

Kinetic theory of spatially inhomogeneous stellar systems without collective effects[★]

P.-H. Chavanis

Laboratoire de Physique Théorique (IRSAMC), CNRS and UPS, Université de Toulouse, 31062 Toulouse Cedex 4, France
e-mail: chavanis@irsamc.ups-tlse.fr

Received 21 October 2012 / Accepted 27 March 2013

ABSTRACT

We review and complete the kinetic theory of spatially inhomogeneous stellar systems when collective effects (dressing of the stars by their polarization cloud) are neglected. We start from the BBGKY hierarchy issued from the Liouville equation and consider an expansion in powers of $1/N$ in a proper thermodynamic limit. For $N \rightarrow +\infty$, we obtain the Vlasov equation describing the evolution of collisionless stellar systems like elliptical galaxies. This corresponds to the mean field approximation. At the order $1/N$, we obtain a kinetic equation describing the evolution of collisional stellar systems like globular clusters. This corresponds to the weak coupling approximation. This equation coincides with the generalized Landau equation derived from a more abstract projection operator formalism. This equation does not suffer logarithmic divergences at large scales since spatial inhomogeneity is explicitly taken into account. Making a local approximation, and introducing an upper cut-off at the Jeans length, it reduces to the Vlasov-Landau equation which is the standard kinetic equation of stellar systems. Our approach provides a simple and pedagogical derivation of these important equations from the BBGKY hierarchy which is more rigorous for systems with long-range interactions than the two-body encounters theory. Making an adiabatic approximation, we write the generalized Landau equation in angle-action variables and obtain a Landau-type kinetic equation that is valid for fully inhomogeneous stellar systems and is free of divergences at large scales. This equation is less general than the recently derived Lenard-Balescu-type kinetic equation since it neglects collective effects, but it is substantially simpler and could be useful as a first step. We discuss the evolution of the system as a whole and the relaxation of a test star in a bath of field stars. We derive the corresponding Fokker-Planck equation in angle-action variables and provide expressions for the diffusion coefficient and friction force.

Key words. gravitation – methods: analytical – globular clusters: general

1. Introduction

In its simplest description, a stellar system can be viewed as a collection of N classical point mass stars in Newtonian gravitational interaction (Spitzer 1987; Heggie & Hut 2003; Binney & Tremaine 2008). As understood early on by Hénon (1964), self-gravitating systems experience two successive types of relaxation, a rapid collisionless relaxation towards a quasi-stationary state (QSS) that is a virialized state in mechanical equilibrium but not in thermodynamical equilibrium, followed by a slow collisional relaxation. One might think that, because of the development of stellar encounters, the system will reach, at sufficiently long times, a statistical equilibrium state described by the Maxwell-Boltzmann distribution. However, it is well-known that unbounded stellar systems cannot be in strict statistical equilibrium¹ because of the permanent escape of high energy stars (evaporation) and the gravothermal catastrophe (core collapse). Therefore, the statistical mechanics of stellar systems is essentially an out-of-equilibrium problem which must be approached through kinetic theories.

The first kinetic equation was written by Jeans (1915). Neglecting encounters between stars, he described the dynamical evolution of stellar systems by the collisionless Boltzmann

equation coupled to the Poisson equation. This purely mean field description applies to large groups of stars such as elliptical galaxies whose ages are much less than the collisional relaxation time. A similar equation was introduced by Vlasov (1938) in plasma physics to describe the collisionless evolution of a system of electric charges interacting by the Coulomb force. The collisionless Boltzmann equation coupled self-consistently to the Poisson equation is often called the Vlasov equation², or the Vlasov-Poisson system.

The concept of collisionless relaxation was first understood by Hénon (1964) and King (1966). Lynden-Bell (1967) developed a statistical theory of this process and coined the term “violent relaxation”. In the collisionless regime, the evolution of the star cluster is described by the Vlasov-Poisson system. Starting from an unsteady or unstable initial condition, the Vlasov-Poisson system develops a complicated mixing process in phase space. Because the Vlasov equation is time-reversible, it never achieves a steady state but develops filaments at smaller and smaller scales. However, the coarse-grained distribution function obtained by locally averaging the fine-grained distribution function over the filaments usually achieves a steady state on a few dynamical times. Lynden-Bell (1967) tried to predict the QSS resulting from violent relaxation by developing a statistical mechanics of the Vlasov equation. He derived a distribution

[★] Appendices are available in electronic form at <http://www.aanda.org>

¹ For reviews of the statistical mechanics of self-gravitating systems see e.g. Padmanabhan (1990), Katz (2003), and Chavanis (2006).

² See Hénon (1982) for a discussion about the name that should be given to this equation.

function formally equivalent to the Fermi-Dirac distribution (or to a superposition of Fermi-Dirac distributions). However, when coupled to the Poisson equation, these distributions have an infinite mass. Therefore, the Vlasov-Poisson system has no statistical equilibrium state (in the sense of Lynden-Bell). This is clear evidence that violent relaxation is incomplete (Lynden-Bell 1967). Incomplete relaxation is caused by inefficient mixing and non-ergodicity. In general, the fluctuations of the gravitational potential $\delta\Phi(\mathbf{r}, t)$ that drive the collisionless relaxation last only for a few dynamical times and die out before the system has mixed efficiently (Tremaine et al. 1986). Understanding the origin of incomplete relaxation, and developing models of incomplete violent relaxation to predict the structure of galaxies, is a very difficult problem (Arad & Johansson 2005). Some models of incomplete violent relaxation have been proposed based on different physical arguments (Bertin & Stiavelli 1984; Stiavelli & Bertin 1987; Hjorth & Madsen 1991; Chavanis et al. 1996; Chavanis 1998; Levin et al. 2008).

On longer timescales, stellar encounters (sometimes referred to as collisions by an abuse of language) must be taken into account. This description is particularly important for small groups of stars such as globular clusters whose ages are of the order of the collisional relaxation time. Chandrasekhar (1942, 1943a,b) developed a kinetic theory of stellar systems to determine the timescale of collisional relaxation and the rate of escape of stars from globular clusters³. To simplify the kinetic theory, he considered an infinite homogeneous system of stars. He started from the general Fokker-Planck equation and determined the diffusion coefficient and the friction force (first and second moments of the velocity increments) by considering the mean effect of a succession of two-body encounters⁴. This approach was based on Jeans (1929) demonstration that the cumulative effect of the weak deflections resulting from the relatively distant encounters is more important than the effect of occasional large deflections produced by relatively close encounters. However, this approach leads to a logarithmic divergence at large scales that is more difficult to remove in stellar dynamics than in plasma physics because of the absence of Debye shielding for the gravitational force⁵. Chandrasekhar & von Neumann (1942) developed a completely stochastic formalism of gravitational fluctuations and showed that the fluctuations of the gravitational force are given by the Holtzmark distribution (a particular Lévy law) in which the nearest neighbor plays a prevalent role. From these results, they argued that the logarithmic divergence has to be cut-off at the interparticle

distance l (see also Jeans 1929; and Spitzer 1940). However, since the interparticle distance is smaller than the Debye length, the same arguments should also apply in plasma physics, which is not the case. Therefore, the conclusions of Chandrasekhar and von Neumann are usually taken with circumspection. In particular, Cohen et al. (1950) argued that the logarithmic divergence should be cut-off at the Jeans length which gives an estimate of the system's size. While in neutral plasmas the effective interaction distance is limited to the Debye length, in a self-gravitating system the distance between interacting particles is only limited by the system's size. Therefore, the Jeans length is the gravitational analogue of the Debye length. These kinetic theories lead to a collisional relaxation time scaling as $t_R \sim (N/\ln N)t_D$, where t_D is the dynamical time and N the number of stars in the system. Chandrasekhar (1949) also developed a Brownian theory of stellar dynamics and showed that, from a qualitative point of view, the results of kinetic theory can be understood very simply in that framework⁶. In particular, he showed that a dynamical friction is necessary to maintain the Maxwell-Boltzmann distribution of statistical equilibrium and that the coefficients of friction and diffusion are related to each other by an Einstein relation which is a manifestation of the fluctuation-dissipation theorem. This relation is confirmed by his more precise kinetic theory based on two-body encounters. It is important to emphasize, however, that Chandrasekhar did not derive the kinetic equation for the evolution of the system as a whole. Indeed, he considered the Brownian motion of a test star in a fixed distribution of field stars (bath) and derived the corresponding Fokker-Planck equation. This equation has been used by Chandrasekhar (1943b), Spitzer & Härm (1958), Michie (1963), King (1965), and more recently by Lemou & Chavanis (2010) to study the evaporation of stars from globular clusters in a simple setting.

King (1960) noted that if we were to describe the dynamical evolution of the cluster as a whole, the distribution of the field stars should evolve in time in a self-consistent manner so that the kinetic equation must be an integro-differential equation. The kinetic equation obtained by King, from the results of Rosenbluth et al. (1957), is equivalent to the Landau equation, although written in a different form. It is interesting to note, for historical reasons, that none of the previous authors seemed to be aware of the work of Landau (1936) in plasma physics. There is, however, an important difference between stellar dynamics and plasma physics. Neutral plasmas are spatially homogeneous because of electroneutrality and Debye shielding. By contrast, stellar systems are spatially inhomogeneous. The above-mentioned kinetic theories developed for an infinite homogeneous system can be applied to an inhomogeneous system

³ Early estimates of the relaxation time of stellar systems were made by Schwarzschild (1924), Rosselland (1928), Jeans (1929), Smart (1938), and Spitzer (1940). On the other hand, the evaporation time was first estimated by Ambartsumian (1938) and Spitzer (1940).

⁴ Later, Cohen et al. (1950), Gasiorowicz et al. (1956), and Rosenbluth et al. (1957) proposed a simplified derivation of the coefficients of diffusion and friction.

⁵ A few years earlier, Landau (1936) had developed a kinetic theory of Coulombian plasmas taking two-body encounters into account. Starting from the Boltzmann (1872) equation, and making a weak deflection approximation, he derived a kinetic equation for the collisional evolution of neutral plasmas. His approach leads to a divergence at large scales that he removed heuristically by introducing a cut-off at the Debye length λ_D (Debye & Hückel 1923) which is the size over which the electric field produced by a charge is screened by the cloud of opposite charges. Later, Lenard (1960) and Balescu (1960) developed a more precise kinetic theory taking collective effects into account. They derived a more elaborate kinetic equation, free of divergence at large scales, in which the Debye length appears naturally. This justifies the heuristic procedure of Landau.

⁶ The stochastic evolution of a star is primarily due to many small deflections produced by relatively distant encounters. The problem of treating particles undergoing numerous weak deflections was originally encountered in relation to the Brownian motion of large molecules which are thermally agitated by the smaller field molecules. The stochastic character of the many small impulses which act on a suspended particle is usually described by the Fokker-Planck equation. Chandrasekhar (1943c) noted the analogy between stellar dynamics and Brownian theory and employed a Fokker-Planck equation to describe the evolution of the velocity distribution function of stars. Indeed, he argued that a given star undergoes many small-angle (large impact parameter) collisions in a time that is small compared with that in which its position or velocity changes appreciably. Since the Newtonian potential makes the cumulative effect of these small momentum transfers dominant, the stochastic methods of the Fokker-Planck equation should be more appropriate for a stellar system than the Boltzmann approach.

only if we make a local approximation. In that case, the collision term is calculated as if the system were spatially homogeneous or as if the collisions could be treated as local. Then, the effect of spatial inhomogeneity is only retained in the advection (Vlasov) term which describes the evolution of the system due to mean field effects. This leads to the Vlasov-Landau equation which is the standard kinetic equation of stellar dynamics. To our knowledge, this equation has been first written (in a different form), and studied, by Hénon (1961). Hénon also exploited the timescale separation between the dynamical time t_D and the relaxation time $t_R \gg t_D$ to derive a simplified kinetic equation for $f(\epsilon, t)$, where $\epsilon = v^2/2 + \Phi(\mathbf{r}, t)$ is the individual energy of a star by unit of mass, called the orbit-averaged Fokker-Planck equation. In this approach, the distribution function $f(\mathbf{r}, \mathbf{v}, t)$, averaged over a short timescale, is a steady state of the Vlasov equation of the form $f(\epsilon, t)$ which slowly evolves in time, on a long timescale, due to the development of collisions (i.e. correlations caused by finite N effects or graininess). Hénon used this equation to obtain a more relevant value for the rate of evaporation from globular clusters, valid for inhomogeneous systems. Cohn (1980) solved the orbit-averaged Fokker-Planck equation numerically to describe the collisional evolution of star clusters. His treatment accounts both for the escape of high energy stars put forward by Spitzer (1940) and for the phenomenon of core collapse resulting from the gravothermal catastrophe discovered by Antonov (1962) and Lynden-Bell & Wood (1968) on the basis of thermodynamics and statistical mechanics. The local approximation, which is a crucial step in the kinetic theory, is supported by the stochastic approach of Chandrasekhar & von Neumann (1942) showing the preponderance of the nearest neighbor. However, this remains a simplifying assumption which is not easily controllable. In particular, as we have already indicated, the local approximation leads to a logarithmic divergence at large scales that is difficult to remove. This divergence would not have occurred if full account of spatial inhomogeneity had been given from the start.

The effect of spatial inhomogeneity was investigated by Severne & Haggerty (1976), Parisot & Severne (1979), Kandrup (1981), and Chavanis (2008a,b). In particular, Kandrup (1981) derived a generalized Landau equation from the Liouville equation by using projection operator technics. This generalized Landau equation is interesting because it takes into account effects of spatial inhomogeneity which were neglected in previous approaches. Since the finite extension of the system is properly accounted for, there is no divergence at large scales⁷. Furthermore, this approach clearly shows which approximations are needed in order to recover the traditional Landau equation. Unfortunately, the generalized Landau equation remains extremely complicated for practical applications.

In addition, this equation is still approximate as it neglects collective effects and considers binary collisions between naked particles. As in any weakly coupled system, the particles engaged in collisions are dressed by the polarization clouds caused by their own influence on other particles. Collisions between dressed particles have quantitatively different outcomes than collisions between naked ones. In the case of plasmas, collective effects are responsible for Debye shielding and they are accounted

⁷ There remains a logarithmic divergence at small scales as in the original Landau (1936) theory due to the neglect of strong collisions. This divergence can be cured heuristically by introducing a cut-off at the Landau length λ_L corresponding to the impact parameter leading to a deflection at 90° (Landau 1936). This divergence does not occur in the theory of Chandrasekhar (1942, 1943a,b) which takes strong collisions into account.

for in the Lenard-Balescu equation. They eliminate the logarithmic divergence that occurs at large scales in the Landau equation. For self-gravitating systems, they lead to anti-shielding and are more difficult to analyze because the system is spatially inhomogeneous⁸. If we consider a finite homogeneous system, and take collective effects into account, one finds a severe divergence of the diffusion coefficient when the size of the domain reaches the Jeans scale (Weinberg 1993). This divergence, which is related to the Jeans instability, does not occur in a stable spatially inhomogeneous stellar system. Some authors like Thorne (1968), Miller (1968), Gilbert (1968, 1970), and Lerche (1971) attempted to take collective effects and spatial inhomogeneity into account. They obtained very complicated kinetic equations that have not found application until now. They managed, however, to show that collective effects are equivalent to increasing the effective mass of the stars, hence diminishing the relaxation time. Since, on the other hand, the effect of spatial inhomogeneity is to increase the relaxation time (Parisot & Severne 1979), the two effects act in opposite directions and may balance each other.

Recently, Heyvaerts (2010) derived from the BBGKY hierarchy issued from the Liouville equation a kinetic equation in angle-action variables that takes both spatial inhomogeneity and collective effects into account. To calculate the collective response, he used Fourier-Laplace transforms and introduced a bi-orthogonal basis of pairs of density-potential functions (Kalnajs 1971a). The kinetic equation derived by Heyvaerts is the counterpart for spatially inhomogeneous self-gravitating systems of the Lenard-Balescu equation for plasmas. Following his work, we showed that this equation could be obtained equivalently from the Klimontovich equation by making the so-called quasi-linear approximation (Chavanis 2012a). We also developed a test particle approach and derived the corresponding Fokker-Planck equation in angle-action variables, taking collective effects into account. This provides general expressions of the diffusion coefficient and friction force for spatially inhomogeneous stellar systems.

In a sense, these equations solve the problem of the kinetic theory of stellar systems since they take into account both spatial inhomogeneity and collective effects. However, the drawback is that they are extremely complicated to solve (in addition of being complicated to derive). In an attempt to reduce the complexity of the problem, we shall derive in this paper a kinetic equation that is valid for spatially inhomogeneous stellar systems but that neglects collective effects. Collective effects may be less crucial in stellar dynamics than in plasma physics. In plasma physics, they must be taken into account in order to remove the divergence at large scales that appears in the Landau equation. In the case of stellar systems, this divergence is removed by the spatial inhomogeneity of the system, not by collective effects. Actually, previous kinetic equations based on the local approximation ignore collective effects and already give satisfactory results. We shall therefore ignore collective effects and derive a kinetic equation (in position-velocity and angle-action

⁸ In a plasma, since the Coulomb force between electrons is repulsive, each particle of the plasma tends to attract to it particles of opposite charge and to repel particles of like charge, thereby creating a kind of cloud of opposite charge which screens the interaction at the scale of the Debye length. In the gravitational case, since the Newton force is attractive, the situation is considerably different. The test star draws neighboring stars into its vicinity and these add their gravitational force to that of the test star itself. The bare gravitational force of the test star is thus augmented rather than shielded. The polarization acts to increase the effective gravitational mass of a test star.

variables) that is the counterpart for spatially inhomogeneous self-gravitating systems of the Landau equation for plasmas. Our approach has three main interests. First, the derivation of this Landau-type kinetic equation is considerably simpler than the derivation of the Lenard-Balescu-type kinetic equation, and it can be done in the physical space without having to introduce Laplace-Fourier transforms nor bi-orthogonal basis of pairs of density-potential functions. This offers a more physical derivation of kinetic equations of stellar systems that may be of interest for astrophysicists. Secondly, our approach is sufficient to remove the large-scale divergence that occurs in kinetic theories based on the local approximation. It represents, therefore, a conceptual progress in the kinetic theory of stellar systems. Finally, this equation is simpler than the Lenard-Balescu-type kinetic equation derived by Heyvaerts (2010), and it could be useful as a first step before considering more complicated effects. Its drawback is to ignore collective effects but this may not be crucial as we have explained (as suggested by Weinberg's work, collective effects become important only when the system is close to instability). This Landau-type equation was previously derived for systems with arbitrary long-range interactions⁹ in various dimensions of space (Chavanis 2007, 2008a,b, 2010) but we think that it is important to discuss these results in the specific case of self-gravitating systems with complements and amplification.

The paper is organized as follows. In Sect. 2, we study the dynamical evolution of a spatially inhomogeneous stellar system as a whole. Starting from the BBGKY hierarchy issued from the Liouville equation, and neglecting collective effects, we derive a general kinetic equation valid at the order $1/N$ in a proper thermodynamic limit. For $N \rightarrow +\infty$, it reduces to the Vlasov equation. At the order $1/N$ we recover the generalized Landau equation derived by Kandrup (1981) from a more abstract projection operator formalism. This equation is free of divergence at large scales since spatial inhomogeneity has been properly accounted for. Making a local approximation and introducing an upper cut-off at the Jeans length, we recover the standard Vlasov-Landau equation which is usually derived from a kinetic theory based on two-body encounters. Our approach provides an alternative derivation of this fundamental equation from the more rigorous Liouville equation. It has therefore some pedagogical interest. In Sect. 3, we study the relaxation of a test star in a steady distribution of field stars. We derive the corresponding Fokker-Planck equation and determine the expressions of the diffusion and friction coefficients. We emphasize the difference between the friction by polarization and the total friction (this difference may have been overlooked in previous works). For a thermal bath, we derive the Einstein relation between the diffusion and friction coefficients and obtain the explicit expression of the diffusion tensor. This returns the standard results obtained from the two-body encounters theory but, again, our presentation is different and offers an alternative derivation of these important results. For that reason, we give a short review of the basic formulae. In Sect. 4, we derive a Landau-type kinetic equation written in angle-action variables and discuss its main properties. This equation, which does not make the local approximation, applies to fully inhomogeneous stellar systems and is free of divergence at large scales. We also develop a test particle approach and derive the corresponding Fokker-Planck equation in angle-action variables. Explicit expressions are given for the diffusion tensor and friction force, and they are compared with previous expressions obtained in the literature.

⁹ For a review on the dynamics and thermodynamics of systems with long-range interactions, see Campa et al. (2009).

2. Evolution of the system as a whole

2.1. The BBGKY hierarchy

We consider an isolated system of N stars with identical mass m in Newtonian interaction. Their dynamics is fully described by the Hamilton equations

$$m \frac{d\mathbf{r}_i}{dt} = \frac{\partial H}{\partial \mathbf{v}_i}, \quad m \frac{d\mathbf{v}_i}{dt} = -\frac{\partial H}{\partial \mathbf{r}_i},$$

$$H = \frac{1}{2} \sum_{i=1}^N m v_i^2 - Gm^2 \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (1)$$

This Hamiltonian system conserves the energy $E = H$, the mass $M = Nm$, and the angular momentum $\mathbf{L} = \sum_i m \mathbf{r}_i \times \mathbf{v}_i$. As recalled in the Introduction, stellar systems cannot reach a statistical equilibrium state in a strict sense. In order to understand their evolution, it is necessary to develop a kinetic theory.

We introduce the N -body distribution function $P_N(\mathbf{r}_1, \mathbf{v}_1, \dots, \mathbf{r}_N, \mathbf{v}_N, t)$ giving the probability density of finding at time t the first star with position \mathbf{r}_1 and velocity \mathbf{v}_1 , the second star with position \mathbf{r}_2 and velocity \mathbf{v}_2 , etc. Basically, the evolution of the N -body distribution function is governed by the Liouville equation

$$\frac{\partial P_N}{\partial t} + \sum_{i=1}^N \left(\mathbf{v}_i \cdot \frac{\partial P_N}{\partial \mathbf{r}_i} + \mathbf{F}_i \cdot \frac{\partial P_N}{\partial \mathbf{v}_i} \right) = 0, \quad (2)$$

where

$$\mathbf{F}_i = -\frac{\partial \Phi_d}{\partial \mathbf{r}_i} = -Gm \sum_{j \neq i} \frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3} = \sum_{j \neq i} \mathbf{F}(j \rightarrow i) \quad (3)$$

is the gravitational force by unit of mass experienced by the i th star due to its interaction with the other stars. Here, $\Phi_d(\mathbf{r})$ denotes the exact gravitational potential produced by the discrete distribution of stars and $\mathbf{F}(j \rightarrow i)$ denotes the exact force by unit of mass created by the j th star on the i th star. The Liouville equation (Eq. (2)), which is equivalent to the Hamilton equations (Eq. (1)), contains too much information to be exploitable. In practice, we are only interested in the evolution of the one-body distribution $P_1(\mathbf{r}, \mathbf{v}, t)$.

From the Liouville equation we can construct the complete BBGKY hierarchy for the reduced distribution functions

$$P_j(\mathbf{x}_1, \dots, \mathbf{x}_j, t) = \int P_N(\mathbf{x}_1, \dots, \mathbf{x}_N, t) d\mathbf{x}_{j+1} \dots d\mathbf{x}_N, \quad (4)$$

where the notation \mathbf{x} stands for (\mathbf{r}, \mathbf{v}) . The generic term of this hierarchy reads

$$\frac{\partial P_j}{\partial t} + \sum_{i=1}^j \mathbf{v}_i \cdot \frac{\partial P_j}{\partial \mathbf{r}_i} + \sum_{i=1}^j \sum_{k=1, k \neq i}^j \mathbf{F}(k \rightarrow i) \cdot \frac{\partial P_j}{\partial \mathbf{v}_i} + (N-j) \sum_{i=1}^j \int \mathbf{F}(j+1 \rightarrow i) \cdot \frac{\partial P_{j+1}}{\partial \mathbf{v}_i} d\mathbf{x}_{j+1} = 0. \quad (5)$$

This hierarchy of equations is not closed since the equation for the one-body distribution $P_1(\mathbf{x}_1, t)$ involves the two-body distribution $P_2(\mathbf{x}_1, \mathbf{x}_2, t)$, the equation for the two-body distribution $P_2(\mathbf{x}_1, \mathbf{x}_2, t)$ involves the three-body distribution $P_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, t)$, and so on.

It is convenient to introduce a cluster representation of the distribution functions. Specifically, we can express the reduced distribution $P_j(\mathbf{x}_1, \dots, \mathbf{x}_j, t)$ in terms of products of distribution functions $P_{j<j'}(\mathbf{x}_1, \dots, \mathbf{x}_{j'}, t)$ of lower order plus a correlation function $P'_j(\mathbf{x}_1, \dots, \mathbf{x}_j, t)$ (see e.g. Eqs. (8) and (9) below). Considering the scaling of the terms in each equation of the BBGKY hierarchy, we can see that there exist solutions of the whole BBGKY hierarchy such that the correlation functions P'_j scale as $1/N^{j-1}$ in the proper thermodynamic limit $N \rightarrow +\infty$ defined in Appendix A. This implicitly assumes that the initial condition has no correlation, or that the initial correlations respect this scaling¹⁰. If this scaling is satisfied, we can consider an expansion of the BBGKY hierarchy in terms of the small parameter $1/N$. This is similar to the expansion of the BBGKY hierarchy in plasma physics in terms of the small parameter $1/\Lambda$, where $\Lambda \gg 1$ represents the number of charges in the Debye sphere (Balescu 2000). However, in plasma physics, the system is spatially homogeneous (because of Debye shielding which restricts the range of interaction) while, for stellar systems, spatial inhomogeneity must be taken into account. This brings additional terms in the BBGKY hierarchy that are absent in plasma physics.

2.2. The truncation of the BBGKY hierarchy at the order $1/N$

The first two equations of the BBGKY hierarchy are

$$\begin{aligned} \frac{\partial P_1}{\partial t}(1) + \mathbf{v}_1 \cdot \frac{\partial P_1}{\partial \mathbf{r}_1}(1) \\ + (N-1) \int \mathbf{F}(2 \rightarrow 1) \cdot \frac{\partial P_2}{\partial \mathbf{v}_1}(1, 2) d\mathbf{x}_2 = 0, \end{aligned} \quad (6)$$

$$\begin{aligned} \frac{1}{2} \frac{\partial P_2}{\partial t}(1, 2) + \mathbf{v}_1 \cdot \frac{\partial P_2}{\partial \mathbf{r}_1}(1, 2) + \mathbf{F}(2 \rightarrow 1) \cdot \frac{\partial P_2}{\partial \mathbf{v}_1}(1, 2) \\ + (N-2) \int \mathbf{F}(3 \rightarrow 1) \cdot \frac{\partial P_3}{\partial \mathbf{v}_1}(1, 2, 3) d\mathbf{x}_3 + (1 \leftrightarrow 2) = 0. \end{aligned} \quad (7)$$

We decompose the two- and three-body distributions in the form

$$P_2(\mathbf{x}_1, \mathbf{x}_2) = P_1(\mathbf{x}_1)P_1(\mathbf{x}_2) + P'_2(\mathbf{x}_1, \mathbf{x}_2), \quad (8)$$

$$\begin{aligned} P_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = P_1(\mathbf{x}_1)P_1(\mathbf{x}_2)P_1(\mathbf{x}_3) + P'_2(\mathbf{x}_1, \mathbf{x}_2)P_1(\mathbf{x}_3) \\ + P'_2(\mathbf{x}_1, \mathbf{x}_3)P_1(\mathbf{x}_2) + P'_2(\mathbf{x}_2, \mathbf{x}_3)P_1(\mathbf{x}_1) + P'_3(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3), \end{aligned} \quad (9)$$

where $P'_j(\mathbf{x}_1, \dots, \mathbf{x}_j, t)$ is the correlation function of order j . Substituting Eqs. (8) and (9) in Eqs. (6) and (7), and simplifying some terms, we obtain

$$\begin{aligned} \frac{\partial P_1}{\partial t}(1) + \mathbf{v}_1 \cdot \frac{\partial P_1}{\partial \mathbf{r}_1}(1) \\ + (N-1) \left[\int \mathbf{F}(2 \rightarrow 1)P_1(2) d\mathbf{x}_2 \right] \cdot \frac{\partial P_1}{\partial \mathbf{v}_1}(1) = \\ - (N-1) \frac{\partial}{\partial \mathbf{v}_1} \cdot \int \mathbf{F}(2 \rightarrow 1)P'_2(1, 2) d\mathbf{x}_2, \end{aligned} \quad (10)$$

¹⁰ If there are non-trivial correlations in the initial state, like primordial binary stars, the kinetic theory will be different from the one developed in the sequel.

$$\begin{aligned} \frac{1}{2} \frac{\partial P'_2}{\partial t}(1, 2) + \mathbf{v}_1 \cdot \frac{\partial P'_2}{\partial \mathbf{r}_1}(1, 2) + \mathbf{F}(2 \rightarrow 1) \cdot \frac{\partial P'_2}{\partial \mathbf{v}_1}(1, 2) \\ + (N-2) \left[\int \mathbf{F}(3 \rightarrow 1)P_1(3) d\mathbf{x}_3 \right] \cdot \frac{\partial P'_2}{\partial \mathbf{v}_1}(1, 2) \\ + \left[\mathbf{F}(2 \rightarrow 1) - \int \mathbf{F}(3 \rightarrow 1)P_1(3) d\mathbf{x}_3 \right] \cdot \frac{\partial P_1}{\partial \mathbf{v}_1}(1)P_1(2) \\ + (N-2) \left[\int \mathbf{F}(3 \rightarrow 1)P'_2(2, 3) d\mathbf{x}_3 \right] \cdot \frac{\partial P_1}{\partial \mathbf{v}_1}(1) \\ - \frac{\partial}{\partial \mathbf{v}_1} \cdot \int \mathbf{F}(3 \rightarrow 1)P'_2(1, 3)P_1(2) d\mathbf{x}_3 \\ + (N-2) \frac{\partial}{\partial \mathbf{v}_1} \cdot \int \mathbf{F}(3 \rightarrow 1)P'_3(1, 2, 3) d\mathbf{x}_3 + (1 \leftrightarrow 2) = 0. \end{aligned} \quad (11)$$

Equations (10) and (11) are exact for all N , but they are not closed. As explained previously, we shall close these equations at the order $1/N$ in the thermodynamic limit $N \rightarrow +\infty$. In this limit $P_1 \sim 1$, $P'_2 \sim 1/N$, and $P'_3 \sim 1/N^2$. On the other hand, we can introduce dimensionless variables such that $|\mathbf{r}| \sim |\mathbf{v}| \sim t \sim m \sim 1$ and $|\mathbf{F}(i \rightarrow j)| \sim G \sim 1/N$ (see Appendix A).

The advection term $\mathcal{V} = \mathcal{V}_0 + \mathcal{V}_{\text{m.f.}}$ in the left-hand side (l.h.s.) of Eq. (10) is of order 1, and the collision term C in the right-hand side (r.h.s.) is of order $1/N$. Let us now consider the terms in Eq. (11) one by one. The first four terms correspond to the Liouville equation. The Liouville operator $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}' + \mathcal{L}_{\text{m.f.}}$ describes the complete two-body problem, including the inertial motion, the interaction between the stars (1, 2) and the mean field produced by the other stars. The terms \mathcal{L}_0 and $\mathcal{L}_{\text{m.f.}}$ are of order $1/N$ while the term \mathcal{L}' is of order $1/N^2$. Therefore, the interaction term \mathcal{L}' can a priori¹¹ be neglected in the Liouville operator. This corresponds to the weak coupling approximation where only the mean field term $\mathcal{L}_0 + \mathcal{L}_{\text{m.f.}}$ is retained. The fifth term in Eq. (11) is a source term \mathcal{S} expressible in terms of the one-body distribution; it is of order $1/N$. If we consider only the mean field Liouville operator $\mathcal{L}_0 + \mathcal{L}_{\text{m.f.}}$ and the source term \mathcal{S} , as we shall do in this paper, we can obtain a kinetic equation for stellar systems that is the counterpart of the Landau equation in plasma physics. The sixth term C is of order $1/N$ and it corresponds to collective effects (i.e. the dressing of the particles by the polarization cloud). In plasma physics, this term leads to the Lenard-Balescu equation. It takes into account dynamical screening and regularizes the divergence at large scales that appears in the Landau equation. In the case of stellar systems, there is no large-scale divergence because of the spatial inhomogeneity of the system. Therefore, collective effects are less crucial in the kinetic theory of stellar systems than in plasma physics. However, this term has been properly taken into account by Heyvaerts (2010) who obtained a kinetic equation of stellar systems that is the counterpart of the Lenard-Balescu equation in plasma physics. The last two terms \mathcal{T} are of the order $1/N^2$ and they will be neglected. In particular, the three-body correlation function P'_3 , of order $1/N^2$, can be neglected at the order $1/N$. In this way, the hierarchy of equations is closed and a kinetic equation involving only two-body encounters can be obtained.

¹¹ Actually, the interaction term becomes large at small scales $|\mathbf{r}_2 - \mathbf{r}_1| \rightarrow 0$, so its effect is not totally negligible (in other words, the expansion in terms of $1/N$ is not uniformly convergent). In particular, the interaction term \mathcal{L}' must be taken into account in order to describe strong collisions with small impact parameter that lead to large deflections very different from the mean field trajectory corresponding to $\mathcal{L}_0 + \mathcal{L}_{\text{m.f.}}$. This is important to regularize the divergence at small scales that appears in the Landau equation. We shall return to this problem in Sect. 2.4.

If we introduce the notations $f = NmP_1$ (distribution function) and $g = N^2m^2P_2'$ (two-body correlation function), we get at the order $1/N$

$$\begin{aligned} \frac{\partial f}{\partial t}(1) + \mathbf{v}_1 \cdot \frac{\partial f}{\partial \mathbf{r}_1}(1) + \frac{N-1}{N} \langle \mathbf{F} \rangle(1) \cdot \frac{\partial f}{\partial \mathbf{v}_1}(1) = \\ - \frac{1}{m} \frac{\partial}{\partial \mathbf{v}_1} \cdot \int \mathbf{F}(2 \rightarrow 1) g(1, 2) \, d\mathbf{x}_2, \\ \frac{1}{2} \frac{\partial g}{\partial t}(1, 2) + \mathbf{v}_1 \cdot \frac{\partial g}{\partial \mathbf{r}_1}(1, 2) + \langle \mathbf{F} \rangle(1) \cdot \frac{\partial g}{\partial \mathbf{v}_1}(1, 2) \\ + \tilde{\mathbf{F}}(2 \rightarrow 1) \cdot \frac{\partial f}{\partial \mathbf{v}_1}(1) f(2) + \mathbf{F}(2 \rightarrow 1) \cdot \frac{\partial g}{\partial \mathbf{v}_1}(1, 2) \\ + \frac{1}{m} \left[\int \mathbf{F}(3 \rightarrow 1) g(2, 3) \, d\mathbf{x}_3 \right] \cdot \frac{\partial f}{\partial \mathbf{v}_1}(1) + (1 \leftrightarrow 2) = 0. \end{aligned} \quad (12)$$

We have introduced the mean force (by unit of mass) created on star 1 by all the other stars

$$\langle \mathbf{F} \rangle(1) = \int \mathbf{F}(2 \rightarrow 1) \frac{f(2)}{m} \, d\mathbf{x}_2 = -\nabla \Phi(1), \quad (14)$$

and the fluctuating force (by unit of mass) created by star 2 on star 1

$$\tilde{\mathbf{F}}(2 \rightarrow 1) = \mathbf{F}(2 \rightarrow 1) - \frac{1}{N} \langle \mathbf{F} \rangle(1). \quad (15)$$

Equations (12) and (13) are exact at the order $1/N$. They form the right basis to develop the kinetic theory of stellar systems at this order of approximation. Since the collision term in the r.h.s. of Eq. (12) is of order $1/N$, we expect that the relaxation time of stellar systems scales as $\sim Nt_D$ where t_D is the dynamical time. As we shall see, the discussion is more complicated because of the presence of logarithmic corrections in the relaxation time and the absence of a strict statistical equilibrium state.

2.3. The limit $N \rightarrow +\infty$: the Vlasov equation (collisionless regime)

In the limit $N \rightarrow +\infty$, for any fixed interval of time $[0, T]$, the correlations between stars can be neglected. Therefore, the mean field approximation becomes exact and the N -body distribution function factorizes in N one-body distribution functions

$$P_N(\mathbf{x}_1, \dots, \mathbf{x}_N, t) = \prod_{i=1}^N P_1(\mathbf{x}_i, t). \quad (16)$$

Substituting this factorization in the Liouville equation, and integrating over $\mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_N$, we find that the smooth distribution function $f(\mathbf{r}, \mathbf{v}, t) = NmP_1(\mathbf{r}, \mathbf{v}, t)$ is the solution of the Vlasov equation

$$\begin{aligned} \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \langle \mathbf{F} \rangle \cdot \frac{\partial f}{\partial \mathbf{v}} = 0, \\ \langle \mathbf{F} \rangle = -\nabla \Phi, \quad \Delta \Phi = 4\pi G \int f \, d\mathbf{v}. \end{aligned} \quad (17)$$

This equation also results from Eq. (12) if we neglect the correlation function $g(1, 2)$ in the r.h.s. and replace $N-1$ by N .

The Vlasov equation describes the collisionless evolution of stellar systems for times shorter than the relaxation time $\sim Nt_D$. In practice $N \gg 1$ so that the domain of validity of the Vlasov equation is huge (see the end of Sect. 2.8). As recalled in the Introduction, the Vlasov-Poisson system develops a process of

phase mixing and violent relaxation leading to a QSS on a very short timescale, of the order of a few dynamical times t_D . Elliptical galaxies are in such QSSs. Lynden-Bell (1967) developed a statistical mechanics of the Vlasov equation in order to describe this process of violent relaxation and predict the QSS achieved by the system. Unfortunately, the predictions of his statistical theory are limited by the problem of incomplete relaxation. Kinetic theories of violent relaxation, which may account for incomplete relaxation, have been developed by Kadomtsev & Pogutse (1970), Severne & Luwel (1980), and Chavanis (1998, 2008a,b).

2.4. The order $O(1/N)$: the generalized Landau equation (collisional regime)

If we neglect strong collisions and collective effects, the first two equations of the BBGKY hierarchy (Eqs. (12) and (13)) reduce to

$$\begin{aligned} \frac{\partial f}{\partial t}(0) + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}}(0) + \frac{N-1}{N} \langle \mathbf{F} \rangle(0) \cdot \frac{\partial f}{\partial \mathbf{v}}(0) = \\ - \frac{1}{m} \frac{\partial}{\partial \mathbf{v}} \cdot \int \mathbf{F}(1 \rightarrow 0) g(0, 1) \, d\mathbf{x}_1, \end{aligned} \quad (18)$$

$$\begin{aligned} \frac{1}{2} \frac{\partial g}{\partial t}(0, 1) + \mathbf{v} \cdot \frac{\partial g}{\partial \mathbf{r}}(0, 1) + \langle \mathbf{F} \rangle(0) \cdot \frac{\partial g}{\partial \mathbf{v}}(0, 1) + (0 \leftrightarrow 1) = \\ - \tilde{\mathbf{F}}(1 \rightarrow 0) \cdot \frac{\partial f}{\partial \mathbf{v}}(0) f(1) + (0 \leftrightarrow 1). \end{aligned} \quad (19)$$

The first equation gives the evolution of the one-body distribution function. The l.h.s. corresponds to the (Vlasov) advection term. The r.h.s. takes into account correlations (finite N effects, graininess, discreteness effects) between stars that develop as a result of their interactions. These correlations correspond to encounters (collisions).

Equation (19) may be viewed as a linear first order differential equation in time. It can be symbolically written as

$$\frac{\partial g}{\partial t} + \mathcal{L}g = \mathcal{S}[f], \quad (20)$$

where $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_{m.f.}$ is a mean field Liouville operator and $\mathcal{S}[f]$ is a source term \mathcal{S} expressible in terms of the one-body distribution function. This equation may be solved by the method of characteristics. Introducing the Green function

$$G(t, t') = \exp \left\{ - \int_{t'}^t \mathcal{L}(\tau) \, d\tau \right\}, \quad (21)$$

constructed with the mean field Liouville operator \mathcal{L} , we obtain

$$\begin{aligned} g(\mathbf{x}, \mathbf{x}_1, t) = - \int_0^t \, d\tau G(t, t-\tau) \\ \times \left[\tilde{\mathbf{F}}(1 \rightarrow 0) \cdot \frac{\partial}{\partial \mathbf{v}} + \tilde{\mathbf{F}}(0 \rightarrow 1) \cdot \frac{\partial}{\partial \mathbf{v}_1} \right] \\ \times f(\mathbf{x}, t-\tau) f(\mathbf{x}_1, t-\tau), \end{aligned} \quad (22)$$

where we have assumed that no correlation is present initially so that $g(\mathbf{x}, \mathbf{x}_1, t=0) = 0$; if correlations are present initially,

it can be shown that they are rapidly washed out¹². Substituting Eq. (22) in Eq. (18), we obtain

$$\begin{aligned} \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \frac{N-1}{N} \langle \mathbf{F} \rangle \cdot \frac{\partial f}{\partial \mathbf{v}} = & \\ \frac{\partial}{\partial v^\mu} \int_0^\tau d\tau \int d\mathbf{r}_1 d\mathbf{v}_1 F^\mu(1 \rightarrow 0) G(t, t-\tau) & \\ \times \left[\tilde{F}^\nu(1 \rightarrow 0) \frac{\partial}{\partial v^\nu} + \tilde{F}^\nu(0 \rightarrow 1) \frac{\partial}{\partial v_1^\nu} \right] & \\ \times f(\mathbf{r}, \mathbf{v}, t-\tau) \frac{f}{m}(\mathbf{r}_1, \mathbf{v}_1, t-\tau). & \end{aligned} \quad (23)$$

In writing this equation, we have adopted a Lagrangian point of view. The coordinates \mathbf{r}_i appearing after the Green function must be viewed as $\mathbf{r}_i(t-\tau) = \mathbf{r}_i(t) - \int_0^\tau d\mathbf{s} \mathbf{v}_i(t-s)$ and $\mathbf{v}_i(t-\tau) = \mathbf{v}_i(t) - \int_0^\tau d\mathbf{s} \langle \mathbf{F} \rangle(\mathbf{r}_i(t-s), t-s)$. Therefore, in order to evaluate the integral of Eq. (23), we must move the stars following the trajectories determined by the self-consistent mean field.

The kinetic equation (Eq. (23)) is valid at the order $1/N$ so it describes the collisional evolution of the system (ignoring collective effects) on a timescale of order Nt_D . Equation (23) is a non-Markovian integro-differential equation. It takes into account delocalizations in space and time (i.e. spatial inhomogeneity and memory effects). Actually, the Markovian approximation is justified in the $N \rightarrow +\infty$ limit because the timescale $\sim Nt_D$ over which $f(\mathbf{r}, \mathbf{v}, t)$ changes is long compared to the correlation time $\tau_{\text{corr}} \sim t_D$ over which the integrand in Eq. (23) has significant support¹³. Therefore, we can compute the correlation function (Eq. (22)) by assuming that the distribution function is frozen at time t . This corresponds to the Bogoliubov ansatz in plasma physics. If we replace $f(\mathbf{r}, \mathbf{v}, t-\tau)$ and $f(\mathbf{r}_1, \mathbf{v}_1, t-\tau)$ by $f(\mathbf{r}, \mathbf{v}, t)$ and $f(\mathbf{r}_1, \mathbf{v}_1, t)$ in Eq. (23) and extend the time integral to $+\infty$, we obtain

$$\begin{aligned} \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \frac{N-1}{N} \langle \mathbf{F} \rangle \cdot \frac{\partial f}{\partial \mathbf{v}} = \frac{\partial}{\partial v^\mu} \int_0^{+\infty} d\tau \int d\mathbf{r}_1 d\mathbf{v}_1 F^\mu(1 \rightarrow 0) & \\ \times G(t, t-\tau) \left[\tilde{F}^\nu(1 \rightarrow 0) \frac{\partial}{\partial v^\nu} + \tilde{F}^\nu(0 \rightarrow 1) \frac{\partial}{\partial v_1^\nu} \right] & \\ \times f(\mathbf{r}, \mathbf{v}, t) \frac{f}{m}(\mathbf{r}_1, \mathbf{v}_1, t). & \end{aligned} \quad (24)$$

Similarly, we can compute the trajectories of the stars by assuming that the mean field is independent on τ and equal to its value at time t so that $\mathbf{r}_i(t-\tau) = \mathbf{r}_i(t) - \int_0^\tau d\mathbf{s} \mathbf{v}_i(t-s)$ and $\mathbf{v}_i(t-\tau) = \mathbf{v}_i(t) - \int_0^\tau d\mathbf{s} \langle \mathbf{F} \rangle(\mathbf{r}_i(t-s), t)$.

The structure of the kinetic equation (Eq. (24)) has a clear physical meaning. The l.h.s. corresponds to the Vlasov advection term arising from mean field effects. The r.h.s. can be viewed as a collision operator $C_N[f]$ taking finite N effects into account. For $N \rightarrow +\infty$, it vanishes and we recover the Vlasov equation. For finite N , it describes the cumulative effect of binary collisions between stars. The collision operator is a sum of two terms,

¹² The term corresponding to $g(\mathbf{x}, \mathbf{x}_1, t=0)$ in the kinetic equation has been called the ‘‘destruction term’’ by Prigogine and Résibois because it describes the destruction of the effect of the initial correlations (Balescu 2000).

¹³ It is sometimes argued that the Markovian approximation is not justified for stellar systems because the force auto-correlation decreases slowly as $1/t$ (Chandrasekhar 1944). However, this result is only true for an infinite homogeneous system (see Appendix G). For spatially inhomogeneous distributions, the correlation function decreases more rapidly and the Markovian approximation is justified (Sewane & Haggerty 1976).

a diffusion term and a friction term. The coefficients of diffusion and friction are given by generalized Kubo formulae, i.e. they involve the time integral of the auto-correlation function of the fluctuating force. The kinetic equation (Eq. (24)) bears some resemblance with the Fokker-Planck equation. However, it is more complicated since it is an integro-differential equation, not a differential equation (see Sect. 3).

Equation (24) may be viewed as a generalized Landau equation. Since the spatial inhomogeneity and the finite extension of the system are properly taken into account, there is no divergence at large scales. There is, however, a logarithmic divergence at small scales which is due to the neglect of the interaction term \mathcal{L}' in the Liouville operator (see Sect. 2.6). At large scales (i.e. for large impact parameters), this term can be neglected and the trajectories of the stars are essentially due to the mean field. Thus, we can make the weak coupling approximation leading to the Landau equation. This approximation describes weak collisions. However, at small scales (i.e. for small impact parameters), we can no longer ignore the interaction term \mathcal{L}' in the Liouville operator and we have to solve the classical two-body problem. This is necessary to correctly describe strong collisions for which the trajectory of the particles deviates strongly from the mean field motion. When the mean field Green function G (constructed with $\mathcal{L}_0 + \mathcal{L}_{\text{m.f.}}$) is replaced by the total Green function G' (constructed with $\mathcal{L}_0 + \mathcal{L}' + \mathcal{L}_{\text{m.f.}}$), taking into account the interaction term, the generalized Landau equation (Eq. (24)) is replaced by a more complicated equation which can be viewed as a generalized Boltzmann equation. This equation is free of divergence (since it takes both spatial inhomogeneity and strong collisions into account), but it is unnecessarily complicated because it does not exploit the dominance of weak collisions over rare strong collisions for the gravitational potential. Indeed, a star suffers a large number of weak distant encounters and very few close encounters. A better practical procedure is to use the generalized Landau equation (Eq. (24)) with a cut-off at small scales in order to take into account our inability to describe strong collisions by this approach (see Sect. 2.6).

The generalized Landau equation (Eq. (24)) was derived by Kandrup (1981) from the Liouville equation by using the projection operator formalism. It can also be derived from the Liouville equation by using the BBGKY hierarchy or from the Klimontovich equation by making a quasilinear approximation (Chavanis 2008a,b).

2.5. The Vlasov-Landau equation

Self-gravitating systems are intrinsically spatially inhomogeneous. However, the collision operator at position \mathbf{r} can be simplified by making a local approximation and performing the integrals as if the system were spatially homogeneous with the density $\rho = \rho(\mathbf{r})$. This amounts to replacing $f(\mathbf{r}_1, \mathbf{v}_1, t)$ by $f(\mathbf{r}, \mathbf{v}_1, t)$ and $\tilde{F}(i \rightarrow j)$ by $\mathbf{F}(i \rightarrow j)$ in Eq. (24). This local approximation is motivated by the work of Chandrasekhar & von Neumann (1942) who showed that the distribution of the gravitational force is a Lévy law (called the Holtzmark distribution) dominated by the contribution of the nearest neighbor. With this local approximation, Eq. (24) becomes

$$\begin{aligned} \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \frac{N-1}{N} \langle \mathbf{F} \rangle \cdot \frac{\partial f}{\partial \mathbf{v}} = & \\ \frac{\partial}{\partial v^\mu} \int_0^{+\infty} d\tau \int d\mathbf{r}_1 d\mathbf{v}_1 F^\mu(1 \rightarrow 0) G_0(t, t-\tau) & \\ \times F^\nu(1 \rightarrow 0) \left(\frac{\partial}{\partial v^\nu} - \frac{\partial}{\partial v_1^\nu} \right) f(\mathbf{r}, \mathbf{v}, t) \frac{f}{m}(\mathbf{r}, \mathbf{v}_1, t), & \end{aligned} \quad (25)$$

where we have used $\mathbf{F}(0 \rightarrow 1) = -\mathbf{F}(1 \rightarrow 0)$. The Green function G_0 corresponds to the free motion of the particles associated with the Liouville operator \mathcal{L}_0 . Using Eqs. (C.1) and (C.2), the foregoing equation can be rewritten as

$$\begin{aligned} \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \frac{N-1}{N} \langle \mathbf{F} \rangle \cdot \frac{\partial f}{\partial \mathbf{v}} = \\ \frac{\partial}{\partial v^\mu} \int_0^{+\infty} d\tau \int d\mathbf{r}_1 d\mathbf{v}_1 F^\mu(1 \rightarrow 0, t) F^\nu(1 \rightarrow 0, t - \tau) \\ \times \left(\frac{\partial}{\partial v^\nu} - \frac{\partial}{\partial v_1^\nu} \right) f(\mathbf{r}, \mathbf{v}, t) \frac{f}{m}(\mathbf{r}, \mathbf{v}_1, t), \end{aligned} \quad (26)$$

where $\mathbf{F}(1 \rightarrow 0, t - \tau)$ is expressed in terms of the Lagrangian coordinates. The integrals over τ and \mathbf{r}_1 can be calculated explicitly (see Appendices C and D). We then find that the evolution of the distribution function is governed by the Vlasov-Landau equation

$$\begin{aligned} \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \frac{N-1}{N} \langle \mathbf{F} \rangle \cdot \frac{\partial f}{\partial \mathbf{v}} = \pi(2\pi)^3 m \\ \times \frac{\partial}{\partial v^\mu} \int k^\mu k^\nu \delta(\mathbf{k} \cdot \mathbf{w}) \hat{u}^2(k) \left(f_1 \frac{\partial f}{\partial v^\nu} - f \frac{\partial f_1}{\partial v_1^\nu} \right) d\mathbf{v}_1 d\mathbf{k}, \end{aligned} \quad (27)$$

where we have noted $\mathbf{w} = \mathbf{v} - \mathbf{v}_1$, $f = f(\mathbf{r}, \mathbf{v}, t)$, $f_1 = f(\mathbf{r}, \mathbf{v}_1, t)$, and where $(2\pi)^3 \hat{u}(k) = -4\pi G/k^2$ represents the Fourier transform of the gravitational potential. Under this form, we see that the collisional evolution of a stellar system is due to a condition of resonance $\mathbf{k} \cdot \mathbf{v} = \mathbf{k} \cdot \mathbf{v}'$ (with $\mathbf{v} \neq \mathbf{v}'$) encapsulated in the δ -function. This δ -function expresses the conservation of energy.

The Vlasov-Landau equation may also be written as (see Appendix C)

$$\begin{aligned} \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \frac{N-1}{N} \langle \mathbf{F} \rangle \cdot \frac{\partial f}{\partial \mathbf{v}} = \\ \frac{\partial}{\partial v^\mu} \int K^{\mu\nu} \left(f_1 \frac{\partial f}{\partial v^\nu} - f \frac{\partial f_1}{\partial v_1^\nu} \right) d\mathbf{v}_1, \end{aligned} \quad (28)$$

$$K^{\mu\nu} = A \frac{w^2 \delta^{\mu\nu} - w^\mu w^\nu}{w^3}, \quad A = 2\pi m G^2 \ln \Lambda, \quad (29)$$

where

$$\ln \Lambda = \int_{k_{\min}}^{k_{\max}} dk/k \quad (30)$$

is the Coulombian logarithm that has to be regularized with appropriate cut-offs (see Sect. 2.6). The r.h.s. of Eq. (28) is the original form of the collision operator given by Landau (1936) for the Coulombian interaction¹⁴. It applies to weakly coupled plasmas. We note that the potential of interaction only appears in the constant A which merely determines the relaxation time. The structure of the Landau equation is independent of the potential. The Landau equation was originally derived from the Boltzmann equation in the limit of weak deflections $|\Delta\mathbf{v}| \ll 1$ (Landau 1936)¹⁵. In the case of plasmas, the system is spatially

¹⁴ The case of plasmas is recovered by making the substitution $Gm^2 \leftrightarrow -e^2$ leading to $A = (2\pi e^4/m^3) \ln \Lambda$.

¹⁵ We note that Eq. (25) with G_0 replaced by the total Green function G' taking into account the interaction term is equivalent to the Boltzmann equation. Indeed, for spatially homogeneous systems, the Boltzmann equation can be derived from the BBGKY hierarchy (Eqs. (10) and (11)) by keeping the Liouville operator $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}'$ describing the two-body problem exactly and the source \mathcal{S} (Balescu 2000). Therefore, the procedure used by Landau which amounts to expanding the Boltzmann equation in the limit of weak deflections is equivalent to the one presented here that starts from the BBGKY hierarchy and neglects \mathcal{L}' in the Liouville operator.

homogeneous and the advection term is absent in Eq. (28). In the case of stellar systems, when we make the local approximation, the spatial inhomogeneity of the system is only retained in the advection term of Eq. (28). This is why this kinetic equation is referred to as the Vlasov-Landau equation. This is the fundamental kinetic equation of stellar systems.

2.6. Heuristic regularization of the divergences

To obtain the Vlasov-Landau equation (Eq. (28)), we have made a local approximation. This amounts to calculating the collision operator at each point as if the system were spatially homogeneous. As a result of this homogeneity assumption, a logarithmic divergence appears at large scales in the Coulombian logarithm (Eq. (30)). In plasma physics, this divergence is cured by the Debye shielding. A charge is surrounded by a polarization cloud of opposite charges which reduces the range of the interaction. When collective effects are properly taken into account, as in the Lenard-Balescu equation, no divergence appears at large scales and the Debye length arises naturally. Heuristically, we can use the Landau equation and introduce an upper cut-off at the Debye length $\lambda_D \sim (k_B T / ne^2)^{1/2}$. For self-gravitating systems, there is no shielding and the divergence is cured by the finite extent of the system. The interaction between two stars is only limited by the size of the system. When spatial inhomogeneity is taken into account, as in the generalized Landau equation (Eq. (24)), no divergence occurs at large scales. Heuristically, we can use the Vlasov-Landau equation (Eq. (28)) and introduce an upper cut-off at the Jeans length $\lambda_J \sim (k_B T / Gm^2 n)^{1/2}$ which gives an estimate of the system's size R .

The Coulombian logarithm (Eq. (30)) also diverges at small scales. As explained previously, this is due to the neglect of strong collisions that produce important deflections. Indeed, for collisions with small impact parameter, the mean field approximation is clearly irrelevant and it is necessary to solve the two-body problem exactly (see Appendix H). Accordingly, the small-scale divergence is cured by the proper treatment of strong collisions. Heuristically, we can use the Landau equation and introduce a lower cut-off at the gravitational Landau length $\lambda_L \sim Gm/v_m^2 \sim Gm^2/(k_B T)$ (the gravitational analogue of the Landau length $\lambda_L \sim e^2/mv_m^2 \sim e^2/k_B T$ in plasma physics) which corresponds to the impact parameter leading to a deflection at 90° .

Introducing a large-scale cut-off at the Jeans length λ_J and a small-scale cut-off at the Landau length λ_L , and noting that $\lambda_L \sim 1/(n\lambda_J^2)$, we find that the Coulombian logarithm scales as $\ln \Lambda \sim \ln(\lambda_J/\lambda_L) \sim \ln(n\lambda_J^3) \sim \ln N$ where $\Lambda \sim n\lambda_J^3 \sim nR^3 \sim N$ is the number of stars in the cluster¹⁶.

2.7. Properties of the Vlasov-Landau equation

The Vlasov-Landau equation conserves the mass $M = \int f d\mathbf{r} d\mathbf{v}$ and the energy $E = \int f \frac{v^2}{2} d\mathbf{r} d\mathbf{v} + \frac{1}{2} \int \rho \Phi d\mathbf{r}$. It also monotonically increases the Boltzmann entropy $S = -\int (f/m) \ln(f/m) d\mathbf{r} d\mathbf{v}$ in the sense that $\dot{S} \geq 0$ (H -theorem). As a result of the local approximation, the proof of these properties is the same¹⁷

¹⁶ Similarly, in plasma physics, introducing a large-scale cut-off at the Debye length λ_D and a small-scale cut-off at the Landau length λ_L , and noting that $\lambda_L \sim 1/(n\lambda_D^2)$, we find that the Coulombian logarithm scales as $\ln \Lambda \sim \ln(\lambda_D/\lambda_L) \sim \ln(n\lambda_D^3)$, where $\Lambda \sim n\lambda_D^3$ is the number of electrons in the Debye sphere.

¹⁷ We recall that the Vlasov advection term conserves mass, energy, and entropy.

as for the spatially homogeneous Landau equation (Baesuc 2000). Because of these properties, we might expect that a stellar system will relax towards the Boltzmann distribution which maximizes the entropy at fixed mass and energy. However, we know that there is no maximum entropy state for an unbounded self-gravitating system (the Boltzmann distribution has infinite mass). Therefore, the Vlasov-Landau equation does not relax towards a steady state and the entropy does not reach a stationary value. Actually, the entropy increases permanently as the system evaporates. But since evaporation is a slow process, the system may achieve a QSS that is close to the Boltzmann distribution. A typical quasi-stationary distribution is the Michie-King model

$$f = Ae^{-\beta m \mathbf{j}^2 / (2r_a^2)} (e^{-\beta m \epsilon} - e^{-\beta m \epsilon_m}), \quad (31)$$

where $\epsilon = v^2/2 + \Phi(\mathbf{r})$ is the energy and $\mathbf{j} = \mathbf{r} \times \mathbf{v}$ the angular momentum. This distribution takes into account the escape of high energy stars and the anisotropy of the velocity distribution. It can be derived, under some approximations, from the Vlasov-Landau equation by using the condition that $f = 0$ if the energy of the star is larger than the escape energy ϵ_m (Michie 1963; King 1965). The Michie-King distribution reduces to the isothermal distribution $f \propto e^{-\beta m \epsilon}$ for low energies. In this sense, we can define a relaxation time for a stellar system. From the Vlasov-Landau equation (Eq. (28)), we find that the relaxation time scales as

$$t_R \sim \frac{v_m^3}{nm^2 G^2 \ln N}, \quad (32)$$

where v_m is the root mean square (rms) velocity of the stars. Introducing the dynamical time $t_D \sim \lambda_J / v_m \sim R / v_m$, we obtain the scaling

$$t_R \sim \frac{N}{\ln N} t_D. \quad (33)$$

The fact that the ratio between the relaxation time and the dynamical time depends only on the number of stars and scales as $N/\ln N$ was first noted by Chandrasekhar (1942).

A simple estimate of the evaporation time was calculated by Ambartsumian (1938) and Spitzer (1940) and gives $t_{\text{evap}} \approx 136 t_R$. More precise values have been obtained by studying the evaporation process in an artificially uniform medium (Chandrasekhar 1943b; Spitzer & Härm 1958; Michie 1963; King 1965; Lemou & Chavanis 2010) or in a more realistic inhomogeneous cluster (Hénon 1961). Since $t_{\text{evap}} \gg t_R$, we can consider that the system relaxes towards a steady distribution given by Eq. (31) on a timescale t_R and that this distribution slowly evolves on a longer timescale as the stars escape¹⁸. The characteristic time in which the system's stars evaporate is t_{evap} . The evaporation is one reason for the evolution of stellar systems. However, as demonstrated by Antonov (1962) and Lynden-Bell

¹⁸ This distribution is not steady in the sense that the coefficients A , β , and r_a slowly vary in time as the stars escape and the system loses mass and energy. However, the distribution keeps the same form during the evaporation process (King 1965). Hénon (1961) previously found a self-similar solution of the orbit-averaged Fokker-Planck equation. He showed that the core contracts as the cluster evaporates and that the central density is infinite (the structure of the core resembles the singular isothermal sphere with a density $\rho \propto r^{-2}$). He argued that the concentration of energy, without concentration of mass, at the center of the system is attributable to the formation of tight binary stars. The invariant profile found by Hénon does not exactly coincide with the Michie-King distribution, but is reasonably close.

& Wood (1968), stellar systems may evolve more rapidly because of gravothermal catastrophe. In that case, the Michie-King distribution changes significantly as a result of core collapse. This evolution has been described by Cohn (1980), and it leads ultimately to the formation of a binary star surrounded by a hot halo. A configuration of this type can have arbitrarily large entropy. Cohn (1980) found that the entropy increases permanently during core collapse, confirming that the Vlasov-Landau equation has no equilibrium state.

Even if the system were confined within a small box so as to prevent both the evaporation and the gravothermal catastrophe, there would be no statistical equilibrium state in a strict sense because there is no global entropy maximum (Antonov 1962). A configuration in which some subset of the particles are tightly bound together (e.g. a binary star), and in which the remaining particles share the energy thereby released, may have an arbitrarily large entropy. However, such configurations, which require strong correlations, are generally reached very slowly (on a timescale much larger than $(N/\ln N)t_D$) as a result of encounters involving many particles. To describe these configurations, one would have to take high order correlations into account in the kinetic theory. These configurations may be relevant in systems with a small number of stars (Chabanol et al. 2000), but when N is large the picture is different. On a timescale of the order of $(N/\ln N)t_D$ the one-body distribution function is expected to reach the Boltzmann distribution which is a local entropy maximum. This state is metastable, but its lifetime is expected to be very large, scaling as e^N , so that it is stable in practice (Chavanis 2005, 2006). In this sense, there exist true statistical equilibrium states for self-gravitating systems confined within a small box. However, we may argue that this situation is highly artificial.

Finally, using very different methods based on ergodic theory, Gurzadyan & Savvidy (1986) argue that, because of collective behaviour, the relaxation time for stellar systems scales like $t_R^* \sim N^{1/3} t_D$ (intermediate between the dynamical time t_D (violent relaxation) and the binary relaxation time $t_R \sim (N/\ln N)t_D$).

2.8. Dynamical evolution of stellar systems: a short review

Using the kinetic theory, we can identify different phases in the dynamical evolution of stellar systems.

A self-gravitating system initially out-of-mechanical equilibrium undergoes a process of violent collisionless relaxation towards a virialized state. In this regime, the dynamical evolution of the cluster is described by the Vlasov-Poisson system. The phenomenology of violent relaxation has been described by Hénon (1964), King (1966), and Lynden-Bell (1967). Numerical simulations that start from cold and clumpy initial conditions generate a QSS that fits the de Vaucouleurs $R^{1/4}$ law for the surface brightness of elliptical galaxies quite well (van Albada 1982). The inner core is almost isothermal (as predicted by Lynden-Bell 1967) while the velocity distribution in the envelope is radially anisotropic and the density profile decreases as r^{-4} . One success of Lynden-Bell's statistical theory of violent relaxation is to explain the isothermal core of elliptical galaxies without recourse to collisions. In contrast, the structure of the halo cannot be explained by Lynden-Bell's theory as it results from an incomplete relaxation. Models of incompletely relaxed stellar systems have been elaborated by Bertin & Stiavelli (1984), Stiavelli & Bertin (1987), and Hjorth & Madsen (1991). These theoretical models nicely reproduce the results of observations and numerical simulations (Londrillo et al. 1991; Trenti et al. 2005). In the simulations, the initial condition needs to be

sufficiently clumpy and cold to generate enough mixing required for a successful application of the statistical theory of violent relaxation. Numerical simulations starting from homogeneous spheres (see e.g. Roy & Perez 2004; Levin et al. 2008; Joyce et al. 2009) show little angular momentum mixing and lead to different results. In particular, they display a larger amount of mass loss (evaporation) than simulations starting from clumpy initial conditions. Clumps thus help the system to reach a universal final state from a variety of initial conditions, which can explain the similarity of the density profiles observed in elliptical galaxies.

On longer timescales, encounters between stars must be taken into account and the dynamical evolution of the cluster is governed by the Vlasov-Landau-Poisson system. The first stage of the collisional evolution is driven by evaporation. Because of a series of weak encounters, the energy of a star can gradually increase until it reaches the local escape energy; in that case the star leaves the system. Numerical simulations (Spitzer 1987; Binney & Tremaine 2008) show that during this regime the system reaches a QSS that slowly evolves in amplitude as a result of evaporation as the system loses mass and energy. This quasi-stationary distribution function is close to the Michie-King model (Eq. (31)). The system has a core-halo structure. The core is isothermal while the stars in the outer halo move in predominantly radial orbits. Therefore, the distribution function in the halo is anisotropic. The density follows the isothermal law $\rho \sim r^{-2}$ in the central region (with a core of almost uniform density) and decreases as $\rho \sim r^{-7/2}$ in the halo (for an isolated cluster with $\epsilon_m = 0$) or tends to zero as $\rho \propto (r_t - r)^{3/2}$ (for a cluster with a finite tidal radius). Because of evaporation, the halo expands while the core shrinks as required by energy conservation. During the evaporation process, the central density increases permanently. At some point in the evolution, the system undergoes an instability related to the Antonov (1962) instability¹⁹ and the gravothermal catastrophe sets in (Lynden-Bell & Wood 1968). This instability is due to the negative specific heat of the inner system that evolves by losing energy and thereby growing hotter. The energy lost is transferred outward by stellar encounters. Hence the temperature always decreases outward, and the center continually loses energy, shrinks, and heats up. This leads to core collapse. Mathematically speaking, core collapse would generate a finite time singularity. When the evolution is modeled by the orbit-averaged Fokker-Planck equation, Cohn (1980) finds that the collapse is self-similar, that the central density becomes infinite in a finite time, and that the density behaves as $\rho \sim r^{-2.23}$. The invariant profile found by Cohn differs from the Michie-King distribution (for which $\rho \sim r^{-2}$) beyond a radius of about $10r_{\text{core}}$. Larson (1970) and Lynden-Bell & Eggleton (1980) find similar results by modeling the evolution of the system by fluid equations. In all cases, the authors find a singular density profile at the collapse time that is integrable at $r = 0$. This means that “the singularity contains no mass”²⁰. In reality, if we come back to the N -body system, there is no singularity and core collapse is arrested by the formation of binary stars caused by three-body collisions. These binaries

can release sufficient energy to stop the collapse and even drive a re-expansion of the cluster in a post-collapse regime (Inagaki & Lynden-Bell 1983; Goodman 1984). Then, in principle, a series of gravothermal oscillations should follow (Bettwieser & Sugimoto 1984).

At the present epoch, small groups of stars such as globular clusters ($N \sim 10^5$, $t_D \sim 10^5$ yr, age $\sim 10^{10}$ yr, $t_R \sim 10^{10}$ yr) are in the collisional regime. They are either in QSSs described by the Michie-King model or experiencing core collapse. By contrast, large clusters of stars like elliptical galaxies ($N \sim 10^{11}$, $t_D \sim 10^8$ yr, age $\sim 10^{10}$ yr, $t_R \sim 10^{19}$ yr) are still in the collisionless regime and their apparent organization is a result of an incomplete violent relaxation.

3. Test star in a thermal bath

3.1. The Fokker-Planck equation

We now consider the relaxation of a test star (i.e. a tagged particle) evolving in a steady distribution of field stars²¹. Because of the encounters with the field stars, the test star has a stochastic motion. We call $P(\mathbf{r}, \mathbf{v}, t)$ the probability density of finding the test star at position \mathbf{r} with velocity \mathbf{v} at time t . The evolution of $P(\mathbf{r}, \mathbf{v}, t)$ can be obtained from the generalized Landau equation (Eq. (24)) by considering that the distribution function of the field stars is fixed. Therefore, we replace $f(\mathbf{r}, \mathbf{v}, t)$ by $P(\mathbf{r}, \mathbf{v}, t)$ and $f(\mathbf{r}_1, \mathbf{v}_1, t)$ by $f(\mathbf{r}_1, \mathbf{v}_1)$ where $f(\mathbf{r}_1, \mathbf{v}_1)$ is the steady distribution of the field stars. This procedure transforms the integro-differential equation (Eq. (24)) into the differential equation

$$\begin{aligned} \frac{\partial P}{\partial t} + \mathbf{v} \cdot \frac{\partial P}{\partial \mathbf{r}} + \frac{N-1}{N} \langle \mathbf{F} \rangle \cdot \frac{\partial P}{\partial \mathbf{v}} = \\ \frac{\partial}{\partial v^\mu} \int_0^{+\infty} d\tau \int d\mathbf{r}_1 d\mathbf{v}_1 F^\mu(1 \rightarrow 0) \\ \times G(t, t - \tau) \left[\tilde{F}^\nu(1 \rightarrow 0) \frac{\partial}{\partial v^\nu} + \tilde{F}^\nu(0 \rightarrow 1) \frac{\partial}{\partial v_1^\nu} \right] \\ \times P(\mathbf{r}, \mathbf{v}, t) \frac{f}{m}(\mathbf{r}_1, \mathbf{v}_1), \end{aligned} \quad (34)$$

where $\langle \mathbf{F} \rangle(\mathbf{r}) = -\nabla\Phi(\mathbf{r})$ is the static mean force created by the field stars with density $\rho(\mathbf{r}_1)$. This equation does not present any divergence at large scales. We can understand the above procedure as follows (see Chavanis 2012b, for more details). Equations (24) and (34) govern the evolution of the distribution function of a test star (described by the coordinates \mathbf{r} and \mathbf{v}) interacting with field stars (described by the running coordinates \mathbf{r}_1 and \mathbf{v}_1). In Eq. (24), all the stars are equivalent so the distribution of the field stars $f(\mathbf{r}_1, \mathbf{v}_1, t)$ changes with time exactly like the distribution of the test star $f(\mathbf{r}, \mathbf{v}, t)$. In Eq. (34), the test star and the field stars are not equivalent since the field stars form a bath. The field stars have a steady given distribution $f(\mathbf{r}_1, \mathbf{v}_1)$ while the distribution of the test star $P(\mathbf{r}, \mathbf{v}, t)$ changes with time.

¹⁹ The collisional evolution of the system can be measured precisely in terms of the scale escape energy $x_0 = 3\Phi(0)/v_m^2(0)$. The onset of instability corresponds to $x_0 \simeq 9.3$ which is the value at which the King model becomes thermodynamically unstable (Cohn 1980).

²⁰ Actually, the number of stars in the core decreases as the collapse proceeds. When $<10^2$ stars remain in the core, strong collisions become important. Goodman (1983) showed that the evolution caused by strong collisions is very similar to that caused by numerous weak ones, except for the rate of evaporation which is greatly increased.

²¹ In plasma physics, this steady distribution is the Boltzmann distribution of statistical equilibrium which is the steady state of the Landau equation. In the case of stellar systems, there is an intrinsic difficulty since no statistical equilibrium state exists in a strict sense: the Boltzmann distribution has infinite mass, and the Vlasov-Landau equation has no steady state because of the escape of high-energy stars. However, we have seen that the system can reach a quasi-steady distribution (e.g. a Michie-King distribution) and that this distribution changes on an evaporation timescale that is long with respect to the collisional relaxation time. Therefore, we can consider that this distribution is steady on the collisional timescale over which the Fokker-Planck approach applies.

If we make a local approximation and use the Vlasov-Landau equation (Eq. (26)), we obtain

$$\begin{aligned} \frac{\partial P}{\partial t} + \mathbf{v} \cdot \frac{\partial P}{\partial \mathbf{r}} + \frac{N-1}{N} \langle \mathbf{F} \rangle \cdot \frac{\partial P}{\partial \mathbf{v}} = \\ \frac{\partial}{\partial v^\mu} \int_0^{+\infty} d\tau \int d\mathbf{r}_1 d\mathbf{v}_1 F^\mu(1 \rightarrow 0, t) F^\nu(1 \rightarrow 0, t - \tau) \\ \times \left(\frac{\partial}{\partial v^\nu} - \frac{\partial}{\partial v_1^\nu} \right) P(\mathbf{r}, \mathbf{v}, t) \frac{f}{m}(\mathbf{r}, \mathbf{v}_1). \end{aligned} \quad (35)$$

Denoting the advection operator by d/dt , Eq. (35) can be written in the form of a Fokker-Planck equation

$$\frac{dP}{dt} = \frac{\partial}{\partial v^\mu} \left(D^{\mu\nu} \frac{\partial P}{\partial v^\nu} - P F_{\text{pol}}^\mu \right) \quad (36)$$

involving a diffusion tensor

$$\begin{aligned} D^{\mu\nu} = \frac{1}{m} \int_0^{+\infty} d\tau \int d\mathbf{r}_1 d\mathbf{v}_1 F^\mu(1 \rightarrow 0, t) \\ \times F^\nu(1 \rightarrow 0, t - \tau) f(\mathbf{r}, \mathbf{v}_1) \end{aligned} \quad (37)$$

and a friction force

$$\begin{aligned} F_{\text{pol}}^\mu = \frac{1}{m} \int_0^{+\infty} d\tau \int d\mathbf{r}_1 d\mathbf{v}_1 F^\mu(1 \rightarrow 0, t) \\ \times F^\nu(1 \rightarrow 0, t - \tau) \frac{\partial f}{\partial v_1^\nu}(\mathbf{r}, \mathbf{v}_1). \end{aligned} \quad (38)$$

If we start directly from Eq. (27), which amounts to performing the integrals over τ and \mathbf{r}_1 in the previous expressions, we obtain the Fokker-Planck equation

$$\begin{aligned} \frac{\partial P}{\partial t} + \mathbf{v} \cdot \frac{\partial P}{\partial \mathbf{r}} + \frac{N-1}{N} \langle \mathbf{F} \rangle \cdot \frac{\partial P}{\partial \mathbf{v}} = \pi(2\pi)^3 m \\ \times \frac{\partial}{\partial v^\mu} \int k^\mu k^\nu \delta(\mathbf{k} \cdot \mathbf{w}) \hat{u}(k)^2 \left(f_1 \frac{\partial P}{\partial v^\nu} - P \frac{\partial f_1}{\partial v_1^\nu} \right) d\mathbf{v}_1 d\mathbf{k}, \end{aligned} \quad (39)$$

with the diffusion and friction coefficients

$$D^{\mu\nu} = \pi(2\pi)^3 m \int k^\mu k^\nu \delta(\mathbf{k} \cdot \mathbf{w}) \hat{u}(k)^2 f_1 d\mathbf{v}_1 d\mathbf{k}, \quad (40)$$

$$F_{\text{pol}}^\mu = \pi(2\pi)^3 m \int k^\mu k^\nu \delta(\mathbf{k} \cdot \mathbf{w}) \hat{u}(k)^2 \frac{\partial f_1}{\partial v_1^\nu} d\mathbf{v}_1 d\mathbf{k}. \quad (41)$$

Using Eq. (40), we can easily establish the identity

$$\frac{\partial D^{\mu\nu}}{\partial v^\nu} = \frac{\partial D}{\partial v^\mu}, \quad (42)$$

where $D = D^{\mu\mu} = \text{tr}(D)$.

The diffusion tensor $D^{\mu\nu}$ results from the fluctuations of the gravitational force caused by the granularities in the distribution of the field stars. It can be derived directly from the formula (see Appendix G)

$$D^{\mu\nu} = \int_0^{+\infty} \langle F^\mu(t) F^\nu(t - \tau) \rangle d\tau, \quad (43)$$

obtained from Eq. (45-a). The friction force F_{pol} results from the retroaction of the field stars to the perturbation caused by the test star, as in a polarization process. It can be derived from a linear response theory (Marochnik 1968; Kalnajs 1971b; Kandrup 1983; Chavanis 2008b). It will be called the friction by polarization to distinguish it from the total friction (see below). Equations (35)–(41) have been derived within the local approximation. More general formulae, valid for fully inhomogeneous stellar systems, are given in Kandrup (1983) and

Chavanis (2008b). The friction force has also been calculated by Tremaine & Weinberg (1984), Bekenstein & Maoz (1992), Maoz (1993), and Nelson & Tremaine (1999) using different approaches.

Since the diffusion tensor depends on the velocity \mathbf{v} of the test star, it is useful to rewrite Eq. (36) in a form that is fully consistent with the general Fokker-Planck equation

$$\frac{dP}{dt} = \frac{\partial^2}{\partial v^\mu \partial v^\nu} (D^{\mu\nu} P) - \frac{\partial}{\partial v^\mu} (P F_{\text{friction}}^\mu), \quad (44)$$

with

$$D^{\mu\nu} = \frac{\langle \Delta v^\mu \Delta v^\nu \rangle}{2\Delta t}, \quad F_{\text{friction}}^\mu = \frac{\langle \Delta v^\mu \rangle}{\Delta t}. \quad (45)$$

By identification, we find that

$$F_{\text{friction}}^\mu = F_{\text{pol}}^\mu + \frac{\partial D^{\mu\nu}}{\partial v^\nu}. \quad (46)$$

Therefore, when the diffusion coefficient depends on the velocity, the total friction is different from the friction by polarization. Substituting Eqs. (40) and (41) in Eq. (46), and using an integration by parts, we find that the diffusion and friction coefficients may be written as

$$\frac{\langle \Delta v^\mu \Delta v^\nu \rangle}{2\Delta t} = \pi(2\pi)^3 m \int k^\mu k^\nu \delta(\mathbf{k} \cdot \mathbf{w}) \hat{u}(k)^2 f_1 d\mathbf{v}_1 d\mathbf{k}, \quad (47)$$

$$\begin{aligned} \frac{\langle \Delta v^\mu \rangle}{\Delta t} = \pi(2\pi)^3 m \int k^\mu k^\nu f_1 \left(\frac{\partial}{\partial v^\nu} - \frac{\partial}{\partial v_1^\nu} \right) \\ \times \delta(\mathbf{k} \cdot \mathbf{w}) \hat{u}(k)^2 d\mathbf{v}_1 d\mathbf{k}. \end{aligned} \quad (48)$$

These expressions can be obtained directly from the equations of motion by expanding the trajectories of the stars in powers of $1/N$ in the limit $N \rightarrow +\infty$ (Chavanis 2008b). We recall that Eqs. (47) and (48) display a logarithmic divergence at small and large scales that must be regularized by introducing proper cut-offs as explained in Sect. 2.6. In astrophysics, the diffusion and friction coefficients of a star were first calculated by Chandrasekhar (1943a) from a two-body encounters theory (see also Cohen et al. 1950; Gasiorowicz et al. 1956; and Rosenbluth et al. 1957). The expressions obtained by these authors are different from those given above, but they are equivalent (see Sect. 3.5). In plasma physics, the diffusion and friction coefficients of a charge were first calculated by Hubbard (1961a) who took collective effects into account, thereby eliminating the divergence at large scales. When collective effects are neglected, his expressions reduce to Eqs. (47) and (48). On the other hand, strong collisions have been taken into account by Chandrasekhar (1943a) in astrophysics and by Hubbard (1961b) in plasma physics. In that case, there is no divergence at small impact parameters in the diffusion and friction coefficients, and the Landau length appears naturally (see Appendix H).

The two forms of the Fokker-Planck equation (Eqs. (36) and (44)) have their own interest. The expression (Eq. (44)) where the diffusion coefficient is placed after the two derivatives $\partial^2(DP)$ involves the total friction force F_{friction} and the expression (Eq. (36)) where the diffusion coefficient is placed between the derivatives $\partial D \partial P$ involves the friction by polarization F_{pol} . Astrophysicists are used to the form in Eq. (44). However, it is the form in Eq. (36) that stems from the Landau equation (Eq. (26)). We shall come back to this observation in Sect. 3.5.

From Eqs. (40) and (41), we easily obtain

$$\frac{\partial D^{\mu\nu}}{\partial v^\nu} = F_{\text{pol}}^\mu. \quad (49)$$

Combining Eq. (46) with Eq. (49), we get

$$\mathbf{F}_{\text{friction}} = 2\mathbf{F}_{\text{pol}}. \quad (50)$$

Therefore, the friction force $\mathbf{F}_{\text{friction}}$ is equal to twice the friction by polarization \mathbf{F}_{pol} (for a test star with mass m interacting with field stars with mass m_f , this factor two is replaced by $(m + m_f)/m$; see Appendix F). This explains the difference of factor 2 in the calculations of Chandrasekhar (1943a) who determined $\mathbf{F}_{\text{friction}}$ and in the calculations of Kalnajs (1971b) and Kandrup (1983) who determined \mathbf{F}_{pol} .

3.2. The Einstein relation

In the central region of the system, the distribution of the field stars is close to the Maxwell-Boltzmann distribution²²

$$f(\mathbf{r}_1, \mathbf{v}_1) = \left(\frac{\beta m}{2\pi}\right)^{3/2} \rho(\mathbf{r}_1) e^{-\beta m \frac{v_1^2}{2}}, \quad (51)$$

where $\beta = 1/(k_B T)$ is the inverse temperature and $\rho(\mathbf{r}_1) \propto e^{-\beta m \Phi(\mathbf{r}_1)}$ is the density given by the Boltzmann law. Therefore, the field stars form a thermal bath. Substituting the identity

$$\frac{\partial f_1}{\partial \mathbf{v}_1} = -\beta m f_1 \mathbf{v}_1, \quad (52)$$

in Eq. (41), using the δ -function to replace $\mathbf{k} \cdot \mathbf{v}_1$ by $\mathbf{k} \cdot \mathbf{v}$, and comparing the resulting expression with Eq. (40), we find that

$$\mathbf{F}_{\text{pol}}^\mu = -\beta m D^{\mu\nu}(\mathbf{v}) v^\nu. \quad (53)$$

The friction coefficient is given by an Einstein relation expressing the fluctuation-dissipation theorem²³. We emphasize that the Einstein relation is valid for the friction force by polarization \mathbf{F}_{pol} , not for the total friction $\mathbf{F}_{\text{friction}}$ (we do not have this subtlety in standard Brownian theory where the diffusion coefficient is constant). Using Eq. (43), we can rewrite Eq. (53) in the form

$$\mathbf{F}_{\text{pol}}^\mu = -\beta m v^\nu \int_0^{+\infty} d\tau \langle F^\mu(t) F^\nu(t - \tau) \rangle. \quad (54)$$

This relation is usually called the Kubo formula. More general expressions of the Kubo formula valid for fully inhomogeneous stellar systems are given in Kandrup (1983) and Chavanis (2008b). Using the Einstein relation, the Fokker-Planck equation (Eq. (36)) takes the form

$$\frac{dP}{dt} = \frac{\partial}{\partial v^\mu} \left[D^{\mu\nu}(\mathbf{v}) \left(\frac{\partial P}{\partial v^\nu} + \beta m P v^\nu \right) \right], \quad (55)$$

where the diffusion coefficient is given by Eq. (40) with Eq. (51). This equation is similar to the Kramers equation in Brownian theory (Kramers 1940) except that the diffusion coefficient is a tensor and that it depends on the velocity of the test star. For an

²² This isothermal distribution cannot hold in the whole cluster since it has an infinite mass.

²³ The collisional evolution of a star, under the effect of two-body encounters, can be understood as an interplay between two competing effects. The fluctuations of the gravitational field induce a diffusion in velocity space which tends to increase the speed of the star. This effect is counterbalanced by a friction (dissipation) which results in a systematic deceleration along the direction of motion. As emphasized by Chandrasekhar (1943a, 1949) in his Brownian theory of stellar motion, the Einstein relation (Eq. (53)) guarantees that the Maxwell distribution (Eq. (51)) is a steady state of the Fokker-Planck equation (Eq. (36)).

isotropic distribution function (e.g. the Maxwellian), it can be put in the form

$$D^{\mu\nu} = \left(D_{\parallel} - \frac{1}{2} D_{\perp} \right) \frac{v^\mu v^\nu}{v^2} + \frac{1}{2} D_{\perp} \delta^{\mu\nu}, \quad (56)$$

where D_{\parallel} and D_{\perp} are the diffusion coefficients in the directions parallel and perpendicular to the velocity of the test star. We note that $D \equiv D^{\mu\mu} = D_{\parallel} + D_{\perp}$. Using $D^{\mu\nu} v^\nu = D_{\parallel} v^\mu$, the friction by polarization (Eq. (53)) can be written as

$$\mathbf{F}_{\text{pol}} = -D_{\parallel} \beta m \mathbf{v}. \quad (57)$$

The friction is proportional and opposite to the velocity of the test star, and the friction coefficient is given by the Einstein relation $\xi = D_{\parallel} \beta m$. The total friction is

$$\mathbf{F}_{\text{friction}} = -2D_{\parallel} \beta m \mathbf{v}. \quad (58)$$

This is the Chandrasekhar dynamical friction.

The steady state of the Fokker-Planck equation (Eq. (55)) is the Maxwell-Boltzmann distribution (Eq. (51)). Since the Fokker-Planck equation admits an H -theorem for the Boltzmann free energy (Risken 1989), one can prove that $P(\mathbf{r}, \mathbf{v}, t)$ converges towards the Maxwell-Boltzmann distribution (Eq. (51)) for $t \rightarrow +\infty$. In other words, the test star acquires the distribution of the field stars (thermalization). We recall, however, that for self-gravitating systems, these results cannot hold everywhere in the cluster since the Maxwell-Boltzmann distribution does not exist globally.

If we assume that the distribution $P(\mathbf{r}, \mathbf{v}, t)$ of the test star is isotropic, the Fokker-Planck equation becomes

$$\frac{dP}{dt} = \frac{\partial}{\partial v} \left[D_{\parallel}(v) \left(\frac{\partial P}{\partial v} + \beta m P v \right) \right], \quad (59)$$

or

$$\frac{dP}{dt} = \frac{1}{v^2} \frac{\partial}{\partial v} \left[v^2 D_{\parallel}(v) \left(\frac{\partial P}{\partial v} + \beta m P v \right) \right]. \quad (60)$$

It can be obtained from an effective Langevin equation

$$\frac{d\mathbf{v}}{dt} = -\nabla\Phi - D_{\parallel}(v) \beta m \mathbf{v} + \frac{1}{2} \frac{\partial D_{\parallel}}{\partial v} + \sqrt{2D_{\parallel}(v)} \mathbf{R}(t), \quad (61)$$

where $\mathbf{R}(t)$ is a Gaussian white noise satisfying $\langle \mathbf{R}(t) \rangle = \mathbf{0}$ and $\langle R^\mu(t) R^\nu(t') \rangle = \delta^{\mu\nu} \delta(t - t')$. Since the diffusion coefficient depends on the velocity, the noise is multiplicative. The force acting on the star consists of two parts, a part derivable from a smoothed distribution of matter $-\nabla\Phi$, and a residual random part attributable to the fluctuations of that distribution. In turn, the random part can be described by a Gaussian white noise multiplied by a velocity-dependent factor plus a friction force equal to the friction by polarization $\mathbf{F}_{\text{pol}} = -D_{\parallel}(v) \beta m \mathbf{v}$ and a spurious drift $\mathbf{F}_{\text{spurious}} = (1/2)(\partial D_{\parallel}/\partial v)$ attributable to the multiplicative noise (we have used the Stratonovich representation). If we neglect the velocity dependence of the diffusion coefficient, we recover the results of Chandrasekhar (1943a, 1949) obtained in his Brownian theory.

3.3. The diffusion tensor for isothermal systems

When the velocity distribution of the field stars is given by the Maxwellian distribution (Eq. (51)), the diffusion tensor (Eq. (40)) may be calculated as follows. If we introduce the representation

$$\delta(x) = \int_{-\infty}^{+\infty} e^{ix} \frac{dt}{2\pi}, \quad (62)$$

for the δ -function in Eq. (40), the diffusion tensor can be rewritten as

$$D^{\mu\nu} = \frac{1}{2}(2\pi)^6 \int_{-\infty}^{+\infty} dt \int d\mathbf{k} k^\mu k^\nu \hat{u}(k)^2 e^{i\mathbf{k}\cdot\mathbf{v}t} \hat{f}(k\mathbf{t}), \quad (63)$$

where \hat{f} is the three-dimensional Fourier transform of the velocity distribution. This equation can be directly obtained from Eq. (37) or Eq. (43) (see Appendix G). This shows that the auxiliary integration variable t in Eq. (63) represents the time. For the Maxwellian distribution (Eq. (51)), $\hat{f}(k\mathbf{t})$ is a Gaussian. If we perform the integration over t (which is the one-dimensional Fourier transform of a Gaussian), we find that the diffusion tensor can be expressed as

$$D^{\mu\nu} = \pi(2\pi)^3 \left(\frac{\beta m}{2\pi}\right)^{1/2} \rho m \int \frac{k^\mu k^\nu}{k} \hat{u}(k)^2 e^{-\beta m \frac{(k\cdot\mathbf{v})^2}{2k^2}} d\mathbf{k}. \quad (64)$$

Alternatively, this expression can be obtained from Eq. (40) by introducing a cartesian system of coordinates for \mathbf{v}_1 with the z -axis taken along the direction of \mathbf{k} , and performing the integration. With the notation $\hat{\mathbf{k}} = \mathbf{k}/k$, Eq. (64) can be rewritten as

$$D^{\mu\nu} = \pi(2\pi)^3 \left(\frac{\beta m}{2\pi}\right)^{1/2} \rho m \int_0^{+\infty} k^3 \hat{u}(k)^2 dk G^{\mu\nu} \left(\sqrt{\frac{\beta m}{2}} \mathbf{v} \right), \quad (65)$$

where

$$G^{\mu\nu}(\mathbf{x}) = \int \hat{k}^\mu \hat{k}^\nu e^{-\hat{\mathbf{k}}\cdot\mathbf{x}} d\hat{\mathbf{k}}. \quad (66)$$

We note that the potential of interaction only appears in a multiplicative constant that fixes the relaxation time (see below). Using Eq. (C.7), we get

$$D^{\mu\nu} = \left(\frac{3}{2\pi}\right)^{1/2} \frac{2\rho m G^2 \ln \Lambda}{v_m} G^{\mu\nu} \left(\sqrt{\frac{3}{2}} \frac{\mathbf{v}}{v_m} \right), \quad (67)$$

where $\ln \Lambda$ is the Coulombian logarithm (Eq. (30)) and $v_m^2 = 3/(\beta m)$ is the mean square velocity of the field stars. The diffusion tensor may be written as

$$D^{\mu\nu} = \frac{2v_m^2}{3t_R} G^{\mu\nu} \left(\sqrt{\frac{3}{2}} \frac{\mathbf{v}}{v_m} \right), \quad (68)$$

where t_R is the local relaxation time defined below in Eq. (76). This relation emphasizes the scaling $D \sim v_m^2/t_R$.

Introducing a spherical system of coordinates with the z -axis in the direction of \mathbf{x} , we can easily compute the integral in Eq. (66). Then, we can write the normalized diffusion tensor in the form

$$G^{\mu\nu} = \left(G_{\parallel} - \frac{1}{2} G_{\perp} \right) \frac{x^\mu x^\nu}{x^2} + \frac{1}{2} G_{\perp} \delta^{\mu\nu}, \quad (69)$$

where

$$G_{\parallel} = \frac{2\pi^{3/2}}{x} G(x), \quad G_{\perp} = \frac{2\pi^{3/2}}{x} [\text{erf}(x) - G(x)], \quad (70)$$

with

$$G(x) = \frac{2}{\sqrt{\pi}} \frac{1}{x^2} \int_0^x t^2 e^{-t^2} dt = \frac{1}{2x^2} \left[\text{erf}(x) - \frac{2x}{\sqrt{\pi}} e^{-x^2} \right]. \quad (71)$$

The error function is defined by

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (72)$$

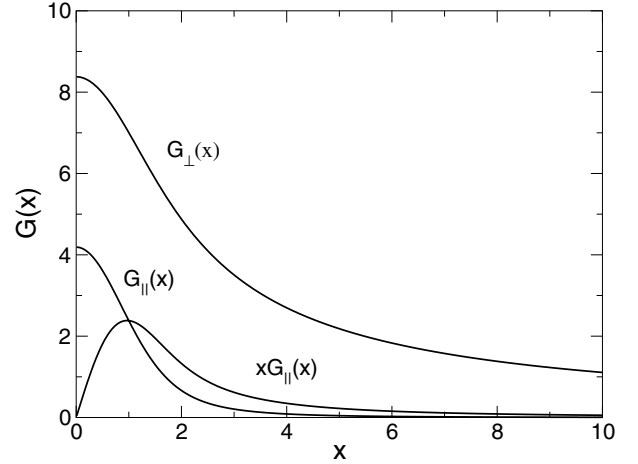


Fig. 1. Normalized diffusion coefficients $G_{\parallel}(x)$, $G_{\perp}(x)$ and friction force $xG_{\parallel}(x)$ for a thermal bath. The friction is maximum for $x \simeq 0.97$, i.e. when the velocity of the test star is approximately equal to 0.792 times the rms velocity of the field stars (its value is $(xG_{\parallel})_{\max} = 2.383$).

We have the asymptotic behaviours

$$G_{\parallel}(0) = \frac{4\pi}{3}, \quad G_{\perp}(0) = \frac{8\pi}{3}, \quad (73)$$

$$G_{\parallel}(x) \sim_{+\infty} \frac{\pi^{3/2}}{x^3}, \quad G_{\perp}(x) \sim_{+\infty} \frac{2\pi^{3/2}}{x}. \quad (74)$$

We note that $G^{\mu\nu}(\mathbf{x}) \simeq G_{\parallel}(0)\delta^{\mu\nu} \simeq \frac{1}{2}G_{\perp}(0)\delta^{\mu\nu}$ when $|\mathbf{x}| \rightarrow 0$. The longitudinal and transverse diffusion coefficients and the friction force are plotted in Fig. 1.

3.4. The relaxation time

We can use the preceding results to estimate the relaxation time of the velocity distribution of the test particle towards the Maxwellian distribution (thermalization). If we set $\mathbf{x} = \sqrt{\beta m/2}\mathbf{v}$, the Fokker-Planck equation (Eq. (55)) can be rewritten as

$$\frac{dP}{dt} = \frac{1}{t_R} \frac{\partial}{\partial x^\mu} \left[G^{\mu\nu}(\mathbf{x}) \left(\frac{\partial P}{\partial x^\nu} + 2Px^\nu \right) \right], \quad (75)$$

where t_R is the local relaxation time

$$t_R = \frac{1}{3} \left(\frac{2\pi}{3} \right)^{1/2} \frac{v_m^3}{\rho m G^2 \ln \Lambda}. \quad (76)$$

The prefactor is equal to 0.482 but, of course, this numerical factor may vary depending on the definition of the relaxation time. The relaxation time is inversely proportional to the local density $\rho(\mathbf{r})$. Therefore, the relaxation time is shorter in regions of high density (core) and longer in regions of low density (halo). Introducing the dynamical time $t_D = \lambda_J/v_m$, we get

$$t_R \sim \frac{n\lambda_J^3}{\ln \Lambda} t_D \sim \frac{N}{\ln N} t_D. \quad (77)$$

We note that the relaxation time of a test particle in a bath is of the same order as the relaxation time of the system as a whole (see Sect. 2.7).

We can also get an estimate of the relaxation time by the following argument (Spitzer 1987). If the diffusion coefficient were constant, the typical velocity of the test star (in one spatial direction) would increase as $\langle (\Delta v)^2 \rangle / 3 \sim 2D_{\parallel}t$. The relaxation

time t_r is the typical time at which the typical velocity of the test star has reached its equilibrium value $\langle v^2 \rangle_{(+\infty)} = 3/(m\beta) = v_m^2$ so that $\langle (\Delta v)^2 \rangle(t_r) = \langle v^2 \rangle_{(+\infty)}$. Since D_{\parallel} depends on v , the description of the diffusion process is more complex. However, the formula $t_r = v_m^2/[6D_{\parallel}(v_m)]$ resulting from the previous arguments with $D_{\parallel} = D_{\parallel}(v_m)$ should provide a good estimate of the relaxation time. Using Eq. (67) and comparing with Eq. (76) we obtain $t_r = K_3 t_R$, where $K_3 = 1/[4G_{\parallel}(\sqrt{3/2})]$. Numerically, $K_3 = 0.13587547\dots$

Finally, we can estimate the relaxation time by $t'_r = \xi^{-1}$ where ξ is the friction coefficient. Using the Einstein relation $\xi = D_{\parallel}\beta m$ (see Eq. (57)) with $D_{\parallel} = D_{\parallel}(v_m)$ we find that $t'_r = 2t_r$.

3.5. The Rosenbluth potentials

It is possible to obtain simple expressions of the diffusion and friction coefficients for any isotropic distribution of the bath. If we start from the expression of the Vlasov-Landau equation (Eq. (28)), we find that the Fokker-Planck equation (Eq. (35)) can be written as

$$\frac{dP}{dt} = \frac{\partial}{\partial v^{\mu}} \int K^{\mu\nu} \left(f_1 \frac{\partial P}{\partial v^{\nu}} - P \frac{\partial f_1}{\partial v_1^{\nu}} \right) dv_1, \quad (78)$$

where $K^{\mu\nu}$ is defined by Eq. (29). Comparing Eq. (78) with Eqs. (36) and (44), the diffusion and friction coefficients are given by

$$D^{\mu\nu} = \int K^{\mu\nu} f_1 dv_1 = A \int f_1 \frac{w^2 \delta^{\mu\nu} - w^{\mu} w^{\nu}}{w^3} dv_1, \quad (79)$$

$$\begin{aligned} F_{\text{friction}}^{\mu} &= 2F_{\text{pol}}^{\mu} = 2 \int K^{\mu\nu} \frac{\partial f_1}{\partial v_1^{\nu}} dv_1 \\ &= 2 \int \frac{\partial K^{\mu\nu}}{\partial v^{\nu}} f_1 dv_1 = -4A \int f_1 \frac{w^{\mu}}{w^3} dv_1. \end{aligned} \quad (80)$$

Using the identities

$$K^{\mu\nu} = A \frac{\partial^2 w}{\partial v^{\mu} \partial v^{\nu}} \quad (81)$$

and

$$\frac{\partial K^{\mu\nu}}{\partial v^{\nu}} = -2A \frac{w^{\mu}}{w^3} = 2A \frac{\partial}{\partial v^{\mu}} \left(\frac{1}{w} \right), \quad (82)$$

the coefficients of diffusion and friction can be rewritten as

$$D^{\mu\nu} = A \frac{\partial^2 g}{\partial v^{\mu} \partial v^{\nu}}(v), \quad (83)$$

$$F_{\text{friction}}^{\mu} = 2F_{\text{pol}}^{\mu} = 4A \frac{\partial h}{\partial v}(v), \quad (84)$$

where

$$g(v) = \int f(v_1) |\mathbf{v} - \mathbf{v}_1| dv_1, \quad h(v) = \int \frac{f(v_1)}{|\mathbf{v} - \mathbf{v}_1|} dv_1, \quad (85)$$

are the so-called Rosenbluth potentials (Rosenbluth et al. 1957). In terms of these potentials, the Fokker-Planck equations (Eqs. (36) and (44)) may be rewritten as

$$\frac{dP}{dt} = A \frac{\partial}{\partial v^{\mu}} \left[\frac{\partial^2 g}{\partial v^{\mu} \partial v^{\nu}} \frac{\partial P}{\partial v^{\nu}} - 2P \frac{\partial h}{\partial v^{\mu}} \right], \quad (86)$$

$$\frac{dP}{dt} = A \left[\frac{\partial^2}{\partial v^{\mu} \partial v^{\nu}} \left(\frac{\partial^2 g}{\partial v^{\mu} \partial v^{\nu}} P \right) - 4 \frac{\partial}{\partial v^{\mu}} \left(P \frac{\partial h}{\partial v^{\mu}} \right) \right]. \quad (87)$$

If the field particles have an isotropic velocity distribution, the Rosenbluth potentials take the particularly simple form (see e.g. Binney & Tremaine 2008)

$$h(v) = 4\pi \left[\frac{1}{v} \int_0^v f(v_1) v_1^2 dv_1 + \int_v^{+\infty} f(v_1) v_1 dv_1 \right], \quad (88)$$

$$\begin{aligned} g(v) &= \frac{4\pi v}{3} \left[\int_0^v \left(3v_1^2 + \frac{v_1^4}{v^2} \right) f(v_1) dv_1 \right. \\ &\quad \left. + \int_v^{+\infty} \left(\frac{3v_1^3}{v} + vv_1 \right) f(v_1) dv_1 \right]. \end{aligned} \quad (89)$$

When $g = g(v)$, the diffusion tensor (Eq. (83)) can be put in the form of Eq. (56) with

$$D_{\parallel} = A \frac{d^2 g}{dv^2}, \quad D_{\perp} = 2A \frac{1}{v} \frac{dg}{dv}. \quad (90)$$

Using Eq. (89), we obtain

$$D_{\parallel} = \frac{8\pi}{3} A \frac{1}{v} \left[\int_0^v \frac{v_1^4}{v^2} f(v_1) dv_1 + v \int_v^{+\infty} v_1 f(v_1) dv_1 \right], \quad (91)$$

$$\begin{aligned} D_{\perp} &= \frac{8\pi}{3} A \frac{1}{v} \left[\int_0^v \left(3v_1^2 - \frac{v_1^4}{v^2} \right) f(v_1) dv_1 \right. \\ &\quad \left. + 2v \int_v^{+\infty} v_1 f(v_1) dv_1 \right]. \end{aligned} \quad (92)$$

On the other hand, when $h = h(v)$, the friction force (Eq. (84)) can be written as

$$F_{\text{friction}}^{\mu} = 2F_{\text{pol}}^{\mu} = 4A \frac{1}{v} \frac{dh}{dv} v. \quad (93)$$

Using Eq. (88), we get

$$F_{\text{friction}}^{\mu} = 2F_{\text{pol}}^{\mu} = -16\pi A \frac{v}{v^3} \int_0^v f(v_1) v_1^2 dv_1. \quad (94)$$

This expression can be obtained directly from Eq. (84) by noting (Binney & Tremaine 2008) that $h(v)$ in Eq. (85) is similar to the gravitational potential $\Phi(\mathbf{r})$ produced by a distribution of mass $\rho(\mathbf{r})$, where \mathbf{v} plays the role of \mathbf{r} and $f(v)$ the role of $\rho(\mathbf{r})$. Therefore, if $f(v)$ is isotropic, Eq. (94) is equivalent to the expression of the gravitational field $\mathbf{F} = -GM(r)\mathbf{r}/r^3$ produced by a spherically symmetric distribution of matter, where $M(r)$ is the mass within the sphere of radius r (Gauss theorem). This formula shows that the friction is due only to field stars with a velocity less than the velocity of the test star. This observation was first made by Chandrasekhar (1943a).

The previous expressions for the diffusion and friction coefficients are valid for any isotropic distribution of the field particles. When $f(v)$ is the Maxwell distribution, we recover the results of Sect. 3.2. If we substitute Eqs. (91)–(93) into Eq. (36) or (44), we get a Fokker-Planck equation describing the evolution of a test particle in a bath with a prescribed distribution $f(v)$ ²⁴. Alternatively, if we come back to the original Landau kinetic equation (Eq. (28)) which can be written as

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial v^{\mu}} \left(D^{\mu\nu} \frac{\partial f}{\partial v^{\nu}} - f F_{\text{pol}}^{\mu} \right), \quad (95)$$

²⁴ Actually, this description is not self-consistent since the distribution $f(v)$ is not steady on the relaxation time t_R unless it is the Maxwell distribution.

assume an isotropic velocity distribution, use $D^{\mu\nu}v^\nu = D_{\parallel}v^\mu$, and substitute the general expressions of the diffusion and friction coefficients (Eqs. (90)–(94)) with $f = f(v, t)$, we obtain the integro-differential equation

$$\frac{\partial f}{\partial t} = A \frac{1}{v^2} \frac{\partial}{\partial v} \left[v^2 \left(\frac{\partial^2 g}{\partial v^2} \frac{\partial f}{\partial v} - 2f \frac{\partial h}{\partial v} \right) \right], \quad (96)$$

or, in more explicit form

$$\frac{\partial f}{\partial t} = 8\pi A \frac{1}{v^2} \frac{\partial}{\partial v} \left[\frac{1}{3} \frac{\partial f}{\partial v} \left(\frac{1}{v} \int_0^v v_1^4 f(v_1, t) dv_1 + v^2 \int_v^{+\infty} v_1 f(v_1, t) dv_1 \right) + f \int_0^v f(v_1, t) v_1^2 dv_1 \right], \quad (97)$$

describing the evolution of the system as a whole. Under this form, Eq. (97) applies to an artificial infinite homogeneous distribution of stars. This equation has been studied by King (1960) in his investigations on the evaporation of globular clusters. Within the local approximation, Eq. (97) also represents the simplification of the collision operator that occurs in the r.h.s. of the Vlasov-Landau equation (Eq. (28)) when the velocity distribution of the stars is isotropic. In that case, we must restore the space variable and the advection term in Eq. (97). From this equation, implementing an adiabatic approximation, we can derive the orbit-averaged Fokker-Planck equation which was used by Hénon (1961) and Cohn (1980) to study the collisional evolution of globular clusters. It reads

$$\frac{\partial q}{\partial \epsilon} \frac{\partial f}{\partial t} - \frac{\partial q}{\partial t} \frac{\partial f}{\partial \epsilon} = 8\pi A \frac{\partial}{\partial \epsilon} \left[f \int_{-\infty}^{\epsilon} f_1 \frac{\partial q_1}{\partial \epsilon_1} d\epsilon_1 + \frac{\partial f}{\partial \epsilon} \left\{ \int_{-\infty}^{\epsilon} f_1 q_1 d\epsilon_1 + q \int_{\epsilon}^{+\infty} f_1 d\epsilon_1 \right\} \right], \quad (98)$$

where

$$q(\epsilon, t) = \frac{1}{3} \int_0^{r_{\max}} [2(\epsilon - \Phi(r, t))]^{3/2} r^2 dr, \quad (99)$$

is proportional to the phase space volume available to stars with an energy less than ϵ .

3.6. Comparison with the two-body encounters theory

In the previous sections, we have derived the standard kinetic equations of stellar systems from the the Liouville equation by using the BBGKY hierarchy. In classical textbooks of astrophysics (Spitzer 1987; Heggie & Hut 2003; Binney & Tremaine 2008), these equations are derived in a different manner. One usually starts from the Fokker-Planck equation (Eq. (44)), makes a local approximation, and evaluates the diffusion tensor $\langle \Delta v^\mu \Delta v^\nu \rangle$ and the friction force $\langle \Delta v^\mu \rangle$ by considering the mean effect of a succession of two-body encounters. This two-body encounters theory was pioneered by Chandrasekhar (1942) and further developed by Rosenbluth et al. (1957). This approach leads directly to the expressions in Eqs. (83) and (84) of the diffusion and friction coefficients of a test star in a bath of field stars. These expressions are then substituted in the Fokker-Planck equation (Eq. (44)). Finally, arguing that the field stars and the test star should evolve in the same manner, the Fokker-Planck equation is transformed into the integro-differential equation (Eq. (97)) describing the evolution of the system as a whole (King 1960). When an adiabatic approximation is implemented (Hénon 1961), one finally obtains the orbit-averaged Fokker-Planck equation (Eq. (98)).

In this paper we have proceeded the other way round. Starting from the Liouville equation, using the BBGKY hierarchy, making a local approximation, and neglecting strong collisions, we have derived the Vlasov-Landau equation (Eq. (28)) describing the evolution of the system as a whole. Then, making a bath approximation²⁵, we have obtained the Fokker-Planck equation in the form of Eq. (36) with the diffusion and friction coefficients given by Eqs. (79) and (80). This equation can then be transformed into Eq. (44) with Eqs. (83) and (84). Our approach emphasizes the importance of the Landau equation in the kinetic theory of stellar systems, while this equation does not appear in the works of Chandrasekhar (1942), Rosenbluth et al. (1957), King (1960), and Hénon (1961), nor in the standard textbooks of stellar dynamics by Spitzer (1987), Heggie & Hut (2003), and Binney & Tremaine (2008).

Actually, the kinetic equation derived by these authors is equivalent to the Landau equation²⁶, but it is written in a different form. They write the Fokker-Planck equation in the form of Eq. (44) with the diffusion coefficient placed after the second derivative ($\partial^2 D$) while the Landau equation (Eq. (28)) is related to the Fokker-Planck equation (Eq. (36)) in which the diffusion coefficient is inserted between the first derivatives ($\partial D \partial$). This difference is important on a physical point of view for two reasons. First, the Landau equation isolates the friction by polarization F_{pol} while the equation derived by Chandrasekhar (1942) and Rosenbluth et al. (1957) involves the total friction F_{friction} . Secondly, the Landau equation has a nice symmetric structure from which we can immediately deduce all the conservation laws of the system (conservation of mass, energy, impulse, and angular momentum) and the H -theorem for the Boltzmann entropy (Baesescu 2000). These properties are less apparent in the equations derived by King (1960) and Hénon (1961) for the evolution of the system as a whole. It is interesting to note that the symmetric structure of the kinetic equation was not realized by early stellar dynamicists while the Landau equation was known long before in plasma physics.

Finally, the approach based on the Liouville equation and on the BBGKY hierarchy is more rigorous than the two-body encounters theory because it relaxes the assumption of locality and does not produce any divergence at large scales²⁷. It leads to the generalized Landau equation (Eq. (24)) that is perfectly well-behaved at large scales contrary to the Vlasov-Landau equation (Eq. (28)) in which a large-scale cut-off has to be introduced in order to avoid a divergence. This divergence is due to the long-range nature of the gravitational potential which precludes a rigorous application of the two-body encounters theory that is valid for potentials with short-range interactions. Actually, the two-body encounters theory is marginally applicable to the gravitational force (it only generates a weak logarithmic divergence at large scales) and this is why it is successful in practice. The generalized Landau equation (Eq. (24)) represents a conceptual improvement of the Vlasov-Landau equation (Eq. (28)) because it goes beyond the local approximation and takes fully into

²⁵ In the BBGKY hierarchy, this amounts to singling out a particular test star and assuming that the distribution of the other stars is fixed.

²⁶ It is important to emphasize, however, that the two-body encounters theory of Chandrasekhar (1942) can take strong collisions into account. Therefore, there is no divergence at small impact parameters and the gravitational Landau length λ_L appears naturally in the Coulombian logarithm $\ln \Lambda$, contrary to the Landau theory which ignores strong collisions (see Appendix H).

²⁷ It is moreover applicable with almost no modification to other systems with long-range interactions for which a two-body encounters theory is not justified (Chavanis 2010).

account the spatial inhomogeneity of the system. Unfortunately, this equation is very complicated to be of much practical use. It can, however, be simplified by using angle-action variables as we show in the next section.

4. Kinetic equations with angle-action variables

4.1. Adiabatic approximation

In order to deal with spatially inhomogeneous systems, it is convenient to introduce angle-action variables (Goldstein 1956; Binney & Tremaine 2008). Angle-action variables have been used by many authors in astrophysics in order to solve dynamical stability problems (Kalnajs 1977; Goodman 1988; Weinberg 1991; Pichon & Cannon 1997; Valageas 2006a) or to compute the diffusion and friction coefficients of a test star in a cluster (Lynden-Bell & Kalnajs 1972; Tremaine & Weinberg 1984; Binney & Lacey 1988; Weinberg 1998, 2001; Nelson & Tremaine 1999; Pichon & Aubert 2006; Valageas 2006b; Chavanis 2007, 2010). By construction, the Hamiltonian H in angle and action variables depends only on the actions $\mathbf{J} = (J_1, \dots, J_d)$ that are constants of the motion; the conjugate coordinates $\mathbf{w} = (w_1, \dots, w_d)$ are called the angles (see Appendix B). Therefore, any distribution of the form $f = f(\mathbf{J})$ is a steady state of the Vlasov equation. According to the Jeans theorem, this is not the general form of Vlasov steady states. However, if the potential is regular, for all practical purposes, any time-independent solution of the Vlasov equation may be represented by a distribution of the form $f = f(\mathbf{J})$ (strong Jeans theorem).

We shall assume that the system has reached a QSS described by a distribution $f = f(\mathbf{J})$ as a result of a violent collisionless relaxation involving only mean field effects. Because of finite N effects, the distribution function f slowly evolves in time. Finite N effects are taken into account in the collision operator appearing in the r.h.s. of Eq. (24). Since this term is of order $1/N$, the effect of collisions (granularities, finite N effects, correlations) is a very slow process that takes place on a relaxation timescale $t_R \sim (N/\ln N)t_D$ (see Sect. 2.7). Therefore, there is a timescale separation between the dynamical time t_D that is the timescale during which the system reaches a steady state of the Vlasov equation through phase mixing and violent collisionless relaxation, and the collisional relaxation time t_R that is the timescale during which the system reaches an almost isothermal distribution because of finite N effects.

Because of this timescale separation, the distribution function is stationary on the dynamical timescale. It will evolve through a sequence of QSSs that are steady states of the Vlasov equation, depending only on the actions \mathbf{J} , slowly changing in time as a result of the cumulative effect of encounters (finite N effects). Indeed, the system re-adjusts itself dynamically at each step of the collisional process. The distribution function averaged over a short dynamical timescale can be approximated by

$$\langle f(\mathbf{r}, \mathbf{v}, t) \rangle \simeq f(\mathbf{J}, t). \quad (100)$$

Therefore, the distribution function is a function $f = f(\mathbf{J}, t)$ of the actions alone that slowly evolves in time under the effect of collisions. This is similar to an adiabatic approximation. The system is approximately in mechanical equilibrium at each stage of the dynamics and the collisions slowly drive it towards an almost isothermal distribution, corresponding to a quasi-thermodynamical equilibrium state.

4.2. Evolution of the system as a whole: a Landau-type equation with angle-action variables

Introducing angle-action variables \mathbf{w} and \mathbf{J} , the generalized Landau equation (Eq. (24)) becomes²⁸

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial \mathbf{J}} \cdot \int_0^{+\infty} d\tau \int d\mathbf{w}_1 d\mathbf{J}_1 F(1 \rightarrow 0) G(t, t - \tau) \times \left[F(1 \rightarrow 0) \cdot \frac{\partial}{\partial \mathbf{J}} + F(0 \rightarrow 1) \cdot \frac{\partial}{\partial \mathbf{J}_1} \right] f(\mathbf{J}, t) \frac{f}{m}(\mathbf{J}_1, t), \quad (101)$$

with

$$F(1 \rightarrow 0) = -m \frac{\partial u}{\partial \mathbf{w}} \left[\mathbf{r}(\mathbf{J}, \mathbf{w}) - \mathbf{r}(\mathbf{J}_1, \mathbf{w}_1) \right]. \quad (102)$$

To obtain Eq. (101), we have averaged Eq. (24) over \mathbf{w} (to simplify the expressions, the average $\langle \cdot \rangle = (2\pi)^{-3} \int d\mathbf{w}$ is implicit), written the scalar products as Poisson brackets, and used the invariance of the Poisson brackets and of the phase space volume element on a change of canonical variables. Introducing the Fourier transform of the potential with respect to the angles

$$A_{k, k_1}(\mathbf{J}, \mathbf{J}_1) = \frac{1}{(2\pi)^{2d}} \int u \left[\mathbf{r}(\mathbf{J}, \mathbf{w}) - \mathbf{r}(\mathbf{J}_1, \mathbf{w}_1) \right] \times e^{-i(k \cdot \mathbf{w} - k_1 \cdot \mathbf{w}_1)} d\mathbf{w} d\mathbf{w}_1, \quad (103)$$

so that

$$u \left[\mathbf{r}(\mathbf{J}, \mathbf{w}) - \mathbf{r}(\mathbf{J}_1, \mathbf{w}_1) \right] = \sum_{k, k_1} A_{k, k_1}(\mathbf{J}, \mathbf{J}_1) e^{i(k \cdot \mathbf{w} - k_1 \cdot \mathbf{w}_1)}, \quad (104)$$

we get

$$F(1 \rightarrow 0) = -im \sum_{k, k_1} A_{k, k_1}(\mathbf{J}, \mathbf{J}_1) k e^{i(k \cdot \mathbf{w} - k_1 \cdot \mathbf{w}_1)}. \quad (105)$$

Substituting this expression in Eq. (101), we obtain

$$\frac{\partial f}{\partial t} = -m \frac{\partial}{\partial \mathbf{J}} \cdot \int_0^{+\infty} d\tau \int d\mathbf{w}_1 d\mathbf{J}_1 \sum_{k, k_1} \sum_{l, l_1} A_{k, k_1}(\mathbf{J}, \mathbf{J}_1) \times k e^{i(k \cdot \mathbf{w} - k_1 \cdot \mathbf{w}_1)} G(t, t - \tau) \left[A_{l, l_1}(\mathbf{J}, \mathbf{J}_1) l e^{i(l \cdot \mathbf{w} - l_1 \cdot \mathbf{w}_1)} \cdot \frac{\partial}{\partial \mathbf{J}} + A_{l_1, l}(\mathbf{J}_1, \mathbf{J}) l_1 e^{i(l_1 \cdot \mathbf{w}_1 - l \cdot \mathbf{w})} \cdot \frac{\partial}{\partial \mathbf{J}_1} \right] f(\mathbf{J}, t) f(\mathbf{J}_1, t). \quad (106)$$

With angle-action variables, the equations of motion of a star determined by the mean field take the very simple form (see Appendix B)

$$\begin{aligned} \mathbf{J}(t - \tau) &= \mathbf{J}(t) = \mathbf{J}, \\ \mathbf{w}(t - \tau) &= \mathbf{w}(t) - \boldsymbol{\Omega}(\mathbf{J}, t)\tau = \mathbf{w} - \boldsymbol{\Omega}(\mathbf{J}, t)\tau, \end{aligned} \quad (107)$$

where $\boldsymbol{\Omega}(\mathbf{J}, t)$ is the angular frequency of the orbit with action \mathbf{J} . As explained previously, we have neglected the variation of the mean field on a timescale of the order of the dynamical time so it is considered to be frozen when we compute the stellar trajectories (adiabatic or Bogoliubov assumption). Substituting these

²⁸ The adiabatic assumption (Eq. (100)) is consistent with the Bogoliubov ansatz of kinetic theory. Since the correlation time of the fluctuations is of the order of the dynamical time or shorter, we can freeze the distribution function at time t to compute the integral over τ (see Sect. 2.4). This distribution function, which is a steady state of the Vlasov equation, defines a set of angle-action variables \mathbf{w} and \mathbf{J} that we can use to perform the integral over τ . Then, the distribution function $f(\mathbf{J}, t)$ evolves with time on a longer timescale according to Eq. (101).

relations in Eq. (106) and making the transformations $\mathbf{l} \rightarrow -\mathbf{l}$ and $\mathbf{l}_1 \rightarrow -\mathbf{l}_1$ in the second term (friction term), we obtain successively

$$\begin{aligned} \frac{\partial f}{\partial t} = & -m \frac{\partial}{\partial \mathbf{J}} \cdot \int_0^{+\infty} d\tau \int d\mathbf{w}_1 d\mathbf{J}_1 \sum_{\mathbf{k}, \mathbf{k}_1} \sum_{\mathbf{l}, \mathbf{l}_1} A_{\mathbf{k}, \mathbf{k}_1}(\mathbf{J}, \mathbf{J}_1) \\ & \times \mathbf{k} e^{i(\mathbf{k}\cdot\mathbf{w} - \mathbf{k}_1\cdot\mathbf{w}_1)} e^{i(\mathbf{l}\cdot\mathbf{w}(\tau) - \mathbf{l}_1\cdot\mathbf{w}_1(\tau))} \\ & \times \left[A_{\mathbf{l}, \mathbf{l}_1}(\mathbf{J}, \mathbf{J}_1) \mathbf{l} \cdot \frac{\partial}{\partial \mathbf{J}} - A_{-\mathbf{l}_1, -\mathbf{l}}(\mathbf{J}_1, \mathbf{J}) \mathbf{l}_1 \cdot \frac{\partial}{\partial \mathbf{J}_1} \right] \\ & \times f(\mathbf{J}, t) f(\mathbf{J}_1, t) \end{aligned} \quad (108)$$

and

$$\begin{aligned} \frac{\partial f}{\partial t} = & -m \frac{\partial}{\partial \mathbf{J}} \cdot \int_0^{+\infty} d\tau \int d\mathbf{w}_1 d\mathbf{J}_1 \sum_{\mathbf{k}, \mathbf{k}_1} \sum_{\mathbf{l}, \mathbf{l}_1} A_{\mathbf{k}, \mathbf{k}_1}(\mathbf{J}, \mathbf{J}_1) \\ & \times \mathbf{k} e^{i(\mathbf{k}\cdot\mathbf{w} - \mathbf{k}_1\cdot\mathbf{w}_1)} e^{i(\mathbf{l}\cdot\mathbf{w} - \mathbf{l}_1\cdot\mathbf{w}_1)} e^{-i(\mathbf{l}\cdot\boldsymbol{\Omega}(\mathbf{J}, t) - \mathbf{l}_1\cdot\boldsymbol{\Omega}(\mathbf{J}_1, t))\tau} \\ & \times \left[A_{\mathbf{l}, \mathbf{l}_1}(\mathbf{J}, \mathbf{J}_1) \mathbf{l} \cdot \frac{\partial}{\partial \mathbf{J}} - A_{-\mathbf{l}_1, -\mathbf{l}}(\mathbf{J}_1, \mathbf{J}) \mathbf{l}_1 \cdot \frac{\partial}{\partial \mathbf{J}_1} \right] \\ & \times f(\mathbf{J}, t) f(\mathbf{J}_1, t). \end{aligned} \quad (109)$$

It is easy to establish that

$$A_{\mathbf{k}_1, \mathbf{k}}(\mathbf{J}_1, \mathbf{J}) = A_{-\mathbf{k}, -\mathbf{k}_1}(\mathbf{J}, \mathbf{J}_1) = A_{\mathbf{k}, \mathbf{k}_1}(\mathbf{J}, \mathbf{J}_1)^*. \quad (110)$$

Therefore, the kinetic equation can be rewritten as

$$\begin{aligned} \frac{\partial f}{\partial t} = & -m \frac{\partial}{\partial \mathbf{J}} \cdot \int_0^{+\infty} d\tau \int d\mathbf{w}_1 d\mathbf{J}_1 \sum_{\mathbf{k}, \mathbf{k}_1} \sum_{\mathbf{l}, \mathbf{l}_1} A_{\mathbf{k}, \mathbf{k}_1}(\mathbf{J}, \mathbf{J}_1) \\ & \times \mathbf{k} e^{i(\mathbf{k}+\mathbf{l})\cdot\mathbf{w}} e^{-i(\mathbf{k}_1+\mathbf{l}_1)\cdot\mathbf{w}_1} e^{-i(\mathbf{l}\cdot\boldsymbol{\Omega}(\mathbf{J}, t) - \mathbf{l}_1\cdot\boldsymbol{\Omega}(\mathbf{J}_1, t))\tau} \\ & \times A_{\mathbf{l}, \mathbf{l}_1}(\mathbf{J}, \mathbf{J}_1) \left(\mathbf{l} \cdot \frac{\partial}{\partial \mathbf{J}} - \mathbf{l}_1 \cdot \frac{\partial}{\partial \mathbf{J}_1} \right) f(\mathbf{J}, t) f(\mathbf{J}_1, t). \end{aligned} \quad (111)$$

Integrating over \mathbf{w}_1 , and recalling that this expression has to be averaged over \mathbf{w} , we obtain

$$\begin{aligned} \frac{\partial f}{\partial t} = & (2\pi)^3 m \frac{\partial}{\partial \mathbf{J}} \cdot \int_0^{+\infty} d\tau \int d\mathbf{J}_1 \sum_{\mathbf{k}, \mathbf{k}_1} |A_{\mathbf{k}, \mathbf{k}_1}(\mathbf{J}, \mathbf{J}_1)|^2 \\ & \times \mathbf{k} e^{-i(\mathbf{k}\cdot\boldsymbol{\Omega}(\mathbf{J}, t) + \mathbf{k}_1\cdot\boldsymbol{\Omega}(\mathbf{J}_1, t))\tau} \\ & \times \left(\mathbf{k} \cdot \frac{\partial}{\partial \mathbf{J}} - \mathbf{k}_1 \cdot \frac{\partial}{\partial \mathbf{J}_1} \right) f(\mathbf{J}, t) f(\mathbf{J}_1, t). \end{aligned} \quad (112)$$

Making the transformation $\tau \rightarrow -\tau$, then $(\mathbf{k}, \mathbf{k}_1) \rightarrow (-\mathbf{k}, -\mathbf{k}_1)$, and adding the resulting expression to Eq. (112), we get

$$\begin{aligned} \frac{\partial f}{\partial t} = & \frac{1}{2} (2\pi)^3 m \frac{\partial}{\partial \mathbf{J}} \cdot \int_{-\infty}^{+\infty} d\tau \int d\mathbf{J}_1 \sum_{\mathbf{k}, \mathbf{k}_1} |A_{\mathbf{k}, \mathbf{k}_1}(\mathbf{J}, \mathbf{J}_1)|^2 \\ & \times \mathbf{k} e^{i(\mathbf{k}\cdot\boldsymbol{\Omega}(\mathbf{J}, t) - \mathbf{k}_1\cdot\boldsymbol{\Omega}(\mathbf{J}_1, t))\tau} \\ & \times \left(\mathbf{k} \cdot \frac{\partial}{\partial \mathbf{J}} - \mathbf{k}_1 \cdot \frac{\partial}{\partial \mathbf{J}_1} \right) f(\mathbf{J}, t) f(\mathbf{J}_1, t). \end{aligned} \quad (113)$$

Finally, using the identity (Eq. (C.11)), we obtain the kinetic equation

$$\begin{aligned} \frac{\partial f}{\partial t} = & \pi (2\pi)^3 m \frac{\partial}{\partial \mathbf{J}} \cdot \sum_{\mathbf{k}, \mathbf{k}_1} \int d\mathbf{J}_1 \mathbf{k} |A_{\mathbf{k}, \mathbf{k}_1}(\mathbf{J}, \mathbf{J}_1)|^2 \\ & \times \delta[\mathbf{k} \cdot \boldsymbol{\Omega}(\mathbf{J}, t) - \mathbf{k}_1 \cdot \boldsymbol{\Omega}(\mathbf{J}_1, t)] \\ & \times \left(\mathbf{k} \cdot \frac{\partial}{\partial \mathbf{J}} - \mathbf{k}_1 \cdot \frac{\partial}{\partial \mathbf{J}_1} \right) f(\mathbf{J}, t) f(\mathbf{J}_1, t). \end{aligned} \quad (114)$$

This kinetic equation was previously derived for systems with arbitrary long-range interactions in various dimensions of space

(Chavanis 2007, 2010) and it is here specifically applied to stellar systems. Since collective effects are neglected, this kinetic equation can be viewed as a Landau-type equation with angle-action variables describing the evolution of spatially inhomogeneous stellar systems. The collisional evolution of these systems is due to a condition of resonance $\mathbf{k} \cdot \boldsymbol{\Omega}(\mathbf{J}, t) = \mathbf{k}_1 \cdot \boldsymbol{\Omega}(\mathbf{J}_1, t)$ (with $(\mathbf{k}_1, \mathbf{J}_1) \neq (\mathbf{k}, \mathbf{J})$) encapsulated in the δ -function. This δ -function expresses the conservation of energy. It can be shown (Chavanis 2007) that the kinetic equation (Eq. (114)) conserves the mass $M = \int f d\mathbf{J}$ and the energy $E = \int f \epsilon(\mathbf{J}) d\mathbf{J}$, and monotonically increases the Boltzmann entropy $S = -\int (f/m) \ln(f/m) d\mathbf{J}$ (H -theorem). However, as explained in Sect. 2.7, this equation does not reach a steady state because of the absence of statistical equilibrium for stellar systems²⁹.

4.3. Relaxation of a star in a thermal bath: a Fokker-Planck equation with angle-action variables

Implementing a test particle approach as in Sect. 3, we find that the equation for $P(\mathbf{J}, t)$, the probability density of finding the test star with an action \mathbf{J} at time t , is

$$\begin{aligned} \frac{\partial P}{\partial t} = & \pi (2\pi)^3 m \frac{\partial}{\partial \mathbf{J}} \cdot \sum_{\mathbf{k}, \mathbf{k}_1} \int d\mathbf{J}_1 \mathbf{k} |A_{\mathbf{k}, \mathbf{k}_1}(\mathbf{J}, \mathbf{J}_1)|^2 \\ & \times \delta[\mathbf{k} \cdot \boldsymbol{\Omega}(\mathbf{J}) - \mathbf{k}_1 \cdot \boldsymbol{\Omega}(\mathbf{J}_1)] \\ & \times \left(\mathbf{k} \cdot \frac{\partial}{\partial \mathbf{J}} - \mathbf{k}_1 \cdot \frac{\partial}{\partial \mathbf{J}_1} \right) P(\mathbf{J}, t) f(\mathbf{J}_1). \end{aligned} \quad (115)$$

The angular frequency $\boldsymbol{\Omega}(\mathbf{J})$ is now a static function determined by the distribution $f(\mathbf{J})$ of the field stars. Equation (115) can be written in the form of a Fokker-Planck equation

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial J^\mu} \left(D^{\mu\nu} \frac{\partial P}{\partial J^\nu} - P F_{\text{pol}}^\mu \right), \quad (116)$$

involving a diffusion tensor

$$\begin{aligned} D^{\mu\nu} = & \pi (2\pi)^3 m \sum_{\mathbf{k}, \mathbf{k}_1} \int d\mathbf{J}_1 k^\mu k^\nu |A_{\mathbf{k}, \mathbf{k}_1}(\mathbf{J}, \mathbf{J}_1)|^2 \\ & \times \delta[\mathbf{k} \cdot \boldsymbol{\Omega}(\mathbf{J}) - \mathbf{k}_1 \cdot \boldsymbol{\Omega}(\mathbf{J}_1)] f(\mathbf{J}_1), \end{aligned} \quad (117)$$

and a friction by polarization

$$\begin{aligned} F_{\text{pol}} = & \pi (2\pi)^3 m \sum_{\mathbf{k}, \mathbf{k}_1} \int d\mathbf{J}_1 \mathbf{k} |A_{\mathbf{k}, \mathbf{k}_1}(\mathbf{J}, \mathbf{J}_1)|^2 \\ & \times \delta[\mathbf{k} \cdot \boldsymbol{\Omega}(\mathbf{J}) - \mathbf{k}_1 \cdot \boldsymbol{\Omega}(\mathbf{J}_1)] \mathbf{k}_1 \cdot \frac{\partial f}{\partial \mathbf{J}_1}(\mathbf{J}_1). \end{aligned} \quad (118)$$

Writing the Fokker-Planck equation in the usual form

$$\frac{\partial P}{\partial t} = \frac{\partial^2}{\partial J^\mu \partial J^\nu} (D^{\mu\nu} P) - \frac{\partial}{\partial J^\mu} (P F_{\text{friction}}^\mu), \quad (119)$$

with

$$D^{\mu\nu} = \frac{\langle \Delta J^\mu \Delta J^\nu \rangle}{2\Delta t}, \quad F_{\text{friction}} = \frac{\langle \Delta \mathbf{J} \rangle}{\Delta t}, \quad (120)$$

²⁹ For self-gravitating systems in lower dimensions of space, or for systems with long-range interactions with a smooth potential like the HMF model, a statistical equilibrium state exists. In this case, it can be shown that the Landau-type equation (Eq. (114)) relaxes towards the Boltzmann distribution on a timescale Nt_D provided there are enough resonances (see Chavanis 2007).

we find that the relation between the friction by polarization and the total friction is

$$F_{\text{friction}}^{\mu} = F_{\text{pol}}^{\mu} + \frac{\partial D^{\mu\nu}}{\partial J^{\nu}}. \quad (121)$$

Substituting Eqs. (117) and (118) in Eq. (121) and using an integration by parts, we find that the diffusion and friction coefficients are given by

$$\frac{\langle \Delta J^{\mu} \Delta J^{\nu} \rangle}{2\Delta t} = \pi(2\pi)^3 m \int dJ_1 f(J_1) \sum_{k, k_1} k^{\mu} k^{\nu} \times |A_{k, k_1}(J, J_1)|^2 \delta(\mathbf{k} \cdot \boldsymbol{\Omega}(J) - \mathbf{k}_1 \cdot \boldsymbol{\Omega}(J_1)), \quad (122)$$

$$\frac{\langle \Delta J \rangle}{\Delta t} = \pi(2\pi)^3 m \int dJ_1 f(J_1) \sum_{k, k_1} \mathbf{k} \left(\mathbf{k} \cdot \frac{\partial}{\partial \mathbf{J}} - \mathbf{k}_1 \cdot \frac{\partial}{\partial \mathbf{J}_1} \right) \times |A_{k, k_1}(J, J_1)|^2 \delta(\mathbf{k} \cdot \boldsymbol{\Omega}(J) - \mathbf{k}_1 \cdot \boldsymbol{\Omega}(J_1)). \quad (123)$$

These expressions can be obtained directly from the Hamiltonian equations of motion by expanding the trajectories of the stars in powers of $1/N$ in the limit $N \rightarrow +\infty$ (Valageas 2006a).

Let us assume that the field stars form a thermal bath with the Boltzmann distribution

$$f(J_1) = A e^{-\beta m \epsilon(J_1)}, \quad (124)$$

where $\epsilon(\mathbf{J})$ is the energy of a star in an orbit with action \mathbf{J} . As we have explained before, this distribution is not defined globally for a self-gravitating system. However, it holds approximately for stars with low energy³⁰. Using the identity $\partial \epsilon / \partial \mathbf{J} = \boldsymbol{\Omega}(\mathbf{J})$ (see Appendix B), we find that

$$\frac{\partial f_1}{\partial J_1} = -\beta m f(J_1) \boldsymbol{\Omega}(J_1). \quad (125)$$

Substituting this relation in Eq. (118), using the δ -function to replace $\mathbf{k}_1 \cdot \boldsymbol{\Omega}(J_1)$ by $\mathbf{k} \cdot \boldsymbol{\Omega}(J)$, and comparing the resulting expression with Eq. (117), we finally get

$$F_{\text{pol}}^{\mu} = -D^{\mu\nu}(\mathbf{J}) \beta m \Omega^{\nu}(\mathbf{J}), \quad (126)$$

which is the appropriate Einstein relation for our problem. For a thermal bath, using Eq. (126), the Fokker-Planck equation (Eq. (116)) can be written as

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial J^{\mu}} \left[D^{\mu\nu}(\mathbf{J}) \left(\frac{\partial P}{\partial J^{\nu}} + \beta m P \Omega^{\nu}(\mathbf{J}) \right) \right], \quad (127)$$

where $D^{\mu\nu}(\mathbf{J})$ is given by Eq. (117) with Eq. (124). Recalling that $\boldsymbol{\Omega}(\mathbf{J}) = \partial \epsilon / \partial \mathbf{J}$, this equation is similar to the Kramers equation in Brownian theory (Kramers 1940). This is a drift-diffusion equation describing the evolution of the distribution $P(\mathbf{J}, t)$ of the test star in an effective potential $U_{\text{eff}}(\mathbf{J}) = \epsilon(\mathbf{J})$ produced by the field stars. For $t \rightarrow +\infty$, the distribution of the test star relaxes towards the Boltzmann distribution (Eq. (124)). This takes place on a typical relaxation time $t_{\text{R}} \sim (N/\ln N)t_{\text{D}}$. Again, this is valid only in the part of the cluster where the Boltzmann distribution holds approximately.

³⁰ This description is also valid for self-gravitating systems in lower dimensions of space, or for other systems with long-range interactions for which a statistical equilibrium state exists (Chavanis 2007, 2010).

4.4. The impact theory and the wave theory

The approach developed in this paper has removed the problematic divergence at large scales that appears in the Vlasov-Landau equation when one makes a local approximation. However, this approach, which is based on a weak coupling approximation, yields a divergence at small scales because it does not take strong collisions into account. Stellar systems are therefore described by two complementary kinetic theories depending on the value of the impact parameter λ (see Appendix H). The two-body encounters theory, or impact theory, based on the Boltzmann equation or on the Fokker-Planck equation is appropriate to describe strong collisions when $\lambda \sim \lambda_{\text{L}}$. In this theory there is no divergence at small scales. However, this theory does not take spatial inhomogeneity and collective interactions between stars into account. As a result, it exhibits a logarithmic divergence at large scales. This implies that the integrand in the Boltzmann equation is valid only for impact parameters sufficiently smaller than the Jeans length ($\lambda \ll \lambda_{\text{J}}$). The wave theory based on the Landau (or Lenard-Balescu) equation written with angle-action variables is appropriate to describe spatial inhomogeneity and collective effects when $\lambda \sim \lambda_{\text{J}}$. In this theory, there is no divergence at large scales. However, this theory does not take into account the curvature of orbits at small impact parameters. As a result, it exhibits a logarithmic divergence at small scales. This implies that the integrand in the Landau (or Lenard-Balescu) equation with angle-action variables is valid only for wavelengths sufficiently larger than the Landau length ($\lambda \gg \lambda_{\text{L}}$). For $N \gg 1$, the range of validity of these two theories greatly overlaps in the intermediate region $\lambda_{\text{L}} \ll \lambda \ll \lambda_{\text{J}}$ corresponding to the domain of validity of the Vlasov-Landau equation. This equation describes only weak collisions and presents logarithmic divergences both at small and large scales. Combining the impact theory and the wave theory, we can motivate (but not rigorously justify) using the Vlasov-Landau equation with a small-scale cut-off at the Landau length and a large-scale cut-off at the Jeans length. The success of this equation is that the divergence of the integral over the impact parameter (or wavenumber) is only logarithmic so that its value depends only weakly on the choice of the cut-offs. However, a better kinetic equation is the Landau (or Lenard-Balescu) equation written with angle-action variables with a small-scale cut-off at the Landau length.

5. Conclusion

Starting from the Liouville equation, using a truncation of the BBGKY hierarchy at the order $1/N$, and neglecting strong collisions and collective effects, we have derived a kinetic equation (Eq. (24)) in physical space that can be viewed as a generalized Landau equation. This equation was previously derived by Kandrup (1981) using projection operator technics. A nice feature of this equation is that it does not present any divergence at large scales since the spatial inhomogeneity of the system and its finite extent are properly accounted for. When a local approximation is implemented, and a cut-off is introduced heuristically at the Jeans length, we recover the Vlasov-Landau equation (Eq. (27)) which is the standard equation of stellar dynamics. On the other hand, using angle-action variables, we have derived a Landau-type equation (Eq. (114)) for fully inhomogeneous stellar systems. We have also developed a test particle approach and derived the corresponding Fokker-Planck equations (Eqs. (39) and (115)) in position-velocity space and angle-action space respectively. Explicit expressions have been given for the diffusion and friction coefficients. We have distinguished the friction by

polarization from the total friction. A limitation of the approach presented here is that it neglects collective effects³¹. More general kinetic equations, corresponding to Lenard-Balescu-type equations taking spatial inhomogeneity and collective effects into account, have been derived recently by Heyvaerts (2010) from the Liouville equation and by Chavanis (2012a) from the Klimontovich equation (these approaches based on the BBGKY hierarchy or on the quasilinear approximation are equivalent but the formalism is different). These kinetic equations are more general than those derived in the present paper, but they are also more complicated (to derive and to solve). Therefore, the equations presented in this paper may be useful as a first step.

We have also discussed the differences between the present approach based on the BBGKY hierarchy and the more classical two-body encounters theory (Chandrasekhar 1942). The two-body encounters theory, which is usually adapted to short-range potentials (Boltzmann 1872; Chapman & Cowling 1939), can take strong collisions into account so it does not yield any divergence at small scales. However, this approach cannot take spatial inhomogeneity into account so it yields a divergence at large scales. This divergence is due to the long-range nature of the gravitational potential. In addition, the two-body encounters theory does not take collective effects into account; these effects are specific to systems with long-range interactions. By contrast, the approach based on the BBGKY hierarchy takes into account the spatial inhomogeneity of the system and collective effects. Therefore, it does not yield any divergence at large scales. However, it fails to take strong collisions into account because of the weak coupling approximation so it yields a divergence at small scales. In a sense, the gravitational potential is intermediate between short-range and long-range potentials because both strong collisions and collective effects must be taken into account. Therefore, the approaches adapted to short-range or long-range potentials are both marginally applicable to self-gravitating systems (they yield a logarithmic divergence at large or small scales respectively). This is why the kinetic equations of stellar dynamics can be obtained in different manners (i.e. from the impact theory or from the wave theory) that turn out to be complementary to each other.

In this paper, we have assumed that the system is isolated from the surrounding. As a result, the source of noise is a result of discreteness (finite N) effects internal to the system. The case where the noise is caused by external sources (perturbations on a galaxy, cosmological environment on dark matter halos) is also interesting. It has been considered by several authors such as Weinberg (2001), Ma & Bertschinger (2004), and Pichon & Aubert (2006) who developed appropriate kinetic theories.

Finally, we would like to stress that the kinetic theories discussed in this paper (leading to complicated equations) do not replace the valuable tools (N -body simulations and Fokker-Planck codes) used by most astronomers. Nevertheless, they complement these numerical techniques and complete our understanding of the gravitational N -body problem. We also hope that the simplified kinetic equations (Eqs. (114), (115), and (127)) derived in this paper (neglecting collective effects) will be amenable to a numerical analysis.

References

Antonov, V. A. 1962, Vest. Leningr. Gos. Univ., 7, 135
 Ambartsumian, V. A. 1938, Ann. Leningrad State Univ., 22, 19

³¹ As discussed in Appendix E, collective effects can substantially decrease the relaxation time for systems at the verge of instability.

- Arad, I., & Johansson, P. H. 2005, MNRAS, 362, 252
 Balescu, R. 1960, Phys. Fluids, 3, 52
 Balescu, R. 2000, Statistical Dynamics: Matter out of Equilibrium (Imperial College Press)
 Bekenstein, J., & Maoz, E. 1992, ApJ, 390, 79
 Bertin, G., & Stiavelli, M. 1984, A&A, 137, 26
 Bettwieser, E., & Sugimoto, D. 1984, MNRAS, 208, 493
 Binney, J., & Lacey, C. 1988, MNRAS, 230, 597
 Binney, J., & Tremaine, S. 2008, Galactic Dynamics (Princeton Series in Astrophysics)
 Boltzmann, L. 1872, Wien, Ber., 66, 275
 Campa, A., Dauxois, T., & Ruffo, S. 2009, Phys. Rep., 480, 57
 Chabanol, M. L., Corson, F., & Pomeau, Y. 2000, Europhys. Lett., 50, 148
 Chandrasekhar, S. 1942, Principles of Stellar Dynamics (University of Chicago press)
 Chandrasekhar, S. 1943a, ApJ, 97, 255
 Chandrasekhar, S. 1943b, ApJ, 97, 263
 Chandrasekhar, S. 1943c, Rev. Mod. Phys., 15, 1
 Chandrasekhar, S. 1944, ApJ, 99, 47
 Chandrasekhar, S. 1949, Rev. Mod. Phys., 21, 383
 Chandrasekhar, S., & von Neumann, J. 1942, ApJ, 95, 489
 Chapman, S., & Cowling, T. G. 1939, The Mathematical Theory of Non-uniform Gases (Cambridge Mathematical Library)
 Chavanis, P. H. 1998, MNRAS, 300, 981
 Chavanis, P. H. 2005, A&A, 432, 117
 Chavanis, P. H. 2006, Int. J. Mod. Phys. B, 20, 3113
 Chavanis, P. H. 2007, Physica A, 377, 469
 Chavanis, P. H. 2008a, Physica A, 387, 787
 Chavanis, P. H. 2008b, Physica A, 387, 1504
 Chavanis, P. H. 2010, J. Stat. Mech., 5, 05019
 Chavanis, P. H. 2012a, Physica A, 391, 3680
 Chavanis, P. H. 2012b, EPJ Plus, 127, 19
 Chavanis, P. H., & Rieutord, M. 2003, A&A, 412, 1
 Chavanis, P. H., Sommeria, J., & Robert, R. 1996, ApJ, 471, 385
 Cohen, R. S., Spitzer, L., & Routly, P. M. 1950, Phys. Rev., 80, 230
 Cohn, H. 1980, ApJ, 242, 765
 Debye, P., & Hückel, E. 1923, Phys. Z., 24, 305
 de Vega, H. J., & Sanchez, N. 2002, Nucl. Phys. B, 625, 409
 Gasiorowicz, S., Neuman, M., & Riddell, J. R. 1956, Phys. Rev., 101, 922
 Gilbert, I. H. 1968, ApJ, 152, 1043
 Gilbert, I. H. 1970, ApJ, 159, 239
 Goldstein, H. 1956, Classical Mechanics (Cambridge MA: Addison-Wesley)
 Goodman, J. 1983, ApJ, 270, 700
 Goodman, J. 1984, ApJ, 280, 298
 Goodman, J. 1988, ApJ, 329, 612
 Gurzadyan, V. G., & Savvidy, G. K. 1986, A&A, 160, 203
 Heggie, D. C. 1975, MNRAS, 173, 729
 Heggie, D., & Hut, P. 2003, The Gravitational Million-Body Problem: A Multidisciplinary Approach to Star Cluster Dynamics (Cambridge University Press)
 Hénon, M. 1961, Ann. Astrophys., 24, 369
 Hénon, M. 1964, Ann. Astrophys., 27, 83
 Hénon, M. 1982, A&A, 114, 211
 Heyvaerts, J. 2010, MNRAS, 407, 355
 Hjorth, J., & Madsen, J. 1991, MNRAS, 253, 703
 Hubbard, J. 1961a, Proc. R. Soc. Lond., 260, 114
 Hubbard, J. 1961b, Proc. R. Soc. Lond., 261, 371
 Inagaki, S., & Lynden-Bell, D. 1983, MNRAS, 205, 913
 Jeans, J. H. 1915, MNRAS, 76, 70
 Jeans, J. H. 1929, Astronomy and Cosmogony (Cambridge Univ. Press)
 Joyce, M., Marcos, B., & Sylos Labini, F. 2009, PRE, 397, 775
 Kac, M., Uhlenbeck, G. E., & Hemmer, P. C. 1963, J. Math. Phys., 4, 216
 Kadomtsev, B. B., & Pogutse, O. P. 1970, PRL, 25, 1155
 Kalnajs, A. 1971a, ApJ, 166, 275
 Kalnajs, A. 1971b, Astro. Space Sci., 13, 279
 Kalnajs, A. 1977, ApJ, 212, 637
 Kandrup, H. 1981, ApJ, 244, 316
 Kandrup, H. 1983, ApSS, 97, 435
 Katz, J. 2003, Found. Phys., 33, 223
 King, I. R. 1960, AJ, 65, 122
 King, I. R. 1965, AJ, 70, 376
 King, I. R. 1966, AJ, 71, 64
 Kramers, H. A. 1940, Physica, 7, 284
 Landau, L. D. 1936, Phys. Z. Sowj. Union, 10, 154
 Larson, R. B. 1970, MNRAS, 147, 323
 Lee, E. P. 1968, ApJ, 151, 687
 Lemou, M., & Chavanis, P. H. 2010, Physica A, 389, 1021
 Lenard, A. 1960, Ann. Phys. (N. Y.), 10, 390

- Lerche, I. 1971, *ApJ*, 166, 207
Levin, Y., Pakter, R., & Rizzato, F. 2008, *PRE*, 78, 021130
Londrillo, P., Messina, A., & Stiavelli, M. 1991, *MNRAS*, 250, 54
Lynden-Bell, D. 1967, *MNRAS*, 136, 101
Lynden-Bell, D., & Eggleton, P. P. 1980, *MNRAS*, 191, 483
Lynden-Bell, D., & Kalnajs, A. J. 1972, *MNRAS*, 157, 1
Lynden-Bell, D., & Wood, R. 1968, *MNRAS*, 138, 495
Ma, C. P., & Bertschinger, E. 2004, *ApJ*, 612, 28
Maoz, E. 1993, *MNRAS*, 263, 75
Marochnik, L. S. 1968, *Sov. Astr.*, 11, 873
Messer, J., & Spohn, H. 1982, *J. Stat. Phys.*, 29, 561
Michie, R. 1963, *MNRAS*, 125, 127
Miller, R. H. 1966, *ApJ*, 146, 831
Monaghan, J. J. 1978, *MNRAS*, 184, 25
Nelson, R. W., & Tremaine, S. 1999, *MNRAS*, 306, 1
Padmanabhan, T. 1990, *Phys. Rep.*, 188, 285
Parisot, J. P., & Severne, G. 1979, *Astro. Space Sci.*, 61, 121
Pichon, C., & Aubert, D. 2006, *MNRAS*, 368, 1657
Pichon, C., & Cannon, R. C. 1997, *MNRAS*, 291, 616
Risken, H. 1989, *The Fokker-Planck equation* (Springer)
Rosenbluth, M., MacDonald, W., & Judd, D. 1957, *Phys. Rev.*, 107, 1
Rosseland, S. 1928, *MNRAS*, 88, 208
Roy, F., & Perez, J. 2004, *MNRAS*, 348, 62
Schwarzschild, K. 1924, *Seeliger Festschrift*, 94
Severne, G., & Haggerty, M. J. 1976, *Astro. Space Sci.*, 45, 287
Severne, G., & Luwel, M. 1980, *Astro. Space Sci.*, 72, 293
Smart, W. M. 1938, *Stellar Dynamics* (Cambridge Univ. Press)
Spitzer, L. 1940, *MNRAS*, 100, 396
Spitzer, L. 1987, *Dynamical Evolution of Globular Clusters* (Princeton Series in Astrophysics)
Spitzer, L., & Härm, R. 1958, *ApJ*, 127, 544
Stiavelli, M., & Bertin, G. 1987, *MNRAS*, 229, 61
Thorne, R. M. 1968, *ApJ*, 151, 671
Tremaine, S., & Weinberg, M. 1984, *MNRAS*, 209, 729
Tremaine, S., Hénon, M., & Lynden-Bell, D. 1986, *MNRAS*, 219, 285
Trenti, M., Bertin, G., & van Albada, T. S. 2005, *A&A*, 433, 57
Valageas, P. 2006a, *A&A*, 450, 445
Valageas, P. 2006b, *PRE*, 74, 6606
van Albada, T. S. 1982, *MNRAS*, 201, 939
Vlasov, A. A. 1938, *Zh. Eksp. i Teor. Fiz.*, 8, 291
Weinberg, M. D. 1991, *ApJ*, 368, 66
Weinberg, M. D. 1993, *ApJ*, 410, 543
Weinberg, M. D. 1998, *MNRAS*, 297, 101
Weinberg, M. D. 2001, *MNRAS*, 328, 311

Appendix A: The thermodynamic limit and the weak coupling approximation

In this Appendix, we discuss the proper thermodynamic limit of self-gravitating systems and justify the weak coupling approximation.

We call v_m the root mean square velocity of the stars and R the system's size. The kinetic energy $K \sim Nmv_m^2$ and the potential energy $U \sim N^2Gm^2/R$ in the Hamiltonian (Eq. (1)) are comparable provided that $v_m^2 \sim GNm/R$ (this fundamental scaling may also be obtained from the virial theorem). As a result, the energy scales as $E \sim Nmv_m^2 \sim GN^2m^2/R$ and the kinetic temperature, defined by $k_B T = mv_m^2/3$, scales as $k_B T \sim GNm^2/R \sim E/N$. Inversely, these relations may be used to define R and v_m as a function of the energy E (conserved quantity in the micro-canonical ensemble) or as a function of the temperature T (fixed quantity in the canonical ensemble).

The proper thermodynamic limit of a self-gravitating system corresponds to $N \rightarrow +\infty$ in such a way that the normalized energy $\epsilon = ER/(GN^2m^2)$ and the normalized temperature $\eta = \beta GNm^2/R$ are of order unity. Of course, the usual thermodynamic limit $N, V \rightarrow +\infty$ with $N/V \sim 1$ is not applicable to self-gravitating systems since these systems are spatially inhomogeneous.

From R and v_m , we define the dynamical time $t_D = R/v_m \sim 1/\sqrt{G\rho}$. If we rescale the distances by R , the velocities by v_m , and the times by t_D , we find that the equations of the BBGKY hierarchy depend on a single dimensionless parameter η/N . This is the equivalent of the plasma parameter in plasma physics. It is usually argued that the correlation functions scale as $P'_n \sim 1/N^{n-1}$. Therefore, when $N \rightarrow +\infty$, we can expand the equations of the BBGKY hierarchy in powers of $1/N \ll 1$. This corresponds to a weak coupling approximation (see below). We note that the thermodynamic limit is well-defined for the out-of-equilibrium problem although there is no statistical equilibrium state (i.e. the density of state and the partition function diverge).

By a suitable normalization of the parameters, we can take $R \sim v_m \sim t_D \sim m \sim 1$. In that case, we must impose $G \sim 1/N$. This is the Kac prescription (Kac et al. 1963; Messer & Spohn 1982). With this normalization, $E \sim N$, $S \sim N$ and $T \sim 1$. The energy and the entropy are extensive but they remain fundamentally non-additive (Campa et al. 2009). The temperature is intensive. This normalization is very convenient since the length, velocity and time scales are of order unity. Furthermore, since the coupling constant G scales as $1/N$, this immediately shows that a regime of weak coupling holds when $N \gg 1$.

Other normalizations of the parameters are possible. For example, Gilbert (1968) considers the limit $N \rightarrow +\infty$ with $R \sim v_m \sim t_D \sim G \sim 1$ and $m \sim 1/N$. In that case, $E \sim 1$, $S \sim N$, and $T \sim 1/N$. On the other hand, de Vega & Sanchez (2002) define the thermodynamic limit as $N \rightarrow +\infty$ with $m \sim G \sim v_m \sim 1$ and $R \sim N$ in order to have $E \sim N$, $S \sim N$, and $T \sim 1$. This normalization is natural since m and G should not depend on N . However, in that case, the dynamical time $t_D = R/v_m \sim N$ diverges with the number of particles. Therefore, this normalization is not very convenient to develop a kinetic theory of stellar systems³². If we impose $G \sim 1$, $E \sim N$, $T \sim 1$ and $t_D \sim 1$, we get $R \sim N^{1/5}$ and $m \sim N^{-2/5}$ (this corresponds to $\rho \sim 1$ since $t_D \sim 1/\sqrt{G\rho}$). We stress that all these scalings are equally valid. The important thing is that ϵ and η are $O(1)$ when $N \rightarrow +\infty$. One

³² If we take $m \sim G \sim R \sim 1$ then $T \sim N$, $E \sim N^2$ and $S \sim N$. In that case, the dynamical time $t_D \sim 1/\sqrt{N}$ tends to zero which is not convenient neither. If we take $m \sim G \sim t_D \sim 1$ then $R \sim N^{1/3}$ leading to $T \sim N^{2/3}$, $E \sim N^{5/3}$ and $S \sim N$.

should choose the most convenient scaling which, in our opinion, is the Kac scaling³³.

We can also present the preceding results in the following manner, by analogy with plasma physics. A fundamental length scale in self-gravitating systems is the Jeans length $\lambda_J = (4\pi G\beta m\rho)^{-1/2}$. This is the counterpart of the Debye length $\lambda_D = (4\pi e^2\beta\rho/m)^{-1/2}$ in plasma physics. Since $v_m^2 \sim GM/R$, we find that $\lambda_J \sim R$. Therefore, the Jeans length is of the order of the system's size. On the other hand, a fundamental time scale in self-gravitating systems is provided by the gravitational pulsation $\omega_G = (4\pi G\rho)^{1/2}$. This is the counterpart of the plasma pulsation $\omega_P = (4\pi e^2\rho/m^2)^{1/2}$ in plasma physics. We can define the dynamical time by $t_D = \omega_G^{-1} = (4\pi G\rho)^{-1/2} = \lambda_J/v_m$. If we rescale the distances by λ_J and the times by t_D , we find that the equations of the BBGKY hierarchy depend on a single dimensionless parameter $1/(n\lambda_J^3)$, where $\Lambda = n\lambda_J^3 \sim N$ gives the number of stars in the Jeans sphere. This is the counterpart of the number of electrons $\Lambda = n\lambda_D^3$ in the Debye sphere in plasma physics. When $\Lambda \gg 1$, we can expand the equations of the BBGKY hierarchy in terms of the small parameter $1/\Lambda$.

Finally, we show that $1/\Lambda$ may be interpreted as a coupling parameter. The coupling parameter Γ is defined as the ratio of the interaction strength at the mean interparticle distance $Gm^2n^{1/3}$ (resp. $e^2n^{1/3}$) to the thermal energy $k_B T$. This leads to $\Gamma = Gm^2n^{1/3}/k_B T = 1/(n\lambda_J^3)^{2/3} = 1/\Lambda^{2/3} \sim 1/N^{2/3}$ (resp. $\Gamma = e^2n^{1/3}/k_B T = 1/(n\lambda_D^3)^{2/3} = 1/\Lambda^{2/3}$). If we define the coupling parameter g as the ratio of the interaction strength at the Jeans (resp. Debye) length Gm^2/λ_J (resp. e^2/λ_D) to the thermal energy $k_B T$, we get $g = 1/\Lambda$ (or $g = E_{\text{pot}}/E_{\text{kin}} = (Gm^2/R)/k_B T = \eta/N$ with the initial variables). Therefore, the expansion of the BBGKY hierarchy in terms of the coupling parameter Γ or g is equivalent to an expansion in terms of the inverse of the number of particles in the Jeans sphere $\Lambda = n\lambda_J^3 \sim N$ (resp. Debye sphere $\Lambda = n\lambda_D^3$). The weak coupling approximation is therefore justified when $\Lambda \gg 1$.

Appendix B: Angle-action variables

In Sect. 4, we have explained that during its collisional evolution a stellar system passes by a succession of QSSs that are steady states of the Vlasov equation slowly changing under the effect of close encounters (finite N effects). The slowly varying distribution function $f(\mathbf{r}, \mathbf{v})$ determines a potential $\Phi(\mathbf{r})$ and a one-particle Hamiltonian $\epsilon = v^2/2 + \Phi(\mathbf{r})$ that we assume to be integrable. Therefore, it is possible to use angle-action variables constructed with this Hamiltonian (Goldstein 1956; Binney & Tremaine 2008). This construction is done adiabatically, i.e. the distribution function, and the angle-action variables, slowly change in time.

A particle with coordinates (\mathbf{r}, \mathbf{v}) in phase space is described equivalently by the angle-action variables (\mathbf{w}, \mathbf{J}) . The Hamiltonian equations for the conjugate variables (\mathbf{r}, \mathbf{v}) are

$$\frac{d\mathbf{r}}{dt} = \frac{\partial\epsilon}{\partial\mathbf{v}} = \mathbf{v}, \quad \frac{d\mathbf{v}}{dt} = -\frac{\partial\epsilon}{\partial\mathbf{r}} = -\nabla\Phi(\mathbf{r}). \quad (\text{B.1})$$

In terms of the variables (\mathbf{r}, \mathbf{v}) , the dynamics is complicated because the potential explicitly appears in the second equation.

³³ If we account for short-range interactions (quantum mechanics, soften potential, hard spheres) there is a new dimensionless parameter in the problem that we shall generically call μ . In that case, the proper thermodynamic limit of self-gravitating systems with short-range interactions corresponds to $N \rightarrow +\infty$ in such a way that ϵ , η , and μ are of order unity. Some examples of scalings are given by Chavanis & Rieutord (2003).

Therefore, this equation $d\mathbf{v}/dt = -\nabla\Phi$ cannot be easily integrated except if $\Phi = 0$, i.e. for a spatially homogeneous system. In that case, the velocity \mathbf{v} is constant and the unperturbed equations of motion reduce to $\mathbf{r} = \mathbf{v}t + \mathbf{r}_0$, i.e. to a rectilinear motion at constant velocity. Now, the angle-action variables are constructed so that the Hamiltonian does not depend on the angles \mathbf{w} . Therefore, the Hamiltonian equations for the conjugate variables (\mathbf{w}, \mathbf{J}) are

$$\frac{d\mathbf{w}}{dt} = \frac{\partial\epsilon}{\partial\mathbf{J}} = \boldsymbol{\Omega}(\mathbf{J}), \quad \frac{d\mathbf{J}}{dt} = -\frac{\partial\epsilon}{\partial\mathbf{w}} = \mathbf{0}, \quad (\text{B.2})$$

where $\boldsymbol{\Omega}(\mathbf{J})$ is the angular frequency of the orbit with action \mathbf{J} . From these equations, we find that \mathbf{J} is a constant and that $\mathbf{w} = \boldsymbol{\Omega}(\mathbf{J})t + \mathbf{w}_0$. Therefore, the equations of motion are very simple in these variables. They extend naturally the trajectories at constant velocity for spatially homogeneous systems. This is why this choice of variables is relevant to develop the kinetic theory. Of course, even if the description of the motion becomes simple in these variables, the complexity of the problem has not completely disappeared. It is now embodied in the relation between the position and momentum variables and the angle and action variables which can be very complicated.

Appendix C: Calculation of $K^{\mu\nu}$

In this Appendix, we compute the tensor $K^{\mu\nu}$ that appears in the Vlasov-Landau equation (Eq. (28)). Within the local approximation, we can proceed as if the system were spatially homogeneous. In that case, the mean field force vanishes, $\langle\mathbf{F}\rangle = \mathbf{0}$, and the unperturbed equations of motion (i.e. for $N \rightarrow +\infty$) reduce to

$$\mathbf{v}(t - \tau) = \mathbf{v}(t) = \mathbf{v}, \quad (\text{C.1})$$

$$\mathbf{r}(t - \tau) = \mathbf{r}(t) - \mathbf{v}(t)\tau = \mathbf{r} - \mathbf{v}\tau, \quad (\text{C.2})$$

corresponding to a rectilinear motion at constant velocity. The collision term in the kinetic equation (Eq. (26)) can be written as

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = \frac{\partial}{\partial v^\mu} \int d\mathbf{v}_1 K^{\mu\nu} \left(\frac{\partial}{\partial v^\nu} - \frac{\partial}{\partial v_1^\nu}\right) f(\mathbf{r}, \mathbf{v}, t) f(\mathbf{r}, \mathbf{v}_1, t), \quad (\text{C.3})$$

with

$$K^{\mu\nu} = \frac{1}{m} \int_0^{+\infty} d\tau \int d\mathbf{r}_1 F^\mu(1 \rightarrow 0, t) F^\nu(1 \rightarrow 0, t - \tau). \quad (\text{C.4})$$

The force by unit of mass created by particle 1 on particle 0 is given by

$$\mathbf{F}(1 \rightarrow 0) = -m \frac{\partial u}{\partial \mathbf{r}}(\mathbf{r} - \mathbf{r}_1), \quad (\text{C.5})$$

where $u(\mathbf{r} - \mathbf{r}') = -G/|\mathbf{r} - \mathbf{r}'|$ is the gravitational potential. The Fourier transform and the inverse Fourier transform of the potential are defined by

$$\hat{u}(\mathbf{k}) = \int e^{-i\mathbf{k}\cdot\mathbf{x}} u(\mathbf{x}) \frac{d\mathbf{x}}{(2\pi)^d}, \quad u(\mathbf{x}) = \int e^{i\mathbf{k}\cdot\mathbf{x}} \hat{u}(\mathbf{k}) d\mathbf{k}. \quad (\text{C.6})$$

For the gravitational interaction:

$$(2\pi)^3 \hat{u}(\mathbf{k}) = -\frac{4\pi G}{k^2}. \quad (\text{C.7})$$

Substituting Eq. (C.6-b) in Eq. (C.5), and writing explicitly the Lagrangian coordinates, we get

$$\mathbf{F}(1 \rightarrow 0, t - \tau) = -im \int \mathbf{k} e^{i\mathbf{k}\cdot(\mathbf{r}(t-\tau) - \mathbf{r}_1(t-\tau))} \hat{u}(\mathbf{k}) d\mathbf{k}. \quad (\text{C.8})$$

Using the equations of motion (C.1) and (C.2), and introducing the notations $\mathbf{x} = \mathbf{r} - \mathbf{r}_1$ and $\mathbf{w} = \mathbf{v} - \mathbf{v}_1$, we obtain

$$\mathbf{F}(1 \rightarrow 0, t - \tau) = -im \int \mathbf{k} e^{i\mathbf{k}\cdot(\mathbf{x} - \mathbf{w}\tau)} \hat{u}(\mathbf{k}) d\mathbf{k}. \quad (\text{C.9})$$

Therefore,

$$K^{\mu\nu} = -m \int_0^{+\infty} d\tau \int d\mathbf{x} \int d\mathbf{k} \int d\mathbf{k}' k_\mu k'_\nu \times e^{i(\mathbf{k} + \mathbf{k}')\cdot\mathbf{x}} e^{-i\mathbf{k}'\cdot\mathbf{w}\tau} \hat{u}(\mathbf{k}) \hat{u}(\mathbf{k}'). \quad (\text{C.10})$$

Using the identity

$$\delta(\mathbf{x}) = \int e^{i\mathbf{k}\cdot\mathbf{x}} \frac{d\mathbf{k}}{(2\pi)^d}, \quad (\text{C.11})$$

and integrating over \mathbf{x} and \mathbf{k}' , we find that

$$K^{\mu\nu} = (2\pi)^3 m \int_0^{+\infty} d\tau \int d\mathbf{k} k^\mu k^\nu e^{i\mathbf{k}\cdot\mathbf{w}\tau} \hat{u}(\mathbf{k})^2. \quad (\text{C.12})$$

Performing the transformation $\tau \rightarrow -\tau$, then $\mathbf{k} \rightarrow -\mathbf{k}$, and adding the resulting expression to Eq. (C.12), we get

$$K^{\mu\nu} = \frac{1}{2} (2\pi)^3 m \int_{-\infty}^{+\infty} d\tau \int d\mathbf{k} k^\mu k^\nu e^{i\mathbf{k}\cdot\mathbf{w}\tau} \hat{u}(\mathbf{k})^2. \quad (\text{C.13})$$

Using the identity (C.11), we finally obtain

$$K^{\mu\nu} = \pi(2\pi)^3 m \int d\mathbf{k} k^\mu k^\nu \delta(\mathbf{k}\cdot\mathbf{w}) \hat{u}(\mathbf{k})^2, \quad (\text{C.14})$$

which leads to Eq. (27).

Introducing a spherical system of coordinates in which the z axis is taken in the direction of \mathbf{w} , we find that

$$K^{\mu\nu} = \pi(2\pi)^3 m \int_0^{+\infty} k^2 dk \int_0^\pi \sin\theta d\theta \times \int_0^{2\pi} d\phi k^\mu k^\nu \delta(kw \cos\theta) \hat{u}(k)^2. \quad (\text{C.15})$$

Using $k_x = k \sin\theta \cos\phi$, $k_y = k \sin\theta \sin\phi$ and $k_z = k \cos\theta$, it is easy to see that only K_{xx} , K_{yy} and K_{zz} can be non-zero. The other components of the matrix $K^{\mu\nu}$ vanish by symmetry. Furthermore,

$$K_{xx} = K_{yy} = \pi^2 (2\pi)^3 m \int_0^{+\infty} k^2 dk \times \int_0^\pi \sin\theta d\theta k^2 \delta(kw \cos\theta) \hat{u}(k)^2 \sin^2\theta. \quad (\text{C.16})$$

Using the identity $\delta(\lambda x) = \frac{1}{|\lambda|} \delta(x)$, we get

$$K_{xx} = K_{yy} = \pi^2 (2\pi)^3 m \frac{1}{w} \int_0^{+\infty} k^3 \hat{u}(k)^2 dk \times \int_0^\pi \sin^3\theta \delta(\cos\theta) d\theta. \quad (\text{C.17})$$

With the change of variables $s = \cos\theta$, we obtain

$$K_{xx} = K_{yy} = \pi^2 (2\pi)^3 m \frac{1}{w} \int_0^{+\infty} k^3 \hat{u}(k)^2 dk \times \int_{-1}^{+1} (1 - s^2) \delta(s) ds, \quad (\text{C.18})$$

so that, finally,

$$K_{xx} = K_{yy} = 8\pi^5 m \frac{1}{w} \int_0^{+\infty} k^3 \hat{u}(k)^2 dk. \quad (\text{C.19})$$

On the other hand,

$$K^{zz} = 2\pi^2 (2\pi)^3 m \frac{1}{w} \int_0^{+\infty} k^3 \hat{u}(k)^2 dk \\ \times \int_0^\pi \sin \theta \cos^2 \theta \delta(\cos \theta) d\theta = 0. \quad (\text{C.20})$$

In conclusion, we obtain

$$K^{\mu\nu} = A \frac{w^2 \delta^{\mu\nu} - w^\mu w^\nu}{w^3}, \quad (\text{C.21})$$

with

$$A = 8\pi^5 m \int_0^{+\infty} k^3 \hat{u}(k)^2 dk. \quad (\text{C.22})$$

Using Eq. (C.7), this leads to Eqs. (28)–(30).

Appendix D: Another derivation of the Landau equation

For an infinite homogeneous system, the distribution function and the two-body correlation function can be written as $f(0) = f(\mathbf{v}, t)$ and $g(0, 1) = g(\mathbf{r} - \mathbf{r}_1, \mathbf{v}, \mathbf{v}_1, t)$. In that case, Eqs. (18) and (19) become

$$\frac{\partial f}{\partial t}(\mathbf{v}, t) = \frac{\partial}{\partial \mathbf{v}} \cdot \int \frac{\partial u}{\partial \mathbf{x}} g(\mathbf{x}, \mathbf{v}, \mathbf{v}_1, t) d\mathbf{x} d\mathbf{v}_1, \quad (\text{D.1})$$

$$\frac{\partial g}{\partial t}(\mathbf{x}, \mathbf{v}, \mathbf{v}_1, t) + \mathbf{w} \cdot \frac{\partial g}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{v}, \mathbf{v}_1, t) = \\ m \frac{\partial u}{\partial \mathbf{x}} \cdot \left(\frac{\partial}{\partial \mathbf{v}} - \frac{\partial}{\partial \mathbf{v}_1} \right) f(\mathbf{v}, t) f(\mathbf{v}_1, t), \quad (\text{D.2})$$

where we have defined $\mathbf{x} = \mathbf{r} - \mathbf{r}_1$ and $\mathbf{w} = \mathbf{v} - \mathbf{v}_1$. Using the Bogoliubov ansatz, we shall treat the distribution function f as a constant, and determine the asymptotic value $g(\mathbf{x}, \mathbf{v}, \mathbf{v}_1, +\infty)$ of the correlation function. Introducing the Fourier transforms of the potential of interaction and of the correlation function, Eq. (D.1) may be replaced by

$$\frac{\partial f}{\partial t} = (2\pi)^3 \frac{\partial}{\partial \mathbf{v}} \cdot \int \mathbf{k} \hat{u}(k) \text{Im} [\hat{g}(\mathbf{k}, \mathbf{v}, \mathbf{v}_1, +\infty)] d\mathbf{k} d\mathbf{v}_1, \quad (\text{D.3})$$

where we have used the reality condition $\hat{g}(-\mathbf{k}) = \hat{g}(\mathbf{k})^*$. On the other hand, taking the Laplace-Fourier transform of Eq. (D.2) and assuming that no correlation is present initially (if there are initial correlations, their effect becomes rapidly negligible), we get

$$\tilde{g}(\mathbf{k}, \mathbf{v}, \mathbf{v}_1, \omega) = im \hat{u}(k) \frac{1}{\omega(\mathbf{k} \cdot \mathbf{w} - \omega)} \\ \times \mathbf{k} \cdot \left(\frac{\partial}{\partial \mathbf{v}} - \frac{\partial}{\partial \mathbf{v}_1} \right) f(\mathbf{v}, t) f(\mathbf{v}_1, t). \quad (\text{D.4})$$

Taking the inverse Laplace transform of Eq. (D.4), and using the residue theorem, we find that the asymptotic value $t \rightarrow +\infty$ of the correlation function, determined by the pole $\omega = 0$, is

$$\hat{g}(\mathbf{k}, \mathbf{v}, \mathbf{v}_1, +\infty) = m \hat{u}(k) \frac{1}{\mathbf{k} \cdot \mathbf{w} - i0^+} \\ \times \mathbf{k} \cdot \left(\frac{\partial}{\partial \mathbf{v}} - \frac{\partial}{\partial \mathbf{v}_1} \right) f(\mathbf{v}, t) f(\mathbf{v}_1, t). \quad (\text{D.5})$$

Using the Plemelj formula

$$\frac{1}{x \pm i0^+} = \mathcal{P} \left(\frac{1}{x} \right) \mp i\pi \delta(x), \quad (\text{D.6})$$

we get

$$\text{Im} [\hat{g}(\mathbf{k}, \mathbf{v}, \mathbf{v}_1, +\infty)] = \pi m \hat{u}(k) \delta(\mathbf{k} \cdot \mathbf{w}) \\ \times \mathbf{k} \cdot \left(\frac{\partial}{\partial \mathbf{v}} - \frac{\partial}{\partial \mathbf{v}_1} \right) f(\mathbf{v}, t) f(\mathbf{v}_1, t). \quad (\text{D.7})$$

Substituting Eq. (D.7) in Eq. (D.3), we obtain the Landau equation (Eq. (27)).

Appendix E: Lenard-Balescu equation for homogeneous stellar systems

If we assume that the system is spatially homogeneous (or make the local approximation), and take collective effects into account, the Vlasov-Landau Eq. (27) is replaced by the Vlasov-Lenard-Balescu equation

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \frac{N-1}{N} \langle F \rangle \cdot \frac{\partial f}{\partial \mathbf{v}} = \pi (2\pi)^3 m \frac{\partial}{\partial v^\mu} \int k^\mu k^\nu \\ \times \delta(\mathbf{k} \cdot \mathbf{w}) \frac{\hat{u}^2(k)}{|\epsilon(\mathbf{k}, \mathbf{k} \cdot \mathbf{v})|^2} \left(f_1 \frac{\partial f}{\partial v^\nu} - f \frac{\partial f_1}{\partial v^\nu} \right) d\mathbf{v}_1 d\mathbf{k}, \quad (\text{E.1})$$

where $\epsilon(\mathbf{k}, \omega)$ is the dielectric function

$$\epsilon(\mathbf{k}, \omega) = 1 + (2\pi)^3 \hat{u}(k) \int \frac{\mathbf{k} \cdot \frac{\partial f}{\partial \mathbf{v}}}{\omega - \mathbf{k} \cdot \mathbf{v}} d\mathbf{v}. \quad (\text{E.2})$$

The Landau equation is recovered by taking $|\epsilon(\mathbf{k}, \mathbf{k} \cdot \mathbf{v})|^2 = 1$. The Lenard-Balescu equation generalizes the Landau equation by replacing the bare potential of interaction $\hat{u}(k)$ by a dressed potential of interaction

$$\hat{u}_{\text{dressed}}(\mathbf{k}, \mathbf{k} \cdot \mathbf{v}) = \frac{\hat{u}(k)}{|\epsilon(\mathbf{k}, \mathbf{k} \cdot \mathbf{v})|}. \quad (\text{E.3})$$

The dielectric function in the denominator takes the dressing of the particles by their polarization cloud into account. In plasma physics, this term corresponds to a screening of the interactions. The Lenard-Balescu equation accounts for dynamical screening since the velocity \mathbf{v} of the particles explicitly appears in the effective potential. However, for Coulombian interactions, it is a good approximation to neglect the deformation of the polarization cloud due to the motion of the particles and use the static results on screening (Debye & Hückel 1923). This amounts to replacing the dynamic dielectric function $|\epsilon(\mathbf{k}, \mathbf{k} \cdot \mathbf{v})|$ by the static dielectric function $|\epsilon(\mathbf{k}, 0)|$. In this approximation, $\hat{u}_{\text{dressed}}(\mathbf{k}, \mathbf{k} \cdot \mathbf{v})$ is replaced by the Debye-Hückel potential $(2\pi)^3 \hat{u}_{\text{DH}}(k) = (4\pi e^2/m^2)/(k^2 + k_{\text{D}}^2)$ corresponding to $u_{\text{DH}}(x) = (e^2/m^2)e^{-k_{\text{D}}r}/r$ in physical space. If we make the same approximation for stellar systems, we find that $\hat{u}_{\text{dressed}}(\mathbf{k}, \mathbf{k} \cdot \mathbf{v})$ is replaced by

$$(2\pi)^3 \hat{u}_{\text{DH}}(k) = -\frac{4\pi G}{k^2 - k_{\text{J}}^2}, \quad (\text{E.4})$$

corresponding to $u_{\text{DH}}(x) = -G \cos(k_{\text{J}}r)/r$ in physical space. In this approximation, the Vlasov-Lenard-Balescu equation (Eq. (E.1)) takes the same form as the Vlasov-Landau equation (Eq. (28)) except that A is now given by $A = 2\pi m G^2 Q$ with

$$Q = \int_{1/R}^{k_{\text{L}}} \frac{k^3}{(k^2 - k_{\text{J}}^2)^2} dk, \quad (\text{E.5})$$

where R is the system's size. We see that Q diverges algebraically, as $(\lambda_J - R)^{-1}$, when $R \rightarrow \lambda_J$ instead of yielding a finite Coulombian logarithm $\ln \Lambda \sim \ln N$ when collective effects are neglected³⁴. This naive approach shows that collective effects (which account for anti-shielding) tend to increase the diffusion coefficient and consequently tend to reduce the relaxation time. This is the conclusion reached by Weinberg (1993) with a more precise approach. However, his approach is not fully satisfactory since the system is assumed to be spatially homogeneous and the ordinary Lenard-Balescu equation is used. In that case, the divergence when $R \rightarrow \lambda_J$ is a manifestation of the Jeans instability that a spatially homogeneous self-gravitating system experiences when the size of the perturbation overcomes the Jeans length. For inhomogeneous systems, the Jeans instability is suppressed so the results of Weinberg should be used with caution. Heyvaerts (2010) and Chavanis (2012a) derived a more satisfactory Lenard-Balescu equation that is valid for spatially inhomogeneous stable self-gravitating systems. This equation does not present any divergence at large scales. However, this equation is complicated and it is difficult to measure the importance of collective effects. The approach of Weinberg (1993), and the arguments given in this Appendix, suggest that collective effects reduce the relaxation time of inhomogeneous stellar systems. This reduction should be particularly strong for a system close to instability because of the enhancement of fluctuations (Monaghan 1978).

These considerations show that it is not possible to make a local approximation and simultaneously take collective effects into account. The usual procedure is to ignore collective effects, make a local approximation, and introduce a large-scale cut-off at the Jeans length. This is the procedure that is usually followed in stellar dynamics (Binney & Tremaine 2008). The only rigorous manner to take collective effects into account is to use the Lenard-Balescu equation written with angle-action variables.

Appendix F: Multi-species systems

It is straightforward to generalize the kinetic theory of stellar systems for several species of stars. The Vlasov-Landau equation (Eq. (27)) is replaced by

$$\frac{df^a}{dt} = \pi(2\pi)^3 \frac{\partial}{\partial v^\mu} \int k^\mu k^\nu \delta(\mathbf{k} \cdot \mathbf{w}) \hat{u}^2(k) \times \sum_b \left(m_b f_1^b \frac{\partial f^a}{\partial v^\nu} - m_a f^a \frac{\partial f_1^b}{\partial v_1^\nu} \right) dv_1 d\mathbf{k}, \quad (\text{F.1})$$

where $f^a(\mathbf{r}, \mathbf{v}, t)$ is the distribution function of species a normalized such that $\int f^a d\mathbf{r} d\mathbf{v} = N_a m_a$, and the sum \sum_b runs over all species. We can use this equation to give a new interpretation of the test particle approach developed in Sect. 3. We make three assumptions: (i) We assume that the system is composed of two types of stars, the test stars with mass m and the field stars with mass m_f ; (ii) we assume that the number of test stars is much lower than the number of field stars; and (iii) we assume that the field stars are in a steady distribution $f(\mathbf{r}, \mathbf{v})$. Because of assumption (ii), the collisions between the field stars and the test stars do not alter the distribution of the field stars so that the field stars

³⁴ In plasma physics, on the contrary, $Q = \int_0^{\lambda_D} k^3 / [(k^2 + k_D^2)^2] dk$ is well-behaved as $k \rightarrow 0$. Its value is $Q = \frac{1}{2} [\ln(1 + \Lambda^2) - \frac{\Lambda^2}{1 + \Lambda^2}]$, where $\Lambda = \lambda_D / \lambda_L$. When collective effects are taken into account, there is no divergence at large scales and the Debye length appears naturally. In the dominant approximation $Q \sim \ln \Lambda$ with $\Lambda = n \lambda_D^3$.

remain in their steady state. The collisions of the test stars among themselves are also negligible, so they only evolve as a result of collisions with the field stars. Therefore, if we call $P(\mathbf{r}, \mathbf{v}, t)$ the distribution function of the test stars (to have notations similar to those of Sect. 3 with, however, a different interpretation), its evolution is given by the Fokker-Planck equation obtained from Eq. (F.1) yielding

$$\frac{dP}{dt} = \pi(2\pi)^3 \frac{\partial}{\partial v^\mu} \int k^\mu k^\nu \delta(\mathbf{k} \cdot \mathbf{w}) \hat{u}^2(k) \times \left(m_f f_1 \frac{\partial P}{\partial v^\nu} - m P \frac{\partial f_1}{\partial v_1^\nu} \right) dv_1 d\mathbf{k}. \quad (\text{F.2})$$

The diffusion and friction coefficients are given by

$$D^{\mu\nu} = \pi(2\pi)^3 m_f \int k^\mu k^\nu \delta(\mathbf{k} \cdot \mathbf{w}) \hat{u}(k)^2 f_1 dv_1 d\mathbf{k}, \quad (\text{F.3})$$

$$F_{\text{pol}}^\mu = \pi(2\pi)^3 m \int k^\mu k^\nu \delta(\mathbf{k} \cdot \mathbf{w}) \hat{u}(k)^2 \frac{\partial f_1}{\partial v_1^\nu} dv_1 d\mathbf{k}. \quad (\text{F.4})$$

We recall that the diffusion coefficient is due to the fluctuations of the gravitational force produced by the field stars, while the friction by polarization is due to the perturbation on the distribution of the field stars caused by the test stars. This explains the occurrence of the masses m_f and m in Eqs. (F.3) and (F.4), respectively.

Using Eq. (46) and noting that

$$\frac{\partial D^{\mu\nu}}{\partial v^\nu} = \frac{m_f}{m} F_{\text{pol}}^\mu, \quad (\text{F.5})$$

we get

$$\mathbf{F}_{\text{friction}} = \left(1 + \frac{m_f}{m} \right) \mathbf{F}_{\text{pol}}. \quad (\text{F.6})$$

If we assume furthermore that $m \gg m_f$, we find that $\mathbf{F}_{\text{friction}} \simeq \mathbf{F}_{\text{pol}}$. However, in general, the friction force is different from the friction by polarization. The other results of Sect. 3 can be easily generalized to multi-species systems.

If the field stars have an isothermal distribution, then from Eq. (F.4), $F_{\text{pol}}^\mu = -\beta m D^{\mu\nu}(\mathbf{v}) v^\nu = -\beta m D_{\parallel}(\mathbf{v}) v^\mu$ and the Fokker-Planck equation (Eq. (F.2)) reduces to

$$\frac{dP}{dt} = \frac{\partial}{\partial v^\mu} \left[D^{\mu\nu}(\mathbf{v}) \left(\frac{\partial P}{\partial v^\nu} + \beta m P v^\nu \right) \right], \quad (\text{F.7})$$

where $D^{\mu\nu}(\mathbf{v})$ is given by Eq. (F.3). Using the results of Sect. 3, we get

$$D^{\mu\nu} = \left(\frac{3}{2\pi} \right)^{1/2} \frac{2\rho_f m_f G^2 \ln \Lambda}{v_{\text{mf}}} G^{\mu\nu} \left(\sqrt{\frac{3}{2}} \frac{\mathbf{v}}{v_{\text{mf}}} \right), \quad (\text{F.8})$$

where $v_{\text{mf}}^2 = 3/(\beta m_f)$ is the velocity dispersion of the field stars and $\rho_f = n_f m_f$ their density. On the other hand, $\ln \Lambda = \ln(\lambda_J / \lambda_L)$, where $\lambda_J \sim (k_B T / G m_f^2 n_f)^{1/2}$ is the Jeans length and $\lambda_L \sim G(m + m_f) / v_{\text{mf}}^2$ is the Landau length. This yields $\Lambda \sim n_f \lambda_J^3 [m_f / (m + m_f)]$. The total friction is $F_{\text{friction}}^\mu = -\beta(m + m_f) D^{\mu\nu}(\mathbf{v}) v^\nu = -\beta(m + m_f) D_{\parallel}(\mathbf{v}) v^\mu$. If the distribution of the field stars is $f \propto e^{-\beta m_f v^2/2}$, the equilibrium distribution of the test stars is $P \propto e^{-\beta m v^2/2} \propto f^{m/m_f}$. When $m \gg m_f$, the evolution is dominated by frictional effects; when $m \ll m_f$ it is dominated by diffusion. The relaxation time scales like $t_R^{\text{bath}} \sim v_m^2 / D$ where v_m^2 is the r.m.s. velocity of the test stars. Therefore, $t_R^{\text{bath}} \sim 1/(D\beta m) \sim 1/\xi$, where ξ is the friction coefficient associated to the friction by polarization. The true friction coefficient is $\xi_* = (1 + m_f/m)\xi$. We get $t_R^{\text{bath}} \sim v_m^2 v_{\text{mf}} / (\rho_f m_f G^2 \ln \Lambda) \sim v_{\text{mf}}^3 / (\rho_f m G^2 \ln \Lambda)$, where we have used $m v_m^2 \sim m_f v_{\text{mf}}^2 \sim 1/\beta$.

Appendix G: Temporal correlation tensor of the gravitational force

The diffusion of the stars is caused by the fluctuations of the gravitational force. For an infinite homogeneous system (or in the local approximation), the diffusion tensor can be derived from Eq. (43) that is well-known in Brownian theory. This expression involves the temporal auto-correlation tensor of the gravitational force experienced by a star. It can be written as

$$\langle F^\mu(t)F^\nu(t-\tau) \rangle = \frac{1}{m} \int d\mathbf{r}_1 d\mathbf{v}_1 F^\mu(1 \rightarrow 0, t) \times F^\nu(1 \rightarrow 0, t-\tau) f(\mathbf{v}_1). \quad (\text{G.1})$$

We first compute the tensor

$$Q^{\mu\nu} = \frac{1}{m} \int F^\mu(1 \rightarrow 0, t) F^\nu(1 \rightarrow 0, t-\tau) d\mathbf{r}_1. \quad (\text{G.2})$$

Proceeding as in Appendix C, we find

$$Q^{\mu\nu} = (2\pi)^3 m \int d\mathbf{k} k^\mu k^\nu e^{i\mathbf{k}\cdot\mathbf{w}\tau} \hat{u}(k)^2. \quad (\text{G.3})$$

Introducing a system of spherical coordinates with the z -axis in the direction of \mathbf{w} , and using Eq. (C.7), we obtain after some calculations

$$Q^{\mu\nu} = 2\pi m G^2 \frac{1}{\tau} \frac{w^2 \delta^{\mu\nu} - w^\mu w^\nu}{w^3}. \quad (\text{G.4})$$

According to Eq. (C.4), we have

$$K^{\mu\nu} = \int_0^{+\infty} Q^{\mu\nu} d\tau. \quad (\text{G.5})$$

Therefore, Eqs. (G.4) and (G.5) lead to Eq. (29) with $\ln \Lambda$ replaced by

$$\ln \Lambda' = \int_{t_{\min}}^{t_{\max}} \frac{d\tau}{\tau}, \quad (\text{G.6})$$

where t_{\min} and t_{\max} are appropriate cut-offs. The upper cut-off should be identified with the dynamical time t_D . On the other hand, the divergence at short times is due to the inadequacy of our assumption of straight-line trajectories to describe very close encounters. If we take $t_{\min} = \lambda_L/v_m$ and $t_{\max} = \lambda_J/v_m \sim t_D$, we find that $\ln \Lambda' = \ln \Lambda$. In Appendix C, we have calculated $K^{\mu\nu}$ by integrating first over time then over space. This yields the Coulombian logarithm (Eq. (30)). Here, we have integrated first over space then over time. This yields the Coulombian logarithm (Eq. (G.6)). As discussed by Lee (1968), these two approaches are essentially equivalent. We remark, however, that the calculations of Sect. C can be performed for arbitrary potentials, while the calculations of this section explicitly use the specific form of the gravitational potential.

According to Eqs. (G.1), (G.2), and (G.4), the force auto-correlation function can be written as

$$\langle F^\mu(t)F^\nu(t-\tau) \rangle = \frac{2\pi m G^2}{\tau} \int d\mathbf{v}_1 \frac{\delta^{\mu\nu} w^2 - w^\mu w^\nu}{w^3} f(\mathbf{v}_1). \quad (\text{G.7})$$

In particular

$$\langle \mathbf{F}(t) \cdot \mathbf{F}(t-\tau) \rangle = 4\pi m G^2 \frac{1}{\tau} \int \frac{f(\mathbf{v}_1)}{|\mathbf{v} - \mathbf{v}_1|} d\mathbf{v}_1. \quad (\text{G.8})$$

The τ^{-1} decay of the auto-correlation function of the gravitational force was first derived by Chandrasekhar (1944) with a

different method. This result has also been obtained, and discussed, by Cohen et al. (1950) and Lee (1968). According to Eqs. (43) and (G.7), the diffusion tensor is given by³⁵

$$D^{\mu\nu} = 2\pi m G^2 \ln \Lambda' \int d\mathbf{v}_1 \frac{\delta^{\mu\nu} w^2 - w^\mu w^\nu}{w^3} f(\mathbf{v}_1). \quad (\text{G.9})$$

This returns Eq. (79) with $\ln \Lambda'$ instead of $\ln \Lambda$.

When $f(\mathbf{v}_1)$ is the Maxwell distribution (Eq. (51)), we can compute the force auto-correlation tensor from Eq. (G.7) by using the Rosenbluth potentials as in Sect. 3.5. Alternatively, combining Eqs. (G.1) and (G.3), we have

$$\langle F^\mu(t)F^\nu(t-\tau) \rangle = (2\pi)^6 m \int d\mathbf{k} k^\mu k^\nu e^{i\mathbf{k}\cdot\mathbf{v}\tau} \hat{u}(k)^2 \hat{f}(k\tau), \quad (\text{G.10})$$

where \hat{f} is the three-dimensional Fourier transform of the distribution function. For the Maxwell distribution (Eq. (51)), we obtain

$$\langle F^\mu(t)F^\nu(t-\tau) \rangle = (2\pi)^3 m \rho \times \int d\mathbf{k} k^\mu k^\nu e^{i\mathbf{k}\cdot\mathbf{v}\tau} \hat{u}(k)^2 e^{-\frac{k^2 \tau}{2\beta m}}. \quad (\text{G.11})$$

Introducing a system of spherical coordinates with the z -axis in the direction of \mathbf{v} , and using Eq. (C.7), we obtain after some calculations

$$\langle F^\mu(t)F^\nu(t-\tau) \rangle = \left(\frac{3}{2\pi}\right)^{1/2} \frac{2\rho m G^2}{v_m} \frac{1}{\tau} G^{\mu\nu} \left(\sqrt{\frac{3}{2}} \frac{v}{v_m}\right), \quad (\text{G.12})$$

where $G^{\mu\nu}(x)$ is defined in Sect. 3.3. In particular,

$$\langle \mathbf{F}(t) \cdot \mathbf{F}(t-\tau) \rangle = \frac{4\pi\rho m G^2}{v\tau} \text{erf}\left(\sqrt{\frac{3}{2}} \frac{v}{v_m}\right). \quad (\text{G.13})$$

Integrating Eq. (G.12) over time, we recover the expression (Eq. (67)) of the diffusion tensor for a Maxwellian distribution with $\ln \Lambda'$ instead of $\ln \Lambda$.

Finally, the auto-correlation tensor of the gravitational field at two different points (at the same time) is

$$\langle F^\mu(\mathbf{0})F^\nu(\mathbf{r}) \rangle = 2\pi n m^2 G^2 \frac{r^2 \delta^{\mu\nu} - r^\mu r^\nu}{r^3}. \quad (\text{G.14})$$

Appendix H: The different kinetic equations

The standard kinetic equations (Boltzmann, Fokker-Planck, Vlasov, Landau, and Lenard-Balescu) and their generalizations can be derived from the BBGKY hierarchy. In this Appendix, we show the connection between these different equations, and discuss their domains of validity, without entering into technical details.

The first two equations of the BBGKY hierarchy may be written symbolically as

$$\frac{\partial f}{\partial t} + (\mathcal{V}_0 + \mathcal{V}_{\text{m.f.}}[f])f = C[g], \quad (\text{H.1})$$

$$\frac{\partial g}{\partial t} + (\mathcal{L}_0 + \mathcal{L}' + \mathcal{L}_{\text{m.f.}}[f])g + C[f, g] + \mathcal{T}[h] = \mathcal{S}[f], \quad (\text{H.2})$$

³⁵ Actually, when the force auto-correlation function decreases as t^{-1} , Eq. (43) is no longer valid and $\langle (\Delta\mathbf{v})^2 \rangle$ behaves as $t \ln t$ (Lee 1968).

where f is the one-body distribution function, g the two-body correlation function, and h the three-body correlation function. In the first equation, $\mathcal{V} = \mathcal{V}_0 + \mathcal{V}_{\text{m.f.}}$ is the Vlasov operator taking into account the free motion \mathcal{V}_0 of the particles and the advection by the mean field $\mathcal{V}_{\text{m.f.}}$. On the other hand, the collision term $C[g]$ describes the effect of two-body correlations on the evolution of the distribution function. In the second equation, $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}' + \mathcal{L}_{\text{m.f.}}$ is a two-body Liouville operator where \mathcal{L}_0 describes the free motion of the particles, \mathcal{L}' describes the exact two-body interaction, and $\mathcal{L}_{\text{m.f.}}$ takes into account the effect of the mean field in the two-body problem. The term $C[f, g]$ describes collective effects and the term $\mathcal{T}[h]$ describes three-body correlations. Finally, $\mathcal{S}[f]$ is a source term depending on the one-body distribution function.

As explained in Sect. 2.2, for $N \gg 1$ we can expand the equations of the BBGKY hierarchy in terms of the small parameter $1/N$. For $N \rightarrow +\infty$, the encounters are negligible and we get the Vlasov equation (Jeans 1915; Vlasov 1938). This corresponds to the mean field approximation. At the order $1/N$, we can neglect three-body correlations ($\mathcal{T} = 0$) and strong collisions ($\mathcal{L}' = 0$) that are of order $1/N^2$. This corresponds to the weak coupling approximation. In fact, since the gravitational potential is singular at $r = 0$, the two-body correlation function $g(\mathbf{r}_1, \mathbf{r}_2)$ becomes large when $|\mathbf{r}_1 - \mathbf{r}_2| \rightarrow 0$ due to the effect of strong collisions. As a result, it is necessary to take the term $\mathcal{L}'g$ into account at small scales. When three-body correlations are neglected ($\mathcal{T} = 0$), Eqs. (H.1) and (H.2) are closed. However, it does not appear possible to solve these equations explicitly without further approximation. The usual strategy is to solve these equations for small, intermediate, and large impact parameters, and then connect these limits.

If we neglect strong collisions ($\mathcal{L}' = 0$), collective effects ($C = 0$), and make a local approximation ($\mathcal{L}_{\text{m.f.}} = 0$), we get the Landau equation (Eqs. (28) and (29)) with

$$A^{\text{Landau}} = 2\pi m G^2 \ln \left(\frac{\lambda_{\text{max}}}{\lambda_{\text{min}}} \right). \quad (\text{H.3})$$

This factor presents a logarithmic divergence at small and large scales. The Landau equation may be derived in different manners that are actually equivalent to the above procedure. Landau (1936) obtained his equation by starting from the Boltzmann equation and using a weak deflection approximation $|\Delta v| \ll 1$. In his calculations, he approximated the trajectories of the particles by straight lines even for collisions with small impact parameters. This is equivalent to starting from the Fokker-Planck equation (Eq. (44)) and calculating the first and second moments of the velocity increments (Eq. (45)) resulting from a succession of binary encounters by using a straight line approximation.

If we neglect collective effects ($C = 0$), make a local approximation ($\mathcal{L}_{\text{m.f.}} = 0$), but take strong collisions into account ($\mathcal{L}' \neq 0$), we get the Boltzmann equation (see Balescu 2000). This equation does not present any divergence at small scales. Since the system is dominated by weak encounters, we can expand the Boltzmann equation for weak deflections $|\Delta v| \ll 1$ while taking into account the effect of strong collisions. This is equivalent to the treatment of Chandrasekhar (1942, 1943a,b) who started from the Fokker-Planck equation (Eq. (44)) and calculated the first and second moments of the velocity increments (Eq. (45)) resulting from a succession of binary collisions by taking strong collisions into account. In the calculation of $\langle \Delta v^\mu \rangle$ and $\langle \Delta v^\mu \Delta v^\nu \rangle$, he used the exact trajectory of the stars (i.e. he solved the two-body problem exactly) and took into account the strong deflections due to collisions with small impact parameters. In the dominant approximation $\ln N \gg 1$, his

approach, completed by Rosenbluth et al. (1957), leads to the Fokker-Planck equation (Eq. (44)) with the expressions of the diffusion tensor and friction force (Eqs. (83) and (84)) expressed in terms of the Rosenbluth potentials (Eq. (85)) with

$$A^{\text{Chandra}} = 2\pi m G^2 \ln \left(\frac{\lambda_{\text{max}} v_m^2}{Gm} \right). \quad (\text{H.4})$$

As shown in Sect. 3.5, the resulting Fokker-Planck equation can be transformed into the Landau equation (Eqs. (28) and (29)). As in the Landau approach, the factor calculated in Eq. (H.4) presents a logarithmic divergence at large scales. However, contrary to the Landau approach, it does not present a logarithmic divergence at small scales since the effect of strong collisions is taken into account explicitly in the Chandrasekhar approach³⁶. As a result, the gravitational Landau length $\lambda_L = Gm/v_m^2$ appears naturally in the calculations of Chandrasekhar. This establishes that the relevant small-scale cut-off in Eq. (H.3) is the Landau length.

The Landau and the Chandrasekhar kinetic theories make the same assumption: binary encounters and an expansion in powers of the momentum transfer. They actually differ in the order in which these are introduced. Landau starts from the Boltzmann equation and considers a weak deflection approximation while Chandrasekhar directly starts from the Fokker-Planck equation, but calculates the coefficients of diffusion and friction with the binary-collision picture. At that level, their theories are equivalent³⁷ since the Fokker-Planck equation can be precisely obtained from the Boltzmann equation in the limit of weak deflections. The crucial difference is that Landau makes a weak coupling assumption and ignores strong collisions (i.e. the bending of the trajectories) while Chandrasekhar takes them into account. On the other hand, both theories ignore spatial inhomogeneity and collective effects.

If we neglect strong collisions ($\mathcal{L}' = 0$) and collective effects ($C = 0$), but take spatial inhomogeneity ($\mathcal{L}_{\text{m.f.}} \neq 0$) into account we get the generalized Landau equation (Eqs. (24) and (114)) derived in this paper (see also Kandrup 1981; and Chavanis 2008a,b). This equation presents a logarithmic divergence at small scales since strong collisions are neglected, but not at large scales since the finite extent of the system is taken into account. This suggests that the relevant large-scale cut-off in Eqs. (H.3) and (H.4) is the Jeans length.

If we neglect strong collisions ($\mathcal{L}' = 0$) but take collective effects ($C \neq 0$) and spatial inhomogeneity ($\mathcal{L}_{\text{m.f.}} \neq 0$) into account, we get the generalized Lenard-Balescu equation derived by Heyvaerts (2010) and Chavanis (2012a).

³⁶ The approach of Chandrasekhar takes into account strong collisions with an impact parameter smaller than the Landau length λ_L that yield a deflection at an angle larger than 90° . As we have seen, this can suppress the divergence at small scales. However, the Chandrasekhar approach does not take into account the possibility of forming binary stars which correspond to bound states with strong correlation between particles. In other words, Chandrasekhar only considers hyperbolic trajectories and not elliptical ones in the two-body problem. The effect of binary stars in the kinetic theory requires a special treatment (Heggie 1975).

³⁷ Actually, Landau studies the evolution of the system as a whole while Chandrasekhar studies the relaxation of a test star in a bath of field stars. Landau obtains an integro-differential equation (the Landau equation) that conserves the energy (microcanonical description) while Chandrasekhar obtains a differential equation (the Fokker-Planck or Kramers equation) that does not conserve the energy and involves a fixed temperature (canonical description). However, these two equations are clearly related as explained in Sect. 3.

The ordinary Landau equation ($\mathcal{L}' = \mathcal{L}_{m.f.} = C = 0$) is intermediate between the Boltzmann equation and the generalized Lenard-Balescu (and generalized Landau) equation. It describes the effect of weak collisions but ignores strong collisions, spatial inhomogeneity, and collective effects. It can be obtained from the generalized Lenard-Balescu equation ($\mathcal{L}' = 0$) by making a local approximation ($\mathcal{L}_{m.f.} = 0$) and neglecting collective effects ($C = 0$), or from the Boltzmann equation ($\mathcal{L}_{m.f.} = C = 0$) by considering the limit of small deflections ($\mathcal{L}' = 0$).

Actually, these kinetic equations describe the effect of collisions at different scales (the scale λ may be interpreted as the impact parameter). For $\lambda \sim \lambda_L$ (small impact parameters), the collisions are strong and we must solve the two-body problem exactly. For $\lambda \sim l$ (intermediate impact parameters), the collisions are weak and we can make a weak coupling approximation. For $\lambda \sim \lambda_J$ (large impact parameters), we must take spatial inhomogeneity and collective effects into account. The Boltzmann equation is valid for $\lambda \ll \lambda_J$. It describes strong

collisions ($\lambda \sim \lambda_L$) and weak collisions ($\lambda \sim l$). The generalized Landau and generalized Lenard-Balescu equations are valid for $\lambda \gg \lambda_L$. They describe weak collisions ($\lambda \sim l$), spatial inhomogeneity and collective effects ($\lambda \sim \lambda_J$). The ordinary Landau equation is valid for $\lambda_L \ll \lambda \ll \lambda_J$. It describes weak collisions. When we go beyond the domains of validity of these equations, divergences occur and appropriate cut-offs must be introduced.

Connecting these different limits, we find that the best description of stellar systems is provided by the generalized Lenard-Balescu equation with a small-scale cut-off at the Landau length. If we neglect collective effects, we get the generalized Landau equation (Eqs. (24) and (114)) with a small-scale cut-off at the Landau length. Finally, if we make a local approximation and neglect collective effects, the Vlasov-Landau equation (Eqs. (28) and (29)) with a small-scale cut-off at the Landau length and a large-scale cut-off at the Jeans length provides a relevant description of stellar systems and has the advantage of the simplicity.