')f_{\rm M}(\mathbf{r}',\mathbf{v}',t), \tag{76}$$

where $\psi(\mathbf{r},\mathbf{r}') = -G/|\mathbf{r}-\mathbf{r}'|$. This is simply the formal solution of the Poisson equation

$$\nabla^2\Phi_{\rm M} = 4\pi G\rho_{\rm M}, \tag{77}$$

where the microscopic mass density is $\rho_{\rm M} = \int \mathrm{d}\mathbf{v}\, f_{\rm M}$. It follows that every particle i satisfies Hamilton equations (30), namely

$$\frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t} = \frac{\partial H_{\rm M}(\mathbf{r}_i,\mathbf{v}_i,t)}{\partial \mathbf{v}_i}, \tag{78a}$$

$$\frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t} = -\frac{\partial H_{\rm M}(\mathbf{r}_i,\mathbf{v}_i,t)}{\partial \mathbf{r}_i}, \tag{78b}$$

where $H_{\rm M}$ is the microscopic one-particle Hamiltonian

$$H_{\rm M}(\mathbf{r},\mathbf{v},t) = \tfrac{1}{2}\mathbf{v}^2 + \Phi_{\rm M}(\mathbf{r},t), \tag{79}$$

Using equations (75)-(79) it is now easy to check that $f_{\rm M}(\mathbf{r},\mathbf{v},t)$ exactly satisfies the *Klimontovich equation*

$$\frac{\mathrm{d}f_{\rm M}}{\mathrm{d}t} = \frac{\partial f_{\rm M}}{\partial t} + \mathbf{v}\cdot\frac{\partial f_{\rm M}}{\partial \mathbf{r}} - \frac{\partial \Phi_{\rm M}}{\partial \mathbf{r}}\cdot\frac{\partial f_{\rm M}}{\partial \mathbf{v}} = 0, \tag{80}$$

which we notice can also be written in a pleasingly coordinate-free form as

$$\frac{\partial f_{\rm M}}{\partial t} + [f_{\rm M}, H_{\rm M}] = 0. \tag{81}$$

This form of the equation is particularly convenient when we wish to perform a canonical coordinate transform, e.g. from position and velocity to angle-action variables. Physically, equation (81) just says that each infinitesimal 'pixel' of phase-space fluid is invariant if one follows it along the trajectory prescribed by the exact Hamiltonian $H_{\rm M}$ (Liouville's theorem).

For later use we mention that, since the transformation to angle-action coordinates (§III C) is canonical, it conserves the infinitesimal volumes $\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v} = \mathrm{d}\boldsymbol{\theta}\mathrm{d}\mathbf{J}$ and so the Klimontovich DF (75) can be expressed as

$$f_{\rm M}(\boldsymbol{\theta},\mathbf{J},t) = m\sum_{i=1}^{N}\delta\big(\boldsymbol{\theta}-\boldsymbol{\theta}_i(t)\big)\delta\big(\mathbf{J}-\mathbf{J}_i(t)\big). \tag{82}$$

B. Angle-averaged and angle-dependent quantities

We now take the formal step of splitting all quantities into an angle-averaged part and an angle-dependent 'fluctuation' (which at this stage is not necessarily small). Thus we write the microscopic DF as

$$f_{\rm M} = f_0 + \delta f, \tag{83}$$

where

$$f_0(\mathbf{J},t) \equiv \int \frac{\mathrm{d}\boldsymbol{\theta}}{(2\pi)^d} f_M(\boldsymbol{\theta},\mathbf{J},t). \tag{84}$$

Obviously if we were to convert back to real space coordinates, then f_0, for a galactic disc, would be axisymmetric (independent of ϕ), while all non-axisymmetric fluctuations would be contained in δf. The equivalent in a homogeneous plasma would be a DF f_0 that depended only on velocity, with all position-dependent fluctuations carried in δf. Similarly, the scalar potential $\Phi_{\rm M}$ as defined in (76) is a linear functional of the DF $f_{\rm M}$, as is the microscopic Hamiltonian $H_{\rm M}$ defined in (79). Thus, in correspondence with (83) we may write

$$H_{\rm M} = H_0 + \delta\Phi. \tag{85}$$

By construction, H_0 is the part of the Hamiltonian that only depends on actions $\mathbf{J}$, while any $\boldsymbol{\theta}$-dependence (and hence any nonaxisymmetry of the potential) is contained in $\delta\Phi$ (see §IV A). In a homogeneous system, $H_0 = v^2/2$.

We note the subtlety that even though all non-axisymmetry of the DF $f_{\rm M}$ is contained in δf, the quantity δf may also contain an axisymmetric piece. Related to this is the fact that f_0 is *not* equivalent to the axisymmetric part of $f_{\rm M}$ (just like $\Phi_0 \neq \overline{\Phi}$ in equation 47). This means that, for example, one cannot in general construct the axisymmetric part of the exact surface density distribution $\overline{\Sigma}_{\rm M}(R) = (2\pi)^{-1}\int \mathrm{d}\phi \int \mathrm{d}\mathbf{v} f_{\rm M}$ of a disk from $f_0(\mathbf{J})$ alone: one also needs to know the θ_R-dependent part of the DF.

Let us briefly consider what happens if there are no non-axisymmetric fluctuations (which in reality would require $N \to \infty$). Setting $\delta f = \delta\Phi = 0$ in (83)-(85) and plugging the results into (81), we get[52]

$$\frac{\partial f_0}{\partial t} + [f_0, H_0] = 0. \tag{86}$$

The left hand side of this equation can be thought of as a 'mean field operator' which advects particles along their mean field trajectories; the equation therefore tells us that in the absence of fluctuations, the phase space density along mean field trajectories is conserved, as we would expect. In fact, since f_0 and H_0 only depend on $\mathbf{J}$, the Poisson bracket is zero (see equation (34)):

$$[f_0, H_0] = 0. \tag{87}$$

This means that $\partial f_0/\partial t = 0$, i.e. f_0 cannot change with time even at a fixed phase space location. Put differently, any f_0 which is only a function of $\mathbf{J}$ will be an equilibrium distribution in the absence of fluctuations, or what is called a *collisionless equilibrium*, or equivalently a quasi-stationary distribution. In a homogeneous stellar system or electrostatic plasma, equation (87) is equivalent to the requirement that $[f_0, \mathbf{v}^2] = 0$, which is trivially true of any f_0 which only depends on velocity. Of course, in this case and in the case of inhomogeneous stellar systems, the fact that f_0 satisfies (87) does not mean that it is a *stable* equilibrium; in general, determining whether some f_0 is linearly stable requires the linear response theory that we will develop in §VI.

Let us now reinstate the fluctuations, in particular allowing f_0 to be time-dependent. Plugging the expansions (83), (85) into (81) and writing the Poisson bracket out explicitly as in (34), we get, without any approximations,

$$\frac{\partial f_0}{\partial t} = \frac{\partial}{\partial \mathbf{J}} \cdot \left(\int \frac{\mathrm{d}\boldsymbol{\theta}}{(2\pi)^d} \delta f \frac{\partial \delta\Phi}{\partial \boldsymbol{\theta}} \right), \tag{88a}$$

$$\frac{\partial \delta f}{\partial t} + \boldsymbol{\Omega} \cdot \frac{\partial \delta f}{\partial \boldsymbol{\theta}} - \frac{\partial \delta\Phi}{\partial \boldsymbol{\theta}} \cdot \frac{\partial f_0}{\partial \mathbf{J}} = -\frac{\partial}{\partial \boldsymbol{\theta}} \cdot \left(\delta f \frac{\partial \delta\Phi}{\partial \mathbf{J}} \right) + \frac{\partial}{\partial \mathbf{J}} \cdot \left(\delta f \frac{\partial \delta\Phi}{\partial \boldsymbol{\theta}} - \int \frac{\mathrm{d}\boldsymbol{\theta}}{(2\pi)^d} \delta f \frac{\partial \delta\Phi}{\partial \boldsymbol{\theta}} \right), \tag{88b}$$

where $\boldsymbol{\Omega}(\mathbf{J}) \equiv \partial H_0 / \partial \mathbf{J}$. Equation (88a) tells us how the (nonlinear) coupling of nonaxisymmetric 'fluctuations' δf and $\delta\Phi$ drives evolution of the angle-averaged part of the DF f_0. Note that this equation takes the form of a continuity equation in action space, as it must, since we are not allowing stars to be created or destroyed. Meanwhile, equation (88b) tells us how the fluctuating part of the DF evolves over time.

If we are doing the problem self-consistently then we must also couple these two equations to the Poisson equation for the fluctuations, namely

$$\delta\Phi(\boldsymbol{\theta}, \mathbf{J}, t) = \int \mathrm{d}\boldsymbol{\theta}' \mathrm{d}\mathbf{J}' \, \psi(\boldsymbol{\theta}, \mathbf{J}, \boldsymbol{\theta}', \mathbf{J}') \delta f(\boldsymbol{\theta}', \mathbf{J}', t). \tag{89}$$

Of course in plasma, the interaction kernel ψ has no dependence on the 'actions' (velocities), which makes the Poisson equation much easier to solve in practice.

C. Fourier expansion

At this point it is convenient to introduce a Fourier representation for our angle-dependent quantities. That is, we use the periodicity of the angle variables $\boldsymbol{\theta}$ to express δf and $\delta\Phi$ as Fourier series, like in equation (35), namely

$$\delta f(\boldsymbol{\theta}, \mathbf{J}, t) = \sum_{\mathbf{n}} \delta f_{\mathbf{n}}(\mathbf{J}, t) \, \mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}}, \tag{90a}$$

$$\delta\Phi(\boldsymbol{\theta}, \mathbf{J}, t) = \sum_{\mathbf{n}} \delta\Phi_{\mathbf{n}}(\mathbf{J}, t) \, \mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\boldsymbol{\theta}}, \tag{90b}$$

Let us now rewrite (88a) and (88b) in Fourier form so that

$$\frac{\partial f_0}{\partial t} = -\frac{\partial}{\partial \mathbf{J}} \cdot \sum_{\mathbf{n}} \mathrm{i}\mathbf{n} \delta\Phi^*_{\mathbf{n}}(\mathbf{J}, t) \delta f_{\mathbf{n}}(\mathbf{J}, t), \tag{91a}$$

$$\frac{\partial \delta f_{\mathbf{n}}}{\partial t} + \mathrm{i}\mathbf{n} \cdot \boldsymbol{\Omega} \delta f_{\mathbf{n}} - \mathrm{i}\mathbf{n} \cdot \frac{\partial f_0}{\partial \mathbf{J}} \delta\Phi_{\mathbf{n}} = \mathrm{i} \sum_{\mathbf{n}'} \mathbf{n}' \cdot \left[\frac{\partial \delta f_{\mathbf{n}-\mathbf{n}'}}{\partial \mathbf{J}} \delta\Phi_{\mathbf{n}'} - \frac{\partial \delta\Phi_{\mathbf{n}-\mathbf{n}'}}{\partial \mathbf{J}} \delta f_{\mathbf{n}'} \right]. \tag{91b}$$

Meanwhile the Fourier-transformed version of (89) reads

$$\delta\Phi_{\mathbf{n}}(\mathbf{J}, t) = (2\pi)^d \sum_{\mathbf{n}'} \int \mathrm{d}\mathbf{J}' \, \psi_{\mathbf{n}\mathbf{n}'}(\mathbf{J}, \mathbf{J}') \, \delta f_{\mathbf{n}'}(\mathbf{J}', t). \tag{92}$$

The (bare) coupling coefficients, $\psi_{\mathbf{n}\mathbf{n}'}(\mathbf{J}, \mathbf{J}')$, encode, in essence, Poisson's equation in angle-action space. In the limit of an homogeneous multi-periodic plasma (§III C 1), these coefficients are greatly simplified and become $\psi_{\mathbf{n}\mathbf{n}'} \propto \delta^{\mathbf{n}'}_{\mathbf{n}} / |\mathbf{n}|^2$. In that limit, note that (i) the only 'resonances' permitted are 1:1 resonances, i.e. $\mathbf{n} = \mathbf{n}'$; and (ii) the coupling between particles does not depend on the actions $(\mathbf{J}, \mathbf{J}')$. The fact that neither (i) or (ii) hold in general for realistic stellar systems makes linear response calculations much harder than in homogeneous plasmas (§VI).

D. Linear and quasilinear approximations

The only assumption we have made in deriving equations (91a)-(92) is that H_0 is integrable, i.e. that there exists a set of global angle-action coordinates $(\boldsymbol{\theta}, \mathbf{J})$ such that $H_0 = H_0(\mathbf{J})$. If this is true then equations (91a)-(92) are exact. Sadly, they are also very difficult to solve.

To render them solvable, we will often make an additional assumption that *angle-dependent fluctuations are small*, i.e. $\delta f \ll f_0$ and $\delta\Phi \ll H_0$. With this, it follows naturally from equations (91a)-(91b) that the fluctuation δf evolves

on the dynamical timescale $t_{\rm dyn} \sim 1/|\mathbf{\Omega}|$, whereas the angle-averaged DF f_0 — whose rate of change is *quadratic* in the small fluctuations — evolves on some much longer timescale $t_{\rm relax} \gg t_{\rm dyn}$. This timescale separation is what allows the system to move through a series of virial equilibria, each of which approximately satisfies equation (86), as we described qualitatively in §II C 3. Of course for self-consistent problems, $H_0[f_0]$ will also be changing on the timescale $t_{\rm relax}$, and in practice every time H_0 changes significantly, we have to recompute the mapping $(\mathbf{r},\mathbf{v}) \to (\boldsymbol{\theta},\mathbf{J})$.

Our assumption of weak fluctuations means that we now have a well-defined programme for solving the system (91a)-(92). On the timescale $\sim t_{\rm dyn}$, we may *linearize* equation (91b) by throwing away the small terms on the right hand side:

$$\frac{\partial \delta f_{\mathbf{n}}}{\partial t} + \mathrm{i}\mathbf{n}\cdot\mathbf{\Omega}\delta f_{\mathbf{n}} - \mathrm{i}\mathbf{n}\cdot\frac{\partial f_0}{\partial \mathbf{J}}\delta\Phi_{\mathbf{n}} = 0, \tag{93}$$

and then solving under the assumption that f_0 is fixed. If we are doing a self-consistent problem we must couple this linearized equation with the Poisson equation (92). The simultaneous solution of these two linear equations is a nontrivial problem, and is the subject of §VI. Supposing we can do this successfully, we can then plug the solutions δf, $\delta\Phi$ into the right hand side of (91a) — or, if appropriate, its ensemble-averaged cousin (118) — and solve for the evolution of f_0 on the much longer timescale $t_{\rm relax}$. Such evolution of f_0 is called *quasilinear*, since it involves the multiplication of two small quantities, each of which is determined using linear theory.

We emphasize that the linear equation (93) is just the kinetic version of the linear perturbation theory that we wrote down for the orbits of individual stars in §IV E. To see this, note that the formal solution to (93) is just

$$\delta f_{\mathbf{n}}(\mathbf{J},t) = \delta f_{\mathbf{n}}(\mathbf{J},0)\,\mathrm{e}^{-\mathrm{i}\mathbf{n}\cdot\mathbf{\Omega}t} + \mathrm{i}\mathbf{n}\cdot\frac{\partial f_0}{\partial \mathbf{J}}\int_0^t \mathrm{d}t'\,\mathrm{e}^{-\mathrm{i}\mathbf{n}\cdot\mathbf{\Omega}(t-t')}\delta\Phi_{\mathbf{n}}(\mathbf{J},t'). \tag{94}$$

The first term on the right hand side represents the ballistic evolution of some initial fluctuation $\delta f(\boldsymbol{\theta},\mathbf{J},0)$ governed by the mean field Hamiltonian H_0, i.e. the *phase mixing of unperturbed orbits* $\mathbf{J} = \text{cst}$, $\boldsymbol{\theta} = \mathbf{\Omega}t + \text{cst}$. Phase mixing causes the initial fluctuation to oscillate rapidly in action (or, more properly, frequency) space for times $t \gg 1/|\mathbf{\Omega}|$ — see Figure 18 for illustration. (Analogously in plasma, phase mixing consists of particles moving along $\mathbf{v} = \text{cst}$,

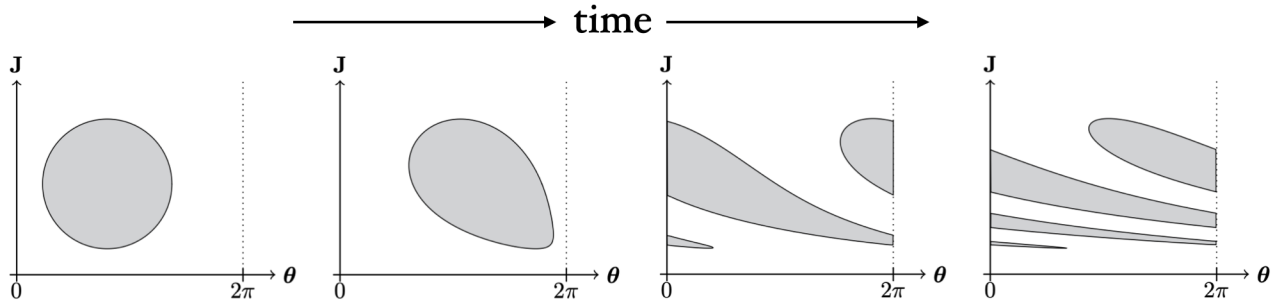

FIG. 18. Schematic illustration of phase mixing in angle-action space as a function of time. The gray blob represents some phase space DF fluctuation δf evolving in the absence of external forces (first term on the right hand side of equation 94). Unperturbed trajectories follow horizontal lines in angle-action space (Figure 10b). Since the orbital frequencies $\mathbf{\Omega}(\mathbf{J})$ depend on $\mathbf{J}$, particles of different actions dephase. This leads to the appearance of ever finer structures in phase space.

$\mathbf{r} = \mathbf{v}t + \text{cst}$ and therefore $\delta f(\mathbf{r},\mathbf{v},0)$ winds up to ever smaller scales in *velocity* space). The second term on the right hand side of (94) tells us how the DF is modified by the potential perturbation $\delta\Phi$ *as evaluated along the unperturbed trajectory*; in other words it is the result of applying an equation like (61) to every star. Put differently, in any linear or quasilinear kinetic theory, the only particle trajectories that are accounted for are the ballistic unperturbed trajectories set by H_0 plus the small sinusoidal wiggles instigated by $\delta\Phi$.

E. Nonlinear and nonperturbative effects

One can imagine iterating the above procedure to higher orders in perturbation theory by plugging (94) back into the right hand side of (91b), calculating $\delta f_{\mathbf{n}}$ to second order in the perturbation strength, and so on[53]. At the level of particle orbits this would correspond to adding 'sinusoidal disturbances on top of sinusoidal disturbances' over and over with ever smaller amplitude. In the case where the perturbations $\delta\Phi$ are stochastic, this approach is usually given the name *weak turbulence theory*. Weak turbulence is the subject of many a plasma monograph such as the excellent

one by Kadomtsev[54], and underlies the resonance broadening theories pioneered by Dupree[55] and others[56–58]. On the other hand, in the case where $\delta\Phi$ is deterministic, the second order calculation leads e.g. to the classic *plasma echo* phenomenon first pointed out by Gould *et al.*[59].

However, no matter how many iterations of this perturbative scheme one performs, it cannot account for fundamentally nonperturbative effects such as particle trapping by strong potential perturbations (§IV F). One way to see this is to note that the perturbative iteration scheme accounts for effects that are of order ε, ε^2, ε^3, etc., where $\varepsilon \sim |\delta\Phi/H_0|$ measures the strength of the potential perturbations. Nonlinear trapping, though, enters the dynamics at order $\varepsilon^{1/2}$ (§IV F), and so cannot be recovered by summing any finite number of perturbative terms. Another way to arrive at the same conclusion is to note that trapped orbits (like the one in Figure 13i) are topologically distinct from untrapped ones (c.f. Figure 13j), and no amount of iterative sinusoidal deformations of the latter can ever reproduce the former.

These fundamentally nonperturbative effects can be very important for the evolution of the system as a whole, but a kinetic theory which accounts for them cannot start from equation (93). In the case where $\delta\Phi$ is stochastic, this regime is termed *strong turbulence*, and formulating an accurate kinetic theory turns out to be very difficult even in the simplest plasmas[56]. Luckily, in galaxies, $\delta\Phi$ is often dominated by a single coherent perturbation, such as a strong bar or spiral arm. In this case a simple kinetic theory incorporating trapping *can* be developed using the pendulum approximation (§IV F), and this will be the subject of §VIII. Moreover, just as we discussed at the end of §IV F 2, if nonlinear effects are somehow quenched or interrupted on a short enough timescale then the problem is *relinearized* and (94) is rendered (approximately) valid. We will see a concrete example of this in §VIII B.

VI. SELF-CONSISTENT LINEAR RESPONSE THEORY

In this section we study the self-consistent response of a system to weak perturbations, using the *linear* approximation (§V D). We assume throughout that the angle-averaged DF f_0 is time-independent, which means that we need to solve equations (92) and (93) simultaneously, but we can ignore equation (91a).

For simplicity we will develop the linear theory in the absence of any possible external perturbations, i.e. we will assume the only potential fluctuations are derived self-consistently internally from the angle-dependent part of the DF δf. However, the case where external perturbations are present is not essentially different (see e.g. Chapter 5 of BT08), and we will show a relevant example in §VI C 1.

A. Laplace transforming the linearized Vlasov–Poisson system

Equations (92) and (93) constitute an initial value problem, and are naturally solved via Laplace transform. Let us therefore define the Laplace transform of a function $h(t)$, and its inverse via

$$\widetilde{h}(\omega) \equiv \int_0^\infty \mathrm{d}t\, h(t)\, \mathrm{e}^{\mathrm{i}\omega t}, \tag{95a}$$

$$h(t) = \frac{1}{2\pi}\int_{\mathrm{i}\sigma-\infty}^{\mathrm{i}\sigma+\infty} \mathrm{d}\omega\, \widetilde{h}(\omega)\, \mathrm{e}^{-\mathrm{i}\omega t}. \tag{95b}$$

Recall that the Laplace transform (95a) is defined for complex ω in the upper-half plane, or more precisely for $\mathrm{Im}\,\omega > \sigma$ where σ is chosen large enough for the integral to converge at $t \to \infty$. In the inverse transform (95b), the 'Bromwich contour' runs along the line $\mathrm{Im}\,\omega = \sigma$ from $\mathrm{Re}\,\omega = -\infty$ to $\mathrm{Re}\,\omega = +\infty$, but can be deformed at will, provided it never crosses over any singularities of $\widetilde{h}(\omega)$.

Laplace transforming equation (93), we get

$$\delta f_{\mathbf{n}}(\mathbf{J},0) - \mathrm{i}(\omega - \mathbf{n}\cdot\boldsymbol{\Omega})\delta\widetilde{f}_{\mathbf{n}}(\mathbf{J},\omega) = \mathrm{i}\mathbf{n}\cdot\frac{\partial f_0}{\partial \mathbf{J}}\delta\widetilde{\Phi}_{\mathbf{n}}(\mathbf{J},\omega), \tag{96}$$

with $\delta f_{\mathbf{n}}(\mathbf{J},0)$ is the initial fluctuation in the DF. The solution to this equation is

$$\delta\widetilde{f}_{\mathbf{n}}(\mathbf{J},\omega) = -\frac{\mathbf{n}\cdot\partial f_0/\partial\mathbf{J}}{\omega - \mathbf{n}\cdot\boldsymbol{\Omega}}\delta\widetilde{\Phi}_{\mathbf{n}}(\mathbf{J},\omega) - \frac{\delta f_{\mathbf{n}}(\mathbf{J},0)}{\mathrm{i}(\omega - \mathbf{n}\cdot\boldsymbol{\Omega})}. \tag{97}$$

(Note that by taking the inverse Laplace transform of this we can recover the formal solution (94)). Meanwhile, the Laplace-transformed version of equation (92) is

$$\delta\widetilde{\Phi}_{\mathbf{n}}(\mathbf{J},\omega) = (2\pi)^d \sum_{\mathbf{n}'}\int \mathrm{d}\mathbf{J}'\, \psi_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}')\delta\widetilde{f}_{\mathbf{n}'}(\mathbf{J}',\omega). \tag{98}$$

Let us now put equations (97) and (98) together in order to eliminate $\delta\widetilde{f}_{\mathbf{n}}(\mathbf{J},\omega)$. The result is the following integral equation for the potential fluctuation $\delta\widetilde{\Phi}_{\mathbf{n}}(\mathbf{J},\omega)$:

$$\delta\widetilde{\Phi}_{\mathbf{n}}(\mathbf{J},\omega) = -(2\pi)^d \sum_{\mathbf{n}'} \int \mathrm{d}\mathbf{J}' \frac{\mathbf{n}'\cdot\partial f_0/\partial \mathbf{J}'}{\omega - \mathbf{n}'\cdot\boldsymbol{\Omega}'} \psi_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}')\, \delta\widetilde{\Phi}_{\mathbf{n}'}(\mathbf{J}',\omega) - (2\pi)^d \sum_{\mathbf{n}'} \int \mathrm{d}\mathbf{J}' \frac{\delta f_{\mathbf{n}'}(\mathbf{J}',0)}{\mathrm{i}(\omega - \mathbf{n}'\cdot\boldsymbol{\Omega}')} \psi_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}'), \tag{99}$$

where $\boldsymbol{\Omega}' \equiv \partial H_0(\mathbf{J}')/\partial \mathbf{J}'$.

B. Dressed wakes and Landau modes

Equation (99) is a linear operator equation of the form

$$|\delta\widetilde{\Phi}(\omega)\rangle = \widehat{\mathsf{M}}(\omega)|\delta\widetilde{\Phi}(\omega)\rangle + |\delta\widetilde{\Phi}_{\mathrm{source}}(\omega)\rangle, \tag{100}$$

where we have written the 'state' of the potential fluctuations at frequency ω as $|\delta\widetilde{\Phi}(\omega)\rangle$ and introduced the abstract linear operator $\widehat{\mathsf{M}}(\omega)$. Note that (100) is not merely a schematic representation of (99); instead, it is what is known in quantum statistical mechanics as a *Dyson equation*[23,60]. One can in fact develop the entire response theory of inhomogeneous systems on the basis of the abstract Dyson equation and operator theory[61–64] without reference to the explicit angle-action representation. We will not go into the details here, but instead just discuss the physics that can be gleaned from equation (100).

On the right hand side of (100) we have the source term $|\delta\widetilde{\Phi}_{\mathrm{source}}(\omega)\rangle$, which corresponds to the second term on the right hand side of (99). This is the part of the potential fluctuations that arises purely from the inevitable, finite-N discreteness noise $\delta f_{\mathbf{n}}(\mathbf{J},0)$, propagated ballistically along mean field trajectories. (Had we included an external potential perturbation in our calculation, it would just be added to this source term). The other term involves the operator $\widehat{\mathsf{M}}(\omega)$ acting on the potential fluctuation $|\delta\widetilde{\Phi}(\omega)\rangle$. It tells us the extent to which potential fluctuations arise because of perturbations in the DF *which were themselves induced by potential fluctuations*. Thus we are discussing *collective effects*: perturbations to the DF which seed their own gravitational fields. Glancing back at (99), we expect these collective effects to be significant if there are fluctuations at frequencies ω which are close to some of the available stellar frequencies $\mathbf{n}'\cdot\boldsymbol{\Omega}'$. These fluctuations will be the all more efficient provided they come from stars at locations in action space $\mathbf{J}'$ where the mean field DF has a large gradient $\partial f_0/\partial\mathbf{J}'$.

How do we solve equation (100)? In homogeneous plasma theory it is easy: in that case, the operator notation we have employed here is unnecessary because $\widehat{\mathsf{M}}(\omega)$ is a scalar (up to proportionality constants, it is equal to 1 minus the dielectric function $\epsilon_{\mathbf{k}}(\omega)$) which 'acts' on the potential fluctuation by simple multiplication. In inhomogeneous systems we are not so lucky. Ultimately this is because, as we warned in §III D, the potential fluctuation $\delta\Phi(\boldsymbol{\theta},\mathbf{J})$ depends on both phase space coordinates *and* momenta, and so $\widehat{\mathsf{M}}(\omega)$ acts upon these fluctuations in a non-trivial long-range fashion. In practice this means we must solve the operator equation by casting $[\mathsf{I} - \widehat{\mathsf{M}}(\omega)]$ in a matrix representation.

The usual technique for doing this is the so-called *biorthogonal basis method* originally developed by Kalnajs[65]. This involves choosing a set of potential and density basis elements, $(\Phi^{(p)},\rho^{(p)})$, which solve the Poisson equation $\nabla^2\Phi^{(p)} = 4\pi G\rho^{(p)}$, but are orthogonal (i.e. $\int\mathrm{d}\mathbf{r}\,\Phi^{(p)*}(\mathbf{r})\,\rho^{(q)}(\mathbf{r}) = 0$ unless $p = q$). Potential fluctuations are then projected onto this basis. The details are rather technical, and unimportant for most of what follows, so we relegate them to Appendix B. The explicit matrix form of the operator $\widehat{\mathsf{M}}(\omega)$ in this basis is

$$\mathsf{M}_{pq}(\omega) = \frac{(2\pi)^d}{\mathcal{E}} \sum_{\mathbf{n}} \int \mathrm{d}\mathbf{J}\, \frac{\mathbf{n}\cdot\partial f_0/\partial\mathbf{J}}{\omega - \mathbf{n}\cdot\boldsymbol{\Omega}}\, \Phi^{(p)*}_{\mathbf{n}}(\mathbf{J})\, \Phi^{(q)}_{\mathbf{n}}(\mathbf{J}), \tag{101}$$

where $\Phi^{(p)}_{\mathbf{n}}(\mathbf{J})$ is the Fourier transform of $\Phi^{(p)}(\mathbf{r})$ in angle space (see 35), and $\mathcal{E}$ is an arbitrary constant with units of energy. Equation (101) is usually called the *response matrix*. There is of course a striking similarity between equation (101) and the corresponding quantity in plasma theory, and we explore this connection in §VI B 1. Moreover, just like in plasma theory, although we derived (101) for $\mathrm{Im}\,\omega > 0$ we can analytically continue this function to the lower-half ω plane by taking the action space integral along the Landau contour. In practice, stellar systems typically sustain only a finite range of resonance frequencies, $\mathbf{n}\cdot\boldsymbol{\Omega}$. This leads to branch cuts in the response matrix[66,67], as is also the case, for example, in magnetized relativistic plasmas[68], in the toroidal ion temperature gradient mode[69,70], or truncated Maxwellians[71]. We will not pursue any details here, but refer the reader to Chapter 5 of BT08 and to Fouvry and Prunet[72] for more information.

Having expressed $\widehat{\mathsf{M}}(\omega)$ as a matrix $\mathsf{M}(\omega)$, the nature of the solution to (100) depends crucially on whether $\det[\mathsf{I} - \mathsf{M}(\omega)] = 0$ or $\det[\mathsf{I} - \mathsf{M}(\omega)] \neq 0$:

- If ω is such that $\det[\mathsf{I}-\mathsf{M}(\omega)] \neq 0$ then the inverse matrix $[\mathsf{I}-\mathsf{M}(\omega)]^{-1}$ exists, and the solution to (100) is simply

$$|\delta\widetilde{\Phi}(\omega)\rangle = [\mathsf{I}-\mathsf{M}(\omega)]^{-1}|\delta\widetilde{\Phi}_{\rm source}(\omega)\rangle. \tag{102}$$

  This is true for the great majority of ω values. Equation (102) equation tells us that at these frequencies, the noise source $|\delta\widetilde{\Phi}_{\rm source}(\omega)\rangle$ gets amplified by collective effects. Sometimes this is referred to the noise being *dressed* by a gravitational *wake*. In the equivalent electrostatic plasma calculation, $\mathsf{I}-\mathsf{M}$ is just the dielectric function $\epsilon_{\mathbf{k}}(\omega)$; for near-equilibrium plasmas the solution (102) tells us that each 'bare' charge is locally screened by the Debye cloud of opposite charges that it gathers round itself. But in the gravitational context there is no screening, and collective effects will often make gravitational perturbations on large scales *stronger*, not weaker. Loosely speaking, whereas Debye clouds tend to *reduce the effective charge*, gravitational wakes tend to *increase the effective mass* hence typically shortening the relaxation time (see §VI C 1 for an example).

- Alternatively, for ω such that $\det[\mathsf{I}-\mathsf{M}(\omega)] = 0$ the inverse of $[\mathsf{I}-\mathsf{M}(\omega)]$ does *not* exist, which tells us that $|\delta\widetilde{\Phi}(\omega)\rangle$ can be non-zero even in the absence of a source term on the right hand side of (100). These are the so-called *Landau modes* of the system[73], and they tend to occur for a discrete set of (generally) complex frequencies $\{\omega_{\rm m}\}$. The system is then able to support a discrete set of collective oscillations which, just like electromagnetic waves in vacuum, do not need to be continually driven by some noise source. Moreover, under fairly weak assumptions one can show that the solution of equation (100) at this frequency will have time-dependence $\sim \mathrm{e}^{\mathrm{i}\omega_{\rm m}t}$ (see Kalnajs[63] for a precise statement). Thus, a system is linearly unstable if there exists a Landau mode whose imaginary part $\mathrm{Im}\,\omega_{\rm m} > 0$. Otherwise, the system is stable.

This confluence of a 'dressed continuum' of fluctuations superposed with a discrete set of 'Landau modes' is characteristic of both plasmas and stellar systems, although the typical responses of these systems to perturbations can be very different, as we will illustrate with some examples below.

1. Example: Jeans instability in a homogeneous stellar system

First, let us address directly the analogy between stellar systems and homogeneous plasmas, by calculating the response matrix $\mathsf{M}(\omega)$ for a fictitious three-dimensional homogeneous stellar system. As noted in §III C 1, in that case, orbits are simple (multi-periodic) straight lines, and we can take $\boldsymbol{\theta} \propto \mathbf{r}$ and $\mathbf{J} \propto \boldsymbol{\Omega} \propto \mathbf{v}$ (see equation 37). The system's instantaneous potential and DF are linked by

$$\Phi(\mathbf{r}) = \int \mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{v}' f(\mathbf{r}',\mathbf{v}')\psi(\mathbf{r},\mathbf{r}'), \tag{103}$$

where $\psi(\mathbf{r},\mathbf{r}') = -G/|\mathbf{r}-\mathbf{r}'|$ is the gravitational pairwise interaction. Making use of the periodicity of this system one has

$$\psi(\boldsymbol{\theta},\boldsymbol{\theta}') = -\frac{G}{L\pi}\sum_{\mathbf{p}\neq\mathbf{0}}\frac{\mathrm{e}^{\mathrm{i}\mathbf{p}\cdot(\boldsymbol{\theta}-\boldsymbol{\theta}')}}{|\mathbf{p}|^2}. \tag{104}$$

Glancing at (B4), we find that, in homogeneous systems, the *natural* basis elements are $\Phi^{(\mathbf{p})} \sim \mathrm{e}^{\mathrm{i}\mathbf{p}\cdot\boldsymbol{\theta}}/|\mathbf{p}|$, namely Fourier plane waves. This is unsurprising, since the system is translationally invariant. In equation (104), the factor $1/|\mathbf{p}|^2$ is naturally reminiscent of the $1/|\mathbf{k}|^2$ factor that arises when Fourier transforming the electrostatic Poisson equation for plasmas.

With this, the response matrix from (101) becomes

$$\mathsf{M}_{\mathbf{pq}}(\omega) = \frac{GL^2}{\pi}\delta^{\mathbf{q}}_{\mathbf{p}}\frac{1}{|\mathbf{p}|^2}\int \mathrm{d}\mathbf{v}\,\frac{\mathbf{p}\cdot\partial f_0/\partial\mathbf{v}}{\overline{\omega}-\mathbf{p}\cdot\mathbf{v}}, \tag{105}$$

where $\overline{\omega} \equiv (2\pi)^{-1}L\omega$. Assuming (again unphysically! — see §II B) that the system's mean DF follows the Maxwellian distribution

$$f_0(\mathbf{v}) = \frac{\rho_0}{(2\pi\sigma^2)^{3/2}}\mathrm{e}^{-|\mathbf{v}|^2/(2\sigma^2)}, \tag{106}$$

where ρ_0 is the system's mean density, and σ is the velocity dispersion, we find from (105) that

$$\mathsf{M}_{\mathbf{pq}}(\omega) = \delta^{\mathbf{q}}_{\mathbf{p}}\frac{1}{|\mathbf{p}|^2}\left(\frac{L}{L_{\rm J}}\right)^2[1+\zeta Z(\zeta)], \tag{107}$$

where $\zeta = \varpi/(\sqrt{2}|\mathbf{p}|\sigma)$ is a dimensionless frequency, Z is the usual plasma dispersion function[74]

$$Z(\zeta) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \mathrm{d}u \frac{\mathrm{e}^{-u^2}}{u - \zeta}, \tag{108}$$

and L_J is the so-called *Jeans length*

$$L_\mathrm{J} = \sqrt{\frac{\pi\sigma^2}{G\rho_0}} \approx 10\,\mathrm{kpc} \left(\frac{\sigma}{100\,\mathrm{km\,s^{-1}}}\right) \left(\frac{\rho_0}{10^8 M_\odot\,\mathrm{kpc}^{-3}}\right)^{-1/2}. \tag{109}$$

The main difference between the expression (107) for a self-gravitating system, and the analogous one for an electrostatic plasma, is the term $1 + \zeta Z(\zeta)$ which in a plasma is replaced by $1 - \zeta Z(\zeta)$.

The dispersion relation for Landau modes is then $\delta^{\mathbf{q}}_{\mathbf{p}} - \mathsf{M}_{\mathbf{pq}}(\omega_\mathrm{m}) = 0$. By looking for solutions with Re $\omega_\mathrm{m} = 0$ one can show that instability occurs (Im $\omega_\mathrm{m} > 0$) if the system is larger than a critical scale, $L > L_\mathrm{J}$ (see §5.2.4 of BT08 for details). Thus, unlike in plasma, a homogeneous Maxwellian distribution of stars is not necessarily stable to linear perturbations, but instead can be made unstable e.g. by making the system dynamically colder (reducing σ) or more dense (increasing ρ_0). The physics underlying this *Jeans instability* is a very simple competition between gravity and velocity dispersion, which we sketch in Figure 19. In a stellar system without any velocity dispersion, a density

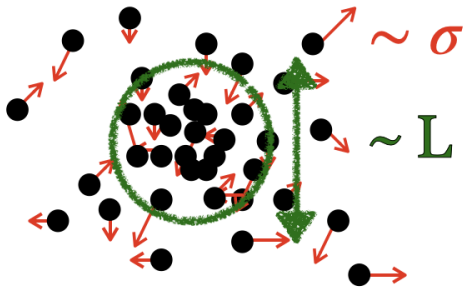

FIG. 19. Jeans instability as a competition between velocity dispersion and gravitational attraction. A density fluctuation ρ_0 over a length scale L will collapse under its own gravity if the collapse timescale $\sim 1/\sqrt{G\rho_0}$ is shorter than the crossing time L/σ where σ is the velocity dispersion.

fluctuation ρ_0 over a length scale L would collapse under its own gravity on the timescale $\sim 1/\sqrt{G\rho_0}$. However, if the stars have some velocity dispersion σ, then they can escape the region of size L on the timescale $\sim L/\sigma$. If the latter timescale is longer than the former, the system cannot avoid collapse: this happens if $L \gtrsim \sigma/\sqrt{G\rho_0}$, which, up to a prefactor, is the Jeans length (109). The Jeans instability is the basic mechanism underlying large-scale structure formation in the Universe[3].

2. Example: Spiral Landau modes in cold stellar disks

Another important example of a Landau mode is a spiral mode in a disk galaxy. In truth, it is still not completely clear what gives rise to the spiral structure that we see in real galaxies like that in the top panel of Figure 1. Rather than attempt a review of this fascinating subject here[77,78], let us restrict ourselves to considering a more artificial system, namely an *isolated, two-dimensional* stellar disk consisting of *equal-mass* stars and *no gas*. In this case, spirals can arise manifestly as linear instabilities (i.e. Landau modes with Im $\omega_\mathrm{m} > 0$) of an underlying DF $f_0(L, J_R)$.

A classic example of such a spiral mode was discovered by Zang[75], which we reproduce in Figure 20 (adapted from Petersen *et al.*[76]). The model Zang employed was a 'Mestel disk', which is a simple two-dimensional disk embedded in a rigid dark matter halo. Just as with the Jeans instability (§VI B 1), one can always make a disk unstable by making it sufficiently dynamically cold and/or dense. In Zang's case, this was achieved by modifying the *active fraction* of the disk, essentially the ratio between the self-gravitating mass in the disk and that in the unresponsive halo. Panel (a) of Figure 20 shows the contours of $\det[\mathsf{I} - \mathsf{M}(\omega)]$ (equation 101) in the complex frequency plane. The yellow cross shows the location where this determinant equals zero, which gives the frequency ω_m of an unstable spiral Landau mode. Precisely, the mode rotates with pattern speed $0.88\,\Omega_0$ and growth rate $0.12\,\Omega_0$, where Ω_0 is a typical dynamical frequency comparable to $\sim 2\pi/t_\mathrm{dyn}$. The spatial structure of the spiral mode is shown in

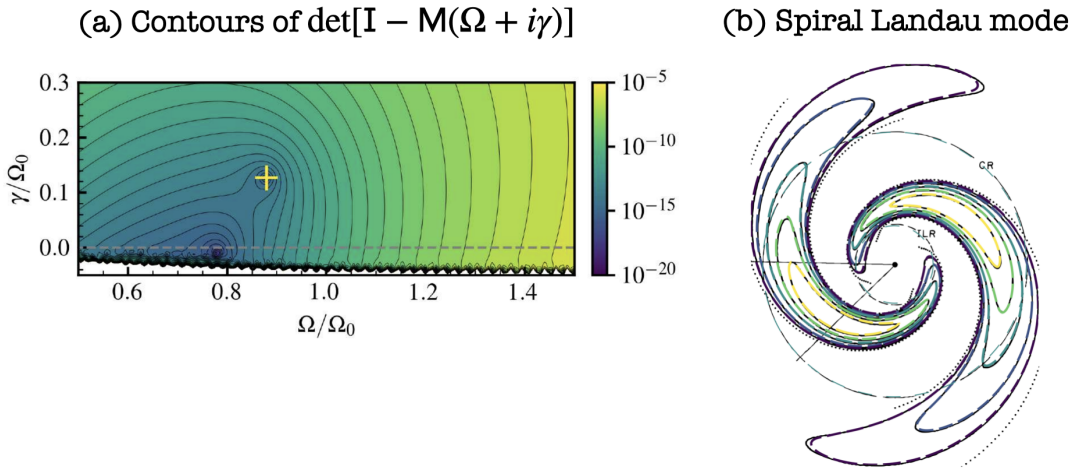

FIG. 20. Determination of an unstable spiral Landau mode in a Mestel disk following Zang [75]. Panel (a) shows the determinant of $\mathsf{I}-\mathsf{M}$ in the complex frequency plane; the yellow cross shows where this function passes through zero. Panel (b) shows contours of the corresponding spiral Landau mode. Figure adapted from Petersen *et al.* [76].

panel (b); this is calculated by finding the eigenvector $\mathbf{X}_\mathrm{m}$ with components X^p_m, such that $\mathsf{M}(\omega_\mathrm{m})\cdot\mathbf{X}_\mathrm{m}=0$ and then writing

$$\delta\Phi_\mathrm{m}(\mathbf{r},t) \propto \mathrm{Re}\left[\sum_p X^p_\mathrm{m}\,\Phi^{(p)}(\mathbf{r})\,\mathrm{e}^{-\mathrm{i}\omega_\mathrm{m}t}\right]. \tag{110}$$

(Since this is a linear eigenmode, the overall normalisation is arbitrary).

The two examples we have given so far both involved smooth mass distributions which were made sufficiently cold and/or dense that they became unstable. However, there are other ways of producing instabilities which are perhaps more familiar to plasma physicists used to the bump-on-tail paradigm[57]. For instance, De Rijcke *et al.* [79] recently performed a linear stability analysis, like the one we have outlined in this section, for a *grooved* Mestel disk. That is, they took a Mestel disk with a stable DF — the $J_R=0$ slice through which is shown with a dotted line in Figure 21a — and then carved a 'groove' in it by removing a portion of the stars in a given range of angular momentum (solid line in Figure 21a). This grooved disk is also unstable to spiral Landau modes, and Figure 21b shows the shape of

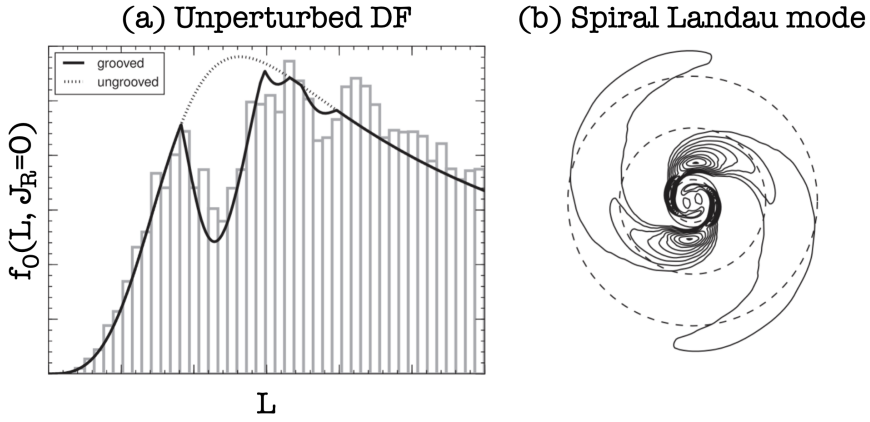

FIG. 21. Groove-induced instability in a Mestel disk. Panel (a) shows the $J_R=0$ cut through the DF in the stable (ungrooved) and unstable (grooved) case. Panel (b) shows the unstable spiral mode as calculated from linear stability analysis of the grooved disk. The solid circle shows the location of the groove for circular orbits, while the three dashed circles show the ILR, CR, and OLR resonance locations of the resulting spiral mode. Figure adapted from De Rijcke *et al.* [79].

the dominant mode.

The above examples suggest — and further experiments tend to confirm — that the Landau mode spectrum of stellar disks tends to be dominated by one, or at most a few, isolated frequencies. This means that spirals in unstable disks

are perhaps more analogous to e.g. Alfvenic eigenmodes in tokamak plasmas than to classic bump-on-tail instabilities in homogeneous plasmas, since the latter tend to harbor a whole spectrum of unstable modes (at least one at every unstable wavenumber $\mathbf{k}$). Because of this, the evolution of unstable stellar disks might be more aptly described by something like the single-wave kinetics of O'Neil [80], Berk and Breizman [81,82], etc. than by the typical many-wave quasilinear theory of Drummond and Pines [83], Kaufman [84]. See §VIII for more.

C. Explicit wake solution

We now suppose that our system is stable, meaning that all Landau modes are damped. An astrophysical example of such a (weakly) damped mode is the (dipole) $\ell = 1$ mode of globular clusters[85], which drive long-lasting sloshing oscillations[86]. We further assume that we may wait long enough that these modes no longer contribute to the response (i.e. we wait for $t \gg |\mathrm{Im}\,\omega_\mathrm{m}|^{-1}$, where ω_m is the system's least strongly-damped Landau mode). In this regime, we can solve (99) directly under the assumption that ω is *not* associated with a Landau mode. Having constructed the solution we will then perform an inverse Laplace transform to get the dressed wake part of the response in the time domain.

The easiest — and most physically enlightening — way to solve (99) is to guess what form the solution should have, and then work out the exact details post-hoc. So, let us look at (99). Supposing for a moment that we were to ignore collective effects ($\widehat{\mathsf{M}} = \mathsf{0}$), we would find

$$[\delta\widetilde{\Phi}_{\mathbf{n}}(\mathbf{J},\omega)]_\mathrm{bare} = -(2\pi)^d \sum_{\mathbf{n}'} \int \mathrm{d}\mathbf{J}' \frac{\delta f_{\mathbf{n}'}(\mathbf{J}',0)}{\mathrm{i}(\omega - \mathbf{n}'\cdot\boldsymbol{\Omega}')} \psi_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}'), \tag{111}$$

which we call the 'bare' solution. Inverse-Laplace transforming this by deforming the Bromwich contour so that it encircles the pole at $\omega = \mathbf{n}'\cdot\boldsymbol{\Omega}'$, we would find

$$[\delta\Phi_{\mathbf{n}}(\mathbf{J},t)]_\mathrm{bare} = (2\pi)^d \sum_{\mathbf{n}'} \int \mathrm{d}\mathbf{J}'\, \delta f_{\mathbf{n}'}(\mathbf{J}',0) \mathrm{e}^{-\mathrm{i}\mathbf{n}'\cdot\boldsymbol{\Omega}'t} \psi_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}'). \tag{112}$$

The physical interpretation of this result is very clear: it is just the (Fourier transform of the) potential at location $\mathbf{r}(\boldsymbol{\theta},\mathbf{J})$ that is created by summing the contributions from every star with initial locations $\mathbf{r}'(0) = \mathbf{r}'(\boldsymbol{\theta}'(0),\mathbf{J}'(0))$, assuming those stars simply travel along mean field trajectories $\boldsymbol{\theta}' = \boldsymbol{\Omega}' t + \boldsymbol{\theta}'(0)$ and $\mathbf{J}' = \mathbf{J}'(0)$.

The ansatz we now make is that the solution to the full equation (99) is exactly of this functional form, except with the bare interaction potential $\psi_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}')$ replaced with some *dressed* (and frequency-dependent) interaction potential $\psi^\mathrm{d}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\omega)$, whose form we do not yet know. That is, we assume the solution has the form

$$\delta\widetilde{\Phi}_{\mathbf{n}}(\mathbf{J},\omega) = -(2\pi)^d \sum_{\mathbf{n}'} \int \mathrm{d}\mathbf{J}' \frac{\delta f_{\mathbf{n}'}(\mathbf{J}',0)}{\mathrm{i}(\omega - \mathbf{n}'\cdot\boldsymbol{\Omega}')} \psi^\mathrm{d}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\omega). \tag{113}$$

This ansatz is actually very natural if one is acquainted with *Rostoker's principle* in plasma kinetics, which essentially says that one can consider a hot plasma as a set of completely uncorrelated particles, provided one replaces the bare Coulombic potential of each particle with its dressed, or *screened*, version[57]. (We will use this idea to derive the Balescu–Lenard kinetic equation for stable stellar systems in §VII B). Plugging the ansatz (113) into (99) we can read off an implicit equation for ψ^d:

$$\psi^\mathrm{d}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\omega) = \psi_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}') - (2\pi)^d \sum_{\mathbf{n}''} \int \mathrm{d}\mathbf{J}'' \frac{\mathbf{n}''\cdot\partial f_0/\partial\mathbf{J}''}{\omega - \mathbf{n}''\cdot\boldsymbol{\Omega}''} \psi_{\mathbf{n}\mathbf{n}''}(\mathbf{J},\mathbf{J}'')\, \psi^\mathrm{d}_{\mathbf{n}''\mathbf{n}'}(\mathbf{J}'',\mathbf{J}',\omega), \tag{114}$$

where $\boldsymbol{\Omega}'' \equiv \partial H_0(\mathbf{J}'')/\partial\mathbf{J}''$. Thus, we will have constructed a valid solution to equation (99) if we can solve equation (114) for ψ^d. While at first sight it may not seem like we have made much progress with this rewriting, it turns out that (114) is actually quite easy to solve if we project the various terms onto the biorthogonal basis elements $\Phi^{(p)}_{\mathbf{n}}(\mathbf{J})$, just as in (101). We leave the (rather boring) details to Appendix B, and here just write down the final answer:

$$\psi^\mathrm{d}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\omega) = -\frac{1}{\mathcal{E}} \sum_{p,q} \Phi^{(p)}_{\mathbf{n}}(\mathbf{J})\, [\mathsf{I} - \mathsf{M}(\omega)]^{-1}_{pq}\, \Phi^{(q)*}_{\mathbf{n}'}(\mathbf{J}'). \tag{115}$$

We can usefully compare this to the bare interaction, which as we show in Appendix B can be expanded as

$$\psi_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}') = -\frac{1}{\mathcal{E}}\sum_p \Phi^{(p)}_{\mathbf{n}}(\mathbf{J})\,\Phi^{(p)*}_{\mathbf{n}'}(\mathbf{J}'). \tag{116}$$

Thus we see that in passing from $\psi \to \psi^{\rm d}$ we have simply 'divided' the bare interaction by $[\mathsf{I}-\mathsf{M}(\omega)]$. This is analogous to dividing by the dielectric function $\epsilon_{\mathbf{k}}(\omega)$ in electrostatic plasma theory.

We can now inverse Laplace transform equation (113). The assumption that Landau modes can be ignored (i.e. $t \gg |{\rm Im}\,\omega_{\rm m}|^{-1}$) means that when we deform the Bromwich contour we need only pick up contributions from the poles at $\omega = \mathbf{n}'\cdot\mathbf{\Omega}'$, and can ignore any singularities of $\psi^{\rm d}$ (i.e. zeroes of $[\mathsf{I}-\mathsf{M}(\omega)]$, see equation 115). The result is

$$\delta\Phi_{\mathbf{n}}(\mathbf{J},t) = (2\pi)^d \sum_{\mathbf{n}'}\int {\rm d}\mathbf{J}'\,\delta f_{\mathbf{n}'}(\mathbf{J}',0)\,{\rm e}^{-{\rm i}\mathbf{n}'\cdot\mathbf{\Omega}'t}\psi^{\rm d}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\mathbf{n}'\cdot\mathbf{\Omega}'), \tag{117}$$

which is the same as (112) with $\psi \to \psi^{\rm d}$. Thus, we have realized Rostoker's idea: the dressed potential fluctuation is equivalent to that carried by the initial fluctuation ($\sim \delta f_{\mathbf{n}'}(\mathbf{J}',0)$), transported along the mean field ($\sim {\rm e}^{-{\rm i}\mathbf{n}'\cdot\mathbf{\Omega}'t}$), but with the Newtonian potential of each star replaced with the dressed potential ($\sim \psi^{\rm d}$) evaluated at the mean field frequency ($\mathbf{n}'\cdot\mathbf{\Omega}'$), with contributions summed over all stars ($\sum_{\mathbf{n}'}\int {\rm d}\mathbf{J}'$). Furthermore, (117) can be plugged into the right hand side of equation (94) to give an explicit solution for the fluctuations in the DF $\delta f_{\mathbf{n}}(\mathbf{J},t)$. Note that the basis functions $\Phi^{(p)}$ play explicit no role in equation (117); they are merely a calculational tool which is *only* used to construct the dressed interaction $\psi^{\rm d}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\omega)$ for a given mean field $f_0(\mathbf{J})$.

1. *Example: Response of a stellar system to a perturbing mass*

So far in this section we have concentrated on the response of a stellar system to internal noise $\delta f(t=0)$. However, an entirely analogous calculation shows that even externally imposed potential fluctuations will be 'dressed' by the system's self-gravity — just as an externally imposed point charge will gather its own Debye cloud in an electrostatic plasma. Here we will not pursue any quantitative details, but instead show a few figures which capture the characteristic features of such a response.

First we present Figure 22, which is adapted from Gibbon [87]. This is a cartoon of a quasineutral electrostatic

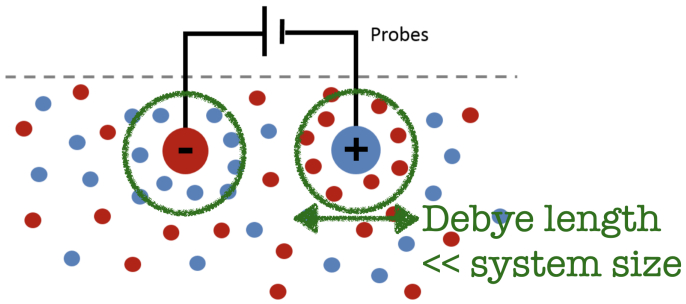

FIG. 22. Cartoon (adapted from Gibbon [87]) of the Debye shielding that occurs when electrical probes are placed inside a quasineutral plasma.

plasma made of blue positive charges and red negative charges. When a pair of electrical probes is inserted into the plasma, they gather around them Debye clouds consisting of particles of the opposite charge. In this way, the electrostatic potential of the probes is screened over a Debye length. This cloud typically much smaller than the scale of the whole system.

Next, we turn to Figure 23, which is adapted from Magorrian [88]. This is an analogous stellar-dynamical scenario, in which a heavy point mass is driven on a straight line orbit at constant speed through a homogeneous stellar system with Maxwellian f_0 (§VI B 1). The size of the box in this case has size $L \approx 0.9 L_{\rm J}$, so the system is (weakly) stable (recall that instability occurs if $L > L_{\rm J}$). The contours in panel (a) show the density response of the system according to linear theory in the case that stars only respond to the point-mass perturber but not to each other. This we call the 'bare' response. By contrast, panel (b) shows the response density in the 'dressed' case in which stars interact both with the perturber and with each other (the contour levels are the same as in panel (a)). Let us make two simple

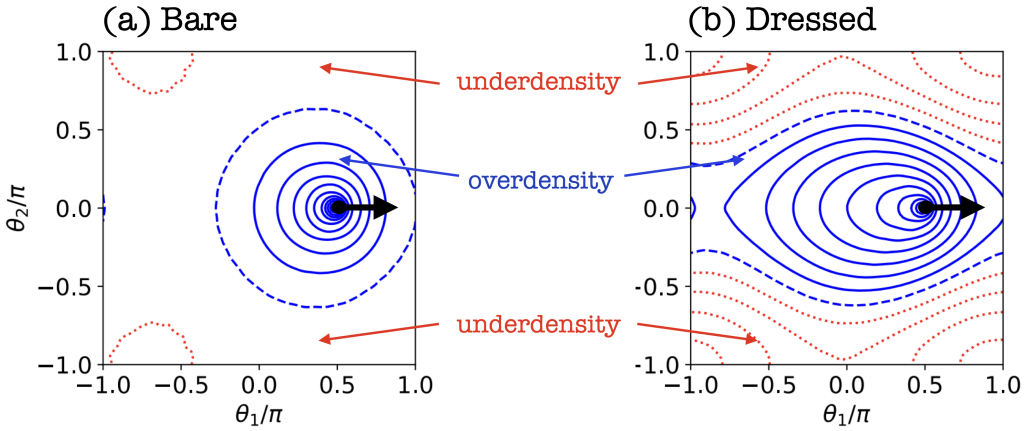

FIG. 23. Response of a Maxwellian box of stars to a point mass perturber, adapted from Magorrian [88]. The point mass (black dot) is driven through the box of size $\sim 0.9 L_{\rm J}$ at constant speed, and the wake response is computed according to linear theory. Solid blue contours denote positive overdensity, and red dotted contours denote negative overdensity, while the blue dashed curve corresponds to zero overdensity. In panel (a) the stars only feel the gravitational potential of the perturber but not of each other, while in panel (b) they interact with both the perturber and each other. The same linearly spaced contour levels are used in both panels.

observations which contrast the behavior of the stellar system shown in these panels with the plasma system sketched in Figure 22. (i) The overdense wake in the gravitational system acts to reinforce the perturbation rather than shield it, since there are no opposite charges. Heuristically therefore one can think of gravitational polarization as a kind of 'anti Debye-shielding'. (ii) The scale of the induced wake, especially in the dressed case, can be comparable to the scale of the system itself, in contrast with an electrostatic plasma where the Debye length is usually much smaller than the system.

Since a homogeneous box is not a particularly realistic model for a stellar system, we finally present Figure 24, which is adapted from D'Onghia *et al.* [89]. This figure shows the response density of a stellar disk to an embedded

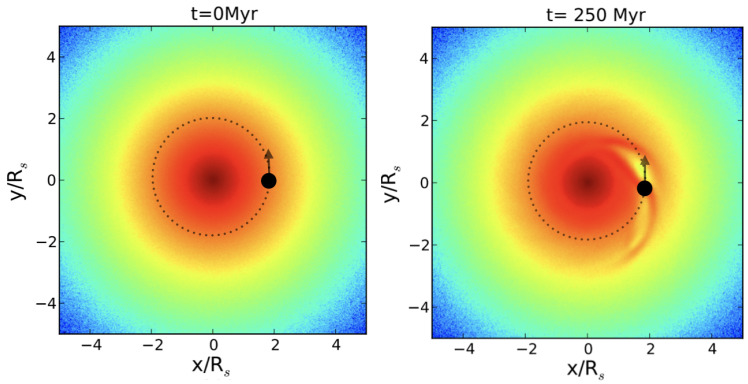

FIG. 24. Response of a stellar disk to a massive perturber driven on a circular orbit (adapted from D'Onghia *et al.* [89]). The perturber mass is much less than 1% of the mass of the system. Still, it manages to produce a significant response over a large swathe of the disk.

point-mass perturber, driven on a circular orbit with constant speed for 1 azimuthal period (250 Myr). The disk is chosen to be fairly dynamically cold, but not so cold as to be unstable. The response was calculated using a full N-body simulation in which stars interact both with the perturber and with each other. The perturbing mass was much less than 1% of the whole system's mass. Despite its relatively small mass, the perturber provokes a response in the disk surface density which extends over a large azimuthal range. This reflects the fact that cold stellar disks are very effective amplifiers of perturbations, as we discussed in §II C 3. As a result, it is typical of weakly stable stellar systems to develop collective responses on a global scale.

VII. QUASILINEAR EVOLUTION AND THE BALESCU–LENARD EQUATION

In many cases we do not know, nor do we really care, about the details of the microscopic DF $f_{\rm M}$ of a particular stellar system. What we often *do* care about, however, is the structure and evolution of some suitably *averaged* DF, $\langle f_{\rm M}\rangle$. If we are interested in the evolution of a system driven by random fluctuations, be they internally or externally generated, then the appropriate average is an *ensemble average* over all possible microscopic realizations of those fluctuations. The central idea of kinetic theory then is to use the statistics of the realizations of the fluctuations to derive an equation for the evolution of $\langle f_{\rm M}\rangle$.

By ergodic arguments, we may expect for an ensemble average of $f_{\rm M}$ over microscopic realizations to be equivalent to an average over angles and over time (on timescales $\ll t_{\rm relax}$). Thus $\langle f_{\rm M}\rangle = f_0$ and the relevant equation describing the evolution is (c.f. equation 91a):

$$\frac{\partial f_0}{\partial t} = -\frac{\partial}{\partial \mathbf{J}} \cdot \sum_{\mathbf{n}} \mathrm{i}\mathbf{n} \, \langle \delta\Phi^*_{\mathbf{n}}(\mathbf{J},t) \delta f_{\mathbf{n}}(\mathbf{J},t)\rangle, \tag{118}$$

In §V D, we spelled out the basic idea of *quasilinear* kinetic theory, which is that if the 'fluctuations' $\delta\Phi$ and δf are sufficiently small, then we can predict the evolution of f_0 on the timescale $t_{\rm relax} \gg t_{\rm dyn}$ by plugging the solutions from linear response theory (§VI) into the right hand side of (118). Quasilinear theory works provided such a timescale separation exists, and provided we can ignore higher order nonlinear and nonperturbative effects like trapping, as discussed in §V E.

In this section we will develop quasilinear kinetic theory more quantitatively. Our study of internally-driven evolution will culminate in the Balescu–Lenard equation which describes the slow evolution of an isolated, stable N-star system. We first describe the basic physics underlying the Balescu–Lenard equation, then simply write it down without a proper derivation and discuss some of its properties (§VII A). Then we take a step back and pursue an alternative (and simpler) derivation of the Balescu–Lenard equation based on Rostoker's principle (§VII B). We give several examples of this equation being used in practice (§VII C), before discussing where the theory breaks down and how it might be improved (§VII D).

Like in §VI, we consider systems in which the fluctuations are internally generated by the finite-N noise, but we note that a very similar formalism can be developed to account for externally-driven quasilinear evolution[90,91].

A. A tale of two fluxes

As a first step, we take the formal solution for the perturbation to the DF in linear theory, equation (94), and plug it into the right hand side of (118). The result is

$$\frac{\partial f_0}{\partial t} = -\frac{\partial}{\partial \mathbf{J}} \cdot \mathbf{F}, \tag{119}$$

where the two contributions to the 'action space flux' $\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2$ on the right hand side are

$$\mathbf{F}_1(\mathbf{J},t) = \sum_{\mathbf{n}} \mathrm{i}\mathbf{n} \, \mathrm{e}^{-\mathrm{i}\mathbf{n}\cdot\boldsymbol{\Omega} t} \langle \delta f_{\mathbf{n}}(\mathbf{J},0) \, \delta\Phi^*_{\mathbf{n}}(\mathbf{J},t)\rangle, \tag{120}$$

and

$$\mathbf{F}_2(\mathbf{J},t) = -\sum_{\mathbf{n}} \mathbf{n}\,\mathbf{n} \cdot \frac{\partial f_0}{\partial \mathbf{J}} \int_0^t \mathrm{d}t' \mathrm{e}^{-\mathrm{i}\mathbf{n}\cdot\boldsymbol{\Omega}(t-t')} \langle \delta\Phi^*_{\mathbf{n}}(\mathbf{J},t) \delta\Phi_{\mathbf{n}}(\mathbf{J},t')\rangle. \tag{121}$$

These equations are not yet written form that is useful for calculation, but their physical meaning is already apparent. In order to make this discussion as clear as possible, let us deal with them in reverse order.

- The second flux $\mathbf{F}_2$ is a diffusion term. This is obvious since it is proportional to the gradient of the DF $f_0(\mathbf{J})$ in action space. That is to say, this term answers the following question: "I am a star with action $\mathbf{J}$, moving along my orbit within a bath of potential fluctuations $\delta\Phi$. If I think of myself as a massless test particle, then on average, how are those potential fluctuations going to nudge me around?" It is natural that the answer involves the correlation of fluctuations sampled at action space location $\mathbf{J}$. Of course, nothing about the expression (121) would formally change had we insisted that $\delta\Phi$ was externally imposed, but here we are going to use the solution for $\delta\Phi$ derived in §VI C assuming the fluctuations are internally generated and that the system is stable. Thus, we say that the flux $\mathbf{F}_2$ describes *orbital diffusion due to dressed discreteness noise*.

- On the other hand, the first flux $\mathbf{F}_1$ provides the answer to the obvious rebuttal: "Wait! I'm not really a test particle, so diffusion cannot be the whole story. I have finite mass, and so I drive potential fluctuations of my own. To what extent is my motion altered by the part of $\delta\Phi$ that I have myself produced?" Physically this term represents *dynamical friction*, which Chandrasekhar intuited back in the 1940s. The key idea is that when any massive body moves through a sea of other bodies, it creates a 'wake' behind it in response (just as in Figures 23-24). This wake then backreacts, or drags, on the massive body, slowing it down. It is therefore natural that the flux $\mathbf{F}_1$ involves the correlation between $\delta\Phi$ at action space location $\mathbf{J}$ and the initial discreteness noise at the same location, $\delta f_{\mathbf{n}}(\mathbf{J},0)$.

Although we have taken all stars to be of equal mass for simplicity, one can show that in a multi-mass system, the more massive a body is the larger the wake it induces, and so the stronger the drag force it feels. This leads for instance to massive bodies losing energy and spiraling in towards the center of their host galaxy. Mathematically, for a subsystem of very massive bodies, the frictional term in (119) is typically much more important than the diffusive term. By contrast, a very light subsystem will contribute only weakly to the total gravitational potential fluctuation, and hence will experience little dynamical friction; for this subsystem, the diffusive term in (119) can dominate over the frictional one. This is also an astrophysically relevant scenario: for instance, a key problem of galactic evolution is to understand the diffusion of stellar orbits (Figure 7) due to interactions with transient spiral waves and/or giant molecular clouds[90]. For the present single-mass system, though, $\mathbf{F}_1$ and $\mathbf{F}_2$ are typically comparable in magnitude.

The next step in deriving a proper quasilinear kinetic equation is to take the solution for the dressed potential from linear theory $\delta\Phi[\delta f(0)]$, equation (117), and plug it into the right hand sides of (120)-(121). One then further assumes that stellar orbits are initially uncorrelated and takes $t \gg t_{\rm dyn}$. The details of this procedure are rather lengthy and do not convey much physical insight, so we refer the reader to e.g. the works by Fouvry and Bar-Or [91], Chavanis [92], and here just write down the final answer, which is that equation (119) takes the Fokker–Planck form

$$\frac{\partial f_0}{\partial t} = \frac{\partial}{\partial \mathbf{J}} \cdot \left[-\mathbf{A}(\mathbf{J}) f_0 + \mathsf{D}(\mathbf{J}) \cdot \frac{\partial f_0}{\partial \mathbf{J}} \right], \tag{122}$$

with the friction vector

$$\mathbf{A}(\mathbf{J}) = \pi (2\pi)^d m \sum_{\mathbf{n},\mathbf{n}'} \mathbf{n} \int \mathrm{d}\mathbf{J}' \delta(\mathbf{n}\cdot\mathbf{\Omega} - \mathbf{n}'\cdot\mathbf{\Omega}') |\psi^{\rm d}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\mathbf{n}'\cdot\mathbf{\Omega}')|^2 \, \mathbf{n}' \cdot \frac{\partial f_0}{\partial \mathbf{J}'}, \tag{123}$$

and diffusion tensor

$$\mathsf{D}(\mathbf{J}) = \pi (2\pi)^d m \sum_{\mathbf{n},\mathbf{n}'} \mathbf{n} \otimes \mathbf{n} \int \mathrm{d}\mathbf{J}' \delta(\mathbf{n}\cdot\mathbf{\Omega} - \mathbf{n}'\cdot\mathbf{\Omega}') |\psi^{\rm d}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\mathbf{n}'\cdot\mathbf{\Omega}')|^2 f_0(\mathbf{J}'). \tag{124}$$

We can put these expression together to arrive at the *Balescu–Lenard equation*

$$\frac{\partial f_0(\mathbf{J})}{\partial t} = \pi (2\pi)^d m \frac{\partial}{\partial \mathbf{J}} \cdot \sum_{\mathbf{n},\mathbf{n}'} \mathbf{n} \int \mathrm{d}\mathbf{J}' \delta(\mathbf{n}\cdot\mathbf{\Omega} - \mathbf{n}'\cdot\mathbf{\Omega}') |\psi^{\rm d}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\mathbf{n}'\cdot\mathbf{\Omega}')|^2 \left(\mathbf{n}\cdot\frac{\partial}{\partial \mathbf{J}} - \mathbf{n}'\cdot\frac{\partial}{\partial \mathbf{J}'} \right) f_0(\mathbf{J}) f_0(\mathbf{J}'). \tag{125}$$

1. Properties of the Balescu–Lenard equation

At first glance the Balescu–Lenard equation might look implausibly complicated, but the physics underlying it is simple. Indeed, with hindsight, one could argue that this is the simplest equation one could write down with all the desired properties (particularly that it must overcome the shortcomings of Chandrasekhar theory that we listed in §II C 3). For instance:

- The equation is expressed in angle-action variables and therefore naturally accounts for the inhomogeneity of the system and its associated nontrivial orbital structure.
- Related to this, unlike in Chandrasekhar's theory, in which we had to stipulate a maximum impact parameter $b_{\rm max}$ to fix the Coulomb logarithm (e.g. equation 17), the Balescu–Lenard equation suffers from no large scale divergence and no such cutoff is needed. The largest scale of fluctuations in the system is accounted for naturally through the lowest-order basis elements $\Phi^{(p)}$ and wavenumbers $\mathbf{n}$.
- It takes the form of a Fokker–Planck equation in action space, with both friction and diffusion terms present, as it must following the physical arguments below equation (121).

TABLE II. Comparison of various pieces entering the Balescu–Lenard equation for homogeneous (equation 126) and inhomogeneous (equation 125) stellar systems.

	Homogeneous	Inhomogeneous
Orbital distribution	$f_0(\mathbf{v},t)$	$f_0(\mathbf{J},t)$
Basis decomposition	$\psi(\mathbf{r},\mathbf{r}') \propto \sum_{\mathbf{k}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}/k^2$	$\psi(\mathbf{r},\mathbf{r}') \propto \sum_{p} \Phi^{(p)}(\mathbf{r})\,\Phi^{(p)*}(\mathbf{r}')$
Response matrix	$\propto \frac{1}{k^2}\int \mathrm{d}\mathbf{v}\,\frac{\mathbf{k}\cdot\partial f_0/\partial\mathbf{v}}{\omega-\mathbf{k}\cdot\mathbf{v}}$	$\propto \sum_{\mathbf{n}}\int \mathrm{d}\mathbf{J}\,\frac{\mathbf{n}\cdot\partial f_0/\partial\mathbf{J}}{\omega-\mathbf{n}\cdot\boldsymbol{\Omega}}\,\Phi^{(p)*}_{\mathbf{n}}(\mathbf{J})\,\Phi^{(q)}_{\mathbf{n}}(\mathbf{J})$
Dressing of perturbations	$\propto 1/\lvert\epsilon_{\mathbf{k}}(\mathbf{k}'\cdot\mathbf{v}')\rvert^2$	$\lvert\psi^{\mathrm{d}}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\mathbf{n}'\cdot\boldsymbol{\Omega}')\rvert^2$
Resonance condition	$\mathbf{k}\cdot(\mathbf{v}-\mathbf{v}')=0$	$\mathbf{n}\cdot\boldsymbol{\Omega}=\mathbf{n}'\cdot\boldsymbol{\Omega}'$

- The right hand side is an antisymmetric functional of $f_0(\mathbf{J})$ and $f_0(\mathbf{J}')$. This tells us that the system is driven by *pairwise interactions*. In other words, the Balescu–Lenard equation describes a collection of orbits at action space location $\mathbf{J}$ interacting with orbits at all other locations $\mathbf{J}'$. It is remarkable that in this approximation, the collective evolution of an amazingly complicated N-star system acts *as if it were driven by uncorrelated two-body encounters* — though they interact not through the usual bare coupling $\psi=-G/|\mathbf{r}-\mathbf{r}'|$ but through the dressed pairwise interaction potential ψ^{d}. This is a direct manifestation of Rostoker's principle, which serves as the basis for the alternative derivation shown in §VII B.
- The Dirac delta function on the right hand side encodes the fact that interactions are not only pairwise but also *resonant*. In other words, the only meaningful interactions occur if some linear combination of frequencies of the star at $\mathbf{J}$ matches a linear combination of the frequencies of the star at $\mathbf{J}'$. This makes sense, since pairs of stars that do not satisfy this condition will undergo an interaction that vanishes on average — each star will feel an equal 'pull' and 'push' from the other as it traverses many mean field orbits, with a net result of zero.
- The equation conserves mass, energy, and satisfies a H-theorem for Boltzmann entropy. In addition, the right hand side exactly vanishes for inhomogeneous Boltzmann DFs (i.e. thermal distributions) of the form $f_0(\mathbf{J}) \propto \mathrm{e}^{-\beta H_0(\mathbf{J})}$ (although these DFs are not realizable in actual galaxies since they contain infinite mass).
- In the limit where collective amplification is ignorable ($\mathsf{M}=\mathsf{0}$, equation 101), we can replace $\psi^{\mathrm{d}}\to\psi$ and trivially recover the *Landau equation* which describes the evolution of an N-body system driven by a 'bare' resonant interactions[93].

Finally, in the limit of a homogeneous system one can use the angle-action mapping (37), and thereby recover (up to constants) the classic Balescu–Lenard collision operator from electrostatic plasma kinetics

$$\frac{\partial f_0(\mathbf{v})}{\partial t} \propto \frac{\partial}{\partial\mathbf{v}}\cdot\sum_{\mathbf{k}}\mathbf{k}\int \mathrm{d}\mathbf{v}'\delta[\mathbf{k}\cdot(\mathbf{v}-\mathbf{v}')]\frac{1}{k^4|\epsilon_{\mathbf{k}}(\mathbf{k}'\cdot\mathbf{v}')|^2}\,\mathbf{k}\cdot\left(\frac{\partial}{\partial\mathbf{v}}-\frac{\partial}{\partial\mathbf{v}'}\right)f_0(\mathbf{v})f_0(\mathbf{v}'). \tag{126}$$

In Table II, we provide a direct comparison between the various terms in the homogeneous equation (126) and the corresponding inhomogeneous equation (125). In particular, for a homogeneous system we have $|\psi^{\mathrm{d}}_{\mathbf{k}\mathbf{k}'}|^2 \propto \delta^{\mathbf{k}'}_{\mathbf{k}}k^{-4}|\epsilon_{\mathbf{k}}|^{-2}$, where ϵ is the dielectric function. This is why the homogeneous equation (126) involves only a single sum over wavenumbers $\mathbf{k}$, rather than the double sum over $\mathbf{n},\mathbf{n}'$ that we had in the inhomogneous equation (125). The factor $\propto|\epsilon_{\mathbf{k}}(\mathbf{k}'\cdot\mathbf{v}')|^{-2}$ in (126) is the classic dielectric shielding factor. In near-Maxwellian plasmas, this factor tends to suppress two-body interactions on large (bigger than Debye) scales, but in stellar systems it can strongly amplify such interactions — see §VII D for a detailed discussion.

2. An historical aside

Of course, the Balescu–Lenard equation was first derived in the context of homogeneous electrostatic plasmas[94,95]. The main mathematical difference in the stellar dynamical case is that one must formulate the equation in angle-action variables.

The basic idea of formulating a quasilinear diffusion equation in action space seems to have been first arrived at by Kaufman[84], Kaufman and Nakayama[96], Kaufman[97], in the context of inhomogeneous plasmas. The idea of doing *gravitational* kinetic theory in angle action variables was introduced around the same time by Kalnajs[63], who was predominantly interested in the linear properties of disk galaxies and their associated spiral modes (§VI). A proper diffusion equation of the form $\partial_t f_0=\partial_{\mathbf{J}}\cdot[\mathsf{D}\cdot\partial_{\mathbf{J}}f_0]$ was not written down in the stellar-dynamical context until a few

years later, by Dekker [98]. For a thorough overview of the history and of linear and quasilinear kinetics in angle-action variables (including their plasma precedents), see Chavanis [99].

The first derivation of the full Balescu–Lenard kinetic equation for stellar systems was in a somewhat opaque 1987 treatise by Luciani and Pellat [100]. The same equation was then arrived at in 2010 by Heyvaerts [101] using the BBGKY hierarchy; then in 2012 by Chavanis [92] using the Klimontovich description; again in 2017 by Heyvaerts *et al.* [102] using a Fokker–Planck description; then in 2018 by Fouvry and Bar-Or [91] using stochastic equations and Novikov's theorem; and finally in 2021 by Hamilton [103] using Rostoker's principle, as we now detail.

B. Alternative derivation of the Balescu–Lenard equation from Rostoker's principle

Rostoker's principle[57,104,105] tells us that we can think of a secularly evolving plasma or stellar system as a collection of N uncorrelated particles undergoing two-body encounters (à la Chandrasekhar), provided we replace the bare Coulombic/Newtonian interaction $\psi(\mathbf{r},\mathbf{r}')$ of each two-body encounter with its dressed counterpart $\psi^{\rm d}$. Roughly speaking, this is because particles are weakly coupled, and because the main effect of collective interactions upon the potential fluctuation spectrum is to dress each particle with its collective wake (equation 117). Here we use this principle to derive the Balescu–Lenard equation (125) in a short and simple way following Hamilton [103]. (Note we do not actually need to deal explicitly with fluctuations here, so we will drop the '0' subscript on the ensemble-averaged DF).

First, consider a 'test' star with coordinates $(\boldsymbol{\theta},\mathbf{J})$ and a 'field' star with coordinates $(\boldsymbol{\theta}',\mathbf{J}')$. Mathematically, Rostoker's principle tells us to forget about all the other stars, and treat the interaction of these two stars as if they were an isolated system with specific two-body Hamiltonian (units of (velocity)^2):

$$h = H_0(\mathbf{J}) + H_0(\mathbf{J}') + m\psi^{\rm d}(\boldsymbol{\theta},\mathbf{J},\boldsymbol{\theta}',\mathbf{J}'), \tag{127}$$

where $H_0(\mathbf{J})$ is the mean field Hamiltonian. Here $m\psi^{\rm d}(\boldsymbol{\theta},\mathbf{J},\boldsymbol{\theta}',\mathbf{J}')$ is the dressed specific potential energy between a star at phase space location $(\boldsymbol{\theta},\mathbf{J})$ and a star at $(\boldsymbol{\theta}',\mathbf{J}')$. (It consists of the usual Newtonian attraction plus collective effects; if we ignore these then $\psi^{\rm d}\to -G/|\mathbf{r}-\mathbf{r}'|$). Let us expand $\psi^{\rm d}$ as a Fourier series in the angle variables:

$$\psi^{\rm d}(\boldsymbol{\theta},\mathbf{J},\boldsymbol{\theta}',\mathbf{J}') = \sum_{\mathbf{n},\mathbf{n}'} \mathrm{e}^{\mathrm{i}(\mathbf{n}\cdot\boldsymbol{\theta}-\mathbf{n}'\cdot\boldsymbol{\theta}')}\,\psi^{\rm d}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\mathbf{n}'\cdot\boldsymbol{\Omega}'). \tag{128}$$

Here $\psi^{\rm d}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\mathbf{n}'\cdot\boldsymbol{\Omega}')$ is the dressed potential interaction we derived in §VI C, and we have put in the correct frequency dependence ($\omega=\mathbf{n}'\cdot\boldsymbol{\Omega}'$), since it is physically clear that these are the only frequencies available to the system. (Alternatively one can cheat: just think of $\psi^{\rm d}/m$ as some function of $\boldsymbol{\theta},\boldsymbol{\theta}',\mathbf{J},\mathbf{J}'$, and then identify the correct functional form at the end by comparison with 125).

Let us treat the two-body interaction as a perturbation to each star's motion. Using Hamilton's equation $\mathrm{d}\mathbf{J}/\mathrm{d}t = -\partial h/\partial\boldsymbol{\theta}$, we find that to zeroth order in this perturbation the test and field stars just follow their mean field trajectories $\boldsymbol{\theta}=\boldsymbol{\theta}_0+\boldsymbol{\Omega}t$ and $\boldsymbol{\theta}'=\boldsymbol{\theta}'_0+\boldsymbol{\Omega}'t$ indefinitely. To first order, similarly to equation (61), the result of their interaction is to nudge each other to new actions $\mathbf{J}+\delta\mathbf{J}$ and $\mathbf{J}'+\delta\mathbf{J}'$ respectively, where

$$\delta\mathbf{J}(\boldsymbol{\theta}_0,\mathbf{J},\boldsymbol{\theta}'_0,\mathbf{J}',\tau) = -m\sum_{\mathbf{n},\mathbf{n}'}\mathrm{i}\mathbf{n}\,\psi^{\rm d}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\mathbf{n}'\cdot\boldsymbol{\Omega}')\,\mathrm{e}^{\mathrm{i}(\mathbf{n}\cdot\boldsymbol{\theta}_0-\mathbf{n}'\cdot\boldsymbol{\theta}'_0)}\,\frac{\mathrm{e}^{\mathrm{i}(\mathbf{n}\cdot\boldsymbol{\Omega}-\mathbf{n}'\cdot\boldsymbol{\Omega}')\tau}-1}{\mathrm{i}(\mathbf{n}\cdot\boldsymbol{\Omega}-\mathbf{n}'\cdot\boldsymbol{\Omega}')}, \tag{129}$$

The result for $\delta\mathbf{J}'$ is the same as (129) except we replace the first factor $\mathrm{i}\mathbf{n}\to-\mathrm{i}\mathbf{n}'$. Note that because the two stars nudge eachother, this pairwise interaction conserves the total energy of the pair. This is different from considering an interaction between a test star and a background star (like in §II C 1).

Next, following Rostoker, we consider the relaxation of our entire system to consist of nothing more than an uncorrelated set of dressed two-body encounters. Then it is easy to write down a master equation for the DF $f(\mathbf{J})$. To do so, we account for (a) test stars being nudged out of the state $\mathbf{J}$ and in to some new state $\mathbf{J}+\Delta\mathbf{J}$, as illustrated in Figure 25a, and (b) test stars being nudged in to the state $\mathbf{J}$ from $\mathbf{J}-\Delta\mathbf{J}$, as in Figure 25b. Processes (a) and (b) are both deterministic and time-reversible, so we must also account for their inverses, which would be represented by the same diagrams but with the arrows pointing in the opposite direction. Let the *transition rate density* be $w(\Delta\mathbf{J},\Delta\mathbf{J}'|\mathbf{J},\mathbf{J}')$. This quantity is defined such that $w(\Delta\mathbf{J},\Delta\mathbf{J}'|\mathbf{J},\mathbf{J}')\,\mathrm{d}\Delta\mathbf{J}\mathrm{d}\Delta\mathbf{J}'\tau$ is the probability that a given test star with initial action $\mathbf{J}$ is scattered to the volume of phase space $\mathrm{d}\Delta\mathbf{J}$ around $\mathbf{J}+\Delta\mathbf{J}$, by a given field star with action $\mathbf{J}'$ that is itself scattered to the volume element $\mathrm{d}\Delta\mathbf{J}'$ around $\mathbf{J}'+\Delta\mathbf{J}'$, in a time interval τ that is much longer than an orbital period but much shorter than the relaxation time. Assuming the system's equilibrium state is

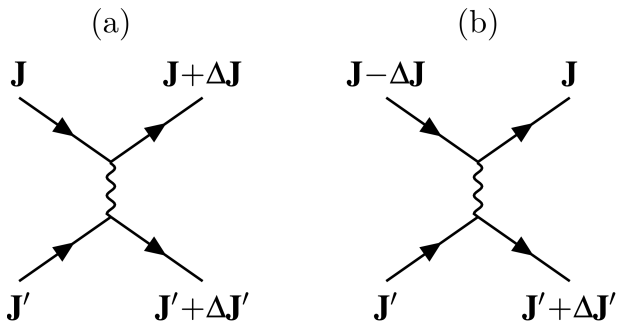

FIG. 25. Two possibilities for 'scattering' in action coordinates. A test star's action changes by $\Delta\mathbf{J}$ during an 'encounter' with a field star whose action changes from $\mathbf{J}'$ to $\mathbf{J}'+\Delta\mathbf{J}'$. The DF $f(\mathbf{J})$ will be decremented if the test star is kicked out of state $\mathbf{J}$ (process (a)), but will be incremented if it is kicked in to state $\mathbf{J}$ from $\mathbf{J}-\Delta\mathbf{J}$ (process (b)). To write down a master equation we must also account for the inverse processes (by reversing the directions of the arrows).

invariant under time reversal, f must satisfy the master equation[106]

$$\frac{\partial f(\mathbf{J})}{\partial t} = \frac{(2\pi)^d}{m}\int \mathrm{d}\mathbf{J}'\mathrm{d}\Delta\mathbf{J}\mathrm{d}\Delta\mathbf{J}'\frac{1}{2}\Big[w(\Delta\mathbf{J},\Delta\mathbf{J}'|\mathbf{J},\mathbf{J}')[-f(\mathbf{J})f(\mathbf{J}')+f(\mathbf{J}+\Delta\mathbf{J})f(\mathbf{J}'+\Delta\mathbf{J}')] \\ + w(\Delta\mathbf{J},\Delta\mathbf{J}'|\mathbf{J}-\Delta\mathbf{J},\mathbf{J}')[f(\mathbf{J}-\Delta\mathbf{J})f(\mathbf{J}')-f(\mathbf{J})f(\mathbf{J}'+\Delta\mathbf{J}')]\Big]. \tag{130}$$

By expanding the integrand for weak interactions, i.e. for $\Delta\mathbf{J},\Delta\mathbf{J}' \ll \mathbf{J},\mathbf{J}'$, up to second order in small quantities, we can show that

$$\frac{\partial f(\mathbf{J})}{\partial t} = \frac{\partial}{\partial \mathbf{J}}\cdot\int \mathrm{d}\mathbf{J}'\left[\mathsf{A}(\mathbf{J},\mathbf{J}')\cdot f(\mathbf{J}')\frac{\partial f}{\partial \mathbf{J}} + \mathsf{B}(\mathbf{J},\mathbf{J}')\cdot f(\mathbf{J})\frac{\partial f}{\partial \mathbf{J}'}\right]. \tag{131}$$

where A is the 3×3 matrix

$$\mathsf{A}(\mathbf{J},\mathbf{J}') = \frac{(2\pi)^d}{2m}\int \mathrm{d}\Delta\mathbf{J}\mathrm{d}\Delta\mathbf{J}' w(\Delta\mathbf{J},\Delta\mathbf{J}'|\mathbf{J},\mathbf{J}')\,\Delta\mathbf{J}\otimes\Delta\mathbf{J} \equiv \frac{(2\pi)^d}{2m}\frac{\langle\Delta\mathbf{J}\otimes\Delta\mathbf{J}\rangle_\tau}{\tau}, \tag{132}$$

where $\langle\Delta\mathbf{J}\otimes\Delta\mathbf{J}\rangle_\tau$ is the expectation value of $\Delta\mathbf{J}\otimes\Delta\mathbf{J}$ after a time interval τ for a given test star action $\mathbf{J}$ and field star action $\mathbf{J}'$. The matrix B is identical to A except we replace $\langle\Delta\mathbf{J}\otimes\Delta\mathbf{J}\rangle_\tau \to \langle\Delta\mathbf{J}\otimes\Delta\mathbf{J}'\rangle_\tau$. In a homogeneous system, $\mathbf{J}$ and $\mathbf{J}'$ can be taken to be the (suitably normalized) linear momenta of the test and field star; then momentum conservation implies $\Delta\mathbf{J} = -\Delta\mathbf{J}'$ so that $\mathsf{A} = -\mathsf{B}$. With this, one can easily recover equation (2.36) of the analogous calculation for a homogeneous plasma by Diamond *et al.*[57].

Now we put the above results together. Since by assumption stars at a given action are randomly distributed in angles, we can calculate the expectation value $\langle\Delta\mathbf{J}\otimes\Delta\mathbf{J}\rangle_\tau$ by averaging over the initial phases $\boldsymbol{\theta}_0,\boldsymbol{\theta}_0'$. (This is analogous to the integral over impact parameters b in Chandrasekhar's theory, as in equation (17)). Thus we have

$$\langle\Delta\mathbf{J}\otimes\Delta\mathbf{J}\rangle_\tau = \int \frac{\mathrm{d}\boldsymbol{\theta}_0}{(2\pi)^d}\frac{\mathrm{d}\boldsymbol{\theta}_0'}{(2\pi)^d}\,\delta\mathbf{J}(\boldsymbol{\theta}_0,\mathbf{J},\boldsymbol{\theta}_0',\mathbf{J}',\tau)\otimes\delta\mathbf{J}(\boldsymbol{\theta}_0,\mathbf{J},\boldsymbol{\theta}_0',\mathbf{J}',\tau), \tag{133}$$

where $\delta\mathbf{J}$ is given in equation (129). Plugging (133) and (129) in to (132) and taking the limit $\tau\to\infty$, and making use of the following two identities

$$\psi^{\mathrm{d}}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\omega_{\mathrm{R}}) = [\psi^{\mathrm{d}}_{-\mathbf{n},-\mathbf{n}'}(\mathbf{J},\mathbf{J}',-\omega_{\mathrm{R}})]^* \quad \text{for } \omega_{\mathrm{R}} \text{ real}, \tag{134a}$$

$$\lim_{\tau\to\infty}\left[|\mathrm{e}^{\mathrm{i}x\tau}-1|^2/x^2\tau\right] = 2\pi\delta(x), \tag{134b}$$

we find the following expression for the matrix A:

$$\mathsf{A}(\mathbf{J},\mathbf{J}') = \pi(2\pi)^d m\sum_{\mathbf{n},\mathbf{n}'}\mathbf{n}\otimes\mathbf{n}\,\delta(\mathbf{n}\cdot\boldsymbol{\Omega}-\mathbf{n}'\cdot\boldsymbol{\Omega}')|\psi^{\mathrm{d}}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\mathbf{n}'\cdot\boldsymbol{\Omega}')|^2. \tag{135}$$

The result for B is identical to (135) except we replace the factor $\mathbf{n}\otimes\mathbf{n}$ with $-\mathbf{n}\otimes\mathbf{n}'$. Putting the explicit formulae for $\mathsf{A}(\mathbf{J},\mathbf{J}')$ and $\mathsf{B}(\mathbf{J},\mathbf{J}')$ in to the kinetic equation (131), we recover the Balescu–Lenard equation (125).

Thus, just as in plasma physics[57], the Balescu–Lenard collision operator can be interpreted as arising from the uncorrelated pairwise scattering of *dressed* stars as they traverse their (otherwise unperturbed) mean field orbits.

C. Examples

Implementing the Balescu–Lenard equation (125) in practice is hard, and has only been attempted a handful of times. Here, we mention some of these implementations, starting with an artificial one-dimensional gravitational system (§VII C 1), then moving on to a hot sphere (§VII C 2) and finally a cold disk (§VII C 3). Other applications include the inhomogeneous Hamiltonian Mean Field model[107], as well as the relaxation of the eccentricities[108] and orientations[109] of stellar orbits around supermassive black holes.

1. *Example: Balescu–Lenard in one-dimensional gravity*

Recently, Roule *et al.*[110] considered the evolution of a one-dimensional 'gravitational' N-body system. Their system is gravitational in the sense that density and potential are linked through Poisson's equation. Yet, because it is one-dimensional the resulting bare interaction between particles is $\psi(x,x') \propto |x-x'|$, and the force between any two particles is proportional to $\mathrm{sgn}(x-x')$. In this simple setup, there is a one-to-one mapping between action J and energy E, and the authors tend to use E to characterize orbits. The initial DF the authors choose for the particles is closely analogous to the isotropic DF that generates the three-dimensional Plummer sphere whose orbits we studied in §III A 1. Crucially for the application of Balescu–Lenard theory, the DF is linearly stable.

Roule *et al.*[110] simulated this one-dimensional N-body system directly many times over with random initial conditions, in order to produce an ensemble of systems over which to average. This was to be compared with the prediction from the Balescu–Lenard equation (125). They also performed simulations 'without collective effects'. To do this they first integrated the mean field motion of random ensembles of massive but non-interacting particles in their chosen model. They then used the fluctuating gravitational potential generated by these particles as the input to a test particle simulation. Since test particles, by definition, do not produce potential fluctuations of their own, they experience no friction (so $\mathbf{F}_1 = \mathbf{0}$ in equation 119) but they do experience diffusion ($\mathbf{F}_2 \neq \mathbf{0}$). These test particle simulations were to be compared to the predictions of 'Landau' theory, i.e. diffusion along the lines of Balescu–Lenard but without collective amplification (§VII A 1).

Figure 26 shows Roule *et al.*[110]'s results. The red crosses in this figure show measurements of the diffusion coefficient

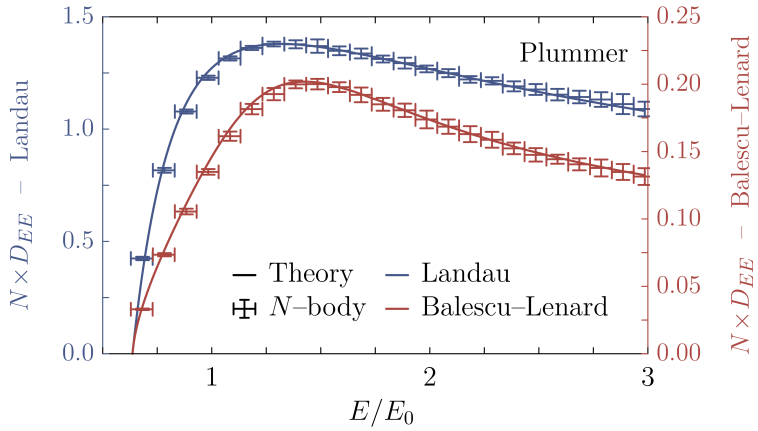

FIG. 26. Diffusion coefficients of particles in a one-dimensional self-gravitating system (§VII C 1). Crosses show the results of N-body simulations, while solid curves show the theoretical results based on the Balescu–Lenard equation (125) (when collective interactions are switched on) or the corresponding Landau equation (when they are switched off). Note that the vertical axis scale is different for red and blue points. Figure taken from Roule *et al.*[110].

in energy of particles over a time long compared to the dynamical time $t_{\rm dyn}$ but short compared to the relaxation time $t_{\rm relax}$. The red solid line shows the theoretical 'Balescu–Lenard' prediction. It demonstrates an excellent fit, confirming that equation (125) is capturing the physics of this system. Meanwhile the blue crosses and blue solid line shows that the 'Landau' theory is an excellent predictor of the diffusion coefficient in the case without collective effects.

Let us emphasize that the red and blue data in Figure 26 are plotted on different y-axes. In particular, the Landau diffusion coefficient is ~ 1 in these units, which would correspond to a relaxation time $t_{\rm relax} \sim N t_{\rm dyn}$, similar to Chandrasekhar's theory (equation 19). What is very notable is that over a large range of energies E, the Balescu–Lenard diffusion coefficient is ten times smaller than the Landau one, corresponding to a relaxation time $t_{\rm relax} \sim$

$10Nt_{\rm dyn}$. In other words, this is a self-gravitating system in which collective effects tend to quench, rather than amplify, the rate of evolution. While this result is unusual, it is not unique. A similar result was obtained by Weinberg[111], who studied the dynamical friction drag on a satellite falling into a spherical galaxy (a similar calculation to the one we will present in §VIII B 1). Weinberg found that when collective effects were included, the wake that the satellite induced in the galaxy was highly symmetric and that this led to a drag which was *weaker* than in the bare case without collective effects. The lesson we draw from these examples is that collective effects can have a major impact upon the evolution of inhomogeneous stellar systems, but it may be difficult to guess in advance what that impact will be.

2. Example: Balescu–Lenard in a hot sphere

On the surface, Balescu–Lenard theory is much more general than Chandrasekhar's theory of two-body relaxation (§II C 1). One might therefore ask (i) whether it contains Chandrasekhar's theory as a limiting case, and (ii) given that Chandrasekhar's theory works very well in hot spheres (§II C 3), does a direct application of the results of Balescu–Lenard theory to those systems give an answer that differs from Chandrasekhar's theory?

The answer to question (i) is yes: one can recover Chandrasekhar's theory from the Balescu–Lenard formalism by ignoring collective effects and confining oneself to interactions on small scales over which the curvature of mean field orbits can be neglected (replacing them with straight lines[93]).

To answer question (ii) requires a lot more work. Fouvry *et al.*[19] investigated it by evaluating the right hand side of equation (125) for an *isochrone sphere*, which is a toy model of a spherical star cluster. Using a known (stable) DF, they calculated the Balescu–Lenard flux $\mathbf{F}$ as a sum of contributions $\mathbf{F}^\ell$ from spherical harmonic quantum numbers $\ell = 0, 1, 2, \ldots$. As illustrated in Figure 27, they found that on small scales $\ell \gg 1$, the modulus $|\mathbf{F}_\ell|$ scales as $\sim \ell^{-1}$. Hence, a sum over ℓ recovers the small scale logarithmic divergence familiar from Chandrasekhar's theory, a result

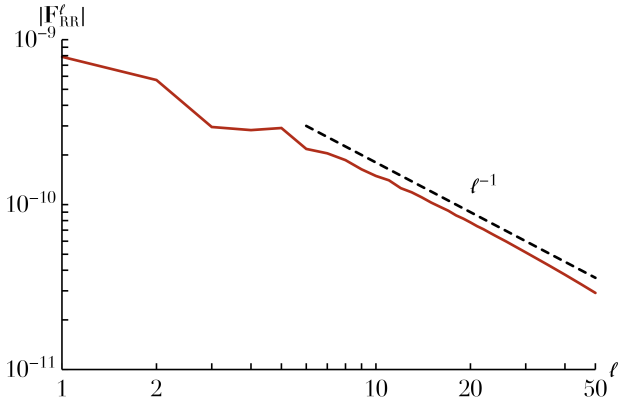

FIG. 27. Balescu–Lenard flux, $\mathbf{F}$ (see equations 119-121), for fluctuations with spherical harmonic quantum number ℓ (an angular scale $\sim \ell^{-1}$). The magnitude of this flux scales as ℓ^{-1} for large $\ell \gg 1$. This confirms that the Balescu–Lenard equation recovers the Coulomb logarithm behavior from Chandrasekhar's theory in the limit of local encounters (§II C 1). Figure taken from Fouvry *et al.*[19].

already obtained by Weinberg[112] when considering dynamical friction (see Figure 5 therein). Such a divergence is always present in a kinetic theory based on weak $\pm 1/r^2$ interactions. This strongly suggests that Chandrasekhar's theory — which is much simpler to implement than the Balescu–Lenard theory — is being recovered by equation (125) on small scales. This motivated Fouvry *et al.*[19] to calculate the Balescu–Lenard flux as accurately as possible on large scales ($\ell \leq \ell_{\rm crit}$, with $\ell_{\rm crit} \sim 12$, say), and to calculate the Chandrasekhar flux on small scales ($\ell > \ell_{\rm crit}$), and then add the two contributions together. The advantage of this over just using Chandrasekhar's theory is that one is not plagued by any large-scale divergence (which is normally cured by imposing a maximum impact parameter $b_{\rm max}$), and in principle one accounts for any resonant and/or collective phenomena that are ignored by Chandrasekhar. The result of this exercise, which is shown in Figure 28, is a slightly improved agreement between the contour map of $\partial f_0/\partial t$ in action space computed from theory as compared to N-body simulations. Of course, as we discussed in §II C 3, the agreement based on Chandrasekhar's theory alone is already good, and as such the large scale Balescu–Lenard 'correction' is small.

In summary, perhaps the most intellectually satisfying approach to spherical cluster kinetics is to couple a a large-scale Balescu–Lenard flux with a small-scale Chandrasekhar flux. But in practice, Chandrasekhar's theory applied to

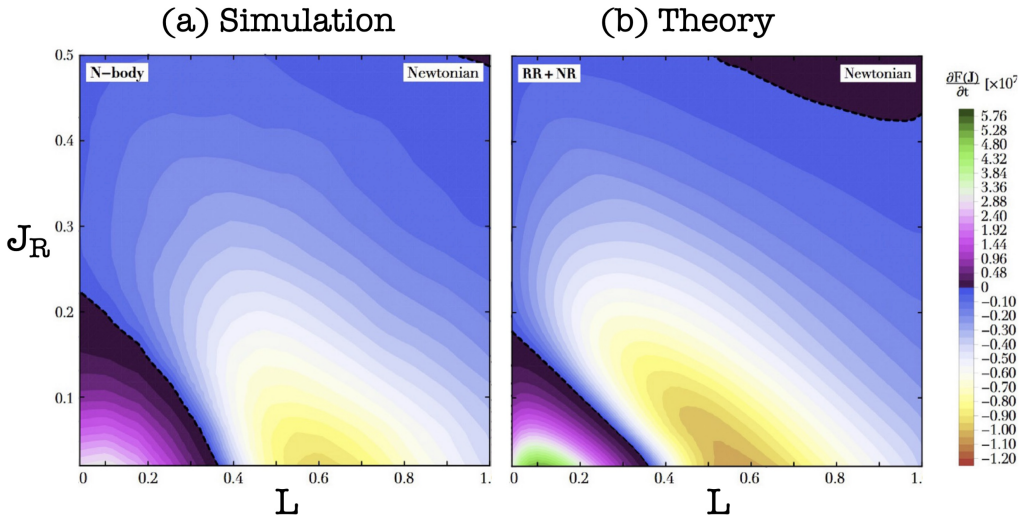

FIG. 28. Contour plot of the evolution rate $\partial f_0/\partial t$ in action space, for a simple model of a spherical cluster as measured from (a) direct N-body simulations and (b) a theory which combines the Balescu–Lenard equation (125) on large scales with the Chandrasekhar theory on small scales. Figure taken from Fouvry *et al.* [19].

all scales (with an additional cutoff at large scales $b_{\max}$) is almost as good a description, and is both much simpler to understand and easier to implement.

3. *Example: Balescu–Lenard in a cold disk*

As a final example, let us discuss the dynamics of an initially stable, razor thin stellar disk, following the N-body simulations of Sellwood [113] and the kinetic predictions from Fouvry *et al.* [114]. The model employed by these authors is the stable Mestel disk, with a DF f_0 whose $J_R = 0$ slice we plotted with a dotted black curve back in Figure 21a. Notably, this DF is very smooth in action space. However, Sellwood [113]'s simulations of this disk shows that after many dynamical times $t_{\rm dyn}$, a grooved/ridged feature always develops in the action space DF (see Figure 29a). The

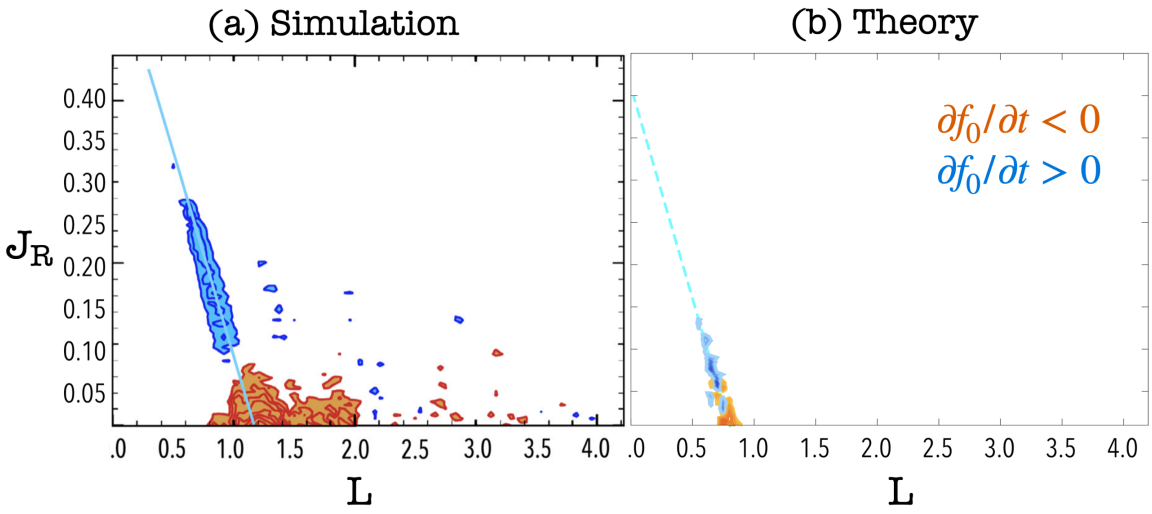

FIG. 29. Change in the DF $f_0(L, J_R)$ for an initially smooth, razor thin stellar disk, (a) as measured in an N-body simulation, (b) as predicted by the Balescu–Lenard equation (125) (figure adapted from Fouvry *et al.* [114]). The DF develops a ridge/groove feature which renders it unstable to spiral Landau modes — see Figure 21.

time taken for this feature to develop increases linearly with the number of particles N employed, suggesting it is driven by the finite-N noise. Remarkably, this newly-modified DF is found in the simulations to be unstable to the

formation of exponentially growing spiral waves.

In fact, this modified DF is precisely the one studied by De Rijcke *et al.*[79] in their linear stability analysis (§VI B 2), the $J_R = 0$ slice through which we showed with a solid black curve in Figure 21a. The corresponding spiral Landau mode they calculated from linear theory (which we showed Figure 21b) matches that extracted by Sellwood[113] from his simulations, confirming that his spiral waves are true linear instabilities of the modified DF. Thus, our stable, isolated stellar disk has driven itself unstable!

We can usefully contrast this behavior with that of a plasma. If we initiated e.g. a finite-N electrostatic plasma with a Maxwellian DF, then the plasma would exhibit thermal fluctuations $\delta\Phi, \delta f \sim N^{-1/2}$, but its mean-field Maxwellian f_0 would simply never evolve. That is because the Maxwellian is the ultimate maximum entropy state in a homogeneous plasma, a perfect thermodynamic equilibrium. But for self-gravitating systems there is no such thermodynamic equilibrium (§II B), and so even a linearly a stable stellar disk is never at rest: instead it gradually carves a groove in itself and drives itself unstable, absent any external influences. It is as if a stable plasma were to spontaneously grow a bump on its own tail!

Before the system becomes unstable, its evolution can be predicted using the Balescu–Lenard equation (125), which after all is supposed to describe the long term evolution of stable systems driven by amplified finite-N noise. Fouvry *et al.*[114] evaluated the right hand side of equation (125) for the same Mestel disk DF used by Sellwood[113]. The resulting prediction is shown in Figure 29b. While there are some important differences between this panel and the result in panel (a), the key physics is quite clear: the grooved/ridged DF is formed because of dressed, resonant coupling between finite-N fluctuations in the disk, which happen to be strongly enhanced along a particular direction in action space (in this case $\mathbf{n} = (2, -1)$, shown with light blue lines in the Figure). Unfortunately, the Balescu–Lenard equation is so cumbersome to implement in practice that nobody has managed to push the theory any further than this.

D. Breakdown of Balescu–Lenard theory; fundamental differences between galaxies and plasmas

As discussed at the beginning of this section and in §V E, any linear or quasilinear kinetic theory will break down if strongly nonlinear effects such as orbit trapping are important to the evolution. We will develop a method for dealing with such systems in §VIII. However, there is another way in which the Balescu–Lenard theory can break down even in the absence of strong nonlinearities. Namely, the Balescu–Lenard equation (125) is derived under a key assumption which is that *we may ignore the direct influence of the Landau modes*. This assumption is often a poor one when applied to real galaxies. This important point is often neglected in the literature, and a discussion of it allows us to highlight some fundamental differences between stellar systems and plasmas, as we will see.

First, Balescu–Lenard theory is designed to describe the evolution of an N-body system in which dressed star-star interactions alone drive the evolution. Mathematically, this assumption enters the Balescu–Lenard theory when one uses equation (117) for the potential fluctuations. This equation is valid on timescales long compared to the damping time of the system's most weakly damped Landau mode $\omega_{\rm m}$, which requires

$$t_{\rm relax} \gg 1/|{\rm Im}\,\omega_{\rm m}|. \tag{136}$$

This requirement should already be making us worried. Recall that for real galaxies $t_{\rm relax}$ may be only of order only a few Gyr, and many stellar systems are capable of harboring weakly damped Landau modes that live for many $t_{\rm dyn}$, which is ~ 250 Myr in the Milky Way. Moreover, the right hand side of (136) is by definition *infinite* during the transition from a stable to unstable DF described in §VII C 3, so the Balescu–Lenard theory is clearly invalid in that regime. To make matters worse, in this regime the Balescu–Lenard equation actually gives nonsensical answers. To see this, note that a Landau mode with Im $\omega_{\rm m} \to 0$ corresponds to a divergence in $\psi^{\rm d}$ (equation 115) at some real frequency, and the Balescu–Lenard equation (125) involves an integral of $|\psi^{\rm d}|^2$ over all real frequencies. Thus, a naive application of the Balescu–Lenard theory to this marginally stable regime will necessarily return an unphysical, divergent result.

How might we do better? Note that in plasma kinetics, the classic quasilinear (QL) theory of wave-particle interactions[83] does just the opposite of the Balescu–Lenard theory: it includes the Landau modes while ignoring the dressed discreteness noise. (Throughout this discussion we use the abbreviation 'QL' to distinguish the classic theory of particles interacting with Landau modes from the generic quasilinear scheme of constructing kinetic equations by multiplying two linear quantities together, see §V D). This assumption *is* justified in certain plasma regimes. In fact, in near-Maxwellian electrostatic plasmas, the Balescu–Lenard and QL regimes are physically separated by the Debye length $\lambda_{\rm D}$.

Let us be more precise. In an electron plasma, the 'dressing' of finite-N noise is accomplished by Debye shielding: for instance, a positive charge tends to accumulate a sphere of negative charge around it, screening its Coulomb potential (Figure 22). This means that on scales larger than $\lambda_{\rm D}$ the amplitude of 'dressed noise' drops to zero and

one can ignore particle-particle interactions altogether, i.e. the contribution to the Balescu–Lenard equation from these scales is negligible. Meanwhile, there exist Landau modes on all scales: these are Langmuir waves, which involve longitudinal oscillations of electrons as they try to maintain the plasma's quasi-neutrality. Langmuir waves oscillate very rapidly — typically their associated timescale, the inverse plasma frequency $\omega_{\rm p}^{-1}$, is the shortest physical timescale in the system. Now, we know from Landau's theory of wave-particle damping that these modes will damp only if they can resonate with electrons with speeds $v \sim \omega_{\rm p}/k$, where k is the wavenumber of the Langmuir oscillation in question. Thus for small k (large scales), the only electrons capable of resonating with the wave are those moving extremely fast, and since these are few in number, small k Langmuir waves tend to be only very weakly damped, and so wave-particle interactions must be taken into account in the kinetic theory via a QL collision operator. Conversely, for very large k the waves are able to resonate with relatively slowly-moving particles (the 'thermal bulk'), of which there are many — hence, large-k Langmuir oscillations are typically very strongly damped. Again it turns out that the separation between these two regimes occurs at the Debye scale $k_{\rm D} = 2\pi/\lambda_{\rm D}$. In short, in an electron plasma the electric potential fluctuations on scales $k \gg k_{\rm D}$ are dominated by collectively-dressed noise and therefore well-described by Balescu–Lenard theory (a 'noise-only' regime), whereas on scales $k \ll k_{\rm D}$ Landau modes dominate and one may use QL theory ('modes-only').

The reader has already anticipated where this argument fails when applied to self-gravitating systems. In this case there is no Debye shielding, so there is nothing to suppress discreteness noise on *any* scale. On small scales (say, the interparticle distance) collective effects are very weak, so the power spectrum just reflects bare Poisson noise. This is why Chandrasekhar's theory works so well there (§VII C 2). On large scales (comparable to the size of the system) there typically exists a combination of weakly-damped Landau modes and collectively-amplified finite-N noise, and crucially *it is near the frequencies of weakly-damped Landau modes that the collective amplification of the finite-N noise tends to be most pronounced.* In other words, the power spectrum of the continuum resulting from dressed noise ($\sim |\psi^{\rm d}_{\mathbf{nn}'}|^2$) is sharply peaked at frequencies $\mathbf{n}' \cdot \mathbf{\Omega}'$ which are very close to the (real part of the) frequency $\omega_{\rm m}$ of weakly-damped Landau modes. This argument shows the logical gap that can arise when naively applying the Balescu–Lenard equation to stellar systems. For in many cases, it is only worth putting in the effort to implement the Balescu–Lenard equation if its predictions are significantly different from those of the much simpler Chandrasekhar two-body relaxation theory, i.e. if there is significant collective dressing of finite-N noise on large scales. But this dressing in turn is typically only significant *if the system harbors weakly damped Landau modes.* It is therefore somewhat self-contradictory to derive Balescu–Lenard theory by arguing these modes fade away rapidly — if they really did so, there might be no interesting collective behaviour to speak of, and therefore no need for the Balescu–Lenard equation! (The 1D system discussed in §VII C 1 is an exception to this rule, since in that case the collective effects tended to *suppress* interactions rather than amplify them, and there are no weakly damped Landau modes.)

It follows that typically in stellar systems, amplified noise and Landau modes cannot naively be decoupled. A comprehensive kinetic theory must therefore account for both types of potential fluctuation self-consistently. Of course, this is also a challenge that arises in plasma kinetics since neither the Balescu–Lenard or QL theory is strictly valid on scales $k \sim k_{\rm D}$, and a general theory must incorporate both effects. Several plasma theorists attempted to formulate such a general theory in the 1960s, notably Rogister and Oberman [115,116], Oberman [117] and Hatori [118,119]. These authors considered a homogeneous electrostatic plasma and derived a set of self-consistent coupled equations for the particle distribution function and the energy in the electric field fluctuations, accounting for both Balescu–Lenard and QL-type effects. It turns out that the divergence of the Balescu–Lenard part of the theory when Im $\omega_{\rm m} \to \pm 0$ is accompanied by a corresponding divergence in the opposite sense in the QL contribution. This renders the generalized theory divergence-free for any Im $\omega_{\rm m}$. (In fact, this cancellation of divergences arises already in linear theory — see Hamilton and Heinemann [120]). A complete, unified kinetic theory of stellar systems would account for both types of potential fluctuations (dressed particle noise and Landau modes), while also taking account of the inhomogeneity of the system using angle-action variables. This is a topic of ongoing research [120,121].

VIII. NONLINEAR KINETIC THEORY

The evolution of a system's mean DF f_0 is always 'nonlinear' in the sense that it depends on the product of two 'fluctuations' $\delta\Phi$ and δf (equation 118). In §VII we discussed a quasilinear approximation to this evolution in which those fluctuations were each calculated using linear theory. Here, we discuss nonlinearities that cannot be captured with a quasilinear approach. We begin in §VIII A by discussing some of the ways in which these nonlinearities develop, and we identify one that is typically most important for stellar systems, namely when orbits are strongly distorted and/or trapped by a single coherent perturbation. Then in §VIII B we write down and solve a kinetic equation which describes this regime, using the pendulum machinery that we developed for individual orbits in §IV F. Finally we give two concrete examples of this nonlinear kinetic theory in practice: first in the context of galactic bar-dark matter halo

interactions and dynamical friction (§VIII B 1), and then regarding the saturation of spiral instabilities (§VIII B 2).

A. Which nonlinear regimes are relevant for stellar systems?

Consider a stellar system that is initially completely quiet ($\delta f = \delta\Phi = 0$), and subject it to some perturbation. Then we know that at least for a short time, the system will be well described by linear theory (§VI), and so e.g. its potential/density response will consist of a combination of amplified wakes and Landau modes. If the system is unstable (and in some cases even if it is weakly stable) the response will grow large enough that the system will eventually cease to be well described by linear theory.

Why, physically, does linear theory break down? Let us list a few of the main possibilities.

- One reason can be that the DF f_0 relaxes quasilinearly, in the manner already discussed in §VII. Since linear theory is based on a constant f_0 (e.g. equation 94), such an evolution will eventually break the initial linear prediction (effectively 'resetting the initial value problem' with a new DF and fluctuation spectrum). This will be the dominant source of nonlinearity if the evolution time of f_0 according to the relevant quasilinear theory — which is naively on the order of $t_{\rm relax} \sim \varepsilon^{-2} t_{\rm dyn}$, where $\varepsilon \sim |\delta\Phi/H_0|$ is the typical fluctuation strength, see equation (118) — is shorter than the timescale over which any alternative nonlinearities can develop.
- One such alternative, which is especially important if the system is unstable, is that nonlinear evolution can arise in the form of direct coupling between Landau modes. That is, the (initially independent) Landau mode solutions attain such large amplitudes that they no longer respond solely to the mean Hamiltonian H_0 and the action space gradient $\partial f_0/\partial \mathbf{J}$ (as would be dictated by linear theory), but instead start to notice each others' gravitational fields. Landau mode coupling can lead to modifications both of the potential fluctuation spectrum and the mean stellar distribution function f_0. For this effect to be significant, the system usually has to harbor at least two Landau modes that reach amplitudes $\varepsilon \sim 1$ on a short enough timescale that they can interact *before* being destroyed by some other process (such as the quasilinear evolution referred to above). The specific nonlinear physics involved here depends on whether (i) only a handful of Landau modes are excited to large amplitudes, or (ii) a broad spectrum of many such modes is excited.

  Case (i), in which only a handful of modes are important, has been studied to some extent in galactic dynamics, mostly in the context of global spiral-spiral interactions[122,123] (in what plasma physicists would refer to as 'three wave coupling'). However, the timescale required for even a rather vigorously-growing global spiral instability to reach a nonlinear amplitude is usually on the order of several Gyr (see §VIII B 2 for a quantitative estimate). This is often comparable to the timescale over which the bulk of the galaxy tends to evolve (§II C 3). This leaves little time for such modes to interact nonlinearly (recall from §I that *a typical disk galaxy is only* ~ 50 *orbits old*!) As a result, spiral-spiral coupling and similar effects seem to be *present but not dominant* in galactic disk evolution.

  Case (ii) on the other hand, in which many modes interact, is usually referred to by plasma physicists as the regime of 'wave turbulence'[54]. That is, plasma physicists tend to think of such a system as consisting of 'particles' (described by f_0) and 'waves' (Landau modes, which dominate the fluctuation spectrum). Then, they construct kinetic theories in which wave-wave coupling plays a dominant role, and in which the evolution of f_0 is deemed secondary or just ignored altogether. Some studies of this flavor have been conducted in galactic dynamics, mainly with reference to tightly wound spiral waves[124–126]). Some authors in the 1970s[127,128] took the plasma analogy so far as to construct versions of spiral structure theory governed by a nonlinear Schrödinger equation, and attempted to describe spirals as solitons/modulational instabilities in the way that Zakharov [129] described Langmuir collapse. However, no serious numerical evidence was ever offered to bolster this formalism. Moreover, as in case (i), it is only approximately correct to freeze the stellar distribution function f_0 in order to isolate the evolution of nonlinearly interacting modes. Even if it were possible, the dynamical youth of galaxies means that the number of wave cycles over which such modes can possibly have interacted is much less than is normally assumed in plasma theory.
- A third possibility is for the system to be 'dynamically degenerate'. For example, with $\partial f_0/\partial \mathbf{J} = 0$ and $\partial H_0/\partial \mathbf{J} = 0$, equation (93) exactly vanishes and the linearized dynamics of fluctuations becomes trivial. The system's evolution is then intrinsically non-linear, and one must resort to statistical closure theory to make progress[56]. Astrophysically such regimes occur when describing the diffusion of stellar orbital orientations around supermassive black holes[130,131] or dynamical friction in harmonic cores in the center of galaxies[132,133]. This regime bears deep connexions with (anomalous) two-dimensional diffusion in magnetized plasmas[134].
- A fourth possibility is that stellar orbits can be trapped, or otherwise strongly distorted, by a single coherent perturbation. For instance, stars can be trapped at the corotation resonance of a rotating bar or spiral (Fig-

ure 13). This effect can be important if the libration time $t_{\rm lib}$ of trapped orbits is shorter than or comparable to the evolution time of the distribution function f_0 in the relevant region of phase space, *and* shorter than the timescale over which some other effect (like incoherent 'collisions' or the sweeping of resonances through phase space) is able to interrupt the libration process, as discussed in §IV F 2.

The last of these nonlinear effects, orbit trapping by a single perturbation, seems to be important in sculpting the evolution of both real and simulated galaxies[41], and there is good evidence that it has occurred in our own Galaxy[135,136]. Therefore, for the rest of this section we will focus only on this type on nonlinearity. We already showed in §IV F how one can describe individual orbits in this regime using the pendulum approximation. Below, we show how a collection of pendulum orbits can be described in a kinetic formalism, although we will ignore self-consistency (i.e. Poisson's equation). We will, however, allow for the possibility that the nonlinearity is interrupted due to stochastic 'collisions', by appending an ad-hoc diffusion operator to our (otherwise collisionless) kinetic equation.

We emphasize, though, that in certain contexts more than one of the above nonlinear effects can be important in different parts of the system and at different times. For instance, in Sellwood and Carlberg [137]'s simulations of grooved unstable stellar disks a spiral Landau mode initially grows exponentially in accordance with linear theory. Once it reaches a large enough amplitude it begins to trap stars nonlinearly at the corotation resonance, which causes its amplitude to saturate (§VIII B 2). At the same time, the spiral interacts quasilinearly with stars at the Lindblad resonances, draining its angular momentum and carving a new groove in the DF at the location of the ILR. This newly-grooved disk is unstable to a new set of linear spiral eigenmodes, and so on in a recurrent cycle[138,139].

B. Kinetic theory in the pendulum approximation

In §IV F, we described the nonlinear dynamics of individual stars orbiting near resonances of a long-lived potential perturbation rotating rigidly with some fixed pattern speed. We did this by isolating a particular resonance, defining an associated set of slow and fast angle-action variables, and averaging out the fast angles. We were left with dynamics in the slow angle-action space (φ, I) governed by the pendulum Hamiltonian h (equation 66), all at a given (fixed) fast action $J_{\rm f}$. We now study how an ensemble of stars behaves near such a resonance, following Hamilton *et al.* [140].

Let us therefore consider a subset of stars that all share the same fast action $J_{\rm f}$. We will describe the density of these stars in the (φ, I) plane using a smooth DF which we call $f(\varphi, I, t)$, defined such that the number of stars in the phase space area element $\mathrm{d}\varphi\,\mathrm{d}I$ surrounding (φ, I) at time t is proportional to $\mathrm{d}\varphi\,\mathrm{d}I\,f(\varphi, I, t)$. Since (φ, I) are canonical variables, the kinetic equation governing f is

$$\frac{\partial f}{\partial t} + [f, h] = C[f], \tag{137}$$

where $[\cdot, \cdot]$ is the Poisson bracket encoding the smooth advection in the (φ, I) plane

$$[f, h] = \frac{\partial f}{\partial \varphi}\frac{\partial h}{\partial I} - \frac{\partial f}{\partial I}\frac{\partial h}{\partial \varphi}, \tag{138}$$

and we have included a possible 'collision operator' $C[f]$, which encodes the effect of any stochastic fluctuations in the potential.

For concreteness, let us consider the interaction of a collection of stars (described by f) interacting with a rigidly rotating stellar 'bar' (a collection of millions of stellar orbits at the center of a galaxy that rotates like a dumbbell). Then for h we can use the pendulum form (66) with $k = 1$. Meanwhile, for $C[f]$ we choose a very simple diffusive form with constant diffusion coefficient D

$$C[f] = D\,\frac{\partial^2 f}{\partial I^2}. \tag{139}$$

This corresponds to every star in our ensemble being kicked with a stochastic white noise forcing in the slow action, $\dot{I} = \eta(t)$, with statistics $\langle \eta(t) \rangle = 0$ and $\langle \eta(t)\eta(t') \rangle = 2D\delta(t - t')$. (For simplicity we will ignore any such forcing in φ and $J_{\rm f}$). The diffusion coefficient D can be calculated from a given theoretical model of, e.g. scattering by passing stars and molecular clouds[90,141,142], dark matter substructure[143–145], spurious numerical heating in simulations[146,147], or can potentially be calibrated from data[16,148] (see below). We emphasize that we make no attempt at self-consistency here. We simply impose a diffusion coefficient by hand, and we do not account at any stage for the self-gravity of the perturbed DF (the stars interact only with the bar, not with each other[149,150]).

Before attempting to solve the kinetic equation (137), one more simplification is in order. Naively, equation (137) seems to depend on three parameters: G, F (equation 66) and D (equation 139). However, we can reduce this to one

effective parameter by introducing certain dimensionless variables. First we note that the typical timescale for a star to diffuse all the way across the resonance (a distance $\sim 2I_\mathrm{h}$ in action space) is the *diffusion time*,

$$t_\mathrm{diff} \equiv \frac{(2I_\mathrm{h})^2}{2D} = \frac{8F}{GD}. \tag{140}$$

Now let us define the dimensionless variables

$$\tau \equiv \sqrt{GF}t = \frac{2\pi t}{t_\mathrm{lib}}, \tag{141a}$$

$$j \equiv \sqrt{\frac{G}{F}}I = \frac{2I}{I_\mathrm{h}}, \tag{141b}$$

$$\Delta \equiv \sqrt{\frac{G}{F^3}}D = \frac{4t_\mathrm{lib}}{\pi t_\mathrm{diff}}, \tag{141c}$$

where the libration time t_lib is given in equation (68) and the island half-width I_h is defined in equation (69). Clearly, τ is a dimensionless measure of time normalized by the libration time, j is a dimensionless measure of the slow action variable relative to the resonant island width, and Δ is the *diffusion strength*, i.e. the ratio of the libration timescale to the diffusion timescale. Treating f as a function of these variables, i.e. writing[151] $f(\varphi, j, \tau)$, equation (131) becomes:

$$\frac{\partial f}{\partial \tau} + j\frac{\partial f}{\partial \varphi} - \sin\varphi \frac{\partial f}{\partial j} = \Delta \frac{\partial^2 f}{\partial j^2}. \tag{142}$$

We see that in these variables the kinetic equation depends on a single parameter, the dimensionless diffusion strength $\Delta \approx t_\mathrm{lib}/t_\mathrm{diff}$. The regime $\Delta \gg 1$ corresponds to very strong diffusion, whereas $\Delta \ll 1$ corresponds to very weak diffusion; the 'collisionless limit' in which stars only interact with the bar and nothing else is $\Delta = 0$. We can estimate Δ heuristically by noting that t_diff is related to the relaxation time t_relax, which is the time required for any stochastic process to change a star's action by order of itself. Then from equation (140) we have

$$\frac{t_\mathrm{diff}}{t_\mathrm{relax}} \sim \left(\frac{I_\mathrm{h}}{J_\mathrm{s,res}}\right)^2 \sim \left|\frac{\Psi_1}{H_0}\right| \sim \varepsilon \ll 1. \tag{143}$$

Here $\varepsilon \sim |\Psi_1/H_0|$ is the dimensionless strength of the bar; for the Milky Way at corotation $\varepsilon \sim 0.02$[135]. Putting equations (141c) and (143) together, we estimate

$$\Delta \sim \frac{t_\mathrm{lib}}{\varepsilon t_\mathrm{relax}} \sim 1 \times \left(\frac{\varepsilon}{0.02}\right)^{-1} \left(\frac{t_\mathrm{lib}}{2\,\mathrm{Gyr}}\right) \left(\frac{t_\mathrm{relax}}{100\,\mathrm{Gyr}}\right)^{-1}. \tag{144}$$

Typically, important Galactic bar resonances like the corotation resonance have libration times of $t_\mathrm{lib} \sim 2$ Gyr[136]. The relaxation time t_relax depends on what we consider to be driving the diffusion, i.e. whether it is two-body relaxation (§II C 1), transient spiral arms (§IV F 2), finite-N granularity noise (§VII), or whatever. In §II C 3, we estimated that the relaxation time of angular momenta in the Milky Way disk is $t_\mathrm{relax} \sim 60$ Gyr, which would give $\Delta \sim 2$. More generally, the estimate (144) tells us that Δ can be of order unity *even for systems with relaxation times much longer than the age of the Universe*. The important point is that it is much easier to diffuse across the width of the resonance than it is to diffuse across the whole of action space, and it is the diffusion across resonances that governs the dynamics[48,49,51].

We can solve equation (142) analytically in the limits $\Delta = 0$ and $\Delta \gg 1$, but it is easier to just solve it numerically for a given Δ. To do this, let us take as our initial condition a simple linear DF

$$f(\varphi, j, \tau = 0) = f_0(j = 0) + \alpha j \equiv f_\mathrm{init}(j), \tag{145}$$

where $\alpha \equiv \partial f_0/\partial j|_{j=0}$ is the gradient of the unperturbed DF at resonance. In Figure 30, we show with colored contours the resulting evolution of the dimensionless auxiliary DF

$$g(\varphi, j, \tau) \equiv \frac{f(\varphi, j, \tau) - f_0(j = 0)}{\alpha}, \tag{146}$$

whose initial value is $g_\mathrm{init}(j) \equiv g(\varphi, j, \tau = 0) = j$. Each row corresponds to a different value of Δ, and within each row, from left to right we plot the solution at different 'times' $\Delta^{1/3}\tau$.

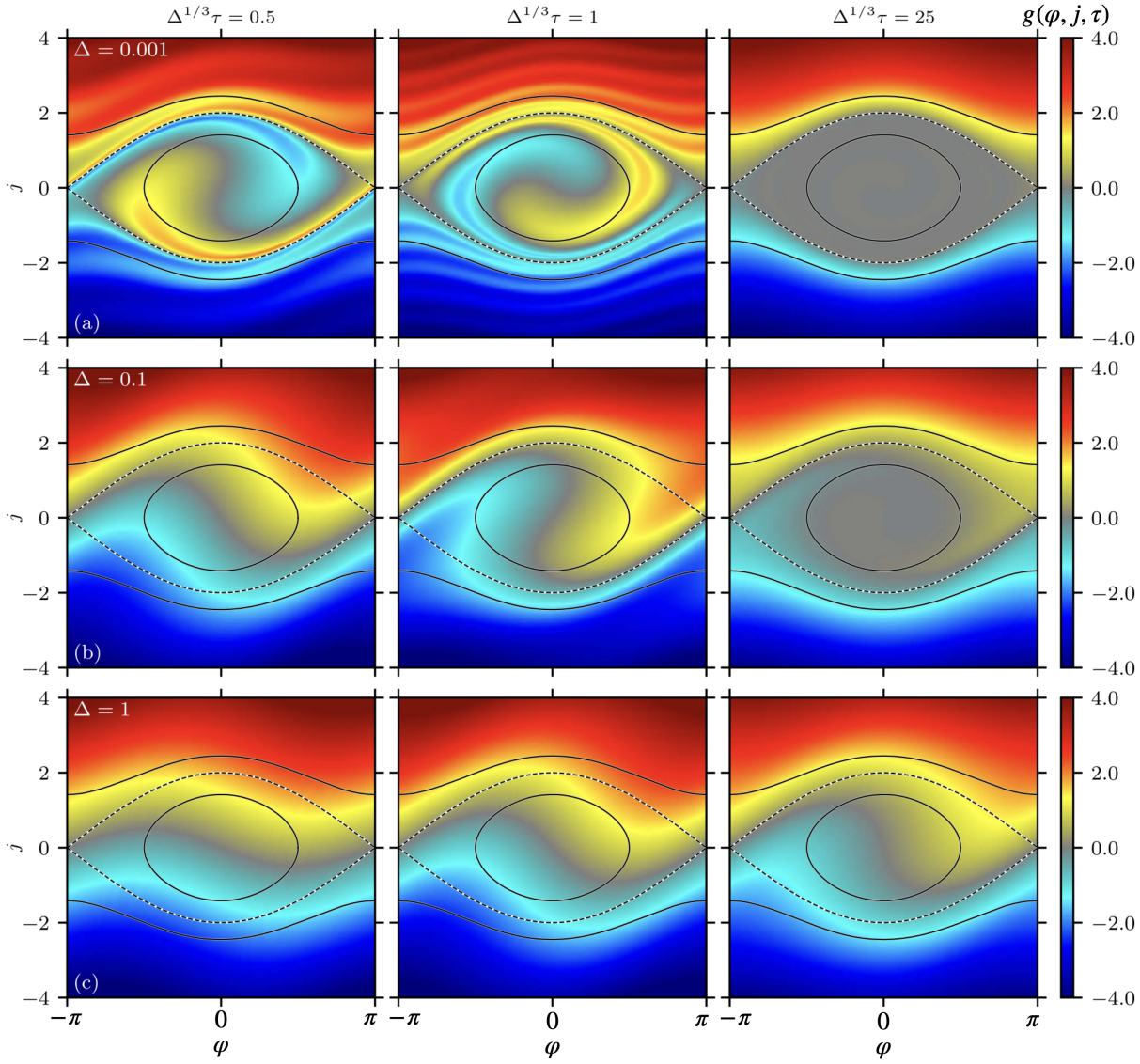

FIG. 30. Contours of $g(\varphi, j, \tau)$ (related to f through equation 146), computed by solving the kinetic equation (142) numerically with the initial condition $g(\tau = 0) = j$. Rows (a)–(c) show the time evolution of the solution for $\Delta = 0.001, 0.1, 1$ respectively. The right column represents the steady-state solution. Adapted from Hamilton *et al.*[140].

First consider row (a), which is in the limit of very weak diffusion ($\Delta = 0.001$). Physically, we expect that in this limit the DF will *phase mix* (c.f. Figure 18) around and within the resonant island. The reason for phase mixing is that in the absence of diffusion ($\Delta = 0$), the trajectories of individual stars trace contours of constant h in the (φ, j) plane. And, since adjacent contours correspond to slightly different libration/circulation periods, the initial DF gets sheared out along these contours, producing ever finer-scale structures in the phase space. If Δ were exactly zero, and in the absence of any coarse-graining, this process would continue indefinitely. However, for our very small but finite value of $\Delta = 0.001$, diffusion is able to wipe out the very smallest-scale features, without needing to invoke any coarse-graining. We see that by the third column, g is approximately zero everywhere inside the separatrix, and is smeared almost evenly on contours of constant h outside of the separatrix. The DF has reached a steady state; crucially, in this state the DF is *symmetric* around $\varphi = 0$ at any fixed j.

Next consider row (c) in Figure 30, which corresponds to rather strong (but not unrealistic! — see equation 144) diffusion, $\Delta = 1$. In this case, the resonance has a much less dramatic effect on g at any time. This again is as expected since the initial linear DF (145) is annihilated by the collision operator $C[f]$ (equation 139). In other words, wherever the bar perturbation induces some curvature in the DF, strong diffusion immediately tries to remove it and hence to restore the linear initial condition. Another way to say this is that the effect of diffusion is to counteract the

phase-mixing of orbits, in particular within the resonant island. Importantly, this means that in steady state the DF is strongly *asymmetric* in φ at a given j, in contrast to row (a). Finally, the case $\Delta = 0.1$ in row (b) is intermediate between these two regimes.

It is also instructive to average the solution $f(\varphi, j, \tau)$ over the slow angle φ and investigate the resulting DF

$$\langle f \rangle_\varphi \equiv \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{d}\varphi \, f(\varphi, j, \tau). \tag{147}$$

In Figure 31 we plot the corresponding auxiliary DF (see equation 146)

$$\langle g \rangle_\varphi = \frac{\langle f \rangle_\varphi - f_0(j=0)}{\alpha}, \tag{148}$$

as a function of j for various Δ values, and at different times (measured in units of the libration time $t_{\rm lib} = 2\pi/\omega_{\rm lib}$). The key point is that for $\Delta \to 0$, the phase mixing of orbits in and around the resonant island causes the angle-averaged

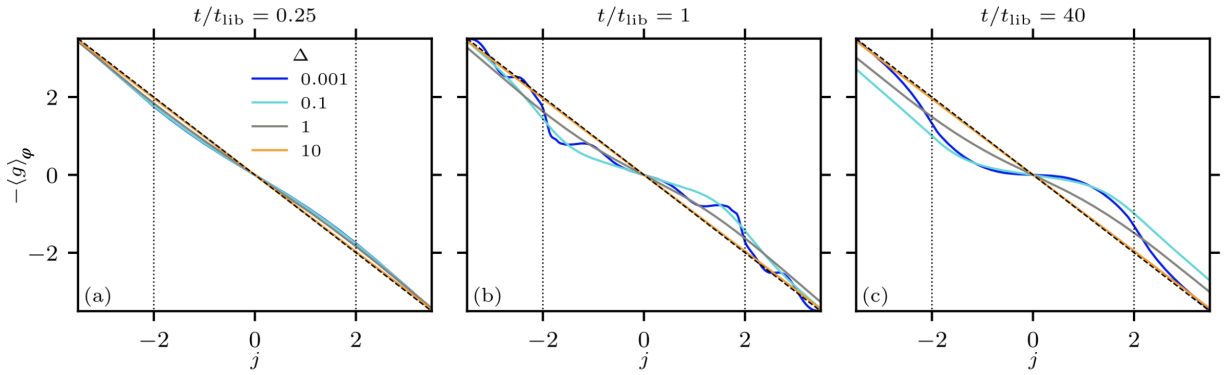

FIG. 31. Time evolution of the slow-angle-averaged DF, $\langle g \rangle_\varphi$ (see equation 148), for the same values of Δ as used in Figure 30. Different panels correspond to different times, and the right hand panel is the steady-state solution. The vertical dotted lines show the maximum extent of the separatrix, $j = \pm 2$ (see equation (141b)). Adapted from Hamilton *et al.* [140].

DF to be *flattened* there (see the blue line in panel c). Diffusion ($\Delta > 0$) suppresses this flattening effect in an attempt to preserve the initial linear profile.

We now show how these results can be applied to two real astrophysical problems.

1. *Example: Dynamical friction on a galactic bar*

As a galactic bar ages, it transfers angular momentum to its host galaxy and consequently its rotation rate slows. The mechanism responsible is dynamical friction: the bar produces a perturbation in the DF of the stars and dark matter particles that surround it (c.f. Figure 23), and that perturbation back-reacts to produce a torque on the bar, draining its angular momentum. Problems of this kind have been the focus of many classic studies in galactic dynamics[41,152], and are strongly analogous to wave-particle interaction problems in plasma (e.g. Landau damping). Here we illustrate the utility of our nonlinear pendulum kinetics, and the important role of the diffusion strength Δ, by calculating the dynamical friction torque on a galactic bar in a simple model.

We assume that center of the bar is at the origin and that its long axis sits in the plane of an initially axisymmetric disk of stars. We let the bar rotate around its center in the ϕ direction with pattern speed $\Omega_{\rm p}$, and consider its gravitational interaction with the stars. One can write down a rather general formula for the torque T felt by the bar as follows. Let the bar have potential $\delta\Phi(\mathbf{r}, t)$. Then from Hamilton's equation (32b) the specific torque felt by an individual star due to the bar is

$$\frac{\mathrm{d}L}{\mathrm{d}t} = -\frac{\partial \delta\Phi}{\partial \theta_\phi}. \tag{149}$$

Let the DF of stars be f (normalized such that $\int \mathrm{d}\theta_\phi \mathrm{d}\theta_R \mathrm{d}L \mathrm{d}J_R f$ is the total mass in stars). Then by Newton's third law, the total torque on the bar is equal to

$$T(t) = \int \mathrm{d}\boldsymbol{\theta} \mathrm{d}\mathbf{J} \, f(\boldsymbol{\theta}, \mathbf{J}, t) \frac{\partial \delta\Phi(\boldsymbol{\theta}, \mathbf{J}, t)}{\partial \theta_\phi}. \tag{150}$$

The challenge is to compute f for a given perturbation $\delta\Phi$, and then perform the integral (150).

The simplest way to calculate f is to resort to linear theory (and ignore diffusion), and use (94). In this case, putting $\delta\Phi_{\mathbf{N}}(\mathbf{J},t) = \delta\Phi_{\mathbf{N}}(\mathbf{J})\,\mathrm{e}^{-\mathrm{i}N_\phi\Omega_\mathrm{p}t}$ and performing the integral over t' in equation (94), one finds a 'linear torque', $T^\mathrm{lin} = \sum_{\mathbf{N}} T^\mathrm{lin}_{\mathbf{N}}$, where the contribution from resonance $\mathbf{N}$ is

$$T^\mathrm{lin}_{\mathbf{N}}(t) \equiv (2\pi)^2 N_\phi \int \mathrm{d}\mathbf{J}\,|\delta\Phi_{\mathbf{N}}(\mathbf{J})|^2 \mathbf{N}\cdot\frac{\partial f_0}{\partial \mathbf{J}}\frac{\sin[(\mathbf{N}\cdot\boldsymbol{\Omega}-N_\phi\Omega_\mathrm{p})t]}{\mathbf{N}\cdot\boldsymbol{\Omega}-N_\phi\Omega_\mathrm{p}}. \tag{151}$$

(We emphasize that in deriving this formula we have ignored collective effects, because we have not accounted for the contribution to $\delta\Phi$ that arises self-consistently from the perturbed stellar distribution. While collective effects are typically very important for galaxy evolution, in the present case it turns out that they do not significantly alter the character of the response[149], and we will continue to neglect them for simplicity). Taking the time-asymptotic limit $t\to\infty$, one arrives at the classic 'LBK torque formula'[152]

$$T^\mathrm{LBK}_{\mathbf{N}} \equiv (2\pi)^2 N_\phi \int \mathrm{d}\mathbf{J}|\delta\Phi_{\mathbf{N}}(\mathbf{J})|^2 \mathbf{N}\cdot\frac{\partial f_0}{\partial \mathbf{J}}\pi\delta\left(\mathbf{N}\cdot\boldsymbol{\Omega}-N_\phi\Omega_\mathrm{p}\right). \tag{152}$$

The LBK formula (152) predicts that angular momentum is transferred to and from the stars exclusively at resonances. It is also directly analogous to the classic formula for the Landau damping rate of a Langmuir wave in an electrostatic plasma[153,154] or Landau damping in more general geometry[84,155]. Importantly, for many realistic DFs $f_0(\mathbf{J})$, the LBK torque is finite and negative, implying a long-term transfer of angular momentum away from the bar and hence a decay in its pattern speed. (As an example, at corotation the sign of the LBK torque is the sign of $\partial f_0/\partial L$, which is typically negative whenever the density profile of the galaxy decreases with radius. This is similar to a plasma velocity distribution whose tail satisfies $\partial f_0/\partial v < 0$). In practice, the time-asymptotic limit may not be valid since the torque can take several dynamical timescales to converge, by which time the pattern speed may have changed significantly[156]. Nevertheless, the LBK formula (152) is a good benchmark against which we can compare the magnitude of the torque arising from more sophisticated calculations.

Of course, since we used the linearized solution (94) to compute f we did not account for the nonlinear orbit trapping that can occur sufficiently close to each resonance, nor did we include any finite Δ. We can include these effects by redoing the calculation using our pendulum formalism (§IV F). Converting to slow-fast angle-action variables $(\boldsymbol{\theta}',\mathbf{J}')$ *around each resonance* $\mathbf{N}$, Fourier expanding the potential $\delta\Phi$ in slow angle space as in equation (64), and expanding the DF as $f = \sum_{\mathbf{k}} f_{\mathbf{k}}(\mathbf{J}',t)\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\boldsymbol{\theta}'}$, one can show that the total 'nonlinear torque' on the bar $T^\mathrm{non\text{-}lin}$ can be written as a sum over $\mathbf{N}$ of contributions (see Hamilton et al.[140] for details[157]):

$$T^\mathrm{non\text{-}lin}_{\mathbf{N}}(t) = 2(2\pi)^2 N_\phi \sum_{k>0} k \int \mathrm{d}\mathbf{J}'\,\mathrm{Im}\left[f_k(\mathbf{J}',t)\Psi^*_k(\mathbf{J}')\right]. \tag{153}$$

Here we have used the shorthand $f_{(0,k)} = f_k$, and similarly for Ψ_k. In using expressions like (153), one must be careful to properly divide phase space into non-overlapping sub-volumes that are dominated by individual resonances $\mathbf{N}$ (see §4.5 of Chiba and Schönrich[42] for a discussion).

In reality, the torque on the bar involves a sum over all possible resonances $\mathbf{N}$ and an integral over all of action space, and each such contribution comes with a slightly different value of Δ. In order to isolate a representative quantity that only depends on a single value of Δ, we will focus henceforth on the *corotation torque density* $S(t)$, which is the contribution to the torque only arising from the corotation resonance $\Omega_\phi = \Omega_\mathrm{p}$ (so $N_\phi = 2$ and $k = 1$, see §IV F 1) and only considering stars with a particular J_R, namely

$$S(t|J_R) = 4(2\pi)^2 \int \mathrm{d}L\,\Psi_1\,\mathrm{Im}\,f_1. \tag{154}$$

For a given Δ we can compute $\mathrm{Im}\,f_1$ using the numerical solution to the kinetic equation (142) that we showed in Figure 30. We will compare this to the benchmark LBK value (see equation 152), which is the result of computing the torque density using linear theory with $\Delta = 0$:

$$S_\mathrm{LBK}(J_R) = 4(2\pi)^2 \int \mathrm{d}L\,|\Psi_1|^2\frac{\partial f_0}{\partial L}\pi\delta[2(\Omega_\phi-\Omega_\mathrm{p})]. \tag{155}$$

In Figure 32a, we show the bar's dimensionless slowing rate $S(t)/S_\mathrm{LBK}$ as a function of time for various values of $\Delta \geq 0$, fixing $J_R = 50$ kpc^2 Gyr^{-1} (which is much less than L_CR, meaning the stars are initially on roughly circular orbits — see Figure 9). Note that S_LBK is negative, so that a positive value of $S(t)/S_\mathrm{LBK}$ means the bar feels a

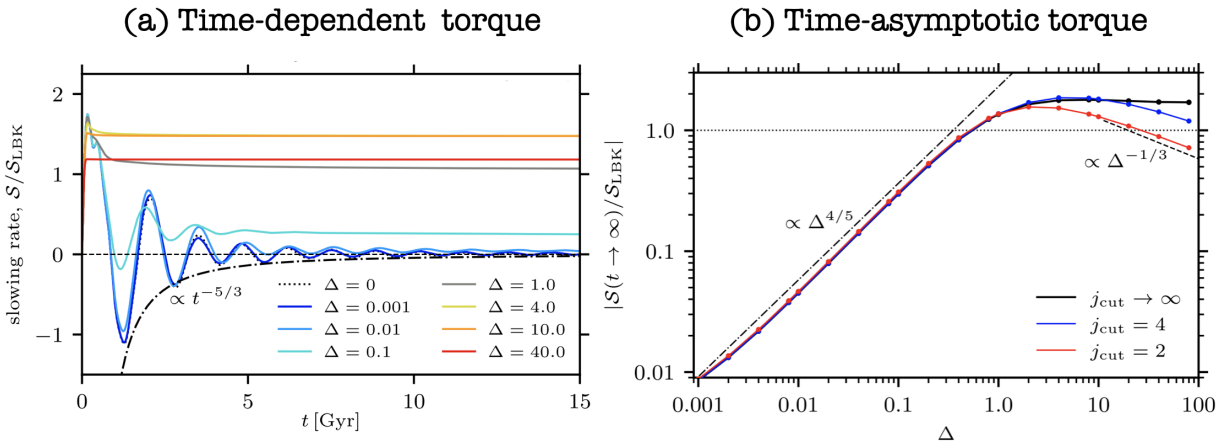

FIG. 32. Adapted from Hamilton *et al.*[140]. (a) Dimensionless torque on the bar (i.e. its 'slowing rate') $S(t)/S_{\rm LBK}$ due to stars at the corotation resonance, with $J_R = 50$ kpc^2 Gyr^{-1} (i.e. roughly circular orbits). The libration time is $t_{\rm lib} \approx 2$ Gyr. Colored lines show results for different $\Delta > 0$, while the dotted line shows the collisioness ($\Delta = 0$) result. (b) Steady-state value of $S(t)/S_{\rm LBK}$ as a function of Δ (note that both axes are on a logarithmic scale, and the different lines just correspond to different choices of integration limits).

negative torque (slowing it down). In the collisionless case ($\Delta = 0$, black dotted line) we see that the slowing rate oscillates on the timescale $\sim t_{\rm lib} \approx 2$ Gyr, and its envelope also decays over time. In the time-asymptotic limit for $\Delta = 0$ we find

$$S(t \to \infty) = 0, \qquad (\Delta = 0), \tag{156}$$

This zero torque result is a consequence of the angle-symmetry of the stellar density distribution ($\partial_\varphi f = 0$ at fixed j) that arises when the DF completely phase mixes within and around the resonant island (Figure 30a). Simply put, in the fully phase-mixed state there are the same number of stars 'pushing' on the bar as 'pulling' on it. This is the direct analogue of O'Neil[80]'s famous result in plasma theory: when nonlinear trapping is included (and diffusion neglected), the Landau damping rate of an electrostatic wave asymptotes to to zero, because of the symmetrization of the 'bouncing' particle distribution in the trough of the wave. However, as one deviates from $\Delta = 0$ to finite $\Delta > 0$, the behavior changes. The slowing rate in the steady state is manifestly positive for all $\Delta > 0$. This is a consequence of the fact that while phase-mixing attempts to abolish the asymmetry of the slow-angle distribution, diffusion replenishes it (e.g. Figure 30b) leading to a finite $\mathrm{Im}\, f_1$ in equation (154). Moreover, for $\Delta \gtrsim 1$, the slowing rate no longer oscillates on the libration timescale, because no trapped star is able to complete a full libration before being de-trapped by diffusion.

To make these statements more quantitative we turn to panel (b) of Figure 32, in which we show the time-asymptotic slowing rate $S(t \to \infty)$ as a function of Δ. We see that $S(t \to \infty)$ grows almost linearly with Δ until $\Delta \sim 1$, after which it saturates at a level similar to that given by the naive LBK theory, even though LBK theory ignores both diffusion and nonlinearity! This is as expected since, roughly speaking, for any $\Delta \gtrsim 1$ there is no nonlinear trapping and the effect of 'diffusion' is just to broaden the delta-function in (152). That is, *diffusion relinearizes the dynamics*. Again, this effect was also appreciated in the plasma theory of electric wave damping: if small-scale collisions are able to knock particles out of the wave trough before they can complete a full bounce orbit, then O'Neil's nonlinear trapping effects are interrupted, and the standard (quasi)-linear Landau damping/growth rate is recovered[48,49].

These results suggest that real galactic bars will always feel some non-zero torque, since finite diffusion will always replenish some asymmetry in the angular distribution of the stars as viewed in the bar frame. They also suggest that for systems with moderate to strong diffusion ($\Delta \gtrsim 1$), the LBK formula (152) provides a robust order-of-magnitude estimate of the frictional torque. Moreover, we have treated the bar's pattern speed, $\Omega_{\rm p}$, as constant here, but very similar arguments apply to the case where the 'interruption' of trapping is not due to collisions (i.e. finite Δ) but rather due to the fact that a changing $\Omega_{\rm p}$ causes resonance locations to sweep through phase space. In that case, the LBK formula holds in what Tremaine and Weinberg[41] called the 'fast regime', namely when the bar slows down rapidly enough that the resonances move by their own width $I_{\rm h}$ on a timescale shorter than $t_{\rm lib}$. A general theory that connects smoothly the fast (LBK or Landau-like) regime and the slow (nonlinear, O'Neil-like) regime was recently completed by Chiba[47].

2. *Example: Saturation of spiral instabilities*

Finally, let us discuss how pendulum kinetics can be used to understand the saturation mechanism of spiral instabilities and predict the resulting saturation amplitude.

We already saw (§VI B 2) that spiral waves can arise in stellar disks as exponentially growing, linear instabilities of the disk's underlying DF f_0, particularly in cases where the disk is sufficiently cold/dense (Figure 20) or grooved (Figure 21). Of course, this exponential growth phase cannot continue forever. Simulations by Sellwood and Carlberg [137] and theoretical work by Hamilton [158] suggests that the saturation occurs because of orbit trapping at the corotation resonance. To understand their argument, we present Figure 33 (adapted from Hamilton [158]).

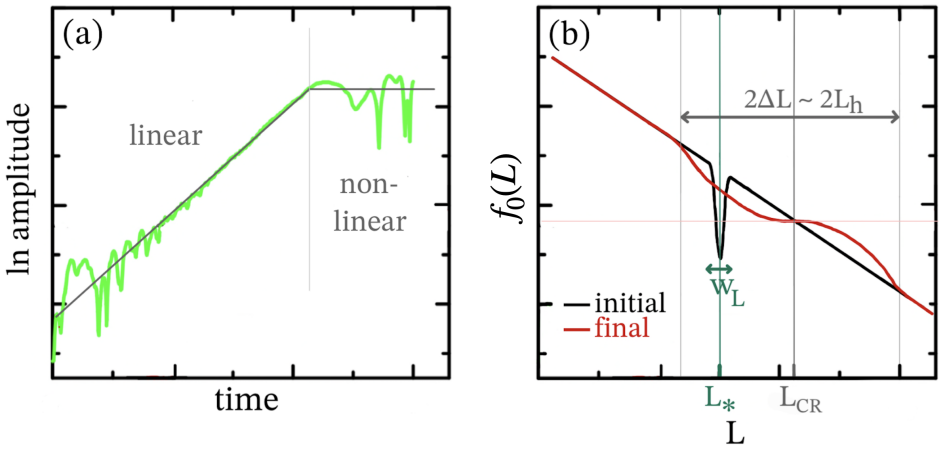

FIG. 33. Taken from Hamilton [158]. (a) Growth and saturation of a spiral mode in an N-body simulation of a razor thin stellar disk (adapted from Sellwood and Carlberg [137], green line). The gray line illustrates the simplification used in this paper: we separate the spiral's life into a linear phase of perfectly exponential growth, and a nonlinearly saturated phase of constant amplitude. (b) Illustration of the change to the axisymmetric DF $f_0(L,t)$ in the vicinity of a corotation resonance. The initial instability is driven by some sharp feature in the DF (here a *groove*, centered on $L = L_*$ with width w_L, see Sellwood and Carlberg [137]). In the saturated state, the DF is flattened around the resonance $L = L_\mathrm{CR}$.

The green line in panel (a) shows the logarithmic amplitude of a spiral density wave with constant pattern speed Ω_p, from an N-body simulation of an initially grooved, razor thin disk. Clearly, once some noisy transients have faded away the mode undergoes very well-defined exponential growth, confirming that it is a linear mode of the initial DF. Eventually, though, the exponential growth stops and the mode settles at a roughly constant amplitude. The gray line illustrates schematically the simplification we will use here, which is that the transition from the linear phase (exponentially growing amplitude) to the nonlinear phase (constant amplitude) happens instantaneously. In panel (b), we sketch the corresponding DF of angular momenta in the disk, $f_0(L)$, for circular orbits ($J_R = 0$). The initial unstable distribution is shown with a black line. The idea is that once the mode has grown to large enough amplitude that it starts to resonantly trap stellar orbits at corotation, the DF in the vicinity of that resonance will be flattened (as shown with the red line), just as we saw in §VIII B 1. This flattening erases the sharp feature in f_0 that originally drove the instability[159].

Hamilton [158] showed that by calculating the angular momentum content of the spiral in the linear (i.e. exponentially growing) and nonlinear (i.e. fully saturated) regimes, and equating the two expressions, one can read off an approximate formula for the saturation amplitude of the spiral potential. This is

$$|\delta\Phi_m| = \left(\frac{128}{9\pi^2}\right)^2 \frac{2}{m} \frac{\beta^2}{|\partial\Omega_\phi/\partial L|_{L_\mathrm{CR}}}, \tag{157}$$

where $\beta = \mathrm{Im}\omega_\mathrm{m}$ is the mode's growth rate. This result can be made more physically enlightening if we use (67) and (68) to express it as

$$\omega_\mathrm{lib} = \frac{256}{9\pi^2}\, m^{1/2}\beta \approx 2.88\, m^{1/2}\beta. \tag{158}$$

In other words, saturation occurs when the nonlinear libration frequency ω_lib is a few times larger than the linear growth rate β (so nonlinear orbit distortions occur before the mode amplitude can grow significantly). Note that

for one-armed spirals, equation (158) is identical to equation (12) of Dewar [160], who studied the saturation of *linear* momentum transfer between electrons and a single longitudinal plasma wave.

One can also relate (157) to the surface density perturbation of the spiral $\delta\Sigma_m$ relative to the background stellar disk Σ_0:

$$\left|\frac{\delta\Sigma_m}{\Sigma_0}\right| = \frac{32768}{81\pi^4}\,\xi^{-1}\cot p\left(\frac{\beta}{\Omega_\mathrm{p}}\right)^2 \tag{159}$$

$$\approx 0.5\times\left(\frac{\xi}{0.5}\right)^{-1}\left(\frac{\cot p}{1.5}\right)\left(\frac{\beta/\Omega_\mathrm{p}}{0.2}\right)^2, \tag{160}$$

where ξ is the active fraction of the disk (§VI B 2) and p is the spiral's pitch angle. Thus, spiral instabilities tend to saturate once their amplitudes reach several tens of percent of the background density.

Last of all, it is enlightening to estimate the timescale t_sat over which saturation of realistic spirals might be achieved. To do this, first recall (equation 158) that saturation occurs when $\omega_\mathrm{lib}(t_\mathrm{sat})\sim\beta$. Second, a basic scaling of equation (68) tells us that at corotation, $\omega_\mathrm{lib}(t)\sim\varepsilon(t)^{1/2}\Omega_\mathrm{p}$, where $\varepsilon(t)\sim|\Phi_m(t)/H_0|$ is the dimensionless amplitude of the mode. Third, in the linear regime we have $\varepsilon(t)=\varepsilon(0)\mathrm{e}^{\beta t}$. Putting these three ingredients together, we get a saturation time

$$t_\mathrm{sat}\sim\frac{2}{\beta}\left[\ln\left(\frac{\beta}{\Omega_\mathrm{p}}\right)+\frac{1}{2}\ln\left(\frac{1}{\varepsilon(0)}\right)\right]. \tag{161}$$

In particular, if the initial fluctuation level is set by Poisson noise alone then $\varepsilon(0)\sim N^{-1/2}$, and we have (very roughly)

$$t_\mathrm{sat}\sim\frac{1}{2\beta}\ln N \tag{162}$$

$$\sim 2.5\,\mathrm{Gyr}\times\left(\frac{T_\mathrm{p}}{250\,\mathrm{Myr}}\right)\left(\frac{\beta/\Omega_\mathrm{p}}{0.2}\right)^{-1}\left(\frac{\ln N}{25}\right), \tag{163}$$

where $T_\mathrm{p}=2\pi/\Omega_\mathrm{p}$ is the orbital period at corotation. Thus, if it has to exponentiate out of Poisson noise, a rather vigorously growing spiral with $\beta/\Omega_\mathrm{p}\gtrsim 0.1$ will saturate after a few Gyr. More realistic galactic noise tends to be of higher amplitude than is implied by Poisson statistics (§VII C 3), though this will only change the timescale t_sat by a logarithmic factor.

IX. FURTHER READING

There are many more connections between plasma theory and stellar dynamics that we have not touched on in this article. Let us mention just a few of them here, alongside references for further reading.

First, the most basic connection between the two fields is the theory of orbits in spatially smooth and constant or time-varying potentials, the regularity/chaoticity of those orbits, and the associated theories of canonical transformations, Hamiltonian perturbation theory, and guiding center motion. These are of course classic problems in the general theory of dynamics[34], and Hamiltonian perturbation theory in angle-action variables has long been a tool of theorists (see Chapter 3 of BT08), but it is only relatively recently that angle-action variables have been adopted across mainstream galactic astronomy[161]. They are now ubiquitous tools with which astronomers study phase space data of stellar systems, particularly from the GAIA satellite[38]. Chaotic orbits are known to exist in triaxial (i.e. non-axisymmetric) galaxies and star clusters (see, e.g., the books by Contopoulos [162], Merritt [163]), but are mostly ignored in kinetic theories, which usually focus on the evolution of near-axisymmetric or near-spherical systems. Guiding center and gyrokinetic theory have close analogues in stellar dynamics, particularly for describing stellar motion near supermassive black holes, where the orbits are closed Keplerian ellipses, which can be averaged over to form Keplerian 'rings' or 'wires'[108,164]. Despite this, the Hamiltonian theories of guiding center motion, averaged Lagrangians, ponderomotive effects and so on developed by the likes of Cary and Brizard [165], Dodin and Fisch [166] in plasma theory have not (yet) received much attention in galactic dynamics.

Second, another common thread between plasma kinetics and stellar dynamics lies in the construction of collisionless equilibria and their stability properties. Though we have barely touched this subject here, developing stable candidate distribution functions is a crucial part of stellar-dynamical modeling, completely analogous to the construction of Vlasov–Poisson or Vlasov–Maxwell[167] equilibria in plasma. Similarly, Nyquist diagrams[168] and Penrose's criterion[169] are just as applicable to stellar systems as they are to plasmas. On the gravitational side, perhaps the best introductions to this subject (written in a language plasma physicists can appreciate) are to be found in §§4-5 of

BT08 and the 1994 book by Palmer [170]. The theory of *nonlinear* stability, formation of BGK-like modes[171], and so on have been developed to some extent in stellar dynamics[172–174], usually with direct reference to previous plasma results.

Third and finally, we mention a major area of stellar-dynamical kinetics that we have only touched on here, the theory of spiral structure. This was one of the original areas in which the galactic-plasma overlap was found to be most fruitful[63,65,98,152,175], since it concerns problems familiar to plasma practitioners such as wave propagation, WKB theory, wave-particle interaction, phase mixing, and so on. Despite this, the theory is far from complete: opinions still differ over the mechanisms that provoke spiral responses, the lifetimes of individual spiral patterns, the importance of gas flows and star formation, and so on — for an array of perspectives, see e.g. reviews by Sellwood and Masters [78], Athanassoula [176], Dobbs and Baba [177].

We hope that our plasma colleagues consider this tutorial article an open invitation to participate in a subject that is still in its theoretical adolescence; in which simulations with realistic numbers of particles are routine[178]; in which observations are astonishing and improving all the time; and, most of all, in which the Universe constantly impresses upon us its complex, kaleidoscopic beauty.

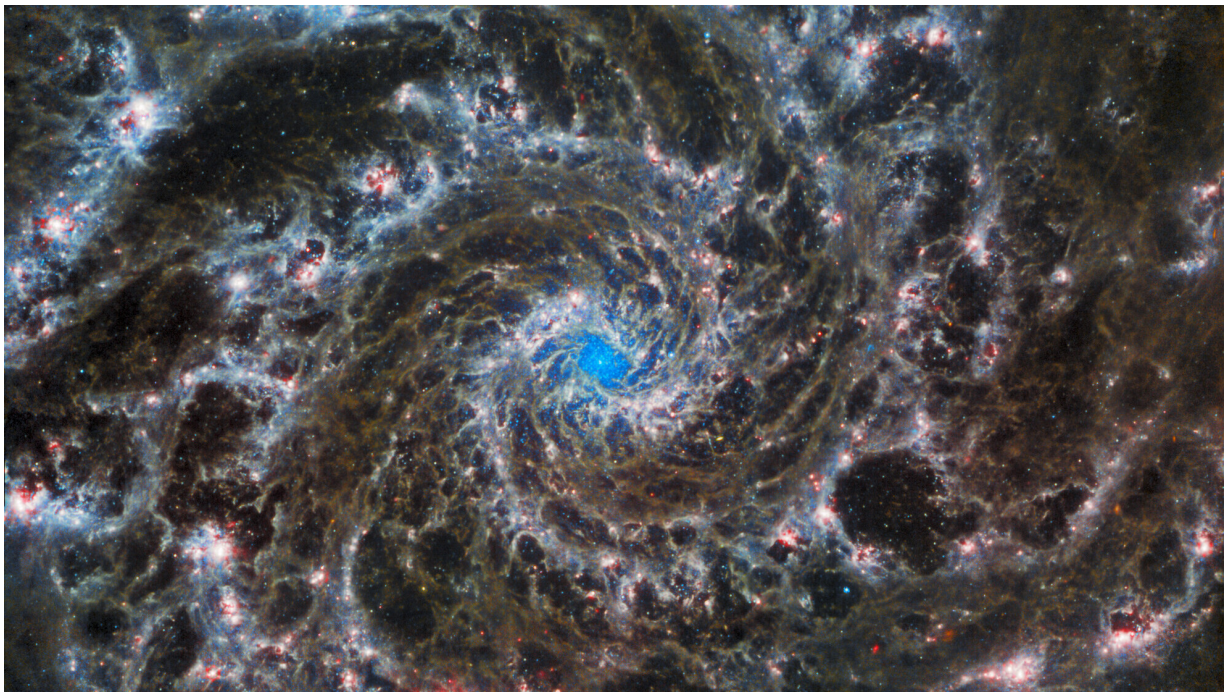

FIG. 34. The Phantom Galaxy. Credit: ESA/Webb, NASA & CSA, J. Lee and the PHANGS-JWST Team.

ACKNOWLEDGMENTS

This article grew out of the notes that we each compiled as lecturers of the Kinetic Theory of Stellar Systems MMathPhys course at the University of Oxford. We thank Prof. James Binney, who designed the course and was its first lecturer, as well as Prof. Paul Dellar and Prof. Alexander Schekochihin who taught the sister courses on kinetic theory of gases and plasmas[179] respectively, and we gratefully acknowledge the hospitality of Merton College, Oxford. We also thank Shaunak Modak and Christophe Pichon for comments on the manuscript. C. H. is supported by the John N. Bahcall Fellowship Fund at the Institute for Advanced Study.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Appendix A: The epicyclic approximation

In this Appendix, we derive the angle-action variables for near-circular orbits in a central potential $\Phi(R)$, using the so-called *epicyclic approximation*. To do this, let us first write down the Lagrangian for motion in this potential

$$\mathcal{L}(\phi, R, \dot{\phi}, \dot{R}) = \frac{1}{2}\dot{R}^2 - \Phi_{\rm eff}(R), \tag{A1}$$

where $\Phi_{\rm eff}(R) \equiv \Phi(R) - L^2/2R^2$ is the effective potential, and $L = R^2\dot{\phi}$ is a constant. The Euler–Lagrange equation, $\partial\mathcal{L}/\partial R = {\rm d}(\partial\mathcal{L}/\partial\dot{R})/{\rm d}t$, then gives

$$\frac{{\rm d}^2 R}{{\rm d}t^2} = -\frac{\partial\Phi_{\rm eff}(R)}{\partial R}. \tag{A2}$$

Of course, perfectly circular motion corresponds to the right hand side of this equation being set to zero, i.e. a minimum of the effective potential. To work out the lowest-order correction to this circular motion, we can expand $\Phi_{\rm eff}(R)$ around $R_{\rm g}$ and find

$$\frac{{\rm d}^2(R - R_{\rm g})}{{\rm d}t^2} = -\Omega_R^2(R - R_{\rm g}) + ... \tag{A3}$$

where $\Omega_R^2 = (\partial^2\Phi_{\rm eff}/\partial R^2)|_{R_{\rm g}}$. This just corresponds to a harmonic oscillator with angular frequency Ω_R. This radial motion may be alternatively described using a pair of radial angle-action coordinates (θ_R, J_R) satisfying

$$J_R = \frac{\dot{R}^2 + \Omega_R^2(R - R_{\rm g})^2}{2\Omega_R} \equiv \frac{E_R}{\Omega_R}, \tag{A4a}$$

$$R = R_{\rm g} + a_R\cos\theta_R, \tag{A4b}$$

where the amplitude of the radial epicycle is

$$a_R \equiv \frac{\sqrt{2E_R}}{\Omega_R}. \tag{A5}$$

Similarly, the azimuthal part of the problem can also be described by angle-action variables:

$$L = R_{\rm g}^2\Omega_\phi, \tag{A6a}$$

$$\phi = \theta_\phi + \frac{2\Omega_\phi a_R}{\Omega_R R_{\rm g}}\sin\theta_R, \tag{A6b}$$

and the frequencies can be written in the following form in terms of Φ, all evaluated at $R = R_{\rm g}$:

$$\Omega_\phi^2 = \frac{1}{R}\frac{\partial\Phi}{\partial R} \equiv \Omega_{\rm c}^2, \tag{A7a}$$

$$\Omega_R^2 = \frac{2\Omega_\phi}{R}\frac{\partial(\Omega_\phi R)}{\partial R} \equiv \kappa^2. \tag{A7b}$$

One can also extend the epicyclic approximation to include vertical harmonic oscillations or higher order corrections, but this will not concern us here. See §3.2.3 of BT08 and Dehnen[180] for more details.

Appendix B: The basis method

To solve equation (114), we assume that we have at our disposal a complete set of *biorthogonal basis elements* $(\Phi^{(p)}(\mathbf{r}), \rho^{(p)}(\mathbf{r}))$, which satisfy

$$\Phi^{(p)}(\mathbf{r}) = \int \mathrm{d}\mathbf{r}'\, \psi(\mathbf{r},\mathbf{r}')\, \rho^{(p)}(\mathbf{r}'), \tag{B1a}$$

$$\int \mathrm{d}\mathbf{r}\, [\Phi^{(p)}(\mathbf{r})]^*\, \rho^{(q)}(\mathbf{r}) = -\mathcal{E}\delta_p^q, \tag{B1b}$$

where $\psi(\mathbf{r},\mathbf{r}') = -G/|\mathbf{r}-\mathbf{r}'|$ and $\mathcal{E}$ is an arbitrary constant with units of energy. The 'completeness' requirement is that we can expand an arbitrary function $\Psi(\mathbf{r}) = \sum_p A_p \Phi^{(p)}(\mathbf{r})$ for some set of constant coefficients A_p. Various such basis sets do exist, and the particular set you should choose depends on the geometry of the system you are interested in. For instance, in a spherical geometry one might take the basis to consist of products of radial Bessel functions and angular spherical harmonics (see, e.g., Kalnajs [65], Petersen *et al.* [76] for more details). The idea is to project (114) onto this basis and then read off the coefficients to get an explicit expression for ψ^{d}.

Thus, let us begin by expanding the pairwise interaction potential $\psi(\mathbf{r},\mathbf{r}') = -G/|\mathbf{r}-\mathbf{r}'|$ for a fixed value of $\mathbf{r}'$ as

$$\psi(\mathbf{r},\mathbf{r}') = \sum_p u_p(\mathbf{r}')\, \Phi^{(p)}(\mathbf{r}), \tag{B2}$$

where the coefficient $u_p(\mathbf{r}')$ has units of $(\text{mass})^{-1}$. To find this coefficient, we multiply both sides of this equation by $\rho^{(q)*}(\mathbf{r})$, integrate over $\mathbf{r}$, and use both properties (B1). The result is

$$\begin{aligned} u_p(\mathbf{r}') &= -\frac{1}{\mathcal{E}} \int \mathrm{d}\mathbf{r}\, \psi(\mathbf{r},\mathbf{r}')\, \rho^{(p)*}(\mathbf{r}) \\ &= -\frac{1}{\mathcal{E}} \Phi^{(p)*}(\mathbf{r}'). \end{aligned} \tag{B3}$$

Plugging this back in to equation (B2) we get

$$\psi(\mathbf{r},\mathbf{r}') = -\frac{1}{\mathcal{E}} \sum_p \Phi^{(p)}(\mathbf{r})\, \Phi^{(p)*}(\mathbf{r}'). \tag{B4}$$

The fact that the bare interaction can be expanded in this 'quasi-separable' form is one of the main advantages of the basis method.

Of course, what we really need is the Fourier transform of this function:

$$\psi_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}') = -\frac{1}{\mathcal{E}} \sum_p \Phi^{(p)}_{\mathbf{n}}(\mathbf{J})\, \Phi^{(p)*}_{\mathbf{n}'}(\mathbf{J}'), \tag{B5}$$

where $\Phi^{(p)}_{\mathbf{n}}(\mathbf{J})$ stands for the Fourier transformed basis elements naturally defined following equation (36). Thus we now have an explicit expression for the functions $\psi_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}')$ which enter the right hand side of equation (99).

We now perform a very similar expansion of the dressed interaction ψ^{d}. By analogy with (B5), we assume that, at fixed ω, this function can be expanded as

$$\psi^{\mathrm{d}}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\omega) = -\frac{1}{\mathcal{E}} \sum_{p,q} \Phi^{(p)}_{\mathbf{n}}(\mathbf{J})\, \mathsf{E}^{-1}_{pq}(\omega)\, \Phi^{(q)*}_{\mathbf{n}'}(\mathbf{J}'), \tag{B6}$$

where $\mathsf{E}(\omega)$ is a dimensionless, as-yet-unknown, matrix which depends only on ω (and we have assumed its inverse exists). Injecting this decomposition into equation (114), we obtain

$$\sum_{p,q} \Phi^{(p)}_{\mathbf{n}}(\mathbf{J})\, \mathsf{E}^{-1}_{pq}(\omega)\, \Phi^{(q)*}_{\mathbf{n}'}(\mathbf{J}') = \sum_{p,q} \Phi^{(p)}_{\mathbf{n}}(\mathbf{J}) \left[\mathsf{M}(\omega)\, \mathsf{E}^{-1}(\omega)\right]_{pq} \Phi^{(q)*}_{\mathbf{n}'}(\mathbf{J}') + \sum_{p,q} \Phi^{(p)}_{\mathbf{n}}(\mathbf{J})\, \mathsf{I}_{pq}\, \Phi^{(q)*}_{\mathbf{n}'}(\mathbf{J}'), \tag{B7}$$

where we introduced the identity matrix I, as well as the system's dimensionless *response matrix* $\mathsf{M}(\omega)$ as in equation (101). Since the basis is biorthogonal, we can identify the elements in equation (B7), and we obtain the matrix relation

$$\mathsf{E}^{-1}(\omega) = \mathsf{M}(\omega)\, \mathsf{E}^{-1}(\omega) + \mathsf{I}, \tag{B8}$$

which immediately tells us that

$$\mathsf{E}(\omega) = \mathsf{I} - \mathsf{M}(\omega). \tag{B9}$$

To conclude, from (B6) we have

$$\psi^{\rm d}_{\mathbf{n}\mathbf{n}'}(\mathbf{J},\mathbf{J}',\omega) = -\frac{1}{\mathcal{E}} \sum_{p,q} \Phi^{(p)}_{\mathbf{n}}(\mathbf{J}) \left[\mathsf{I} - \mathsf{M}(\omega)\right]^{-1}_{pq} \Phi^{(q)*}_{\mathbf{n}'}(\mathbf{J}'). \tag{B10}$$

Thus we have done what we anticipated in §VI A, which is essentially to invert an operator $\mathsf{I} - \mathsf{M}$, with M encoding collective effects. We can easily recover the bare interaction, equation (B5), by ignoring collective effects, i.e. by setting $\mathsf{M}(\omega) = \mathsf{0}$ in (B10).

[1] J. Binney and M. Merrifield, *Galactic astronomy* (Princeton Univ. Press, 1998).
[2] J. Falcón-Barroso and J. H. Knapen, *Secular evolution of galaxies* (Cambridge University Press, 2012).
[3] H. Mo, F. Van den Bosch, and S. White, *Galaxy formation and evolution* (Cambridge Univ. Press, 2010).
[4] J. Binney and S. Tremaine, *Galactic Dynamics: Second Edition* (Princeton Univ. Press, 2008).
[5] T. Alexander, ARA&A **55**, 17 (2017).
[6] D. Lynden-Bell, MNRAS **136**, 101 (1967).
[7] R. J. Ewart, A. Brown, T. Adkins, and A. A. Schekochihin, JPP **88**, 925880501 (2022).
[8] A. Campa, T. Dauxois, D. Fanelli, and S. Ruffo, *Physics of long-range interacting systems* (OUP Oxford, 2014).
[9] J. G. Hills and C. A. Day, Astrophys. Lett. **17**, 87 (1976).
[10] S. Chandrasekhar, ApJ **97**, 255 (1943).
[11] E. M. Lifshitz and L. P. Pitaevskii, *Physical Kinetics* (Elsevier, 1981).
[12] D. C. Heggie and P. Hut, *The gravitational million-body problem* (Cambridge Univ. Press, 2003).
[13] Actually, the argument we have made here does not really apply to the geometry of a thin disk. Accounting for the disky geometry, the estimate on the right hand side of (19) is reduced by a factor like $\sim 10^4$ (see equation 5 of Sellwood [181] and the surrounding discussion). Nevertheless, even this shortened two-body relaxation time is much, much longer than the age of the Universe.
[14] K. J. Joshi, F. A. Rasio, and S. Portegies Zwart, ApJ **540**, 969 (2000).
[15] M. Hénon, ApSS **14**, 151 (1971).
[16] N. Frankel, J. Sanders, Y.-S. Ting, and H.-W. Rix, ApJ **896**, 15 (2020).
[17] M. Hénon, in *Dynamics of the Solar Systems*, Vol. 69, edited by A. Hayli (1975) p. 133.
[18] M. Giersz and D. C. Heggie, MNRAS **268**, 257 (1994).
[19] J.-B. Fouvry, C. Hamilton, S. Rozier, and C. Pichon, MNRAS **508**, 2210 (2021).
[20] D. Lynden-Bell and R. Wood, MNRAS **138**, 495 (1968).
[21] A. H. Boozer, RMP **76**, 1071 (2004).
[22] F. Zonca, L. Chen, S. Briguglio, G. Fogaccia, G. Vlad, and X. Wang, New J. Phys. **17**, 013052 (2015).
[23] L. Chen and F. Zonca, RMP **88**, 015008 (2016).
[24] J. Bovy, ApJS **216**, 29 (2015).
[25] A. M. Price-Whelan, JOSS **2** (2017), 10.21105/joss.00388.
[26] E. Vasiliev, MNRAS **482**, 1525 (2019).
[27] J. Binney and D. Spergel, ApJ **252**, 308 (1982).
[28] J. Binney and D. Spergel, MNRAS **206**, 159 (1984).
[29] M. Hénon and C. Heiles, AJ **69**, 73 (1964).
[30] H. Goldstein, *Classical mechanics* (Addison-Wesley, 1950).
[31] M. D. Weinberg, AJ **108**, 1398 (1994).
[32] M. D. Weinberg, AJ **108**, 1403 (1994).
[33] M. D. Weinberg, AJ **108**, 1414 (1994).
[34] A. J. Lichtenberg and M. A. Lieberman, *Regular and chaotic dynamics*, Vol. 38 (Springer, 2013).
[35] V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, 1989).
[36] J. A. Bittencourt, *Fundamentals of plasma physics* (Springer, 2013).
[37] J. A. S. Hunt, M. W. Bub, J. Bovy, J. T. Mackereth, W. H. Trick, and D. Kawata, MNRAS **490**, 1026 (2019).
[38] W. H. Trick, J. Coronado, and H.-W. Rix, MNRAS **484**, 3291 (2019).
[39] J. L. Sanders and J. Binney, MNRAS **457**, 2107 (2016).
[40] J. Binney, MNRAS **495**, 895 (2020).
[41] S. Tremaine and M. D. Weinberg, MNRAS **209**, 729 (1984).
[42] R. Chiba and R. Schönrich, MNRAS **513**, 768 (2022).
[43] J. A. Sellwood and J. Binney, MNRAS **336**, 785 (2002).
[44] J. M. Petit and M. Henon, Icarus **66**, 536 (1986).
[45] M. Hasegawa and K. Nakazawa, A&A **227**, 619 (1990).
[46] P. A. Sturrock, Phys. Rev. **141**, 186 (1966).
[47] R. Chiba, MNRAS **525**, 3576 (2023).
[48] G. L. Johnston, Phys. Fluids **14**, 2719 (1971).
[49] S. P. Auerbach, Phys. Fluids **20**, 1836 (1977).
[50] V. N. Duarte, N. N. Gorelenkov, R. B. White, and H. L. Berk, PoP **26**, 120701 (2019).
[51] E. A. Tolman and P. J. Catto, JPP **87**, 855870201 (2021).
[52] Equation (86) is usually called the Collisionless Boltzmann Equation in stellar dynamics, or the Vlasov equation in plasma kinetics. But in research papers one is equally likely to see either of those names given to the Klimontovich equation (81), or to something else entirely, so caution is required[182].

[53]C. Pichon and D. Aubert, MNRAS **368**, 1657 (2006).
[54]B. B. Kadomtsev, *Plasma turbulence* (New York: Academic Press, 1965).
[55]T. H. Dupree, Phys. Fluids **9**, 1773 (1966).
[56]J. A. Krommes, Phys. Rep. **360**, 1 (2002).
[57]P. H. Diamond, S.-I. Itoh, and K. Itoh, *Modern Plasma Physics* (Cambridge Univ. Press, Cambridge, 2010).
[58]J. A. Krommes, JPP **81**, 205810601 (2015).
[59]R. W. Gould, T. M. O'Neil, and J. H. Malmberg, PRL **19**, 219 (1967).
[60]L. Al'tshul and V. Karpman, JETP **20**, 1043 (1965).
[61]A. J. Layzer, Phys. Rev. **129**, 897 (1963).
[62]R. Bălescu, J. Math. Phys. **4**, 1009 (1963).
[63]A. J. Kalnajs, ApJ **166**, 275 (1971).
[64]T. Kato, *Perturbation theory for linear operators* (Springer, 2013).
[65]A. J. Kalnajs, ApJ **205**, 745 (1976).
[66]J. Barré, A. Olivetti, and Y. Y. Yamaguchi, J. Phys. A **44**, 405502 (2011).
[67]J. Barré and Y. Y Yamaguchi, J. Phys. A **46**, 225501 (2013).
[68]D. B. Melrose, *Instabilities in Space and Laboratory Plasmas* (Cambridge University Press, 1986).
[69]J. Y. Kim, Y. Kishimoto, W. Horton, and T. Tajima, PoP **1**, 927 (1994).
[70]T. Kuroda, H. Sugama, R. Kanno, M. Okamoto, and W. Horton, J. Phys. Soc. Japan **67**, 3787 (1998).
[71]S. D. Baalrud, PoP **20**, 012118 (2013).
[72]J.-B. Fouvry and S. Prunet, MNRAS **509**, 2443 (2021).
[73]Note that in this article we use the nomenclature 'Landau mode' regardless of whether the growth rate of the mode is positive, negative or zero.
[74]B. D. Fried and S. D. Conte, *The Plasma Dispersion Function* (Academic Press, New-York, 1961).
[75]T. A. Zang, *The stability of a model galaxy*, Ph.D. thesis, MIT (1976).
[76]M. S. Petersen, M. Roule, J.-B. Fouvry, C. Pichon, and K. Tep, arXiv, 2311.10630 (2023).
[77]J. Kormendy and J. Kennicutt, Robert C., ARA&A **42**, 603 (2004).
[78]J. A. Sellwood and K. L. Masters, ARA&A **60**, 73 (2022).
[79]S. De Rijcke, J.-B. Fouvry, and C. Pichon, MNRAS **484**, 3198 (2019).
[80]T. O'Neil, Phys. Fluids **8**, 2255 (1965).
[81]H. L. Berk and B. N. Breizman, Phys. Fluids **2**, 2235 (1990).
[82]H. L. Berk and B. N. Breizman, *Enhancement of particle-wave energy exchange by resonance sweeping* (Technical report, DOE/ET/53088-722-Rev., 1996).
[83]W. Drummond and D. Pines, *Non-linear stability of plasma oscillation* (Technical report, General Atomic Div., 1961).
[84]A. N. Kaufman, Phys. Fluids **15**, 1063 (1972).
[85]M. D. Weinberg, ApJ **421**, 481 (1994).
[86]D. C. Heggie, P. G. Breen, and A. L. Varri, MNRAS **492**, 6019 (2020).
[87]P. Gibbon, arXiv, 2007.04783 (2020).
[88]J. Magorrian, MNRAS **507**, 4840 (2021).
[89]E. D'Onghia, M. Vogelsberger, and L. Hernquist, ApJ **766**, 34 (2013).
[90]J. Binney and C. Lacey, MNRAS **230**, 597 (1988).
[91]J.-B. Fouvry and B. Bar-Or, MNRAS **481**, 4566 (2018).
[92]P.-H. Chavanis, Physica A **391**, 3680 (2012).
[93]P.-H. Chavanis, A&A **556**, A93 (2013).
[94]R. Bălescu, Phys. Fluids **3**, 52 (1960).
[95]A. Lenard, Ann. Phys. **10**, 390 (1960).
[96]A. N. Kaufman and T. Nakayama, Phys. Fluids **13**, 956 (1970).
[97]A. N. Kaufman, Phys. Fluids **14**, 387 (1971).
[98]E. Dekker, Phys. Rep. **24**, 315 (1976).
[99]P.-H. Chavanis, Eur. Phys. J. Plus **139**, 51 (2024).
[100]J. F. Luciani and R. Pellat, ApJ **317**, 241 (1987).
[101]J. Heyvaerts, MNRAS **407**, 355 (2010).
[102]J. Heyvaerts, J.-B. Fouvry, P.-H. Chavanis, and C. Pichon, MNRAS **469**, 4193 (2017).
[103]C. Hamilton, MNRAS **501**, 3371 (2021).
[104]N. Rostoker, Phys. Fluids **7**, 479 (1964).
[105]N. Rostoker, Phys. Fluids **7**, 491 (1964).
[106]The factor of 1/2 accounts for the fact that, when we integrate over all possible $\mathbf{J}'$, $\Delta\mathbf{J}$ and $\Delta\mathbf{J}'$, the contribution from each diagram in Figure 25 accounts for all possible interactions, and both diagrams are equivalent. The 'accounting trick' of using both diagrams and then halving the answer leads very simply to the Balescu–Lenard equation.
[107]F. P. C. Benetti and B. Marcos, PRE **95**, 022111 (2017).
[108]B. Bar-Or and J.-B. Fouvry, ApJL **860**, L23 (2018).
[109]J.-B. Fouvry, B. Bar-Or, and P.-H. Chavanis, PRE **99**, 032101 (2019).
[110]M. Roule, J.-B. Fouvry, C. Pichon, and P.-H. Chavanis, PRE **106**, 044118 (2022).
[111]M. D. Weinberg, MNRAS **239**, 549 (1989).
[112]M. D. Weinberg, ApJ **300**, 93 (1986).
[113]J. A. Sellwood, ApJ **751**, 44 (2012).
[114]J.-B. Fouvry, C. Pichon, J. Magorrian, and P.-H. Chavanis, A&A **584**, A129 (2015).
[115]A. Rogister and C. Oberman, JPP **2**, 33 (1968).
[116]A. Rogister and C. Oberman, JPP **3**, 119 (1969).
[117]C. Oberman, "Advanced kinetic theory," in *Physics of Hot Plasmas* (Springer, 1970) pp. 42–102.
[118]T. Hatori, Phys. Fluids **11**, 1584 (1968).

[119]T. Hatori, Phys. Fluids **12**, 1652 (1969).
[120]C. Hamilton and T. Heinemann, MNRAS **525**, 4161 (2023).
[121]C. Hamilton and T. Heinemann, arXiv, 2011.14812 (2020).
[122]M. Tagger, J.-F. Sygnet, E. Athanassoula, and R. Pellat, ApJL **318**, L43 (1987).
[123]J.-F. Sygnet, M. Tagger, E. Athanassoula, and R. Pellat, MNRAS **232**, 733 (1988).
[124]A. Toomre, ApJ **158**, 899 (1969).
[125]A. Toomre, ARA&A **15**, 437 (1977).
[126]J. W. K. Mark, ApJ **212**, 645 (1977).
[127]C. Norman, MNRAS **182**, 457 (1978).
[128]D. ter Haar, Phys. Scr. **20**, 291 (1979).
[129]V. E. Zakharov, Zh. Eksp. Teor. Fiz. **35**, 908 (1972).
[130]B. Kocsis and S. Tremaine, MNRAS **448**, 3265 (2015).
[131]J.-B. Fouvry, B. Bar-Or, and P.-H. Chavanis, ApJ **883**, 161 (2019).
[132]J. I. Read, T. Goerdt, B. Moore, A. P. Pontzen, J. Stadel, and G. Lake, MNRAS **373**, 1451 (2006).
[133]K. Kaur and N. C. Stone, MNRAS **515**, 407 (2022).
[134]J. B. Taylor and B. McNamara, Phys. Fluids **14**, 1492 (1971).
[135]R. Chiba, J. K. S. Friske, and R. Schönrich, MNRAS **500**, 4710 (2020).
[136]R. Chiba and R. Schönrich, MNRAS **505**, 2412 (2021).
[137]J. A. Sellwood and R. G. Carlberg, MNRAS **517**, 2610 (2022).
[138]J. A. Sellwood and F. D. Kahn, MNRAS **250**, 278 (1991).
[139]J. A. Sellwood and R. G. Carlberg, MNRAS **489**, 116 (2019).
[140]C. Hamilton, E. A. Tolman, L. Arzamasskiy, and V. N. Duarte, ApJ **954**, 12 (2023).
[141]A. Jenkins and J. Binney, MNRAS **245**, 305 (1990).
[142]R. S. De Simone, X. Wu, and S. Tremaine, MNRAS **350**, 627 (2004).
[143]J. Peñarrubia, MNRAS **490**, 1044 (2019).
[144]A. A. El-Zant, J. Freundlich, F. Combes, and A. Halle, MNRAS **492**, 877 (2020).
[145]B. Bar-Or, J.-B. Fouvry, and S. Tremaine, ApJ **871**, 28 (2019).
[146]M. D. Weinberg and N. Katz, MNRAS **375**, 425 (2007).
[147]A. D. Ludlow, S. M. Fall, J. Schaye, and D. Obreschkow, arXiv, 2105.03561 (2021).
[148]Y.-S. Ting and H.-W. Rix, ApJ **878**, 21 (2019).
[149]M. D. Weinberg, MNRAS **213**, 451 (1985).
[150]D. Dootson and J. Magorrian, arXiv, 2205.15725 (2022).
[151]Note that f still depends parametrically on $\mathbf{N}$ and $J_{\rm f}$, though we suppress the explicit dependence here to keep the notation clean.
[152]D. Lynden-Bell and A. J. Kalnajs, MNRAS **157**, 1 (1972).
[153]L. D. Landau, Zh. Eksp. Teor. Fiz. **16**, 574 (1946).
[154]S. Ichimaru, Phys. Rev. **140**, B226 (1965).
[155]R. W. Nelson and S. Tremaine, MNRAS **306**, 1 (1999).
[156]M. D. Weinberg, arXiv, astro-ph/0404169 (2004).
[157]Note, however, that Hamilton et al.[140] were working in three dimensions, since they were considering the interaction of a bar with a spherical halo. Here we are working in two dimensions, so our DF f is normalized differently.
[158]C. Hamilton, MNRAS, in press, stae305 (2024).
[159]In linear theory, the angular momentum transfer at each resonance $\mathbf{N}$ is governed by equation (152). The right hand side of this is clearly zero if one such resonance dominates and the gradient of the DF vanishes there. The reason that corotation is the dominant resonance is that the disk is initially rather cold. The corotation resonance acts most strongly upon perfectly circular orbits, whereas Lindblad resonances require significant radial epicyclic motion to have any effect.
[160]R. L. Dewar, Phys. Fluids **16**, 431 (1973).
[161]J. Binney, New Astron. Rev. **57**, 29 (2013).
[162]G. Contopoulos, *Order and chaos in dynamical astronomy* (Springer, 2002).
[163]D. Merritt, *Dynamics and Evolution of Galactic Nuclei* (Princeton Univ. Press, 2013).
[164]S. Sridhar and J. Touma, MNRAS **303**, 483 (1999).
[165]J. R. Cary and A. J. Brizard, RMP **81**, 693 (2009).
[166]I. Dodin and N. Fisch, PRE **77**, 036402 (2008).
[167]P. J. Channell, Phys. Fluids **19**, 1541 (1976).
[168]H. Nyquist, Bell Syst. Tech. J. **11**, 126 (1932).
[169]O. Penrose, Phys. Fluids **3**, 258 (1960).
[170]P. Palmer, *Stability of collisionless stellar systems* (Springer, 1994).
[171]I. B. Bernstein, J. M. Greene, and M. D. Kruskal, Phys. Rev. **108**, 546 (1957).
[172]A. M. Fridman and V. L. Polyachenko, *Physics of gravitating systems I: Equilibrium and stability* (Springer, 2012).
[173]J. Goodman, ApJ **329**, 612 (1988).
[174]P. O. Vandervoort, MNRAS **339**, 537 (2003).
[175]R. L. Dewar, Phys. Fluids **15**, 712 (1972).
[176]E. Athanassoula, Phys. Rep. **114**, 319 (1984).
[177]C. Dobbs and J. Baba, Publ. Astron. Soc. Aust. **31**, e035 (2014).
[178]L. Wang, R. Spurzem, S. Aarseth, K. Nitadori, P. Berczik, M. B. N. Kouwenhoven, and T. Naab, MNRAS **450**, 4070 (2015).
[179]A. A. Schekochihin, Lecture Notes for the Oxford MMathPhys programme (2022).
[180]W. Dehnen, AJ **118**, 1190 (1999).
[181]J. A. Sellwood, RMP **86**, 1 (2014).
[182]M. Hénon, A&A **114**, 211 (1982).