

Symplectic coarse graining approach to the dynamics of spherical self-gravitating systems

Luca Barbieri^{1,2,3*}, Pierfrancesco Di Cintio^{4,1}, Guido Giachetti^{5,6}, Alicia Simon-Petit^{1,2}, Lapo Casetti^{1,2,3}

¹*Dipartimento di Fisica e Astronomia and CSDC, Università di Firenze, via G. Sansone 1, I-50019 Sesto Fiorentino, Italy*

²*INFN - Sezione di Firenze, via G. Sansone 1, I-50019 Sesto Fiorentino, Italy*

³*INAF - Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, I-50125 Firenze, Italy*

⁴*ISC-CNR, Via della Lastruccia 10, I-50019 Sesto Fiorentino, Italy*

⁵*SISSA, via Bonomea 265, I-34136 Trieste, Italy*

⁶*INFN - Sezione di Trieste, via Valerio 2, I-34127 Trieste, Italy*

Accepted... Received...; in original form...

ABSTRACT

We investigate the evolution of the phase-space distribution function around slightly perturbed stationary states and the process of violent relaxation in the context of the dissipationless collapse of an isolated spherical self-gravitating system. By means of the recently introduced symplectic coarse graining technique, we obtain an effective evolution equation that allows us to compute the scaling of the frequencies around a stationary state, as well as the damping times of Fourier modes of the distribution function, with the magnitude of the Fourier k -vectors themselves. We compare our analytical results with N -body simulations.

Key words: galaxies: formation – galaxies: evolution – gravitation – methods: analytical – methods: numerical

1 INTRODUCTION

Self-gravitating systems in a non-equilibrium, “cold” (i.e., with small or vanishing kinetic energy K) state are known to undergo a process of “violent relaxation” (hereafter VR, Lynden-Bell 1967), where particle energies are redistributed within few dynamical times t_{dyn} by the strong variations of the self-consistent gravitational potential, followed by a more gentle phase-mixing transient characterized by damped collective oscillations. Such dynamics drives the system towards a long-lived non-thermal (quasi-)equilibrium state, frequently dubbed in the statistical mechanics community as quasi-stationary state (hereafter, QSS; see e.g. Campa et al. 2014; Caglioti & Rousset 2007). The latter, when describing the system in the continuum limit in terms of its phase-space single-particle distribution function $f(\mathbf{r}, \mathbf{p}, t)$, are interpreted as stationary solutions of the collisionless Boltzmann¹ equation (hereafter CBE)

$$\frac{\partial}{\partial t} f + \frac{\mathbf{p}}{m} \cdot \nabla_{\mathbf{r}} f - m \nabla \Phi \cdot \nabla_{\mathbf{p}} f = 0, \quad (1)$$

coupled to the Poisson equation linking mass density ρ and potential Φ

$$\nabla^2 \Phi(\mathbf{r}) = -4\pi G \rho(\mathbf{r}) = -4\pi G \int f \, d\mathbf{p}. \quad (2)$$

Since Lynden-Bell’s seminal paper, VR has been extensively studied with both analytical (Saslaw 1970; Shu 1978; Garcia Lambas et al. 1985; Ossipkov 1985; Tremaine et al. 1986, 1987; Madsen 1987; Kandrup 1987, 1989; Soker 1996; Kull et al. 1997; Kandrup 1998a,b; Mangalam et al. 1999; Arad & Lynden-Bell 2005; Treumann & Baumjohann 2013; Chavanis 2006a,b; Giachetti & Casetti 2019; Beraldo e Silva et al. 2019b; Santini et al. 2021) and numerical approaches (Villumsen 1984; Severne & Luwel 1986; Mineau et al. 1990; Eftymiopoulos & Voglis 2001; Arad & Johansson 2005; Levin et al. 2008a; Di Cintio & Ciotti 2011). Remarkably, due to the same $1/r^2$ behaviour of the interaction, some authors have also investigated the process of VR in plasma physics, in particular in the context of charged particle beams (Kandrup et al. 2001, 2005; Levin et al. 2008b; Pakter et al. 2009; Teles et al. 2009, 2010).

Slowly decaying power-law long-range interactions (of which the $1/r^2$ force is a special case, Di Cintio et al. 2013) have in general non-trivial equilibrium states (Bachelard et al. 2011; Campa et al. 2014; Defenu et al. 2015; Di Cintio et al. 2015; Giachetti et al. 2021) as well as complex evolution paths towards such equilibria. A large body of literature has indeed been devoted to the understanding of VR in systems governed by long-range interactions (see e.g. Antoniazzi et al. 2007; Chavanis 2008; Gabrielli et al. 2010; Casetti & Gupta 2014; Teles et al. 2015; Marcos et al. 2017; Campa & Chavanis 2017;

* E-mail: luca.barbieri@unifi.it

¹ We note that in the jargon of plasma physics, as well as in most of the statistical mechanics literature, the CBE is typically referred to, with a little language abuse, as *Vlasov equation*, see e.g. Hénon (1982). In the present work we stick to the denomination of collisionless Boltzmann equation as it is standard in the astrophysical literature.

Fouvry et al. 2020 and references therein, for an extensive review, see also Levin et al. 2014 or Campa et al. 2014).

From the astrophysical point of view, VR is commonly indicated as the process responsible for the quasi-homology of the structural and phase-space properties of systems such as elliptical galaxies or the end products of gravitational N -body simulations with cold initial conditions (van Albada 1982; Bertin & Stiavelli 1984; Bertin & Trenti 2003; Trenti & Bertin 2005; Trenti et al. 2005; Nipoti et al. 2006; Barber et al. 2012; Di Cintio et al. 2013; Sylos Labini 2013).

In the picture emerging from all these studies, given the time reversibility of the CBE (1), it is still not completely clear what is the dynamical mechanism that damps the collective oscillations during a VR process. Some authors interpret the VR as a non-linear Landau damping (see e.g. Kandrup 1998a, and also Mouhot & Villani 2011). In particular Barré et al. (2010, 2011) analyzed the linear Landau mechanism in the context of simple one dimensional toy models starting from non-homogeneous perturbed CBE equilibrium states. Barré & Yamaguchi (2013) later generalized the non-homogeneous Landau damping formalism to the multi-dimensional case with instantaneously integrable mean field Hamiltonian, concluding that the decay of such perturbed inhomogeneous states is characterized by long time power-law transients towards the QSS.

The analysis of N -body simulations indicates that after the early violent transient, the systems generally evolve in phase-space at smaller and smaller scales. Such behaviour is preserved in the continuum limit, where the associated distribution function f behaves as an incompressible fluid in the one particle phase-space, evolving through a conservative filamentation process. In the common wisdom, it is generally accepted that for asymptotically large times the dynamics at small scales (in both the N -body and continuum pictures), does not affect the evolution of the macroscopic observables. For this reason, in order to formulate a theory of VR, it is sufficient to introduce a coarse grained treatment of the phase-space distribution f , as in Lynden-Bell’s original formulation.

The main challenge in current theories of VR is to produce (or at least prove the existence of) an effective evolution equation linking a given far from equilibrium initial state to its corresponding final QSS. Levin et al. (2008a) and Teles et al. (2011) presented a semi-analytical procedure to evaluate f in the QSS for a class of highly simplified initial conditions (the so-called multiple “water-bags”) coarse grained as a sum of Heavyside theta functions. The distribution function recovered with such method matches fairly well its counterpart extracted from N -body simulations of one dimensional models², when the conditions are such that a characteristic core-halo structure is formed. In addition, the frequency of the first collective oscillations is also reproduced. However, no effective damping mechanism is included in their formalism and therefore the finer structure of f in the QSS is lost. Moreover, the evolution of a broader class of initial conditions can not be treated.

More recently Giachetti et al. (2020) introduced a novel

coarse graining procedure that preserves the symplectic structure of the phase-space³. Using this formalism, they were able to formulate an explicit effective evolution equation for the coarse grained f of one-dimensional systems with general long-range interactions, without making assumptions on the specific form of the initial state f_0 . The theory allows to compute the damping times $\tau_{\mathbf{k}}$ associated to the Fourier modes $f_{\mathbf{k}}(t)$ of the coarse-grained distribution function. Such prediction was successfully tested against numerical simulations of one-dimensional toy models.

In this work we extend the effective evolution equation derived by Giachetti et al. (2020) to higher-dimensional systems whose mean-field Hamiltonian is instantaneously integrable. In this framework we investigate the VR process of a spherical gravitational dissipationless collapse.

The paper is structured as follows: in Section 2 we briefly review symplectic coarse graining as discussed in Giachetti et al. (2020). In Section 3 we generalize it to more dimensions and we derive the effective coarse-grained evolution equation for the phase-space distribution and we evaluate the frequencies and the damping times of the Fourier modes of $f_{\mathbf{k}}$. In section 4 we discuss the numerical simulation set-up and the convergence tests. In Section 5 we discuss our results. Finally, Section 6 summarizes our findings. In order to increase the readability of the paper, the proofs of some results have been reported in Appendices.

2 SYMPLECTIC COARSE GRAINING

In this section we review the symplectic coarse graining procedure introduced by Giachetti et al. (2020). The main idea is to describe VR using an evolution equation which includes an effective dissipative term. The effective evolution equation is based on the following assumptions:

- (i) The normalization of f and the total energy must be conserved during the evolution.
- (ii) The symplectic assumption: The effective evolution equation must be covariant under a canonical change of coordinates or, in other words, it must preserve the underlying symplectic structure of phase-space. This means that the evolution operator has to be a combination of quantities which are invariant under a change of canonical coordinates. The Poisson brackets of any pair of phase-space functions are covariant under a canonical change of coordinates, and therefore, can be used to construct the evolution operator.
- (iii) The mean-field Hamiltonian $H[f]$ must keep the same functional form after the coarse graining procedure:

$$H[f] \rightarrow H[\tilde{f}], \quad (3)$$

where \tilde{f} is the coarse-grained distribution function.

- (iv) The Poisson brackets in the evolution operator must contain only functions of the mean field Hamiltonian $\lambda(H)$, and act linearly on \tilde{f} , as in the CBE (1).

From a physical point of view, assumptions (iii) and (iv) ensure that the dynamics is still ruled by the mean-field potential.

² We note that Martelloni et al. (2016) implemented a semi-analytical procedure to recover the QSS distribution function from water-bag initial states, using a generalized maximum entropy principle.

³ We note that Chavanis & Bouchet (2005) already formulated a coarse grained evolution equation for f though not evidently preserving the symplectic structure.

From all the assumptions above, it follows that the effective evolution equation for \tilde{f} is

$$\frac{\partial \tilde{f}}{\partial t} = \mathcal{L}_H \tilde{f}, \quad (4)$$

where the linear operator \mathcal{L}_H is given by

$$\mathcal{L}_H = \sum_{n=1}^{+\infty} \{\lambda_1(H), \{\lambda_2(H), \{\dots \{\lambda_n(H), \cdot\} \dots\}\}\} \quad (5)$$

and the λ 's are generic (and unknown) functions. Using the Poisson bracket formula

$$\{\lambda_k(H), \cdot\} = \frac{d\lambda_k(H)}{dH} \{H, \cdot\}, \quad (6)$$

it is possible to extract the $\lambda_k(H)$ functions from the nested Poisson brackets (5), obtaining

$$\frac{\partial \tilde{f}}{\partial t} = \{H, \tilde{f}\} + \sum_{n=2}^{+\infty} \mu_n(H) \{H, \{H, \{\dots, \{H, \tilde{f}\} \dots\}\}\}, \quad (7)$$

where $\mu_n(H) = \prod_{i=1}^n d\lambda_i(H)/dH$. Giachetti et al. (2020) made the further assumption that $\mu_1(H) = 1$ in order to reproduce the CBE when the coarse graining scale goes to zero. By doing so, the effective evolution equation for \tilde{f} features Poisson-brackets chains with an odd or an even number of brackets. The chains with odd numbers do not break the time-reversal invariance and simply renormalize the dynamics. The chains with an even number of brackets instead break the time-reversal invariance, thus allowing for a damping mechanism. Moreover, it can be shown that Eq. (7) obeys physical constraints such as energy and normalization conservation, as required by assumption (i). In addition, it can be shown that Equation (7) at the lowest-order truncation

$$\frac{\partial \tilde{f}}{\partial t} = \{H, \tilde{f}\} + \mu_2(H) \{H, \{H, \tilde{f}\}\}, \quad (8)$$

with the assumption $\mu_2(H) \geq 0$, does satisfies the constraints on the monotonic time evolution of convex Casimir invariants, as required by the general result by Tremaine et al. (1986).

The functions μ_k in the above equation can be calculated explicitly for one-dimensional systems. The equation derived from a symplectic coarse-graining for one-dimensional systems in Giachetti et al. (2020) has the general structure (8) with a coefficient $\mu_2(H)$ given by

$$\mu_2(H) = \frac{1}{24} \Delta t (\Delta J)^2 \left(\frac{d\omega(H)}{dH} \right)^2, \quad (9)$$

where Δt and ΔJ are the coarse graining scales and $\omega(H)$ is the frequency of the one-dimensional motion.

3 EFFECTIVE EQUATION FOR INSTANTANEOUSLY INTEGRABLE SYSTEMS

3.1 The coarse graining procedure

We now generalize the symplectic coarse graining procedure to instantaneously integrable systems in a generic physical dimension d . Let us first consider a time-dependent mean-field Hamiltonian with d degrees of freedom assumed to be integrable at every instant of time

$$H = \frac{\mathbf{p}^2}{2} + U(\mathbf{r}, t). \quad (10)$$

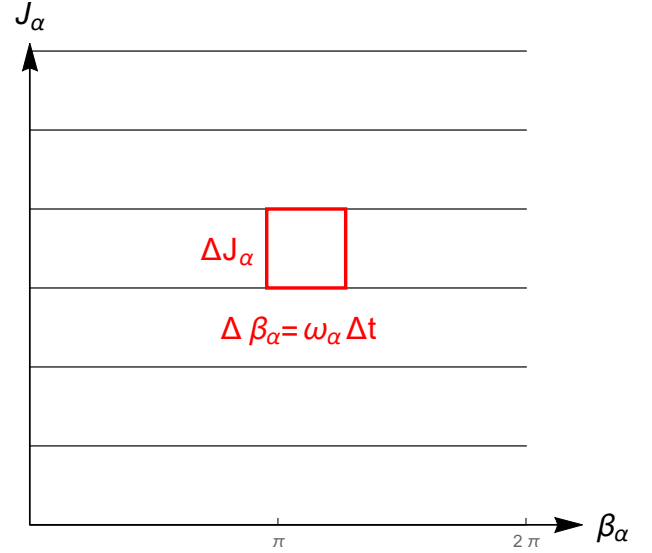


Figure 1. Dynamics of a generic couple (β_α, J_α) . The thin black lines represent the dynamical flow. The square area highlighted by the heavy red lines marks the corresponding coarse-graining volume $\Delta\Gamma_\alpha$.

Due to the time dependence, in principle we can not apply the canonical transformation to action-angle variables. Here we assume that for all the Hamiltonians of interest it is possible to choose a time interval Δt where the time dependence can be neglected. During that time interval, we perform a canonical transformation to action-angle coordinates, to be updated at the subsequent Δt . For this reason, we have dubbed this set of coordinates *comoving action-angle coordinates*.

Let us now integrate the CBE in that time interval Δt . By definition of Δt , we can neglect the time dependence of H and therefore the evolution from t to $t + \Delta t$ becomes

$$f(\mathbf{p}, \mathbf{r}, t + \Delta t) = \exp(\Delta t \{H, \cdot\}) f(\mathbf{p}, \mathbf{r}, t) = U_{\Delta t} f(\mathbf{p}, \mathbf{r}, t). \quad (11)$$

We perform the coarse graining upon the above formulation of the evolution equation. By doing so, the flow associated to the CBE is substituted by its coarse-grained version

$$\tilde{U}_{\Delta t} = \frac{1}{\Delta\Gamma} \int_{\Delta\Gamma} d\Gamma \exp(\Delta t \{H, \cdot\}) = \langle \exp(\Delta t \{H, \cdot\}) \rangle, \quad (12)$$

where $\Delta\Gamma$ is the coarse graining volume in phase-space. Being the system instantaneously integrable, it is possible to perform a canonical change of variables from (\mathbf{r}, \mathbf{p}) to comoving action-angle coordinates $(\boldsymbol{\beta}, \mathbf{J})$. In action-angle space the dynamics is described by the following set of d couples of differential equations

$$\begin{aligned} \dot{J}_\alpha &= 0 \\ \dot{\beta}_\alpha &= \omega_\alpha(\mathbf{J}) \\ \forall \alpha &= 1, \dots, d \end{aligned} \quad (13)$$

where $\boldsymbol{\beta} = (\beta_1, \dots, \beta_d)$ and $\mathbf{J} = (J_1, \dots, J_d)$. The dynamics governed by Eq. (13) is sketched in Figure 1, where for any couple of variables (β_α, J_α) the dynamical flow is represented by straight lines parallel to the axes β_α . In order to preserve the symplectic structure of integrable systems we have chosen the volume element so that it follows the dynamical flow and thus, for any couple (β_α, J_α) , we have chosen the basis of the

coarse-graining volume element to be parallel to the axes β_α with side $\Delta\beta_\alpha$ defined by the motion along such axes,

$$\Delta\beta_\alpha = \omega_\alpha(\mathbf{J})\Delta t. \quad (14)$$

Moreover, the dynamics suggests to choose the height orthogonal to the axes β_α with side ΔJ_α . By doing so, the volume element for a couple of variables (β_α, J_α) is $\Delta\Gamma_\alpha = \Delta J_\alpha \Delta\beta_\alpha$. Repeating the procedure for any couple (β_α, J_α) the total coarse graining volume is finally given by

$$\Delta\Gamma = \prod_{\alpha=1}^d \Delta\Gamma_\alpha = \Delta\beta\Delta\mathbf{J}. \quad (15)$$

Having defined the coarse graining volume, let us then perform the coarse graining using comoving action-angle variables. We first define the coarse-grained quantities using the uniform average over the coarse-graining volume

$$\tilde{A}(\tilde{\mathbf{J}}, \tilde{\beta}) = \langle A(\mathbf{J}, \beta) \rangle = \frac{1}{\Delta\Gamma} \int_{\Delta\Gamma} A(\mathbf{J}, \beta) d\Gamma, \quad (16)$$

$$\langle \exp(-\Delta t \omega(\mathbf{J}) \cdot \nabla_\beta) \rangle_{\Delta\Gamma} \tilde{f}(\tilde{\mathbf{J}}, \tilde{\beta}, t) = \exp\left(\sum_{[m_i]=0}^{+\infty} \frac{(-1)^{\sum_{i=1}^d m_i} (\Delta t)^{\sum_{i=1}^d m_i} \langle \langle \omega_1^{m_1} \dots \omega_d^{m_d} \rangle \rangle}{m_1! \dots m_d!} \partial_{\tilde{\beta}_1}^{m_1} \dots \partial_{\tilde{\beta}_d}^{m_d} \right) \tilde{f}(\tilde{\mathbf{J}}, \tilde{\beta}, t). \quad (19)$$

In Eq. (19) the notation $[m_i]$ means that the sum is performed over the set $m_1 \dots m_d$ from 0 to $+\infty$ without the null terms. Finally, the double bracket $\langle \langle \omega_1^{m_1} \dots \omega_d^{m_d} \rangle \rangle$

where A is a generic function of the phase-space variables. In the equation above $\tilde{\mathbf{J}}$ and $\tilde{\beta}$ are coarse-grained variables (the variables corresponding to the center of the coarse-graining volume) and \mathbf{J} and β are the original ones. We now perform the coarse graining of the CBE (11) using the coarse-grained evolution operator defined in (12) obtaining

$$\tilde{f}(\tilde{\mathbf{J}}, \tilde{\beta}, t + \Delta t) = \langle \exp(-\Delta t \{H, \cdot\}) \rangle \tilde{f}(\tilde{\mathbf{J}}, \tilde{\beta}, t) = \langle \exp(-\Delta t \omega(\mathbf{J}) \cdot \nabla_\beta) \rangle \tilde{f}(\tilde{\mathbf{J}}, \tilde{\beta}, t). \quad (17)$$

We now have to transform equation (17) into a partial differential equation. In order to do that, we use the time translation generator ∂_t writing the left-hand side of equation (17) as

$$\tilde{f}(\tilde{\mathbf{J}}, \tilde{\beta}, t + \Delta t) = \exp(\Delta t \partial_t) \tilde{f}(\tilde{\mathbf{J}}, \tilde{\beta}, t). \quad (18)$$

We can then write the right-hand side of equation (17) using the cumulant expansion as

is the cumulant of order $\sum_{i=1}^d m_i$. We are now able to write down the differential equation for the coarse-grained distribution function equating (18) with (19) as

$$\partial_t \tilde{f}(\tilde{\mathbf{J}}, \tilde{\beta}, t) = \sum_{[m_i]=0}^{+\infty} \frac{(-1)^{\sum_{i=1}^d m_i} (\Delta t)^{\sum_{i=1}^d (m_i-1)} \langle \langle \omega_1^{m_1} \dots \omega_d^{m_d} \rangle \rangle}{m_1! \dots m_d!} \partial_{\tilde{\beta}_1}^{m_1} \dots \partial_{\tilde{\beta}_d}^{m_d} \tilde{f}(\tilde{\mathbf{J}}, \tilde{\beta}, t). \quad (20)$$

To clarify the meaning of this equation, let us write it down separating the different orders in Δt

$$\partial_t \tilde{f}(\tilde{\mathbf{J}}, \tilde{\beta}, t) = - \sum_{i=1}^d \langle \langle \omega_i(\mathbf{J}) \rangle \rangle \partial_{\tilde{\beta}_i} \tilde{f} + \sum_{k=2}^{+\infty} \frac{(\Delta t)^{k-1} (-1)^k}{k!} \sum_{i_1 \dots i_k=1}^d \langle \langle \omega_{i_1}(\mathbf{J}) \dots \omega_{i_k}(\mathbf{J}) \rangle \rangle \partial_{\tilde{\beta}_1}^{m_1} \dots \partial_{\tilde{\beta}_d}^{m_d} \tilde{f}(\tilde{\mathbf{J}}, \tilde{\beta}, t). \quad (21)$$

We note that, as expected from the general discussion in Sec. 2, the coarse graining procedure yields two types of additional terms with respect to the CBE evolution: those with an odd number of derivatives with respect to angles do not break the time-reversal invariance, while those with an even number of derivatives do break such an invariance. As a consequence, summing up all the extra terms, time-reversal invariance is broken and this allows for an effective damping of \tilde{f} . This result extends to multidimensional cases what was already found by Giachetti et al. (2020) for the one-dimensional case.

We now want to understand the differences between the one-dimensional case already studied in Giachetti et al. (2020) and its d -dimensional extension introduced here. In order to do that, we have to write Eq. (21) in a coordinate-independent fashion. In this case, it is not easy to understand the exact covariant structure of the equation at an arbitrary order in Δt ,

therefore we limit ourselves to consider the first non-trivial order,

$$\partial_t \tilde{f}(\tilde{\mathbf{J}}, \tilde{\beta}, t) = - \sum_{i=1}^d \langle \langle \omega_i(\mathbf{J}) \rangle \rangle \partial_{\tilde{\beta}_i} \tilde{f} + \frac{\Delta t}{2} \sum_{i,j=1}^d \langle \langle \omega_i(\mathbf{J}) \omega_j(\mathbf{J}) \rangle \rangle \partial_{\tilde{\beta}_i} \partial_{\tilde{\beta}_j} \tilde{f}(\tilde{\mathbf{J}}, \tilde{\beta}, t). \quad (22)$$

Equation (22) has the structure of a Fokker-Planck equation where the first term represents the drift and the second the diffusion. Let us now calculate the first and the second cumulants of ω . This calculation can be done expanding $\omega_i(\mathbf{J})$ up to second order in $\Delta\mathbf{J} = (\Delta J_1, \dots, \Delta J_d)$ around the action coordinates of the center of the coarse graining volume $\tilde{\mathbf{J}}$ and

then retaining only the first orders, obtaining

$$\begin{aligned} \langle \langle \omega_i(\mathbf{J}) \rangle \rangle &= \frac{1}{\Delta \mathbf{J}} \int_{\Delta \mathbf{J}} \omega_i(\mathbf{J}) d\mathbf{J} = \omega_i(\tilde{\mathbf{J}}) \\ \langle \langle \omega_i(\mathbf{J}) \omega_j(\mathbf{J}) \rangle \rangle &= \langle \omega_i(\mathbf{J}) \omega_j(\mathbf{J}) \rangle - \langle \omega_i(\mathbf{J}) \rangle \langle \omega_j(\mathbf{J}) \rangle = \\ &= \frac{1}{12} \sum_{\alpha=1}^d \frac{\partial \omega_i(\tilde{\mathbf{J}})}{\partial \tilde{J}_\alpha} \frac{\partial \omega_j(\tilde{\mathbf{J}})}{\partial \tilde{J}_\alpha} (\Delta J_\alpha)^2. \end{aligned} \quad (23)$$

The substitution of (23) into (22) yields

$$\begin{aligned} \partial_t \tilde{f}(\tilde{\mathbf{J}}, \tilde{\boldsymbol{\beta}}, t) &= - \sum_{i=1}^d \omega_i(\tilde{\mathbf{J}}) \partial_{\tilde{\beta}_i} \tilde{f} + \\ &+ \frac{\Delta t}{24} \sum_{i,j,\alpha=1}^d \frac{\partial \omega_i(\tilde{\mathbf{J}})}{\partial \tilde{J}_\alpha} \frac{\partial \omega_j(\tilde{\mathbf{J}})}{\partial \tilde{J}_\alpha} (\Delta J_\alpha)^2 \partial_{\tilde{\beta}_i} \partial_{\tilde{\beta}_j} \tilde{f}. \end{aligned} \quad (24)$$

Finally, using the definition of the frequencies $\omega_i = \partial H / \partial \tilde{J}_i$, we can exchange the index i with the index α in the derivative of the frequencies obtaining

$$\begin{aligned} \partial_t \tilde{f}(\tilde{\mathbf{J}}, \tilde{\boldsymbol{\beta}}, t) &= - \sum_{i=1}^d \omega_i(\tilde{\mathbf{J}}) \partial_{\tilde{\beta}_i} \tilde{f} + \\ &+ \frac{\Delta t}{24} \sum_{i,j,\alpha=1}^d \frac{\partial \omega_\alpha(\tilde{\mathbf{J}})}{\partial \tilde{J}_i} \frac{\partial \omega_\alpha(\tilde{\mathbf{J}})}{\partial \tilde{J}_j} (\Delta J_\alpha)^2 \partial_{\tilde{\beta}_i} \partial_{\tilde{\beta}_j} \tilde{f}. \end{aligned} \quad (25)$$

From this expression it is possible to extract a fully covariant form of the evolution equation of \tilde{f} using the definition of Poisson brackets once for the first term and twice for the second term, obtaining

$$\partial_t \tilde{f} = \{H, \tilde{f}\} + \frac{\Delta t}{24} \sum_{\alpha=1}^d (\Delta J_\alpha)^2 \{\omega_\alpha(\mathbf{H}), \{\omega_\alpha(\mathbf{H}), \tilde{f}\}\}, \quad (26)$$

where we have expressed $\omega_\alpha(\tilde{\mathbf{J}}) = \omega_\alpha(\mathbf{H})$ using the fact that the actions $\tilde{\mathbf{J}}$ are functions of the constants of motion \mathbf{H} through the relation $\tilde{\mathbf{J}} = \tilde{\mathbf{J}}(\mathbf{H})$. Now that the equation is fully covariant, using the following identity

$$\{g(\mathbf{H}), \cdot\} = \sum_{i=1}^d \frac{\partial g(\mathbf{H})}{\partial H_i} \{H_i, \cdot\}, \quad (27)$$

we can move the frequencies out of the double Poisson bracket, obtaining

$$\partial_t \tilde{f} = \{H, \tilde{f}\} + \frac{\Delta t}{24} \sum_{i,j,\alpha=1}^d (\Delta J_\alpha)^2 \frac{\partial \omega_\alpha}{\partial H_i} \frac{\partial \omega_\alpha}{\partial H_j} \{H_i, \{H_j, \tilde{f}\}\}. \quad (28)$$

Before considering any physical implication of this equation we note one important thing: the equation does *not* have the general structure (8) as found by Giachetti et al. (2020) following the general assumptions discussed in the previous Section. Indeed, Equation (28) contains not only the Hamiltonian H but all the constants of the motion \mathbf{H} . This is because our coarse graining is constructed in order to preserve the symplectic structure of (instantaneously) integrable systems. In other words, the coarse graining procedure retains all the information that define integrability. The motion of an integrable system is confined on a submanifold of phase-space, i.e., a d -dimensional torus \mathbb{T}^d defined by the intersections of the constants of motion \mathbf{H} (not only the Hamiltonian H). Therefore the extra terms which arise from the coarse graining

must contain all these constants. In conclusion, the structure that was supposed to be valid in the general case is not valid for integrable systems, because of their peculiar geometric structure; we expect the general form (8) to hold when the only integral of motion is the Hamiltonian itself.

To further analyze our effective evolution equation, we consider again its expression in comoving action-angle variables (22) as a Fokker-Planck-like equation with a diffusion mechanism along the torus where the diffusion matrix contains the derivatives of the frequencies. This means that the damping can take place if and only if the frequencies depend⁴ on $\tilde{\mathbf{J}}$. From (13), it is clear that a test particle located on the surface of the inner torus moves with different frequencies with respect to a test particle located on the surface of the outer torus. As a consequence, the damping is generated by a differential rotation between two neighbouring tori. In one dimension, the only independent constant of motion is the Hamiltonian itself, so Equation (28) reduces to (8) with the diffusion coefficient (9), as found in Giachetti et al. (2020). In this case, the mechanism of damping is a differential rotation between level curves of the mean-field Hamiltonian.

3.2 The Jeans theorem, conservation laws and the evolution of Casimirs

Having obtained equation (28), the first thing to look at are its stationary solutions. Since we are dealing with a coarse-grained version of the CBE it is natural to ask which stationary solutions have survived the coarse-graining procedure. The stationary solutions of (28) must satisfy the condition

$$\{H_i, \tilde{f}\} = 0; \quad i = 1, \dots, d. \quad (29)$$

In other words, the stationary solutions are all the possible functions of the constants of the motion. This is exactly the Jeans theorem for the stationary solutions of the CBE, that are therefore preserved also in its coarse-grained version. Note that, however, the coarse-grained \tilde{f} can not vary on length scales smaller than the coarse graining scale; therefore, the stationary solutions of the effective equation (28) are a only subset of the solutions of the full CBE.

Since the main goal of this paper is to apply the symplectic coarse graining to the physically relevant case of three-dimensional spherically symmetric isolated self-gravitating systems, there are two other fundamental conservation laws to be proven (besides the conservation of total energy and of normalization of \tilde{f}): the conservation of total momentum and of total angular momentum, that follow from the translational and rotational invariance of the Hamiltonian, respectively. A coarse graining procedure implies a loss of information below the coarse-graining scale $\Delta \Gamma$ and these symmetries are not necessarily preserved. Nevertheless, considering mean-field Hamiltonians that retain the central symmetry at every instant of time, we can show that our coarse graining procedure preserves total momentum and total angular momentum. The proof can be found in Appendix A.

Tremaine et al. (1987) have shown that after a generic coarse graining procedure the convex Casimirs, i.e., the integrals over

⁴ Indeed, for a system of harmonic oscillators there is no damping, since the frequencies are independent of the actions.

the phase-space of smooth convex functions of the distribution function, must decrease in time⁵. Our coarse-graining procedure has to fulfill such constraint too. In Appendix A we prove that the time derivative of a generic convex Casimir evaluated along the flow generated by Eq. (28) is negative. In Appendix B we give an alternative proof of the argument by Tremaine et al. (1987) that generalizes the one reported in Giachetti et al. (2020) to a generic system of particles in any dimension and with a generic mean-field Hamiltonian (not necessarily instantaneously integrable).

3.3 Frequencies and damping times of Fourier modes

The direct solution of the effective equation (28) is (at least) as difficult as the solution of the CBE, so that, to test the effectiveness of our approach, it is useful to extract from such an equation some general prediction that can be compared with results obtained by N -body numerical simulations. In the following we shall derive two predictions of this kind. First, we shall extend to the d -dimensional case a prediction obtained by Giachetti et al. (2020) for one-dimensional systems, that is, the dependence of the damping time of the coarse-grained distribution function on the coarse-graining scale during the VR process. Second, we shall derive a dispersion relation for the frequencies of the oscillations of the coarse-grained distribution function around a stationary state. Both these prediction will be then tested against N -body simulations.

3.3.1 Damping time of Fourier modes

In order to recover a damping time scale for the coarse-grained distribution function from Eq. (28), we start by rewriting it in comoving action-angle coordinates and treating the resulting formulation as an effective Fokker-Planck equation

$$\begin{aligned} \partial_t \tilde{f}(\mathbf{J}, \boldsymbol{\beta}, t) = & - \sum_{i=1}^d \omega_i(\mathbf{J}) \partial_{\beta_i} \tilde{f} + \\ & + \frac{\Delta t}{24} \sum_{i,j,\alpha=1}^d (\Delta J_\alpha)^2 \frac{\partial \omega_\alpha(\mathbf{J})}{\partial J_i} \frac{\partial \omega_\alpha(\mathbf{J})}{\partial J_j} \partial_{\beta_i} \partial_{\beta_j} \tilde{f}(\mathbf{J}, \boldsymbol{\beta}, t). \end{aligned} \quad (30)$$

To extract a time scale, we perform a double Fourier transform in t and $\boldsymbol{\beta}$, obtaining the following dispersion relation:

$$\tilde{\omega} = -\mathbf{n} \cdot \boldsymbol{\omega}(\mathbf{J}) + i \frac{\Delta t}{24} \sum_{\alpha=1}^d (\Delta J_\alpha)^2 (\mathbf{n} \cdot \nabla_{\mathbf{J}} (\omega_\alpha(\mathbf{J})))^2, \quad (31)$$

where \mathbf{n} is the vector of the Fourier modes of \tilde{f} and $\boldsymbol{\omega}$ is the vector of the associated frequencies. The dispersion relation has a real part due to the underlying collisionless evolution, associated to an oscillatory behaviour, and an imaginary part due to the extra terms appearing after the coarse graining, that introduce a diffusive behaviour. The damping time is then obtained by taking the reciprocal of the term arising from the diffusive term in the definition of $\tilde{\omega}$, as

$$\tau = \frac{1}{\frac{\Delta t}{24} \sum_{\alpha=1}^d (\Delta J_\alpha)^2 (\mathbf{n} \cdot \nabla_{\mathbf{J}} (\omega_\alpha(\mathbf{J})))^2}, \quad (32)$$

⁵ Without any coarse graining all the Casimirs are constants of motion.

It is now natural to ask what is the difference between this case and the one-dimensional case discussed in Giachetti et al. (2020). Let us first write down Δt and ΔJ_α in a coordinate-independent fashion. Redefining $\Delta\Gamma$, the coarse graining volume element in phase-space, as

$$\Delta\Gamma = \prod_{i=1}^d \Delta\beta_i \Delta J_i \propto (\Delta t \Delta J)^d \quad (33)$$

where we have used the conditions (15) and (14), we can set $\Delta t \propto (\Delta\Gamma)^{\frac{1}{2d}}$ and $\Delta J_\alpha \propto (\Delta\Gamma)^{\frac{1}{2d}}$, so that Eq. (32) becomes

$$\tau \propto \frac{1}{(\Delta\Gamma)^{\frac{3}{2d}} \sum_{\alpha=1}^d (\mathbf{n} \cdot \frac{\partial \omega(\mathbf{J})}{\partial J_\alpha})^2}. \quad (34)$$

The expression above yields a finite τ provided at least one of the terms in the sum over the d degrees of freedom at the denominator does not vanish, that is, the frequencies depend on at least one of the actions. In this case, the dependence of τ on the coarse graining volume is

$$\tau \propto (\Delta\Gamma)^{-\frac{3}{2d}}. \quad (35)$$

However, if none of the frequencies depend on the actions, τ given by Eq. (35) diverges, coherently with the fact that in this case no damping can occur. This time scale τ can be thought of as the time after which the dynamics of the fine-grained phase-space distribution function f has moved to scales smaller than $\Delta\Gamma$ in phase-space, so that our coarse-grained description is no longer able to capture it. We note that the dimension d enters in the definition of τ in a rather simple way, and the result found for the one-dimensional case by Giachetti et al. (2020) as $\tau \propto \Delta\Gamma^{-3/2}$ is recovered by setting $d = 1$ in Eq. (35).

A simple way to probe the evolution of the system at different scales in phase space is to evaluate the time dependence of the Fourier components of the distribution function itself, $\tilde{f}_{\mathbf{k}}$, where $\mathbf{k} = (\mathbf{k}_p, \mathbf{k}_r)$, with \mathbf{k}_r and \mathbf{k}_p its components along any couple of canonical coordinates (\mathbf{r}, \mathbf{p}) as

$$\tilde{f}(\mathbf{k}, t) = \int e^{i(\mathbf{k}_p \cdot \mathbf{p} + \mathbf{k}_r \cdot \mathbf{r})} \tilde{f}(\mathbf{p}, \mathbf{r}, t) d\mathbf{p} d\mathbf{r}. \quad (36)$$

A given $\tilde{f}_{\mathbf{k}}$ probes a region in phase-space of size proportional to $\|\mathbf{k}\|^{-1}$, where $\|\mathbf{k}\| = (\|\mathbf{k}_p\|^2 + \|\mathbf{k}_r\|^2)^{1/2}$, in the direction parallel to \mathbf{k} itself, while the size of the probed region in the directions orthogonal to the wave vector does not depend on the wave vector. Therefore, the volume of phase space probed by $\tilde{f}_{\mathbf{k}}$ scales with \mathbf{k} as $\|\mathbf{k}\|^{-1}$, and to describe the evolution of $\tilde{f}_{\mathbf{k}}$ with our coarse graining procedure one has to choose $\Delta\Gamma \propto \|\mathbf{k}\|^{-1}$. Inserting this dependence in Eq. (35) we obtain that $\tilde{f}_{\mathbf{k}}$ should damp out on a time scale

$$\tau_{\mathbf{k}} \propto \|\mathbf{k}\|^{\frac{3}{2d}}. \quad (37)$$

3.3.2 Frequencies of Fourier modes around a QSS

To derive our second prediction, let us consider the effective equation expressed in Cartesian coordinates $\mathbf{x} = (\mathbf{r}, \mathbf{p})$, as

$$\frac{\partial \tilde{f}}{\partial t} = \sum_{n=1}^{2d} A_n(\mathbf{x}, t) \frac{\partial \tilde{f}}{\partial x_n} + \sum_{m,n=1}^{2d} D_{m,n}(\mathbf{x}, t) \frac{\partial^2 \tilde{f}}{\partial x_m \partial x_n}, \quad (38)$$

where the vector \mathbf{A} is defined

$$\mathbf{A} = \begin{pmatrix} A_n^q(\mathbf{x}, t) \\ A_n^p(\mathbf{x}, t) \end{pmatrix} \quad (39)$$

such that its components are

$$\begin{aligned}
 A_n^q(\mathbf{x}, t) &= -p_n + \\
 \frac{1}{24}\Delta t \sum_{\alpha=1}^d (\Delta J_\alpha)^2 \sum_{m=1}^{2d} &\left(-\frac{\partial\omega_\alpha(\mathbf{H})}{\partial q_m} \frac{\partial^2\omega_\alpha(\mathbf{H})}{\partial p_m \partial p_n} + \right. \\
 &\left. + \frac{\partial\omega_\alpha(\mathbf{H})}{\partial p_m} \frac{\partial^2\omega_\alpha(\mathbf{H})}{\partial q_m \partial q_n} \right) \\
 A_n^p(\mathbf{x}, t) &= \frac{\partial U}{\partial q_n} + \\
 \frac{1}{24}\Delta t \sum_{\alpha=1}^d (\Delta J_\alpha)^2 \sum_{m=1}^{2d} &\left(\frac{\partial\omega_\alpha(\mathbf{H})}{\partial q_m} \frac{\partial^2\omega_\alpha(\mathbf{H})}{\partial p_m \partial q_n} + \right. \\
 &\left. - \frac{\partial\omega_\alpha(\mathbf{H})}{\partial p_m} \frac{\partial^2\omega_\alpha(\mathbf{H})}{\partial q_m \partial q_n} \right).
 \end{aligned} \tag{40}$$

The diffusion matrix is

$$\mathbf{D} = \begin{pmatrix} D_{n,m}^{qq} & D_{n,m}^{qp} \\ D_{n,m}^{pq} & D_{n,m}^{pp} \end{pmatrix} \tag{41}$$

expressed in terms of its diffusion coefficients given by

$$\begin{aligned}
 D_{n,m}^{qq} &= \frac{1}{24}\Delta t \sum_{\alpha=1}^d (\Delta J_\alpha)^2 \frac{\partial\omega_\alpha(\mathbf{H})}{\partial p_m} \frac{\partial\omega_\alpha(\mathbf{H})}{\partial p_n} \\
 D_{n,m}^{pp} &= \frac{1}{24}\Delta t \sum_{\alpha=1}^d (\Delta J_\alpha)^2 \frac{\partial\omega_\alpha(\mathbf{H})}{\partial q_m} \frac{\partial\omega_\alpha(\mathbf{H})}{\partial q_n} \\
 D_{n,m}^{qp} &= D_{n,m}^{pq} = \frac{1}{24}\Delta t \sum_{\alpha=1}^d (\Delta J_\alpha)^2 \frac{1}{2} \left(-\frac{\partial\omega_\alpha(\mathbf{H})}{\partial q_m} \frac{\partial\omega_\alpha(\mathbf{H})}{\partial p_n} + \right. \\
 &\left. + \frac{\partial\omega_\alpha(\mathbf{H})}{\partial p_m} \frac{\partial\omega_\alpha(\mathbf{H})}{\partial q_n} \right).
 \end{aligned} \tag{42}$$

As expected, Eq. (38) has the structure of a nonlinear Fokker-Planck equation with a drift vector \mathbf{A} and a diffusion matrix \mathbf{D} . The non-linearity comes from the fact that both \mathbf{A} and \mathbf{D} depend on the coarse grained distribution function \tilde{f} through the mean-field Hamiltonian H . All the diffusive terms in Eq. (38) are, as it should be, time even, in the sense that they break the time-reversal invariance. The drift terms, by contrast, are partly time odd and partly time even. All those terms that do not depend on the coarse-graining scales are time odd, as one would expect from their dependence on the ‘‘Vlasov-like’’ term in (26). The extra terms arising from the coarse-graining procedure are partly time odd and partly time even. In particular, all the contributions from $A_n^p(\mathbf{x}, t)$ are time even, the first addendum of the component $A_n^q(\mathbf{x}, t)$ is time even and the second is time odd, that is

$$\begin{aligned}
 &\sum_{n=1}^{2d} A_{n,te}^q(\mathbf{x}, t) \frac{\partial f}{\partial q_n} = \\
 &= \frac{1}{24}\Delta t \sum_{\alpha=1}^d (\Delta J_\alpha)^2 \sum_{m,n=1}^{2d} \left(\frac{\partial\omega_\alpha(\mathbf{H})}{\partial p_m} \frac{\partial^2\omega_\alpha(\mathbf{H})}{\partial q_m \partial q_n} \right) \frac{\partial f}{\partial q_n}.
 \end{aligned} \tag{43}$$

Since we want to consider the (small) oscillations around a stationary solution of Eq. 1, we can neglect the diffusive terms and all the other time even terms, as they would contribute to the mechanism of relaxation toward a stationary solution but not to the oscillations. With such assumption, the effective

equation is simplified and reduces to

$$\frac{\partial \tilde{f}}{\partial t} = \sum_{n=1}^d A_{n,te}^q(\mathbf{x}, t) \frac{\partial f}{\partial q_n}. \tag{44}$$

Moreover, as we are considering a distribution function evolving around a CBE stationary solution we can neglect the time dependence in $\mathbf{A}(\mathbf{x}, t)$. Since the latter is only through \tilde{f} we are now left with a linear advection equation. From the dimensional analysis of the drift equations we then have

$$\omega_{\|\mathbf{k}\|} = -\mathbf{A}_{n,te}^q \left(\Delta t \sum_{\alpha=1}^d \Delta J_\alpha^2 \right) \cdot \mathbf{k}, \tag{45}$$

where \mathbf{k} is the wave vector in phase-space and where the drift vector $\mathbf{A}_{n,te}^q$ depends linearly on the coarse-graining factor $\Delta t \sum_{\alpha=1}^d \Delta J_\alpha^2$.

Following the same procedure used to derive Eq. (37), we can now impose the condition

$$\Delta t \propto (\Delta\Gamma)^{\frac{1}{2d}} \quad \Delta J_\alpha \propto (\Delta\Gamma)^{\frac{1}{2d}} \quad \Delta\Gamma \propto \|\mathbf{k}\|^{-1} \tag{46}$$

and thus

$$\Delta t \Delta J_\alpha^2 = \|\mathbf{k}\|^{-\frac{3}{2d}}. \tag{47}$$

Using this condition the dispersion relation becomes

$$\omega_{\|\mathbf{k}\|} = \omega_0 + \omega_1 \|\mathbf{k}\|^{(1-\frac{3}{2d})}, \tag{48}$$

that yields the scaling of the frequencies of oscillations around a CBE equilibrium solution $\omega_{\|\mathbf{k}\|}$ with $\|\mathbf{k}\|$.

The scaling laws given by Eqs. (48) and (37) can be tested using numerical simulations.

4 NUMERICAL METHODS

We now turn to the numerical test of the predictions derived in the previous Section.

4.1 The code

In the numerical simulation presented in this work we employed a particle mesh code (see e.g. Hockney & Eastwood 1988; Londrillo & Messina 1990) that solves the Poisson equation

$$\nabla^2 \Phi(\mathbf{r}) = -4\pi G \rho(\mathbf{r}) \tag{49}$$

on a grid in spherical coordinates (r, ϑ, φ) made of $N_r \times N_\vartheta \times N_\varphi$ grid points, and interpolates the acceleration $\mathbf{a}_i = -\nabla\Phi$ at each particle position \mathbf{r}_i .

In order to speed up the calculations and enforce the preservation of the spherical symmetry, we solve only the radial part of Eq. (49), averaging out the contributions in ϑ and φ , similarly to what was done in Pattabiraman et al. (2013); Rodriguez et al. (2018); Di Cintio et al. (2021). By doing so, in practice, the particles’ equations of motion become

$$\ddot{\mathbf{r}}_i = -\frac{GM(r_i)}{r_i^3} \mathbf{r}_i, \tag{50}$$

where $M(r_i)$ is the mass within the particle radial coordinate r_i . When needed, the potential $\Phi(r_i)$ on particle i can be obtained, after having sorted the radial coordinates, as

$$\Phi(r_i) = -G \left(\frac{M(r_i)}{r_i} + \sum_{j=i+1}^N \frac{m_j}{r_j} \right). \tag{51}$$

We note that with such simplification, in addition to the preservation of the system’s spherical symmetry, each particle preserves exactly its angular momentum vector $\mathbf{L}_i = \mathbf{r}_i \times \mathbf{p}_i$, and no radial orbit instability can take place even for highly radially anisotropic initial conditions (Ciotti et al. 2007).

We note also that, imposing in this way the conservation of the spherical symmetry bares a strong resemblance to what done by the so-called *shell models* (Hénon 1964) that solve the dynamics of a system of infinitely thin equal mass rotating spherical shells, under the effect of their mutual and self-gravity and the effective force due to their angular momentum \mathbf{L}_i . The latter systems have an effective three-dimensional phase-space since the state of motion of each “particle” (i.e., shell) is given by its radius r and its radial and tangential velocities $v_{\text{rad}}, v_{\text{tan}}$.

The particles’ equations of motion (50) are solved with the arbitrary order symplectic scheme by Kinoshita et al. (1991) with a fixed timestep Δt . In order to find the optimal value for Δt , we apply the usual criterion (see e.g. Dehnen & Read 2011)

$$\Delta t \equiv \frac{\eta}{\sqrt{\max \nabla^2 \Phi}}, \quad (52)$$

where η is a dimensionless control parameter of the order 10^{-2} . With such choice, we obtain that $0.006 \lesssim \Delta t \lesssim 0.013$, in units of the system’s initial dynamical time defined by

$$t_{\text{dyn}} = \sqrt{\frac{r_s^3}{GM}}. \quad (53)$$

In the expression above r_s is the scale radius of the initial mass distribution, typically of the order of the radius containing half of the system’s mass M . In all the simulations presented in this paper, we have fixed the order of the symplectic integrator to three and used a fixed radial grid of $N_r = 2 \times 10^3$ points extended from 0 to $100r_s$. Particles that occasionally cross the upper border of the grid can be either removed from the simulation (if their individual energy $\mathcal{E}_i > 0$) or further propagated under the effect of the radial acceleration $a_{\text{rad},i} = -GM_{\text{int}}/r_i$, where M_{int} is the mass inside the region covered by mesh.

As a rule, we have extended all our simulations up to $T_{\text{max}} = 1500t_{\text{dyn}}$.

We stress the fact that, using a particle-mesh code rather than a tree- or direct N -body code allows to use a larger number of particles while keeping a relatively low computational time. Moreover, when performing N -scaling sets of simulations, it allows to fix the resolution (i.e. the size of the grid) while varying N . We should point out that, on one hand the accuracy with which a particle-mesh code recovers the gravitational potential is limited by the grid, while on the other hand the potential evaluation in direct codes is affected by the choice of the softening length. In both cases, one may argue that the central (and denser) regions of the systems are not well represented when not using adaptive softening or grid refinement. In this work however, we are interested mainly in the collective behaviour at a coarse scale and therefore neglect the issues that may arise from the resolution. As a test, we performed some additional direct N -body simulations for $N = 10^4$ and identical initial conditions as those used for the particle-mesh runs (see below) finding a rather good agreement between the evolution of the collective quantities such as the virial ratio, the velocity dispersion and anisotropy parameter.

4.2 Initial conditions

We have performed a set of simulations with system size in the interval $10^4 \leq N \leq 3 \times 10^6$. Following Di Cintio et al. (2013) the positions of the N particles are sampled initially from a Plummer (1911) distribution, i.e.,

$$\rho(r) = \frac{3Mr_s^2}{4\pi(r^2 + r_s^2)^{5/2}}. \quad (54)$$

The particles’ initial velocities are always sampled from a position-independent isotropic Maxwell-Boltzmann distribution of unit σ and then renormalized to obtain the desired value of the virial ratio

$$\alpha_{\text{vir}} \equiv 2K/|W|, \quad (55)$$

where K and W are the total kinetic energy and virial function, respectively. The latter is evaluated for a discrete system of N particles as (see e.g. Binney & Tremaine 2008)

$$W = \sum_{i=1}^N m_i(\mathbf{r}_i \cdot \mathbf{a}_i). \quad (56)$$

In addition, in order to compute the oscillation frequencies around a stationary state to be compared with Eq. (48), we have also evolved some systems with equilibrium initial conditions sampled from the isotropic phase-space distribution obtained from Eqs. (1-2) for the density profile (54) via the standard Eddington (1916) (Abel) integral inversion

$$f(\mathcal{E}) = \frac{1}{\sqrt{8\pi^2}} \frac{d}{d\mathcal{E}} \int_{\mathcal{E}}^0 \frac{d\rho}{d\Phi} \frac{d\Phi}{\sqrt{\Phi - \mathcal{E}}}. \quad (57)$$

In the equation above $\mathcal{E} = v^2/2 + \Phi(r)$ is the one-particle energy per unit mass, corresponding to the mean-field Hamiltonian H and where the density ρ and potential Φ are required to be monotonic functions of the radial coordinate r . Note that here the oscillations around the stationary state (57) are those induced by finite size effects of the discrete N -particle distribution, rather than by a perturbation of a continuum stationary f .

4.3 Diagnostics

N -body systems can be always thought of as discretizations of a continuum distribution function. In the discrete case, the putative f can be formally defined as a sum of delta functions, that is,

$$f(\mathbf{r}, \mathbf{p}) \equiv \frac{1}{N} \sum_{i=1}^N \delta^3(\mathbf{r} - \mathbf{r}_i) \delta^3(\mathbf{p} - \mathbf{p}_i). \quad (58)$$

With such a definition, the Fourier mode of given vector $\mathbf{k} = (x, y, z, p_x, p_y, p_x)$ in Eq. (36) becomes

$$f(\mathbf{k}, t) = \frac{1}{N} \sum_{i=1}^N e^{i(\mathbf{k}_r \cdot \mathbf{r}_i(t) + \mathbf{k}_p \cdot \mathbf{p}_i(t))}, \quad (59)$$

where $\mathbf{r}_i(t)$ and $\mathbf{p}_i(t)$ are the time-dependent positions and momenta of the i -th particle.

For all simulations we have computed the amplitude of the fluctuations of the numerically recovered $f_{\mathbf{k}}$ and their frequencies. The time signal of the real part of $f_{\mathbf{k}}$ was rescaled as

$$\bar{f}_{\mathbf{k}} = ||f_{\mathbf{k}}| - \langle f_{\mathbf{k}} \rangle| \quad (60)$$

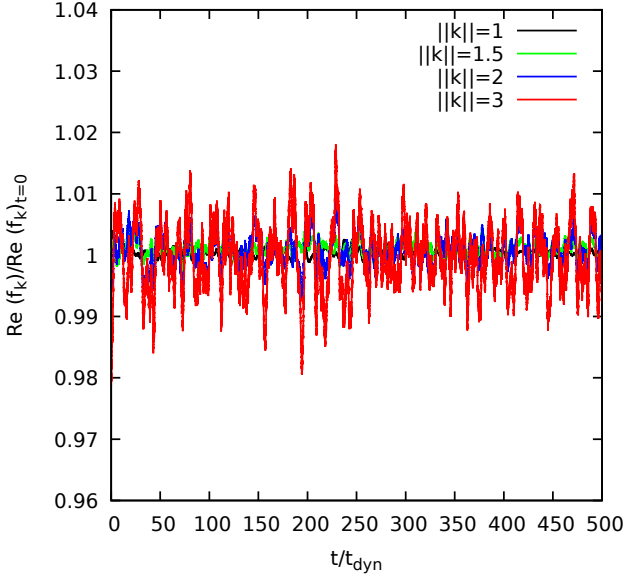


Figure 2. Evolution of the real part of Fourier modes of the numerically recovered distribution function for an equilibrium Plummer model with $N = 10^5$ and for different values $\|\mathbf{k}\| = 1$ (black), $\|\mathbf{k}\| = 1.5$ (green), $\|\mathbf{k}\| = 2$ (blue) and $\|\mathbf{k}\| = 3$ (red).

where

$$\langle f_{\mathbf{k}} \rangle = \frac{1}{T_{\max}} \int_0^{T_{\max}} |f_{\mathbf{k}}| dt \quad (61)$$

and T_{\max} is, again, the total integration time. We then fixed a window between the two times t_a and t_b and evaluated the integral mean $\bar{f}_{\mathbf{k}}$ of $f_{\mathbf{k}}$ and its standard deviation as

$$\begin{aligned} \langle \bar{f}_{\mathbf{k}} \rangle &= \frac{1}{t_b - t_a} \int_{t_a}^{t_b} \bar{f}_{\mathbf{k}}(t) dt; \\ \sigma_{\bar{f}_{\mathbf{k}}}^2 &= \frac{1}{t_b - t_a} \int_{t_a}^{t_b} (\bar{f}_{\mathbf{k}}(t) - \langle \bar{f}_{\mathbf{k}} \rangle)^2 dt. \end{aligned} \quad (62)$$

Typically, in order to minimize the dependence on the initial conditions and the effect of numerical relaxation at late times, we choose $t_a = 10t_{\text{dyn}}$ and $t_b = 10^3 t_{\text{dyn}}$ for all values of N .

To evaluate the damping times $\tau_{\|\mathbf{k}\|}$ associated to a given Fourier mode of f in the N -body simulations with non-equilibrium initial conditions we proceeded as follows. We first define the quantity

$$g_{\mathbf{k}}(t) = \int_0^{T_{\max}} |f_{\mathbf{k}}| dt. \quad (63)$$

The latter, once the system is settled in a QSS after the VR and the subsequent collisionless damping of the collective oscillations, is expected to grow linearly with time, while the first stages of its evolution would be characterized by several slope changes. We fit the linear trend over the whole integration time and define the putative numerical damping time $\tilde{\tau}_{\|\mathbf{k}\|}$ as the time after which the residuals of the linear fit are smaller than a given threshold η . By reducing η iteratively between 10^{-3} and 10^{-7} we found empirically its optimal value at around 10^{-5} as the value below which $\tilde{\tau}_{\|\mathbf{k}\|}$ does not appreciably depend on the specific choice of the initial density profile, number of particles and initial virial ratio.

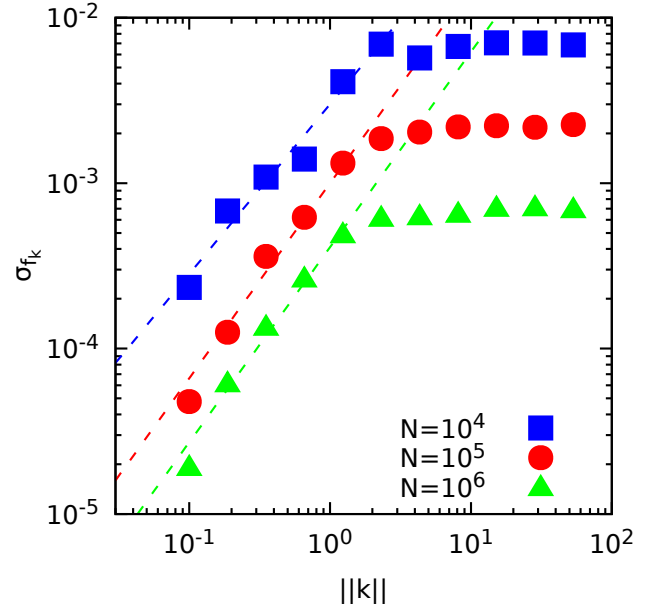


Figure 3. Standard deviation of the fluctuations in $f_{\mathbf{k}}$ as a function of $\|\mathbf{k}\|$ for equilibrium Plummer models and $N = 10^4$ (squares), 10^5 (circles) and 10^6 (triangles).

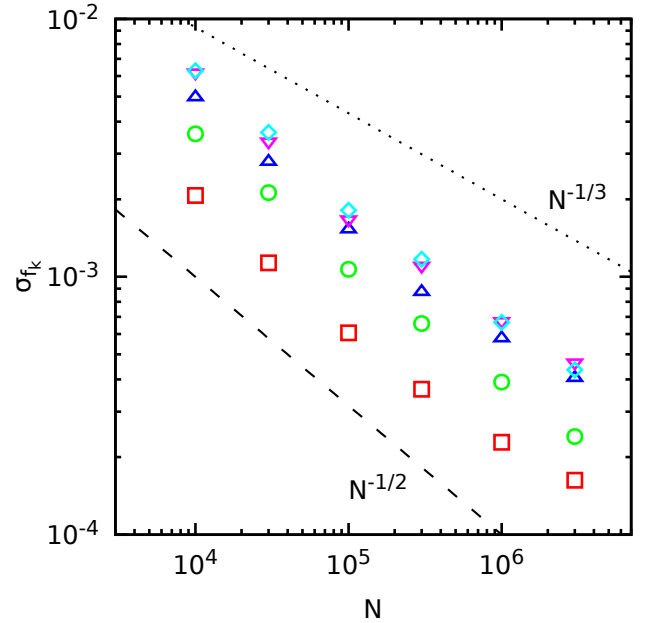


Figure 4. Standard deviation of the fluctuations in $f_{\mathbf{k}}$ as a function of system size N for equilibrium Plummer models and for $\|\mathbf{k}\| \approx 0.1$ (squares), 0.3 (circles), 0.5 (upward triangles), 1 (downward triangles) and 3 (diamonds). The dashed and dotted lines mark the $N^{-1/2}$ and $N^{-1/3}$ trends, respectively.

5 NUMERICAL SIMULATIONS AND RESULTS

5.1 Equilibrium systems and convergence test

Before studying the evolution of the (Fourier modes of the) numerically recovered distribution function of non-equilibrium systems, we have tested its numerical stability for isotropic equilibrium models sampled from Eq. (57) and evaluated its oscillation frequency. For fixed N we find that, as one would

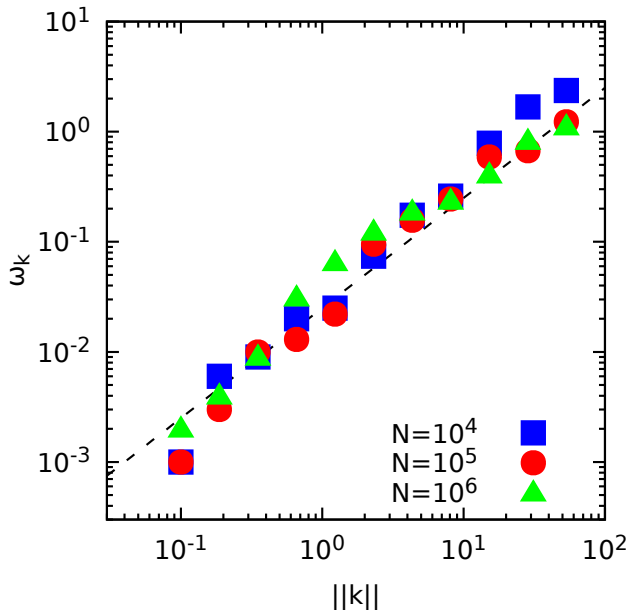


Figure 5. Frequencies $\omega_{\mathbf{k}}$ of the fluctuations of $f_{\mathbf{k}}$ as a function of $\|\mathbf{k}\|$ for equilibrium Plummer models and $N = 10^4$ (squares), 10^5 (circles) and 10^6 (triangles). The dashed line is proportional to $\|\mathbf{k}\|^{1/2}$ up to an additive constant, as predicted by Eq. (48) for $d = 3$.

expect, the amplitude of the fluctuation of the Fourier modes $f_{\mathbf{k}}$ increases for increasing $\|\mathbf{k}\|$, as shown in Fig. 2 for an equilibrium Plummer model with $N = 10^5$ and $\|\mathbf{k}\| = 1, 1.5, 2$ and 3 . This is because, for larger values of the wave number, the discrete Fourier transform (36) probes scales in phase-space occupied by a smaller and smaller fraction of the total number of simulation particles N , thus being plagued by stronger and stronger discreteness effects and Poissonian noise.

In Figure 3 we show the trend of the standard deviation $\sigma_{f_{\mathbf{k}}}$ as defined in Eq. (62), as a function of $\|\mathbf{k}\|$ for Plummer models with total number of particles $N = 10^4, 10^5$ and 10^6 , respectively. Remarkably, at larger wave vectors (small scales in phase-space) the values of $\sigma_{f_{\mathbf{k}}}$ saturate for all N , while for $3 \times 10^{-2} < \|\mathbf{k}\| < 3$ are seemingly well described by a power-law trend with N as marked by the three dashed lines. The reason of such plateau, is due to the fact that at these scales the number of neighbouring particles probed by a given \mathbf{k} becomes vanishing small even for large N and therefore the associated $f_{\mathbf{k}}$ are prone to almost the same discreteness effects. We observed also that for values of $\|\mathbf{k}\| \lesssim 0.03$ (not shown here), the amplitude of the fluctuations saturates again. This is not surprising since at low values of the wave vector the discrete Fourier transform of f is evaluated for scales encompassing the whole phase-space occupied by the system.

In general, at fixed wave-number $\|\mathbf{k}\|$, for increasing N the amplitudes of the fluctuations in $f_{\mathbf{k}}$ decrease. This is apparent from Figure 4 where we reported the standard deviation of the fluctuations of $f_{\mathbf{k}}$ as a function of the number of particles in the interval $10^4 \leq N \leq 3 \times 10^6$ for five choices of $\|\mathbf{k}\|$ between 0.1 and 3. Qualitatively, the trend of $\sigma_{f_{\mathbf{k}}}$ with N is intermediate between a $N^{-1/2}$ and $N^{-1/3}$ power-law behaviour, as shown in the figure by the dashed and dotted lines, respectively. Curiously, the same $N^{-\alpha}$ scaling with $1/3 < \alpha < 1/2$ has

been also reported by Di Cintio & Casetti (2019, 2020b,a) for the maximal Lyapunov exponent and the amplitude of particle-correlations in N -body simulations of collisionless equilibrium models. We speculate that such behaviour could be traced back to a $1/\sqrt{N}$ Poissonian noise associated to the choice of the initial conditions mixing with the underlying $N^{-1/3}$ trend of the collective quantities towards the continuum limit theorized by several studies on the discreteness effects in the large N many body problems (cfr. Gurzadian & Savvidy 1986; Gurzadyan & Kocharyan 2009; Ovod & Ossipkov 2014). A formal derivation of such trend for $\sigma_{f_{\mathbf{k}}}$ is however beyond the scope of the present paper and will be explored elsewhere.

Finally, we have numerically evaluated the frequencies of the oscillations in $f_{\mathbf{k}}$ for the same N -body systems sampled from equilibrium distributions discussed above, by performing a standard Fourier analysis of their time series. In Fig. 5 we show the frequencies of Fourier modes $\omega_{\|\mathbf{k}\|}$ as a function of $\|\mathbf{k}\|$ for equilibrium Plummer models with $N = 10^4, 10^5$ and 10^6 . Remarkably, we observe a trend of $\omega_{\|\mathbf{k}\|}$ of the form

$$\omega_{\mathbf{k}} = \omega_0 + \omega_1 \|\mathbf{k}\|^{1/2} \quad (64)$$

for all the values of N that we have tested (dashed line in Figure 5). Despite the fact that the finite size effects induce such oscillations around an equilibrium distribution, the power-law trend is seemingly independent on N , pointing to the fact that its origin could be a mean-field effect. For this reason, since our effective equation is a coarse-grained description of the CBE, it should also be able to capture this behaviour. As we can see, for $d = 3$ Eq. (48) yields a $\|\mathbf{k}\|^{1/2}$ trend that perfectly matches the one obtained from the N -body simulations.

5.2 Non-equilibrium systems

We have performed a set of N -body simulations starting from initial conditions characterized by $0.2 \leq \alpha_{\text{vir}} \leq 0.9$, Plummer density profile and N ranging from 10^4 up to 3×10^6 . We have followed the evolution of the numerically recovered Fourier modes of the phase-space distribution function for different values of N at fixed initial virial ratio and density profiles. Remarkably, the value attained in the QSS by $f_{\mathbf{k}}$ does not sensibly depend on N nor does the structure of its initial fluctuations during the VR phase as shown in Figure 6 for three values $\|\mathbf{k}\|$ around 1, and different choices of N for a model starting with $\alpha_{\text{vir}} = 0.4$. In Fig. 7 we show the trend of the damping time of the Fourier modes $\tau_{\mathbf{k}}$ as a function of $\|\mathbf{k}\|$. As we can see, the trend saturates for large values of $\|\mathbf{k}\|$ and for small values of $\|\mathbf{k}\|$ regardless of the value of N . For large values of $\|\mathbf{k}\|$, since we are probing small scales in the phase-space, the saturation is due to the finite number N of particles in the system. For small values of $\|\mathbf{k}\|$, since we are probing large scales in the phase-space, the saturation is due to fact that the system occupies a finite volume of the phase-space, and therefore the notion of coarse graining itself becomes immaterial. In practice, for low values of $\|\mathbf{k}\|$ one is probing a phase-space volume enclosing the whole system. Moreover, at such large scales, the numerical evaluation of Eq. (59) is dominated by the contribution of particles on highly elongated orbits that reach large radii with vanishing kinetic energy (velocity) or small radii with large values of radial velocity. Between the two saturation thresholds $\tau_{\mathbf{k}}$ has an increasing trend with $\|\mathbf{k}\|$ that might be described by a

power law, although the window exhibiting such a power-law behaviour is small (although increasing with N). The exponent of the power law extracted from the simulations for intermediate values of $\|\mathbf{k}\|$ seems close to $5/2$, that is, much larger than the theoretical prediction of $1/2$ given by Eq. (37). We do not have a conclusive explanation for the mismatch between the theoretically predicted exponent and the numerically estimated one. We may nonetheless speculate that it might be a finite-size effect, given that a larger number of particles yields a larger power-law window, so that also the numerically estimated exponent might depend on N , although the difference between $1/2$ and $5/2$ seems too large to be due only to such an effect. Another possibility is that the theory does not accurately describe the violent relaxation process ongoing in our simulations. Indeed, the latter is quite fast, occurring on a typical timescale of the order of the dynamical time and with large oscillations of the mean-field potential, so that the assumption of the theory that there is a wide separation between the coarse-graining time scale Δt and the time scale of variation of the mean-field interaction potential, in order to consider the system as integrable for times of the order of Δt , might not be met in this case. This would be coherent with the fact that the theory works well when applied to the equilibrium simulations, where the mean-field interaction is stationary. Finally, one could also argue that the truncation of the effective equation to the second order is the cause of the mismatch between the theoretical prediction and the numerical result. However, as we can see from the most general formulation of the effective evolution equation (21), higher-order terms are expected to become smaller and smaller thus suggesting that the second order approximation should describe well the coarse-grained dynamics. Anyway, it remains to verify how fast do such higher order terms decay in time with k and therefore the possibility that the second order truncation could be at the origin of the failure of our approach in predicting the k -dependence of the damping times can not be ruled out completely.

6 CONCLUSION AND OUTLOOK

In this paper we have extended the formalism of symplectic coarse graining recently introduced by Giachetti et al. (2020). We have obtained an effective equation for a coarse-grained distribution function \tilde{f} valid in any physical dimension d with a mean-field (instantaneously) integrable Hamiltonian. The equation is able to explain the damping mechanism as an effective diffusion along the torus \mathbb{T}^d defined by the condition of integrability.

Our effective evolution equation does not exactly fulfill the general structure argued for in Giachetti et al. (2020) and this is due to the peculiar geometric structure of integrable systems. We have also shown that our formalism preserves macroscopic conservation laws such as those of the total energy and of the normalization of \tilde{f} , and that fulfills the Jeans theorem. Moreover, in agreement with the general argument by Tremaine et al. (1987), it makes convex Casimirs decrease in time. The fact that our equation is valid in any physical dimension d allows us to apply it to the astrophysically relevant case of self-gravitating systems in three dimensions, even though the constraint of instantaneous integrability of the mean-field Hamiltonian binds us to consider three-dimensional systems

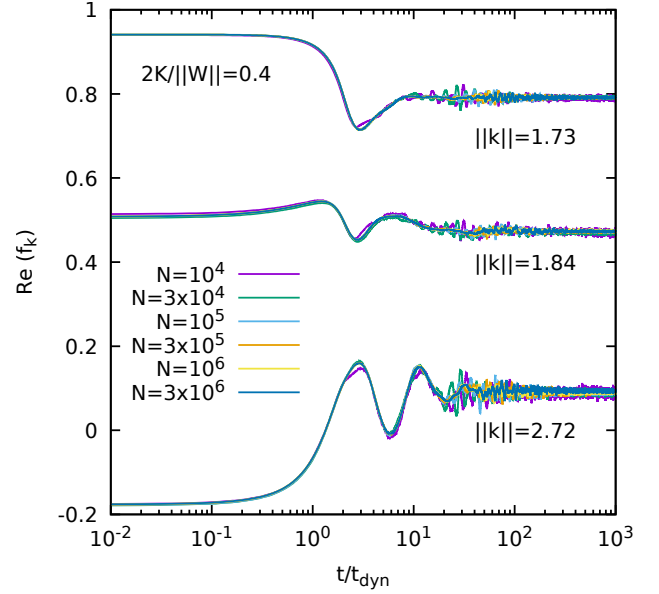


Figure 6. Evolution of the real part of three Fourier modes of the numerically recovered distribution function corresponding to $\|\mathbf{k}\| = 1.73, 1.84$ and 2.72 , for models with initial virial ratio $2K/||W|| = 0.4$ and different numbers of particles in the interval $10^4 \leq N \leq 3 \times 10^6$.

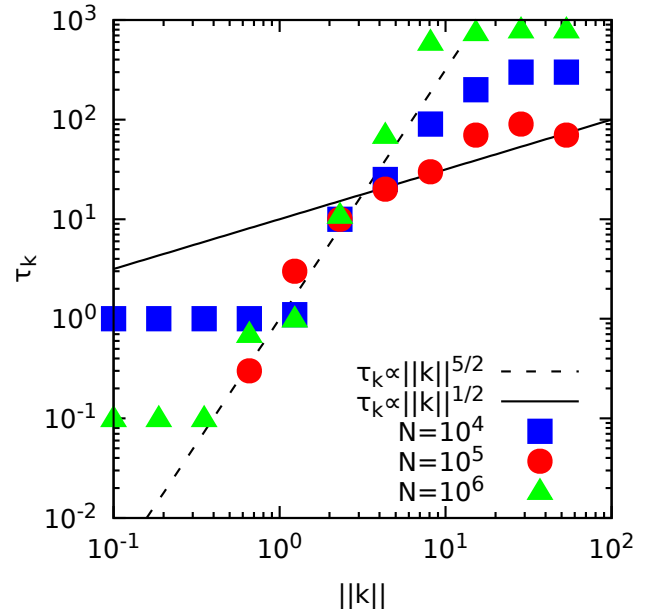


Figure 7. Damping times $\log \tau_{\mathbf{k}}$ as function of the magnitude of the wave vector \mathbf{k} for systems starting with $\alpha_{\text{vir}} = 0.4$, Plummer density profile and $N = 10^4$ (squares), 10^5 (circles) and 10^6 (triangles). The heavy solid line marks the theoretical $\|\mathbf{k}\|^{1/2}$ trend while the dashed line marks the $\|\mathbf{k}\|^{5/2}$ trend recovered from the N -body simulations, respectively.

with enforced spherical symmetry throughout their evolution. In this perspective, in the context of three-dimensional central fields, we proved that our formulation of an effective evolution equation preserves other two fundamental quantities: total angular momentum and total momentum.

Finally, in order to test our theory with numerical simula-

tions, we have extracted two analytical predictions from our theory. First, a prediction on the dependence on the coarse-graining scale of the damping time of the Fourier modes of the distribution function for a system evolving from an out of the equilibrium initial state; second, a prediction on the dependence on the wave vector of the frequencies of oscillation of the distribution function around a stationary solution. We have performed N -body simulations of equilibrium and nonequilibrium spherical self-gravitating systems. In the equilibrium case, we found a very good agreement between the theoretical prediction on the dependence of the frequencies of oscillations around equilibrium on the wave vector and their numerically recovered counterparts. On the contrary, the scaling of the damping time of the Fourier modes of \tilde{f} with k in numerical simulations appears to be plagued by finite N and finite volume effects, even though a seemingly power-law trend is found in a rather narrow window of values of k . The exponent in the relation (37) is found to be $1/2$ for $d = 3$, while that extracted from the simulations approaches $5/2$ for the largest values of N we considered. Possible explanations of this mismatch have been discussed before and may involve finite N effects or the fact that a hypothesis of the theory is not satisfied in the numerically computed cases.

As a natural follow-up of this work, from a theoretical point of view, it would be interesting to generalize the effective equation for the coarse-grained distribution function to non-integrable mean-field Hamiltonians in order to study generic gravitational systems without restrictions imposed by integrability. In this case, since the \mathbb{T}^d structure is broken, we expect a structure of the equation similar to the one-dimensional case, as postulated in Giachetti et al. (2020) (at the lowest truncation order), as given in Eq. (8), where the diffusion coefficient $\mu_2(H)$ should depend on the fundamental time scales of non-integrable systems. In particular, the presence of chaos (see e.g. El-Zant et al. 2019; Beraldo e Silva et al. 2019a; Di Cintio & Casetti 2020b) is expected to come into play, therefore $\mu_2(H)$ should also depend on time scales of the order of the reciprocal of the Lyapunov exponents of the dynamical trajectories of the mean-field Hamiltonian. Another natural follow-up would be a direct comparison between the numerical integration of the CBE and that of the effective evolution equation for the coarse-grained distribution function.

ACKNOWLEDGMENTS

We acknowledge support from Fondazione Cassa di Risparmio di Firenze under the project *HIPERCHEL* for the use of high performance computing resources at the university of Firenze. L.B. acknowledges financial support from the Fondazione Cassa di Risparmio di Firenze under the project *THE SWITCH*. L.B., L.C., P.F.D.C., and A.S.-P. acknowledge financial support from the MIUR-PRIN2017 project *Coarse-grained description for non-equilibrium systems and transport phenomena (CO-NEST)* n. 201798CZL. We thank the Referee for his/her comments that helped improving the presentation of our results.

DATA AVAILABILITY

The simulation data underlying this article will be shared on reasonable request addressed to the corresponding author.

References

- Antoniazzi A., Fanelli D., Barré J., Chavanis P.-H., Dauxois T., Ruffo S., 2007, *Phys. Rev. E*, 75, 011112
- Arad I., Johansson P. H., 2005, *MNRAS*, 362, 252
- Arad I., Lynden-Bell D., 2005, *MNRAS*, 361, 385
- Bachelard R., Dauxois T., de Ninno G., Ruffo S., Staniscia F., 2011, *Phys. Rev. E*, 83, 061132
- Barber J. A., Zhao H., Wu X., Hansen S. H., 2012, *MNRAS*, 424, 1737
- Barré J., Yamaguchi Y., 2013, *Journal of Physics A Mathematical General*, 46, 225501
- Barré J., Olivetti A., Yamaguchi Y. Y., 2010, *Journal of Statistical Mechanics: Theory and Experiment*, 2010, 08002
- Barré J., Olivetti A., Yamaguchi Y. Y., 2011, *Journal of Physics A Mathematical General*, 44, 405502
- Beraldo e Silva L., de Siqueira Pedra W., Valluri M., Sodr e L., Bru J.-B., 2019a, *ApJ*, 870, 128
- Beraldo e Silva L., de Siqueira Pedra W., Valluri M., 2019b, *ApJ*, 872, 20
- Bertin G., Stiavelli M., 1984, *A&A*, 137, 26
- Bertin G., Trenti M., 2003, *ApJ*, 584, 729
- Binney J., Tremaine S., 2008, *Galactic Dynamics: Second Edition*. Princeton University Press
- Caglioti E., Rousset F., 2007, *Journal of Statistical Physics*, 129, 241
- Campa A., Chavanis P.-H., 2017, *Journal of Statistical Mechanics: Theory and Experiment*, 5, 053202
- Campa A., Dauxois T., Fanelli D., Ruffo S., 2014, *Physics of Long-Range Interacting Systems*. Oxford University Press, Oxford
- Casetti L., Gupta S., 2014, *European Physical Journal B*, 87, 91
- Chavanis P.-H., 2006a, *Physica A Statistical Mechanics and its Applications*, 359, 177
- Chavanis P.-H., 2006b, *Physica A Statistical Mechanics and its Applications*, 365, 102
- Chavanis P. H., 2008, in Campa A., Giansanti A., Morigi G., Sylos Labini F., eds, *American Institute of Physics Conference Series Vol. 970, Dynamics and Thermodynamics of Systems with Long Range Interactions: Theory and Experiments*. pp 39–90 ([arXiv:0904.2729](https://arxiv.org/abs/0904.2729)), doi:10.1063/1.2839131
- Chavanis P. H., Bouchet F., 2005, *A&A*, 430, 771
- Ciotti L., Nipoti C., Londrillo P., 2007, in *Collective Phenomena in Macroscopic Systems*. pp 177–186 ([arXiv:astro-ph/0701826](https://arxiv.org/abs/astro-ph/0701826)), doi:10.1142/9789812778901_0018
- Defenu N., Trombettoni A., Codello A., 2015, *Phys. Rev. E*, 92, 052113
- Dehnen W., Read J. I., 2011, *European Physical Journal Plus*, 126, 55
- Di Cintio P., Casetti L., 2019, *MNRAS*, 489, 5876
- Di Cintio P., Casetti L., 2020a, in Bragaglia A., Davies M., Sills A., Vesperini E., eds, *IAUS proceedings series Vol. 351, Star Clusters: From the Milky Way to the Early Universe*. pp 426–429 ([arXiv:1907.12774](https://arxiv.org/abs/1907.12774)), doi:10.1017/S1743921319006744
- Di Cintio P., Casetti L., 2020b, *MNRAS*, 494, 1027
- Di Cintio P., Ciotti L., 2011, *International Journal of Bifurcation and Chaos*, 21, 2279
- Di Cintio P., Ciotti L., Nipoti C., 2013, *MNRAS*, 431, 3177
- Di Cintio P., Ciotti L., Nipoti C., 2015, *Journal of Plasma Physics*, 81, 495810504
- Di Cintio P., Pasquato M., Kim H., Yoon S.-J., 2021, *A&A*, 649, A24
- Eddington A. S., 1916, *MNRAS*, 76, 572
- Efthymiopoulos C., Voglis N., 2001, *A&A*, 378, 679
- El-Zant A. A., Everitt M. J., Kassem S. M., 2019, *MNRAS*, 484, 1456
- Fouvry J.-B., Chavanis P.-H., Pichon C., 2020, *Phys. Rev. E*, 102, 052110
- Gabrielli A., Joyce M., Marcos B., 2010, *Phys. Rev. Lett.*, 105,

210602
 Garcia Lambas D., Mosconi M. B., Sersic J. L., 1985, *Ap&SS*, 113, 89
 Giachetti G., Casetti L., 2019, *Journal of Statistical Mechanics: Theory and Experiment*, 4, 043201
 Giachetti G., Santini A., Casetti L., 2020, *Physical Review Research*, 2, 023379
 Giachetti G., Defenu N., Ruffo S., Trombettoni A., 2021, *Phys. Rev. Lett.*, 127, 156801
 Gurzadian V. G., Savvidy G. K., 1986, *A&A*, 160, 203
 Gurzadyan V. G., Kocharyan A. A., 2009, *A&A*, 505, 625
 Hénon M., 1964, *Annales d’Astrophysique*, 27, 83
 Hénon M., 1982, *A&A*, 114, 211
 Hockney R. W., Eastwood J. W., 1988, *Computer simulation using particles*. Taylor & Francis
 Kandrup H. E., 1987, *MNRAS*, 225, 995
 Kandrup H. E., 1989, *Phys. Rev. A*, 40, 7265
 Kandrup H. E., 1998a, *ApJ*, 500, 120
 Kandrup H. E., 1998b, *Annals of the New York Academy of Sciences*, 858, 28
 Kandrup H., Kishek R., Bohn C., 2001, *Chaotic mixing in charged-particle beams and galaxies*, NASA STI/Recon Technical Report N
 Kandrup H. E., Bohn C. L., Kishek R. A., O’Shea P. G., Reiser M., Sideris I. V., 2005, *Annals of the New York Academy of Sciences*, 1045, 12
 Kinoshita H., Yoshida H., Nakai H., 1991, *Celestial Mechanics and Dynamical Astronomy*, 50, 59
 Kull A., Treumann R. A., Böhringer H., 1997, *ApJ*, 484, 58
 Levin Y., Pakter R., Rizzato F. B., 2008a, *Phys. Rev. E*, 78, 021130
 Levin Y., Pakter R., Teles T. N., 2008b, *Phys. Rev. Lett.*, 100, 040604
 Levin Y., Pakter R., Rizzato F. B., Teles T. N., Benetti F. P. C., 2014, *Phys. Rep.*, 535, 1
 Londrillo P., Messina A., 1990, *MNRAS*, 242, 595
 Lynden-Bell D., 1967, *MNRAS*, 136, 101
 Madsen J., 1987, *ApJ*, 316, 497
 Mangalam A., Nityananda R., Sridhar S., 1999, *ApJ*, 524, 623
 Marcos B., Gabrielli A., Joyce M., 2017, *Phys. Rev. E*, 96, 032102
 Martelloni G., Martelloni G., de Buyl P., Fanelli D., 2016, *Phys. Rev. E*, 93, 022107
 Mineau P., Feix M. R., Rouet J. L., 1990, *A&A*, 228, 344
 Mouhot C., Villani C., 2011, *Acta Math.*, 207, 29
 Nipoti C., Londrillo P., Ciotti L., 2006, *MNRAS*, 370, 681
 Ossipkov L. P., 1985, *Astronomicheskij Tsirkulyar*, 1399, 1
 Ovod D. V., Ossipkov L. P., 2014, *Astronomy Reports*, 58, 746
 Pakter R., Levin Y., Teles T. N., 2009, in *APS Division of Plasma Physics Meeting Abstracts*. p. NP8.011
 Pattabiraman B., Umbreit S., Liao W.-k., Choudhary A., Kalogera V., Memik G., Rasio F. A., 2013, *ApJS*, 204, 15
 Plummer H. C., 1911, *MNRAS*, 71, 460
 Rodriguez C. L., Pattabiraman B., Chatterjee S., Choudhary A., Liao W.-k., Morscher M., Rasio F. A., 2018, *Computational Astrophysics and Cosmology*, 5, 5
 Santini A., Giachetti G., Casetti L., 2021, *Violent relaxation in the Hamiltonian mean field model: II. Non-equilibrium phase diagrams* ([arXiv:2108.11882](https://arxiv.org/abs/2108.11882))
 Saslaw W. C., 1970, *MNRAS*, 150, 299
 Severne G., Luwel M., 1986, *Ap&SS*, 122, 299
 Shu F. H., 1978, *ApJ*, 225, 83
 Soker N., 1996, *ApJ*, 457, 287
 Sylos Labini F., 2013, *MNRAS*, 429, 679
 Teles T. N., Pakter R., Levin Y., 2009, *Applied Physics Letters*, 95, 173501
 Teles T. N., Pakter R., Levin Y., 2010, *Physical Review Accelerators and Beams*, 13, 114202
 Teles T. N., Levin Y., Pakter R., 2011, *MNRAS*, 417, L21

Teles T. N., Gupta S., Di Cintio P., Casetti L., 2015, *Phys. Rev. E*, 92, 020101
 Tremaine S., Hénon M., Lynden-Bell D., 1986, *MNRAS*, 219, 285
 Tremaine S., Hénon M., Lynden-Bell D., 1987, *MNRAS*, 227, 543
 Trenti M., Bertin G., 2005, *A&A*, 429, 161
 Trenti M., Bertin G., van Albada T. S., 2005, *A&A*, 433, 57
 Treumann R. A., Baumjohann W., 2013, *A&A*, 558, A40
 Villumsen J. V., 1984, *ApJ*, 284, 75
 van Albada T. S., 1982, *MNRAS*, 201, 939

APPENDIX A: CONSERVATION LAWS AND CONVEX CASIMIRS

We now prove some properties of the dynamics dictated by Eq. (28).

A1 Conservation of the normalization

Let us prove the conservation of normalization. The normalization is defined by the functional

$$N[\tilde{f}] = \int \tilde{f} d\mathbf{x}, \quad (\text{A1})$$

where the integral is over all phase-space and \mathbf{x} is a generic system of canonical coordinates. Taking the time derivative of equation (A1) and using Eq. (26) we get

$$\begin{aligned} \frac{dN[\tilde{f}]}{dt} &= \int \frac{\partial \tilde{f}}{\partial t} d\mathbf{x} = \\ & \int (\{H, \tilde{f}\} + \frac{\Delta t}{24} \sum_{\alpha=1}^n (\Delta J_{\alpha})^2 \{\omega_{\alpha}(\mathbf{H}), \{\omega_{\alpha}(\mathbf{H}), \tilde{f}\}\}) d\mathbf{x} = 0, \end{aligned} \quad (\text{A2})$$

where in the last step we have used the general formula valid for Poisson brackets

$$\int \{f, g\} d\mathbf{x} = 0, \quad (\text{A3})$$

with f and g generic functions defined in phase space.

A2 Conservation of total energy

To prove that the energy of the system is conserved we consider its functional expression

$$E[\tilde{f}] = \frac{1}{2} \int \mathbf{p}^2 \tilde{f}(\mathbf{r}, \mathbf{p}) d\mathbf{x} + \frac{1}{2} \int \tilde{f}(\mathbf{r}', \mathbf{p}') V(\mathbf{r} - \mathbf{r}') \tilde{f}(\mathbf{r}', \mathbf{p}') d\mathbf{x}', \quad (\text{A4})$$

that satisfies the rule

$$H[\tilde{f}] = \frac{\delta E[\tilde{f}]}{\delta \tilde{f}}, \quad (\text{A5})$$

where $H[\tilde{f}]$ is the mean field Hamiltonian (10). Then, we calculate the time evolution of the energy E as

$$\frac{dE[\tilde{f}]}{dt} = \int \frac{\delta E[\tilde{f}]}{\delta \tilde{f}} \frac{\partial \tilde{f}}{\partial t} d\mathbf{x} = \int H[\tilde{f}] \frac{\partial \tilde{f}}{\partial t} d\mathbf{x}, \quad (\text{A6})$$

where in the second step we have used (A5). Finally, we substitute equation (28) into (A6), obtaining

$$\begin{aligned} \frac{dE[\tilde{f}]}{dt} &= \int H(\{H, \tilde{f}\}) \\ &+ \frac{\Delta t}{24} \sum_{i,j,\alpha=1}^n (\Delta J_\alpha)^2 \frac{\partial \omega_\alpha}{\partial H_i} \frac{\partial \omega_\alpha}{\partial H_j} \{H_i, \{H_j, \tilde{f}\}\} d\mathbf{x} \quad , \end{aligned} \quad (\text{A7})$$

where the first addendum is the contribution of the CBE evolution to the time derivative of E and so must be zero because CBE conserves the energy. The second addendum in (A7), instead, is the contribution of the extra term to the time derivative of E and it gives a zero contribution too. This can be shown as follows:

$$\begin{aligned} &\frac{\Delta t}{24} \sum_{i,j,\alpha=1}^n (\Delta J_\alpha)^2 \int H \frac{\partial \omega_\alpha}{\partial H_i} \frac{\partial \omega_\alpha}{\partial H_j} \{H_i, \{H_j, \tilde{f}\}\} d\mathbf{x} = \\ &\frac{\Delta t}{24} \sum_{i,j,\alpha=1}^n (\Delta J_\alpha)^2 \int \frac{\partial(H\omega_\alpha)}{\partial H_i} \frac{\partial \omega_\alpha}{\partial H_j} \{H_i, \{H_j, \tilde{f}\}\} d\mathbf{x} \\ &- \frac{\Delta t}{24} \sum_{i,j,\alpha=1}^n (\Delta J_\alpha)^2 \int \omega_\alpha \frac{\partial H}{\partial H_i} \frac{\partial \omega_\alpha}{\partial H_j} \{H_i, \{H_j, \tilde{f}\}\} d\mathbf{x} = \\ &\frac{\Delta t}{24} \sum_{i,j,\alpha=1}^n (\Delta J_\alpha)^2 \int \frac{\partial(H\omega_\alpha)}{\partial H_i} \frac{\partial \omega_\alpha}{\partial H_j} \{H_i, \{H_j, \tilde{f}\}\} d\mathbf{x} \quad (\text{A8}) \\ &- \frac{\Delta t}{48} \sum_{i,j,\alpha=1}^n (\Delta J_\alpha)^2 \int \frac{\partial H}{\partial H_i} \frac{\partial(\omega_\alpha)^2}{\partial H_j} \{H_i, \{H_j, \tilde{f}\}\} d\mathbf{x} = \\ &\frac{\Delta t}{24} \sum_{\alpha=1}^n (\Delta J_\alpha)^2 \int \{H\omega_\alpha, \{\omega_\alpha, \tilde{f}\}\} d\mathbf{x} \\ &- \frac{\Delta t}{48} \sum_{\alpha=1}^n (\Delta J_\alpha)^2 \int \{H, \{(\omega_\alpha)^2, \tilde{f}\}\} d\mathbf{x} = 0 \quad , \end{aligned}$$

where in the first and in the second step we have simply used the product rule for derivative, in the third one we have used (27) and in the last one we have used (A3).

A3 Time evolution of convex Casimirs

The last thing left to prove is that convex Casimirs decrease in time because of the presence of the extra terms in the effective equation (28). To prove that, we calculate the time derivative of a Casimir. Consider an arbitrary convex Casimir functional

$$\mathcal{C}[\tilde{f}] = \int C(\tilde{f}) d\mathbf{x}, \quad (\text{A9})$$

where C is any smooth convex function. The time derivative of $\mathcal{C}[\tilde{f}]$ is given by

$$\begin{aligned} \frac{d\mathcal{C}[\tilde{f}]}{dt} &= \int d\mathbf{x} \frac{\partial C(\tilde{f})}{\partial \tilde{f}} \frac{\partial \tilde{f}}{\partial t} = \\ &\int d\mathbf{x} \frac{\partial C(\tilde{f})}{\partial \tilde{f}} (\{H, \tilde{f}\}) \quad (\text{A10}) \\ &+ \frac{\Delta t}{24} \sum_{\alpha=1}^n (\Delta J_\alpha)^2 \{\omega_\alpha(\mathbf{H}), \{\omega_\alpha(\mathbf{H}), \tilde{f}\}\} d\mathbf{x} \quad . \end{aligned}$$

In the last line of Eq. (A10), the first addendum is just the contribution of the CBE evolution to the time variation of a generic Casimir, so it must be zero. The second addendum,

instead, is the time variation of the Casimir due to the coarse graining. Using the Leibniz formula for Poisson brackets,

$$\{f, gh\} = g\{f, h\} + h\{f, g\}, \quad (\text{A11})$$

we have then

$$\begin{aligned} &\frac{\Delta t}{24} \sum_{\alpha=1}^n (\Delta J_\alpha)^2 \int d\mathbf{x} \frac{\partial C(f)}{\partial f} \{\omega_\alpha(\mathbf{H}), \{\omega_\alpha(\mathbf{H}), \tilde{f}\}\} d\mathbf{x} = \\ &\frac{\Delta t}{24} \sum_{\alpha=1}^n (\Delta J_\alpha)^2 \int d\mathbf{x} \{\omega_\alpha(\mathbf{H}), \tilde{f}\} \left\{ \frac{\partial C}{\partial f}, \omega_\alpha \right\} \\ &\quad + \left\{ \frac{\partial C}{\partial f} \{\omega_\alpha, f\}, \omega_\alpha \right\} = \\ &- \frac{\Delta t}{24} \sum_{\alpha=1}^n (\Delta J_\alpha)^2 \int d\mathbf{x} \frac{\partial^2 C}{\partial f^2} (\{\omega_\alpha(\mathbf{H}), \tilde{f}\})^2 \leq 0, \end{aligned} \quad (\text{A12})$$

where in the first step we have used Leibniz rule for Poisson brackets (A11), while in the second step we have used Eq. (A3), the antisymmetry of Poisson brackets and the following identity

$$\left\{ \frac{\partial C(f)}{\partial f}, g \right\} = \frac{\partial^2 C(f)}{\partial f^2} \{f, g\}. \quad (\text{A13})$$

In the last step we have used the convexity of the function $C(f)$.

We have proved that the effective equation conserves energy and normalization and makes convex Casimirs decrease in time. This is true even if the structure of the effective equation (28) is different from the one obtained in Giachetti et al. (2020).

A4 Central force Hamiltonians

Before proving the conservation of central and rotational symmetry and the conservation of total momentum and total angular momentum we briefly summarize the fundamental properties of central force Hamiltonians that will be useful to our purpose. A standard Hamiltonian in three-dimensional space of the form

$$H = \frac{\mathbf{p}^2}{2} + U(\mathbf{r}) \quad (\text{A14})$$

where \mathbf{r} is the position vector, is a central Hamiltonian if the potential $U(\mathbf{r})$ has a central symmetry

$$U(\mathbf{r}) = U(\|\mathbf{r}\|) \quad . \quad (\text{A15})$$

Thanks to this property, we can study the problem using spherical coordinates where $r = \|\mathbf{r}\|$. The Hamiltonian becomes

$$H = \frac{p_r^2}{2} + \frac{1}{2r^2} \left(p_\theta^2 + \frac{p_\phi^2}{\sin^2 \theta} \right) + U(r), \quad (\text{A16})$$

where p_r is the canonical momentum conjugated to r , p_θ is the canonical momentum conjugated to θ and p_ϕ is the canonical momentum conjugated to ϕ . The square of the angular momentum \mathbf{L}^2 , the projection of the angular momentum on the z -axis L_z and the Hamiltonian H are independent constants of motion in involution. Indeed, defining

$$H_1 = p_\phi = L_z, \quad (\text{A17})$$

$$H_2 = p_\theta^2 + \frac{p_\phi^2}{\sin^2 \theta} = \mathbf{L}^2, \quad (\text{A18})$$

$$H_3 = H, \quad (\text{A19})$$

we have

$$\{H_i, H_j\} = 0 \quad \forall i, j = 1, 2, 3. \quad (\text{A20})$$

Thanks to this property the central field Hamiltonians are integrable, and there exists a canonical transformation to action-angle variables such that the transformed Hamiltonian only depends on the actions. The latter are (see e.g. Binney & Tremaine (2008))

$$\begin{aligned} J_\phi &= L_z, \\ J_\theta &= \|\mathbf{L}\| - L_z, \\ J_r &= \frac{1}{\pi} \int_{r_-}^{r_+} dr \sqrt{\left(2(H - V(r)) - \frac{\mathbf{L}^2}{r^2}\right)} = J_r(H, \mathbf{L}^2), \end{aligned} \quad (\text{A21})$$

where r_- and r_+ are defined as the zeros of the argument of the square root as

$$2r^2(E - V(r)) - \mathbf{L}^2 = 0. \quad (\text{A22})$$

Physically r_- and r_+ are the maximum and the minimum radii of an orbit. The Hamiltonian turns out to be a function of J_r and of $J_\theta + J_\phi$,

$$H = h(J_r, \mathbf{L}^2) = h(J_r, (J_\theta + J_\phi)^2). \quad (\text{A23})$$

Being the frequencies defined as

$$\omega_i(\mathbf{H}) = \frac{\partial H}{\partial J_i}, \quad (\text{A24})$$

a simple calculation shows that they only depend on H and \mathbf{L}^2 and do not depend on L_z , a fact related to the rotational invariance of central Hamiltonians. This property will turn out to be relevant to prove the conservation of the total angular momentum in the coarse-grained dynamics.

A5 Preservation of central symmetry and conservation of total momentum

To prove the conservation of total momentum it is enough to prove that the effective equation (28) preserves central symmetry at each time step. So let us consider a time-dependent Hamiltonian that presents a central symmetry at all times,

$$H = \frac{\mathbf{p}^2}{2} + U(\mathbf{r}, t) \quad U(\mathbf{r}, t) = U(-\mathbf{r}, t) \quad \forall t. \quad (\text{A25})$$

We require that, when performing the central symmetry transformation in phase space

$$\begin{aligned} \mathbf{r} &\rightarrow -\mathbf{r}, \\ \mathbf{p} &\rightarrow -\mathbf{p}, \end{aligned} \quad (\text{A26})$$

the initial coarse-grained distribution is left invariant

$$\tilde{f}(\mathbf{r}, \mathbf{p}, 0) = \tilde{f}(-\mathbf{r}, -\mathbf{p}, 0) \quad (\text{A27})$$

Under these hypotheses, we prove that the central symmetry condition for \tilde{f} is preserved by the flow generated by the effective equation. First, we rewrite the effective equation for three-dimensional systems as

$$\frac{\partial \tilde{f}}{\partial t} = \mathcal{L}_{\mathbf{H}} \tilde{f}, \quad (\text{A28})$$

where

$$\mathcal{L}_{\mathbf{H}} = \{H, \cdot\} + \sum_{\alpha=1}^3 \frac{1}{24} \Delta t (\Delta J_\alpha)^2 \{\omega_\alpha(\mathbf{H}), \{\omega_\alpha(\mathbf{H}), \cdot\}\}. \quad (\text{A29})$$

Then, we prove that the operator $\mathcal{L}_{\mathbf{H}}$ is invariant under the central symmetry transformations (A26). Performing a central symmetry operation in phase-space Poisson brackets are left invariant; this is not surprising, because a central symmetry transformation in phase space is a canonical transformation and therefore it leaves Poisson brackets invariant. Moreover, for central field Hamiltonian, we have shown that the frequencies of motion depend only on H and $\|\mathbf{L}\|^2$; as a consequence, the frequencies are left invariant by the central symmetry transformation. Combining all these facts, the operator $\mathcal{L}_{\mathbf{H}}$ acting on \tilde{f} is left invariant by a central symmetry transformation in phase space. Now we can finally prove that the effective evolution preserves the central symmetry of \tilde{f} at every instant of time. In fact, using the Dyson formula it is possible to write down the formal solution of (A28) as

$$\begin{aligned} \tilde{f}(\mathbf{r}, \mathbf{p}, t) &= T \exp\left(\int_0^t dt \mathcal{L}_{\mathbf{H}}(t)\right) \tilde{f}(\mathbf{r}, \mathbf{p}, 0) \\ &= \tilde{f}(\mathbf{r}, \mathbf{p}, 0) \\ &+ \sum_{n=1}^{+\infty} \int_0^t \int_0^{t_1} \cdots \int_0^{t_{n-1}} dt_1 \cdots dt_n \mathcal{L}_{\mathbf{H}}(t_1) \cdots \mathcal{L}_{\mathbf{H}}(t_n) \tilde{f}(\mathbf{r}, \mathbf{p}, 0), \end{aligned} \quad (\text{A30})$$

where T stands for the time-ordered product, that is, $t > t_1 > \cdots > t_n$. Using this formula combined with Eq. (A27) and the fact that $\mathcal{L}_{\mathbf{H}}$ stays central-symmetric we find that $\tilde{f}(t)$ is central-symmetric for all times, that is,

$$\tilde{f}(\mathbf{r}, \mathbf{p}, t) = \tilde{f}(-\mathbf{r}, -\mathbf{p}, t). \quad (\text{A31})$$

The conservation of total momentum immediately follows. Indeed, the total momentum in the continuum picture is

$$\mathbf{P}[\tilde{f}] = \int d\mathbf{x} \mathbf{p} \tilde{f}(\mathbf{x}, t), \quad (\text{A32})$$

and the the last integral is zero at every instant of time because of Eq. (A31), so that an initially zero total momentum stays zero for all times.

A6 Preservation of rotational symmetry and conservation of total angular momentum

Now we prove that the flow generated by the effective equation (A28) preserves rotational symmetry for all times. We consider a time-dependent Hamiltonian that is central at all times, i.e.,

$$H(\mathbf{p}, \mathbf{r}, t) = \frac{\mathbf{p}^2}{2} + U(\mathbf{r}, t) \quad (\text{A33})$$

with

$$U(\mathbf{r}, t) = U(\|\mathbf{r}\|, t) \quad \forall t. \quad (\text{A34})$$

Under a rotational symmetry transformation defined as

$$\begin{aligned} \mathbf{r} &\rightarrow R\mathbf{r}, \\ \mathbf{p} &\rightarrow R\mathbf{p}, \end{aligned} \quad (\text{A35})$$

where R is a generic rotation matrix in three dimensions, we have that H is invariant, that is,

$$H(\mathbf{p}, \mathbf{r}, t) = H(R\mathbf{p}, R\mathbf{r}, t) \quad \forall t. \quad (\text{A36})$$

Moreover, we require that under a rotational symmetry transformation also the initial coarse-grained distribution is left invariant

$$\tilde{f}(\mathbf{r}, \mathbf{p}, 0) = \tilde{f}(R\mathbf{r}, R\mathbf{p}, 0). \quad (\text{A37})$$

Let us now consider the effective equation in three dimensions (A28). Since the rotational symmetry (A35) is a canonical transformation the Poisson brackets are left invariant. Moreover, also the frequencies of motion ω_i are left invariant, because both the Hamiltonian and the modulus of angular momentum are left invariant. Using all these conditions combined with Eq. (A37) and with the Dyson formula (A30) we obtain that the rotational symmetry of \tilde{f} is preserved at all times

$$\tilde{f}(\mathbf{r}, \mathbf{p}, t) = \tilde{f}(R\mathbf{r}, R\mathbf{p}, t). \quad (\text{A38})$$

Now we prove the conservation of the total angular momentum. In the continuum picture, the latter is

$$\mathbf{L}[\tilde{f}] = \int d\mathbf{x} \mathbf{r} \times \mathbf{p} \tilde{f}(\mathbf{x}, t). \quad (\text{A39})$$

The time derivative of \mathbf{L} is

$$\begin{aligned} \frac{d\mathbf{L}[\tilde{f}]}{dt} &= \int d\mathbf{x} \frac{\delta \mathbf{L}}{\delta \tilde{f}} \frac{\partial \tilde{f}}{\partial t} = \int d\mathbf{x} \mathbf{r} \times \mathbf{p} \frac{\partial \tilde{f}}{\partial t} = \int d\mathbf{x} (\mathbf{r} \times \mathbf{p}) \{H, \tilde{f}\} \\ &+ \sum_{\alpha=1}^3 \frac{1}{24} \Delta t (\Delta J_{\alpha})^2 \int d\mathbf{x} (\mathbf{r} \times \mathbf{p}) \{\omega_{\alpha}(\mathbf{H}), \{\omega_{\alpha}(\mathbf{H}), \tilde{f}\}\}. \end{aligned} \quad (\text{A40})$$

The first addendum in the second term of the last equality is the contribution of the CBE evolution to the change in time of \mathbf{L} that has to be zero. The second term can be evaluated using the formula (A3) combined with the fact that \tilde{f} goes to zero at infinity and the Leibniz rule for Poisson brackets (A11) two times, obtaining

$$\begin{aligned} &\int d\mathbf{x} (\mathbf{r} \times \mathbf{p}) \{\omega_{\alpha}(\mathbf{H}), \{\omega_{\alpha}(\mathbf{H}), \tilde{f}\}\} = \\ &= \int d\mathbf{x} \tilde{f} \{\omega_{\alpha}(\mathbf{H}), \{\omega_{\alpha}(\mathbf{H}), \mathbf{r} \times \mathbf{p}\}\}. \end{aligned} \quad (\text{A41})$$

Finally, using Eq. (27) twice, the time derivative of the total angular momentum becomes

$$\frac{d\mathbf{L}[\tilde{f}]}{dt} = \sum_{i,j=1}^3 \int d\mathbf{x} \tilde{f} \frac{\partial \omega_{\alpha}(\mathbf{H})}{\partial H_i} \frac{\partial \omega_{\alpha}(\mathbf{H})}{\partial H_j} \{H_i, \{H_j, \mathbf{r} \times \mathbf{p}\}\}. \quad (\text{A42})$$

We recall that $\mathbf{H} = (\mathbf{L}^2, L_z, H)$ and $\mathbf{r} \times \mathbf{p} = \sum_{i=1}^3 L_i \hat{x}_i$. Using the Poisson brackets commutation relations

$$\{\mathbf{L}^2, L_i\} = 0 \quad \forall i = x, y, z \quad (\text{A43})$$

$$\{H, L_i\} = 0 \quad \forall i = x, y, z \quad (\text{A44})$$

$$\{L_i, L_j\} = \epsilon_{ijk} L_k, \quad (\text{A45})$$

it is easy to see that the only surviving Poisson bracket is the one involving L_z , so that

$$\frac{d\mathbf{L}[\tilde{f}]}{dt} = \int d\mathbf{x} \left(\frac{\partial \omega_{\alpha}(\mathbf{H})}{\partial L_z} \right)^2 \{L_z, \{L_z, \mathbf{r} \times \mathbf{p}\}\}. \quad (\text{A46})$$

The latter expression is zero if $\omega_{\alpha}(\mathbf{H})$ does not depend on L_z . As we have seen in the previous subsection, this is always true for central Hamiltonians and then the total angular momentum is conserved.

APPENDIX B: CASIMIRS EVOLUTION AFTER COARSE GRAINING

Here we will prove with a different approach the general result, already proved by Tremaine et al. (1986), that convex Casimirs decrease in time under a coarse graining of the distribution function in phase space. Our proof is an extension to multidimensional systems of the procedure formulated by Giachetti et al. (2020) for the one-dimensional case. As before, a convex Casimir is defined as

$$\mathcal{C}[\tilde{f}] = \int d\mathbf{x} C(\tilde{f}) \quad (\text{B1})$$

where C is a convex function and \tilde{f} is a coarse-grained distribution defined as

$$\tilde{f}(\mathbf{x}, t) = \frac{1}{\Delta\Gamma} \int_{\Delta\Gamma} d\mathbf{y} f(\mathbf{y}, t), \quad (\text{B2})$$

where $\Delta\Gamma$ is the coarse-graining cell and \mathbf{x} are the phase-space coordinates of the center of the cell. We will prove that convex Casimirs decrease in time, that is

$$\mathcal{C}[\tilde{f}(\mathbf{x}, t + \Delta t)] \leq \mathcal{C}[\tilde{f}(\mathbf{x}, t)]. \quad (\text{B3})$$

The definition (B2) implies that we lose information on scales smaller than the coarse-graining one by averaging the distribution function f . In practice, after a certain time during the evolution, the coarse graining becomes a mean of different values of adjacent small filaments. How can we describe the coarse-grained evolution? To answer this question we first write down the instantaneous evolution of f according to the CBE, using the exponential operator, as

$$f(\mathbf{x}, t + \Delta t) = \exp(-\Delta t \mathbf{v}(\mathbf{x}, t) \cdot \nabla_{\mathbf{x}}) f(\mathbf{x}, t) = U_{\Delta t} f(\mathbf{x}, t), \quad (\text{B4})$$

where we have assumed that for all the Hamiltonian fields $\mathbf{v}(\mathbf{x}, t)$ of interest it is possible to choose a time interval Δt where the time dependence can be neglected. Then, we define the coarse-grained evolution operator as the mean of $U_{\Delta t}$ over the coarse-graining volume,

$$\tilde{U}_{\Delta t} = \frac{1}{\Delta\Gamma} \int d\mathbf{y} U_{\Delta t} = \frac{1}{\Delta\Gamma} \int \exp(-\Delta t \mathbf{v}(\mathbf{y}, t) \cdot \nabla_{\mathbf{x}}) d\mathbf{y}. \quad (\text{B5})$$

In practice, this means that we mean the variation of the norm and of the direction of the Hamiltonian field for scales smaller than the coarse graining volume. Then, the instantaneous evolution for the coarse-grained version of the distribution function becomes

$$\tilde{f}(\mathbf{x}, t + \Delta t) = \tilde{U}_{\Delta t} \tilde{f}(\mathbf{x}, t). \quad (\text{B6})$$

From (B6), we can prove (B3). Let us expand the exponential operator up to first order in Δt , obtaining

$$\tilde{f}(\mathbf{x}, t + \Delta t) = \frac{1}{\Delta\Gamma} \int d\mathbf{y} (\tilde{f}(\mathbf{y}, t) - \Delta t \mathbf{v}(\mathbf{y}, t) \cdot \nabla_{\mathbf{x}} \tilde{f}(\mathbf{y}, t)). \quad (\text{B7})$$

Now, since we are working with a first order precision in Δt we can replace the directional derivative along \mathbf{v} in (B7) with a finite difference approximation,

$$-\Delta t \mathbf{v}(\mathbf{y}, t) \cdot \nabla_{\mathbf{x}} \tilde{f}(\mathbf{x}, t) = \tilde{f}(\mathbf{x} - \mathbf{v}(\mathbf{y}, t) \Delta t, t) - \tilde{f}(\mathbf{x}, t), \quad (\text{B8})$$

so that

$$\begin{aligned} \tilde{f}(\mathbf{x}, t + \Delta t) &= \frac{1}{\Delta \Gamma} \int d\mathbf{y} \tilde{f}(\mathbf{x} - \mathbf{v}(\mathbf{y}, t) \Delta t, t) = \\ &= \langle \tilde{f}(\mathbf{x} - \mathbf{v}(\mathbf{y}, t) \Delta t, t) \rangle_{\Delta \Gamma}. \end{aligned} \quad (\text{B9})$$

Now, using the fact that for every convex function we have

$$C(\langle x \rangle) \leq \langle C(x) \rangle, \quad (\text{B10})$$

replacing x with $\tilde{f}(\mathbf{x}, t + \Delta t)$ in (B10) and using (B9) we have

$$C(\tilde{f}(\mathbf{x}, t + \Delta t)) \leq \langle C(\tilde{f}(\mathbf{x} - \mathbf{v}(\mathbf{y}, t) \Delta t, t)) \rangle_{\Delta \Gamma}. \quad (\text{B11})$$

To calculate the evolution of a generic Casimir we have to integrate this inequality over the phase space, obtaining

$$\begin{aligned} \mathcal{C}[\tilde{f}(\mathbf{x}, t + \Delta t)] &= \int d\mathbf{x} C(\tilde{f}(\mathbf{x}, t + \Delta t)) \leq \\ &\leq \frac{1}{\Delta \Gamma} \int \int d\mathbf{y} d\mathbf{x} C(\tilde{f}(\mathbf{x} - \mathbf{v}(\mathbf{y}, t) \Delta t, t)). \end{aligned} \quad (\text{B12})$$

Finally, using the fact that the integrals with respect to \mathbf{x} are extended over the whole phase space, we can make the following change of variables

$$\mathbf{y}' = \mathbf{y} \quad \mathbf{x}' = \mathbf{x} - \mathbf{v}(\mathbf{y}, t) \Delta t \quad (\text{B13})$$

whose Jacobian is

$$J_{(\mathbf{x}, \mathbf{y})}(\mathbf{x}', \mathbf{y}') = \begin{pmatrix} J_{\mathbf{x}\mathbf{x}'} & J_{\mathbf{y}\mathbf{x}'} \\ J_{\mathbf{x}\mathbf{y}'} & J_{\mathbf{y}\mathbf{y}'} \end{pmatrix} = \begin{pmatrix} \mathbb{I} & J_{\mathbf{y}\mathbf{v}} \\ \mathbb{0} & \mathbb{I} \end{pmatrix} \quad (\text{B14})$$

so that it has a unitary determinant

$$|\det(J_{(\mathbf{x}, \mathbf{y})}(\mathbf{x}', \mathbf{y}'))| = 1. \quad (\text{B15})$$

Thus we obtain

$$\mathcal{C}[\tilde{f}(\mathbf{x}, t + \Delta t)] \leq \frac{1}{\Delta \Gamma} \int \int d\mathbf{y}' d\mathbf{x}' C(\tilde{f}(\mathbf{x}', t)) = \mathcal{C}[\tilde{f}(\mathbf{x}, t)], \quad (\text{B16})$$

that proves the relation (B3). In conclusion, this result explicitly shows that the coarse graining introduces a time arrow in the system, making convex Casimirs decrease during the evolution.