



Gas flow in barred potentials - II. Bar-driven spiral arms

Mattia C. Sormani, James Binney, John Magorrian

► To cite this version:

Mattia C. Sormani, James Binney, John Magorrian. Gas flow in barred potentials - II. Bar-driven spiral arms. Monthly Notices of the Royal Astronomical Society, 2015, 451, pp.3437-3452. <10.1093/mnras/stv1135>. <insu-03644879>

HAL Id: insu-03644879

<https://insu.hal.science/insu-03644879v1>

Submitted on 28 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Gas flow in barred potentials – II. Bar-driven spiral arms

Mattia C. Sormani,¹★ James Binney¹ and John Magorrian^{1,2}

¹*Rudolf Peierls Centre for Theoretical Physics, 1 Keble Road, Oxford OX1 3NP, UK*

²*Institut d'Astrophysique de Paris, 98bis boulevard Arago, F-75014 Paris, France*

Accepted 2015 May 15. Received 2015 May 15; in original form 2015 April 18

ABSTRACT

Spiral arms that emerge from the ends of a galactic bar are important in interpreting observations of our and external galaxies. It is therefore important to understand the physical mechanism that causes them. We find that these spiral arms can be understood as kinematic density waves generated by librations around underlying ballistic closed orbits. This is even true in the case of a strong bar, provided the librations are around the appropriate closed orbits and not around the circular orbits that form the basis of the epicycle approximation. An important consequence is that it is a potential's orbital structure that determines whether a bar should be classified as weak or strong, and not crude estimates of the potential's deviation from axisymmetry.

Key words: ISM: kinematics and dynamics – galaxies: kinematics and dynamics.

1 INTRODUCTION

Many grand design barred spirals exhibit spiral arms starting at and extending out from the end of the bar. How are these spiral arms generated? The natural interpretation is that they are driven by the bar. The spiral arms and the bar must also be rotating at the same pattern speed if the starting points of the arms always coincide with the bar's ends. Such spiral arms are relevant for the interpretation of observations of both our Galaxy and external galaxies. For example, it is likely that some features present in spectral-line emission of atomic and molecular gas in the inner Galaxy, for instance the '3 kpc arm', are produced by such arms (e.g. Dame & Thaddeus 2008). It is therefore important to have a physical understanding of why such spiral arms arise.

On the theoretical side, the possibility of bar-driven spiral arms is now well established. The first investigations used analytical methods and were carried out by Feldman & Lin (1973) and Lin & Lau (1975). Then, in the following years, Sanders & Huntley (1976), Sanders (1977) and Huntley, Sanders & Roberts (1978) performed pioneering numerical experiments and discussed the physical mechanism responsible for the generation of the spiral arms in terms of closed orbits. Their work was later extended independently by Wada (1994) and Lindblad & Lindblad (1994), and more recently by Piñol-Ferrer, Lindblad & Fathi (2012). These authors constructed phenomenological analytical models under the epicyclic approximation with the purpose of understanding the spiral arms. In their models gas parcels follow weakly oval orbits, whose major axes orientation changes with radius giving rise to kinematic density waves a-la Lindblad, and hence to spiral arms.

An alternative viewpoint which does not consider the spiral arms as density waves has also been discussed in a series of papers by Romero-Gómez et al. (2006, 2007), Athanassoula et al. (2009a), Athanassoula, Romero-Gómez & Masdemont (2009b), Athanassoula et al. (2010). Their theory, which is more directly applicable to stars than to gas, is based on the observation that orbits in the vicinity of unstable Lagrangian points can be trapped into invariant manifolds whose morphology can reproduce the spiral arms.

Many authors have noted the presence of bar-driven spiral arms in simulations, and discussed them in a more or less descriptive way (see e.g. Athanassoula 1992; Englmaier & Gerhard 1999; Bissantz, Englmaier & Gerhard 2003; Rodriguez-Fernandez & Combes 2008). For a recent review see also section 2.3 of Dobbs & Baba (2014). Note that in all works cited above the gas is driven by an external potential generated by the bar and is not self-gravitating, and therefore the spiral arms are not spiral density waves in the sense of Lin & Shu (1964). We do not discuss self-gravity in this work.

The viewpoint that spiral arms can be understood as density waves is nowadays often assumed, but has never been tested with sufficient detail. On modern computers, it is very cheap to run simulations able to test the predictions and the limits of the phenomenological models cited above. Moreover, the literature only addresses the weak bar regime under the epicyclic approximation, and does not discuss how the picture should be extended to the strong bar case.

In this paper we investigate the physical mechanism responsible for the generation of the spiral arms and in particular we resume the discussion of how they can be understood as kinematic density waves. The structure is as follows. In Section 2 we briefly review previous work aimed at understanding in a phenomenological fashion the spiral arms in the epicyclic approximation. Then we run grid-based, isothermal, non-self gravitating 2D hydrodynamic simulations in an externally imposed rigidly rotating barred

*E-mail: mattia.sormani@physics.ox.ac.uk

potential, addressing both the weak and the strong bar case. The numerical methods employed are explained in Section 3. In Section 4 we compare the results of the simulations with the phenomenological models available in the literature. In Section 5 we discuss in more detail the weak bar case, and in Section 6 how the results for the weak bar case should be extended to the strong bar case. Finally, we summarize our findings in Section 7.

2 REVIEW OF PREVIOUS WORK

In this section we briefly review previous work aimed at understanding the bar-driven spiral arms in the epicycle approximation. In Section 2.1 we discuss the equations describing ballistic closed orbits in the epicycle approximation (see e.g. section 3.3.3 in Binney & Tremaine 2008). In Section 2.2 we discuss how various authors have modified the ballistic equations in order to describe gaseous parcels, and the phenomenological models based on these equations. These phenomenological models will be in later sections compared with the results of hydro-simulations.

2.1 Ballistic closed orbits in the epicycle approximation

Consider a rigidly rotating external potential of the form

$$\Phi(R, \theta) = \Phi_0(R) + \Phi_1(R, \theta), \quad (1)$$

where Φ_0 is an axisymmetric potential and Φ_1 is a small but otherwise arbitrary perturbation. The potential is assumed to be rigidly rotating at pattern speed Ω_p . R, θ are polar coordinates in the rotating frame, with $\theta = 0$ corresponding to the positive horizontal axis, and θ increasing clockwise. The equation of motion for a ballistic particle in this potential is

$$\ddot{\mathbf{x}} = -\nabla\Phi + \Omega_p^2 \mathbf{x} - 2\Omega_p (\hat{\mathbf{e}}_z \times \dot{\mathbf{x}}), \quad (2)$$

where $\hat{\mathbf{e}}_z$ is the unit vector perpendicular to the plane. The first term on the RHS represents gravitational forces, the second term is the centrifugal force and the third is the Coriolis force.

We want to find closed orbits in the above potential under the epicycle approximation. In the absence of the perturbation Φ_1 the potential is axisymmetric and the only stable closed orbits are circular orbits. These circular orbits will be described in polar form as

$$R(t) = R_0, \quad (3)$$

$$\theta(t) = \theta_0(t) = \Omega_f t, \quad (4)$$

where

$$\Omega_f = \Omega - \Omega_p, \quad (5)$$

and $\Omega = \Omega(R_0)$ is the angular velocity for circular motion at radius R_0 in the potential Φ_0 . Now consider the situation in the presence of the small perturbation Φ_1 . We want orbits that are closed in the rotating frame.¹ We expect that far from resonances closed orbits will be weak oval deformations of the circular orbits found in Φ_0 . We look for closed orbits of the form

$$R(t) = R_0 + R_1(t), \quad (6)$$

$$\theta(t) = \theta_0(t) + \theta_1(t), \quad (7)$$

¹ Recall that the notion of closure of an orbit is frame dependent. Orbits that are closed in the inertial frame are in general not closed in the rotating frame and vice versa.

where all quantities with subscript 1 are to be considered small. By expressing equation (2) in polar components, then substituting equations (6), (7) and (1) into it and approximating to first order in quantities with subscript 1, the equations of motion take the following form:

$$\ddot{R}_1 + \left(\frac{d^2\Phi_0}{dR^2} - \Omega^2 \right)_{R_0} R_1 - 2R_0\Omega\dot{\theta}_1 = - \left(\frac{\partial\Phi_1}{\partial R} \right)_{R_0, \theta_0(t)}, \quad (8)$$

$$\ddot{\theta}_1 + 2\Omega \frac{\dot{R}_1}{R_0} = - \frac{1}{R_0^2} \left(\frac{\partial\Phi_1}{\partial\theta} \right)_{R_0, \theta_0(t)}, \quad (9)$$

where the derivatives of Φ_1 are evaluated along the unperturbed trajectory $[R_0, \theta_0(t)]$ and are therefore to be considered given functions of time. Exploiting the fact that $\theta_0(t)$ is a linear function of time, equation (9) can be integrated immediately to obtain $\dot{\theta}_1$, which can then be substituted into equation (8). This gives

$$\ddot{R}_1 + \kappa_0^2 R_1 = f_1(t), \quad (10)$$

where

$$\kappa_0^2 = \left(R \frac{d\Omega^2}{dR} + 4\Omega^2 \right)_{R_0}, \quad (11)$$

$$f_1(t) = - \left[\left(\frac{\partial\Phi_1}{\partial R} \right) + \frac{2\Omega}{\Omega_f R_0} \Phi_1 \right]_{R_0, \theta_0(t)}. \quad (12)$$

Equation (10) is the equation of a forced harmonic oscillator. The natural frequency of this oscillator is κ_0 , the usual epicycle frequency that can be calculated from Φ_0 . The driving force is $f_1(t)$, and as one would expect it is due to the perturbing potential Φ_1 and reduces to zero when this is turned off. $f_1(t)$ is a periodic function of time. If we expand it in its Fourier components with respect to time, we will find the same components contained in the Fourier expansion of Φ_1 with respect to θ . This is not surprising, as the perturbed potential encountered by the guiding centre of the ballistic particle is given by $\Phi_1(R_0, \theta_0(t))$ where $\theta_0(t)$ increases linearly with time.

The general solution of equation (10) could in principle be easily written down. For each solution, we could eliminate t from equations (6) and (7) to find the orbit in the form $R = R(\theta)$. Not all solutions of equation (10) correspond in general to closed orbits in the rotating frame. Some particular solutions describe closed orbits, while the others in general describe non-closed loop orbits. The first type of solution are the ballistic closed orbits in the epicycle approximation.

For concreteness, let us now consider the particular form of the potential used by Sanders (1977) and Wada (1994):

$$\Phi_0(R) = - \frac{av_0^2}{\sqrt{R^2 + a^2}}, \quad (13)$$

$$\Phi_1(R, \theta) = \Phi_b(R) \cos(2\theta), \quad \Phi_b(R) = -\epsilon \frac{(av_0 R)^2}{(R^2 + a^2)^2}, \quad (14)$$

where a, v_0, ϵ are constant parameters. The axisymmetric part Φ_0 is a Kuzmin–Toomre potential, and the perturbation Φ_1 was introduced by Sanders (1977). Following Wada (1994), we choose $a = 1$ kpc and $v_0 = 100$ km s^{−1} and a pattern speed of $\Omega_p = 10$ km s^{−1} kpc^{−1}. Different values of ϵ will be considered in the following. Fig. 1 shows the circular speed curve and the behaviour of $\Omega - n\kappa_0/2$, from which we can determine the location of the resonances, for this choice of values of the parameters. For this potential, equation (10) becomes

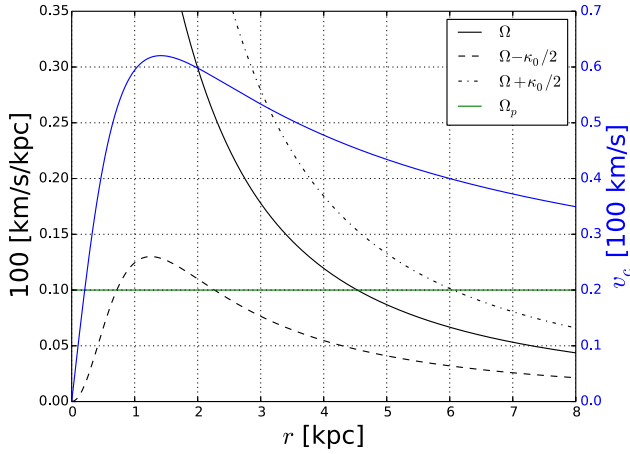


Figure 1. Blue curve: circular speed curve for the potential in equation (13) with values of parameters indicated in the text. Black curves: behaviour of $\Omega - \kappa_0/2$ for the same potential. The horizontal green line indicates the assumed value of the pattern speed.

$$\ddot{R}_1 + \kappa_0^2 R_1 = f_w \cos(2\Omega_f t), \quad (15)$$

where

$$f_w(R_0) = - \left[\left(\frac{d\Phi_b}{dR} \right) + \frac{2\Omega}{\Omega_f R_0} \Phi_b \right]_{R_0}. \quad (16)$$

The forcing is now a periodic function of time with frequency $2\Omega_f$. By solving (15), it is found that ballistic closed orbits have the major axis always either perpendicular or parallel to the major axis of the bar.² The orientation changes abruptly at each inner Lindblad resonance (ILR), at Corotation (CR), and at each outer Lindblad resonance (OLR). Note however that the above analysis is not valid in the vicinity of these points: at all Lindblad resonances the natural frequency of the oscillator $2\Omega_f$ tends to the forcing frequency κ_0 , while at CR the forcing f_w becomes infinite. The orientation of the ballistic closed orbits will be important in the next subsection.

2.2 Gaseous closed orbits in the epicycle approximation: phenomenological models

In the previous section we have seen that for the particular potential given by equations (13) and (14) the significant ballistic closed orbits are always oriented either perpendicularly or parallel to the major axis of the bar, with abrupt changes of orientation at each resonance. However, the discussion was only valid for ballistic particles. How should we modify it to describe the motion of a gaseous parcel which is part of a continuous fluid? Sanders & Huntley (1976) suggested that in the gaseous case the orientation shift between horizontally and vertically aligned orbits happens gradually, and that

² These correspond to orbits obtained by taking the limit $\lambda \rightarrow 0$ of the solution that neglect transients of equation (17), which are also the orbits with $C_1 = 0$ in the notation of section 3.3.3 in Binney & Tremaine (2008). It should be noted that other closed orbits solutions are possible at special radii, i.e. when the ratio $\kappa_0/(\Omega - \Omega_p)$ happens to be a rational number, and correspond to orbits with $C_1 \neq 0$. However these are isolated cases that are not part of a family of closed orbits parenting non-closed orbits, and therefore are not of interest in this paper.

this gives rise to kinematic spiral arms a-la Lindblad. Moreover, they noted that the addition of a dissipation term to equation (15),

$$\ddot{R}_1 + 2\lambda \dot{R}_1 + \kappa_0^2 R_1 = f_w \cos(2\Omega_f t), \quad (17)$$

can remove the divergence of the radial oscillation amplitude at the Lindblad resonances, and that the solution of equation (17) where transients³ are neglected always describes a closed orbit. They also noted that at each ILR, the closed orbit solution so obtained has a major axis inclined at $\pi/4$ independently of the value of λ . Since $\pi/4$ is halfway between horizontally and vertical elongated orbits, this corroborated the conjecture that in the gaseous case we expect a gradual shift from vertically elongated to horizontally elongated when we cross a Lindblad resonance, which is the key to the generation of spiral arms in their model.

Wada (1994) and independently Lindblad & Lindblad (1994), Pinol-Ferrer et al. (2012) developed this idea further by taking the dissipation term more seriously, considering it as a phenomenological model that describes the trajectory of gaseous particles over extended regions. The approaches of Wada (1994) and Lindblad & Lindblad (1994) differ slightly in the way they implement the dissipation term. The phenomenological model of Wada (1994) is based exactly on equation (17). This amounts to add a dissipation term for the motion of the particle in the radial direction only. Lindblad & Lindblad (1994) added a dissipation term that is proportional to the difference between the velocity of the particle and the local circular speed in the underlying axisymmetric potential. Here, we consider in more detail the variant of Wada (1994), but the results of the Lindblad & Lindblad (1994) variant are qualitatively similar, and a more detailed account of the results using this variant can be found in Piñol-Ferrer et al. (2012).

Let us now briefly review Wada's model. The solution of equation (17) that excludes transient terms yields closed orbits of the following form:

$$R(\theta) = R_0 + A \cos 2(\theta - \phi), \quad (18)$$

where the amplitude is

$$A = \frac{|f_w|}{F}, \quad (19)$$

with

$$F = \sqrt{[\kappa_0^2 - (2\Omega_f)^2]^2 + (4\lambda\Omega_f)^2}. \quad (20)$$

The phase ϕ in (18) is given by the solution of

$$\begin{aligned} \sin 2\phi &= \frac{4\lambda|\Omega_f|}{F} \text{sign}(f_w), \\ \cos 2\phi &= \frac{\kappa_0^2 - (2\Omega_f)^2}{F}, \end{aligned} \quad (21)$$

which lies in the range $-\pi \leq 2\phi \leq \pi$. It encodes the information about the orientation of the major axis of the closed orbit, and is the inclination of the major axis of the orbit with respect to the horizontal axis. From equations (21) we see that the quadrant to which 2ϕ belongs is determined by (i) the sign of f_w , which flips at CR, and (ii) by the sign of $\kappa_0^2 - 2\Omega_f^2$, which flips at each Lindblad resonance. The phase ϕ in equation (18) is determined for all orbits

³ By transients we mean the part of the solution that vanishes in the limit $t \rightarrow \infty$. In this case we have a driven and damped harmonic oscillator. The general solution is the sum of a decaying exponential and an oscillatory term, and the decaying exponential is the transient.

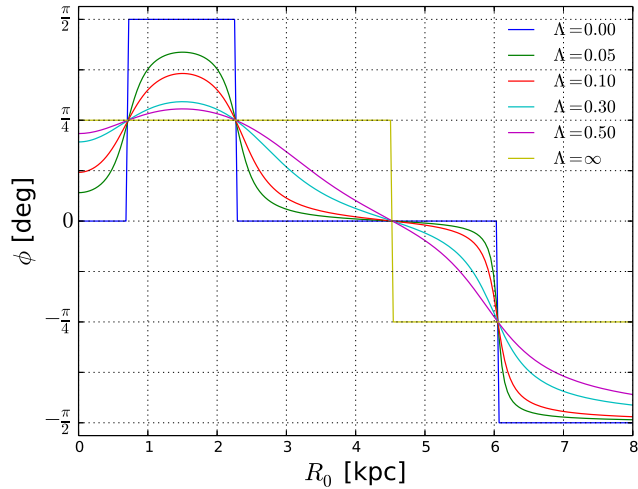


Figure 2. Predictions of Wada’s model for the phase ϕ as a function of guiding-centre radius R_0 for the potential in equation (13). ϕ represents the orientation of the major axis of a closed orbit. Curves corresponding to different values of the damping parameter Λ are shown. The possible values allowed by the phenomenological model are enclosed between the blue and yellow lines.

once $\lambda = \lambda(R_0)$ is given at each radius. Since λ has the dimension of a frequency, it is convenient to express it as a multiple of a characteristic frequency at that radius. Following Wada (1994), we express λ as a multiple of the epicycle frequency at that radius, namely

$$\lambda = \Lambda \kappa_0. \quad (22)$$

Fig. 2 reproduces fig. 2 of Wada (1994) and shows his model predictions for the phase ϕ for different constant values of Λ . In the limit $\Lambda \rightarrow 0$ we recover the ballistic case, while in the limit $\Lambda \rightarrow \infty$ all orbits are circular. The curves $\Lambda = 0$ and $\Lambda = \infty$ bound the possible values that the ϕ can assume in its model. Note that this plot does not depend on the strength of the bar potential ϵ , because the phase ϕ does not depend on the magnitude of f_w but only on its sign.

A is the amplitude of radial oscillations. Fig. 3 shows the amplitude A of the radial oscillations predicted by the model for various values of Λ and two different values of ϵ . As can be seen from this figure, A always diverges at CR, regardless of the value of Λ , but diverges at the Lindblad resonances only in the absence of the damping term. Note also that, for given ϵ , away from resonances the value of A is limited from above: the maximum value that A can reach at given R_0 is limited by its value for $\Lambda = 0$. Finally, note that for a given potential, this theory has only one adjustable parameter at each radius, the damping λ . Both ϕ and A at each radius depend on a single phenomenological parameter in this theory.

Using the theory outlined above, it is possible to construct explicit models of the spiral arms. Fig. 4, top panels, reproduces the model shown in fig. 4 of Wada (1994). It shows a nested sequence of closed orbits for values of the parameters $\epsilon = 0.05$ and $\Lambda = 0.05$. The bottom panels show another model, for a weaker bar $\epsilon = 0.005$ and $\Lambda = 0.02$.

3 NUMERICAL METHODS

In our simulations, we assume that the gas is a fluid governed by the Euler equations complemented by the equation of state of an

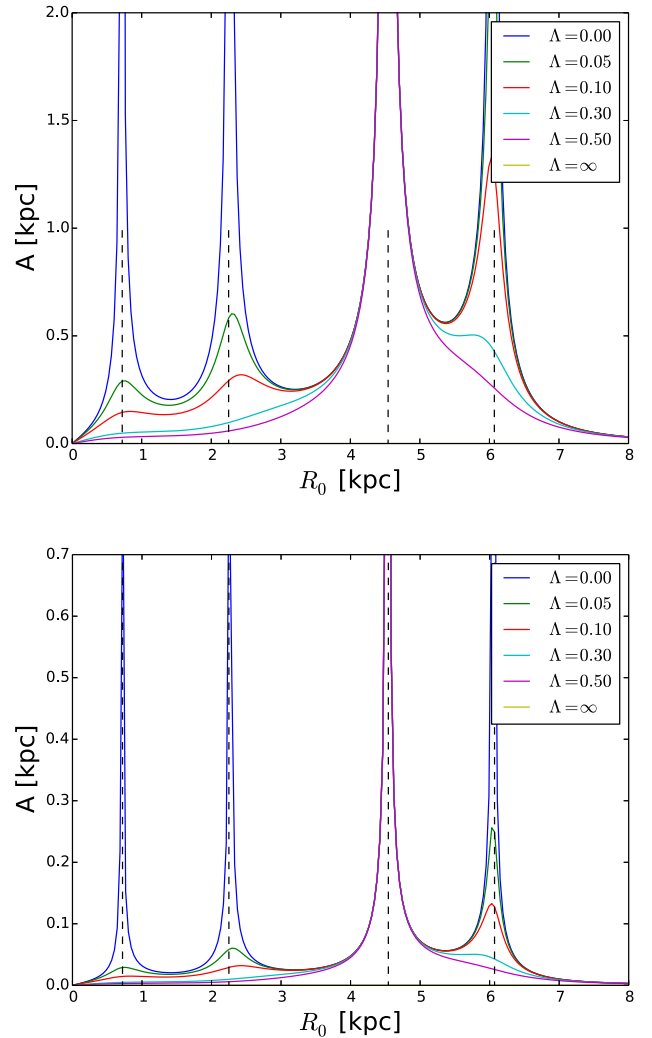


Figure 3. Amplitude of the radial oscillations predicted by Wada’s model for two different values of the bar strength ϵ . The top panel is for $\epsilon = 0.05$, and the bottom panel is for $\epsilon = 0.005$. In each panel, curves corresponding to different values of the damping parameter Λ are shown. In absence of a damping term, $\Lambda = 0$, the amplitude diverges at all Lindblad resonances and at corotation. The presence of a damping term makes the amplitude finite at the Lindblad resonances, but not at corotation.

isothermal ideal gas. Then we run two-dimensional hydrodynamical simulations in an externally imposed, rigidly rotating barred potential. The output of each simulation consists in snapshots of the velocity and surface density distributions $\rho(\mathbf{x})$ and $\mathbf{v}(\mathbf{x})$ at chosen times.

We use a grid-based, Eulerian code based on the second-order flux-splitting scheme developed by van Albada, van Leer & Roberts (1982) and later used by Athanassoula (1992), Weiner & Sellwood (1999) and others to study gas dynamics in bar potentials. We used the same implementation of the code as was used by Sormani, Binney & Magorrian (2015), which also implements the recycling law of Athanassoula (1992).

We used a grid $N \times N$ to simulate a square 16 kpc on a side. N depends on the resolution of the simulation. The initial conditions of the runs are as follows. We start with gas in equilibrium on circular orbits in an axisymmetrized potential and, to avoid transients, turn on the non-axisymmetric part of the potential gradually. We use

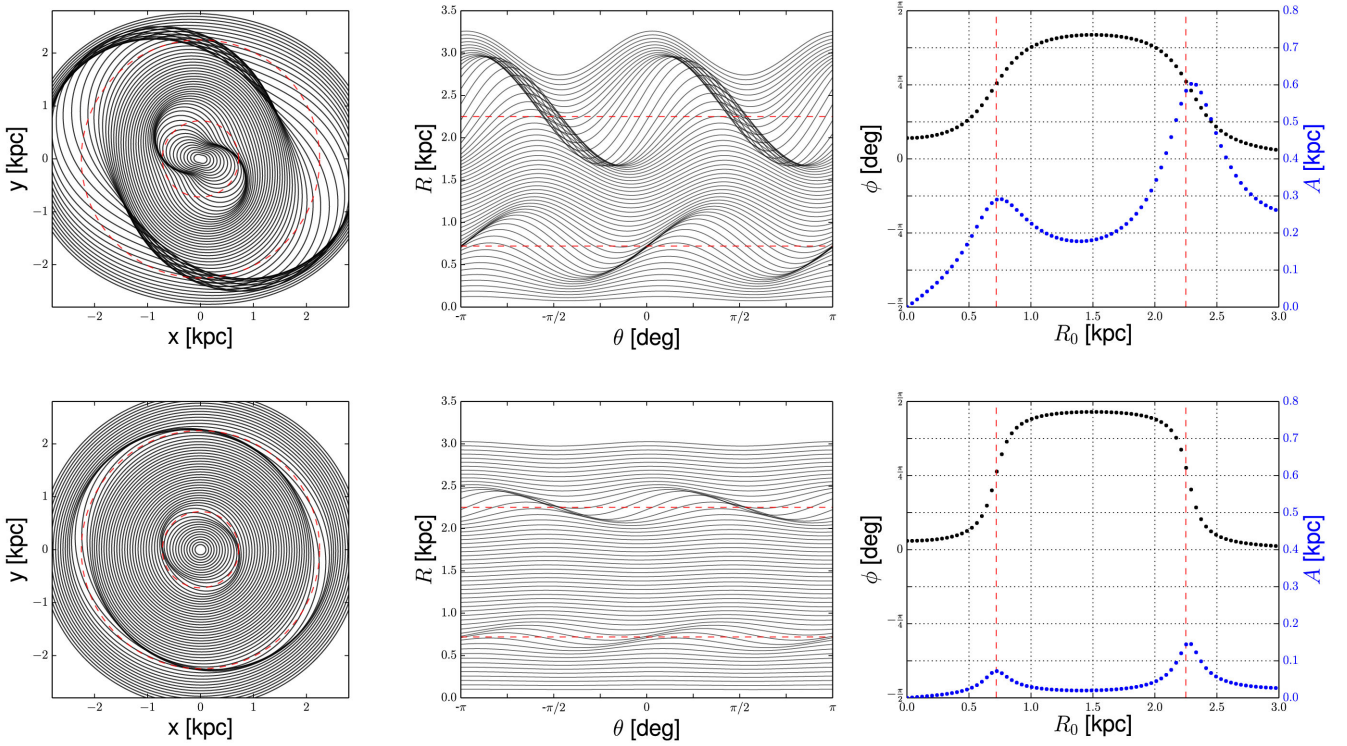


Figure 4. Top: reproduction of Wada (1994) predictions for a bar strength of $\epsilon = 0.05$. The damping parameter is $\Lambda = 0.05$. Bottom: another model produced using Wada (1994) theory, for $\epsilon = 0.005$ and $\Lambda = 0.02$.

outflow boundary conditions: gas can freely escape the simulated region, after which it is lost forever. The potential well is sufficiently deep, however, to prevent excessive quantities of material to escape.

The recycling law introduces a term in the continuity equation that was originally meant to take into account in a simple way the effects of star formation and stellar mass loss. The equation governing this process is

$$\frac{\partial \rho}{\partial t} = \alpha(\rho_0^2 - \rho^2), \quad (23)$$

where $\alpha = 0.3 \text{ M}_\odot \text{ pc}^{-2} \text{ Gyr}^{-1}$ is a constant and ρ_0 is the initial surface density, which is taken to be $\rho_0 = 1 \text{ M}_\odot \text{ pc}^{-2}$. In practice, the only effect of the recycling law is to prevent too much gas accumulating in the very centre and to replace gas lost at the boundary due to the outflow boundary conditions. It does not affect the morphology of the results, so our results do not change if we disable the recycling law.

4 PHENOMENOLOGICAL MODELS VERSUS HYDRO-SIMULATIONS

What happens if we run simulations in the potential of equations (13) and (14)? Do the results of a hydro-simulation resemble the predictions of the phenomenological models? Fig. 5 answers this question. It shows the results of a simulation in which $\epsilon(t)$ is a slowly changing function of time. At $t = 0$ we have $\epsilon = 0$, then ϵ increases linearly with time until it reaches Wada's value of $\epsilon = 0.05$ at $t = 12.3 \text{ Gyr}$, and for later times it is kept constant at this value. The potential evolves so slowly that the gas flow configuration evolves almost adiabatically, adjusted to the instantaneous underlying potential. In other words, at each time t the gas configuration is almost the same as the steady-state configuration that

would be obtained by freezing the potential and then waiting for the gas to settle down into a steady state.

At early times, when ϵ is very small, the results of the hydro-simulation do qualitatively resemble the predictions of the phenomenological models, as can be seen by comparing the top row in Fig. 5 with Fig. 4. At later times, when ϵ is greater, this is not true: for example panels in the bottom row do not look like Wada (1994) predictions. Inspection of streamlines shows that for very small values of ϵ the gas is flowing on almost circular orbits, as the epicycle approximation requires, but when ϵ becomes larger the gas flows instead on very horizontally elongated orbits, and the epicycle approximation is not valid anymore. How can we explain this behaviour? Why does the gas flow on very elongated orbits despite the fact that the perturbation is apparently very small? As we shall now see, the key to answer these questions lies in the orbital structure of the underlying potential.

Fig. 6 shows how the orbital structure of the underlying potential changes as ϵ is varied. Each row refers to a particular value of ϵ . Let us first consider the top row, which shows the case $\epsilon = 0$, when the bar perturbation is turned off and the potential is exactly axisymmetric. Consequently, orbits that close in our rotating frame are either circular or they are orbits for which the precession frequency happens to coincide with our chosen ‘pattern speed’, which is actually of no physical significance when $\epsilon = 0$. The top left panel of Fig. 6 shows several families of closed orbits: each dot represents a closed orbit in terms of its value of the Jacobi constant E_j and the coordinate at which the orbit cuts the vertical axis. All the orbits shown in this diagram (known as characteristic diagram; see Contopoulos & Grosbol 1989) have the property of cutting the vertical axis at right angles. The line of dark dots shows the circular orbits. Since for $\epsilon = 0$ the potential is axisymmetric, these are the only stable closed orbits.

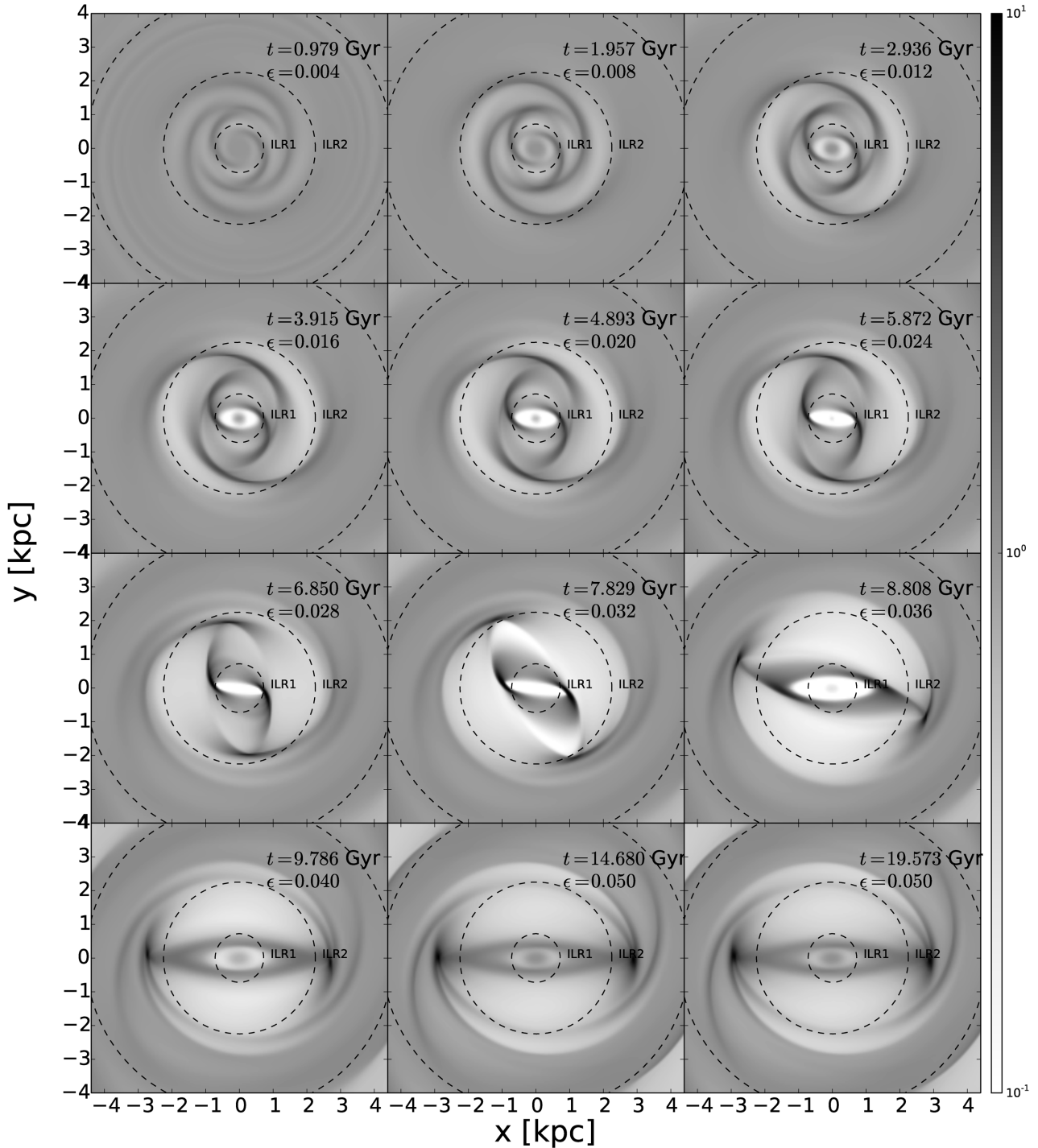


Figure 5. Time sequence for a simulation in the potential defined by equations (13) and (14). Each panel shows the density map at a particular time. The bar strength ϵ increases slowly and linearly with time until it reaches the value of $\epsilon = 0.05$, and is then frozen. The gas adjusts almost adiabatically to the instantaneous underlying potential. The sound speed is $c_s = 2.5 \text{ km s}^{-1}$ and the spatial resolution is $dx = 20 \text{ pc}$. The dotted circles indicate the two ILRs and the CR. In each panel, the evolutionary time t and the instantaneous value of ϵ are shown. The colour bar is in units of $M_\odot \text{ pc}^{-2}$.

The red crosses in the top left panel of Fig. 6 show eccentric orbits that close in the rotating frame. Such orbits exist only between the two ILRs. They spring out from the sequence of nearly circular orbits at a radius of the innermost inner Lindblad resonance ILR1, and then merge into it again at a bigger radius corresponding to the other inner Lindblad resonance, ILR2. The red crosses that are above the dark line are elongated perpendicularly to the bar, while

those below it are elongated along the bar. In fact, since the potential is axisymmetric, the red cross above the dark line and that vertically below it represent orbits that differ only in a 90° rotation of the major axis. Since the potential is axisymmetric, it is actually possible to find equivalent orbits for any orientation, but these are not shown in the diagram as they do not cut the vertical axis at right angles. The central panel of Fig. 6 shows how the orbits appear in the xy plane.

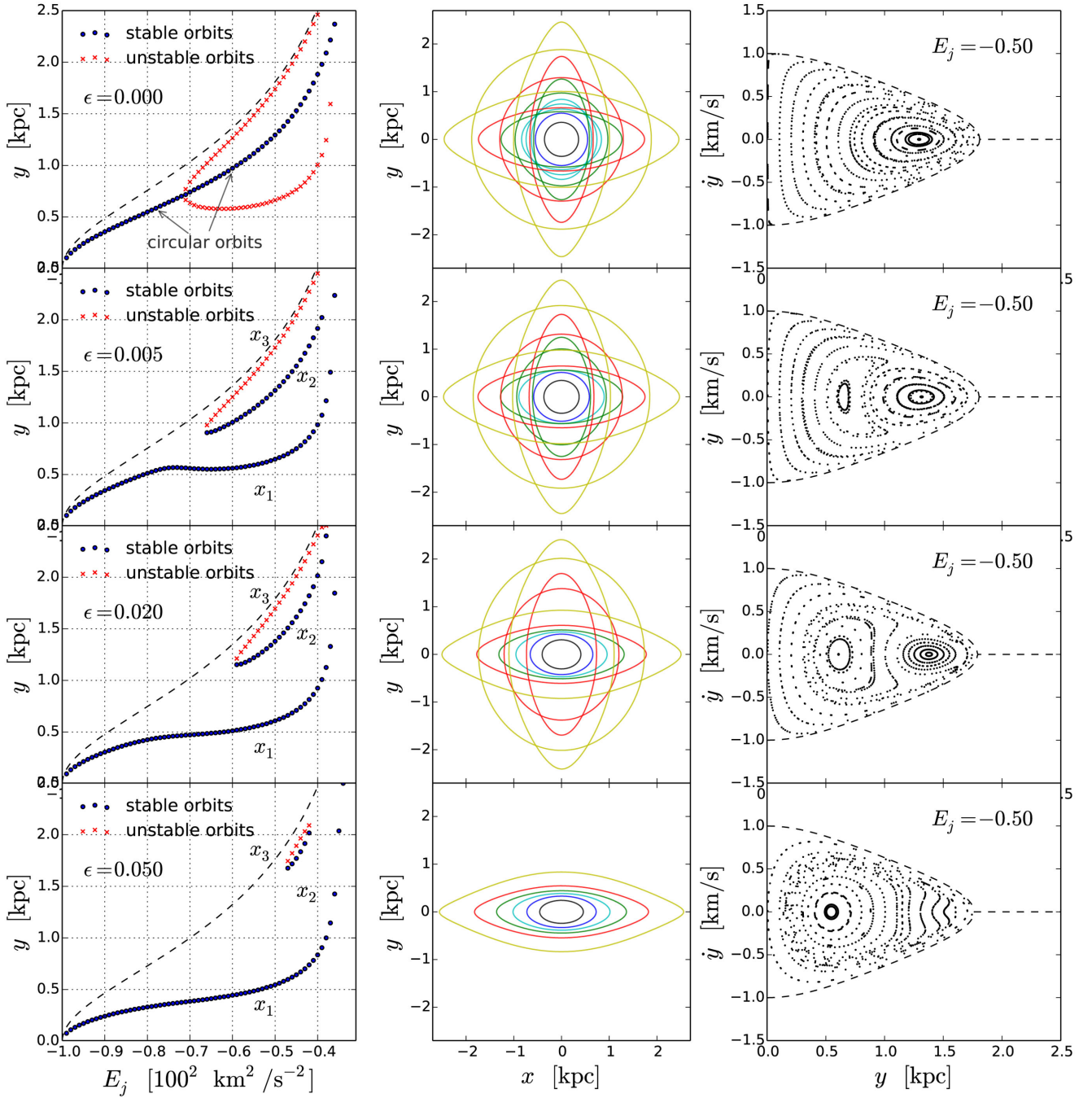


Figure 6. Orbital structure of the potential given by equation (13) for different values of ϵ . Rows from top to bottom show the cases $\epsilon = 0.0, 0.005, 0.02, 0.05$, respectively. In the left column, the Jacobi constant E_j of a closed orbit is plotted against the value at which the orbit cuts the vertical axis in potentials with different values of ϵ . Each dot represents a closed orbit, and all closed orbits cut the vertical axis with purely horizontal velocity. The middle column shows the closed orbits in xy space, with orbits of the same colour having the same energy. The orbits shown are those corresponding to values of E_j equal to $\{-0.9, -0.8, -0.7, -0.6, -0.5, -0.4\}$. The vertical slices of the grid in the left panels are drawn at these energies. The correspondence between energy and colours can be identified from the bottom panel in the central column: here greater energy (i.e. less negative) corresponds to bigger orbits, and the same colour coding applies to other panels. The right column shows surfaces of section for $E_j = -0.5$.

These orbits are coloured differently according to their value of the Jacobi constant E_j .

We now consider what happens when a small perturbation is turned on, so the potential is no longer exactly axisymmetric. The left panel of the second row of Fig. 6 shows that for $\epsilon = 0.005$, a bifurcation is present approximately at the location of ILR1. The

line of dark points that marks the circular orbits in the axisymmetric case is now split. The part inside ILR1 merged with what used to be the orbits elongated along the bar in the axisymmetric case (which have become stable) to form a continuous line. In fact, this continuous line constitutes the x_1 family in the notation of Contopoulos & Grosbøl (1989). The part that used to be circular

orbits between the two ILRs now forms a closed loop with the orbits elongated perpendicularly to the bar. In fact, the former circular orbits between the ILRs have become the stable x_2 family, while the orbits elongated perpendicularly to the bar have become the x_3 family and have remained unstable, as opposed to the x_1 orbits that have become stable. Note also that the x_3 orbits are not the x_1 orbits rotated as in the axisymmetric case, because the potential is no longer axisymmetric.

When in the bottom two rows of Fig. 6 we further increase ϵ , the x_2 - x_3 loop shrinks, and has almost disappeared when $\epsilon = 0.05$. Note that the shape of the orbits, visible in the central column, does change as we increase ϵ , but not dramatically. What changes significantly is the fraction of the volume of phase space that is occupied by non-closed orbits that librate around x_1 orbits rather than around x_2 orbits. The right column of Fig. 6 illustrates this point by showing surfaces of section⁴ for a value of Jacobi constant $E_j = -0.5$ (for a definition of surfaces of section see for example chapter 3 in Binney & Tremaine 2008). When $\epsilon = 0$, all non-closed orbits are parented by the circular orbit, which corresponds to the centre of the ‘eye’ in the top-right panel. When $\epsilon = 0.005$, two eyes are present in the surfaces of section; the centre of the left one corresponds to the x_1 orbit, and the centre of the right one to the x_2 orbit (which replaces the circular orbit). Some orbits are now parented by the x_1 orbit, and some others by the x_2 . As we increase ϵ , the x_1 orbits become predominant in the surface of section, and the fraction of orbits parented by it increases until the x_2 orbit disappears for this energy.

It is now easy to go back to Fig. 5 and interpret the results. When ϵ is small, the gas is circulating on x_2 orbits in the outer parts and on weakly elongated x_1 orbits in the inner part (inside the ILR1).⁵ When ϵ is increased, the volume of phase space occupied by orbits parented by x_2 orbits is reduced. x_2 orbits start disappearing at small radii, and as they disappear the gas has no other choice than to settle on to x_1 orbits. When $\epsilon = 0.05$ the x_2 orbits are gone almost everywhere, and the gas is flowing everywhere on x_1 orbits, including the region where the x_1 orbits are highly elongated. In this regime, the epicycle approximation obviously cannot work as the x_1 orbits are not weakly deformed circular orbits. If we call weak a bar that can be well described under the epicycle approximation, then the $\epsilon = 0.05$ case should be classified as a strong bar, despite the fact that the non-axisymmetric part of the potential might naively appear small compared to the axisymmetric part.

We have now two questions.

(i) How do the phenomenological theories reviewed above compare with the hydro-simulations in the weak bar case, that is, when ϵ is extremely small and the epicycle approximation is valid?

(ii) What can we say in the strong bar case, when the gas is flowing on very elongated orbits and the epicycle approximation is not valid?

Sections 5 and 6 investigate these two questions respectively.

5 WEAK BAR CASE

In this section we want to investigate how the phenomenological models reviewed above compare in detail with the results of a hydro-simulation in the weak bar case, when the epicycle approximation is valid. To do this, we study the case $\epsilon = 0.005$. Note that this is a value 10 times smaller than the value $\epsilon = 0.05$ considered by Wada (1994) and Piñol-Ferrer et al. (2012). As we have seen in the previous section, the $\epsilon = 0.05$ case should actually be considered a strong bar case, because of its orbital structure.

When the sound speed is not too high, the morphology of the density distribution qualitatively resembles the predictions of Wada (1994) shown in Fig. 4. This can be seen from Fig. 7, which shows results of hydro-simulations with $\epsilon = 0.005$ and different values of the sound speed c_s . In these simulations the bar strength is frozen after the value $\epsilon = 0.005$ is reached, and the snapshots considered are all at $t = 2.9$ Gyr, long after the bar is fully grown to this value and the gas has long settled down into a steady-state configuration. The left column shows the density distribution. The density distributions for $c_s = 0.625 \text{ km s}^{-1}$ and $c_s = 1.25 \text{ km s}^{-1}$ are very similar: it appears as if the gas configuration is tending towards some limiting configuration as $c_s \rightarrow 0$ at fixed spatial resolution, and that this limiting configuration has much in common with the predictions of the phenomenological models.

A second important prediction of the phenomenological model that is confirmed by the hydro-simulations is that the spiral arms can be understood as kinematic density waves. Fig. 8 investigates the instantaneous streamlines, that is, the streamlines calculated from a frozen snapshot of the velocity field. Since the gas is in a steady state, these are not much different from the real streamlines followed by gas parcels during the time evolution. We see that in all cases streamlines are almost closed curves in the sense that in general the mismatch after a single loop is much less than the extent of a radial oscillation during the loop.

Another verified prediction is that streamlines are well described by closed orbits of the form of equation (18). The middle column in Fig. 8 shows in full lines the same streamlines as the left column in the $R\theta$ plane, and in dashed lines a simple fit to each streamline using equation (18) where A and ϕ are considered independent free parameters. The streamlines are well fitted by a function of this form, and if we were to reproduce a nested sequence of streamlines as in Fig. 4 using the best-fitting values of these two parameters, we would obtain kinematic density waves that reproduce those obtained in the simulation.

However, in the phenomenological models, A and ϕ both depend on a single parameter, the dissipation λ , while in the fitting procedure above A and ϕ are treated as independent parameters. Moreover, in the phenomenological models some values of A and ϕ are forbidden: ϕ must stay within the blue and yellow line in Fig. 2, and the amplitude A away from resonance has an upper bound. Therefore it is not obvious that dissipation parameter λ required by the phenomenological model can be tuned for each streamline as to produce the values of A and ϕ given by a hydro-simulation. In other words, the values obtained from a hydro-simulation for A and ϕ need not to come from the same value of λ , or could be outside the regions obtainable for any value of λ . Indeed, from the right column of Fig. 8 it is clear that in the hydro-simulation A and ϕ are often outside the regions allowed by the phenomenological models for the same potential. Therefore not all hydro-simulations can be reproduced by an appropriate gauging of the dissipation. In fact, we chose to fit using A and ϕ as two independent parameters instead of using directly the dissipation λ as the only free parameter because

⁴ However, one must be careful not to confuse the full phase-space volume occupied by a group of orbits with the area in a surface of section filled by the same orbits (see Binney, Gerhard & Hut 1985).

⁵ Note that there must be a transition region in between, and spiral arms are generated as the orientation of orbits changes to make this transition. When the bar perturbation is stronger and orbits are more elongated, the transition can be mediated by shocks (see Sormani et al. 2015) rather than by a soft spiral arm overdensity as in this case. Note also that in the present case the transition is $x_2 \rightarrow x_1$ outwards-in, while the transition discussed in Sormani et al. (2015) is $x_1 \rightarrow x_2$.

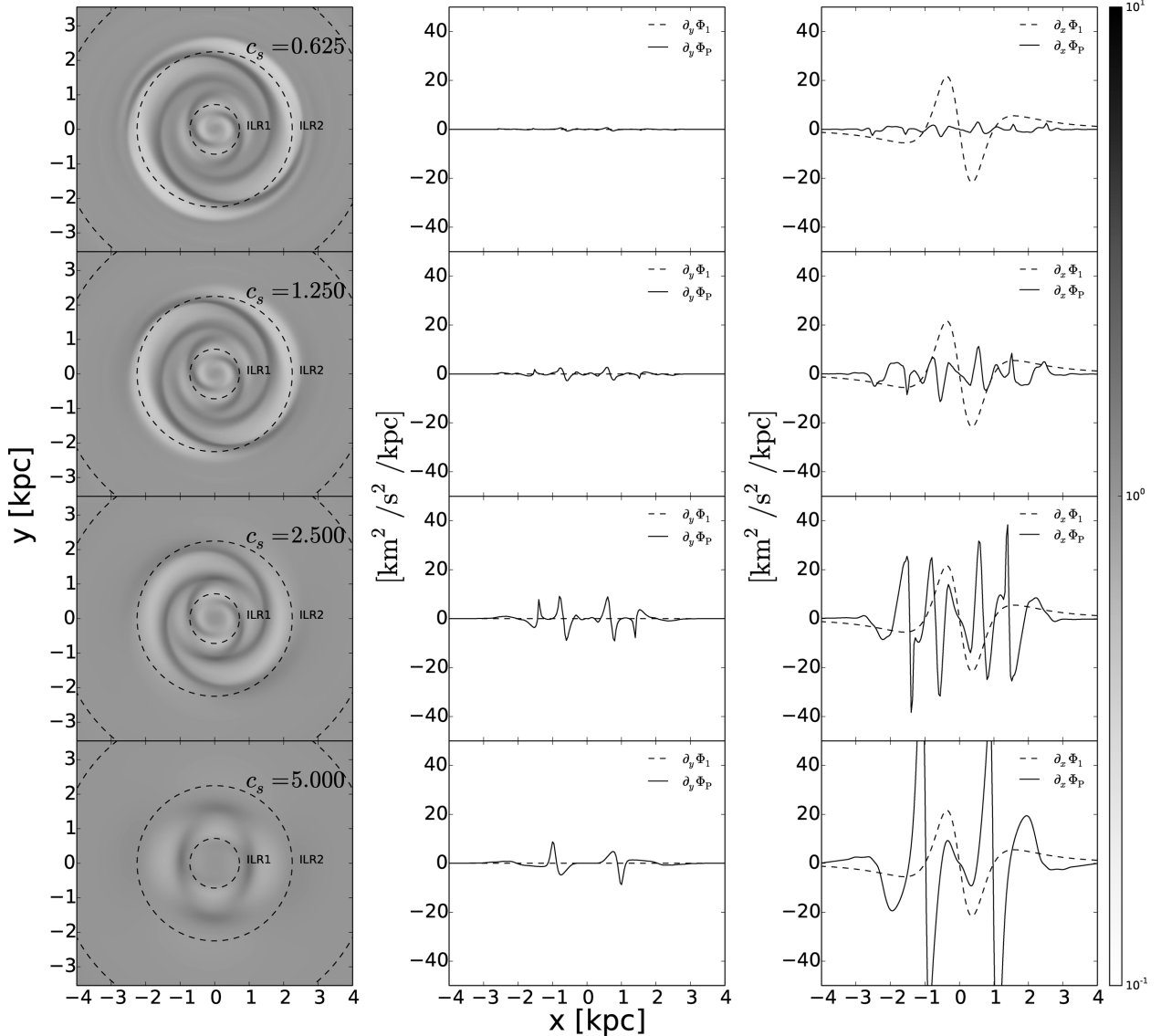


Figure 7. Results of a simulation in Wada’s potential in the case of a weak bar ($\epsilon = 0.005$) for different values of the sound speed c_s . Each panel in the left column shows the steady-state density long after the bar has reached the final value of $\epsilon = 0.005$ – compare with Fig. 4. The middle column shows the y components of pressure forces (full lines) and bar perturbation gravitational forces (dashed line) from the potential in equation (13) with $\epsilon = 0.005$ calculated along the horizontal axis. The right column shows the x components of the same forces, again calculated along the horizontal axis. Note that pressure forces can be derived from a pressure potential Φ_p as described in the text.

A and ϕ lie outside the allowed regions too often. However, in the limit of very low sound speed, the results of the hydro-simulations for A and ϕ are not far from the predictions of the phenomenological model in Fig. 4. The values of ϕ in the top two rows of Fig. 8 are more similar than those in other rows to the black dots in the bottom right panel of Fig. 4, and the values of A have also much in common with the blue dots in the same panel, although many are still forbidden because they lie just above the curve $A = 0$ in the bottom panel of Fig. 3. Another interesting prediction that is verified in this limit is that the orientation of the major axis of the closed orbits at the two ILRs is $\pi/4$.

Why are the results of the phenomenological models well reproduced in the limit of vanishing sound speed, and less well when the sound speed is higher? To understand this, consider the equations of motion of a gaseous parcel flowing in the hydro-simulation. If

the fluid were really an ideal isothermal fluid, these would be

$$\frac{D\mathbf{v}}{Dt} = -\nabla\Phi - c_s^2 \frac{\nabla\rho}{\rho}, \quad (24)$$

where D indicates the material derivative, the first term on the right-hand side is the gravitational force and the second is the pressure force. However, in a real hydrodynamical simulation at fixed spatial resolution, some unavoidable amount of numerical viscosity, which is not included in equation (24), will always be present. This numerical viscosity decreases as we increase the resolution, and tends to zero as the resolution goes to infinity. The equations of motion of a parcel of gas are therefore something like

$$\frac{D\mathbf{v}}{Dt} = -\nabla\Phi - c_s^2 \frac{\nabla\rho}{\rho} - F_{\text{dx}}, \quad (25)$$

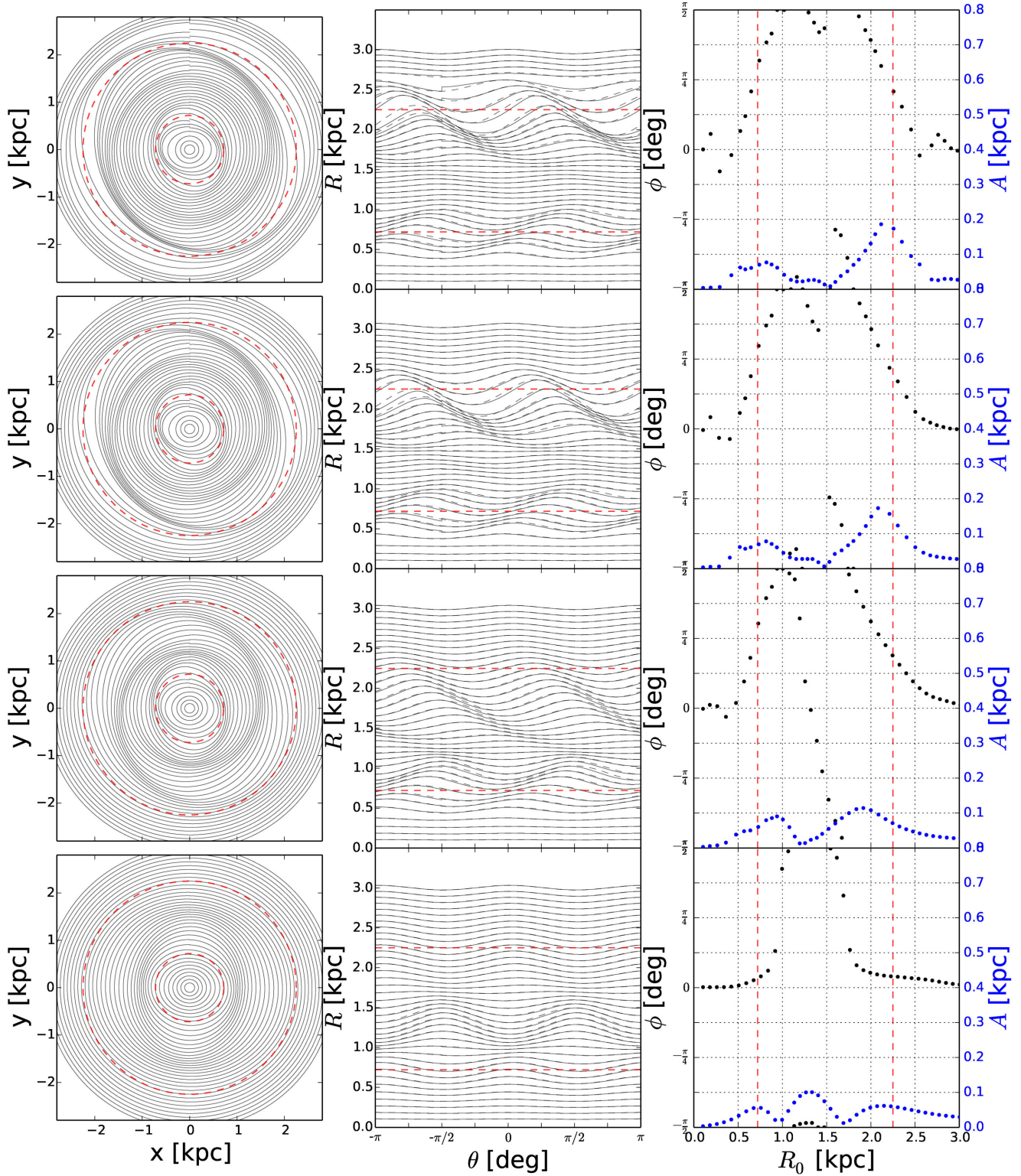


Figure 8. Analysis of the instantaneous streamlines for the snapshots shown in Fig. 7. The left column shows a sequence of nested streamlines in the xy plane. The middle column shows in full lines the same streamlines in the $R\theta$ plane, and in dashed lines the result of a simple fit using equation (10), where A and ϕ are independent parameters. The right panel shows the best-fitting values of A and ϕ as a function of radius. In each panel the dashed red lines mark the two inner Lindblad resonances.

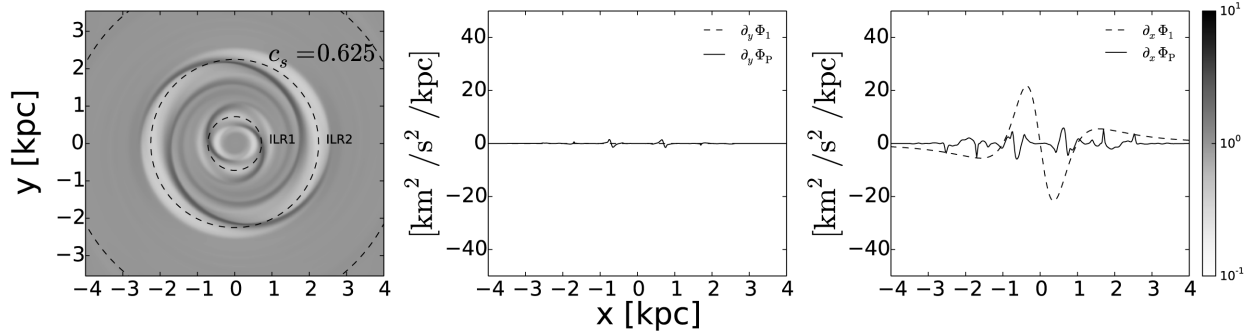


Figure 9. Same as the top row of Fig. 7, but for a simulation with twice the spatial resolution. Here, $\mathrm{d}x = 10$ pc.

where F_{dx} indicates the force arising from the finite resolution, which we will loosely refer to as the viscous force. Hence, a parcel in a hydro-simulation is subject to three different forces: gravitational, pressure and viscosity.

The phenomenological models completely neglect pressure forces. Instead, they only account for the gravitational forces and, phenomenologically, for the viscous forces F_{dx} . Note that the pressure force can be derived from a ‘pressure potential’

$$\Phi_{\mathrm{P}} = c_s^2 \log(\rho/\rho_0), \quad (26)$$

where ρ_0 is an arbitrary number. In a steady-state configuration, the density does not change with time, and the pressure forces can be derived from a static potential. Hence, a gaseous parcel can be considered to move in a potential that is the sum of the gravitational plus the pressure potential. This static pressure potential is not easily derived a priori, and we can only calculate it once we are given the steady-state configuration of the gas after solving the hydrodynamical equations.

The central and right columns in Fig. 7 show the pressure and perturbation potential forces along the horizontal axis. The central column shows forces in the y direction; the gravitational forces in this direction are zero for the chosen form of the perturbation potential. The right column shows forces in the horizontal direction. We see that at high sound speed, when the phenomenological models are less accurate, the magnitude of the pressure forces is in general greater than the magnitude of the perturbation potential forces. At $c_s = 1.25$ the pressure contribution becomes smaller than the bar perturbation contribution, and at $c_s = 0.625$ the pressure forces are negligible compared to the bar perturbation contribution. This is why the phenomenological models above work in this limit. The phenomenological models have terms that take into account both the effects of viscosity and the gravitational potential, but they do not take into account the effects of pressure. In the limit where pressure is negligible, they work. If in the phenomenological models we could replace the gravitational potential with the sum of the gravitational plus the pressure one, then they would describe the results of the simulations at finite pressure. The problem is that in general the pressure potential is not easily obtained a priori.

Before moving on to discuss the strong bar case, we have a couple of remarks left. From a theoretical point of view, when the sound speed tends to zero the gas is always supersonic and we expect all small perturbations to cause shocks. Hence, we expect that when c_s tends to zero, we would find shocks anywhere where there is a small velocity gradient. At a shock, the gradient of the density diverges but the forces due to pressure could remain finite as they depend on the product of the sound speed (which goes to zero) with the gradient

of the pressure (which goes to infinity). The reason why shocks are not formed in our simulations when the sound speed tends to zero is that at finite resolution the numerical viscosity prevents them from forming. In the limit of vanishing sound speed, the numerical viscosity becomes more important than the pressure force. Indeed, we have checked that increasing the numerical resolution leads to steady states in which there are sharper density variations (see Fig. 9). Note that the pressure forces in Fig. 9 are stronger than the pressure forces for a simulation in the same potential, with the same value of the sound speed but at lower resolution (top row of Fig. 7). Hence, we conclude that the phenomenological models work in reproducing the result of a simulation in the limit of low sound speed and of finite numerical viscosity, which means finite numerical resolution.

6 STRONG BAR CASE

In the previous section, we have seen that in the weak bar case the spiral arms can be understood as kinematic density waves, generated by weakly oval closed orbits. What can we say about the strong bar case? In order to investigate a realistic example of a strong bar we consider a simulation we used in Sormani et al. (2015), performed in the Binney et al. (1991) barred potential. The simulation used here is the one with sound speed $c_s = 10 \mathrm{km s}^{-1}$ and spatial resolution $\mathrm{d}x = 10$ pc.

In Sormani et al. (2015), spiral arms were present and noted, but not discussed in detail. It was shown that in the region where spiral arms are present the velocity field was very well approximated by x_1 orbits. To discuss the spiral arms, we will need to look at the tiny deviations of the hydro-velocity field from the ballistic x_1 orbits field.

The left panel in Fig. 10 shows a density snapshot from the simulation in Sormani et al. (2015), taken after the bar is fully grown and the gas has reached an approximate steady state. Spiral arms emerging from the end of the bar are clearly identifiable. These spiral arms are stationary in the frame corotating with the bar. The central panel of Fig. 10 shows what happens when we nest together many instantaneous streamlines. We see that, to a very good approximation, streamlines are closed curves that when nested together generate the spiral arms. In the right panel ballistic closed orbits are shown for comparison. Fig. 11 analyses in more detail a selection of instantaneous streamlines in this snapshot. In the top panel, five streamlines are shown in dashed lines. These are very nearly closed. In the same panel, the ballistic closed orbits that cut the y -axis at the same value of the streamlines are shown in full lines. It can be seen that streamlines are librations around underlying closed x_1 orbits. In the middle panel, the same streamlines and orbits

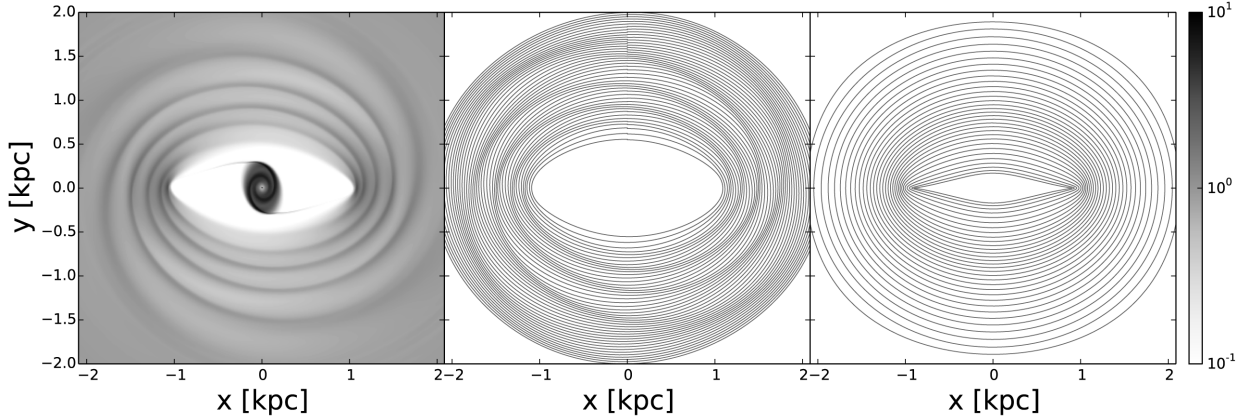


Figure 10. In the left panel, the density distribution for a simulation from SBM2015 with $c_s = 10 \text{ km s}^{-1}$ and spatial resolution $dx = 10 \text{ pc}$. This snapshot is taken after the bar is fully grown and the final steady state is reached. The middle panel shows a nested sequence of instantaneous streamlines. The right panel shows ballistic closed x_1 orbits in the same underlying potential.

are shown in the $R\theta$ plane. In the bottom panel, the radial difference between the two orbits is shown in the $R\theta$ plane. The librations do not have the simple $\cos(2\theta)$ structure that we encountered in Section 5.

The above considerations strongly suggest that also in the strong bar case the spiral arms can be understood as kinematic density waves, generated by small librations around underlying closed orbits. In the epicycle approximation, the librations are around circular orbits. Here, the gas parcels librate around x_1 orbits. It is natural to ask whether the phenomenological models can be extended to the strong bar case, where now the perturbations should be made around the x_1 orbits and not around circular orbits. In order to investigate this, let us go back to the equations of motion in the rotating frame. For a gaseous parcel in a hydro-simulation, equation (2) is modified to

$$\ddot{\mathbf{x}} = -\nabla(\Phi + \Phi_P) + \Omega_p^2 \mathbf{x} - 2\Omega_p (\hat{\mathbf{e}}_z \times \dot{\mathbf{x}}) + F_{dx}, \quad (27)$$

where Φ_P is the ‘potential’ that gives the pressure force introduced above and F_{dx} is the viscous force. We can do a Floquet analysis to investigate librations around ballistic closed orbits (see e.g. Binney & Tremaine 1987). Let us write $\Phi = \Phi_0 + \Phi_1$, where this time Φ_0 is not necessarily axisymmetric. Let $\mathbf{x}_c(t)$ be a ballistic closed orbit in the potential Φ_0 . We can write a libration around this closed orbit as $\mathbf{x}(t) = \mathbf{x}_c(t) + \mathbf{x}_1(t)$. By substituting this last equation into (27), expanding to first order in quantities with subscript 1 and cancelling the zeroth-order terms, we obtain that the equation of motion for the libration is

$$\ddot{\mathbf{x}}_1 = -[(\mathbf{x}_1 \cdot \nabla)\nabla\Phi_0 + \nabla\Phi_1 + \nabla\Phi_P]_{\mathbf{x}_c(t)} + \Omega_p^2 \mathbf{x}_1 - 2\Omega_p (\hat{\mathbf{e}}_z \times \dot{\mathbf{x}}_1) + F_{dx}. \quad (28)$$

In this equation, all the derivatives of the potentials are to be calculated along the unperturbed closed orbit and are therefore known functions of time. Equations (9) for the epicycle approximation can be considered as a particular case of equation (28), obtained assuming that Φ_0 is axisymmetric, $\Phi_P = 0$ and there are no viscosity forces (which are then added heuristically in the phenomenological models). In trying to generalize the analysis of Section 2 to the case when Φ is a strongly barred potential, we are faced with the difficulty that there is in general no natural way of splitting Φ into a Φ_0 where we can easily calculate closed orbits and a perturbation Φ_1 that only weakly deforms closed orbits. The only natural choice is to

set $\Phi_1 = 0$. If this choice is made, the whole perturbing potential is given by Φ_P , which is not known a priori. In the phenomenological models in the epicycle approximation we were able to neglect Φ_P compared to Φ_1 in the low pressure case, but here the same cannot be done if $\Phi_1 = 0$.

Suppose that Φ_P is somehow known. If we assume $\Phi_1 = 0$ and define a suitable phenomenological viscous term F_{dx} in equation (28), would the resulting equation correctly describe librations that give rise to spiral arms? To investigate this, we can do as follows. We first extract the density from the snapshot shown in the left panel of Fig. 10. From this density, we derive the pressure potential using equation (26) and $c_s = 10 \text{ km s}^{-1}$. Then we pretend we do not know how Φ_P was obtained and find librations using the following equation:

$$\mathbf{x}_1 = -[(\mathbf{x}_1 \cdot \nabla)\nabla\Phi_0 + \nabla\Phi_P]_{\mathbf{x}_c(t)} + \Omega_p^2 \mathbf{x}_1 - 2\Omega_p (\hat{\mathbf{e}}_z \times \dot{\mathbf{x}}_1) - 2\lambda \dot{\mathbf{x}}_1. \quad (29)$$

In this last equation, we have introduced a phenomenological dissipation term analogous to the models in the epicycle approximation. The dissipation is proportional to the difference between the total velocity and the local x_1 velocity field. Φ_0 , Φ_P and Ω_p are now known, and given a value of λ we can solve equation (29) to find the librations. Among all possible solutions to this equation, we want the solutions $\mathbf{x}_1(t)$ that are periodic with the same period of the underlying closed orbits. This amounts to finding the solution where transients are gone, dissipated away. The method that we used to solve the differential equation (29) is described in detail in Appendix A.

Fig. 12 shows a nested sequence of librations around x_1 orbits calculated using equation (29) and a value of $\lambda = 50 \text{ km s}^{-1} \text{ kpc}^{-1}$. Spiral arms identical to those present in the hydro-simulation of Fig. 10 are recovered. This confirms that the spiral arms can be understood as kinematic density waves also in the strong bar case. Equation (29) well describes the librations if the correct Φ_P is provided.

However, in the above analysis we have cheated in the sense that we have obtained Φ_P from the result of a full hydrodynamical calculation. Is there a simple way of deriving a suitable Φ_P , without solving the full hydrodynamical problem? We have not found

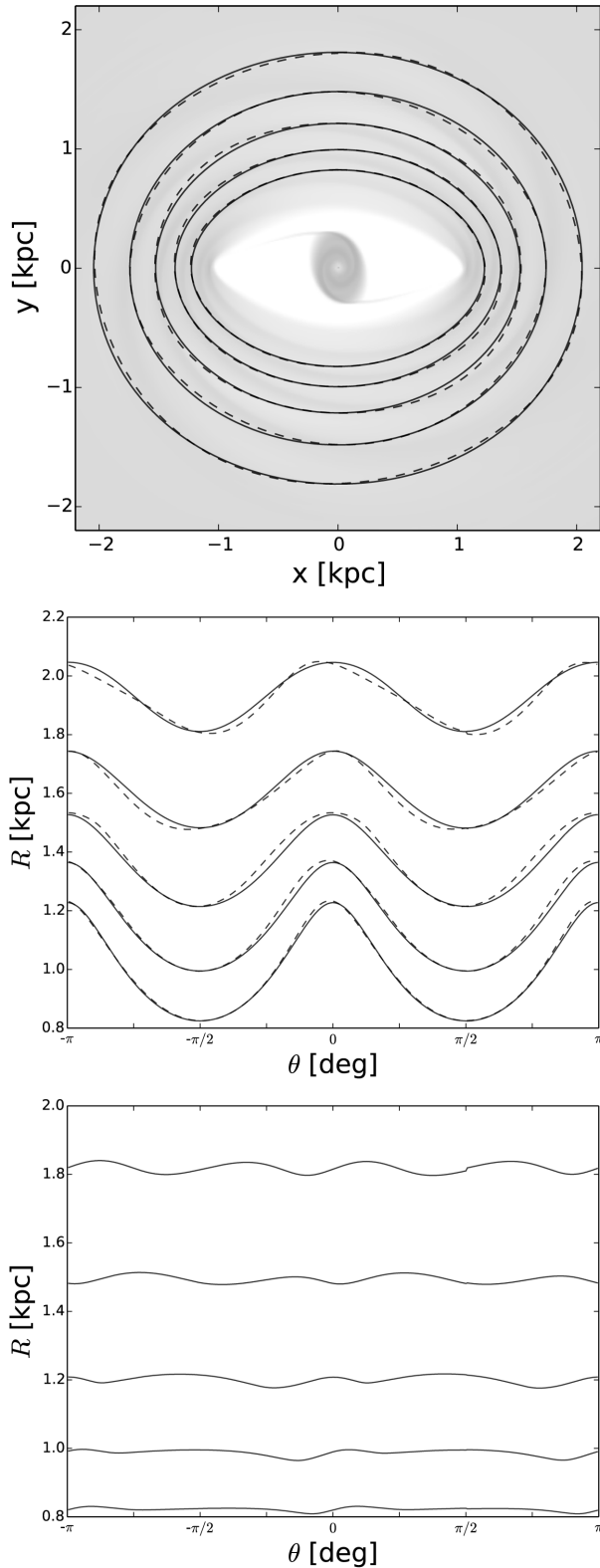


Figure 11. A closer look at some of the streamlines of Fig. 10. In the top panel, streamlines (dashed lines) and ballistic x_1 orbits that cut the y -axis at the same value (full lines). In background is visible the density distribution of Fig. 10 for ease of comparison. The middle panel shows the same streamlines and orbits in the $R\theta$ plane. The bottom panel shows the difference $R_{\text{streamline}}(\theta) - R_{x_1}(\theta)$, where the zero is shifted for clarity.

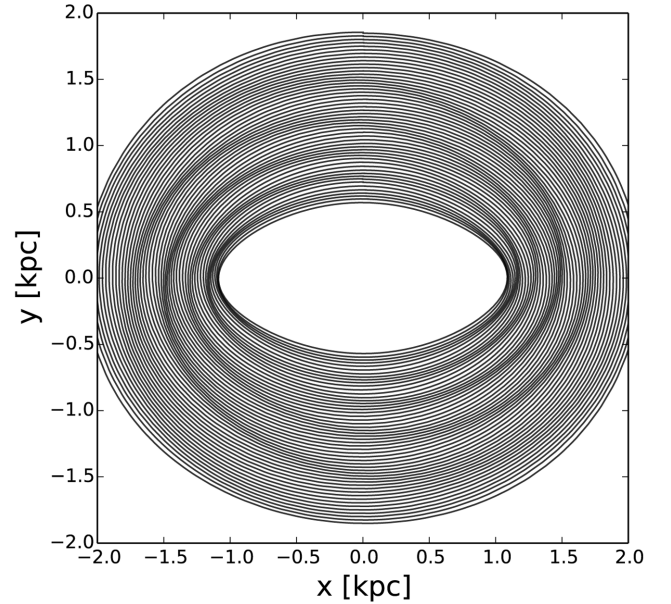


Figure 12. Nested librations around closed x_1 orbits generated using equation (29). The potential Φ_0 is the potential used in SBM2015, the pressure potential Φ_P is derived from the density map in the left panel of 10, and $\lambda = 50 \text{ km s}^{-1} \text{ kpc}^{-1}$.

a better way. We considered the possibility of obtaining Φ_P from the underlying closed orbits. Since the x_1 orbits are highly elongated, we have tried to infer a density map from orbit crowding, and from this the pressure potential. However, the pressure forces obtained in this way were much weaker than the one obtained in the hydro-simulation for a value of the sound speed of $c_s = 10 \text{ km s}^{-1}$, and were not capable of reproducing spiral arms even for a small value of the dissipation. For a higher value of the sound speed we obtained segments of spiral arms that were not coherent and did not produce any ‘grand design’. Hence, we conclude that while the spiral arms in the strong bar case can still be explained by kinematic density waves generated by librations around the appropriate closed orbits, calculation of these librations requires knowledge of the pressure forces and hence solution of the full hydrodynamical problem. Indeed, the morphology of the spiral arms found in Sormani et al. (2015) depends significantly on the value of pressure, and a phenomenological model that wants to reproduce the results of a full hydrodynamical calculation must take pressure explicitly into account.

7 CONCLUSION

In this paper, we have investigated bar-driven spiral arms in the absence of self-gravity. Our focus was on understanding the physical mechanism involved. We concluded that, both in the weak and in strong bar cases, the spiral arms can be understood as kinematic density waves generated by librations around underlying ballistic closed orbits. In the weak bar case, the librations can be considered deviations from circular orbits. In the strong bar case, the librations are to be considered deviations from the appropriate underlying closed orbits, which can be highly elongated. In the strong bar case the epicycle approximation is not valid. In fact, whether a bar can be considered weak or strong is determined by the validity of the epicycle approximation. In Section 4 we argued that a bar is weak or strong accordingly as the orbital structure is or is not obtainable from

the epicycle approximation. Bars that might naively be considered weak, for example the case $\epsilon = 0.05$, should be considered strong.

A parcel in a hydro-simulation is subject to three different forces: gravitational, pressure and viscosity. We have tested the phenomenological models available in the literature aimed at explaining the spiral arms in the weak bar case, when the epicycle approximation is valid. We found that the key ingredient not taken into account by these models is pressure. Therefore they work well in regimes where pressure can be neglected, such as simulations in a weak bar at finite resolution and vanishing sound speed. We have also discussed how the phenomenological models should be extended to the strong bar case. When the pressure forces are known, these extensions work very well in explaining the spiral arms in the strong bar case. Unfortunately, the pressure forces are in general known only after solving the full hydrodynamical problem and are not known a priori. Thus, while the phenomenological models provide insight into the physical mechanism that generates the spiral arms, they appear to be of little practical use.

ACKNOWLEDGEMENTS

MCS acknowledges the support of the Clarendon Scholarship Fund and is indebted to Steven N. Shore for helpful discussions. JB and JM were supported by Science and Technology Facilities Council by grants R22138/GA001 and ST/K00106X/1. JM acknowledges support from the ‘Research in Paris’ programme of Ville de Paris. The research leading to these results has received funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 321067.

REFERENCES

- Athanassoula E., 1992, MNRAS, 259, 345
 Athanassoula E., Romero-Gómez M., Bosma A., Masdemont J. J., 2009a, MNRAS, 400, 1706
 Athanassoula E., Romero-Gómez M., Masdemont J. J., 2009b, MNRAS, 394, 67
 Athanassoula E., Romero-Gómez M., Bosma A., Masdemont J. J., 2010, MNRAS, 407, 1433
 Binney J., Tremaine S., 1987, Galactic dynamics. Princeton Univ. Press, p. 175.
 Binney J., Tremaine S., 2008, Galactic Dynamics, 2nd edn. Princeton Univ. Press, Princeton
 Binney J., Gerhard O. E., Hut P., 1985, MNRAS, 215, 59
 Binney J., Gerhard O. E., Stark A. A., Bally J., Uchida K. I., 1991, MNRAS, 252, 210
 Bissantz N., Englmaier P., Gerhard O., 2003, MNRAS, 340, 949
 Contopoulos G., Grosbol P., 1989, A&AR, 1, 261
 Dame T. M., Thaddeus P., 2008, ApJ, 683, L143
 Dobbs C., Baba J., 2014, Publ. Astron. Soc. Aust., 31, 35
 Englmaier P., Gerhard O., 1999, MNRAS, 304, 512
 Feldman S. I., Lin C. C., 1973, Studies Appl. Math., 52, 1
 Huntley J. M., Sanders R. H., Roberts W. W., Jr, 1978, ApJ, 221, 521
 Lin C. C., Lau Y. Y., 1975, SIAM J. Appl. Math., 29, 352
 Lin C. C., Shu F. H., 1964, ApJ, 140, 646
 Lindblad P. O., Lindblad P. A. B., 1994, in King I. R., ed., ASP Conf. Ser. Vol. 66, Physics of the Gaseous and Stellar Disks of the Galaxy. Astron. Soc. Pac., San Francisco, p. 29
 Piñol-Ferrer N., Lindblad P. O., Fathi K., 2012, MNRAS, 421, 1089
 Rodríguez-Fernández N. J., Combes F., 2008, A&A, 489, 115
 Romero-Gómez M., Masdemont J. J., Athanassoula E., Garc’a-Gómez C., 2006, A&A, 453, 39
 Romero-Gómez M., Athanassoula E., Masdemont J. J., Garc’a-Gómez C., 2007, A&A, 472, 63

Sanders R. H., 1977, ApJ, 217, 916

Sanders R. H., Huntley J. M., 1976, ApJ, 209, 53

Sormani M. C., Binney J., Magorrian J., 2015, MNRAS, 449, 2421

van Albada G. D., van Leer B., Roberts W. W., Jr, 1982, A&A, 108, 76

Wada K., 1994, PASJ, 46, 165

Weiner B. J., Sellwood J. A., 1999, ApJ, 524, 112

APPENDIX A: SOLVING THE FLOQUET EQUATION

In this Appendix we show how to solve equation (29). This equation arose during a Floquet analysis to find librations around closed orbits in a strongly barred potential. The equation to be solved is:

$$\ddot{\mathbf{x}}_1 = -[(\mathbf{x}_1 \cdot \nabla) \nabla \Phi_0 + \nabla \Phi_P]_{x_c(t)} + \Omega_p^2 \mathbf{x}_1 - 2\Omega_p (\hat{e}_z \times \dot{\mathbf{x}}_1) - 2\lambda \dot{\mathbf{x}}_1 \quad (\text{A1})$$

Expanding into components, rearranging and renaming we can rewrite this equation as the following system:

$$\begin{aligned} \ddot{x}_1 &= -x_1 \kappa_{xx}(t) - y_1 \kappa_{xy}(t) + F_x(t) + \Omega_p^2 x_1 - 2\Omega_p \dot{y}_1 - 2\lambda \dot{x}_1, \\ \ddot{y}_1 &= -x_1 \kappa_{xy}(t) - y_1 \kappa_{yy}(t) + F_y(t) + \Omega_p^2 y_1 + 2\Omega_p \dot{x}_1 - 2\lambda \dot{y}_1, \end{aligned} \quad (\text{A2})$$

where

$$\begin{aligned} \kappa_{xx}(t) &= [\partial_x^2 \Phi_0]_{x_c(t)}, \\ \kappa_{xy}(t) &= [\partial_{xy}^2 \Phi_0]_{x_c(t)}, \\ \kappa_{yy}(t) &= [\partial_y^2 \Phi_0]_{x_c(t)}, \\ F_x(t) &= -[\partial_x \Phi_P]_{x_c(t)}, \\ F_y(t) &= -[\partial_y \Phi_P]_{x_c(t)}. \end{aligned} \quad (\text{A3})$$

In equations (A2) it is assumed that $\lambda > 0$ and that all the κ s and F s are given periodic functions of time with period T :

$$\begin{aligned} \kappa_{xx}(t) &= \kappa_{xx}(t + T), \\ \kappa_{xy}(t) &= \kappa_{xy}(t + T), \\ \kappa_{yy}(t) &= \kappa_{yy}(t + T), \\ F_x(t) &= F_x(t + T), \\ F_y(t) &= F_y(t + T). \end{aligned} \quad (\text{A4})$$

The goal is to solve equations (A2) to find $\mathbf{x}_1(t) = (x_1(t), y_1(t))$. We want to find the solution of this equation that neglects transients, i.e. the solution to which all solutions tend in the limit $t \rightarrow \infty$. This solution is the one such that $\mathbf{x}_1(t)$ is periodic with the same periodicity of the F s and κ s. This is also a requirement if we want to obtain a closed orbit, and we have to assume such periodicity if our phenomenological model is to make sense.⁶ Let us therefore

⁶ Note that equation (A1) is conceptually similar to the following simpler equation

$$\ddot{x} + \kappa(t)x + \lambda \dot{x} = F(t) \quad (\text{A5})$$

If κ were constant and not time dependent, equation (A5) would be the equation of a damped and driven harmonic oscillator. The general solution of this equation is a transient that decays exponentially plus a periodic term with the same periodicity as $F(t)$. Our problem is more general, and κ is a function of time. We are interested in the solution that neglects the transients, which is the generalization of the solution of the damped

expand all time-dependent quantities as Fourier series, assuming they are periodic with period T :

$$\begin{aligned}
 x_1(t) &= \sum_n X_n e^{in\omega t}, \\
 y_1(t) &= \sum_n Y_n e^{in\omega t}, \\
 \kappa_{xx}(t) &= \sum_n K_{xx;n} e^{in\omega t}, \\
 \kappa_{xy}(t) &= \sum_n K_{xy;n} e^{in\omega t}, \\
 \kappa_{yy}(t) &= \sum_n K_{yy;n} e^{in\omega t}, \\
 F_x(t) &= \sum_n F_{x;n} e^{in\omega t}, \\
 F_y(t) &= \sum_n F_{y;n} e^{in\omega t}.
 \end{aligned} \tag{A6}$$

Fortunately, the product of the two Fourier series gives us another Fourier series of the same type. Substituting the Fourier expansions into equations (A2) and then equating coefficients term by term we arrive at

$$\begin{aligned}
 -[(n\omega)^2 + \Omega_p^2] X_n &= -\sum_m K_{xx;m} X_{n-m} - \sum_m K_{xy;m} Y_{n-m} \\
 &\quad - (in\omega) [2\lambda X_n + 2\Omega_p Y_n] + F_{x;n}, \\
 -[(n\omega)^2 + \Omega_p^2] Y_n &= -\sum_m K_{yy;m} Y_{n-m} - \sum_m K_{yx;m} X_{n-m} \\
 &\quad - (in\omega) [2\lambda Y_n - 2\Omega_p X_n] + F_{y;n}.
 \end{aligned} \tag{A7}$$

The last two equations are a linear algebraic system of equations in the unknowns X_n and Y_n . All the other Fourier coefficients, K s and F s, are assumed to be known. To solve this system let us rewrite it in the form

$$\mathbb{A}\mathbf{X} = \mathbf{F}, \tag{A8}$$

where \mathbb{A} is an infinite matrix and

$$\mathbf{X} = \begin{pmatrix} \vdots \\ X_{-1} \\ Y_{-1} \\ X_0 \\ Y_0 \\ X_1 \\ Y_1 \\ \vdots \end{pmatrix}. \tag{A9}$$

and driven harmonic oscillator that neglects the decaying exponential. This solution is the one such that $x(t)$ is periodic with the same periodicity of $F(t)$ and $k(t)$.

\mathbf{F} is

$$\mathbf{F} = \begin{pmatrix} \vdots \\ F_{x;-1} \\ F_{y;-1} \\ F_{x;0} \\ F_{y;0} \\ F_{x;1} \\ F_{y;1} \\ \vdots \end{pmatrix}. \tag{A10}$$

To write \mathbb{A} , let us divide it into a part \mathbb{K} containing the Fourier coefficients of the κ 's and a part \mathbb{D} that does not:

$$\mathbb{A} = \mathbb{D} + \mathbb{K}. \tag{A11}$$

\mathbb{D} is a block-diagonal matrix: it has 2×2 blocks along the diagonal. Each block is

$$\mathbb{D}_n = \begin{pmatrix} -(n\omega)^2 + 2\lambda(in\omega) + \Omega_p & (in\omega)2\Omega_p \\ -(in\omega)2\Omega_p & -(n\omega)^2 + 2\lambda(in\omega) + \Omega_p \end{pmatrix}. \tag{A12}$$

This is the block referring to the vector (X_n, Y_n) . The entire matrix \mathbb{D} is then

$$\mathbb{D} = \begin{pmatrix} \ddots & & & & & & \\ & \mathbb{D}_{-1} & & & & & \\ & & \mathbb{D}_0 & & & & \\ & & & \mathbb{D}_1 & & & \\ & & & & \ddots & & \end{pmatrix}. \tag{A13}$$

The matrix \mathbb{K} is

$$\mathbb{K} = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ \cdots & \mathbb{K}_0 & \mathbb{K}_1 & \mathbb{K}_2 & \mathbb{K}_3 & \mathbb{K}_4 & \cdots \\ \cdots & \mathbb{K}_{-1} & \mathbb{K}_0 & \mathbb{K}_1 & \mathbb{K}_2 & \mathbb{K}_3 & \cdots \\ \cdots & \mathbb{K}_{-2} & \mathbb{K}_{-1} & \mathbb{K}_0 & \mathbb{K}_1 & \mathbb{K}_2 & \cdots \\ \cdots & \mathbb{K}_{-3} & \mathbb{K}_{-2} & \mathbb{K}_{-1} & \mathbb{K}_0 & \mathbb{K}_1 & \cdots \\ \cdots & \mathbb{K}_{-4} & \mathbb{K}_{-3} & \mathbb{K}_{-2} & \mathbb{K}_{-1} & \mathbb{K}_0 & \cdots \\ \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{A14}$$

where each \mathbb{K}_n is a 2×2 block given by

$$\mathbb{K}_n = \begin{pmatrix} K_{xx;n} & K_{xy;n} \\ K_{yx;n} & K_{yy;n} \end{pmatrix}. \tag{A15}$$

Since the linear algebraic system represented by equation (A8) is infinite, we need to truncate it in order to be able to solve it. On the diagonal of the matrix \mathbb{K} there is \mathbb{K}_0 . Suppose we truncate the Fourier expansion of all κ s at order \mathbb{K}_0 . Then the matrix \mathbb{A} is block diagonal and the system is very easy to solve by considering only line pairs at a time. Now suppose we truncate the Fourier expansion of $\kappa(t)$ at some higher order. In this case, the matrix \mathbb{A} is not block diagonal, but is a block band matrix that has elements on the sides of the diagonal up to the order of truncation. For example, if we truncate at $\mathbb{K}_{\pm 1}$, then we have a (block) band of width 3. To find the

solution we need to truncate the system at some finite order and in general we cannot solve the system exactly as in the case in which we truncate at \mathbb{K}_0 , where equations are decoupled in pairs. Thus there are two truncations involved. The first is where to truncate the Fourier expansions of κ and F . This truncation has to be done at some order where κ and F are well represented by their Fourier expansions. The second is the size of truncation of the system \mathbb{A} , and must be done after the first truncation has been performed. This

second truncation must be done for a size sufficiently high that the solution is converged and is not affected by a further increase in the size of the system. We have found that this convergence takes place quite rapidly.

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