

# A dynamical mechanism for the origin of nuclear rings

Mattia C. Sormani <sup>1</sup>★, Emanuele Sobacchi,<sup>2,3</sup> Francesca Fragkoudi,<sup>4,5</sup>  
Matthew Ridley <sup>6</sup>, Robin G. Treß,<sup>1</sup> Simon C. O. Glover<sup>1</sup> and Ralf S. Klessen<sup>1,7</sup>

<sup>1</sup>Zentrum für Astronomie, Institut für theoretische Astrophysik, Universität Heidelberg, Albert-Ueberle-Str. 2, D-69120 Heidelberg, Germany

<sup>2</sup>Physics Department, Ben-Gurion University, PO Box 653, Beer-Sheva 84105, Israel

<sup>3</sup>Department of Natural Sciences, The Open University of Israel, 1 University Road, PO Box 808, Raanana 4353701, Israel

<sup>4</sup>Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, D-85748 Garching bei München, Germany

<sup>5</sup>GEPI, Observatoire de Paris, PSL Université, CNRS, 5 Place Jules Janssen, F-92190 Meudon, France

<sup>6</sup>Rudolf Peierls Centre for Theoretical Physics, 1 Keble Road, Oxford OX1 3NP, UK

<sup>7</sup>Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Universität Heidelberg, Im Neuenheimer Feld 205, D-69120 Heidelberg, Germany

Accepted 2018 August 14. Received 2018 July 17; in original form 2018 February 17

## ABSTRACT

We develop a dynamical theory for the origin of nuclear rings in barred galaxies. In analogy with the standard theory of accretion discs, our theory is based on shear viscous forces among nested annuli of gas. However, the fact that gas follows non-circular orbits in an external barred potential has profound consequences: it creates a region of reverse shear in which it is energetically favourable to form a stable ring that does not spread despite dissipation. Our theory allows us to approximately predict the size of the ring given the underlying gravitational potential. The size of the ring is loosely related to the location of the Inner Lindblad Resonance in the epicyclic approximation, but the predicted location is more accurate and is also valid for strongly barred potentials. By comparing analytical predictions with the results of hydrodynamical simulations, we find that our theory provides a viable mechanism for ring formation if the effective sound speed of the gas is low ( $c_s \lesssim 1 \text{ km s}^{-1}$ ), but that nuclear spirals/shocks created by pressure destroy the ring when the sound speed is high ( $c_s \simeq 10 \text{ km s}^{-1}$ ). We conclude that whether this mechanism for ring formation is relevant for real galaxies ultimately depends on the effective equation of state of the interstellar medium (ISM). Promising confirmation comes from simulations in which the ISM is modelled using state-of-the-art cooling functions coupled to live chemical networks, but more tests are needed regarding the role of turbulence driven by stellar feedback. If the mechanism is relevant in real galaxies, it could provide a powerful tool to constrain the gravitational potential, in particular the bar pattern speed.

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

## 1 INTRODUCTION

Rings of gas and stars are a common morphological feature of galaxies (e.g. Buta & Combes 1996; Knapen 2005; Comerón et al. 2010; Comerón 2013; Buta 2017a,b). They are most often associated with a stellar bar and in the parlance of galaxy morphology are classified into three main types, primarily according to their size: nuclear rings, inner rings, and outer rings.

Nuclear rings are the smallest type of ring and are commonly found at the centre of barred galaxies. They are usually sites of intense star formation. The radii of nuclear rings range from a few

tens of parsec to  $\sim 3$  kpc (e.g. Knapen 2005; Comerón et al. 2010; Comerón 2013). Our Galaxy is likely to have such a ring with radius  $\simeq 150$  pc, which corresponds to the Central Molecular Zone (e.g. Liszt 2009; Molinari et al. 2011; Kruijssen, Dale & Longmore 2015; Henshaw et al. 2016; Sormani et al. 2018). Nuclear rings are a gas reservoir for the accretion disc that surrounds the supermassive black hole that is present at the centre of most galaxies, although it is not clear how the gas migrates from the nuclear ring ( $r \sim 100$  pc) down to the accretion discs at much smaller radii ( $r \sim 10^{-3}$  pc; see e.g. Phinney 1994; Combes 2001 and also Li, Sellwood & Shen 2017).

It is currently unclear what exactly determines the size of nuclear rings (e.g. van de Ven & Chang 2009; Comerón et al. 2010). It is clear that there must be some connection between the size of the ring

\* E-mail: [mattia.sormani@alumni.sns.it](mailto:mattia.sormani@alumni.sns.it)

and the Inner Lindblad Resonance (ILR) of the underlying potential (e.g. Buta & Combes 1996), and that gas in the ring must flow on a family of non-circular orbits mildly elongated perpendicular to the bar called  $x_2$  orbits (e.g. Regan & Teuben 2003; Maciejewski 2004; Kim et al. 2012; Sormani, Binney & Magorrian 2015a). These two facts are not contradictory, since the size and extent of the  $x_2$  orbital family is correlated to the location of the ILR(s) (e.g. Contopoulos & Grosbøl 1989; Athanassoula 1992a). Virtually all modern explanations for the formation of rings broadly agree with these two statements, but they differ in the details of how the associations are made.

The most widely accepted theory, which we shall call resonant theory, suggests that the reason why nuclear rings are associated with the ILR is that a spiral pattern is induced by the bar such that it experiences net torques from the barred potential that change sign at each resonance (e.g. Combes 1988, 1996; Buta & Combes 1996). Thus, as Buta & Combes (1996) describe it, ‘rings are formed by gas accumulation at the Lindblad resonances, under the continuous action of gravity torques from the bar pattern’. However, the concept of the ILR is well defined only for a weak bar. One may extend it to the case of a strong bar based on the  $x_2$  orbital family (see for example the discussion in section 3.2 of Athanassoula 1992a), but then the  $x_2$  orbits are in a sense all ‘resonant’, and the resonant theory is unable to predict which orbits will be populated by gas.

The sound speed of the gas  $c_s$  is an important parameter of the flow. By running 2D hydrodynamical simulations in an externally imposed barred potential, Kim et al. (2012; see also Patsis & Athanassoula 2000) find that the size and morphology of the nuclear region depends on  $c_s$ , even if the underlying gravitational potential and therefore the ILR position stays exactly the same. It is not clear how this finding can be reconciled with the resonant theory.

There is also evidence that viscosity plays a key role in the formation of rings. This is indicated by the fact that simulations using grid codes, which model the gas as a continuous isothermal fluid and typically do not include explicit viscous terms, can produce a variety of morphologies ranging from nuclear rings to nuclear spirals (e.g. Maciejewski 2004; Kim et al. 2012; Li, Shen & Kim 2015; Sormani et al. 2015a; Sormani, Binney & Magorrian 2015b; Fragkoudi, Athanassoula & Bosma 2017), while simulations that use the sticky-particle technique, which models the gas as inelastically colliding test particles and therefore contain some explicit amount of dissipation, almost invariably produce nuclear rings, and almost never nuclear spirals (e.g. Schwarz 1981; Combes & Gerin 1985; Byrd et al. 1994; Jenkins & Binney 1994; Rautiainen & Salo 2000; Rautiainen, Salo & Laurikainen 2002; Rautiainen, Salo & Buta 2004; Rodriguez-Fernandez & Combes 2008). Thus viscosity seems to favour the development of rings, and how this can be reconciled with the resonant theory is also unclear.<sup>1</sup>

<sup>1</sup>According to the resonant theory, the key to ring formation is the generation of a spiral pattern on which gravity torques can act. Since as we discuss below in Section 4.2 (see also Sormani et al. 2015b) viscosity does not play a major role in the generation of the spiral pattern, according to the resonant theory viscosity should play little role in the formation of rings. Indeed, the spiral pattern is generated by pressure, not viscosity. This poses another puzzling question: if the idea underlying the resonant theory is correct, we would expect spirals created by pressure to always eventually develop into rings because of gravity torques. Yet, simulations show that this often does not happen, with the gas maintaining a quasi-steady nuclear spiral pattern that does not develop into rings (e.g. Kim et al. 2012; Li et al. 2015).

Hence, two key parameters are the sound speed of the gas and the amount of viscosity. In order to understand the formation of rings we need to understand exactly what role they play. Unfortunately, to the best of our knowledge there are no studies that explore the effects of viscosity on gas flowing on non-circular orbits in a non-axisymmetric galactic potential,<sup>2</sup> with the exception of the sticky-particles simulations that, however, are difficult to analyse because the viscosity they include is not the same as the classic Navier–Stokes viscosity.

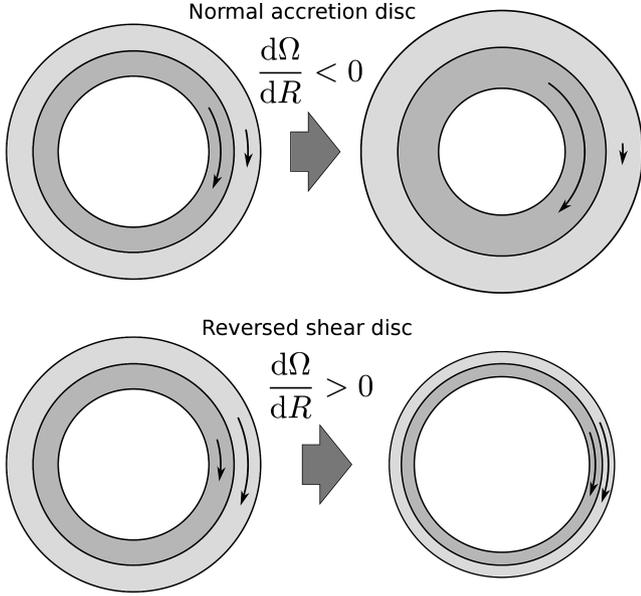
In this paper, we aim to study the effects of viscosity and pressure on a disc of gas flowing on non-circular orbits in an externally imposed barred potential. We find that viscosity has consequences that can be significantly different from the circular case, which may be related to the origin of nuclear rings. We also aim to understand the interplay between viscosity and pressure, the latter being the other key parameter of the flow. Finally, we want to investigate whether the ring size can be predicted more precisely, i.e. beyond establishing a loose connection between the ring size and the location of the ILR.

The paper is structured as follows. In Section 2 we discuss a simple effect based on the viscous theory of accretion discs that is apparently only of academic interest, but will turn out to be a useful toy model to understand the mechanism discussed in Section 3. In Section 3 we discuss the key mechanism that forms the basis of our picture for the formation of nuclear rings. We show that viscosity can create a trapping region where it is energetically favourable to form a ring. However, in common with viscous accretion disc theory in its simplest form, throughout this section we make one critical assumption: we neglect pressure. Hence in Sections 4.1 and 4.2 we compare the analytical predictions of Section 3 with the results of idealized isothermal simulations of a disc of gas flowing in an externally imposed barred potential that include various amounts of explicit viscosity and pressure. In Section 4.3 we compare with a much more realistic simulation of the entire bar region of the Milky Way that models the interstellar medium (ISM) using state-of-the-art cooling functions and chemical networks rather than a simple isothermal approximation. In Section 5 we discuss (i) our results and approximations, (ii) what have we learned about the effects of viscosity and pressure, (iii) relation to other works, and (iv) relevance for real galaxies and possible applications. Section 6 sums up.

## 2 A CONSIDERATION FROM THE ELEMENTARY THEORY OF ACCRETION DISCS

In the elementary theory of accretion discs we learn that, as a consequence of viscous torques between adjacent annuli of gas, mass flows inwards and angular momentum flows outwards (e.g. Shakura & Sunyaev 1973; Lynden-Bell & Pringle 1974; Pringle 1981). This situation is schematically depicted in the top panel of Fig. 1: the inner ring rotates *faster* than the outer ring, thus friction between the two will try to slow down the inner ring and speed up the outer ring. Angular momentum is transferred from the inner ring to the outer ring. So, if the inner ring loses angular momentum, but is forced to remain on a circular orbit, it must move inwards to the radius corresponding to its new angular momentum, which increases

<sup>2</sup>A similar problem has been studied extensively in the context of planetary ring dynamics (see Section 5.5.2).



**Figure 1.** When  $d\Omega/dR < 0$ , as in a Kepler potential, the elementary theory of viscous accretion discs tells us that a ring of gas spreads in time. In a hypothetical situation where  $d\Omega/dR > 0$ , the ring would tighten in time. There is no accretion in this case and one gets a ring-forming disc.

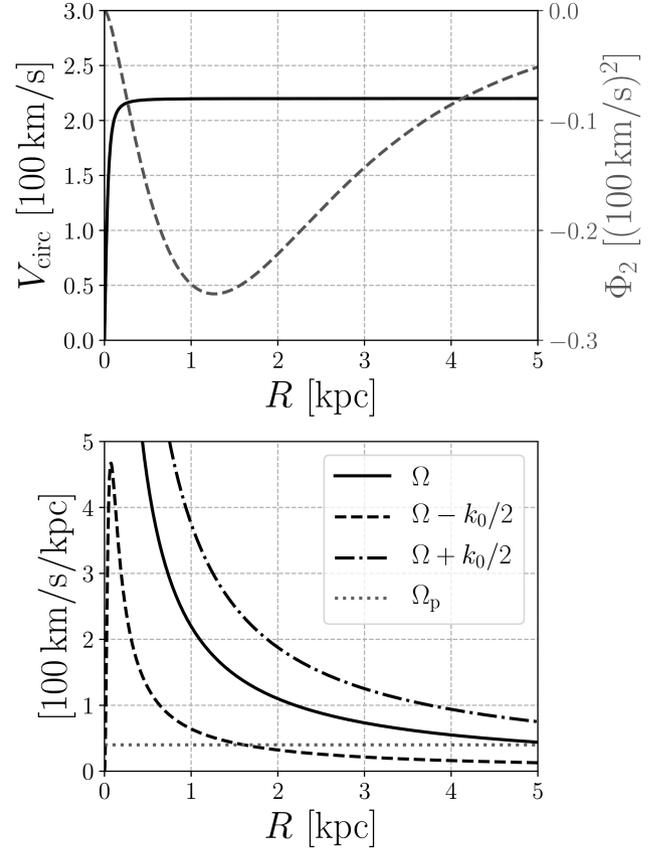
its circular velocity.<sup>3</sup> The outer ring moves outwards, decreasing its circular velocity. The net result is that the ring spreads. The equations for the time evolution of the surface density  $\Sigma(R, t)$  have much in common with the heat (diffusion) equation.

The previous result relies on the fundamental assumption that  $d\Omega/dR < 0$ , i.e. that the angular velocity  $\Omega(R)$  is a decreasing function of radius. This is true for the Kepler potential and in all practical cases in galaxies. However, let us ignore this for a moment and suppose that  $d\Omega/dR > 0$ . What would happen in this case? The situation is depicted in the lower panel of Fig. 1. The inner ring rotates *slower* than the outer ring, thus friction between the two will try to speed up the inner ring and slow down the outer ring. Angular momentum is transferred from the outer ring to the inner ring. So, if everything is forced to remain on a circular orbit, the inner ring must move outwards, and the outer ring moves inwards. Therefore, we reach the conclusion that *if  $d\Omega/dR > 0$ , a ring of finite width becomes tighter and tighter*. There is no accretion in this case – instead, a disc with some density fluctuations in the initial surface density will break up into rings. The equations for the time evolution of the surface density  $\Sigma(R, t)$  would look like the heat equation, but with heat flowing from cold to hot!<sup>4</sup>

These considerations seem only of academic interest, since  $d\Omega/dR > 0$  never actually happens in the rotation curves of real galaxies. However, we will see that this ‘reversed shear’ disc serves as a useful toy model to understand what actually happens when gas follows non-circular orbits in a barred potential. The same idea has also been successfully applied to explain the confinement and

<sup>3</sup>This is the ‘donkey effect’, see Lynden-Bell & Kalnajs (1972) or Box 3.3 in Binney & Tremaine (2008).

<sup>4</sup>Indeed, the analysis of Lynden-Bell & Pringle (1974) – that shows that the most energetically favourable situation is reached when most of the gas is driven to the centre while all angular momentum has been transported to infinity by an infinitesimal amount of mass – is not valid in the case  $d\Omega/dR > 0$  (see their last equation on page 606 and related remarks).



**Figure 2.** *Top panel:* the circular velocity curve  $v_{\text{circ}}(R) = \sqrt{Rd\Phi_0/dR}$  (full black line) and the quadrupole  $\Phi_2(R)$  (dashed line) of the potential described in Section 3. *Bottom panel:* the curves  $\Omega$  and  $\Omega \pm k_0/2$ , where  $k_0^2 = (2\Omega/R)d(\Omega R^2)/dR$  is the epicyclic frequency. The red dashed line indicates the value of the bar pattern speed. The ILR is located at the intersection between the horizontal  $\Omega_p$  line and the  $\Omega - k_0/2$  curve.

narrow edges of rings in the context of planetary ring dynamics (see Section 5.5.2).

### 3 A MECHANISM FOR THE FORMATION OF NUCLEAR RINGS

Our starting point is the well known fact that gas flowing in the central regions of a barred potential does not follow circular orbits, but a family of closed orbits elongated perpendicular to the bar called  $x_2$  orbits (e.g. Contopoulos & Grosbøl 1989; Sormani et al. 2015a).

As a concrete example, let us consider a simple analytical barred potential of the following form:

$$\Phi(R, \theta) = \Phi_0(R) + \Phi_2(R) \cos(2\theta), \quad (1)$$

where  $R, \theta$  are standard polar coordinates. For the axisymmetric part we take a simple logarithmic potential:

$$\Phi_0(R) = \frac{v_0^2}{2} \log(R^2 + R_c^2), \quad (2)$$

where  $v_0 = 220 \text{ km s}^{-1}$  and  $R_c = 0.05 \text{ kpc}$ . For the quadrupole  $\Phi_2$  we adopt the same function that Sormani, Binney & Magorrian (2015c) used to simulate the gas flow in model Milky Way galaxies, which is described in detail in Appendix A. The potential rigidly rotates with a constant pattern speed of  $\Omega_p = 40 \text{ km s}^{-1} \text{ kpc}^{-1}$ . Fig. 2

plots the circular velocity curve, the quadrupole, and the resonance diagram in the epicyclic approximation.

The left-hand panel of Fig. 3 shows  $x_2$  orbits for this potential. These illustrate some fairly general characteristics of the  $x_2$  orbital family. Orbits become more and more elongated as we move to larger values of the semimajor axis  $a$ . The  $x_2$  family has a finite radial extent, and there are no  $x_2$  orbits beyond a maximum value of the semimajor axis. This puts an obvious upper limit on the maximum size of the ring, i.e. it must be less than or equal to the size of the largest  $x_2$  orbit, which for this particular potential is  $a_{\max} \simeq 1.5$  kpc. Note that the ILR is at  $R_{\text{ILR}} \simeq 1.6$  kpc (see intersection of  $\Omega - k_0/2$  curve and  $\Omega_p$  in Fig. 2).

Imagine constructing a steady state configuration of gas flowing on such orbits. To do that, consider a disc of gas comprised by nested annuli, each annulus covering the region between an  $x_2$  orbit and an adjacent one. Ignoring pressure and viscous forces, a steady state can be constructed by taking the surface density in each annulus to be inversely proportional to the speed along the orbit  $v$  and the width of the annulus  $db$ , that is

$$\rho_{\text{annulus}} = \frac{\alpha}{vdb}, \quad (3)$$

where  $\alpha$  is a constant that can be chosen independently for each annulus, and determines the mass of the annulus.

Note that this construction is possible only for orbits with semimajor axis  $a \lesssim 1.35$  kpc (green and orange in Fig. 3): orbits with semimajor axis greater than this value (shown in grey) cross adjacent orbits, so nested annuli cannot be constructed. Therefore, gas on these orbits would shock and plunge onto inner orbits. We shall call orbits in the grey area the ‘forbidden orbits’.

In the absence of pressure and viscous forces, steady states constructed this way will survive forever. What happens if we introduce some small viscosity, so that adjacent annuli can exert torques on each other, like they do in accretion discs? Let us calculate the torque that an annulus exerts on an adjacent one. Viscosity creates forces between layers of fluid that move relative to each other. The viscous force per unit length and unit surface density acting on an orbit contour can be written as

$$f_i = n_j \sigma_{ij}, \quad (4)$$

where  $\hat{\mathbf{n}}$  is the normal to the contour (see Fig. 4) and

$$\sigma_{ij} = \nu \left( \partial_i v_j + \partial_j v_i - \frac{2}{3} \delta_{ij} (\nabla \cdot \mathbf{v}) \right) \quad (5)$$

is the viscous stress tensor, where  $\nu$  is the coefficient of kinematic viscosity with dimensions  $[\nu] = \text{length}^2 \times \text{time}^{-1}$ . The torque is the cross-product of the radius and the force. Therefore, the torque per unit length and unit surface density applied to the contour is

$$\mathbf{k} = \mathbf{r} \times \mathbf{f}, \quad (6)$$

where  $\mathbf{r}$  is the radial vector from the centre, and since our problem is 2D we can write  $\mathbf{k} = k \hat{\mathbf{z}}$ . Equation (6) gives the torque on a contour. It can be seen as the flux of angular momentum through the contour. To find the net torque acting over a region bounded by a closed contour we need to integrate equation (6) over the contour.

For circular orbits equation (6) reduces to

$$k_{\text{circ}} = -\nu R^2 \left( \frac{d\Omega}{dR} \right). \quad (7)$$

Thus we see that regions of ‘normal shear’ ( $d\Omega/dR < 0$ ), in which angular momentum is flowing outwards, correspond to  $k_{\text{circ}} > 0$ , while regions of ‘reversed shear’ ( $d\Omega/dR > 0$ ), in which angular

momentum is flowing inwards, correspond to  $k_{\text{circ}} < 0$ . This correspondence between the sign of  $k$  and the direction of the angular momentum flux remains true also in the non-circular case. Thus, regions of  $k < 0$  are regions of ‘reversed shear’.

The top right-hand panel in Fig. 3 plots the quantity  $k$  along the  $x_2$  orbits as a function of the polar angle  $\theta$ . For orbits in the green region,  $k > 0$  everywhere. This region behaves like a ‘normal’ accretion disc with  $d\Omega/dR < 0$ . Orange orbits instead have portions around  $\theta = 0$  and  $\theta = \pi$  (which correspond to the  $x$  axis) where  $k < 0$ . These are portions of reversed shear where an inner orbit tries to slow down an outer orbit, as in the circular case with  $d\Omega/dR > 0$ . The reversed shear portions become larger as we move to outer orbits, and eventually the value of  $k$  diverges on the  $x$  axis as we enter the ‘forbidden region’ (grey orbits). This suggests that some orbits within the orange region will be unstable for the formation of rings, in analogy with the considerations in Section 2. A ring of gas flowing on these orbits would survive without spreading in time.

Intuitively, the reversal as we go from the  $y$  to the  $x$  axis happens for the following reason. Consider two adjacent orbits in the orange region. At the point where they intercept the  $y$  axis, the angular velocity of the outer one is smaller than the angular velocity of the inner one, as in a normal accretion disc. However, because the elongation of the orbits increases as we move outwards, the outer orbit reduces its distance to the centre more than the inner orbit does as they approach the  $x$  axis. The greater loss of gravitational potential energy of the outer orbit boosts its angular velocity and makes it greater than the angular velocity of the inner orbit as they cross the  $x$  axis, thus effectively producing a viscous force that tries to increase the velocity of the inner orbit, mimicking what would happen in the circular case for  $d\Omega/dR > 0$ .

The above considerations lead to the following conclusions:

- (i) The fact that  $x_2$  orbits are non-circular implies that there are regions of ‘reversed shear’ (orange in Fig. 3).
- (ii) It is likely that this leads to the existence of regions that are unstable for the formation of rings and where a ring could survive without spreading in time despite dissipation.

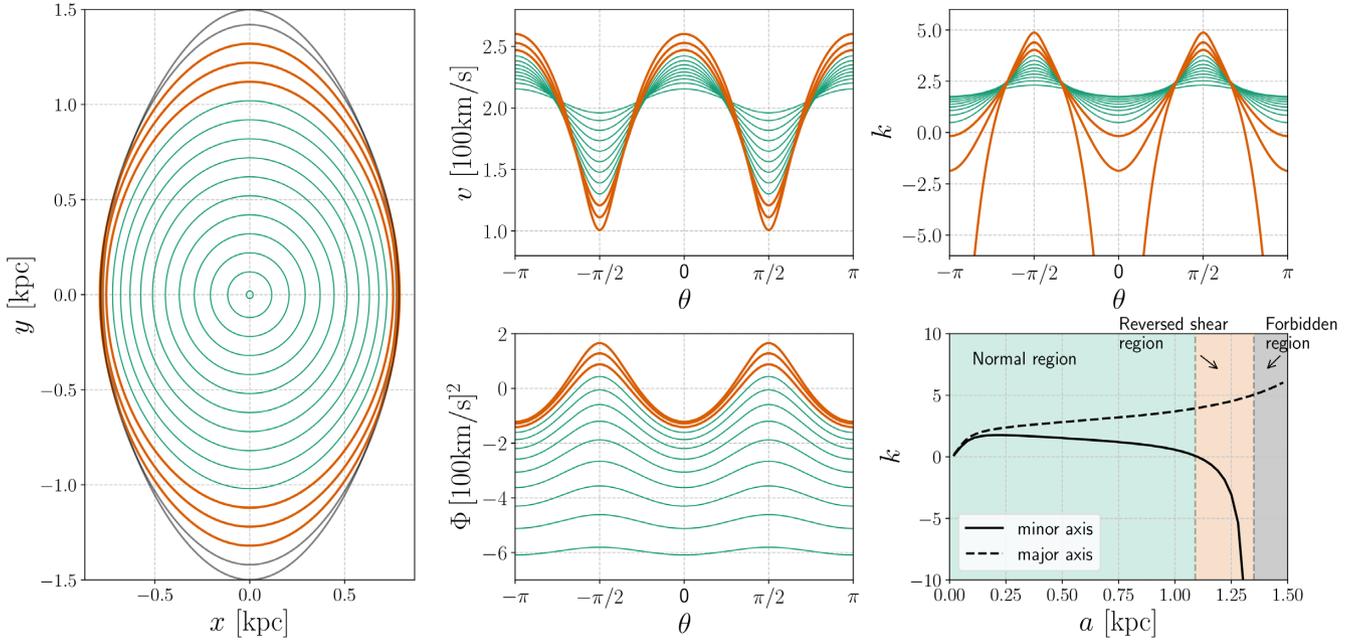
### 3.1 Predicting the size of the ring

In the last section we have argued that regions which are unstable for the formation of rings must be contained within the ‘reversed shear’ region (orange in Fig. 3), which is defined as the region where some reverse shear is present along  $x_2$  orbits. This gives a possible range where a ring may form. However, it is unlikely that rings can form throughout the whole orange region, because the ‘reverse shear’ needs not only to be present, but also to be dominant. Thus, it is more likely that rings will form only in a subregion within the orange region. Can we be more precise and narrow down this subregion, so as to predict more precisely where a ring should form?

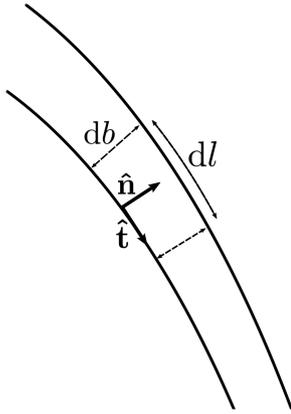
The considerations of the previous section only apply locally along an orbit. But along the same  $x_2$  orbit, there can be both portions of normal and reversed shear. To determine the average behaviour of the gas and whether the net transport of angular momentum is directed outwards or inwards, one needs to integrate  $\mathbf{k}$  along a full orbit contour, i.e. to calculate

$$\mathbf{K} = \int \rho \mathbf{k} dl, \quad (8)$$

where  $dl$  is the length of the element of the contour (see Fig. 4). Since we are in two dimensions we can write  $\mathbf{K} = K \hat{\mathbf{z}}$ . The quantity  $K$  can be thought of as the flux of angular momentum through the contour defined by an entire orbit. We use the convention that  $K > 0$



**Figure 3.** *Left-hand panel:*  $x_2$  orbits calculated in the barred potential described in Section 3. The bar major axis is horizontal. Green indicates orbits in which shear is always as in a ‘normal’ accretion disc, while orange indicates orbits in which portions of ‘reversed shear’ are present. Grey indicates ‘forbidden’ orbits (see text). *Top-middle panel:* modulus of the velocity calculated in a frame rotating at  $\Omega_p$  for the same orbits shown in the left-hand panel. *Bottom-middle panel:* gravitational potential along the orbits. *Top right-hand panel:* the value of  $k$  along orbits, i.e. the torque per unit length and per unit surface density that is exerted on orbit contours (see equation 6). Units are normalized so that  $v = 1$ , and are  $[k] = 100 \text{ km s}^{-1}$ . *Bottom right-hand panel:* values of  $k$  along the minor and major axis as a function of the semimajor axis of the orbits. When the full line is negative, a portion of the orbit with reversed shear exists.

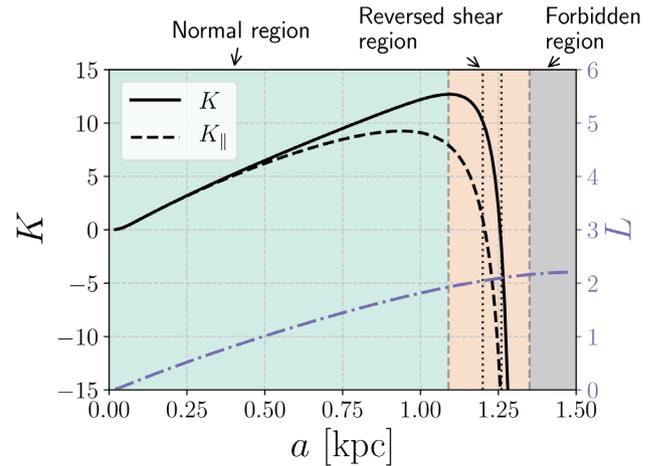


**Figure 4.** Schematic diagram showing a portion of two adjacent orbits.

means angular momentum flux is directed outwards. Therefore, the difference  $\delta K = K(a) - K(a + da)$  gives the net angular momentum gained by the annulus comprised by the orbits with semimajor axis  $a$  and  $a + da$ . When  $K$  is increasing (decreasing), the annulus loses (gains) angular momentum. Therefore, where  $K$  is increasing (decreasing) matter moves inwards (outwards). Hence, the curve  $K(a)$  potentially contains interesting information about locations where a ring could form. However, as we shall now see, to complicate things is the fact that  $K$  depends on  $\rho$ , as well on the velocity field  $\mathbf{v}$ .

The solid line in Fig. 5 plots the curve  $K(a)$  calculated under the following assumptions:

- (i) The velocity field is exactly that of  $x_2$  orbits.
- (ii)  $v = \text{constant}$ .



**Figure 5.** *Full black line:* the flux of angular momentum through orbit contours as a function of the semimajor axis (see equation 8). *Black dashed line:* the flux of angular momentum calculated taking into account only the force tangential to  $x_2$  orbits and neglecting the component perpendicular to them (see equation 9). Both  $K$  curves are calculated under the assumption that  $v = \text{constant}$  and  $\rho = \text{constant}$  along the  $y$  axis. Units are normalized such that  $\rho = v = 1$  and are  $[K] = 100 \text{ km s}^{-1} \text{ kpc}$ . *Blue dashed line:* the angular momentum per unit mass of an infinitely thin annulus comprised between two  $x_2$  orbits, defined as  $L = \int (\mathbf{v} \times \mathbf{r}) dm / \int dm$  where  $dm$  is a mass element along the annulus and  $\mathbf{v}$  is the velocity in an inertial frame, as a function of its semimajor axis. Units are  $[L] = 100 \text{ km s}^{-1} \text{ kpc}$ .

- (iii) The density distribution is of the type described by equation (3) with the mass scaling (i.e. the values of  $\alpha$ ) chosen such that  $\rho = \text{constant}$  along the  $y$  axis.

From this figure we infer that if a disc is set up with this initial configuration, material in the region where  $dK/da > 0$  will move inwards, while material where  $dK/da < 0$  will move outwards. Hence, we expect a gap to open approximately between the green and orange region, and a ring to form somewhere in the orange region. However, these considerations only apply at the initial instant. It is not easy to determine the subsequent development of the  $K(a)$  curve in detail because (i) the density distribution will change, (ii) viscous (and pressure, if present) forces will deform the orbits, so that the velocity field will deviate from that of  $x_2$  orbits. Both these facts will affect the curve  $K(a)$ .

While a redistribution of  $\rho$  changes the shape of the curve  $K(a)$ , the point  $K = 0$  seems of particular significance. It is clear from the definition (8) that this point must remain fixed if the density is redistributed while being forced to remain a steady state of the type described by (3). Under this assumption, also the regions where the curve  $K(a)$  is positive (negative) do not change. Hence, we must always have  $dK/da < 0$  at the point where  $K = 0$  and we expect an accumulation of mass just outside the  $K = 0$  orbit. Thus we might expect that a ring should form with approximately the size and shape of the orbit with semimajor axis  $a_{\text{ring}}$  such that  $K(a_{\text{ring}}) = 0$ .

However, the numerical experiments that we have performed show that this usually overestimates the size of the ring. The discrepancy is due to the fact that in the presence of viscous and pressure forces, even in the most controlled and idealized situations it is not exactly true that the density remains exactly of the type described by (3) and that the velocity remains exactly that of  $x_2$  orbits.

Numerical simulations suggest that a related quantity that seems to do a better job in predicting the size of the ring is

$$\mathbf{K}_{\parallel} = \int \rho \mathbf{k}_{\parallel} dl, \quad (9)$$

where  $\mathbf{k}_{\parallel} = \mathbf{r} \times \mathbf{f}_{\parallel}$  is the torque calculated using only the component of the viscous force that is parallel to the  $x_2$  orbits and neglecting the component perpendicular to the orbit, i.e.

$$\mathbf{f}_{\parallel} = (\mathbf{f} \cdot \hat{\mathbf{t}}) \hat{\mathbf{t}}. \quad (10)$$

See Fig. 4 for the definition of  $\hat{\mathbf{t}}$ . The dashed black line in Fig. 5 shows  $K_{\parallel}$  for comparison with  $K$ . Numerical simulations suggest that a ring forms where  $K_{\parallel} = 0$ . Note that in the circular case  $\mathbf{k} = \mathbf{k}_{\parallel}$ , since the force perpendicular to orbits is directed radially and produces no torque. Note also that in the non-circular case,  $\mathbf{f}$  depends on the factor 2/3 on the right side of (5), which is related to the ratio of shear to bulk viscosity, while  $\mathbf{f}_{\parallel}$  is independent of this value.

What is the justification for neglecting the component of the force perpendicular to the orbits? This component acts in the direction perpendicular to the orbits, so it works by pushing the contours of the orbits to deform them. This force does not try to ‘speed up’ or ‘slow down’ orbits, and does not correspond to our intuitive intuition of a viscous force that is related to layers of fluid sliding relative to each other. Thus it may be that this component acts, in essence, as some sort of pressure, and only causes a slight readjustment in the shape of the orbit until it becomes unimportant. This is speculative and ultimately based on the empirical fact that based on our numerical experiments the location  $K_{\parallel} = 0$  seems to be a good predictor in predicting where a ring forms (see Section 4.3). More tests with different potentials are needed to verify whether this is really true, and if so why.

### 3.2 Viscous versus gravitational torques

It is instructive to compare gravitational and viscous torques. The right-hand panel in Fig. 6 shows what is sometimes referred to as the ‘butterfly diagram’, i.e. it shows the gravitational torque per unit mass, given by

$$\tau_g = \frac{\partial \Phi}{\partial \theta}. \quad (11)$$

The left-hand panel shows the analogous quantity related to viscous forces, i.e. it shows the viscous torques per unit mass:

$$\tau_v = \frac{1}{\rho} \left\{ \partial_x [\rho (x\sigma_{yx} - y\sigma_{xx})] + \partial_y [\rho (x\sigma_{yy} - y\sigma_{xy})] \right\}. \quad (12)$$

A derivation of equation (12) is given in Appendix B. The quantity  $\tau_v$  differs from  $k$  (see equation 6) in that the latter is the torque exerted on contours, while the former is the net angular momentum acting on fluid elements and corresponds to  $k$  integrated over an infinitesimally small closed contour. Unlike the gravitational torques, the viscous torques depend on the density and velocity distributions. The left-hand panel in Fig. 6 is constructed under the same assumptions used to construct the  $K$  curves in Fig. 5, namely  $\rho = \text{constant}$  along the  $y$  axis and  $v = \text{constant}$ . The numbers in the colour bar are calculated assuming a value of  $v = 10^{-4} \times 100 \text{ km s}^{-1} \text{ kpc}$ , which is the same used for simulations in Section 4.1. This is a fairly low value which is a factor of 30 below estimates of turbulent viscosity in the nuclear regions of real galaxies (see Section 5.1).

Blue regions in the left-hand panel correspond to ‘normal shear’, while the red sides in the left-hand panel are regions of ‘reversed shear’. Gas orbiting in the blue region of the left-hand panel loses angular momentum due to viscous torques, while gas in the red regions gains angular momentum. While the instantaneous value of gravitational torques is usually much larger than the instantaneous value of viscous torques, the time average of the former over an  $x_2$  orbit is always zero, while the average of the latter is usually non-zero.

## 4 NUMERICAL EXPERIMENTS

### 4.1 Isothermal viscous rings

The first test that we perform is the analogue of the standard viscous ring spreading problem (e.g. fig. 1 of Pringle 1981) but instead of the Kepler potential we use our barred gravitational potential. The simulations are 2D, isothermal, and include an explicit viscosity term. The gas moves in a rigidly rotating externally imposed barred potential and self-gravity of the gas is neglected. Thus, the equations of motion in an inertial frame are the continuity equation and the Navier–Stokes equation

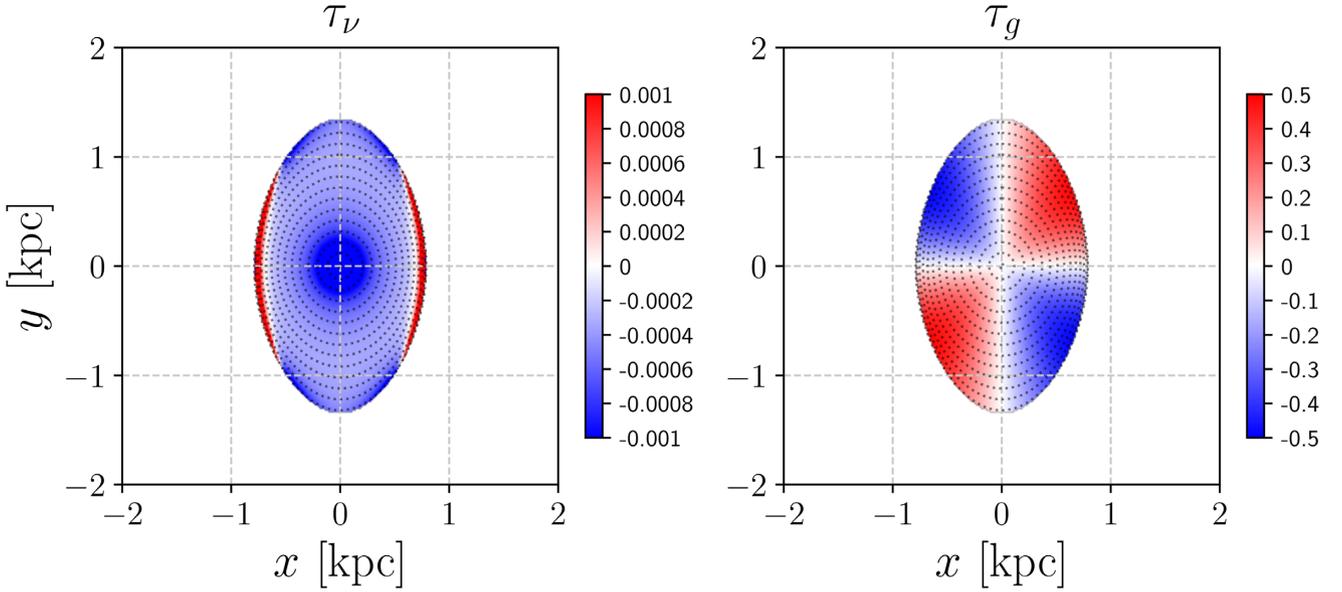
$$\partial_t \rho + \partial_i (\rho v_i) = 0, \quad (13)$$

$$\partial_t v_i + (v_j \partial_j) v_i = -\frac{1}{\rho} \partial_i P - \frac{1}{\rho} \partial_j (\rho \sigma_{ij}) - \partial_i \Phi, \quad (14)$$

where  $\sigma_{ij}$  is given by equation (5),  $\Phi$  is the external gravitational potential, and  $P = c_s^2 \rho$  where  $c_s$  is a constant.

#### 4.1.1 Numerical code

We use the public code PLUTO (Mignone et al. 2007). This is a free software for the numerical solution of systems of conservation laws targeting high Mach number flows in astrophysical fluid dynamics. We use a 2D static polar grid in the region  $R \times \theta = [0.1, 2.0] \text{ kpc} \times$



**Figure 6.** Comparison of viscous and gravitational torques. *Left-hand panel:* viscous torques calculated using equation (12) assuming that  $\nu = 10^{-4} \times 100 \text{ km s}^{-1} \text{ kpc}$  and that  $\rho$  is of the type given by equation (3) with mass normalization such that  $\rho = \text{constant}$  along the  $y$  axis. *Right-hand panel:* gravitational torques per unit mass, see equation (11). Dots delineate the  $x_2$  orbits (left-hand panel of Fig. 3). Units are  $[\tau] = (100 \text{ km s}^{-1})^2$ .

$[0, 2\pi]$ . The grid is uniformly spaced in both  $R$  and  $\theta$  with  $950 \times 1440$  points, which corresponds to a resolution of  $2 \text{ pc}$  in radius and  $0.25 \text{ deg}$  in polar angle. We use the following parameters: RK2 time-stepping, no dimensional splitting, HLL Riemann solver, and the VAN ALBADA flux limiter. Boundary conditions are reflective on the inner boundary at  $R = 0.1 \text{ kpc}$  and outflow on the outer boundary at  $R = 2.0 \text{ kpc}$ . Viscosity is treated with the EXPLICIT scheme.

#### 4.1.2 Initial conditions

The initial density distribution is taken to be of the type described by equation (3). The function  $\alpha$  is chosen so that the density distribution along the  $y$  axis is the sum of one or more gaussians of the following type:

$$\rho_0(a) = \exp \left[ - \left( \frac{a - a_0}{\Delta a} \right)^2 \right], \quad (15)$$

where  $\Delta a = 50 \text{ pc}$  and  $a_0$  is a parameter that controls where the ring is centred. The initial velocity distribution is taken to be that of  $x_2$  orbits in runs with non-axisymmetric potentials, or the circular velocity  $v_{\text{circ}}(R) = \sqrt{R d\Phi_0/dR}$  in runs with axisymmetric potentials. In regions where  $x_2$  orbits do not exist, we have used a nearest-neighbour interpolation to assign the velocity of the closest  $x_2$  orbit. This makes no difference in practice because those regions were always initially void of gas, and the density was set to a negligibly small value of  $\rho = 10^{-12}$ . The density units are arbitrary since the equations of motions (13) and (14) are invariant under density rescaling.

#### 4.1.3 Results

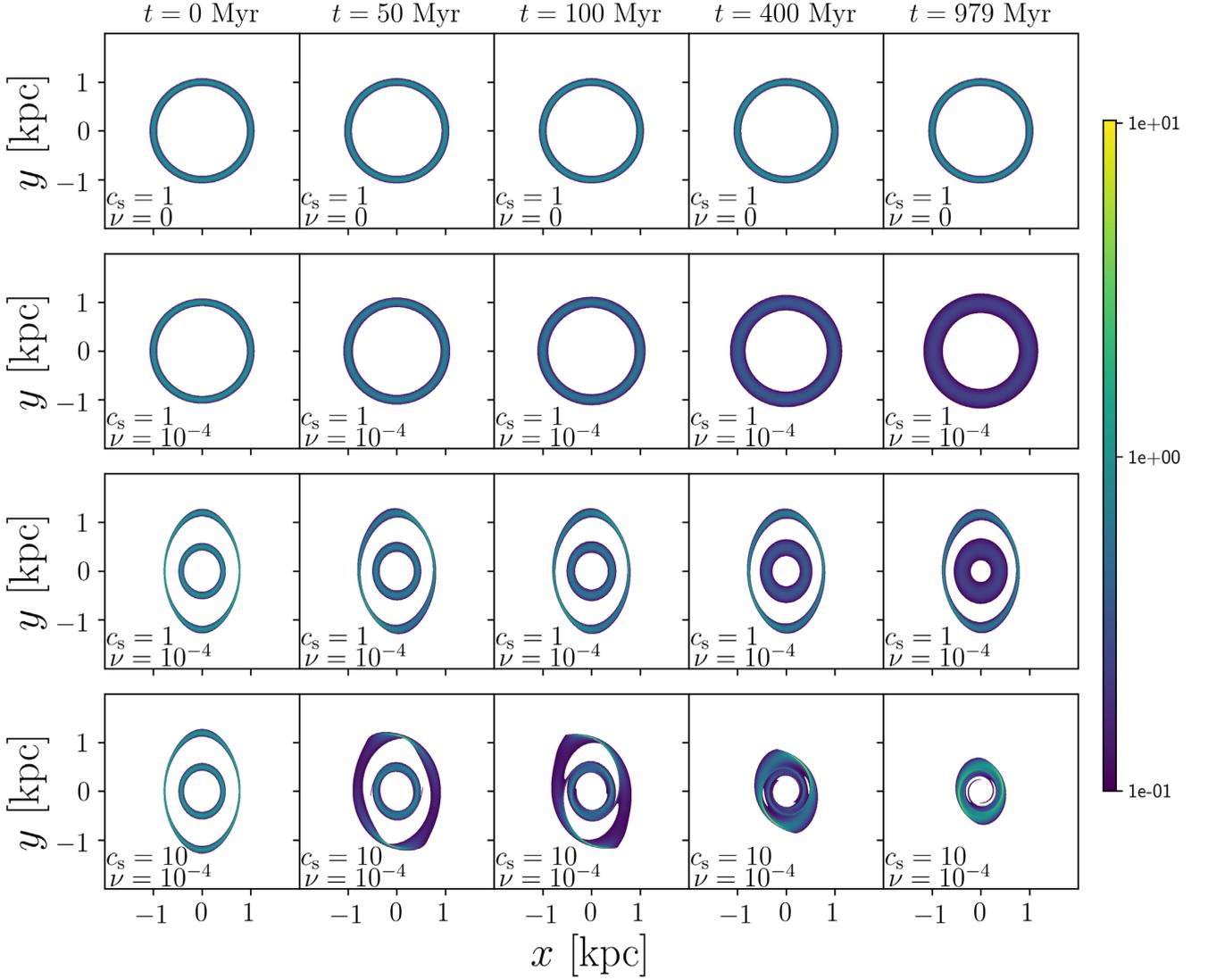
Fig. 7 shows the time evolution of the 2D surface density for four different simulations with different potentials (barred or axisymmetric) and different values of  $\nu$  and  $c_s$ . Fig. 8 shows cuts of the surface density along the  $x$  axis for the same snapshots.

The first row shows the evolution of a ring in an axisymmetric potential when  $\nu = 0$ . The ring is centred on  $a_0 = 1 \text{ kpc}$  and the potential is given by equation (2) (see the circular velocity curve in Fig. 2). The sound speed is  $c_s = 1 \text{ km s}^{-1}$ . As expected, in the absence of viscosity the ring does not spread. Not much happens. This also shows that the numerical viscosity is negligibly small.

The second row shows what happens when we turn on the viscosity in the axisymmetric case. Everything is the same as in the first row except that viscosity now has a value  $\nu = 10^{-4} \times 100 \text{ km s}^{-1} \text{ kpc} \simeq 3 \times 10^{24} \text{ cm}^2 \text{ s}^{-1}$ . Viscosity makes the ring spread, exactly as we would expect in analogy to the standard viscous ring spreading problem (compare for example the second row in Fig. 8 with fig. 1 of Pringle 1981). The value of viscosity used here is fairly low, roughly a factor of 30 lower than estimates for real galaxies based on turbulent viscosity (see Section 5.1). As we will discuss in Section 5.1, this shows that a relatively low value of  $\nu$  can already produce significant effects over a time-scale of  $1 \text{ Gyr}$ .

The third row shows the evolution of two rings in the same barred potential used to construct Fig. 3. The axisymmetric part  $\Phi_0$  of this potential is exactly the same as in the first two rows, but in addition now also the quadrupole  $\Phi_2$  is present (see the blue dashed line in Fig. 2 and equation 1). The two rings are initially centred on  $a_0 = 0.4 \text{ kpc}$  and  $a_0 = 1.2 \text{ kpc}$ , respectively, i.e. the inner one is in the middle of the green region in Fig. 3 while the outer one is in the middle of the orange region. The inner ring spreads, similar to the circular case (second row). The outer ring instead, after an initial adjustment, does not spread and at  $t = 979 \text{ Myr}$  still has the same width it had at  $t = 50 \text{ Myr}$  (see third row in Fig. 8). This is in remarkable agreement with the idea described in Section 3.

The fourth row shows a simulation that is identical to that of the third row except that the sound speed is increased to  $c_s = 10 \text{ km s}^{-1}$ . This time, the outer ring does not survive. Pressure forces quickly create spiral shocks, which destroy the outer ring and make material on the outer ring plunge inwards and merge with the inner ring. Thus whether the mechanism discussed in Section 3 can work depends



**Figure 7.** Simulations of gaseous rings for different potentials and values of  $c_s$  and  $\nu$ . Each panel shows the gas surface density and time increases from left to right as indicated at the top of the figure. *First row:* one ring is placed in a purely axisymmetric potential (equation 2) and viscosity is turned off,  $\nu = 0$ . The ring does not spread. *Second row:* same as the first row, but viscosity is set to a value  $\nu = 10^{-4} \times 100 \text{ km s}^{-1} \text{ kpc} \simeq 3 \times 10^{24} \text{ cm}^2 \text{ s}^{-1}$ . The ring spreads due to viscous torques. *Third row:* two rings are placed in a barred potential given by (1) in the presence of viscosity. The inner ring (placed in the green region in Fig. 3) spreads, similar to the axisymmetric case in the second row, while the outer ring (placed in the orange region in Fig. 3) does not spread despite the presence of viscosity. *Fourth row:* same as third row, but for a higher value of the sound speed  $c_s$ . The ring is destroyed by spirals created by pressure. In all simulations the initial conditions are given such that in the absence of pressure and viscous forces it would be a steady state. Gas circulates clockwise and the bar major axis is horizontal.

on the relative importance of pressure forces. In order to understand the effects of pressure we have run a further isothermal simulation that is described in the next section. The effects of pressure will be discussed in Section 5.2.

## 4.2 Isothermal disc

In order to better understand the effects of pressure, we have run a further isothermal simulation in which we have set up a filled disc of gas on  $x_2$  orbits. The numerical set-up is identical to that of Section 4.1, except for the initial conditions. The sound speed of this simulation is  $c_s = 10 \text{ km s}^{-1}$  while the viscosity is turned off,  $\nu = 0$ .

### 4.2.1 Initial conditions

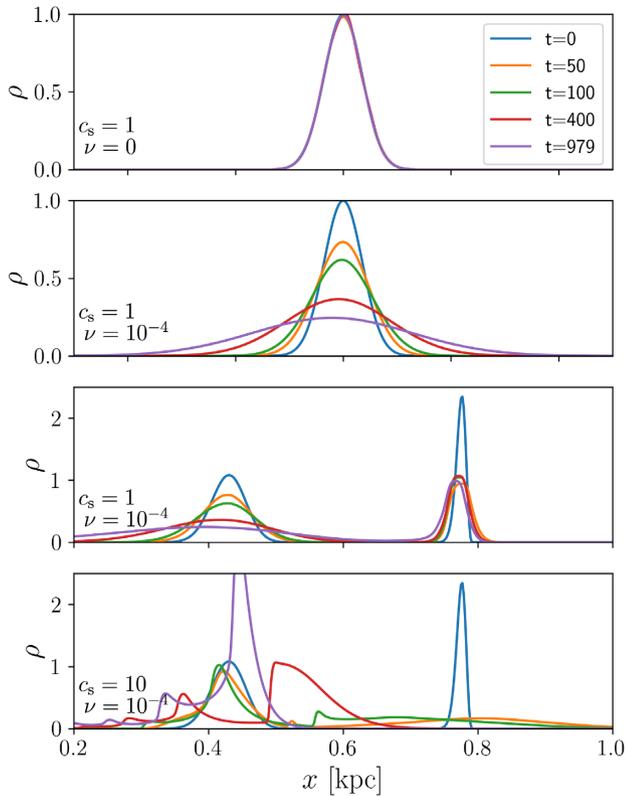
The initial density distribution is such that in the absence of pressure and viscous forces, it would be a steady state of the type described by equation (3). The normalization  $\alpha$  is chosen so that the density distribution along the  $y$  axis is

$$\rho_0(a) = \begin{cases} 1 & \text{if } a \leq 1.35 \text{ kpc,} \\ 0 & \text{if } a > 1.35 \text{ kpc.} \end{cases} \quad (16)$$

The value  $a = 1.35 \text{ kpc}$  corresponds to the orbit at which the forbidden region starts (cf. the grey area in Fig. 3).

### 4.2.2 Results

Fig. 9 shows the time evolution of the surface density, while Fig. 10



**Figure 8.** Density cuts along the  $x$  axis for the same simulations shown in Fig. 7.

shows in more detail the snapshot at  $t = 50$  Myr. Pressure forces cause the shape of the gas streamlines to deviate from  $x_2$  orbits almost immediately, creating spiral shocks shortly after the start of the simulation. That these are really shocks as opposed to generic density waves can be seen from Fig. 11, which shows cuts of the density and the velocity along the red dashed orbit shown in Figs 9 and 10. These cuts show that density and velocity have discontinuous jumps, which is the defining characteristics of shocks. They also show that the shock front moves upstream with respect to the overall flow (i.e. it moves to the left in Fig. 11). At the resolution of the simulation shown in Fig. 9 the spiral shocks appear smooth, but we have verified that at higher resolution the shock fronts develop wiggles and break up as a consequence of the wobble instability (Wada & Koda 2004; Kim, Kim & Kim 2014; Sormani et al. 2017). This instability is not important for the discussion in this paper.

A notable feature of the simulation that is clearly visible in Fig. 9 is that the disc shrinks with time. This happens because for an extended period of time, which lasts approximately for  $0 \lesssim t \lesssim 800$  Myr, the overall surface density distribution is not symmetric with respect to the  $y$  axis but is tilted so that much more material lies in the two quadrants ( $x < 0, y > 0$ ) and ( $x > 0, y < 0$ ) than in the other two quadrants (e.g. it is clear that at  $t = 50$  Myr the major axis of the gas disc makes an angle with the  $y$  axis and points towards negative  $x$ ). Since gravitational torques from the bar potential act in such a way as to remove angular momentum from material in these two quadrants, the total angular momentum steadily decreases during the tilted phase and the disc shrinks. After this phase the disc returns to a more symmetric configuration (see the final panel at  $t = 979$  Myr in Fig. 9) and the angular momentum stabilizes to a value approximately 60 per cent of the initial value. Now the bar is unable to remove further angular momentum, and the final state

has a fat ring with  $R \sim 0.4$  kpc, but this ring has nothing to do with the mechanism described in Section 3.<sup>5</sup> The physical reasons that cause the tilt and hence the removal of angular momentum can be traced back to the spiral shocks (see Section 5.2).

Note that based on the resonant theory, we might have naively expected exactly the opposite, i.e. that the gas *gains* angular momentum, because the spiral pattern is mostly in the quadrants ( $x > 0, y > 0$ ) and ( $x < 0, y < 0$ ), where the torque is positive. This shows that expectations based on the naive interpretation of the torques acting on the spiral pattern should be taken with great caution, and explains why predictions of the resonant theory are often not verified in actual simulations (Kim et al. 2012).

We have repeated this simulation using a value of the viscosity  $\nu = 10^{-4} \times 100 \text{ km s}^{-1} \text{ kpc}$ , but the results are almost indistinguishable. The effects of pressure in this case completely overwhelm the effects of viscosity, which is negligible. This also shows that what creates the spirals is pressure, not viscosity. We continue the discussion of the effects of pressure in Section 5.2.

### 4.3 Non-isothermal simulation

The simulations discussed in the previous two sections are idealized numerical experiments and are far from realistic. As a more realistic example let us consider here the simulations of Sormani et al. (2018). These authors used high resolution, 3D hydrodynamical simulations performed using the moving-mesh code AREPO (Springel 2010) to simulate the central regions of the Milky Way. In these simulations gas moves in a multicomponent externally imposed rotating barred gravitational potential which is tuned to reproduce the characteristics of our Galaxy. The simulation box includes the whole region  $R \leq 10$  kpc and thus the entire bar region, not just the nuclear region.

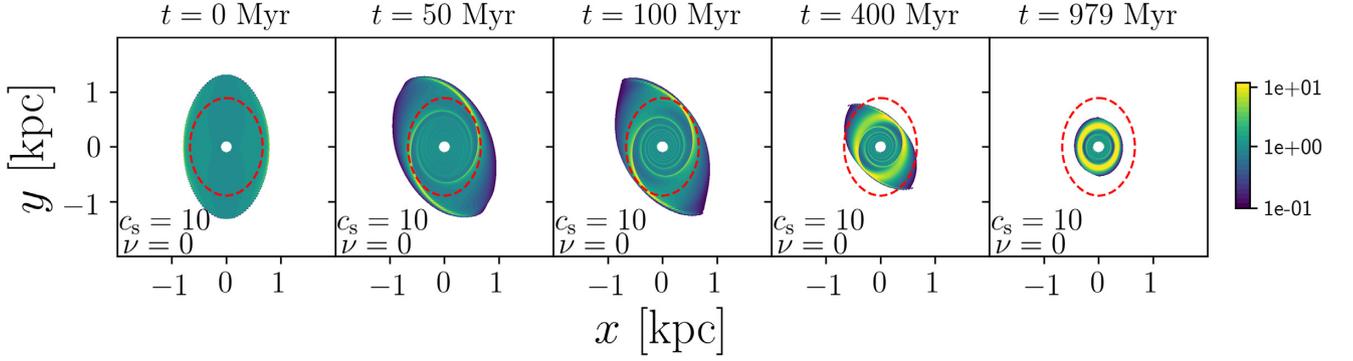
A key feature of the simulations of Sormani et al. (2018) is that they include a time-dependent chemical network that keeps track of the chemical composition of the gas (essentially hydrogen, carbon, and oxygen chemistry). The chemical network is deeply coupled to the heating and cooling processes of the ISM. Thus, instead of assuming a simple equation of state, the temperature of the gas is calculated using a cooling function that takes into account several sources of heating and cooling, which include contributions that depend on the chemical composition of the gas and the instantaneous rate of chemical reactions (e.g. H collisional excitation and ionization,  $\text{H}^+$  recombination,  $\text{H}_2$  formation, etc), and external contributions (e.g. heating from the average UV interstellar radiation field, cosmic rays).<sup>6</sup> The net result is a two-phase ISM made of a cold component with  $T \simeq 100\text{K}$  (which corresponds to  $c_s \lesssim 1 \text{ km s}^{-1}$ ) and a warm component with  $T \simeq 10^4\text{K}$  (which corresponds to  $c_s \simeq 10 \text{ km s}^{-1}$ ).

The large-scale dynamics in these simulations follows the typical dynamics of gas flowing in a barred potential (e.g. Sellwood & Wilkinson 1993). They are set up as follows.<sup>7</sup> To avoid transients,

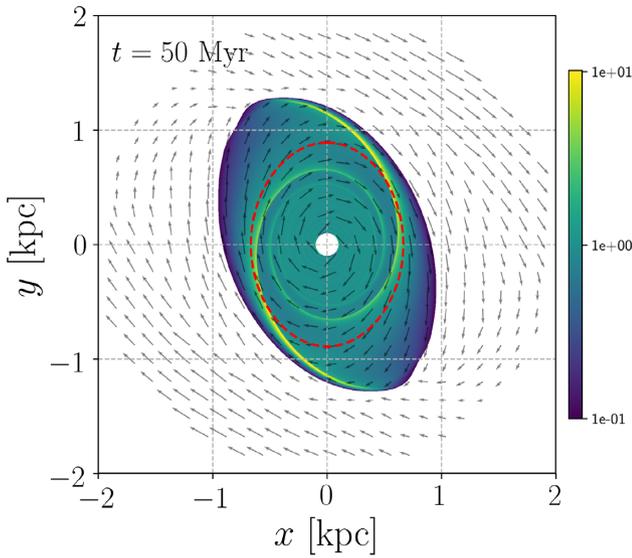
<sup>5</sup>The final state will slowly shrink over time due to viscous torques acting as in a standard accretion disc, but this shrinking is extremely slow because numerical viscosity is extremely small (first row of Fig. 7).

<sup>6</sup>See tables 1 and 2 in Sormani et al. (2018) for the full list of reactions included in the chemical model and processes included in the cooling function.

<sup>7</sup>Movies are available at [http://www.ita.uni-heidelberg.de/~mattia/videos/asymmetry/high/rhproj\\_barframe.mp4](http://www.ita.uni-heidelberg.de/~mattia/videos/asymmetry/high/rhproj_barframe.mp4) (whole bar region) and [http://www.ita.uni-heidelberg.de/~mattia/videos/asymmetry/high/rhprojz\\_oom\\_barframe.mp4](http://www.ita.uni-heidelberg.de/~mattia/videos/asymmetry/high/rhprojz_oom_barframe.mp4) (zoom on the nuclear region).



**Figure 9.** Evolution of the surface density for an idealized isothermal simulation with  $c_s = 10 \text{ km s}^{-1}$ ,  $\nu = 0$ . Initial conditions are given such that  $\rho = 1$  along the  $y$  axis and that in the absence of pressure and viscous forces it would be a steady state. The dashed red line shows the orbit examined in more detail in Fig. 11.



**Figure 10.** Velocity field superimposed on the surface density for the snapshot at  $t = 50 \text{ Myr}$  of the simulation shown in Fig. 9.

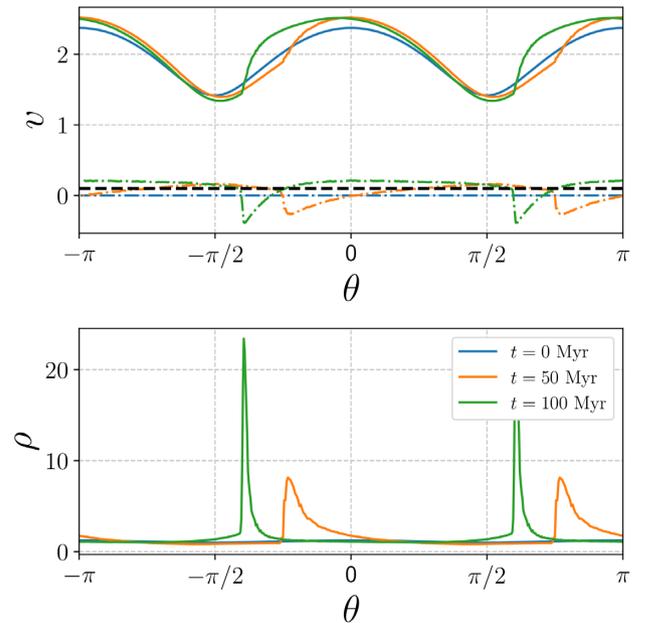
the gas is initially on circular orbits and the bar is introduced gradually during the first 150 Myr. As the bar grows, material in the outer parts ( $R \gtrsim 3 \text{ kpc}$ ) deviates from circular orbits and starts following  $x_1$  orbits, which is a family of orbits elongated parallel to the bar. Two narrow large-scale shocks are formed along which gas plunges in a dynamical time from  $R \sim 3 \text{ kpc}$  down to the nuclear region, where it settles on  $x_2$  orbits.

A nuclear ring forms in these simulations with semimajor axis  $\sim 400 \text{ pc}$ . This is shown in Fig. 12 (see also last panel in fig. 7 of Sormani et al. 2018). It has the following striking characteristics:

- (i) The ring is extremely dense and thin.
- (ii) The ring does not spread over time.
- (iii) The ring is extremely long lived and its radius does not change in time even if the simulation is continued to  $t \simeq 1 \text{ Gyr}$ .

Thus the ring seems to have all the characteristics expected from the mechanism described in Section 3. Indeed this is not a coincidence, since the study in this paper was inspired by the simulations of Sormani et al. (2018). Note also that in the simulations of Sormani et al. (2018) self-gravity of the gas is ignored, so the ring self-gravity cannot play a role in determining its thinness.

Figs 13, 14 and 15 show the  $x_2$  orbits, the curve  $K$  and the torques



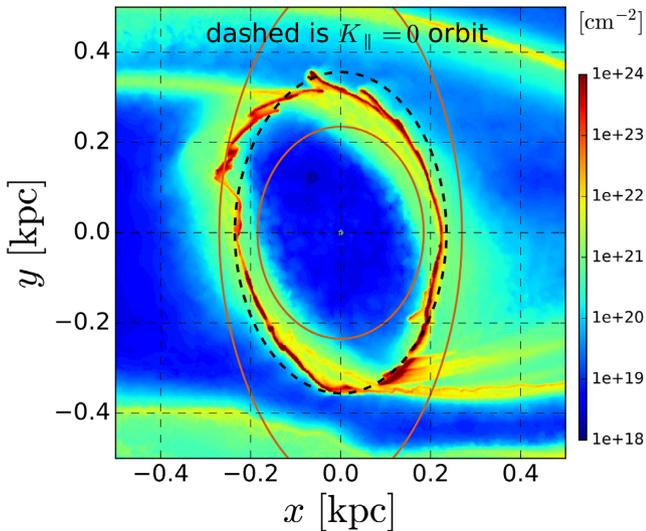
**Figure 11.** Velocity and density along the dashed  $x_2$  orbit in Figs 9 and 10. In the top panel, the full lines are the component of the velocity parallel to the orbit, the dot-dashed lines are the components of the velocity perpendicular to the orbit, and the black dashed line indicates the value of the sound speed.

for the potential used in Sormani et al. (2018). In accordance with the discussion in Section 3, the ring forms in the middle of the reversed shear region (orange in Fig. 13 and delimited by two orange orbits in Fig. 12). Fig. 12 shows that the  $K_{11} = 0$  orbit fits very well with both the size and the shape of the ring. This is a promising confirmation of the considerations in Section 3.1, but more tests with different potentials are needed to be certain that this is not a coincidence.

## 5 DISCUSSION

### 5.1 Viscosity

We have assumed that viscosity is present in our theory, but so far we have said nothing about the source of this viscosity. Let us first give a rough estimate of the minimum amount of viscosity needed to produce significant effects. Viscous stresses can significantly affect the dynamics of the system if the viscous time-scale,  $t_v \approx R^2/\nu$  (where  $R$  is the typical dimension of the ring), is shorter than the



**Figure 12.** Total surface density at  $t = 294$  Myr of the ‘high’ simulation of Sormani et al. (2018). The dashed line corresponds to the  $K_{\parallel} = 0$  orbit. The two orange orbits are the orbits that delimit the orange region in Fig. 13. Gas rotates clockwise.

Hubble time,  $t_H = 14$  Gyr. This gives a minimum coefficient of

$$\nu_{\min} \approx \frac{R^2}{T_H} = 2 \times 10^{25} \left( \frac{R}{1 \text{ kpc}} \right)^2 \left( \frac{t_H}{14 \text{ Gyr}} \right)^{-1} \text{ cm}^2 \text{ s}^{-1}. \quad (17)$$

What is the viscosity of the real ISM? The molecular viscosity is negligible. But there are random turbulent motions of interstellar clouds, maintained by stellar feedback and MHD instabilities (e.g. Sellwood & Balbus 1999; Mac Low & Klessen 2004). These give rise to turbulent viscosity. Using the following typical parameters, namely (i) a cloud size  $a \sim 10$  pc, (ii) a velocity dispersion  $c \sim 8 \text{ km s}^{-1}$ , (iii) a fraction  $f \sim 0.1$  of the volume filled by clouds, Lynden-Bell & Pringle (1974) estimated the turbulent viscosity coefficient as

$$\nu \approx \frac{1}{3} c \frac{a}{f} = 8 \times 10^{25} \text{ cm}^2 \text{ s}^{-1}. \quad (18)$$

Since  $\nu > \nu_{\min}$ , we conclude that viscous stresses play an important role in shaping the gas distribution in the central  $\sim 1$  kpc of the Galaxy.

Note that the value given by (17) is roughly a factor of 10 greater than the value  $\nu \simeq 3 \times 10^{24} \text{ cm}^2 \text{ s}^{-1}$  assumed for the simulations in Section 4.1, which was already sufficient to cause a significant spreading of an axisymmetric ring over a time-scale of 1 Gyr (see Fig. 7). This suggests viscosity may in general be more effective than naive estimates such as the one indicated above. This fact is also implicitly contained in the definition of the dimensionless parameter  $\tau = 12\nu t R_0^{-2}$  of Pringle (1981; see his equation 2.13), which plays the role of time in the analytical solution of the viscous spreading problem.<sup>8</sup> The numerical factor 12 means that viscosity acts 12 times faster in that problem than the estimate  $t_v \sim R_0^2/\nu$  would suggest.

Two more facts corroborate the conclusion that viscosity must be dynamically significant in the nuclear regions of galaxies. First, as already noted by Lynden-Bell & Pringle, the turbulent viscosity in galactic centres is likely to be even higher than their estimate due to the higher velocity dispersion of molecular clouds there (e.g. Heyer

& Dame 2015). Secondly, the above discussion neglects another source of dissipation, namely shocks. At the locations of ‘reversed shear’ along  $x_2$  orbits, gas streamlines converge, creating a zone of enhanced interaction where material is compressed and shocked. Shocks are a natural way of producing dissipation, and it may be that this contributes in driving the gas towards the local minimum of the energy represented by the ring configuration.

## 5.2 Understanding the effects of pressure

The results of Sections 4.1 and 4.2 suggest that at low sound speed ( $c_s = 1 \text{ km s}^{-1}$ ) a ring created by the mechanism described in Section 3 can survive, while at high sound speed ( $c_s = 10 \text{ km s}^{-1}$ ) the ring cannot survive because it is destroyed by spiral shocks. What is the origin of these spiral shocks?

To understand this, it is useful to consider a very simplified toy problem. Gas on an  $x_2$  orbit revolves periodically around the galactic centre and feels a periodic potential that is an oscillating function of  $\theta$  (see the middle-bottom panel of Fig. 3). The gas is faster (slower) where the potential is low (high) (see the middle-top panel of Fig. 3). Very crudely, we may model this gas as a 1D isothermal fluid that moves in a periodic potential given by

$$\Phi(x) = -2 \cos(2\pi x). \quad (19)$$

The equations of motion are then the 1D continuity and Euler equations:

$$\partial_t \rho + \partial_x (\rho v) = 0 \quad (20)$$

$$\partial_t v + v \partial_x v = -\frac{\partial_x P}{\rho} - \partial_x \Phi, \quad (21)$$

where  $P = c_s^2 \rho$ . In the absence of the pressure term ( $c_s = 0$ ), it is easy to check that the following is a steady state of the system:

$$v(x) = \sqrt{2(B - \Phi(x))} \quad (22)$$

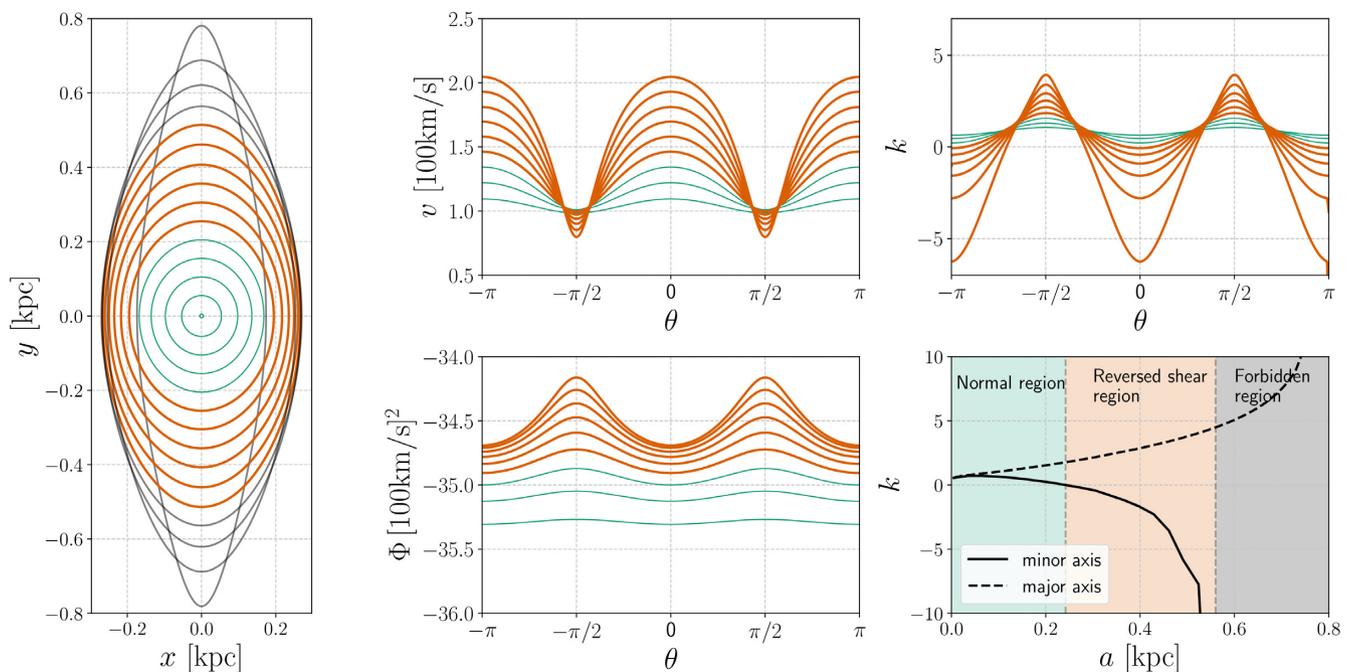
$$\rho(x) = 1/v(x), \quad (23)$$

where  $B$  is a constant. This corresponds to a purely ballistic state, i.e. it is the toy problem version of gas on an  $x_2$  orbit.

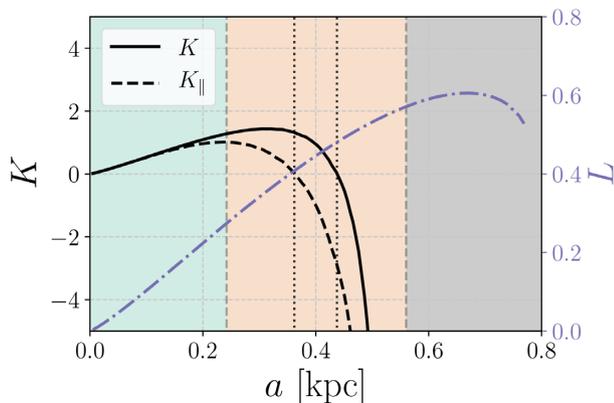
What happens when we introduce pressure? Suppose we release the system from an initial state given by equations (22) and (23). If  $c_s = 0$  nothing happens because the initial state is a steady state. But if  $c_s > 0$ , the system is out of equilibrium and not exactly in a steady state. Fig. 16 shows the subsequent evolution of the system if the initial state is the one with  $B = 2.1$  and the sound speed is  $c_s = 0.02$ . In this case, pressure forces are negligible and the system remains approximately in the initial state. Fig. 17 repeats the same experiment with a value of the sound speed  $c_s = 0.2$ . Now, pressure forces are significant and quickly lead to the development of a shock (see Fig. 17). The shock initially develops close to the point where  $v$  is minimum ( $\Phi$  is maximum) and subsequently moves upstream. The reason why the shock develops is that as the gas tries to ‘climb up’ from the potential well, pressure pushes backwards and slows it down, making the gas ‘stall’ (Prendergast 1983). The condition for the shock to develop is that the sound speed (black dashed line in Figs 16 and 17) is close enough to the minimum of  $v$ . Thus, the shocks occur if the sound speed is high enough (but also low enough that the gas is supersonic) to get close to the minimum of  $v$ . In a system with fixed  $c_s$ , a shock occurs if the oscillations of  $\Phi$  are large enough.

Now we can understand what happens in the case of Fig. 11. The value  $c_s = 10 \text{ km s}^{-1}$  is close enough to the minimum value of  $v$

<sup>8</sup>The solution given by Pringle (1981) is the Green function of the problem.



**Figure 13.** Same as Fig. 3 but for the gravitational potential of Sormani et al. (2018).



**Figure 14.** Same as Fig. 5 but for the gravitational potential of Sormani et al. (2018).

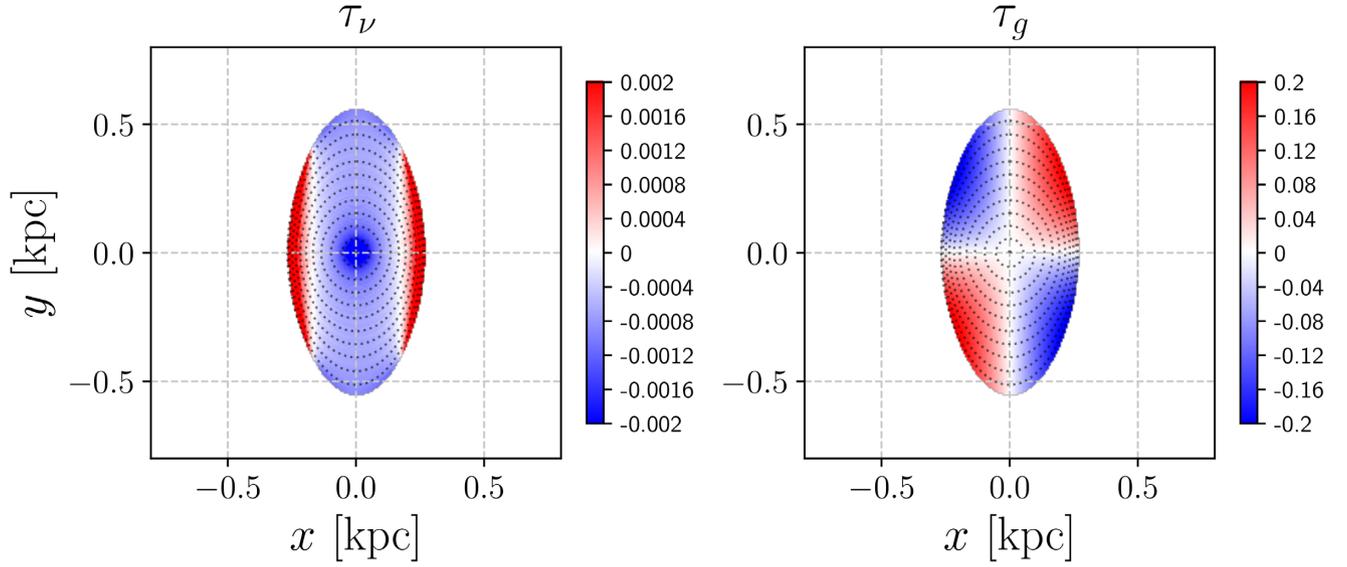
along the  $x_2$  orbit that a shock is triggered. Shocked gas is slowed down, and starts falling to the centre. This is why material in the spiral pattern moves *inwards*, despite the fact that the gravitational torques from the bar potential may suggest the opposite.<sup>9</sup> After reaching its closest approach to the centre during its first infall, gas gains speed (the green curve in Fig. 11 has a lower minimum and a higher maximum velocity than the  $x_2$  orbit velocity given by the blue curve). Thus while close to the shock gas is falling in, it is actually moving out far from the shock (dashed lines in Fig. 11). Meanwhile, the shock moves upstream, as in the toy model. This causes the overall gas distribution to tilt, causing the removal of angular momentum discussed in Section 4.2.

The shock in the simulations of Fig. 9 first appears in the outer parts, and then quickly propagates inwards. This happens because oscillations of  $\Phi$  are stronger in the outer parts (see bottom-middle

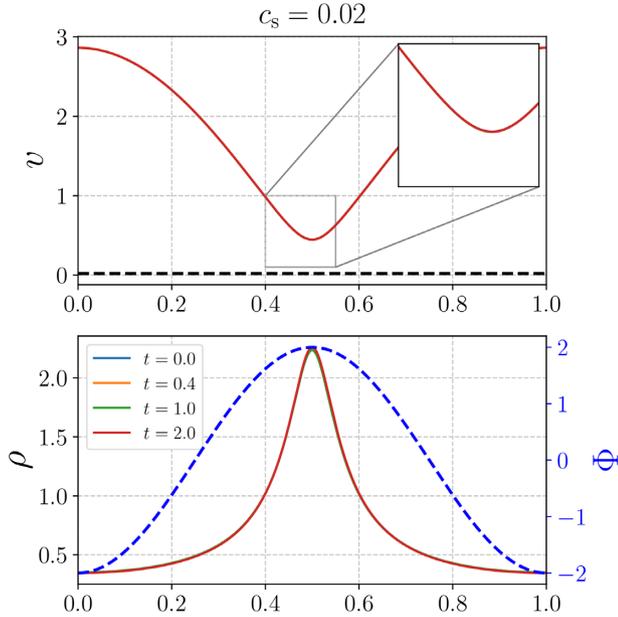
panel of Fig. 3). As we move inwards, at some point the oscillations of  $\Phi$  are not strong enough to trigger a shock, but merely a density wave. Thus, there is a critical radius  $R_c$  above which oscillations of  $\Phi$  are strong enough to trigger a shock (Roberts 1969). The conclusion is that the disc in Fig. 9 shrinks until it is small enough that shocks cannot be triggered anymore because the oscillations of  $\Phi$  are small.

This also explains why if one increases the sound speed the nuclear disc shrinks, a trend which has been observed by many authors (e.g. Englmaier & Gerhard 1997; Patsis & Athanassoula 2000; Kim et al. 2012; Sormani et al. 2015a). When the sound speed is increased, smaller  $\Phi$  oscillations are sufficient to trigger a shock (raising the black dashed in Figs 16 and 17 favours the triggering of shocks). Hence, as the sound speed increases the value of  $R_c$  decreases, shocks can be produced further in and the disc shrinks further according to the considerations in the last paragraph. Other explanations of why the  $x_2$  disc is smaller at higher sound speed also rely on shocks and the loss of angular momentum (Kim et al. 2012; Sormani et al. 2015a). However, these explanation focus on the loss of angular momentum at the tips of the bar, far from the nuclear region. What is different and interesting in the explanation proposed here is that the part of the shock that matters in setting the size of the nuclear disc is that already within the  $x_2$  region. Indeed, the simulation in Fig. 9 does not contain the full larger-scale off-set shock that typically develops in a gas flow in a barred potential, but just its last portion. We have confirmed this interpretation by repeating the simulation shown in Fig. 9 with different values of the sound speed. These show the same trend observed by other authors, i.e. the final size of the disc is smaller at higher sound speeds, but they do not contain the larger scale flow, only the  $x_2$  disc. For example, when  $c_s = 5 \text{ km s}^{-1}$ , spirals develop but the final disc is bigger than in the last panel of Fig. 9. When  $c_s = 1 \text{ km s}^{-1}$ , the pressure is too low to generate spiral shocks – instead, a gap is created approximately between the green and orange region in Fig. 3 as expected based on the considerations in Section 3, and a ring is formed.

<sup>9</sup>Indeed, it is known that spiral shocks can drive disc accretion (e.g. Arzamasov & Rafikov 2018).



**Figure 15.** Same as Fig. 6 but for the gravitational potential of Sormani et al. (2018).



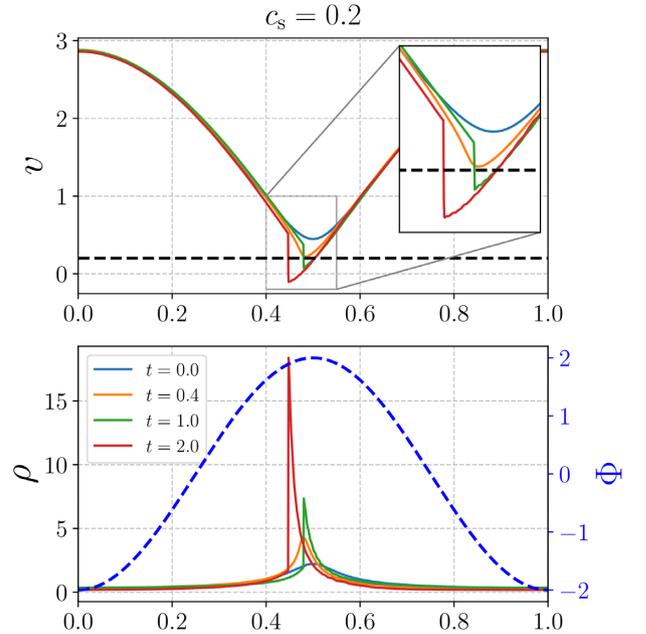
**Figure 16.** Evolution of density and velocity in the 1D toy problem described in Section 5.2 for a low value of the sound speed,  $c_s = 0.02$ . No shocks are formed.

### 5.3 Gravitational torques versus viscous torques

In the description of the mechanism of Section 3 we have assumed that the time average of the net gravitational torque on the disc,

$$\langle \mathcal{T}_g \rangle = \int \langle \rho \rangle \left( \frac{\partial \Phi}{\partial \theta} \right) dx dy, \quad (24)$$

can be considered small. This assumption is correct for a collection of ballistic particles on  $x_2$  orbits, for which  $\mathcal{T}_g$  oscillates periodically in time, and for a set-up such as the one described in Section 3 and used as initial condition in the idealized simulations of Section 4.1, because the density distribution is symmetric with respect to the axis. However, is this assumption justified in a more realistic situation such as the one described in Section 4.3?

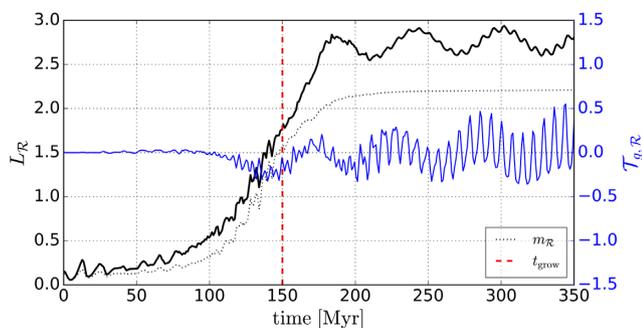


**Figure 17.** Evolution of density and velocity in the 1D toy problem described in Section 5.2 for a high value of the sound speed,  $c_s = 0.2$ . A shock is formed that moves upstream.

The black solid curve in Fig. 18 shows the total angular momentum  $L_{\mathcal{R}}$  possessed by gas in the simulations of Sormani et al. (2018) in the region defined as

$$\mathcal{R} := \left\{ 0.2^2 < \left( \frac{x}{0.4 \text{ kpc}} \right)^2 + \left( \frac{y}{0.5 \text{ kpc}} \right)^2 < 1 \right\} \quad (25)$$

i.e. a region comprised by two ellipses of semimajor axis of 0.1 and 0.5 kpc. This region roughly includes the ring shown in Fig. 12, while excluding outer material flowing on  $x_1$  orbits and the innermost material. The blue solid line is the total gravitational torque  $\mathcal{T}_{g,\mathcal{R}}$  acting on gas inside this region, and the dotted black curve  $m_{\mathcal{R}}$  is the total gas mass.



**Figure 18.** Full black line: total angular momentum contained in the region  $\mathcal{R}$  (defined by equation 25) of the ‘high’ simulation of Sormani et al. (2018), as a function of time. Full blue line: total gravitational torque exerted on material in the region  $\mathcal{R}$ . Dotted line: total mass contained in the region  $\mathcal{R}$  (values should be read on the left-hand side axis). The vertical red dashed line indicates the time at which the bar is fully grown. Units are as follows:  $[L] = 10^8 M_\odot 100 \text{ km s}^{-1} \text{ kpc}$ ,  $[T] = 10^8 M_\odot (100 \text{ km s}^{-1})^2$ ,  $[m] = 3 \times 10^8 M_\odot$ .

In the initial phase ( $t \lesssim 180 \text{ Myr}$ ), the angular momentum in the region  $\mathcal{R}$  increases steadily.<sup>10</sup> This is mainly due to the fact that the nuclear region is accreting gas brought to  $\mathcal{R}$  by the large-scale bar shocks (see the dotted line in Fig. 18), and not to the fact that the angular momentum per unit mass increases. After the initial phase, accretion almost stops, the mass in this region stays approximately constant, and the angular momentum oscillates quasi-periodically. The gravitational torques oscillate and their time average is clearly much lower than the amplitude of the oscillations.<sup>11</sup> An estimate obtained by averaging the data in Fig. 18 over a period  $t = [200, 320] \text{ Myr}$ , corresponding to approximately two of the longer oscillations, gives  $\langle L_{\mathcal{R}} \rangle / \langle T_{g,\mathcal{R}} \rangle \simeq 3 \text{ Gyr}$ . This is an estimate of the time-scale over which gravitational torques have net effects in this region, and it is much longer than the duration of the simulation.

In the phase  $t > 180 \text{ Myr}$ , as can be seen from the movies linked above, the gas distribution is made of molecular clouds that approximately follow  $x_2$  orbits (which act as guiding centres) on top of which there are significant excursions. A lot of turbulent motions are taking place, but over time these turbulent motions are dissipated, eventually leading to the narrow ring shown in Fig. 12. Thus, the gas does not really move in annuli smoothly exerting torques on each other as assumed in Section 3. However, it is likely that the formation of the ring does not depend too much on the details of the dissipation process (in this case, turbulent dissipation). The ring configuration is a local minimum of the energy, and the system can be driven to it regardless of the details of the dissipation process.

It is clearly seen by looking at the  $L_{\mathcal{R}}$  curve in Fig. 18 for  $t > 180 \text{ Myr}$  that there are two main time-scales of the oscillations, a short one with period  $T_1 \simeq 7 \text{ Myr}$  and a longer one with period  $T_2 \simeq 60 \text{ Myr}$ . The smaller time-scale  $T_1$  corresponds to half a period of an  $x_2$  orbit. Thus, these oscillations resemble those that we would obtain if we had a bunch of particles flowing on  $x_2$  orbits. The longer time-scale  $T_2$  corresponds to a very interesting phenomenon, that is

<sup>10</sup>Recall that the bar is introduced gradually during the first 150 Myr in the simulations of Sormani et al. (2018) in order to avoid transients.

<sup>11</sup>The instantaneous gravitational torques are not zero because the gas distribution is not symmetric about the Galactic centre (showing this was one of the main goals of Sormani et al. 2018). Instead, the gas distribution is made up of a few giant molecular clouds that revolve around the centre approximately following  $x_2$  orbits.

the major axis of the  $x_2$  disc/ring oscillates so that its angle with the  $y$  axis changes periodically.

The picture that emerges from the above considerations is one in which the gravitational torques are important in bringing down gas from the outer parts down to the nuclear region (with the mediation of shocks). Once material is in the nuclear region, it settles on  $x_2$  orbits and from that point viscosity becomes important and a ring of the type described in Section 3 is developed.

## 5.4 Relevance for real galaxies

### 5.4.1 Effective equation of state of the ISM

The results of Sections 4.1 and 4.2 (see also the discussions of Sections 5.2 and 5.3) suggest that the mechanism described in Section 3 is a viable way of ring formation if the gas sound speed is low ( $c_s \lesssim 1 \text{ km s}^{-1}$ ), but not if the gas sound speed is high ( $c_s \simeq 10 \text{ km s}^{-1}$ ). In the chemistry simulations of Section 4.3 the mechanism is able to work because the cold phase of the two-phase medium effectively behaves as a low sound speed gas. Thus, whether our theory can be applied to real galaxies depends on the ‘effective’ equation of state of the ISM.

What is the equation of state of the ISM? This is a very difficult question because the ISM is a complex dynamical entity and it is not easy to justify a set of equations that captures its dynamics. It is made of several components that are tightly coupled together. The four most important of these, in terms of their impact on the overall dynamical evolution, are gas, dust, cosmic rays, and magnetic fields. In addition, the ISM is highly inhomogeneous and shows structure on all observable scales, with star formation occurring in dense molecular clouds, and it is subject to several complex feedback loops driven by the energy and momentum input from stars in the form of radiation, winds, and supernovae.

There are two very different approaches to this complexity in the literature. The first one is to simply ignore it. If one looks at the properties of the atomic component of the ISM, as revealed by its 21 cm emission, then one finds a remarkably constant velocity dispersion in the range  $5\text{--}20 \text{ km s}^{-1}$  quite independent of galactic radius and galaxy type (e.g. Walter et al. 2008). This has led many studies of ISM dynamics on galactic scales to adopt an ‘isothermal prescription’ and to take the gas to be isothermal with a sound speed of  $c_s \simeq 10 \text{ km s}^{-1}$ , where the sound speed is *not* related to the actual kinetic temperature of the gas, but corresponds to an ‘effective’ temperature that is supposed to account for the turbulent motions in the ISM on unresolved scales (e.g. Roberts 1969; Cowie 1980; Athanassoula 1992b; Englmaier & Gerhard 1997; Regan & Teuben 2003; Maciejewski 2004; Li, Mac Low & Klessen 2005, 2006; Kim et al. 2012; Sormani et al. 2015a; Fragkoudi et al. 2017). The mechanism of ring formation presented in this paper generally does not apply if the ISM is modelled using this isothermal prescription, because the sound speed typically assumed ( $c_s \simeq 5\text{--}20 \text{ km s}^{-1}$ ) is too high.<sup>12</sup> However, as discussed above this treatment of the ISM is very crude. A similar simplification comes from the use of sticky-particle techniques, which model the gas as an ensemble of inelastically colliding particles (e.g. Schwarz 1981; Combes & Gerin 1985; Byrd et al. 1994; Jenkins & Binney 1994; Rautiainen & Salo 2000; Rautiainen et al. 2002, 2004; Rodriguez-

<sup>12</sup>Although it may apply to some simulations which assumed low sound speeds, e.g. the lowest sound speed simulations of Patsis & Athanassoula (2000) and Kim et al. (2012) which form a ring.

Fernandez & Combes 2008). The underlying assumption is that the ISM is mostly composed of long-lived dense molecular clouds, that are essentially described as ‘bricks’, that only interact collisionally (e.g. Tan 2000). Clearly this approach does not do justice to the multiphase nature and chemical complexity of the observed ISM (Klessen & Glover 2016).

The second approach is to face the observed complexity and construct a model that takes into account as many physical and chemical processes as possible and also includes a realistic treatment of stellar feedback. With the advent of new numerical modelling techniques and the continuously increasing capabilities provided by modern supercomputing centres this approach has become very popular in recent years. This has started with detailed modelling of a small portion of the galaxy or an individual molecular cloud, and we have now reached the point where it is feasible to apply this detailed modelling to simulations of an entire galaxy (see e.g. Smith et al. 2014; Bournaud et al. 2015; Iffrig & Hennebelle 2015; Walch et al. 2015; Geen et al. 2016; Girichidis et al. 2016; Hu et al. 2017; Kim & Ostriker 2017; Peters et al. 2017; Safranek-Shrader et al. 2017; Seifried et al. 2017; Hennebelle 2018; Hopkins et al. 2018; Pillepich et al. 2018). The simulations of Sormani et al. (2018) follow this tradition and arguably provide a more realistic treatment of the ISM than the isothermal prescription, and the mechanism presented in this paper seems to work in this case. This is probably due to the fact that indeed a large fraction of the mass in the nuclear region (called the Central Molecular Zone or CMZ in the Milky Way) is relatively cold with a sound speed smaller than  $1 \text{ km s}^{-1}$ .

This is a reassuring confirmation, but the approach of Sormani et al. (2018) still lacks several important ingredients. Most importantly, it lacks the small scale turbulence of the ISM, which is believed to be driven by supernova explosions and other stellar feedback mechanisms (e.g. Klessen & Glover 2016). This is evident in the fact that the ring in Fig. 12 is too thin compared to real rings. It also does not include magnetic fields, which may also affect the dynamics significantly. A useful test would be to run simulations similar to those of Sormani et al. (2018) but including the effects of stellar feedback, to see whether the size of the ring is affected or not. If the results of the chemistry simulations are confirmed, it would support the idea that the mechanism is relevant to real galaxies.

We conclude that whether or not the mechanism presented in this paper is relevant for real galaxies depends on the nature of the effective equation of state of the ISM. The question is complex and not settled yet, but future simulations may provide answers.

#### 5.4.2 Morphology of nuclear regions

The discussion of Section 5.2 suggests that the morphology of the nuclear region depends on the effective sound speed of the gas, i.e. that it depends on the importance of pressure forces. Lower pressure leads to more ring-like morphologies, while higher pressure leads to more spiral-like morphologies. The effective pressure of the ISM in real galaxies might be related to different temperature regimes of a multiphase medium, but may also contain contributions from turbulent pressure (hence related to the velocity dispersion of clouds) and magnetic pressure. This may provide insight into observations of external galaxies where the nuclear region displays different morphologies in different tracers (Izumi et al. 2013; Fathi et al. 2013; Izumi et al. 2015; Espada et al. 2017, see also García-Burillo et al. 2005). Different morphologies may reflect different temperature regimes of a multiphase medium, or different magnetic field strengths.

#### 5.4.3 Constraining the pattern speed

The pattern speed  $\Omega_p$  is the single most important parameter of a bar, but is notoriously difficult to measure (e.g. Tremaine & Weinberg 1984; Rautiainen, Salo & Laurikainen 2008; Aguerri et al. 2015). Thus methods that allow to constrain it are very valuable. The size of the ring in our theory is sensitive to the underlying gravitational potential. In particular, it is sensitive to the pattern speed of the bar: a higher pattern speed produces a smaller ring as it shrinks the extent of the  $x_2$  family of periodic orbits. Hence if the mass distribution of a barred potential is known but the pattern speed is not, our theory can be used to constrain  $\Omega_p$  by adjusting its value until the reversed shear region matches the observed size of the ring. The size of the ring predicted by our theory also depends on other properties of the potentials (e.g. a stronger bar generally produces smaller rings), which could therefore also be constrained.

#### 5.4.4 Accretion onto the central black hole

The ring configuration is a configuration of ‘local minimum’ of the energy. Hence, to get out of this ring and reach the *global* minimum of the energy (which corresponds to most of the mass in the centre) gas needs to overcome an ‘energy barrier’. The question of how the gas may overcome this energy barrier might be related to how the gas migrates from the nuclear ring ( $r \sim 100 \text{ pc}$ ) down to an accretion disc that fuels a supermassive black hole ( $r \sim 10^{-3} \text{ pc}$ ). Since nuclear rings are sites of intense star formation, we might speculate that stellar feedback ‘launches’ parcels of gas that overcome the barrier and can then plunge towards the black hole (Davies et al. 2007).

### 5.5 Relation to other works

#### 5.5.1 Galactic dynamics

As we mentioned in the introduction the main theory for the formation of nuclear rings is the resonant theory, which suggests that they are related to the ILR of the underlying gravitational potential (e.g. Combes 1988, 1996; Buta & Combes 1996). In our model the size of the ring is also indirectly related to the ILR, as it is well known that the radius of the ILR correlates with the radial extent of the  $x_2$  orbit family (Contopoulos & Grosbol 1989; Athanassoula 1992a), and therefore with the size of the reversed shear region (see Section 3). The main difference between our model and the resonant theory is that the latter relies only on gravitational torques, while our model also relies on viscous torques (see the discussion in Section 5.3). In Sections 4.2 and 5.2 we have also argued that, when examined in more detail, the predictions of the resonant theory seem not to be verified in simulations because naive application of the idea that torques act on the spiral pattern may lead to incorrect conclusions.

An alternative theory based on viscous torques was proposed by Icke (1979) and Lesch et al. (1990) and recently revisited by Krumholz & Kruijssen (2015). This theory is based on the idea that in a circular accretion disc (with  $d\Omega/dR < 0$ ) gas accumulates at the point where  $\Omega(R)$  becomes flat (as in solid body rotation) because this reduces the shear responsible for the inward mass transfer. The predictions from this theory are different from the predictions of the mechanism proposed in Section 3. First, the location where the rotation curve turns solid body is, in general, very different from the location of the reversed shear region: as can be seen from equation (2) the rotation curve in Fig. 2 only turns to solid body around  $R_c = 0.05 \text{ kpc}$ , while the reversed shear region is at  $R \sim$

1.2 kpc. Secondly, according to such theories the size of the ring does not depend on the pattern speed of the bar, in contrast to what is predicted by the mechanism of Section 3 and what is observed in simulations (e.g. Athanassoula 1992b; Li et al. 2015; Sormani et al. 2015c). For similar reasons, this theory cannot be the explanation for the ring that forms in the simulations of Sormani et al. (2018): the ring does not form near the point at which the rotation curve turns solid body.

Li et al. (2015; see also Kim et al. 2012) ran various 2D isothermal simulations of gas flow in barred potentials with varying bar strengths and pattern speeds. These authors use the ‘isothermal prescription’ (see Section 5.4), hence they are in the high sound speed regime in which our mechanism does not apply (see Section 5.2). However, they argue that the size of the ring is determined by the amount of angular momentum lost as the gas enters and travels along the bar shocks, which is different from the interpretation we have given in Section 5.2. They also argue that a stronger bar removes more angular momentum, leading to smaller rings. The same correlation between bar strength and size of the ring is expected in our theory, but for a different reason: a stronger bar changes the orbital structure of the potential, which results in smaller  $x_2$  orbits (e.g. Contopoulos & Grosbol 1989; Athanassoula 1992a). This correlation was also observed in the sticky-particle simulations of Salo et al. (1999) and is consistent with the observational result that ‘stronger bars host smaller rings’ (Comerón et al. 2010).

Binney et al. (1991; see also Sormani et al. 2015a) predicts that the major axis of the ring should be equal to the minor axis of the cusped  $x_1$  orbit. This predicts a ring whose size is generally significantly smaller than that of Section 3, although the details depend on the gravitational potential. Indeed, it would predict a ring whose size is much smaller than the one developed in the simulations of Sormani et al. (2018). Note also that Fig. 12 suggests that the bar shocks touch the ring near the point of its maximum extension in the  $y$  direction, but tests we have performed with different potentials suggest that this is not generally the case: sometimes the shock can intersect the ring at lower or values of  $y$ , and sometimes it can intersect the  $y$  axis further out, flying by the ring without touching it. This suggests that the point where the shock touches the ring is not what determines the size of the ring.

### 5.5.2 Planetary ring dynamics

The problem studied in this paper has many analogies with the confinement of narrow rings in the context of planetary ring dynamics (see e.g. Longaretti 2018 for a recent review). In particular, the idea of reverse shear discussed in Section 3 is equivalent to the idea of angular momentum flux reversal that has been developed by Borderies–Goldreich–Tremaine in a series of papers (e.g. Borderies, Goldreich & Tremaine 1982, 1983a,b, 1989; Borderies 1989) to explain the confinement and the sharp edges of planetary rings. These authors showed that under the influence of a perturbing potential, created, e.g. by a satellite, the particle streamlines can be sufficiently distorted so that the angular momentum flux is directed inwards, which is very similar to the picture we described in Section 3 (in our notation, the angular momentum flux is directed inwards when  $k < 0$ ). In the context of planetary rings this idea has been explored also using  $N$ -body simulations, both global (Hänninen & Salo 1994, 1995; see also Goldreich, Rappaport & Sicardy 1995 for a theoretical discussion of their results) and limited to a local comoving region (Mosqueira 1996).

An obvious difference between the galactic and the planetary problems is the order of magnitude of the perturbations. While for a strong barred potential such as the one considered in Section 3 particle orbits deviate strongly from circular motions (see Fig. 3), in the case of planetary ring dynamics the deviations from circularity are typically very small, with eccentricities of the order of  $10^{-5}$  (e.g. Borderies 1989). This reflects the fact that the resonance zone, i.e. the region where the effect of the resonances is non-negligible, is much narrower in the planet–satellite case. Another major difference between the two problems is the mechanism that generates viscosity, namely collisions between particles in the planetary case and turbulent viscosity in the galactic case.

## 6 SUMMARY

We have developed a dynamical theory to explain why nuclear rings arise and to predict their location given the underlying gravitational potential. The main conclusions of this paper are as follows:

(i) When the gas follows non-circular orbits such as  $x_2$  orbits, viscous torques between adjacent annuli of gas can drive the gas towards a ring configuration (see Section 3), in contrast to what is found in the elementary theory of circular viscous accretion discs in which all the gas is driven to the centre. The ring configuration is a local minimum of the energy.

(ii) The sound speed and the amount of viscosity are two key parameters of the gas flow in the nuclear regions of galaxies.

(iii) When the sound speed is low ( $c_s \lesssim 1 \text{ km s}^{-1}$ ), a ring may be formed according to the mechanism of item (i). The location of this ring correlates with the ILR and can be approximately predicted based on the underlying gravitational potential: it must be within the reversed shear region (orange in Fig. 3). We have speculated that the ring size might be predicted even more precisely as the orbit where  $K_{\parallel} = 0$ , but this requires further verification (see Section 3.1).

(iv) When the sound speed is high ( $c_s \simeq 10 \text{ km s}^{-1}$ ), such a ring cannot survive because spiral shocks are developed, regardless of the presence of viscosity. In this case, the size of the nuclear region is determined by the critical radius  $R_c$  at which the oscillations of the gravitational potential become strong enough to induce a shock. The value of  $R_c$  decreases at higher sound speeds, leading to smaller nuclear regions, according to the discussion in Section 5.2.

(v) Whether the mechanism described in Section 3 is relevant for the formation of nuclear rings in real galaxies depends on the ‘effective’ equation of state of the ISM. This is a complex question that is not settled yet, but simulations will provide great help in the near future.

Putting everything together, we argued that a plausible scenario of ring formation (see Section 5.3) in a real galaxy involves two phases: in the first phase gravitational torques bring the gas from outside the bar region down to the nuclear region. Then, once gas is within the nuclear region, gravitational torques become less important and turbulent dissipation, while it settles on  $x_2$  orbits, will drive the gas towards the stable ring configuration described in Section 3.

## ACKNOWLEDGEMENTS

The authors thank the anonymous referee for pointing out the connection between this work and the work by Borderies–Goldreich–Tremaine in the context of planetary ring dynamics. The authors also thank James Binney, David Cole, Payel Das, Cornelis Dullemond, Sacha Hony, Zhi Li, Yuri Lyubarsky, Diederik Kruijssen, Mattis Magg, John Magorrian, Eva Schinnerer, and Ralph Schoenrich for

useful comments and discussions. MCS thanks Andrea Mignone and Claudio Zanni for help with the PLUTO code. MCS, RGT, SCOG, and RSK acknowledge support from the Deutsche Forschungsgemeinschaft via the Collaborative Research Centre (SFB 881) “The Milky Way System” (subprojects B1, B2, and B8) and the Priority Program SPP 1573 “Physics of the Interstellar Medium” (grant numbers KL 1358/18.1, KL 1358/19.2, and GL 668/2-1). RSK furthermore thanks the European Research Council for funding in the ERC Advanced Grant STARLIGHT (project number 339177). ES acknowledges support from the Israeli Science Foundation under Grant No. 719/14. The simulations in this paper were run using the bwForCluster MLS&WISO; the authors acknowledge support by the state of Baden-Württemberg through bwHPC and the German Research Foundation (DFG) through grant INST 35/1134-1 FUGG.

## REFERENCES

- Aguerre J. A. L. et al., 2015, *A&A*, 576, A102  
 Arzamasskiy L., Rafikov R. R., 2018, *ApJ*, 854, 84  
 Athanassoula E., 1992a, *MNRAS*, 259, 328  
 Athanassoula E., 1992b, *MNRAS*, 259, 345  
 Binney J., Tremaine S., 2008, *Galactic Dynamics*. 2nd edn, Princeton Univ. Press, Princeton  
 Binney J., Gerhard O. E., Stark A. A., Bally J., Uchida K. I., 1991, *MNRAS*, 252, 210  
 Borderies N., 1989, *Celest. Mech. Dyn. Astron.*, 46, 207  
 Borderies N., Goldreich P., Tremaine S., 1982, *Nature*, 299, 209  
 Borderies N., Goldreich P., Tremaine S., 1983a, *AJ*, 88, 226  
 Borderies N., Goldreich P., Tremaine S., 1983b, *AJ*, 88, 1560  
 Borderies N., Goldreich P., Tremaine S., 1989, *Icarus*, 80, 344  
 Bournaud F., Daddi E., Weiß A., Renaud F., Mastropietro C., Teyssier R., 2015, *A&A*, 575, A56  
 Buta R., Combes F., 1996, *Fundam. Cosm. Phys.*, 17, 95  
 Buta R. J., 2017a, *MNRAS*, 471, 4027  
 Buta R. J., 2017b, *MNRAS*, 470, 3819  
 Byrd G., Rautiainen P., Salo H., Buta R., Crocher D. A., 1994, *AJ*, 108, 476  
 Combes F., 1988, in Pudritz R. E., Fich M., eds, *NATO Advanced Science Institutes (ASI) Series C*, Vol. 232. Springer Nature, Switzerland, p. 475  
 Combes F., 1996, in Buta R., Crocker D. A., Elmegreen B. G., eds, *ASP Conference Series*, Vol. 91, IAU Colloq. 157: Barred Galaxies. Astron. Soc. Pac., San Francisco, p. 286  
 Combes F., 2001, in Aretxaga I., Kunth D., Mújica R., eds, *Advanced Lectures on the Starburst-AGN*. World Scientific Publishing, Singapore, p. 223  
 Combes F., Gerin M., 1985, *A&A*, 150, 327  
 Comerón S., 2013, *A&A*, 555, L4  
 Comerón S., Knapen J. H., Beckman J. E., Laurikainen E., Salo H., Martínez-Valpuesta I., Buta R. J., 2010, *MNRAS*, 402, 2462  
 Contopoulos G., Grosbol P., 1989, *A&AR*, 1, 261  
 Cowie L. L., 1980, *ApJ*, 236, 868  
 Davies R. I., Müller Sánchez F., Genzel R., Tacconi L. J., Hicks E. K. S., Friedrich S., Sternberg A., 2007, *ApJ*, 671, 1388  
 Englmaier P., Gerhard O., 1997, *MNRAS*, 287, 57  
 Espada D. et al., 2017, *ApJ*, 843, 136  
 Fathi K. et al., 2013, *ApJ*, 770, L27  
 Fragkoudi F., Athanassoula E., Bosma A., 2017, *MNRAS*, 466, 474  
 García-Burillo S., Combes F., Schinnerer E., Boone F., Hunt L. K., 2005, *A&A*, 441, 1011  
 Geen S., Hennebelle P., Tremblin P., Rosdahl J., 2016, *MNRAS*, 463, 3129  
 Girichidis P. et al., 2016, *ApJ*, 816, L19  
 Goldreich P., Rappaport N., Sicardy B., 1995, *Icarus*, 118, 414  
 Hänninen J., Salo H., 1994, *Icarus*, 108, 325  
 Hänninen J., Salo H., 1995, *Icarus*, 117, 435  
 Hennebelle P., 2018, *A&A*, 611, A24  
 Henshaw J. D. et al., 2016, *MNRAS*, 457, 2675  
 Heyer M., Dame T. M., 2015, *ARA&A*, 53, 583  
 Hopkins P. F. et al., 2018, *MNRAS*, 477, 1578  
 Hu C. Y., Naab T., Glover S. C. O., Walch S., Clark P. C., 2017, *MNRAS*, 471, 2151  
 Icke V., 1979, *A&A*, 78, 21  
 Iffrig O., Hennebelle P., 2015, *A&A*, 576, A95  
 Izumi T. et al., 2013, *PASJ*, 65, 100  
 Izumi T. et al., 2015, *ApJ*, 811, 39  
 Jenkins A., Binney J., 1994, *MNRAS*, 270, 703  
 Kim C. G., Ostriker E. C., 2017, *ApJ*, 846, 133  
 Kim W. T., Seo W. Y., Stone J. M., Yoon D., Teuben P. J., 2012, *ApJ*, 747, 60  
 Kim W. T., Kim Y., Kim J. G., 2014, *ApJ*, 789, 68  
 Klessen R. S., Glover S. C. O., 2016, in Revaz Y., Jablonka P., Teyssier R., Mayer L., eds, *Star Formation in Galaxy Evolution: Connecting Numerical Models to Reality*, Saas-Fee Advanced Course, Vol. 43. Springer-Verlag, Berlin Heidelberg, p. 85  
 Knapen J. H., 2005, *A&A*, 429, 141  
 Kruijssen J. M. D., Dale J. E., Longmore S. N., 2015, *MNRAS*, 447, 1059  
 Krumholz M. R., Kruijssen J. M. D., 2015, *MNRAS*, 453, 739  
 Lesch H., Biermann P. L., Crusius A., Reuter H. P., Dahlem M., Barteldrees A., Wielebinski R., 1990, *MNRAS*, 242, 194  
 Li Y., Mac Low M.-M., Klessen R. S., 2005, *ApJ*, 626, 823  
 Li Y., Mac Low M.-M., Klessen R. S., 2006, *ApJ*, 639, 879  
 Li Z., Shen J., Kim W.-T., 2015, *ApJ*, 806, 150  
 Li Z., Sellwood J. A., Shen J., 2017, *ApJ*, 850, 67  
 Liszt H. S., 2009, preprint ([arXiv:0905.1412](https://arxiv.org/abs/0905.1412))  
 Longaretti P.-Y., 2018, in Tiscareno M. S., Murray C. D., eds, *Theory of Narrow Rings and Sharp Edges, in Planetary Ring Systems: Properties, Structure, and Evolution*, Cambridge Planetary Science. Cambridge Univ. Press, Cambridge, p. 225  
 Lynden-Bell D., Kalnajs A. J., 1972, *MNRAS*, 157, 1  
 Lynden-Bell D., Pringle J. E., 1974, *MNRAS*, 168, 603  
 Mac Low M.-M., Klessen R. S., 2004, *Rev. Mod. Phys.*, 76, 125  
 Maciejewski W., 2004, *MNRAS*, 354, 892  
 Mignone A., Bodo G., Massaglia S., Matsakos T., Tesileanu O., Zanni C., Ferrari A., 2007, *ApJS*, 170, 228  
 Molinari S. et al., 2011, *ApJ*, 735, L33  
 Mosqueira I., 1996, *Icarus*, 122, 128  
 Patsis P. A., Athanassoula E., 2000, *A&A*, 358, 45  
 Peters T. et al., 2017, *MNRAS*, 467, 4322  
 Phinney E. S., 1994, in Shlosman I., ed., *Mass-Transfer Induced Activity in Galaxies*. Cambridge Univ. Press, Cambridge, p. 1  
 Pillepich A. et al., 2018, *MNRAS*, 473, 4077  
 Prendergast K. H., 1983, in Athanassoula E., ed., *Proceeding of the Symposium, Besancon, France, August 9-13, 1982*, IAU Symposium, Vol. 100, *Internal Kinematics and Dynamics of Galaxies*. Dordrecht, D. Reidel Publishing Co., p. 215  
 Pringle J. E., 1981, *ARA&A*, 19, 137  
 Rautiainen P., Salo H., 2000, *A&A*, 362, 465  
 Rautiainen P., Salo H., Laurikainen E., 2002, *MNRAS*, 337, 1233  
 Rautiainen P., Salo H., Buta R., 2004, *MNRAS*, 349, 933  
 Rautiainen P., Salo H., Laurikainen E., 2008, *MNRAS*, 388, 1803  
 Regan M. W., Teuben P., 2003, *ApJ*, 582, 723  
 Roberts W. W., 1969, *ApJ*, 158, 123  
 Rodriguez-Fernandez N. J., Combes F., 2008, *A&A*, 489, 115  
 Safrank-Shrader C., Krumholz M. R., Kim C. G., Ostriker E. C., Klein R. I., Li S., McKee C. F., Stone J. M., 2017, *MNRAS*, 465, 885  
 Salo H., Rautiainen P., Buta R., Purcell G. B., Cobb M. L., Crocker D. A., Laurikainen E., 1999, *AJ*, 117, 792  
 Schwarz M. P., 1981, *ApJ*, 247, 77  
 Seifried D. et al., 2017, *MNRAS*, 472, 4797  
 Sellwood J. A., Balbus S. A., 1999, *ApJ*, 511, 660  
 Sellwood J. A., Wilkinson A., 1993, *Rep. Prog. Phys.*, 56, 173  
 Shakura N. I., Sunyaev R. A., 1973, *A&A*, 24, 337  
 Smith R. J., Glover S. C. O., Clark P. C., Klessen R. S., Springel V., 2014, *MNRAS*, 441, 1628  
 Sormani M. C., Binney J., Magorrian J., 2015a, *MNRAS*, 449, 2421  
 Sormani M. C., Binney J., Magorrian J., 2015b, *MNRAS*, 451, 3437

- Sormani M. C., Binney J., Magorrian J., 2015c, *MNRAS*, 454, 1818  
Sormani M. C., Sobacchi E., Shore S. N., Treß R. G., Klessen R. S., 2017, *MNRAS*, 471, 2932  
Sormani M. C., Treß R. G., Ridley M., Glover S. C. O., Klessen R. S., Binney J., Magorrian J., Smith R., 2018, *MNRAS*, 475, 2383  
Springel V., 2010, *MNRAS*, 401, 791  
Tan J. C., 2000, *ApJ*, 536, 173  
Tremaine S., Weinberg M. D., 1984, *ApJ*, 282, L5  
van de Ven G., Chang P., 2009, *ApJ*, 697, 619  
Wada K., Koda J., 2004, *MNRAS*, 349, 270  
Walch S. et al., 2015, *MNRAS*, 454, 238  
Walter F., Brinks E., de Blok W. J. G., Bigiel F., Kennicutt R. C., Jr., Thornley M. D., Leroy A., 2008, *AJ*, 136, 2563

## APPENDIX A: ANALYTIC EXPRESSION FOR THE QUADRUPOLE POTENTIAL

The quadrupole  $\Phi_2$  used in Section 3 is generated by the following density distribution:

$$\rho_2(r, \phi, \theta) = \frac{A}{4\pi G} \left( \frac{v_0 e}{r_q} \right)^2 \exp\left(-\frac{2r}{r_q}\right) \sin^2 \phi \cos(2\theta), \quad (\text{A1})$$

where  $A$  is a dimensionless parameter,  $G$  is the gravitational constant,  $v_0$  and  $r_q$  are parameters, and  $(r, \theta, \phi)$  are spherical coordinates where  $r$  is the radial distance,  $\phi$  is the polar angle, and  $\theta$  is the azimuthal angle. The plane  $\phi = \pi/2$  coincides with the plane of the galaxy. In the main text we always assume  $\phi = \pi/2$  since everything is 2D. The values of the parameters used in the main text are  $A = 0.4$ ,  $v_0 = 220 \text{ km s}^{-1}$ , and  $r_q = 1.5 \text{ kpc}$ .

The potential corresponding to this density distribution can be found analytically. Since the density  $\rho_2$  is proportional to the real part of the spherical harmonic  $Y_2^2$ , which is an eigenfunction of the Laplacian operator, the potential must be of the following form:

$$\Phi_2(r, \theta, \phi) = \Phi_2(r) \sin^2 \phi \cos(2\theta). \quad (\text{A2})$$

Taking the Laplacian of this equation and using the Poisson equation we find

$$\nabla^2 \Phi_2(r, \theta, \phi) = \left[ \Phi_2'' + \frac{2}{r} \Phi_2' - \frac{6}{r^2} \Phi_2 \right] \sin^2 \phi \cos(2\theta) \quad (\text{A3})$$

$$= 4\pi G \rho_2. \quad (\text{A4})$$

Hence  $\Phi_2(r)$  must satisfy

$$\Phi_2'' + \frac{2}{r} \Phi_2' - \frac{6}{r^2} \Phi_2 = A \left( \frac{v_0 e}{r_q} \right)^2 \exp\left(-\frac{2r}{r_q}\right). \quad (\text{A5})$$

We can reduce this to a dimensionless equation by defining the following dimensionless variables:

$$x = \frac{r}{r_q}, \quad (\text{A6})$$

$$F = -\frac{\Phi_2}{A(v_0 e)^2}. \quad (\text{A7})$$

Equation (A5) becomes

$$F''(x) + \frac{2F'(x)}{x} - \frac{6F(x)}{x^2} = -e^{-2x}. \quad (\text{A8})$$

To find the potential corresponding to the density distribution (A1), we need to solve this equation with boundary conditions

$$F(0) = F(\infty) = 0. \quad (\text{A9})$$

The solution is

$$F(x) = \frac{3 - e^{-2x} (2x^4 + 4x^3 + 6x^2 + 6x + 3) + 4x^5 E_1(2x)}{20x^3}, \quad (\text{A10})$$

where  $E_1(x)$  is the exponential integral function, a special function defined as

$$E_1(x) \equiv \int_x^\infty \frac{e^{-t}}{t} dt. \quad (\text{A11})$$

## APPENDIX B: CONSERVATION OF ANGULAR MOMENTUM FOR A FLUID

Consider a fluid governed by the continuity equation and the compressible Navier–Stokes equation, which in conservative form are

$$\partial_t \rho + \partial_j (\rho v_j) = 0 \quad (\text{B1})$$

$$\partial_t (\rho v_i) + \partial_j (\rho v_i v_j + \delta_{ij} P) = -\partial_j (\rho \sigma_{ij}) - \rho \partial_i \Phi, \quad (\text{B2})$$

where  $\Phi$  is the external gravitational potential, the viscous tensor is given by (5), and we use the convention that repeated indices means summation. From (B1) and (B2) one can prove the following statement for the conservation of angular momentum:

$$\begin{aligned} \partial_t (\rho L_i) + \partial_j (\rho L_i v_j + \epsilon_{ipm} x_p \delta_{mj} P) \\ = \partial_j (\rho \epsilon_{ipm} x_p \sigma_{mj}) - \rho \epsilon_{ipm} x_p \partial_m \Phi, \end{aligned} \quad (\text{B3})$$

where  $L_i$  is the angular momentum per unit mass,

$$L_i = \epsilon_{ijk} x_j v_k. \quad (\text{B4})$$

The first term on the right-hand side of (B3) represent the flux of angular momentum due to viscous forces, while the second represents the change in angular momentum due to gravitational forces. After some manipulation and using equation (B1), equation (B3) can be recast in the following form:

$$\frac{D\mathbf{L}}{Dt} = \boldsymbol{\tau}_P + \boldsymbol{\tau}_v + \boldsymbol{\tau}_g, \quad (\text{B5})$$

where  $D$  is the convective derivative

$$\frac{D}{Dt} = \partial_t + \mathbf{v} \cdot \nabla \quad (\text{B6})$$

and the torques per unit mass due to pressure, viscous, and gravitational forces are given by

$$[\boldsymbol{\tau}_P]_i = -\frac{1}{\rho} \partial_j (\rho \epsilon_{ipm} x_p \delta_{mj} P), \quad (\text{B7})$$

$$[\boldsymbol{\tau}_v]_i = \frac{1}{\rho} \partial_j (\rho \epsilon_{ipm} x_p \sigma_{mj}), \quad (\text{B8})$$

$$[\boldsymbol{\tau}_g]_i = -\epsilon_{ipm} x_p \partial_m \Phi. \quad (\text{B9})$$

The physical meaning of equation (B6) is that if we follow a fluid element, its angular momentum changes due to three contributions: pressure, viscous, and gravitational torques. In two dimensions, equation (B8) reduces to equation (12).

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