Unstable spiral modes in disk-shaped galaxies

(density waves/galactic dynamics/spiral arms)

Y. Y. LAU, C. C. LIN, AND JAMES W.-K. MARK

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Mass. 02139

Contributed by C. C. Lin, February 9, 1976

ABSTRACT The mechanisms for the maintenance and the excitation of trailing spiral modes of density waves in disk-shaped galaxies, as proposed by Lin in 1969 and by Mark recently, are substantiated by an analysis of the gas-dynamical model of the galaxy. The self-excitation of the unstable mode is caused by waves propagating outwards from the corotation circle, which carry away angular momentum of a sign opposite to that contained in the wave system inside that circle. Specifically, a simple dispersion relationship is given as a definite integral, which allows the immediate determination of the pattern frequency and the amplification rate, once the basic galactic model is known.

1. Introduction

The development of the theory of normal galactic spirals has been, for a long time, based on the hypothesis that there exists a pattern of density waves with a quasi-stationary spiral structure (QSSS hypothesis) (1, 2). On the other hand, Toomre's study (3) of the radial propagation of waves with a relatively tight winding shows that a pure short trailing wave system would move away from the corotation region (i.e., the region where the angular velocities of the wave pattern and of the material are about the same). Since the propagation time of these waves over the whole galactic disk is found to be of the order of one billion years, Toomre pointed out that the quasi-stationary hypothesis was questionable, unless a mechanism for replenishment could be provided.

Lin (4) responded to this challenge by suggesting the following picture. When these short trailing density waves reach the central regions of the galactic disk, they could induce a feedback process either by outward-propagating long trailing waves or directly by a bar-like structure formed through the action of these short trailing waves. The latter process has been calculated by Feldman and Lin (5) and by Lin and Lau (6). The driving mechanism due to long waves has been calculated by Mark (7–10) in the stellar case and confirmed by Lin and Lau (6) in the gas dynamical model. It is found that there is in general an amplification of these waves by "stimulated emission" at the corotation circle. Mark (10–13) then proposed that unstable modes can actually be obtained by combining the feedback and the amplification mechanisms. Estimations were given of the rates of amplification.

In this note, we shall present the principal results of one simple analysis, following closely the mechanisms proposed. It yields, in the asymptotic theory, explicit solutions that represent unstable modes. These solutions exist in broad classes of disk models with reasonable distributions of density and rotational velocity. If the disk is locally stable everywhere and the corotation circle is in a *marginally* stable region, then the *e*-folding time *T* for energy is given by the formula

$$T = \tau / \ln 2 = 1.44\tau \quad [11]$$

where τ is the time required for the waves to travel one complete cycle. In a typical cycle, a packet of short trailing waves begins at the corotation circle. It propagates inwards to a central turning point (to be defined more precisely below), where it is turned back outwards (as long trailing waves) to return to the corotation circle. The mechanism for this amplification process will now be described before we present the analytical theory.

Other mechanisms related to the generation and maintenance of spiral modes will be briefly discussed in *Section 4*.

2. The models and the mechanisms

For our present purposes, we shall adopt a gas-dynamical model, since our primary aim is to demonstrate the basic mechanism, and the gas-dynamical model has the advantage of simplicity. Indeed, a similar theory has also been developed by Mark (manuscript submitted) for the stellar dynamical model under certain reasonable simplifying approximations, some of which are not needed here. The fundamental dispersion relationship is then slightly different from [3.6]. However, a careful comparison shows that, at least for the type of models under consideration, there is no qualitative difference between the results for a model based on gas dynamics and one based on stellar dynamics. Thus, a consideration of the gas dynamical model also helps to clarify the role of the resonant regions: the behavior of individual corotating (or resonant) stars is not important.

To provide a feedback process from central regions, we shall, for the moment, consider galactic disk models without inner Lindblad resonance. We shall further assume that the stability parameter (14, 15) Q (defined in Eq. 3.4 below) increases steadily towards the center (12), so that the disk is essentially "immobile" there. This is a good dynamical simulation of the central regions in real galaxies where there may be a nuclear bulge. It can then be shown that the energy of the incoming short waves is essentially fully turned back by refraction. (See *Section 4* below for other possibilities.)

The amplification process occurs essentially near the corotation region. Consider neutral waves for the moment. As the waves complete the cycle described above, the (radially) outgoing wave, at unit energy flux, induces an inward propagating wave from the corotation circle with twice as much energy flux, since there is also an outgoing wave induced, which carries away energy of an opposite sign, at unit flux, to the outer regions of the galactic disk (6–10). Thus, there is doubling of energy of the waves over a period of time τ , and hence the formula [1.1] given above. Naturally, this formula has also been deduced by straightforward mathematical analysis. (See *Section 3* below.) Somewhat different numerical values will be obtained if the disk is not marginally stable in the corotation region (see ref. 10, which also contains results for cases $Q \neq 1$ in the corotation region).

Thus, if we consider the wave system in the whole region

within the corotation circle, it carries negative total energy. As a mode, it is amplified by the outward transfer of energy of an opposite sign across the corotation circle. In terms of angular momentum, this amounts to the transfer of angular momentum outwards by the trailing wave system, in line with the general trend suggested by Lynden-Bell and Kalnajs (16). However, they attributed the generation mechanism to the resonant stars. In our analysis, this effect is neglected, since it has been shown (7–10) that it only leads to a minor modification of the mechanism discussed here.

The fate of the outward propagating waves is easily understood in the stellar dynamical model. They are simply absorbed at the (first-order) *outer* Lindblad resonance. In the gas-dynamical model, it is less clear. However, in a real galaxy, there are even other ways for disposing of the *positive* wave energy (e.g., by turbulent dissipation) besides absorption at the outer Lindblad resonance. The adoption of the radiation boundary condition in the gas dynamical model is therefore just a mathematical device to simulate the real galaxy without explicitly introducing dissipative mechanisms.

3. Analytical results

The basic equations governing density waves in a self-gravitating gaseous disk may be found in a convenient form in the paper by Lin and Lau (6). If we retain only the dominant terms in the Euler equations of motion and apply an asymptotic formula to the Poisson equation, we obtain the following differential equation for the enthalpy perturbation h (see ref. 6, p. 362):

$$\frac{d^2 u}{dr^2} + k_3^2 u = 0 \quad [3.1]$$

where

$$u = h \exp \left\{ i \int^r dr \Sigma / 2 \right\} \quad [3.2]$$

$$k_3^2 = \frac{\Sigma^2}{4} - \left(\frac{\kappa}{a} \right)^2 (1 - \nu^2) \quad [3.3]$$

$$\Sigma = \left(\frac{\kappa}{a} \right) \frac{2}{Q}, \text{ and } Q = \frac{\kappa a}{\pi G \sigma} \quad [3.4]$$

In these equations, r is the distance from the galactic center, κ is the epicyclic frequency, σ is the surface density, a is the acoustic speed, $\nu = (\omega - m\Omega)/\kappa$ where ω is the mode frequency, m is the number of the spiral arms, and Ω is the circulation frequency (circular velocity in angular measure). Note that Σ is proportional to the surface density. At large distances r , where we assume $Q = 1$, we have $k_3^2 = (\kappa/a)^2 \nu^2$. Note that k_3^2 has a *double zero* at the corotation radius r_{co} , defined by $\nu(r_{co}) = 0$.

As mentioned in Section 2, an unstable normal mode may be expected, if there is a mechanism for turning incoming waves back near the galactic center. We see here that this type of “reflection” near the galactic center is indeed possible if there is a location $r = r_{ce}$ when k_3^2 has a simple zero. This simple zero can occur in a model in which Q steadily increases as $r \rightarrow 0$. In that case k_3^2 has the appearance shown in Fig. 1.

We consider a mode which satisfies the boundary conditions

$$u = 0 \text{ at } r = 0, \text{ and} \quad [3.5a]$$

$$u \sim \exp \left[-i \int_{r_{co}}^r dr \frac{\kappa}{a} \nu \right], \text{ for } r \gg r_{co}. \quad [3.5b]$$

The first condition follows from the condition of regularity, and the second is one of outgoing waves. These conditions then lead to the following dispersion relation for the determination of ω :

$$\exp \left[-i \oint k dr \right] = -\sqrt{2} \quad [3.6]$$

FIG. 1. Schematic diagram showing the variation of k_3^2 (the square of the reduced wave number defined in Eq. [3.3] in the text) as a function of radial distance r .

where the closed path denotes the cycle ($r_{ce} \rightarrow r_{co} \rightarrow r_{ce}$). In [3.6], $k = -\Sigma/2 \pm k_3$, where the branch of k_3 is so defined that $\text{Re } k_3 > 0$. The positive sign is used as r varies from r_{ce} to r_{co} , and the negative sign is used in completing the integral from r_{co} back to r_{ce} . Equating the magnitude and phase separately in Eq. [3.6], we obtain the following pair of approximate dispersion relations, in the limit of small growth rate:

$$\oint k dr = (2n + 1)\pi, (n = 0, 1, 2, \dots) \quad [3.7a]$$

and

$$e^{i\omega_i \tau} = \sqrt{2} \quad [3.7b]$$

where

$$\tau = \oint \frac{dr}{|c_g|} \equiv \oint \frac{dr}{|\partial \omega / \partial k|} \quad [3.7c]$$

is the transit time for a group of waves to complete a cycle ($r_{ce} \rightarrow r_{co} \rightarrow r_{ce}$). Eq. [3.7b] clearly indicates that the e -folding time is on the order of the time required for a group of waves to propagate across the galactic disk, which is of the order of one billion years.

Solutions have been found for the eigenvalues of ω on the basis of Eq. [3.7] for certain specific models. Clearly, the rate of amplification calculated depends on the amplitude distribution of the wave, which is not properly described by Eq. [3.1]. However, this does not invalidate the above formulas [3.7]. The principle of conservation of density of wave action holds (3, 17, 18) both for the exact theory and for the approximate form [3.1]. Since the propagation of the waves is well described by Eq. [3.1], the same amplification rate is obtained in both cases. The accurate amplitude distribution must be obtained by going to the next approximation. Better still, it should be calculated by numerical integration, by using the exact linearized differential equations of motion and the integral form for the Poisson equation.

It should be emphasized that the dispersion relationship [3.6] is approximate. The accurate treatment of the eigenvalue problem might actually lead to a very narrow band of continuous spectrum.

Example. If we take a rotation curve given by $\Omega = 250/(r + 1)$ km/sec-kpc, where r is in kiloparsecs (kpc), $a = 45$ km/sec, and $Q = 1$ for $r \geq 5$ but increases roughly like $5/r$ for smaller values of r , we have a reasonable model for our purposes. For this model, there is a solution with $n = 0$ in Eq. [3.7], with $r_{co} = 8.5$ kpc and $r_{ce} = 3.5$ kpc, approximately. The short wavelength approximation is well satisfied even at $r = 3$ –4 kpc, since $|kr| = |\Sigma/2 \pm k_3|r \gtrsim 4$. It is even better satisfied at outer parts.

For spiral waves with $|kr|$ of the order of 4, we have also performed numerical calculations to check the accuracy of our asymptotic solution of the Poisson equation by comparing it with the potential obtained by an exact calculation based on the

Poisson integral formula. The results are found to be in agreement with each other to within 5% as long as $|kr| > 4$.

4. Discussion

(a) **Modal Analysis.** As mentioned above, the asymptotic theory demonstrated in this note can be improved and put on a more solid foundation by numerical calculations based on the exact linearized equations.

Another improvement to be made is the inclusion of *inner Lindblad resonance*. Certainly, the return of wave energy from the central region would no longer be perfect in this case. But this would presumably only lead to a reduction of the amplification, at least in many cases. Basically, we still have the same physical picture: the waves propagating outwards from the corotation circle amplify the system of waves inside it by carrying away angular momentum of an opposite sign. There is of course the possibility that the energy may be completely absorbed at the inner Lindblad resonance. In that case, the above mechanism would not work, and we would need a different process. [See (b) below.]

Zang and Toomre (Ph.D. Dissertation of Thomas Zang, 1976, Massachusetts Institute of Technology) have found self-excited modes for galactic models with uniform linear velocity and various distributions of the mobile component of mass density. Their models do exhibit an inner Lindblad resonance. The underlying mechanism is not explicitly displayed by their analysis. It is suggested here that one component of the sustaining mechanism could be of the type described in this note. However, the length scales and the amplification rates obtained by Zang and Toomre seem both to be too large for our calculations to be properly applied. There are probably other amplifying mechanisms in their particular galactic models.

(b) **Some Other Mechanisms.** We would now briefly refer to two other mechanisms for generating unstable modes. One is the interaction with a halo whenever it exists (13), the other is Jeans instability in the outer regions (19). In case Jeans instability is present one could have a distinct mode, even when the mechanism discussed in this paper is absent. If both mechanisms were present, there would be a question of proper phase matching. In general, the spiral pattern may not be perfect. Since real galaxies do not usually have perfect patterns, one should regard our theoretical studies as providing the various possible mechanisms, and leave the choice of the preferred mechanisms and their relative importance to natural forces in action.

(c) **Reflection from Galactic Center.** In the previous section, we considered the reflection of waves from the galactic central regions as a result of the increase of the parameter Q near the galactic center. If the differential equation developed for the asymptotic treatment of these density waves *remains* essentially valid near the center, then we find that there is a simple modification of our theory possible in order to include the effect of cylindrical geometry. We replace the coefficient k_3^2 in [3.1] by

$$\bar{k}_3^2 = k_3^2 - \frac{m^2}{r^2} \quad [4.1]$$

This function is definitely of the form shown in Fig. 1 (with a slight modification near corotation), whether Q increases with decreasing r . The feedback process always works. Indeed, certain dynamical factors may even further enhance this geo-

metrical effect. However, the whole discussion is uncertain because the differential equation used may not be valid at the origin, since the usual asymptotic solution of Poisson's equation is not accurate there. Further investigation is needed.

(d) **Analogy with Plasma Physics.** The mechanism discussed in this paper is similar to a type of instability known in plasmas. In 1969, Berk *et al.* (20) found that instability may be excited by an exchange of energy between different regions in a non-uniform plasma in a mirror machine geometry. The consideration of the suppression of the unstable modes of this type led to a criterion on the critical length of the mirror device (21). The same critical length scale has been approximately reproduced by Lau (manuscript submitted) in a simplified form of the theory, which is mathematically similar to that presented in this note.

(e) **The Role of Resonant Stars.** Even though we stressed the role of the resonant stars in the framework of the linear theory, they may not play a crucial role in the maintenance mechanism of spiral structure in real galaxies. For one thing, after the spiral structure is generated, these stars might be displaced from the condition of resonance after a few hundred million years. If so, the absorption of waves at outer Lindblad resonance in the analytical theory must then be taken over by physically significant dissipative processes (see Section 2). The lack of absorption at inner Lindblad resonance can only aid the reflection process from central regions described above.

We wish to thank Dr. A. Toomre for helpful discussions on several aspects of the problem. This work is supported in part by a grant from the National Science Foundation.

1. Lin, C. C. & Shu, F. H. (1964) *Astrophys. J.* **140**, 646–655.
2. Lin, C. C. & Shu, F. H. (1968) in *Astrophysics and General Relativity*, ed. Chretien, M. (Gordon and Breach, New York), Vol. 2, pp. 235–328.
3. Toomre, A. (1969) *Astrophys. J.* **158**, 899–913.
4. Lin, C. C. (1970) in *Proc. International Astronomical Union Symposium No. 38* (Basel), pp. 377–390.
5. Feldman, S. I. & Lin, C. C. (1973) *Stud. Appl. Math.* **52**, 1–20.
6. Lin, C. C. & Lau, Y. Y. (1975) *SIAM J. Appl. Math.* **29**, 352–370.
7. Mark, J. W.-K. (1974) in *The Formation and Dynamics of Galaxies*, ed. Shakeshaft, J. R. (D. Reidel Publ. Co., Boston, Mass.), pp. 417–423.
8. Mark, J. W.-K. (1975) in *La Dynamique des Galaxies Spirales*, ed. Weliachew, L. (Centre National de la Recherche Scientifique, Paris), pp. 67–79.
9. Mark, J. W.-K. (1976) *Astrophys. J.* **203**, 81–96.
10. Mark, J. W.-K. (1976) *Astrophys. J.*, in press (April 15 issue).
11. Mark, J. W.-K. (1974) *Bull. Am. Astron. Soc.* **6**, 462.
12. Mark, J. W.-K. (1975) *Bull. Am. Astron. Soc.* **7**, 395.
13. Mark, J. W.-K. (1976) *Astrophys. J.*, in press (May 15 issue).
14. Toomre, A. (1964) *Astrophys. J.* **139**, 1217–1238.
15. Julian, W. H. & Toomre, A. (1966) *Astrophys. J.* **146**, 810–830.
16. Lynden-Bell, D. & Kalnajs, A. J. (1972) *Mon. Not. R. Astron. Soc.* **157**, 1–30.
17. Shu, F. H. (1970) *Astrophys. J.* **160**, 99–112.
18. Mark, J. W.-K. (1974) *Astrophys. J.* **193**, 539–559.
19. Lau, Y. Y., Lin, C. C., Mark, J. W.-K. & LaBudde, R. A. (1975) presented at the *Seventh Annual Meeting of the Division of Dynamical Astronomy*, American Astronomical Society (Pasadena, Calif., Dec. 11, 12).
20. Berk, H. L., Pearlstein, L. D., Callen, J. D., Horton, C. W. & Rosenbluth, M. N. (1969) *Phys. Rev. Lett.* **22**, 876–879.
21. Berk, H. L., Pearlstein, L. D. & Cordey, J. (1972) *Phys. Fluids* **15**, 891–910.