Density waves in debris discs and galactic nuclei

Mir Abbas Jalali[1,2]★ and Scott Tremaine[2]★

[1] *Sharif University of Technology, Azadi Avenue, PO Box 11155-9567 Tehran, Iran*
[2] *School of Natural Sciences, Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA*

ABSTRACT

We study the linear perturbations of collisionless near-Keplerian discs. Such systems are models for debris discs around stars and the stellar discs surrounding supermassive black holes at the centres of galaxies. Using a finite-element method, we solve the linearized collisionless Boltzmann equation and Poisson's equation for a wide range of disc masses and rms orbital eccentricities to obtain the eigenfrequencies and shapes of normal modes. We find that these discs can support large-scale 'slow' modes, in which the frequency is proportional to the disc mass. Slow modes are present for arbitrarily small disc mass so long as the self-gravity of the disc is the dominant source of apsidal precession. We find that slow modes are of two general types: parent modes and hybrid child modes, the latter arising from resonant interactions between parent modes and singular van Kampen modes. The most prominent slow modes have azimuthal wavenumbers $m = 1$ and $m = 2$. We illustrate how slow modes in debris discs are excited during a fly-by of a neighbouring star. Many of the non-axisymmetric features seen in debris discs (clumps, eccentricity, spiral waves) that are commonly attributed to planets could instead arise from slow modes; the two hypotheses can be distinguished by long-term measurements of the pattern speed of the features.

Key words: methods: numerical – planets and satellites: formation – protoplanetary discs – galaxies: kinematics and dynamics – galaxies: nuclei.

1 INTRODUCTION

Debris discs are planetesimal discs that are detected through thermal infrared emission or scattered starlight from dust formed in recent planetesimal collisions. The bolometric luminosity from detectable debris discs is typically $\gtrsim 10^{-5}$ of the stellar luminosity; the inferred dust masses are typically $\lesssim 1 M_\oplus$ and the ages of the host stars range from 10 Myr to 10 Gyr (see Wyatt 2008 for a review).

A variety of features in debris discs have been interpreted as evidence for planets. These include structures in the β Pictoris disc, including a warp (Heap et al. 2000), a system of tilted rings (Wahhaj et al. 2003) and a bright clump (Telesco et al. 2005); clumps in the discs around Vega (Wyatt 2003), ϵ Eridani (Greaves et al. 2005), η Corvi (Wyatt et al. 2005) and HD 107146 (Corder et al. 2009); the eccentricity of the discs around HR 4796A and Fomalhaut (Telesco et al. 2000; Kalas, Graham & Clampin 2005); spiral structure in the disc around HD 141569 (Clampin et al. 2003); and sharp inner or outer edges in the discs around Fomalhaut and HD 92945 (Kalas et al. 2005; Golimowski et al. 2011).

Detailed dynamical models have shown that most or all of these features can be produced by planets (see Wyatt 2009 for a review).

Moreover, in the case of β Pictoris (Lagrange et al. 2010), and perhaps Fomalhaut (Kalas et al. 2008), planets have been detected that may indeed be responsible for some or all of these features. Nevertheless, it is important to ask what long-lived structures could arise in debris discs *without* planets.

In this paper we examine the possibility that low-mass discs can support long-lived normal modes maintained by the self-gravity of the disc. Normally it is assumed that debris discs cannot support such modes because of their small masses; all localized disturbances are dispersed by the Keplerian shear. However, a special feature of Keplerian orbits is that eccentric orbits do not precess. Thus the evolution of eccentric disturbances in a debris disc is governed by the non-Keplerian forces, however small these may be. In this paper we shall focus on the non-Keplerian forces arising from the self-gravity of the disc. We neglect other possible perturbations for a variety of reasons. We ignore gravitational forces from planets because our principal goal is to understand the properties of discs in the simplest case, when no planets are present. We ignore radiation pressure, even though this affects the dynamics of the dust that dominates the thermal infrared emission and the scattered light; our justification is that the large planetesimals that generate the dust are unaffected by radiation pressure but we recognize that the distribution of (invisible) parent bodies and (visible) dust is likely to be different. We ignore gas drag since old debris discs contain little

★E-mail: mjalali@sharif.edu (MAJ); tremaine@ias.edu (ST)

or no gas, and since the planetesimals are likely to be large enough to be insensitive to drag. We ignore collisions between planetesimals because they are likely to be rare; indeed such collisions probably drive the long-term erosion of the disc in which case the collision time cannot be much less than the disc age.

Debris discs are distinct from protoplanetary discs: the latter comprised mostly gas, not dust or planetesimals; they are much younger (typically less than a few Myr) and more massive (0.001–0.1 $M_\odot$) than debris discs (see Williams & Cieza 2011 for a review). Protoplanetary discs are depleted by various processes, including photoevaporation, accretion on to the host star, condensation of refractory elements into dust grains and then planetesimals, and stellar winds. Eventually they are likely to evolve into planetesimal/debris discs. Although our analysis here is restricted to collisionless systems, many of our results – in particular the existence of stable, slow, modes supported by the self-gravity of the disc – also apply to protoplanetary gas discs and may explain some non-axisymmetric features of these discs.

To summarize we treat debris discs as collisionless systems composed of particles influenced only by the gravity of the host star and the self-gravity of the disc. Their dynamics is therefore similar to the dynamics of discs of stars orbiting near the supermassive black holes found in the centres of most galaxies. Examples of these include the disc(s) of young stars found in the central parsec of the Milky Way (Genzel, Eisenhauer & Gillessen 2010), the two discs – one of young stars at ∼0.1 pc and one of old stars at ∼1 pc – found at the centre of M31 (Bender et al. 2005), and the stellar discs that are inferred to form in the outer parts of quasar accretion discs (Goodman 2003).

The properties of the normal modes of low-mass near-Keplerian discs were investigated by Tremaine (2001, hereafter T01), who found that (i) the frequency of the mode is proportional to the ratio μ of the masses of the disc and central star, but the shape of the mode is independent of μ so long as $\mu \ll 1$ (hence these are called 'slow' modes); (ii) all slow modes are stable; (iii) in discs with rms eccentricity $e_{\mathrm{rms}} \ll 1$ all slow modes have azimuthal wavenumber $m = 1$, i.e. they are lopsided.

The results in T01 are based on linear normal-mode calculations for discs composed of particles in circular orbits, with softened self-gravity used to mimic the effects of the velocity dispersion or non-zero eccentricities of the particles. These calculations are supplemented by analytic results using the WKB (short-wavelength) approximation, which assumes that the wavelengths of the normal modes are small compared to the radius. The WKB results appear to provide a useful guide even though this short-wavelength approximation is not realistic for some of the disc modes. In this paper the effects of the velocity dispersion are computed directly, and we examine discs with a range of rms eccentricities e_{rms}, from nearly zero ('cold' discs) to ∼0.35 ('warm' discs). Our numerical results are derived using a finite-element method (FEM) for studying the linear normal modes of collisionless self-gravitating discs, as described in Jalali (2010). In particular, we intend to address the following questions: (i) what are the properties of the frequency spectra of near-Keplerian discs? (ii) are there any unstable modes? (iii) are there isolated oscillatory modes in the spectrum that survive Landau damping? (iii) what are the differences between the spectra of cold and warm discs? (iv) how can stable density waves be excited in such discs?

We introduce a family of axisymmetric near-Keplerian discs in Section 2 and construct their equilibrium phase-space distribution functions (DFs) in Section 3. We obtain the governing equations of the perturbed dynamics in Section 4 and explain the numerical

solution procedure in Section 5. We present the frequency spectra of our discs in Section 6 and discuss the characteristics of eigenmodes in warm and cold discs. We describe how these waves can be excited by tidal forces in Section 7. The reader who is mainly interested in the application of our results to debris discs and galactic nuclei can focus on Figs 7 and 10 and the discussion in Section 8.

2 THE MODEL

We introduce a simple model of annular discs around massive objects by subtracting two Toomre (1963) discs with $n = 1$ and $n = 2$; the resulting surface density is

$$
\begin{aligned}
S_{\mathrm{d}}(r) &= \frac{3M_{\mathrm{d}}}{4\pi b^2} \left\{ \frac{1}{[1 + (r/b)^2]^{3/2}} - \frac{1}{[1 + (r/b)^2]^{5/2}} \right\}, \\
&= \frac{3M_{\mathrm{d}}}{4\pi b^2} \frac{(r/b)^2}{[1 + (r/b)^2]^{5/2}},
\end{aligned}
\tag{1}
$$

where M_{d} is the disc mass, b is a length scale and r is the radial distance to the central star. The potential corresponding to the surface density S_{d} is

$$
\Phi_{\mathrm{d}}(r) = -\frac{GM_{\mathrm{d}}}{2b} \frac{1 + 2(r/b)^2}{[1 + (r/b)^2]^{3/2}},
\tag{2}
$$

with G being the gravitation constant. For a central star of mass $M_\star$, the total potential governing the motion of particles is

$$
\Phi_0(r) = -\frac{GM_\star}{r} + \Phi_{\mathrm{d}}(r).
\tag{3}
$$

We define

$$
\mu = \frac{M_{\mathrm{d}}}{M_\star}, \quad R = r/b,
\tag{4}
$$

and work with the dimensionless unperturbed potential

$$
V_0(R) \equiv \frac{b\Phi_0}{GM_\star} = -\frac{1}{R} - \frac{\mu}{2} \frac{1 + 2R^2}{(1 + R^2)^{3/2}},
\tag{5}
$$

and density

$$
\Sigma_0(R) \equiv \frac{b^2 S_{\mathrm{d}}}{M_\star} = \frac{3\mu}{4\pi} \frac{R^2}{(1 + R^2)^{5/2}}.
\tag{6}
$$

The top panel in Fig. 1 shows the radial profile of Σ_0/μ.

The velocity of circular orbits, $v_{\mathrm{c}}(R)$, is determined from

$$
v_{\mathrm{c}}^2(R) = R \frac{\mathrm{d}V_0}{\mathrm{d}R} = \frac{1}{R} + \frac{\mu}{2} \frac{R^2(2R^2 - 1)}{(1 + R^2)^{5/2}}.
\tag{7}
$$

The second term on the right-hand side of (7) becomes negative for $R^2 < 1/2$. This means that our discs cannot exist in the absence of a central point mass. More precisely, $v_{\mathrm{c}}^2 \geq 0$ at all R if and only if $\mu \leq 5^{5/2}$; this is not a limitation in practice since protoplanetary discs are expected to have $\mu \ll 1$.

We restrict ourselves to razor-thin discs since the vertical structure of thin discs should not strongly affect their large-scale response. Using the polar coordinates (R, ϕ) and their corresponding generalized momenta (p_R, p_ϕ), the Hamiltonian function governing the motion of particles reads

$$
\mathcal{H}_0(p_R, p_\phi, R) \equiv E = \frac{p_R^2}{2} + \frac{p_\phi^2}{2R^2} + V_0(R).
\tag{8}
$$

Since ϕ is a cyclic coordinate, its conjugate momentum p_ϕ is a constant of motion in the unperturbed disc. The orbital energy E is another integral of motion. Canonical perturbation theories describing the motion of particles, and the perturbed collisionless Boltzmann equation (CBE), are substantially simplified by using the action
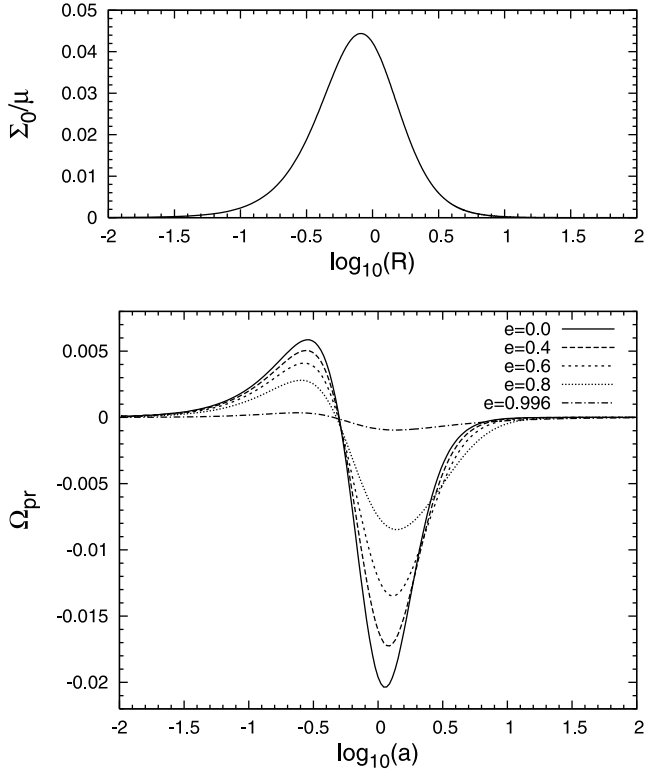
Figure 1. Top: surface-density profile of the composite Toomre disc (equation 6). Bottom: variation of the precession rate Ω_{pr} for $\mu = 0.1$. For radial orbits (eccentricity $e = 1$) $\Omega_{pr} = 0$.

variables $\boldsymbol{J} = (J_R, J_\phi)$ and their conjugate angles $\boldsymbol{w} = (w_R, w_\phi)$ with

$$J_R = \frac{1}{2\pi} \oint p_R \, dR, \quad J_\phi = \frac{1}{2\pi} \oint p_\phi \, d\phi = p_\phi. \quad (9)$$

These integrals are taken along the orbits, which consist of slowly precessing Kepler ellipses when $\mu \ll 1$. The unperturbed Hamiltonian $\mathcal{H}_0$ depends only on the actions, not the angles. The action $J_\phi = L$ is the magnitude of the angular-momentum vector. In the angle-action space, the equations of motion become

$$\dot{\boldsymbol{w}} = \boldsymbol{\Omega}(\boldsymbol{J}) = \frac{\partial \mathcal{H}_0(\boldsymbol{J})}{\partial \boldsymbol{J}}, \quad \dot{\boldsymbol{j}} = 0, \quad (10)$$

and the orbital frequencies $\boldsymbol{\Omega}(\boldsymbol{J}) = (\Omega_R, \Omega_\phi)$ are computed from

$$\frac{2\pi}{\Omega_R(\boldsymbol{J})} = \oint \frac{dR}{p_R(R, \boldsymbol{J})}, \quad \frac{\Omega_\phi(\boldsymbol{J})}{\Omega_R(\boldsymbol{J})} = \frac{J_\phi}{2\pi} \oint \frac{dR}{R^2 p_R(R, \boldsymbol{J})}. \quad (11)$$

In the limit $\mu \to 0$, the potential is Keplerian and we have $\Omega_R = \Omega_\phi = a^{-3/2}$ with a being the orbital semi-major axis. For $0 < \mu \ll 1$ the radial and azimuthal frequencies are no longer equal, but their difference $\Omega_{pr} = \Omega_\phi - \Omega_R$ is small, and that is the precession rate of the line of apsides. The Taylor expansion of Ω_{pr} begins with terms of $\mathcal{O}(\mu)$ because Ω_{pr} vanishes for Keplerian orbits. Consequently, for $\mu \ll 1$, the precession rate is proportional to the disc mass. For nearly circular orbits, the precession rate is given analytically by

$$\Omega_{pr} = \frac{3\mu}{4} \frac{R^{3/2}(1 - 4R^2)}{(1 + R^2)^{7/2}} + \mathcal{O}(\mu^2). \quad (12)$$

Instead of the actions one may use the semi-major axis $a(\boldsymbol{J})$ and eccentricity $e(\boldsymbol{J})$ defined by

$$a = \frac{R_{min}(\boldsymbol{J}) + R_{max}(\boldsymbol{J})}{2}, \quad e = \frac{R_{max}(\boldsymbol{J}) - R_{min}(\boldsymbol{J})}{R_{min}(\boldsymbol{J}) + R_{max}(\boldsymbol{J})}, \quad (13)$$

where $R_{min}(\boldsymbol{J})$ and $R_{max}(\boldsymbol{J})$ are the minimum and maximum distances of particles from the central star. These definitions are consistent with the standard Keplerian definitions when the disc mass vanishes. In the bottom panel of Fig. 1, we have plotted the variation of Ω_{pr} versus a for $\mu = 0.1$ and several choices of e. It is seen that the precession rate of orbits – of any eccentricity – has a positive peak within the region where Σ_0 is rising, and then switches sign and remains negative in the outer regions. The precession rate crosses through zero near $a = 0.5$ at all eccentricities. The maximum precession rate for nearly circular orbits and $\mu \ll 1$ is given by equation (12) as $\omega_0 = 0.05861\mu$, which occurs at $R = 0.2859$. In Section 6, we shall show that the pattern speeds of stable waves are closely related to ω_0.

3 PHASE-SPACE DISTRIBUTION FUNCTION

Particle orbits in collisionless discs are not necessarily circular. We therefore construct phase-space DFs that enable us to distribute non-circular orbits in the disc. We seek DFs of the form (Sawamura 1988; Pichon & Lynden-Bell 1996)

$$f_0(\mathcal{E}, L) = L^{2K+2} g_K(\mathcal{E}), \quad \mathcal{E} = -E, \quad (14)$$

where $0 \leq L \leq L_c(\mathcal{E})$, $L_c(\mathcal{E})$ is the angular momentum of a circular orbit with energy $E = -\mathcal{E}$, and K is a positive integer. To reproduce the surface density the DF must satisfy the relation

$$\Sigma_0(R) = 2 \int_0^\Psi d\mathcal{E} \int_0^{L_{max}} \frac{f_0(\mathcal{E}, L) \, dL}{\sqrt{L_{max}^2 - L^2}}, \quad \Psi = -V_0, \quad (15)$$

where $L_{max} = R[2(\Psi - \mathcal{E})]^{1/2}$. Substituting (14) into (15) and performing the integral over L we find

$$2^{K+1} B\left(K + \frac{3}{2}, \frac{1}{2}\right) \int_0^\Psi d\mathcal{E} \, g_K(\mathcal{E})(\Psi - \mathcal{E})^{K+1} = \frac{\Sigma_0(R)}{R^{2K+2}}, \quad (16)$$

where $B(p, q)$ is the beta function. Taking the $(K + 2)$th-order derivative of both sides of (16) with respect to Ψ gives an explicit analytic form of g_K,

$$g_K(\Psi) = \frac{1}{\sqrt{\pi} 2^{K+1} \Gamma(K + 3/2)} \frac{d^{K+2}}{d\Psi^{K+2}} \frac{\Sigma_0(R)}{R^{2K+2}}. \quad (17)$$

One needs to know explicitly the function $R(\Psi)$ before doing the derivatives on the right-hand side of (17). Since μ is small in the discs we are considering, we utilize a perturbation method to compute R in terms of Ψ. Let us define $u = 1/R$ and rewrite (5) in the form

$$\Psi = u + \mu Q(u), \quad Q(u) = \frac{1}{2} \frac{u(2 + u^2)}{(1 + u^2)^{3/2}}. \quad (18)$$

We now assume a formal series expansion for u in terms of μ as (Bellman 1964)

$$u(\Psi) = u_0(\Psi) + \mu u_1(\Psi) + \mu^2 u_2(\Psi) + \cdots, \quad (19)$$

and substitute this into (18). The functions $u_j(\Psi)$ are recursively determined by putting equal to zero the coefficients of μ^j ($j = 0, 1, 2, \ldots$). The recursion begins with $u_0 = \Psi$. Up to the third-order terms, we obtain

$$u_1 = -Q(u_0),$$

$$u_2 = -u_1 Q'(u_0),$$

$$u_3 = -u_2 Q'(u_0) - \frac{1}{2} u_1^2 Q''(u_0), \quad (20)$$

where $Q'(u) = dQ/du$. The series for u converges rapidly so keeping the terms of $\mathcal{O}(\mu^2)$ is quite sufficient for computing $R(\Psi) = 1/u(\Psi)$ in discs with $\mu \leq 0.1$.
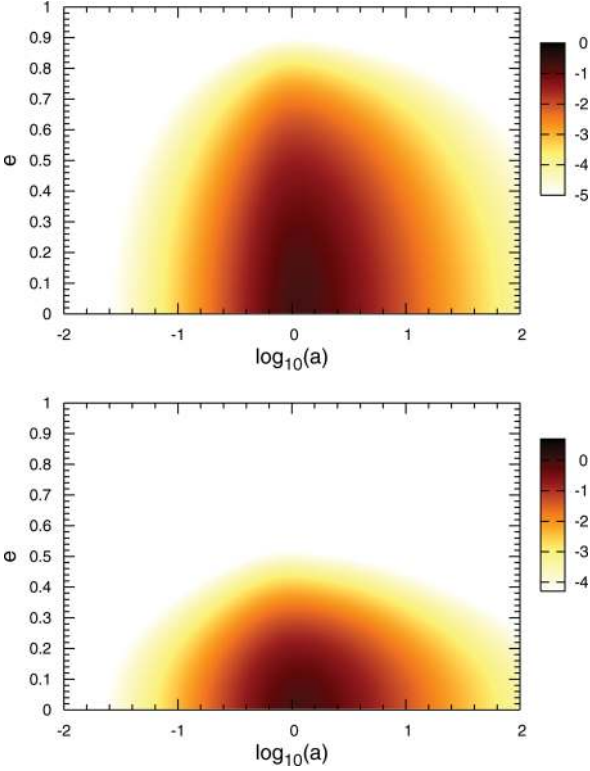
Figure 2. Contours of $\log_{10}[f_0(a, e)/\mu]$ for $\mu = 0.1$. The maximum surface density $\Sigma_0(R)$ occurs at $R = 0.8165$. Therefore, the highest phase-space density appears in the vicinity of $\log_{10}(a) \simeq 0$. Top: $K = 5$. Bottom: $K = 29$.

The functions $g_K(\mathcal{E})$ admit negative values for $K = 0, 1$, but they are positive-definite and therefore physical for plausible values of $\mu < 1$ when $K \geq 2$. We have plotted the contours of $\log_{10}(f_0/\mu)$ using (a, e) as independent variables in Fig. 2 for $K = 5$ and $K = 29$. The mean and rms eccentricity of the disc, $\bar{e}$ and e_{rms}, are given by

$$\bar{e} = \frac{\int e f_0(\boldsymbol{J}) \, \mathrm{d}^2 \boldsymbol{J}}{\int f_0(\boldsymbol{J}) \, \mathrm{d}^2 \boldsymbol{J}} = \frac{\Gamma(\frac{3}{2})\Gamma(\frac{5}{2} + K)}{\Gamma(3 + K)} + \mathcal{O}(\mu)$$

$$e_{\mathrm{rms}}^2 = \frac{\int e^2 f_0(\boldsymbol{J}) \, \mathrm{d}^2 \boldsymbol{J}}{\int f_0(\boldsymbol{J}) \, \mathrm{d}^2 \boldsymbol{J}} = \frac{2}{2K + 5} + \mathcal{O}(\mu). \tag{21}$$

Larger values of K correspond to colder discs. For $\mu \ll 1$ the mean eccentricity $\bar{e} = 0.329$ for $K = 5$ and $\bar{e} = 0.159$ for $K = 29$. When $K \gg 1$ the DF at a given energy or semi-major axis approaches the Schwarzschild or Rayleigh DF,

$$f_0(e^2)de^2 \propto \exp(-e^2/e_0^2)de^2, \quad e_0^{-2} = K + 1/2. \tag{22}$$

In this limit the mean and rms eccentricity are related to e_0 by $\bar{e} = \sqrt{\pi}e_0/2$, $e_{\mathrm{rms}} = e_0$.

A necessary condition for stability to small-scale axisymmetric disturbances is that Toomre's $Q > 1$; here $Q = \sigma_R \Omega_R/(3.36\Sigma_0)$ where σ_R is the radial velocity dispersion. The models in this paper with $\mu \ll 1$ have $Q > 0.5/\mu$ everywhere and thus are stable in this sense. The top two panels of Fig. 3 show the rms eccentricity and σ_R as functions of radius; for $\mu \ll 1$ these are independent of μ. The bottom panel shows μQ which is also independent of μ for $\mu \ll 1$. Note in particular that the rms eccentricity is almost independent of radius.
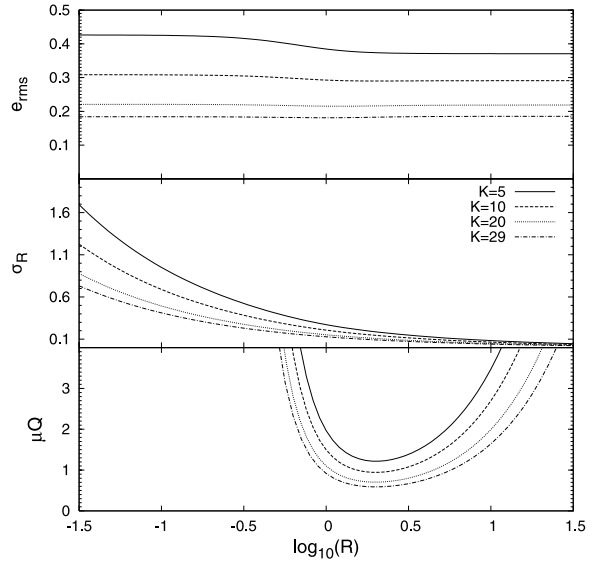
Figure 3. The rms eccentricity, radial velocity dispersion and μQ (the disc/star mass ratio times Toomre's stability parameter Q) as functions of radius. When $\mu \ll 1$ all three plots are independent of μ; the curves are from numerical models with $\mu = 0.1$.

4 PERTURBED DYNAMICS

We assume that gas drag, collisions and other non-gravitational effects are negligible so the disc can be treated as a collisionless fluid. We impose small-amplitude disturbances to the surface density, potential and DF:

$$\Sigma(R, \phi, t) = \Sigma_0(R) + \epsilon \Sigma_1(R, \phi, t), \tag{23}$$

$$V(R, \phi, t) = V_0(R) + \epsilon V_1(R, \phi, t) + \epsilon V_e(R, \phi, t), \tag{24}$$

$$f(\boldsymbol{w}, \boldsymbol{J}, t) = f_0(\boldsymbol{J}) + \epsilon f_1(\boldsymbol{w}, \boldsymbol{J}, t), \tag{25}$$

where $\epsilon \ll 1$ and V_e is an external perturbing potential, perhaps induced by a binary companion, an encounter with a passing star or the tidal field of the birth cluster. The perturbed surface density Σ_1 and its corresponding potential V_1 are related through Poisson's integral:

$$\begin{aligned} V_1(R, \phi, t) = &-G \iint \frac{\Sigma_1(R', \phi', t)R' \, \mathrm{d}R' \, \mathrm{d}\phi'}{\sqrt{R^2 + R'^2 - 2RR' \cos(\phi - \phi')}} \\ &+ GR \iint \frac{\Sigma_1(R', \phi', t)\cos(\phi - \phi') \, \mathrm{d}R' \, \mathrm{d}\phi'}{R'}, \end{aligned} \tag{26}$$

and we consider self-consistent density perturbations so that

$$\Sigma_1 = \int f_1 \, \mathrm{d}^2 \boldsymbol{v}. \tag{27}$$

The second term on the right-hand side of (26) is the indirect potential perturbation that arises because we are working in a non-inertial reference frame centred on the star. It is non-zero only for $m = 1$ perturbations since perturbations with $m \neq 1$ leave the centre of mass of the disc unchanged. For a particle with actions $\boldsymbol{J}$, the radial distance R and $\exp(\mathrm{i}m\phi)$ can be expanded as Fourier series in the angle variables,

$$R = \sum_{l=-\infty}^{+\infty} \xi_l(\boldsymbol{J})\mathrm{e}^{\mathrm{i}l w_R}, \quad \mathrm{e}^{\mathrm{i}m\phi} = \mathrm{e}^{\mathrm{i}m w_\phi}\sum_{l=-\infty}^{+\infty} \eta_l(\boldsymbol{J})\mathrm{e}^{\mathrm{i}l w_R}. \tag{28}$$

Any function of R and ϕ that is 2π-periodic in ϕ can thus be expressed in the $(\boldsymbol{w}, \boldsymbol{J})$ coordinates. For the Hamiltonian function that governs the motion of particles, we write

$$\mathcal{H} = \mathcal{H}_0(\boldsymbol{J}) + \epsilon V_1(\boldsymbol{w}, \boldsymbol{J}, t) + \epsilon V_{\mathrm{e}}(\boldsymbol{w}, \boldsymbol{J}, t), \tag{29}$$

where $\mathcal{H}_0$ is defined in equation (8). Therefore, the perturbed equations of motion become

$$\dot{\boldsymbol{w}} = \frac{\partial \mathcal{H}}{\partial \boldsymbol{J}} = \boldsymbol{\Omega}(\boldsymbol{J}) + \epsilon \frac{\partial}{\partial \boldsymbol{J}} (V_1 + V_{\mathrm{e}}), \tag{30}$$

$$\dot{\boldsymbol{J}} = -\frac{\partial \mathcal{H}}{\partial \boldsymbol{w}} = -\epsilon \frac{\partial}{\partial \boldsymbol{w}} (V_1 + V_{\mathrm{e}}). \tag{31}$$

It is obvious that the actions vary slowly in the perturbed disc. Subtracting the evolutionary equations of w_ϕ and w_R gives the apsidal precession rate in the perturbed disc,

$$\dot{w}_\phi - \dot{w}_R = \Omega_{\mathrm{pr}}(\boldsymbol{J}) + \epsilon \left(\frac{\partial}{\partial J_\phi} - \frac{\partial}{\partial J_R} \right) (V_1 + V_{\mathrm{e}}). \tag{32}$$

Since $\Omega_{\mathrm{pr}} = \mathcal{O}(\mu)$, for low-mass discs ($\mu \ll 1$) this equation contains two small parameters, μ and ϵ.

The DF in the perturbed disc obeys the CBE,

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + [f, \mathcal{H}] = 0, \tag{33}$$

where $[\cdot, \cdot]$ denotes a Poisson bracket. Here we confine ourselves to the linearized equation:

$$\frac{\partial f_1}{\partial t} + [f_1, \mathcal{H}_0] + [f_0, V_1] = -[f_0, V_{\mathrm{e}}]. \tag{34}$$

The remainder of this paper is devoted to the study of solutions of this equation and their application to collisionless near-Keplerian discs.

5 THE FINITE-ELEMENT METHOD

The dynamics and stability of collisionless discs are usually studied by one of two numerical methods: (i) N-body simulations (e.g. Sellwood 1987); (ii) expansion of the perturbed gravitational potential in a set of basis functions, followed by the evaluation of a matrix representing the response of the disc to a given imposed potential (e.g. Kalnajs 1977). Neither of these methods, however, is ideal for investigation of the oscillations and response of low-mass near-Keplerian discs, for several reasons: (i) slow oscillations are stable (T01) and therefore more difficult to detect than growing modes; (ii) slow oscillations have low frequencies, and thus N-body simulations must be followed for many dynamical times; (iii) low-mass discs also support short-wavelength fast (i.e. frequency independent of μ) oscillations and these cannot be resolved without a large set of basis functions; (iv) we shall find that some slow oscillations have nearly singular components. Here, we adopt an FEM and reduce the linearized CBE to a system of ordinary differential equations that describes the temporal evolution of the disc, both the eigenfrequency spectrum of an isolated disc and the response of a disc to external perturbations. We use a C_0 FEM (all functions are continuous, but not necessarily differentiable at boundaries between elements) in the configuration space.

In this section, we briefly review the principles of FEM modelling. For a general introduction see Zienkiewicz, Taylor & Zhu (2005). Detailed descriptions of the application of an FEM to collisionless self-gravitating systems can be found in Jalali (2010) for perturbed systems and in Jalali & Tremaine (2011) for equilibrium models.

5.1 Finite ring elements in the configuration space

We split the configuration space into N ring elements. The nth element is characterized by its nodes at R_n and R_{n+1}, and by a linear interpolating vector $\boldsymbol{G}_n(R)$ defined by

$$\boldsymbol{G}_n = \begin{bmatrix} G_{1,n} & G_{2,n} \end{bmatrix}, \quad G_{1,n} = 1 - \bar{R}, \quad G_{2,n} = \bar{R}, \tag{35}$$

where $\bar{R} = (R - R_n)/\Delta R_n$ and $\Delta R_n = R_{n+1} - R_n$. Since we are interested only in linear perturbations, disturbances of different azimuthal wavenumber m are independent. For the wavenumber m, the potential V_1 and the surface density Σ_1 are thus computed from

$$V_1(R, \phi, t) = \mathrm{Re} \sum_{n=1}^{N} H_n(R) \boldsymbol{G}_n(R) \cdot \boldsymbol{a}_n(t) \mathrm{e}^{\mathrm{i}m\phi}, \tag{36}$$

$$\Sigma_1(R, \phi, t) = \mathrm{Re} \sum_{n=1}^{N} H_n(R) \boldsymbol{G}_n(R) \cdot \boldsymbol{b}_n(t) \mathrm{e}^{\mathrm{i}m\phi}. \tag{37}$$

The function $H_n(R)$ is unity for $R_n \le R \le R_{n+1}$ and zero otherwise. The column vectors

$$\boldsymbol{a}_n = \begin{bmatrix} a_{n1} & a_{n2} \end{bmatrix}^{\mathrm{T}}, \quad \boldsymbol{b}_n = \begin{bmatrix} b_{n1} & b_{n2} \end{bmatrix}^{\mathrm{T}}$$

contain the nodal potentials and densities, respectively. According to the definition of $\boldsymbol{G}_n(R)$, Σ_1 is equal to $\mathrm{Re}\, b_{n1}\exp(\mathrm{i}m\phi)$ at $R = R_n$ and to $\mathrm{Re}\, b_{n2}\exp(\mathrm{i}m\phi)$ at $R = R_{n+1}$. Similarly, the nodal potentials at these radii involve a_{n1} and a_{n2}. The perturbed surface density and its corresponding potential are continuous and differentiable inside elements and the continuity of these functions at the boundaries of elements (nodes of rings) implies

$$a_{n2} = a_{n+1,1}, \quad b_{n2} = b_{n+1,1}. \tag{38}$$

This means that for a given m we have $N_{\mathrm{t}} = N + 1$ independent nodal potentials/densities.

The angle-action representation of the perturbed potential V_1 reads

$$V_1(\boldsymbol{w}, \boldsymbol{J}, t) = \mathrm{Re} \sum_{l=-\infty}^{+\infty} \tilde{h}_{1,l}(\boldsymbol{J}, t) \, \mathrm{e}^{\mathrm{i}(lw_R + mw_\phi)}, \tag{39}$$

where

$$\tilde{h}_{1,l}(\boldsymbol{J}, t) = \sum_{n=1}^{N} \boldsymbol{\Psi}_l(n, \boldsymbol{J}) \cdot \boldsymbol{a}_n(t), \tag{40}$$

$$\boldsymbol{\Psi}_l(n, \boldsymbol{J}) = \frac{1}{2\pi} \oint H_n(R) \boldsymbol{G}_n \, \mathrm{e}^{\mathrm{i}m(\phi - w_\phi)} \mathrm{e}^{-\mathrm{i}lw_R} \, \mathrm{d}w_R. \tag{41}$$

The external disturbance V_{e} can also be expressed in terms of angle and action variables. To compute the perturbed DF $f_1(\boldsymbol{w}, \boldsymbol{J}, t)$, we use Fourier series of angle variables and write

$$f_1(\boldsymbol{w}, \boldsymbol{J}, t) = \mathrm{Re} \sum_{n=1}^{N} \sum_{l=-\infty}^{\infty} \boldsymbol{E}_l(n, \boldsymbol{J}) \cdot \boldsymbol{z}_l^n(t) \, \mathrm{e}^{\mathrm{i}(lw_R + mw_\phi)}, \tag{42}$$

where

$$\boldsymbol{E}_l(n, \boldsymbol{J}) = \begin{bmatrix} E_{l1}(n, \boldsymbol{J}) & E_{l2}(n, \boldsymbol{J}) \end{bmatrix} \tag{43}$$

is an interpolating vector in the action space (to be specified in Section 5.3) and

$$\boldsymbol{z}_l^n = \begin{bmatrix} z_{l1}^n & z_{l2}^n \end{bmatrix}^{\mathrm{T}} \tag{44}$$

is a column vector of to-be-determined DFs whose elements should satisfy the continuity condition

$$z_{l2}^n = z_{l1}^{(n+1)}. \tag{45}$$

Equation (42) calculates the distribution of perturbed orbits based on their passage through ring elements in the configuration space. If an orbit stays only inside the nth element, its DF becomes

$$\hat{f}_n(\boldsymbol{w}, \boldsymbol{J}, t) = \text{Re} \sum_{l=-\infty}^{\infty} \boldsymbol{E}_l(n, \boldsymbol{J}) \cdot z_l^n(t) \, e^{i(l w_R + m w_\phi)}. \quad (46)$$

In general, eccentric orbits may visit more than one ring element. The summation over n in (42) takes this behaviour into account.

5.2 Projected evolutionary equations

We use the conditions (38) and assemble the nodal densities $\boldsymbol{b}_n(t)$ and potentials $\boldsymbol{a}_n(t)$ in the global N_t-dimensional vectors $\boldsymbol{d}(t)$ and $\boldsymbol{p}(t)$, respectively. Similarly, $z_l^n(t)$ are collected in $\boldsymbol{z}_l(t)$. We now take the inner product of (34) with

$$e^{-i(l' w_R + m w_\phi)} [\boldsymbol{E}_{l'}(n', \boldsymbol{J})]^T,$$

and integrate the resulting systems of equations over the angle-action space to obtain the Galerkin-weighted residual form of (34) as

$$\mathbf{U}_1(l) \cdot \frac{\mathrm{d}\boldsymbol{z}_l(t)}{\mathrm{d}t} = -i\mathbf{U}_2(l) \cdot \boldsymbol{z}_l(t) + i\mathbf{U}_3(l) \cdot \boldsymbol{p}(t) + i\boldsymbol{Z}_l(t). \quad (47)$$

Here $\mathbf{U}_1$, $\mathbf{U}_2$ and $\mathbf{U}_3$ are constant square matrices of dimension $N_t \times N_t$, and $\boldsymbol{Z}_l(t)$ is an N_t-dimensional column vector, which is the Galerkin projection of $-[f_0, V_e]$.

One can also verify that the Galerkin projections of (26) and (27), respectively, become

$$\boldsymbol{p}(t) = \mathbf{C} \cdot \boldsymbol{d}(t), \quad \boldsymbol{d}(t) = \sum_l \mathbf{F}(l) \cdot \boldsymbol{z}_l(t). \quad (48)$$

The constant matrices $\mathbf{C}$ and $\mathbf{F}(l)$ are of dimension $N_t \times N_t$. We combine (47) and (48) to express $\boldsymbol{p}(t)$ in terms of $\boldsymbol{z}_l(t)$, and transform (47) to a non-homogeneous ordinary differential equation for $\boldsymbol{z}_l(t)$:

$$\frac{\mathrm{d}\boldsymbol{z}_l(t)}{\mathrm{d}t} = -i\mathbf{U}_1^{-1}(l) \cdot \mathbf{U}_2(l) \cdot \boldsymbol{z}_l(t) + i\mathbf{U}_1^{-1}(l) \cdot \boldsymbol{Z}_l(t)$$

$$+ \sum_{l'=-\infty}^{+\infty} i\mathbf{U}_1^{-1}(l) \cdot \mathbf{U}_3(l) \cdot \mathbf{C} \cdot \mathbf{F}(l') \cdot \boldsymbol{z}_{l'}(t), \quad (49)$$

for $l, l' = 0, \pm 1, \pm 2, \dots$. By defining

$$\boldsymbol{z}(t) = \begin{bmatrix} \dots & \boldsymbol{z}_{-2}^T & \boldsymbol{z}_{-1}^T & \boldsymbol{z}_0^T & \boldsymbol{z}_{+1}^T & \boldsymbol{z}_{+2}^T & \dots \end{bmatrix}^T, \quad (50)$$

and collecting the elements of $\mathbf{U}_1^{-1}(l) \cdot \boldsymbol{Z}_l(t)$ (for all $l = 0, \pm 1, \pm 2, \cdots$) in the global forcing vector $\boldsymbol{F}(t)$, the system (49) can be cast into the standard form of linear evolutionary equations:

$$\frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{z}(t) = -i\mathbf{A} \cdot \boldsymbol{z}(t) + i\boldsymbol{F}(t). \quad (51)$$

In the absence of external disturbances, $\boldsymbol{F}(t) = 0$, the corresponding homogeneous equation admits a solution of the form $\boldsymbol{z}(t) = \exp(-i\omega t)\boldsymbol{c}$ that yields the linear eigensystem:

$$\mathbf{A} \cdot \boldsymbol{c} = \omega \boldsymbol{c}. \quad (52)$$

We find the spectrum of ω using Hessenberg transformation of $\mathbf{A}$ followed by QR factorization. The eigenvector conjugate to a given eigenfrequency ω_j is then computed using the singular value decomposition

$$\mathbf{A} - \omega_j \mathbf{I} = \mathbf{V}_1^T \cdot \mathbf{W} \cdot \mathbf{V}_2, \quad (53)$$

where $\mathbf{W}$ is a diagonal matrix whose elements are the singular values of $\mathbf{A} - \omega_j \mathbf{I}$, and $\mathbf{I}$ is the identity matrix. The column of $\mathbf{V}_2$ corresponding to the smallest singular value is the eigenvector $\boldsymbol{z}^{(j)}$ associated with ω_j.

5.3 Interpolating functions in the action space

In our C_0 FEM analysis, the local interpolating vector functions $\boldsymbol{G}_n(R)$ (also known as shape functions) can reconstruct the spatial profile of any oscillatory wave whose wavelength is sufficiently large compared to the sizes of elements. However, we must also interpolate f_1 in the action space, which requires defining the interpolating vectors $\boldsymbol{E}_l(n, \boldsymbol{J})$ (equation 42). To do this we use arbitrary dynamic solutions of the linearized CBE, which should be an adequate representation of the DF for the purposes of interpolation. In the angle-action space, and using equation (39), one can show

$$e^{im\phi} \boldsymbol{G}_n(R) = \tilde{V}_1(n, \boldsymbol{J}, \boldsymbol{w}) = \sum_{l=-\infty}^{+\infty} \boldsymbol{\Psi}_l(n, \boldsymbol{J}) \, e^{i l w_R + i m w_\phi}. \quad (54)$$

We assume $\partial f_1 / \partial t = -i\gamma f_1$, substitute $\tilde{V}_1(n, \boldsymbol{J}, \boldsymbol{w})$ into the linearized CBE and solve the resulting equation to obtain the interpolating vector in action space (cf. equation 42)

$$\boldsymbol{E}_l(n, \boldsymbol{J}) = \frac{l \partial f_0 / \partial J_R + m \partial f_0 / \partial J_\phi}{l \Omega_R + m \Omega_\phi - \gamma} \boldsymbol{\Psi}_l(n, \boldsymbol{J}). \quad (55)$$

The physical eigenfrequency ω will thus be equal to $\omega_c + \gamma$ with ω_c being the computed eigenvalue of (52). Varying γ tests the robustness of our numerical methods since the results should be independent of γ. Our tests show that in general our results are insensitive to variations of γ. Nevertheless, the choice $\gamma \gtrsim \max[\Omega_{pr}(\boldsymbol{J})]$ offers better performance, particularly in colder models. We originally used $\gamma = 0$, corresponding to the use of static CBE solutions as interpolating vectors in action space, but with this choice we found occasional spurious growing modes.

We remark that we do not generate a finite-element mesh in action space for two reasons: (i) to reduce the size of the Galerkin-weighted evolutionary equations; and (ii) to avoid creating spurious growing modes. The second of these properties has a straightforward mathematical explanation: the number of reachable eigenmodes in the configuration space is equal to the number of independent nodal variables, which is $N + 1$ in our FEM analysis. However, the eigenvalue problem (52) has been formulated in the phase space and the number of computed eigenmodes is equal to $(N + 1) \times (l_{\max} - l_{\min} + 1)$ where $l_{\min} < 0$ and $l_{\max} > 0$ are the lower and upper bounds in the l sums. Since the nodal densities $\boldsymbol{d}$ are related to z_l through equation (48), there will be $(N + 1) \times (l_{\max} - l_{\min})$ computed eigenmodes more than $N + 1$ eigenmodes that the dimension of $\boldsymbol{d}$ determines. The extra modes should therefore overlap in groups of $(l_{\max} - l_{\min})$ members to avoid spurious modes. This happens in our numerical calculations performed in Section 6 when the Fourier expansions over w_R converge inside all ring elements, as is expected for the reconstruction of V_1 and f_1 in the (w_R, w_ϕ)-subspace. Generating a finite-element mesh, let us say with N_a nodes in the two-dimensional $\boldsymbol{J}$-space, will result in $2N_a \times (l_{\max} - l_{\min} + 1)$ modes, but assuming the convergence of Fourier series, only $(N + 1)$ groups of them will correspond to eigenmodes in the configuration space. Consequently, $2N_a - N + 1$ computed modes will be spurious, and our calculations show that such spurious modes are growing. Working with $2N_a = N + 1$ will not help because it does not necessarily guarantee the convergence of FEM model in the action space.

Only few modes out of $N + 1$ possible states in the configuration space (see Section 6) are physical. The rest are either singular, or do not satisfy the boundary conditions as $R \to \infty$. Note that for frequencies ω that lie between the maximum and minimum of the precession frequency Ω_{pr} the singular modes may be van Kampen modes (restricted to the surface in action space on which $\omega = \Omega_{pr}$)

which would be damped by the Landau mechanism. However, not all modes with frequency in this range are necessarily van Kampen modes, since the mode may not be produced by the orbits associated with that resonance. Thus discrete modes may overlap in frequency space with continuous modes.

6 PROGRADE WAVES

Our finite-element mesh is uniform in log radius,

$$R_n = 10^{-\alpha_1 + \alpha_2 y(n, N)}, \tag{56}$$

$$y(n, N) = \frac{1}{2(N+1)} + \frac{n-1}{N+1}, \quad n = 1, 2, \ldots, N. \tag{57}$$

The numerical computation of the Fourier coefficients $\boldsymbol{\Psi}_l(n, \boldsymbol{J})$ and then the interpolating vectors $\boldsymbol{E}_l(n, \boldsymbol{J})$ needs a mesh in the (a, e) space. Such a grid is not arbitrary because there must be at least one orbit that visits the nth ring element in configuration space. We fulfil this requirement using the following two-dimensional grid

$$(a_i, e_j) = [R_i, y(j, M_e)],$$

where the grid points along the a-direction exactly coincide with the boundary nodes of the mesh in the configuration space, and there are $j = 1, 2, \ldots, M_e + 1$ grid points in the e-direction. A circular orbit is indeed assigned to each boundary of a ring element. This is particularly helpful in cold discs where one must interpolate the population of circular orbits. The parameters α_1 and α_2 are chosen so that the computed disc mass $4\pi^2 \int f(\boldsymbol{J}) \, d\boldsymbol{J}$ using the grid points in the (a, e)-space agrees with the actual disc mass within 1 per cent.

In this paper we focus on slow modes with azimuthal wavenumber $m = 1$. Slow modes exist with larger m, so long as the calculation includes Fourier terms with index $l = -m$. In particular, we have found a number of isolated, non-singular $m = 2$ modes; these are present only if we use a fine FEM mesh, since they are more compact and have shorter wavelengths than the $m = 1$ modes. The wavelengths of $m = 2$ modes shrink to zero as the disc becomes colder (see Appendix). This behaviour is expected since the only large-scale slow modes in cold low-mass discs have $m = 1$. We found no unstable modes, which is also expected for low-mass discs (T01).

We began our calculations with $N = M_e = 70$ and $l = -1$, and increased the number of Fourier terms and ring elements until the eigenfrequencies of stable modes found from (52) converged to a fractional accuracy of 10^{-4}. Typically this required computing all Fourier terms with $-2 \le l \le 3$ and a grid with $N = 160$ and $M_e = 140$ ($N = 180$ and $M_e = 140$ for the models with the lowest rms eccentricity, corresponding to $K = 29$). We have also experimented with including terms with larger values of $|l|$ but these had only a small effect on our results. Taking grid points in the regions with tiny values of $f(a, e)$ (see Fig. 2) leads to large errors in the properties of the calculated density waves because the FEM discretization errors become larger than the absolute magnitudes of physical quantities. We evade this difficulty by generating the FEM mesh only in the annular region $0.01 \le R \le 100$ using the parameters $(\alpha_1, \alpha_2) = (2, 4)$ in equation (57).

All non-singular eigenmodes with $m = 1$ were found to be prograde ($\omega > 0$). We find two general types of modes: a parent family that is already present when only the $l = -1$ Fourier component is included in the calculation, and a child family that bifurcates from the parent family as more l-terms are included. The eigenfrequencies of child modes are very close to those of their parent mode

(typically within 1–2 per cent). They emerge from resonant interactions between two approximate modes that are weakly coupled: the parent modes and singular van Kampen modes. For $l = 0$ and $l = +1$ the singular components of the child modes correspond to the corotation resonance (CR) and outer Lindblad resonance (OLR), respectively. The coupling between slow and van Kampen modes is probably due mostly to highly eccentric orbits that are perturbed by the gravity from both waveforms. The main evidence for this is that as the mean eccentricity $\bar{e}$ shrinks, child modes collapse to singular modes and disappear.

We denote the maximum precession rate of circular orbits by $\omega_0 = \max[\Omega_{\mathrm{pr}}(\boldsymbol{J})]$; from equation (12) $\omega_0 = 0.05861\mu + \mathcal{O}(\mu^2)$. We then plot our results using the normalized frequency $\bar{\omega} = \omega/\omega_0$. Fig. 4 shows the eigenfrequency spectra of prograde $m = 1$ parent modes for the mass ratios $\mu = 0.025$ and $\mu = 0.05$, and for four
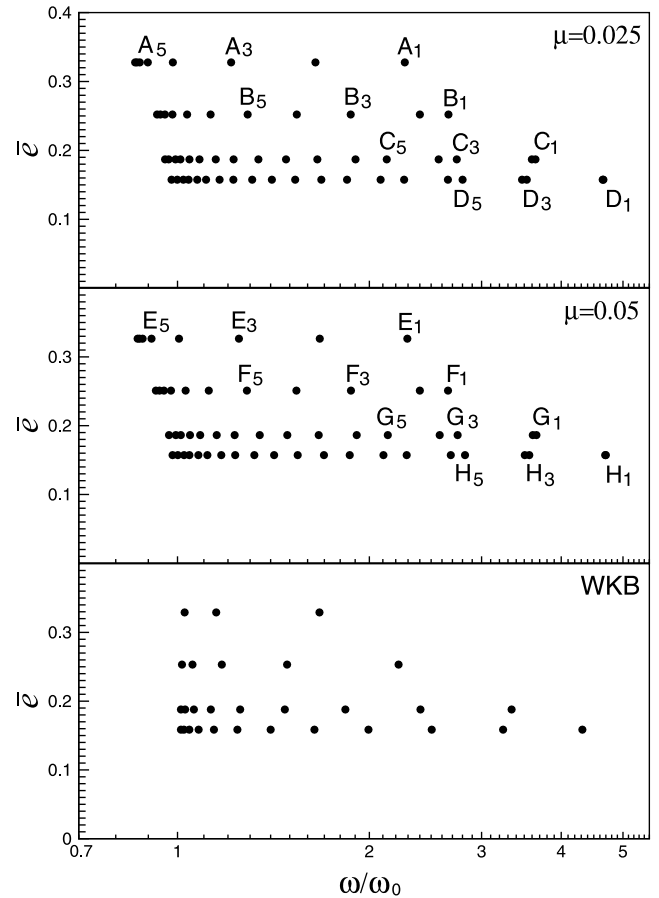
Figure 4. Eigenfrequency spectra of stable, prograde density waves in near-Keplerian discs. Only parent modes are shown. The vertical axis is the mean eccentricity $\bar{e}$ and the horizontal axis is $\bar{\omega} = \omega/\omega_0$ where $\omega_0 = 0.05861\mu + \mathcal{O}(\mu^2)$. These models correspond to DFs of the form (14) with $K = 5, 10, 20, 29$. The calculation includes Fourier terms $l = -2, -1, 0, 1$. Note the logarithmic scale of the horizontal axis. Top: $\mu = 0.025$ and $\omega_0 = 0.00146$. Middle: $\mu = 0.05$ and $\omega_0 = 0.00293$. The eigenfrequencies of modes D_1 and D_2 are very close and indistinguishable in the plots. They are $\bar{\omega}_{D_1} = 4.659$ and $\bar{\omega}_{D_2} = 4.647$. Similarly, we have $\bar{\omega}_{H_1} = 4.700$ and $\bar{\omega}_{H_2} = 4.689$. Note the similarity of the spectra in the top and middle diagrams despite the change of a factor of 2 in the disc mass μ; this feature is characteristic of slow modes. Mode shapes associated with the labelled frequencies have been plotted in Figs 5, 6 and 7. Bottom: eigenfrequency spectra derived from the WKB approximation described in the Appendix. Each plotted point represents a degenerate leading/trailing pair of modes.

values of the mean eccentricity $\bar{e}$. The frequencies of child modes are not shown to avoid overcrowding the diagrams. Although the maximum precession rate ω_0 is proportional to μ, the spectra of $\bar{\omega}$ agree to within 1 per cent in the models with $\mu = 0.025$ and $\mu = 0.05$. This scaling shows that our results can be directly applied to all discs with mass ratios $\mu \ll 1$, in particular to the tiny mass ratios $\mu \lesssim \mathcal{O}(10^{-3})$ characteristic of debris discs.

Fig. 4 shows that the modes become more closely spaced as their frequency decreases and the minimum frequency in each spectrum is an accumulation point. This implies the existence of prograde waves with arbitrarily short wavelengths. There is also a nice correlation between the precession rate of the most eccentric orbit in the model (see Figs 1 and 2) and the lowest frequency in the spectrum. Models with highly eccentric orbits have an accumulation point of lower frequency. Fig. 4 shows that the number of modes increases with decreasing $\bar{e}$. In the limit of $\bar{e} \to 0$, however, dispersion-supported waves (or *p* modes) cannot exist according to WKB results (T01).

The frequency spacing between modes B_1 and B_2 is larger than the spacing between C_1 and C_2, which in turn is larger than the spacing between D_1 and D_2 (which is so small that the two points are indistinguishable in the figure). Similar behaviour is seen in the F, G and H families in the middle panel of the figure. In the limit $\bar{e} \to 0$, the parent modes tagged with the numbers $2j + 1$ and $2j + 2$ ($j = 0, 1, 2, \ldots$) become degenerate. In the language of T01, they form a degenerate leading/trailing pair of *p* modes (see Appendix). The pairing process begins from modes with highest pattern speeds, for the resonant cavities of those modes are fed mostly by near-circular orbits, which are the only population used in the WKB analysis. The child modes of degenerate pairs also disappear because their supporting eccentric orbits disappear as $\bar{e} \to 0$. Modes with $\bar{\omega} \to 1$ and sufficiently large $\bar{e}$ engage highly eccentric orbits and thus lead to more complex dynamics. Eccentric orbits are indeed the backbones of discs, and when perturbed, they affect a vast radial domain while near-circular orbits have only a local influence on developing patterns.

We now examine the shapes of the modes. After finding ω, we calculate its corresponding eigenvector c, and use this to compute the nodal potentials p and nodal densities d from (48). Defining

$$X(R) = \text{Re} \sum_{n=1}^{N} H_n(R) \mathbf{G}_n(R) \cdot \mathbf{b}_n, \tag{58}$$

$$Y(R) = \text{Im} \sum_{n=1}^{N} H_n(R) \mathbf{G}_n(R) \cdot \mathbf{b}_n, \tag{59}$$

one can compute the perturbed density patterns

$$\Sigma_1(R, \phi, t) = X(R) \cos(m\phi - \omega t) - Y(R) \sin(m\phi - \omega t) \tag{60}$$

for a single wavenumber m. Note that $\mathbf{b}_n$ are extracted from the elements of d using the following formula:

$$\mathbf{b}_n = [\, d_n \quad d_{n+1} \,]^{\text{T}}, \quad n = 1, 2, \ldots, N. \tag{61}$$

Fig. 5 shows the profile of $X(R)$ for the labelled parent modes of Fig. 4. Not only are the normalized frequencies of the modes A_j and E_j identical, but also their mode shapes are very similar. These remarks apply to the pairs (B_j, F_j), (C_j, G_j) and (D_j, H_j) as well, and demonstrate that the waveforms are independent of μ so long as $\mu \ll 1$, as one would expect for slow modes. The figure also shows that the wavelength of oscillations increases with the pattern speed ω in a given disc, and decreases as the mean eccentricity of the disc shrinks. The number of nodes increases as the frequency decreases. An interesting property of the waves showing multiple nodes is that

their density peaks are approximately equally spaced in logarithmic scales.

The child modes are hybrid modes that inherit the features of their parents in the central regions of the disc, but have a spike at the location of singular modes that couple to them. Fig. 6 displays the parent mode D_8 of frequency $\bar{\omega} = 2.083$ (see Fig. 4) and its children $D_{8,\text{CR}}$ with $\bar{\omega} = 2.0765$ and $D_{8,\text{OLR}}$ with $\bar{\omega} = 2.0728$, which contain singular van Kampen modes at the CR and OLR, respectively. In low-mass discs, these resonances are at large radii where the surface density is small, so the singular component of a child mode involves only a small fraction of the mass involved in the parent mode. As the disc mass shrinks to zero the child modes merge with the parent mode. The reason is that the eigenfrequency of the parent mode is proportional to the disc mass so with very small disc masses the CR and OLR are at extremely large radii where the surface density is negligible. Thus the distinction between parent and child modes is unimportant for low-mass discs such as debris discs.

Fig. 7 displays shaded contour plots of the pattern of $\Sigma_1(R, \phi, t)$ for some models with $\mu = 0.025$ (mode shapes corresponding to $\mu = 0.05$ are similar). It is seen that the wave packets are more radially compact in the colder ($K = 29$) model than warmer ($K = 5, 10$) ones.

The properties of these modes can be explored using the short-wavelength or WKB approximation described in the Appendix. The validity of this approximation requires $k > h/R$ where k is the wavenumber and h is a dimensionless number of order unity. If two adjacent nodes of a wave are at R_1 and R_2 then $\int_{R_1}^{R_2} k \, dR = \pi$ so the condition for validity of the WKB approximation may be written as $\pi > h \log R_2/R_1$ or $\log_{10} R_2/R_1 < 1.36/h$. Inspection of Fig. 6 shows that for $h = 1$ this condition is satisfied by the majority of modes we have computed.

The bottom panel of Fig. 4 shows the WKB frequency spectrum. Each point corresponds to a degenerate pair of modes, one composed of leading spiral waves and the other of trailing. These modes arise from waves in the resonant cavities defined by the closed frequency contours in Fig. A1. The WKB approximation correctly reproduces several striking features of the FEM frequency spectra: (i) all modes are prograde ($\omega > 0$); (ii) both the number of modes and the maximum frequency grow as the mean eccentricity $\bar{e}$ of the disc shrinks; (iii) there is an accumulation point of modes near $\omega/\omega_0 = 1$ in the FEM spectra and at $\omega/\omega_0 = 1$ in the WKB spectra; (iv) there is also reasonable quantitative agreement between the frequencies derived by the two methods, at least for the discs with the lowest mean eccentricity. The WKB analysis in the Appendix fails to find the child modes for two reasons: (i) it is based on the epicycle approximation, which assumes that the eccentricity is small and thus neglects the highly eccentric orbits that couple the slow and van Kampen modes; (ii) it is based on the approximation that the disc mass $\mu \to 0$, and in this limit the pattern speed of the slow mode goes to zero so the OLR and CRs are at very large radii where the disc surface density is negligible.

7 EXCITATION OF OSCILLATORY WAVES

Protostars live in the harsh environments of their birth clusters. Simulations of the Orion nebula (Scally & Clarke 2001) show that about 10 per cent of stars can have encounters closer than 100 au within 10^7 years. Such encounters can excite waves in planetesimal/debris discs. Encounters were invoked as a possible explanation for the asymmetries in the β Pictoris debris disc by Kalas & Jewitt (1995) and Larwood & Kalas (2001) but these authors treated the
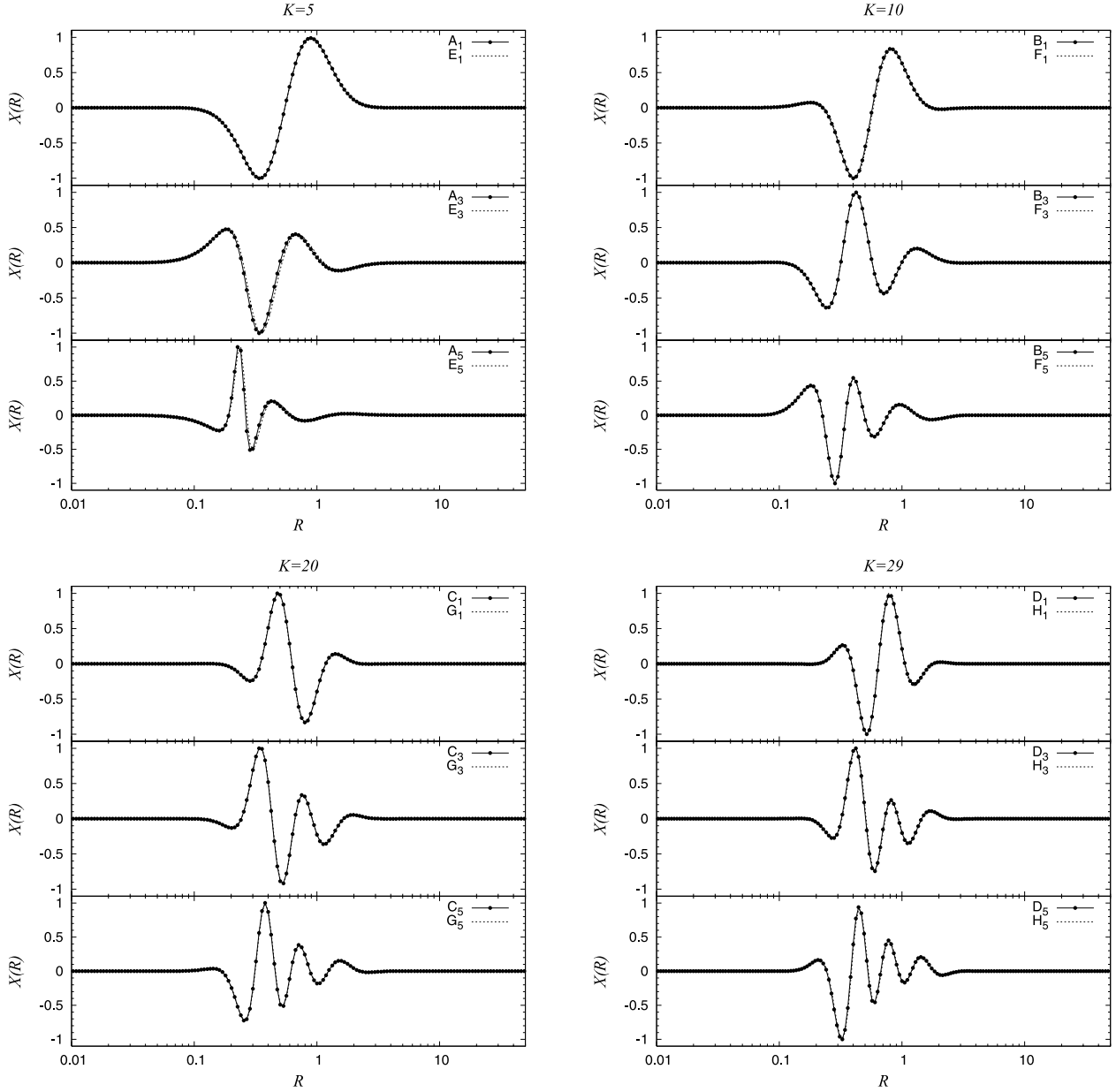
Figure 5. Perturbed density components $X(R)$ (cf. equation 60) for some stable modes in near-Keplerian discs of $\mu = 0.025$ (solid lines) and $\mu = 0.05$ (dotted lines). In several panels the dotted line is not visible because it lies under the solid line. There are $N = 160$ ring elements in the configuration space for $K = 5$, 10 and 20, and $N = 180$ elements for $K = 29$. Filled circles mark the locations of element nodes. In all panels, the maximum of $|X(R)|$ has been normalized to unity.

debris disc as a collection of test particles, which can give misleading results since the self-gravity of the disc dominates the apsidal precession.

Since our goal here is only to illustrate this process we confine ourselves to in-plane parabolic encounters. Consider a disc particle orbiting around a star of mass $M_\star$, and assume a perturber of mass M_p. As in earlier sections, we scale all lengths so that the disc length scale b is unity, and denote the normalized position vectors of the particle and perturber (with respect to the host star) by $\boldsymbol{R}$ and $\boldsymbol{R}_p$, respectively. The equation of motion for a disc particle is

$$\frac{\mathrm{d}^2 \boldsymbol{R}}{\mathrm{d}t^2} = -\nabla \left[\boldsymbol{a}_\star \cdot \boldsymbol{R} + V_0(\boldsymbol{R}) + V_1(\boldsymbol{R}, t) + V_p(\boldsymbol{R}, \boldsymbol{R}_p) \right], \qquad (62)$$

where $\boldsymbol{a}_\star$ is the acceleration vector of the host star in an inertial frame, V_0 is the unperturbed gravitational potential due to $M_\star$ and the self-gravity of the disc, V_1 is the perturbed self-gravitational potential of the disc, and $V_p = -(M_p/M_\star)/|\boldsymbol{R}_p - \boldsymbol{R}|$ is the potential field of the perturber. The gradient ∇ is taken over the $\boldsymbol{R}$ space, and the normalized time t is related to the actual time t_{actual} through $t/t_{\mathrm{actual}} = (GM_\star/b^3)^{1/2}$.

We assume that $\boldsymbol{a}_\star$ is due to the encounter; thus we ignore the cluster's tidal field. Consequently,

$$\boldsymbol{a}_\star \cdot \boldsymbol{R} = \left(\frac{M_p}{M_\star} \right) \frac{\boldsymbol{R}_p \cdot \boldsymbol{R}}{R_p^3}, \quad R_p = |\boldsymbol{R}_p|. \qquad (63)$$
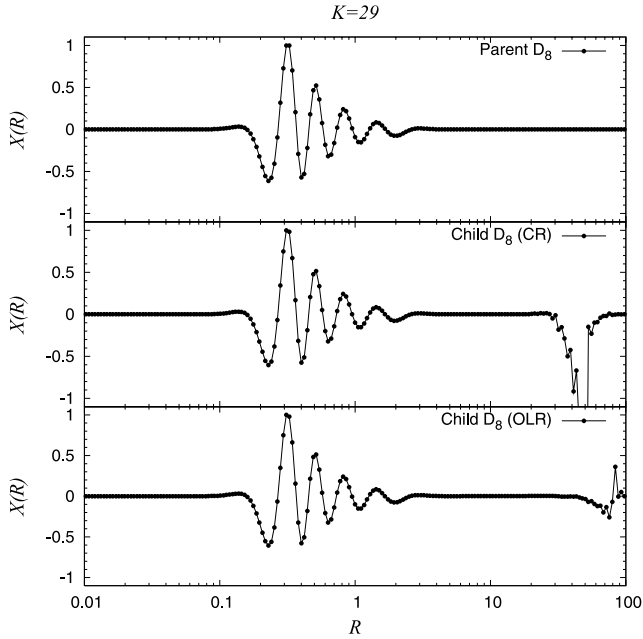
Figure 6. Perturbed density components $X(R)$ for the parent mode D_8 and two of its associated child modes $D_{8,CR}$ and $D_{8,OLR}$, in which the parent mode is coupled to singular modes at the CR and OLR. The model parameters are $\mu = 0.025$ and $K = 29$.

For a distant encounter, $R \ll R_p$, the potential V_p can be expanded as the following series:

$$V_p = -\left(\frac{M_p}{M_\star}\right)\frac{1}{R_p}\sum_{i=0}^{\infty}\left(\frac{R}{R_p}\right)^i P_i\left[\cos\left(\phi - \phi_p\right)\right],$$

$$\cos\left(\phi - \phi_p\right) = \frac{\boldsymbol{R_p} \cdot \boldsymbol{R}}{R_p R}, \tag{64}$$

where ϕ and ϕ_p are, respectively, the azimuths of the disc particle and perturber measured from an inertial reference line, and P_i are Legendre polynomials. The effective potential due to a flying-by

perturber thus reads

$$V_e = \boldsymbol{a_\star} \cdot \boldsymbol{R} + V_p,$$

$$= -\left(\frac{M_p}{M_\star}\right)\frac{1}{R_p}\sum_{i=2}^{\infty}\left(\frac{R}{R_p}\right)^i P_i\left[\cos\left(\phi - \phi_p\right)\right], \tag{65}$$

where we have dropped the $i = 0$ term in (64) because it makes no contribution to the force, and the $i = 1$ term has been cancelled by $\boldsymbol{a_\star} \cdot \boldsymbol{R}$.

Modes having azimuthal wavenumber $m = 1$ can only be excited by those $i \geq 3$ terms of V_e that produce $\cos\phi$ and $\sin\phi$ factors. Modes with $m = 2$ are excited by the $i = 2$ term of V_p, which is much larger than the $i \geq 3$ terms for distant perturbers ($R/R_p \ll 1$); however, we have found (Section 6) that slow modes with $m = 2$ have wavelengths that are generally smaller than those of $m = 1$ modes, even for discs with a relatively large mean eccentricity $\bar{e}$, and which shrink to zero as $\bar{e} \rightarrow 0$. Thus $m = 2$ modes couple less effectively to smooth perturbing potentials. We conclude that the dominant slow mode excited by an external perturber may have either $m = 1$ or $m = 2$.

For brevity, we shall examine only $m = 1$ modes here. The dominant term of (65) for $m = 1$ perturbations is

$$V_e \simeq -\frac{3}{8}\left(\frac{M_p}{M_\star}\right)\left(\frac{1}{R_p}\right)\left(\frac{R}{R_p}\right)^3 \cos\left(\phi - \phi_p\right). \tag{66}$$

This can be expressed in the angle-action variables as (cf. equation 39)

$$V_e \simeq \text{Re} \sum_{l=-\infty}^{+\infty} Q(t)\tilde{h}_{e,l}(\boldsymbol{J})e^{i(lw_R + w_\phi)}, \tag{67}$$
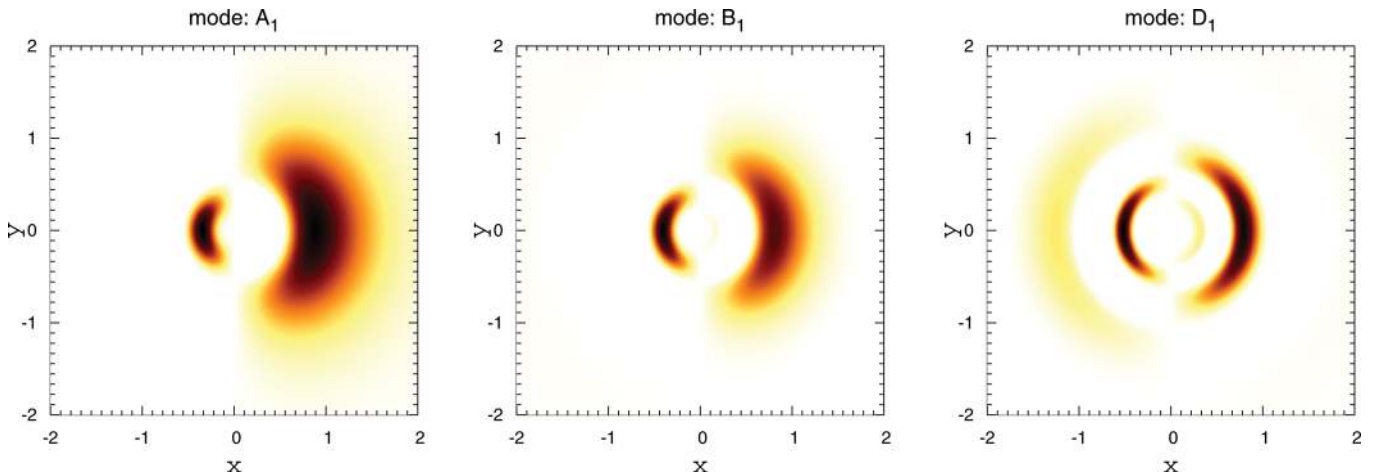
where

$$Q(t) = -\frac{3}{8}\left(\frac{M_p}{M_\star}\right)\frac{1}{[R_p(t)]^4}e^{-i\phi_p(t)} \tag{68}$$

is the time-varying part of the external perturbation, and

$$\tilde{h}_{e,l}(\boldsymbol{J}) = \frac{1}{2\pi}\oint R^3 \cos\left[lw_R + (w_\phi - \phi)\right] \, dw_R. \tag{69}$$

For a parabolic encounter with minimum distance $R_{p,\text{min}}$ and gravity parameter $\bar{M} = 1 + M_p/M_\star$, the true anomaly ϕ_p and radial distance R_p are computed through the following equations:

$$t(\phi_p) = \frac{\sqrt{2}}{\omega_p}\left[\tan\left(\frac{\phi_p}{2}\right) + \frac{1}{3}\tan^3\left(\frac{\phi_p}{2}\right)\right], \tag{70}$$

Figure 7. The patterns of oscillatory waves in the configuration space for a near-Keplerian disc with the mass ratio $\mu = 0.025$. We have displayed only the positive part of $\Sigma_1(R, \phi, t)$ at $t = 0$. Maximum densities of all panels have been normalized to unity, and the contour levels range from 0 to 1. The point mass sits at (0, 0). Left: $K = 5$. Middle: $K = 10$. Right: $K = 29$.

$$R_{\mathrm{p}} = \frac{2R_{\mathrm{p,min}}}{1 + \cos(\phi_{\mathrm{p}})}, \quad \bar{M} = \omega_{\mathrm{p}}^2 R_{\mathrm{p,min}}^3. \tag{71}$$

The effect of the external perturbation on the evolution of f_1 inside the nth element is determined by the Galerkin projection of the Poisson bracket $-[f_0, V_{\mathrm{e}}]$ as

$$4\pi^2 \mathrm{i} \mathbf{Z}_{l'}^{n'}(t) = -\int\int \mathbf{E}_{l'}^{\mathrm{T}}(n', \mathbf{J})[f_0, V_{\mathrm{e}}]$$
$$\times \mathrm{e}^{-\mathrm{i}(l' w_R + w_\phi)} \, \mathrm{d}^2 \mathbf{J} \, \mathrm{d}^2 \mathbf{w}. \tag{72}$$

Substituting from (67) into (72) and performing the integral over the angle space give the two-vector

$$\mathbf{Z}_l^n(t) = Q(t) \int \left(l \frac{\partial f_0}{\partial J_R} + \frac{\partial f_0}{\partial J_\phi} \right) \tilde{h}_{e,l}(\mathbf{J}) \mathbf{E}_l^{\mathrm{T}}(n, \mathbf{J}) \, \mathrm{d}^2 \mathbf{J}, \tag{73}$$

whose components ($Z_{1,l}^n$ and $Z_{2,l}^n$) are, respectively, the contribution of the disturbing force to the inner and outer nodes of the nth ring element in the configuration space. The disturbance at the jth ring node thus reads

$$Z_{j,l}(t) = \begin{cases} Z_{1,l}^j, & j = 1, \\ Z_{1,l}^j + Z_{2,l}^{j-1}, & 1 < j < N+1, \\ Z_{2,l}^N, & j = N+1, \end{cases} \tag{74}$$

and we obtain

$$\mathbf{Z}_l(t) = \begin{bmatrix} Z_{1,l} & Z_{2,l} & \ldots & Z_{N,l} & Z_{(N+1),l} \end{bmatrix}^{\mathrm{T}}. \tag{75}$$

Defining $\mathbf{F}_l = \mathbf{U}_1^{-1}(l) \cdot \mathbf{Z}_l$, the global forcing vector is assembled as

$$\mathbf{F}(t) = \begin{bmatrix} \ldots & \mathbf{F}_{-1}^{\mathrm{T}}(t) & \mathbf{F}_0^{\mathrm{T}}(t) & \mathbf{F}_{+1}^{\mathrm{T}}(t) & \ldots \end{bmatrix}^{\mathrm{T}},$$
$$= Q(t)\mathbf{g}, \tag{76}$$

where $\mathbf{g}$ is a constant vector. For each Fourier number l, we have N_t unknown DFs collected in the vector $z_l(t)$. The Fourier series in terms of w_R is usually truncated at some $l_{\min} < 0$ and $l_{\max} > 0$. We thus have $\mathcal{N} = (l_{\max} - l_{\min} + 1) \times N_t$ unknown DFs that we collect in the $\mathcal{N}$-vector $z(t)$. Similarly, $\mathbf{F}(t)$ and $\mathbf{g}$ are $\mathcal{N}$-dimensional vectors.

Any excited wave is a superposition of all eigenmodes of (52):

$$z(t) = \sum_{j=1}^{\mathcal{N}} q_j(t) z^{(j)}, \tag{77}$$

but not all eigenvectors $z^{(j)}$ are physical. Increasing the number of elements increases the accuracy of the eigenvalues and eigenvectors describing the isolated oscillatory modes but also adds spurious and/or singular modes. Such non-physical modes can contribute noise to the calculated disc response. To keep only physical modes, we introduce

$$\mathbf{q} = \begin{bmatrix} q_1 & q_2 & \ldots & q_{\mathcal{N}} \end{bmatrix}^{\mathrm{T}},$$
$$\mathbf{M} = \begin{bmatrix} z^{(1)} & z^{(2)} & \ldots & z^{(\mathcal{N})} \end{bmatrix},$$

and express (77) in the matricial form $z = \mathbf{M} \cdot \mathbf{q}$. This is substituted into (51) to obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{q} = -\mathrm{i}\mathbf{J} \cdot \mathbf{q} + \mathrm{i}Q(t)\mathbf{M}^{-1} \cdot \mathbf{g}, \tag{78}$$

where $\mathbf{J} = \mathbf{M}^{-1} \cdot \mathbf{A} \cdot \mathbf{M}$ is a diagonal matrix – or a Jordan form if there are degenerate eigenvalues (Perko 2001) – whose elements are the eigenfrequencies of (52). The diagonalizing matrix $\mathbf{M}$ is often called the modal matrix.

We call the $\mathcal{N}$-dimensional vector $\boldsymbol{\zeta} = \mathbf{M}^{-1} \cdot \mathbf{g}$ the forcing vector and rewrite (78) in terms of its components:

$$\frac{\mathrm{d}}{\mathrm{d}t}q_j(t) = -\mathrm{i}\omega_j q_j(t) + \mathrm{i}Q(t)\zeta_j, \quad j = 1, 2, \ldots, \mathcal{N}. \tag{79}$$

Equations associated with non-physical ω_j can now be dropped from (79) and we find both the homogeneous and particular solutions,

$$q_j(t) = \mathrm{e}^{-\mathrm{i}\omega_j(t-t_0)} q_j(t_0) + \mathrm{i}\zeta_j \int_{t_0}^t Q(\tau) \mathrm{e}^{-\mathrm{i}\omega_j(t-\tau)} \, \mathrm{d}\tau, \tag{80}$$

for $j = 1, 2, \ldots, \mathcal{N}_{\mathrm{p}}$ with $\mathcal{N}_{\mathrm{p}}$ being the number of physical modes. Fly-by perturbations begin at $t = t_0 = -\infty$ ($\phi_{\mathrm{p}} = -\pi$) with $q_j(t_0) = 0$ ($\forall j$). Consequently, using the orbit equations (70) and (71), and defining $\beta = \omega_{\mathrm{p}}/\omega_j$, we arrive at

$$q_j(t, \phi_{\mathrm{p}}) = -\mathrm{i}\frac{3\sqrt{2}(\bar{M}-1)}{64\bar{M}^{4/3}}\omega_j^{5/3} \zeta_j Q_j \mathrm{e}^{-\mathrm{i}\omega_j t}, \tag{81}$$

$$Q_j(\phi_{\mathrm{p}}, \beta) = \beta^{5/3} \int_{-\pi}^{\phi_{\mathrm{p}}} (1 + \cos\xi)^2 \, \mathrm{e}^{\mathrm{i}\omega_j t(\xi) - \mathrm{i}\xi} \, \mathrm{d}\xi, \tag{82}$$

which leaves behind the permanent oscillation

$$q_j(t) = -\mathrm{i}\frac{3\sqrt{2}(\bar{M}-1)}{64\bar{M}^{4/3}}\omega_j^{5/3} \zeta_j Q_j(\pi, \beta) \mathrm{e}^{-\mathrm{i}\omega_j t}, \tag{83}$$

when the encounter ends at $\phi_{\mathrm{p}} = +\pi$. The integrands of the real and imaginary parts of $Q_j(\pi, \beta)$ are, respectively, even and odd functions of ϕ_{p} over the interval $[-\pi, +\pi]$. One thus obtains $\mathrm{Im}[Q_j(\pi, \beta)] = 0$. The real part of $Q_j(\pi, \beta)/\beta^{5/3}$ is positive-definite, and therefore, a necessary and sufficient condition for the excitation of the jth oscillatory mode is that the corresponding component of the forcing vector $\zeta_j \neq 0$. The asymptotic forms are

$$Q_j(\pi, \beta) \approx 2\pi\beta^{5/3}, \tag{84}$$

for $\beta \gg 1$, and

$$Q_j(\pi, \beta) \approx \frac{2^{15/4}\pi^{1/2}}{3}\beta^{1/6} \exp\left(-\frac{2\sqrt{2}}{3\beta}\right), \tag{85}$$

for $\beta \ll 1$. The exponential decay for small ω_{p} is due to adiabatic invariance.

We have computed $\boldsymbol{\zeta}$ for all parent modes of Fig. 4, and have plotted its components ζ_j in Fig. 8 for two $\mu = 0.05$ models with different mean eccentricities. The results for child modes and other models are similar. In our models ζ_j is larger for modes with low frequencies. This can be understood as a competition between two effects seen in Fig. 5: (i) as the mode frequency decreases, the number of its nodes increases, so the coupling of the mode to a smooth external field is reduced; (ii) as the frequency decreases, the outermost peak of $X(R)$ shifts to a larger radius and hence contributes more to the term $-[f_0, V_{\mathrm{e}}]$ in (34) – recall that $V_{\mathrm{e}} \sim R^3$. In general
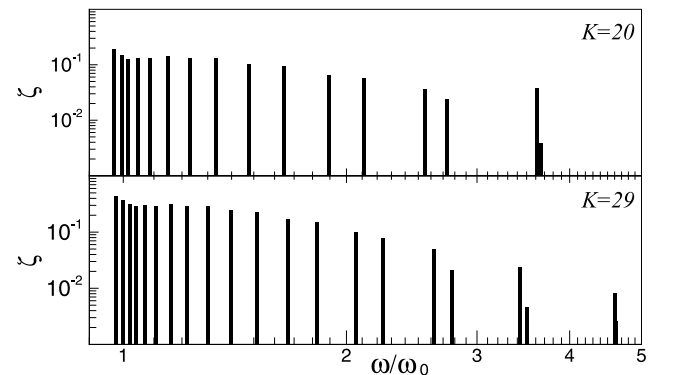
Figure 8. The forcing vector components ζ_j for the modes of the discs having $\mu = 0.05$ and $K = 20$ and 29. The corresponding frequency spectra have been plotted in Fig. 4.
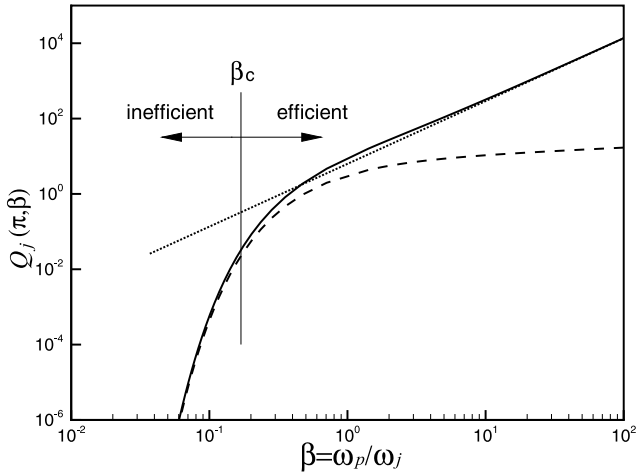
Figure 9. The profile of $Q_j(\pi, \beta)$ (solid line) together with its analytic asymptotes when $\beta \to 0$ (dashed line; equation 85) and $\beta \to \infty$ (dotted line; equation 84). The excitation is efficient for $\beta > \beta_c$ with $\beta_c = 0.169367$ (see text for definition).

the second effect wins, causing the coupling, as measured by ζ, to be larger for low-frequency modes.

We also find that the range of ζ is similar for all K models (Fig. 8). This shows that the response of a near-Keplerian disc is not sensitive to its mean eccentricity: $m = 1$ slow modes in warm and cold discs have an equal chance of being excited by encounters.

The excitation efficiency of modes is determined by the function $Q_j(\pi, \beta)$, which has been plotted in Fig. 9. The excitation of mode j is inefficient for $\omega_p \lesssim \beta_c \omega_j$ where we have defined the critical frequency ratio $\beta_c = 0.169\,367$ as the point where $Q_j(\pi, \beta)$ drops to 10 per cent of its value predicted by its $\beta \to \infty$ asymptote (equation 84). For a given mass parameter $\bar{M}$, a perturber can only excite mode j efficiently if its orbit has periastron $R_{\mathrm{p,min}} \lesssim [\bar{M}/(\beta_c^2 \omega_j^2)]^{1/3}$. Faster modes have larger ω_j, and therefore need closer encounters to be excited.

In Fig. 10, we illustrate the excitation of mode D_1 at four azimuths during the fly-by of a disc with mass ratio $\mu = 0.025$, for an encounter with the parameter $\omega_p = \sqrt{2}\omega_{D_1} = 0.0096538$. The interaction begins at $\phi_p = -\pi$ ($t = -\infty$) and ends at $\phi_p = +\pi$ ($t = +\infty$). We plot the positive part of the response density:

$$\Sigma_{1,D_1}(R, \phi, t) = \mathrm{Re}\left\{ q_j[t, \phi_p(t)]\, e^{im\phi - i\omega_j t} \right.$$

$$\left. \times \sum_{n=1}^{N} H_n(R) \mathbf{G}_n(R) \cdot \mathbf{b}_n^{(j)} \right\}, \tag{86}$$

where j corresponds to mode D_1 and the vectors $\mathbf{b}_n^{(j)}$ are extracted from $\mathbf{d}^{(j)} = \sum_l \mathbf{F}(l) \cdot \mathbf{z}_l^{(j)}$ as we did in equation (61).

We remark that stable modes always rotate with a constant angular velocity, but the perturber-star centreline has a variable angular velocity. Therefore the perturber may lead or lag the maximum response in azimuth, and the maximum response may occur some time after closest approach.

In general, of course, the close passage of a perturber will excite multiple modes. The main visual difference between a single-mode and multi-mode response is the occurrence of long-period beating patterns in the latter case. We have constructed animations of the evolution of the multi-mode pattern during an encounter and the beating can be quite striking

(http://www.youtube.com/watch?v=ZTyXK7H6Q8E). This animation is for a model with $K = 10$ and $\mu = 0.025$, and the 11 modes with the highest frequencies are participating in the response.

8 APPLICATION TO DEBRIS DISCS AND GALACTIC NUCLEI

We have shown that low-mass, near-Keplerian, collisionless discs can support stable, long-lived, large-scale slow modes. The most prominent of these are expected to have azimuthal wavenumbers $m = 1$ and $m = 2$.

8.1 Debris discs

The existence of slow modes implies that debris discs can support waves in the planetesimal population that provides most of the disc mass, and suggests that collisions in this non-axisymmetric distribution could generate non-axisymmetric dust distributions that would be visible in thermal emission or scattered light.

Non-axisymmetric structures in debris discs are normally assumed to be produced by planets, but our results imply that some or even most of these structures may be density waves. Specific examples include the following.

(i) β Pictoris: scattered starlight reveals that this star is surrounded by a debris disc extending to $\gtrsim 1000$ au. The disc is brighter on one side than the other, perhaps due to an $m = 1$ slow mode, and also contains brightness enhancements that could be due to shorter-wavelength density waves. The disc exhibits warps or tilted rings at various radii (Heap et al. 2000; Wahhaj et al. 2003); although the present paper examines only in-plane slow modes, there should also be slow bending modes, and these provide a possible explanation for the warps. There is a $10M_J$ planet orbiting at ~ 10 au in the β Pic system (Lagrange et al. 2010) but it is far from clear that this is the cause of the warps and other features; several authors have argued that the asymmetries provide evidence for two or even three planets (Freistetter, Krivov & Löhne 2007; Currie et al. 2011) but it is implausible to invoke a new planet for every feature.

(ii) Fomalhaut: this star is surrounded by a ring of dust with a sharp inner edge at 130 au. The centre of the ring is offset by 15 au from the host star, implying an eccentricity of 0.11; the ring is narrowest at apastron, implying that the eccentricity declines with radius (Kalas et al. 2005). Quillen (2006) stressed that these features could be produced by a planet orbiting just inside the ring, and a possible planet was subsequently discovered (Kalas et al. 2008). The eccentricity of the ring could be forced by the planet or a slow density wave, depending on whether the planet mass or ring mass is larger. The sharp inner edge of the ring is most likely due to the planet.

(iii) Vega: observations at a variety of wavelengths between $350\,\mu m$ and 1.3 mm reveal a face-on dust ring of radius ~ 100 au, dominated by two clumps (see Marsh et al. 2006 for a summary of the data, and beware that Piétu et al. 2011 question the reality of non-axisymmetric structure in the disc). The clumps are usually ascribed to dust trapped in a resonance with an unseen planet (e.g. Kuchner & Holman 2003; Wyatt 2003), but $m = 1$ and $m = 2$ slow density waves provide an alternative explanation. Within a few years we may be able to distinguish these hypotheses by measurements of the motion of these clumps relative to the host star: the expected angular speed of the planet is $\sim 1^\circ$ yr^{-1} while slow modes should have negligible pattern speeds.
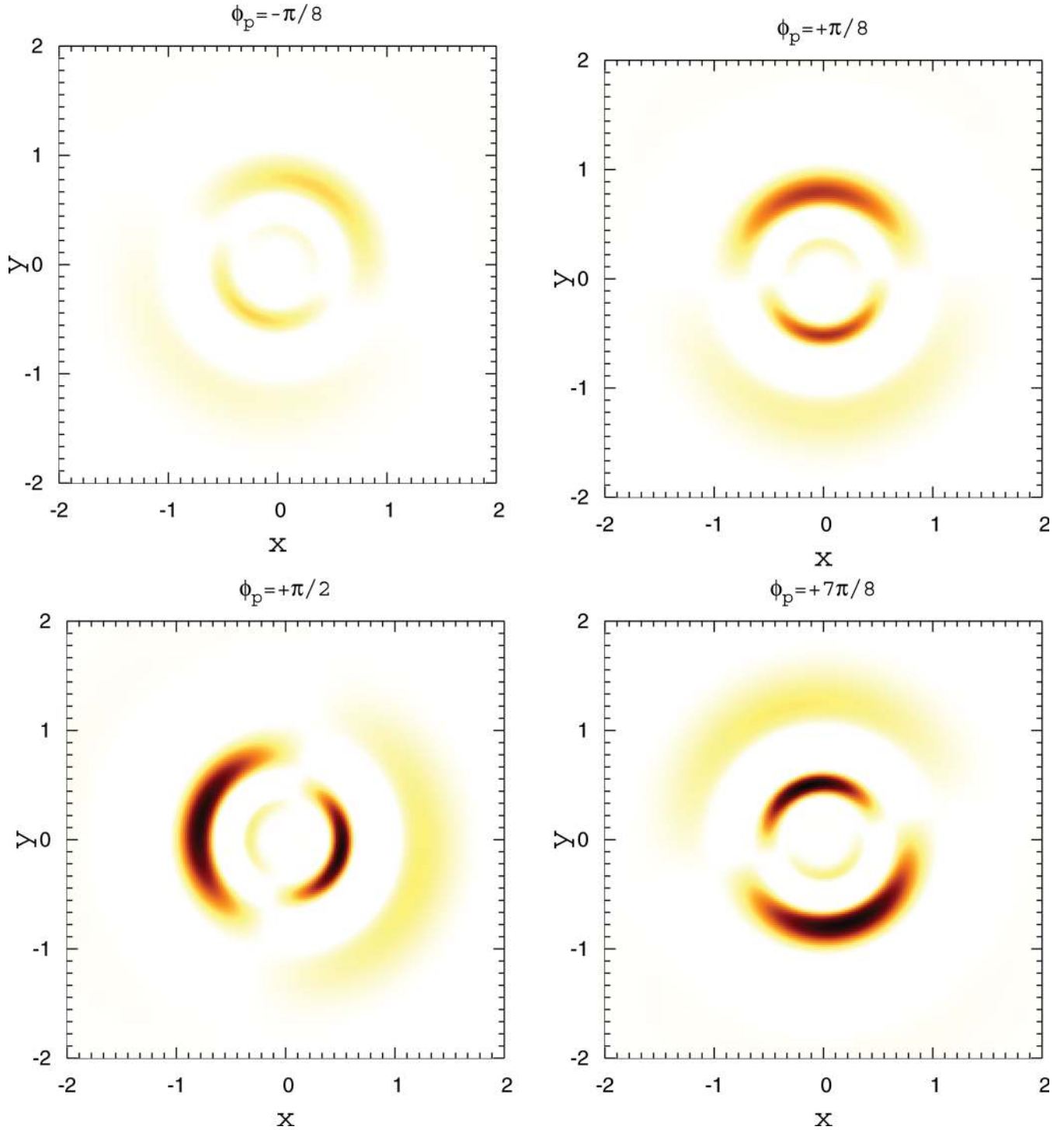
Figure 10. The evolution of mode D_1 in a disc with mean eccentricity $\bar{e} = 0.159$ as a perturber on a parabolic orbit encounters the star–disc system. ϕ_p is the azimuthal angle of the flying-by star; the line $\phi_p = 0$ coincides with the x-axis and with the direction of periastron. We have $\phi_p = (-\pi, +\pi)$ for $t = (-\infty, +\infty)$. The contours show the positive part of the response density. The periastron distance of the perturber is $R_{p,min} = 188.58(1 + M_p/M_\star)^{1/3}(M_d/10^{-3}M_\star)^{-2/3}$.

(iv) ϵ Eridani: a nearly face-on ring of dust surrounds this star at ∼60 au. The disc exhibits several clumps and a lopsided bright-ness distribution in images at $450\,\mu m$ and $850\,\mu m$ (Greaves et al. 2005). Some but not all of these peaks may be background sources. Models in which the clumps are due to resonances with a planet are described by Ozernoy et al. (2000), Quillen & Thorndike (2002) and Deller & Maddison (2005). These features could be due to slow

modes, but the presence of density maxima at several azimuths would require that more than one mode was present. Resonance models predict angular velocities around the host star of about $1°\,yr^{-1}$.

(v) HR 4796A: there is an edge-on debris ring ∼80 au from the host star. One ansa of the ring is brighter, hotter and at smaller radius than the other (Telesco et al. 2000; Moerchen et al. 2011).

This asymmetry is most naturally explained by an eccentric dust ring ($e \simeq 0.06$); at periastron the dust is closer to the star and therefore hotter and brighter ('pericentre glow', Wyatt et al. 1999). The ring eccentricity is usually assumed to be excited by secular perturbations from a nearby planet but an $m = 1$ slow mode of the disc is an alternative. The mode might be excited by the companion star HR 4796B, currently at a projected separation of ~500 au.

(vi) AB Aurigae: near-infrared images reveal a debris disc of over 1000 au radius. The disc shows spiral arms at radii of several hundred au, some of which are also seen at submillimetre wavelengths, as well as rings, gaps and clumps at smaller radii (Hashimoto et al. 2011). As in the case of other systems, the features could be due to planets or slow modes, and these hypotheses can be distinguished by proper-motion measurements.

(vii) η Corvi: the disc surrounding this star appears at 450 μm as two equally bright peaks equidistant from the host star at 100 au; these can be modelled either as the ansae of an edge-on axisymmetric ring or as a more face-on disc containing dust trapped in a resonance with a planet (Wyatt et al. 2005). A third possibility is an $m = 2$ slow mode in a face-on disc.

(viii) HD 141569A: Clampin et al. (2003) have detected strong spiral structure in the debris disc around this star, at about 400 au radius. They suggest that the spiral may be excited by tides from nearby stars. There is also a gap in the disc at about 250 au radius; both of these features might be due to planets but only if planets can form at radii exceeding 200 au. Wyatt (2005) has suggested that the spiral could be caused by a Jupiter-mass planet on an eccentric orbit ($e \simeq 0.2$, $a \simeq 250$ au) but slow modes provide a more economical explanation, especially given the difficulties of forming planets at such large distances.

(ix) HD 100546: this disc exhibits an apparent dark hole and bright clump at about 30 au from the host star (Quanz et al. 2011). These features could be due to an orbiting planet or a slow density wave. The Keplerian motion at this radius is about $3°$ yr^{-1}. At much larger radii, ~250 au, the disc exhibits spiral structure (Grady et al. 2001). Possible explanations include a planet at several hundred au from the star or density waves excited by a passing star. The latter possibility was discussed by Quillen et al. (2005) but dismissed because their estimated lifetime for the spiral structure was only ~10^4 yr and no suitable nearby star could be found; the results of the present paper imply that the structure could last for a much longer time – perhaps as long as the 10 Myr age of the star – so the chance of a suitable encounter in the past is much larger.

(x) HD 61005: this star is surrounded by an asymmetric edge-on debris disc of radius ~60 au. The asymmetry can be modelled as a mean eccentricity of 0.05, but there are no planets more massive than ~3 Jupiter masses close to the ring (Buenzli et al. 2010).

(xi) HD 15115: this star hosts an edge-on debris disc; the dominant thermal emission from the disc arises at radii ~35 au but the disc is visible to much larger radii. The surface brightness of the east side of the disc is about 1 mag fainter than the west side at a given radius and the surface-brightness distribution perpendicular to the disc midplane is asymmetric on the west side (Kalas, Fitzgerald & Graham 2007); both features can arise naturally from an $m = 1$ distortion.

(xii) HD 107146: there is a dust ring at 100 au that exhibits clumps and a lopsided brightness distribution in 1.3 mm images (Corder et al. 2009). These might be due either to a planetary resonance or to slow density waves; however, 880 μm observations with similar resolution do not confirm the existence of the clumps (Hughes et al. 2011).

8.2 Discs in galactic nuclei

The results of this paper also illuminate our understanding of stellar discs in galactic nuclei. They can be applied directly to such discs if the apsidal precession is dominated by the self-gravity of the disc, rather than relativistic effects or the gravitational field from a spherical stellar population in the nucleus.

The apparent 'double' nucleus of M31 is most likely a stellar disc that has been distorted by a large-amplitude slow mode (see Peiris & Tremaine 2003, Salow & Statler 2004, and references therein). Such modes arise naturally in *N*-body simulations (Jacobs & Sellwood 2001). They can be excited by gas inflow and star formation in the central few parsec of the galaxy (Hopkins & Quataert 2010a) or by instabilities induced by a small population of counter-rotating stars (Touma 2002). Slow modes may also play a central role in feeding supermassive black holes (Hopkins & Quataert 2010b).

9 DISCUSSION

The finite element formulation has enabled us to explore the modal spectrum of low-mass near-Keplerian collisionless discs, and to calculate the corresponding mode shapes for a wide range of initial radial dispersions (rms eccentricities). Our method also yields moments of the DF, which provide the evolutionary equations for energy and angular-momentum transport in perturbed discs, and allows the accurate representation of modes that contain a singular resonant component (e.g. Fig. 6).

We find that near-Keplerian discs support 'slow' modes, that is, modes for which the eigenfrequency or pattern speed is proportional to the disc mass. WKB analysis shows that these modes are closely related to the *p* modes found by Tremaine (2001) in cold near-Keplerian discs with softened gravity. Both our numerical results and analytic arguments imply that there are no unstable slow modes. All slow modes in the discs we have examined are prograde (positive pattern speed). Slow modes can exist with arbitrary azimuthal wavenumber m, but modes with $m = 1$ and $m = 2$ have the largest scale and are the easiest to excite by an external perturber.

The eigenmodes of the linearized CBE bifurcate from the degenerate leading/trailing modes predicted by the WKB theory. The modes are degenerate for cold discs and split into close (in frequency) pairs as the mean eccentricity grows, until for $\bar{e} \gtrsim 0.2$ there is no apparent pairing in frequency space (see Fig. 4).

Some of the non-axisymmetric structures that are commonly observed in debris discs, such as clumps, lopsided rings, and spiral arms, may be due to slow modes, perhaps excited by the fly-by of a passing star or binary companion. These features are normally ascribed to hypothetical massive planets embedded in the disc. The two hypotheses can be distinguished in some discs by monitoring the motion of these features over decade time-scales: many features associated with planets should orbit the host star at a pattern speed that is not far from the Keplerian angular speed of the planet, whereas slow modes should have negligible pattern speeds. Structures induced by modes and planets may also be distinguishable in the future by high-resolution far-infrared observations by interferometers or large single-dish telescopes (e.g. ALMA or CCAT).

Future theoretical work should include the exploration of slow bending modes and of the behaviour of slow modes in thick discs that resemble the discs seen in galactic nuclei.

REFERENCES

Bellman R., 1964, Perturbation Techniques in Mathematics, Physics and Engineering. Holt, Rinehart and Winston, Inc., New York
Bender R. et al., 2005, ApJ, 631, 280
Binney J., Tremaine S., 2008, Galactic Dynamics, 2nd edn. Princeton Univ. Press, Princeton
Buenzli E. et al., 2010, A&A, 524, L1
Clampin M. et al., 2003, AJ, 126, 385
Corder S. et al., 2009, ApJ, 690, L65
Currie T., Thalmann C., Matsumura S., Madhusudhan N., Burrows A., Kuchner M., 2011, ApJ, 736, L33
Deller A. T., Maddison S. T., 2005, ApJ, 625, 398
Freistetter F., Krivov A. V., Löhne T., 2007, A&A, 466, 389
Genzel R., Eisenhauer F., Gillessen S., 2010, Rev. Modern Phys., 82, 3121
Golimowski D. A. et al., 2011, AJ, 142, 30
Goodman J., 2003, MNRAS, 339, 937
Grady C. A. et al., 2001, AJ, 122, 3396
Greaves J. S. et al., 2005, ApJ, 619, L187
Hashimoto J. et al., 2011, ApJ, 729, L17
Heap S. R., Lindler D. J., Lanz T. M., Cornett R. H., Hubeny I., Maran S. P., Woodgate B., 2000, ApJ, 539, 435
Hopkins P. F., Quataert E., 2010a, MNRAS, 405, L41
Hopkins P. F., Quataert E., 2010b, MNRAS, 407, 1529
Hughes A. M., Wilner D. J., Andrews S. M., Williams J. P., Su K. Y. L., Murray-Clay R. A., Qi C., 2011, ApJ, 740, 38
Jacobs V., Sellwood J. A., 2001, ApJ, 555, L25
Jalali M. A., 2010, MNRAS, 404, 1519
Jalali M. A., Tremaine S., 2011, MNRAS, 410, 2003
Kalas P., Jewitt D., 1995, AJ, 110, 794
Kalas P., Graham J. R., Clampin M., 2005, Nat, 435, 1067
Kalas P., Fitzgerald M. P., Graham J. R., 2007, ApJ, 661, L85
Kalas P. et al., 2008, Sci, 322, 1345
Kalnajs A. J., 1977, ApJ, 212, 637
Kuchner M. J., Holman M. J., 2003, ApJ, 588, 1110
Lagrange A.-M. et al., 2010, Sci, 329, 57
Larwood J. D., Kalas P. G., 2001, MNRAS, 323, 402
Marsh K. A., Dowell C. D., Velusamy T., Grogan K., Beichman C. A., 2006, ApJ, 646, L77
Moerchen M. M., Churcher L. J., Telesco C. M., Wyatt M., Fisher R. S., Packham C., 2011, A&A, 526, A34
Ozernoy L. M., Gorkavyi N. N., Mather J. C., Taidakova T. A., 2000, ApJ, 537, L147
Peiris H. V., Tremaine S., 2003, ApJ, 599, 237
Perko L., 2001, Differential Equations and Dynamical Systems, 3rd edn. Springer-Verlag, New York
Pichon C., Lynden Bell D., 1996, MNRAS, 282, 1143
Piétu V., di Folco E., Guilloteau S., Gueth F., Cox P., 2011, A&A, 531, L2
Quanz S. P., Schmid H. M., Geissler K., Meyer M. R., Henning T., Brandner W., Wolf S., 2011, ApJ, 738, 23
Quillen A. C., 2006, MNRAS, 372, L14
Quillen A. C., Thorndike S., 2002, ApJ, 578, L149
Quillen A. C., Varnière P., Minchev I., Frank A., 2005, AJ, 129, 2481
Salow R. M., Statler T. S., 2004, ApJ, 611, 245
Sawamura M., 1988, PASJ, 40, 279
Scally A., Clarke C., 2001, MNRAS, 325, 449
Sellwood J. A., 1987, ARA&A, 25, 151
Telesco C. M. et al., 2000, ApJ, 530, 329
Telesco C. M. et al., 2005, Nat, 433, 133
Toomre A., 1963, ApJ, 138, 385
Touma J. R., 2002, MNRAS, 333, 583
Tremaine S., 2001, AJ, 121, 1776 (T01)
Wahhaj Z., Koerner D. W., Ressler M. E., Werner M. W., Backman D. E., Sargent A. I., 2003, ApJ, 584, L27
Williams J. P., Cieza L. A., 2011, ARA&A, 49, 67
Wyatt M. C., 2003, ApJ, 598, 1321
Wyatt M. C., 2005, A&A, 440, 937
Wyatt M. C., 2008, ARA&A, 46, 339
Wyatt M. C., 2009, in Mann I., Nakamura A. M., Mukai T., eds, Small Bodies in Planetary Systems. Springer-Verlag, Berlin, p. 37
Wyatt M. C., Dermott S. F., Telesco C. M., Fisher R. S., Grogan K., Holmes E. K., Piña R. K., 1999, ApJ, 527, 918
Wyatt M. C., Greaves J. S., Dent W. R. F., Coulson I. M., 2005, ApJ, 620, 492
Zienkiewicz O. C., Taylor R. L., Zhu J. Z., 2005, The Finite Element Method: Its Basis and Fundamentals, 6th edn. Elsevier Butterworth-Heinemann, Amsterdam

APPENDIX A: WKB ANALYSIS

The dispersion relation for a collisionless disc can be computed analytically in the WKB or short-wavelength approximation (Binney & Tremaine 2008)

$$1 = \frac{2\pi G \Sigma_0 |k|}{\Omega_R^2} \mathcal{G}\left[\frac{\omega - m\Omega_\phi}{\Omega_R}, \left(\frac{\sigma_R k}{\Omega_R}\right)^2\right],$$

$$\text{where} \quad \mathcal{G}(s, \chi) = \frac{2}{\chi}e^{-\chi}\sum_{n=1}^{\infty}\frac{I_n(\chi)}{1 - s^2/n^2}. \tag{A1}$$

Here Ω_R and Ω_ϕ are the orbital frequencies (equation 11), $I_n(\chi)$ is a modified Bessel function, m is the azimuthal wavenumber, σ_R is the radial velocity dispersion and the perturbed surface density is assumed to vary as $\Sigma_1(R, \phi, t) \propto \exp[i(m\phi + \int^r k(r')dr' - \omega t)]$. The dispersion relation (A1) is valid if $\sigma_R \ll \Omega_\phi R$ and $|k|R \gg 1$.

For low-mass discs, $\mu \ll 1$, the precession frequency is $\Omega_{\text{pr}} = \Omega_\phi - \Omega_R = \mathcal{O}(\mu)$ (equation 12) and the surface density Σ_0 is also $\mathcal{O}(\mu)$. For slow modes ω is $\mathcal{O}(\mu)$, and since $s = (\omega - m\Omega\phi)/\Omega_R$ is nearly $-m$, the denominator $1 - s^2/n^2$ becomes small for $n = m$ and the solution of the dispersion relation for a given m is obtained by keeping the dominant $n = m$ term of the summation in (A1). Moreover $\sigma_R^2 = \frac{1}{2}e_{\text{rms}}^2 R^2 \Omega_R^2$ where e_{rms} is the rms eccentricity defined in equation (21) – strictly, this is the rms eccentricity at a given radius rather than of the whole disc but in our models the rms eccentricity is almost independent of radius (Fig. 3). The dispersion relation for slow modes then simplifies to

$$\omega = m\Omega_{\text{pr}} + \frac{m\pi\Sigma_0 |k|}{\Omega_R}\mathcal{F}_m\left[(e_{\text{rms}}kR)^2/2\right]$$

$$\text{where} \quad \mathcal{F}_m(\chi) = \frac{2}{\chi}e^{-\chi}I_m(\chi). \tag{A2}$$

Here all quantities are written in the dimensionless units of Section 2. Note that as $z \to 0$, $I_m(z) \to z^m/(2^m m!)$. Thus as the rms eccentricity shrinks to zero the wavenumber of a slow mode must vary as $k \sim e_{\text{rms}}^{2(1-m)/(2m-1)}$. In other words, for $m = 1$ slow modes have $|kR| \sim 1$ even for cold discs – in this case the use of the WKB approximation for slow modes is not formally justified, but the results provide a useful qualitative guide to the behaviour of the frequency spectra that we find using FEM (see further discussion at the end of Section 6). For $m > 1$ slow modes exist but with wavelengths that shrink as e_{rms} declines. For cold discs $\mathcal{F}_m(\chi) = 0$ for $m > 1$ so the disc supports only singular modes at the resonances $\omega = m\Omega_{\text{pr}}$.
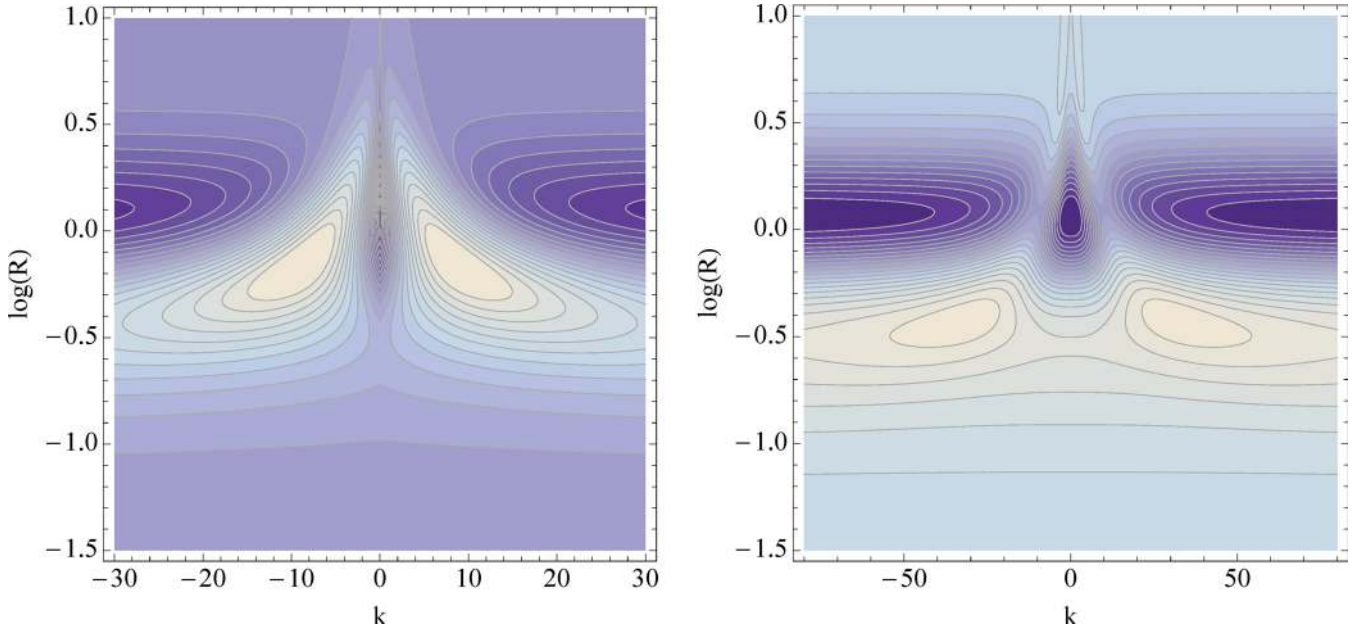
Figure A1. Contours of constant ω in the WKB approximation, from equation (A2) for $m = 1$ (left) and $m = 2$ (right). The disc has mean eccentricity $\bar{e} = 0.159$. The horizontal axis is the wavenumber and the vertical axis is the log of the radius (base 10). Lighter shades correspond to higher values of ω.

In the WKB approximation, disturbances in the disc can be decomposed into wavepackets that propagate at the group velocity $d\omega/dk$ along contours of constant frequency ω. These contours are illustrated in Fig. A1 for a disc with mean eccentricity $\bar{e} = 0.159$ and $m = 1, 2$. Wavepackets that propagate along open contours eventually wind up ($|k| \to \infty$) and disappear. Discrete normal modes can arise for closed contours if the appropriate resonance condition is satisfied. Consider the case $m = 1$ (left-hand panel of Fig. A1). For the closed contours centred on $|k| = 10$, $R = 0.6$ the resonance condition is (T01)[1]

$$\oint dk \, dR = 2\pi\left(n - \frac{1}{2}\right), \tag{A3}$$

where $n = 1, 2, 3, \ldots$ and the integral is taken over the area in (k, R) space enclosed by the contour. These modes, called p modes by T01, occur in degenerate pairs, one composed of leading and one of trailing waves ($k < 0$ and $k > 0$, respectively).

Equation (A3) can be solved numerically to find the frequencies of the p modes. These are plotted in the bottom panel of Fig. 4 for $m = 1$, and are in good qualitative agreement with the frequencies calculated by FEM. There is similar agreement between equation (A3) and FEM for $m = 2$ modes.

---

[1] The factor of 2 on the left-hand side of equation (56) in T01 is incorrect.

This paper has been typeset from a TeX/LaTeX file prepared by the author.