

Manuscrit d'habilitation à diriger les recherches

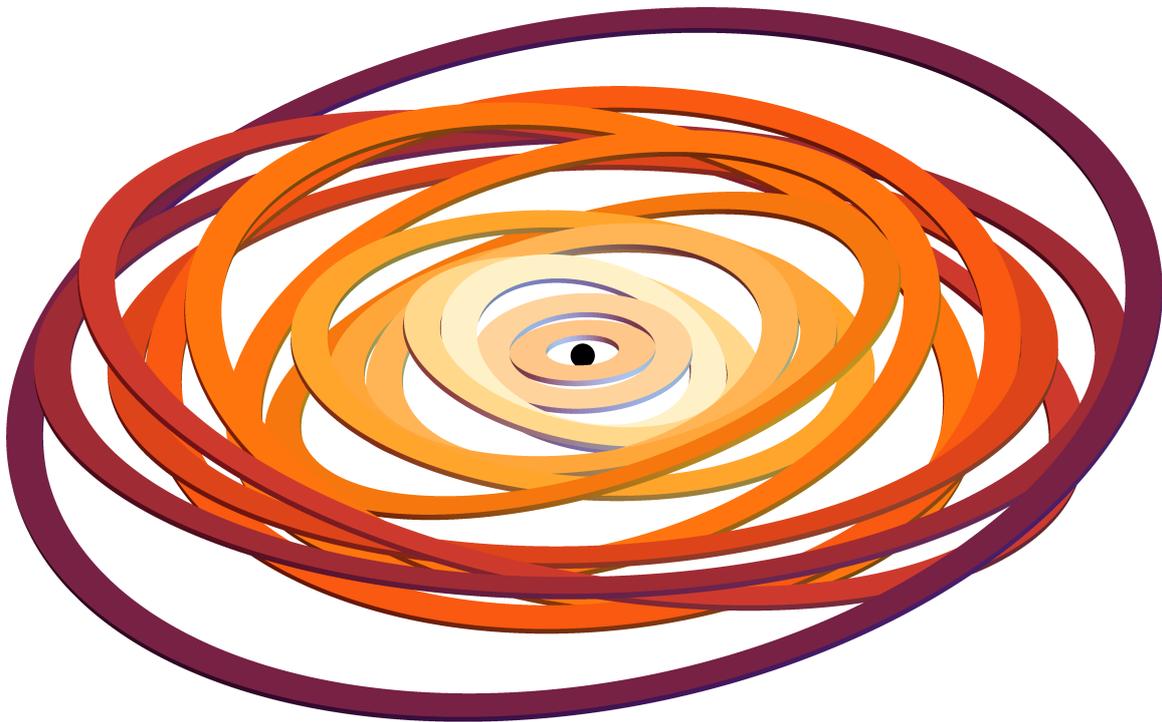
Sorbonne Université

Stellar Dynamics in Galactic Nuclei

Jean-Baptiste Fouvry

fouvry@iap.fr

Institut d'Astrophysique de Paris



Thèse d'habilitation soutenue le 03/10/2022 devant le jury composé de

M.	Stéphane Colombi	Président
M.	Yuri Levin	Rapporteur
M.	Nicholas Stone	Rapporteur
M.	Eugene Vasiliev	Rapporteur
Mme.	Ann Marie Madigan	Examinatrice
Mme.	Smadar Naoz	Examinatrice
M.	Martin Weinberg	Examineur

Abstract

Most galaxies harbor a supermassive black hole in their centre around which orbits a stellar cluster, the galactic nucleus. The unique proximity of the Milky-Way's central black hole, SgrA*, offers an extraordinary opportunity to study such a crowded environment. Although galactic nuclei are among the densest stellar systems in the universe, the steep potential well generated by the central black hole allows for efficient orbital interactions between the stars. Ultimately, this drives the relaxation of the stellar orbits through an intricate hierarchy of dynamical processes. In this manuscript, we focus on two such processes: scalar resonant relaxation through which the stellar orbital eccentricities relax, and vector resonant relaxation through which the stellar orbital orientations get reshuffled. For both processes, we report on recent developments in kinetic theory, devise efficient numerical methods to simulate these dynamics, and present first quantitative applications of these frameworks to constrain SgrA*'s stellar content.

Résumé

La plupart des galaxies abritent en leur centre un trou noir supermassif autour duquel orbite un amas stellaire, le noyau galactique. La proximité unique de SgrA*, le trou noir supermassif au centre de la Voie Lactée, offre une opportunité extraordinaire d'étudier un environnement aussi peuplé. Bien que les noyaux galactiques soient parmi les systèmes stellaires les plus denses de l'univers, le profond puits de potentiel généré par le trou noir central garantit des interactions orbitales très efficaces entre les étoiles. Cela permet la relaxation des orbites stellaires par le biais d'une complexe hiérarchie de processus dynamiques. Dans ce manuscrit, nous nous concentrons sur deux de ces processus: la relaxation résonante scalaire grâce à laquelle les orbites stellaires relaxent en excentricité, et la relaxation résonante vectorielle grâce à laquelle les orientations des orbites stellaires sont redistribuées. Pour ces deux processus, nous décrivons les récents développements de la théorie cinétique, concevons des schémas numériques efficaces pour simuler ces dynamiques, et présentons des premières applications quantitatives de ces paradigmes pour contraindre les propriétés de l'amas stellaire orbitant autour de SgrA*.

Table of contents

1	Stellar dynamics in galactic nuclei	4
1.1	Hamiltonian	4
1.2	Orbit-average	5
1.3	Dynamical processes	6
1.4	Relaxation timescales	7
2	Scalar Resonant Relaxation	9
2.1	Kinetic theory	9
2.1.1	Mean-field dynamics	9
2.1.2	Eccentricity relaxation	9
2.1.3	Resonant diffusion coefficients	10
2.1.4	Non-resonant diffusion coefficients	11
2.1.5	Prospects	12
2.2	Eccentricity relaxation around SgrA*	13
2.2.1	Model's assumptions	13
2.2.2	Methodology	14
2.2.3	Results	14
2.2.4	Prospects	15
2.3	Simulations	16
2.3.1	Equations of motion	16
2.3.2	Rates of changes	18
2.3.3	Time integration	22
2.3.4	Application to SgrA*	25
2.3.5	Discussion	25
2.3.6	Prospects	25
3	Vector Resonant Relaxation	27
3.1	Statistics of VRR	27
3.1.1	VRR Hamiltonian	27
3.1.2	Noise correlation	29
3.1.3	Random walk	29
3.1.4	Noise self-consistency	32
3.1.5	Prospects	33
3.2	Neighbour separation	33
3.2.1	Piecewise Markovian prediction	36
3.2.2	Application	37
3.2.3	Prospects	38
3.3	Thermodynamical equilibrium	39
3.3.1	Equilibrium configurations	39
3.3.2	Self-consistency	40
3.3.3	A typical equilibrium distribution	41
3.3.4	Impact of the initial conditions	42
3.3.5	Prospects	44
3.4	Simulations	45
3.4.1	Equations of motion	45
3.4.2	Rates of change	46
3.4.3	Time integration	47
3.4.4	Application	47
3.4.5	Discussion	48
3.4.6	Prospects	48
3.5	Kinetic theory in axisymmetric clusters	48
3.5.1	Axisymmetric limit	48
3.5.2	Kinetic equation	49
3.5.3	Application	50
3.5.4	Prospects	52
4	Conclusion	53
A	Balescu–Lenard equation	55

1 Stellar dynamics in galactic nuclei

The work presented in this section is based on Fouvry et al. (2022); Lisa Collaboration (2022).

Supermassive black holes (BHs) are ubiquitous in external galaxies (Magorrian et al., 1998; Genzel et al., 2010; Kormendy & Ho, 2013; Graham, 2016) and their induced feedback plays a critical role in regulating galaxy formation through cosmic ages (Heckman & Best, 2014). These supermassive BHs are generically surrounded by a dense population of stars, which forms the nuclear star cluster (Neumayer et al., 2020). Recent outstanding observations keep providing us with new information on these crowded regions. These include detailed census of stellar populations around SgrA* in the centre of the Milky-Way (MW) (Ghez et al., 2008; Gillessen et al., 2017); the characterisation of SgrA*'s clockwise stellar disc (Bartko et al., 2009; Lu et al., 2009; Yelda et al., 2014); the first observation of the relativistic precession of the star S2 within our own Galactic centre (Gravity Collaboration et al., 2020); the observation of a cool accretion disc around SgrA* (Murchikova et al., 2019); the first image of the shadow of SgrA* (Event Horizon Telescope Collaboration et al., 2022); and the mergers of binary BHs recently detected via gravitational wave (GW) emission (Abbott et al., 2019) that may have originated in galactic nuclei.

These various successes will soon be complemented with ever finer resolution around SgrA* (Gravity Collaboration et al., 2017), as well as much larger stellar populations permitted by the planned upgrade on VLTI/GRAVITY (Eisenhauer, 2019; Gravity Collaboration et al., 2021) and upcoming thirty-meter class telescopes such as TMT (Do et al., 2019) and ELT/MICADO (Davies et al., 2018; Pott et al., 2018). Such a wealth of observational information offers new venues to investigate the details of stellar dynamics around supermassive BHs, as well as to probe the possible presence of intermediate mass black holes (IMBHs) in these dense regions (see, e.g., Portegies Zwart & McMillan, 2002).

Although galactic nuclei are the densest stellar systems of the universe, the gravitational potential therein still remains nonetheless dominated by the central supermassive BH. Because of the steep potential it generates, the dynamics of galactic nuclei encompasses a wide range of relaxation processes which act on radically different timescales (Rauch & Tremaine, 1996; Hopman & Alexander, 2006; Alexander, 2017). In this first section, we introduce the key equations on which detailed descriptions of these processes may be devised.

1.1 Hamiltonian

We consider a set of N stars of individual masses m_i orbiting a supermassive BH of mass M_\bullet , assuming that $M_\star = \sum_i m_i \ll M_\bullet$. Within a given inertial frame, we denote the location of the BH with \mathbf{R}_\bullet and the locations of the stars with \mathbf{R}_i . This system's total Hamiltonian is

$$H = \frac{\mathbf{P}_\bullet^2}{2M_\bullet} + \sum_{i=1}^N \left[\frac{\mathbf{P}_i^2}{2m_i} - \frac{GM_\bullet m_i}{|\mathbf{R}_i - \mathbf{R}_\bullet|} \right] + \sum_{i=1}^N m_i \Phi_{\text{GR}}^i - \sum_{i<j}^N \frac{Gm_i m_j}{|\mathbf{R}_i - \mathbf{R}_j|}, \quad (1.1)$$

with the canonical momenta $\mathbf{P}_\bullet = M_\bullet \dot{\mathbf{R}}_\bullet$ and $\mathbf{P}_i = m_i \dot{\mathbf{R}}_i$. In Eq. (1.1), we introduced, for convenience, the effective potential, Φ_{GR} , to account for the (conservative) relativistic corrections induced by the central BH, namely the Schwarzschild and Lense–Thirring precessions (see, e.g., Merritt, 2013).

Because the central BH dominates the potential, it is appropriate to rewrite Eq. (1.1) within new canonical coordinates centred on the BH. An appropriate choice is the democratic coordinates (Duncan et al., 1998)

$$\begin{cases} \mathbf{r}_\bullet = \frac{1}{M_\bullet + M_\star} \left[M_\bullet \mathbf{R}_\bullet + \sum_{i=1}^N m_i \mathbf{R}_i \right], \\ \mathbf{r}_i = \mathbf{R}_i - \mathbf{R}_\bullet, \end{cases} \quad ; \quad \begin{cases} \mathbf{p}_\bullet = \mathbf{P}_\bullet + \sum_{i=1}^N \mathbf{P}_i, \\ \mathbf{p}_i = \mathbf{P}_i - \frac{m_i}{M_\bullet + M_\star} \left[\mathbf{P}_\bullet + \sum_{i=1}^N \mathbf{P}_i \right], \end{cases} \quad (1.2)$$

which keep track of the position of the system's barycentre and the position of the stars w.r.t. the BH. Without loss of generality, we set $(\mathbf{r}_\bullet, \mathbf{p}_\bullet) = (0, 0)$ and Eq. (1.1) ultimately becomes

$$H = \sum_{i=1}^N \left[\frac{\mathbf{p}_i^2}{2m_i} - \frac{GM_\bullet m_i}{|\mathbf{r}_i|} \right] + \sum_{i=1}^N m_i \Phi_{\text{GR}}^i - \sum_{i<j}^N \frac{Gm_i m_j}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2M_\bullet} \left[\sum_{i=1}^N \mathbf{p}_i \right]^2. \quad (1.3)$$

This writing is satisfactory because the first term is N independent Keplerian Hamiltonians, the second is the relativistic corrections induced by the central BH, the third term is the pairwise interactions between the stars, while the last term is the kinetic energy of the central BH. As discussed next, once averaged over the fast Keplerian motions, the first and last terms become irrelevant constants.

1.2 Orbit-average

Because the BH is supermassive, the dominant terms in Eq. (1.3) are the interactions of the stars with the BH. These correspond to a collection of 2-body Keplerian problems, which are explicitly integrable. Therefore, to easily characterise this fast motion, one performs a new change of canonical coordinates towards appropriate angle-action coordinates, $(\boldsymbol{\theta}, \mathbf{J})$, namely the orbital elements or Delaunay variables (Binney & Tremaine, 2008).

As illustrated in the left panel of Fig. 1.1, we define the orbital elements as

$$(\boldsymbol{\theta}, \mathbf{J}) = (M, \omega, \Omega, \Lambda, L, L_z), \quad (1.4)$$

with M the mean anomaly, i.e. the angle associated with the fast orbital motion, ω the argument of the pericentre, and Ω the longitude of the ascending node. The conjugate actions (Λ, L, L_z) are given by

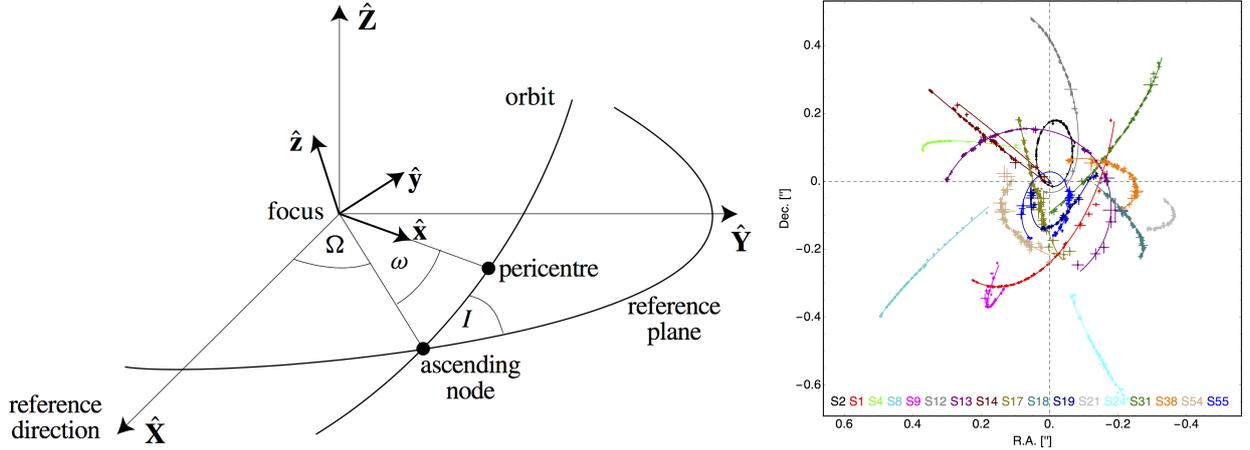


Figure 1.1: Keplerian orbits around a supermassive BH. Left: From Murray & Dermott (1999). Illustration of the Keplerian orbital elements. Right: From Gillessen et al. (2017). Detailed observations of the Keplerian dynamics of the S-stars in the vicinity of SgrA* (~ 10 mpc). At leading order, the orbits take the form of closed ellipses.

$$\Lambda = m\sqrt{GM_{\bullet}a}; \quad L = \Lambda\sqrt{1-e^2}; \quad L_z = L\cos(I), \quad (1.5)$$

with a the orbit's semi-major axis, e its eccentricity, and I its inclination. Here, L stands for the norm of the angular momentum vector, L_z its projection along the z -axis, and Λ the circular angular momentum.

Within Delaunay variables, Eq. (1.3) simply becomes

$$H = H_{\text{Kep}} + H_{\text{GR}} + H_{\star} + \frac{1}{2M_{\bullet}} \left[\sum_{i=1}^N \mathbf{p}_i \right]^2, \quad (1.6)$$

with

$$H_{\text{Kep}} = - \sum_{i=1}^N \frac{m_i^3 (GM_{\bullet})^2}{2\Lambda_i^2}; \quad H_{\text{GR}} = \sum_i m_i \Phi_{\text{GR}}^i; \quad H_{\star} = - \sum_{i<j}^N \frac{Gm_i m_j}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (1.7)$$

Here, H_{Kep} describes the dynamics of N Keplerian orbits, H_{GR} is the BH's relativistic corrections, while H_{\star} captures the pairwise interactions among the stars.

The Hamiltonian H_{Kep} is said to be degenerate because it only depends on the actions $\{\Lambda_i\}_i$. As a result, under this Hamiltonian, all variables but M_i are conserved, and the mean anomaly evolves with the (fast) Keplerian orbital frequency

$$\Omega_{\text{Kep}}(a) = \sqrt{\frac{GM_{\bullet}}{a^3}}. \quad (1.8)$$

This Keplerian motion can equivalently be described using the eccentric anomaly, E , or the true anomaly, f . These follow from Kepler's equation with

$$M = E - e \sin(E); \quad r = a(1 - e \cos(E)) = \frac{a(1 - e^2)}{1 + e \cos(f)}; \quad \frac{dM}{dE} = \frac{r}{a}; \quad \frac{dM}{df} = \frac{r^2}{a^2} \frac{1}{\sqrt{1 - e^2}}, \quad (1.9)$$

and offer simple expressions of the orbital radius, i.e. the distance from the central BH. The right panel of Fig. 1.1 illustrates such Keplerian motions, as currently observed around SgrA*.

If the BH is supermassive, i.e. if $M_\bullet \gg M_\star$, the dynamics is dominated by H_{Kep} and the evolution of all other variables is much slower. We may then average the stellar dynamics over these fast orbital motions, through the so-called secular approximation. One introduces the orbit-averaged Hamiltonian

$$\langle H \rangle = \int_0^{2\pi} \frac{dM_1}{2\pi} \dots \frac{dM_N}{2\pi} H, \quad (1.10)$$

where $\langle \cdot \rangle$ stands for an average over all unperturbed stellar orbits. Phrased differently, stars are smeared out along their unperturbed Keplerian orbits, hence becoming massive wires (see, e.g., Touma et al., 2009).

Since $\langle H \rangle$ is independent of the M_i , the associated actions Λ_i are conserved and H_{Kep} in Eq. (1.6) averages to a constant. Upon expanding, the last term in Eq. (1.6) consists of products of the orbit-averaged stellar momenta $\langle \mathbf{p}_i \rangle$ and the orbit-averaged stellar kinetic energies $\langle \mathbf{p}_i^2 \rangle$. In our barycentric frame, one has $\langle \mathbf{p}_i \rangle = 0$ and $\langle \mathbf{p}_i^2 \rangle = (m_i^2 GM_\bullet / \Lambda_i)^2$ by virtue of the virial theorem. Because constant terms depending only on Λ_i do not induce any dynamics, they can be omitted. Equation (1.6) becomes

$$\langle H \rangle = \langle H_{\text{GR}} \rangle + \langle H_\star \rangle. \quad (1.11)$$

The orbit-averaged Hamiltonian from Eq. (1.11) describes the dynamics of N gravitationally coupled Keplerian wires, also subject to relativistic precession. Each wire is characterised by the five quantities $\{\omega, \Omega, \Lambda, L, L_z\}$, of which Λ is conserved for the orbit-averaged dynamics. Equation (1.11) is the cornerstone on which all the coming sections will build on.

1.3 Dynamical processes

Let us now discuss the key dynamical timescales in galactic nuclei, as illustrated in Fig. 1.2.

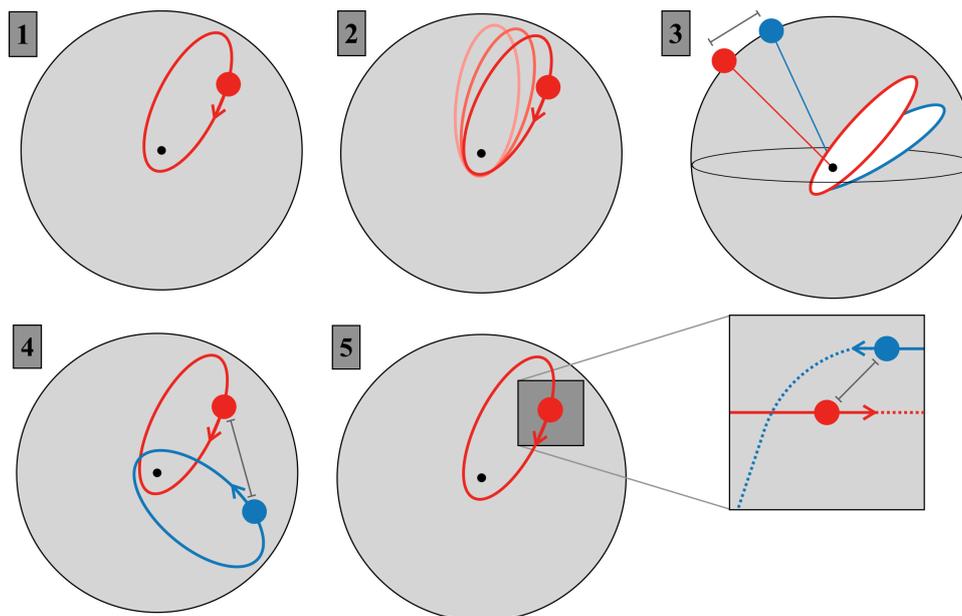


Figure 1.2: Hierarchy of timescales in a galactic nucleus: (1) the dynamical time; (2) the precession time; (3) the vector resonant relaxation time; (4) the scalar resonant relaxation time; (5) the non-resonant relaxation time.

- **Dynamical time.** On account of its mass, the central BH dominates the nucleus' mean potential. At leading order, it therefore imposes a Keplerian motion to any star orbiting within the BH's sphere of influence. These motions take the form of closed ellipses which can be described by their orbital elements, as defined in Eq. (1.4), and are currently monitored for the S-cluster around SgrA*, see the right panel of Fig. 1.1. This Keplerian timescale being so short, e.g., ~ 16 yr for S2 (Gillessen et al., 2017), one is led to smearing out the stars along the Keplerian ellipses, as in Eq. (1.11).
- **Precession time.** On longer timescales, the gravitational potential self-consistently generated by the stellar cluster as well as the relativistic corrections imposed by the BH, namely the Schwarzschild precession (see, e.g., Merritt, 2013) cause the ellipses to precess in their planes. This drives the evolution of ω , e.g., $\sim 3 \times 10^4$ yr for S2 (Gravity Collaboration et al., 2020).
- **Vector resonant relaxation.** Subsequently, because of the non-spherically symmetric fluctuations in the galactic nucleus, as well as because of the relativistic corrections induced by a spinning BH, namely the

Lense–Thirring precession (see, e.g., Merritt, 2013), the ellipses' orbital orientations get reshuffled, e.g., ~ 1 Myr for S2 (Kocsis & Tremaine, 2011). This process is called vector resonant relaxation (VRR) (Kocsis & Tremaine, 2015). In that limit, the orbit's angular momenta change in orientations, $\hat{\mathbf{L}} = \mathbf{L}/|\mathbf{L}|$, without changing in magnitude $|\hat{\mathbf{L}}|$ (equivalently in e), nor in energy Λ (equivalently in a).

- **Scalar resonant relaxation.** On longer timescales, resonant torques between in-plane precessing orbits lead to an efficient diffusion of the ellipses' eccentricity. This process is called scalar resonant relaxation (SRR) (Rauch & Tremaine, 1996), as the quantity that diffuses is the norm of the orbits' angular momenta (i.e. e), e.g., ~ 10 Myr for S2 (Bar-Or & Fouvy, 2018).
- **Non-resonant relaxation.** Finally, on even longer timescales, rather than being driven by the interaction between Keplerian ellipses, a star will see its evolution being driven by nearby pairwise interactions. It is only through the cumulative contributions from these localised scatterings that stars can ultimately relax in Keplerian energy (i.e. in a), through a process called non-resonant relaxation (NR) (Bahcall & Wolf, 1976; Cohn & Kulsrud, 1978; Shapiro & Marchant, 1978; Bar-Or et al., 2013; Vasiliev, 2017). It is generically the slowest relaxation timescale in galactic nuclei, e.g., ~ 1 Gyr for S2 (Kocsis & Tremaine, 2011).

1.4 Relaxation timescales

We now present a first heuristic estimation of the timescales associated with each of these processes (inspired by Rauch & Tremaine, 1996; Alexander, 2017). The fastest timescale in a galactic nucleus is the Keplerian motion induced by the central BH. Following Eq. (1.8), we therefore introduce the Keplerian time

$$T_{\text{Kep}} \simeq \sqrt{\frac{a^3}{GM_{\bullet}}}. \quad (1.12)$$

Because of their self-generated mass potential (and neglecting relativistic corrections for simplicity), Keplerian wires also process within their plane on the precession time

$$T_{\text{p}} \simeq Q T_{\text{Kep}}, \quad (1.13)$$

with $Q = M_{\bullet}/M_{\star} \gg 1$, where $M_{\star} \simeq Nm$ stands for the total enclosed stellar mass.

We now use the angular momentum vector, $\mathbf{L} = L\hat{\mathbf{L}}$, of an arbitrary star as the tracer of relaxation. Following Eq. (1.5), its typical amplitude is $L \simeq m\sqrt{GM_{\bullet}a}$, while the typical torque between two stars is Gm^2/a . Assuming that the system is on average symmetric, only Poisson shot noise from the other stars can source the dynamics of $\hat{\mathbf{L}}$. As such, the typical instantaneous torque felt by a given star is then

$$\tau = \sqrt{N} \frac{Gm^2}{a}. \quad (1.14)$$

Once this torque estimated, we introduce the ballistic time, T_{bal} , as the time needed for \mathbf{L} to change by a non-trivial amount for a given and fixed torque τ . One has

$$T_{\text{bal}} \simeq \frac{L}{\tau} \simeq \sqrt{N} Q T_{\text{Kep}}. \quad (1.15)$$

We also introduce the coherence time, T_{c} , as the time required for τ to change by a non-trivial amount. Phrased differently, recalling that τ is a stochastic variable, the coherence time corresponds to the characteristic decay time of the correlation function $t \mapsto \langle \tau(t) \tau(0) \rangle$, i.e. how long one should wait for the system to have reshuffled enough so that $\tau(t)$ is not correlated anymore with its initial value, $\tau(0)$. It is important to stress that T_{c} depends on the considered relaxation process, as we now detail:

- For VRR, the randomisation is due to the stars' out-of plane orientation dynamics. Since there is no mean potential that drives it, it is only the noise, i.e. the stochastic torque τ itself, that can get the stars to move out of their planes. As such, the associated coherence time, $T_{\text{c}}^{\text{VRR}}$, is proportional to the ballistic time, T_{bal} .
- For SRR, the loss of coherence in the noise arises as soon as wires have sufficiently precessed within their orbital planes. The decoherence is therefore driven by the "constructive" in-plane precession, so that the coherence time, $T_{\text{c}}^{\text{SRR}}$, scales like the precession time, T_{p} .
- For NR, nearby velocity deflections typically last T_{Kep} , so that the noise decorrelates every T_{Kep} , i.e. after every collision. Yet, since one must integrate over all impact parameters, this boosts the coherence time by an amount $\ln(\Lambda)$, the Coulomb logarithm (see, e.g., §7.4.4 in Binney & Tremaine, 2008).

All in all, we obtain the following coherence times

$$T_c^{\text{VRR}} \simeq T_{\text{bal}}; \quad T_c^{\text{SRR}} \simeq T_p; \quad T_c^{\text{NR}} \simeq \ln(\Lambda) T_{\text{Kep}}. \quad (1.16)$$

We are now set to estimate the typical timescale for each of these relaxation processes. We consider a generic diffusion equation sourced by a time-correlated noise. Up to $t \simeq T_c$ of each process, one is in the ballistic regime. The torque can be considered constant, and the angular momentum typically increases by the amount

$$\Delta L|_{\text{bal}} \simeq \tau T_c \simeq L \frac{T_c}{T_{\text{bal}}}, \quad (1.17)$$

following Eq. (1.15). On longer timescales, $t \gg T_c$, the random walk undergone by \mathbf{L} can be mimicked as an usual Brownian motion. More precisely, every T_c , \mathbf{L} undergoes a random kick of amplitude $\Delta L|_{\text{bal}}$. Adding these kicks in quadrature, after a time t the angular momentum has therefore changed by an amount

$$\Delta L(t) \simeq \Delta L|_{\text{bal}} \sqrt{\frac{t}{T_c}} \simeq L \sqrt{\frac{t}{T_{\text{bal}}^2/T_c}}. \quad (1.18)$$

Overall, we therefore get that the typical diffusion time, T , required for the angular momentum to change by a non-trivial amount reads

$$T = \frac{T_{\text{bal}}^2}{T_c}. \quad (1.19)$$

This is an important relation. Equation (1.16) finally gives us the estimates of the timescales

$$T_p \simeq Q T_{\text{Kep}}; \quad T_{\text{VRR}} \simeq \sqrt{N} Q T_{\text{Kep}}; \quad T_{\text{SRR}} \simeq N Q T_{\text{Kep}}; \quad T_{\text{NR}} \simeq N \frac{Q^2}{\ln(\Lambda)} T_{\text{Kep}}, \quad (1.20)$$

with $Q = M_{\bullet}/M_{\star}$. We recover the expected ordering of timescales $T_{\text{Kep}} \ll T_p \ll T_{\text{VRR}} \ll T_{\text{SRR}} \ll T_{\text{NR}}$, as hinted in Fig. 1.2. A goal of the coming sections is to develop appropriate kinetic frameworks to obtain more precise and quantitative estimates for these different dynamical processes, in particular SRR (§2) and VRR (§3).

2 Scalar Resonant Relaxation

The work presented in this section is based on Bar-Or & Fouvry (2018); Tep et al. (2021); Fouvry et al. (2022).

In this section, our focus is SRR, the most efficient mechanism to change a star’s angular momentum in a galactic nucleus. We first place this process on firm theoretical grounds (§2.1) and then investigate its role in explaining the thermal distribution of stellar eccentricities observed around SgrA* (§2.2). Finally, we devise an efficient multipole integration scheme for the dynamics of Keplerian wires around a supermassive BH (§2.3).

2.1 Kinetic theory

Following the seminal work from Rauch & Tremaine (1996), SRR lacked a formal self-consistent description for many years. Previous attempts at modeling this process used either ad hoc methods (Hopman & Alexander, 2006; Eilon et al., 2009; Madigan et al., 2011; Antonini & Merritt, 2013; Merritt, 2015), or N -body simulations (Perets et al., 2009; Merritt et al., 2011; Hamers et al., 2014). Only recently have several studies (Bar-Or & Alexander, 2014; Sridhar & Touma, 2016; Fouvry et al., 2017) laid the foundations for a self-consistent theory of SRR. In this section, we build upon these works, following Bar-Or & Fouvry (2018). We show that for an isotropic spherical system, SRR can quantitatively be described as a diffusion process, and derive the associated diffusion coefficients from first principles.

2.1.1 Mean-field dynamics

We consider a test star of mass m_* on a Keplerian orbit of semi-major axis a and eccentricity e around a supermassive BH of mass M_\bullet , embedded within a spherically symmetric stellar cluster of mean density $\rho(r)$ and isotropic velocity distribution. We track the long-term evolution of the star’s eccentricity through the dimensionless angular momentum

$$h = \sqrt{1 - e^2}. \quad (2.1)$$

Following an orbit-average over the fast Keplerian motion (Eq. 1.10), the test wire undergoes on longer timescales an in-plane precession described by

$$\frac{d\omega}{dt} = \Omega_p(a, h) = \Omega_{GR}(a, h) + \Omega_M(a, h). \quad (2.2)$$

Here, $\Omega_p(a, h)$ describes the total precession frequency of the wire’s pericentre. It is sourced by the relativistic corrections from the central BH, i.e. the Schwarzschild precession (Merritt, 2013), $\Omega_{GR}(a, h)$, as well as by the mass precession imposed by the mean background stellar cluster, $\Omega_M(a, h)$.

More precisely, the relativistic precession reads

$$\Omega_{GR}(a, h) = 3 \frac{r_g}{a} \frac{1}{h^2} \Omega_{Kep}(a), \quad (2.3)$$

with the (fast) Keplerian frequency, $\Omega_{Kep}(a)$ (Eq. 1.8), as well as the gravitational radius $r_g = GM_\bullet/c^2$. This precession is said to be prograde as one always has $\Omega_{GR}(a, h) > 0$. The mass precession, induced by the background stellar cluster, is given by (see, e.g., Tremaine, 2005)

$$\Omega_M(a, h) = \frac{\Omega_{Kep}(a)}{\pi M_\bullet e} \int_0^\pi df M_\star[<r(f)] \cos(f), \quad (2.4)$$

with f the true anomaly and $M_\star(<r)$ the stellar mass enclosed within the radius r . This precession is retrograde, i.e. $\Omega_M(a, h) < 0$, for any positive-definite spherical mass distribution.

2.1.2 Eccentricity relaxation

The diffusion of h is sourced by the potential fluctuations generated by the background stellar cluster. This cluster is expected to be dynamically old, i.e. has been orbiting around SgrA* for a time much longer than the SRR time. As a consequence, we may assume that all its orbital elements are fully relaxed. To be more precise, let us introduce the background cluster’s distribution function (DF), $F = F(\Lambda, L, L_z, \omega, \Omega)$ so that $F d\Lambda dL dL_z d\omega d\Omega$ is proportional to the stellar mass enclosed within the phase space volume element $d\Lambda dL dL_z d\omega d\Omega$. We then assume that: (i) the cluster is spherically symmetric, i.e. $F = F(\Lambda, L)$; (ii) the cluster has a thermal distribution

of angular momentum (Bar-Or & Alexander, 2016), i.e. $F = F(\Lambda)$. Assumption (ii) entails two crucial consequences. First, SRR does not drive any dynamical friction (Bar-Or & Fouvy, 2018), and the efficiency of a star's relaxation is independent of its mass. Second, SRR is unaffected by gravitational polarisation (Bar-Or & Fouvy, 2018), i.e. one may safely limit oneself to the (bare) coupling coefficients, (see Eq. 2.10).

In that limit, the test star's eccentricity follows a diffusion equation of the form (Bar-Or & Alexander, 2016)

$$\frac{\partial P(h, t|a)}{\partial t} = \frac{1}{2} \frac{\partial}{\partial h} \left[h D_{hh}(a, h) \frac{\partial}{\partial h} \left(\frac{P(h, t|a)}{h} \right) \right], \quad (2.5)$$

where $P(h, t|a)$ is the probability distribution function (PDF) of the test star's angular momentum, h , for a given semi-major axis a , as a function of time, normalised so that $\int dh P(h, t|a) = 1$. In Eq. (2.5), we neglect the diffusion in a (hence $D_{aa} = 0$), which is minor compared to that in angular momentum (Bar-Or & Alexander, 2016): the test star diffuses at fixed semi-major axis. We also neglect the BH's loss-cone region, that would drive a small exiting flux at low h . The steady state of Eq. (2.5) is given by the thermal solution, $P(h|a) = 2h$, i.e. the eccentricity PDF also followed by the background stars.

The diffusion coefficient in angular momentum, $D_{hh}(a, h)$, is the sum of two contributions

$$D_{hh}(a, h) = D_{hh}^{\text{RR}}(a, h) + D_{hh}^{\text{NR}}(a, h), \quad (2.6)$$

where $D_{hh}^{\text{RR}}(a, h)$ captures the contribution from SRR, while $D_{hh}^{\text{NR}}(a, h)$ is associated with NR. We now detail the content of each of these coefficients.

2.1.3 Resonant diffusion coefficients

An important source of eccentricity relaxation stems from long-range resonant couplings between in-plane precessing wires. When sourced by a spherically symmetric background, the SRR diffusion coefficients read

$$D_{hh}^{\text{RR}}(a, h) = \frac{4\pi G^2}{L_c^2} \sum_{k=1}^{+\infty} \sum_{\substack{k'=-\infty \\ k' \neq 0}}^{+\infty} k^2 \int da' dh' F_{\text{tot}}(a', h') |A_{kk'}(a, h, a', h')|^2 \delta_{\text{D}}[k\Omega_{\text{p}}(a, h) - k'\Omega_{\text{p}}(a', h')], \quad (2.7)$$

with $L_c = \sqrt{GM_{\bullet} a}$ the star's specific circular angular momentum, following Eq. (1.5). We do not reproduce here the derivation of Eq. (2.7) and refer to Bar-Or & Fouvy (2018) for a detailed presentation.

In Eq. (2.7), we introduced the DF, $F_{\text{tot}}(a', h')$, describing the orbital distribution of the background cluster, whose potential fluctuations are responsible for the long-term diffusion of stellar eccentricities. It is defined as

$$F_{\text{tot}}(a, h) = \sum_i m_i^2 N_i(a) f_i(h|a), \quad (2.8)$$

where the sum over i runs over all the sub-populations of the background cluster. Each population is characterised by an individual mass, m_i , while $N_i(a)$ is the number of stars per unit semi-major axis a , and $f_i(h|a)$ is the conditional PDF of h for a given a normalised so that $\int dh f_i(h|a) = 1$. With such a convention, our assumption of a thermal eccentricity distribution simply translates to $f_i(h|a) = 2h$.

Equation (2.7) involves the coupling coefficients $|A_{kk'}(a, h, a', h')|^2$. They describe the efficiency of the gravitational coupling between the wires (a, h) and (a', h') through the resonance (k, k') . They read

$$|A_{kk'}(a, h, a', h')|^2 = 16\pi^2 \sum_{\ell=0}^{\ell_{\text{max}}} \frac{|y_{\ell}^n|^2 |y_{\ell}^{n'}|^2}{(2\ell+1)^3} |K_{kk'}^{\ell}(a, h, a', h')|^2, \quad (2.9)$$

with the harmonic truncation ℓ_{max} , and the constant coefficients $y_{\ell}^n = Y_{\ell}^n(\frac{\pi}{2}, \frac{\pi}{2})$, where the complex spherical harmonics are defined with the convention $\int d\hat{\mathbf{r}} |Y_{\ell}^n(\hat{\mathbf{r}})|^2 = 1$. This equation also involves the in-plane SRR (bare) coupling coefficients, $K_{kk'}^{\ell}$, which read

$$K_{kk'}^{\ell}(a, h, a', h') = \int_0^{\pi} \frac{dM}{\pi} \int_0^{\pi} \frac{dM'}{\pi} \cos(kf) \cos(k'f') \frac{\text{Min}[r, r']^{\ell}}{\text{Max}[r, r']^{\ell+1}}, \quad (2.10)$$

with M, f (resp. M', f') the mean and true anomaly of orbit (a, h) (resp. (a', h')), as defined in Eq. (1.9). In regard to symmetries, one finds: (i) $K_{kk'}^{\ell}$ is non-zero only when $|k|, |k'| \leq \ell$, as well as $(\ell - k)$ and $(\ell - k')$ both even; (ii) $|A_{kk'}|^2 = |A_{\pm k \pm k'}|^2$, i.e. the strength of the coupling is independent of the sign of the resonance numbers.

In Eq. (2.10), the Min-Max terms stem from the usual Legendre expansion of the Newtonian interaction potential, see Eq. (2.34). The computation of $K_{kk'}^{\ell}$ is the overall bottleneck of the SRR calculation. Following

Eq. (1.9), it is convenient to perform the integral from Eq. (2.10) w.r.t. the true anomalies f and f' . Fortunately, because its integrand is quasi-separable, the two-dimensional integral from Eq. (2.10) can efficiently be performed using a standard multipole approach. We postpone to §2.3.2.2 the presentation of this algorithm.

Once the coupling coefficients estimated, we rely on Eq. (2.7) to estimate the diffusion coefficients. For a given wire (a, h) and a given resonance pair (k, k') , this amounts to finding all the wires (a', h') for which the resonance condition $k\Omega_p(a, h) = k'\Omega_p(a', h')$ is satisfied. We refer to §A in Tep et al. (2021) for details on how to efficiently solve this resonance condition. In the left panel of Fig. 2.1, we illustrate the typical dependence of the precession frequency, $\Omega_p(a, h)$, in orbital space. In that figure, we find that for circular orbits and large semi-

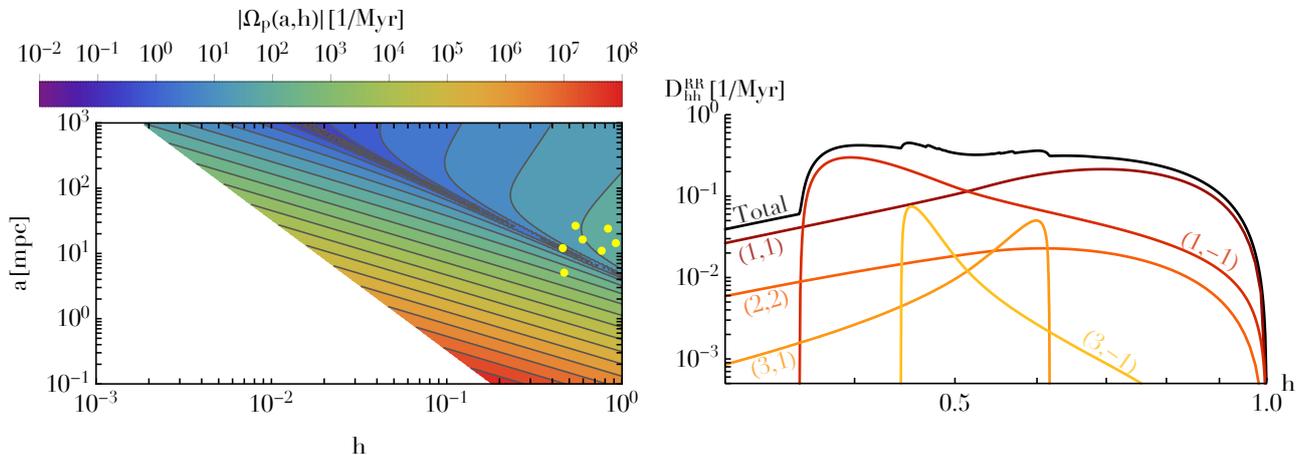


Figure 2.1: From Tep et al. (2021). Left: Precession frequencies, $|\Omega_p(a, h)|$, for the Top-Heavy model from Eq. (2.11). The yellow dots represent the orbital parameters of the 7 S-stars (S1, S2, S4, S6, S8, S9, S12) (Gillessen et al., 2017), whose main-sequence ages have recently been constrained (Habibi et al., 2017). The white region on the left represents the central BH's loss cone. Right: For the same model, contribution to the SRR diffusion coefficients of the different resonances (k, k') , for a given semi-major axis, $a = 10$ mpc, computed with $\ell_{\max} = 10$.

major axis, i.e. $h \rightarrow 1$ and $a \gg 1$, the precession is dominated by the mass precession and is therefore retrograde, i.e. $\Omega_p < 0$, while for low h and low a , the precession is dominated by relativistic precession and is therefore prograde, i.e. $\Omega_p > 0$. Lines of constant precession frequencies correspond to the resonant lines along which the diffusion coefficients from Eq. (2.7) are computed. Finally, the precession of very eccentric orbits is dominated by the diverging relativistic corrections. This is responsible for the "Schwarzschild barrier" (Merritt et al., 2011; Bar-Or & Alexander, 2016) that explains the drastic reduction of the SRR diffusion coefficients as wires get more and more eccentric.

In the right panel of Fig. 2.1, we illustrate the contributions from the various resonance pairs (k, k') to the total SRR diffusion coefficients. These are typically dominated by $(k, k') = (1, 1)$ (for small and large h) and $(k, k') = (1, -1)$ (for intermediate h). For intermediate eccentricities, higher order resonances also contribute. We refer to Fig. 2.10 for a quantitative comparison of Eq. (2.7) with measurements in numerical simulations.

We conclude this section by pointing out a few key connexions between Eq. (2.7) and the generic inhomogeneous Balescu–Lenard (BL) equation, reproduced in §A:

- Following the average from Eq. (1.10), the (fast) action Λ is conserved by SRR. This would occur in Eq. (A.1) if one replaces the mean-field Hamiltonian, H_0 , by its average, $\langle H_0 \rangle = \int d\theta_f H_0$, over some fast angle θ_f .
- The galactic nucleus was assumed to be spherically symmetric, so that $\Omega_z = \partial \langle H \rangle / \partial L_z = 0$. To remove this degeneracy, one averages the system's DF w.r.t. L_z , see, e.g., Hamilton et al. (2018) for the case of globular clusters. This additional symmetry is responsible for (i) the factors h present in the Fokker–Planck (FP) Eq. (2.5); (ii) the development in spherical harmonics in Eq. (2.9).
- Because collective effects are absent in nuclei with thermal distribution of eccentricities, the coupling coefficients from Eq. (2.10) are formally equivalent to the generic Landau coupling coefficients from Eq. (A.7). There, the true anomaly, f , is the angle somewhat naturally associated with the pericentre precession.

2.1.4 Non-resonant diffusion coefficients

As detailed in introduction, a star also relaxes in eccentricity via NR. In that case, it is the slow build-up of nearby scatterings that ultimately drives the diffusion of the star's orbital parameters. In order to evaluate the associated diffusion coefficient, $D_{hh}^{\text{NR}}(a, h)$, we use the exact same approach as in Bar-Or & Alexander (2016). Because it does not involve any resonance condition, NR is much less demanding to compute than SRR. We do not reproduce here any detailed expressions, but simply report on the key steps of that calculation:

- At a given phase space location (\mathbf{r}, \mathbf{v}) , we compute the local velocity diffusion coefficients, $\langle \delta \mathbf{v} \rangle(\mathbf{r}, \mathbf{v})$ and

$\langle(\delta\mathbf{v})^2\rangle(\mathbf{r}, \mathbf{v})$, as given in Eq. (7.83a) in Binney & Tremaine (2008)¹. This calculation is greatly simplified by our assumption that the background cluster is fully relaxed, i.e. the cluster follows an isotropic DF, $F = F(\Lambda)$. In practice, we set the Coulomb logarithm of a given background family to $\ln(\Lambda_i) = \ln(M_\bullet/m_i)$.

- We translate these velocity diffusion coefficients into energy and angular momentum ones, e.g., $\langle\delta E\rangle(\mathbf{r}, \mathbf{v})$.
- Local kicks accumulate as the star moves along its Keplerian wire. Following an orbit-average, one obtains then the associated orbit-averaged diffusion coefficients, e.g., $\langle\Delta E\rangle = \int \frac{dM}{2\pi} \langle\delta E\rangle$.
- Having obtained the first- and second-order diffusion coefficients in the test star's energy and angular momentum, we finally obtain the associated diffusion coefficients in (a, h) -space through an appropriate change of variables (Risken, 1989). In particular, we get $\langle(\Delta h)^2\rangle(a, h) = D_{hh}^{\text{NR}}(a, h)$.

In the left panel of Fig. 2.2, we illustrate the NR coefficients for a typical model of SgrA*. Contrary to SRR,

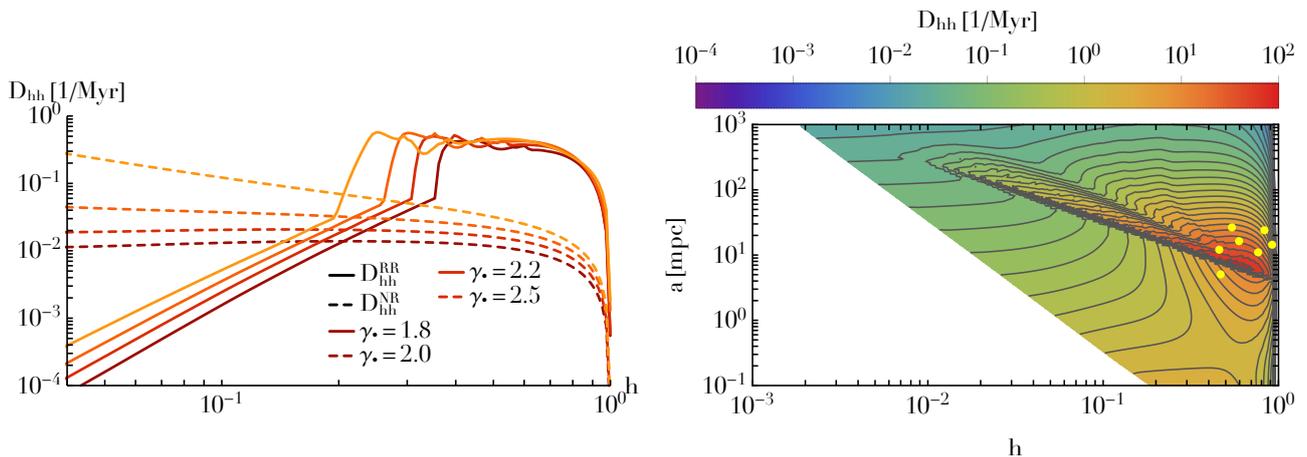


Figure 2.2: From Tep et al. (2021). Left: SRR (full lines) and NR (dashed lines) diffusion coefficients for the same model as in Fig. 2.1, with $a = 10$ mpc and $\ell_{\text{max}} = 10$. Different colors correspond to different values of γ , the power-law index of the IMBH's population, see Eq. (2.11). Right: Total diffusion coefficients, $D_{hh}(a, h)$, in orbital space, using the same model and convention as in Fig. 2.1.

NR's efficiency is mostly independent of the stars' eccentricities. Reassuringly, we find that both coefficients go to 0 as $h \rightarrow 1$ (circular orbits), while the SRR ones get drastically reduced for very eccentric orbits (Merritt et al., 2011; Bar-Or & Alexander, 2016). As such, for small enough h , NR dominates over SRR.

In the right panel of Fig. 2.2, we illustrate the overall dependence of the total diffusion coefficients from Eq. (2.6), i.e. both the SRR and NR contributions, in the whole (a, h) orbital space. In that figure, one clearly notes the presence of resonance lines associated with SRR: this rugged but accurate aspect can be linked to that of the isocontours of the resonance frequencies from the left panel of Fig. 2.1. One also notes that the bulk of the currently observed S-stars lie in a region of orbital space where the diffusion of eccentricities is dominated by SRR. It is therefore essential to account for these resonant mechanisms to accurately describe the dynamical fate of the S-stars' eccentricities. This is the topic of §2.2.

2.1.5 Prospects

Stellar relaxation around a supermassive BH is a classical problem of dynamics. Understanding this process is crucial for (i) the long-term steady-state stellar distribution of galactic nuclei and mass segregation therein (see, e.g., Gillessen et al., 2017); (ii) short-term transient phenomena such as tidal disruption events (TDEs), GW emissions, and hypervelocity stars (Hills, 1988; Koposov et al., 2019); (iii) and the distribution of unique source populations such as young stars, X-ray binaries, and radio pulsars, to name a few (Alexander, 2017). All these phenomena depend upon relaxation in energy and in angular momentum.

The relaxation in energy is well described by NR, in particular for an isotropic DF. There, the only poorly determined quantity is the Coulomb logarithm, which has only a small effect on diffusion. Despite the approximations made in the derivation of these diffusion coefficients (e.g., Nelson & Tremaine, 1999), these NR coefficients are in good agreement with the ones measured in direct N -body simulations (Bar-Or et al., 2013).

Relaxation in angular momentum is dominated by SRR (Rauch & Tremaine, 1996; Hopman & Alexander, 2006), as demonstrated in the numerical works of Eilon et al. (2009); Merritt et al. (2011). The foundation for a concrete kinetic theory of SRR was put forward independently by Bar-Or & Alexander (2014); Sridhar & Touma (2016); Fouvry et al. (2017), via essentially equivalent approaches. In this section, we built upon Bar-Or & Alexander (2014); Fouvry & Bar-Or (2018) and presented a calculation of the SRR diffusion coefficients from

¹All these expressions can be recovered from Eq. (A.1) using the homogeneous assumption $(\boldsymbol{\theta}, \mathbf{J}) \mapsto (\mathbf{r}, \mathbf{v})$ (see, e.g., Chavanis, 2013b).

first principles, without any free parameters. This concludes the long journey, started by Rauch & Tremaine (1996), of bringing SRR to the same level of completeness as NR.

We close this section by mentioning a few venues for future improvements.

- The treatment presented here is limited to SRR in a spherical and isotropic background distribution, for which collective effects can be ignored (see, e.g., Nelson & Tremaine, 1999). Building upon Tremaine (2005); Polyachenko et al. (2020), one should explore the role of gravitational polarisation and dynamical friction in anisotropic systems.
- Vasiliev (2015) has shown how to apply NR in non-spherically symmetric systems. It would be physically enlightening to relax the assumption of spherical symmetry and investigate its impact on SRR.
- Here, we assumed that the central BH dominates the potential. This assumption breaks down close to the BH's radius of influence, where the contribution of the underlying stellar population is comparable to that of the BH. Some of these limitations could be mitigated in future studies, and tested against future N -body simulations that now approach realistic number of stars in galactic nuclei (Panamarev et al., 2019).
- The ability to calculate SRR diffusion coefficients provides us with the opportunity to make more realistic estimates of the astrophysical impact of SRR in galactic nuclei. As SRR can efficiently drive the evolution of angular momentum, it may contribute to the supply rate of stellar objects into the BH's loss-cone, e.g., via extreme mass ratio inspirals (EMRIs) and TDEs. Yet, this contribution is significant only if the loss-cone is close to the orbital region where SRR dominates over NR. This should be investigated in detail.

2.2 Eccentricity relaxation around SgrA*

In this section, we leverage the previous characterisation of SRR to constrain hidden features of SgrA*'s galactic centre, focusing on the relaxation of stellar eccentricities. Indeed, as emphasised in Gillessen et al. (2017) (see fig. 13 therein), the S-cluster orbiting SgrA* (for $a \simeq 10$ mpc) has been observed with a significantly relaxed distribution of eccentricities. Detailed spectroscopic observations (Habibi et al., 2017) also provide us with well constrained main-sequence ages for these same stars.

The age of the young B-stars at a distance of ~ 10 mpc from the BH is comparable to the resonant relaxation time, implying that SRR may have played an important role in their dynamical structure (Hopman & Alexander, 2006). Hence, any credible diffusion mechanism has to be efficient enough to drive a significant eccentricity relaxation of the S-stars within their lifetime – see Chen & Amaro-Seoane (2014) for an alternate scenario. Equivalently, one may use this constraint as a dynamical probe to further characterise the properties of the unresolved old stellar and a putative dark cluster, which both drive the relaxation of the S-stars themselves (see, e.g., Generozov & Madigan, 2020). Following Merritt et al. (2009); Antonini & Merritt (2013), Monte-Carlo simulations have similarly shown that the presence of stellar mass BHs around SgrA* reduces the resonant relaxation time near the present-day location of the S-stars to ~ 10 Myr, which is of the order of the age of the S-stars (Habibi et al., 2017). In this section, we use kinetic theory to constrain the range of cluster models that are compatible with the observational requirement of having significantly relaxed S-stars' eccentricities.

2.2.1 Model's assumptions

In Eq. (2.6), we quantified the two main processes (SRR and NR) through which stars can relax in eccentricities in galactic nuclei. We now use these processes as dynamical probes in the context of the recent observation of the S-stars' eccentricities within SgrA*.

For the observational data, we use the orbital parameters listed in Gillessen et al. (2017). Specifically, we use the (a, h) coordinates for 7 of those stars (S1, S2, S4, S6, S8, S9, S12), for which Habibi et al. (2017) provides us with main-sequence ages. These ages are a measure of the total time that the diffusion has had to operate in Eq. (2.5). For simplicity, we assume that on these timescales, NR did not drive any significant diffusion of the S-stars' energies, i.e. the stars' semi-major axes, a , are kept fixed.

Regarding the initial conditions for the stars' eccentricities, we investigate two possible scenarii: (i) binary tidal disruptions (Hills, 1988; Gould & Quillen, 2003), i.e. large initial eccentricities; (ii) an episode of disc formation (Alexander, 2005; Levin, 2006; Kozlov et al., 2019), i.e. small initial eccentricities. In practice, we assume that the S-stars are initialised following a Gaussian distribution centred around $h(t=0) = 0.2$ (with width 0.02) to mimic binary disruptions (Generozov & Madigan, 2020), or $h(t=0) = 0.9$ (same width) to mimic in-situ formation. For alternative scenarii, we refer to Madigan et al. (2009); Perets et al. (2007).

Let us now also make key assumptions regarding the background old stellar cluster. Following the convention from Eq. (2.8), we assume that it is composed of various sub-populations of different individual masses m_i with a total mass $M_i (< a_0)$ enclosed within a physical radius a_0 . In addition, we also assume that each population follows a thermal distribution in eccentricity, i.e. $f_i(h|a) = 2h$ in Eq. (2.8), and infinite power-law

distribution in semi-major axes, so that $M_i(<a) = M_i(<a_0)(a/a_0)^{3-\gamma_i}$. We also treat the S-stars as test stars, i.e. they do not contribute to the system's mean potential and do not interact with one another.

Assuming a two-family background composed of stars and another heavy sub-populations (e.g., IMBHs), we then have a total of 7 free parameters for the available models, namely the power indices $(\gamma_*, \gamma_\bullet)$, the individual masses (m_*, m_\bullet) , the total enclosed masses $(M(<a_0), M_\bullet(<a_0))$, as well as the initial eccentricity of the S-stars, $h_0 = h(t=0)$. These models are complemented with the observed constraints on the 7 considered S-stars, namely their observed semi-major axis, eccentricity, and main-sequence age.

We start our investigation with the two-family model of Generozov & Madigan (2020). Using $a_0 = 0.1$ pc, it contains both stars and IMBHs following

$$\begin{cases} m_* = 1 M_\odot, \\ m_\bullet = 50 M_\odot, \end{cases} \quad \begin{cases} M_*(<a_0) = 7.9 \times 10^3 M_\odot, \\ M_\bullet(<a_0) = 38 \times 10^3 M_\odot, \end{cases} \quad \begin{cases} \gamma_* = 1.5, \\ \gamma_\bullet = 1.8, \end{cases} \quad (2.11)$$

where the stars' parameters follow Schödel et al. (2017). We also note that such a model is compatible with the current constraints associated with S2's pericentre shift (Gravity Collaboration et al., 2020).

2.2.2 Methodology

Having picked a set of initial conditions for the S-stars, and a model for the background clusters, we are now in a position to compute the associated eccentricity diffusion coefficients. In order to determine whether or not such a model is compatible with the observational constraint of a significant eccentricity relaxation of the S-stars, we proceed as follows.

We first compute the SRR and NR diffusion coefficients for each a of the considered S-stars. As the semi-major axes are conserved, we can integrate Eq. (2.5) forward in time separately for each S-star, using the method of lines (see Tep et al., 2021, for details). Once these integrations performed, we compare the reached PDF in h to the observed data of the S-cluster, determining whether or not the background model allowed for an efficient enough relaxation of the S-stars' eccentricities.

Let us denote a model with α , i.e. the collection of the 7 parameters of the background clusters and the S-stars' initial eccentricity. The likelihood of a given model is then defined as

$$L(\alpha) = \prod_k P(h_k, T_k | a_k), \quad (2.12)$$

where $k=1, \dots, 7$ runs through the 7 S-stars mentioned before, and T_k is the star's main-sequence age. To compare different models to one another, we finally introduce the likelihood ratio

$$\lambda_R(\alpha) = 2 \ln(L_{\max}/L(\alpha)) > 0, \quad (2.13)$$

with L_{\max} the maximum value of $L(\alpha)$ within the range of models probed. The smaller λ_R the more likely the model. As detailed in §F of Tep et al. (2021), these ratios can also be translated into confidence levels.

2.2.3 Results

As an illustration of the present method, we first consider the Top-Heavy model from Eq. (2.11), and let the individual masses m_* and m_\bullet vary. We impose the natural constraint $m_\bullet \geq m_*$ while fixing the total enclosed masses $M_*(<a_0)$ and $M_\bullet(<a_0)$. This is presented in Fig. 2.3 with the associated confidence levels, for the two scenarii of initial conditions.

As expected, in Fig. 2.3, we recover that the larger the individual masses, the larger the underlying Poisson shot noise, and therefore the more efficient the diffusion process and the faster the relaxation of the S-stars. Conversely, Fig. 2.3 shows that models with small individual masses cannot explain the S-cluster's eccentricity PDF. As such, a relatively massive set of background sources orbiting around SgrA* is required to trigger a fast enough orbital diffusion of the observed stars within their lifetime. Finally, we can compare both panels of Fig. 2.3 that correspond to different initial conditions for the S-stars. As already highlighted in Fig. 2.2, the smaller h_0 , i.e. the more eccentric the stellar initial conditions, the slower the relaxation of the S-stars.

The global shape of the likelihood contours presented in Fig. 2.3 clearly illustrates the known dynamical degeneracy in flipping the stars and IMBHs of the same mass, as the efficiency of relaxation is solely connected to the amplitude of the Poisson shot noise generated by the cluster as a whole. Finally, we note that all the likelihood landscapes present an absolute minimum. This suggests that, having diffused only a finite time, the observed eccentricity distribution of the S-stars is not fully thermal.

In order to increase the observed stellar sample and tighten the inferred model's constraints, we present in Fig. 2.4 the same measurements as in Fig. 2.3 but using 30 additional S-stars (as in fig. 13 of Gillessen et al. (2017)) Their individual ages was fixed to $T = 7.1$ Myr, i.e. the average age of the 7 S-stars whose ages have

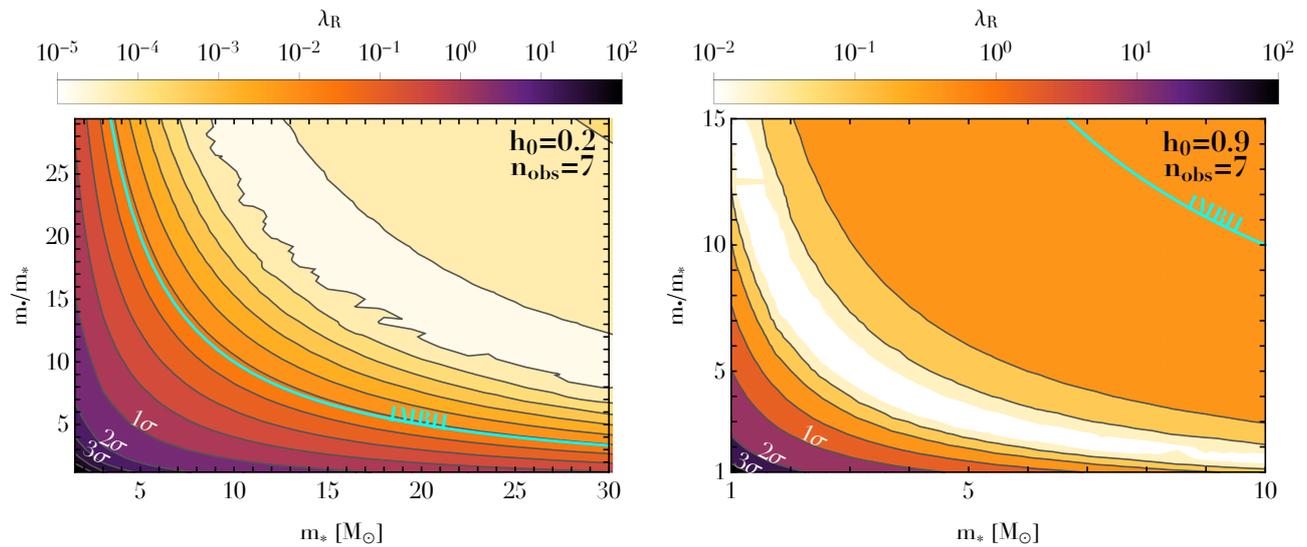


Figure 2.3: From Tep et al. (2021). Confidence regions for the Top-Heavy model (Eq. 2.11) as one varies the individual masses (m_* , m_\bullet) using the likelihood ratio from Eq. (2.13) applied to the 7 observed S-stars. We assume either a large initial eccentricity ($h_0=0.2$, left), or a small one ($h_0=0.9$, right). The cyan line corresponds to $m_\bullet=100 M_\odot$, above which the heavy objects are usually considered to be IMBHs.

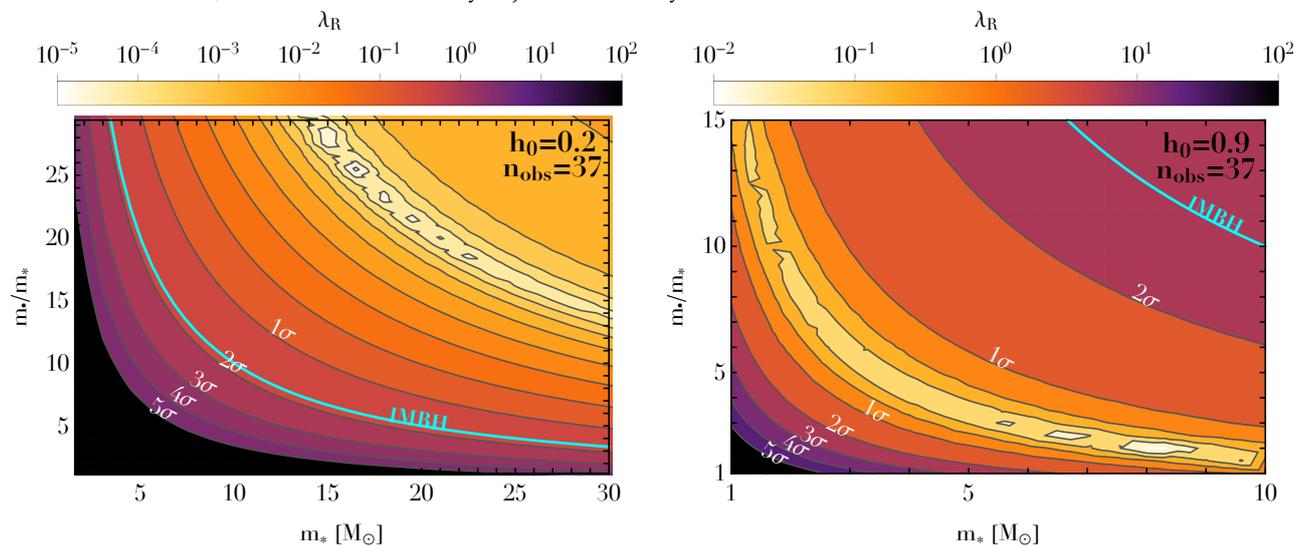


Figure 2.4: From Tep et al. (2021). Same as Fig. 2.3, except that the sample of stars has been expanded with 30 other S-stars (Gillessen et al., 2017), assuming a common age $T=7.1$ Myr.

been measured (Habibi et al., 2017). As expected, we recover that a larger sample of observed stars leads to narrower contours around the likelihood maximum, making the presence of a second population of massive objects all the more mandatory. Finally, we also note that since the expanded sample of 37 stars contains stars with semi-major axes larger than that of the initial 7 S-stars, i.e. stars whose relaxation is slower, the location of the likelihood maximum in Fig. 2.4 gets displaced to larger masses as one increases the observed stellar sample.

We stop here the quantitative application of SRR to constrain SgrA* and refer to §3.4 in Tep et al. (2021) for prospective experiments regarding future surveys (Do et al., 2019).

2.2.4 Prospects

In this section, we used the explicit kinetic theory of SRR to assess the dynamical structure of SgrA*'s stellar cluster. In particular, we emphasised how the recent observations of the thermal distribution of eccentricities of the S-stars in conjunction with computations of the eccentricity diffusion coefficients can be effectively leveraged to place constraints on putative IMBHs around SgrA*. Finally, investigating a simple two-population model, we showed how the presence of a heavy sub-population is mandatory to source an efficient enough relaxation of the S-stars' eccentricities.

Let us conclude this section by discussing some venues for future developments.

- First, our model for the old stellar cluster is over-simplistic. We assumed for example that the background cluster is spherically symmetric. However, in §3.3 we show that in systems with a large mass spectrum, VRR

can lead to equilibria distribution where the massive components follow a strongly anisotropic structure. Surely, such a structure must affect the efficiency of the eccentricity relaxation it sources.

- We assumed that the S-stars' semi-major axes were fixed throughout their eccentricity diffusion. One should investigate whether any additional diffusion in a , sourced by NR, could affect the present constraints.
- Here, we focused our interest on the innermost S-stars ($a \simeq 10$ mpc), which are known to have significantly relaxed in eccentricity. Admissible models must then source an efficient enough eccentricity diffusion. A similar approach may be used for the S-stars further out, which have only partially relaxed in eccentricity, so that admissible models must source a slow enough diffusion in these outer regions. Leveraging both constraints simultaneously, one should ultimately be in a position to effectively bracket models of clusters.
- Finally, future observations will undoubtedly prove useful in placing these investigations on firmer grounds. First, GRAVITY is currently tracking in detail the trajectory of S2 (Gravity Collaboration et al., 2020). Any deviations of S2 from its expected mean-field trajectory will bear imprints from the potential fluctuations on the scale of S2's orbit. Similarly, the observation of stars on scales even smaller than S2's will carry essential information on SgrA*'s stellar cluster even closer to the central BH. Finally, on larger scales, future observations by upcoming thirty-meter telescopes (Pott et al., 2018; Davies et al., 2018; Do et al., 2019) will allow for a finer characterisation of the S-stars current distribution, $P(a, h, t)$. The dependence of P w.r.t. a strongly depends on the stellar formation mechanism, while its dependence w.r.t. h constrains the scale, i.e. the a , at which the S-stars diffuse less and less efficiently towards the thermal eccentricity distribution.

2.3 Simulations

In galactic nuclei, the central BH creates a gigantic range of timescales: from a few years for the fast Keplerian motion (even a few minutes for stars near the event horizon) up to a Hubble time for NR in the MW's nucleus. Direct numerical simulations of galactic nuclei remain therefore very challenging. These were first performed with grid methods (see, e.g., Jacobs & Sellwood, 2001; Kazandjian & Touma, 2013). And, it is only recently that an effective direct simulation of a galactic nucleus with $N = 10^6$ has been presented (Panamarev et al., 2019), and even in this case most of the stars lie outside the central BH's sphere of influence. Conversely, simulations in the very relativistic regime are still limited to a small number of particles, $N \simeq 10^{2-3}$, should they use direct N -body methods (Merritt et al., 2011) or effective ones (see, e.g., Madigan et al., 2011; Hamers et al., 2014).

To circumvent these difficulties, one has to resort to additional assumptions. Traditionally, the secular approximation (Eq. 1.10) smears out the stars along their underlying fast Keplerian motion, in other words, one replaces stars with Keplerian wires. This is in particular at the heart of Gauss' method (Touma et al., 2009) which provides an efficient algorithm to compute the force between two such wires. The main benefit from this approach is that any explicit average over the fast orbital motion offers a reduction in the range of timescales in the system. These methods can then explore longer timescales beyond the reach of naive direct methods. Yet, most of these approaches suffer from relying on the computation of all the individual forces between objects, i.e. these methods come with a complexity scaling like $\mathcal{O}(N^2)$ with N the total number of stars.

In this section, we show how a multipole expansion (Hénon, 1964; Aarseth, 1967; Hénon, 1973; van Albada & van Gorkom, 1977; Fry & Peebles, 1980; Villumsen, 1982; White, 1983; McGlynn, 1984; Meiron et al., 2014; Dehnen, 2014) yields a numerical scheme to integrate the long-term dynamics of Keplerian wires with a complexity scaling like $\mathcal{O}(NK\ell_{\max}^2)$, with K a parameter independent of N , and ℓ_{\max} the maximum harmonics considered in the multipole expansion. We also show that the system is equivalent to $2N$ classical spin vectors, and devise integration schemes that exactly comply with the associated geometric constraints.

2.3.1 Equations of motion

The orbit-averaged Hamiltonian from Eq. (1.11) describes the dynamics of N gravitationally coupled Keplerian wires. Each wire is characterised by the five quantities $(\omega, \Omega, \Lambda, L, L_z)$ of which Λ is conserved through the orbit-averaged dynamics. Of course, the equations of evolutions for the four other coordinates can be obtained from Hamilton's canonical equations of motion. Such calculations are rather involved because they involve computing the corresponding derivatives of $r(\omega, \Omega, L, L_z)$ at fixed (M, Λ) . Furthermore, the associated equations can become degenerate, e.g., at $I=0$ (equatorial orbits), $e=0$ (circular orbits), or $e=1$ (radial orbits). An equivalent alternative is to work directly with the forces acting on the wires, as we now pursue.

Rather than integrating the motion w.r.t. the orbital elements, we keep track of the wire dynamics through the dimensionless vectors

$$\mathbf{h} = \frac{\mathbf{r} \times \mathbf{p}}{\Lambda}; \quad \mathbf{e} = \frac{\mathbf{p} \times (\mathbf{r} \times \mathbf{p})}{m^2 GM_\bullet} - \hat{\mathbf{r}}, \quad (2.14)$$

where a hat denotes a unit vector as usual. Here, \mathbf{h} is the angular momentum vector scaled to the circular angular momentum at the same energy, and \mathbf{e} is the eccentricity vector, which points in the direction of the wires' pericentre and has magnitude $|\mathbf{e}| = e$.

While only four orbital elements ω, Ω, L, L_z evolve, the two vectors (\mathbf{h}, \mathbf{e}) have six dynamical variables in total. The two associated degeneracies are captured by the two identities

$$\mathbf{h} \cdot \mathbf{e} = 0; \quad \mathbf{h}^2 + \mathbf{e}^2 = 1. \quad (2.15)$$

2.3.1.1 Orbit-averaged rates of change

Using Eq. (2.14), one can compute $\dot{\mathbf{h}}$ and $\dot{\mathbf{e}}$ from derivatives w.r.t. the canonical variables \mathbf{r} and \mathbf{p} . Focusing only on the self-interacting term $\langle H_\star \rangle$ in Eq. (1.11), one obtains (see §3.1 in Fouvry et al., 2022)

$$\dot{\mathbf{h}}_i = \frac{1}{\Lambda_i} \sum_{j \neq i} G m_i m_j \langle \mathbf{r}_i \times \mathbf{F}_{ij} \rangle, \quad (2.16a)$$

$$\dot{\mathbf{e}}_i = \frac{a_i}{\Lambda_i^2} \sum_{j \neq i} G m_i m_j \langle \mathbf{F}_{ij} \times (\mathbf{r}_i \times \mathbf{p}_i) + \mathbf{p}_i \times (\mathbf{r}_i \times \mathbf{F}_{ij}) \rangle, \quad (2.16b)$$

with $\langle \cdot \rangle$ the orbit-average over the unperturbed orbits, and the shorthand $\mathbf{F}_{ij} = \mathbf{F}(\mathbf{r}_i - \mathbf{r}_j)$ with $\mathbf{F}(\mathbf{r}) = \partial|\mathbf{r}|^{-1}/\partial\mathbf{r}$.

The expressions from Eq. (2.16) satisfy the constraints from Eq. (2.15) for any conservative field \mathbf{F}_{ij} (Touma et al., 2009). However, carrying out the orbit averages requires numerical treatment. The resulting numerical average inevitably carries a small error with the consequence that the constraints from Eq. (2.15) are no longer exactly honoured. While in general small numerical errors in conserved quantities are not a serious problem, this particular situation is awkward, since the conditions from Eq. (2.15) are essential for the interpretation of $(\mathbf{h}_i, \mathbf{e}_i)$ as wires. It is therefore paramount to construct numerical expressions for $\dot{\mathbf{h}}_i$ and $\dot{\mathbf{e}}_i$ that, despite their discretisation errors, keep the constraints from Eq. (2.15) valid to machine precision.

To achieve that goal, we note that Eq. (2.16) is equivalent to the Milankovitch's equations (Milankovitch, 1939; Tremaine et al., 2009; Rosengren & Scheeres, 2014)

$$\dot{\mathbf{h}}_i = -\frac{1}{\Lambda_i} \left(\mathbf{h}_i \times \frac{\partial \langle H_\star \rangle}{\partial \mathbf{h}_i} + \mathbf{e}_i \times \frac{\partial \langle H_\star \rangle}{\partial \mathbf{e}_i} \right), \quad (2.17a)$$

$$\dot{\mathbf{e}}_i = -\frac{1}{\Lambda_i} \left(\mathbf{h}_i \times \frac{\partial \langle H_\star \rangle}{\partial \mathbf{e}_i} + \mathbf{e}_i \times \frac{\partial \langle H_\star \rangle}{\partial \mathbf{h}_i} \right), \quad (2.17b)$$

with the averaged Hamiltonian $\langle H_\star \rangle$ expressed as a function of $\mathbf{h}_i, \mathbf{e}_i$ (and the constant Λ_i).

These relations are equivalent to the canonical equations of motion, but much more useful. First, the vectors \mathbf{h} and \mathbf{e} are well-defined for all orbits, even with zero eccentricities. Equation (2.17) also ensures that the constraints from Eq. (2.15) and the conservation of the total energy are explicitly satisfied for any $\langle H_\star \rangle$, i.e.

$$\frac{d(\mathbf{h}_i^2 + \mathbf{e}_i^2)}{dt} = \dot{\mathbf{h}}_i \cdot \mathbf{h}_i + \dot{\mathbf{e}}_i \cdot \mathbf{e}_i = 0, \quad (2.18a)$$

$$\frac{d(\mathbf{h}_i \cdot \mathbf{e}_i)}{dt} = \dot{\mathbf{h}}_i \cdot \mathbf{e}_i + \dot{\mathbf{e}}_i \cdot \mathbf{h}_i = 0, \quad (2.18b)$$

$$\frac{d \langle H_\star \rangle}{dt} = \sum_i \left\{ \frac{\partial \langle H_\star \rangle}{\partial \mathbf{h}_i} \cdot \dot{\mathbf{h}}_i + \frac{\partial \langle H_\star \rangle}{\partial \mathbf{e}_i} \cdot \dot{\mathbf{e}}_i \right\} = 0. \quad (2.18c)$$

These properties suggest an alternative, fully conservative approach to numerically computing the rates $\dot{\mathbf{h}}_i$ and $\dot{\mathbf{e}}_i$. Instead of discretising the orbit averages in Eq. (2.16), we first discretise the orbit average for $\langle H_\star \rangle$ and subsequently use Milankovitch's relations (Eq. 2.17) to obtain $\dot{\mathbf{h}}_i$ and $\dot{\mathbf{e}}_i$.

2.3.1.2 Reformulation as a spin system

Following Klein (1924), we introduce the vectors

$$\mathbf{b}_{i+} = \mathbf{h}_i + \mathbf{e}_i; \quad \mathbf{b}_{i-} = \mathbf{h}_i - \mathbf{e}_i, \quad (2.19)$$

so that the identities from Eq. (2.15) simply become

$$|\mathbf{b}_{i\pm}| = 1. \quad (2.20)$$

This is an important rewriting since the system of N wires is now fully described by $2N$ unit vectors $\mathbf{b} = \mathbf{b}_{i\pm}$. The dynamics of these vectors is simply

$$\dot{\mathbf{b}}_{\pm} = \dot{\mathbf{h}} \pm \dot{\mathbf{e}}. \quad (2.21)$$

Since the vectors \mathbf{b} remain on the unit sphere, this equation can be expressed as a precession

$$\dot{\mathbf{b}} = \mathbf{B}(\mathbf{b}) = \boldsymbol{\Omega} \times \mathbf{b}. \quad (2.22)$$

The precession vectors $\boldsymbol{\Omega}$ are not uniquely determined through Eq. (2.21) since a component parallel to \mathbf{b} does not affect $\dot{\mathbf{b}}$. The most conservative choice for $\boldsymbol{\Omega}$ – in the sense that $|\boldsymbol{\Omega}|^2$ is minimised – is therefore

$$\boldsymbol{\Omega} = \mathbf{b} \times \dot{\mathbf{b}}, \quad (2.23)$$

for which $\boldsymbol{\Omega} \cdot \mathbf{b} = 0$, i.e. \mathbf{b} moves along a great circle for constant $\boldsymbol{\Omega}$.

One can easily rewrite Milankovitch's equations (2.17) in terms of the vectors \mathbf{b} , to obtain

$$\dot{\mathbf{b}} = \frac{2}{\Lambda} \frac{\partial \langle H_\star \rangle}{\partial \mathbf{b}} \times \mathbf{b}. \quad (2.24)$$

Comparing with Eq. (2.22), one could be tempted to identify $2\Lambda^{-1} \partial \langle H_\star \rangle / \partial \mathbf{b}$ with $\boldsymbol{\Omega}$. However, the gradient of the Hamiltonian is not unique since, owing to the constraints from Eq. (2.20), $\langle H_\star \rangle$ is only determined up to additive terms of the form $(1 - \mathbf{b}^2) \tilde{H}$, with \tilde{H} an arbitrary function. Such terms alter the gradient but have no effect on $\dot{\mathbf{b}}$, reflecting the gauge invariance of Eq. (2.24).

We are now set to compute the rates of change for each wire (§2.3.2.1), and to perform the time integration of their dynamics (§2.3.3). Throughout these sections, we test our algorithm in three systems, namely a simple analytical "Pair" Hamiltonian of two wires, Kozai–Lidov oscillations of $N = 2$ stars (whose orbits do not overlap radially) around a central BH, and a $N = 10^4$ stellar cluster mimicking SgrA*. We refer to Fouvry et al. (2022) for the detailed description of these setups.

2.3.2 Rates of changes

To compute the rates of change $\dot{\mathbf{h}}$ and $\dot{\mathbf{e}}$, we proceed in two steps: (i) we replace in Eq. (1.10) the averaging integrals with discrete sums (§2.3.2.1); (ii) we approximate the forces using spherical harmonics (§2.3.2.2). We refer to §B in Fouvry et al. (2022) for explicit expressions of the contributions to $\dot{\mathbf{h}}$ and $\dot{\mathbf{e}}$ from $\langle H_{\text{GR}} \rangle$, namely the Schwarzschild and Lense–Thirring precessions.

2.3.2.1 Discretised rates of change

In order to numerically calculate the orbit averages (Eq. 1.10), we approximate them by discrete midpoint sums over K positions along each orbit², namely

$$m_i \int_0^{2\pi} \frac{dM_i}{2\pi} f(M_i) \rightarrow \sum_{k=1}^K \mu_{ik} f(M_{ik}), \quad (2.25)$$

for any function $f(M)$. In other words, each eccentric wire of mass m_i is replaced by K nodes of masses μ_{ik} . Placing the nodes' mean anomalies, M_{ik} , uniformly (Gürkan & Hopman, 2007) is not the best idea, not only because it requires solving Kepler's Eq. (1.9) for each node, but also because it poorly samples the pericentric passage, in particular for large eccentricity, which hampers the convergence w.r.t. K . Instead, we place the nodes uniformly in eccentric anomaly

$$M_{ik} = E_k - e_i \sin(E_k) \quad \text{with} \quad E_k = (k - \frac{1}{2}) \Delta E, \quad (2.26)$$

with $\Delta E = 2\pi/K$. As such, the nodes are placed symmetrically w.r.t. the pericentre, but there is no node exactly at that position. To complete the discretisation from Eq. (2.25), we finally specify the nodes' masses

$$\mu_{ik} = \frac{m_i}{2\pi} \frac{dM_i}{dE} \Delta E = \frac{m_i}{K} (1 - e_i \cos(E_k)). \quad (2.27)$$

The discretised potential energy of node k on wire i subsequently reads

$$\phi_{ik} = \mu_{ik} \psi_{ik} \quad \text{with} \quad \psi_{ik} = - \sum_{j \neq i}^N \sum_{l=1}^K \frac{G \mu_{jl}}{|\mathbf{r}_{ik} - \mathbf{r}_{jl}|}. \quad (2.28)$$

The total interaction energy between the wires is then computed as

$$\langle H_\star \rangle = \frac{1}{2} \sum_{i=1}^N \sum_{k=1}^K \phi_{ik}. \quad (2.29)$$

²Adopting a different value of K for each wire, e.g., depending on eccentricity, is a straightforward extension.

It is key here to emphasise that $\langle H_\star \rangle$ depends on \mathbf{h}_i and \mathbf{e}_i through the nodes positions \mathbf{r}_{ik} but also through their masses μ_{ik} . This latter dependence occurs through the eccentricity e_i in Eq. (2.27) because of our sampling of the nodes in eccentric rather than mean anomaly. Via Milankovitch's Eq. (2.17), the dependence through μ_{ik} induces the rates of change (see §A in Fouvry et al., 2022)

$$\dot{\mathbf{h}}_i|_\mu = 0, \quad (2.30a)$$

$$\dot{\mathbf{e}}_i|_\mu = \frac{1}{\Lambda_i} (\mathbf{h}_i \times \hat{\mathbf{e}}_i) \sum_{k=1}^K \bar{\mu}_i \psi_{ik} \cos(E_k) \quad (2.30b)$$

with $\bar{\mu}_i = m_i/K$ the mean node mass on the wire. We emphasise that these rates of changes would have been missed if one had mistakenly directly discretised the forces acting on the wires, following Eq. (2.16).

The rates of change induced by the dependence of $\langle H_\star \rangle$ through \mathbf{r}_{ik} read (see §A in Fouvry et al., 2022)

$$\dot{\mathbf{h}}_i|_{\mathbf{r}} = \frac{a_i}{\Lambda_i} [\hat{\mathbf{e}}_i \times (\mathbf{C}_i - e_i \mathbf{F}_i) + (\mathbf{h}_i \times \hat{\mathbf{e}}_i) \times \mathbf{S}_i], \quad (2.31a)$$

$$\dot{\mathbf{e}}_i|_{\mathbf{r}} = \frac{a_i}{\Lambda_i} \left[\mathbf{F}_i \times \mathbf{h}_i - e_i \mathbf{S}_i + \frac{1}{e_i} [\mathbf{S}_i \cdot \hat{\mathbf{e}}_i - \mathbf{C}_i \cdot (\mathbf{h}_i \times \hat{\mathbf{e}}_i)] \hat{\mathbf{e}}_i \right], \quad (2.31b)$$

with $\mathbf{F}_i = \sum_k \mathbf{f}_{ik}$, $\mathbf{C}_i = \sum_k \mathbf{f}_{ik} \cos(E_k)$, and $\mathbf{S}_i = \sum_k \mathbf{f}_{ik} \sin(E_k)$. Here, we introduced

$$\mathbf{f}_{ik} = -\frac{\partial \phi_{ik}}{\partial \mathbf{r}_{ik}} = \sum_{j \neq i}^N \sum_{l=1}^K G \mu_{ik} \mu_{jl} \mathbf{F}(\mathbf{r}_{ik} - \mathbf{r}_{jl}), \quad (2.32)$$

which is the gravitational force acting on node k of wire i and generated by the nodes of all other wires.

Upon close inspection, one can rewrite Eq. (2.31a) as

$$\dot{\mathbf{h}}_i|_{\mathbf{r}} = \frac{1}{\Lambda_i} \sum_{k=1}^K \mathbf{r}_{ik} \times \mathbf{f}_{ik}, \quad (2.33)$$

which is identical to the discretisation of Eq. (2.16a) and implies the conservation of the total angular momentum, $\mathbf{L}_{\text{tot}} = \sum_i \Lambda_i \mathbf{h}_i$. This was expected from Noether's theorem since our discretisation of the wires is invariant under spatial rotations. We emphasise that the rates from Eq. (2.31) satisfy exactly $\dot{\mathbf{h}}_i \cdot \mathbf{h}_i + \dot{\mathbf{e}}_i \cdot \mathbf{e}_i = 0$ and $\dot{\mathbf{h}}_i \cdot \mathbf{e}_i + \dot{\mathbf{e}}_i \cdot \mathbf{h}_i = 0$, as required.

From Eqs. (2.30) and (2.31), we note that the computation of $\dot{\mathbf{h}}_i$ and $\dot{\mathbf{e}}_i$ requires $\mathcal{O}(N^2 K^2)$ computations. As shown in the coming section, this can be reduced to $\mathcal{O}(NK \ell_{\text{max}}^2)$ when approximating $|\mathbf{r}_{ik} - \mathbf{r}_{jl}|$ by its expansion in spherical harmonics up to order ℓ_{max} . To this end, it is advantageous to introduce the decomposition

$$\mathbf{f}_{ik} = \mathbf{f}_{ik}^{\text{all}} - \mathbf{f}_{ik}^{\text{self}} \quad \text{with} \quad \mathbf{f}_{ik}^{\text{all}} = \sum_{\{j,l\} \neq \{i,k\}}^{N,K} G \mu_{ik} \mu_{jl} \mathbf{F}(\mathbf{r}_{ik} - \mathbf{r}_{jl}), \quad \mathbf{f}_{ik}^{\text{self}} = \sum_{l \neq k}^K G \mu_{ik} \mu_{il} \mathbf{F}(\mathbf{r}_{ik} - \mathbf{r}_{il}). \quad (2.34)$$

Here, $\mathbf{f}_{ik}^{\text{all}}$ captures the contributions from all other nodes on all wires, while $\mathbf{f}_{ik}^{\text{self}}$ captures the contributions from the other nodes on the same wire.

We end this section by stressing that our algorithm for $\mathbf{f}_{ik}^{\text{all}}$ and $\mathbf{f}_{ik}^{\text{self}}$ must compute the exact same wire self-gravity so that when computing the difference $\mathbf{f}_{ik} = \mathbf{f}_{ik}^{\text{all}} - \mathbf{f}_{ik}^{\text{self}}$ these erroneous non-physical contributions exactly cancel. This is important, since the wire self-gravity can be substantial: it diverges logarithmically with the number K of nodes in the case of the exact (unsoftened) Newtonian gravity.

2.3.2.2 Multipole expansion

In order to accelerate the computation of the $\mathcal{O}(N^2 K^2)$ pairwise node interactions, we expand the Newtonian interaction kernel, $|\mathbf{r} - \mathbf{r}'|^{-1}$, in spherical harmonics via

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \sum_{\ell=0}^{+\infty} \sum_{m=-\ell}^{\ell} \begin{cases} U_{\ell m}(\mathbf{r}) T_{\ell m}(\mathbf{r}') & \text{if } r < r', \\ T_{\ell m}(\mathbf{r}) U_{\ell m}(\mathbf{r}') & \text{if } r' < r. \end{cases} \quad (2.35)$$

Here, we introduced the real-valued upper and lower solid spherical harmonics, $U_{\ell m}(\mathbf{r}) \propto r^\ell Y_{\ell m}(\hat{\mathbf{r}})$ and $T_{\ell m}(\mathbf{r}) \propto r^{-\ell+1} Y_{\ell m}(\hat{\mathbf{r}})$, with detailed conventions and recurrences spelled out in §C of Fouvry et al. (2022).

At this stage, we note that there exist essentially two flavours of code based on spherical harmonic expansion: (i) codes in which the radial forces are evaluated using a basis-function expansion (see, e.g., Hernquist

& Ostriker, 1992; Saha, 1993); (ii) codes in which the radial forces associated with a given spherical harmonic are evaluated exactly (see, e.g., Villumsen, 1982) using formulas involving the usual factor $r_{\min}^{\ell}/r_{\max}^{\ell+1}$, here absorbed into the definitions of $U_{\ell m}$ and $T_{\ell m}$. Our goal here is to follow this second avenue and tailor it to the case of Keplerian wires. The expansion from Eq. (2.35) is isotropic, even when truncated at order ℓ_{\max} . Hence, by virtue of Noether's theorem, it conserves the system's total angular momentum \mathbf{L}_{tot} . However, this expansion is not invariant under translation, so that the total linear momentum is typically not conserved in spherical harmonic codes. Fortunately, as already highlighted in Eq. (1.10), Keplerian wires do not induce any force on the central object. This ensures, by design, the conservation of the system's total linear momentum.

Inserting the multipole expansion from Eq. (2.35) into Eq. (2.34), we have

$$\mathbf{f}_{ik}^{\text{all}} = \sum_{j,l}^{N,K} \sum_{\ell,m} G \mu_{ik} \mu_{jl} \begin{cases} U_{\ell m}(\mathbf{r}_{jl}) \nabla T_{\ell m}(\mathbf{r}_{ik}) & \text{if } r_{jl} < r_{ik}, \\ T_{\ell m}(\mathbf{r}_{jl}) \nabla U_{\ell m}(\mathbf{r}_{ik}) & \text{if } r_{jl} > r_{ik}, \end{cases} \quad (2.36)$$

with an analogous expression for $\mathbf{f}_{ik}^{\text{self}}$. The crucial difference between this expression and Eq. (2.34) is that the dependence on the positions \mathbf{r}_{ik} and \mathbf{r}_{jl} has been factorised. Hence, the gradient terms depending on \mathbf{r}_{ik} can be taken outside of the sum over nodes after splitting it into an inner and outer part. This gives

$$\mathbf{f}_{ik}^{\text{all}} = \sum_{\ell,m} \mu_{ik} P_{\ell m}(r_{ik}) \nabla T_{\ell m}(\mathbf{r}_{ik}) + \sum_{\ell,m} \mu_{ik} Q_{\ell m}(r_{ik}) \nabla U_{\ell m}(\mathbf{r}_{ik}). \quad (2.37)$$

In that expression, we introduced the prefix sums

$$P_{\ell m}(r) = \sum_{j,l:r_{jl} < r} G \mu_{jl} U_{\ell m}(\mathbf{r}_{jl}); \quad Q_{\ell m}(r) = \sum_{j,l:r_{jl} > r} G \mu_{jl} T_{\ell m}(\mathbf{r}_{jl}), \quad (2.38)$$

which correspond to the multipoles of the distribution of all nodes inside and outside of radius r , respectively.

We point out that the expansion from Eq. (2.35) can also be used to define a cluster's total truncated total energy, $E_{\text{tot}}(\ell_{\max}, K)$, via Eq. (1.11), which we compute efficiently using the same algorithm as for the wire forces. In the limit $\ell_{\max} \rightarrow +\infty$, i.e. in the absence of any multipole expansion, Eq. (1.11) provides us with the full total energy, $E_{\text{tot}}^{\text{direct}}$, which can be computed via a direct $\mathcal{O}(N^2 K^2)$ sum over all node-node couplings.

To compute the rates of change in Eq. (2.37), we begin by sorting the nodes into ascending order of radius³. Since the nodes on each wire are already sorted when created, this sorting can be made in $\mathcal{O}(NK \ln(N))$ operations and is completely subdominant in the total operation budget of our code. Once the nodes sorted in radius, the multipoles $P_{\ell m}$ and $Q_{\ell m}$ can straightforwardly be computed by increasing and decreasing recurrences. Using $\alpha = \{i, k\}$ to denote the wires, we follow the algorithm

- 1 Sort all nodes in radius r_{α} .
- 2 Set $\mathbf{f}_{\alpha}^{\text{all}} = 0$ for all nodes α .
- 3 Set $P_{\ell m} = 0$.
- 4 For all nodes α in order of increasing radius r_{α} :
 - 4.1 For all (ℓ, m) :
 - 4.1.1 $\mathbf{f}_{\alpha}^{\text{all}} \leftarrow \mathbf{f}_{\alpha}^{\text{all}} + \mu_{\alpha} P_{\ell m} \nabla T_{\ell m}(\mathbf{r}_{\alpha})$.
 - 4.1.2 $P_{\ell m} \leftarrow P_{\ell m} + G \mu_{\alpha} U_{\ell m}(\mathbf{r}_{\alpha})$.
- 5 Set $Q_{\ell m} = 0$.
- 6 For all nodes α in order of decreasing radius r_{α} :
 - 6.1 For all (ℓ, m) :
 - 6.1.1 $\mathbf{f}_{\alpha}^{\text{all}} \leftarrow \mathbf{f}_{\alpha}^{\text{all}} + \mu_{\alpha} Q_{\ell m} \nabla U_{\ell m}(\mathbf{r}_{\alpha})$.
 - 6.1.2 $Q_{\ell m} \leftarrow Q_{\ell m} + G \mu_{\alpha} T_{\ell m}(\mathbf{r}_{\alpha})$.

The total operation count of the algorithm (2.39) is $\mathcal{O}(NK \ell_{\max}^2)$ with ℓ_{\max} the maximum harmonic order considered. The same algorithm, but restricted to the nodes from only one wire, calculates $\mathbf{f}_{\alpha}^{\text{self}}$ for that wire, and again requires $\mathcal{O}(NK \ell_{\max}^2)$ operations for all wires. This is the main benefit for the present method which, at the cost of discretised orbit averages and (truncated) harmonic expansions, has a computational complexity scaling linearly with the total number of wires.

The computations of $\{\dot{\mathbf{h}}, \dot{\mathbf{e}}\}_{\text{all}}$ and $\{\dot{\mathbf{h}}, \dot{\mathbf{e}}\}_{\text{self}}$ are completely independent (and of similar computational difficulty), so that they can be performed in parallel. Similarly, the recurrences over $P_{\ell m}$ and $Q_{\ell m}$, i.e. steps 4 and 6 in (2.39), are independent one from another (and of similar computational difficulty), so that they can also be performed in parallel. As a consequence, one can easily parallelise the previous algorithm over four cores. Further parallelisation requires the parallel computation of the prefix sums from Eq. (2.38). This is more complex, but not impossible, e.g., using the parallel scan algorithm from Ladner & Fischer (1980).

³For self-gravity to be correctly removed, it is mandatory for the radius sortings used to compute $\mathbf{f}_{ik}^{\text{all}}$ and $\mathbf{f}_{ik}^{\text{self}}$ to be consistent.

2.3.2.3 Convergence of the force calculations

In Fig. 2.5, we illustrate the dependence of the relative error in the truncated total energy, $E_{\text{tot}}(\ell_{\text{max}}, K)$, as a function of K . Similar plots for $\dot{\mathbf{h}}$ and $\dot{\mathbf{e}}$ can be found in Fouvy et al. (2022). For the Kozai–Lidov system, for

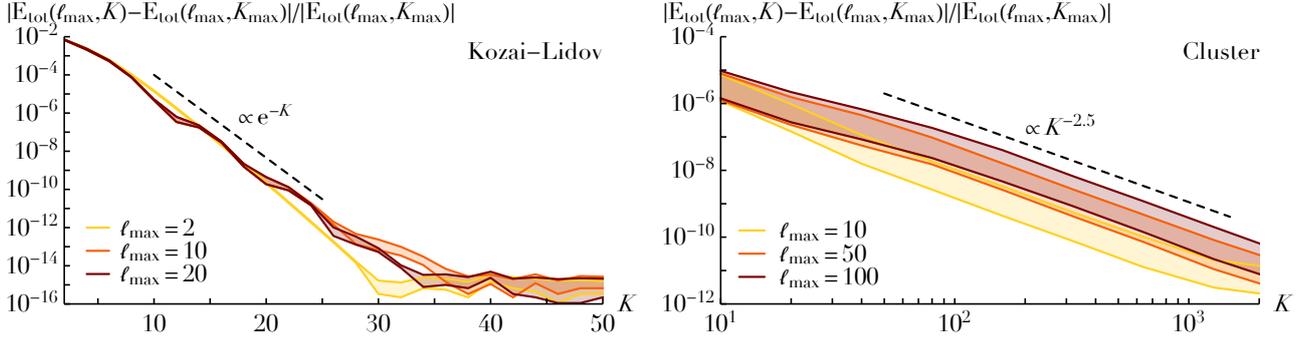


Figure 2.5: From Fouvy et al. (2022). Relative error in the truncated total energy, $E_{\text{tot}}(\ell_{\text{max}}, K)$, as a function of the number of nodes K for fixed multipole truncation order ℓ_{max} . Errors are computed w.r.t. to calculations with $K_{\text{max}} = 100$ for the Kozai–Lidov simulations (left panel) and $K_{\text{max}} = 20\,480$ for cluster simulations (right panel). Colored regions correspond to the 16% and 84% levels over 500 independent realisations.

a fixed value of ℓ_{max} , we find that the error decreases exponentially as e^{-K} . This is a direct consequence of the absence of any radial overlap between the two stellar orbits at play. In that case, the double orbit-average integral from Eq. (1.10) naturally splits into the product of two 1D integrals, which the multipole algorithm from Eq. (2.38) catches. Importantly, the integrands of each of these 1D integrals are 2π -periodic w.r.t. the eccentric anomaly. In that case, the midpoint sampling from Eq. (2.26) ensures an exponential convergence w.r.t. the number of nodes K (Trefethen & Weideman, 2014).

For the clusters composed of $N = 10^4$ stars, given that the orbits exhibit a wide range of eccentricities ($0 \leq e \leq 0.99$), radial overlaps are ubiquitous and the exponential convergence w.r.t. K does not hold anymore. More precisely, for a fixed value of ℓ_{max} , the error on $E_{\text{tot}}(\ell_{\text{max}}, K)$ decreases like $K^{-2.5}$. For the rates of change, we find that $\dot{\mathbf{h}}$ converges in $K^{-2.5}$, while $\dot{\mathbf{e}}$ converges in $K^{-1.5}$ (see fig. 1 in Fouvy et al. (2022)). We believe that this difference in the rates of convergence stems from the fact that the discontinuity in Eq. (2.36) is only in the radial force, which does not enter the equation for $\dot{\mathbf{h}}$.

In Fig. 2.6, we investigate the relative errors in $E_{\text{tot}}(\ell_{\text{max}}, K)$ as a function of ℓ_{max} for a fixed number of nodes K . Similar plots for $\dot{\mathbf{h}}$ and $\dot{\mathbf{e}}$ can be found in Fouvy et al. (2022). For the Kozai–Lidov simulations, we find that

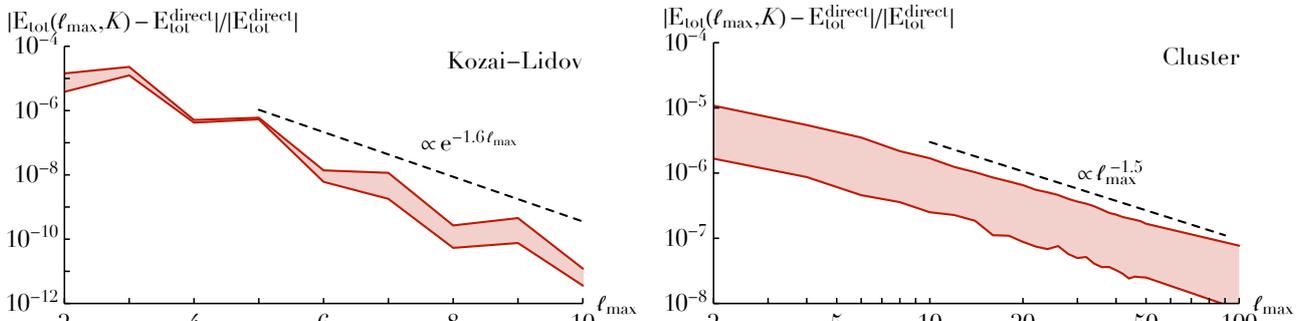


Figure 2.6: From Fouvy et al. (2022). Relative error in the truncated total energy, $E_{\text{tot}}(\ell_{\text{max}}, K)$, as a function of the multipole truncation order ℓ_{max} for a fixed number of nodes ($K = 100$ for Kozai–Lidov and $K = 20\,480$ for the cluster), using the same convention as in Fig. 2.5. The full total energy, $E_{\text{tot}}^{\text{direct}}$, was computed by a direct summation of Eq. (1.7) using $K = 20\,480$ for the Kozai–Lidov simulations and $K = 1\,000$ for the clusters.

the error decreases exponentially in $e^{-1.6 \ell_{\text{max}}}$. Such a rapid convergence stems from the absence of any radial overlap between the two stellar orbits. More precisely, glancing back at Eq. (2.35), the radial dependence of a typical term of the Legendre expansion is $(r_{\text{min}}/r_{\text{max}})^\ell \leq \eta^\ell$, with $\eta = 1/5$ the largest distance ratio that occurs as the inner and outer stars run through their orbits. As a consequence, when truncated at order ℓ_{max} , the typical error is expected to scale like $\eta^{\ell_{\text{max}}} \simeq e^{-1.6 \ell_{\text{max}}}$, matching the numerical measurement from Fig. 2.6.

In that same figure, we note that the convergence w.r.t. ℓ_{max} is much slower for the $N = 10^4$ clusters. This is a consequence of the very slow convergence of the Legendre expansion from Eq. (2.35) in any regime where the wires overlap in radius. Here, we empirically find that $E_{\text{tot}}(\ell_{\text{max}}, K)$ converges roughly like $\ell_{\text{max}}^{-1.5}$. This is compatible with the asymptotic scalings predicted in §B5 of Kocsis & Tremaine (2015) in the context of VRR. Indeed, for radially overlapping non-coplanar orbits, which mostly compose the present $N = 10^4$ clusters, a given harmonic ℓ gives a contribution to the total truncated energy scaling like $\ell^{-2.5}$. Once summing over all the harmonics $0 \leq \ell \leq \ell_{\text{max}}$, this ultimately leads to a relative error in $E_{\text{tot}}(\ell_{\text{max}}, K)$ scaling like $\ell_{\text{max}}^{-1.5}$. We finally

emphasise that the truncation of the present multipole approach effectively softens the interaction potential, thereby avoiding the singularities that would otherwise occur when wires cross.

2.3.3 Time integration

In our application, the state of the system is described by the N sets $y = \{\mathbf{b}_+, \mathbf{b}_i\}_i$. Given a timestep τ , an integration scheme is a mapping $y_n \rightarrow y_{n+1}$ where subscripts denote different time slices. Standard Runge–Kutta integration of the three-dimensional vectors \mathbf{b} generically leads to violations of the constraints $|\mathbf{b}|=1$. Instead, we resort here to integration schemes that, by construction, exactly comply with these constraints. For convenience, we define the velocity field, $\mathbf{B}(y) = \dot{\mathbf{h}} \pm \dot{\mathbf{e}}$, so that the equations of motion simply read $\dot{\mathbf{b}} = \mathbf{B}(y)$.

2.3.3.1 Explicit Lie group methods

For a fixed value of the precession vector $\boldsymbol{\Omega} = \mathbf{b} \times \mathbf{B}$ (see Eq. 2.23), and an initial condition \mathbf{b}_0 , Eq. (2.22) can be integrated exactly for a duration t , via Rodrigues' rotation formula, to the new location

$$\begin{aligned} \mathbf{b}(t) &= \cos(\Omega t) \mathbf{b}_0 + \sin(\Omega t) \widehat{\boldsymbol{\Omega}} \times \mathbf{b}_0 + [1 - \cos(\Omega t)] (\mathbf{b}_0 \cdot \widehat{\boldsymbol{\Omega}}) \widehat{\boldsymbol{\Omega}} \\ &= \phi[t\boldsymbol{\Omega}] \circ \mathbf{b}_0, \end{aligned} \quad (2.40)$$

with $|\mathbf{b}(t)| = |\mathbf{b}_0|$. In order to exactly preserve the constraints $|\mathbf{b}| = 1$, explicit integration schemes analogous to Runge–Kutta methods can be devised via Lie group methods (see, e.g., §IV.8 in Hairer et al., 2006) by concatenating appropriately rotations with precession vectors obtained at various intermediate stages. In practice, we use Munthe–Kaas integrators for their simplicity (Munthe-Kaas, 1999). To do so, one rewrites the dynamics as

$$\dot{\mathbf{b}}(t) = \phi[\mathbf{U}(t)] \circ \mathbf{b}_0, \quad (2.41)$$

which imposes that $\mathbf{U}(t)$ evolves according to

$$\dot{\mathbf{U}} = d\phi_{\mathbf{U}}^{-1} \circ \{\boldsymbol{\Omega}(\phi[\mathbf{U}] \circ \mathbf{b}_0)\}, \quad (2.42)$$

with $\boldsymbol{\Omega}(\mathbf{b}) = \mathbf{b} \times \mathbf{B}(\mathbf{b})$ following Eq. (2.23). Here, $d\phi_{\mathbf{U}}^{-1}$ is the inverse of the differential of the rotation map (see Theorem 3 in Munthe-Kaas, 1999). It generically reads

$$d\phi_{\mathbf{U}}^{-1} \circ \boldsymbol{\Omega} = \sum_{k=0}^{+\infty} \frac{B_k}{k!} \text{ad}_{\mathbf{U}}^k \circ \boldsymbol{\Omega}, \quad (2.43)$$

with B_k the Bernoulli numbers and $\text{ad}_{\mathbf{U}}^k$ the k -th power of the adjoint operator. In the present context, this operator simply reads

$$\text{ad}_{\mathbf{U}}^0 \circ \boldsymbol{\Omega} = \boldsymbol{\Omega}; \quad \text{ad}_{\mathbf{U}}^k \circ \boldsymbol{\Omega} = \mathbf{U} \times (\text{ad}_{\mathbf{U}}^{k-1} \circ \boldsymbol{\Omega}). \quad (2.44)$$

A Munthe–Kaas scheme proceeds then by using a classical Runge–Kutta scheme applied to $\mathbf{U}(t)$ and "correcting" the intermediate precession vectors with $d\phi_{\mathbf{U}}^{-1}$ appropriately truncated (Theorem IV.8.5 in Hairer et al., 2006). For a second-order scheme, one can use $d\phi_{\mathbf{U}}^{-1} \circ \boldsymbol{\Omega} = \boldsymbol{\Omega}$, from which one constructs the two-stage explicit midpoint rule (coined MK2). Starting from an initial state \mathbf{b}_n , it proceeds via

$$\begin{aligned} \boldsymbol{\Omega}_1 &= \boldsymbol{\Omega}(\mathbf{b}_n), \\ \mathbf{b}_2 &= \phi\left[\frac{\tau}{2}\boldsymbol{\Omega}_1\right] \circ \mathbf{b}_n, \\ \boldsymbol{\Omega}_2 &= \boldsymbol{\Omega}(\mathbf{b}_2), \\ \mathbf{b}_{n+1} &= \phi[\tau\boldsymbol{\Omega}_2] \circ \mathbf{b}_n. \end{aligned} \quad (2.45)$$

Naturally, the rotations are performed over all the elements of y simultaneously.

To construct higher order methods, one has to account for the fact that the rotations from Eq. (2.40) do not commute, i.e. better approximations of $d\phi_{\mathbf{U}}^{-1}$ have to be used. Here, we present a four-stage scheme (coined

MK4) based on the classical fourth-order Runge–Kutta scheme. It reads

$$\begin{aligned}
\tilde{\Omega}_1 &= \Omega(\mathbf{b}_n), \\
\mathbf{b}_2 &= \phi[\frac{\tau}{2}\tilde{\Omega}_1] \circ \mathbf{b}_n, \\
\tilde{\Omega}_2 &= d\phi_{\frac{\tau}{2}\tilde{\Omega}_1}^{-1} \circ \Omega(\mathbf{b}_2), \\
\mathbf{b}_3 &= \phi[\frac{\tau}{2}\tilde{\Omega}_2] \circ \mathbf{b}_n, \\
\tilde{\Omega}_3 &= d\phi_{\frac{\tau}{2}\tilde{\Omega}_2}^{-1} \circ \Omega(\mathbf{b}_3), \\
\mathbf{b}_4 &= \phi[\tau\tilde{\Omega}_3] \circ \mathbf{b}_n, \\
\tilde{\Omega}_4 &= d\phi_{\tau\tilde{\Omega}_3}^{-1} \circ \Omega(\mathbf{b}_4), \\
\mathbf{b}_{n+1} &= \phi[\tau(\frac{1}{6}\tilde{\Omega}_1 + \frac{1}{3}\tilde{\Omega}_2 + \frac{1}{3}\tilde{\Omega}_3 + \frac{1}{6}\tilde{\Omega}_4)] \circ \mathbf{b}_n,
\end{aligned} \tag{2.46}$$

where $d\phi_{\mathbf{U}}^{-1}$ is truncated at second order in \mathbf{U} so that Eq. (2.43) becomes

$$d\phi_{\mathbf{U}}^{-1} \circ \Omega = \Omega - \frac{1}{2}\mathbf{U} \times \Omega + \frac{1}{12}\mathbf{U} \times (\mathbf{U} \times \Omega). \tag{2.47}$$

The schemes from Eqs. (2.45) and (2.46) are (i) explicit; (ii) conserve $|\mathbf{b}|=1$ exactly⁴; (iii) require two or four computations of the derivatives, respectively; (iv) are, respectively, second- and fourth-order accurate. Because they are direct translations of usual Runge–Kutta methods, it is straightforward to design Munthe–Kaas schemes of higher order.

2.3.3.2 Symplectic scheme

Following McLachlan et al. (2014), we also consider an integrator based on the spherical midpoint method (coined MD2). It is based on the implicit relation

$$\frac{\mathbf{b}_{n+1} - \mathbf{b}_n}{\tau} = \mathbf{B} \left(\frac{\mathbf{b}_{n+1} + \mathbf{b}_n}{|\mathbf{b}_{n+1} + \mathbf{b}_n|} \right). \tag{2.48}$$

This implicit relation is solved via the fixed-point iteration

$$\mathbf{b}_{n+1}^{(0)} = \mathbf{b}_n; \quad \mathbf{b}_{n+1}^{(k+1)} = \mathbf{b}_n + \tau \mathbf{B} \left(\frac{\mathbf{b}_n + \mathbf{b}_{n+1}^{(k)}}{|\mathbf{b}_n + \mathbf{b}_{n+1}^{(k)}|} \right), \tag{2.49}$$

terminating when $\|\mathbf{b}_{n+1}^{(k+1)} - \mathbf{b}_{n+1}^{(k)}\| \leq \epsilon$, for a given tolerance ϵ . Naturally, these iterations are performed for all the elements of y simultaneously. In practice, we impose $\epsilon = 2.2 \times 10^{-15}$, which typically requires 5–12 iterations.

The scheme from Eq. (2.48) is (i) second-order accurate; (ii) implicit; (iii) exactly conserves the constraints from Eq. (2.20)⁵; (iv) symplectic (McLachlan et al., 2014); (v) and conserves $\mathbf{L}_{\text{tot}} = \sum_i \Lambda_i \mathbf{h}_i$, as it is a global linear invariant. Because MD2 is a symmetric scheme, one can use symmetric composition techniques (see, e.g., §V.3.2 in Hairer et al., 2006) to devise higher order symplectic integrators.

The MK2, MK4, MD2 schemes conserve the constraints for the vectors (\mathbf{h}, \mathbf{e}) to machine precision, which as far as we know, none of the previous studies (e.g., Touma et al., 2009; Hamers & Portegies Zwart, 2016) have managed.

2.3.3.3 Convergence of the time integrations

As a first check of the sanity of the present algorithm, we compare in the left panel of Fig. 2.7 the time-evolution of the Kozai–Lidov system as predicted by the multipole approach and a direct integration of the associated 3-body problem. Both methods predict similar oscillations of the inner star⁶. On average, the multipole approach, owing to its explicit orbit-average, used a timestep $\sim 10^5$ times larger than the one of the direct integration.

In order to further assess the performance of the integration schemes, we now consider a simple analytical two-body "Pair" Hamiltonian (see §D in Fouvry et al., 2022). In the right panel of Fig. 2.7, we illustrate the dependence of the relative error in E_{tot} after a finite time as a function of τ . We find that MK2 and MD2 are second-order accurate, with a finite-time error in $\mathcal{O}(\tau^2)$, while MK4 is fourth-order accurate.

In Fig. 2.8, for the same Hamiltonian, we investigate the long-time trends for the errors in E_{tot} and \mathbf{L}_{tot} . As expected, the explicit schemes show errors growing like t . There are two main improvements in the symplectic

⁴We systematically re-normalise via $\mathbf{b} \leftarrow \mathbf{b}/|\mathbf{b}|$ after every evaluation of Eq. (2.40) to prevent a drift of $|\mathbf{b}|$ from round-off errors.

⁵The fixed-point outcome from Eq. (2.49) is systematically renormalised via $\mathbf{b} \leftarrow \mathbf{b}/|\mathbf{b}|$ to prevent a drift of $|\mathbf{b}|$ from round-off errors.

⁶When represented in Fig. 2.7, the MK2, MK4 and MD2 integrations are indistinguishable.

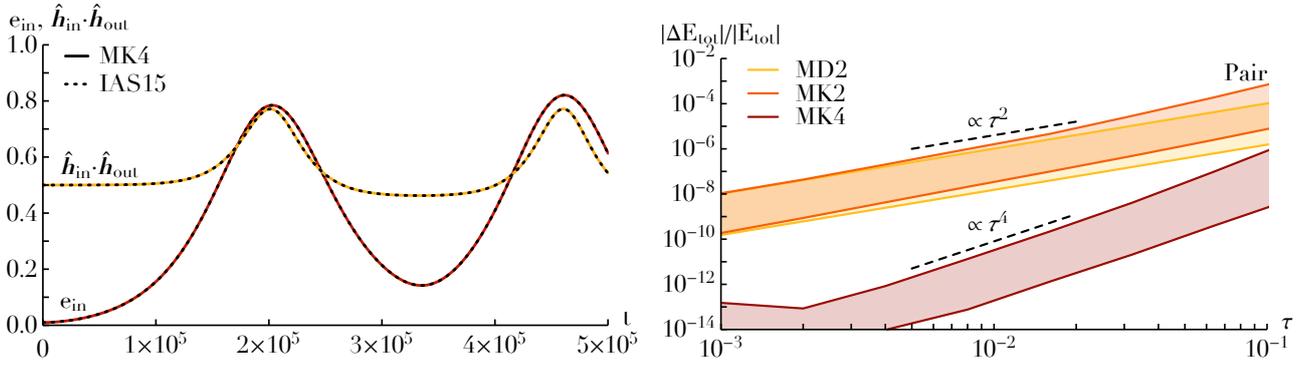


Figure 2.7: From Fouvy et al. (2022). Left: Illustration in the Kozai–Lidov simulations of the evolution of the inner star’s eccentricity, e_{in} , as well as its inclination w.r.t. the outer star, $\hat{\mathbf{h}}_{\text{in}} \cdot \hat{\mathbf{h}}_{\text{out}}$. Full lines correspond to the multipole integration (with the MK4 scheme from Eq. 2.46), while dashed lines were obtained by direct integration using the IAS15 integrator (Rein & Spiegel, 2015). Right: Relative error in the total energy E_{tot} w.r.t. its initial value after a fixed time, as a function of the considered integration timestep. Colored regions are the 16% and 84% levels over 500 independent realisations of the "Pair" Hamiltonian.

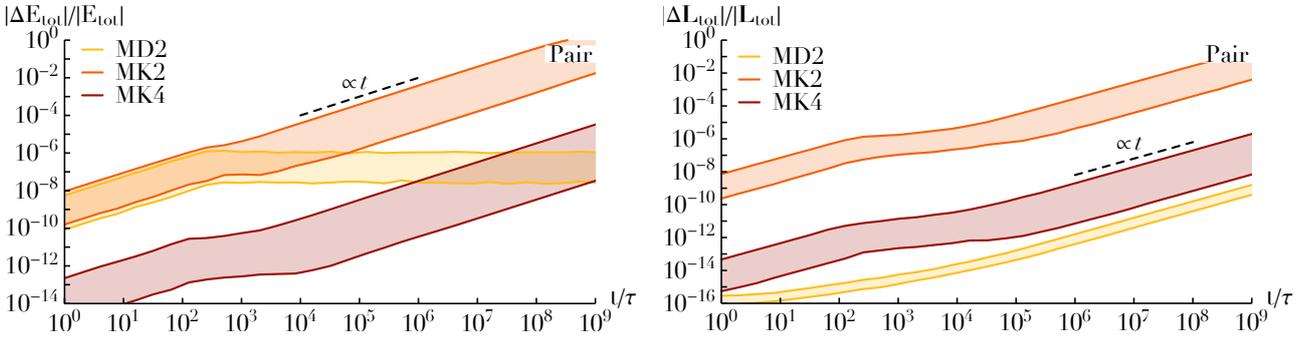


Figure 2.8: From Fouvy et al. (2022). Relative error in the total energy E_{tot} (left) and angular momentum \mathbf{L}_{tot} (right) w.r.t. their initial values, as a function of the number of timesteps, t/τ , using the same convention as in the right panel of Fig. 2.7.

scheme: (i) the error in E_{tot} is bounded on long timescales; (ii) the error in \mathbf{L}_{tot} only grows via the accumulation of round-off errors. Figure 9 of Fouvy et al. (2022) presents similar measurements for the cluster simulations, and reaches the same conclusions.

Finally, in Fig. 2.9, we illustrate the error-cost relation in the cluster simulations. The cost of a given simula-

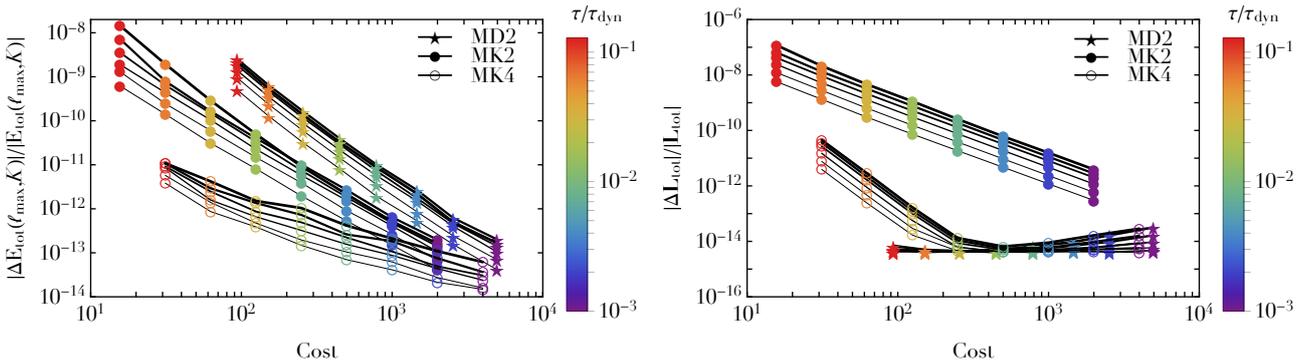


Figure 2.9: From Fouvy et al. (2022). Error-cost relation for the cluster simulations with $N = 10^4$ particles. Simulations were performed with (i) the symplectic scheme (MD2; star symbols), or the explicit ones (MK2 and MK4; circle symbols); with (ii) different integration timesteps $\tau/\tau_{\text{dyn}} \propto 2^{-k} \times 10^{-3}$ ($0 \leq k \leq 7$; different colors); (iii) up to different finite times, $T/\tau_{\text{dyn}} = 0.128 \times 2^{-k}$ ($0 \leq k \leq 5$): the larger T , the thicker the line. The cost of a given simulation is estimated via $\langle \# \text{forces} \rangle / (\tau/\tau_{\text{dyn}})$ with $\langle \# \text{forces} \rangle$ the average number of force evaluations per integration timestep, e.g., two for MK2. Median errors in $E_{\text{tot}}(\ell_{\text{max}}, K)$ (left) and \mathbf{L}_{tot} (right) are estimated over 128 realisations. See fig. 10 in Fouvy et al. (2022) for detailed definitions.

tion is estimated via $\langle \# \text{forces} \rangle / (\tau/\tau_{\text{dyn}})$, with $\langle \# \text{forces} \rangle$ the average number of force evaluations per integration timestep and τ_{dyn} the shortest dynamical time of the clusters at hand (see §F in Fouvy et al., 2022). While the explicit schemes always perform a constant number of force evaluations per timestep, the symplectic scheme requires from 5 evaluations (smallest τ) up to 12 (largest τ). As one reduces τ , the error in $E_{\text{tot}}(\ell_{\text{max}}, K)$ keeps shrinking illustrating that our discretisation scheme indeed corresponds to a Hamiltonian system, see Eq. (2.18c). Finally, for the present rather short integrations, at a given cost, the explicit scheme outperforms

the symplectic one regarding the conservation of $E_{\text{tot}}(\ell_{\text{max}}, K)$, while the converse holds regarding L_{tot} .

2.3.4 Application to SgrA*

We conclude this section by using the precedent method to simulate a typical $N = 10^4$ cluster around SgrA*. In particular, we use these simulations to validate the SRR diffusion coefficients derived in Eq. (2.7). Details for the measurements and predictions are given in §F of Fouvry et al. (2022).

In Fig. 2.10, we present the (finite-time) eccentricity diffusion coefficients, $D_{hh} = \{[\Delta h(T)]^2\}/T$, where $\{\cdot\}$ stands for an ensemble-average, and we used a finite time $T = 25$ kyr, along with $\ell_{\text{max}} = 10$ and $K = 100$. This

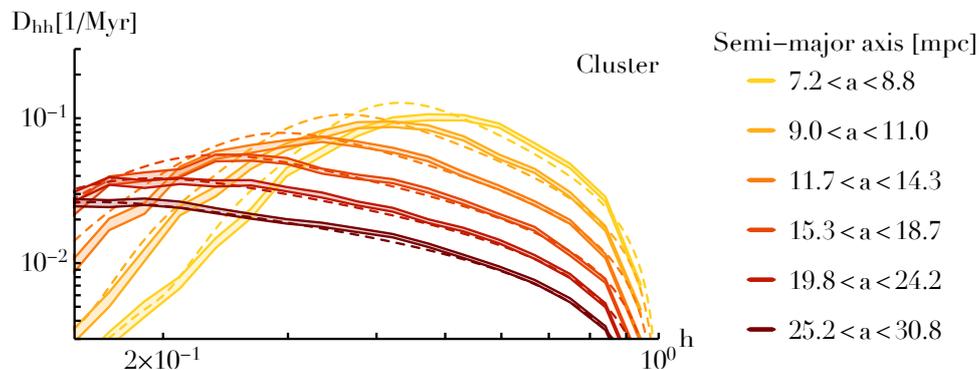


Figure 2.10: From Fouvry et al. (2022). (Finite-time) diffusion coefficients $D_{hh} = \{[\Delta h(T)]^2\}/T$, with $h = \sqrt{1-e^2}$, in a galactic nucleus mimicking SgrA*. Wires are split in different bins of semi-major axes (a) in milliparsecs, represented with different colors. The dashed curves correspond to the SRR theoretical predictions from Eq. (2.7) corrected for finite-time effects (see Eq. 24 in Bar-Or & Fouvry, 2018). The thick curves correspond to the numerical measurements performed in the multipole simulations and their associated bootstrap errors.

figure shows a good agreement between the diffusion coefficients inferred from the numerical simulations and the ones inferred from kinetic theory. It is likely that the remaining (slight) mismatch stems from the difficulty of measuring/predicting long-term diffusion coefficients on such a short finite-time, T , see Eq. (24) in Bar-Or & Fouvry (2018). Overall, Fig. 2.10 illustrates the sanity of the present algorithm which reproduced the intricate resonant dynamical interactions of Keplerian wires in a galactic nucleus. Of course, Fig. 2.10 is only a first illustration of the physical mechanisms that can be investigated with the multipole method.

2.3.5 Discussion

The present multipole approach differs from the alternate Gauss' method (Touma et al., 2009) which does not rely on a harmonic expansion. In a nutshell, Gauss' method computes all the pairwise interactions between the wires, i.e. it has a computational complexity scaling quadratically in the number of wires. More precisely, it scales like $\mathcal{O}(N^2 K_G)$ with K_G the typical number of points used to perform the second orbit-average, which cannot be performed analytically. Gauss' method has several strengths: (i) it is very well suited to perform time integrations of orbit-averaged problems with a small number of wires; (ii) because the computation of the orbit-averaged force between two wires is a numerically expensive calculation, i.e. a computational complexity in $\mathcal{O}(K_G)$, this method can significantly benefit from parallelisation; (iii) because it relies on the computation of pairwise interactions, Gauss' method lends itself nicely to multi-timestep approach contrary to the multipole method; (iv) Gauss' method is straightforward to apply in the case of a softened interaction (Touma et al., 2009), which the present multipole approach implicitly introduces through its harmonic truncation. However, one the main drawback of Gauss' method is that its computational difficulty gets prohibitively large as N gets larger, an aspect that the present multipole method alleviates by design.

2.3.6 Prospects

In this section, we described a multipole expansion method to perform orbit-averaged simulations of Keplerian wires in galactic nuclei. This algorithm has two main advantages: (i) by performing an explicit orbit average over the fast Keplerian motion, it greatly reduces the range of dynamical times in the system; (ii) by relying on a multipole expansion of the pairwise interaction, it benefits from a computational complexity scaling linearly with the total number of particles.

The present code is only a first step towards the development of ever faster integration methods for the dynamics of galactic nuclei, here seen as archetypes of systems that suffer from global resonance conditions in their orbital frequencies. Let us conclude by mentioning some venues for future works.

- As already discussed, the present algorithm can naturally be parallelised over four cores. Further parallelisation requires the computation of Eq. (2.38) in parallel. This asks for further developments.
- Even after the orbit average over the fast Keplerian motion induced by the central BH, the range of dynamical times in a typical nucleus still remains very significant. This is a consequence of the divergence of the relativistic precession for very eccentric wires, as well as of the wide ranges of semi-major axes and individual masses these systems comprise. Even with the present method, long time integrations still remain challenging. One likely solution for some of these difficulties would be to implement different timesteps for different wires. This is no easy feat, since to compute the prefix sums from Eq. (2.38), it is essential to scroll through the entire population of nodes.
- The present algorithm has problems with nodes at very similar radii. Indeed, the multipole expansion from Eq. (2.35) fails to converge in the limit $r_i \rightarrow r_j$ even if $|\mathbf{r}_i - \mathbf{r}_j| > 0$, i.e. even if the nodes are spatially well separated. We expect that such an issue can be cured by reverting to direct summation of the forces for nodes with near-identical radii. This only requires a slight modification of the present algorithm. This change would also call for the inclusion of force softening to avoid artificially large forces between nodes that are much closer to one another than their respective wire-companions. Overall, such improvements are expected to increase the numerical accuracy and stability of the scheme.

3 Vector Resonant Relaxation

The work presented in this section is based on Fouvry et al. (2019b,c); Giral Martínez et al. (2020); Magnan et al. (2021); Fouvry (2022).

In this section, we focus on VRR, the process through which stars' orbital orientations vary. VRR plays a crucial role in warping accretion (Bregman & Alexander, 2012) and stellar discs (Kocsis & Tremaine, 2011), enhancing binary mergers rate (Hamers et al., 2018), dissolving stellar discs (Bartko et al., 2009; Lu et al., 2009; Yelda et al., 2014), or aligning IMBHs' orientations (Szölgvény & Kocsis, 2018). Here, our goal is first to characterise the stochastic VRR Hamiltonian driving the random walk of stellar orientations (§3.1) and the separation of nearby orbits (§3.2). As for numerical approaches, we detail how entropy methods allow for the characterisation of the VRR's non-trivial thermodynamical equilibria (§3.3) and how multipole methods offer efficient integration schemes (§3.4). Finally, we tailor the BL kinetic theory to the axisymmetric limit of VRR (§3.5).

3.1 Statistics of VRR

In this first section, we set out to characterise VRR in the limit of an isotropic distribution of stars. To do so, we first present in §3.1.1 the fundamental equations of VRR. In §3.1.2, we characterise the properties of the system's potential fluctuations, as inferred from estimates of correlations at the initial time. This allows us then to describe in §3.1.3 the random walk of a test particle's orientation and to develop an effective stochastic equation of motion to efficiently mimic it. Finally, in §3.1.4, we detail the fundamental self-consistency existing between the potential fluctuations and the orientations' random walks they induce.

3.1.1 VRR Hamiltonian

We consider a set of $N \gg 1$ stars orbiting a supermassive BH, that we call the background bath of particles. For example, they represent SgrA*'s unresolved stellar populations (low mass stars, IMBHs) which all contribute to the clumpiness of the gravitational potential. Provided that one considers the dynamics of stars on timescales longer than the in-plane precession, but shorter than the relaxation time for eccentricity (by SRR) and energy (by NR), one can orbit average the stars along their respective mean anomalies and pericentre phases: stars are smeared out into annuli (left panel of Fig. 3.1). Each annuli is characterised by the conserved parameters

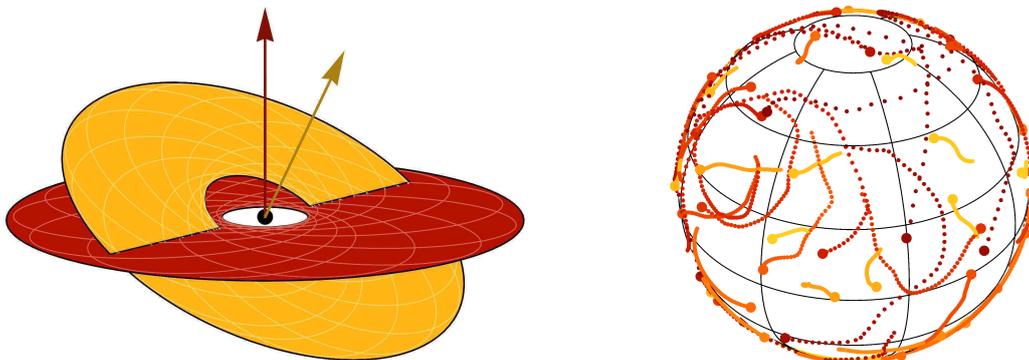


Figure 3.1: Left: From Giral Martínez et al. (2020). Illustration of the interaction between two annuli with different orbital parameters $\mathbf{K} = (m, a, e)$. The coupling between the two annuli depends on the relative angle, $\cos(\phi)$, between the two normals. Right: From Fouvry et al. (2019c). Typical correlated random walks in orientations on the unit sphere undergone by stellar orientations during VRR. Particles are colored according to their semi-major axis (from red for small a to yellow for large a). Given that the angular momentum $L(\mathbf{K})$ is an increasing function of a , particles with larger a see their orientations evolve slower.

$$\mathbf{K} = (m, a, e), \quad (3.1)$$

where m is the star's individual mass, a its semi-major axis, and e its eccentricity. Importantly, the norm

$$L(\mathbf{K}) = m\sqrt{GM_{\bullet}a(1-e^2)}. \quad (3.2)$$

of the angular momentum vector of each annuli is conserved. In that limit, the only remaining dynamical quantity is the instantaneous orbital orientation that we denote $\hat{\mathbf{L}}$. As such, VRR describes the dynamics of a set of long-range coupled unit vectors $\hat{\mathbf{L}}_i$.

Following Eq. (1.11), the system's total Hamiltonian reads¹

$$H_{\text{VRR}} = - \sum_{i < j}^N \left\langle \frac{Gm_i m_j}{|\mathbf{r}_i(t) - \mathbf{r}_j(t')|} \right\rangle_{t, t'}, \quad (3.3)$$

where the sum over (i, j) runs over all the pairs of particles and the double orbit-average $\langle \cdot \rangle_{t, t'}$ operates over the fast Keplerian motion and in-plane precession of both particles with $\mathbf{r}_i(t)$ and $\mathbf{r}_j(t')$ describing their respective unperturbed in-plane trajectories. Building upon a harmonic expansion, Kocsis & Tremaine (2015) shows that the system's total Hamiltonian can be rewritten as

$$\begin{aligned} H_{\text{VRR}} &= - \sum_{i < j}^N \sum_{\substack{\ell > 2 \\ \text{even}}}^{+\infty} \frac{2\ell + 1}{4\pi} \mathcal{H}_\ell[\mathbf{K}_i, \mathbf{K}_j] P_\ell(\hat{\mathbf{L}}_i \cdot \hat{\mathbf{L}}_j) \\ &= - \sum_{i < j}^N \sum_{\substack{\ell > 2 \\ \text{even}}}^{+\infty} \sum_{m = -\ell}^{\ell} \mathcal{H}_\ell[\mathbf{K}_i, \mathbf{K}_j] Y_{\ell m}(\hat{\mathbf{L}}_i) Y_{\ell m}(\hat{\mathbf{L}}_j), \end{aligned} \quad (3.4)$$

with the real surface spherical harmonics, $Y_{\ell m}(\hat{\mathbf{L}})$, normalised so that $\int d\hat{\mathbf{L}} |Y_{\ell m}(\hat{\mathbf{L}})|^2 = 1$. Equation (3.4) involves the (symmetric) coupling coefficients, $\mathcal{H}_\ell[\mathbf{K}_i, \mathbf{K}_j]$, which only depend on the stars' conserved parameters. They generically read

$$\mathcal{H}_\ell[\mathbf{K}, \mathbf{K}'] = \frac{4\pi Gmm'}{2\ell + 1} |P_\ell(0)|^2 \int_0^\pi dM \int_0^\pi dM' \frac{\text{Min}[r, r']^\ell}{\text{Max}[r, r']^{\ell+1}}; \quad \mathcal{J}_\ell[\mathbf{K}, \mathbf{K}'] = \frac{1}{L[\mathbf{K}]} \mathcal{H}_\ell[\mathbf{K}, \mathbf{K}'], \quad (3.5)$$

with P_ℓ the Legendre polynomials, and M and M' the mean anomalies of both orbits. We also introduced the "specific" coupling coefficients, $\mathcal{J}_\ell[\mathbf{K}, \mathbf{K}']$. We emphasise two important symmetries of these coefficients: (i) $\mathcal{H}_\ell[\mathbf{K}, \mathbf{K}'] = \mathcal{H}_\ell[\mathbf{K}', \mathbf{K}]$; (ii) $\mathcal{H}_\ell[\mathbf{K}, \mathbf{K}'] = \mathcal{J}_\ell[\mathbf{K}, \mathbf{K}'] = 0$ for any odd ℓ . In order not to have to invert Kepler's equation, it is convenient to perform the integrals from Eq. (3.5) w.r.t. the eccentric anomalies E and E' (Eq. 1.9). Naturally, benefiting from the quasi-separability of the Min-Max ratio, one can use the same multipole approach as in §2.3.2.2 to compute efficiently these coefficients. Figure 3.2 illustrates the typical shape of these coupling coefficients. We note that a given particle is strongly coupled to particles that share similar conserved

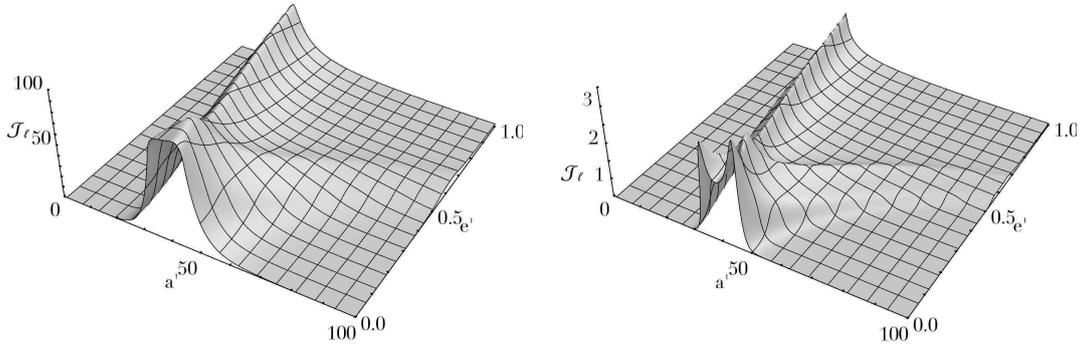


Figure 3.2: From Giral Martínez et al. (2020). Illustration of the coupling coefficients, $\mathcal{J}_\ell[\mathbf{K}, \mathbf{K}']$, for $\ell=10$ (left) and $\ell=40$ (right). Here, we chose $(a, e) = (40, 0.15)$, and the vertical axes have been commonly arbitrarily rescaled. One can note that a given orbit mostly couples with orbits with similar semi-major axes.

parameters, in particular similar semi-major axes. Moreover, as ℓ increases, the amplitude of the coupling coefficient, \mathcal{J}_ℓ , reduces (see §B of Kocsis & Tremaine, 2015, for detailed scalings).

The instantaneous state of the bath is characterised by its discrete DF

$$F^{\text{b}}(\hat{\mathbf{L}}, \mathbf{K}, t) = \sum_{i=1}^N \delta_{\text{D}}(\hat{\mathbf{L}} - \hat{\mathbf{L}}_i(t)) \delta_{\text{D}}(\mathbf{K} - \mathbf{K}_i), \quad (3.6)$$

which follows the normalisation $\int d\hat{\mathbf{L}} d\mathbf{K} F^{\text{b}} = N$. Describing VRR amounts to describing the time-evolution of $F^{\text{b}}(\hat{\mathbf{L}}, \mathbf{K}, t)$.

The bath's DF can naturally be expanded in spherical harmonics as

$$F^{\text{b}}(\hat{\mathbf{L}}, \mathbf{K}, t) = \sum_{\alpha} F_{\alpha}^{\text{b}}(\mathbf{K}, t) Y_{\alpha}(\hat{\mathbf{L}}), \quad (3.7)$$

¹In Eq. (3.3), we neglected the Lense-Thirring precession, as it is negligible for the S-cluster (see, e.g., fig. 1 in Kocsis & Tremaine, 2011)

with the shorthand notation $\alpha = \{\ell, m\}$. In Eq. (3.7), we introduced the "magnetisations", $F_\alpha^b(\mathbf{K}, t)$, as

$$F_\alpha^b(\mathbf{K}, t) = \int d\hat{\mathbf{L}} F^b(\hat{\mathbf{L}}, \mathbf{K}, t) Y_\alpha(\hat{\mathbf{L}}). \quad (3.8)$$

When expanded in spherical harmonics, the system's evolution follows (Fouvry et al., 2019c)

$$\frac{\partial F_\alpha^b(\mathbf{K}, t)}{\partial t} = - \sum_{\gamma, \delta} \int d\mathbf{K}' Q_{\alpha\gamma\delta}[\mathbf{K}, \mathbf{K}'] F_\gamma^b(\mathbf{K}', t) F_\delta^b(\mathbf{K}, t), \quad (3.9)$$

with the time-independent coupling tensors $Q_{\alpha\gamma\delta}[\mathbf{K}, \mathbf{K}'] = \mathcal{J}_{\ell_\gamma}[\mathbf{K}, \mathbf{K}'] E_{\alpha\gamma\delta}$, and $E_{\alpha\gamma\delta}$ the (real) Elsasser coefficients (James, 1973). See §B of Fouvry et al. (2019c) for a detailed presentation of these coefficients. Equation (3.9) is an exact writing of the fundamental evolution equation for VRR. Its difficulty stems from being a quadratic matrix differential equation in the fields $F_\alpha^b(\mathbf{K}, t)$. The right panel of Fig. 3.1 illustrates typical trajectories on the unit sphere during VRR: the motions are large-scale and time-correlated.

3.1.2 Noise correlation

Rather than describing the exact fate of all the particles, one can aim at characterising the statistical properties of the magnetisations $F_\alpha^b(\mathbf{K}, t)$. This was the purpose of Fouvry et al. (2019c).

Because they are generated by $N \gg 1$ particles, it is reasonable to assume that $F_\alpha^b(\mathbf{K}, t)$ are Gaussian random fields. Assuming that the system's dynamics is stationary in time, we may then fully characterise these fields by their correlation function

$$C_{\alpha\beta}^b(\mathbf{K}, \mathbf{K}', t - t') = \langle F_\alpha^b(\mathbf{K}, t) F_\beta^b(\mathbf{K}', t') \rangle, \quad (3.10)$$

where $\langle \cdot \rangle$ stands for the average over realisations of the system, i.e. over the initial conditions of the bath. Importantly, we assume that the background bath is isotropic on average, so that $\langle F_\alpha^b \rangle = 0$ for any $\ell_\alpha > 0$.

The bath correlation function is even w.r.t. time, and generically decreases on a timescale larger than some coherence time T_c . As a result, as a first approximation, it is reasonable to replace $C_{\alpha\beta}^b(\mathbf{K}, \mathbf{K}', t - t')$ by a Gaussian function tailored to match the function's behaviour for $t \ll T_c$. More precisely, one can compute the expectations for $C_{\alpha\beta}^b(\mathbf{K}, \mathbf{K}', 0)$ and $\partial^2 C_{\alpha\beta}^b(\mathbf{K}, \mathbf{K}', t) / \partial t^2|_{t=0}$. Such a calculation can be carried out explicitly using Eq. (3.9), and our assumption that the system is isotropic on average, making the computation of ensemble averages tractable. We refer to §D of Fouvry et al. (2019b) for details on this calculation. One obtains

$$C_{\alpha\beta}^b(\mathbf{K}, \mathbf{K}', t - t') = \delta_{\alpha\beta} \delta_D(\mathbf{K} - \mathbf{K}') C_{\ell_\alpha}^b(\mathbf{K}, t) \quad \text{with} \quad C_\ell^b(\mathbf{K}, t) = n(\mathbf{K}) e^{-\frac{A_\ell}{2}(t/T_c(\mathbf{K}))^2}. \quad (3.11)$$

with the coefficient $A_\ell = \ell(\ell + 1)$. Here, $n(\mathbf{K})$ is the DF of the stars' parameters, normalised to $\int d\hat{\mathbf{L}} d\mathbf{K} n(\mathbf{K}) = N$.

Equation (3.11) fully characterises the statistical properties of the potential fluctuations generated by the background particles. We note that the amplitude of the temporal correlation is proportional to the background stellar density, $n(\mathbf{K})$. Because the system is isotropic on average, these correlations are diagonal when expanded in spherical harmonics, and only depend on the index ℓ . Finally, Eq. (3.11) also involves, $T_c(\mathbf{K})$, the coherence time of the noise that reads

$$\frac{1}{T_c^2(\mathbf{K})} = \int d\mathbf{K}' n(\mathbf{K}') \sum_\ell B_\ell \mathcal{J}_\ell^2[\mathbf{K}, \mathbf{K}'], \quad (3.12)$$

with the constant coefficient $B_\ell = \ell(\ell + 1)(2\ell + 1)/(8\pi)$. The coherence time characterises the typical time one has to wait for the bath to reshuffle enough so that its fluctuations become statistically independent. In the left panel of Fig. 3.3, we compare the prediction from Eq. (3.11) with numerical measurements (see §C of Fouvry et al., 2019c, for details).

3.1.3 Random walk

We have just characterised the noise fluctuations resulting from the coupled motions of N particles. Assuming that its statistics is given by the correlation function from Eq. (3.11), we now investigate the stochastic dynamics of one given test particle embedded in that noisy environment. Such a random walk is illustrated in the left panel of Fig. 3.4.

Throughout this section, we place ourselves in the test particle limit, i.e. the test particle is of zero mass so that its dynamics is fully determined by the time-dependent density of the background particles, and we neglect any backreaction of the test particle onto the background. We denote the parameters of the test particle with \mathbf{K}_t and its instantaneous orientation with $\hat{\mathbf{L}}_t$. Similarly to Eq. (3.6), we introduce the single particle DF

$$F^t(\hat{\mathbf{L}}, t) = \delta_D(\hat{\mathbf{L}} - \hat{\mathbf{L}}_t(t)), \quad (3.13)$$

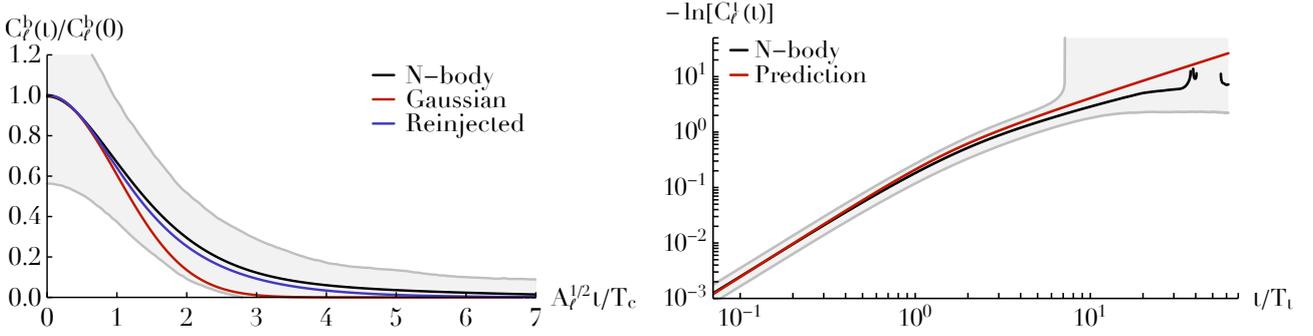


Figure 3.3: From Fouvry et al. (2019c). Left: Correlation of the $\ell=2$ noise fluctuations (Eq. 3.10). The black line is the numerical measurement with the gray lines illustrating the 10% and 90% spreads over the available realisations. The red line is the Gaussian prediction from Eq. (3.11), while the purple line is the updated prediction using the self-consistency relation from Eq. (3.27). Right: Correlated random walk of test particles, $C_l^t(t)$ (Eq. 3.18). This figure illustrates the two successive regimes of diffusion, namely ballistic ($\propto t^2$ for $t \ll T_t$) followed by diffusive ($\propto t$ for $t \gg T_t$). The remaining mismatch stems from non-ergodic effects, see §F in Fouvry et al. (2019c).

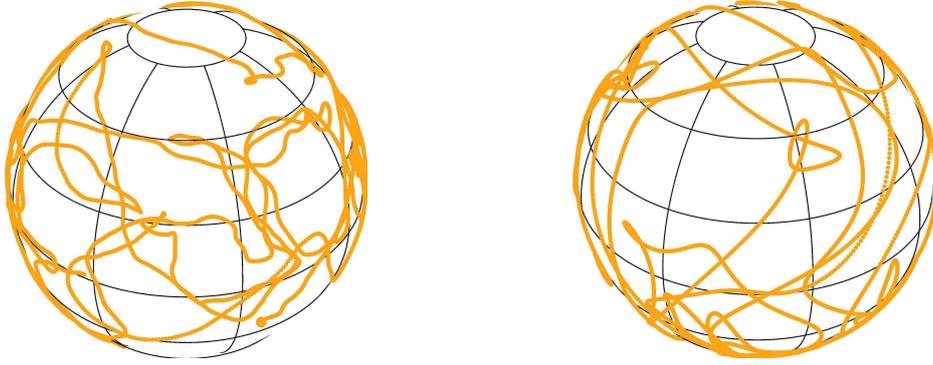


Figure 3.4: From Fouvry et al. (2019c). Left: Typical correlated random walk in orientation undergone by a test particle during VRR in self-consistent N -body simulations. Right: Random walk generated from the effective Langevin Eq. (3.22), for a test particle with the same parameters as in the left panel.

which is expanded in spherical harmonics via

$$F^t(\widehat{\mathbf{L}}, t) = \sum_{\alpha} Y_{\alpha}(\widehat{\mathbf{L}}) F_{\alpha}^t(t) \quad \text{with} \quad F_{\alpha}^t(t) = Y_{\alpha}(\widehat{\mathbf{L}}_t(t)). \quad (3.14)$$

In practice, since the spherical harmonics $\ell=1$ are such that $F_{(1,m)}^t(t) \sim \widehat{\mathbf{L}}_t(t)$, characterising the random walk of the test particle requires the knowledge of $F_{\alpha}^t(t)$ for $\ell_{\alpha}=1$.

The evolution equation for $F_{\alpha}^t(t)$ immediately follows from Eq. (3.9), and reads

$$\frac{\partial F_{\alpha}^t(t)}{\partial t} = - \sum_{\delta} Q_{\alpha\delta}^t(t) F_{\delta}^t(t) \quad \text{with} \quad Q_{\alpha\delta}^t(t) = \sum_{\gamma} \int d\mathbf{K} Q_{\alpha\gamma\delta}[\mathbf{K}_t, \mathbf{K}] F_{\gamma}^b(\mathbf{K}, t). \quad (3.15)$$

Here, $F_{\gamma}^b(\mathbf{K}, t)$ are the background magnetisations, as defined in Eq. (3.8), while $Q_{\alpha\delta}^t(t)$ is a time-dependent forcing driving the test particle's dynamics.

Equation (3.15) takes the form of a time-dependent linear matrix differential equation for the test particle's magnetisations, $F_{\alpha}^t(t)$. In order to guarantee well-behaved asymptotics for large times, we approach the resolution of Eq. (3.15) using Magnus series (see Blanes et al., 2009, for a review). In that framework, one generically solves for the motion of the test particle as

$$F_{\alpha}^t(t) = \sum_{\delta} [e^{\Omega(t)}]_{\alpha\delta} F_{\delta}^t(0), \quad (3.16)$$

where the matrix $\Omega(t)$ is constructed from the series expansion $\Omega(t) = \sum_{k \geq 1} \Omega^{(k)}(t)$, whose first terms are

$$\Omega^{(1)}(t) = - \int_0^t ds \mathbf{Q}^t(s), \quad (3.17a)$$

$$\Omega^{(2)}(t) = \frac{1}{2} \int_0^t ds_1 \int_0^{s_1} ds_2 [\mathbf{Q}^t(s_1), \mathbf{Q}^t(s_2)], \quad (3.17b)$$

with $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$ the matrix commutator.

Similarly to Eq. (3.10), we are interested in computing the correlation

$$C_{\alpha\beta}^t(t-t') = \langle F_\alpha^t(t) F_\beta^t(t') \rangle. \quad (3.18)$$

Using the Magnus series from Eq. (3.16), we can rewrite this correlation function as

$$C_{\alpha\beta}^t(t) = \sum_{\delta} \langle [e^{\Omega(t)}]_{\alpha\delta} \rangle \langle F_\delta^t(0) F_\beta^t(0) \rangle. \quad (3.19)$$

To obtain Eq. (3.19), we used two key assumptions. (i) Within the test particle limit, the independence hypothesis (Corrsin, 1959) allows us to separate the average over the test particle's initial condition and the background particles. (ii) We assumed that the noise is stationary in time, hence fixing our time origin in $t'=0$.

Assuming that the test particle is initially uniformly distributed on the unit sphere, one naturally has $\langle F_\delta^t(0) F_\beta^t(0) \rangle = \delta_{\delta\beta}/(4\pi)$. To compute the remaining average in Eq. (3.19), we restrict ourselves to terms that are at most second order in \mathbf{Q}^t . In practice, this is done by truncating the Magnus series in Eq. (3.17) at second order in the noise, and using the cumulant theorem to compute ensemble averages. We refer to SE in Fouvry et al. (2019c) for details and only give here the final result. One gets $C_{\alpha\beta}^t(t) = \delta_{\alpha\beta} C_{\ell_\alpha}^t(t)$ with

$$\begin{aligned} C_\ell^t(t) &= \frac{1}{4\pi} \exp \left\{ -\frac{1}{2} A_{\ell_\alpha} \sum_{\ell} B_\ell \int d\mathbf{K} \mathcal{J}_\ell^2[\mathbf{K}_t, \mathbf{K}] \int_0^t dt_1 \int_0^t dt_2 C_\ell^b(\mathbf{K}, t_1 - t_2) \right\} \\ &= \frac{1}{4\pi} \exp \left\{ -\frac{1}{2} A_{\ell_\alpha} \sum_{\ell} B_\ell \int d\mathbf{K} n(\mathbf{K}) \mathcal{J}_\ell^2[\mathbf{K}_t, \mathbf{K}] \frac{2T_c^2(\mathbf{K})}{A_\ell} \chi[\sqrt{A_\ell/2}(t/T_c(\mathbf{K}))] \right\}. \end{aligned} \quad (3.20)$$

To get the second line, we replaced the bath correlation, $C_\ell^b(\mathbf{K}, t)$, with the Gaussian ansatz from Eq. (3.11) and introduced the dimensionless function

$$\chi(\tau) = \int_0^\tau d\tau_1 \int_0^\tau d\tau_2 e^{-(\tau_1 - \tau_2)^2} = e^{-\tau^2} - 1 + \sqrt{\pi} \tau \operatorname{erf}(\tau) \simeq \begin{cases} \tau^2 & \tau \ll 1, \\ \sqrt{\pi} \tau & \tau \gg 1. \end{cases} \quad (3.21)$$

In Eq. (3.20), it is interesting to note that the function χ captures the expected two regimes of relaxation. On short timescales, $t \ll T_c$, the correlation C_ℓ^t decays like a Gaussian and the motion of the test particle is ballistic, i.e. $(\Delta \hat{\mathbf{L}})^2 \propto t^2$. On such short timescales, the test particle's random walk is analogous to the one induced by a time-independent fluctuation. On long timescales, $t \gg T_c$, the correlation decays exponentially and the motion of the test particle is diffusive, i.e. $(\Delta \hat{\mathbf{L}})^2 \propto t$. On these timescales, the test particle's random walk is analogous to the one induced by fluctuations that are δ_D -correlated time, as in classical Brownian motion. In the right panel of Fig. 3.3, we compare Eq. (3.20) with numerical measurements (see Fouvry et al., 2019b, for details).

Because it involves a sum over the harmonics ℓ and an integral over \mathbf{K} , Eq. (3.20) remains difficult to implement. We conclude this section by presenting a simpler toy model that generates stochastic motions with correlation properties similar to the ones of Eq. (3.20). We assume that the motion of the test particle is generated by an effective dipolar ($\ell=1$) Gaussian noise. More precisely, it follows the Langevin equation

$$\frac{d\hat{\mathbf{L}}_t}{dt} = \Gamma_t \boldsymbol{\eta}(t) \times \hat{\mathbf{L}}_t \quad \text{with} \quad \begin{cases} \langle \eta_i(t) \rangle = 0, \\ \langle \eta_i(t) \eta_j(t') \rangle = \delta_{ij} e^{[(t-t')/T_t]^2}. \end{cases} \quad (3.22)$$

with $\boldsymbol{\eta}(t)$ a 3D Gaussian noise. Our goal is now to tailor the amplitude, Γ_t , and coherence time, T_t , to match the behaviour of the test particle correlation from Eq. (3.20) on short and long timescales (Hamers et al., 2018).

Following the same steps as in Eq. (3.16), one can easily compute the correlation function, $C_{\alpha\beta}^{\text{toy}}(t)$, of a test particle whose dynamics is imposed by Eq. (3.22). One finds $C_{\alpha\beta}^{\text{toy}}(t) = \delta_{\alpha\beta} C_\ell^{\text{toy}}(t)$ with

$$C_\ell^{\text{toy}}(t) = \frac{1}{4\pi} \exp \left\{ -\frac{1}{2} A_{\ell_\alpha} \Gamma_t^2 T_t^2 \chi[t/T_t] \right\}. \quad (3.23)$$

By matching the ballistic and diffusive regimes of Eq. (3.23) with the ones of Eq. (3.20), we get the relations

$$\Gamma_t^2 = \frac{1}{T_c^2(\mathbf{K})}, \quad (3.24a)$$

$$T_t^2 = \frac{1}{\Gamma_t^2} \sum_{\ell} \frac{B_\ell}{\sqrt{A_\ell/2}} \int d\mathbf{K} n(\mathbf{K}) \mathcal{J}_\ell^2[\mathbf{K}_t, \mathbf{K}] T_c(\mathbf{K}). \quad (3.24b)$$

where the bath coherence time, $T_c(\mathbf{K})$, was already introduced in Eq. (3.12).

Equation (3.23) is an important result, as it offers a simple analytical description of the statistical properties of the random walk of a test particle's orientation, jointly induced by all the background particles. The test particle's random walk is characterised by the two quantities (Γ_t, T_t) that both depend on the test particle's parameters, \mathbf{K}_t . On the one hand, the torque amplitude, Γ_t , controls the amplitude of the test particle's initial ballistic motion. As such, the torque time, $1/\Gamma_t$, represents the typical time it would take the test particle to explore the unit sphere for a given and frozen value of the background noise. On the other hand, the coherence time, T_t , describes the typical time one has to wait for the torque amplitude to reach a statistically different value, i.e. to decorrelate itself. It is therefore the timescale after which the test particle leaves the ballistic regime and enters the diffusive one. These two regimes are visible in the right panel of Fig. 3.3.

One strength of the present formalism is that, following Eq. (3.24), one now has at one's disposal explicit expressions for these two parameters. These coefficients are time-independent and can easily be computed for various cluster models (by varying $n(\mathbf{K})$) and test particles (by varying \mathbf{K}_t). In addition, because the time correlation of the Gaussian noise, $\boldsymbol{\eta}(t)$, is explicitly known, it is straightforward to generate noise realisations (Romero & Sancho, 1999). These time series can then be used in Eq. (3.22) to generate effective random walks. This is illustrated in the right panel of Fig. 3.4.

3.1.4 Noise self-consistency

In the previous sections, we proceeded in two successive steps. First, in §3.1.2, we used estimates of the derivatives of the correlation function of the noise at the initial time to obtain Eq. (3.11), a simple ansatz for the time-dependence of the correlation function of the noise generated by N background particles. Then, in §3.1.3, we used this noise as a source term driving the stochastic dynamics of a test particle, leading to Eq. (3.20). Yet, if the test particle is now taken to be one background particle, its orientation random walk and the background fluctuations sourcing it have to be self-consistent. This is what we now explore.

We start from Eq. (3.10) and replace $F_\alpha^b(\mathbf{K}, t)$ by its discrete definition from Eq. (3.6). One gets

$$C_{\alpha\beta}^b(\mathbf{K}, \mathbf{K}', t-t') = \sum_{i,j} \left\langle \delta_D(\mathbf{K}-\mathbf{K}_i) \delta_D(\mathbf{K}'-\mathbf{K}_j) Y_\alpha(\widehat{\mathbf{L}}_i(t)) Y_\beta(\widehat{\mathbf{L}}_j(t)) \right\rangle. \quad (3.25)$$

We now assume that each background particle can be treated as a test particle, and that their long-term motions are decorrelated from one another. Only the contributions from $i=j$ remain, and, paying a careful attention to normalisation, Eq. (3.25) becomes

$$C_{\alpha\beta}^b(\mathbf{K}, \mathbf{K}', t) = \delta_D(\mathbf{K}-\mathbf{K}') 4\pi n(\mathbf{K}) C_{\alpha\beta}^t(\mathbf{K}, t), \quad (3.26)$$

where $C_{\alpha\beta}^t(\mathbf{K}, t)$ is the test particle's correlation defined in Eq. (3.18). Luckily, we have already solved for this correlation function for an arbitrary background perturbation in Eq. (3.20). Reinjecting into Eq. (3.26), we get

$$C_\ell^b(\mathbf{K}, t) = n(\mathbf{K}) \exp \left\{ -\frac{1}{2} A_\ell \sum_{\ell'} B_{\ell'} \int d\mathbf{K}' \mathcal{J}_{\ell'}^2[\mathbf{K}, \mathbf{K}'] \int_0^t dt_1 \int_0^t dt_2 C_{\ell'}^b(\mathbf{K}', t_1-t_2) \right\}, \quad (3.27)$$

where $C_\ell^b(\mathbf{K}, t)$ was introduced in Eq. (3.11). Equation (3.27) takes the form of a self-consistent integral equation satisfied by the correlation function, $C_\ell^b(\mathbf{K}, t)$.

Following Taylor & McNamara (1971), this self-consistency can be further clarified by defining

$$R_\ell^b(\mathbf{K}, t) = \frac{1}{2} \int_0^t dt_1 \int_0^t dt_2 C_\ell^b(\mathbf{K}, t_1-t_2). \quad (3.28)$$

from which we get

$$\frac{d^2 R_\ell^b(\mathbf{K}, t)}{dt^2} = n(\mathbf{K}) \exp \left\{ -A_\ell \sum_{\ell'} B_{\ell'} \int d\mathbf{K}' \mathcal{J}_{\ell'}^2[\mathbf{K}, \mathbf{K}'] R_{\ell'}^b(\mathbf{K}', t) \right\}. \quad (3.29)$$

Equation (3.29) is the important result of this section. It highlights the intrinsic non-linearity of the VRR problem and its mandatory self-consistency. Yet, as it couples both different harmonics (via $\sum_{\ell'}$) and different orbital parameters (via $\int d\mathbf{K}'$), such a differential equation is too intricate to be easily solved explicitly.

Fortunately, one can still proceed iteratively to obtain improved approximations of the noise correlation function. To do so, one starts from the Gaussian dependence obtained in Eq. (3.11). This (motivated) ansatz is then reinjected in the r.h.s. of Eq. (3.27) to obtain a new approximation of the noise correlation function that has both a ballistic and diffusive part. Such a procedure is illustrated in Fig. 3.3, where we show how one can better match the late-time properties of the system's noise through this iterative process.

3.1.5 Prospects

In this section, we illustrated how one can describe self-consistently the statistical properties of the stochastic evolution of a star's orientation in a galactic nucleus during VRR. The main difficulty of the present derivation lies in the system being fundamentally degenerate, i.e. having a vanishing mean-field Hamiltonian, $\langle H_{\text{VRR}} \rangle = 0$ (§3.1.1). As a result, the system is non-Markovian, i.e. correlated in time and orientation.

Placing ourselves in the limit of an isotropic distribution of orientations, we circumvented some of these difficulties in §3.1.2 by characterising the statistical properties of the potential fluctuations present in the cluster. In §3.1.3, we used this result to describe the random walk of a test particle embedded in this stochastic system, recovering both the ballistic and diffusive regimes. In §3.1.4, we highlighted the fundamental self-consistency that exists between the spontaneous fluctuations in the system and the associated random walk in orientations.

Of course, the present statistical description is only a first step towards a complete theory of VRR, and we conclude this section by listing below some possible tracks for future developments.

- In §3.1.3, we described the dynamics of one test star's orientation. In order to investigate how a substructure like a disc slowly dissolves (Kocsis & Tremaine, 2011), the next logical step is to describe the joint dynamics of a pair of nearby test stars. This is the focus of §3.2.
- If one were to replace the test star by a binary, then the exact statistical properties of the VRR random walk in orientation can lead to enhanced rates of mergers for the binary itself (Hamers et al., 2018). As such, ever more reliable and precise description of the VRR dynamics should allow for more accurate prescriptions on the rate of binary mergers expected to originate from galactic nuclei.
- Here, we relied extensively on the isotropic assumption, and as such neglected any effects associated with anisotropic clustering in orientation (see, e.g., Szölgvény et al., 2021). Predicting such non-trivial equilibrium distributions is the topic of §3.3.
- Finally, here we focused our interest on galactic nuclei, i.e. systems dominated by a central mass. Provided that one updates accordingly the coupling coefficients, $\mathcal{J}_\ell[\mathbf{K}, \mathbf{K}']$, similar investigations can naturally be pursued in the context of spherical globular clusters (see, e.g., Meiron & Kocsis, 2018; Roupas, 2020).

3.2 Neighbour separation

In the previous section, we characterised the correlated random walks undergone by a test particle's orientation when embedded in an isotropic bath of background particles. Rather than focusing on a single test particle, let us now tackle the dual problem of embedding within the same bath two test particles sharing similar orientations and orbital parameters. Our goal is to obtain precise estimates of the efficiency with which these two particles diffuse away from one another. We call this process "neighbour separation". This process is the one that can drive the spontaneous dilution of stellar discs in galactic nuclei (see, e.g., Kocsis & Tremaine, 2011), as illustrated in Fig. 3.5. Ultimately, such a dilution should prove paramount to constrain SgrA*'s dynamical

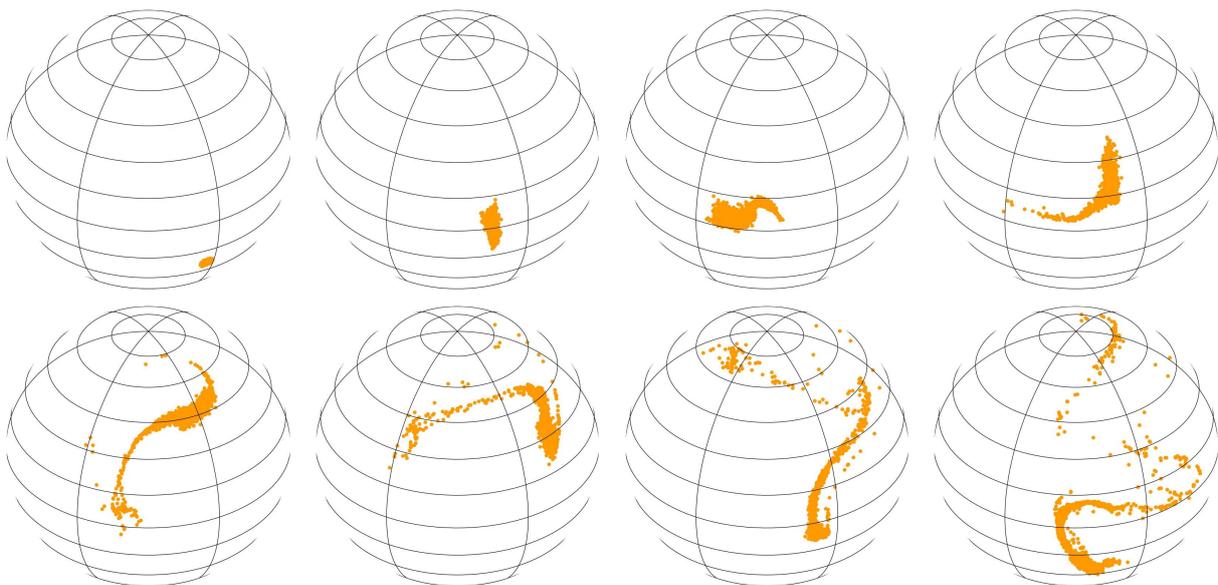


Figure 3.5: From Giral Martínez et al. (2020). Time sequence (from left to right and top to bottom) of the spontaneous dilution of a population of neighbouring test particles differing initially both in orientations and orbital parameters. Characterising the efficiency of "neighbour separation" is mandatory to constrain the survivability of SgrA*'s observed clockwise stellar disc.

content from its observed clockwise stellar disc (Bartko et al., 2009; Yelda et al., 2014; Gillessen et al., 2017).

We are interested in the efficiency with which two nearby particles diffuse away from one another, as illustrated in the left panel of Fig. 3.6. Compared with §3.1.3, the main difficulty now stems from the fact that these

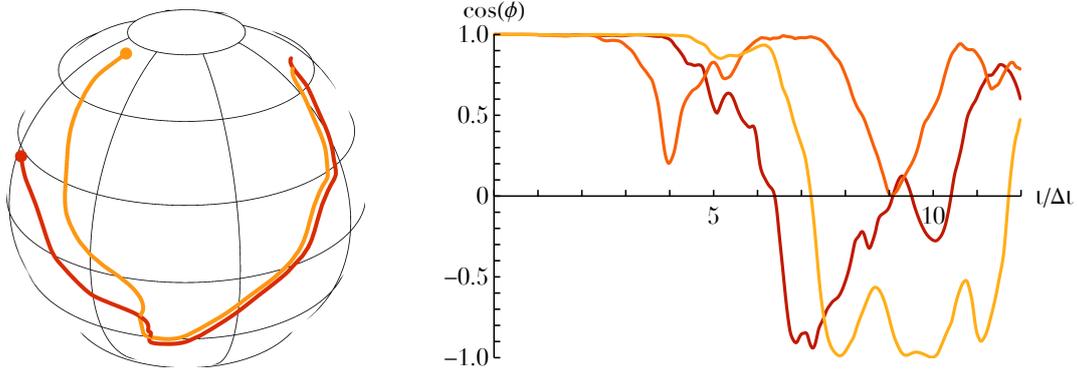


Figure 3.6: From Giral Martínez et al. (2020). Left: Separation between two neighbour test particles in a N -body simulation. Their paths follow each other closely initially but eventually diverge away from one other. Right: Typical random walks of the angle $\cos(\phi)$ between pairs of test particles (with the same orbital parameters) in N -body simulations. The orientations typically decorrelate after a few Δt , as defined in Eq. (3.40).

two neighbour particles, even though of zero mass, are not independent as they evolve within the same background noise. As a result, in contrast to Eq. (3.18), we are not interested in the correlation of the harmonics at different times for the same particle, but rather in the correlation of the harmonics of two different particles at the same time. We therefore define the "neighbour" correlation function

$$C_{\alpha\beta}^n(t) = \langle F_{\alpha}^1(t) F_{\beta}^2(t) \rangle, \quad (3.30)$$

where, following Eq. (3.14), $F_{\alpha}^i(t)$ ($i=1, 2$) is the instantaneous magnetisation of particle i . For an isotropic background noise and isotropic initial conditions, the addition theorem for spherical harmonics offers a simple interpretation of Eq. (3.30). Indeed, one has

$$C_{\alpha\beta}^n(t) = \frac{\delta_{\alpha\beta}}{4\pi} \langle P_{\ell_{\alpha}}(\cos(\phi(t))) \rangle, \quad (3.31)$$

with P_{ℓ} the Legendre polynomials and $\phi(t)$ the instantaneous angle between the two particles. Since $P_1(x) = x$, the case $\ell_{\alpha} = 1$ is of prime importance as it directly gives the angular separation between the two test particles. Higher order harmonics are connected to higher order moments of $\cos(\phi)$. Our goal here is to characterise the random walk undergone by $\phi(t)$, as illustrated in the right panel of Fig. 3.6. In Eq. (3.30), we emphasise that the test particle limit is assumed, so that any self-gravity among the test particles is neglected. This is an important limitation of the upcoming calculation, as self-gravity within a stellar disc (see figs. 6 and 7 in Kocsis & Tremaine, 2011) enhances the disc's coherence and reduces the efficiency with which it can dissolve. Accounting for self-gravity is, surely, no easy task, and deserves further investigations.

In order to compute the correlation from Eq. (3.30), we follow the same approach as in Eq. (3.16). We write

$$C_{\alpha\beta}^n(t) = \sum_{\gamma, \delta} \langle [e^{\Omega_1(t)}]_{\alpha\gamma} [e^{\Omega_2(t)}]_{\beta\delta} \rangle \langle F_{\gamma}^1(0) F_{\delta}^2(0) \rangle. \quad (3.32)$$

In that expression, we followed Eq. (3.17) and introduced $\Omega_i(t)$ the Magnus matrix specific to particle i . We also enforced, once again, that the statistics of the background noise is independent of the initial locations of the test particles. This allows us to separate the averages in Eq. (3.32).

We assume that the initial conditions for the test particles are isotropic: while the location of the two test particles are correlated with one another, the distribution of any given test particle is statistically uniform on the unit sphere. The second average from Eq. (3.32) can then generically be written as

$$\langle F_{\gamma}^1(0) F_{\delta}^2(0) \rangle = \frac{\delta_{\alpha\beta}}{4\pi} D_{\gamma}, \quad (3.33)$$

where, by isotropy, the diagonal matrix \mathbf{D} sees its entries D_{γ} depend only on ℓ_{γ} . In the case where the test particles are launched with a fixed initial angular separation, ϕ_0 , one finds $D_{\ell} = P_{\ell}(\cos(\phi_0))$. This further simplifies to $\mathbf{D} = \mathbf{I}$, when the two test particles are systematically launched with the exact same initial orientations.

Having characterised the test particles' initial conditions, we now go back to the first term from Eq. (3.32). Owing to symmetries, we write

$$C_{\alpha\beta}^n(t) = \langle [e^{\Omega_1(t)} \mathbf{D} e^{-\Omega_2(t)}]_{\alpha\beta} \rangle. \quad (3.34)$$

At this stage, we emphasise the two origins of the separation undergone by the test particles: (i) their initial angular separation, as captured by \mathbf{D} ; (ii) their difference in orbital parameters, as captured by $\Omega_1 \neq \Omega_2$.

The main difficulty in the computation of the ensemble average from Eq. (3.34) comes from the non-commutativity of the matrix exponential. Similarly to Eq. (3.19), we perform a truncation at second-order in the bath fluctuations. Following $\mathcal{S}\mathcal{D}$ of Giral Martínez et al. (2020), one ultimately gets

$$C_{\alpha\beta}^m(t) = \delta_{\alpha\beta} D_{\ell_\alpha} C_{\ell_\alpha}^+(t) C_{\ell_\alpha}^-(t). \quad (3.35)$$

In that expression, $C_{\ell}^+(t)$ (resp. $C_{\ell}^-(t)$) captures the separation of the two test particles sourced by their differences in initial orientations (resp. in conserved parameters). They read

$$C_{\ell}^+(t) = \exp \left\{ - \sum_{\ell, \ell_\gamma} \frac{D_{\ell_\alpha} - D_{\ell_\gamma}}{(2\ell_\alpha + 1)D_{\ell_\alpha}} (E_{\ell_\alpha \ell \ell_\gamma}^L)^2 \int d\mathbf{K} n(\mathbf{K}) \mathcal{J}_\ell[\mathbf{K}_1, \mathbf{K}] \mathcal{J}_\ell[\mathbf{K}_2, \mathbf{K}] \frac{2T_c^2(\mathbf{K})}{A_\ell} \chi[\sqrt{A_\ell/2}(t/T_c(\mathbf{K}))] \right\}, \quad (3.36a)$$

$$C_{\ell}^-(t) = \exp \left\{ - \frac{1}{2} A_{\ell_\alpha} \sum_{\ell} B_\ell \int d\mathbf{K} n(\mathbf{K}) (\mathcal{J}_\ell[\mathbf{K}_1, \mathbf{K}] - \mathcal{J}_\ell[\mathbf{K}_2, \mathbf{K}])^2 \frac{2T_c^2(\mathbf{K})}{A_\ell} \chi[\sqrt{A_\ell/2}(t/T_c(\mathbf{K}))] \right\}, \quad (3.36b)$$

with $E_{\ell_\alpha \ell \ell_\gamma}^L$ the "isotropic" Elsasser coefficients (see $\mathcal{S}\mathcal{B}$ in Fouvry et al., 2019c). Writing Eq. (3.36) as exponentials guarantees a physically admissible behaviour at late times, where correlations must tend to zero.

These generic expressions become clearer for small angular separations. If the two test particles are initially separated by a small angle ϕ_0 , one has $D_\ell = P_\ell(\cos(\phi_0))$, and at first order in $(1 - \cos(\phi_0))$, Eq. (3.36) becomes

$$C_{\ell}^+(t) = \exp \left\{ - \frac{1}{2} A_\ell (1 - \cos(\phi_0)) \Psi^+[\mathbf{K}_1, \mathbf{K}_2, t] \right\}; \quad C_{\ell}^-(t) = \exp \left\{ - \frac{1}{2} A_\ell \Psi^-[\mathbf{K}_1, \mathbf{K}_2, t] \right\}. \quad (3.37)$$

The function Ψ^+ (resp. Ψ^-) captures the effects from orientation (resp. parameter) separation. They read

$$\Psi^+[\mathbf{K}_1, \mathbf{K}_2, t] = \sum_{\ell} B_\ell (A_\ell - 2) \int d\mathbf{K} n(\mathbf{K}) \mathcal{J}_\ell[\mathbf{K}_1, \mathbf{K}] \mathcal{J}_\ell[\mathbf{K}_2, \mathbf{K}] \frac{2T_c^2(\mathbf{K})}{A_\ell} \chi[\sqrt{A_\ell/2}(t/T_c(\mathbf{K}))], \quad (3.38a)$$

$$\Psi^-[\mathbf{K}_1, \mathbf{K}_2, t] = \sum_{\ell} B_\ell \int d\mathbf{K} n(\mathbf{K}) (\mathcal{J}_\ell[\mathbf{K}_1, \mathbf{K}] - \mathcal{J}_\ell[\mathbf{K}_2, \mathbf{K}])^2 \frac{2T_c^2(\mathbf{K})}{A_\ell} \chi[\sqrt{A_\ell/2}(t/T_c(\mathbf{K}))]. \quad (3.38b)$$

The prediction from Eq. (3.37) is self-similar in the sense that the only dependence on the considered harmonics ℓ is carried by A_ℓ , which is factored in the exponent. Equation (3.38) also offers a simple interpretation of the dynamical mechanisms driving the separation of neighbours, as we now discuss.

A first source of separation originates from any misalignment in the test particles' orientations, as captured by Ψ^+ . This term does not vanish even when the two test particles have the same orbital parameters. The smaller the initial separation, i.e. the smaller $(1 - \cos(\phi_0))$, the longer it takes neighbours to separate. In particular, C_{ℓ}^+ becomes unity when $\cos(\phi_0) \rightarrow 1$, i.e. when the two test particles share the exact same initial orientation. We note that Ψ^+ also involves an extra $(A_\ell - 2)$ because separation are sensitive to tides, not forces.

On top of this first effect, a second source of separation stems from the difference in the test particles' parameters, as captured by Ψ^- . This transpires in the coupling factor $(\mathcal{J}_\ell[\mathbf{K}_1, \mathbf{K}] - \mathcal{J}_\ell[\mathbf{K}_2, \mathbf{K}])^2$ which highlights that the test particles couple to the bath differently. The closer \mathbf{K}_1 and \mathbf{K}_2 , the slower the induced separation, with C_{ℓ}^- becoming unity when $\mathbf{K}_1 = \mathbf{K}_2$, i.e. when the underlying annuli of the two test particles are identical.

We emphasise once again the generality of Eq. (3.35) that captures jointly three contributions to neighbour separation: (i) the orbital distribution of background particles, via $n(\mathbf{K})$; (ii) the orbital differences of the two test particles, via \mathbf{K}_1 and \mathbf{K}_2 ; (iii) the difference in their initial orientation, via \mathbf{D} . In all these expressions, time is measured in units of the coherence time $T_c(\mathbf{K})$ (Eq. 3.12), which sets the typical timescale for that process.

We may now test the prediction of Eq. (3.35) against tailored numerical simulations (see $\mathcal{S}\mathcal{F}$ in Giral Martínez et al., 2020, for the setup) using the method from $\mathcal{S}\mathcal{3}$. Our main result is illustrated in the left panel of Fig. 3.7. Because of the very slow separation of nearby particles, Eq. (3.35) does not provide a good match to the late-time decorrelation of test particles with initially very similar orientations. In that same figure, we also illustrate that particles with different orbital parameters separate faster, as expected.

Despite this underwhelming comparison, we note however that on the coherence time, i.e. for $t \lesssim \Delta t$ (to be defined in Eq. 3.40), the analytical prediction matches the numerical measurements. This is expected because Eq. (3.35) stems from a second-order Taylor expansion around $t=0$. Yet, in its present form, Eq. (3.35) is unable to describe the separation of test particles sharing very similar orientations and orbital parameters, as this correlation follows a plateau that only decays on very long timescales. Fortunately, this failure of the analytical prediction can be alleviated by using an appropriate "piecewise" Markovian approach, as we now explore.

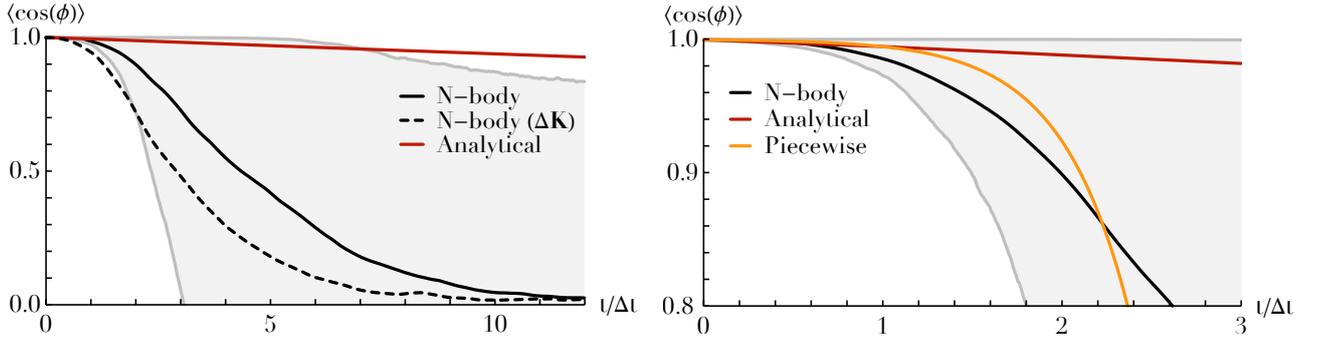


Figure 3.7: From Giral Martínez et al. (2020). Left: Average angular separation, $\langle \cos(\phi) \rangle$, between nearby test particles with the same orbital parameters. The "Analytical" prediction corresponds to the direct application of Eq. (3.35) while the curve " $\Delta\mathbf{K}$ " illustrates the separation in numerical simulations for test particles with different orbital parameters. Right: Zoom of the left panel, highlighting the "Piecewise" prediction from Eq. (3.45).

3.2.1 Piecewise Markovian prediction

Regardless of the initial similarity between the test particles, Eq. (3.35) works well on short timescales, that it when $t \lesssim T_c(\mathbf{K})$. As a result, a reasonable way to fix the late-time behaviour of our prediction is to construct a sequence of short-time predictions, following for each of them Eq. (3.35). Let us therefore pick a timelapse Δt (whose precise value is specified later on) and construct a piecewise prediction of $\langle \cos(\phi) \rangle$ in timelapses of duration Δt . Doing so, we therefore construct a sequence

$$\cos(\phi_0) \rightarrow \dots \rightarrow \cos(\phi_n) = \cos(\phi(t=n\Delta t)) \quad (3.39)$$

of angular separation between the two test particles. In order to construct the prediction for a given timelapse, say $n\Delta t \rightarrow (n+1)\Delta t$, we follow Eq. (3.35) using $\cos(\phi_n)$ as the initial angular separation between the two test particles and $t = \Delta t$ as the duration for which Eq. (3.35) is pushed forward in time.

We emphasise that such a protocol is not equivalent to making a single prediction for the whole time series using only $\cos(\phi_0)$ as the initial separation in Eq. (3.35). Because the current angular separation, $\cos(\phi_n)$, is now used in Eq. (3.35) as an initial condition for the n -th timelapse, the piecewise approach neglects any correlation that exists in the background noise between various timelapses. Indeed, in Eq. (3.32), we assumed that the initial conditions of the test particles are independent from the background bath. In the present piecewise approach, we neglect any such correlations. This is our Markovian assumption, a key ingredient.

In a sense, the piecewise approach relies on ideas similar to explicit integration methods, such as Euler method. Rather than approximating the solution with a single perturbative expansion around its initial conditions, we construct the solution step by step. The key difference here is that Eq. (3.15) is stochastic, i.e. the noise driving the separation is time-correlated. By proceeding with successive timelapses, we unavoidably neglect part of this correlation. If one were to take Δt arbitrarily small, the VRR fluctuations would be treated as uncorrelated in time. This drastically differs from the large-scale coherent flows imposed by VRR.

The choices for Δt are therefore limited. On the one hand, to capture most of the noise correlation, one needs $\Delta t \gtrsim T_c$. On the other hand, in order for each timelapse to be accurately predicted, one needs to take Δt as small as possible. Given these two constraints, a natural choice is to take Δt of the order of T_c . For a given pair of particles, we therefore define the separation timestep as

$$\Delta t = \text{Max}[T_c(\mathbf{K}_1), T_c(\mathbf{K}_2)]. \quad (3.40)$$

Having decided upon a timelapse Δt , we can now construct typical sequences of angular separation as in Eq. (3.39). This requires us to estimate the statistics of the transition $\cos(\phi_n) \rightarrow \cos(\phi_{n+1})$. Building upon our Markovian assumption, we write

$$\langle \cos(\phi_{n+1}) \rangle = \int d(\cos(\phi_n)) \rho_n(\cos(\phi_n)) \langle \cos(\phi_{n+1}) | \cos(\phi_n) \rangle. \quad (3.41)$$

In that expression, $\langle \cos(\phi_{n+1}) | \cos(\phi_n) \rangle$ follows from Eq. (3.35) evaluated after a time $t = \Delta t$ for two test particles exactly separated by the angle $\cos(\phi_n)$ at the start of the timelapse. In Eq. (3.41), we also introduced ρ_n as the PDF of $\cos(\phi_n)$. In practice, this PDF is not known, so that without further approximation, the integral from Eq. (3.41) cannot be explicitly computed.

The simplest way around this difficulty is to rely on a first-order development of $\langle \cos(\phi_{n+1}) | \cos(\phi_n) \rangle$ near $\cos(\phi_n) \simeq 1$, i.e. in the limit of small angular separations. This is a calculation that we already started in Eq. (3.37). As detailed in §G of Giral Martínez et al. (2020), one obtains the linear relation

$$\langle \cos(\phi_{n+1}) | \cos(\phi_n) \rangle = \xi_0 + \xi_1 \cos(\phi_n). \quad (3.42)$$

The constant coefficients, ξ_k , follow the simple expression

$$\begin{cases} \xi_0 = -e^{-\psi^-} \psi^+, \\ \xi_1 = e^{-\psi^-} (1 + \psi^+), \end{cases} \quad \text{with} \quad \begin{cases} \psi^+ = \Psi^+[\mathbf{K}_1, \mathbf{K}_2, \Delta t], \\ \psi^- = \Psi^-[\mathbf{K}_1, \mathbf{K}_2, \Delta t], \end{cases} \quad (3.43)$$

following Eq. (3.38). We are now in a position to compute the average from Eq. (3.41) to obtain

$$\langle \cos(\phi_{n+1}) \rangle = \xi_0 + \xi_1 \langle \cos(\phi_n) \rangle. \quad (3.44)$$

This is the key relation of the piecewise approach, which offers an explicit relation between successive values of $\langle \cos(\phi_n) \rangle$. Equation (3.44) is an arithmetico-geometric relation, which can easily be solved. Given an initial condition $\langle \cos(\phi_0) \rangle$, the sequence of expectations for the subsequent timelapses, $t = n\Delta t$, $n \geq 0$, reads

$$\langle \cos(\phi_n) \rangle = q + (\xi_1)^n (\langle \cos(\phi_0) \rangle - q), \quad (3.45)$$

with $q = \xi_0 / (1 - \xi_1)$. While it depends on the discrete time index n , Eq. (3.45) can still be viewed as a continuous function of time, provided one makes the replacement $n \rightarrow t / \Delta t$.

In the right panel of Fig. 3.7, we compare the piecewise prediction from Eq. (3.45) against numerical measurements. While the match is not ideal, the piecewise prediction still manages to describe the slow initial separation of the neighbour particles, while the straightforward application of the analytical prediction from Eq. (3.35) failed at it. Because Eq. (3.45) was obtained through a perturbative expansion for small separations, it can only be confidently used for separations satisfying $\langle \cos(\phi) \rangle \gtrsim 0.8$. Fortunately, this corresponds to $\sim 37^\circ$, which is more than enough to investigate the dynamical dilution of SgrA*'s clockwise disc, whose typical angular separation is 16° (Gillessen et al., 2017).

3.2.2 Application

We conclude this section with a brief application of "neighbour separation" to a galactic nucleus mimicking SgrA*. We illustrate in particular how this process allows one to probe the underlying kinematic properties of the unresolved nucleus by requiring consistency with a given level of neighbour dilution.

Following Gillessen et al. (2017), we take the mass of the central BH to be $M_\bullet = 4.3 \times 10^6 M_\odot$, and its radius of influence to be 2 pc. For simplicity, we assume that the background old stellar cluster is made of a single-mass population of individual mass, $m_\star = 1 M_\odot$, with a thermal distribution of eccentricity, $P(e|a) = 2e$ and a power-law distribution for the semi-major axes, $f_a(a) \propto a^{2-\gamma}$. See SJ in Giral Martínez et al. (2020) for details.

In such a scale-free configuration, one can show that the bath's coherence time, $T_c(\mathbf{K})$ (Eq. 3.12), follows the simple dependence (see Eq. 33 in Giral Martínez et al., 2020)

$$T_c(a, e) \simeq 8.8 \times \frac{1}{\Omega_{\text{Kep}}(a)} \frac{1}{\sqrt{N(<a)}} \frac{M_\bullet}{\sqrt{\langle m^2 \rangle}} \sqrt{1-e^2}, \quad (3.46)$$

with $\Omega_{\text{Kep}}(a)$ the Keplerian frequency from Eq. (1.8), $N(<a) \propto a^{3-\gamma}$ the average number of stars within a sphere of radius a from the centre and $\sqrt{\langle m^2 \rangle}$ the cluster's mass spectrum. From Eq. (3.46), we recover the expected trends that the larger the spread in mass and the larger the eccentricity, the shorter the coherence time.

To mimic SgrA*'s clockwise stellar disc (see, e.g., Gillessen et al., 2017), we consider test stars that share the same orbital parameters, $a_t = 50$ mpc and $e_t = 0.1$. We crudely assume these stars are $T_\star = 10$ Myr old (Habibi et al., 2017). For simplicity, we do not account for any separation stemming from differences in the test particles' orbital parameters, which would further reduce the timescales predicted below. From Gillessen et al. (2017), we know that the current angular dispersion of the clockwise disc is approximately $\langle \phi_{\text{disc}} \rangle \simeq 16^\circ$, i.e. $\langle \cos(\phi_{\text{disc}}) \rangle \simeq 0.96$. This is within the range of applicability of the piecewise prediction (right panel of Fig. 3.7). For a given model and initial condition, we define the dilution time, T_{dil} , via the implicit relation

$$\langle \cos(\phi[t = T_{\text{dil}}]) \rangle = \langle \cos(\phi_{\text{disc}}) \rangle. \quad (3.47)$$

Once the average angular separation between the test stars has reached such a large value, we may consider that the stellar disc has effectively been dissolved by the VRR fluctuations. For a background model to be accepted, one must ensure that $T_{\text{dil}} \simeq T_\star$.

The left panel of Fig. 3.8 illustrates the variation of the dilution time as one varies the power index of the background stellar cusp, γ_\star , as well as the initial angular separation, $\langle \cos(\phi_0) \rangle$. Naturally, the smaller the initial separation, $\langle \cos(\phi_0) \rangle$, the longer it takes for the initially coherent disc to dissolve. The same figure also shows that the cuspiest the density profile, the faster the dilution. This is a direct consequence of Eq. (3.46): the larger γ_\star , the larger $N(<a)$, therefore the smaller the coherence time, $T_c(\mathbf{K})$, hence the faster the dilution. Of course, the numerical values used here are somewhat ad hoc and would definitely require more careful selections.

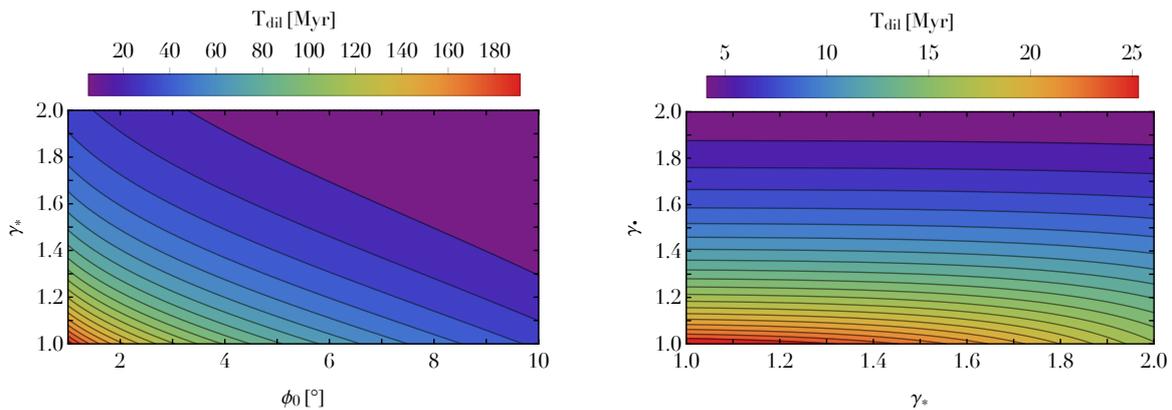


Figure 3.8: From Giral Martínez et al. (2020). Dependence of the diffusion time, T_{dil} , of a tentative clockwise stellar disc embedded in a background stellar cluster. Left: As a function of the initial angular separation ϕ_0 and the power-law index, γ_* , of the stellar population without any IMBHs. Right: As a function of the power index γ_* (resp. γ_\bullet) of the stellar (resp. IMBH) population.

Yet, Fig. 3.8 shows how the present formalism could be used to place constraints on the parameters of the unresolved background stellar cluster (here through its index γ_*) as well as on the formation channels of stars in galactic nuclei (here through the size of their initial angular dispersion, ϕ_0).

Let us finally consider the case where the background cluster is composed not only of old stars but also of IMBHs of individual mass $m_\bullet = 100 M_\odot$. We assume that the total enclosed mass within the sphere of influence remains the same with a total IMBH mass fraction of 20%. In practice, within the old unresolved stellar cluster, NR has had the time to lead to mass segregation so that one expects for the old stars and the IMBHs not to share the same power index (Bahcall & Wolf, 1977). This is briefly explored in the right panel of Fig. 3.8 where we illustrate the dependence of the dilution time as one varies independently the power-law indices of both populations. In that figure, we recover that cuspiest profiles lead to more efficient separations, the effect being most visible for the IMBHs because of the larger Poisson fluctuations they generate.

3.2.3 Prospects

In this section, we illustrated how to quantitatively describe the two-point statistical properties of the neighbour separation induced by VRR. We placed ourselves in the limit of an isotropic distribution of background stars, whose potential fluctuations slowly stir neighbour particles away from one another. A key result was obtained in Eq. (3.35) which highlighted the two main effects sourcing the separation of nearby particles, namely their difference in orientations and in conserved orbital parameters. Relying on a development for small angular separation and a Markovian assumption, this estimator was subsequently improved in Eq. (3.45) to construct a piecewise prediction for the rate of neighbour separation. Finally, in §3.2.2, for the sake of illustration, we presented a first application of this formalism to determine the ease with which young stellar discs, such as the one observed within SgrA*, can spontaneously dissolve under the effects of the stochastic VRR dynamics. In particular, we emphasised how the initial distribution of angular separation, the profile of the background unresolved cluster, as well as the possible presence of IMBHs all influence the efficiency of a disc’s dilution.

The present section has only addressed some aspects of what a definite theory of VRR should achieve. Let us now conclude by listing interesting future developments.

- We relied extensively on isotropy both for the perturbing background fluctuations as well as for the test particles’ initial conditions. Any effects associated with anisotropic clusterings in orientations was neglected. As explored in §3.3, non-spherical structures can spontaneously arise in long-term thermodynamical equilibria therefore affecting the statistics of the fluctuations and the efficiency of neighbour separation.
- As discussed in Eq. (3.30), we neglected the self-gravity between neighbouring test particles. Accounting for this additional interaction should increase the coherence of the test particles’ population and therefore reduce the efficiency of a disc’s dilution. Yet, given the complexity of the present stochastic dynamics, accounting for this self-gravitating coupling surely sounds like an analytical and technical challenge.
- We focused on an isotropic description of the dilution process. Strikingly, simulations such as in Fig. 3.2 exhibit clear elongations of the tracer distribution. It would be of interest to quantify, e.g., through three-point correlations, the rate with which these elongations arise. This could be leveraged as a supplementary observational constraint for future observations of SgrA* (Do et al., 2019).
- As long as one updates accordingly the coupling coefficients, \mathcal{J}_ℓ , VRR also occurs in globular clusters (see, e.g., Meiron & Kocsis, 2018), ultimately leading to the mixing (in orientation) of neighbouring stars. In

the light of GAIA's exquisite data, this stochastic process could prove important to understand how co-eval stars mix in the crowded environment of globular clusters.

- Building upon §3.2.2, a natural next step is to quantitatively apply this framework to SgrA*, in particular regarding two constraints: (i) the innermost S-stars seem to have an isotropic distribution of orientations; (ii) some of the outer stars seem to belong to a disc-like structure (see, e.g., Gillessen et al., 2017). As a consequence, the neighbour separation must be, on the one hand, efficient enough to mix the inner stars and, on the other hand, inefficient enough to allow for the outer disc's survival. Ultimately such a program should allow one to place constraints on SgrA*'s possible population of IMBHs.

3.3 Thermodynamical equilibrium

Sufficiently close to the central BH, VRR can occur on a timescale much shorter than the stellar ages themselves (see fig. 1 in Kocsis & Tremaine, 2015). As such, one may expect that the observed distribution of stellar orientations corresponds to some statistical equilibrium. Determining the outcome of this long-term rearrangement has been the focus of recent efforts (Roupas et al., 2017; Takács & Kocsis, 2018; Szölgvény & Kocsis, 2018; Touma et al., 2019; Tremaine, 2020a,b; Gruzinov et al., 2020; Máthé et al., 2022). These various works offered new clues on the fascinating properties of these long-range interacting systems such as negative temperature, negative heat capacity, and phase transition. In particular, Szölgvény & Kocsis (2018) showed that in clusters with a wide range of stellar populations (i.e. various masses and semi-major axes), VRR can lead to the spontaneous formation of a narrow disc through the angular segregation of the most massive stars.

This is the problem that we consider in this section. Our goal is to investigate the impact of a cluster's orbital diversity, e.g., its mass distribution or initial orientation distribution, on its long-term distribution of orientations. Rather than directly integrating the VRR evolution equations to let the cluster relax on its own (see §3.4), we focus on predicting the cluster's endstates by directly maximising its entropy. More precisely, we develop and implement an explicit and efficient optimisation procedure to infer a cluster's thermodynamical equilibrium. Such a roadmap was already started out by Roupas et al. (2017) in the limit of a single-population system with a quadrupolar interaction. This was later improved by Takács & Kocsis (2018) which, while still restricting themselves to a single-population considered harmonic expansion beyond the quadrupole. Here, we expand these works to multi-population clusters and emphasise the critical role played by the diversity of orbits and masses to allow for non-trivial anisotropic equilibria.

3.3.1 Equilibrium configurations

Rather than focusing on describing the details of $\partial F^b/\partial t$, see Eq. (3.6), one may solely focus on characterising the expected long-term equilibrium configurations reached through VRR. This is given by

$$F^{\text{eq}}(\widehat{\mathbf{L}}, \mathbf{K}) = \lim_{t \rightarrow +\infty} \langle F^b(\widehat{\mathbf{L}}, \mathbf{K}, t) \rangle, \quad (3.48)$$

where $\langle \cdot \rangle$ stands for an ensemble average over independent realisations of the system. Following Eq. (3.4), we note that the Hamiltonian of VRR is such that (i) the phase space domain for $\widehat{\mathbf{L}}$ is of finite volume; (ii) for $\mathbf{K} \neq \mathbf{K}'$, there is no divergence of the VRR interaction energy for $\widehat{\mathbf{L}} = \widehat{\mathbf{L}}'$. As such, the VRR dynamics does not present the typical peculiarities of the statistical mechanics of self-gravitating systems such as the evaporation of stars, the formation of binaries, the gravothermal instability, and the absence of a true statistical equilibrium (see, e.g., Chavanis, 2006). As a consequence, VRR systems always possess well-defined statistical equilibria.

On long timescales, the only information retained by the cluster are its invariants. These are:

- The number density of stars with orbital parameters \mathbf{K}

$$N(\mathbf{K}) = \int d\widehat{\mathbf{L}} F^{\text{eq}}(\widehat{\mathbf{L}}, \mathbf{K}), \quad (3.49)$$

that satisfies the normalisation $\int d\mathbf{K} N(\mathbf{K}) = N$, the total number of particles.

- The total energy

$$E_{\text{tot}} = \frac{1}{2} \int d\widehat{\mathbf{L}} d\mathbf{K} F^{\text{eq}}(\widehat{\mathbf{L}}, \mathbf{K}) \varepsilon(\widehat{\mathbf{L}}, \mathbf{K}), \quad (3.50)$$

where $\varepsilon(\widehat{\mathbf{L}}, \mathbf{K})$ stands for the energy of a particle of orientation $\widehat{\mathbf{L}}$ and orbital parameter \mathbf{K} , as given by

$$\varepsilon(\widehat{\mathbf{L}}, \mathbf{K}) = - \sum_{\alpha} \int d\mathbf{K}' \mathcal{H}_{\alpha}[\mathbf{K}, \mathbf{K}'] F_{\alpha}^{\text{eq}}(\mathbf{K}') Y_{\alpha}(\widehat{\mathbf{L}}), \quad (3.51)$$

where the magnetisations $F_{\alpha}^{\text{eq}}(\mathbf{K})$ follow from Eq. (3.8).

- The total angular momentum

$$\mathbf{L}_{\text{tot}} = \int d\widehat{\mathbf{L}} d\mathbf{K} L(\mathbf{K}) \widehat{\mathbf{L}} F^{\text{eq}}(\widehat{\mathbf{L}}, \mathbf{K}), \quad (3.52)$$

with $L(\mathbf{K})$ the norm of the angular momentum, given by Eq. (3.2).

In the microcanonical ensemble (i.e. for an isolated cluster), for a given set of invariants $\{N(\mathbf{K}), E_{\text{tot}}, \mathbf{L}_{\text{tot}}\}$, the admissible equilibrium configurations are obtained by extremising Boltzmann's entropy

$$S = -k_{\text{B}} \int d\widehat{\mathbf{L}} d\mathbf{K} F^{\text{eq}}(\widehat{\mathbf{L}}, \mathbf{K}) \ln [F^{\text{eq}}(\widehat{\mathbf{L}}, \mathbf{K})], \quad (3.53)$$

with k_{B} the Boltzmann constant, under the previous conservation constraints. Such a maximisation generically yields (see, e.g., Roupas et al., 2017)

$$F^{\text{eq}}(\widehat{\mathbf{L}}, \mathbf{K}) = N(\mathbf{K}) \frac{e^{-\beta \varepsilon(\widehat{\mathbf{L}}, \mathbf{K}) + L(\mathbf{K}) \gamma \cdot \widehat{\mathbf{L}}}}{\int d\widehat{\mathbf{L}}' e^{-\beta \varepsilon(\widehat{\mathbf{L}}', \mathbf{K}) + L(\mathbf{K}) \gamma \cdot \widehat{\mathbf{L}}'}}, \quad (3.54)$$

where β and γ are the Lagrange multipliers respectively associated with the total energy and total angular momentum conservation. See §A in Magnan et al. (2021) for a derivation of Eq. (3.54).

For a given galactic nucleus, the total angular momentum vector \mathbf{L}_{tot} provides us with one specific direction which we always chose to be the $+z$ axis. Unfortunately, this does not imply that the thermodynamical equilibria of VRR necessarily have an axial symmetry around \mathbf{L}_{tot} , since a spontaneous symmetry breaking could occur, e.g., through the flaring of the stellar disc (see fig. 6 in Kocsis & Tremaine, 2011). Nonetheless, for simplicity, we assume here that VRR does not exhibit any such symmetry breaking. We restrict ourselves to axisymmetric DFs, i.e. systems such that

$$\forall \mathbf{K}, \forall \ell, \forall m \neq 0, F_{\ell m}^{\text{eq}}(\mathbf{K}) = 0. \quad (3.55)$$

From there, we can make the simplifications $\mathbf{L}_{\text{tot}} \rightarrow L_{\text{tot}} > 0$, and $\gamma \rightarrow \gamma > 0$. This assumption greatly reduces the number of spherical harmonics to consider and therefore significantly alleviates the numerical complexity. However, we stress that this assumption is not always physically motivated. It is verified in the single-population and single-harmonic case (see fig. 1 in Roupas et al., 2017) or for null inverse temperature case in multi-population systems (see §C3 in Magnan et al., 2021), but it does not hold in some other regimes, as emphasised by the finding of a warped VRR disc (see fig. 6 in Kocsis & Tremaine, 2011). Overall, the axisymmetric assumption constitutes an important limitation of the present approach, and deserves to be challenged in future studies. To comply with this approach, the numerical average of Eq. (3.48) is carried out over realisations which are all rotated to have their \mathbf{L}_{tot} aligned along $+z$.

3.3.2 Self-consistency

While Eq. (3.54) is fully generic, it comes with the difficulty that one needs to impose self-consistency to it. Indeed, $F^{\text{eq}}(\widehat{\mathbf{L}}, \mathbf{K})$ involves the one-particle energy, $\varepsilon(\widehat{\mathbf{L}}, \mathbf{K})$, which via Eq. (3.51) and the magnetisations from Eq. (3.8) involves $F^{\text{eq}}(\widehat{\mathbf{L}}, \mathbf{K})$ itself. Within the present microcanonical ensemble, imposing the cluster's total energy, angular momentum, and orbital distribution ultimately sets up the cluster's temperature (via β), rate of rotation (via γ), and shape (via $F_{\ell 0}^{\text{eq}}(\mathbf{K})$)².

To effectively solve such a generic problem, we discretise the cluster as a finite set of stellar populations, indexed by k and described by the orbital parameters \mathbf{K}_k . We then write $F^{\text{eq}}(\widehat{\mathbf{L}}, \mathbf{K}) = \sum_k F_k(\widehat{\mathbf{L}}) \delta_{\text{D}}(\mathbf{K} - \mathbf{K}_k)$ with $F_k(\widehat{\mathbf{L}})$ the distribution of orientation of the k^{th} population and $N_k = \int d\widehat{\mathbf{L}} F_k(\widehat{\mathbf{L}})$ (see Eq. 3.49) its number of stars. After this discretisation, a cluster's configuration is fully characterised by its set of order parameters

$$\boldsymbol{\theta} = (\beta, \gamma, \{F_{\ell, k}\}), \quad (3.56)$$

with $F_{\ell, k} = \int d\widehat{\mathbf{L}} Y_{\ell 0}(\widehat{\mathbf{L}}) F_k(\widehat{\mathbf{L}})$ the axisymmetric ℓ -magnetisation of the k^{th} population, see Eq. (3.8).

For a given initial condition, i.e. a given $(E_{\text{tot}}, L_{\text{tot}}, \{N_k\})$, the cluster's equilibrium configurations are obtained for the parameters $\boldsymbol{\theta}$ that are joint roots of the consistency functions

$$\begin{aligned} C_E &= (E_{\text{tot}} - E_{\text{tot}}[F^{\text{eq}}(\boldsymbol{\theta})])/E_{\text{tot}}; \\ C_L &= (L_{\text{tot}} - L_{\text{tot}}[F^{\text{eq}}(\boldsymbol{\theta})])/L_{\text{tot}}; \\ \forall \ell, \forall k, C_{F_{\ell, k}} &= (F_{\ell, k} - F_{\ell, k}[F^{\text{eq}}(\boldsymbol{\theta})])/(N_k y_{\ell}), \end{aligned} \quad (3.57)$$

²There are infinitely many harmonics, ℓ , but in practice we found it acceptable to stop at $\ell_{\text{max}} = 10$, see §D2 in Magnan et al. (2021).

with $y_\ell = \sqrt{(2\ell+1)/(4\pi)}$ and $F^{\text{eq}} = F^{\text{eq}}(\boldsymbol{\theta})$ following Eq. (3.54). To place these constraints on equal footing, the cost functions $\mathbf{C} = (C_E, C_\ell, C_{F_{\ell,k}})$ are dimensionless and rescaled to be of order unity. In order to find roots of the highly multi-dimensional function $\boldsymbol{\theta} \mapsto \mathbf{C}[\boldsymbol{\theta}]$, similarly to Takács & Kocsis (2018), we use Newton–Raphson’s method (see, e.g., Press et al., 2007). Starting from a configuration $\boldsymbol{\theta}_n$, we proceed with the iteration

$$\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n - \mathbf{J}^{-1}[\boldsymbol{\theta}_n] \mathbf{C}[\boldsymbol{\theta}_n], \quad (3.58)$$

where $(\mathbf{J}[\boldsymbol{\theta}])_{ij} = \partial C_i / \partial \theta_j$ is the Jacobian of the cost functions. Importantly, because \mathbf{C} (see Eq. 3.57) has a simple analytical form, its Jacobian is straightforward to compute, as detailed in §B of Magnan et al. (2021).

To finalise our algorithm, it only remains to specify an appropriate choice for the initial condition $\boldsymbol{\theta}_0$ of the iteration process. Given the technical difficulty of setting up an optimisation routine with thousands of degree of freedom, we limit ourselves to only predicting the clusters’ global thermodynamical equilibria. In practice, for systems with $L_{\text{tot}} \neq 0$, these systems are expected to be found along the series of equilibria associated with the branch that has a solution for $\beta = 0$ (see Fig. 3.11). As a consequence, our iterative approach is to first find the equilibrium for $\beta = 0$ and L_{tot} given. This is a much easier problem because there are no self-gravitating contribution to account for. Once this starting point obtained, we move along the associated series of equilibria by iteratively changing the target E_{tot} , while keeping the same L_{tot} . Although such an iterative approach allows us to smoothly deal with optimisation problems with a large number of degrees of freedom, it cannot be used to determine any of the clusters’ other thermodynamical equilibria (unstable or metastable), as it is, by design, limited to the dominant branch that goes through $\beta = 0$. We refer to §C of Magnan et al. (2021) for details.

3.3.3 A typical equilibrium distribution

We are now set to use our entropy optimisation algorithm to investigate typical equilibrium distribution of orientations in galactic nuclei. Our fiducial model is the same as in Szölgyén & Kocsis (2018). We assume that stars are formed through a series of distinct episodes of star formation or infall events. More precisely, for a given realisation we generate a total of 16 discs, each composed of 512 stars. For each disc, its orientation is drawn uniformly on the unit sphere, while stars within a disc all lie within the same small polar cap. Finally, for a given star, its orbital parameters $\mathbf{K} = (m, a, e)$ are drawn from power-law PDFs proportional to (m^{-2}, a^0, e) within the domains $m_{\text{max}}/m_{\text{min}} = 100$, $a_{\text{max}}/a_{\text{min}} = 100$ and $(e_{\text{min}}, e_{\text{max}}) = (0, 0.3)$. We recall that realisations are always rotated to have their \mathbf{L}_{tot} aligned along $+z$. We illustrate in the left panel of Fig. 3.9 a typical realisation of such a protocol and refer to §3.1 in Magnan et al. (2021) for further details.

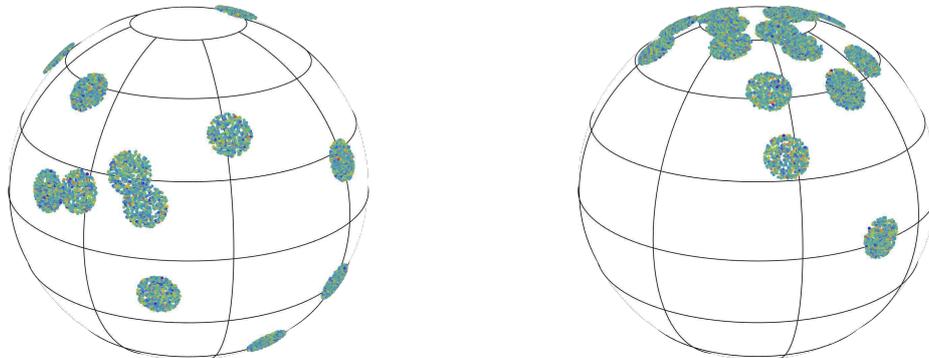


Figure 3.9: From Magnan et al. (2021). Left: Typical initial condition for the stellar orientations. Each of the $N_{\text{disc}} = 16$ dense patches of stars are scattered uniformly on the unit sphere and mimic distinct episodes of star formation or infall events. Stars are colored according to the norm of their angular momentum vector (Eq. 3.2), with the smaller norm in blue and the largest one in red. Right: Same protocol as for the left panel except that the orientations of the discs are drawn from a von Mises–Fischer PDF with concentration $\kappa = 5$ (Eq. 3.62).

Once an initial stellar distribution has been drawn, one can compute its two key invariants, E_{tot} and $|\mathbf{L}_{\text{tot}}|$. In practice, we keep track of these two invariants through the dimensionless quantities

$$E = -\frac{E_{\text{tot}}}{N^2 G m_{\text{min}}^2 / a_{\text{min}}}; \quad s = \frac{|\mathbf{L}_{\text{tot}}|}{\sum_i L(\mathbf{K}_i)}, \quad (3.59)$$

which we respectively call the cluster’s binding energy and spin. In the left panel of Fig. 3.12, we present the typical distribution of (E, s) for a large number of clusters drawn from our fiducial procedures. Clusters exhibit a wide range of binding energy and spin which in turn impacts the diversity of thermodynamical equilibria.

Let us now investigate in detail the particular case of a cluster with invariants $(E, s) = (2 \times 10^{-3}, 0.2)$, i.e. a typical cluster realisation. We then maximise entropy and determine this cluster’s equilibrium configuration,

$F^{\text{eq}}(\widehat{\mathbf{L}}, \mathbf{K})$. For a given realisation, we define $\widehat{L}_z = \widehat{\mathbf{L}} \cdot \widehat{\mathbf{L}}_{\text{tot}}$ and estimate the system's DF, $F^{\text{eq}}(m, \widehat{L}_z)$, integrated over all semi-major axes and eccentricities. To highlight non-spherically symmetric structures, we finally define

$$\widehat{F}_{\text{eq}}(m, \widehat{L}_z) = \frac{F^{\text{eq}}(m, \widehat{L}_z)}{\text{Max}_{\widehat{L}_z} [F^{\text{eq}}(m, \widehat{L}_z)]}, \quad (3.60)$$

such that each mass bin has its maximum equal to unity. The resulting equilibrium is illustrated in Fig. 3.10.

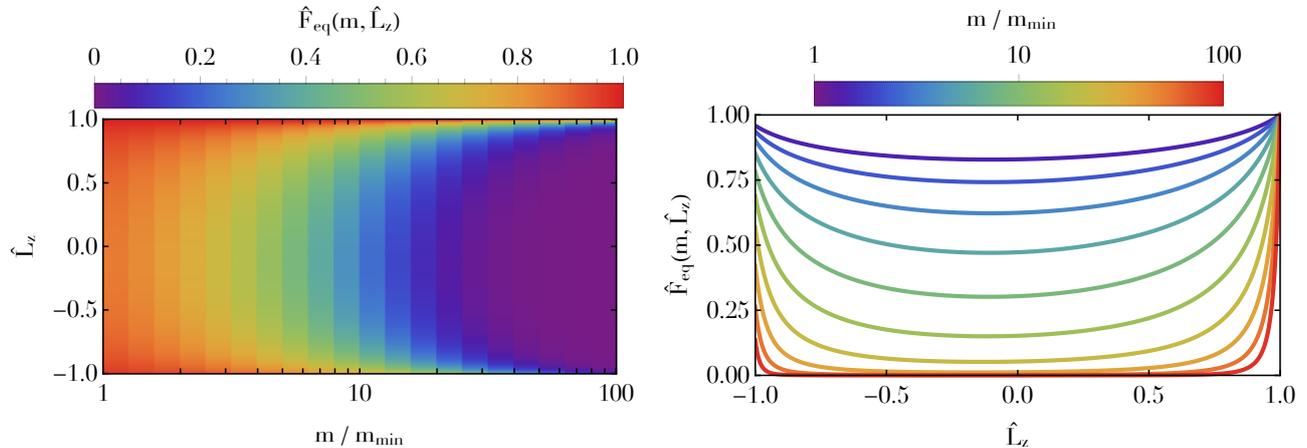


Figure 3.10: From Magnan et al. (2021). Left: Thermodynamical equilibrium, $\widehat{F}_{\text{eq}}(m, \widehat{L}_z)$ (Eq. 3.60), as a function of mass and orientation with $(E, s) = (2 \times 10^{-3}, 0.2)$. Right: Same as the left panel, where each colored line corresponds to a different mass bin. The orientation's anisotropy between the different masses is clearly visible.

As is clearly visible, heavy particles tend to have their unit angular momentum vector, $\widehat{\mathbf{L}}$, oriented towards the North pole, which implies that the associated stars tend to orbit near the equatorial plane, a conclusion already reached in Szölglyén & Kocsis (2018) using a Monte-Carlo approach. Having the heavy particles, e.g., the IMBHs orbit close to the same orbital plane can drastically impact the rate of their pairwise mergers in galactic nuclei. For the heaviest stars, we also note the presence of an additional over-density near the South pole indicating the presence of a counter-rotating equatorial disc. As expected from the conservation of angular momentum, this component is naturally less populated than the main prograde disc.

As one considers lighter particles, the anisotropy fades away and the lightest particles do not show any strong sign of spontaneous orientation alignment. Because we have the full equilibrium distribution, $F^{\text{eq}}(\widehat{\mathbf{L}}, \mathbf{K})$, at our disposal, we may also study its dependence w.r.t. the semi-major axis and eccentricity. In that regard, we find a (weaker) dependence of F^{eq} w.r.t. a , so that the disc predicted in Fig. 3.10 is slightly thinner for medium semi-major axes (Máthé et al., 2022). Finally, we find that, in essence, F^{eq} is independent of e . This was to be expected since the coupling coefficients, $\mathcal{H}_\ell[\mathbf{K}, \mathbf{K}']$ (Eq. 3.5), and the norm of the angular momentum, $L(\mathbf{K})$ (Eq. 3.2), only weakly depend on e , especially for the chosen quasi-circular orbits, $0 \leq e \leq 0.3$.

3.3.4 Impact of the initial conditions

As highlighted in the left panel of Fig. 3.12, even when drawn from the same protocol, clusters exhibit a significant diversity in their binding energies and spins, i.e. a diversity in their invariants. Let us therefore investigate the dependence of the equilibria w.r.t. these invariants.

In Fig. 3.11, we present series of equilibria (caloric curves) giving the inverse temperature, β , as a function of the normalised total energy, $-E$, for various total angular momentum (i.e. various s). The temperature generically increases with the energy, except in a small region of negative specific heat. For any value of the spin, there also exists a region of negative temperature. Both effects were already pointed out in Roupas et al. (2017); Takács & Kocsis (2018). We refer to §C1 in Magnan et al. (2021) for further discussion. We note however that these behaviours arise outside of the astrophysical regime of the left panel of Fig. 3.12. In the right panel of Fig. 3.11, we present the dependence of the angular frequency γ w.r.t. the total angular momentum s for fixed values of the binding energy E . The angular frequency is always found to increase with the cluster's spin.

Within the domain of invariants spanned by the left panel of Fig. 3.12, all equilibria are found to be qualitatively similar, but differ in the strength of their anisotropy. To characterise the level of anisotropy, we define the angular size, θ_c , of the northern polar cap that contains a fraction f_c of the heavy particles' unit angular

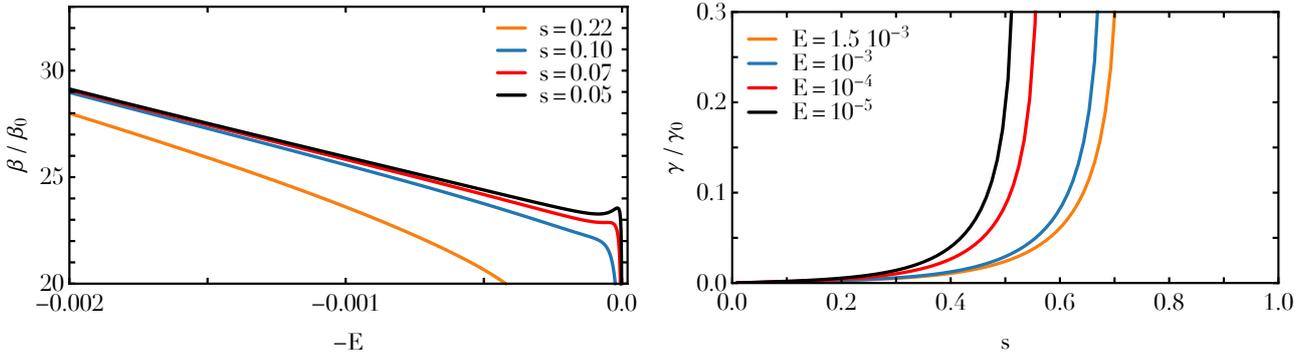


Figure 3.11: From Magnan et al. (2021). Left: Equilibrium inverse temperature, β , as a function of the cluster’s normalised total energy, $-E$, with $\beta_0^{-1} = NGm_{\min}^2/a_{\min}$. Right: Equilibrium angular frequency, γ , as a function of the cluster’s spin, s , for different binding energies, with $\gamma_0^{-1} = m_{\min}\sqrt{GM_{\bullet}a_{\min}}$. For a given s , our approach to determine equilibria is to start from the simple problem at $\beta = 0$ and iteratively move up to the target energy.

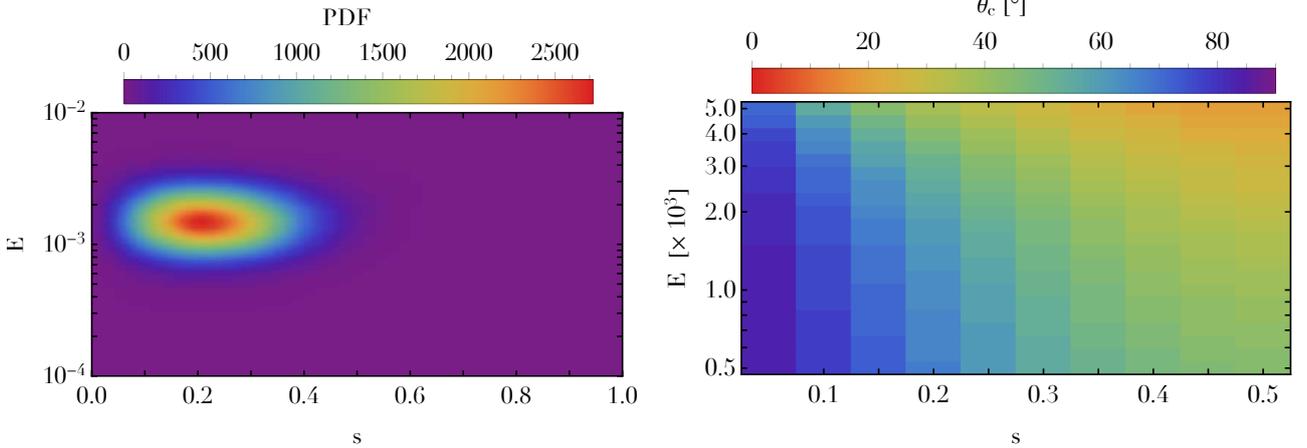


Figure 3.12: From Magnan et al. (2021). Left: Distribution of the spin and binding energy, (s, E) (Eq. 3.59), of clusters drawn following the fiducial protocol from the left panel of Fig. 3.9. Right: Segregation angle, θ_c , as a function of cluster’s invariants, (s, E) , the total spin and binding energy.

momentum $\hat{\mathbf{L}}$. More precisely, θ_c follows from the implicit constraint

$$\frac{\int_{m \geq m_c} d\mathbf{K} \int_{\hat{L}_z \geq \cos(\theta_c)} d\hat{\mathbf{L}} F^{\text{eq}}(\hat{\mathbf{L}}, \mathbf{K})}{\int_{m \geq m_c} d\mathbf{K} \int d\hat{\mathbf{L}} F^{\text{eq}}(\hat{\mathbf{L}}, \mathbf{K})} = f_c, \quad (3.61)$$

noting that anti-aligned stars are not accounted for. Equivalently, θ_c is also the angular size of the equatorial disc that contains a fraction f_c of the heavy particles. In practice, we used the mass threshold $m_c = 10 m_{\min}$ with the fraction $f_c = 50\%$. We illustrate in the right panel of Fig. 3.12 the dependence of θ_c w.r.t. the cluster’s invariants (s, E) . We observe the crucial role played by these two invariants in driving the spontaneous alignment of the orbital orientations of the heaviest particles. The more bound the cluster and the larger its total angular momentum, the thinner the disc of heavy particles in the relaxed equilibrium.

Let us now step out and explore the impact of the distribution of the orbital stellar parameters themselves on the clusters’ thermal equilibria. Keeping all other parameters the same, we draw (m, a) according to power law distributions $(m^{\gamma_m}, a^{\gamma_a})$ but varying the power law indices (γ_m, γ_a) between different realisations. Their impact on the average segregation rate of the heavy particles, $\langle \theta_c \rangle$, is illustrated in the left panel of Fig. 3.13. We find that the slope γ_m has the strongest effects, with clusters containing fewer heavy stars segregating in a disc twice thinner than the fiducial cluster. The semi-major axis power law index also impacts the segregation rate: clusters with larger semi-major axes display a stronger mass segregation. Such trends may possibly be used in the future to estimate indirectly the stellar and compact objects’ initial mass function in galactic centres.

To conclude this section, we briefly investigate the impact of the dispersion of the stars’ initial orientations on the thermodynamical equilibria. Rather than drawing the initial orientations of the discs uniformly on the unit sphere, we draw them along some biased direction to reflect the preferential infall of new star-forming gas along specific directions, e.g., imposed by past gaseous accretion events. More precisely, we assume that the orientation of each disc is drawn according to the Von Mises–Fisher PDF Wood (1994)

$$P(\hat{\mathbf{L}}) = \frac{\kappa}{4\pi \sinh(\kappa)} e^{\kappa \hat{L}_z}, \quad (3.62)$$

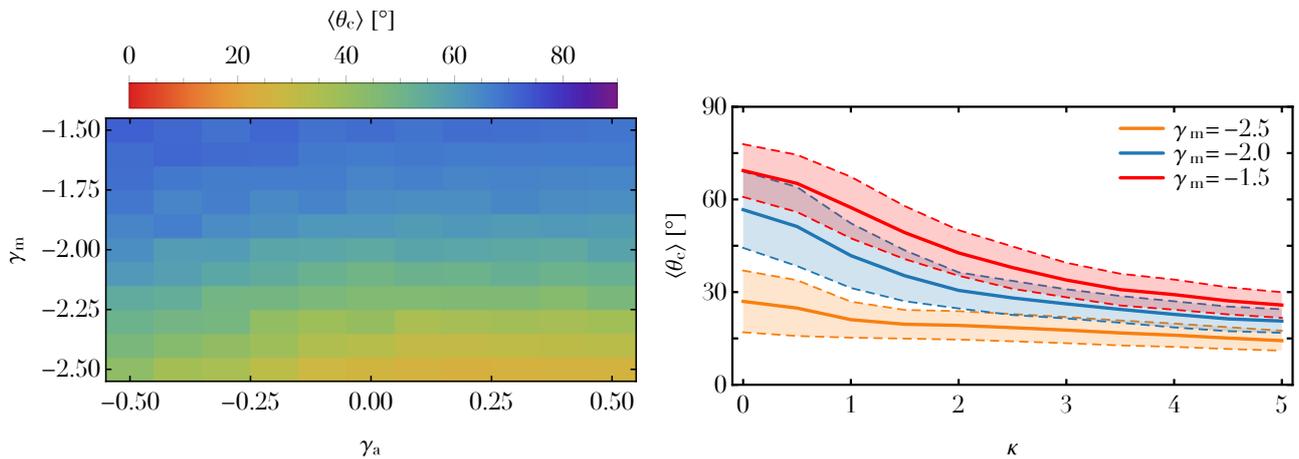


Figure 3.13: From Magnan et al. (2021). Left: Dependence of the average segregation angle, $\langle \theta_c \rangle$, with the power law indices (γ_a, γ_m) of the initial distributions in semi-major axes and masses. Right: Dependence of $\langle \theta_c \rangle$ w.r.t. the concentration parameter κ (Eq. 3.62) for three different mass distributions.

with κ the PDF's concentration. The case $\kappa = 0$ (resp. $\kappa = 5$) is illustrated in the left (resp. right) panel of Fig. 3.9. The larger κ , the smaller the spread of the PDF on the unit sphere.

In the right panel of Fig. 3.13, we illustrate the dependence of the average segregation angle, $\langle \theta_c \rangle$, as a function of κ . Narrower initial distributions of orientations lead to stronger segregations of the heavy stars at equilibrium. This is expected since the more concentrated the distribution of initial orientation, the larger the values of s and E , which both correlate positively with the segregation strength (Fig. 3.12). To emphasise this conclusion, the right panel of Fig. 3.13 illustrates the thickness of the asymptotic massive disc as a function of the initial mass function slope, γ_m . Similarly to the left panel of Fig. 3.13, at fixed initial anisotropy, the steeper the slope, the stronger the mass segregation. Overall, investigations as in Fig. 3.13 should prove useful to place some constraints on the origin of SgrA*'s surrounding stellar cluster, in particular in the light of the "paradox of youth" (Ghez et al., 2003; Genzel et al., 2010).

3.3.5 Prospects

In this section, we relied on a maximum entropy method to assess the internal stellar structure of galactic nuclei and their underlying distribution of orbital orientations. Building upon Roupas et al. (2017); Takács & Kocsis (2018), we expanded their approach to multi-populations clusters and showed how to jointly account for the constraints of energy, angular momentum, stellar mass, and orbital parameter conservation to efficiently characterise a cluster's expected thermodynamical equilibrium. In particular, as pointed out in Szölygyén & Kocsis (2018) using a Monte-Carlo approach, we recovered in Fig. 3.10 the spontaneous alignment of the heavier stellar components such as IMBHs. The more bound the stellar cluster and the larger its spin, the thinner the asymptotic disc obtained via VRR. Benefiting from the versatility of the present method, we used it to explore the dependence of these structures w.r.t. the cluster's initial conditions such as the distribution of masses and semi-major axes or the initial spread in orientations (Fig. 3.13). Ultimately, these explorations should pave the way for systematic explorations of the impact of a given formation model on the long-term fate of galactic nuclei.

As for future prospects, we may mention:

- We assumed from the start that stellar eccentricities were conserved during VRR, hence neglecting the contributions from SRR whose signature is clearly visible in the S-cluster distribution, see §2.3.4. As hinted in Szölygyén et al. (2021), allowing the eccentricities to also vary during the orientations' relaxation can impact the system's equilibrium. Similarly, stars can also change of semi-major axes through NR whose possible impact on the outcome of VRR deserves to be investigated in detail.
- As highlighted in Fig. 3.10, VRR may lead to highly non-spherical distribution of orientations. Naturally, such a non-trivial end-state distribution must impact the efficiency with which newly formed stellar discs can spontaneously dilute, as explored in §3.2.
- We limited ourselves to axisymmetric equilibria. At a significantly larger numerical cost, one could extend the present approach to non-axisymmetric distributions, i.e. self-consistently account for the magnetisations $F_{\ell m}^{\text{eq}}$ with $m \neq 0$. This was already partially tackled in Roupas et al. (2017), yet in the limit of a single-population system with a quadrupolar interaction. Fortunately, in that case, as soon as $|\mathbf{L}_{\text{tot}}| \neq 0$, the global maximum of entropy were always found to be axisymmetric.
- In Eq. (3.4), we averaged the cluster's Hamiltonian over both the (fast) Keplerian motion and the (slower) in-plane precessions. Lifting this second average, it would be of interest to determine self-consistent equi-

libria for Keplerian elliptic wires around a supermassive BH (Gruzinov et al., 2020; Tremaine, 2020b) and characterise the impact of a cluster's invariants on the amount of orientation anisotropy.

- The generic VRR Hamiltonian from Eq. (3.4) also applies to globular clusters, as long as they are assumed to remain on average spherically symmetric. In that context, one could use stellar metallicities or ages as additional tracers of the system's relaxation.
- By design, entropy maximisation can only predict the overall end-state of VRR, i.e. its thermodynamical limit. It cannot estimate of the time required for the equilibrium to be reached. To circumvent this issue, one could: (i) Use direct time integrations of the system's dynamics to let the cluster relax. This is touched upon in §3.4; (ii) Derive an explicit kinetic theory for the cluster, assuming a sufficiently symmetric orbital distribution. This is explored in §3.5 for a single-population axisymmetric cluster.
- Finally, following recent surveys (Gillessen et al., 2017; Habibi et al., 2017), planned upgrade on the VLTI (Eisenhauer, 2019; Gravity Collaboration et al., 2021), as well as the future thirty-meter class telescopes such as ELT (Pott et al., 2018; Davies et al., 2018) and TMT (Do et al., 2019), the astrophysical community will soon have at its disposal a wealth of orbitally-resolved stars around SgrA*. Characterising their orbital distribution will place stringent constraints on the properties of the (likely present) un-observed IMBHs and the impact of their distribution of orientations on their in-situ merger rates.

3.4 Simulations

The numerical study of VRR is challenging because its Hamiltonian (Eq. 3.4) requires explicit orbit-averages that smear out the stars along their in-plane motions. This greatly complicates the computation of the system's pairwise interactions as it now involve pairs of long-range coupled annuli rather than pairs of stars. In this section, we walk in the footsteps of §2.3 and set out to illustrate how a multipole expansion (Hénon, 1964, 1973; van Albada & van Gorkom, 1977; Villumsen, 1982; Meiron et al., 2014; Dehnen, 2014) tailored to the VRR Hamiltonian offers an efficient numerical scheme that can integrate the long-term evolution of annuli with a complexity scaling like $\mathcal{O}(NK\ell_{\max}^2)$ with N the total number of stars, K a number of sampling nodes independent of N , and ℓ_{\max} the maximum harmonics considered in the multipole expansion.

3.4.1 Equations of motion

Following our double orbit-average the canonical variables for a given annulus are (Eq. 1.5)

$$(\Omega, L_z) = (\phi, L[\mathbf{K}] \cos(\vartheta)), \quad (3.63)$$

with $L[\mathbf{K}]$ the norm of the angular momentum vector (Eq. 3.2) and (ϑ, ϕ) the usual spherical coordinates. We also dropped the index i to shorten the notation. Importantly, the VRR dynamics is submitted to the constraint

$$|\widehat{\mathbf{L}}| = 1. \quad (3.64)$$

Hamilton's equation naturally read $\dot{\Omega} = \partial H_{\text{VRR}} / \partial L_z$ and $\dot{L}_z = -\partial H_{\text{VRR}} / \partial \Omega$ with H_{VRR} given by Eq. (3.4). The conservation of $L[\mathbf{K}]$ ultimately leads to the evolution equation for $\widehat{\mathbf{L}}$

$$\frac{d\widehat{\mathbf{L}}}{dt} = \frac{1}{L[\mathbf{K}]} \frac{\partial H_{\text{VRR}}}{\partial \widehat{\mathbf{L}}} \times \widehat{\mathbf{L}} \quad (3.65)$$

Equation (3.65) is the VRR equivalent of Milankovitch's Eq. (2.24). We emphasise that H_{VRR} benefits from a gauge invariance associated with the constraints from Eq. (3.64). As a consequence, similarly to Eq. (2.22), we rewrite Eq. (3.65) as

$$\frac{d\widehat{\mathbf{L}}}{dt} = \boldsymbol{\Omega}(\widehat{\mathbf{L}}) \times \widehat{\mathbf{L}}, \quad (3.66)$$

with the conservative choice $\boldsymbol{\Omega} = \widehat{\mathbf{L}} \times (d\widehat{\mathbf{L}}/dt)$.

Starting from Eq. (3.4), one easily gets an explicit expression for $d\widehat{\mathbf{L}}_i/dt$, namely

$$\begin{aligned} \frac{d\widehat{\mathbf{L}}_i}{dt} &= -\frac{1}{L[\mathbf{K}_i]} \sum_{j=1}^N \sum_{\substack{\ell \geq 2 \\ \text{even}}}^{+\infty} \frac{2\ell+1}{4\pi} \mathcal{H}_\ell[\mathbf{K}_i, \mathbf{K}_j] P'_\ell(\widehat{\mathbf{L}}_i \cdot \widehat{\mathbf{L}}_j) \widehat{\mathbf{L}}_j \times \widehat{\mathbf{L}}_i \\ &= -\frac{1}{L[\mathbf{K}_i]} \sum_{j=1}^N \sum_{\substack{\ell \geq 2 \\ \text{even}}}^{+\infty} \sum_{m=-\ell}^{\ell} \mathcal{H}_\ell[\mathbf{K}_i, \mathbf{K}_j] Y_{\ell m}(\widehat{\mathbf{L}}_j) \mathbf{X}_{\ell m}(\widehat{\mathbf{L}}_i), \end{aligned} \quad (3.67)$$

with $\mathcal{H}_\ell[\mathbf{K}, \mathbf{K}']$ introduced in Eq. (3.5) and the real vector spherical harmonics $\mathbf{X}_{\ell m}(\widehat{\mathbf{L}}) = (\partial Y_{\ell m}(\widehat{\mathbf{L}})/\partial \widehat{\mathbf{L}}) \times \widehat{\mathbf{L}}$.

Of course, the similarities of the present setup with the case considered in §2.3 are numerous. We point out two important simplifications in the VRR case. First, one could be concerned that Eq. (3.67) includes the non-physical self-interaction contribution from the case $j=i$. Fortunately, this is not a problem, because this particular term does not contribute to the dynamics since $\widehat{\mathbf{L}}_i \times \widehat{\mathbf{L}}_i = 0$. As a consequence, contrary to Eq. (2.34), one does not need to filter out these self-interaction components. This simplifies the numerical implementation.

Second, because the stellar parameters $\mathbf{K} = (m, a, e)$ are conserved by the VRR dynamics, the nodes' masses (see Eq. 2.25) are kept exactly fixed during the evolution. As a consequence, subtle contributions as in Eq. (2.30) are absent from the VRR problem. Phrased differently, the discretisation of the VRR Hamiltonian and the computation of the rates of change commute with one another.

3.4.2 Rates of change

In order to numerically compute Eq. (3.67), we follow the same approach as in Eq. (2.26) by discretising the radial integral from Eq. (3.5) with K nodes uniformly sampled in eccentric anomaly. One gets

$$\frac{d\widehat{\mathbf{L}}_i}{dt} = -\frac{4\pi G}{L_i} \sum_{j=1}^N \sum_{k,l=1}^K \sum_{\ell,m} \gamma_\ell Y_{\ell m}(\widehat{\mathbf{L}}_j) \mathbf{X}_{\ell m}(\widehat{\mathbf{L}}_i) \mu_{ik} \mu_{jl} \frac{\text{Min}[r_{ik}, r_{jl}]^\ell}{\text{Max}[r_{ik}, r_{jl}]^{\ell+1}}, \quad (3.68)$$

with the shortened notations $L_i = L(\mathbf{K}_i)$, $\gamma_\ell = |P_\ell(0)|^2 / (2\ell + 1)$, and the effective mass $\mu_{ik} = m_i r_{ik} / a_i$ (see Eq. 2.27). This can be cast as

$$\frac{d\widehat{\mathbf{L}}_i}{dt} = -\frac{4\pi G}{L_i} \sum_{k=1}^K \mu_{ik} \boldsymbol{\omega}_{ik} \quad \text{with} \quad \boldsymbol{\omega}_{ik} = \sum_{j,l}^{N,K} \gamma_\ell Y_{\ell m}(\widehat{\mathbf{L}}_j) \mathbf{X}_{\ell m}(\widehat{\mathbf{L}}_i) \mu_{jl} \frac{\text{Min}[r_{ik}, r_{jl}]^\ell}{\text{Max}[r_{ik}, r_{jl}]^{\ell+1}}. \quad (3.69)$$

Notations can be further condensed by writing

$$\boldsymbol{\omega}_{ik} = (\mathbf{P}_{ik} + \mathbf{Q}_{ik}) \cdot \mathbf{X}_{ik}, \quad (3.70)$$

where the scalar product stands for the summation over all harmonic indices. We also introduced

$$\mathbf{X}_{ik} = \left\{ \gamma_\ell \mathbf{X}_{\ell m}(\widehat{\mathbf{L}}_i) \right\}_{(\ell,m)}, \quad (3.71)$$

as well as the prefix sums

$$\mathbf{P}_{ik} = \{P_{ik}^{\ell m}\}_{(\ell,m)} = \left\{ \sum_{j,l; r_{jl} < r_{ik}} \mu_{jl} Y_{\ell m}(\widehat{\mathbf{L}}_j) \frac{r_{jl}^\ell}{r_{ik}^{\ell+1}} \right\}_{(\ell,m)}; \quad \mathbf{Q}_{ik} = \{Q_{ik}^{\ell m}\}_{(\ell,m)} = \left\{ \sum_{j,l; r_{jl} > r_{ik}} \mu_{jl} Y_{\ell m}(\widehat{\mathbf{L}}_j) \frac{r_{ik}^\ell}{r_{jl}^{\ell+1}} \right\}_{(\ell,m)}, \quad (3.72)$$

following Eq. (2.38). We point out that we did not include the contribution $\{j, l\} = \{i, k\}$. Such a choice is allowed because nodes of the same wire do not source any evolution on one another (see Eq. 3.67). Removing this contribution allows for the definition of \mathbf{P}_n and \mathbf{Q}_n to remain somewhat symmetric.

The last step of the algorithm is, in essence, the same as in §2.3.2.2. More precisely, one first sorts the nodes, $\alpha = \{i, k\}$, by increasing radii. The $P_n^{\ell m}$ and $Q_n^{\ell m}$ are then determined through the recurrences

$$P_1^{\ell m} = 0; \quad P_{\alpha+1}^{\ell m} = \left(\frac{r_\alpha}{r_{\alpha+1}} \right)^{\ell+1} \left(P_\alpha^{\ell m} + \frac{\mu_\alpha}{r_\alpha} Y_\alpha^{\ell m} \right) \quad \text{for } 1 \leq \alpha < NK, \quad (3.73a)$$

$$Q_{NK}^{\ell m} = 0; \quad Q_{\alpha-1}^{\ell m} = \left(\frac{r_{\alpha-1}}{r_\alpha} \right)^\ell \left(Q_\alpha^{\ell m} + \frac{\mu_\alpha}{r_\alpha} Y_\alpha^{\ell m} \right) \quad \text{for } 1 < \alpha \leq NK. \quad (3.73b)$$

We conclude this section by discussing a few properties of the present method. First, the computations of \mathbf{P}_n and \mathbf{Q}_n can each be performed in a single pass, ensuring an overall complexity of the algorithm in $\mathcal{O}(NK\ell_{\max}^2)$. These two recurrences are also independent from one another so that they can natively be performed in parallel. This ensures a straightforward parallelisation over two cores. The relations from Eq. (3.73) are also expected to be numerically well-behaving because their respective geometric prefactors, $(r_\alpha/r_{\alpha+1})^{\ell+1}$ and $(r_{\alpha-1}/r_\alpha)^\ell$, are always smaller than unity. A final appeal of this approach is its memory imprint that scales linearly with N . In particular, nodes associated with the same annulus share the exact same orbital orientation: the spherical harmonics and their gradients only need to be computed once for each annulus.

In Fig. 3.14, we illustrate the dependence of the rates of change $d\widehat{\mathbf{L}}/dt$ as a function of K and ℓ_{\max} for the same cluster as in Fig. 3.15. For a fixed value of ℓ_{\max} , we find that $d\widehat{\mathbf{L}}/dt$ converges in $K^{-2.5}$. This matches

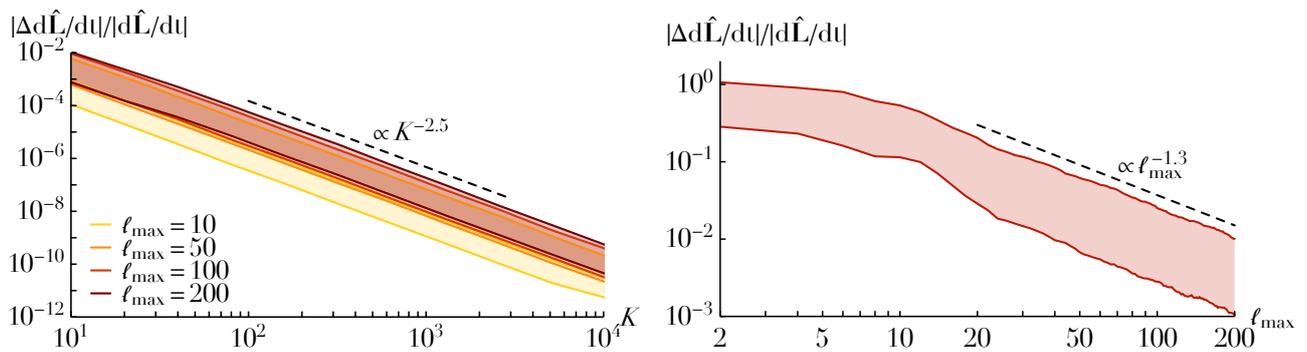


Figure 3.14: From Fouvry (2022). Relative error in $d\widehat{\mathbf{L}}/dt$ for one realisation of the system considered in Fig. 3.15 with $N = 2048$. Colored regions are the 16% and 84% levels among the stars. Left: Relative error as function of K , the number of nodes, for fixed values of ℓ_{\max} w.r.t. calculations with $K = 20480$. Right: Relative error as a function of ℓ_{\max} , the maximum harmonic index, for a fixed value $K = 20480$ w.r.t. calculations with $\ell_{\max} = 400$.

the rate obtained for E_{tot} in Fig. 2.5. This comes as no surprise since Eq. (3.68) does not suffer from any discontinuities. For a fixed value of K , we find that $d\widehat{\mathbf{L}}/dt$ converges roughly like $\ell_{\max}^{-1.3}$. This is compatible with the asymptotic scalings derived in §B of Kocsis & Tremaine (2015). Indeed, for radially overlapping non-coplanar orbits, which mostly compose the present cluster, a given harmonic ℓ gives a contribution of order ℓ^{-2} (see Eq. (B75) in Kocsis & Tremaine, 2015). Once summing over all harmonics $0 \leq \ell \leq \ell_{\max}$, this ultimately leads to an error scaling like ℓ_{\max}^{-1} in qualitative agreement with Fig. 3.14.

3.4.3 Time integration

In Eq. (3.66), we have show that the time evolution of the unit angular momentum vector takes the simple form

$$\frac{d\widehat{\mathbf{L}}_i}{dt} = \boldsymbol{\Omega}_i \times \widehat{\mathbf{L}}_i, \quad (3.74)$$

with $\boldsymbol{\Omega}_i = \widehat{\mathbf{L}}_i \times (d\widehat{\mathbf{L}}_i/dt)$. This precession equation ensures that $|\widehat{\mathbf{L}}_i| = 1$ throughout the evolution. Fortunately, this corresponds to the exact same geometrical constraints than the ones of the SRR problem (see §2.3.1.2). As a consequence, it is straightforward to apply the integration schemes from §2.3.3 to the present VRR case. We refer to §2.3.3.3 for the presentation of these integrators and the associated convergence tests.

3.4.4 Application

To conclude this section, we use direct time integrations to recover the spontaneous formation of stellar discs in galactic nuclei, as already investigated in §3.3 using there a maximum entropy method. We consider the exact same cluster as in Fig. 3.3.3. More precisely, we set $N = 2048$, $\ell_{\max} = 10$, $K = 20$. Defining $\tau_{\text{dyn}} = [a_{\text{min}}^3 / (Gm_{\text{min}})]^{1/2}$, we set the integration timestep to $\tau = 5 \times 10^{-4} \tau_{\text{dyn}}$ and integrate up to $T = 10^3 \tau_{\text{dyn}}$ using the MK2 scheme (see Eq. 2.45) for a total of 100 independent realisations. With such choices, the final relative errors in \mathbf{L}_{tot} (resp. E_{tot}) are typically of the order of 10^{-10} (resp. 10^{-7}).

In Fig. 3.15, we illustrate the ensemble-averaged outcome of these time integrations. It is reassuring to note

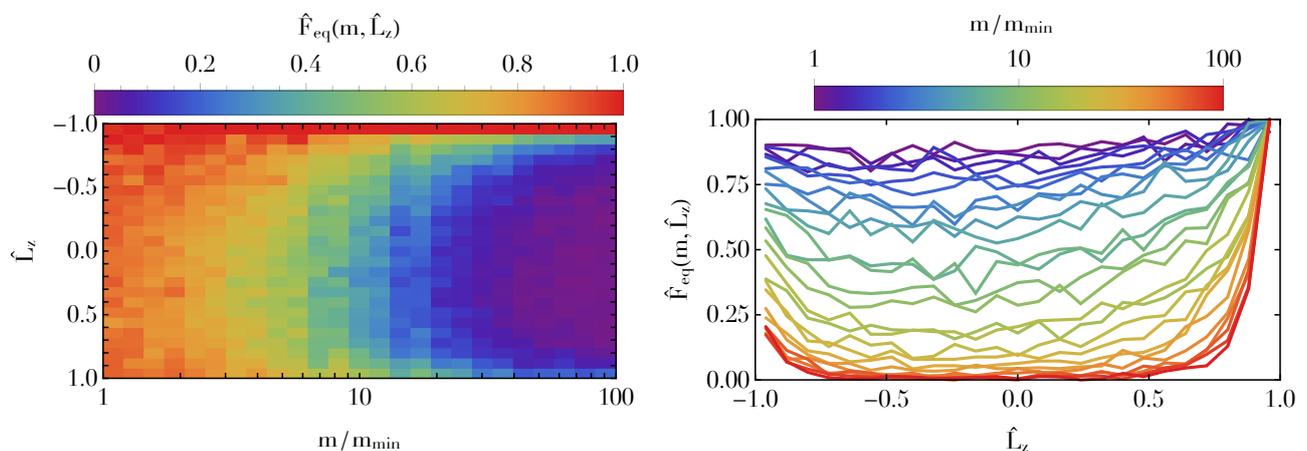


Figure 3.15: Same as Fig. 3.10 but here the relaxed DF, $\widehat{F}_{\text{eq}}(m, \widehat{L}_z)$, is obtained by directly integrating the VRR equations of motion via a multipole method. The similarities with Fig. 3.10 is striking and one recovers the spontaneous alignment of the massive stars within the same orbital plane.

that we recover in Fig. 3.15 the exact same results as in Fig. 3.10, albeit more noisy because of the Poisson shot noise from every individual N -body realisations. We refer to §3.3.3 for a physical discussion of these equilibria.

3.4.5 Discussion

The present multipole approach differs from an alternative integration scheme presented in Kocsis & Tremaine (2015). In a nutshell, their method relies on rewriting the VRR Hamiltonian as a large collection of two-body problems. As such, stellar orientations are integrated forward in time by resolving, in the appropriate order, series of two-body precessions. A key gain from this splitting approach is that each two-body problem can be integrated exactly for an arbitrary duration. This ensures the exact conservation of each angular momentum vector as well as of the cluster's total angular momentum vector. This also makes the overall scheme symplectic. Rates of change are computed via the first line of Eq. (3.67), i.e. by evaluating Legendre polynomials and not spherical harmonics. However, to do so, Kocsis & Tremaine (2015) has to pre-compute initially all the coupling coefficients, $\{\mathcal{J}_\ell[\mathbf{K}_i, \mathbf{K}_j]\}_{(i,j)}$, that must be stored for the entire duration of the simulation. This is one of the main drawback of that approach whose computational complexity and memory imprint scale like $\mathcal{O}(N^2 \ell_{\max})$. These two problems are fully alleviated by the present multipole approach. Fortunately, Kocsis & Tremaine (2015) could still make some of these difficulties less cumbersome because their algorithm can be parallelised somewhat easily and can also benefit from multi-timesteps schemes.

3.4.6 Prospects

In this section, we detailed how one may use a multipole approach to perform efficient double orbit-averaged simulations of stellar orientations in galactic nuclei. The present approach has three main advantages: (i) By relying on explicit orbit-averages, the cluster's fast dynamical timescales are naturally integrated upon. This drastically reduces the range of dynamical times sustained by the system; (ii) The multipole method has a computational complexity and memory imprint that both scale linearly with the total number of stars; (iii) The integration schemes exactly conserve the system's intrinsic geometric constraints, while the symplectic integrator MD2 (see Eq. 2.48) also conserves the total angular momentum vector to round-off errors.

Yet, the algorithm described here deserves further investigations in various directions:

- The present code naturally parallelises over two cores. Further parallelisations are more delicate as they require to parallelise the recurrences from Eq. (3.73). One could rephrase those as traditional prefix sums and use standard parallel-scan algorithms (Ladner & Fischer, 1980). Similarly, to accommodate even wider ranges of dynamical timescales, it would be worthwhile to adapt the algorithm to multi-timesteps schemes.
- Throughout this manuscript, we focused on galactic nuclei. Yet, VRR also occurs in spherically symmetric globular clusters (Meiron & Kocsis, 2018). The associated Hamiltonian is formally identical to Eq. (3.3) and only asks for an appropriate update of the effective mass $\mu(r)$ (see Eq. 2.27).

3.5 Kinetic theory in axisymmetric clusters

Long-range interacting systems, such as galactic nuclei, generically undergo an evolution in two stages. First, a fast (collisionless) violent relaxation (Lynden-Bell, 1967) during which the system reaches a quasistationary state (i.e. a steady state of the collisionless dynamics) dynamically frozen for the mean-field dynamics. Then, as a result of the finite number of particles, the system undergoes a slow (collisional) relaxation that drives it towards its thermodynamical equilibrium. This second stage is generically described by the BL equation (see §A) recently generalised to inhomogeneous systems (Heyvaerts, 2010; Chavanis, 2012). By design, these formalisms account simultaneously for inhomogeneity (i.e. non trivial orbital structures), collective effects (i.e. spontaneous amplification of perturbations) and non-local resonant couplings.

In this section, we present an explicit application of the BL framework to VRR. First, in §3.5.1, in order to simplify the problem at hand, we present our model of an axisymmetric single-population system. In §3.5.2, we derive the appropriate BL equation for that system. Finally, in §3.5.3, we apply this formalism to recover the diffusion coefficients as well as the scaling of the relaxation rate w.r.t. the number of particles.

3.5.1 Axisymmetric limit

We consider a population of N annuli sharing all the same orbital parameter \mathbf{K} . Starting from the generic VRR Hamiltonian from Eq. (3.4), the system's specific Hamiltonian takes the short form

$$H = \sum_{i < j}^N \mu U(\hat{\mathbf{L}}_i \cdot \hat{\mathbf{L}}_j) + \sum_{i=1}^N U_{\text{ext}}(\hat{\mathbf{L}}_i), \quad (3.75)$$

where all stars share the same individual mass $\mu = M_*/N$, with M_* the total stellar mass. Equation (3.75) involves the pairwise interaction potential $U(\widehat{\mathbf{L}}, \widehat{\mathbf{L}}') = U(\widehat{\mathbf{L}} \cdot \widehat{\mathbf{L}}')$, which we develop in Legendre polynomials as

$$U(\widehat{\mathbf{L}} \cdot \widehat{\mathbf{L}}') = - \sum_{\ell} \alpha_{\ell} P_{\ell}(\widehat{\mathbf{L}} \cdot \widehat{\mathbf{L}}'). \quad (3.76)$$

Here, the coefficients α_{ℓ} follow naturally from the coupling coefficients, $\mathcal{H}_{\ell}[\mathbf{K}, \mathbf{K}]$, of Eq. (3.4) and depend on the particular choice of the orbit. Finally, in Eq. (3.75), we introduced $U_{\text{ext}}(\widehat{\mathbf{L}})$ as an imposed external potential, e.g., to represent the Lense–Thirring precession (see, e.g., Merritt, 2013) imposed by the central BH.

The phase space for $\widehat{\mathbf{L}}$ is two-dimensional. A natural choice of (specific) canonical coordinates is then $\mathbf{w} = (\phi, \cos(\vartheta) = u)$ with (ϑ, ϕ) the usual spherical coordinates. The equations of motion of particle i read

$$\frac{d\phi_i}{dt} = \frac{\partial H}{\partial u_i}; \quad \frac{du_i}{dt} = - \frac{\partial H}{\partial \phi_i}. \quad (3.77)$$

Given its generality, the Hamiltonian from Eq. (3.75) encompasses a wide class of long-range interacting systems: (i) $U = -\alpha_1 P_1$ describes classical Heisenberg spins (see, e.g., Gupta & Mukamel, 2011; Barré & Gupta, 2014); (ii) $U = -\alpha_2 P_2$ is the Maier–Saupe model for liquid crystals (Maier & Saupe, 1958; Roupas et al., 2017); (iii) $U = -\sum_{\ell} \alpha_{2\ell} P_{2\ell}$ corresponds to VRR in single-population clusters (e.g., Takács & Kocsis, 2018).

In the coming section, we place ourselves within the axisymmetric limit and assume that the system’s mean-field Hamiltonian is independent of ϕ . In that case, we show how the inhomogeneous BL equation straightforwardly applies to this regime and derive the associated kinetic equation.

3.5.2 Kinetic equation

We assume that the present single-population cluster is characterised by a mean DF, $F(\widehat{\mathbf{L}})$, normalised to $\int d\widehat{\mathbf{L}} F = M_*$. Following Eq. (3.75), the mean specific Hamiltonian of a particle in that system generically reads

$$\begin{aligned} H_0(\widehat{\mathbf{L}}) &= \int d\widehat{\mathbf{L}}' U(\widehat{\mathbf{L}} \cdot \widehat{\mathbf{L}}') F(\widehat{\mathbf{L}}') + U_{\text{ext}}(\widehat{\mathbf{L}}) \\ &= \sum_{\ell} h_{\ell} P_{\ell}(u) + U_{\text{ext}}(u), \end{aligned} \quad (3.78)$$

where we assumed that the system’s DF and external potential were axisymmetric, i.e. $F(\widehat{\mathbf{L}}) = F(u)$ and $U_{\text{ext}}(\widehat{\mathbf{L}}) = U_{\text{ext}}(u)$, and introduced the coefficients $h_{\ell} = -2\pi\alpha_{\ell} \int du P_{\ell}(u) F(u)$. Because $H_0(\widehat{\mathbf{L}}) = H_0(u)$, the Poisson bracket satisfies $[F(u), H(u)] = 0$, i.e. any axisymmetric DF is a steady state for the mean field dynamics. The mean Hamiltonian, $H_0 = H_0(u)$ is also integrable, since the action $J = u$ is conserved along the unperturbed orbits, while the associated angle $\theta = \phi$ is 2π -periodic and evolving with the frequency $\Omega(u) = dH_0/du$.

Investigating the system’s long-term evolution amounts then to describing the slow distortions of the system’s mean DF, $F(u)$, assumed to remain axisymmetric and linearly stable throughout its evolution. Benefiting from the generic framework of long-range kinetic theory (see §A), deriving the kinetic equation for $F(u)$ is immediate. One only needs to proceed by analogies, as we now detail.

The interaction potential from Eq. (3.76) can generically be written under the separable form

$$U(\widehat{\mathbf{L}}, \widehat{\mathbf{L}}') = - \sum_p \psi^{(p)}(\widehat{\mathbf{L}}) \psi^{(p)*}(\widehat{\mathbf{L}}'), \quad (3.79)$$

where the natural potential basis elements are indexed by $p = (\ell, m)$. They read

$$\psi^{(p)}(\widehat{\mathbf{L}}) = C_{\ell} Y_{\ell}^m(\widehat{\mathbf{L}}), \quad (3.80)$$

where $Y_{\ell}^m(\widehat{\mathbf{L}})$ are the complex spherical harmonics, normalised so that $\int d\widehat{\mathbf{L}} Y_{\ell}^m Y_{\ell'}^{m'*} = \delta_{\ell\ell'} \delta_{mm'}$ and we introduced $(C_{\ell})^2 = 4\pi\alpha_{\ell}/(2\ell+1)$. Following Eq. (A.2), it is straightforward to check that the basis elements, $\psi^{(p)}$, form a biorthogonal basis. As defined in Eq. (A.5), their Fourier transform w.r.t. the angle $\theta = \phi$ reads

$$\psi_k^{(p)}(u) = \int \frac{d\phi}{2\pi} e^{-ik\phi} C_{\ell} Y_{\ell}^m = \delta_k^m c_{\ell}^m(u), \quad (3.81)$$

with $c_{\ell}^m(u) \propto P_{\ell}^m(u)$ spelled out explicitly in Fouvry et al. (2019b).

We can now inject Eq. (3.81) into the generic expression of the system’s response matrix (Eq. A.6) to obtain

$$M_{pp'}(\omega) = 2\pi \delta_m^{m'} \int du \frac{m \partial F / \partial u}{\omega - m \Omega(u)} c_{\ell}^{m*}(u) c_{\ell'}^{m'}(u). \quad (3.82)$$

We refer to §A for a discussion of $\mathbf{M}(\omega)$. Following Eq. (A.4), the dressed susceptibility coefficients read

$$\psi_{kk'}^d(u, u', \omega) = -\delta_k^{k'} \sum_{\ell, \ell' \geq |k|} c_\ell^k(u) [\mathbf{I}_k - \mathbf{M}_k(\omega)]_{\ell\ell'}^{-1} c_{\ell'}^{k*}(u), \quad (3.83)$$

where we introduced the matrices

$$[\mathbf{I}_k]_{\ell\ell'} = \delta_\ell^{\ell'}; \quad [\mathbf{M}_k(\omega)]_{\ell\ell'} = \mathbf{M}_{[\ell, k], [\ell', k]}(\omega). \quad (3.84)$$

Having characterised the system's linear response, we can now tailor the BL framework (Eq. A.1) to our case. Indeed, assuming that the system is linearly stable and that its frequency profile is non-degenerate (i.e. $\partial\Omega/\partial u = 0$ only in isolated points), the system's long-term evolution is given by

$$\frac{\partial F(u)}{\partial t} = 2\pi^2 \mu \frac{\partial}{\partial u} \left[\int du' |\psi_{\text{tot}}^d(u, u', \Omega(u))|^2 \delta_D[\Omega(u) - \Omega(u')] \left(\frac{\partial}{\partial u} - \frac{\partial}{\partial u'} \right) F(u) F(u') \right], \quad (3.85)$$

with the total dressed susceptibility coefficients

$$|\psi_{\text{tot}}^d(u, u', \omega)|^2 = 2 \sum_{k \geq 1} k |\psi_{kk}^d(u, u', k\omega)|^2. \quad (3.86)$$

Collective effects can be switched off by imposing $\mathbf{M}(\omega) = 0$, i.e. by replacing the dressed susceptibility coefficients, $\psi_{kk'}^d(u, u', \omega)$, with their bare analogs, $\psi_{kk'}(u, u')$ (Eq. A.7). This gives the inhomogeneous Landau equation (Chavanis, 2013b).

Finally, we recall that Eq. (3.85) can be recast as a Fokker-Planck equation via

$$\frac{\partial F}{\partial t} = -\frac{\partial}{\partial u} [D_1(u) F(u)] + \frac{1}{2} \frac{\partial^2}{\partial u^2} [D_2(u) F(u)], \quad (3.87)$$

with the first- and second-order diffusion coefficients

$$D_1(u) = \frac{1}{2} \frac{\partial D_2}{\partial u} + 2\pi^2 \mu \int du' |\psi_{\text{tot}}^d|^2 \delta_D[\Omega(u) - \Omega(u')] \frac{\partial F}{\partial u'}, \quad (3.88a)$$

$$D_2(u) = (2\pi)^2 \mu \int du' |\psi_{\text{tot}}^d|^2 \delta_D[\Omega(u) - \Omega(u')] F(u'). \quad (3.88b)$$

In practice, for a given value of u , one carries out the integrals $\int du'$ in Eq. (3.88b) by finding the resonant actions u_* satisfying $\Omega(u_*) = \Omega(u)$, which allows for the replacement $\delta_D(\Omega(u) - \Omega(u')) = \sum_{u_*} \delta_D(u' - u_*) / |\partial\Omega/\partial u'|_{u'=u_*}$. Because of its prefactor $\mu = M_*/N$, the BL equation describes the system's long-term self-consistent evolution at first order in $1/N$. Of course, the similarities between Eq. (3.88b) and the SRR diffusion coefficients derived in Eq. (2.7) are obvious: both are archetypes of long-range resonant dynamics. In the present axisymmetric case, we emphasise in Eq. (3.85) that (i) the orbital space is one-dimensional, as visible in the one-dimensional integral $\int du'$; (ii) symmetry limits resonances to $k = k'$, as imposed by the Kronecker symbol in Eq. (3.83).

As a consequence, if ever the system's mean frequency profile, $u \mapsto \Omega(u)$, is monotonic, the resonance condition $\delta_D[\Omega(u) - \Omega(u')]$ only allows for local resonances, i.e. $u' = u$. This immediately leads to a zero flux and $\partial F/\partial t = 0$ in Eq. (3.85). In that case, the system cannot relax under $1/N$ effects and undergoes a kinetic blocking (Chavanis & Lemou, 2007; Chavanis, 2013a; Barré & Gupta, 2014; Lourenço & Rocha Filho, 2015). As such, the system can only relax under weaker finite- N effects associated with higher-order correlations. We refer to Fouvry et al. (2019a, 2020) for an explicit derivation of a $1/N^2$ kinetic equation for $1D$ homogeneous systems with an arbitrary potential of interaction in the absence of collective effects. Still, even if the total flux vanishes, the diffusion coefficients $D_1(u)$ and $D_2(u)$ remain non-zero, and do not suffer from this "conspiracy". Conversely, if the frequency profile is non-monotonic, non-local resonances, $u' \neq u$, are allowed, the flux is non-zero, and the system can relax at order $1/N$. We illustrate these various effects in the coming section.

3.5.3 Application

We refer to Fouvry et al. (2019b) for details on the evaluation of Eq. (3.85) and the associated numerical simulations. We also emphasise that the test scenarios investigated below do not correspond directly to Hamiltonians that are permitted by the astrophysical VRR problem, but are rather generically appropriate for the long-term dynamics of long-range interacting particles on a sphere in the axisymmetric limit.

Following Gupta & Mukamel (2011), we first consider a system driven by interactions of the form

$$U(x) = -\alpha_1 P_1(x); \quad U_{\text{ext}}(u) = D_{\text{ext}} u^2, \quad (3.89)$$

with α_1 and D_{ext} given in Fouvry et al. (2019b). Following Eq. (3.78), we find that $\Omega(u) = dH_0/du$ is a first degree polynomial in u , i.e. the frequency profile is monotonic. As for the DF, we consider the waterbag

$$F(u) = C \Theta[\sin(a) - |u|], \quad (3.90)$$

with $\theta[x]$ the Heaviside function and C a normalisation constant. In practice, we carefully pick $\sin(a)$ so that the system is linearly stable (Gupta & Mukamel, 2011). The gradient of the DF from Eq. (3.90) involves Dirac deltas which makes the computation of the response matrix immediate as one can get rid of the integral from Eq. (3.82) (see §B in Fouvry et al., 2019b). Yet, because of these infinite gradients, the system also supports neutral modes, i.e. modes with a zero growth rate (Chavanis et al., 2005). These lead to localised divergences in the system's dressed diffusion coefficients. In the left panel of Fig. 3.16, we illustrate the prediction for such diverging coefficients as well as measurements from direct N -body simulations (see §C of Fouvry et al., 2019b). Overall, we recover a good agreement between the kinetic theory and the numerical simulations.

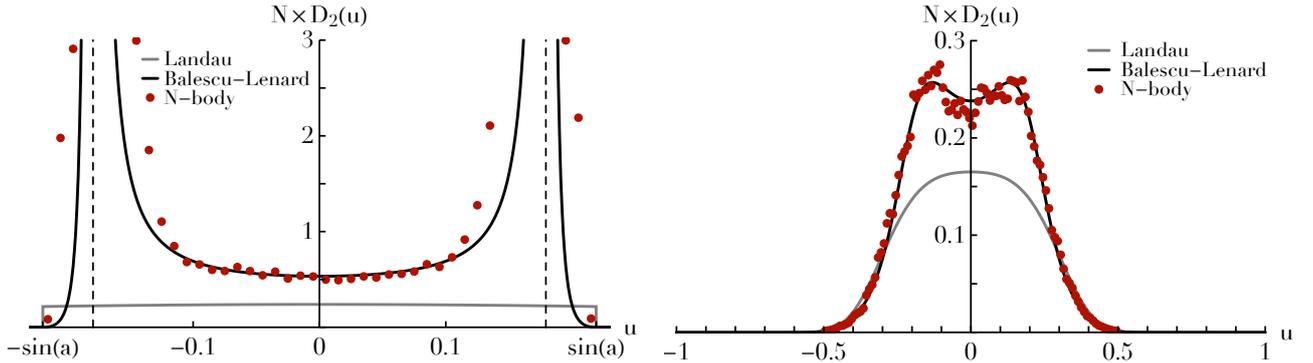


Figure 3.16: From Fouvry et al. (2019b). Second-order diffusion coefficient, $N \times D_2(u)$, as predicted by Eq. (3.88b) in the absence (Landau) or presence (BL) of collective effects and compared with N -body simulations (that naturally include collective effects). Left: For the waterbag DF from Eq. (3.90). As a result of the presence of neutral modes (see §B in Fouvry et al., 2019b), the BL diffusion coefficients locally diverge, as indicated by the vertical dashed lines. Right: Same as the left panel but for the smooth DF from Eq. (3.91).

Keeping the same interactions as in Eq. (3.89), one can avoid neutral modes by considering the smooth DF

$$F(u) = C e^{-(u/\sigma)^4}, \quad (3.91)$$

We confirm in §D of Fouvry et al. (2019b) that such a DF is linearly stable. In the right panel of Fig. 3.16, we illustrate the diffusion coefficients for that system and recover a good agreement.

Glancing back at Eq. (3.85), we argued that a system with a monotonic frequency profile undergoes a kinetic blocking and cannot relax under $1/N$ effects. We illustrate this effect in Fig. 3.17 for the waterbag DF from Eq. (3.90). More precisely, inspired by Lourenço & Rocha Filho (2015), we estimate the dependence of the

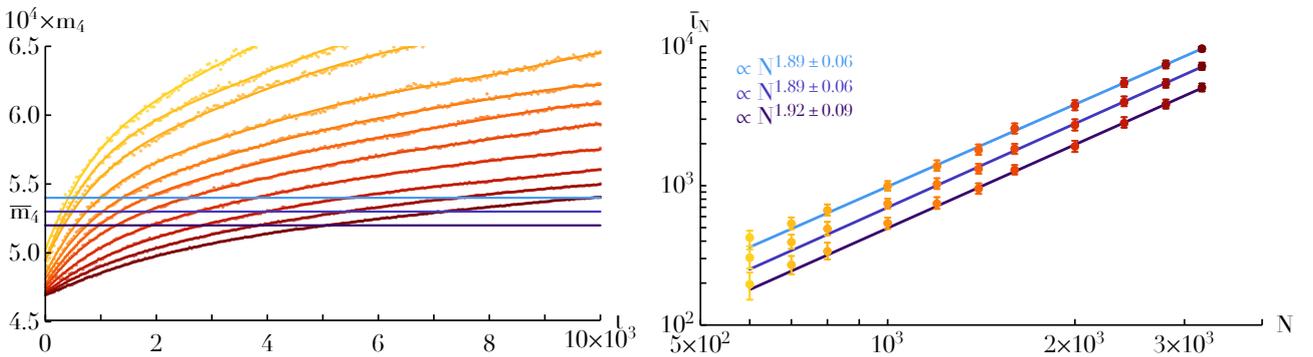


Figure 3.17: From Fouvry et al. (2019b). Dependence of the relaxation rate with the number of particles for the waterbag DF from Eq. (3.90) that has a monotonic frequency profile yielding a kinetic blocking. Left: Time dependence of $m_4(N, t)$ for simulations with $N \in \{6, 7, 8, 10, 12, 14, 16, 20, 24, 28, 32\} \times 10^2$ (from light to dark colors) averaged over 100 realisations (dots) and the associated fits (curves). The horizontal lines represent the values \bar{m}_4 used to measure the crossing times. Right: Dependence of the crossing time \bar{t}_N with the number of particles for different \bar{m}_4 (light to dark colors). Errors bars are estimated by bootstrap resamplings over the available realisations. The lines correspond to power-law fits.

relaxation rate w.r.t. N using the quantity $m_4(N, t) = \{(u - \{u\})^4\}$ with $\{x\}$ the average over all the particles of a given realisation. For a given value of N , the time series of $m_4(N, t)$ is averaged over the available realisations, as illustrated in the left panel of Fig. 3.17. Finally, for a given threshold value \bar{m}_4 , we determine the crossing

time \bar{t}_N such that $m_4(N, \bar{t}_N) = \bar{m}_4$. Should the BL equation (3.85) have a non-vanishing flux, one expects the scaling $\bar{t}_N \propto N$. The dependence of $N \mapsto \bar{t}_N$ for the waterbag DF is illustrated in the right panel of Fig. 3.17. In the range $6 \times 10^2 \leq N \leq 32 \times 10^2$, we measure the scaling $\bar{t}_N \propto N^{1.92 \pm 0.09}$, which is expected to even further converge to $\bar{t}_N \propto N^2$ for larger values of N (Lourenço & Rocha Filho, 2015). The present system indeed suffers from a kinetic blocking: non-local resonant couplings are impossible within a monotonic frequency profile.

In order to recover the $1/N$ scaling predicted by the BL equation, we must consider a model in which higher harmonics ($\ell=3$ or higher) contribute to the pairwise interaction. We consider therefore

$$U(x) = -\alpha_1 P_1(x) - \alpha_3 P_3(x); \quad U_{\text{ext}}(u) = D_{\text{ext}} u^2, \quad (3.92)$$

with α_1 , α_3 and D_{ext} spelled out in Fouvry et al. (2019b). In that case, Eq. (3.78) gives that $\Omega(u)$ is a non-monotonic second degree polynomial in u . As for the axisymmetric DF, we take

$$F(u) = C e^{-(u-u_0)^2/(2\sigma^2)}, \quad (3.93)$$

with σ given in Fouvry et al. (2019b). In the left panel of Fig. 3.18, we illustrate this DF along with its associated density profile. Importantly, we point out that in the vicinity of the DF's maximum, the frequency profile

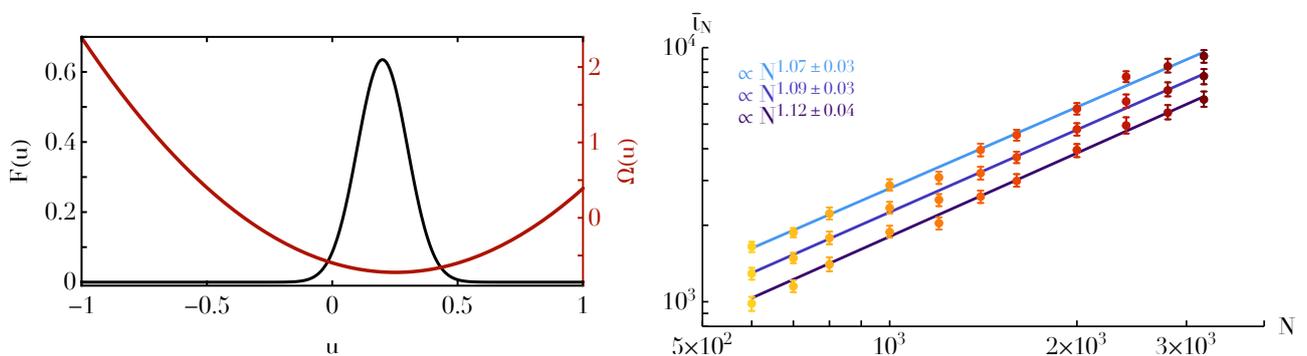


Figure 3.18: From Fouvry et al. (2019b). Left: Illustration of the DF from Eq. (3.93) and the associated non-monotonic frequency profile, $\Omega(u)$. In the region of the DF's maximum, the resonance condition $\Omega(u') = \Omega(u)$ has two solutions, allowing for non-local resonant couplings. Right: Same as in Fig. 3.17, but for the DF from Eq. (3.93) that has a non-monotonic frequency profile, preventing any kinetic blocking.

is non-monotonic allowing for non-local resonances that ultimately drive a non-zero flux in Eq. (3.85). In the right panel of Fig. 3.18, we estimate the scaling of the relaxation rate of the DF from Eq. (3.93) w.r.t. the number of particles. In the range $6 \times 10^2 \leq N \leq 32 \times 10^2$, we find $\bar{t}_N \propto N^{1.1 \pm 0.06}$ in sensible agreement with the prediction from the BL equation. Because this system supports non-local orbital resonances, it relaxes much more efficiently than kinetically-blocked systems.

3.5.4 Prospects

In this section, we focused our interest on describing the long-term process of VRR in the limit of an axisymmetric distribution of orbits. Once one has recognised that this system's evolution equations are formally identical to the ones governing a generic long-range interacting system, the derivation of the associated kinetic equation becomes straightforward. In the present case, the reduced number of dimensions of phase space imposes additional geometric constraints to the system's dynamics, in particular allowing only for resonances $k = k'$. This greatly simplifies the underlying kinetic equation. We illustrated how, in the presence of a monotonic frequency profile, a kinetic blocking occurs and the system cannot relax under $1/N$ effects. This blocking gets lifted with a non-monotonic frequency profile, for which non-local resonant couplings are possible.

As for future developments, we mention:

- In the presence of a monotonic frequency profile, the system is kinetically blocked and its dynamics is only sourced by $1/N^2$ effects. Building upon Fouvry et al. (2020), it would surely be enlightening to derive an appropriate $1/N^2$ kinetic equation describing three-body correlations in $1D$ inhomogeneous systems. Surely, this is no easy endeavour if one also aims at accounting for collective effects.
- The present resonant theory assumes that there does not exist any global resonance condition satisfied by all the orbits. While such an assumption is generically valid in galactic nuclei, it is incorrect within a constant density core where all mean-field orbits are (i) exactly closed; (ii) undergo no precession; and (iii) share the exact same orbital frequency. Such fundamentally degenerate systems deserve further investigations.

4 Conclusion

Galactic nuclei are fascinating self-gravitating systems because of the wealth of dynamical processes they encompass. It is indeed rare to encounter a system where dynamical times range over more than 16 orders of magnitude from a couple of minutes (close to the event horizon) up to a Hubble time (e.g., for the relaxation of semi-major axes). Describing stellar dynamics in galactic nuclei requires therefore to setup appropriate approximations to describe the processes of interest. This was the focus of this manuscript.

In §2, we focused on SRR, i.e. the relaxation of stellar eccentricities. We presented a first-principles derivation of the appropriate kinetic theory (§2.1) emphasising its fundamental connexion with the BL equation. We subsequently applied this framework to SgrA* (§2.2) using the observed thermal distribution of the S-stars' eccentricities to constrain SgrA*'s background stellar cluster. Finally, we devised an efficient multipole integration scheme (§2.3) to simulate the dynamics of Keplerian wires around a supermassive BH, while accounting for this system's key geometric constraints.

In §3, we turned our interest to VRR, i.e. the relaxation of stellar orbital orientations. We first developed a stochastic framework to describe the statistics of the potential fluctuations driving the random walks of the stellar orientations (§3.1). We expanded this formalism to the process of "neighbour separation" (§3.2) highlighting its importance to constrain the survivability of SgrA*'s clockwise stellar disc. We showed how entropy methods can predict the non-trivial thermodynamical equilibria of VRR (§3.3). We also devised an efficient structure-preserving multipole integration scheme to simulate the dynamics of Keplerian annuli around a supermassive BH (§3.4). Finally, leveraging once again the versatility of the BL equation, we developed an explicit kinetic theory for VRR in axisymmetric clusters emphasising these systems' possible kinetic blockings (§3.5).

We conclude this manuscript by mentioning three main venues for future works that could pave the way for important developments in our understanding of stellar dynamics in galactic nuclei.

Kinetic theory

- *Linear response and collective effects.* In §2.1, we characterised SRR in the limit of a thermal background eccentricity distribution. When this limit is not satisfied, galactic nuclei amplify fluctuations via collective effects and source a resonant dynamical friction. What is the importance of linear response?
- *Non-axisymmetry.* In §3.1 and §3.2, we characterised the statistics of VRR in the limit of a spherically symmetric background distribution, while in §3.3, we restricted the entropy method to axisymmetric equilibria. What is the precise impact of non-axisymmetries on VRR?
- *Simultaneous relaxations.* Following the estimates from Eq. (1.20), we treated VRR, SRR, and NR as separate processes. Of course, this is only an assumption and their respective timescale separation is only approximate. What is the impact of joint diffusions in $\hat{\mathbf{L}}$, h , and a when all processes operate?

Numerical methods

- *More efficient schemes.* Even though they offer a complexity scaling linearly in the total number of stars, the multipole schemes presented in §2.3 and §3.4 still suffer from the wide range of timescales present in galactic nuclei. How can one introduce multi-timesteps and further parallelise these methods?
- *Time integration of kinetic equations.* In §2.1 and §3.5, we presented effective resonant kinetic theories. However, we limited ourselves to only computing the associated diffusion coefficients for some given initial conditions. How can one integrate self-consistently these non-linear kinetic equations forward in time?

Application to SgrA*

- *VRR and clockwise stellar disc.* In §3.2, we characterised how efficiently stars with similar orientations diffuse away from another through VRR. How much can one constrain SgrA*'s background stellar content from the observation of SgrA*'s clockwise stellar disc?
- *Leveraging constraints together.* When constraining SgrA*'s stellar cluster, we treated the constraints stemming from SRR and VRR as independent. What are the models of SgrA* that are simultaneously compatible with these seemingly different dynamical constraints?
- *Preparing future observations.* Future observations of SgrA*'s galactic nucleus will offer ever more precise estimates of the PDF, $P(\hat{\mathbf{L}}, h, a)$, of the young stellar cluster. How should one use such detailed statistical information to constrain SgrA*'s cluster and the likely IMBHs therein?

REFERENCES

References

- Aarseth, S. 1967, in *Les Nouvelles Méthodes de la Dynamique Stellaire*, 47
 Abbott, B. P., et al. 2019, *Phys. Rev. X*, 9, 031040
 Alexander, T. 2005, *Phys. Rep.*, 419, 65
 Alexander, T. 2017, *ARA&A*, 55, 17
 Antonini, F., & Merritt, D. 2013, *ApJ*, 763, L10
 Bahcall, J. N., & Wolf, R. A. 1976, *ApJ*, 209, 214
 —. 1977, *ApJ*, 216, 883
 Bar-Or, B., & Alexander, T. 2014, *Class. Quantum Grav.*, 31, 244003
 —. 2016, *ApJ*, 820, 129
 Bar-Or, B., & Fouvry, J.-B. 2018, *ApJ*, 860, L23
 Bar-Or, B., Kupi, G., & Alexander, T. 2013, *ApJ*, 764, 52
 Barré, J., & Gupta, S. 2014, *J. Stat. Mech.*, 2, 02017
 Bartko, H., Martins, F., Fritz, T. K., et al. 2009, *ApJ*, 697, 1741
 Binney, J., & Tremaine, S. 2008, *Galactic Dynamics: Second Edition* (Princeton University Press)
 Blanes, S., Casas, F., Oteo, J. A., & Ros, J. 2009, *Phys. Rep.*, 470, 151
 Bregman, M., & Alexander, T. 2012, *ApJ*, 748, 63
 Chavanis, P.-H. 2006, *Int. J. Mod. Phys. B*, 20, 3113
 —. 2012, *Physica A*, 391, 3680
 —. 2013a, *Eur. Phys. J. B*, 128, 126
 —. 2013b, *A&A*, 556, A93
 Chavanis, P.-H., & Lemou, M. 2007, *Eur. Phys. J. B*, 59, 217
 Chavanis, P.-H., Vatteville, J., & Bouchet, F. 2005, *Eur. Phys. J. B*, 46, 61
 Chen, X., & Amaro-Seoane, P. 2014, *ApJ*, 786, L14
 Cohn, H., & Kulsrud, R. M. 1978, *ApJ*, 226, 1087
 Corrsin, S. 1959, *Atmospheric Diffusion and Air Pollution* (Academic Press)
 Davies, R., et al. 2018, in *Ground-based and Airborne Instrumentation for Astronomy VII*, Vol. 10702, 1070215
 Dehnen, W. 2014, *CompAC*, 1, 1
 Do, T., et al. 2019, *BAAS*, 51, 530
 Duncan, M. J., Levison, H. F., & Lee, M. H. 1998, *AJ*, 116, 2067
 Eilon, E., Kupi, G., & Alexander, T. 2009, *ApJ*, 698, 641
 Eisenhauer, F. 2019, in *The Very Large Telescope in 2030*, 30
 Event Horizon Telescope Collaboration, et al. 2022, *ApJ*, 930, L12
 Fouvry, J.-B. 2022, *MNRAS*, in prep
 Fouvry, J.-B., & Bar-Or, B. 2018, *MNRAS*, 481, 4566
 Fouvry, J.-B., Bar-Or, B., & Chavanis, P.-H. 2019a, *Phys. Rev. E*, 100, 052142
 —. 2019b, *Phys. Rev. E*, 99, 032101
 —. 2019c, *ApJ*, 883, 161
 Fouvry, J.-B., Chavanis, P.-H., & Pichon, C. 2020, *Phys. Rev. E*, 102, 052110
 Fouvry, J.-B., Dehnen, W., Tremaine, S., & Bar-Or, B. 2022, *ApJ*, in press
 Fouvry, J.-B., Pichon, C., & Magorrian, J. 2017, *A&A*, 598, A71
 Fouvry, J.-B., & Prunet, S. 2022, *MNRAS*, 509, 2443
 Fry, J. N., & Peebles, P. J. E. 1980, *ApJ*, 236, 343
 Generozov, A., & Madigan, A.-M. 2020, *ApJ*, 896, 137
 Genzel, R., Eisenhauer, F., & Gillessen, S. 2010, *Rev. Mod. Physics*, 82, 3121
 Ghez, A. M., Duchêne, G., Matthews, K., et al. 2003, *ApJ*, 586, L127
 Ghez, A. M., Salim, S., Weinberg, N. N., et al. 2008, *ApJ*, 689, 1044
 Gillessen, S., Plewa, P. M., Eisenhauer, F., et al. 2017, *ApJ*, 837, 30
 Giral Martínez, J., Fouvry, J.-B., & Pichon, C. 2020, *MNRAS*, 499, 2714
 Gould, A., & Quillen, A. C. 2003, *ApJ*, 592, 935
 Graham, A. W. 2016, in *Astrophysics and Space Science Library*, Vol. 418, *Galactic Bulges* (Springer), 263
 Gravity Collaboration, et al. 2017, *A&A*, 602, A94
 —. 2020, *A&A*, 636, L5
 —. 2021, *A&A*, 645, A127
 Gruzinov, A., Levin, Y., & Zhu, J. 2020, *ApJ*, 905, 11
 Gupta, S., & Mukamel, D. 2011, *J. Stat. Mech.*, 3, 03015
 Gürkan, M. A., & Hopman, C. 2007, *MNRAS*, 379, 1083
 Habibi, M., Gillessen, S., Martins, F., et al. 2017, *ApJ*, 847, 120
 Hairer, E., Lubich, C., & Wanner, G. 2006, *Geometric numerical integration: Second Edition* (Springer)
 Hamers, A. S., Bar-Or, B., Petrovich, C., & Antonini, F. 2018, *ApJ*, 865, 2
 Hamers, A. S., & Portegies Zwart, S. F. 2016, *MNRAS*, 459, 2827
 Hamers, A. S., Portegies Zwart, S. F., & Merritt, D. 2014, *MNRAS*, 443, 355
 Hamilton, C. 2021, *MNRAS*, 501, 3371
 Hamilton, C., Fouvry, J.-B., Binney, J., & Pichon, C. 2018, *MNRAS*, 481, 2041
 Heckman, T. M., & Best, P. N. 2014, *ARA&A*, 52, 589
 Hénon, M. 1964, *Annales d'Astrophysique*, 27, 83
 —. 1973, *A&A*, 24, 229
 Hernquist, L., & Ostriker, J. P. 1992, *ApJ*, 386, 375
 Heyvaerts, J. 2010, *MNRAS*, 407, 355
 Hills, J. G. 1988, *Nature*, 331, 687
 Hopman, C., & Alexander, T. 2006, in *J. Phys. Conf. Ser.*, Vol. 54, 321
 Jacobs, V., & Sellwood, J. A. 2001, *ApJ*, 555, L25
 James, R. W. 1973, *Proc. Royal Soc. Lond.*, 331, 469
 Kalnajs, A. J. 1976, *ApJ*, 205, 745
 Kazandjian, M. V., & Touma, J. R. 2013, *MNRAS*, 430, 2732
 Klein, O. 1924, *Zeitschrift für Physik*, 22, 109
 Kocsis, B., & Tremaine, S. 2011, *MNRAS*, 412, 187
 —. 2015, *MNRAS*, 448, 3265
 Koposov, S. E., Boubert, D., Li, T. S., et al. 2019, *MNRAS*, 491, 2465
 Kormendy, J., & Ho, L. C. 2013, *ARA&A*, 51, 511
 Ladner, R. E., & Fischer, M. J. 1980, *J. ACM*, 27, 831–838
 Levin, Y. 2006, *MNRAS*, 374, 515
 Lisa Collaboration, including Fouvry, J.-B. 2022, arXiv, 2203.06016
 Lourenço, C. R., & Rocha Filho, T. M. 2015, *Phys. Rev. E*, 92, 012117
 Lu, J. R., Ghez, A. M., Hornstein, S. D., et al. 2009, *ApJ*, 697, L463
 Lynden-Bell, D. 1967, *MNRAS*, 136, 101
 Madigan, A.-M., Hopman, C., & Levin, Y. 2011, *ApJ*, 738, 99
 Madigan, A.-M., Levin, Y., & Hopman, C. 2009, *ApJ*, 697, L44
 Magnan, N., Fouvry, J.-B., Pichon, C., & Chavanis, P.-H. 2021, arXiv, 2111.09011
 Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, *AJ*, 115, 2285
 Maier, W., & Saupe, A. 1958, *ZNatA*, 13, 564
 Máthé, G., Szölgvény, Á., & Kocsis, B. 2022, arXiv, 2202.07665
 McGlynn, T. A. 1984, *ApJ*, 281, 13
 McLachlan, R. I., Modin, K., & Verrier, O. 2014, *Phys. Rev. E*, 89, 061301
 Meiron, Y., & Kocsis, B. 2018, *ApJ*, 855, 87
 Meiron, Y., Li, B., Holley-Bockelmann, K., & Spurzem, R. 2014, *ApJ*, 792, 98
 Merritt, D. 2013, *Dynamics and Evolution of Galactic Nuclei* (Princeton Univ. Press)
 —. 2015, *ApJ*, 804, 128
 Merritt, D., Alexander, T., Mikkola, S., & Will, C. M. 2011, *Phys. Rev. D*, 84, 044024
 Merritt, D., Gualandris, A., & Mikkola, S. 2009, *ApJ*, 693, L35
 Milankovitch, M. 1939, *Bull. Serb. Acad. Sci. Math. Nat. A*, 6, 1
 Munthe-Kaas, H. 1999, *Appl. Numer. Math.*, 29, 115
 Murchikova, E. M., Phinney, E. S., Pancoast, A., & Blandford, R. D. 2019, *Nature*, 570, 83
 Murray, C., & Dermott, S. 1999, *Solar System Dynamics* (Cambridge Univ. Press)
 Nelson, R. W., & Tremaine, S. 1999, *MNRAS*, 306, 1
 Neumayer, N., Seth, A., & Böker, T. 2020, *A&A Rev.*, 28, 4
 Panamarev, T., Just, A., Spurzem, R., et al. 2019, *MNRAS*, 484, 3279
 Perets, H. B., Gualandris, A., Kupi, G., Merritt, D., & Alexander, T. 2009, *ApJ*, 702, 884
 Perets, H. B., Hopman, C., & Alexander, T. 2007, *ApJ*, 656, 709
 Polyachenko, E. V., Berczik, P., Just, A., & Shukhman, I. G. 2020, *MNRAS*, 492, 4819
 Portegies Zwart, S. F., & McMillan, S. L. W. 2002, *ApJ*, 576, 899
 Pott, J. U., et al. 2018, in *Ground-based and Airborne Instrumentation for Astronomy VII*, Vol. 10702, 1070290
 Press, W., et al. 2007, *Numerical Recipes 3rd Ed.* (Cambridge Univ. Press)
 Rauch, K. P., & Tremaine, S. 1996, *New Astron.*, 1, 149
 Rein, H., & Spiegel, D. S. 2015, *MNRAS*, 446, 1424
 Risken, H. 1989, *The Fokker-Planck Equation* (Berlin: Springer)
 Romero, A. H., & Sancho, J. M. 1999, *J. Comput. Phys.*, 156, 1
 Rosengren, A. J., & Scheeres, D. J. 2014, *Celest. Mech. Dyn. Astron.*, 118, 197
 Roupas, Z. 2020, *J. Phys. A*, 53, 045002
 Roupas, Z., Kocsis, B., & Tremaine, S. 2017, *ApJ*, 842, 90
 Saha, P. 1993, *MNRAS*, 262, 1062
 Schödel, R., et al. 2017, *A&A*, 609, A27
 Shapiro, S. L., & Marchant, A. B. 1978, *ApJ*, 225, 603
 Sridhar, S., & Touma, J. R. 2016, *MNRAS*, 458, 4143
 Szölgvény, Á., & Kocsis, B. 2018, *Phys. Rev. Lett.*, 121, 101101
 Szölgvény, Á., Máthé, G., & Kocsis, B. 2021, *ApJ*, 919, 140
 Takács, Á., & Kocsis, B. 2018, *ApJ*, 856, 113
 Taylor, J. B., & McNamara, B. 1971, *Physics of Fluids*, 14, 1492
 Tep, K., Fouvry, J.-B., Pichon, C., et al. 2021, *MNRAS*, 506, 4289
 Touma, J., Tremaine, S., & Kazandjian, M. 2019, *Phys. Rev. Lett.*, 123, 021103
 Touma, J. R., Tremaine, S., & Kazandjian, M. V. 2009, *MNRAS*, 394, 1085
 Trefethen, L. N., & Weideman, J. A. C. 2014, *SIAM Review*, 56, 385
 Tremaine, S. 2005, *ApJ*, 625, 143
 —. 2020a, *MNRAS*, 491, 1941
 —. 2020b, *MNRAS*, 493, 2632
 Tremaine, S., Touma, J., & Namouni, F. 2009, *AJ*, 137, 3706
 van Albada, T. S., & van Gorkom, J. H. 1977, *A&A*, 54, 121
 Vasiliev, E. 2015, *MNRAS*, 446, 3150
 —. 2017, *ApJ*, 848, 10
 Villumsen, J. V. 1982, *MNRAS*, 199, 493
 White, S. D. M. 1983, *ApJ*, 274, 53
 Wood, A. T. 1994, *Commun. Stat.-Simul. Comput.*, 23, 157
 Yelda, S., et al. 2014, *ApJ*, 783, 131

A Balescu–Lenard equation

In this Appendix, we reproduce the main results regarding the inhomogeneous BL equation. In recent years, this equation has been derived via various complementary approaches. We mention in particular: (i) Heyvaerts (2010) via the BBGKY hierarchy; (ii) Chavanis (2012) through the Klimontovich equation; (iii) Fouvry & Bar-Or (2018) via Novikov theorem; (iv) Hamilton (2021) via Rostoker’s principle. We do not present any detailed derivation of the BL equation and limit ourselves to presenting it under its most generic form. Once this generic formalism available, it can subsequently be tailored to galactic nuclei as in §2.1 to derive the master equation of SRR around a supermassive BH or as in §3.5 for VRR in axisymmetric galactic nuclei,

We consider a generic Hamiltonian system in $2d$ dimensions and write the phase space coordinates as $\mathbf{w} = (\mathbf{q}, \mathbf{p})$. We consider a system composed of N particles of individual mass $m = M_{\text{tot}}/N$ embedded within a given external potential, $U_{\text{ext}}(\mathbf{w})$, and coupled one to another through a long-range interaction potential, $U(\mathbf{w}, \mathbf{w}')^1$. In the absence of perturbations, the system is assumed to be in an integrable stable steady state for which there exist some angle-action coordinates $\mathbf{w} = (\boldsymbol{\theta}, \mathbf{J})$ (Binney & Tremaine, 2008). Within these coordinates, Jeans’ theorem allows us to write the system’s DF as $F(\mathbf{w}, t) = F(\mathbf{J}, t)$, which we normalise as $\int d\mathbf{w} F = M_{\text{tot}}$. Similarly, the system’s mean Hamiltonian, $H_0(\mathbf{w}) = U_{\text{ext}}(\mathbf{w}) + \int d\mathbf{w}' U(\mathbf{w}, \mathbf{w}') F(\mathbf{w}')$ is integrable. It reads $H_0(\mathbf{w}) = H_0(\mathbf{J})$ with the associated orbital frequencies $\boldsymbol{\Omega}(\mathbf{J}) = \partial H_0 / \partial \mathbf{J}$.

As a result of the finite number of particles, the system’s orbital structure gets slowly distorted. To first order in $1/N$, this dynamics is generically given by the inhomogeneous BL equation

$$\begin{aligned} \frac{\partial F(\mathbf{J})}{\partial t} = & -\pi(2\pi)^d m \frac{\partial}{\partial \mathbf{J}} \cdot \left[\sum_{\mathbf{k}, \mathbf{k}'} \mathbf{k} \int d\mathbf{J}' |\psi_{\mathbf{k}\mathbf{k}'}^{\text{d}}(\mathbf{J}, \mathbf{J}', \mathbf{k} \cdot \boldsymbol{\Omega}(\mathbf{J}))|^2 \right. \\ & \left. \times \delta_{\text{D}}[\mathbf{k} \cdot \boldsymbol{\Omega}(\mathbf{J}) - \mathbf{k}' \cdot \boldsymbol{\Omega}(\mathbf{J}')] \left(\mathbf{k}' \cdot \frac{\partial}{\partial \mathbf{J}'} - \mathbf{k} \cdot \frac{\partial}{\partial \mathbf{J}} \right) F(\mathbf{J}) F(\mathbf{J}') \right]. \end{aligned} \quad (\text{A.1})$$

The BL equation is the master equation describing the irreversible long-term relaxation of inhomogeneous long-range interacting systems driven by resonant couplings between collectively dressed intrinsic Poisson fluctuations. It conserves the total mass, the total energy, and satisfies an H -theorem for Boltzmann entropy (see, e.g., Chavanis, 2012).

Equation (A.1) involves the dressed coupling coefficients, $\psi_{\mathbf{k}\mathbf{k}'}^{\text{d}}(\mathbf{J}, \mathbf{J}', \omega)$, that describe the efficiency with which orbits \mathbf{J} and \mathbf{J}' may resonantly couple to one another through the resonance $(\mathbf{k}, \mathbf{k}')$ at the frequency ω . Importantly, these coefficients also account for the system’s spontaneous collective polarisation, i.e. orbits are dressed. To be computed, these coefficients require the introduction of a biorthogonal basis of potentials and densities ($\psi^{(p)}(\mathbf{w}), \rho^{(p)}(\mathbf{w})$) that we define with the convention (Kalnajs, 1976)

$$\psi^{(p)}(\mathbf{w}) = \int d\mathbf{w}' U(\mathbf{w}, \mathbf{w}') \rho^{(p)}(\mathbf{w}'); \quad \int d\mathbf{w} \rho^{(p)}(\mathbf{w}) \psi^{(q)*}(\mathbf{w}) = -\delta_{pq}. \quad (\text{A.2})$$

Once basis elements at our disposal, the pairwise interaction, $U(\mathbf{w}, \mathbf{w}')$, is cast under the separable form

$$U(\mathbf{w}, \mathbf{w}') = - \sum_p \psi^{(p)}(\mathbf{w}) \psi^{(p)*}(\mathbf{w}'). \quad (\text{A.3})$$

The dressed susceptibility coefficients, $\psi_{\mathbf{k}\mathbf{k}'}^{\text{d}}(\mathbf{J}, \mathbf{J}', \omega)$, then read

$$\psi_{\mathbf{k}\mathbf{k}'}^{\text{d}}(\mathbf{J}, \mathbf{J}', \omega) = - \sum_{p,q} \psi_{\mathbf{k}}^{(p)}(\mathbf{J}) [\mathbf{I} - \mathbf{M}(\omega)]_{pq}^{-1} \psi_{\mathbf{k}'}^{(q)*}(\mathbf{J}'), \quad (\text{A.4})$$

where we introduced the Fourier transform of the basis elements w.r.t. the dynamical angle $\boldsymbol{\theta}$ as

$$\psi_{\mathbf{k}}^{(p)}(\mathbf{J}) = \int \frac{d\boldsymbol{\theta}}{(2\pi)^d} \psi^{(p)}(\mathbf{w}) e^{-i\mathbf{k} \cdot \boldsymbol{\theta}}. \quad (\text{A.5})$$

Equation (A.4) also involves the identity matrix \mathbf{I} and the response matrix

$$M_{pq}(\omega) = (2\pi)^d \sum_{\mathbf{k}} \int d\mathbf{J} \frac{\mathbf{k} \cdot \partial F / \partial \mathbf{J}}{\omega - \mathbf{k} \cdot \boldsymbol{\Omega}(\mathbf{J})} \psi_{\mathbf{k}}^{(p)*}(\mathbf{J}) \psi_{\mathbf{k}}^{(q)}(\mathbf{J}). \quad (\text{A.6})$$

¹For $3D$ self-gravitating systems, one usually has $U_{\text{ext}} = \frac{1}{2} \mathbf{v}^2$ (kinetic energy) and $U(\mathbf{x}, \mathbf{x}') = -G/|\mathbf{x} - \mathbf{x}'|$ (Newtonian interaction).

This matrix fully captures the system's linear response. Because it involves a resonant denominator, its effective numerical implementation is a challenge (see, e.g., Fouvry & Prunet, 2022). A system is said to be linearly unstable if there exists a complex frequency, $\omega = \omega_0 + i\eta$ (with $\eta > 0$), for which $\mathbf{M}(\omega)$ admits an eigenvalue equal to 1. In that case, the system supports an unstable mode of pattern speed ω_0 and growth rate η (see, e.g., §5.3 of Binney & Tremaine, 2008). Phrased differently, $\mathbf{M}(\omega)$ generically describes the system's linear stability around the mean integrable state $F(\mathbf{J})$ through the linearised collisionless Vlasov equation, $\partial\delta F/\partial t + [\delta F, H_0] + [F, \delta H(\delta F)]$ with $[f, h] = \frac{\partial f}{\partial \boldsymbol{\theta}} \cdot \frac{\partial h}{\partial \mathbf{J}} - \frac{\partial f}{\partial \mathbf{J}} \cdot \frac{\partial h}{\partial \boldsymbol{\theta}}$ the Poisson bracket.

In the absence of collective effects², one may neglect $\mathbf{M}(\omega)$ in Eq. (A.4). In that limit, the dressed coupling coefficients, $\psi_{\mathbf{k}\mathbf{k}'}^{\text{d}}(\mathbf{J}, \mathbf{J}', \omega)$, become the (frequency-independent) bare coupling coefficients, $\psi_{\mathbf{k}\mathbf{k}'}(\mathbf{J}, \mathbf{J}')$. These are the Fourier transform of the pairwise interaction w.r.t. the canonical angles (Chavanis, 2013b) and read

$$\psi_{\mathbf{k}\mathbf{k}'}(\mathbf{J}, \mathbf{J}') = \int \frac{d\boldsymbol{\theta}}{(2\pi)^d} \frac{d\boldsymbol{\theta}'}{(2\pi)^d} U(\mathbf{w}, \mathbf{w}') e^{-i(\mathbf{k}\cdot\boldsymbol{\theta} - \mathbf{k}'\cdot\boldsymbol{\theta}')}. \quad (\text{A.7})$$

Equation (A.7) is the generic version of the SRR (resp. VRR) coupling coefficients from Eq. (2.10) (resp. Eq. 3.5).

²This is in particular the case for SRR in galactic nuclei with a thermal distribution of eccentricities, $P(h|a) = 2h$, see §2.1.