

# SECULAR EVOLUTION DRIVEN BY MASSIVE ECCENTRIC DISKS/RINGS: AN APSIDALLY ALIGNED CASE

IRINA DAVYDENKOVA<sup>1</sup> & ROMAN R. RAFIKOV<sup>2,3</sup>

*Draft version September 11, 2018*

## ABSTRACT

Massive eccentric disks (gaseous or particulate) orbiting a dominant central mass appear in many astrophysical systems, including planetary rings, protoplanetary and accretion disks in binaries, and nuclear stellar disks around supermassive black holes in galactic centers. We present an analytical framework for treating the nearly Keplerian secular dynamics of test particles driven by the gravity of an eccentric, apsidally aligned, zero-thickness disk with arbitrary surface density and eccentricity profiles. We derive a disturbing function describing the secular evolution of coplanar objects, which is explicitly related (via one-dimensional, convergent integrals) to the disk surface density and eccentricity profiles without using any ad hoc softening of the potential. Our analytical framework is verified via direct orbit integrations, which show it to be accurate in the low-eccentricity limit for a variety of disk models (for disk eccentricity  $\lesssim 0.1 - 0.2$ ). We find that free precession in the potential of a disk with a smooth surface density distribution can naturally change from prograde to retrograde within the disk. Sharp disk features — edges and gaps — are the locations where this tendency is naturally enhanced, while the precession becomes very fast. Radii where free precession changes sign are the locations where substantial (formally singular) growth of the forced eccentricity of the orbiting objects occurs. Based on our results, we formulate a self-consistent analytical framework for computing an eccentricity profile for an aligned, eccentric disk (with a prescribed surface density profile) capable of precessing as a solid body under its own self-gravity.

*Subject headings:* accretion, accretion disks — protoplanetary disks — planets and satellites: rings

## 1. INTRODUCTION

Astrophysical disks orbiting in the gravitational potential of a dominant central mass  $M_c$  often possess nonaxisymmetric shape. The nonaxisymmetric distortion can be modeled as a manifestation of *disk eccentricity*. In this picture, to zeroth order, different components of the disk — parcels of gas in fluid (collisional) disks, or particles (e.g. stars) in collisionless disks — move on eccentric Keplerian orbits in the field of a central mass. Even if the mass of the disk  $M_d$  is much smaller than  $M_c$ , the self-gravity of the disk can still play a very important role in its dynamics, as well as the orbital evolution of external objects, by driving precession and causing an exchange of angular momentum between different parts of the disk on long (secular) timescales.

Such eccentric disks are encountered in a variety of astrophysical contexts — galactic, stellar, and planetary. One of the closest examples is provided by the eccentric planetary rings, such as the  $\epsilon$ ,  $\alpha$ , and  $\beta$  rings<sup>4</sup> of Uranus (Elliot & Nicholson 1984), as well as the Titan and Maxwell ringlets of Saturn (Porco 1990). These particulate collisional rings are very narrow, essentially representing limiting cases of eccentric disks with the

spread in semimajor axes of their constituent particles  $\Delta a \lesssim e_r a_r$ , where  $a_r$  and  $e_r$  are the mean semimajor axis and eccentricity of the rings. As demonstrated by Goldreich & Tremaine (1979), self-gravity of the rings, coupled with collisional effects (Chiang & Culter 2003; Chiang & Goldreich 2000; Mosqueira & Estrada 2002; Pan & Wu 2016), can counter differential precession driven by the planetary oblateness, allowing rings to precess as a solid body while maintaining a coherent eccentric shape.

A significant number of stellar binaries are known to host exoplanets, orbiting either the whole system (Doyle et al. 2011; Welsh et al. 2014) or one of the binary components (Chauvin et al. 2011; Hatzes et al. 2003). The formation and early dynamics of such planets are significantly complicated by the fact that the nonaxisymmetric binary potential excites the nonzero eccentricity of the protoplanetary disks (Kley et al. 2008; Miranda et al. 2017; Regály et al. 2011), in which the building blocks of these planets — planetesimals — orbit. It has been recently shown (Rafikov 2013a,b; Rafikov & Silsbee 2015a,b; Silsbee & Rafikov 2015a,b) that the gravitational effect of such an eccentric protoplanetary disk plays a key role in planetesimal dynamics for young binaries.

Spectroscopic observations of accretion disks in cataclysmic variables using the technique of Doppler tomography (Marsh & Horne 1988) suggest that a certain type of variability in these systems — the so-called “superhump” (Horne 1984) — is caused by the precession of an eccentric accretion disk around the white dwarf (Lubow 1991). Asymmetric evolving emission lines found in the spectra of compact disks of gaseous debris around some metal-rich white dwarfs (Gänsicke et al. 2006) also

<sup>1</sup> Université de Genève, 2-4 rue du Lièvre, c.p. 64, 1211 Genève 4, Switzerland; Irina.Davydenkova@unige.ch

<sup>2</sup> Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK; rrr@damtp.cam.ac.uk

<sup>3</sup> Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540, USA

<sup>4</sup> The latter two are also significantly inclined with respect to the equatorial plane of the planet.

provide evidence for their nonzero eccentricity (Dennihiy et al. 2018; Miranda & Rafikov 2018).

Finally, optical emission from the central region of the M31 galaxy exhibits a double-nucleus morphology (Bacon et al. 2001; Bender et al. 2005). The best interpretation of the existing photometric and spectroscopic data, first proposed by Tremaine (1995), points at the existence of a highly eccentric ( $e_d \sim 0.5$ ) stellar disk orbiting the central supermassive black hole. A number of models, both purely kinematic, i.e. not accounting for the disk self-gravity in maintaining its coherence (Brown & Magorrian 2013; Peiris & Tremaine 2003), and as fully dynamic (Bacon et al. 2001; Salow & Statler 2001, 2004; Sambhus & Sridhar 2002), have been put forward to understand this system. Our own Galaxy harbors an eccentric disk of young stars orbiting the central supermassive black hole (Bartko et al. 2009; Levin & Beloborodov 2003), whose gravity may affect its own dynamics (Nayakshin et al. 2006).

In many of these systems, disk gravity plays the dominant role in disk dynamics, as well as in the orbital evolution of nearby objects (e.g. planetesimals in protoplanetary disks in binaries). This motivated a number of past analytical (Chiang & Goldreich 2000; Goldreich & Tremaine 1979) and numerical (Bacon et al. 2001; Nayakshin et al. 2006; Salow & Statler 2004; Sambhus & Sridhar 2002) studies aimed at clarifying the details of the dynamics driven by the gravity of an eccentric disk. Such calculations inevitably require an efficient method for computing the potential  $\Phi_d$  of an eccentric disk at every point. Moreover, since  $M_d \ll M_c$ , the disk-driven evolution is typically rather slow, justifying the use of a *secular* approximation, in which the disk potential is averaged over the orbital motion of an object under consideration. A direct calculation of such an averaged potential, or a *secular disturbing function* as it is known in celestial mechanics, in general requires evaluation of three-dimensional integrals (see equation (A1)), which is impractical in many applications.

Silsbee & Rafikov (2015b) presented a calculation of a secular disturbing function for a particular model of a *radially extended* (i.e. having  $\Delta a \sim a$ ), apsidally aligned eccentric disk. They assumed that both the surface density and the eccentricity of the disk vary as *power laws* of the semimajor axis  $a$  of the mass elements comprising the disk. Their resultant disturbing function does not involve multidimensional integration and can be used for efficient analysis of disk-driven orbital dynamics. In particular, it was employed to provide a self-consistent treatment of the secular evolution of planetesimals orbiting in a massive eccentric protoplanetary disk within (or around) a young stellar binary (Rafikov & Silsbee 2015a,b; Silsbee & Rafikov 2015a).

The goal of our present work is to provide a natural but important generalization of the results of Silsbee & Rafikov (2015b). Here we develop a general analytical framework for computing a secular disturbing function for an apsidally aligned<sup>5</sup>, eccentric disk with *arbitrary* radial profiles of the disk surface density  $\Sigma$  and

<sup>5</sup> The assumption of apsidal alignment is retained here for simplicity. It is relaxed in a subsequent work of Davydenkova & Rafikov (in prep.).

eccentricity  $e_d$ . We show that this disturbing function can be reduced to a combination of one-dimensional integrals over the radial profiles of  $\Sigma$  and  $e_d$ , enabling application of our results to a broad range of practical problems (e.g. computation of the structure of a rigidly precessing eccentric disk, see §7). We also provide numerical verification of our analytical results using direct orbit integrations.

Our work is organized as follows. We describe our methodology and outline the results of the disturbing function calculation in §2; the details of its derivation can be found in Appendix A. We describe the strategy for numerical verification of our analytical calculations in §3 and then present our findings in §4. Having tested our analytical framework, we then describe several of its applications in §5, including the derivation of a self-consistent method for calculating the eccentricity distribution of an eccentric, apsidally aligned disk (with a prescribed surface density distribution) that can precess as a solid body while maintaining its overall shape (§6). We discuss our results in §7 and provide a brief summary in §8.

## 2. SECULAR DISTURBING FUNCTION

Our goal is to calculate the secular (i.e. orbit-averaged) gravitational potential felt by a test particle orbiting in the combined gravitational field of a central mass  $M_d$  and an eccentric disk (fluid or particulate). This particle can be an external object or it can be one of the mass elements comprising the disk (see §6). The test particle moves on an eccentric orbit coplanar with the disk, with semimajor axis  $a_p$ , eccentricity  $e_p$ , and apsidal angle  $\varpi_p$ .

The disk is purely two-dimensional (i.e. it has zero vertical thickness) and is not warped (i.e. lies in a single plane). It is eccentric and apsidally aligned in the sense that the trajectories of its constituent mass elements (fluid or particle) are *confocal Keplerian ellipses*, which are *apsidally-aligned* in the direction making an angle  $\varpi_d$  with respect to the reference direction. We define  $r_d$  to be the distance from the common focus of the eccentric orbits of the constituent particles and  $\varphi_d$  to be the polar angle with respect to the disk apsidal line; see Figure 1 for illustration.

For every trajectory with a semimajor axis  $a$ , we can define the disk *surface density at the periastron*  $\Sigma_d(a)$  and the eccentricity of the fluid trajectory  $e_d(a)$ , which we will simply call *disk eccentricity*. The disk mass distribution can also be characterized by the mass per unit semimajor axis  $\mu(a)$ . Using the basic properties of Keplerian dynamics, one can show that  $\Sigma_d$  and  $\mu$  are directly related via

$$\mu(a) = 2\pi a \Sigma_d(a) \left[ \frac{1 + e_d(a)}{1 - e_d(a)} \right]^{1/2} [1 - e_d(a)(1 + \zeta)], \quad (1)$$

where  $\zeta \equiv d \ln e_d(a) / d \ln a$ . In general,  $\Sigma_d(a)$  (or  $\mu(a)$ ) and  $e_d(a)$  can be arbitrary functions of the semimajor axis  $a_d$ , as long as  $e_d(a)$  varies slowly enough for the particle trajectories to be noncrossing<sup>6</sup>. In this work, we choose  $\Sigma_d$  (rather than  $\mu$ ) to characterize the distribution of mass in the disk.

Note that in the secular approximation relied upon in this study, the energy and semimajor axes of particles

<sup>6</sup> This requires  $de_d/d \ln a < (1 - e_d)$  (Ogilvie 2001; Statler 2001).

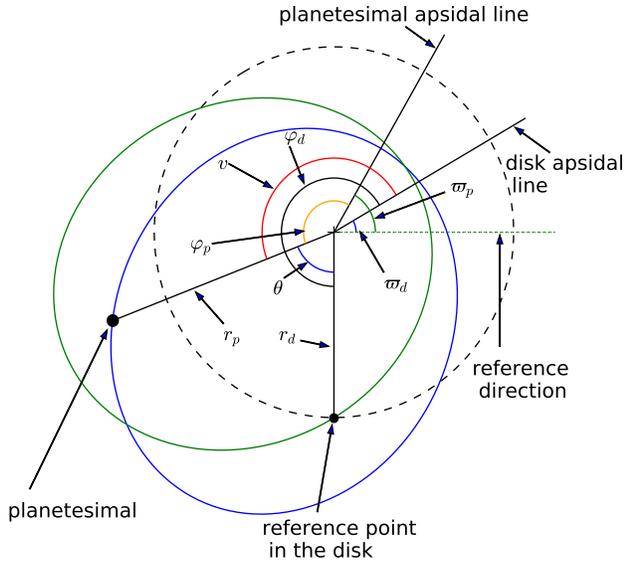


FIG. 1.— Geometry of the problem, showing elliptical trajectories of the test particle (blue) and a mass element of the disk (green). See text for details.

(or fluid elements) comprising the disk are the integrals of motion. As a result, the amount of disk mass per unit semi-major axis  $\mu(a)$  is strictly conserved even if the disk shape changes. Consequently, according to equation (1), if  $e_d$  does not change in time, then  $\Sigma_d(a)$  is also independent of time (this will be important in §6).

Statler (2001) and Ogilvie (2001) provided an expression for the two-dimensional surface density  $\Sigma(r_d, \varphi_d)$  of an eccentric disk in terms of  $e_d(a)$  and disk mass distribution  $\mu(a)$ . For our present purposes, it is more convenient to write  $\Sigma$  as a function of  $a$  and  $\varphi_d$ , relating it to  $\Sigma_d(a)$ . Using the calculation of  $\Sigma(a, \varphi_d)$  in Statler (1999), for an apsidally aligned disk, we can write

$$\Sigma(a, \varphi_d) = \Sigma_d(a) \frac{1 - e_d^2 - \zeta e_d(1 + e_d)}{1 - e_d^2 - \zeta e_d [e_d + \cos E(\varphi_d)]}, \quad (2)$$

where  $E(\varphi_d)$  is the eccentric anomaly ( $E = \varphi_d = 0$  at pericenter) and  $\Sigma_d$  and  $e_d$  are functions of  $a$ .

Even though the expression (2) holds for arbitrary  $e_d$ , in the rest of the paper, we will take the eccentricities of both the disk and test particle to be small,  $e_d(r) \ll 1$  and  $e_p \ll 1$ . This is needed for our secular theory (formulated at the lowest order in eccentricity) to provide an accurate description of orbital dynamics. As a consequence of this approximation, equation (1) also yields  $\Sigma_d(a) \approx \mu(a)/(2\pi a)$  to lowest order in  $e_d$ . Thus, even if  $e_d \ll 1$  varies in time,  $\Sigma_d(a)$  should still be conserved to  $O(e_d)$  accuracy in the course of secular evolution.

### 2.1. Secular (Orbit-averaged) potential of the disk

Our calculation of the orbit-averaged disturbing function  $R_d$  due to an eccentric disk uses the general mathematical procedure outlined in a seminal paper of Heppenheimer (1980) for the calculation of the disturbing function due to an *axisymmetric* disk. In this approach, the expansion of the disturbing function in

terms of a small parameter — test particle eccentricity — proceeds differently from the classical Laplace-Lagrange theory (Murray & Dermott 1999). The resultant expression for  $R_d$  does not contain non-integrable singularities at the particle semimajor axis, resulting in a convergent expression for the disturbing function. In other words, this calculation does not require introduction of an ad hoc *softening* of the potential. This method was later used by Ward (1981) to study the stability of the early Solar System perturbed by an axisymmetric protoplanetary disk.

Silsbee & Rafikov (2015b) extended the method of Heppenheimer (1980) to the case of *non-axisymmetric, apsidally aligned, eccentric* disks with  $\Sigma$  given by equation (2). However, their work was restricted to disks with power-law profiles of  $\Sigma_d(a)$  and  $e_d(a)$ . Here we generalize this calculation even further to cover the disks with *arbitrary* behaviors of  $\Sigma_d(a)$  and  $e_d(a)$ .

As a result of a rather lengthy derivation, the mathematical details of which are presented in Appendix A, we arrive at the following expression for the disturbing function due to an apsidally aligned disk:

$$R_d = a_p^2 n_p \left[ \frac{1}{2} A_d e_p^2 + B_d e_p \cos(\varpi_p - \varpi_d) \right], \quad (3)$$

where  $n_p = (GM_c/a_p^3)^{1/2}$  is the particle mean motion, with the coefficients  $A_d$  and  $B_d$  (having dimensions of  $[s^{-1}]$  and discussed in more detail in §2.2) given by the expressions (A25)-(A29). The key underlying simplifications making this calculation possible are that  $e_d \ll 1$ , as well as  $de_d/d \ln a \ll 1$ ; see equation (A2).

Introducing a two-component eccentricity vector  $\mathbf{e}_p = (k_p, h_p) = e_p(\cos \varpi_p, \sin \varpi_p)$  for a test particle, as well as the auxiliary vector

$$\mathbf{B}_d = B_d(\cos \varpi_d, \sin \varpi_d), \quad (4)$$

the expression (3) can be rewritten as

$$R_d = a_p^2 n_p \left[ \frac{1}{2} A_d \mathbf{e}_p^2 + \mathbf{e}_p \cdot \mathbf{B}_d \right] = a_p^2 n_p \left[ \frac{A_d}{2} (k_p^2 + h_p^2) + B_d (k_p \cos \varpi_d + h_p \sin \varpi_d) \right] \quad (5)$$

Note that  $\mathbf{B}_d$  is different from the disk eccentricity vector  $\mathbf{e}_d = e_d(\cos \varpi_d, \sin \varpi_d)$ ; in fact,  $\mathbf{B}_d$  involves a convolution of  $e_d(a)$  with a complicated kernel (see equation (A28)) over the radial extent of the disk.

The mathematical structure of  $R_d$  in equation (3) is identical to that of the conventional Laplace-Lagrange secular disturbing function (Murray & Dermott 1999). The difference between them lies in the explicit dependence of the coefficients  $A_d$  and  $B_d$  on  $a_p$ . The behavior of  $A_d(a_p)$  and  $B_d(a_p)$  is determined, eventually, by the radial profiles of the disk surface density  $\Sigma_d$  and eccentricity  $e_d$ . The exploration of this behavior is the major focus of our study.

### 2.2. Mathematical properties of $A_d$ and $B_d$

The expression (A25) for  $A_d$  consists of two parts. One of them,  $A_d^{\text{bulk}}$ , involves integral convolution over the disk semimajor axis  $a$  of a (prescribed) disk surface density  $\Sigma_d(a)$ , as well as its radial derivatives up to second order, all of which enter linearly; see equation (A26). The

convolution kernel involves a Laplace coefficient  $b_{1/2}^{(0)}(\alpha)$ , which features a weak (logarithmic) singularity arising as  $\alpha = a/a_p \rightarrow 1$  (i.e. for disk annuli close to the particle orbit). This singularity is, however, fully integrable and vanishes upon integration over the radial extent of the disk (see Appendix A). As always, the Laplace coefficients are defined as

$$b_s^{(j)}(\alpha) = \pi^{-1} \int_0^{2\pi} \frac{\cos(j\psi)d\psi}{(1 - 2\alpha \cos \psi + \alpha^2)^s}. \quad (6)$$

In addition, whenever the disk has sharp edges (or discontinuous transitions in surface density), the expression (A25) for  $A_d$  features *boundary terms*  $A_d^{\text{edge}}$ , given by the equation (A27). These terms involve  $b_{1/2}^{(0)}(\alpha)$  and its derivative  $b_{1/2}^{(0)'}(\alpha)$ , both evaluated at  $\alpha = a_p/a_e$ , where  $a_e$  is the semimajor axis of the disk edge. As the semimajor axis of a test particle  $a_p$  approaches the disk edge, both  $b_{1/2}^{(0)}(\alpha)$  and  $b_{1/2}^{(0)'}(\alpha)$  diverge:  $b_{1/2}^{(0)}(\alpha) \sim \ln|1 - \alpha|$ , while  $b_{1/2}^{(0)'}(\alpha) \sim |1 - \alpha|^{-1}$ , where in this case  $\alpha = a_e/a_p \rightarrow 1$ . As a result, boundary terms, as well as  $A_d$ , diverge at the sharp disk edge. This singularity of the secular disturbing function is further explored in §4.

The divergence of  $A_d$  at the disk edge does not arise if  $\Sigma_d(a)$  goes to zero at the boundary in a sufficiently smooth fashion. Indeed,  $b_{1/2}^{(0)'}(\alpha)$  in the boundary terms in equation (A27) is multiplied by  $\Sigma_d$  at the edge;  $b_{1/2}^{(0)}(\alpha)$  is multiplied by both  $\Sigma_d$  and its radial derivative  $\Sigma_d'$ . As a result, boundary terms vanish whenever  $\Sigma_d \propto |a - a_e|^\kappa$  and  $\kappa > 1$  near the edge at  $a_e$ . Thus, in disks with  $\Sigma_d$  smoothly (faster than linearly in  $|a - a_e|$ ) turning to zero at a finite semimajor axis  $a_e$  the coefficient  $A_d$  has no boundary contributions and, hence, does not diverge at the boundary. The same is also true for disks without edges, which have their surface density smoothly declining to zero as  $a \rightarrow 0$  and  $a \rightarrow \infty$ . Both possibilities will be explored further in §4.

Similarly, the expression (A25) for  $B_d$  involves radial integrals over the product of  $\Sigma_d(a)$  and the disk eccentricity  $e_d(a)$ , as well as their derivatives up to second order (contribution  $B_d^{\text{bulk}}$  given by equation (A28)), in addition to the boundary terms  $B_d^{\text{edge}}$  represented by equation (A28). Once again,  $B_d^{\text{edge}}$  vanishes whenever  $\Sigma_d$  goes to zero at the disk edges sufficiently rapidly, e.g. as  $\Sigma_d \sim |a - a_e|^\kappa$  with  $\kappa > 1$ .

Other features of  $A_d(a_p)$  and  $B_d(a_p)$  behavior and their effect on secular dynamics will be discussed in §4-7.

### 3. NUMERICAL VERIFICATION

Having derived the analytical framework embodied in equations (3) and (A25)-(A29), we also provide its numerical verification. We do this by comparing the eccentricity evolution of test particles computed based on our analytical results with the results of direct orbit integration in the potential of an eccentric disk (computed numerically). Details of both approaches, as well as the disk models used in this comparison, are outlined below.

#### 3.1. (Semi-)analytical orbital evolution

One approach to calculating orbital evolution in the potential of an eccentric disk uses Lagrange equations for the evolution of orbital elements,  $\dot{k}_p = -(a_p^2 n_p)^{-1} \partial R_d / \partial h_p$ ,  $\dot{h}_p = (a_p^2 n_p)^{-1} \partial R_d / \partial k_p$ , where we use the secular expression (3) for the disturbing function  $R_d$ . As a result one, finds

$$\dot{k}_p = -A_d h_p - B_d \sin \varpi_d, \quad \dot{h}_p = A_d k_p + B_d \cos \varpi_d, \quad (7)$$

with the general solution given by the superposition of the free and forced eccentricity vectors (Murray & Dermott 1999):

$$\mathbf{e}_p = \mathbf{e}_{\text{free}}(t) + \mathbf{e}_{\text{forced}}, \quad (8)$$

$$\mathbf{e}_{\text{free}}(t) = e_{\text{free}}(\cos(A_d t + \varpi_0), \sin(A_d t + \varpi_0)), \quad (9)$$

$$\mathbf{e}_{\text{forced}} = e_{\text{forced}}(\cos \varpi_d, \sin \varpi_d). \quad (10)$$

Here the forced eccentricity is

$$e_{\text{forced}}(a) = -\frac{B_d(a)}{A_d(a)}, \quad (11)$$

and constants  $e_{\text{free}} > 0$  and  $\varpi_0$  are such that at  $t = 0$ , equation (8) satisfies the initial condition  $\mathbf{e}_p(0) = (k(0), h(0)) = e_p(0)(\cos \varpi_p(0), \sin \varpi_p(0))$ , where  $e_p(0)$  and  $\varpi_p(0)$  are the initial eccentricity and periastron angle of an object.

In particular, an object starting on a circular orbit ( $e_p(0) = 0$ ) has  $e_{\text{free}} = e_{\text{forced}}$ , and its motion is described by

$$e_p(t) = \left| \frac{2B_d}{A_d} \sin \frac{A_d t}{2} \right|, \quad (12)$$

$$\tan \varpi_p(t) = \tan \left( \frac{A_d t}{2} + \varpi_d + \frac{\pi}{2} \right),$$

where  $\varpi_p$  stays in the range  $(0, 2\pi)$ . We will use this solution for numerical verification of our semi-analytical results, although any alternative solution corresponding to different initial conditions would work just as well.

Full determination of  $\mathbf{e}_p$  requires calculating  $A_d$  and  $B_d$  for the given profiles of the disk surface density  $\Sigma_d(a)$  and eccentricity  $\varpi_d(a)$ . We do this with the help of equations (A25)-(A29) by numerically evaluating the corresponding integrals. Despite the weak (logarithmic) singularity of the integrands, the integrals themselves are convergent and are calculated directly without resorting to any type of softening of the integrand near its singular point (which is often done in studies of disk dynamics; see, e.g. (Hahn 2003; Touma et al. 2009; Tremaine 2001)). The lack of an ad hoc parameter (softening length) in our calculation is its important distinctive feature.

Once  $A_d$  and  $B_d$  are calculated, equations (12) (or, more generally, equations (8)-(11)) provide a full semi-analytical description of the particle motion in the disk potential.

#### 3.2. Direct orbit integration

An alternative method to compute particle motion uses the direct orbit integrator MERCURY (Chambers 1999), which employs the Bulirsch-Stoer algorithm (Press et al. 1992). This integration uses as its input the gravitational acceleration  $\mathbf{g}$  due to the disk potential, which is

calculated on a grid of (245, 100) points<sup>7</sup> in  $(r, \varphi)$ . Values of  $\mathbf{g}$  on the grid are then interpolated to provide accurate accelerations everywhere in the disk.

At each grid point,  $\mathbf{g}$  is computed by direct summation of  $\nabla\Phi_d$  (see equation (A1)) over the disk surface, with the surface density given by equation (2) in its *exact* form, i.e. not performing small- $e_d$  expansion. This 2-dimensional integral is convergent in the Cauchy principal value sense, even though the integrand diverges at the point where acceleration is calculated. To avoid this mathematical singularity, our numerical evaluation is performed at a small height ( $10^{-3}$  AU or smaller) above the disk; we make sure that the result is convergent with respect to the value of this height. Note that this procedure is not equivalent to the introduction of softening in our theoretical calculation.

We integrate particle orbits starting with  $e_p = 0$ . This allows us to directly compare the results with the solution (12) following from our semi-analytical calculation, thus directly verifying the accuracy of our framework for the secular evolution in the potential of an eccentric disk.

### 3.3. Summary of the disk models used

In our comparison effort, we use the following model distributions of the disk eccentricity  $e_d(a)$ :

$$e_p(r) = e_0 \left( 1 + \frac{a_2 - a}{a_2 - a_1} \right), \quad (13)$$

$$e_p(r) = e_0 \frac{(a - a_1)^2 (a - a_2)^2}{(a_2 - a_1)^4}, \quad (14)$$

and surface density at periastron  $\Sigma_d(a)$ ,

$$\Sigma_d(a) = \Sigma_0 \frac{(a - a_1)^2 (a - a_2)^2}{(a_2 - a_1)^4}, \quad (15)$$

$$\Sigma_d(a) = \Sigma_0 \exp \left[ \frac{4 - [(a/a_c) + (a_c/a)]^2}{0.18} \right], \quad (16)$$

$$\Sigma_d(a) = \Sigma_0 \left[ 1 + 4 \sin \left( \pi \frac{a - a_1}{a_2 - a_1} \right) \right]. \quad (17)$$

with  $a_1 = 0.1$  AU,  $a_2 = 5$  AU,  $a_c = 1.5$  AU and  $e_0$  and  $\Sigma_0$  being the normalization factors that we vary in our models. In our calculations, we always assume the mass of the central star to be  $1 M_\odot$ .

Figure 2 illustrates the behavior of these profiles of  $\Sigma_d$  and  $e_d$ . In the majority of our calculations, we use the linear eccentricity profile (13), corresponding to a disk that is everywhere eccentric. Disks with the eccentricity profile (14), studied in §4.3.2, have circular inner and outer edges but are eccentric in between.

The surface density profile (15) corresponds to a disk in which  $\Sigma_d$  smoothly goes to zero at the edges at  $a_1$  and  $a_2$ . Based on the discussion in §2.2, we expect  $A_d$  and  $B_d$  to not diverge at the edges of such a disk, which is verified in §4.1.1. Profile (16) describes a disk without edges (§4.1.2), which has its surface density exponentially decreasing for both  $a \ll a_c$  and  $a \gg a_c$ . This disk also does not feature boundary terms (as it has no boundaries). Finally, the  $\Sigma_d$  profile (17) describes a disk with discontinuous drops of the surface density at

<sup>7</sup> We also tried denser grids but have not found any difference in the outcomes.

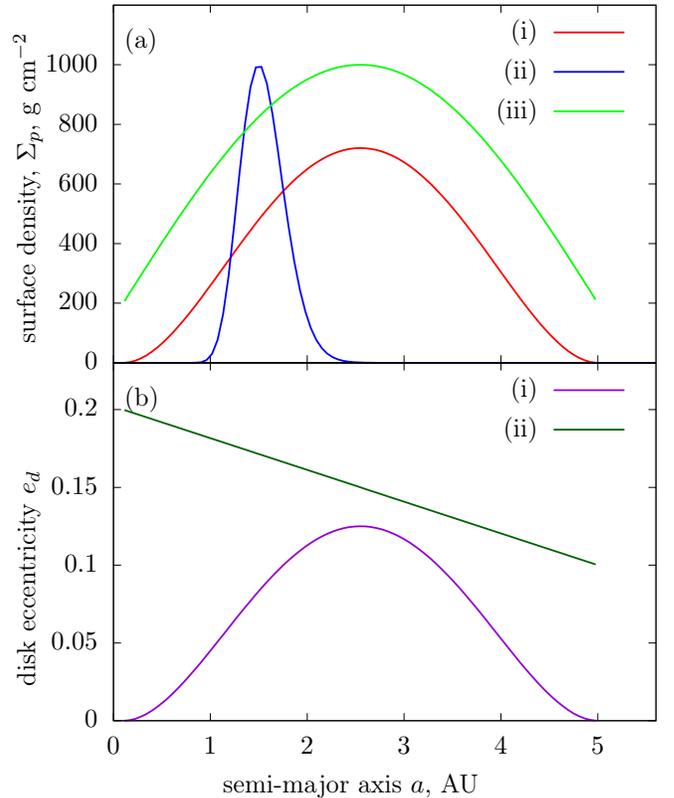


FIG. 2.— Profiles of the disk surface density  $\Sigma_d(a)$  and eccentricity  $e_d(a)$  considered in this work. Panel (a) shows the following  $\Sigma_d(a)$  profiles: (i) is given by formula (15) with  $\Sigma_0 \approx 1100 \text{ g cm}^{-2}$ , (ii) - by formula (16) with  $\Sigma_0 = 100 \text{ g cm}^{-2}$ , (iii) - by formula (17) with  $\Sigma_0 = 20 \text{ g cm}^{-2}$ . Panel (b) shows the  $e_d(a)$  profiles as follows: (i) is given by formula (14) with  $e_0 = 2$  and (ii) by formula (13) with  $e_0 = 0.1$ .

the edges. In this disk, we expect  $A_d$  and  $B_d$  to diverge at the boundaries, which is verified in §4.1.3.

Figure 3 illustrates the 2D distributions of the surface density obtained via equation (2) with some of the  $\Sigma_d$  and  $e_d$  profiles listed above. One can see that  $\Sigma(r, \varphi)$  can have a rather complicated structure, depending on the particular disk model used. The contours of constant  $\Sigma(r, \varphi)$  (on the left) are very different from the elliptical trajectories (on the right) of the mass elements giving rise to the surface density distribution in the disk.

## 4. RESULTS

We start our comparison by providing an illustration of the orbital evolution caused by the gravity of an underlying eccentric disk. Figure 4 displays the variation of the test particle eccentricity  $e_p$  (on the left) and apsidal angle  $\varpi_p$  (on the right) in time, computed at several values of the semimajor axis  $a_p$  for a particular disk model —  $\Sigma_d(a)$  given by equation (15) with  $\Sigma_0 \approx 1100 \text{ g cm}^{-2}$  and  $e_d(a)$  given by equation (13) with  $e_0 = 10^{-2}$ . One can see that the agreement between the direct integration (green) and our secular prediction (red) is very good at all radii. Both  $e_p(t)$  and  $\varpi_p(t)$  follow the predictions of (12) very closely, agreeing both in the amplitude of the eccentricity oscillations and in their phase (or period).

From the behavior of  $\varpi_p(t)$  at different semimajor

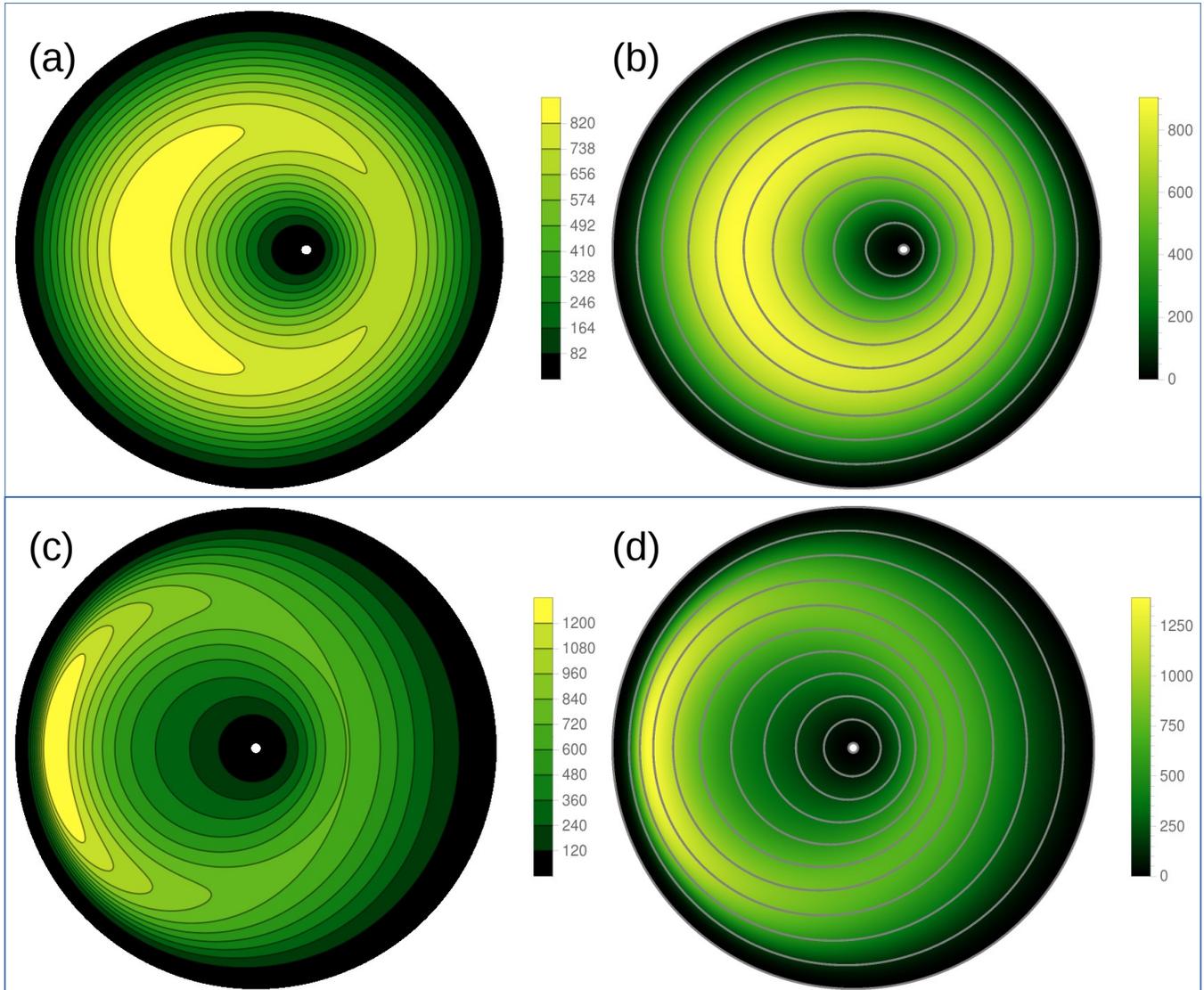


FIG. 3.— Maps of the disk surface density  $\Sigma(r, \varphi)$  (color indicates the amplitude of  $\Sigma$ ) with overlaid contours of  $\Sigma$  (thin curves; left panels) and eccentric trajectories of the mass elements comprising the disk (grey; right panels). The disk is always oriented with the apsidal lines pointing to the right. Top panels (a,b) are drawn for  $\Sigma_d(a)$ , given by equation (15) with  $\Sigma_0 \approx 1100 \text{ g cm}^{-2}$ , and  $e_d(a)$  given by (13) with  $e_0 = 0.2$  (this high  $e_0$  was chosen to better illustrate the  $\Sigma$  distribution). Bottom panels (c,d) are drawn for the same  $\Sigma_d(a)$  but a different disk eccentricity profile (14) with  $e_0 = 0.5$ . Note the substantial difference in  $\Sigma(r, \varphi)$  caused by just the difference in the  $e_d(a)$  profiles.

axes, one can immediately see that free precession of the particle orbit can be *prograde* in some parts of the disk and *retrograde* in others. The change of sign of the precession occurs for orbits fully enclosed within the disk. It is also clear that not only the sign but also the period of the precession is a function of the location in the disk.

Evolutionary time series of  $e_p(t)$ , such as the one presented in Figure 4, allow one to easily measure the maximum amplitude of the eccentricity oscillations,  $e_p^m$ , which should be compared to the theoretical values of  $2B_d/A_d$ ; see equation (12). Similarly, the period  $P_{\text{sec}}$  of each  $e_p$  oscillation yields the corresponding free precession rate as  $\dot{\omega}_{\text{sec}} = 2\pi/P_{\text{sec}}$ , which in this Figure agrees very well with the theoretical  $A_d$ . The value of  $\dot{\omega}_{\text{sec}}$  can also be independently inferred from the slope of the numerical  $\dot{\omega}(t)$  curves, which should be equal to

$A_d/2$ , see equation (12).

Close inspection of Figure 4 reveals additional features in the  $e_p(t)$ ,  $\varpi_p(t)$  evolution curves beyond the large-scale secular oscillations, which are well described by the solution (12). These features manifest themselves as small-amplitude, short-period oscillations around the purely secular solution, most pronounced near the outer edge of the disk. The period of these oscillations is equal to the local orbital period  $2\pi/n_p$ . Their amplitude scales linearly with the disk mass, depends on the semimajor axis of test particle  $a_p$ , but is independent of the disk eccentricity (we observe oscillations with the same amplitude even when the disk eccentricity is zero). We interpret these oscillations as resulting from the nonelliptical shape of the particle orbits in the combined potential of the star and the disk. As the

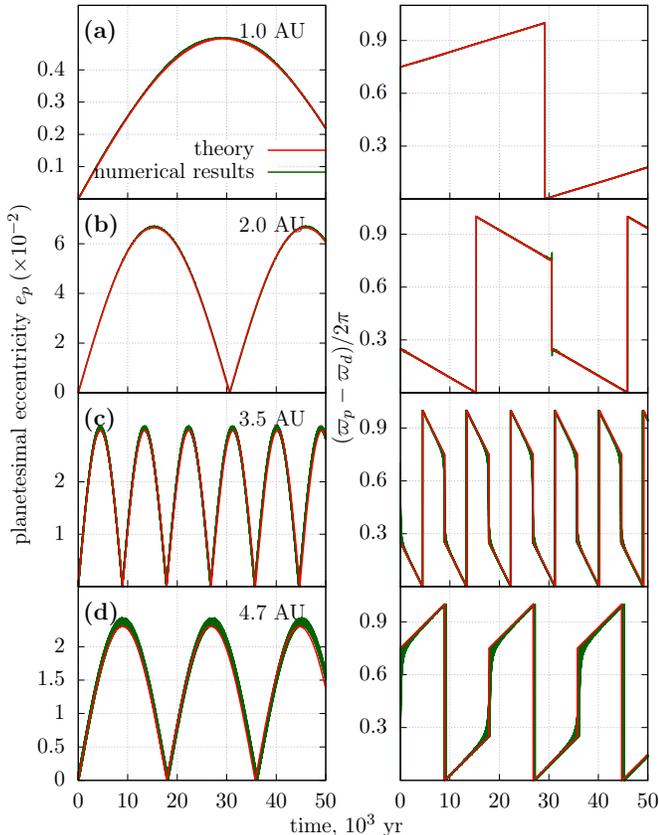


FIG. 4.— Verification of our analytical calculation of the disk disturbing function  $R_d$  using direct numerical integration of particle orbits in the (numerically computed) disk potential with MERCURY. The disk model has  $\Sigma_d(a)$  given by (15) with  $\Sigma_0 \approx 1100 \text{ g cm}^{-2}$  and  $e_d(a)$  given by (13) with  $e_0 = 0.01$ . The time evolution of the particle eccentricity  $e_p$  (left) and apsidal angle  $\varpi_p$  (right) is shown for different values of the particle semi-major axis  $a_p$ , as labeled on the panels. In all cases, test particles start with zero eccentricity, which results in pronounced secular oscillations of  $e_p$ , well described by the solution (12). The true eccentricity vector precesses at a steady rate, resulting in  $\varpi_d$  evolution characterized by the solution (12).

numerical integration outputs osculating orbital elements (effectively fitting a pure ellipse at every point of a truly nonelliptical trajectory), oscillations of  $e_p$  (as well as  $\varpi_p$  and  $a_p$ ) on a local dynamical timescale naturally arise. We did verify that the angular momentum of a test particle is strictly conserved through these oscillation cycles when  $e_d = 0$ . A similar effect was discussed by Georgakarakos (2002) in application to hierarchical triples.

In our subsequent presentation, we will focus on the behavior of  $e_p^m$  and  $\dot{\varpi}_{\text{sec}}$ , derived from the data similar to that shown in Figure 4, in different disk models.

#### 4.1. Radial variation for different $\Sigma_d$ models

We now test the accuracy of our secular theory in disks with different  $\Sigma_d(a)$  profiles. This allows us to examine the different possible behaviors of the secular coefficients  $A_d$  and  $B_d$  as well as to see how sensitive the agreement with the numerical results is to the different features of the  $\Sigma_d(a)$  distributions.

##### 4.1.1. $\Sigma_d$ smoothly vanishing at the edges

We start by looking at the secular effect of a disk, in which  $\Sigma_d$  smoothly goes to zero at the boundaries, i.e. the one with  $\Sigma_d$  given by the equation (15). In Figure 5 we display the theoretical radial profiles of  $A_d$ ,  $B_d$ , and  $2B_d/A_d$  for such a disk. The curves of  $A_d(a_p)$  are compared with the numerically determined  $\dot{\varpi}_{\text{sec}}$  (blue dots), while the theoretical  $2B_d(a_p)/A_d(a_p)$  is compared against the numerical  $e_p^m$ . This particular calculation uses the linear eccentricity profile  $e_d(a)$  given by the equation (13) with relatively low  $e_0 = 0.01$ , so that the maximum value of the disk eccentricity (reached at its inner edge) is 0.02. Thus, the requirement  $e_d \ll 1$  necessary for the secular results (3), (A25)-(A29) to apply is fulfilled.

One can see an almost perfect match between the numerical results and theory, both in terms of the amplitude of the disk-induced particle eccentricity as well as the period of the associated orbital precession. Theoretical calculation easily reproduces even the very subtle features of the  $e_p^m(a)$  behavior, including the variations happening on very short radial scales manifesting themselves as sharp features in Figure 5. There are several features to note in this figure.

First,  $A_d$  has a different sign in different parts of the disk: precession of the free eccentricity vector is *prograde* near the boundaries of the disk, while it is *retrograde* away from the edges. This change of sign of  $\dot{\varpi}_{\text{sec}}$  was clear already in Figure 4 (which is drawn for the same disk model as Figure 5, at the locations highlighted with circles in the latter), but the Figure 5 provides a much more detailed representation of the different characteristics of the secular effect of the disk.

Second,  $B_d$  also changes sign as  $a_p$  varies. As a result of sign variations of both  $A_d$  and  $B_d$ , the forced eccentricity vector  $\mathbf{e}_{\text{forced}}$  can be *aligned* with the apsidal line of the disk in some intervals of  $a_p$ , and *anti-aligned* with it in others.

Third, both numerical and analytical  $e_p^m$  exhibit formal singularity at two distinct locations in this disk,  $a \approx 1.54 \text{ AU}$  and  $\approx 4.39 \text{ AU}$ . The origin of these singularities can be traced to the expression (11) for the forced eccentricity (recall that our theory predicts  $e_p^m = 2e_{\text{forced}}$ ), which has  $A_d$  in the denominator, and the fact that  $A_d$  crosses zero (as the sense of  $\mathbf{e}_{\text{free}}$  precession changes) at these locations, see Figure 5b. We will discuss these singularities in more detail in §4.2.

Fourth, at certain values of the semimajor axis (at  $\approx 0.92 \text{ AU}$  and  $\approx 4.39 \text{ AU}$ ) the disk-induced forced eccentricity vanishes. This happens because  $B_d$  changes sign at these locations so that  $B_d \rightarrow 0$ ; see Figure 5c and equation (11). Interestingly, one of the semimajor axes where  $e_p^m \rightarrow 0$  ( $a_p \approx 4.39 \text{ AU}$ ) is located in the immediate proximity of the singularity of  $e_p^m$ . This occurs because in this part of the disk,  $A_d$  and  $B_d$  go through zero at almost the same (but still slightly different) values of  $a_p$ .

Fifth, for the surface density profile (15), we find that  $A_d$  and  $B_d$  remain finite everywhere in the disk, including the boundaries. This is in line with the expectations outlined in §2.2 for the disks with  $\Sigma_d$  smoothly going to zero at the edges, which do not result in divergent boundary terms in  $A_d$  and  $B_d$ .

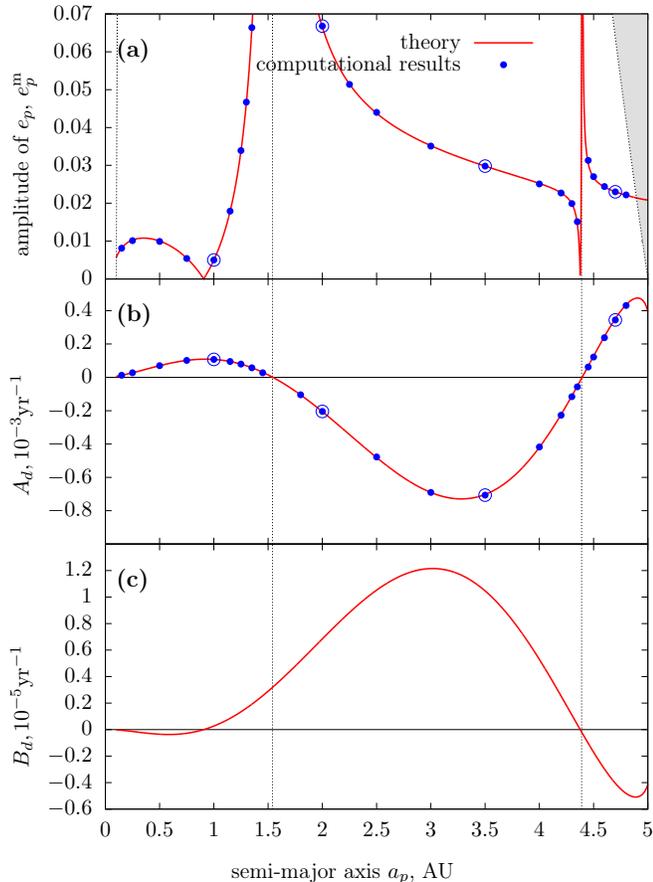


FIG. 5.— Characterization of secular oscillations of test particles (with orbits fully enclosed within the disk) driven by the disk gravity. Shown as a function of the semimajor axis of test particle  $a_p$  are: (a) amplitude  $e_p^m$  of the eccentricity oscillations (blue dots) compared with theoretical  $2e_{\text{forced}} = 2B_d/A_d$  (red curve), (b) the frequency of secular oscillations  $\tilde{\omega}_{\text{sec}}$  (blue dots) compared against  $A_d$  (red curve) and (c) the coefficient  $B_d$  of the theoretical secular disturbing function. This calculation uses the disk model with  $\Sigma_d$  distribution (15) and  $\Sigma_0 \approx 1100 \text{ g cm}^{-2}$ , so that  $\Sigma_d$  goes to zero at finite semimajor axes (0.1 AU and 5 AU), but smoothly, so that the boundary terms in the expressions for  $A_d$  and  $B_d$  do not arise; see §2.2. The radial profile of the disk eccentricity  $e_d(a)$  is given by equation (13) with  $e_0 = 0.01$ . Circled dots correspond to the four values of  $a_p$  used in Figure 4. Vertical dotted lines mark the locations where  $A_d = 0$ . Gray regions near the disk edges correspond to the edge noncrossing constraints  $a_p(1 + e_p^m) < a_2$  at apastron and  $a_p(1 - e_p^m) > a_1$  at periastron. One can see excellent agreement between the secular theory and the direct numerical integration.

#### 4.1.2. Disk without boundaries

Next we consider secular dynamics in the potential of a Gaussian ring with  $\Sigma_d$  given by the equation (16), which does not feature well-defined edges. We plot the behavior of the corresponding  $e_p^m$ ,  $A_d$ , and  $B_d$  in Figure 6.

One can see that many features of secular dynamics present in the case studied in §4.1.1 are present here as well: both  $A_d$  and  $B_d$  are finite, they vary with  $a_p$  and change sign, while  $e_p^m$  exhibits both singularities and nulls. This is likely related to the fact that these two  $\Sigma_d$  profiles are morphologically similar: they exhibit no discontinuities, are single-peaked, and smoothly decay to zero away from the peak. The only obvious difference is

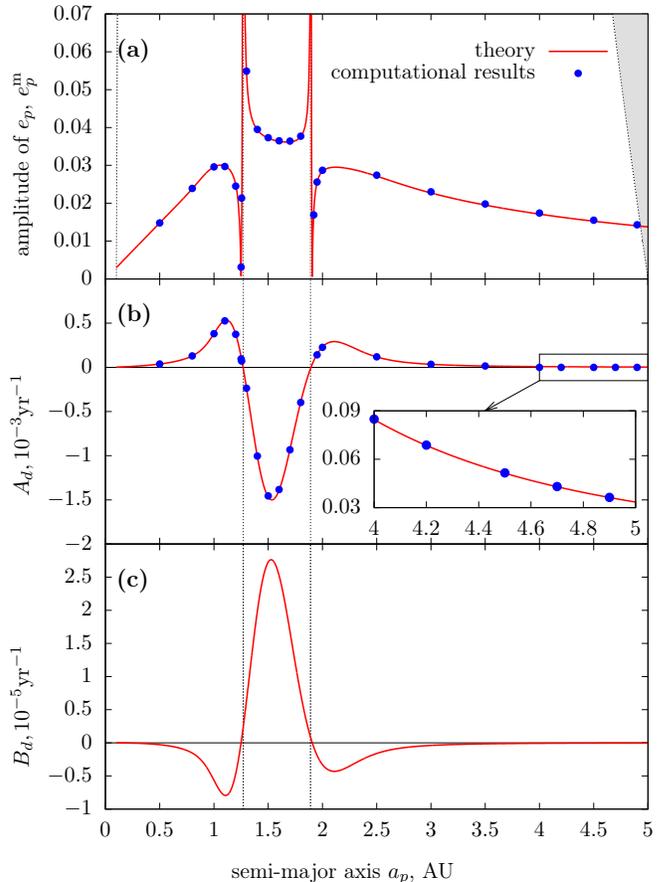


FIG. 6.— Same as Figure 5, but now for a disk (Gaussian ring) model with different  $\Sigma_d(a)$  distribution (16), with  $\Sigma_0 = 100 \text{ g cm}^{-2}$ ; this model does not have boundaries at finite radii. The disk eccentricity profile is still given by equation (13) with  $e_0 = 0.01$ . The inset in panel (b) shows the behavior of  $A_d$  far from the main body of the ring.

that in the Gaussian case in Figure 6 the nulls of  $e_p^m$  are located very close to *both* singularities of  $e_p^m$ . However, the significance of this difference is unclear.

#### 4.1.3. $\Sigma_d$ sharply truncated at the edges

We also examined secular behavior for the disk model in which surface density displays a discontinuous drop to zero at the inner and outer edges, namely, the one represented by equation (17). Sharply truncated  $\Sigma$  distributions are rather typical in planetary rings and other astrophysical disks, and it is important to study the effect of their gravity on the dynamics of embedded objects.

Figure 7 illustrates secular dynamics in the potential of such a disk. It again shows the variation of sign of  $A_d$ , which results in the emergence of two singularities of  $e_p^m$  — one very close to the inner edge of the disk at  $a_p = 0.11 \text{ AU}$  and another at  $a_p \approx 1.28 \text{ AU}$ . However, for this disk model,  $B_d$  does not change sign — it always stays positive. As a result, there are no nulls of  $e_p^m$  for orbits in the potential of such a disk.

An even more dramatic difference with the cases explored in §4.1.1-4.1.2 is the behavior of  $A_d$  and  $B_d$  near the disk edges. Unlike the two previous cases,

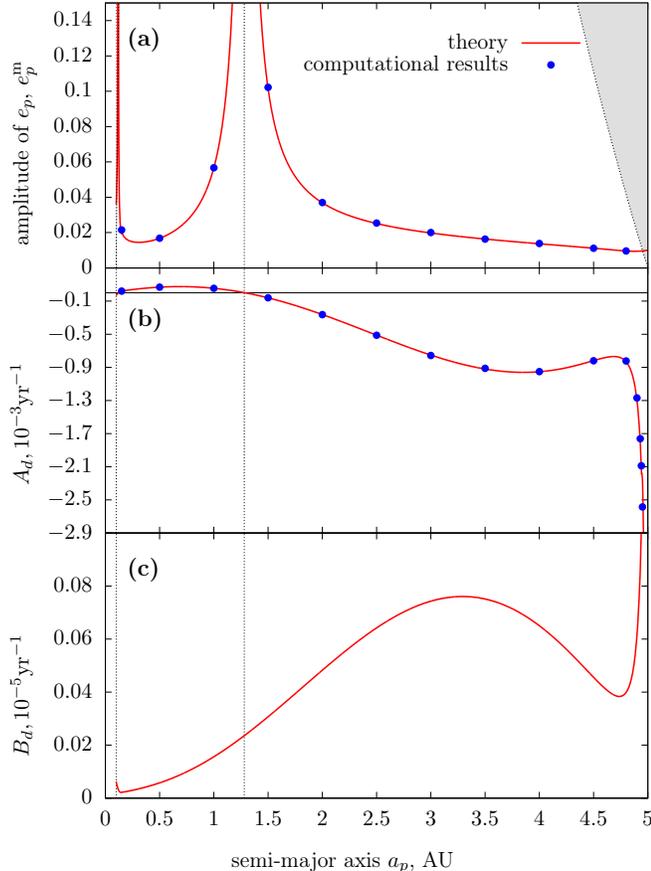


FIG. 7.— Same as Figure 5, but now for the disk model with  $\Sigma_d(a)$  distribution (17) and  $\Sigma_0 = 20 \text{ g cm}^{-2}$ , which has sharp edges at finite semimajor axes (0.1 AU and 5 AU). The radial profile of the disk eccentricity is still given by equation (13), but now with the overall normalization  $e_0 = 0.005$ . Note the sharp increase of the amplitude of  $A_d$  near the outer edge of the disk, accurately matched by our theory.

in which both coefficients remained finite everywhere at all radii, in a sharply truncated disk,  $A_d$  and  $B_d$  exhibit singularity as the disk edge is approached. This is very clearly seen at the outer<sup>8</sup> boundary of the disk in Figure 7, where both theoretical curves and the results of orbit integrations exhibit divergent behavior. Unfortunately, we cannot probe this divergence in great detail numerically, as the orbits of test particles start crossing the edge of the disk.

At the same time, the test particle eccentricity  $e_p^m$  remains finite at the disk edges, even though both  $A_d$  and  $B_d$  are singular there. This is because both coefficients of the disturbing function (3) diverge in similar fashion at the edge, so that  $e_{\text{forced}}$  remains finite there.

#### 4.2. Singularities of $e_p^m$

A common feature for all  $\Sigma_d$  distributions examined in §4.1 is the emergence of multiple singularities of  $e_p^m$ . At these locations,  $e_p^m$  formally diverges, and the assumption  $e_p \ll 1$  used in deriving our secular disturbing function (3)-(5) breaks down. The way in which this happens is

<sup>8</sup> Numerical issues prevent us from demonstrating the divergent behavior of  $A_d$  and  $B_d$  at the inner edge of the disk.

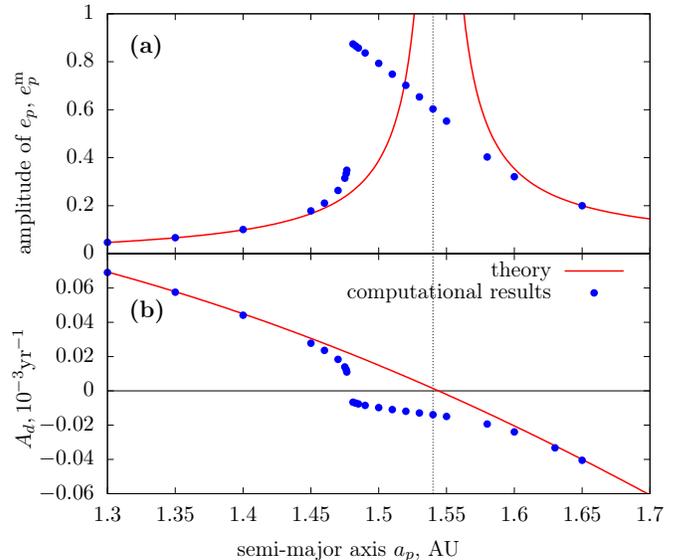


FIG. 8.— Zoom-in on a part of Figure 5 in the vicinity of a secular singularity at  $a \approx 1.5$  AU. One can see how particle motion starts to deviate from our lowest-order secular theory as  $e_p$  grows to values of order unity.

illustrated in Figure 8, where we compare the numerical and analytical results in the vicinity of one of the singularities (near  $a_p = 1.54$  AU) in a disk with  $\Sigma_d(a)$  given by equation (15); see Figure 5.

One can see that as the theoretical singularity is approached, the behavior of  $\dot{\omega}_{\text{sec}}$  starts to deviate from the prediction (A25)-(A27). As a result,  $\dot{\omega}_{\text{sec}}$  goes through zero at a location slightly different from the one where  $A_d = 0$ . Note that even at the point where  $\dot{\omega}_{\text{sec}} = 0$  particle eccentricity remains finite (even though it reaches values close to 1). This means that equation (11) is no longer valid when  $e_p \sim 1$  and that additional, higher-order terms become important in addition to the lowest-order secular potential contribution (3). This discrepancy could be at least partly ameliorated by including higher-order (in  $e_p$ ) terms in the calculation of  $R_d$ , as was done recently in Sefilian & Touma (2018) for the case of power-law disks.

As evidenced by Figures 5 and 6,  $e_p^m$  singularities often occur in the immediate vicinity of nulls of  $B_d$ . This results in a characteristic shape of these singularities, with  $e_p^m$  sharply dropping to zero in close proximity to the singularity. This leads to a dramatic difference in the eccentricities of particles with almost identical semimajor axes, resulting in their orbits crossing. Such locations thus provide a natural environment for particle collisions.

It is not clear why in some cases conditions  $A_d = 0$  and  $B_d = 0$  get realized at almost the same value of  $a$ . Inspection of the integrands in equations (A26), (A28) does not reveal an obvious reason for that to be the case. Interestingly, we find such "null-singularity" pairs only in the two disks without sharp edges (§4.1.1-4.1.2), for which the boundary terms  $A_d^{\text{edge}}$  and  $B_d^{\text{edge}}$  in equations (A25) vanish. The disk with sharply truncated  $\Sigma_d$  (see §4.1.3 and Figure 7) has singularities without neighboring nulls of  $B_d$  (in fact,  $B_d$  does not change sign in this disk). Whether this outcome is due to

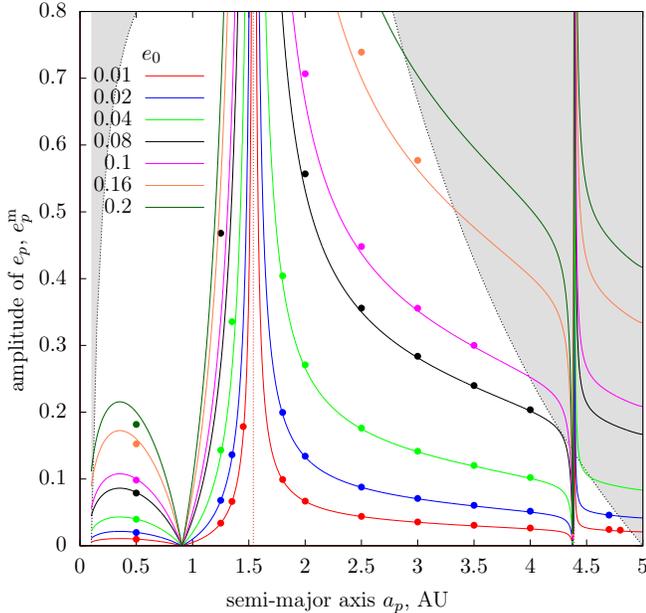


FIG. 9.— Agreement between our secular theory and direct orbit integrations as a function of the disk eccentricity amplitude. Shown is the amplitude of secular oscillations found in direct orbit integrations, as well as theoretical values of  $2e_p^{\text{forced}}$ , shown as a function of both the distance from the star (horizontal axis) and the overall normalization  $e_0$  of the disk eccentricity  $e_d$  (different curves). This calculation assumes surface density profile in the form (15) with  $\Sigma_0 \approx 1100 \text{ g cm}^{-2}$  and the eccentricity profile given by equation (13). Continuous curves of different colors (corresponding to different values of eccentricity normalization  $e_0$  in equation (13), as shown in the panel) represent analytical results based on our secular theory. Dots of different colors display the corresponding numerical results. Other notation is the same as in Figure 5.

the nontrivial boundary terms in this disk model is not clear. These curious properties of  $e_p^m$  singularities deserve further investigation.

#### 4.3. Sensitivity to the disk eccentricity $e_d$

Next we examine how our secular theory fares against changes of the disk eccentricity. First, we explore the effect of uniformly varying just the amplitude of eccentricity (4.3.1), keeping the radial profile of  $e_d(a)$  the same. We then look at the effect of a different radial profile of  $e_d(a)$  on the agreement between the theory and numerical calculations (§4.3.2).

##### 4.3.1. Variation of the disk eccentricity amplitude

In Figure 9, we plot  $e_p^m(a)$  for the disk model with  $\Sigma_d$  and  $e_d$  given by equations (15) and (13), where we set  $\Sigma_0 \approx 1100 \text{ g cm}^{-2}$  but vary the eccentricity normalization  $e_0$  as indicated in the Figure (which is very similar to Figure 5a).

One can see that our theory works surprisingly well in predicting  $e_p^m(a)$  even when particle eccentricity reaches values in excess of  $e_p = 0.5$ , when one would naively expect the description based on the lowest-order secular disturbing function (3) to break down. For the  $e_d$  profile (13), the maximum value of  $e_d$ , reached at the inner boundary of the disk, is  $2e_0$ . Thus, even for disks with the inner eccentricities reaching  $e_d(a_1) = 0.4$  our theory

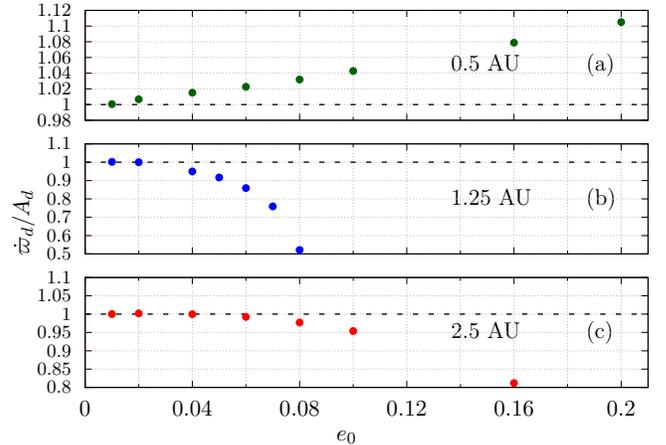


FIG. 10.— Relative deviation between the numerical ( $\tilde{\omega}_{\text{sec}}$ ) and theoretical ( $A_d$ ) precession rates  $\tilde{\omega}_{\text{sec}}/A_d$ , plotted as a function of the eccentricity amplitude  $e_0$  of the linear disk eccentricity profile given by equation (13). Different panels correspond to different semimajor axes of the test particle: (a)  $a_p = 0.5$ , (b) 1.25, (c) 2.5 AU. The calculation assumes  $\Sigma_d$  model (15) with  $\Sigma_0 \approx 1100 \text{ g cm}^{-2}$ .

performs quite well.

The amplitude of eccentricity oscillations  $e_p^m(a)$  is just one metric by which the performance of our theory can be judged. Another obvious one is the free precession rate  $\tilde{\omega}_{\text{sec}}$ . In Figure 10 we plot the ratio of  $\tilde{\omega}_{\text{sec}}$  to its analytical counterpart  $A_d$  as a function of the disk eccentricity normalization  $e_0$ . In the framework of our secular calculation, the precession rate should not depend on  $e_0$  and be equal to  $A_d$ .

Figure 10 shows that this is not really the case and  $\tilde{\omega}_{\text{sec}}$  does deviate from  $A_d$  when  $e_0$  is nonnegligible. The agreement between these two frequencies is somewhat less impressive than for  $e_p^m(a)$ , with  $\tilde{\omega}_{\text{sec}}$  deviating from  $A_d$  by tens of percent already for  $e_0 = 0.1$ . Note that the discrepancy between the numerical and analytical secular frequencies is a strong function of the semimajor axis  $a_p$ , with the largest deviation occurring in the vicinity of the  $\tilde{\omega}_{\text{sec}}$  singularity (just interior of it) at 1.3 AU. Thus, one should be careful when applying our lowest-order secular calculation to characterize the orbital evolution of test particles at certain locations in the disk with  $e_d \gtrsim 0.1$ .

##### 4.3.2. Variation of the disk eccentricity profile

Next we study the effect of changing the radial profile of the disk eccentricity  $e_d(a)$ . Figure 11 is the analogue of Figure 9 but made for a disk with an eccentricity profile (14).

The first thing to note is that the radial profile of the theoretical  $e_p^m(a)$  still shows two clear singularities at the same radial locations as in Figure 9. This is easy to understand, since  $e_p^m$  diverges at radii where  $A_d \rightarrow 0$ , and  $A_d$  is independent of the disk eccentricity profile (it depends only on the  $\Sigma_d(a)$  profile). As a result, singularities of  $e_p^m$  stay at fixed locations even though  $e_d(a)$  varies.

Second, the detailed shape of  $e_p^m(a)$  is notably different from that shown in Figures 5 & 9, despite the fact that the  $\Sigma_d(a)$  profile is the same in both cases. While

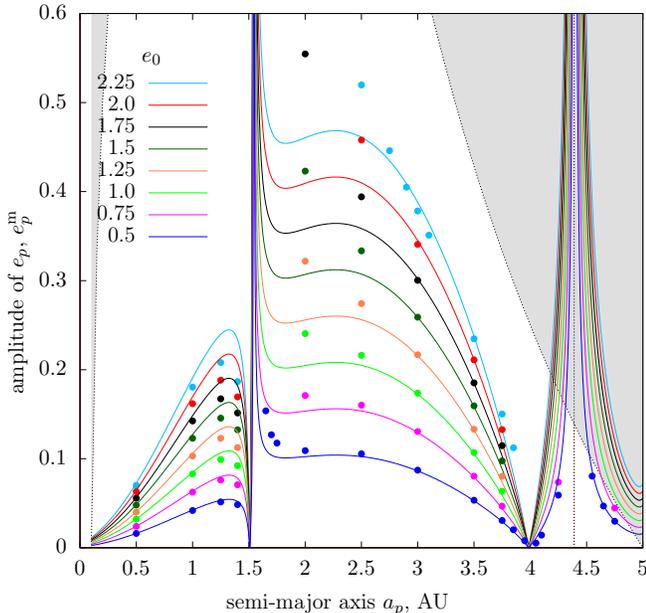


FIG. 11.— Same as Figure 9 but now for a different eccentricity profile given by equation (14), with different  $e_0$  corresponding to this profile. Comparing the results with Figure 9 we see that variation of the  $e_d(a)$  profile changes the behavior of the maximum particle eccentricity, although the general topology of the curves remains roughly the same. The agreement between theory and direct orbit integrations is somewhat worse in this case compared to Figure 9.

previously  $e_p^m$  was dropping to zero right next to the outer singularity at  $a_p \approx 4.39$  AU, in Figure 11 this happens near the inner singularity at  $\approx 1.5$  AU. Another null of  $e_p^m$ , more distant from the singularity, has also swapped its location and now lies in the outer part of the disk in Figure 11.

Third, the agreement between the theory and direct orbit integrations is somewhat worse for the disk with an  $e_d(a)$  profile (14). This is especially noticeable to the right of the inner singularity. For example, the  $e_p^m$  differs by  $\approx 50\%$  from the theoretical expectation at  $a_p = 2$  AU for  $e_0 = 1.75$  (black curve, which corresponds to the maximum eccentricity in the disk of  $\approx 0.1$ ). Away from the region  $1.5 - 2$  AU the agreement is generally better, even though the particle eccentricity excited by the disk potential can be quite high,  $e_p \gtrsim 0.1 - 0.2$ .

We hypothesize that the reduced accuracy with which our secular theory predicts secular dynamics in the case of a disk with an  $e_d(a)$  profile (14) is caused by the fact that this disk features a rather nonaxisymmetric surface density distribution  $\Sigma(r, \varphi)$ . Indeed, Figure 3c,d show that the disk with  $e_d(a)$  given by (14) features a well-defined concentration of mass in the outer region near the apastra of the constituent trajectories. As a result, for certain values of  $a$  surface density  $\Sigma(r_d, \varphi_d)$  exhibits substantial variations *along the eccentric trajectories* of the mass elements comprising the disk, see Figure 3d. According to equation (2), this is only possible if  $\zeta_{e_d}$  is not small for these values of  $a$ , meaning that the assumption  $|\zeta_{e_d}| \ll 1$  underlying our expansion (A2) is not fulfilled, which explains the deviations seen for this particular disk model.

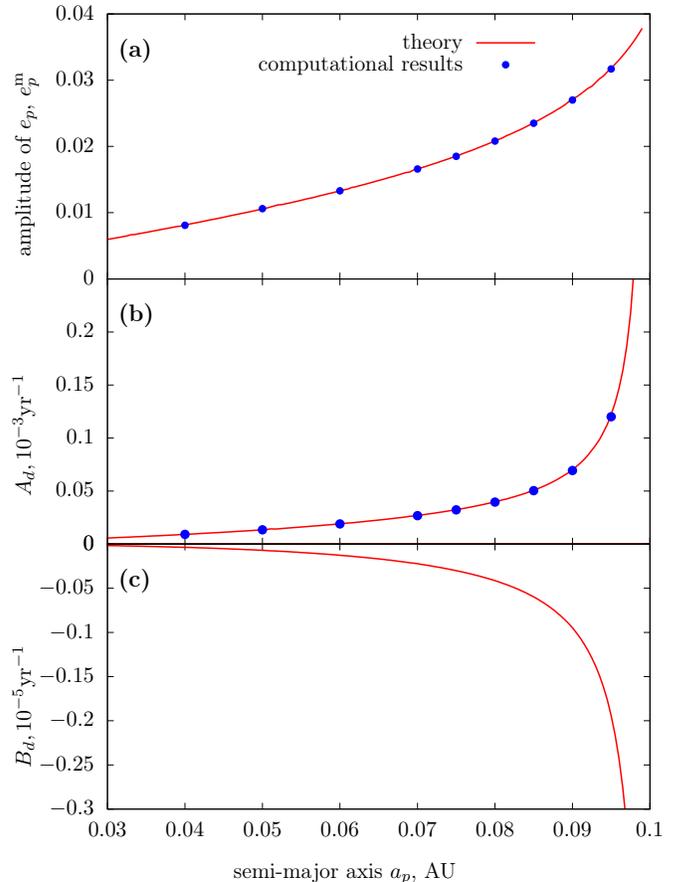


FIG. 12.— Characterization of the secular behavior in the potential of a disk with the same parameters as in Figure 7, but now computed for a test particle orbiting *inside* the inner hole of the disk at 0.1 AU (i.e. outside the main body of the disk with sharp edges). One can see that our theory still works very well even beyond the radial extent of the disk.

By contrast, the disk with the eccentricity profile (13) shows a more axisymmetric distribution of  $\Sigma(r, \varphi)$ , see Figure 3a,b, meaning smaller  $|\zeta_{e_d}|$  and higher accuracy of the expansion (A2). Thus, we expect that our secular theory should perform better for disks without highly nonuniform azimuthal features in  $\Sigma(r_d, \varphi_d)$ .

#### 4.4. Motion outside the disk

Our derivation of the secular disturbing function in Appendix A explicitly assumes that the particle orbits fully *within* the disk. In other words, in the case of a disk with  $\Sigma_d$  dropping to zero at some finite  $a_1$  and  $a_2$  the particle semimajor axis  $a_p$  satisfies  $a_1 < a_p < a_2$ , and its eccentric orbit does not cross the boundaries of the disk. Obviously, if the disk has no edges, as in, e.g., the case of a Gaussian ring (16), then the particle orbit is always fully enclosed within the disk.

However, close examination of our derivation of  $R_d$  in Appendix A demonstrates that it should also apply equally well *outside* the disk with sharp edges, as long as the particle orbit does not cross the disk boundaries. For example, in the case of a particle orbiting outside the outer edge of the disk ( $a_p > a_2$ ), first, one drops the contribution of the outer disk (i.e.  $a > a_p$  as there is no disk material there) when computing  $R_d$ , and, second,

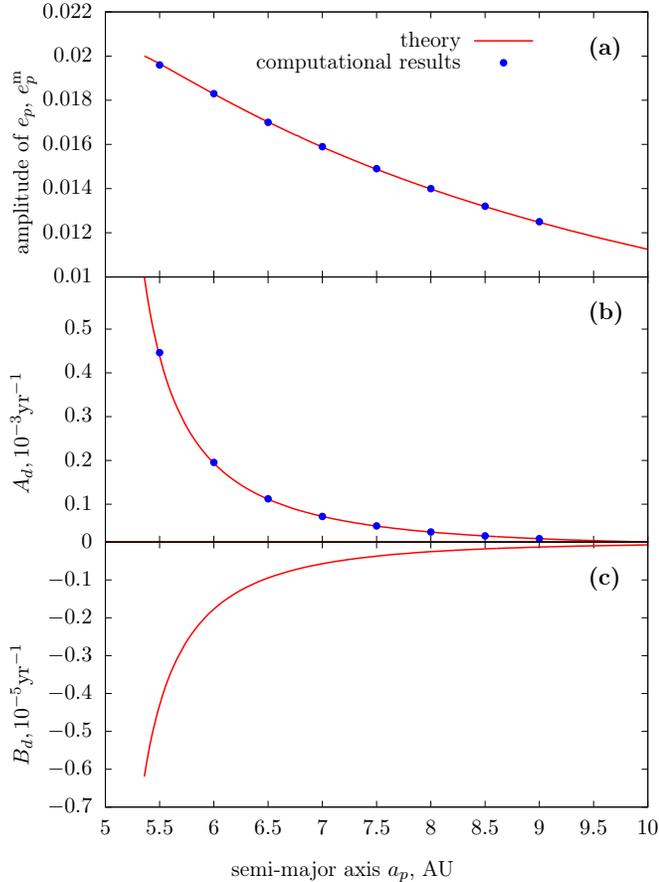


FIG. 13.— Same as Figure 12, but now for a test particle orbiting *outside* the outer edge of the disk at 5 AU.

the integration in the inner disk runs not up to  $\alpha = 1$  but only up to  $\alpha = a_2/a_p < 1$ . As a result, the resultant expressions (A25)-(A29) for the coefficients  $A_d$  and  $B_d$  apply without modification, even when  $a_p < a_1 < a_2$  or  $a_1 < a_2 < a_p$ .

To verify this claim, in Figures 12 & 13, we show radial profiles of  $e_p^m$  and  $\dot{\omega}_{\text{sec}}$ , as well as those of  $2e_{\text{forced}}$ ,  $A_d$  and  $B_d$  for particles orbiting inside and outside (correspondingly) the radial extent of the disk. Disk parameters are the same as in Figure 7; in particular,  $\Sigma_d$  is given by (17) with sharp edges at  $a_1 = 0.1$  AU and  $a_2 = 5$  AU.

One can see that, as in the previous cases illustrated in Figures 5-7, there is excellent agreement between our theory and direct orbit integrations, as long as the disk and particle eccentricities are low. Both the theoretical and the numerical values of the particle eccentricity extrapolated to the disk edges match the corresponding values extrapolated from inside the disk; see Figure 7. At the same time, both  $A_d$  and  $B_d$  diverge as  $a_p \rightarrow a_1 - 0$  and  $a_p \rightarrow a_2 + 0$ , mirroring the singularity of these coefficients identified previously (§2.2).

These results extend the applicability of our calculation of  $R_d$  to (coplanar) particles having *arbitrary* semimajor axis relative to the disk, as long as their orbits do not cross the edge of the disk where  $\Sigma_d$  discontinuously drops to zero. However, it can be

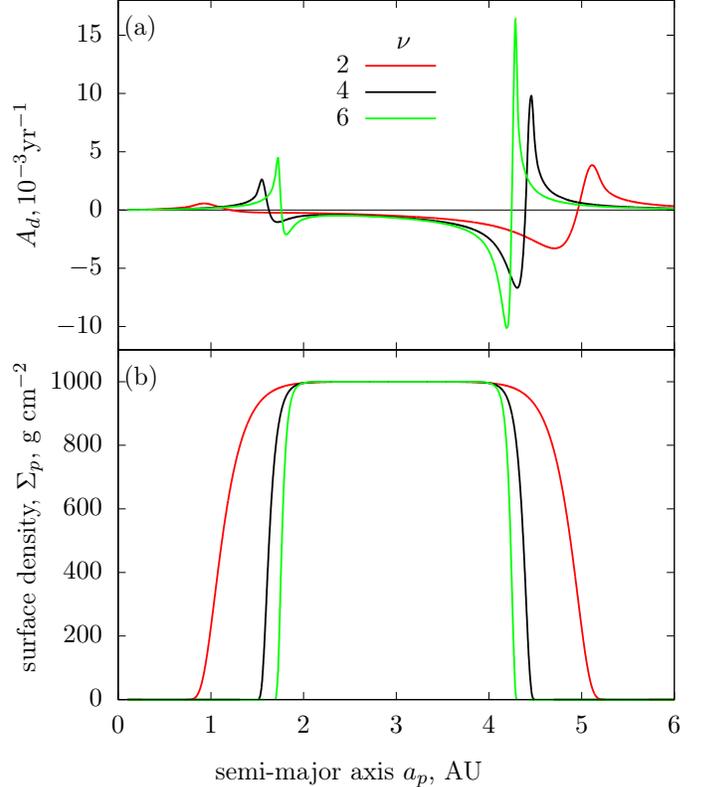


FIG. 14.— Illustration of the divergence of the free precession rate  $A_d$  (panel (a)) near the sharp edges of the disk. A massive circular ring with three different profiles of  $\Sigma_d$  given by equation (B1) and shown in panel (b), varying in sharpness of  $\Sigma_d$  drop near the edges (regulated by the parameter  $\nu$ , as shown in panel (a)), is considered. One can see that the more abrupt the variation of  $\Sigma_d$  is at the edges, the higher is the amplitude of  $A_d$  that is reached in these regions. In the limit of a discontinuity in  $\Sigma_d$  one would find  $A_d \rightarrow \infty$ , in agreement with our theory.

shown that even this latter constraint can be removed, extending the applicability of our results even further. We do not dwell on this point here<sup>9</sup>, deferring it to a future study.

## 5. APPLICATIONS

The results presented in §4 demonstrate the validity and accuracy of the secular theory developed in this work in the low- $e$  limit. This motivates us to use this theory to further explore several aspects of secular motion in the potential of an eccentric disk.

### 5.1. Edge effects

Some astrophysical disks are known to have very sharp edges. For example, using *Voyager 2* occultation data, Graps et al. (1995) demonstrated that the  $\epsilon$  ring of Uranus has  $\Sigma_d$  steeply dropping to zero at the ring boundaries. The edges of the Saturn rings are also known to be very sharp (Tiscareno 2013). Outside the Solar System, eclipse data reveal the sharpness of the edge of the circumbinary ring around the young star KH 15D (Winn et al. 2006).

<sup>9</sup> Verification of this statement by direct orbit integrations can be tricky because of the formal logarithmic divergence of the acceleration at the edge of the disk, see §5.1.

Our calculations predict that sharp edges result in the divergence of the secular potential of a razor-thin disk, leading to a divergence in the precession rate near these locations (§4.1.3). This outcome, resulting from the nonvanishing boundary terms, was previously pointed out in Silsbee & Rafikov (2015a) for truncated power-law disks, and now we generalize it for other models of eccentric disks with edges. This prediction is nicely confirmed by the direct integration of particle orbits in a particular disk model with sharp edge; see Figure 7 and §4.1.3.

The divergence of  $\dot{\omega}_{\text{sec}}$  (or  $A_d$ ) near the sharp edge of a zero-thickness disk can be traced to the fact that the (in-plane) gravitational acceleration in this region behaves as  $g_d \propto \ln|\Delta r|$ , where  $\Delta r$  is the separation from the edge. Specializing to the case of an axisymmetric disk, one then finds the free precession rate (Fontana & Marzari 2016)

$$\dot{\omega}_{\text{sec}} = -\frac{n}{2ra_K} \frac{dg_d}{dr} \propto (\Delta r)^{-1} \quad (18)$$

(where  $g_K = GM_*/r^2$ ) near the edge at the leading order. The divergent behavior  $\dot{\omega}_{\text{sec}} \propto (\Delta r)^{-1}$  coincides with the scaling of the boundary terms in the expression (§A27) for  $A_d^{\text{edge}}$ , see (2.2). Analogous singularities should arise at any radius in the disk where  $\Sigma_d(a)$  exhibits a discontinuity.

A disk with  $\Sigma_d$  dropping to zero smoothly over a narrow but finite range of  $a$  would not have  $A_d$  diverging there, as the boundary terms ( $A_d^{\text{edge}}$  and  $B_d^{\text{edge}}$ ) vanish for smooth  $\Sigma_d$  profiles. Nevertheless,  $A_d$  still exhibits a nontrivial behavior in this region. This is illustrated in Figure 14 where we plot  $\dot{\omega}_{\text{sec}} = A_d(a)$  (in the top) for rings with several  $\Sigma_d(a)$  profiles<sup>10</sup> of different degree of steepness near the boundaries (shown in the bottom). One can see that near the edge,  $\dot{\omega}_{\text{sec}}$  exhibits very rapid variation, changing from large negative values just inside the edge to large positive values just outside the edge.

This behavior can be understood by noticing that equation (A26) for  $A_d^{\text{bulk}} = A_d$  contains a second derivative of the surface density  $\Sigma_d$  inside the integral. Very close to the boundary, where  $\Sigma_d$  is still high,  $\Sigma_d''(a)$  is large and negative. Due to the logarithmic singularity of  $b_{1/2}^{(0)}(\alpha)$  as  $\alpha \rightarrow 1$ , the radial convolution in equation (A26) enhances the contribution of this region (with large  $\Sigma_d'' < 0$ ) to  $A_d$ , resulting in *rapid retrograde* precession in this part of the disk. On the other hand, just slightly away from the boundary, where  $\Sigma_d$  is low,  $\Sigma_d''(a)$  is large and positive, driving *fast prograde* precession there.

As the sharpness of the  $\Sigma_d$  profile near the boundaries increases, so does the magnitude of  $\Sigma_d''(a)$ . As a result, the amplitude of  $\dot{\omega}_{\text{sec}} = A_d$  on both sides of the edge grows. In the limit of the infinitely sharp  $\Sigma_d$  transition at the edge, the behavior of  $A_d$  becomes singular, with the change of sign at the edge, in agreement with the expectation (18). In our formalism, this is accounted for mathematically by the appearance of the boundary terms (A27), whereas  $\Sigma_d''$  in the integral (A26) remains finite.

In a disk with small but finite vertical thickness  $h \ll a$  the behavior of the coefficients of  $R_d$  would be slightly

<sup>10</sup> The explicit expression for  $\Sigma_d(a)$  is given by the equation (B1).

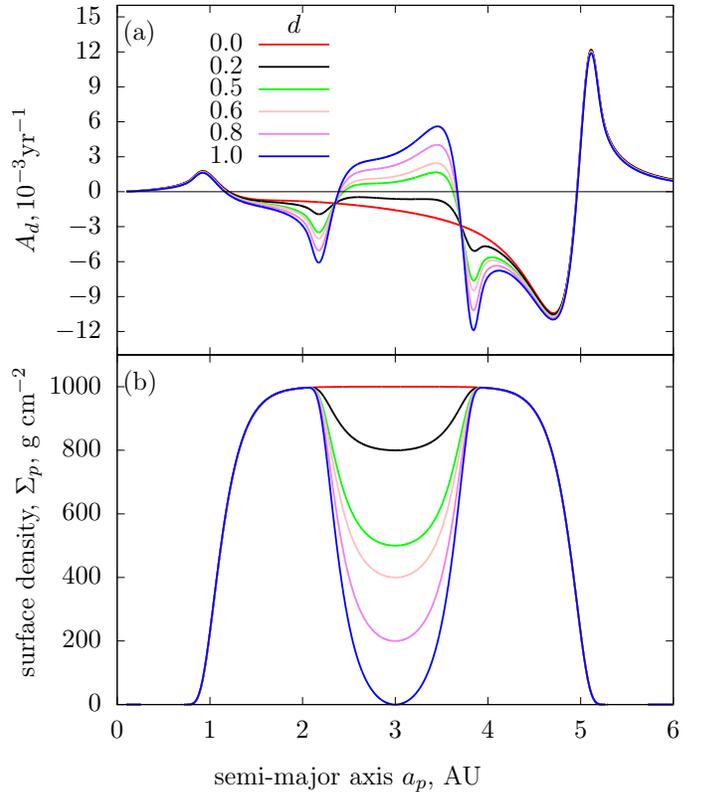


FIG. 15.— Variation of the free eccentricity precession rate  $A_d$  (panel (a)) in a disk with a gap. An underlying profile (B1) of  $\Sigma_d$  is modified by imposing a gap of different relative depth  $d$  (values are shown in panel (a)) according to the prescription (B2), as illustrated in panel (b). The width of the gap  $w = 1.5$  AU is fixed. Note the evolution of the precession rate in the gap from negative to positive as the gap depth is increased.

different. In such a disk, the rise of  $A_d$  and  $B_d$  would saturate at a finite value  $\propto h^{-1}$  as the edge is approached. This transition is easy to understand by setting  $\Delta r \sim h$  in equation (18).

The divergence of  $A_d$  near the sharp edges can have important implications for, e.g., the dynamics of planetary rings. Even though  $\Sigma_d$  remains finite at the edge, our results demonstrate that particle orbits should precess very rapidly and at a rapidly changing rate as their semimajor axes get closer to the edge. This naturally leads to particle orbit crossing resulting in their collisions, helping redistribute angular momentum near the disk edge (Chiang & Culter 2003; Chiang & Goldreich 2000).

## 5.2. Free precession in disks with gaps

Disk gaps present another example of sharp surface density gradients. Gaps could form naturally as a result of gas clearing by the gravitational torques due to massive planets orbiting within the disk.

Ward (1981) has looked at the effect of gaps on the free precession rate of test particles in axisymmetric disks with a power-law profile of  $\Sigma_d$ , finding  $\dot{\omega}_{\text{sec}}$  to be negative within the disk, but positive within the gap. Ward (1981) modeled the gap by setting  $\Sigma_d$  to zero within a range of semimajor axes, which makes his result quite natural. Indeed, a disk with such a gap can be viewed as

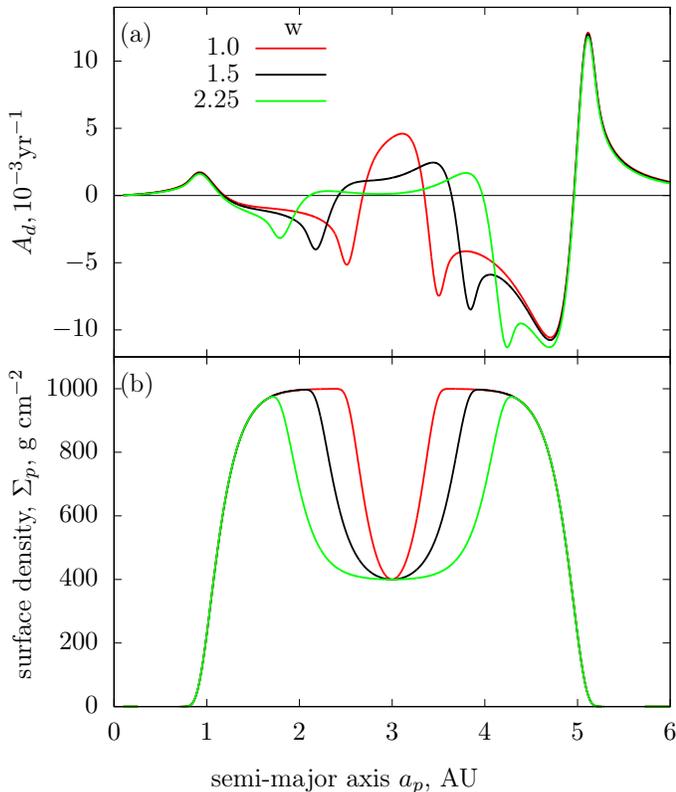


FIG. 16.— Same as Figure 15 but now showing the effect of the gap width  $w$  (indicated in panel (a)) on the behavior of  $A_d$ . The gap depth in equation (B2) is kept fixed at  $d = 0.6$ .

a combination of two disjoint disks with sharp edges. An object orbiting within a gap is thus moving exterior to an inner truncated disk and interior to an outer truncated disk. Our results in §4.1.3, 4.4, 5.1 demonstrate that interaction with each of the disks drives *prograde* free precession of such an object, with the combined effect being simply a linear superposition (also prograde) of the two contributions.

Using our theoretical formalism, we can explore how the results of Ward (1981) change for more realistic (i.e. less sharp) gap profiles. We consider a disk, which is a combination of a wide ring with flat  $\Sigma_d$  given by equation (B1), and a gap of width  $w$  and relative depth  $d$  (changing from  $d = 0$  for no gap to  $d = 1$  for  $\Sigma_d = 0$  in the gap center), with the profile given by equation (B2). In Figure 15, we show how  $A_d = \dot{\omega}_{\text{sec}}$  varies as we change the gap depth. One can see that  $A_d$  inside the gap region, which is negative in the absence of a  $\Sigma_d$  depression, gradually decreases in magnitude, crosses zero, and becomes large and positive as the gap depth is increased. At the edges of the gap,  $A_d$  exhibits a nontrivial structure reminiscent of that seen in Figure 14. If our gap had a more abrupt drop of  $\Sigma_d$  at its edges, we would have converged to the case explored by Ward (1981).

Figure 16 looks at the effect of variation of the gap width, keeping its depth constant. One can see that narrower gaps have a more pronounced effect on free precession in the center of the gap. This is because the edges of such gaps are sharper and also closer to

the orbit of a precessing object. In this case, the intuition developed in §14 again suggests that  $A_d$  should be large and positive, as is observed in Figure 16. To summarize, the effect of a gap on free precession seems to be determined more by the *sharpness* of the  $\Sigma_d$  gradient rather than by the gap width or depth separately.

### 5.3. Free precession in smooth disks

The nontrivial behavior of  $\dot{\omega}_{\text{sec}}$  discussed in §5.1-5.2 is caused by the localized sharp features in  $\Sigma_d(a)$ . However, our results in §4.1.1-4.1.3, confirmed by direct orbit integration, clearly demonstrate that  $A_d$  can also easily change sign within a disk with a *smooth* distribution of  $\Sigma_d$ . Since, according to the solution (9),  $A_d = \dot{\omega}_{\text{sec}}$  is the rate of precession of the free eccentricity vector, this implies that free precession of a test particle can be *both prograde and retrograde*, depending on its location within the disk. The possibility of a change of sign of free precession may appear rather surprising, given that the simple truncated power-law models usually employed to study the gravitational effect of thin disks tend to predict  $\dot{\omega}_{\text{sec}} < 0$ , i.e. *retrograde* free precession (Batygin et al. 2011; Heppenheimer 1980; Ward 1981). Prograde precession is (locally) possible in disks featuring gaps, i.e. sharp drops of  $\Sigma_d$  (Ward 1981; see also §5.2), but the disks explored in §4.1.1-4.1.2 have rather smooth profiles of  $\Sigma_d(a)$ .

On the other hand, already in Silsbee & Rafikov (2015a) it was shown that even the truncated power-law disks can exhibit *prograde* free precession far from the disk edges for certain values of the density slope  $p = -d \ln \Sigma_d / d \ln a$ , e.g. when  $p < 0$  so that  $\Sigma_d$  grows with  $a$ . Given that near the edges (but still within the disk) one expects  $A_d < 0$  (see §5.1), the direction of free precession must change sign at some location inside such a disk. In light of this observation, our current results simply show that the change of sign of free precession is a rather common phenomenon for arbitrary profiles of  $\Sigma_d$ .

Unfortunately, predicting the direction of the free precession (i.e. the sign of  $A_d$ ) at a specific semimajor axis  $a$  in a disk with a given profile  $\Sigma_d(a)$  is not an easy task. Even in a disk without sharp boundaries, when the boundary term  $A_d^{\text{edge}}$  vanishes and  $A_d = A_d^{\text{bulk}}$  is represented by the equation (A26), it is generally not straightforward to predict a priori the sign of this integral term. Indeed, a smooth, continuous  $\Sigma_d(a)$  gradually decaying to zero at finite (e.g. given by equation (15)) or infinite (e.g. profile (16)) boundaries would necessarily have  $\Sigma_d''(a)$  changing sign within the disk. Integration over  $a$  provides a nontrivial, nonlocal mapping between the global behavior of  $\Sigma_d''(a)$  and the value (and sign) of  $A_d$ .

As discussed in §4.1 and 4.2, a change of sign of  $A_d$  inside the disk is also important because it gives rise to very high (formally divergent) values of the test particle eccentricity at radii where  $A_d = 0$ , as long as the disk eccentricity  $e_d$  is nonzero. Previously, a similar effect — a localized singularity of  $e_p$  — was identified in studies of planet formation within stellar binaries, both analytically (Rafikov 2013a,b; Rafikov & Silsbee 2015a,b; Silsbee & Rafikov 2015a,b) and numerically (Meschiari 2014). Its origin could be traced to a secular resonance, caused by the cancellation of prograde precession due to the binary companion and retrograde precession driven

by the disk gravity in presence of the nonzero binary torque (and disk torque, if the disk is eccentric). The importance of these singularities for planet formation in binaries lies in the fact that high values of  $e_p$  lead to very energetic collisions between planetesimals, resulting in their destruction and hampering planetary accretion. In this work, we show that the same mechanism works even without a binary companion — disk torque is always present when  $e_d \neq 0$  and results in divergent  $e_p$  whenever  $A_d \rightarrow 0$ , which naturally happens in our disks.

As shown in Rafikov & Silsbee (2015b) and Silsbee & Rafikov (2015b), an opposite effect is also possible in disks in binaries — at certain locations, the eccentricities of test particles affected by the combined potential of a companion and an eccentric disk become very small, as a result of the cancellation of the corresponding torques<sup>11</sup>. Again, in our case, this happens even without a binary companion — at the locations where  $B_d = 0$ , one naturally finds  $e_p \rightarrow 0$ , as a result of the cancellation of torques arising from different parts of the same disk. This can be seen, e.g., at  $\approx 0.9$  AU in Figures 5, 9 and at  $\approx 4$  AU in Figure 11. At these locations, the relative velocities of colliding objects naturally become very small, promoting their agglomeration (rather than fragmentation) and growth.

## 6. SELF-CONSISTENT MODELS OF SELF-GRAVITATING, RIGIDLY PRECESSING DISKS

We now use our results to assess the possibility of constructing self-consistent models of long-lived, self-gravitating eccentric disks orbiting massive central objects. Such models could describe, for example, the eccentric nuclear stellar disks around supermassive black holes observed in the centers of some galaxies.

We will assume that the surface density distribution in the disk is given by equation (2), which essentially implies that for each semimajor axis, there is a single, unique value of the disk eccentricity  $e_d(a)$ , and that eccentric orbits of particles at all semimajor axes have the same orientation  $\varpi_d$ . This orientation cannot be fixed in time, as the disk's own non-Newtonian potential causes the orbits of individual particles to precess. In other words,  $\varpi_d = \varpi_d(t)$ . Given the distribution of the surface density  $\Sigma_d(a)$  at the pericenter, the question we ask is whether one can determine the profile of  $e_d(a)$  that the disk must have to precess coherently as a solid body at a constant rate  $\dot{\varpi}_d$  (i.e.  $\varpi_d(t) = \dot{\varpi}_d t$ ). This arrangement, obviously, requires  $\dot{\varpi}_d$  to be independent of  $a$ , since otherwise, differential precession would lead to disk twisting (apsidal misalignment of different parts of the disk), destroying its coherence.

Introducing for convenience the complex eccentricity  $E_p = k_p + ih_p = e_p e^{i\varpi_p}$  we can combine equations (7) into a single evolution equation for  $E_p$ :

$$-i\dot{E}_p = A_d E_p + B_d e^{i\varpi_d}. \quad (19)$$

This equation is valid for any object, including the particles or fluid elements comprising the disk and contributing to its potential. Provided that solid-body precession is the only secular effect of the disk self-gravity, i.e. that the disk remains stationary in the

frame precessing at the rate  $\varpi_d$ , we look for solutions with  $\dot{e}_p = 0$  (i.e.  $e_p(a, t) = e_p(a)$ ) and  $\varpi_p = \varpi_d = \dot{\varpi}_d t$ . Also, by our assumption, at each point the disk eccentricity  $e_d$  is the same as the eccentricity of its constituent particles passing through this point, meaning that we need to identify  $e_p = e_d$ . Plugging the ansatz  $E_p = e_d(a) e^{i\varpi_d t}$  into the equation (19), one arrives at the following master equation:

$$[\dot{\varpi}_d - A_d(\Sigma_d, a)] e_d(a) = B_d(\Sigma_d, e_d, a). \quad (20)$$

This equation represents a self-consistent mathematical framework for determining the radial profile of  $e_d$  that an eccentric disk needs to have to be able to precess as a solid body (without changing its shape) under the action of its own self-gravity. The precession rate  $\dot{\varpi}_d$  plays the role of an *eigenvalue* of the problem. Equations (A25)-(A29) provide explicit dependencies of  $A_d$  and  $B_d$  on  $\Sigma_d$ ,  $e_d$ , and  $a$ . The dependence is such that (20) is an *integral equation* for  $e_d$ . It is linear in  $e_d$  and is essentially a Fredholm equation of the second type. Solving this integral equation, we obtain a set of eigenvalues (precession rates  $\dot{\varpi}_d$ ), as well as the corresponding eigenfunctions (radial profiles of  $e_d$ ; the normalization of  $e_d$  remains unconstrained because of the linear nature of equation (20)).

This calculation uses the radial distribution of  $\Sigma_d$  as an input. As mentioned in §2, when  $e_d(a)$  does not change in the course of evolution (or whenever  $e_d \ll 1$ ), the radial profile of the surface density at periastron  $\Sigma_d(a)$  remains fixed in the course of secular evolution.

We defer the detailed exploration of the equation (20) for disks with different  $\Sigma_d$  profiles to future work. We will simply note here that some of our findings — divergent behavior of  $A_d$  and  $B_d$  near the sharp disk edges, the changes of signs of  $A_d$  and  $B_d$  inside the disk, etc. — make finding the solutions of this equation rather nontrivial.

## 7. DISCUSSION

Our results allow one to efficiently compute the effect of the gravity of an eccentric disk on the secular evolution of astrophysical objects coplanar with the disk. This work provides a natural generalization of the earlier calculation of Silsbee & Rafikov (2015a), in which the secular potential was computed for eccentric disks with  $\Sigma_d(a)$  and  $e_d(a)$  given by power laws of  $a$  only. Even prior to that, Heppenheimer (1980) and Ward (1981) derived the disturbing function for a particular case of *axisymmetric* power-law disks. Our present results extend these calculations for *arbitrary* behaviors of  $\Sigma_d(a)$  and  $e_d(a)$ , allowing a much broader range of applications of our framework.

The accuracy with which our lowest-order theory works depends on both the behavior of the disk eccentricity and the particle eccentricity. The results of §4.3 demonstrate that our secular theory becomes inaccurate when  $e_d \gtrsim 0.2$  or so; the actual value of  $e_d$  at which this happens depends on both the location in the disk (see Figure 10) and the radial profile of  $e_d$  (see Figure 11). Moreover, our results clearly show that even for nearly circular disks, the behavior of  $e_p$  can be singular at certain locations (at least for  $\Sigma_d$  profiles considered in this work), resulting in very high  $e_p \sim 1$  even at semimajor axes where

<sup>11</sup> This often requires a particular relative orientation of the disk and binary apsidal lines (Rafikov & Silsbee 2015b).

$e_d(a) \ll 1$ . As evidenced by Figure 8, this naturally leads to deviations from our theory at these locations, even for low- $e_d$  disks.

Real astrophysical disks have rather different values of  $e_d$ . For example, the stellar disk in M31 has a substantial eccentricity,  $e_d \sim 0.5$  (Brown & Magorrian 2013; Peiris & Tremaine 2003; Tremaine 1995). Our secular theory is unlikely to provide a good description of the secular dynamics in this system, as even its topology of the phase space should look very different from that corresponding to the disturbing function (3). A higher-order extension of our approach, such as that presented recently in Sefilian & Touma (2018) to generalize the Silsbee & Rafikov (2015a) calculations to fourth order in  $e_p$ , may provide a better tool for studying disks with nonnegligible  $e_d$ .

On the other hand, disk eccentricity can be low enough in gaseous protoplanetary disks in binary stellar systems. Simulations demonstrate that under certain circumstances (moderately high binary eccentricity), a disk orbiting one of the binary components can have rather low eccentricity, at the level of several percent (Marzari et al. 2009; Müller & Kley 2012; Regály et al. 2011). The same is true for circumbinary protoplanetary disks on AU scales (Meschiari 2014). In such systems, our theory should be well suited for describing both the effect of the disk gravity on planetesimal motion and planet formation in binary systems (Rafikov & Silsbee 2015a,b), as well as for studying the self-consistent dynamics of the gaseous component of the disk (Ogilvie 2001) driven by its self-gravity, pressure forces, etc.

Eccentric planetary rings typically have  $e_d \sim 10^{-3} - 10^{-2}$  (Elliot & Nicholson 1984), which is also low enough for our theory to work well in describing the effect of the ring gravity on the secular motion of the adjacent objects (including the ring particles themselves; see §6).

However, when applying our framework to real planetary rings, a word of caution is in order. One of the underlying assumptions used in deriving the linearized equation (A2) is that  $\zeta e_d = de_d/d \ln a \ll 1$ . At the same time the  $\epsilon$  ring of Uranus exhibits a change of eccentricity  $\Delta e_d \approx 7.11 \times 10^{-4}$  over the width  $\Delta a \approx 58.1$  km of the ring with a mean semimajor axis  $a = 51,149$  km (French et al. 1991). We can evaluate  $\zeta e_d \approx a \Delta e_d / \Delta a \approx 0.6$ , which is certainly not small compared to unity. This means that  $\Sigma$  exhibits strong variation along the eccentric orbits of ring particles. As a result, both the expansion (A2) and our resultant framework (3), (A25)-(A29) may become inaccurate when applied to rings with  $\zeta e_d \sim 1$ . A similar issue was previously discussed in §4.3.2.

The cost involved in calculating secular evolution according to our approach is relatively low — it requires computation of only one-dimensional integrals involving  $\Sigma_d(a)$  and  $e_d(a)$  (to obtain coefficients  $A_d$  and  $B_d$ ). This is to be contrasted with the direct approach to computing secular potential, embodied by equation (A1), in which one first needs to carry out the two-dimensional integration over the full disk to obtain the potential at every point and then one additional integration to average it over the particle trajectory. Our procedure is clearly less numerically intensive and reproduces direct calculations very well in the low- $e$  limit. This allowed us to use it for exploring different characteristics of secular

motion in the disk potential, which we did in §5.1-6.

The secular dynamics of self-gravitating disks is often explored by modeling them as collections of narrow adjacent rings coupled via the *softened* secular gravitational potential (Hahn 2003; Touma et al. 2009; Tremaine 2001). In this approach, the secular potential is approximated by the modified version of the classical Laplace-Lagrange disturbing function (Murray & Dermott 1999), which is regularized via softening to avoid the singularity that arises when the semimajor axes of the rings overlap. This procedure inevitably introduces an ad hoc *softening parameter* into the calculation, which inevitably leads to ambiguity of the results, since a physical justification for a particular choice of softening is not obvious.

Our approach, ascending to the framework developed in Heppenheimer (1980), Ward (1981), and Silsbee & Rafikov (2015a), does not suffer from this ambiguity. Even though the integrand of the expressions (A26), (A28) for  $A_d$  and  $B_d$  contains Laplace coefficients  $b_{1/2}^{(0)}(\alpha)$  and  $b_{1/2}^{(1)}(\alpha)$ , which diverge as  $\alpha \rightarrow 1$  (i.e. for particles with semimajor axes inside the radial extent of the disk), their singularity is weak (logarithmic in  $\alpha$ ). As a result, the integrals in equations (A26), (A28) are fully convergent *without* introducing any ad hoc softening. This makes our calculation of  $R_d$  more robust and self-consistent compared to some other treatments. The only possible divergences that remain in our case may arise at the sharp edges of the disk and are truly physical in their nature (see §5.1).

Our calculation of the disturbing function explicitly assumes the disk to have zero thickness and to be coplanar with the particle orbit. It is easy to show that our results would not change qualitatively if the disk were to have a small vertical thickness  $h$ , as long as  $h$  is small compared to the semimajor axis  $a$  in the disk. Quantitative corrections to our results due to the nonzero  $h$  should be of order  $O(h/a)$  at most. The only substantial difference will arise close to the edges of the disks with sharply truncated  $\Sigma_d$ , see §5.1.

We also explicitly assumed the surface density of the disk  $\Sigma$  to have a specific form (2), valid when the mass elements comprising the disk move around the central mass on eccentric orbits, which are apsidally aligned and have a unique value of eccentricity for a given semimajor axis. This assumption is natural for fluid disks, in which orbit crossing is impossible in the absence of shocks, as well as for highly collisional planetary rings. However, particulate disks of low optical depth (e.g. nuclear stellar disks) may have a more complicated form of  $\Sigma$  because orbit crossings are possible in such systems. We can naturally extend our approach to treat such cases by using the additive nature of gravity: if a more general disk structure can be represented as a superposition of multiple (sub)disks each with  $\Sigma$  in the form (2) but different  $\Sigma_d(a)$ ,  $e_d(a)$  and apsidal orientation, then the resultant disturbing function  $R_d$  will be a sum of individual contributions in the form (3) produced by each of the subdisks.

## 8. SUMMARY

We explored the secular effect of a massive, razor-thin, eccentric, apsidally aligned disk on the motion of

coplanar objects in the combined potential of such a disk (considered as a perturbation) and a central point mass. This problem is of great importance for many astrophysical disks (both gaseous and particulate), ranging from planetary rings to nuclear stellar disks in centers of galaxies harboring supermassive black holes. Our main findings are briefly summarized below.

- We developed a general analytical framework for computing the secular disturbing function due to the gravity of an aligned eccentric disk. This disturbing function has the conventional Laplace-Lagrange form with coefficients that contain one-dimensional integrals over the semimajor axis only and do not involve softening of any form. It is valid for arbitrary radial profiles of the disk surface density  $\Sigma_d$  and eccentricity  $e_d$ , and works for coplanar objects orbiting both inside and outside the disk (for disks with edges).
- We verified the accuracy of our analytical calculation using direct orbit integrations, finding excellent agreement in the limit of low eccentricity (both disk and particle) for all radial profiles of  $\Sigma_d$  and  $e_d$  that we considered. Our framework is accurate at the  $\lesssim 10\%$  level for disk eccentricities  $e_d \lesssim 0.1 - 0.2$ . However, these figures strongly depend on both the location in the disk and the disk model, e.g. eccentricity profile  $e_d(a)$ .
- Our calculations demonstrate that free precession of particle orbits in the potential of a smooth eccentric disk can naturally change from prograde to retrograde, and vice versa. Thus, the retrograde free precession previously found for certain types of disks (with a power-law dependence of density

on radius) does not hold in general. At the locations where the free precession rate changes sign, forced particle eccentricities reach very high values (formally diverge).

- Sharp features in the disk surface density distribution, such as disk edges or gaps, inevitably change the sense of free precession (e.g. prograde in the gaps, retrograde in the disk). In disks with sharp edges (where  $\Sigma_d$  discontinuously drops to zero), the free precession rate formally diverges at the edge.
- Using our results, we formulated a general framework for computing the eccentricity profile of a disk (with a prescribed surface density profile  $\Sigma_d(a)$ ), which is required for it to be precessing rigidly due to its self-gravity alone, while maintaining stationary structure in the frame of precession.

In the future, we plan to extend our approach for understanding the secular dynamics of *twisted* (i.e. *apsidally misaligned*) eccentric disks (Davydenkova & Rafikov, in prep.), as well as for exploring the inclination dynamics in the case of a non-coplanar (warped) disk. Exploration of the modal solutions (i.e. long-lived, rigidly precessing disk models discussed in §6) based on our secular framework is another logical application of our results.

We are grateful to Ryan Miranda, Antranik Sefilian, Kedron Silsbee and Scott Tremaine for useful discussions. Financial support for this study has been provided by the NSF via grant AST-1409524 and NASA via grant 15-XRP15-2-0139.

## REFERENCES

- Bacon, R., Emsellem, E., Combes, F., et al. 2001, *A&A*, 371, 409  
 Bartko, H., Martins, F., Fritz, T. K., et al. 2009, *ApJ*, 697, 1741  
 Batygin, K., Morbidelli, A., & Tsiganis, K. 2011, *A&A*, 533, A7  
 Bender, R., Kormendy, J., Bower, G., et al. 2005, *ApJ*, 631, 280  
 Brown, C. K., & Magorrian, J. 2013, *MNRAS*, 431, 80  
 Chambers, J. E. 1999, *MNRAS*, 304, 793  
 Chauvin, G., Beust, H., Lagrange, A.-M., & Eggenberger, A. 2011, *A&A*, 528, A8  
 Chiang, E. I., & Culter, C. J. 2003, *ApJ*, 599, 675  
 Chiang, E. I., & Goldreich, P. 2000, *ApJ*, 540, 1084  
 Denny, E., Clemens, J. C., Dunlap, B. H., et al. 2018, *ArXiv e-prints*, arXiv:1801.05443  
 Doyle, L. R., Carter, J. A., Fabrycky, D. C., et al. 2011, *Science*, 333, 1602  
 Elliot, J. L., & Nicholson, P. D. 1984, in *IAU Colloq. 75: Planetary Rings*, ed. R. Greenberg & A. Brahic, 25–72  
 Fontana, A., & Marzari, F. 2016, *A&A*, 589, A133  
 French, R. G., Nicholson, P. D., Porco, C. C., & Marouf, E. A. 1991, *Dynamics and structure of the Uranian rings*, ed. J. T. Bergstralh, E. D. Miner, & M. S. Matthews, 327–409  
 Gänsicke, B. T., Marsh, T. R., Southworth, J., & Rebassa-Mansergas, A. 2006, *Science*, 314, 1908  
 Georgakarakos, N. 2002, *MNRAS*, 337, 559  
 Goldreich, P., & Tremaine, S. 1979, *AJ*, 84, 1638  
 Graps, A. L., Showalter, M. R., Lissauer, J. J., & Kary, D. M. 1995, *AJ*, 109, 2262  
 Hahn, J. M. 2003, *ApJ*, 595, 531  
 Hatzes, A. P., Cochran, W. D., Endl, M., et al. 2003, *ApJ*, 599, 1383  
 Heppenheimer, T. A. 1980, *Icarus*, 41, 76  
 Horne, K. 1984, *Nature*, 312, 348  
 Kley, W., Papaloizou, J. C. B., & Ogilvie, G. I. 2008, *A&A*, 487, 671  
 Levin, Y., & Beloborodov, A. M. 2003, *ApJ*, 590, L33  
 Lubow, S. H. 1991, *ApJ*, 381, 259  
 Marsh, T. R., & Horne, K. 1988, *MNRAS*, 235, 269  
 Marzari, F., Scholl, H., Thébault, P., & Baruteau, C. 2009, *A&A*, 508, 1493  
 Meschiari, S. 2014, *ApJ*, 790, 41  
 Miranda, R., Muñoz, D. J., & Lai, D. 2017, *MNRAS*, 466, 1170  
 Miranda, R., & Rafikov, R. R. 2018, *ArXiv e-prints*, arXiv:1802.00464  
 Mosqueira, I., & Estrada, P. R. 2002, *Icarus*, 158, 545  
 Müller, T. W. A., & Kley, W. 2012, *A&A*, 539, A18  
 Murray, C. D., & Dermott, S. F. 1999, *Solar system dynamics*  
 Nayakshin, S., Dehnen, W., Cuadra, J., & Genzel, R. 2006, *MNRAS*, 366, 1410  
 Ogilvie, G. I. 2001, *MNRAS*, 325, 231  
 Pan, M., & Wu, Y. 2016, *ApJ*, 821, 18  
 Peiris, H. V., & Tremaine, S. 2003, *ApJ*, 599, 237  
 Porco, C. C. 1990, *Advances in Space Research*, 10, 221  
 Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical recipes in FORTRAN. The art of scientific computing*  
 Rafikov, R. R. 2013a, *ApJ*, 764, L16  
 —. 2013b, *ApJ*, 765, L8  
 Rafikov, R. R., & Silsbee, K. 2015a, *ApJ*, 798, 69  
 —. 2015b, *ApJ*, 798, 70  
 Regály, Z., Sándor, Z., Dullemond, C. P., & Kiss, L. L. 2011, *A&A*, 528, A93  
 Salow, R. M., & Statler, T. S. 2001, *ApJ*, 551, L49  
 —. 2004, *ApJ*, 611, 245  
 Sambhus, N., & Sridhar, S. 2002, *A&A*, 388, 766

Sefilian, A. A., & Touma, J. R. 2018, ArXiv e-prints, arXiv:1804.06859  
 Silsbee, K., & Rafikov, R. R. 2015a, ApJ, 808, 58  
 —. 2015b, ApJ, 798, 71  
 Statler, T. S. 1999, ApJ, 524, L87  
 —. 2001, AJ, 122, 2257  
 Tiscareno, M. S. 2013, Planetary Rings, ed. T. D. Oswalt, L. M. French, & P. Kalas, 309  
 Touma, J. R., Tremaine, S., & Kazandjian, M. V. 2009, MNRAS, 394, 1085  
 Tremaine, S. 1995, AJ, 110, 628

—. 2001, AJ, 121, 1776  
 Ward, W. R. 1981, Icarus, 47, 234  
 Welsh, W. F., Orosz, J. A., Carter, J. A., & Fabrycky, D. C. 2014, in IAU Symposium, Vol. 293, Formation, Detection, and Characterization of Extrasolar Habitable Planets, ed. N. Haghighipour, 125–132  
 Winn, J. N., Hamilton, C. M., Herbst, W. J., et al. 2006, ApJ, 644, 510

## APPENDIX

## CALCULATION OF THE SECULAR POTENTIAL OF THE DISK

In this section, we present a calculation of the disturbing function  $R_d$  due to an eccentric disk with a surface density in the form (2) in the low-eccentricity limit. We do not assume any specific forms for the functions  $\Sigma_d(a)$  and disk eccentricity  $e_d(a)$  apart from requiring them to be twice differentiable. Following the recipe presented in Silsbee & Rafikov (2015b), we first write the disturbing function  $R_d$  due to the disk  $\mathbb{D}$  as

$$R_d = R(\mathbb{D}) = -\langle \Phi_d(r_d, \varphi_d) \rangle, \quad \Phi_d(r_d, \varphi_d) = -G \int_{\mathbb{D}} \frac{\Sigma(r_d, \varphi_d) r_d dr_d d\theta}{\sqrt{r_p^2 + r_d^2 - 2r_p r_d \cos \theta}}, \quad (\text{A1})$$

where brackets  $\langle \dots \rangle$  represent time averaging over planetesimal orbital motion,  $\mathbf{r}_p$  is the instantaneous planetesimal radius vector ( $r_p = |\mathbf{r}_p|$ , making an angle  $\varphi_p$  with the planetesimal apsidal line),  $\mathbf{r}_d$  is the radius vector of a surface element of the disk ( $r_d = |\mathbf{r}_d|$ , making an angle  $\varphi_d$  with the disk apsidal line), and  $\theta = \varphi_d + \varpi_d - \varphi_p - \varpi_p$  is the angle between the two; see Figure 1.

Note that the indirect potential vanishes identically for a disk composed of mass elements moving on purely Keplerian orbits around the central mass. This is easy to understand based on Kepler’s second law, which in our case means that the mass elements on opposite sides of a given elliptical trajectory (with respect to its focus) exert equal and opposite forces on the central mass. As a result, mass elements with orbits within a given semimajor axis interval  $da$  exert no net force on the central mass and do not cause its reflex motion, meaning that the indirect potential is zero.

We need to evaluate the integral (A1) to second order in eccentricity, i.e. keeping the terms  $O(e_p^2)$  and  $O(e_p e_d)$ . At the same time, we do not need to keep the terms not involving  $e_p$  (e.g.  $O(e_d)$  or  $O(e_d^2)$ ), as those will vanish upon substitution into the Lagrange equations (Murray & Dermott 1999). Nor do we need to keep terms of higher order in  $e_d$  than those listed above. Given that some of the variables entering the expression (2) are functions of  $a_p$  and not  $r_p$ , we need to express them through  $r_p$  via  $a_p \approx r_p(1 + e_d(r_p) \cos \varphi_d) + O(e_d^2)$ . With this in mind, we can expand the expression (2) to lowest order in  $e_d \ll 1$  and  $de_d/d \ln r_d \ll 1$  as (Silsbee & Rafikov 2015b)

$$\Sigma(r_d, \varphi_d) \approx \Sigma_d(r_d) + r_d \frac{d}{dr_d} [\Sigma_d(r_d) e_d(r_d)] \cos \varphi_d - \Sigma_d(r_d) e_d(r_d). \quad (\text{A2})$$

As will become obvious later, the integral (A1) over the last term in equation (A2) does not contribute to  $R_d$  at the required level of accuracy, so we drop it from the consideration from now on.

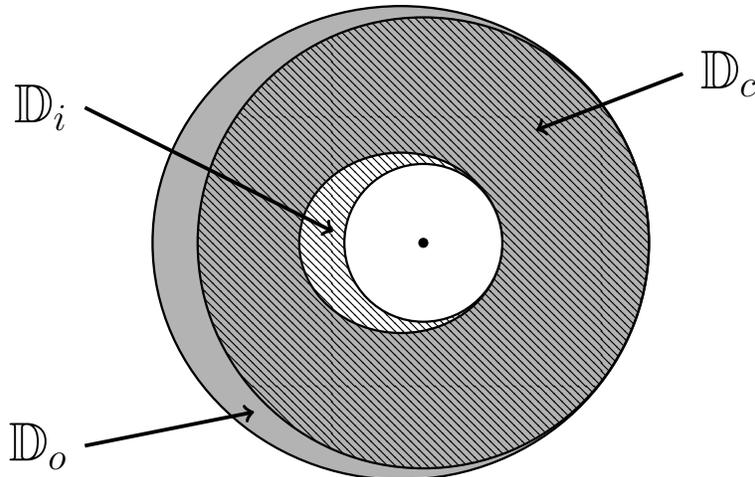


FIG. A17.— Decomposition of the disk into three distinct regions ( $\mathbb{D}_c$ ,  $\mathbb{D}_i$ ,  $\mathbb{D}_o$ ) over which potential is integrated. The gray area corresponds to the integration area, that is, the physical disk  $\mathbb{D} = \mathbb{D}_o + \mathbb{D}_c - \mathbb{D}_i$ . The hatched region  $\mathbb{D}_c$  is circular. See text for details.

Following (Silsbee & Rafikov 2015b), we split the expression (A1) into three integrals over the regions  $\mathbb{D}_c$ ,  $\mathbb{D}_i$  and  $\mathbb{D}_o$ , as shown in Figure A17. We define  $\mathbb{D}_c$  to be a circular region with an inner radius equal to the periastron distance of the inner edge of the disk,  $r_{d,\text{in}} = a_1(1 - e_d(a_1))$ , and the outer radius equal to the periastron distance of the outer edge of the disk  $r_{d,\text{out}} = a_2(1 - e_d(a_2))$ . Region  $\mathbb{D}_i$  is the *crescent* region confined between the inner circle of  $\mathbb{D}_c$  and the inner ellipse of the integration region  $\mathbb{D}$ ;  $\mathbb{D}_o$  is the *crescent* region confined between the outer circle of  $\mathbb{D}_c$  and the outer ellipse of the integration region  $\mathbb{D}$ . In other words,

$$\mathbb{D} = \mathbb{D}_c - \mathbb{D}_i + \mathbb{D}_o, \quad \text{and} \quad R(\mathbb{D}) = R(\mathbb{D}_c) - R(\mathbb{D}_i) + R(\mathbb{D}_o). \quad (\text{A3})$$

We now consider the three contributions to  $R(\mathbb{D})$  separately.

*Calculation of the part of the disturbing function due to the circular region  $\mathbb{D}_c$*

We start by computing the integral over the circular annulus  $\mathbb{D}_c$ . When using the approximation (A2), it is natural to divide  $R(\mathbb{D}_c)$  into two parts: one containing  $\cos \varphi_d$  (we call it  $I_1$ ) and the one independent of  $\varphi_d$  ( $I_2$ ). We evaluate the  $\varphi_d$ -independent contribution first:

$$I_1 = \left\langle G \int_{r_{d,\text{in}}}^{r_{d,\text{out}}} dr_d \Sigma_d(r_d) r_d \int_0^{2\pi} \frac{d\theta}{\sqrt{r_p^2 + r_d^2 - 2r_p r_d \cos \theta}} \right\rangle, \quad (\text{A4})$$

This integral has a singularity at  $r_p = r_d$ , so we break it into two parts,  $r_p < r_d$  and  $r_p > r_d$ , while also rewriting the inner integral as a Laplace coefficient  $b_{1/2}^{(0)}$ , see definition (6):

$$I_1 = \pi G \left\langle \int_{r_{d,\text{in}}}^{r_p} \Sigma_d(r_d) \frac{r_d}{r_p} b_{1/2}^{(0)} \left( \frac{r_d}{r_p} \right) dr_d + \int_{r_p}^{r_{d,\text{out}}} \Sigma_d(r_d) b_{1/2}^{(0)} \left( \frac{r_p}{r_d} \right) dr_d \right\rangle. \quad (\text{A5})$$

Now let us make a change of variables in both integrals so as to get rid of the time dependence in the Laplace coefficients, entering via  $r_p$ . In the first integral, we set  $\alpha = r_d/r_p < 1$ , while in the second, we use  $\alpha = r_p/r_d < 1$ . We will also approximate  $r_{d,\text{in}} \approx a_1$  and  $r_{d,\text{out}} \approx a_2$ , which, as we will see, introduces error of negligible order. Now we have

$$I_1 = I_1^{(0)} + I_1^{(1)} = \pi G \left\langle \int_{a_1/r_p}^1 r_p \Sigma_d(\alpha r_p) \alpha b_{1/2}^{(0)}(\alpha) d\alpha + \int_{r_p/a_2}^1 r_p \Sigma_d(r_p/\alpha) \alpha^{-2} b_{1/2}^{(0)}(\alpha) d\alpha \right\rangle. \quad (\text{A6})$$

Our goal is to get rid of triangular brackets, that is to time average the integrals. The time dependence is hidden in  $r_p$  and has the form  $r_p(t) = a_p(1 - e_p \cos E(t))$  where  $e_p \ll 1$  and  $E$  is the eccentric anomaly of the particle orbit. Since we are only interested in terms which are no more than quadratic in eccentricity (of both particle and the disk), we next expand the integrals in a Taylor series over  $r_p - a_p = -a_p e_p \cos E(t)$  up to second order and average over the planetesimal orbit using the relations  $\langle \cos E \rangle = -e_p/2$ ,  $\langle \cos^2 E \rangle = 1/2$ . Note that the dependence on  $r_p$  is also present in the limits of integration, which, upon this Taylor expansion, give rise to the *nonintegral* boundary terms. We omit the tedious process of writing out the Taylor expansion and subsequent averaging and present just the result:

$$\begin{aligned} I_1^{(0)} &\approx \pi G a_p \frac{e_p^2}{2} \int_{a_1/a_p}^1 \left[ \frac{d}{dr_p} (\alpha r_p \Sigma_d(\alpha r_p)) + \frac{a_p}{2} \frac{d^2}{dr_p^2} (\alpha r_p \Sigma_d(\alpha r_p)) \right] \Big|_{r_p=a_p} b_{1/2}^{(0)}(\alpha) d\alpha + \\ &+ \pi G \frac{e_p^2}{4} \left[ -\frac{a_1^3}{a_p^2} \Sigma_d(a_1) b_{1/2}^{(0)'} \left( \frac{a_1}{a_p} \right) + \frac{a_1^3}{a_p} \Sigma_d'(a_1) b_{1/2}^{(0)} \left( \frac{a_1}{a_p} \right) + \frac{a_1^2}{a_p} \Sigma_d(a_1) b_{1/2}^{(0)} \left( \frac{a_1}{a_p} \right) \right], \end{aligned} \quad (\text{A7})$$

where we dropped the terms independent of  $e_p$  — such terms vanish when substituted into the Lagrange evolution equations. For  $I_1^{(1)}$  we similarly have

$$\begin{aligned} I_1^{(1)} &\approx \pi G a_p \frac{e_p^2}{2} \int_{a_p/a_2}^1 \left[ \frac{1}{r_p} \frac{d}{dr_p} (r_p \Sigma_d(\alpha^{-1} r_p)) + \frac{1}{2} \frac{d^2}{dr_p^2} (r_p \Sigma_d(\alpha^{-1} r_p)) \right] \Big|_{r_p=a_p} \alpha^{-2} b_{1/2}^{(0)}(\alpha) d\alpha \\ &- \pi G \frac{e_p^2}{4} \left[ 2a_2 \Sigma_d(a_2) b_{1/2}^{(0)} \left( \frac{a_p}{a_2} \right) + a_p \Sigma_d(a_2) b_{1/2}^{(0)'} \left( \frac{a_p}{a_2} \right) + a_2^2 \Sigma_d'(a_2) b_{1/2}^{(0)} \left( \frac{a_p}{a_2} \right) \right]. \end{aligned} \quad (\text{A8})$$

Now we turn to calculation of the integral contribution  $I_2$  containing  $\varphi_d = \theta + v$ , where we defined  $v = \varphi_p + \varpi_p - \varpi_d$ . Let us introduce a new auxiliary function,  $Q(r_d) = d[e(r_d)\Sigma_d(r_d)]/dr_d$ . Then,

$$I_2 = G \left\langle \int_{r_{d,\text{in}}}^{r_{d,\text{out}}} Q(r_d) r_d^2 \int_0^{2\pi} \frac{\cos \theta \cos v d\theta}{\sqrt{r_p^2 + r_d^2 - 2r_p r_d \cos \theta}} dr_d \right\rangle, \quad (\text{A9})$$

where we expanded  $\cos(\theta + v)$  and took into account that the term proportional to  $\sin\theta$  evaluates to zero.

Next, we divide the integral over  $dr_d$  into two parts, one for  $r_d < r_p$  and another for  $r_d > r_p$ , just like it was done for  $I_1$ . We again approximate  $r_{d,\text{in}} \approx a_1$  and  $r_{d,\text{out}} \approx a_2$ , as the linear corrections to these expressions would lead to a contribution, which is third order in eccentricity. We then, again, make a change of variables in both resulting integrals, setting  $\alpha = r_d/r_p < 1$  in the first one and  $\alpha = r_p/r_d < 1$  in the second. As a result, using the definition of the Laplace coefficient  $b_{1/2}^{(1)}$ , we obtain the following expression:

$$I_2 = I_2^{(0)} + I_2^{(1)} = \pi G \left\langle \int_{a_1/r_p}^1 \cos v r_p^2 Q(\alpha r_p) \alpha^2 b_{1/2}^{(1)}(\alpha) d\alpha + \int_{r_p/a_2}^1 \cos v r_p^2 Q(\alpha^{-1} r_p) \alpha^{-3} b_{1/2}^{(1)}(\alpha) d\alpha \right\rangle. \quad (\text{A10})$$

There are two main differences in our subsequent expansion over  $r_p - a_p$  compared to the previous case. First,  $Q(r_d) \sim O(e_d)$  and, thus, in the expansion, we only need to retain terms up to first order in  $e_p$ . Second, now we have the time dependence of the integrand also through  $\cos v(t) = \cos(\Delta\varpi + \varphi_p(t)) = \cos(\Delta\varpi + E(t) + e_p \sin E(t))$ , where we defined  $\Delta\varpi = \varpi_p - \varpi_d$  and used the relation  $\varphi_p = E + e_p \sin E$ . The easiest way to deal with the calculation is to consider  $r_p$  as a function of  $e_p$ ,  $r_p = a_p(1 - e_p \cos E)$ , and to expand the integrals directly over  $e_p$ . We start with  $I_2^{(0)}$ :

$$I_2^{(0)} = \pi G \left\langle \int_{a_1/(a_p(1-e_p \cos E))}^1 \cos(\Delta\varpi + E + e_p \sin E) r_p^2(e_p, E) Q(\alpha r_p(e_p, E)) \alpha^2 b_{1/2}^{(1)}(\alpha) d\alpha \right\rangle. \quad (\text{A11})$$

We need the zeroth and first-order (in  $e_p \ll 1$ ) terms of the Taylor expansion of this expression at  $e_p = 0$ . We will omit the technicalities of writing out the expansion and present the result:

$$I_2^{(0)} \approx \pi G \int_{a_1/a_p}^1 a_p^2 Q(\alpha a_p) \alpha^2 b_{1/2}^{(1)}(\alpha) d\alpha \langle \cos(\Delta\varpi + E) \rangle - \pi G e_p \frac{a_1^3}{a_p} b_{1/2}^{(1)}\left(\frac{a_1}{a_p}\right) Q(a_1) \langle \cos(\Delta\varpi + E) \cos(E) \rangle \quad (\text{A12})$$

$$- \pi G e_p \int_{a_1/a_p}^1 \left( a_p^2 Q(\alpha a_p) \langle \sin(\Delta\varpi + E) \sin(E) \rangle + a_p \frac{d}{dr_p} (r_p^2 Q(\alpha r_p)) \Big|_{r_p=a_p} \langle \cos(\Delta\varpi + E) \cos(E) \rangle \right) \alpha^2 b_{1/2}^{(1)}(\alpha) d\alpha.$$

Using the averages of the trigonometric functions of  $E$ ,

$$\langle \cos(\Delta\varpi + E) \rangle = -\frac{e_p}{2} \cos \Delta\varpi, \quad \langle \sin(\Delta\varpi + E) \sin(E) \rangle = \langle \cos(\Delta\varpi + E) \cos(E) \rangle = \frac{1}{2} \cos \Delta\varpi, \quad (\text{A13})$$

based on  $\langle \sin E \rangle = 0$ ,  $\langle \sin E \cos E \rangle = 0$ , and  $\langle \sin^2 E \rangle = 1/2$ , the final result for  $I_2^{(0)}$  becomes

$$I_2^{(0)} \approx -\pi G \frac{e_p}{2} \cos \Delta\varpi \left[ \int_{a_1/a_p}^1 \left( 4Q(\alpha a_p) + a_p \frac{d}{dr_p} (Q(\alpha r_p)) \Big|_{r_p=a_p} \right) a_p^2 \alpha^2 b_{1/2}^{(1)}(\alpha) d\alpha + a_1^2 b_{1/2}^{(1)}\left(\frac{a_p}{a_1}\right) Q(a_1) \right]. \quad (\text{A14})$$

Following a similar procedure, for  $I_2^{(1)}$  we have

$$I_2^{(1)} \cong -\pi G \frac{e_p}{2} \cos \Delta\varpi \left[ \int_{a_p/a_2}^1 \left( 4Q\left(\frac{a_p}{\alpha}\right) + a_p \frac{d}{dr_p} \left(Q\left(\frac{r_p}{\alpha}\right)\right) \Big|_{r_p=a_p} \right) \frac{a_p^2}{\alpha^3} b_{1/2}^{(1)}(\alpha) d\alpha - a_2^2 b_{1/2}^{(1)}\left(\frac{a_p}{a_2}\right) Q(a_2) \right]. \quad (\text{A15})$$

Equations (A7), (A8), (A14), and (A15) provide the desired contribution to the disturbing function from the circular annulus  $\mathbb{D}_c$  since  $R(\mathbb{D}_c) = I_1^{(0)} + I_1^{(1)} + I_2^{(0)} + I_2^{(1)}$ .

#### Calculation of the Part of the Disturbing Function Due to $\mathbb{D}_i$

Now let us turn to the inner crescent  $\mathbb{D}_i$ . As the eccentricity of our disk is small, the width of this crescent is already  $O(e_d)$ , thus, we only need to keep the terms up to first order in  $e_p$  in our expansion. Following the recipe in Silsbee & Rafikov (2015b), we write

$$R(\mathbb{D}_i) = - \left\langle G \int_{a_1(1-e_1)}^{a_1(1+e_1)} \frac{r_d}{r_p} \Sigma(r_d) dr_d \int_{\xi-\Delta\varpi-\varphi_p}^{2\pi-\xi-\Delta\varpi-\varphi_p} \frac{d\theta}{\sqrt{1+\alpha^2-2\alpha \cos \theta}} \right\rangle, \quad (\text{A16})$$

where  $\xi(r_d) = \arccos((a_1 - r_d)/e_1 a_1)$  determines the azimuthal extent of the crescent for a given  $r_d$  (Silsbee & Rafikov 2015b), and  $\alpha = r_d/r_p$ .

For the function in the inner integral, we can use the Fourier expansion,

$$(1 + \alpha^2 - 2\alpha \cos \theta)^{-1/2} = \frac{1}{2} b_{1/2}^{(0)}(\alpha) + \sum_{j=1}^{\infty} b_{1/2}^{(j)}(\alpha) \cos(j\theta), \quad (\text{A17})$$

letting us to calculate the inner integral as

$$\begin{aligned} & \int_{\xi - \Delta\varpi - \varphi_p}^{2\pi - \xi - \Delta\varpi - \varphi_p} \left( \frac{1}{2} b_{1/2}^{(0)}(\alpha) + \sum_{j=1}^{\infty} b_{1/2}^{(j)}(\alpha) \cos(j\theta) \right) d\theta = (\pi - \xi(r_d)) b_{1/2}^{(0)}(\alpha) + \sum_{j=1}^{\infty} \frac{1}{j} b_{1/2}^{(j)}(\alpha) \sin(j\theta) \Big|_{\xi - \Delta\varpi - \varphi_p}^{2\pi - \xi - \Delta\varpi - \varphi_p} \\ & = (\pi - \xi(r_d)) b_{1/2}^{(0)}(\alpha) - \sum_{j=1}^{\infty} \frac{2}{j} b_{1/2}^{(j)}(\alpha) \sin(j\xi(r_d)) \cos(j(\Delta\varpi + \varphi_p)). \end{aligned} \quad (\text{A18})$$

The next step is to remember that we only need to keep terms up to  $O(e_d)$  in equation (A16) and to notice that the interval over which radial integration is performed is already  $O(e_d)$ . As a result, in (A16) we can set  $r_d \approx a_1$ ,  $\Sigma_d(r_d) \approx \Sigma_d(a_1)$  and take them out of the integral. We can also set  $\alpha \approx a_1/r_p$  in equation (A18) for the same reason. As a result, we find

$$\begin{aligned} R(\mathbb{D}_i) & \approx -G\Sigma(a_1) \left\langle \frac{a_1}{r_p} \int_{a_1(1-e_1)}^{a_1(1+e_1)} \left[ (\pi - \xi(r_d)) b_{1/2}^{(0)} \left( \frac{a_1}{r_p} \right) - \sum_{j=1}^{\infty} \frac{2}{j} b_{1/2}^{(j)} \left( \frac{a_1}{r_p} \right) \sin(j\xi(r_d)) \cos(j(\Delta\varpi + \varphi_p)) \right] dr_d \right\rangle \\ & = -\pi G e_1 a_1^2 \Sigma(a_1) \left\langle r_p^{-1} b_{1/2}^{(0)} \left( \frac{a_1}{r_p} \right) - r_p^{-1} b_{1/2}^{(1)} \left( \frac{a_1}{r_p} \right) \cos(\Delta\varpi + \varphi_p) \right\rangle, \end{aligned} \quad (\text{A19})$$

where we have used the fact that

$$\int_{a_1(1-e_1)}^{a_1(1+e_1)} \sin(j\xi(r_d)) dr_d = e_1 a_1 \int_0^\pi \sin(j\xi) \sin(\xi) d\xi = \delta_{j1} \frac{\pi}{2} e_1 a_1, \quad (\text{A20})$$

where  $\delta_{ij}$  is the Kronecker delta.

Now we just need to expand (A19) up to first order in  $e_p$  and then orbit average the resultant expression, which is straightforward using the expressions (A13):

$$\begin{aligned} & \left\langle r_p^{-1} \left[ b_{1/2}^{(0)} \left( \frac{a_1}{r_p} \right) - b_{1/2}^{(1)} \left( \frac{a_1}{r_p} \right) \cos(\Delta\varpi + \varphi_p) \right] \right\rangle \approx a_p^{-1} \left\langle b_{1/2}^{(0)} \left( \frac{a_1}{a_p} \right) - b_{1/2}^{(1)} \left( \frac{a_1}{a_p} \right) \cos(\Delta\varpi + E) \right\rangle \\ & + a_p e_p \left\langle \cos E \left[ a_p^{-2} b_{1/2}^{(0)} \left( \frac{a_1}{a_p} \right) + a_p^{-1} a_1 a_p^{-2} b_{1/2}^{(0)'} \left( \frac{a_1}{a_p} \right) \right] \right\rangle - \left\langle a_p e_p \cos E a_p^{-2} b_{1/2}^{(1)} \left( \frac{a_1}{a_p} \right) \cos(\Delta\varpi + E) \right. \\ & \left. + a_p e_p \cos E a_p^{-1} a_1 a_p^{-2} b_{1/2}^{(1)'} \left( \frac{a_1}{a_p} \right) \cos(\Delta\varpi + E) - e_p a_p^{-1} b_{1/2}^{(1)} \left( \frac{a_1}{a_p} \right) \sin E \sin(\Delta\varpi + E) \right\rangle \\ & = a_p^{-1} b_{1/2}^{(0)} \left( \frac{a_1}{a_p} \right) + \frac{e_p}{2} a_p^{-1} \left[ b_{1/2}^{(1)} \left( \frac{a_1}{a_p} \right) - \frac{a_1}{a_p} b_{1/2}^{(1)'} \left( \frac{a_1}{a_p} \right) \right] \cos \Delta\varpi. \end{aligned} \quad (\text{A21})$$

As a result, by dropping the first term not containing  $e_p$ , one finally finds

$$R(\mathbb{D}_i) \approx -\pi G \Sigma(a_1) e_1 \frac{e_p}{2} \frac{a_1^2}{a_p} \left[ b_{1/2}^{(1)} \left( \frac{a_1}{a_p} \right) - \frac{a_1}{a_p} b_{1/2}^{(1)'} \left( \frac{a_1}{a_p} \right) \right] \cos \Delta\varpi. \quad (\text{A22})$$

#### Calculation of the Part of the Disturbing Function Due to $\mathbb{D}_o$

This case has only minor differences from the previous one, so we will omit the details:

$$\begin{aligned} R(\mathbb{D}_o) & = \left\langle G \int_{a_2(1-e_2)}^{a_2(1+e_2)} \Sigma(r_d) dr_d \int_{\xi - \Delta\varpi - \varphi_p}^{2\pi - \xi - \Delta\varpi - \varphi_p} \frac{d\theta}{\sqrt{1 + \alpha^2 - 2\alpha \cos \theta}} \right\rangle \\ & \approx \pi G e_2 a_2 \Sigma(a_2) \left\langle b_{1/2}^{(0)} \left( \frac{r_p}{a_2} \right) - b_{1/2}^{(1)} \left( \frac{r_p}{a_2} \right) \cos(\Delta\varpi + \varphi_p) \right\rangle \\ & \approx \pi G e_2 a_2 \Sigma(a_2) \left[ b_{1/2}^{(0)} \left( \frac{a_p}{a_2} \right) + e_p b_{1/2}^{(1)} \left( \frac{a_p}{a_2} \right) \cos \Delta\varpi + \frac{e_p}{2} a_p a_2^{-1} b_{1/2}^{(1)'} \left( \frac{a_p}{a_2} \right) \cos \Delta\varpi \right]. \end{aligned} \quad (\text{A23})$$

Thus, again dropping the first,  $e_p$ -independent term, we eventually arrive at

$$R(\mathbb{D}_o) \approx \pi G \Sigma(a_2) e_2 \frac{e_p}{2} a_2 \left[ 2b_{1/2}^{(1)} \left( \frac{a_p}{a_2} \right) + \frac{a_p}{a_2} b_{1/2}^{(1)'} \left( \frac{a_p}{a_2} \right) \right] \cos \Delta\varpi. \quad (\text{A24})$$

*The full result*

The full expression resulting from evaluating the integral (A1) is given by  $R_d = I_1^{(0)} + I_1^{(1)} + I_2^{(0)} + I_2^{(1)} + R(\mathbb{D}_i) + R(\mathbb{D}_o)$  and can be written out explicitly using equations (A7), (A8), (A14), (A15), (A22), and (A24). When doing this, we introduce two important modifications.

First, it is possible to combine the different terms corresponding to the outer and inner disks into a single expression using certain properties of the Laplace coefficients. For example, we can combine  $I_1^{(0)}$  and  $I_1^{(1)}$  by changing the variable  $\alpha \rightarrow \alpha^{-1}$  in equation (A8) for  $I_1^{(1)}$  and then manipulate the result using the fact that  $b_s^{(j)}(\alpha^{-1}) = \alpha^{2s} b_s^{(j)}(\alpha)$ . This procedure turns  $I_1^{(1)}$  in equation (A8) into the expression analogous to equation (A7) but with the limits of integration running from 1 to  $a_2/a_p$ . This allows us to combine it with  $I_1^{(0)}$  to get an expression analogous to equation (A7) with integration running from  $a_1/a_p < 1$  to  $a_2/a_p > 1$  (for orbits *within* the disk,  $a_1 < a_p < a_2$ ). The integrand of this final expression has a weak (logarithmic) singularity at  $\alpha = 1$ , but the integral itself is fully convergent. This procedure is then repeated to combine  $I_2^{(0)}$  and  $I_2^{(1)}$ . In addition, differentiating the identity  $b_s^{(j)}(\alpha^{-1}) = \alpha^{2s} b_s^{(j)}(\alpha)$  we obtain the relation for  $b_s^{(j)'}(\alpha^{-1})$ , which allows us to also merge  $R(\mathbb{D}_i)$  and  $R(\mathbb{D}_o)$  into a single expression.

Second, in all these expressions, we switch from  $d/dr_p$  to derivatives with respect to the argument (denoted by prime), e.g.  $dQ(\alpha r_p)/dr_p = \alpha Q'(a)|_{a=\alpha r_p}$ .

As a result of performing these steps, one finds that  $R_d$  can be written in the form (3) with the coefficients  $A_d$  and  $B_d$  given by the following expressions:

$$A_d = A_d^{\text{bulk}} + A_d^{\text{edge}}, \quad B_d = B_d^{\text{bulk}} + B_d^{\text{edge}}, \quad (\text{A25})$$

$$A_d^{\text{bulk}} = \frac{\pi G}{2a_p n_p} \int_{a_1/a_p}^{a_2/a_p} \alpha b_{1/2}^{(0)}(\alpha) (a^2 \Sigma_d(a))'' \Big|_{a=\alpha a_p} d\alpha, \quad (\text{A26})$$

$$A_d^{\text{edge}} = -\frac{\pi G}{2a_p n_p} \left[ \left( \frac{a}{a_p} \right)^2 b_{1/2}^{(0)} \left( \frac{a}{a_p} \right) (a \Sigma_d(a))' - \left( \frac{a}{a_p} \right)^3 b_{1/2}^{(0)'} \left( \frac{a}{a_p} \right) \Sigma_d(a) \right] \Bigg|_{a=a_1}^{a=a_2}, \quad (\text{A27})$$

$$B_d^{\text{bulk}} = -\frac{\pi G}{2n_p} \int_{a_1/a_p}^{a_2/a_p} \alpha^2 b_{1/2}^{(1)}(\alpha) [4(\Sigma_d(a)e_d(a))' + a(\Sigma_d(a)e_d(a))''] \Big|_{a=\alpha a_p} d\alpha, \quad (\text{A28})$$

$$B_d^{\text{edge}} = \frac{\pi G}{2a_p^3 n_p} \left[ a^3 b_{1/2}^{(1)} \left( \frac{a}{a_p} \right) (\Sigma_d(a)e_d(a))' + a^2 \left( b_{1/2}^{(1)'} \left( \frac{a}{a_p} \right) - \frac{a}{a_p} b_{1/2}^{(1)} \left( \frac{a}{a_p} \right) \right) \Sigma_d(a)e_d(a) \right] \Bigg|_{a=a_1}^{a=a_2}, \quad (\text{A29})$$

where  $F(a)|_{a=a_1}^{a=a_2} \equiv F(a=a_2) - F(a=a_1)$ . If we consider a disk model with power-law profiles of  $\Sigma_d(a) \propto a^{-p}$  and  $e_d(a) \propto a^{-q}$ , the results (A25)-(A29) reduce to the corresponding formulae in Silsbee & Rafikov (2015b).

GAP

The ring profile used in §5.1 to illustrate the effects of sharp disk edges on the free precession rate is given by  $\Sigma_d(a) = \Sigma_0(a, \nu)$  with

$$\Sigma_0(a, \nu) = 100 \times \exp \left[ 1 - \exp \left( \frac{1}{0.1 + 5^6 ((x-3)^\nu + 1)^{-6}} \right) \right] \text{ g cm}^{-2}, \quad (\text{B1})$$

where  $x = a/\text{AU}$  and  $\nu$  is a parameter controlling the sharpness of the disk edges; see Figure 14.

The presence of a gap of radial width  $w$  and relative depth  $d < 1$  is modeled in §5.2 using the surface density profile  $\Sigma_d(a) = \Sigma_0(a, 2) \times \text{Gap}(a, d, w)$ , where

$$\text{Gap}(a, d, w) = 1 - d \times \exp \left[ 1 - e^{u(x, w) - u(3, w)} \right], \quad u(x, w) = \left[ 0.1 + \left( \frac{w}{(x-3)^2 + 1} \right)^5 \right]^{-1}. \quad (\text{B2})$$