A new approach to the problem of modes in the Mestel disk

M. Demleitner and B. Fuchs

Astronomisches Recheninstitut, Mönchhofstr. 12-14, 69120 Heidelberg, Germany

Abstract. We examine the modes admitted by the Mestel disk, a disk with a globally flat rotation curve. In contrast to previous analyses of this problem by Zang (1976) and Evans & Read (1998a, 1998b), we approximate the orbits to obtain almost closed expressions for the kernel of the integral equation governing the behaviour of the modes. Otherwise we, like them, follow Kalnajs' programme to simultaneously solve the Boltzmann and Poisson equations. We investigate the modes admitted by both the self-consistent and a cut-out Mestel disk, the difference being that in the latter, a part of the matter in the disk is immobilised. This breaks the self-similarity and produces a pronouncedly different picture, both technically and in terms of the disk properties. The self-consistent disk is governed by a Cauchy integral equation, the cut-out disk by an integral equation that can be treated as a Fredholm equation of the second kind. In general, our approximation reproduces the results of the previous works remarkably well, yielding quantities mostly within 5% of the values reported by Zang and Evans & Read and thus also the basic result that in a "standard" cut-out disk, only one-armed modes are unstable at the limit of axisymmetric stability. In the self-consistent disk, relatively compact expressions for the kernel allow an intuitive understanding of most of the properties of neutral (non-rotating, non-growing) modes there. We finally show that self-consistent Mestel disks do not admit growing or rotating modes in this sort of stellar-dynamical analysis.

Key words. galaxies: spiral – galaxies: structure – stellar dynamics

1. Introduction

It was not long after Kalnajs (1971, 1977, and references therein) and others had laid out a programme to perform stability analyses for galactic disks that it was applied to the Mestel disk by Zang (1976; henceforth cited as Zang) in a Ph.D. Thesis supervised by Alar Toomre. This disk, characterised by an infinitesimal height and a globally constant rotation curve, lends itself as a model for disks that show spiral structure mainly for two reasons. Firstly, its structure is so simple that one of its distribution functions has been made into a textbook example (Binney & Tremaine 1987); secondly, and probably more importantly, flat rotation curves seem to be rather prevalent among (high surface brightness) spiral galaxies (e.g., Persic & Salucci 1991), and it has long been suspected that flat rotation curves and strong, "modal-looking" spiral structure are somehow linked (e.g., Biviano et al. 1991).

One of the drawbacks of the Mestel disk is that the orbits in it cannot be expressed in elementary functions. Zang circumvented this problem essentially through numerical integration and some elegant devices exploiting the self-similarity of the Mestel disk (i.e., orbits of a given eccentricity look like scaled copies of each other,

regardless of the angular momentum of a particle travelling on the orbit) to curb down on the computational effort. While the self-similarity helps to simplify the calculation, it also implies that a self-consistent Mestel disk does not possess any scales. In particular, the absence of a time scale already leads to the expectation that discrete modes will be hard to construct.

Zang found that the Mestel disk is a rather bewildering construct. His formalism leads to a singular integral equation that is quite difficult to treat numerically. However, the structure of its kernel as well as physical arguments suggest that there are no discrete modes apart from rather exotic non-rotating and non-growing ones, implying that either no growing modes exist or modes become unstable regardless of their growth rate and pattern speed. Thus, his investigations concentrated on a variant of the Mestel disk in which a certain part of the disk is declared "immobile", the cut-out disk. This tampered system is furnished with a length scale, admits discrete modes and becomes treatable with his numeric technique.

But even the cut-out disk behaves contrary to the preconceptions of the mid-70s. While Zang would have liked to see a mildly unstable two-armed mode, typically the only mode that is growing in an axisymmetrically stable disk is a one-armed mode, and for this, a stability limit cannot be found. Even when he made the inner cut-out

---

Send offprint requests to: M. Demleitner,
e-mail: `msdemlei@ari.uni-heidelberg.de`

so steep that $m = 2$ modes emerged in axisymmetrically stable disks in a way that Toomre (1977) describes as unrealistic, these one-armed modes, again in Toomre's words continued to "plague" him.

Despite these somewhat unexpected findings and the fact that it was only fragmentarily published in Toomre (1977) and Toomre (1981), Zang's thesis became a frequently-cited classic and the Mestel disk developed a life of its own. Recently, Evans & Read (1998a, 1998a; see also Read (1997; henceforth cited as ERI and ERII)) generalised Zang's work to the entire family of physically plausible disks with surface densities following a power law, and Goodman & Evans (1999) once more tackled the issue of rotating modes in the untampered (self-consistent) Mestel disk.

This work set out as an attempt to avoid the numerical integrations that had to be employed by Zang and ERI by approximating the orbits and eventually deriving a closed kernel of the integral equation governing the problem. Among the merits of the availability of such a kernel is that an analytic expression for the kernel might allow a decision whether or not the self-consistent Mestel disk admits rotating modes. Also, it might facilitate an analysis of the processes at work at various resonances of the modes in the cut-out disk in the context of a full disk without having to resort to somewhat daring techniques to extend WKBJ approximations across resonances (e.g., Mark 1976).

The organisation of this paper is as follows: in Sect. 2, we derive the orbits and the Hamiltonian of the disk. Section 3 basically follows Kalnajs' programme that will in the end yield an (almost) closed expression for the kernel of an integral equation governing the modes in the disk. To assess the validity of this kernel, we will compare the results reported in ERII with ours in Sect. 4. Finally, we revisit the issue of modes in the self-consistent disk in Sect. 5.

2. The orbits

The equations of motion in the Mestel disk are

$$
\begin{aligned}
\ddot{r} &= r\dot{\theta}^2 - 1/r \\
L_z &= r^2\dot{\theta} = \text{const.}
\end{aligned}
\tag{1}
$$

Here, r and θ denote polar coordinates, and L_z is the z-component of the angular momentum per unit mass. Units have been chosen such that the circular velocity $v_c = 2\pi G\Sigma_0 = 1$, where G is the gravitational constant, and Σ_0 is the surface density at a reference radius that can be set to unity because of the self-similarity of the disk.

We seek closed expressions for the solution of the Eq. (1). Since the exact solutions cannot be expressed in elementary functions, we need to employ some approximations. A straightforward linearisation in configuration space – the epicyclic approximation – fails for the Mestel disk, since the Hamiltonian derived from it does not contain contributions from non-circular motion components. Therefore we linearise in Fourier space,

a procedure generally known as the method of harmonic balance (Bogoliubov & Mitropolski 1961).

Inserting one of the Eq. (1) into the other yields

$$
\ddot{r} = L_z^2/r^3 - 1/r.
\tag{2}
$$

We now substitute $r(t) = R_0(1 + \zeta\cos(\kappa t))$, making R_0 the reference radius of a given orbit, ζ a measure for its deviation from circularity, and κ an epicyclic frequency. After this, the first Fourier coefficients of (2) with respect to t are

$$
a_0 = \frac{L_z^2(1 + \zeta^2/2) - R_0^2(1 - \zeta^2)^2}{R_0^3(1 - \zeta^2)^{5/2}}
\tag{3}
$$

$$
a_1 = -\frac{2R_0^2(1 - \zeta^2)^2(\sqrt{1 - \zeta^2} - 1) + 3L_z^2\zeta^2}{\zeta R_0^3(1 - \zeta^2)^{5/2}}
\tag{4}
$$

$$
b_1 = 0,
\tag{5}
$$

where a_i and b_i denote the cosine and sine coefficients, respectively.

Dropping all higher Fourier coefficients and substituting the result back into (2), one finds by comparing the terms independent of t

$$
L_z = \frac{\sqrt{2}R_0(1 - \zeta^2)}{\sqrt{\zeta^2 + 2}},
\tag{6}
$$

and by comparing terms proportional to $\cos(\kappa t)$

$$
\kappa = \frac{\sqrt{2}}{\zeta R_0}\sqrt{1 - 2\frac{\sqrt{1 - \zeta^2}}{\zeta^2 + 2}}.
\tag{7}
$$

The circular frequency of the guiding centre Ω results from a Fourier transformation of the second Eqs. in (1) with (6) inserted and is given by

$$
\Omega = \frac{L_z}{R_0^2(1 - \zeta^2)^{3/2}}.
\tag{8}
$$

We can now derive the Hamiltonian for a mass point in the Mestel disk in action-angle coordinates by simply inserting (6), (7), and (8) into a Hamiltonian written in Cartesian coordinates. Assuming orbits not too far from circularity, one can linearise with respect to ζ. After again dropping all Fourier components of order 2 or higher, one is left with

$$
H(J, L_z) = \frac{1}{2} + \ln(L_z) + \frac{\sqrt{2}J}{L_z},
\tag{9}
$$

where $J = \kappa\zeta^2 R_0^2/2$ is the energy in the epicyclic motion divided by κ and thus the second integral. This clearly is the simplest Hamiltonian that could have anything to do with a Mestel disk and indeed the Hamiltonian that would result from a classic epicyclic approximation if it worked. As a matter of fact, at this point one has $\kappa = \sqrt{2}/R_0$ and $\Omega = 1/R_0$, reproducing the result of an epicyclic analysis. This Hamiltonian has already been used by Collett et al. (1997).

The set of approximated action-angle coordinates L_z, J, $w_1 := \kappa t$, $w_2 := \Omega t$ is correct to first order in ζ, and the canonical equations are correct to first order in ζ as well. As is demonstrated in Demleitner (2000), the Jacobian of the transformation between Cartesian coordinates and this set is $1 + O(\zeta^2)$, showing that the transformation is indeed canonical to the required order.

Because of the two approximations (only keeping the lowest Fourier coefficients and linearising with respect to ζ), it is not a priori clear that the disk described by this transformation and the resulting Hamiltonian (9) does indeed behave like a Mestel disk. It is one of the purposes of this work to show a posteriori, at is were, that this is the case.

3. Derivation of the matrix equation

With the Hamiltonian and the transformation equations, we can now follow the programme outlined in the series of papers by Kalnajs cited above. The first step is to establish a base for the potential perturbation. We choose

$$
\begin{aligned}
\phi_{\alpha,m} = \; & L_z^{i\alpha-1/2}\mathrm{e}^{imw_2-i\omega t} \\
& \mathrm{e}^{\left(i\sqrt{2J/L_z}\left(\alpha\cos(w_1)-\sqrt{2}m\sin(w_1)\right)\right)}.
\end{aligned} \tag{10}
$$

Here, m is the number of circumferential maxima of the spiral pattern (the number of arms), and α parametrises how closely the spiral is wound, where a spiral with larger α is more trailing provided that $\omega > 0$. This expression basically is the well-known logarithmic spiral $\mathrm{e}^{i\alpha\ln(r)+im\theta-i\omega t}$ transformed into our action-angle coordinates with harmonics in the angle coordinates higher than one disregarded. The factor $L_z^{-1/2}$ was added to ensure definiteness in the integration of the angular momentum. This analogy motivates following ERI in naming α the logarithmic wave number and m the angular harmonic number. We assume that all potential perturbations can be represented as expansions of the form

$$
\sum_{m=0}^{\infty} \int_{-\infty}^{\infty} \Phi_m(\alpha)\phi_{\alpha,m}(w_1,w_2,J,L_z)\,\mathrm{d}\alpha. \tag{11}
$$

The next step in Kalnajs' programme is to solve the linearised Boltzmann equation

$$
\frac{\partial f_1}{\partial t} + [f_1, H] = -[f_0, \phi_{\alpha,m}], \tag{12}
$$

written with Poisson brackets $[\,.\,,.\,]$ for an elementary perturbations $\phi_{\alpha,m}(w_1,w_2,J,L_z)$ of the form (10). Here, f_1 is the unknown perturbation of the distribution function, and f_0 an equilibrium distribution function; for f_0, we use the distribution function given by Binney & Tremaine (1987), which, in our units and coordinates, and after applying Stirling's formula, is

$$
\begin{aligned}
f_0(L_z, J) \;=\; & \frac{F}{L_z} \exp\left(-\frac{\sqrt{2}J}{\sigma^2 L_z}\right) \quad \text{with} \\
& F = \frac{\sqrt{1-2\sigma^2}^{\,1-1/\sigma^2}}{2\sqrt{2}\mathrm{e}\pi^2\sigma^2 \mathrm{G}}.
\end{aligned} \tag{13}
$$

Expanding the Poisson brackets in (12) yields

$$
\begin{aligned}
\frac{\partial f_1}{\partial t} + \frac{\sqrt{2}}{L_z}\frac{\partial f_1}{\partial w_1} + \frac{1}{L_z}\frac{\partial f_1}{\partial w_2} &= -i\frac{FL_z^{i\alpha}\mathrm{e}^{imw_2-i\omega t}}{\sigma^2 L_z^{5/2}} \\
& \left(X(w_1)-m\sigma^2\right)\exp\left(iX(w_1)-\frac{\sqrt{2}J}{\sigma^2 L_z}\right)
\end{aligned} \tag{14}
$$

with an auxiliary function $X(w_1)=\sqrt{\sqrt{2}J/L_z}\big(\alpha\cos(w_1)-\sqrt{2}m\sin(w_1)\big)$. On the right-hand side of this equation, a term of order ζ^2 has been neglected.

To solve this equation, we perform a separation of variables by

$$
f_1(w_1,w_2,t) = g_1(w_1)\mathrm{e}^{imw_2-i\omega t} \tag{15}
$$

and arrive at an inhomogeneous ordinary differential equation, the homogeneous solution of which is

$$
\tilde{g}_1(w_1) = \mathrm{e}^{i(\omega L_z-m)w_1/\sqrt{2}}. \tag{16}
$$

The general solution can be obtained via a variation of constants, where the integration constant is fixed by the condition that the distribution function must be 2π-periodic, $g_1(w_1) = g_1(w_1 + 2\pi)$.

Before writing down the result, we note that in (14) the terms with w_1 in the exponent and its coefficient are quite similar to each other. This invites a partial integration in the resulting expression, which after some algebra leads to

$$
\begin{aligned}
g_1(w_1) = \; & C\sqrt{\frac{L_z}{\sqrt{2}J}} \exp\left(-\frac{\sqrt{2}J}{\sigma^2 L_z}\right) \bigg(i\exp\big(iX(w_1)\big) \\
& -\frac{1}{\mathrm{e}^{-2\pi i\eta}-1}\left(\sqrt{2}\eta+m\sigma^2\right) \\
& \int_0^{2\pi} \exp\big(iX(w_1+w_1')-i\eta w_1'\big)\,\mathrm{d}w_1'\bigg).
\end{aligned} \tag{17}
$$

In this expression, we have abbreviated $\eta = (\omega L_z - m)/\sqrt{2}$.

Substituting (17) into (15) yields the perturbation of the distribution function under an elementary potential perturbation $\phi_{\alpha,m}$. Since (12) is linear in $\phi_{\alpha,m}$, the response to an arbitrary perturbation $\Phi_1^{\mathrm{imposed}}$ of the form (11) can immediately be computed by summing over all α and m. Thus, the solution of (12) can be written as

$$
f_1^{\mathrm{response}} = \mathcal{L}\Phi_1^{\mathrm{imposed}} \tag{18}
$$

with a linear integral operator $\mathcal{L}$.

Further pursuing Kalnajs' programme, we attempt to find a linear combination of elementary perturbations such that f_1^{response} causes the $\Phi_1^{\mathrm{imposed}}$ we started with. To do this, the Poisson equation has to be solved simultaneously to the Boltzmann equation. Since $\mathcal{L}$ and the Laplacian occurring in the Poisson equation do not commute, their eigenfunctions are different, and a straightforward attempt to solve the two equations simultaneously will result in an integro-differential equation that will be very hard to solve indeed.

To avoid this, one converts the system of equations to a matrix equation. The idea here is to compute potential-density pairs, i.e., two sets of mutually orthogonal basis functions $\phi_{\alpha,m}$ and $\mu_{\alpha,m}$ for each of the perturbations in density and potential where each pair $(\phi_{\alpha,m}, \mu_{\alpha,m})$ satisfies the Poisson equation. When one expands the perturbations in the potential and the surface density in these functions, the Poisson equation simply reads $\Phi_m(\alpha) = M_m(\alpha)$ in the coefficients $\Phi_m(\alpha)$ and $M_m(\alpha)$ of these expansions.

The basis $\phi_{\alpha,m}$ is already fixed by (10), and the corresponding density base function can be found using the scheme of Clutton-Brock (1972). For our calculation, however, the only important property of the potential-density pair is the scalar product of $\phi_{\alpha,m}$ and $\mu_{\alpha,m}$. This scalar product is of course invariant under canonical transformations, and to the the required order Clutton-Brock's result

$$\int_0^{2\pi} d\theta \int_0^\infty r \, dr \, \phi_{\alpha,m}^* \mu_{\alpha',m'} = 4\pi^2 K_m(\alpha)\delta_{m,m'}\delta(\alpha - \alpha') \quad (19)$$

for logarithmic spirals holds. Here,

$$K_m(\alpha) = \frac{\left|\Gamma\left(\frac{3}{4} + \frac{1}{2}m + \frac{1}{2}i\alpha\right)\right|^2}{\pi G \left|\Gamma\left(\frac{1}{4} + \frac{1}{2}m + \frac{1}{2}i\alpha\right)\right|^2} \quad (20)$$

is a quantity closely related to ERI's Kalnajs factor.

After expanding the potential and surface density perturbations in (18), one can project both sides on $\phi_{\alpha',m'}$. Exploiting (19) and inserting the (now trivial) Poisson equation, (18) takes the form

$$4\pi^2 K_{m'}(\alpha')\Phi_{m'}^{\text{response}}(\alpha') = \int_{-\infty}^\infty d\alpha \int_0^\infty dL_z \int_0^\infty dJ$$
$$\int_0^{2\pi} dw_1 \sum_{m=0}^\infty \int_0^{2\pi} dw_2 \, \Phi_m^{\text{imposed}}(\alpha)(\mathcal{L}\phi_{\alpha,m})\phi_{\alpha',m'}^*. \quad (21)$$

Equating Φ^{response} and Φ^{imposed}, an integral equation results, the solutions of which, if they exist, are self-consistent modes.

The integrand of (21) can be written as a sum $I_1 + I_2$ of a part without a further integral,

$$I_1 = \Phi_{\alpha,m}\frac{FL_z^{i(\alpha-\alpha')-2}}{4\pi^2\sigma^2 K_{m'}(\alpha')}e^{i(m-m')w_2 - \frac{\sqrt{2}J}{\sigma^2 L_z}}$$
$$e^{i\sqrt{\frac{\sqrt{2}J}{L_z}}\left((\alpha-\alpha')\cos(w_1) - \sqrt{2}(m-m')\sin(w_1)\right)}, \quad (22)$$

and one in which a further integration over the auxiliary angle w_1' has to be performed,

$$I_2 = \int_0^{2\pi} dw_1' \, i\Phi_{\alpha,m}\frac{\sqrt{2}FL_z^{i(\alpha-\alpha')-2}(m\sigma^2 + \sqrt{2}\eta)}{8\pi^2\sigma^2 K_{m'}(\alpha')(e^{-2\pi i\eta} - 1)}$$
$$e^{i(m-m')w_2 - i\eta w_1' - \frac{\sqrt{2}J}{\sigma^2 L_z} + \sqrt{2}\,m'\sin(w_1)}$$
$$e^{i\sqrt{\frac{\sqrt{2}J}{L_z}}\left(\alpha\cos(w_1'+w_1) - \alpha'\cos(w_1) - \sqrt{2}\,m\sin(w_1'+w_1)\right)}. \quad (23)$$

The quadratures over w_2, w_1, and J, carried out in that order, are rather straightforward. Details can be found in

Demleitner (2000). When integrating over w_1' one has to evaluate an expression of the type

$$\int_0^{2\pi} e^{i\tau w_1'}e^{-\nu\cos(v-w_1')} \, dw_1', \quad (24)$$

which can be done by expanding the second factor of the integrand into modified Bessel functions I_k using

$$e^{-\nu\cos(v-w_1')} = I_0(\nu) + 2\sum_{k=1}^\infty I_k(\nu)\cos\left(k(v - w_1')\right) \quad (25)$$

(Abramowitz & Stegun 1972). Unfortunately, the series introduced by this operation does not in general collapse and will be present in the kernel.

The integration over the angular momentum is the most involved one, resembling the calculations to compute the angular momentum function in Zang's formalism. For the summands of the series resulting from the w_1' integration, one has to treat integrals of the structure

$$\int_{-\infty}^\infty \frac{e^{(i\gamma+\lambda+1)u}}{2\left(2\xi^2 - (1-e^u)^2\right)} \, du, \quad (26)$$

where

$$u = \ln(\omega L_z/m) \quad (27)$$

(note that this substitution only works for nonzero m and ω), $\gamma = \alpha - \alpha'$ and $\xi = k/m$. The integer k is a summation index varying between 0 and ∞, and the complete integral requires this expression to be evaluated for $\lambda \in \{-1, 0, 1\}$.

We will compute this integral using the residue theorem, and thus the first step is to find the zeroes of the denominator of the integrand. These are at

$$u_j^{(1,2)} = \ln(1 \pm \sqrt{2}\xi) + 2\pi ij \quad (28)$$

for integer j. These zeroes are well-defined, since ξ, as the ratio of two integers, can never be equal to $1/\sqrt{2}$. Let for the moment $\xi < 1/\sqrt{2}$ and $\gamma > 0$.

There are three technical problems in the application of the residue theorem to this integral. The first is that at least one pole is on the real axis and hence on the integration contour. The tempting idea to take the (existing) Cauchy principal value leads to erroneous results. As Landau (1946) noted in a plasma-physical context, the correct reasoning is that if a system supports a non-growing perturbation at present and started in equilibrium, the perturbation must have been a growing one at some time in the past. Since it has evolved continuously, the integration contours of the problem must have been continuously deformed as well and thus cannot have jumped across poles. A growing perturbation corresponds to an ω with a positive imaginary part, which with the substitution (27) means that the integration contour moves into the upper half-plane. Therefore, as $\Im\omega \to 0$, one has to leave the zeroes on the real axis out of the contour, as is sketched in Fig. 1.

The second problem is that as the size of the integration path $R \to \infty$ (cf. Fig. 1) one cannot continuously
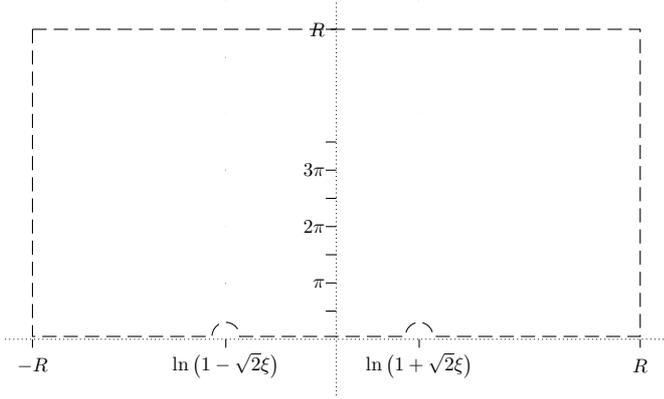
Fig. 1. Poles of the integrand of (26) and a sketch of the integration path. The diamonds in the left column mark the positions of the poles at $\Re(u) = \ln(1 - \sqrt{2}\xi)$ in the case $\xi > 1/\sqrt{2}$

deform the contour since the poles are in the way, and the same poles make it hard to establish that the contribution of the line at $\Im u = iR$ vanishes. The standard technique to avoid this is to write down the series of residues and show that integral over the difference between this series and the integrand vanishes as $R \to \infty$. Since this difference has no poles, the contour can be readily deformed.

The last problem is that for $\lambda = \pm 1$ the contributions at $\Re u = \pm R$ do not necessarily vanish. Indeed, one finds that these "border terms"

$$
= \begin{cases}
-\frac{i}{4\gamma\xi^2 - 2} \lim_{x \to -\infty} \frac{e^{i\gamma x}}{\gamma}; & \lambda = -1, \ u \to -\infty \\
-\frac{i}{2} \lim_{x \to -\infty} \frac{e^{i\gamma x}}{\gamma}; & \lambda = 1, \ u \to \infty .
\end{cases} \quad (29)
$$

Of course, the limits occurring in these expressions do not exist. However, it should be remembered that this integral is being evaluated as part of the kernel of an integral equation. Therefore, the limits can (and have to) be taken in the sense of a distribution limit, $\lim_{x \to \pm\infty} e^{i\gamma x}/\gamma = \pm i\pi\delta(\gamma)$.

With these tools, one can go about evaluating the residues of the integrand, which can easily be done using $\mathrm{res}(f, z) = \lim_{u \to z} f(u)(u - z)$ and an application of de l'Hospital's rule. We will give the end result below.

A few words about the other cases are in order. First, if $\xi > 1/\sqrt{2}$, the poles at $u_j^{(1)}$ move away from the real axis (cf. Fig. 1). One can keep the expression resulting for $\xi < 1/\sqrt{2}$ by choosing an appropriate branch of the logarithm in the result. Basically, one has to see to it that the $u_j^{(1)}$ in Fig. 1 "shift down" rather than "up" by $i\pi$. To ensure this, one sets $\ln(-1) = -i\pi$, and with $\ln(1) = 0$ a natural branch cut is along the positive imaginary axis, so that $\frac{i\pi}{2} \geq \ln(e^{i\phi}) > -3\pi i/2$.

If $\gamma < 0$, one has to integrate through the lower half-plane, and consequently the poles on the real axis are added into the residue sum. The resulting expression is, of course, identical to the one for $\gamma > 0$.

After carrying out all integrations and the summation of m (which is trivial because the integration over w_2

contributes a $\delta_{m,m'}$, thus decoupling different angular harmonics), (21) takes the form

$$
\Phi_{m'}(\alpha') = \mathcal{F}(\alpha')\Phi_{m'}(\alpha') + \int_{-\infty}^{\infty} \mathcal{K}(\alpha, \alpha')\Phi_{m'}(\alpha) \, d\alpha. \quad (30)
$$

The summand with $\mathcal{F}$ originates from the terms with δ-functions and computes to

$$
\mathcal{F}(\alpha') = \frac{\sqrt{1 - 2\sigma^2}^{\,1 - 1/\sigma^2}}{2\sqrt{2}\sigma^2} \frac{\left|\Gamma\left(\frac{1}{4} + \frac{1}{2}m + \frac{1}{2}i\alpha'\right)\right|^2}{\left|\Gamma\left(\frac{3}{4} + \frac{1}{2}m + \frac{1}{2}i\alpha'\right)\right|^2}
$$
$$
\times \left(1 - (1 + \sigma^2)e^{-\sigma^2(\alpha'^2 + 2m'^2)/2}\left(I_0\left(\frac{\sigma^2}{2}(\alpha'^2 + 2m'^2)\right)\right.\right.
$$
$$
\left.\left. + \sum_{k=1}^{\infty} \frac{2m'^2}{m'^2 - 2k^2} I_k\left(\frac{\sigma^2}{2}(\alpha'^2 + 2m'^2)\right)\right)\right). \quad (31)
$$

The kernel itself becomes

$$
\mathcal{K}(\alpha, \alpha') = -i \frac{\omega^{-i(\alpha - \alpha')}\sqrt{1 - 2\sigma^2}^{\,1 - 1/\sigma^2}}{2\sigma^2 e^{\sigma^2(\alpha^2 + \alpha'^2 + 4m'^2)/4 + 1}\left(e^{2\pi(\alpha - \alpha')} - 1\right)}
$$
$$
\frac{\left|\Gamma\left(\frac{1}{4} + \frac{1}{2}m + \frac{1}{2}i\alpha'\right)\right|^2}{\left|\Gamma\left(\frac{3}{4} + \frac{1}{2}m + \frac{1}{2}i\alpha'\right)\right|^2}\left(m'^{i(\alpha - \alpha')}\sigma^2 I_0(\tau)\right.
$$
$$
- \sum_{k=1}^{\infty} I_k(\tau)\left((\sqrt{2}k - m'\sigma^2)\Lambda^k e^{(i(\alpha - \alpha') - 1)\ln(m' - \sqrt{2}k)}\right.
$$
$$
\left.\left. - (\sqrt{2}k + m'\sigma^2)\Lambda^{-k}e^{(i(\alpha - \alpha') - 1)\ln(m' + \sqrt{2}k)}\right), \quad (32)
$$

where we have abbreviated

$$
\tau = \frac{\sigma^2}{2}\sqrt{(\alpha^2 + 2m'^2)(\alpha'^2 + 2m'^2)}
$$
$$
\Lambda = \frac{\alpha\alpha' + \sqrt{2}m'(\sqrt{2}m' + i(\alpha - \alpha'))}{\sqrt{(2m'^2 + \alpha^2)(2m'^2 + \alpha'^2)}}. \quad (33)
$$

Formally, (30) is an integral equation of the second kind. However, $\mathcal{K}$ is not compact – the infinite integration range and the singularity on the $\alpha = \alpha'$ diagonal prevent this –, and thus the solution theory for it is significantly different from what one has in the Fredholm case. We will return to this issue later. For now, let us just follow Zang in noting that ω factors out of the kernel, and thus the value of ω is irrelevant for the solvability of this equation.

The special case of $\omega = 0$ has not been treated so far. It is interesting for two reasons: For one, when searching for the limit of axisymmetric stability, it suffices to consider $\omega = 0$, since axisymmetric patterns cannot rotate anyway and considering growth is not necessary for establishing the stability margin, and for two, "neutral" modes (for which $\omega = 0$ even when $m \neq 0$) were the only modes the existence of which ERII and Zang report for the self-consistent disk.

The expression (32) cannot be taken to the limit $\omega \to 0$, because $\omega^{i(\alpha - \beta)}$ in general has an essential singularity at $\omega = 0$, which according to Picard's theorem implies that the image of an arbitrarily small neighbourhood of 0 is the entire complex plane with the possible exception of one point. Instead, one has to substitute $\omega = 0$ before

the integration over L_z. This simplifies the terms quite significantly, resulting in

$$\mathcal{K}_0(\alpha') = 2\mathcal{F}(\alpha'). \tag{34}$$

The integral equation has now collapsed into the algebraic equation

$$\begin{aligned} \Phi_{m'}(\alpha') &= \mathcal{K}_0(\alpha')\Phi_{m'}(\alpha'), \quad \text{or} \\ 0 &= \mathcal{K}_0(\alpha') - 1, \end{aligned} \tag{35}$$

which clearly simplifies solving the equation by at least an order of magnitude.

4. The cut-out disk

The Mestel disk has three rather bizarre properties: its mass is infinite, the surface density is not bounded in the centre, and it does not possess a natural time-scale. All these shortcomings can be cured by a relatively simple measure that was already employed by Zang. The idea is to immobilise a radially varying part of the mass, declaring it as, say, halo or bulge. Thus one keeps the simple structure of the disk, but cuts the singularity in the centre and possibly the long mass tail out to infinity as far as the Boltzmann equation is concerned. Also, the cut-out will break the self-similarity and thus introduces a time scale.

As an added benefit, the ratio of active to total mass and thus (at radially constant velocity dispersion) Toomre's stability parameter Q (Toomre 1964) becomes a function of the radius, opening the possibility to introduce a Q-barrier in the disk. In such a Q-barrier, incoming waves are refracted away from the centre and thereby evade their absorption at the inner Lindblad resonance. In common semi-phenomenological theories of spiral structure (Bertin et al. 1989), this Q-barrier serves as a part of a feedback loop, to be closed by, e.g., the corotation at which swing amplification (Toomre 1981) or related mechanisms would not only provide reflection but even amplification of wave packets.

In the shearing sheet (Goldreich & Lynden-Bell 1965; Julian & Toomre 1966), it is possible to show rigorously that the introduction of something quite similar to a Q-barrier leads to the emergence of stationary modes where no such modes are present in the classic border-less analyses (Fuchs 2001).

Technically, the immobilisation of the mass is done by multiplying the distribution function with a cut-out function $H(r)$, with $0 \leq H(r) \leq 1$, while keeping the Hamiltonian and thus the rotation curve unchanged. Since the main objective of our investigation of the cut-out disk was an assessment of the quality of our approximation, we used the cut-out function used by ERI (who in turn basically used what Zang had suggested). In our case, the introduction of the cut-out has the effect of eliminating the border terms in the L_z integration, but only if one uses Zang's doubly cut-out disk (i.e., the cut-out function tends to zero for both $r \to 0$ and $r \to \infty$). Zang's

family of cut-out functions is parametrised by two positive integers M and N, the cut-out indices, describing the steepness of the cut-outs in the outskirts and centre of the disk, respectively. Since both of these integers appear as the order of poles of the integrand in the L_z integration, and the residues at these poles get very messy indeed as their order increases, we settled for the case $M = N = 2$, resulting in a cut-out function

$$H(L_z) = \frac{L_z^2 L_c^2}{(L_z^2 + L_0^2)(L_z^2 + L_c^2)}, \tag{36}$$

where L_0 and L_c are two parameters qualifying the locations of the inner and outer cut-outs; following ERII, we usually use $L_0 = 1$ and $L_c = 10$.

It should be noted that for further analytic exploration of the cut-out disk, this cut-out is not favourable, since it still leads to a very cumbersome kernel. A cut-out of $H(L_z) = 2L_z L_0/(L_0^2 + L_z^2)$ results in a much more compact kernel and reproduces the overall behaviour the cut-out disk (due to its lower active surface density, the resulting disk has an even higher stability than one with Zang's cut-out, though).

Since the cut-out function (36) only depends on L_z, most of the calculations done for the self-consistent disk just carry over. The integration over L_z becomes somewhat easier, since there are no border terms, but this advantage is more than offset by the sheer length of the terms and the occurrence of higher-order poles. The resulting integral equation has an "almost" compact kernel, with no singularity on the diagonal and a relatively sharp decline of the kernel as α and α' approach infinity. Consequently, ERI's naive approach of simply Fredholm-discretizing the integral equation and cutting off at some appropriate α still works well. The only difference to ERI's method of solving this integral equation was that we found a spline-based quadrature to converge faster than the integration scheme proposed by them.

Skipping the details – which can be found in Demleitner (2000) –, let us briefly compare the stability of disks described with our approximation with Read's findings. Since ERII mostly report results for singly cut-out disks which are not favourable for an analysis using our tools, we give the values from Read's (1997) more comprehensive tables. Our disks become unstable to axisymmetric ($m = 0$) perturbations at $\sigma_{\min} = 0.274$; Read only gives values for $L_c = 100$, where we find $\sigma_{\min} = 0.349$ versus Read's $\sigma_{\min} = 0.327$. This agreement to within roughly 5% is remarkable considering the relatively high temperature of this disk. The higher stability of the disk with lower L_c is mostly due to its lower active surface density.

The one-armed ($m = 1$) modes investigated in ERII behave somewhat erratically, in that at vanishing growth rate and $\Re\omega \to 0$ there seem to be modes for almost arbitrary σ. That, however, is a consequence of the fact that ERII performed these computations without an outer cut-out. Introducing an outer cut-out, the disk seems to become stable at $\sigma = 0.57$, although clearly at that point

the assumption of a razor-thin disk becomes quite unrealistic and $v_c \gg \sigma$ is severely violated.

In the – at least historically – most interesting case of two-armed ($m = 2$) modes, our disks become unstable at $\sigma_{\min} = 0.214$ with a pattern speed $\Omega_p = \omega/m = 0.548$, whereas Read finds $\sigma_{\min} = 0.205$ and $\Omega_p = \omega/m = 0.542$. Thus, not only is the qualitative result that the $M = N = 2$ cut-out disk is stable with respect to two-armed perturbations reproduced, the quantitative agreement again is far better than one might expect given the rather severe assumptions that go into our model.

Read (1997) does not give stability limits for the $m = 3$ and $m = 4$ perturbations in doubly cut-out disks, but a comparison of the eigenvalue plots leads to the expectation that our results – $\sigma_{\min} = 0.141$ at $\Omega_p = 0.61$ for $m = 3$ and $\sigma_{\min} = 0.124$ at $\Omega_p = 0.588$ for $m = 4$ – would again reproduce the results of the more exact calculations quite well. The $m = 4$ stability limit has to be taken with a grain of salt, since our formalism reproduces modes at high velocity dispersion as $\omega \to 0$ in accordance with Read's findings. It is remarkable, though, that these modes are not quenched by an outer cut out, contrary to what is found for $m = 0$.

The properties of the modes computed from the kernels are, by and large, identical to those discussed by ERII, and our model even shows more subtle features like the rapid overtaking of eigenvalues when $m = 4$.

As to comparison to simulations, the fact that the expressions get very unwieldy for larger M and N is somewhat unfortunate. Sellwood & Evans (2001) point out that for the global mode to dominate over particle noise in simulations, it is highly desirable to maintain a high amplification, implying steep cut-outs and thus relatively large cut-out indices. The values chosen by Sellwood & Evans for their "standard" disk, $N = 4$ and $M = 6$, are beyond the reasonable range of applicability for our approach. However, they compute a doubly cut-out model that does not possess any global modes at $Q = 1$ (remember that Q refers to the axisymmetric stability of the self-consistent disk) and state that the $N = 1$ and $M = 2$ disk with $L_0 = 1$ becomes stable with respect to $m = 1$ perturbations below $L_c \approx 5$. The rotation curve they use in this model is falling but only very slightly so, so that a disk with a completely flat rotation curve should behave analogously. Indeed, ignoring the "rogue" eigenvalues (Demleitner 2000) – that become relevant in this model –, we find that even the $m = 1$ mode gets quenched below $L_c = 4.7$.

Summarising, our approximation does well in reproducing the results from the previous calculations for the cut-out disks. For now, the value of this insight is rather limited, since one still has to resort to numerical integration to solve the integral equation, and the numerics are computationally about as expensive as those in Zang's formalism, since the expression for the kernel is very large. Still, using a cut-out more suitable to our formalism results in more compact and manageable expressions that might facilitate a closer look at the interplay between the

Q-barrier, the various resonances, and density waves in a full-disk analytical model.

5. The self-consistent disk

The axisymmetric stability of the self-consistent disk can be determined by solving the algebraic Eq. (35) with a kernel still somewhat simpler than (34) that is the result of setting $m = 0$ in (22) and (23), yielding $\sigma_{\min} = 0.3860$ with $\alpha = 3.51$, within 2% of Read's (1997) values of $\sigma_{\min} = 0.3780$ and $\alpha = 3.46$.

Since the integral equation describing the self-consistent disk for non-axisymmetric modes is highly singular, Zang's formalism could not provide solutions for the full problem. However, it did yield a family of rather exotic modes characterised by $\omega = 0$ and consisting of a single logarithmic spiral – the neutral modes. It should be noted that these single logarithmic spirals are scale-free, a property that sets them apart from from any other disturbance of the form (11). In our formalism, they can be computed by solving the algebraic Eq. (35) with the kernel (34). The resulting behaviour again agrees very well with the findings of ERII and Zang and has to be called rather peculiar: one-armed perturbations cannot be stabilised at all, two-armed modes set in below $\sigma = 0.155$ (or $\sigma = 0.149$ according to Read 1997), three-armed ones are extremely stable and do not set in down to at least $\sigma = 0.05$, whereas four-armed neutral modes again cannot be stabilised at all.

The neutral modes come in pairs, i.e., if a mode with logarithmic wave number α exists, so does one with $-\alpha$. This is a consequence of the anti-spiral theorem (Lynden-Bell & Ostriker 1967), since these neutral modes are not limits of growing modes. The marginal modes in the cut-out disk and possibly existing modes with nonzero ω in the self-consistent disk, on the other hand, are limits of growing modes by the application of Landau's rule and thus not subject to the anti-spiral theorem.

ERII, from their perspective of analysing the entire family of physically viable power-law disks, briefly hypothesised that the stability properties might be related to the emergence of closed orbits of a given symmetry; these would, according to their reasoning, absorb power from perturbations and stabilise the disks when present. In our model, no closed orbits at all exist, since the ratio of epicyclic and orbital frequencies is fixed and irrational. As the model reproduces the stability properties found by ERII nonetheless, one has to look for a different explanation.

Turning to (34), one sees that the contribution from the terms except the one with the sum is fairly independent of m, whereas the sum does significantly change its value with m. The term to watch here is the $m'^2/(m'^2 - 2k^2)$. When $m' \approx \sqrt{2}k$, the corresponding summand is much larger than its neighbours since the ratio $I_k(x)$ and $I_{k+1}(x)$ basically is k (plus some constant) and everything else does not depend on k; when m is not too large, that

"resonant" summand will dominate the value of the entire series. The sign of this resonant summand is determined by the side relative to the nearest k on which $m' - \sqrt{2}k$ reaches its minimum – for the $m = 3$ series, the near-resonance is at $k = 2$ and thus one has a positive contribution with $m' - \sqrt{2}k = 0.17$ that enters into the kernel negatively (enhancing stability), whereas the resonance for the $m = 4$ occurs for $m' - \sqrt{2}k = -0.24$ and thus desta-bilises the disk. The next azimuthal harmonic that has an m close to a $\sqrt{2}k$ is 7, and it is again negative; not surprisingly, the $m = 7$ harmonic is quite unstable as well.

Unfortunately, k has no immediate meaning in physical terms, although (17) suggests that the k-th terms of the series in (34) should determine the response in the kth harmonic in the epicyclic angle w_1. In the denominator of the resonance term $m'^2 - 2k^2$, the factor $\sqrt{2}$ is simply the epicyclic frequency, so that we seem to see a resonance between the frequency with which a body in the disk en-counters a forcing, $m\Omega$, and a harmonic of the epicyclic frequency, $k\kappa$, which is a natural frequency with which a body wants to respond. If the forcing is near resonant with the response suggestion multiplied by the epicyclic frequency but has a slightly higher angular symmetry, the response is damped, whereas in the reverse case it is am-plified, a behaviour common in forced oscillations. This resembles the situation at a Lindblad resonance, with the difference that for one, the resonance is not exact (if it were, the series would evaluate to infinity), and, for sec-ond, it occurs on the entire disk.

Let us now turn to the general case of rotating modes. As has been hinted at above, the main discomfort with the notion of modes in the singular disk is that in its governing equation ω factors out, so that without further devices a disk is either stable or becomes unstable at all growth rates and pattern speeds at once. This seems unlikely enough to dismiss the possibility of such modes. On the other hand, this would imply that even a completely cold disk would be stable to all non-axisymmetric perturbations, which again would be quite an exotic situation.

Quite recently, Goodman & Evans (1999) tackled the problem employing the Jeans equations. The basic results of their work are that the Mestel disk's weirdness starts in its centre, and once one fixes boundary conditions there, the Mestel disk becomes quite tame. Without that bound-ary condition, however, the problem is not well-posed. By the requirement that the centre does not absorb or emit energy, they curb the two-dimensional continuum of solu-tions to a spectrum of one-dimensional continua, in which the ratio between growth rate and pattern speed is fixed and only the phase of ω remains unknown. Even this am-biguity can be removed by fixing the phase shift during the reflection of a wave at the disk's centre, thus reducing the spectrum to a countable set of discrete points; how-ever, there is nothing in the equilibrium model that would allow the definition of this phase change.

With a closed expression for the kernel of the Mestel disk, one is in a position to reconsider the stellar-dynamical equivalent of Goodman & Evans' problem using

the tools of the theory of singular integral equations: does the integral Eq. (30) admit continuous solutions, and if so, for what disk parameters?

The kernel (32) governing the self-consistent disk has a Cauchy-type singularity on the diagonal, implying that the operator $\int \mathcal{K}(\alpha, \alpha')\Phi(\alpha)\,d\alpha$ is not compact, so that the familiar Fredholm theory cannot be applied. For the treatment of integral equations of this type, Muskhelishvili (1953) and his coworkers have developed a theory. A nice summary is found in the more modern work of Polyanin & Manzhirov (1998). The following discussion is based on these works.

The integral Eq. (30) has the form

$$0 = a(\alpha')\Phi(\alpha') + \frac{1}{i\pi}\int_L \frac{\mathbf{K}(\alpha, \alpha')}{\alpha - \alpha'}\,\Phi(\alpha)\mathrm{d}\alpha, \tag{37}$$

where $\mathbf{K}$ is at least Hölder continuous, L is the real axis (integration over the real axis is henceforth implied for integral signs), and, of course, the integral is to be un-derstood as a Cauchy principal value. In the present case, $a(\alpha') = \mathcal{F}(\alpha') - 1$ and $\mathbf{K}(\alpha, \alpha') = i\pi(\alpha - \alpha')\mathcal{K}(\alpha, \alpha')$. The multiplication with $(\alpha - \alpha')$ ensures that the kernel is bounded on the diagonal, since the single zero in the de-nominator stemming from $\exp(2\pi(\alpha - \alpha')) - 1$ now cancels out. The only part of the kernel that could jeopardise the condition of Hölder continuity is the square root, but since only α^2 is present in the arguments of the square roots, $\mathbf{K}$ does indeed satisfy a Hölder condition.

It is convenient to rewrite (37) to the equivalent form

$$0 = a(\alpha')\Phi(\alpha') + \frac{1}{i\pi}b(\alpha')\int \frac{\Phi(\alpha)}{\alpha - \alpha'}\mathrm{d}\alpha$$
$$+ \frac{1}{i\pi}\int \mathbf{K}_r(\alpha, \alpha')\Phi(\alpha)\mathrm{d}\alpha, \tag{38}$$

where

$$b(\alpha) = \mathbf{K}(\alpha', \alpha') \quad \text{and}$$
$$\mathbf{K}_r(\alpha, \alpha') = \frac{\mathbf{K}(\alpha, \alpha') - \mathbf{K}(\alpha', \alpha')}{\alpha - \alpha'}. \tag{39}$$

After this manipulation, $\mathbf{K}_r$ is a compact operator, whereas

$$\hat{\mathbf{K}}\Phi := a(\alpha')\Phi(\alpha') + \frac{1}{i\pi}b(\alpha')\int \frac{\Phi(\alpha)}{\alpha - \alpha'}\,\mathrm{d}\alpha = 0, \tag{40}$$

known as the dominant or characteristic part of the inte-gral equation, decides whether or not the integral equation has a nontrivial solution. For our equation, $b(\alpha') = \mathcal{F}(\alpha')$.

To find a solution of (40), let us introduce an auxiliary function

$$\Psi(z) = \frac{1}{2\pi i}\int \frac{\Phi_m(\alpha)}{(\alpha - z)}\,\mathrm{d}\alpha, \tag{41}$$

defined for complex z. If $\Phi_m(\alpha)$ is continuous – this is henceforth assumed –, this function is piecewise analytic in the upper and lower half planes but may be discontinuous across the real axis. The merit of this auxiliary function is that it allows a transformation of the problem of the

solution of the integral Eq. (40) into a Riemann boundary value problem.

This can be done by applying the Sokhotski-Plemelj formulae, that, in their formulation for the real axis, state that when Φ and Ψ are linked by (41),

$$\Psi^+(\alpha') + \Psi^-(\alpha') = \frac{1}{i\pi} \int \frac{\Phi(\alpha)}{\alpha - \alpha'} \, d\alpha$$
$$\Psi^+(\alpha') - \Psi^-(\alpha') = \Phi(\alpha') \tag{42}$$

hold for all $\alpha' \in \mathbb{R}$. Here, $\Psi^+(\alpha')$ and $\Psi^-(\alpha')$ are the limiting values of $\Psi(z)$ as $z \to \alpha'$ from above and below the real axis, respectively.

Inserting (42) into (40) and collecting the terms yields

$$\Psi^+(\alpha') = \frac{a(\alpha') - b(\alpha')}{a(\alpha') + b(\alpha')} \Psi^-(\alpha') =: D(\alpha')\Psi^-(\alpha'). \tag{43}$$

Going back to the integral equation for the singular disk, one has

$$D(\alpha') = -1/(2\mathcal{F}(\alpha') - 1). \tag{44}$$

This represents a Riemann boundary value problem, asking (in the special case relevant here) for two functions Ψ^+ and Ψ^- that are analytic on the upper and lower half planes, respectively, and that satisfy $\Psi^+(\alpha') = D(\alpha')\Psi^-(\alpha')$ for all real α'.

For reasons that will become clear below, we for the time being demand that D has neither zeroes nor poles on the real axis. Comparing (35) and (44), one sees that this will be the case exactly when neutral modes are admitted by the disk in question, which strongly suggests that (30) has no solution in the regular case of a bounded and nonzero D, and at least one solution otherwise. With this concept mind, let us return to the analysis of the boundary value problem, for now keeping our attention on the case of a regular D.

The $D(\alpha')$ treated here is the restriction of an analytic function (in particular, by the assumption that it has no poles on the real axis), and thus is guaranteed to be Hölder continuous. Furthermore, $D(\alpha')$ is a positive real function for real α'. This ensures that its index, the increment of its argument over the integration contour,

$$\text{Ind}(D) = \frac{1}{2\pi} \int d\arg D(\alpha), \tag{45}$$

is zero. This in turn implies that $\ln D$ is a well-defined (single-valued) function. Then one can take the logarithm of both sides of (43), to arrive at

$$\ln \Psi^+(\alpha') - \ln \Psi^-(\alpha') = \ln D(\alpha'). \tag{46}$$

Using the second of the Sokhotski-Plemelj formulae (42), it follows that with

$$G(z) = \frac{1}{2\pi i} \int \frac{\ln D(\alpha)}{\alpha - z} \, d\alpha \tag{47}$$

one can write down a solution to the boundary value problem (43)

$$X^+(\alpha') = e^{G^+(\alpha')} \qquad X^-(\alpha') = e^{G^-(\alpha')}, \tag{48}$$

provided that the problem does have a solution.

Thus one has $D = X^+/X^-$, and the boundary value problem (43) can be written as

$$\frac{\Psi^+(\alpha')}{X^+(\alpha')} = \frac{\Psi^-(\alpha')}{X^-(\alpha')}. \tag{49}$$

Now, the functions $G^\pm$ by construction have no pole in their respective domains, and therefore $X^\pm$ have no zeroes. If $\Phi^\pm$ are defined via an integral as in (41), this implies that they are analytic in their respective half-planes. Thus, the functions on the two sides of (49) are analytic in their respective half-planes, and they are identical on the real axis. By Morera's theorem this implies that the right-hand side and the left hand side represent the same analytic function, $S(z)$, say. Now, what are the properties of $S(z)$?

We already pointed out that $S(z)$ is analytic on the entire complex plane. At infinity, one has $\Psi^\pm(\infty) = \pm\frac{1}{2}\Phi(\infty)$. Since we require any solution of the original integral Eq. (30) to vanish at infinity, one has $\Psi^\pm(\infty) = 0$ as well. Since $D \to 1$ for large α, G approaches 0 and $X^\pm$ tends to 1. This implies that S vanishes at infinity. Since S is an entire function, it follows from Liouville's theorem that S must be identically zero. By the first Sokhotski-Plemelj formula one now has $\Phi(\alpha') = S(\alpha)(X^+ + X^-)$, and hence the homogenous integral equation has no nontrivial solution. This is a special case of the general result that is proven somewhat more rigorously in the literature cited above: A homogenous Cauchy integral equation with an index of zero is unsolvable.

Loosening the conditions on D, we now allow it to have "poles" at some points α_i, which, for the sake of simplicity, are supposed to be poles of order one. The term "pole" has been put in quotes since it is not necessary that D is analytic in $\mathbb{C} \setminus \{\alpha_i\}$, only that $D(\alpha_i)$ behaves like $(\alpha' - \alpha_i)^{-1}$ at the exceptional points.

Writing

$$D(\alpha') = \frac{D'(\alpha')}{\prod_{i=1}^n (\alpha' - \alpha_i)}, \tag{50}$$

the boundary condition (43) takes the form

$$\Psi^+(\alpha') = \frac{D'(\alpha')}{\prod_{i=1}^n (\alpha' - \alpha_i)} \Psi^-(\alpha'), \tag{51}$$

where D' now satisfies the conditions we required of D in (43). Solving the boundary value problem for this D' and substituting the result back into (51), the equivalent of (49) now is

$$\frac{\Psi^+(\alpha')}{X^+(\alpha')} = \frac{\Psi^-(\alpha')}{X^-(\alpha') \prod_{i=1}^n (\alpha' - \alpha_i)}, \tag{52}$$

where X and G are again defined by (48) and (47), except that now D is replaced by D'. Let us again call the analytic function defined in the left and right sides of (52) $S(z)$.

By the right side of (52), S now may have poles of order 1 at α_i but still has to vanish at infinity. Applying the generalised Liouville theorem, it follows that any solution of the integral equation must be of the form

$$\Phi(\alpha) = \sum \frac{c_i}{\alpha - \alpha_i} \left(X^+(\alpha) + X^-(\alpha) \prod (\alpha - \alpha_i) \right). \quad (53)$$

Evidently, solutions of this type cannot vanish at ∞, since such solutions will at best converge to $X^-(\infty)$ times a nonzero constant as $\alpha' \to \infty$.

Thus, provided that all singularities of $S(z)$ are poles[1], the self-consistent disk admits no growing modes at all. The obvious explanation for this bizarre absence of growing modes even in very cold disks is that growing modes would introduce scales into the scale-free disk.

Stepping back to the mathematics for a moment, the fact that $\mathcal{F}$ is purely real on the real axis is crucial for that result, since even a minute complex contribution could provide for modes. For example, adding an $i\epsilon$ to $\mathcal{F}$ would not change the result that the index (45) of the problem (43) vanishes when $\mathcal{F}(\alpha') < 1/2$ for all α', so that there still would be no modes above the stability limit of the neutral modes. Below that stability limit, however, the index would be nonzero and the integral equation would admit nontrivial solutions.

Such an imaginary contribution would necessitate that the real part of the kernel (32) shows a singularity on the diagonal. We cannot offer an explanation how this might be linked to the scale-freeness of the disk at this point.

6. Conclusions

In this paper we have derived an analytic expression for the kernel of the Mestel disk. In the cut-out disk, it reproduces the behaviour described by Zang and ERII both qualitatively and quantitatively surprisingly well. It turns out that the expressions for the kernel are quite large in the cut-out case, at least when one uses the classic cut-out functions suggested by Zang, so the computational effort involved in finding modes is about as large with our formalism as it is with the older seminumerical scheme. On the other hand, it seems that with carefully selected cut-out functions, the kernel could be simple enough for further analytic work.

The kernel governing modes in the self-consistent disk is relatively compact. It allowed us to interpret the stability behaviour of neutral modes in terms of global resonances between the orbital (which coincides with the frequency of excitation in this case) and epicyclic frequencies. In the longstanding question of rotating modes we could establish that no rotating modes can exist above the stability limit of the neutral modes. It might seem surprising that rotating modes should know about the stability limit of the non-rotating modes, considering that the ratio of

orbital and epicyclic frequency is no longer independent of the radius for rotating modes. However, due to the lack of a length scale in the Mestel disk, any mode is arbitrarily close to a non-rotating one.

Even below the stability limit of the neutral modes, no rotating or growing modes exist in the perfectly scale-free disk under reasonable assumptions like continuity of the coefficient function in the expansion of the potential perturbation (11). This bizarre property is probably best explained by an inability of the disk to break its self-similarity and thus by a completely artificial feature. In conclusion, it seems to us that the self-consistent disk's peculiar properties make its investigation something of a pedagogical (Goodman & Evans 1999) exercise, whereas the cut-out disk might yet prove to be a valuable testbed for wave dynamics in a fairly realistic full-disk setting.

References

Abramowitz, M., & Stegun, I. 1972, Handbook of Mathematical Functions, U.S. Government Printing Office

Bertin, G., Lin, C. C., Lowe, S. A., & Thurstans, R. P. 1989, ApJ, 338, 78

Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton University Press)

Biviano, A., Girardi, M., Giuricin, G., Mardirossian, F., & Mezzetti, M. 1991, ApJ, 376, 458

Bogoliubov, N. N., & Mitropolski, Y. A. 1961, Asymptotic Methods in the Theory of Non-Linear Oscillations (Gordon and Breach)

Clutton-Brock, 1972, Ap&SS, 17, 292

Collett, J. L., Dutta, S. N., & Evans, N. W. 1997, MNRAS, 285, 49

Demleitner, M. 2000, A new approach to the problem of modes in Mestel disks, Dissertation, Universität Heidelberg

Evans, N. W., & Read, J. C. A. 1998a, MNRAS, 300, 83

Evans, N. W., & Read, J. C. A. 1998b, MNRAS, 300, 106

Fuchs, B. 2001, Density Waves in the Shearing Sheet, II: Global Modes, A&A, in preparation

Goldreich, P., & Lynden-Bell, D. 1965, MNRAS, 130, 125

Goodman, J., & Evans, N. W. 1999, MNRAS, 309, 599

Julian, W. H., & Toomre, A. 1966, ApJ, 146, 810

Kalnajs, A. J. 1971, ApJ, 166, 275

Kalnajs, A. J. 1977, ApJ, 212, 637

Landau, L. D. 1946, J. Phys. USSR 10, 25, reprinted in: Ter Haar, D. 1969, Men of Physics: L. D. Landau II (Pergamon Press)

Lynden-Bell, D., & Ostriker, J. P. 1967, MNRAS, 136, 293

Mark, J. W.-K. 1976, ApJ, 205, 363

Muskhelishvili, N. I. 1953, Singular integral equations, Noordhoff, based on the second Russian edition published in 1946

---

[1] This assumption is hard to prove, but generally true, which is why Pipkin (1991) calls it Gentlemen's Theorem Number 1 – a theorem no gentleman would question.

Persic, M., & Salucci, P. 1991, ApJ, 368, 60

Pipkin, A. C. 1991, A course on integral equations (Springer)

Polyanin, A. D., & Manzhirov, A. V. 1998, Handbook of integral equations (CRC Press)

Read, J. C. A. 1997, The Stability of Model Disk Galaxies, Ph.D. Thesis, University of Oxford

Toomre, A. 1964, ApJ, 139, 1218

Toomre, A. 1977, ARA&A, 15, 437

Toomre, A 1981, in The structure and evolution of normal galaxies (Cambridge University Press), 111

Zang, T. A. 1976, The Stability of a Model Galaxy, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA