

# Dynamics of Spiral Structure

Shera Jafaritabar<sup>1</sup>

Heidelberg University  
e-mail: shera.jafaritabar@stud.uni-heidelberg.de

July 14, 2024

## ABSTRACT

**Context.** Integrates approaches combining theoretical models, simulation, and observational data covering the complex dynamics of spiral structures in galaxies.

**Aims.** This study aims to investigate the dynamics of spiral structures in galaxies, focusing on three primary mechanisms: swing amplification, tidal interactions, and bar interactions. By exploring these mechanisms, the study seeks to elucidate the formation of spiral arms, the persistence and lifetime of spiral arms, and the classification of galaxies and their morphologies in each of the mechanisms.

**Methods.** The study utilizes a combination of theoretical analysis for the dispersion relations and the stability criterion. Swing amplification is examined through its role in converting leading density waves into trailing ones, amplified by resonance at the corotation radius theoretically, and further compared to the simulation results of Toomre (1981). Tidal interactions are studied using the paper and the N-body simulations of Semczuk et al. (2017) to model the effects of gravitational interactions with nearby satellite galaxies through repeated cycles. Bar interactions are explored through hydrodynamical simulations of Huntley et al. (1978) that investigate the response of gas to bar-like perturbations, highlighting the alignment of spiral structures with the bar's rotation.

**Results.** The study finds that swing amplification is one of the best explanations for the formation of transient, multi-armed, or grand-design spiral patterns, and it is facilitated by differential rotation and self-gravity within the galactic disk. Tidal interactions are also very effective in producing grand-design spiral structures, going through repeated close encounters with the satellite galaxy. This can further explain the presence of grand-design spirals in the universe despite their transient nature. Bar-driven spirals are confirmed to align closely with the bar's rotation pattern, contributing to the formation of well-defined bisymmetric spiral arms, and they are believed to be long-lived. Observations, such as radial mixing and gas presence in spirals support the transient picture in spirals, and simulations also develop transient multi-armed spirals if exposed to higher sectoral harmonics. The study also highlights the importance of various theoretical topics such as Lindblad resonances, corotation resonances, and dispersion relations in the propagation and amplification of spiral density waves.

**Conclusions.** Even with the observational discrepancies in each of the tidal and bar-driven spirals, they remain the main process for grand-design galaxies with nearby satellites, and barred grand-design galaxies. The study finds swing amplification to be the main process for the formation of transient, grand-design or multi-armed spiral patterns.

**Key words.** spiral structure - swing amplification - bar driven spirals - tidally driven spirals

## 1. Introduction

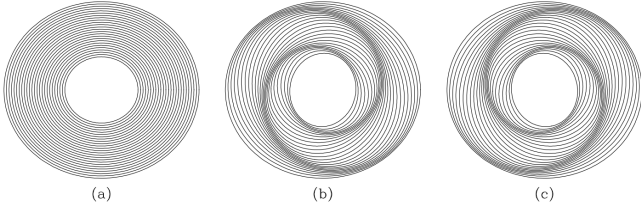
Spiral galaxies are the predominant galaxy type in the local universe, with around two-thirds of all galaxies being spirals (Semczuk et al. 2017). The first classification of the spiral galaxy was Hubble (1926) classifying them from S0/Sa for early-type galaxies, to Sb/Sc for late-type galaxies. Hubble's classification was based on their color; calling *redder* galaxies early type and *bluer* galaxies as late type, which is no longer correct. It also does not include all the ways that spiral galaxies are different from each other. The spiral arms also exhibit diverse morphologies characterized by the number of arms, leading to classifications ranging from the well-defined "grand design" or two-armed spirals to the more fragmented multi-armed and "flocculent" spirals (Sellwood & Masters 2022).

The appearance of spiral arms in galaxies is also significantly influenced by the wavelength of observation, with older stars emitting more light in red and infrared (IR) bands, making red and near-infrared (NIR) images more reflective of the underlying stellar population, while younger, brighter stars dominate images in bluer bands. Consequently, spiral arms appear smoother and more continuous in NIR images. The Spitzer Survey of Stel-

lar Structure in Galaxies (S4G) revealed that galaxies appearing flocculent in optical light often maintain this structure in mid-infrared (MIR) observations, though some flocculent galaxies in blue light exhibit underlying grand design spirals in NIR images (Sellwood & Masters 2022). Buta et al. (2015) confirmed that MIR classifications of most spiral galaxies align with B-band images, albeit with slightly earlier Hubble-type classifications due to more prominent bulges in MIR.

Few mechanisms can be responsible for producing spiral arms, as I will cover the 3 most important ones in this paper, namely: Swing Amplification, Tidal, and Bar interactions. Each of these theories are successful in explaining certain characteristics of spirals, and they all have different lifetime (Sellwood 2011).

This paper will first talk about the nature of spiral arms and then will go over some fundamental tools that we will be needing for explaining some of the theories. After covering the three theories for the spiral arms formation, it will further be discussed whether the spiral patterns are transient or long-lasting (quasi-stationary).



**Fig. 1.** Closed orbits in a galaxy with  $\Omega - \frac{\kappa}{2}$  creating spiral patterns. Figure adapted from Binney & Tremaine (2008).

## 2. Spiral Arms: Material Arms or Kinematic Density Waves?

Differential rotation gives a very easy way to produce spiral arms, with the stars closer to the center, completing a bigger fraction of the orbit compared to the outer stars in the same amount of time. However, if the arms are purely material arms, the arms should wind up and the structure will disappear. The other problem that this can bring is over time, differential rotation can cause many turns while in observations, most of the galaxies don't have more than two turns (Toomre 1977).

On the other hand, we can think of spiral structures as kinematic density waves. In the kinematic density waves, the gas and stars flow through them but not with a constant pattern speed (which is characteristic of material arms), but their pattern speed decreases as the radius increases. When they are perturbed, they induce elliptical stellar orbits, which are generally not closed. By picking a rotating frame that is closed after each time we encounter a spiral arm, or in other words, after  $\kappa/m$ , we can create a closed orbit. Setting  $m = 2$  and assuming that  $\Omega - \frac{\kappa}{2}$  is constant, we can have what we can see in Figure (1) panel a, and by rotating the axes of the ellipse, we can create leading or trailing spiral density waves in figure (1) panel b and c respectively (Binney & Tremaine 2008).

However, in real galaxies  $\Omega - \frac{\kappa}{2}$  is not constant, so orbits are not exactly closed. Binney & Tremaine (2008) concluded that for galaxies with similar circular speed curve like the Milky Way galaxy, density waves unwind slower than material arms. This also can be one explanation for grand-design spirals being the most common type among spiral galaxies.

## 3. Preliminaries

Before starting to go over the main theories in the spiral structure, a brief introduction on some of the fundamental topics that are widely used in spiral dynamics shall be given. Firstly the dispersion relations derived by Lin & Shu (1966a) and the hypotheses used to derive that will be discussed. Further Lin-Shu dispersion relation shall be compared to the Lin-Shu-Kalnajs dispersion relation, and the stability criterion will be discovered. In the end, a brief introduction on Lindblad resonances and how the wave will propagate through them will be given.

### 3.1. Hypothesis

Lin & Shu (1964) proposed that spiral arms are density waves rather than material arms, and subsequently derived the dispersion relation for a fluid disc. Following that, Lin & Shu (1966b) and Kalnajs (1965) derived the dispersion relation for a stellar disk. For the derivation of these dispersion relations, several approximations and hypotheses were used. In this section, I will provide a brief review of these assumptions.

#### 3.1.1. Thin Disc, Linear Perturbation, and Axisymmetric Responses

The first assumption that is used to simplify the analysis is that all theoretical models assume a razor-thin disk, which significantly reduces the complexity of the calculations. The dynamics of a smooth stellar fluid are governed by the collision-less Boltzmann equation (CBE), which describes the evolution of the distribution function (DF) of stars, and the Poisson equation, relating gravitational potential and density distribution. By using the linear perturbation assumption, they linearized the equations of continuity, the equation of motion (Euler equation), the equation of state, and the Poisson equation assuming that the unperturbed disc is axisymmetric and has no radial motions (Shu 2016; Binney 2020).

#### 3.1.2. WKB Approximation

WKB approximation, or tight-winding approximation was introduced to help the problem of gravity being a long-range force. By assuming a small pitch angle, distant perturbations can be neglected; meaning that matter's response to gravity is localized (Shu 2016; Binney & Tremaine 2008).

#### 3.1.3. Quasi Stationary Spiral

The ubiquity of spirals in galaxies led many astronomers to favor long-lived spiral patterns because they would not require constant regeneration (Sellwood 2011). The quasi-stationary spiral structure (QSSS) hypothesis suggests that spiral arms are stable, long-lived features that appear nearly stationary in a rotating frame of reference, meaning that they rotate with fixed pattern speed (Shu 2016). This concept is pivotal in understanding the enduring nature of spiral arms in galaxies, as first proposed by Lin & Shu (1964). The QSSS hypothesis can be formulated by considering the gravitational potential in a disk of infinitesimal thickness. In cylindrical coordinates  $(r, \phi, z)$ , where the galactic center is at  $r = 0$ , and the mid-plane galactic disk is at  $z = 0$ , the gravitational potential in the mid-plane of the disk can be written as the sum of a stationary axisymmetric part (which is the bulge, disk, and halo) and a non-stationary and a non-axisymmetric part (spiral gravitational perturbation) as (Shu 2016):

$$V(r, \phi, z = 0, t) = V_0(r) + V_1(r)e^{i(\omega t - m\phi)}, \quad (1)$$

where  $V_0(r)$  is the stationary axisymmetric part and  $V_1(r)e^{i(\omega t - m\phi)}$  is the non-stationary and non-axisymmetric part.

The QSSS hypothesis assumes that spiral perturbations are small relative to the axisymmetric state, allowing them to be treated as small oscillations. These perturbations can be Fourier-analyzed in time and angle (Binney 2020). If  $\omega$  is real, we will write it as  $\omega = m\Omega_p - i\gamma$ , where  $\gamma$  is the growth rate. For  $\gamma \ll m|\Omega_p|$  it will lead to the formulation of a nearly stationary structure in a rotating frame with pattern angular speed  $\Omega_p$  (Shu 2016).

We will have spirals if the complex radial part has an amplitude and a phase  $V(r) = A(r)e^{i\Phi(r)}$ . Trailing and leading spirals are when the wave number  $k = \Phi'(r)$  is negative or positive. Unequal superposition of trailing and leading spirals causes the structure to form spirals (Shu 2016).

### 3.2. Dispersion Relation

Using the 3 hypotheses in the previous section, Lin & Shu (1964) derived the dispersion relation. The dispersion relation of a stel-

lar disk can help us understand spiral arms since most of the mass of the disc is made of stars. However, before getting to the dispersion relation for stars, I first introduce the dispersion relation for a fluid disc. The dispersion relation for linear tight winding perturbation in a thin fluid disc, Lin-Shu (LS) dispersion relation, is given as (Shu 2016):

$$(\omega - m\Omega)^2 = \kappa^2 + c_s^2 k^2 - 2\pi G |k| \Sigma_0, \quad (2)$$

where  $c_s$ ,  $\Omega$ ,  $\kappa$  and  $\Sigma_0$  are sound speed, angular frequency, epicyclic frequency, and surface density respectively, and all of them are a function of galactic-centric radius  $R$ . Also  $\omega = m\Omega_P$  where  $\Omega_P$  is the pattern speed (Shu 2016).

In the LS dispersion relation,  $\omega$  is the angular velocity in the inertial frame, and  $\omega - m\Omega$  is the angular velocity in a rotating frame of a star at  $R$ . On the right-hand side of this equation, the first term ( $\kappa^2$ ) is the rotation effect that stabilizes the disk, the second term expresses the effect of pressure, and while it is positive, it stabilizes the fluid against perturbation. The last term is what incorporates the self-gravity of the disc, promoting growth instabilities (Shu 2016). If the term,  $-2\pi G |k| \Sigma_0$  becomes large enough to make the term  $(\omega - m\Omega)^2$  non-negative, the disk will become unstable. Positive or negative  $(\omega - m\Omega)^2$  can determine whether the perturbation is stable or unstable and its amplitude will grow exponentially (Shu 2016; Toomre 1977). Now using equation (2) we can derive that:

$$k_{\pm} = \frac{\pi G \Sigma_0}{\kappa c_s^2} \pm \sqrt{\frac{\pi G^2 \Sigma_0^2}{c_s^2} - \frac{\kappa^2}{c_s^2}}. \quad (3)$$

Now setting the term under the square root to zero in equation (3), we will derive the Toomre's parameter  $Q$  which works as a criterion for the stability of the disc; for  $Q > 1$ , the disc is stable, and for  $Q < 1$  disc is unstable (Shu 2016):

$$Q = \frac{\kappa c_s}{\pi G \Sigma_0}. \quad (4)$$

Another stability criterion is a critical wavenumber where the disk is stable for all the wavenumbers smaller than  $k_{\text{crit}}$ . When  $(\omega - m\Omega)^2 = 0$  for a cold disc ( $c_s = 0$ ),  $k_{\text{crit}}$  will be given as  $k_{\text{crit}} = \frac{\kappa^2}{2\pi G \Sigma_0}$ . Following that we can state that all perturbations with wavenumber  $|k| > k_{\text{crit}}$  or wavelength  $\lambda < \lambda_{\text{crit}}$  are stable, where (Shu 2016; Binney & Tremaine 2008):

$$\lambda < \lambda_{\text{crit}} = \frac{2\pi}{k_{\text{crit}}} = \frac{4\pi^2 G \Sigma}{\kappa^2}. \quad (5)$$

Kalnajs (1965) found the dispersion relation for the frequency of axisymmetric waves in a 2D stellar disk. He also introduced a new parameter,  $F$ , which is a reduction factor taking into account the weakening of self-gravity by random stellar motion which depends on  $Q$ ,  $k$ , and  $\omega$ . Specifically,  $F \leq 1$ , with  $F = 1$  representing a cold disk that has zero random motion ( $Q = 0$ ). For a stable disk, we need  $\omega^2 \geq 0$ , which requires  $F$  to remain sufficiently small. The dispersion relation is given by (Toomre 1977; Shu 2016):

$$\omega^2 = \kappa^2 - 2\pi G \Sigma_0 |k| F. \quad (6)$$

Lin & Shu (1966b) replaced the frequency with the Doppler-shifted frequency at which stars encounter an  $m$ -fold symmetric spiral,  $\omega - m\Omega$ , (Shu 2016). Equating this to Equation (6), we will find Lin-Shu Kalnajs dispersion relation (LSK) (Toomre 1977; Shu 2016):

$$(\omega - m\Omega)^2 = \kappa^2 - 2\pi G \Sigma_0 |k| F, \quad (7)$$

In general, the behavior of LSK dispersion relation is similar to LS dispersion for smaller wavenumber, but is different for larger wavenumbers. For the short wave regime, LSK dispersion relation approaches  $\frac{(\omega - m\Omega)^2}{\kappa^2} = 1$  but LS dispersion relation approaches  $\frac{(\omega - m\Omega)^2}{\kappa^2} > 1$ . The difference is that for fluids, pressure will become large at smaller wavelength, and on the other hand, for a collisionless stellar disc this force does not exist, and frequencies of the perturbations cannot become larger than the epicyclic frequency  $\kappa$  (Binney & Tremaine 2008; Shu 2016).

Toomre's parameter also would take a different form compared to equation (4) (Toomre 1977; Sellwood & Masters 2022):

$$\sigma_R \geq \sigma_{R,\text{crit}} \approx \frac{3.358 G \Sigma}{\kappa} \rightarrow Q \equiv \frac{\sigma_R}{\sigma_{R,\text{crit}}} \geq 1. \quad (8)$$

For a stellar thin disk, the constant 3.358 results from an exact Gaussian velocity distribution. Comparing to equation (4), we see that the constant  $\pi$  is replaced by 3.358, and the sound speed in the gas is replaced by velocity dispersion  $\sigma_R$  (Sellwood & Masters 2022; Toomre 1977). For  $Q \geq 1$ ,  $F$  remains small enough to ensure that  $\omega^2 \geq 0$ , and therefor maintaining disk stability. while for  $Q < 1$  we can have the formation of spirals (Shu 2016; Sellwood & Masters 2022). There also is a forbidden area around the vicinity of the CR; it is a region where the disk has a large  $Q$ , but the dispersion relation has no real solution, where due to pressure and random motions the density waves diminish (Binney & Tremaine 2008).

### 3.3. Propagation of The Waves

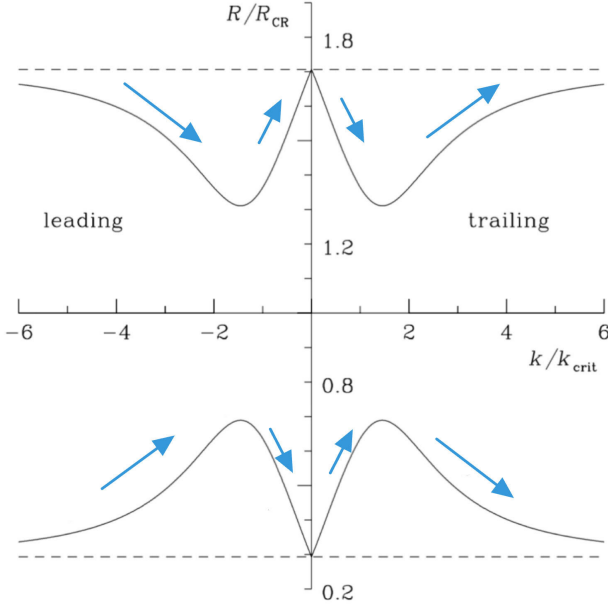
So far we have talked about waves being quasi-stationary, however, it is not very realistic and a wave would propagate radially with some group velocity once induced in a disk. The group velocity in a fluid using LS dispersion relation can be derived to be (Shu 2016; Binney & Tremaine 2008):

$$\frac{\partial \omega}{\partial k} = \pm \frac{c_s^2 |k| - \pi G \Sigma_0}{\omega - m\Omega}, \quad (9)$$

with negative and positive signs showing leading or trailing spiral.  $c_s^2 k - \pi G \Sigma_0$  is positive for short wavelength and negative for long wavelengths, and also  $\omega - m\Omega$  is negative and positive for  $R$  smaller and larger than Corotation radius (CR). Binney & Tremaine (2008) graphed the relation between wavenumber and radius in a stellar disc with  $Q = 1.2$  in Figure (2). When the group velocity is positive it propagates outwards and vice-versa, and it changes its direction once it hits the forbidden regions. In figure (2) We can see that, short-leading and long-trailing spirals will approach the CR, while short-trailing and long-leading spirals will move away from the CR (Binney & Tremaine 2008). We also can see that it will change its direction once it hits the forbidden regions.

The behavior of the wave around inner and outer Lindblad resonance (ILR and OLR) can also further be noted. Long stellar density waves ( $\frac{|k|}{k_{\text{crit}}} \ll 1$ ) are reflected at Lindblad resonances and short waves ( $\frac{|k|}{k_{\text{crit}}} \gg 1$ ) will be absorbed due to Landau damping (Binney & Tremaine 2008). This way, both long and stellar density waves cannot pass the Lindblad resonances, so the region for stellar density waves is restricted between the ILR and OLR (Shu 2016).

As the density wave packet travels inward, it becomes more tightly wrapped and eventually absorbed. All the stars in a disk experience a force as the wave passes over them, but their orbits typically change adiabatically. Lindblad resonance occurs when



**Fig. 2.** Propagation diagram for tight-winding stellar density waves following the LSK dispersion relation, for  $m = 2$ , in a stellar disk with  $Q = 1.2$ . The Horizontal dashed lines are the inner Linblad resonance at  $R = 0.293R_{CR}$  and outer Linblad resonance at  $R = 1.707R_{CR}$  (Binney & Tremaine 2008).

the forcing frequency  $\omega - m\Omega = \pm\kappa$ , with the negative sign being ILR, where stars overtake the wave, and the positive sign for OLR, where the wave overtakes the stars. At these resonances, ILR can absorb short leading density waves, leading to a concentration of material that forms an inner spiral structure, and similarly, and OLR, density waves get reflected that help maintain the outer spiral arms (Binney & Tremaine 2008; Toomre 1977). The corotation radius similarly represents the location where the spiral pattern speed equals the local rotational velocity of the disc. At this radius, stars co-rotate with the spiral pattern and enhance the structure with their interaction with the density wave (Goldreich & Tremaine 1978).

## 4. Spiral Structure Theory

### 4.1. Swing Amplification

Swing amplification is the mechanism that is crucial in the formation of spiral structures in galaxies. Swing amplification was first proposed by Goldreich & Lynden-Bell (1965) and further developed by Toomre (1981). It operates by converting leading density waves into trailing waves, and amplifying them significantly as they cross CR (Toomre 1981; Sellwood & Masters 2022).

Swing amplification relies on shearing flows, epicyclic motions, and the self-gravity of the disc (Binney 2020). The process begins with a small short leading perturbation in the density of the galactic disc. In a galactic disc, inner regions rotate faster than outer regions, which is called differential rotation (Shu 2016). Small density fluctuations in the disk, create initial disturbances, and as the disc rotates, these disturbances are sheared, by the inner parts of the disturbance moving faster than the outer parts, which results in a trailing pattern (Binney 2020). So it can be said that as the leading perturbation moves through the disc, the spiral pattern winds up, which stretches the pertur-

bation into a trailing pattern. This process so far is the “swing” part of the swing amplification (Toomre 1981; Sellwood & Masters 2022).

Stars in a galactic disc follow epicycles, and it’s in the same direction as the spiral arms are being sheared by differential rotation. During the transition from leading to trailing, as the leading wave gets stretched, it moves through the disc and eventually reaches the CR. At the CR, because the pattern speed of the spiral wave matches the angular rotation speed of the stars, their orbital frequencies match, and they resonate with each other. This resonance amplifies the wave and will increase its density contrast, and this is the *amplification* part. After passing through CR, the wave continues shearing and transforms into an amplified trailing wave. This amplified wave can maintain its enhanced density and sustain the spiral structure over a longer period (Toomre 1981; Sellwood & Masters 2022).

The self-gravity of the galactic disc can also help amplification of density waves. As the leading perturbation is stretched into a trailing one, its self-gravity helps to increase the density contrast even further. This self-gravitating amplification is essential for maintaining the spiral structure over time, particularly in regions where the mass of the disc is significant relative to the total mass supporting the galaxy’s rotation curve (Binney 2020; Toomre 1981).

As the amplified trailing wave propagates through the disk, it can induce new perturbations through its interaction with the gas and stars. Additionally, it also can create local gravitational instability and density variations leading to the formation of new leading waves. These new waves will go through the same process, and this creates a feedback loop where new perturbations continually get amplified as they pass through the corotation radius (Binney 2020; Shu 2016).

Sellwood & Carlberg (2014) identified that due to the resonant absorption at Lindblad resonances (which was previously explained in the swing amplification section), permanently modifying the DF is permanently modified because this absorption creates steep gradients in the DF around the resonance. The gradient will act as barriers that reflect subsequent waves, and when a swing-amplified wave encounters this modified DF, it is partially reflected back towards corotation and undergoes further amplification. This feedback loop increases the prominence of the wave, progressively modifying the DF more strongly with each cycle (Binney 2020).

In the next paragraphs, I will introduce a mathematical framework that lead to the two dimensionless parameters, that were introduced by Toomre (1981). In a local region away from the center, the curvature can be ignored and the equation of motion in the radial and rotational direction of galaxies will be written as (Toomre 1981):

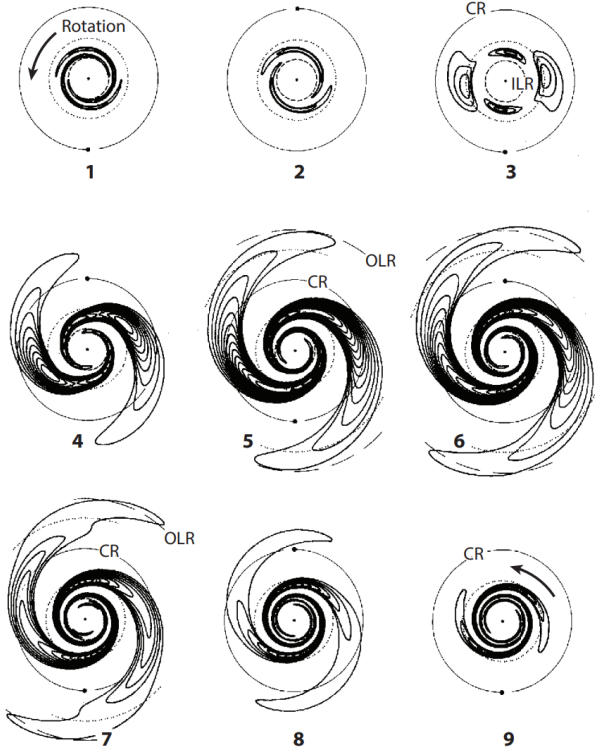
$$\ddot{x} - 2\Omega_0\dot{y} - 4\Omega_0A_0x = g \sin \gamma, \quad (10)$$

$$\ddot{y} + 2\Omega_0\dot{x} = g \cos \gamma, \quad (11)$$

Where  $x = R - R_0$ ,  $y = R_0(\phi - \Omega_0t)$ , and  $A_0 = \frac{1}{2}R_0\left(\frac{d\Omega}{dR}\right)_0$  is the Oort’s constant at  $R_0$ .  $\gamma$  is also the angle between the spiral arm and the radial direction of the galaxy, where negative and positive correspond to leading or trailing structures.

Then by using normal displacements of stars,  $\xi = x \sin \gamma + y \cos \gamma$  Toomre (1981) found:

$$\ddot{\xi} + S(\gamma)\xi = 0 \quad (12)$$



**Fig. 3.** evolution of a leading wave packet with  $m = 2$  in a disk with  $Q = 1.5$ ,  $X = 2$ , and a reduced surface density of 0.5. Numbers indicate the time sequence. Abbreviations: CR, corotation resonance, ILR/OLR: Inner/outer Lindblad resonance. Figure adapted from Toomre (1981).

where  $S$  is the spring rate and is given as:

$$S(\gamma) = \kappa^2 - 8\Omega_0 A_0 \cos^2 \gamma + 12A_0^2 \cos^4 \gamma - 2\pi G \Sigma_0 k F \quad (13)$$

$$= \left( 1 - \frac{2\Gamma}{2-\Gamma} \cos^2 \gamma + \frac{3}{2} \frac{\Gamma^2}{2-\Gamma} \cos^4 \gamma - \frac{F}{X} \sec \gamma \right) \kappa^2 \quad (14)$$

and  $\Gamma$  and  $X$  are defined as:

$$\Gamma = -\frac{d \ln \Omega}{d \ln R}, \quad (15)$$

$$X = \frac{k_{\text{crit}} R}{m} = \frac{\kappa^2 R}{2\pi G \Sigma_0 m}. \quad (16)$$

The effectiveness of swing amplification is often described using the dimensionless parameter  $X$ , which characterizes the extent of the perturbation's interaction with the disc's differential rotation. Swing amplification happens the most effectively when  $X$  varies between 1 and 2 (Sellwood & Masters 2022; Toomre 1981).  $\Gamma$  is also the shear rate of the disk, and usually varies from 0 which is a uniform rotation, to 1.5 for a Kepler potential.  $\Gamma = 1$  also indicates a flat rotation curve. One immediate result we can see here, for  $\Gamma = 0$ , galaxy is a rigid body and the spring rate is always positive, meaning that spiral arms don't amplify. We can conclude that swing amplification cannot work without differential rotation. (Sellwood & Masters 2022; Toomre 1981).

Toomre (1981) utilized linear perturbation theory to study the numerical evolution of a leading wave packet with  $m = 2$  in a disk characterized by  $Q = 1.5$ ,  $X = 2$ , and a reduced surface density of 0.5 meaning that only half of the central gravitational attraction comes from the disk, while a rigid halo contributes the

other half. Their result is illustrated in Figure (3), which demonstrates what has been discussed so far: Initially, the wave is in a leading configuration, but as time progresses, it unwinds and transitions into a more open pattern (as in frame 3). Eventually, the wave becomes a trailing pattern that increasingly winds more tightly (shown in frame 9). There also is a significant amplification of the wave amplitude, so that the trailing wave in frame 9 exhibits an amplitude approximately twenty times greater than that of the initial leading wave in frame 1. Intermediate frames (4, 5, and 6) display even stronger transient spiral patterns, highlighting the amplification process (Binney & Tremaine 2008).

The swing amplification phenomenon is not accounted for by the WKB approximation (I will elaborate on this further in the next section). In the Mestel disk scenario, the rate of change of the pitch angle of the wave is given by the equation:

$$\cot \alpha = -Rt \frac{d\Omega}{dR} = 2At, \quad (17)$$

$$\frac{d\alpha}{dt} = -\frac{2A}{1 + 4A^2 t^2}. \quad (18)$$

Where,  $A$  is the Oort's constant at radius  $R$ . When the arm is tightly wound, its rotation rate ( $\frac{d\alpha}{dt}$ ) is slow, however, as the arm swings from leading to trailing, the rotation rate reaches a maximum value of  $2A = \Omega$ . This value is comparable to the average angular speed of stars around their epicycles,  $\kappa$ , which in a Mestel disk is equal to  $\sqrt{2}\Omega$ . This synchronization enhances the gravitational interaction between the spiral and the stars, leading to the observed amplification (Binney & Tremaine 2008).

The number of dominating spiral arms in simulation, using equation (16), and setting  $X \approx 2$  (which is fair for an effective swing amplification), can be estimated to be:

$$m = \frac{\kappa^2 R}{4\pi G \Sigma_0}. \quad (19)$$

Swing amplification can be extended to explain multi-armed spirals and is not exclusive to grand-design spirals. Several numerical simulations have been instrumental in demonstrating the effectiveness of swing amplification in forming spiral structures. Simulations by Sellwood & Carlberg (1984) and subsequent studies have shown that spiral arms formed through swing amplification are transient but recurrent features, continuously forming, dissipating, and reforming due to the ongoing interaction of density waves with the disc's differential rotation (Sellwood & Masters 2022).

#### 4.1.1. Shearing Sheet Spirals

Having explored the concept of swing amplification, we now turn our attention to shearing sheet theory which offers a complementary perspective to swing amplification. As the wave packet is traveling away from the Lindblad resonances toward CR, it unwinds as it goes, and eventually, the tight winding approximation on which LSK dispersion is built on fails, and that will be when we need to use the shearing sheet. Shearing sheet theory provides a local frame to analyze the behavior of density waves within a small rectangular patch of the galactic disk (Binney 2020).

Binney (2020) rederives the results from analysis of the shearing sheet and swing amplification of Julian and Toomre (JT) and using the axisymmetric limit of the shearing sheet, rederives the LSK dispersion relation and Toomres stability criterion that was discussed in the *Dispersion Relation* section. They use  $(x, y)$  as the center of their patch where  $x$  is the radial coordinate (where  $R + x$  is a general radius), and  $y = R\phi$  where  $\phi$  is the

angle between the patch's center and the point  $(x, y)$ . By using the Lagrangian of the rotating frame, they read the momentum and find the Hamiltonian. Then they find two constant of motion, and extend the constant  $\Omega$  (which is the frequency of our patch) above to  $r = R + x$ , writing  $\Phi$  and  $H$  as (Binney 2020):

$$\Phi(R + x) \approx \Phi(R) + R\Omega^2 x + \frac{1}{2}\Omega(\Omega - 4A)x^2, \quad (20)$$

$$H_x = \frac{1}{2}p_x^2 + \frac{\kappa^2}{2}(x - \bar{x})^2 H_y = \frac{A}{B}\Delta_y^2 \quad (21)$$

where  $\Delta_y = p_y - R\Omega \approx 2\Omega x + \dot{y}$ ,  $A$  is the first Oort's constant,  $B \equiv A - \Omega$  is the second Oort's constant, and  $\kappa^2 = 4\Omega(\Omega - A) = -4\Omega B$ .

If the pattern is shearing with particles on a circular orbit, we have a continuous and time-dependant evolution of wave vector which is against QSSS. By assuming that waves evolve due to shear, Binney (2020) considered perturbation in the surface density to be as  $\Sigma_1(x, t) = \tilde{\Sigma}_1(t) \exp(ik \cdot x)$ , and for a particle on a circular orbit,  $k \cdot x$  is constant at the location of every particle, and as a result,  $k$  has to be a function of time:  $k(t) = (2At, 1)k_y$ . From here we can see that in a shearing system, non-axisymmetric waves must wind up. Binney (2020) then constructed a solution using this assumption to the linearized CBE.

LSK theory predicts no theory for wave modes near the corotation radius (for  $Q > 1$ ) because of the quasi-stationary assumption breaking down as the wave is winding due to shear. Julian & Toomre (1966), abbreviated as JT66, provided solution filling the gap for LSK theory. Semiczuk et al. (2017) rederived these equations, and got similar results as Julian & Toomre (1966) but with a different normalisation for JT Kernel. The JT equation derived by Semiczuk et al. (2017) reads as:

$$\begin{aligned} \tilde{\Sigma}_1(t) &= e^{-ik \cdot x} \int dp_x dv_\phi f_1 \\ &= \tilde{\Sigma}_{ti}(t) + \int_{ti}^t \kappa dt' K(t, t') [\tilde{\Sigma}_e(t') + \tilde{\Sigma}_1(t')] \end{aligned} \quad (22)$$

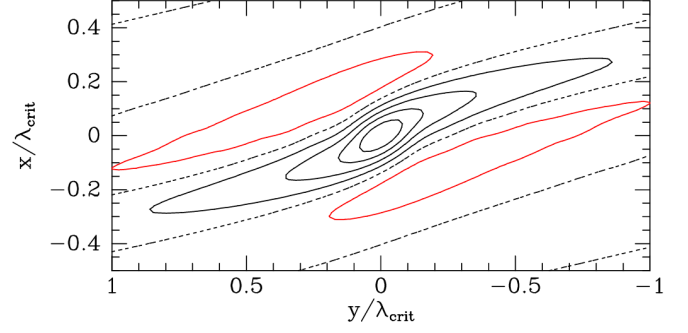
Where surface density is  $\Sigma = \Sigma_e + \Sigma_l$ , and  $\Sigma_e$  is the external density that generates an external gravitation field.  $\tilde{\Sigma}_{ti}(t)$  is also the density at  $t$  generated by the initial condition without self-gravity. And in the end,  $K(t, t')$  is also named to be the JT Kernel.

JT66 Kernels are not perfect normal modes as one would experience in electrodynamics since the general JT Kernel is not invariant under time translation due to the shearing sheet. However, in the special case of axisymmetric limit, meaning that by taking the limit of  $k_y \rightarrow 0$  and  $|t| \rightarrow \infty$ , and letting  $k_x = 2Atk_y$  to be a constant, waves do not wind up and JT is actually normal modes, identical to the dispersion relation of LSK waves with  $m = 0$ .

Binney (2020) wrote the LSK dispersion relation as:

$$0 = 1 - \frac{k_{\text{crit}}}{k_x} e^{-\chi} \sum_{n=-\infty}^{\infty} I_n(\chi) \frac{(n+1)^2 - s^2}{(n+1)} \quad (23)$$

Where  $s = ip/\kappa$ , and for  $s^2 > 0$ , if the system crosses the  $y = 0$ , then the system has a mode of frequency  $s$ , with the associated value of  $\frac{k_x}{k_{\text{crit}}}$ . The system is stable and will oscillate indefinitely if the modes happen only for  $s^2 > 0$ , and Toomre's stability criterion  $Q > 1$  can be found by finding the value of  $\sigma$  where the curve crosses  $y = 0$ . When the perturbing density has a small value of  $k_y$  ( $k_y \lesssim 0.5k_{\text{crit}}$ ), there are wavepackets on both sides of the CR that move inward and outward toward the ILR



**Fig. 4.** Overdensity created by a cloud with a Gaussian surface density located at the origin in a disc with  $Q = 1.4$ . The contour values are 2, 1.5, 0.5, 0, and -0.5 times  $M/(0.1\lambda_{\text{crit}})^2$ . The zero contours are dashed, and the negative contours are red. Axes are scaled by the critical wavelength  $\lambda_{\text{crit}}$ . Figure is adapted from Binney (2020)

and OLR, which is consistent with the LSK dispersion, however, when perturbing density has a larger value of  $k_y$ , disc responds more energetically, and the result is less localized and spreads out more and not confined to a certain region. In this situation, it will also decay more rapidly, and this behavior also cannot be explained by the LSK dispersion relation. The problem lies in the waves winding up and making  $k$  time dependant, which is again, not included in the LSK dispersion relation (Binney 2020).

An orbiting mass can cause a succession of endless  $\delta$  function stimuli. By considering a succession of broadband stimuli, Binney (2020) found that the surface density can be written as:

$$\Sigma_e(x) = \frac{M e^{-|x|^2/2\Delta^2}}{2\pi\Delta^2} \quad (24)$$

Where wavenumbers have an amplitude of:

$$\hat{\Sigma}_e(k) \equiv \int d^2x e^{-ik \cdot x} \Sigma_e(x) = M e^{-|k|^2\Delta^2/2} \quad (25)$$

By replacing it into equation (22) he found the perturbation caused by an orbiting mass:

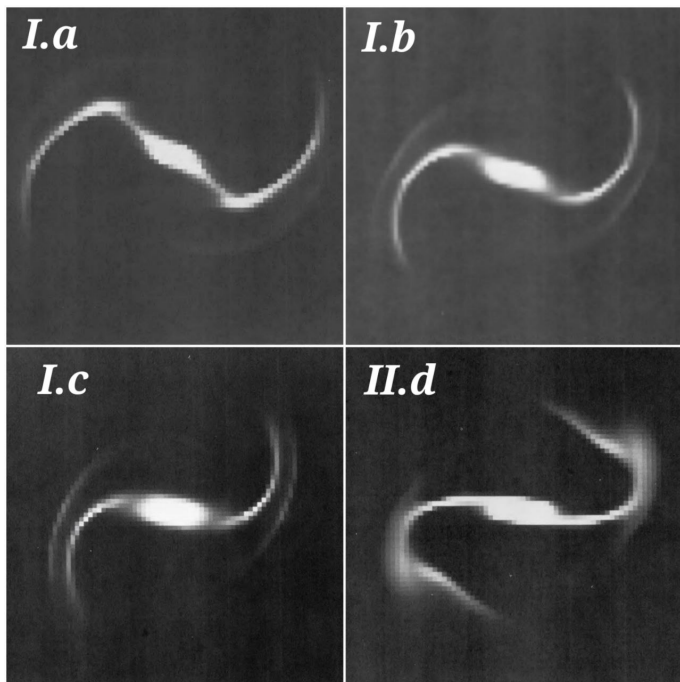
$$\hat{\Sigma}_1(k, t) = \int_{ti}^t dt' K(t, t') [M e^{-|k|^2\Delta^2/2} + \hat{\Sigma}_1(k, t')] \quad (26)$$

Any object on a circular orbit in a galactic disc has an effective mass greater than its actual mass due to the gravitational wake it generates, which holds for both massive objects like GMCs and stars with nearly circular orbits. Sundelius (1991) showed that these enhanced densities (wakes), as can be seen in Figure (4), are detectable in simulations by stacking images centered on different stars. The formation of wakes and the resulting effective mass enhancement accelerates the relaxation of the galactic disc by increasing Poisson noise (Binney 2020).

Over time the wavepacket becomes less distinct and eventually fades away. Although the wavepacket fades in real space, its effects persist in velocity space. As a result, even though the disturbances of shearing sheets have a short lifetime, Gaia mission can probe the structures at velocity space (Binney 2020).

#### 4.2. Bar Driven Spirals

Barred spiral galaxies are among the most common types of galaxies, with approximately two-thirds of all spirals having bar



**Fig. 5.** photographic representation of gas density in oval (model I) and prolate (model II) bar-like perturbation. The intensity scale is logarithmic. Each photographic square represents one cell of (80 x 80) numerical grid. Figures adapted from Huntley et al. (1978)

structure, suggesting a potential link between bars and spirals Huntley et al. (1978). Although the nature of their connection remains ambiguous, in this section I will go over the paper "*Bar-Driven Spiral Waves in Disk Galaxies*" by Huntley et al., which is one of the earliest papers that studied this topic.

The idea of bar and spiral arms having a common pattern speed is very intuitive from observation, since most spiral arms connect to the end of the bars. This idea was supported by hydrodynamical simulation such as the one done by Huntley et al. (1978) where they studied the response of gas to two types of a steady bar-like perturbation: oval and prolate. The results of their simulations can be seen in Figure (5). The panels *I.a*, *I.b*, and *I.c* are for the oval distortion models but each with different pattern speeds ( $\Omega_p$ ) (Huntley et al. 1978):

- Model Ia: Characterized by a low pattern speed, this model exhibited two inner Lindblad resonances. The resulting gas bar extended to the outer of these resonances and led the perturbing potential by about 45 degrees. The spiral structure began rather abruptly beyond the second inner resonance and this sudden formation is due to the resonance condition altering the gas dynamics significantly at these radii, causing a distinct transition from the bar to the spiral arms
- Model Ib: With an intermediate pattern speed, this model had only one inner Lindblad resonance. The central gas bar became more aligned with the perturbing potential, and the spiral arms trailed more smoothly from the ends of the gas bar, indicating a smaller shock.
- Model Ic: Featuring the highest pattern speed, this model had no inner Lindblad resonance. The gas orbits interior to corotation were aligned along the perturbation axis, resulting in trailing spiral arms beginning beyond corotation. The shocks were less pronounced and the alignment was smoother than the previous 2 models.

The prolate spheroidal model, which is the panel II.d in Figure (4), demonstrated that the central gas bar resulting from it had an extremely straight and narrow density maximum. The gas-density maxima on either side of the disk center were parallel to and offset from the minimum of the perturbing potential. This alignment suggested that the gas-density maxima led the rotation of the bar-like perturbation with the velocity field in these models showing a highly elliptical pattern of gas flow.

Bar-driven spirals are believed to be long-lived patterns (Sellwood 2011) and have no recurrent mechanism. As examples of barred spirals, one can name nearby galaxies like NGC 1300 and NGC 1365 which have regular bisymmetric spiral arms connected to the ends of their bars (Sellwood & Masters 2022). However, research by Speights & Rooke (2016) has shown that the spirals in NGC 1365 have a lower pattern speed than the bar, indicating that the bar cannot be the only driven response. While bar-driven spirals can only account for grand-design barred galaxies, there are also multiple barred galaxies that have more than 2 spiral arms, e.g, NGC 2336, and M 83 as seen in NIR images in NED (Sellwood & Masters 2022). These two discrepancies, along with the observed spirals without bars, indicate that bars cannot be the only interaction that can create the spiral pattern.

### 4.3. Tidal Driven Spirals

One of the main scenarios in the formation of grand-design spirals is their interaction with a satellite nearby. In this section, I will review the N-Body simulations performed by Semczuk et al. (2017), which will further give us insights into this scenario.

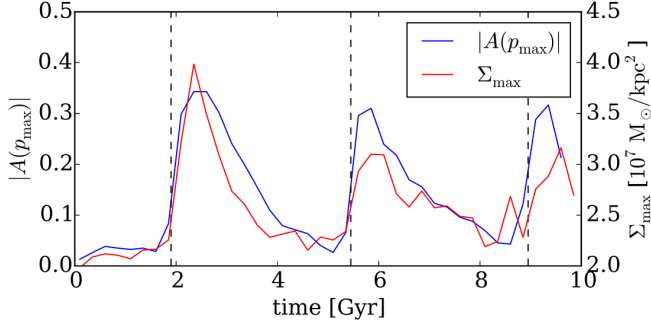
The magnitude of tidal perturbation in a galaxy due to an external body can be quantified by the dimensionless parameter  $S$ , as defined by Elmegreen et al. (1991):

$$S = \left( \frac{M_{pt}}{M_{gal}} \right) \left( \frac{R_{gal}}{d} \right) \left( \frac{\Delta T}{T} \right)^3 \quad (27)$$

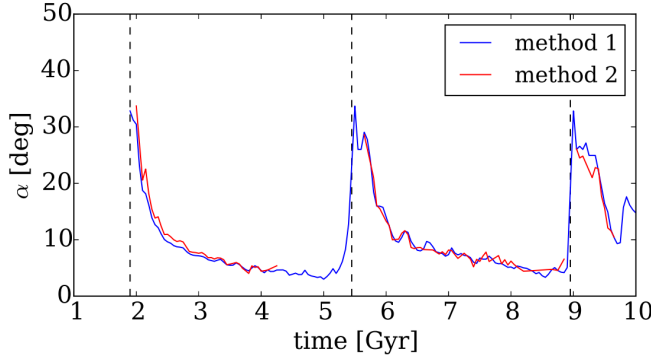
where  $M_{gal}$  and  $R_{gal}$  are the mass and characteristic size of the perturbed galaxy,  $M_{pt}$  is mass of the perturber,  $d$  is the closest approach distance between the bodies,  $\Delta T$  is the interaction time for the perturber to move one radian around the galaxy, and  $T$  is the time for stars in the galaxy's outer disk to complete one radian, expressed as  $T = (R_{gal}^3 / GM_{gal})^{1/2}$ . Semczuk et al. (2017) indicated that spiral arms in galaxies can be triggered by smaller companions when  $S$  ranges between 0.01 and 0.25 Semczuk et al. (2017).

Semczuk et al. (2017) used GADGET-2 N-body simulations to model a Milky Way-like galaxy orbiting a Virgo-like galaxy cluster with a Navarro-Frenk-White dark matter halo and an exponential stellar disk. The galaxy is placed on four different orbits with different apocentric and pericentric distances but similar eccentricities, and they tracked the galaxy's evolution over 10 billion years. The primary focus of them was however on orbit 4, which exhibited the most persistent spiral arms.

The simulations of Semczuk et al. (2017) revealed that the formation of spiral arms in galaxies is triggered by pericenter passages, where tidal forces from a cluster cause the stars in the galaxy's disk to form tidal tails. These stars remain mostly bound to the galaxy, and the resulting structure winds inward to create spiral arms, which dissipate and reform at subsequent pericenter passages. This process usually forms a two-armed, grand-design pattern that can be approximated by logarithmic spirals. Using



**Fig. 6.** Time evolution of the arm strength,  $|A(p_{\max})|$  as blue line, and surface density  $\Sigma_{\max}$  as red line made in orbit O4. Dashed lines are the pericenter passages. Figure is adapted from Semiczuk et al. (2017)



**Fig. 7.** Time evolution of the pitch angle,  $\alpha$  using two different methods of 1, and 2 in blue and red lines. Dashed lines are the pericenter passages. Figure is adapted from Semiczuk et al. (2017)

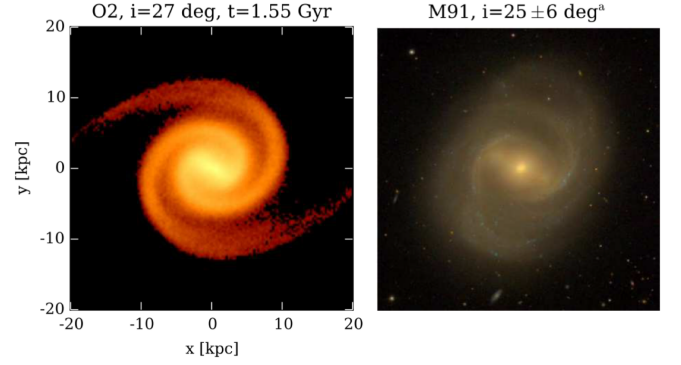
this, the surface distribution can be written as:

$$A(m, p) = \frac{1}{N_s} \sum_j \exp[i(m\phi_j + p \ln R_j)] \quad (28)$$

Where  $N_s$  is the number of stars,  $(\phi_j, R_j)$  are the polar coordinates of the  $j$ th star,  $m$  is the number of spirals, and since the resulted galaxy is mostly grand spirals;  $m = 2$ .  $p$  is a parameter related to pitch angle, and by finding the  $p_{\max}$  that maximizes  $A(2, p) = A(p)$ , they found the pitch angle ( $\alpha$ ) using  $\tan \alpha = 2/p_{\max}$ . They also set  $9\text{kpc} \leq R \leq 15\text{kpc}$ , since for  $R < 9\text{kpc}$ , they observed the pitch angle to be  $90^\circ$ , meaning that the bar was contaminating the result.  $|A(p_{\max})|$  can be thought of as a parameter to measure the arms' strength.

Semiczuk et al. (2017) also used a secondary method to determine the pitch angle by fitting logarithmic spirals directly to the surface density distribution. Using least squared method,  $\phi = B \ln R + C$ , the pitch angle was found to be  $\tan \alpha = 1/|B|$ . Then after selecting points of maximum density at various radii and fitting spirals to these points, the study confirms that  $\Sigma_{\max} = \Sigma(\Phi_{\max})$  has a similar evolution to  $|A(p_{\max})|$ , as can be seen in Figure (6), they both get to a maximum value after the pericenter passages and then their value decrease until the next passages. Semiczuk et al. (2017) noted that the peaks of the arms' strength happen around 0.5 Gyr after the pericenter passage. This is due to the tidal features being the strongest in the outer parts of the disk and needing some time for the effect to get to the inner part and the ring, where the measurement is performed.

Also as shown in Figure (7), the pitch angle  $\alpha$  of the spiral arms in a specific radial range decreases from  $\alpha \approx 30^\circ$  to below  $\alpha \approx 10^\circ$ , indicating that the arms wind up over time. They



**Fig. 8.** The right panel is the surface density maps of the simulated galaxy modified to correspond to M91 galaxy. The right panel is the SDSS image of the M91 galaxy. Figure is adapted from Semiczuk et al. (2017)

also observed that the pitch angle peaks at the pericenter, which demonstrates that the strongest arm formation happens shortly after each passage and diminishes thereafter. Unlike the arm's strength, the maximums happen right after each passage.

Semiczuk et al. (2017) measured the pattern speed of the spiral arms using two methods; The first involves calculating the cross-correlation of the perturbed surface density at different times, while the second finds the maximum density at fixed radii across different epochs. Both methods yield consistent results, indicating that the pattern speed decreases with radius, which suggests that it is not a quasi-stationary density wave, and it is a kinematic density wave Semiczuk et al. (2017). In general resonances is given as  $m(\Omega - \Omega_p) = \pm \frac{\kappa}{n}$ , with  $n = 1$ , for Lindblad resonances, and  $m = 2$  for grand design spirals. As a result, for a tidal-driven grand design spiral, the pattern speed will be  $\Omega_p = \Omega - \kappa/2$  for ILR, and  $\Omega_p = \Omega + \kappa/2$  for OLR. (Toomre 1977). Semiczuk et al. (2017) find the pattern speed to be close to inner Lindblad resonance, confirming further arms are kinetic density waves.

In the end Semiczuk et al. (2017) made a Comparison with Observations by comparing its findings with observed spiral galaxies in the Virgo cluster. They identified nine galaxies that do not show signs of interaction with satellites or other galaxies, ensuring that the observed spiral structures are likely due to cluster tidal forces. The comparison reveals good agreement between the shapes and behaviors of the simulated and observed spiral arms, as can be seen in Figure (8).

Many simulations over the years have proved the fact that tidal interactions can very likely be responsible for grand-design spirals that are near a companion, with M51 and M81 being the clear examples of this phenomenon. However, we observe some grand-design galaxies that are not in the vicinity of any other companions. One solution that has been proposed is that maybe dark matter subhalos could also be responsible, but this would require multiple and non-disruptive passages of subhalos which seems impossible (Sellwood & Masters 2022). As much as tidal interactions are successful in explaining grand-design spirals, it is unable to explain multi-armed or flocculent galaxies (Semiczuk et al. 2017). Regarding their lifetime, some simulations also show that Tidal-driven spirals have a lifetime of 1Gyr, but taking the fact that they go through repeated cycles of tidal interaction (as observed in Semiczuk et al. (2017)) into account, it can be enough to explain the presence of  $m = 2$  spirals in our universe (Sellwood 2011).

## 5. Spiral Arms: Long-Lived or Transient?

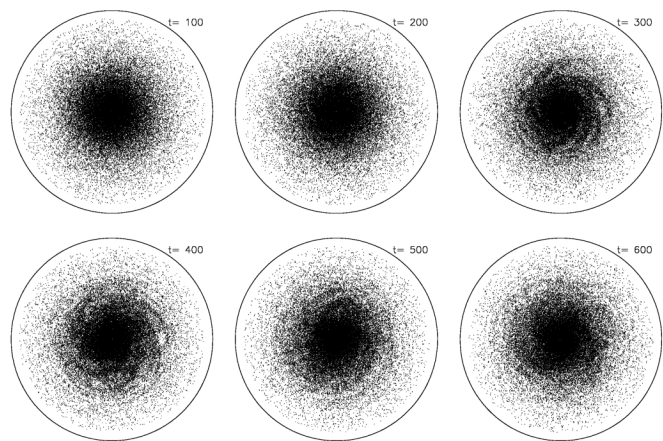
To create and sustain spirals, angular momentum needs to be taken from the stars in the inner part and given to the stars in the outer part which is done through the resonances. Transient spirals are more effective at redistributing angular momentum because they involve broad, time-dependent resonances that affect large parts of the galaxy, whereas long-lived spirals have narrow, localized effects confined to specific resonances, which makes them less influential in driving the internal evolution of galaxy discs. The lifetime of spiral patterns determines how quickly and effectively they can transport angular momentum, scatter stars into different orbits, and mix the material within the galaxy. Transient spirals, by being short-lived and recurring, are more effective at driving these changes and thus have a greater impact on the evolution of galaxy discs compared to long-lived, quasi-steady patterns (Sellwood 2011).

Observational evidence from near-infrared photometric images like Schweizer (1976) does not provide conclusive evidence on their lifetimes. The velocity distribution of stars in the Milky Way suggests a preference for transient spiral patterns, since studies such as *Hipparcos mission* (Dehnen 1998) show non-circular motions and irregular velocity distribution suggesting that the stellar motion have been influenced by short-lived spirals. If the spirals were long-lived we would expect a more uniform velocity distribution (Binney & Tremaine 2008; Sellwood 2011).

Sellwood (2011) adopts a model from the BLLT survey, which is similar to the model proposed by Fall & Efstathiou (1980), that is a modified exponential disc model with a central density dip and including a small pseudo-isothermal halo. Sellwood & Carlberg (2014) ran two series of simulations: one with only  $m = 2$  disturbance forces (bisymmetric mode) and another including higher sectoral harmonics with  $m$  varying between 2 and 8. The results showed that when only  $m = 2$  disturbance forces were allowed, the disc remained stable with no significant non-axisymmetric features or heating over 40 disc rotations. However, when higher harmonics were included, the disc quickly developed multi-arm spiral patterns (as observed in figure (9)). The disc also underwent significant heating, indicating that swing amplification of these higher harmonics led to the observed spiral structures and increased random kinetic energy in the disc. The reason for the different behavior is for the swing amplification effectiveness parameter  $X$  (equation 16), which lies between 1 and 2.5 for  $m \geq 4$ , leading to strong and efficient amplification, while it has higher values than 2.5 for  $m = 2$ . This suggested that the BLLT model's predicted bisymmetric spiral mode is not robust under realistic conditions (multiple harmonic disturbances), implying that such spirals are likely transient rather than long-lived.

The simulations by Sellwood (2011) demonstrate that galaxy models, which are designed to support long-lived, bisymmetric spiral modes, quickly develop short-lived, multi-arm spiral patterns due to swing amplification. These transient patterns rapidly heat the disc and alter the conditions necessary for sustaining long-lived spirals. New stars are born in gas clouds, on nearly circular orbits. Gas clouds as time goes by dissipate their random motion through collisions. New stars being added to the disk at a rate of a few stars per year, "cools" the spirals, and allows it to be responsive to perturbations. This is consistent with the observation that gas-rich galaxies often exhibit spiral patterns while gas-poor S0 galaxies do not (Sellwood 2011).

The spread in metallicities among stars of the same age cannot be explained by simple chemical evolution models that lack



**Fig. 9.** Time evolution of BLT model for sectoral harmonics bigger than 2. Only one particle in 400 particles is plotted and the outer circle has a radius of  $R = 6.5$  with one rotation at  $R = 2$  being 13 time units. The figure is adapted from Sellwood (2011).

radial mixing. Instead, recurrent transient spiral patterns induce radial churning, causing stars to migrate across the disc and leading to the observed metallicity spread. This process happens due to the spirals being transient and recurrent, ensuring ongoing mixing without significant heating of the disc (Sellwood 2011).

## 6. Conclusion

This review has explored the primary mechanisms behind the formation and evolution of spiral arms in galaxies, and mainly focusing on swing amplification, tidal and bar interactions. After confirming the very fast winding up of the material arms, inspired by the idea of kinematic density waves, Lin & Shu (1964); Kalnajs (1965); Lin & Shu (1966b) derived the LS dispersion relation for fluids, and the LSK dispersion relation for a stellar disc. Using the LS and LSK dispersion relations, Toomre's parameter  $Q$ , as in equations (4) and (5) were derived respectively as stability criteria (Shu 2016).

The swing amplification process was also introduced in this review, that was developed by Toomre (1981). Swing Amplification driven by differential rotation and self-gravity, is when a short leading wave, is reflected to short trailing waves at the CR, and its density is enhanced. Swing amplification relies on the continuous regeneration of spiral arms through the interaction of density waves with the galactic disc, causing a feedback loop and has been shown to produce transient, multi-armed, or grand-design spiral patterns (Toomre 1977). Briefly, after swing amplification, shearing sheets were introduced as a more comprehensive framework for when LSK dispersion relation fails due to the spiral arms winding up and the failure of quasi-stationary assumption (Binney 2020).

Tidal interactions with companion galaxies or dark matter subhalos are another possible driver of spiral arm formation, with M51 and M81 being two strong evidence for it. Tidal interactions tend to produce grand-design spirals and are unlikely to produce multi-armed spirals (Semczuk et al. 2017). Tidal interactions cause transient spiral arms, with multiple passages and the repeated cycle of forming spirals and unwinding (Sellwood 2011).

Bar interactions also can contribute to the formation of spiral arms. Bar-driven spirals are mostly grand-design spirals with a common pattern speed between the bar and the spiral. There also

have been some discrepancies, either by finding multi-armed barred galaxies, e.g. NGC 2336, and M 83, or by finding different pattern speed, e.g. NGC 1365, suggesting the bars can not be the only factor in spiral formation (Sellwood & Masters 2022; Huntley et al. 1978). Regarding their lifetime, Bar-driven spirals are suspected to be long-lived patterns (Sellwood 2011).

While certain mechanisms, like bar interactions, suggest the possibility of long-lived spiral patterns, the simulations by Sellwood (2011), showed the prevalence of transient spirals indicated by swing amplification. He found that all long-lived grand-design modes quickly turned into transient multi-armed spirals. Observational evidence such as the lack of spiral arms in S0 galaxies, or radial mixing, further points into the transient picture of spirals (Sellwood 2011).

In conclusion, the formation and sustainability of spiral arms in galaxies are governed by a combination of mechanisms that can produce both long-lived and transient structures. The contradictions between these mechanisms and the varying observational evidence highlight the complexity of spiral galaxy dynamics and underscores the need for continued research to fully understand the nature of these fascinating cosmic features.

## References

- Binney, J. 2020, MNRAS, 496, 767  
 Binney, J. & Tremaine, S. 2008, *Galactic Dynamics: Second Edition*  
 Buta, R. J., Sheth, K., Athanassoula, E., et al. 2015, ApJS, 217, 32  
 Dehnen, W. 1998, AJ, 115, 2384  
 Elmegreen, D. M., Sundin, M., Elmegreen, B., & Sundelius, B. 1991, A&A, 244, 52  
 Fall, S. M. & Efstathiou, G. 1980, MNRAS, 193, 189  
 Goldreich, P. & Lynden-Bell, D. 1965, MNRAS, 130, 125  
 Goldreich, P. & Tremaine, S. 1978, ApJ, 222, 850  
 Huntley, J. M., Sanders, R. H., & Roberts, W. W., J. 1978, ApJ, 221, 521  
 Julian, W. H. & Toomre, A. 1966, ApJ, 146, 810  
 Kalnajs, A. J. 1965, PhD thesis, Harvard University, Massachusetts  
 Lin, C. C. & Shu, F. H. 1964, ApJ, 140, 646  
 Lin, C. C. & Shu, F. H. 1966a, Proceedings of the National Academy of Science, 55, 229  
 Lin, C. C. & Shu, F. H. 1966b, Proceedings of the National Academy of Science, 55, 229  
 Schweizer, F. 1976, ApJS, 31, 313  
 Sellwood, J. A. 2011, MNRAS, 410, 1637  
 Sellwood, J. A. & Carlberg, R. G. 1984, ApJ, 282, 61  
 Sellwood, J. A. & Carlberg, R. G. 2014, The Astrophysical Journal, 785, 137  
 Sellwood, J. A. & Masters, K. L. 2022, ARA&A, 60 [arXiv:2110.05615]  
 Semczuk, M., Łokas, E. L., & del Pino, A. 2017, ApJ, 834, 7  
 Shu, F. H. 2016, ARA&A, 54, 667  
 Speights, J. C. & Rooke, P. C. 2016, The Astrophysical Journal, 826, 2  
 Sundelius, B. 1991, Dynamics of Disc Galaxies  
 Toomre, A. 1977, ARA&A, 15, 437  
 Toomre, A. 1981, in *Structure and Evolution of Normal Galaxies*, ed. S. M. Fall & D. Lynden-Bell, 111–136