

Secular instabilities of Keplerian stellar discs

Karamveer Kaur,¹★ Mher V. Kazandjian,² S. Sridhar¹ and Jihad R. Touma³

¹Raman Research Institute, Sadashivanagar, Bangalore 560 080, India

²Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands

³Department of Physics, American University of Beirut, PO Box 11-0236, Riad El-Solh, Beirut 11097 2020, Lebanon

Accepted 2018 February 12. Received 2018 February 12; in original form 2017 October 28

ABSTRACT

We present idealized models of a razor-thin, axisymmetric, Keplerian stellar disc around a massive black hole, and study non-axisymmetric secular instabilities in the absence of either counter-rotation or loss cones. These discs are prograde mono-energetic waterbags, whose phase-space distribution functions are constant for orbits within a range of eccentricities (e) and zero outside this range. The linear normal modes of waterbags are composed of sinusoidal disturbances of the edges of distribution function in phase space. Waterbags that include circular orbits (*polarcaps*) have one stable linear normal mode for each azimuthal wavenumber m . The $m = 1$ mode always has positive pattern speed and, for polarcaps consisting of orbits with $e < 0.9428$, only the $m = 1$ mode has positive pattern speed. Waterbags excluding circular orbits (*bands*) have two linear normal modes for each m , which can be stable or unstable. We derive analytical expressions for the instability condition, pattern speeds, growth rates, and normal mode structure. Narrow bands are unstable to modes with a wide range in m . Numerical simulations confirm linear theory and follow the non-linear evolution of instabilities. Long-time integration suggests that instabilities of different m grow, interact non-linearly, and relax collisionlessly to a coarse-grained equilibrium with a wide range of eccentricities.

Key words: Galaxy: centre – galaxies: kinematics and dynamics – galaxies: nuclei.

1 INTRODUCTION

Dense clusters of stars orbit massive black holes (MBHs) in galactic nuclei. The best studied cases are the nuclear star clusters of the Milky Way and M31, each of which possesses a low-mass (or Keplerian) stellar disc around the MBH. Since the black hole’s gravity dominates the force on stars, Toomre $Q \gg 1$, so an axisymmetric Keplerian disc is expected to be linearly stable to axisymmetric perturbations on Keplerian orbital time-scales. Even when a disc is stable to all modes on these short time-scales, it may be unstable to modes that grow over the much longer *secular* time-scale of apse precession. Secular instabilities must necessarily be non-axisymmetric with the azimuthal wavenumber $m \neq 0$ (Sridhar & Touma 2016, hereafter *ST1*). A good example is the $m = 1$ instability of counter-rotating discs, which may be applicable to the nuclear disc of M31 (Touma 2002; Kazandjian & Touma 2013). Stellar discs with distribution functions (DFs) even in the angular momentum and empty loss cones (i.e. DF is zero at zero angular momentum) may be unstable to $m = 1$ modes (Tremaine 2005). Mono-energetic discs dominated by nearly radial orbits, could be prone to loss cone instabilities of all m , if there is some amount

of counter-rotating stars (Polyachenko, Polyachenko & Shukhman 2007).

A natural question is the following: can prograde, axisymmetric discs support secular instabilities, even when counter-rotation and loss cone are absent? The answers available in the literature pertain to the stability of razor-thin discs. Tremaine (2001) and Jalali & Tremaine (2012) proved that a Schwarzschild DF is stable to modes of all m in the tight-winding limit. This was generalized by *ST1* who proved that a DF, which is a strictly monotonic function of the angular momentum at fixed semimajor axis (i.e. at fixed Keplerian energy), is stable to modes of all m . However, these results are insufficient to address the general question, which could be relevant to the history of the clockwise disc of young stars at the centre of the Milky Way. If these stars formed in a fragmenting, circular gas disc around the MBH (Levin & Beloborodov 2003), then the initial stellar orbits should have small eccentricities and the same sense of rotation (i.e. no counter-rotation) about the MBH. But Yelda et al. (2014) found that the mean eccentricity of the stellar orbits is $\bar{e} \simeq 0.27$. Is this largish value the result of secular instabilities? The goal of this paper is to present the simplest models of stellar discs orbiting MBHs, whose instabilities can be studied explicitly. This is done by combining analytical methods from *ST1* with numerical simulations derived from Touma, Tremaine & Kazandjian (2009).

In Section 2 the problem is stated within the framework of *ST1*. Using their stability result as a guide we motivate the search for

★ E-mail: karamveer@rri.res.in

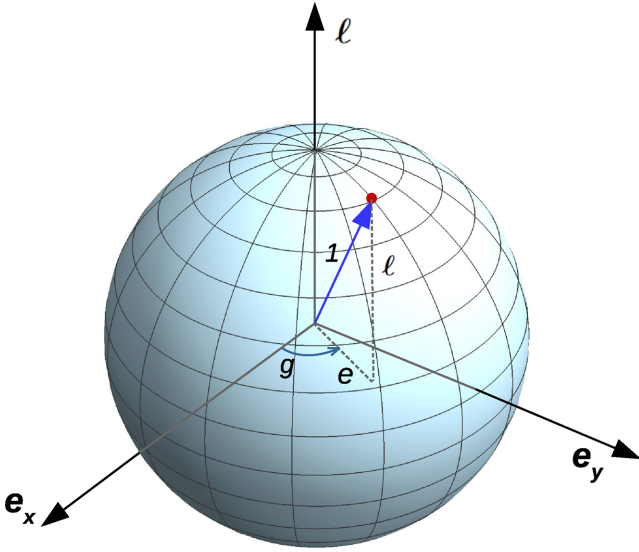


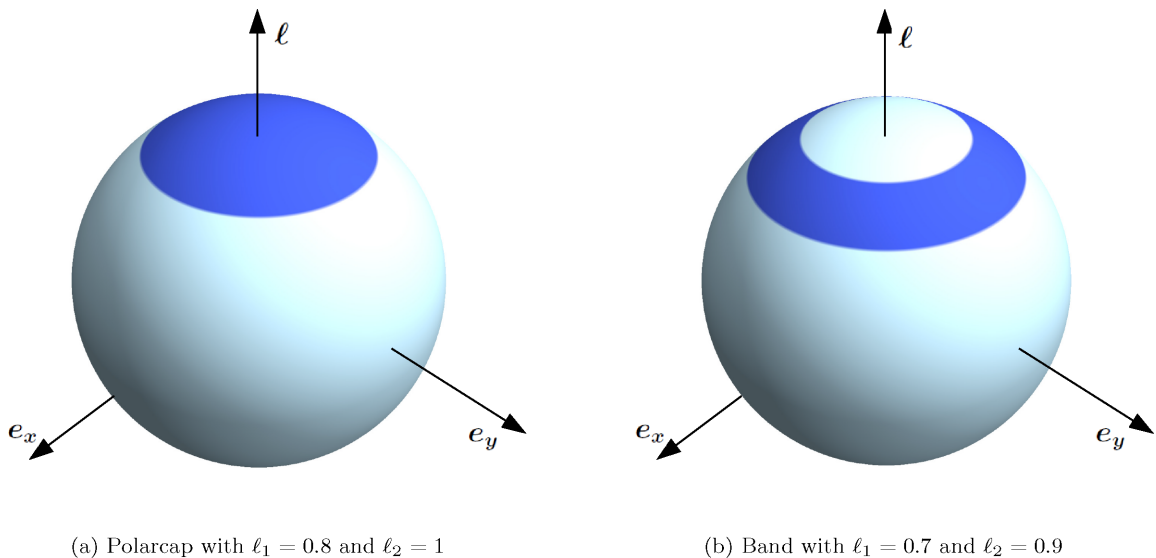
Figure 1. Phase space of a mono-energetic disc. Each star in the disc is represented by point on the unit sphere (shown in red), with canonical coordinates (ℓ, g) . The latitudes are lines of constant ℓ , and longitudes are lines of constant g . The projection of (ℓ, g) onto the equatorial plane gives the eccentricity vector $\mathbf{e} = (e_x, e_y)$.

DFs that are either non-monotonic or not strictly monotonic in the angular momentum. This leads in Section 3 to mono-energetic discs, which are composed of stars with equal semimajor axes. The phase space of a mono-energetic disc is a sphere (see Fig. 1), and secular gravitational interactions between stars have an explicit logarithmic form. Drawing on earlier work in plasma physics we introduce the simplest of prograde, axisymmetric DFs, which correspond to ‘waterbags’. The phase-space distribution function of a waterbag is constant for orbits whose eccentricities (e) lie within a certain range, and zero outside this range. These are of two types of waterbags: *polarcaps*, which include circular orbits, and *bands*, which exclude circular orbits – see Fig. 2. The linear stability analysis of these systems leads to normal modes that are composed of sinusoidal disturbances of the edges of distribution function in the phase space. For each $m \neq 0$, a polarcap has one stable normal mode, whereas

a band has two normal modes that may be stable or unstable. In Section 4 we present numerical simulations of an unstable and a stable band; these give an immediate graphical picture, both in real space and phase space, of linear and non-linear evolution. The linear stability problem for a band is formulated and solved in Section 5. Section 6 explores instabilities further, drawing detailed comparisons between linear theory and numerical simulations, as well as following the long-time evolution of an unstable band. We conclude in Section 7.

2 SECULAR DYNAMICS OF KEPLERIAN STELLAR DISCS

Our model system is a razor-thin flat stellar disc of total mass M , composed of many stars, orbiting an MBH of mass $M_\bullet \gg M$. Since the mass ratio $\varepsilon = M/M_\bullet \ll 1$, the dominant gravitational force on the stars is the inverse-square Newtonian force of the MBH. The limiting case of negligible stellar self-gravity, $\varepsilon \rightarrow 0$, reduces to the problem of each star orbiting the MBH independently on a fixed Keplerian ellipse with period $T_{\text{kep}} = 2\pi(a^3/GM_\bullet)^{1/2}$, where a = semimajor axis. When $0 < \varepsilon \ll 1$, self-gravity is small but its effects build up over the long secular times, $T_{\text{sec}} = \varepsilon^{-1}T_{\text{kep}} \gg T_{\text{kep}}$. **ST1** describes the average behaviour of dynamical quantities over times T_{sec} , by systematically averaging over the fast Keplerian orbital phase – a method that goes back to Gauss. The secular orbit of each star in the disc is represented by a Gaussian ring, which is a Keplerian ellipse with the MBH at one focus, of fixed semimajor axis, whose eccentricity and apsidal longitude can evolve over times T_{sec} . Hence the natural measure of time in secular theory is $\tau = \varepsilon \times \text{time}$, the ‘slow’ time variable. The state of a Gaussian ring at any time τ can be specified by giving its three-dimensional Delaunay coordinates, $\mathcal{R} = \{I, L, g\}$, where $I = \sqrt{GM_\bullet a}$ = constant that is a measure of the Keplerian energy, L is the specific angular momentum that is restricted to the range $-I \leq L \leq I$, and $0 \leq g < 2\pi$ is the longitude of the periapsis. Ring space (or \mathcal{R} -space) is topologically equivalent to \mathbb{R}^3 , with I the ‘radial coordinate’, $\arccos(L/I)$ the ‘colatitude’, and g the ‘azimuthal angle’. A disc composed of $N \gg 1$ stars, each of mass $m_* = M/N$, is a collection of N points in \mathcal{R} -space. The simplest description of a stellar disc uses the single-ring probability DF, $F(\mathcal{R}, \tau) = F(I, L, g, \tau)$, which is normalized



(a) Polarcap with $\ell_1 = 0.8$ and $\ell_2 = 1$

(b) Band with $\ell_1 = 0.7$ and $\ell_2 = 0.9$

Figure 2. Two types of prograde waterbags.

as

$$\int d\mathcal{R} F(\mathcal{R}, \tau) = \int dI dL dg F(I, L, g, \tau) = 1. \quad (1)$$

Over times much shorter than the resonant relaxation times, $T_{\text{res}} = NT_{\text{sec}}$, the graininess of the ring–ring interactions has negligible effects and the stellar system can be thought of as collisionless. Formally, the collisionless limit corresponds to assuming that the system is composed of an infinite number of stars, each of infinitesimal mass, the whole having a mass M equal to the total stellar mass: $N \rightarrow \infty$, $m_* \rightarrow 0$ with $M = Nm_*$ held constant. Then each star is like a test ring, whose motion is governed by the secular Hamiltonian, $\Phi(I, L, g, \tau)$, which is equal to the (scaled) self-gravitational disc potential:¹

$$\Phi(I, L, g, \tau) = \int dI' dL' dg' \Psi(I, L, g, I', L', g') \times F(I', L', g', \tau), \quad (2)$$

where

$$\Psi(I, L, g, I', L', g') = -GM_* \oint \oint \frac{dw}{2\pi} \frac{dw'}{2\pi} \frac{1}{|\mathbf{r} - \mathbf{r}'|} \quad (3)$$

is the (scaled) interaction potential between two rings.² Ring orbits are determined by the Hamiltonian equations of motion:

$$I = \sqrt{GM_* a} = \text{constant}, \quad \frac{dL}{d\tau} = -\frac{\partial\Phi}{\partial g}, \quad \frac{dg}{d\tau} = \frac{\partial\Phi}{\partial L}. \quad (4)$$

This is a Hamiltonian flow in \mathcal{R} -space that is restricted to the $I = \text{constant}$ two-sphere. The flow carries with it the DF, whose evolution is governed by the secular collisionless Boltzmann equation (CBE):

$$\frac{\partial F}{\partial \tau} + [F, \Phi]_{L,g} = 0, \quad \text{where } [F, \Phi]_{L,g} = \frac{\partial F}{\partial g} \frac{\partial \Phi}{\partial L} - \frac{\partial F}{\partial L} \frac{\partial \Phi}{\partial g} \quad (5)$$

is the two-dimensional Poisson bracket in (L, g) -space. Φ itself depends on F through the \mathcal{R}' -space integral of equation (2). Therefore equation (5), together with the secular Hamiltonian of equation (2), defines the self-consistent initial value problem of the secular time evolution of the DF, given an arbitrarily specified initial DF $F(I, L, g, 0)$. A general property of this time evolution is the following: since the I of any ring is constant in time, the probability for a ring to be in $(I, I + dI)$ is a conserved quantity. In other words the probability distribution function in one-dimensional I -space, defined by

$$P(I) = \int dL dg F(I, L, g, \tau), \quad (6)$$

is independent of τ , as can be verified directly using the CBE of equation (5).

2.1 Axisymmetric equilibria and linear stability

Secular equilibria are DFs that are time-independent and self-consistent solutions of the CBE. They can be constructed using the

secular Jeans theorem of [ST1](#), which states that F must be function of the isolating integrals of motion of the secular Hamiltonian. An axisymmetric equilibrium DF is independent of g and can be written as $F = (2\pi)^{-1} F_0(I, L)$, because I and L are two isolating integrals of motion of the axisymmetric Hamiltonian, $\Phi_0(I, L)$. Equation (2) gives Φ_0 self-consistently in terms of F_0 :³

$$\Phi_0(I, L) = \int dI' dL' F_0(I', L') \oint \frac{dg'}{2\pi} \Psi(I, L, g, I', L', g'). \quad (7)$$

The equations of motion (4) for a ring become very simple in an axisymmetric disc:

$$I = \text{constant}, \quad L = \text{constant}, \quad \frac{dg}{d\tau} \equiv \Omega_0(I, L) = \frac{\partial\Phi_0}{\partial L}. \quad (8)$$

The semimajor axis and eccentricity of a ring are constant, with the apsidal longitude precessing at the constant angular frequency $\Omega_0(I, L)$.

The time evolution of perturbations to an axisymmetric equilibrium DF can be studied by considering the total DF to be $F = (2\pi)^{-1} F_0(I, L) + F_1(I, L, g, \tau)$, where the perturbation F_1 contains no net mass:

$$\int dI dL dg F_1(I, L, g, \tau) = 0. \quad (9)$$

If $\Phi_1(I, L, g, \tau)$ is the self-gravitational potential due to F_1 , then the total Hamiltonian is $\Phi = \Phi_0(I, L) + \Phi_1(I, L, g, \tau)$. By substituting for F and Φ in the CBE (equation 5), and using $[F_0, \Phi_0]_{L,g} = 0$, we can derive the equation governing the time evolution of F_1 . For small perturbations $|F_1| \ll F_0$ this is the linearized collisionless Boltzmann equation (LCBE):

$$\frac{\partial F_1}{\partial \tau} + \Omega_0 \frac{\partial F_1}{\partial g} = \frac{1}{2\pi} \frac{\partial F_0}{\partial L} \frac{\partial \Phi_1}{\partial g}, \quad (10a)$$

$$\Phi_1(I, L, g, \tau) = \int dI' dL' dg' \Psi(I, L, g, I', L', g') \times F_1(I', L', g', \tau). \quad (10b)$$

The LCBE is a linear (partial) integro-differential equation for F_1 , and determines the linear stability of the axisymmetric DF, $F_0(I, L)$.

An axisymmetric perturbation $F_1(I, L, \tau)$ gives rise to a $\Phi_1(I, L, \tau)$ that is also independent of g . Then the LCBE (equation 10a) implies $\partial F_1 / \partial \tau = 0$, whose physical solution is $F_1 = 0$, because an axisymmetric perturbation cannot change the angular momentum of a star. Hence it is only non-axisymmetric, or g -dependent, perturbations that are of interest in secular theory. Since τ and g appear in the LCBE only as $(\partial/\partial\tau)$ and $(\partial/\partial g)$ we can look for linear modes of the form $F_1 \propto \exp[i(mg - \omega\tau)]$, where $m \neq 0$ is the azimuthal wavenumber. Using only the general symmetric properties of $\Psi(\mathcal{R}, \mathcal{R}')$, the following result was proved in [ST1](#) for DFs that are strictly monotonic functions of L .

Stationary, axisymmetric discs with DFs $F_0(I, L)$ are neutrally stable (i.e. ω is real) to secular perturbations of all m when $\partial F_0 / \partial L$ is of the same sign (either positive or negative) everywhere in its domain of support, $-I \leq L \leq I$ and $I_{\text{min}} \leq I \leq I_{\text{max}}$.

As noted in [ST1](#) these secularly stable DFs can have both prograde and retrograde populations of stars because $-I \leq L \leq I$. The discs have net rotation and include physically interesting cases, such as a secular analogue of the well-known Schwarzschild DF. To investigate secular instabilities, the above stability result motivates

³ $\Psi(I, L, g, I', L', g')$ depends on the apses only in the combination $|g - g'|$, so the integral over g' is independent of g .

¹ [ST1](#) include relativistic effects of the MBH and tidal forces due to external gravitational fields, but these are not considered in this paper.

² Here $\mathbf{r} = (x, y)$ and $\mathbf{r}' = (x', y')$ are the position vectors of the two stars with respect to the MBH – see section 4.1 of [ST1](#) for details of the transformation from \mathbf{r} and \mathbf{r}' to the corresponding Delaunay variables. Here w and w' are the mean anomalies of the stars representing the Keplerian orbital phase on their respective Gaussian rings.

us to look at axisymmetric discs with DFs, $F_0(I, L)$, that are either non-monotonic or not strictly monotonic functions of L at fixed I .

A general way to proceed would be to develop stability theory, using only the symmetry properties of $\Psi(\mathcal{R}, \mathcal{R}')$, as ST1 did. But the goal of this paper is more specific: We wish to construct the simplest class of disc models that permits quantitative study of the onset and growth of linear non-axisymmetric instabilities. In order to do this we must be able to calculate physical quantities such as the apse precession frequency $\Omega_0(I, L)$, using equations (7) and (8). Hence we need to use explicit forms for Ψ , for a physically motivated model of a stellar disc.

3 MONO-ENERGETIC DISCS

3.1 Collisionless Boltzmann equation

$\Psi(I, L, g, I', L', g')$ depends on the apses only in the combination $|g - g'|$, and can be developed in a Fourier series in $(g - g')$. When the spread in the semimajor axes of the disc stars is comparable to the mean disc radius, the Fourier coefficients are, in general, complicated functions of (I, L, I', L') – although for numerical calculations it is straightforward to calculate them on any grid in this four-dimensional space. Analytical forms are readily available if restrictions are placed on L and L' , such as both the rings being near-circular and well separated (the ‘Laplace–Lagrange’ limit of planetary dynamics) or both rings being very eccentric, corresponding to the ‘spoke’ limit of Polyachenko et al. (2007). But secular dynamics and statistical mechanics are really about the exchange of angular momentum of stars at fixed semimajor axes, so it seems preferable if we do not place such severe restrictions on L or L' . Let us consider discs with a small spread in semimajor axes; since this is equivalent to a small spread in Keplerian orbital energies, the disc may be called nearly mono-energetic. Having nearly the same semimajor axes, any two rings either cross each other or come very close to each other, so $\Psi(\mathcal{R}, \mathcal{R}')$ can be large, even infinite, in magnitude. For nearly circular rings the dominant contribution, which is a logarithmic singularity, was worked out by Borderies, Goldreich & Tremaine (1983).

In a nearly mono-energetic disc most pairs of rings intersect each other. It is useful to consider the strictly mono-energetic limit, $I = I_0 = \sqrt{GM_\bullet a_0}$, when every ring intersects every other ring. Since all rings have the same semimajor axis a_0 , they also have the same Keplerian orbital period, $T_{\text{kep}} = 2\pi(a_0^3/GM_\bullet)^{1/2}$. Hence it is convenient to use a dimensionless slow time variable, $t = \tau/T_{\text{kep}} = \text{time}/T_{\text{sec}}$, to study the dynamics of mono-energetic discs. The state of a ring at time t can be specified by giving its periape, g , and the dimensionless specific angular momentum $\ell = L/I_0$. Since $-1 \leq \ell \leq 1$, the motion of any ring is restricted to the unit sphere (Fig. 1) on which $\ell = \cos(\text{colatitude})$ and $g = \text{azimuthal-angle}$ are canonical coordinates. For a mono-energetic disc F takes the form

$$F(I, L, g, \tau) = \frac{\delta(I - I_0)}{I_0} f(\ell, g, t). \quad (11)$$

Then equation (1) implies the following normalization for f :

$$\int d\ell dg f(\ell, g, t) = 1. \quad (12)$$

Hence $f(\ell, g, t)$ is the (dimensionless) DF for mono-energetic discs on the (ℓ, g) phase space of Fig. 1. The eccentricity of a ring, $e = \sqrt{1 - \ell^2}$, is equal to the length of the projection of the corresponding position vector on the sphere’s equatorial plane. The eccentricity vector (or Lenz vector) is defined as $\mathbf{e} = (e_x, e_y)$ with $e_x = e \cos g$

and $e_y = e \sin g$. We can think of (e_x, e_y, ℓ) as a right-handed Cartesian coordinate system, with the ring phase space realized as the unit sphere, $e_x^2 + e_y^2 + \ell^2 = 1$.

The formula of Borderies et al. (1983) for the ring–ring interaction potential, $\psi(\ell, \ell', g - g') = \Psi(I_0, I_0\ell, g, I_0, I_0\ell', g')$, takes the following attractive form given in Touma & Tremaine (2014):

$$\psi(\ell, \ell', g - g') = \frac{GM_\bullet}{a_0} \left\{ -\frac{4}{\pi} \log 2 + \frac{1}{2\pi} \log |e - e'|^2 \right\}. \quad (13)$$

This expression for ψ is, strictly speaking, valid only when $e, e' \ll 1$. But Touma & Tremaine (2014) have shown that this formula for ψ serves as a good approximation for all values of e and e' , and used this fact to study axisymmetric and non-axisymmetric secular thermodynamic equilibria; they also provide an improved fitting formula but we do not use this. Henceforth we take equation (13) as the basic ‘law of interaction’, between any two rings in a mono-energetic disc. Using equation (11) in equation (2) we see that the mean-field self-gravitational potential $\varphi(\ell, g, t) = \Phi(I_0, I_0\ell, g, \tau)$ is given in explicit form as

$$\begin{aligned} \varphi(\ell, g, t) &= \int d\ell' dg' \psi(\ell, \ell', g - g') f(\ell', g', t) \\ &= -\frac{4GM_\bullet}{\pi a_0} \log 2 + \frac{GM_\bullet}{2\pi a_0} \int d\ell' dg' \log |e - e'|^2 \\ &\quad \times f(\ell', g', t). \end{aligned} \quad (14)$$

We have already cast the independent variables (ℓ, g, t) in dimensionless form. Equations (4), governing the dynamics of a ring, can now be written in the following dimensionless form:

$$\frac{d\ell}{dt} = -\frac{\partial H}{\partial g}, \quad \frac{dg}{dt} = \frac{\partial H}{\partial \ell}, \quad (15)$$

where

$$\begin{aligned} H(\ell, g, t) &= \frac{T_{\text{kep}}}{I_0} \varphi(\ell, g, t) \\ &= \int d\ell' dg' \log |e - e'|^2 f(\ell', g', t) + \text{constant} \end{aligned} \quad (16)$$

is the dimensionless secular Hamiltonian. These equations of motion imply the natural Poisson Bracket on the (ℓ, g) unit sphere:

$$[f, H] = \frac{\partial f}{\partial g} \frac{\partial H}{\partial \ell} - \frac{\partial f}{\partial \ell} \frac{\partial H}{\partial g}. \quad (17)$$

Substituting equation (11) in equation (5) we obtain the following CBE governing the self-consistent evolution of the DF:

$$\frac{\partial f}{\partial t} + [f, H] = 0. \quad (18)$$

Equations (16)–(18) provide a complete, dimensionless description of the collisionless dynamics of mono-energetic Keplerian discs.

3.2 Linear stability of axisymmetric equilibria

In the study of axisymmetric equilibria and their linear, non-axisymmetric perturbations it is useful to have at hand the Fourier expansion of the ring–ring interaction potential, $\log |e - e'|^2$, that appears in the definition of the Hamiltonian in equation (16). From equation (C.2) of Touma & Tremaine (2014) we have

$$\begin{aligned} \log |e - e'|^2 &= \log [e^2 - 2ee' \cos(g - g') + e'^2] \\ &= \log(e_{>}^2) - 2 \sum_{m=1}^{\infty} \frac{1}{m} \left(\frac{e_{<}}{e_{>}} \right)^m \cos [m(g - g')], \end{aligned} \quad (19)$$

where $e_{<} = \min(e, e')$ and $e_{>} = \max(e, e')$.

Any DF of the form $f = (2\pi)^{-1} f_0(\ell)$, which is normalized as $\int_{-1}^1 d\ell f_0(\ell) = 1$, represents an axisymmetric equilibrium. Using equation (19) in equation (16), we have the corresponding axisymmetric Hamiltonian:

$$\begin{aligned} H_0(\ell) &= \int_{-1}^1 d\ell' \log(e_{>}^2) f_0(\ell') \\ &= \int_0^{|\ell|} d\ell' \log(1 - \ell'^2) \{f_0(\ell') + f_0(-\ell')\} \\ &\quad + \log(1 - \ell^2) \int_{|\ell|}^1 d\ell' \{f_0(\ell') + f_0(-\ell')\}, \end{aligned} \quad (20)$$

where we have dropped a constant term. The apse precession frequency is given as

$$\Omega_0(\ell) = \frac{dH_0}{d\ell} = -\frac{2\ell}{1 - \ell^2} \int_{|\ell|}^1 d\ell' \{f_0(\ell') + f_0(-\ell')\}. \quad (21)$$

Some general properties of Ω_0 are (i) since the product $\ell \Omega_0(\ell) \leq 0$, the apse precession of a ring is always opposite to the faster Keplerian orbital motion; (ii) as $\ell \rightarrow 0$ we have $\Omega_0(\ell) \rightarrow -2\ell$, so highly eccentric rings precess very slowly; (iii) in the limit of circular rings $\ell \rightarrow \pm 1$, and $\Omega_0(\ell) \rightarrow \mp \{f_0(1) + f_0(-1)\}$ goes to a finite limit.

When the axisymmetric equilibrium is perturbed the total DF is $f(\ell, g, t) = (2\pi)^{-1} f_0(\ell) + f_1(\ell, g, t)$, and the corresponding self-consistent Hamiltonian is $H_0(\ell) + H_1(\ell, g, t)$. Substituting these in the mono-energetic CBE (equation 18) and linearizing, we obtain the LCBE governing the evolution of f_1 :

$$\frac{\partial f_1}{\partial t} + \Omega_0(\ell) \frac{\partial f_1}{\partial g} = \frac{1}{2\pi} \frac{df_0}{d\ell} \frac{\partial H_1}{\partial g}, \quad (22)$$

where

$$H_1(\ell, g, t) = \int d\ell' dg' \log|e - e'|^2 f_1(\ell', g', t). \quad (23)$$

We seek solutions of the form $f_1(\ell, g, t; m) = \text{Re}\{f_{1m}(\ell) \exp[i(mg - \omega_m t)]\}$ and $H_1(\ell, g, t) = \text{Re}\{H_{1m}(\ell) \exp[i(mg - \omega_m t)]\}$, where, without loss of generality, we take m to be a positive integer. Equation (23) gives $H_{1m} = -2\pi/m \int_{-1}^1 d\ell' (e_{<}/e_{>})^m f_{1m}(\ell')$. Then the LCBE reduces to the following equation:

$$[\omega_m - m\Omega_0(\ell)] f_{1m}(\ell) = \frac{df_0}{d\ell} \int_{-1}^1 d\ell' \left(\frac{e_{<}}{e_{>}}\right)^m f_{1m}(\ell'), \quad (24)$$

which is an integral eigenvalue problem, for the eigenvalues ω_m and corresponding eigenfunctions $f_{1m}(\ell)$. This equation is a special case of equation (75) of ST1, which is valid for a general disc. Proceeding in a manner similar to ST1, it is straightforward to prove the stability result: *all DFs $f_0(\ell)$ that are strictly monotonic functions of ℓ are linearly stable*. This raises again the question of the stability of DFs that are not strictly monotonic in ℓ . Since this question is now posed in the context of equation (24) – which is given in explicit form – we can proceed to explore it quantitatively. Among all the DFs that are not strictly monotonic functions of ℓ , the simplest are probably the ‘waterbag’ DFs that are discussed below.

3.3 Waterbags and the linear stability problem

A mono-energetic waterbag is a region of the unit sphere phase space of Fig. 1 within which the DF takes a constant positive value

and is zero outside this region.⁴ Time evolution that is governed by the CBE of equations (16)–(18) conserves both the area of the region and the value of the DF. Hence the dynamical problem reduces to following the evolution of the contour(s) bounding the region. Analogous to the contour dynamics of fluid vortices on a sphere (Dritschel 1988), the deformation of the contour(s) defining a waterbag stellar disc can be very complicated.

3.3.1 Axisymmetric equilibria

An axisymmetric mono-energetic waterbag has a DF, $f_0(\ell)$, that takes a constant positive value for $\ell \in [\ell_1, \ell_2]$, and is zero outside this interval. Since our primary interest in this paper concerns the stability of discs in which stars orbit the MBH in the same sense, we assume that $0 \leq \ell_1 < \ell_2 \leq 1$. The normalized DF for such a ‘prograde waterbag’ is

$$f_0(\ell) = \begin{cases} \frac{1}{\ell_2 - \ell_1} & \text{for } \ell_1 \leq \ell \leq \ell_2, \\ 0 & \text{otherwise.} \end{cases} \quad (25)$$

There are two different cases, corresponding to $\ell_2 = 1$ (polarcap) and $\ell_2 < 1$ (band) – see Fig. 2. It can be seen that bands have DFs that are non-monotonic in ℓ , whereas polarcaps have DFs that are not strictly monotonic in ℓ . Hence the stability result, stated below equation (24), does not apply to either of these systems. But their stability properties can be determined completely, as we show below.

The waterbag DF describes a circular annular disc composed of stars with eccentricities $e = \sqrt{1 - \ell^2} \in [e_2, e_1]$, where $e_i = \sqrt{1 - \ell_i^2}$ for $i = 1, 2$. The inner and outer radii of the disc $r_{\min} = a_0(1 - e_1)$ and $r_{\max} = a_0(1 + e_1)$ are determined by the most eccentric rings in the disc. The normalized surface density profile, $\Sigma_0(r)$, is obtained by integrating $f_0(\ell)$ over the velocities, as is done in Appendix A. This gives

$$\Sigma_0(r) = \begin{cases} \frac{\sin^{-1}[\ell_2/\ell_0(r)] - \sin^{-1}[\ell_1/\ell_0(r)]}{2\pi^2 a_0^2 (\ell_2 - \ell_1)}, & |r - a_0| \leq a_0 e_2, \\ \frac{\cos^{-1}[\ell_1/\ell_0(r)]}{2\pi^2 a_0^2 (\ell_2 - \ell_1)}, & a_0 e_2 < |r - a_0| \leq a_0 e_1, \\ 0, & a_0 e_1 < |r - a_0|, \end{cases} \quad (26)$$

where $\ell_0(r) = \sqrt{2r/a_0 - r^2/a_0^2}$. Surface density profiles are plotted in Fig. 3(a) for the polarcap and band of Fig. 2, and also a broad-band ($\ell_1 = 0.1, \ell_2 = 0.9$), whose stability is studied later. We note that the $\Sigma_0(r)$ profiles of a polarcap and a band are very different: the former has a single maximum at the centre of the disc, whereas the latter has a characteristic double-horned shape.

The apse precession frequency $\Omega_0(\ell)$ can be determined by using equation (25) in equation (21). For a polarcap,

$$\Omega_0(\ell) = \begin{cases} -\frac{2\ell}{(1 - \ell^2)}, & 0 \leq |\ell| \leq \ell_1, \\ -\frac{2\ell}{(1 + |\ell|)(1 - \ell_1)}, & \ell_1 < |\ell| \leq 1, \end{cases} \quad (27)$$

⁴ The ‘waterbag’ model was originally developed for the Vlasov equation by Berk & Roberts (1970).

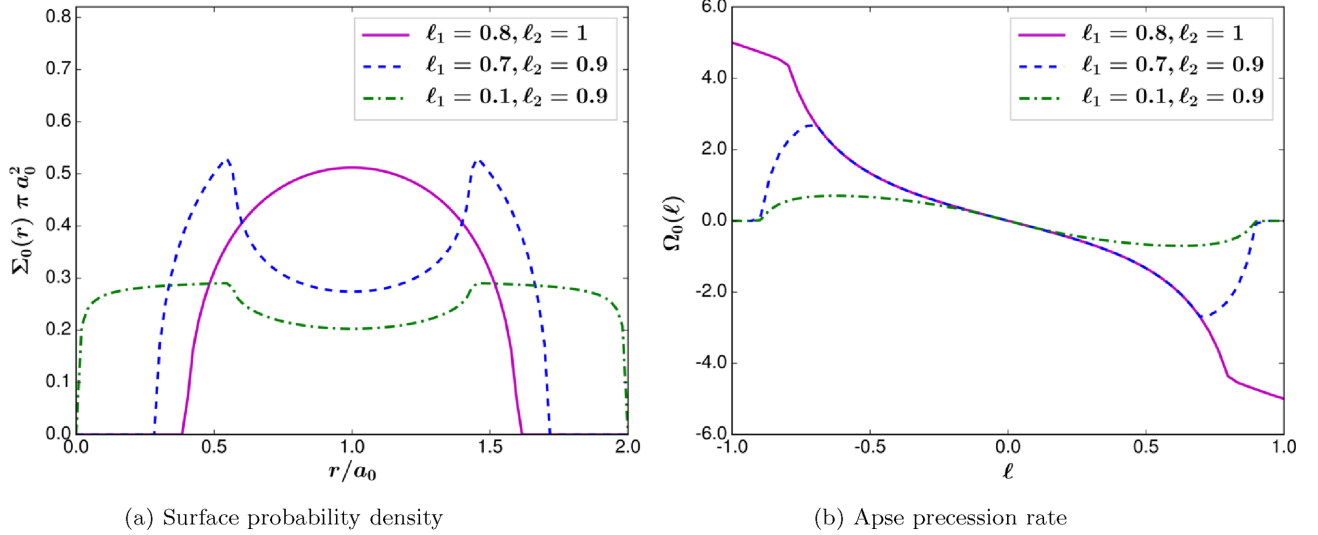


Figure 3. Physical features of waterbags: solid and dashed lines are for the polarcap and band of Fig. 2, respectively. The broken dashed line is for a broad-band, to be studied later.

and for a band,

$$\Omega_0(\ell) = \begin{cases} -\frac{2\ell}{(1-\ell^2)}, & 0 \leq |\ell| \leq \ell_1, \\ -\frac{2\ell}{(1-\ell^2)} \left(\frac{\ell_2 - |\ell|}{\ell_2 - \ell_1} \right), & \ell_1 < |\ell| \leq \ell_2, \\ 0, & \ell_2 < |\ell| \leq 1. \end{cases} \quad (28)$$

Even though the waterbag itself occupies only the interval $[\ell_1, \ell_2]$ we calculate $\Omega_0(\ell)$ for all $\ell \in [-1, 1]$, because it gives the apse precession frequency of any test ring that may be introduced into the system. Ω_0 is an antisymmetric function of ℓ , as can be seen in Fig. 3(b). For a polarcap Ω_0 is non-zero when $\ell = \pm 1$, whereas for a band $\Omega_0(\ell)$ vanishes for all $|\ell| > \ell_2$.

3.3.2 Stability to non-axisymmetric modes

An arbitrary collisionless perturbation of a waterbag can be described as a deformation of its boundaries. From Fig. 2 we see that a polarcap has just one boundary at $\ell = \ell_1$ whereas a band has two boundaries, at $\ell = \ell_1$ and $\ell = \ell_2$. Non-axisymmetric perturbations of the boundaries can be resolved as a Fourier series in the apsidal longitude g . Fig. 4 shows an $m = 3$ deformation of the polarcap and band of Fig. 2, where m is the azimuthal wavenumber of perturbation. *Polarcaps* are linearly stable to all non-axisymmetric modes. In order to prove this we note that, for a polarcap, $df_0/d\ell = (1 - \ell_1)^{-1} \delta(\ell - \ell_1)$. Substituting this in the integral equation (24) we obtain

$$[\omega_m - m\Omega_0(\ell)] f_{1m}(\ell) = \frac{\delta(\ell - \ell_1)}{1 - \ell_1} \int_{-1}^1 d\ell' \left(\frac{e_{<}}{e_{>}} \right)^m f_{1m}(\ell'), \quad (29)$$

where $\Omega_0(\ell)$ is given by equation (27). The physical solution is $f_{1m}(\ell) = A_m \delta(\ell - \ell_1)$, where A_m is a complex amplitude. Using this in equation (29) we obtain the eigenvalue,

$$\omega_m = m\Omega_0(\ell_1) + \frac{1}{1 - \ell_1}. \quad (30)$$

Since ω_m is real for all $m = 1, 2, \dots$ and $0 \leq \ell_1 < 1$, all normal modes are stable and purely oscillatory. For each m there is a normal

mode with

$$f_1(\ell, g, t; m) = \text{Re} \{ A_m \delta(\ell - \ell_1) \exp [im(g - \lambda_P t)] \}, \quad (31)$$

where

$$\lambda_P(m, \ell_1) = \frac{\omega_m}{m} = -\frac{2\ell_1}{(1 - \ell_1^2)} + \frac{1}{m(1 - \ell_1)} \quad (32)$$

is the precession frequency of the m -lobed, sinusoidal deformation of the polarcap boundary. The first term on the right-hand side is just the apse precession frequency in the unperturbed polarcap, and is negative. The second term comes from the self-gravity of the deformation, which is positive. The competition between these two terms results in the following interesting features of $\lambda_P(m, \ell_1)$, as can be seen in Fig. 5.

(i) For a polarcap with given ℓ_1 , λ_P is a decreasing function of m . This is because the self-gravity of the deformed edge is smaller for bigger m , due to mutual cancellation from its lobes and dips. In the limit $m \rightarrow \infty$ this vanishes altogether and $\lambda_P \rightarrow \Omega_0(\ell_1)$.

(ii) The $m = 1$ mode always has prograde precession, with $\lambda_P = 1/(1 + \ell_1)$.

(iii) Modes with $m = 2, 3, \dots$ precess in a prograde sense for $0 \leq \ell_1 < 1/(2m - 1)$, and in a retrograde sense for $1/(2m - 1) < \ell_1 \leq 1$. λ_P vanishes when a polarcap is such that $\ell_1 = 1/(2m - 1)$ for some m ; then it has a stationary time-independent deformation with m lobes.

(iv) For $\ell_1 > 1/3$, only the $m = 1$ mode has positive pattern speed.

Bands have richer stability properties because, for each m , there are two normal modes (as shown in Section 5). Each of these is composed of sinusoidal disturbances of the two edges of phase-space DF – see the lower panels of Fig. 4 for a representation of an $m = 3$ mode. For bands $df_0/d\ell = \{\delta(\ell - \ell_1) - \delta(\ell - \ell_2)\}/\Delta\ell$, where $\Delta\ell = (\ell_2 - \ell_1)$. Substituting this in equation (24) we obtain the following integral equation:

$$[\omega_m - m\Omega_0(\ell)] f_{1m}(\ell) = \frac{\delta(\ell - \ell_1) - \delta(\ell - \ell_2)}{\Delta\ell} \int_{-1}^1 d\ell' \left(\frac{e_{<}}{e_{>}} \right)^m \times f_{1m}(\ell'), \quad (33)$$

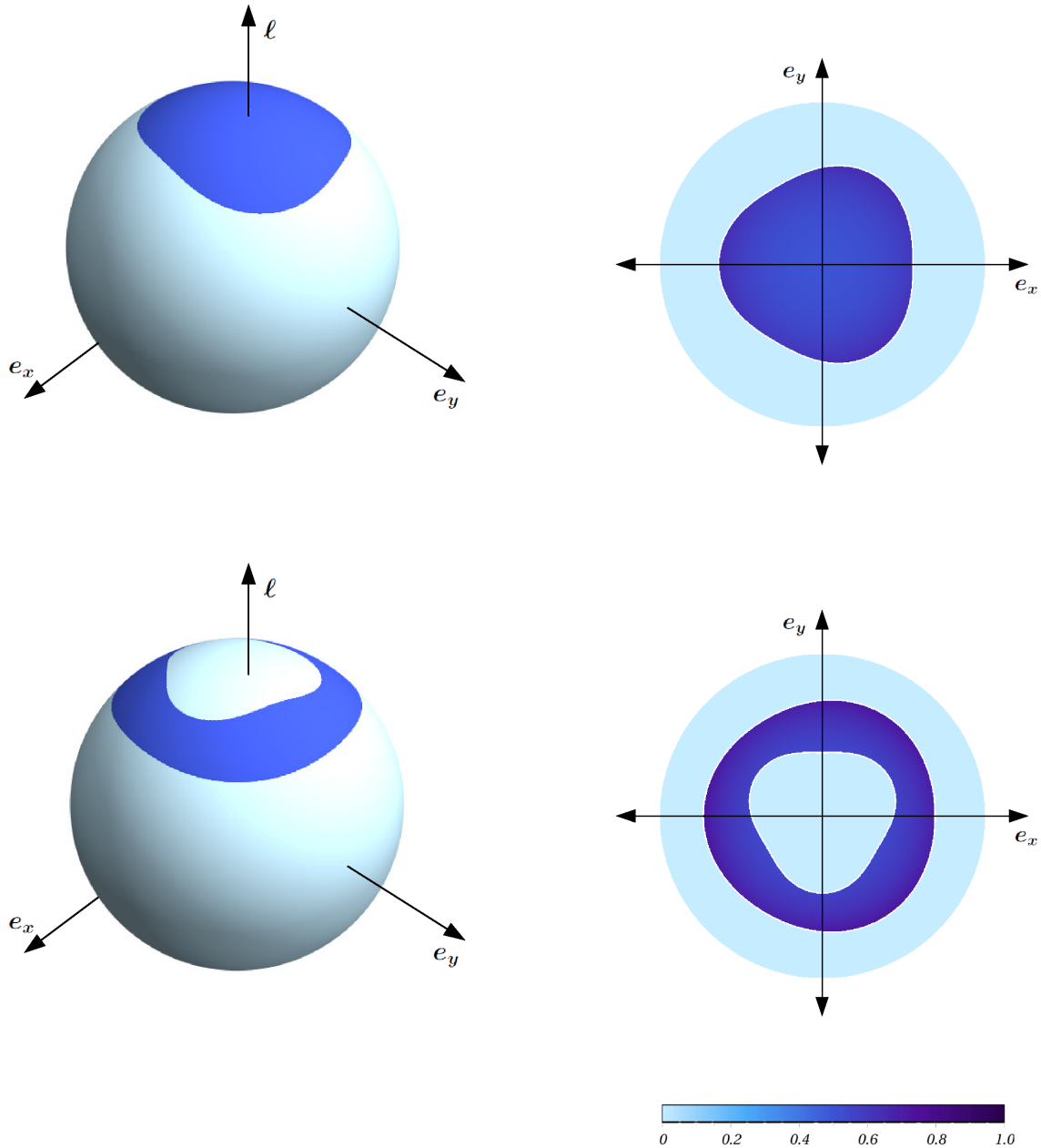


Figure 4. $m = 3$ normal mode for polarcap and band. The panels on the left show the deformed polarcap (upper panel) and band (lower panel) DFs. The panels on the right are for the corresponding probability densities, $n(e_x, e_y) = \ell^{-1} \times \text{DF}$, in the (e_x, e_y) plane. Since the DF is constant within the deformed boundaries, $n \propto 1/\sqrt{1-e^2}$.

where $\Omega_0(\ell)$ is given by equation (28). Hence the eigenfunctions are of the form

$$f_{lm}(\ell) = A_{m1} \delta(\ell - \ell_1) + A_{m2} \delta(\ell - \ell_2), \quad (34)$$

where A_{m1} and A_{m2} are complex amplitudes. When equation (34) for $f_{lm}(\ell)$ is substituted in equation (33) the integral equation reduces to a 2×2 matrix eigenvalue problem. This is the simplest linear stability problem in secular dynamics that can be studied analytically in detail – see Section 5. Before doing this we present numerical simulations of an unstable band and a stable band, so the reader may have an immediate picture of the time evolution going beyond the linear evolution of small disturbances.

4 NUMERICAL EXPLORATION OF WATERBAG STABILITY

We performed N -ring numerical simulations of waterbag bands, for a range of system parameters (ℓ_1, ℓ_2) . The full list is given in Table 1. The last entry has $\ell_2 = 1$, so is a polarcap and not a band. It is included in the table as a limiting case of a class of broad-bands. Here we discuss the stability of the two bands whose $\Sigma_0(r)$ and $\Omega_0(\ell)$ profiles feature in Fig. 3: one is the band `waterbag_1_s0` with $(\ell_1 = 0.7, \ell_2 = 0.9)$, and the other is the broad-band `waterbag_2_s0` with $(\ell_1 = 0.1, \ell_2 = 0.9)$.

We simulate a planar system of N rings, each of which has the same semimajor axis a_0 and mass m_* , orbiting an MBH of mass M_* . The total disc mass $M = Nm_*$ is chosen to be much smaller than

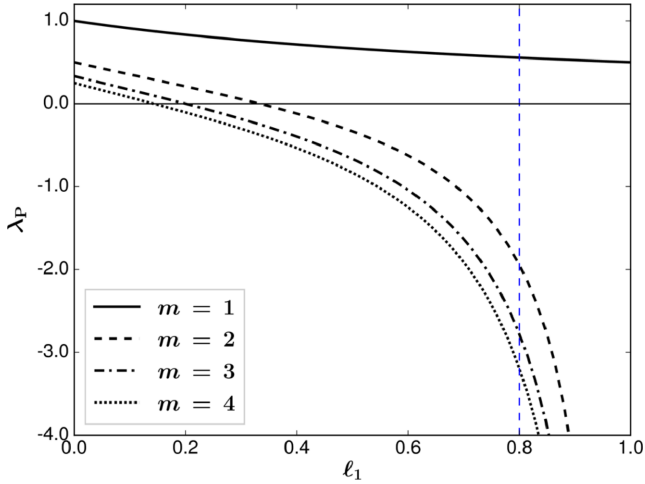


Figure 5. Precession frequency of normal modes of polarcaps. The intersections of the vertical dashed line with the λ_P curves gives the spectrum of the normal modes of the polarcap of Fig. 2. Only the $m = 1$ normal mode has positive precession for all values of ℓ_1 .

Table 1. List of all numerical simulations. The upper five cases correspond to Set I and the lower ones to Set II. The total duration of each simulation, T_{end} , is given in units of Gyr; it is of order a few secular times and differs from case to case.

System name	ℓ_1	ℓ_2	T_{end}	Stable?
waterbag_1_s0	0.7	0.9	2.5	No
waterbag_2_s0	0.1	0.9	9.4	Yes
waterbag_3_s0	0.8	0.9	10.0	No
waterbag_4_s0	0.85	0.9	6.17	No
waterbag_5_s0	0.7	0.97	8.79	Yes
waterbag_ℓ1_0.8_ℓ2_0.81	0.8	0.81	1.8	No
waterbag_ℓ1_0.8_ℓ2_0.82	0.8	0.82	10.0	No
waterbag_ℓ1_0.8_ℓ2_0.83	0.8	0.83	12.5	No
waterbag_ℓ1_0.8_ℓ2_0.84	0.8	0.84	13.3	No
waterbag_ℓ1_0.8_ℓ2_0.85	0.8	0.85	1.65	No
waterbag_ℓ1_0.8_ℓ2_0.86	0.8	0.86	34.2	No
waterbag_ℓ1_0.8_ℓ2_0.87	0.8	0.87	0.28	No
waterbag_ℓ1_0.8_ℓ2_0.88	0.8	0.88	5.9	No
waterbag_ℓ1_0.8_ℓ2_0.89	0.8	0.89	5.9	No
waterbag_ℓ1_0.8_ℓ2_0.90	0.8	0.90	41.2	No
waterbag_ℓ1_0.8_ℓ2_0.91	0.8	0.91	20.0	No
waterbag_ℓ1_0.8_ℓ2_0.92	0.8	0.92	10.8	No
waterbag_ℓ1_0.8_ℓ2_0.93	0.8	0.93	6.4	No
waterbag_ℓ1_0.8_ℓ2_0.94	0.8	0.94	44.0	No
waterbag_ℓ1_0.8_ℓ2_0.95	0.8	0.95	38.7	No
waterbag_ℓ1_0.8_ℓ2_0.96	0.8	0.96	18.4	No
waterbag_ℓ1_0.8_ℓ2_0.97	0.8	0.97	5.1	No
waterbag_ℓ1_0.8_ℓ2_0.98	0.8	0.98	211	Yes
waterbag_ℓ1_0.8_ℓ2_0.99	0.8	0.99	16.3	Yes
waterbag_ℓ1_0.8_ℓ2_1.00	0.8	1.00	19.0	Yes

M_* , so $\varepsilon = M/M_* \ll 1$ and the secular time-scale, $T_{\text{sec}} = \varepsilon^{-1} T_{\text{kep}}$, is much longer than the Kepler orbital period. Each ring can be thought of as a point on the unit sphere phase space of Fig. 1, with coordinates (ℓ^i, g^i) for $i = 1, 2, \dots, N$. The projection of the points on to the equatorial plane gives N eccentricity vectors, $\mathbf{e}^i = e^i (\cos g^i \hat{x} + \sin g^i \hat{y})$, where $e^i = \sqrt{1 - (\ell^i)^2}$ is the eccentricity.

Then the normalized secular energy of the whole system is

$$\mathcal{H} = \frac{1}{N} \sum_{\substack{i,j \\ j>i}} \log |\mathbf{e}^i - \mathbf{e}^j|^2, \quad (35)$$

which serves as the N -ring Hamiltonian for secular dynamics on the sphere:

$$\frac{dg^i}{dt} = \frac{\partial \mathcal{H}}{\partial \ell^i}, \quad \frac{d\ell^i}{dt} = -\frac{\partial \mathcal{H}}{\partial g^i} \quad (\text{for } i = 1, 2, \dots, N), \quad (36)$$

where $t = \text{time}/T_{\text{sec}}$ is, as earlier, the dimensionless time variable. The Hamiltonian equations can be rewritten compactly as

$$\frac{d\mathbf{e}^i}{dt} = \frac{2}{N} \sum_{\substack{j=1 \\ j \neq i}}^N \frac{(\mathbf{e}^i - \mathbf{e}^j) \times \ell^j}{|\mathbf{e}^i - \mathbf{e}^j|^2}, \quad (37)$$

where $\ell^i = \ell^i \hat{z}$. These vectorial equations are similar to those presented in Touma et al. (2009), with the difference that our interaction Hamiltonian is unsoftened and logarithmic. The equations have been solved using a Bulirsch–Stoer integrator, with relative and absolute tolerances equal to 10^{-8} . Our fiducial system has the following parameters.

- (i) The disc is composed of $N = 1000$ rings.
- (ii) Semimajor axis of each ring is $a_0 = 1$ pc.
- (iii) Black hole mass $M_* = 10^7 M_\odot$, giving a Kepler orbital period $T_{\text{kep}} = 0.03$ Myr.
- (iv) Disc mass $M = 10^3 M_\odot$, so $\varepsilon = 10^{-4}$ and the secular time-scale $T_{\text{sec}} = 0.3$ Gyr.

The typical relative energy and angular momentum errors for the simulations listed in Table 1 are $\sim 10^{-6}$.

The evolution of the two bands, waterbag_1_s0 and waterbag_2_s0, is shown in Figs 6 and 7, respectively. The upper two panels are for the surface mass density in the x - y plane, and the lower two panels show the rings represented as 1000 points on the (e_x, e_y) plane.⁵ We begin with initial conditions corresponding to the two bands of Fig. 3. The following overall features can be noticed.

- (i) For waterbag_1_s0 a non-axisymmetric $m = 3$ instability grows; it is seen very clearly around 0.3 Gyr and, by ~ 0.6 Gyr, there are distinct signs of non-linear evolution.
- (ii) In contrast the broad-band waterbag_2_s0 is seen to be stable over a time-scale of 5 Gyr.

Dynamical behaviour can be characterized in more detail by looking at mode amplitudes, $a_m(t)$, which were evaluated by computing fast Fourier transforms over annuli of the projected mass density. These are plotted in Fig. 8(a) for waterbag_1_s0 and Fig. 8(b) for waterbag_2_s0. The main features are as follows.

- (i) For waterbag_1_s0 the initially unstable mode has $m = 3$, and this remains dominant until about 0.6 Gyr. Later there is growth of other modes, especially, $m = 1$ and $m = 2$.

(ii) Modes of all m maintain a low amplitude for waterbag_2_s0. We note that sampling noise, which is unavoidable in the initial conditions, was such that a $m = 2$ mode had a greater initial amplitude than the other modes (see Fig. 8b). The $m = 2$ mode is seen to be stable and precessing in Fig. 7. Interactions of some stars with the $m = 2$ mode has, presumably, scattered them in phase space. Whereas a study of this mode–particle scattering is

⁵ Since we are dealing with prograde discs, all the points have positive ℓ^i .

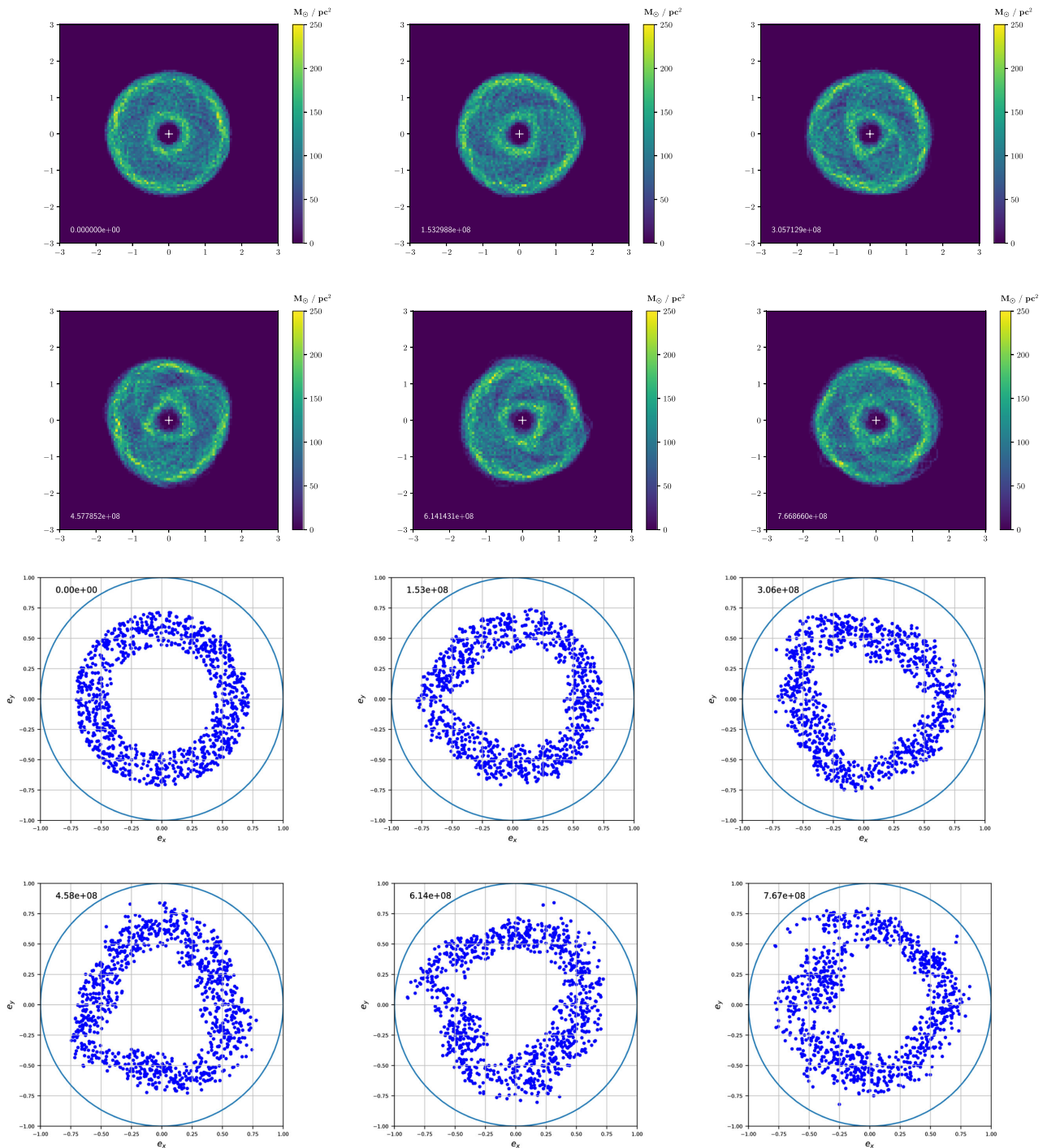


Figure 6. Evolution of the unstable band `waterbag_1_s0`. Upper two rows show the surface density in real space (with distances measured in parsec), and the lower two rows show the distribution in the eccentricity plane at the same respective time. The $m = 3$ mode is clearly visible as three overdensity lumps in the surface density plots and as a triangular feature in the eccentricity plane. Note that the time (in years) is indicated within the subfigures.

beyond the scope of this paper, simulations with a larger number of particles will help clarify the nature of this process.

In the next section we present a detailed account of the linear stability of bands. We will also discuss how linear theory accounts for the behaviour of `waterbag_1_s0` and `waterbag_2_s0`.

5 LINEAR STABILITY OF BANDS

A normal mode of a waterbag band has the form $f_1(\ell, g, t; m) = \text{Re}\{f_{1m}(\ell)\exp[i(mg - \omega_m t)]\}$, where ω_m is a complex eigenfrequency. Since a normal mode is composed of sinusoidal disturbances of the two edges of the phase-space DF, the corresponding eigenfunction is of the form, $f_{1m}(\ell) = A_{m1} \delta(\ell - \ell_1) + A_{m2} \delta(\ell - \ell_2)$, where A_{m1} and A_{m2} are complex amplitudes – see equation (34).

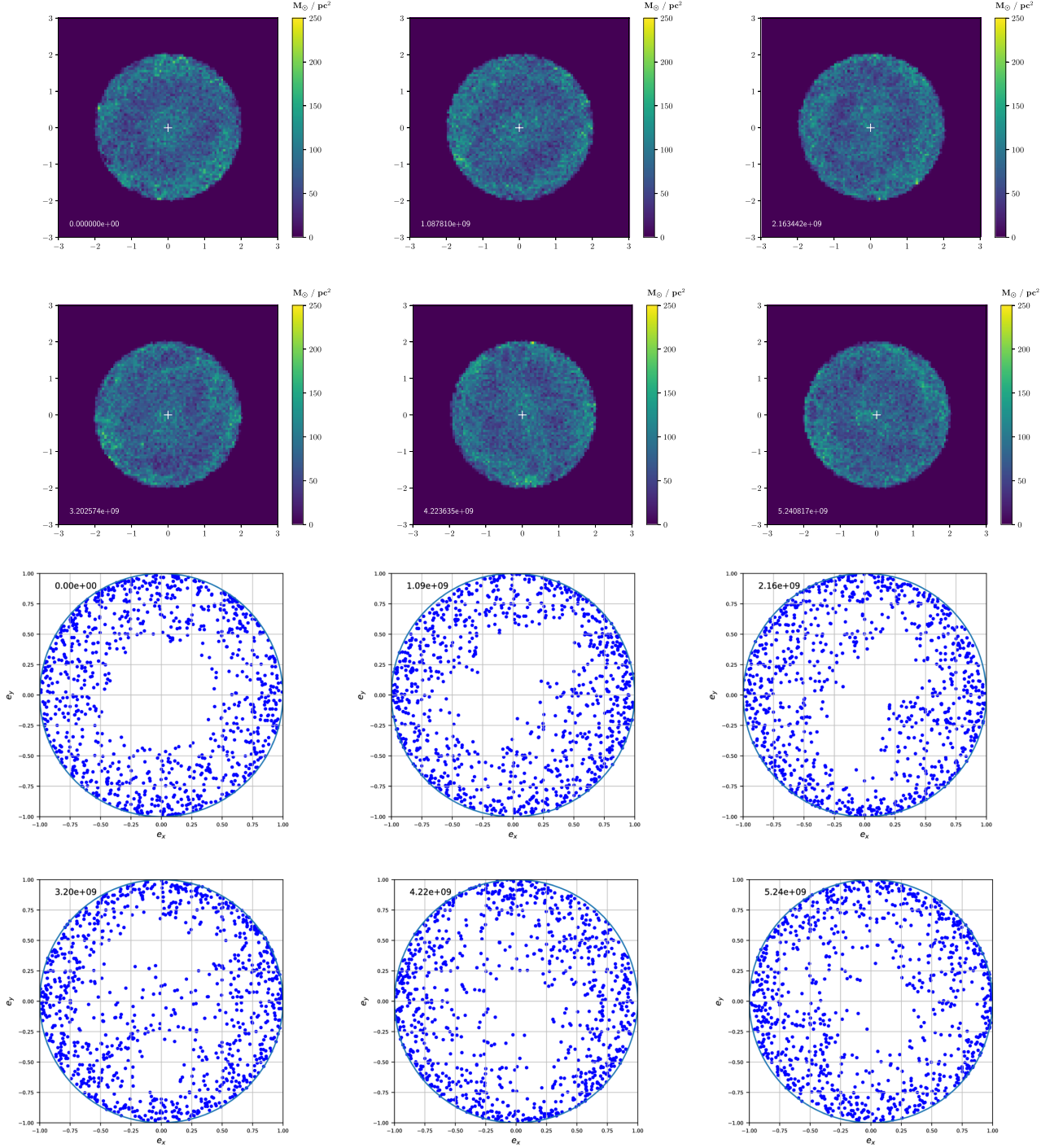


Figure 7. Evolution of the stable broad-band `waterbag_2_s0`. Upper two rows show the surface density in real space (with distances measured in parsec), and the lower two rows show the distribution in the eccentricity plane at the same respective time. Note that the time (in years) is indicated within the subfigures.

When this is substituted in the integral equation (33), it reduces to the following 2×2 matrix eigenvalue problem:

$$\begin{pmatrix} \frac{1}{\Delta\ell} + m\Omega_0(\ell_1) & \frac{1}{\Delta\ell} \left(\frac{e_2}{e_1}\right)^m \\ -\frac{1}{\Delta\ell} \left(\frac{e_2}{e_1}\right)^m & -\frac{1}{\Delta\ell} + m\Omega_0(\ell_2) \end{pmatrix} \begin{pmatrix} A_{m1} \\ A_{m2} \end{pmatrix} = \omega_m \begin{pmatrix} A_{m1} \\ A_{m2} \end{pmatrix}. \quad (38)$$

Here $\Delta\ell = (\ell_2 - \ell_1)$, and equation (28) gives $\Omega_0(\ell_1) \equiv \Omega_1 = -2\ell_1/(1 - \ell_1^2)$ and $\Omega_0(\ell_2) = 0$. The solutions for the eigenfrequency and the ratio of edge disturbance amplitudes are

$$\omega_m^\pm = \frac{m\Omega_1}{2} \pm \frac{1}{\Delta\ell} \sqrt{\left[1 + \frac{m\Delta\ell\Omega_1}{2}\right]^2 - \left(\frac{e_2}{e_1}\right)^{2m}}, \quad (39a)$$

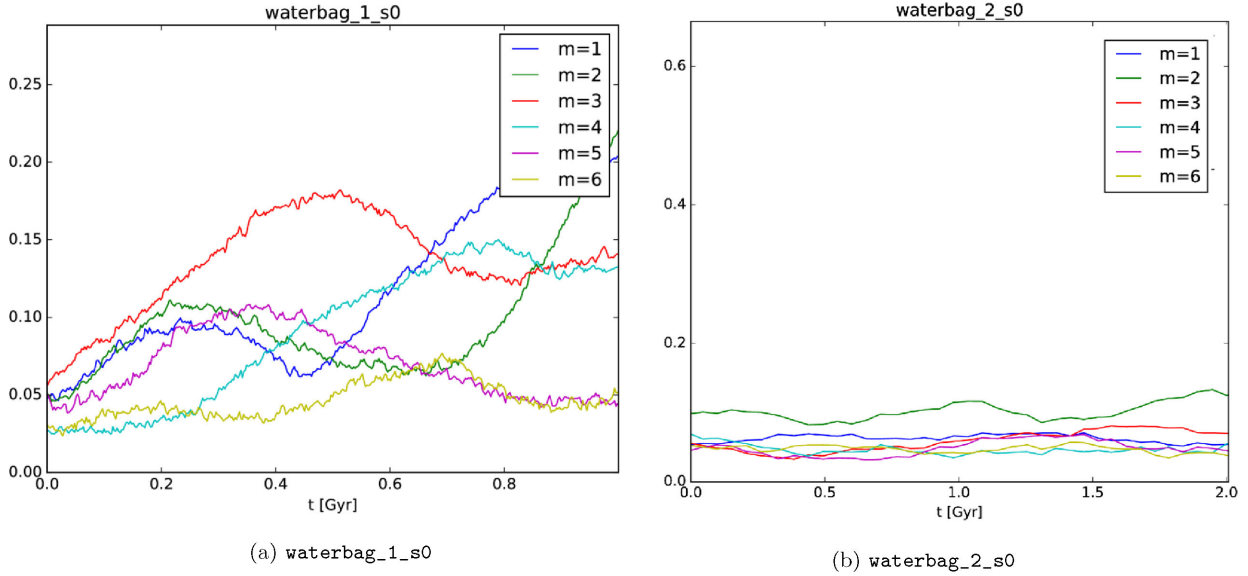


Figure 8. Evolution of mode amplitudes $a_m(t)$.

$$\left(\frac{A_{m2}}{A_{m1}}\right)^\pm = -\left[1 + \frac{m \Delta \ell \Omega_1}{2}\right] \left(\frac{e_1}{e_2}\right)^m \pm \sqrt{\left[1 + \frac{m \Delta \ell \Omega_1}{2}\right]^2 \left(\frac{e_1}{e_2}\right)^{2m} - 1}. \quad (39b)$$

A number of properties of linear modes are as follows.

(i) For each $m = 1, 2, \dots$ there are two normal modes denoted by ‘ \pm ’. Each normal mode is made up of two edge disturbances corresponding to the DF boundaries $\ell = \ell_1$ and $\ell = \ell_2$.

(ii) The eigenfrequencies, ω_m^\pm , are either real or complex conjugates of each other. If they are both real then both the normal modes are stable with pattern speed $\lambda_p^\pm = \omega_m^\pm/m$. When the eigenfrequencies are complex conjugates, then one normal mode grows exponentially (an instability) and the other decays exponentially, with both modes having the same pattern precession frequency.

(iii) From equation (39a) we see that the condition for instability is

$$\left(\frac{1 - \ell_2^2}{1 - \ell_1^2}\right)^{m/2} > \left|1 - \frac{m(\ell_2 - \ell_1)\ell_1}{1 - \ell_1^2}\right|. \quad (40)$$

(iv) It can be verified that the above inequality cannot be satisfied for any $0 \leq \ell_1 < \ell_2 < 1$, when $m = 1, 2$. So all bands have stable $m = 1$ and $m = 2$ modes, and only modes with $m = 3, 4, \dots$ can be unstable.

(v) The unstable band *waterbag_1_s0* has $\ell_1 = 0.7$ and $\ell_2 = 0.9$. The stable broad-band *waterbag_2_s0* has $\ell_1 = 0.1$ and $\ell_2 = 0.9$. Using these values of (ℓ_1, ℓ_2) in equation (40) it can be verified that (a) *waterbag_1_s0* has precisely two unstable modes, for $m = 3$ and $m = 4$; (b) for *waterbag_2_s0* modes of all m are stable. This is in agreement with the numerical simulations discussed in Section 4.

(vi) The inequality condition (40) defines a region of instability in the (ℓ_1, ℓ_2) parameter plane, for each value of m . These are displayed in Fig. 9 for $m = 3, 4, 5, 6$. As m increases the crescent-like region of instability expands.

5.1 Structure of normal modes

Stable modes. When inequality condition (40) is not satisfied the two normal mode eigenfrequencies ω_m^\pm , given by equation (39a), are both real with corresponding pattern speeds $\lambda_p^\pm = \omega_m^\pm/m$. The DF of the normal modes is

$$f_1^\pm(\ell, g, t; m) = \text{Re}\left\{A_{m1}^\pm \exp[im(g - \lambda_p^\pm t)] \delta(\ell - \ell_1) + A_{m2}^\pm \exp[im(g - \lambda_p^\pm t)] \delta(\ell - \ell_2)\right\}. \quad (41)$$

The four complex amplitudes, A_{m1}^\pm and A_{m2}^\pm , are related by equation (39b), which implies that $(A_{m2}/A_{m1})^\pm$ are real whenever ω_m^\pm are real. When the ratio is positive/negative, the normal mode is an in-phase/out-of-phase combination of the two sinusoidal edge disturbances. Moreover the product $(A_{m2}/A_{m1})^+ (A_{m2}/A_{m1})^- = 1$, which implies (i) if the ‘+’ mode is an in-phase (or out-of-phase) combination of the two edge disturbances so is the ‘-’ mode, and vice versa; (ii) if disturbance at one of the edges makes a dominant contribution to the ‘+’ mode, then the other edge disturbance makes a dominant contribution to the ‘-’ mode. To summarize, a stable ‘ \pm ’ mode is either an in-phase or out-of-phase superposition of the edge disturbances, with generally unequal amplitudes. The pattern speeds, λ_p^\pm , of the ‘ \pm ’ modes are generally unequal.

Unstable modes. When inequality condition (40) is satisfied the two normal mode eigenfrequencies ω_m^\pm , given by equation (39a), are complex conjugates of each other. We write $\omega_m^\pm = m\lambda_p \pm i\omega_1$, where λ_p is the pattern speed and $\omega_1 > 0$ can be thought as the growth rate of the ‘+’ mode, or as the damping rate of the ‘-’ mode; we will refer to ω_1 as the growth rate. Equation (39a) gives

$$\lambda_p = \frac{\Omega_1}{2} = -\frac{\ell_1}{1 - \ell_1^2}, \quad (42a)$$

$$\omega_1 = \sqrt{\frac{1}{\Delta \ell^2} \left(\frac{1 - \ell_2^2}{1 - \ell_1^2}\right)^m - \left(\frac{1}{\Delta \ell} - \frac{m \ell_1}{1 - \ell_1^2}\right)^2}. \quad (42b)$$

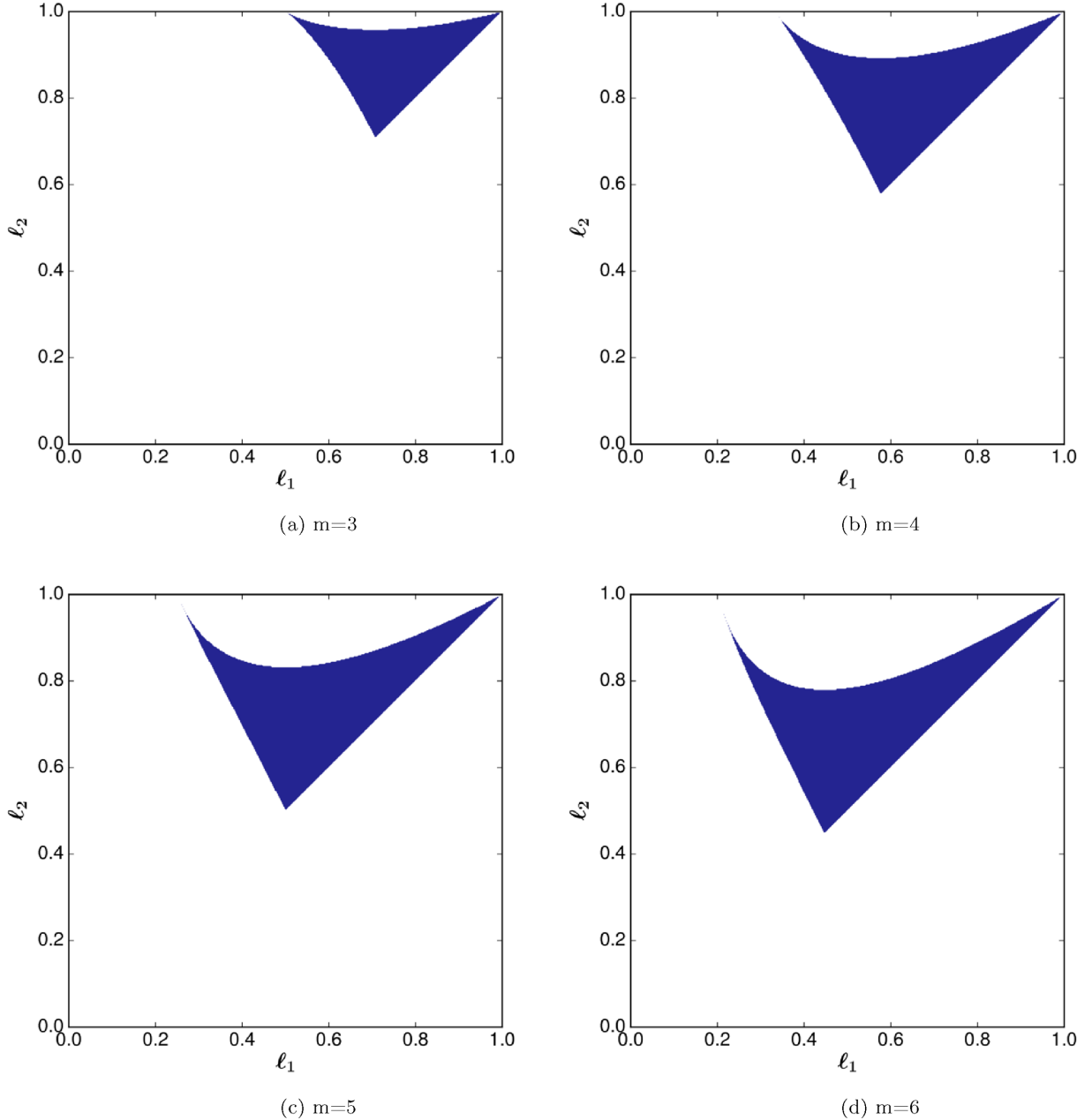


Figure 9. Instability region in (ℓ_1, ℓ_2) plane for $m = 3, 4, 5, 6$.

The pattern speed is negative and depends only on ℓ_1 . On the other hand the growth rate depends on all of (ℓ_1, ℓ_2, m) .

Equations (39a) and (39b) imply that whenever ω_m^\pm are complex conjugates, $(A_{m2}/A_{m1})^\pm$ are also complex conjugates. Moreover magnitude of the amplitude ratio $|(A_{m2}/A_{m1})^\pm| = 1$, so we can write $(A_{m2}/A_{m1})^\pm = \exp[\pm i m \theta_m]$, where

$$\theta_m = \frac{1}{m} \cos^{-1} \left[\left(\frac{1 - \ell_1^2}{1 - \ell_2^2} \right)^{m/2} \left(\frac{m \ell_1 \Delta \ell}{1 - \ell_1^2} - 1 \right) \right], \quad (43)$$

where θ_m is the relative phase shift between the two edge disturbances composing a normal mode. Then the DF of the growing and damping normal modes of a given m is given by the following

superposition of the two edge disturbances:

$$f_1^\pm(\ell, g, t; m) = \exp[\pm \omega_1 t] \operatorname{Re} \left\{ A_m^\pm \exp[im(g - \lambda_P t)] \delta(\ell - \ell_1) + A_m^\pm \exp[im(g \pm \theta_m - \lambda_P t)] \delta(\ell - \ell_2) \right\}, \quad (44)$$

where A_m^\pm is a complex amplitude that is common to both edge disturbances. In contrast to a stable mode, an unstable ‘ \pm ’ mode is a superposition of the edge disturbances with a relative phase shift but equal amplitudes, and a pattern speed $\lambda_P = \Omega_1/2$ that is the same for both ‘ \pm ’ modes.

In order to get an idea of the dependence of the growth rate as a function of the parameters (ℓ_1, ℓ_2, m) , we plot in Fig. 10 the growth rate as a function of m for different values of $\Delta \ell$ and ℓ_2 . For

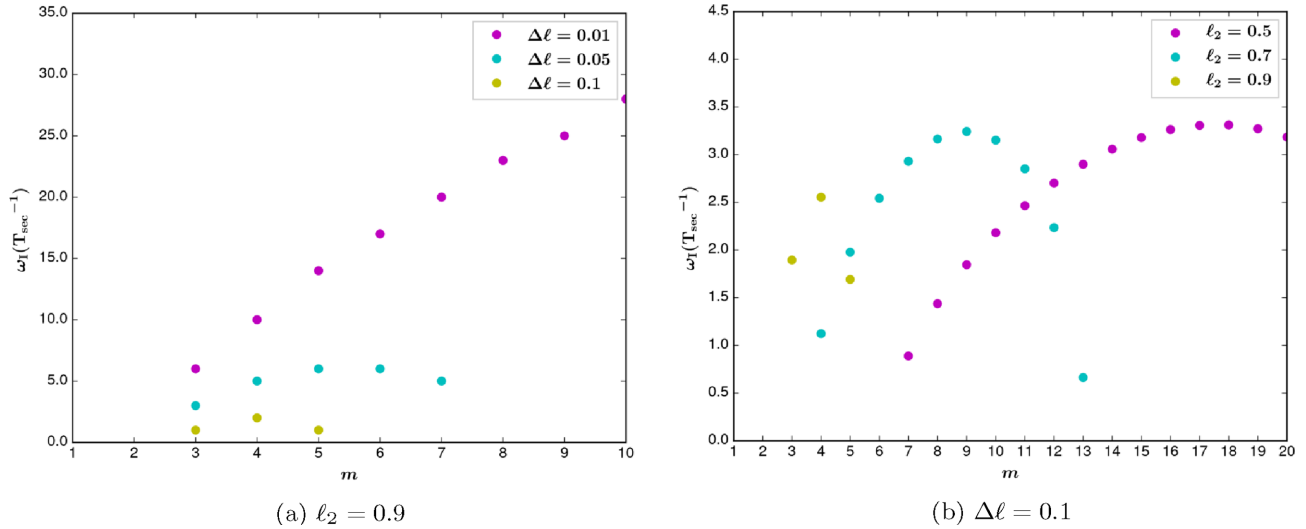


Figure 10. Growth rate ω_1 variation with m . Left-hand panel (a) corresponds to waterbags with fixed $\ell_2 = 0.9$. Right-hand panel (b) for waterbags of fixed thickness $\Delta\ell = 0.1$.

fixed $\ell_2 = 0.9$ and three different values of $\Delta\ell$, we see that bands with smaller $\Delta\ell$ are unstable over a larger range of m , with higher maximum growth rates occurring at larger m . For fixed $\Delta\ell = 0.1$ and three different values of ℓ_2 , the maximum growth rates are similar but occur at smaller m for larger ℓ_2 .

We note that *waterbag_1_s0* has unstable modes for $m = 3, 4$ with the $m = 3$ mode having the higher growth rate, $\omega_1 \sim 0.72 T_{\text{sec}}^{-1} \simeq 2.4 \text{Gyr}^{-1}$; this is consistent with the initial growth of the $m = 3$ mode in Figs 6 and 8(a). In the next section we present a more detailed comparison of numerical experiments with linear theory.

6 EVOLUTION OF INSTABILITIES

We ran a suite of numerical simulations of waterbag bands, with parameters listed in the Table 1. The primary goal is to put the linear theory of the previous section to stringent tests, and is explored through the upper (Set I) and lower (Set II) groups shown in Table 1.

(i) *Set I* consists of five cases, of which two – the unstable band *waterbag_1_s0* and the stable band *waterbag_2_s0* – have already been discussed.

(ii) *Set II* is a detailed test of the linear theory prediction of the transition from instability to stability of a band with fixed $\ell_1 = 0.8$, as ℓ_2 is varied over a range of values.

Then we give a taste of the long-term evolution of an unstable band that goes well beyond the applicability of linear theory. Here the point of interest is in the collisionless relaxation to a state with a wide spread in eccentricities.

6.1 Set I

Of the five cases in Set I, *waterbag_1_s0* and *waterbag_2_s0* have been discussed earlier. *waterbag_5_s0* is stable according to linear theory, and the simulation results confirmed this, showing stable evolution similar to *waterbag_2_s0*. We now consider two new unstable bands, *waterbag_3_s0* and *waterbag_4_s0*. In Table 2 we list

Table 2. Theoretical predictions for the unstable bands of Set I.

System name	Unstable m	Fastest growing mode		
		m_0	$\omega_{1, \text{max}}$ (Gyr^{-1})	λ_{p0} (rad Gyr^{-1})
<i>waterbag_1_s0</i>	3, 4	3	2.4	-4.57
<i>waterbag_3_s0</i>	3, 4, 5	4	8.5	-7.41
<i>waterbag_4_s0</i>	3–7	6	20.6	-10.21

the predictions of linear theory for these two bands, including also *waterbag_1_s0* whose instability was discussed earlier. For each band all its unstable modes are identified, and the growth rate ($\omega_{1, \text{max}}$) and pattern speed (λ_{p0}) of the most unstable mode (m_0) are computed using equations (42b) and (42a). *Simulations of waterbag_3_s0*. From Fig. 11 we see that a $m = 4$ pattern emerges by $\sim 0.06 \text{Gyr}$, which is in agreement with linear theory. Non-linear interactions, mainly with the unstable $m = 5$ mode, lead to distortions of the pattern. This can be seen clearly in Fig. 13(a) that plots the mode amplitudes a_m versus time: the $m = 4$ mode has the maximum amplitude until $\sim 0.2 \text{Gyr}$, after which the $m = 5$ mode begins to dominate.

Simulations of waterbag_4_s0. From Fig. 12 we see that a $m = 6$ pattern emerges by $\sim 0.03 \text{Gyr}$, which is in agreement with linear theory. Non-linear interactions with other unstable modes lead to distortions of the pattern. This can be seen clearly in Fig. 13(b) that plots the mode amplitudes a_m versus time: the $m = 6$ mode dominates until $\sim 0.2 \text{Gyr}$, after which there seems to be non-linear interactions among many modes.

Table 3 shows the general agreement between linear theory and simulations.

6.2 Set II

The narrowest band in Table 1 is *waterbag_ℓ1_0.8_ℓ2_0.81*, with $\Delta\ell = 0.01$. According to linear theory this band is unstable to a wide range of modes with $m = 3\text{--}57$, with $m = 36$ having the fastest growth rate. Fig. 14 shows the evolution

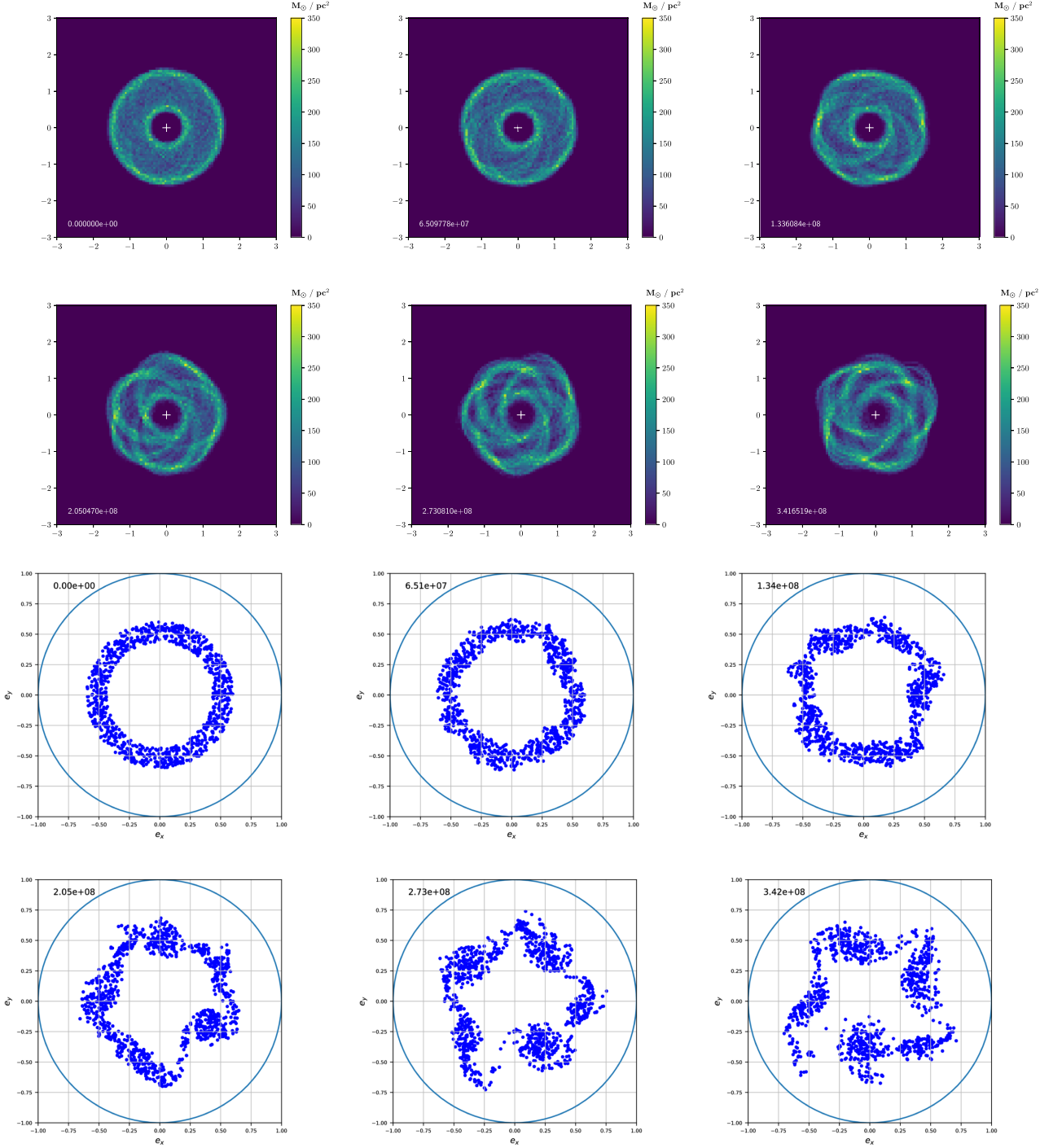


Figure 11. Similar to Fig. 6, but for `waterbag_3_s0`. A $m = 4$ pattern emerges by ~ 0.06 Gyr.

of this narrow band, whose initial evolution shows an instability dominated by $m \sim 36$ mode, in agreement with linear theory.

Linear theory also predicts a transition from instability to stability when the lower boundary is held fixed at $\ell_1 = 0.8$ and the band is made broader by increasing ℓ_2 . This transition occurs at $\ell_2 = \ell_{\text{crit}} \simeq 0.963$: bands with $\ell_2 < \ell_{\text{crit}}$ are unstable to various modes whereas

broader bands with $\ell_{\text{crit}} < \ell_2 < 1$ are stable for all m . In order to test this precise prediction, we ran a total of 20 simulations increasing ℓ_2 in steps of 0.01, from 0.81 to 1, and looked for signs of instabilities. From the last column of Table 1 we see that the simulations confirm linear theory, with the small difference that the transition seems to happen when ℓ_2 crosses 0.97, instead of the predicted value of 0.963.

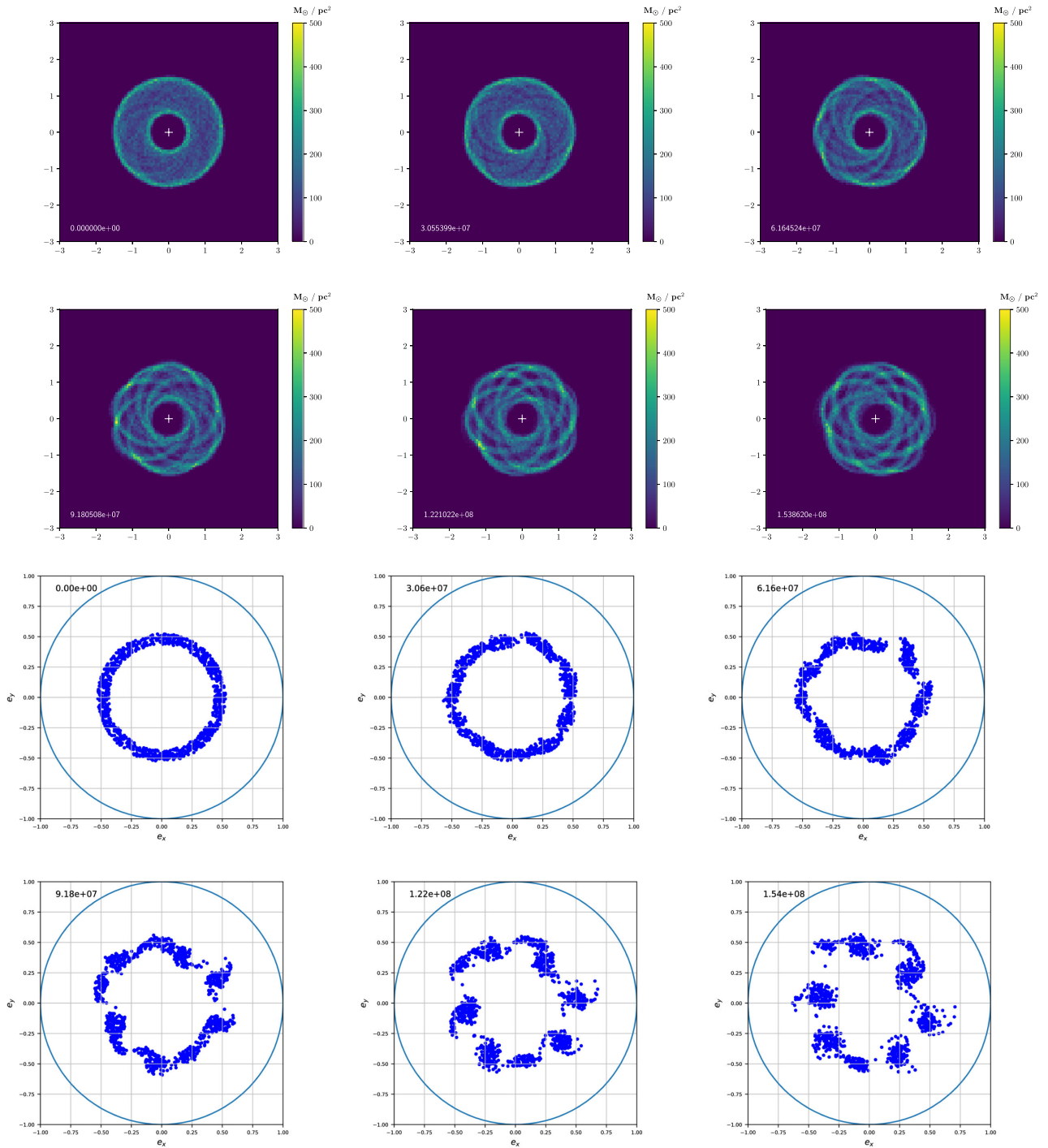


Figure 12. Similar to Fig. 6, but for *waterbag_4_s0*. A $m = 6$ pattern emerges by ~ 0.03 Gyr.

6.3 Collisionless relaxation

As instabilities unfold and non-linear interactions between modes dominate, what can we expect of evolution over long times? We have earlier in this section followed the short-time evolution of the unstable band *waterbag_3_s0*, with its initial growth of a dominant $m = 4$ mode over ~ 0.06 Gyr, followed by the rise of a

$m = 5$ mode around ~ 0.2 Gyr lasting until at least ~ 0.34 Gyr. What happens after this? Here we follow the evolution for ~ 4 Gyr.

Fig. 15 shows both the initial and final states of *waterbag_3_s0*. When compared with the intermediate states of Fig. 11, the final state appears more axisymmetric. The final state also has a wider range of eccentricities than the initial state. It consists of a nearly circular high-density ring, surrounded by

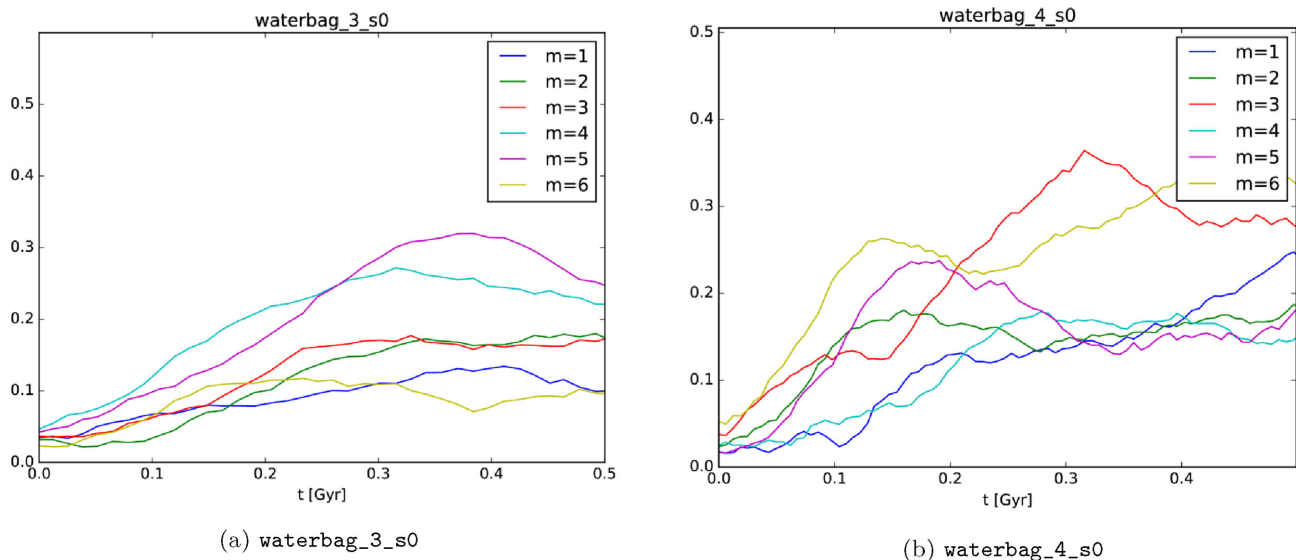


Figure 13. Evolution of mode amplitudes a_m . (a) waterbag_3_s0, (b) waterbag_4_s0.

Table 3. Comparison between linear theory and simulations for the unstable bands of Set I.

System name	Fastest growing mode		Agreement
	m_0 (theory)	m_0 (simulations)	
waterbag_1_s0	3	3	Yes
waterbag_3_s0	4	4	Yes ^a
waterbag_4_s0	6	6	Yes ^a

^aThere is good agreement for waterbag_3_s0 for $t < 0.2$ Gyr, and for waterbag_4_s0 for $0.05 < t < 0.15$ Gyr.

a lower density halo of particles with a wide range of eccentricities. The strong non-axisymmetric instabilities that plagued the initial state seem to have saturated, leaving behind a relaxed, coarse-grained state that is approximately axisymmetric and steady in time. The secular precessional time-scale for the initial state is $T_{\text{sec}} \sim 0.8$ Gyr, so the total duration of the run, 4 Gyr is about $5 T_{\text{sec}}$. This is too short a duration for a collisional process like resonant relaxation to be effective. Hence what we have witnessed must be collisionless relaxation, where non-axisymmetric instabilities provide the pathway for transition from one axisymmetric state to another.

7 CONCLUSIONS

Mono-energetic waterbags are the simplest models of low-mass stellar discs around a MBH. We studied, analytically and numerically, the stability of initial states that are prograde and axisymmetric. These waterbags have a DF, $f_0(\ell)$, which is constant when $0 \leq \ell_1 \leq \ell \leq \ell_2 \leq 1$, and zero when ℓ is outside this range. There are two types of waterbags: polarcaps with $\ell_2 = 1$ and bands with $\ell_2 < 1$. The linear stability problem can be solved simply: for each m the growth rates of instabilities, pattern speeds of stable and unstable

modes and the complete normal mode structure have been determined explicitly as functions of (ℓ_1, ℓ_2) , the waterbag parameters.

(i) Polarcaps have one stable normal mode for each m , with the noteworthy feature that the $m = 1$ mode always has positive pattern speed. For a polarcap consisting of orbits with eccentricities $e < 0.9428$, only the $m = 1$ mode has a positive pattern speed.

(ii) Bands have two normal modes for each m , and can be either stable or unstable. Very narrow bands (with $\ell_1 \simeq \ell_2$) are unstable to modes with a wide range in m , whereas broad-bands approaching a polarcap (with $\ell_2 \simeq 1$) are stable.

The evolution of instabilities was also explored through numerical simulations, which can explore both linear and non-linear regimes. A variety of numerical experiments were performed by which we demonstrated good agreement with linear theory. Long-time integration showed the growth of instabilities of different m , which interacted with each other non-linearly, then saturated and later relaxed collisionlessly into a quasi-steady state, which has a wider range of eccentric orbits than the initial state. This suggests secular non-axisymmetric instabilities could provide pathways for stars to exchange angular momentum via the mean self-gravitational field, and spread out in eccentricities.

It is straightforward to extend our study to include external gravitational sources (such as nuclear density cusps or distant perturbers) and general relativity, as described in ST1. But one clearly needs to go well beyond our simple models in order to study real systems, like the disc of young stars at the Galactic Centre. We need to consider more general DFs and include orbits with a range of semimajor axes and inclinations. But self-gravitational dynamics poses difficult problems and secular dynamics is still in its infancy, so we need to build the tools step-by-step; describing the collisionless relaxation of even an unstable band remains a challenge for dynamists.

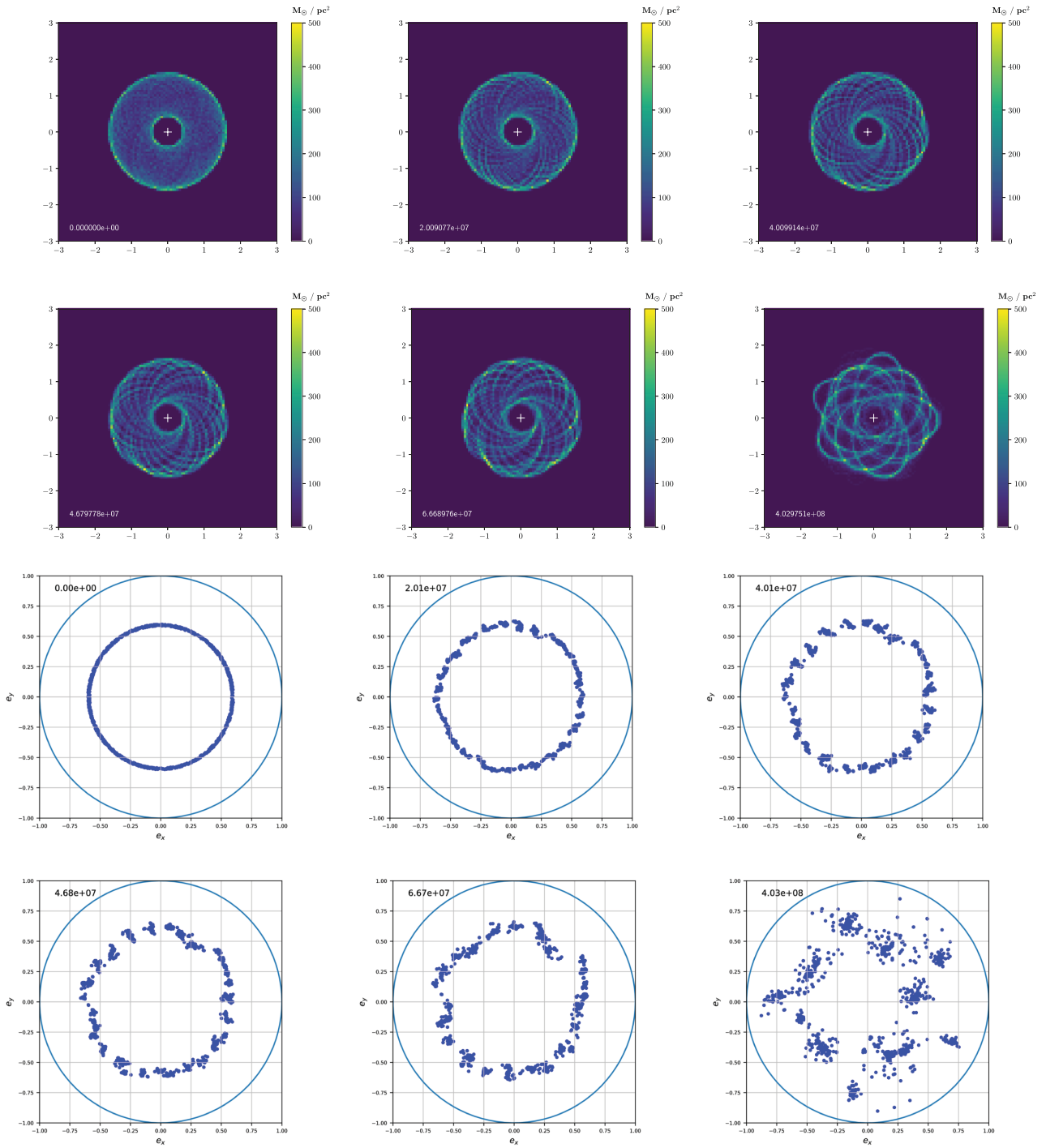


Figure 14. Similar to Fig. 6, but for `waterbag_l1_0.8_l2_0.81`. A high m pattern emerges by ~ 0.02 Gyr.

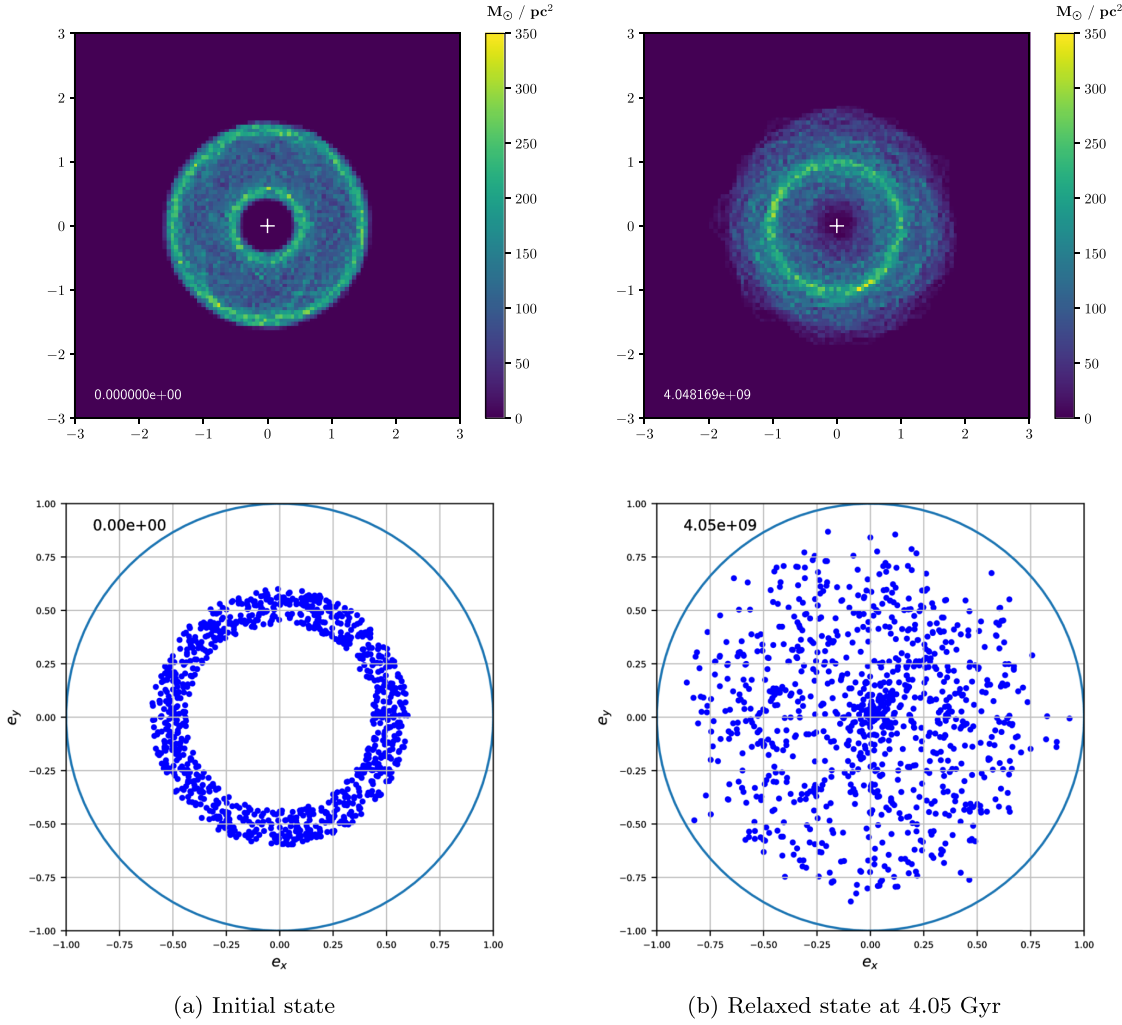


Figure 15. Collisionless relaxation of waterbag_3_s0.

REFERENCES

- Berk H., Roberts K., 1970, *Methods Comput. Phys.*, 9, 87
 Borderies N., Goldreich P., Tremaine S., 1983, *Icarus*, 53, 84
 Dritschel D. G., 1988, *J. Comput. Phys.*, 79, 477
 Jalali M. A., Tremaine S., 2012, *MNRAS*, 421, 2368
 Kazandjian M. V., Touma J. R., 2013, *MNRAS*, 430, 2732
 Levin Y., Beloborodov A. M., 2003, *ApJ*, 590, L33
 Polyachenko E. V., Polyachenko V. L., Shukhman I. G., 2007, *MNRAS*, 379, 573
 Sridhar S., Touma J. R., 2016, *MNRAS*, 458, 4129 (ST1)
 Touma J. R., 2002, *MNRAS*, 333, 583
 Touma J., Tremaine S., 2014, *J. Phys. A*, 47, 292001
 Touma J. R., Tremaine S., Kazandjian M. V., 2009, *MNRAS*, 394, 1085
 Tremaine S., 2001, *AJ*, 121, 1776
 Tremaine S., 2005, *ApJ*, 625, 143
 Yelda S., Ghez A. M., Lu J. R., Do T., Meyer L., Morris M. R., Matthews K., 2014, *ApJ*, 783, 131

APPENDIX A: SURFACE PROBABILITY DENSITY

The surface probability density function is obtained by integrating the disc DF over velocity space:

$$\Sigma(\mathbf{r}) = \int d\mathbf{u} \hat{f}(\mathbf{r}, \mathbf{u}), \quad (\text{A1})$$

where the DF $\hat{f}(\mathbf{r}, \mathbf{u})$ is written as a function of \mathbf{r} and \mathbf{u} , which are the position vector and velocity of a star, respectively, in the MBH's rest frame. For a razor-thin disc, the four-dimensional phase volume, $d\mathbf{r} d\mathbf{u} = I dw dI dg d\ell$. Hence the DF of an axisymmetric mono-energetic disc (not necessarily a waterbag) is related to the DF, $f_0(\ell)$, of Section 3.2, as follows:

$$\hat{f}(\mathbf{r}, \mathbf{u}) = \frac{f_0(\ell)}{4\pi^2 I_0} \delta(I - I_0). \quad (\text{A2})$$

Then

$$\Sigma_0(r) = \frac{1}{4\pi^2 I_0} \int du d\phi u f_0(\ell) \delta(I - I_0), \quad (\text{A3})$$

where u is the speed, ϕ is the angle between \mathbf{u} and \mathbf{r} , and $I_0 = \sqrt{GM_\bullet a_0}$. Since the discs we consider have only prograde orbits, $\ell > 0$, which implies that $0 \leq \phi \leq \pi$. We now express the (scaled) Delaunay variables, $\{\ell, I\}$, in terms of $\{u, \phi\}$:

$$I = \left(\frac{2}{GM_\bullet r} - \frac{u^2}{(GM_\bullet)^2} \right)^{-1/2}, \quad (\text{A4a})$$

$$\ell = L/I = I^{-1} r u \sin \phi. \quad (\text{A4b})$$

Hence

$$\delta(I - I_0) = \frac{\delta(u - u_0)}{|dI/du|_{u_0}} = \left(\frac{GM_\bullet}{a_0^3} \right)^{1/2} \frac{\delta(u - u_0)}{u_0}, \quad (\text{A5})$$

where

$$u_0(r) = \begin{cases} \sqrt{GM_\bullet \left(\frac{2}{r} - \frac{1}{a_0} \right)} & \text{for } r \leq 2a_0, \\ 0 & \text{for } r > 2a_0 \end{cases} \quad (\text{A6})$$

is the speed at radius r of an orbit with semimajor axis a_0 . Substituting equation (A5) in equation (A3) and using equations (A4b) and (A6), the surface density for a general mono-energetic DF:

$$\Sigma_0(r) = \frac{1}{4\pi^2 a_0^2} \int d\phi f_0(\ell_0(r) \sin \phi), \quad (\text{A7})$$

where

$$\ell_0(r) = \frac{r u_0(r)}{I_0} = \begin{cases} \sqrt{\frac{2r}{a_0} - \frac{r^2}{a_0^2}} & \text{for } r \leq 2a_0, \\ 0 & \text{for } r > 2a_0. \end{cases} \quad (\text{A8})$$

For the waterbag DF of equation (25), $f_0(\ell) = 1/\Delta\ell = 1/(\ell_2 - \ell_1) = \text{constant}$ for $0 \leq \ell_1 \leq \ell \leq \ell_2 \leq 1$ and is zero outside this range. This implies that $\Sigma_0(r)$ is non-zero only when $|r - a_0| \leq a_0 e_1$. Within this range of radii,

$$\Sigma_0(r) = \frac{1}{4\pi^2 a_0^2 \Delta\ell} \Delta\phi(r), \quad (\text{A9})$$

where $\Delta\phi(r)$ is the range in ϕ for which

$$\frac{\ell_1}{\ell_0(r)} \leq \sin \phi \leq \frac{\ell_2}{\ell_0(r)}. \quad (\text{A10})$$

All we need to do now is to determine $\Delta\phi(r)$. There are two cases to consider.

Case 1: $\ell_2 \leq \ell_0(r)$. Using equation (A8), this condition is equivalent to $|r - a_0| \leq a_0 e_2$. Then $\Delta\phi(r) = 2(\phi_2 - \phi_1)$, where $\phi_1(r) = \sin^{-1}[\ell_1/\ell_0(r)]$ and $\phi_2(r) = \sin^{-1}[\ell_2/\ell_0(r)]$.

Case 2: $\ell_1 \leq \ell_0(r) \leq \ell_2$. Using equation (A8), this condition is equivalent to $a_0 e_2 \leq |r - a_0| \leq a_0 e_1$. Then $\Delta\phi(r) = 2(\pi/2 - \phi_1)$.

Substituting these expressions for $\Delta\phi(r)$ in equation (A9), we obtain equation (26) for the surface probability density of an axisymmetric mono-energetic waterbag.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.