THE ASTROPHYSICAL JOURNAL, Vol. 160, April 1970

ON THE DENSITY-WAVE THEORY OF GALACTIC SPIRALS. I. SPIRAL STRUCTURE AS A NORMAL MODE OF OSCILLATION

FRANK H. SHU
State University of New York at Stony Brook
Received 1969 August 5; revised 1969 October 2

ABSTRACT

An exact formulation of the linearized problem, including appropriate boundary conditions, is developed to explore whether extensive galactic density waves of spiral form are permissible normal modes of oscillation for a stellar disk. An "anti-spiral theorem," of the type reported previously by Lynden-Bell and Ostriker for neutral modes in a *gaseous* disk, holds here with limited validity—namely, whenever the effects of stellar resonances can be ignored.

I. INTRODUCTION

Various writers starting with B. Lindblad have proposed that the spiral structure found in galaxies whose basic mass distribution is flat and axisymmetric can be understood in terms of a density wave which is maintained by the gravitational field of the disk stars. Much of Lindblad's work on normal spirals (see, e.g., Lindblad 1963) leaned heavily on the behavior of individual stars; consequently, he was not able to discuss quantitatively the properties of the collective modes of stellar oscillations. To remedy this shortcoming, Lin (1966), Lin and Shu (1964, 1966, 1967), Kalnajs (1965), and Shu (1968) employed in their analyses the statistical theory of stellar dynamics. The observational support for the theory was subsequently summarized by Lin, Yuan, and Shu (1969).

a) Remaining Difficulties with the Theory

Notwithstanding the good agreement with observations, considerable difficulties remain. The validity of gravitational theories for *spirals of extensive structure* is challenged by the proof of Lynden-Bell and Ostriker (1967) that patterns of a spiral form do not generally exist as normal modes of oscillation for a differentially rotating and nondissipative gaseous system. Insight concerning this difficulty is obtained from the demonstration by Toomre (1969) that waves of a spiral form (whether in a gaseous disk or in a stellar disk) propagate in the radial direction with appreciable group velocity. Even existing spiral waves must eventually disappear if they are neither replenished nor returned.

b) Objectives of the Present Paper

Apart from interactions with interstellar magnetic fields, the dynamics of stars differs from the dynamics of gas (and dust) in that stars can resonate with an oscillating gravitational field without continual interruptions from collisions. It is demonstrated in the present paper (see also Shu 1968) that this difference in a stellar system limits the applicability of an "anti-spiral theorem."

A by-product of this work is that the exact dynamical results obtained here allow the properties of galactic density waves to be systematically studied in the WKBJ approximation (Shu 1970, hereinafter referred to as Paper II). The principal new contribution of Paper II is the explicit demonstration that the "action density" of spiral waves propagates with the group velocity found by Toomre.

II. STELLAR RESPONSE TO A PERIODIC DISTURBANCE

We begin our study by computing the planar response of a stellar disk exposed to the action of an unsteady gravitational field. We set aside for the time being the question of whether this field is self-consistent.

The time development of the phase density for a galaxy of encounterless stars is provided by Liouville's equation

$$\frac{\partial \psi}{\partial t} + [\psi, H] = 0 , \tag{1}$$

where the second term on the left represents the Poisson bracket expressed in suitable coordinates. For a disk of infinitesimal thickness, we choose polar coordinates (ϖ, θ) with conjugate momenta per unit mass (p_ϖ, p_θ). In these coordinates, the Hamiltonian (per unit mass) is given by

$$H = \tfrac{1}{2}(p_\varpi^2 + p_\theta^2/\varpi^2) + \mathfrak{V}(\varpi, \theta, t) , \tag{2}$$

where $\mathfrak{V}(\varpi, \theta, t)$ is the value of the total gravitational potential in the plane of the disk. The Poisson bracket now has the appearance

$$[\psi, H] = p_\varpi \frac{\partial \psi}{\partial \varpi} + \frac{p_\theta}{\varpi^2}\frac{\partial \psi}{\partial \theta} + \left(\frac{p_\theta^2}{\varpi^3} - \frac{\partial \mathfrak{V}}{\partial \varpi}\right)\frac{\partial \psi}{\partial p_\varpi} - \frac{\partial \mathfrak{V}}{\partial \theta}\frac{\partial \psi}{\partial p_\theta} . \tag{3}$$

Suppose that the gravitational potential can be regarded as a superposition of a small-amplitude perturbation onto a basic time-independent value possessing axial symmetry:

$$\mathfrak{V} = \mathfrak{V}_0(\varpi) + \mathfrak{V}_1(\varpi, \theta, t) , \tag{4a}$$

$$H = H_0(\varpi, p_\varpi, p_\theta) + \mathfrak{V}_1(\varpi, \theta, t) , \quad H_0 = \tfrac{1}{2}(p_\varpi^2 + p_\theta^2/\varpi^2) + \mathfrak{V}_0(\varpi) . \tag{4b}$$

The phase density ψ_0 corresponding to the basic state is required to satisfy the condition

$$\frac{\partial \psi_0}{\partial t} + [\psi_0, H_0] = 0 . \tag{5}$$

The theory of characteristics provides the most general form for ψ_0 as an arbitrary function of the four constants of motion (cf. Vandervoort 1967):

$$E_0 = \tfrac{1}{2}(p_\varpi^2 + p_\theta^2/\varpi^2) + \mathfrak{V}_0(\varpi) , \tag{6a}$$

$$J = p_\theta , \tag{6b}$$

$$T = t - \int^{\varpi} \{2[E_0 - \mathfrak{V}_0(\varpi')] - J^2/\varpi'^2\}^{-1/2} d\varpi' , \tag{6c}$$

$$P = \theta - \int^{\varpi} \frac{J}{\varpi'^2} \{2[E_0 - \mathfrak{V}_0(\varpi')] - J^2/\varpi'^2\}^{-1/2} d\varpi' . \tag{6d}$$

In general, the phase integrals T and P are nonisolating, and phase mixing (Lynden-Bell 1962) will tend to establish a steady-state coarse-grained distribution which can be regarded as a function solely of the isolating integrals E_0 and J. Hence, our attention can be limited to basic states whose phase distributions have the form

$$\psi_0 = \begin{cases} F_0(E_0, J) & \text{for } E_0 < 0 \text{ and } J \geq 0 , \\ 0 & \text{for } E_0 \geq 0 \text{ or } J < 0 . \end{cases} \tag{7}$$

The condition that ψ_0 vanishes for positive values of the energy E_0 assures the presence of only bound stars. The condition that ψ_0 vanishes for negative values of the angular momentum J restricts all stars to rotating in the same direction. The latter restriction is purely a mathematical device adopted to simulate the predominant sense of rotation in actual disk galaxies. In practice, little physical consequence is introduced by the inclusion of a few stars which are not bound or which rotate in the "wrong" direction.

From equation (1) is obtained, in the linearized approximation, the relation which governs the time development of the perturbations:

$$\frac{\partial \psi_1}{\partial t} + [\psi_1, H_0] = -[\psi_0, \mathfrak{B}_1] . \tag{8}$$

We choose to regard the right-hand side as given and to solve for ψ_1. The solution is formally obtained in Lagrangian coordinates by noting that the left-hand side of equation (8) represents the total time rate of change of ψ_1 when we follow the unperturbed orbit of a star whose coordinates in phase space at time t are $(\varpi, \theta, p_\varpi, p_\theta)$. Hence, the integration of equation (8) over time leads to

$$\psi_1(t) - \psi_1(t_0) = -\int_{t_0}^{t} [\psi_0, \mathfrak{B}_1]_0 \, dt' = \int_{t_0}^{t} \left(\frac{\partial \mathfrak{B}_1}{\partial \varpi} \frac{\partial \psi_0}{\partial p_\varpi} + \frac{\partial \mathfrak{B}_1}{\partial \theta} \frac{\partial \psi_0}{\partial p_\theta} \right)_0 dt' , \tag{9}$$

where the subscript zero on the brackets has been used to indicate that the integration is to be performed along the unperturbed orbit. In equation (9) we have written $\psi_1(t)$ to represent $\psi_1(\varpi(t), \theta(t), p_\varpi(t), p_\theta(t), t)$.

It is convenient to transform to the characteristic variables E_0, J. In the transformation $(\varpi, p_\varpi, p_\theta) \rightarrow (\varpi, E_0, J)$, the partial derivatives transform as follows:

$$\frac{\partial}{\partial \varpi} = \frac{\partial}{\partial \varpi} + \left(-p_\theta^2/\varpi^3 + \frac{\partial \mathfrak{B}_0}{\partial \varpi} \right) \frac{\partial}{\partial E_0} , \tag{10a}$$

$$\frac{\partial}{\partial p_\varpi} = \Pi_0 \frac{\partial}{\partial E_0} , \tag{10b}$$

$$\frac{\partial}{\partial p_\theta} = \frac{\partial}{\partial J} + \frac{J}{\varpi^2} \frac{\partial}{\partial E_0} . \tag{10c}$$

In equation (10b), Π_0 denotes the following function of (ϖ, E_0, J):

$$p_\varpi = \Pi_0(\varpi, E_0, J) = \{2[E_0 - \mathfrak{B}_0(\varpi)] - J^2/\varpi^2\}^{1/2} . \tag{11}$$

In terms of these variables and when ψ_0 is given by equation (7), equation (9) takes the form

$$\psi_1(t) - \psi_1(t_0) = \int_{t_0}^{t} \left\{ \left(\Pi_0 \frac{\partial \mathfrak{B}_1}{\partial \varpi} + \frac{J}{\varpi^2} \frac{\partial \mathfrak{B}_1}{\partial \theta} \right) \frac{\partial F_0}{\partial E_0} + \frac{\partial \mathfrak{B}_1}{\partial \theta} \frac{\partial F_0}{\partial J} \right\}_0 dt' . \tag{12}$$

The first term on the right-hand side may be written in a more useful form by noting that along unperturbed orbits

$$\frac{d\varpi}{dt} = \Pi_0 , \quad \frac{d\theta}{dt} = \frac{J}{\varpi^2} , \tag{13}$$

and hence

$$\Pi_0 \frac{\partial \mathfrak{B}_1}{\partial \varpi} + \frac{J}{\varpi^2} \frac{\partial \mathfrak{B}_1}{\partial \theta} = \frac{d\mathfrak{B}_1}{dt} - \frac{\partial \mathfrak{B}_1}{\partial t} . \tag{14}$$

A partial integration of equation (12) may now be effected to give

$$\psi_1(t) = \frac{\partial F_0}{\partial E_0}\mathfrak{B}_1(t) + \psi_1(t_0) - \frac{\partial F_0}{\partial E_0}\mathfrak{B}_1(t_0) + \int_{t_0}^{t}\left\{-\frac{\partial \mathfrak{B}_1}{\partial t}\frac{\partial F_0}{\partial E_0} + \frac{\partial \mathfrak{B}_1}{\partial \theta}\frac{\partial F_0}{\partial J}\right\}_0 dt' . \quad (15)$$

Again, we have used the shorthand notation to write $\mathfrak{B}_1(t)$ for $\mathfrak{B}_1(\varpi(t), \theta(t), t)$ and $F_1(t)$ for $F_1(\varpi(t), \theta(t), E_0, J, t)$.

We look for periodic solutions of the form

$$\mathfrak{B}_1(\varpi, \theta, t) = V(\varpi)e^{i(\omega t - m\theta)} , \quad (16a)$$

$$\psi_1(\varpi, \theta, E_0, J, t) = \phi(\varpi, E_0, J)e^{i(\omega t - m\theta)} . \quad (16b)$$

Substituting equations (16) into equation (15), we obtain the requirement

$$\phi(\varpi(t), E_0, J) = \frac{\partial F_0}{\partial E_0} V(\varpi(t)) + e^{-i[\omega t - m\theta(t)]}\left\{\psi_1(t_0) - \frac{\partial F_0}{\partial E_0}\mathfrak{B}_1(t_0) - i\left(\omega\frac{\partial F_0}{\partial E_0} + m\frac{\partial F_0}{\partial J}\right)\int_{t_0}^{t} V(\varpi(t'))e^{i[\omega t' - m\theta(t')]}dt'\right\} . \quad (17)$$

Only single-valued solutions for ϕ are physically acceptable. The left-hand side of equation (17) is a function only of $(\varpi(t), E_0, J)$ and should not depend explicitly on $(\theta(t), t)$. For bound stars $\varpi(t)$ is bounded by the values of the turning points (ϖ_1, ϖ_2) and is a periodic function of t—with period $2\tau_{12}(E_0, J)$, say. Since ϕ depends only on $\varpi(t)$ and not on $\theta(t)$ or t, ϕ must also be a function of t with period $2\tau_{12}$.

Suppose that the azimuthal angle traversed in one radial period is $2\theta_{12}(E_0, J)$. We now replace t by $t + 2\tau_{12}$ in equation (17) and use the fact that

$$\theta(t + 2\tau_{12}) = \theta(t) + 2\theta_{12} \quad (18)$$

to obtain, after some manipulation,

$$\phi(t) = \phi(t + 2\tau_{12}) = \frac{\partial F_0}{\partial E_0} V(\varpi(t)) + e^{-2i(\omega\tau_{12} - m\theta_{12})}\left[\!\left[\phi(t) - \frac{\partial F_0}{\partial E_0} V(\varpi(t)) - i\left(\omega\frac{\partial F_0}{\partial E_0} + m\frac{\partial F_0}{\partial J}\right)\int_{t}^{t+2\tau_{12}} V(\varpi(t')) \times \exp\{i\omega(t' - t) - im[\theta(t') - \theta(t)]\}dt'\right]\!\right] . \quad (19)$$

Regarding equation (19) as a trap for $\phi(t)$, we may solve to obtain

$$\phi(t) = \frac{\partial F_0}{\partial E_0} V(\varpi(t)) - \left(\omega\frac{\partial F_0}{\partial E_0} + m\frac{\partial F_0}{\partial J}\right)\frac{e^{-i(\omega\tau_{12} - m\theta_{12})}}{2\sin(\omega\tau_{12} - m\theta_{12})} \times \int_{t}^{t+2\tau_{12}} V(\varpi(t'))\exp\{i\omega(t' - t) - im[\theta(t') - \theta(t)]\}dt' . \quad (20)$$

We identify $\varpi(t)$ as ϖ and define

$$\tau = t' - t - \tau_{12} , \quad (21a)$$

$$\theta_*(\tau) = \theta(t') - \theta(t) - \theta_{12} , \quad (21b)$$

$$\varpi_*(\tau) = \varpi(t') , \quad (21c)$$

to recover from equation (20) the solution for ϕ in the Eulerian description:

$$\phi(\varpi, E_0, J) = \frac{\partial F_0}{\partial E_0} V(\varpi) + f(\varpi, E_0, J)\,, \tag{22a}$$

$$f(\varpi, E_0, J) = -\frac{\omega \partial F_0/\partial E_0 + m\partial F_0/\partial J}{2\sin(\omega\tau_{12} - m\theta_{12})} \int_{-\tau_{12}}^{\tau_{12}} V(\varpi_*(\tau)) e^{i[\omega\tau - m\theta_*(\tau)]} d\tau\,. \tag{22b}$$

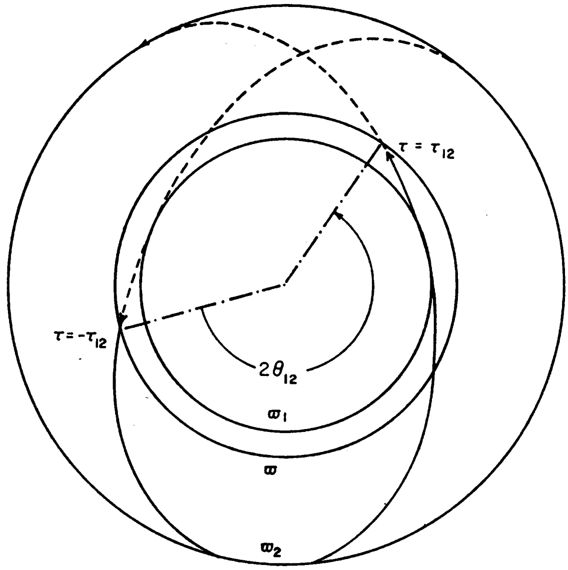

FIG. 1.—Typical stellar orbit. Solid part of the orbit represents one complete period of radial oscillation, the star having started from radial distance ϖ and having returned after having reached the maximum radial distance ϖ_2 and the minimum ϖ_1.

In the above, the functions $\varpi_*(\tau)$ and $\theta_*(\tau)$ are functions of (ϖ, E_0, J, τ) uniquely determined by the following equations of motion with associated periodicity conditions:

$$\frac{d\varpi_*}{d\tau} = \Pi_0(\varpi_*, E_0, J) \quad \text{with} \quad \varpi_* = \varpi \quad \text{at} \quad \tau = \pm\,\tau_{12}(E_0, J)\,, \tag{23a}$$

$$\frac{d\theta_*}{d\tau} = J/\varpi_*^2(\tau) \quad \text{with} \quad \theta_* = \pm\,\theta_{12}(E_0, J) \quad \text{at} \quad \tau = \pm\,\tau_{12}(E_0, J)\,. \tag{23b}$$

Figure 1 illustrates schematically a typical orbit described by the above set of equations.

When ω is real in equation (22b), the possibility for resonance arises for the response of the stars whose energies and angular momenta are such that $\sin(\omega\tau_{12} - m\theta_{12})$ vanishes. Resonance results because in each cycle of their (unperturbed) peculiar motion the stars feel the same phase of the perturbation gravitational field. The simplest form of resonance arises because of corotation, $\omega\tau_{12} - m\theta_{12} = 0$. More important, perhaps, for the theory of spiral structure are the resonances, $\omega\tau_{12} - m\theta_{12} = \pm\pi$ (cf. Paper II). For small peculiar velocities, resonant stars of the latter type follow the "dispersion orbits" described by Lindblad (1959); for this reason, we refer to such resonances as Lindblad resonances.

III. FORMULATION OF THE COMPLETE PROBLEM

We shall now formulate the problem of the normal modes of oscillation which occur in the plane of a differentially rotating self-gravitating sheet. In what follows, we assume for simplicity that the galaxy is composed of stars of identical mass m_*.

a) The Stellar Density Response

To deduce the stellar surface-density response from our solution for the phase density, we need the Jacobian of the transformation $(p_\varpi, p_\theta) \to (E_0, J)$:

$$\left|\frac{\partial(p_\varpi, p_\theta)}{\partial(E_0, J)}\right| = \left|\frac{\partial p_\varpi}{\partial E_0}\right| = 1/|\Pi_0(\varpi, E_0, J)| . \tag{24}$$

Since the stellar surface-density response σ_{*1} is obtained by integration of $m_* \psi_1$ over all allowable velocities $(p_\varpi, p_\theta/\varpi)$, it has the form

$$\sigma_{*1}(\varpi, \theta, t) = S_*(\varpi) e^{i(\omega t - m\theta)} , \tag{25}$$

where $S_*(\varpi)$ is given through

$$\varpi S_*(\varpi) = m_* \int\int \left\{ 2 \frac{\partial F_0}{\partial E_0} V(\varpi) + [f]_{\Pi_0(\varpi, E_0, J)>0} + [f]_{\Pi_0(\varpi, E_0, J)<0} \right\} \frac{dE_0 dJ}{|\Pi_0|} . \tag{26}$$

It is necessary to distinguish explicitly between $\Pi_0 > 0$ and $\Pi_0 < 0$ since the transformation from p_ϖ to E_0 is not one-to-one; for a given value of E_0, Π_0 can be either positive or negative. This distinction is not necessary for the term involving $\partial F_0/\partial E_0$ since it is even in $p_\varpi = \Pi_0$.

As a result of the property of Newton's laws under time reversal, the solutions of equation (23), for given E_0 and J, have the properties

$$\left\{ \begin{matrix} \varpi_*(\tau) \\ \theta_*(\tau) \end{matrix} \right\}_{\Pi_0(\varpi, E_0, J)>0} = \left\{ \begin{matrix} \varpi_*(-\tau) \\ -\theta_*(-\tau) \end{matrix} \right\}_{\Pi_0(\varpi, E_0, J)<0} . \tag{27}$$

With f given by equation (22b), equation (26) can now be transformed into the form

$$\varpi S_*(\varpi) = 2m_* \int\int_{\Pi_0>0} \left\{ \frac{\partial F_0}{\partial E_0}(E_0, J) V(\varpi) + f_*(\varpi, E_0, J) \right\} \frac{dE_0 dJ}{\Pi_0(\varpi, E_0, J)} , \tag{28}$$

where f_* is called the "essential part" of the perturbation amplitude of the distribution function and is given by

$$f_*(\varpi, E_0, J) = - \frac{\omega \partial F_0/\partial E_0 + m \partial F_0/\partial J}{2 \sin(\omega \tau_{12} - m\theta_{12})} \int_{-\tau_{12}}^{\tau_{12}} V(\varpi_*(\tau)) \cos[\omega\tau - m\theta_*(\tau)] d\tau . \tag{29}$$

For real ω, difficulty arises in the interpretation of the (E_0, J) integrations in equation (28) whenever there are resonant stars whose orbits are such that $(\omega\tau_{12} - m\theta_{12})$ is an integral multiple of π. In this case, the poles of $[\sin(\omega\tau_{12} - m\theta_{12})]^{-1}$ lie along the path of integration, and some ambiguity arises whether only the principal value of the integral is to be evaluated or whether "imaginary contributions" from the poles need to be included as well.[1] For our present purpose—the demonstration of an "anti-spiral theorem" when resonant stars are absent—it is sufficient to assume that such ambiguities in interpretation do not arise.

[1] See Stix (1962), pp. 118–124, 131–148, for a complete discussion of the analogous problem in the case of electrostatic plasma oscillations, in particular, for Landau's resolution of the difficulty by reference to the initial-value problem.

b) Integral Equation for Self-sustained Oscillations

The gravitational potential $\mathfrak{B}_1 = V(\varpi)e^{i(\omega t - m\theta)}$ associated with an imposed surface density $\sigma_1 = S(\varpi)e^{i(\omega t - m\theta)}$ is obtained from potential theory as

$$V(\varpi) = -G\int_0^\infty S(a)ada \oint \frac{e^{im\phi}d\varphi}{\sqrt{(\varpi^2 + a^2 - 2\varpi a \cos \varphi)}} \tag{30a}$$

$$= -2\pi G\int_0^\infty H_m(\varpi, a)aS(a)da ,$$

$$H_m(\varpi, a) = H_m(a, \varpi) = \frac{1}{\varpi + a} h_m(\zeta) , \tag{30b}$$

$$h_m(\zeta) = \frac{2}{\pi}\int_0^{\pi/2} \frac{\cos 2m\chi}{\sqrt{(1 - \zeta \cos^2 \chi)}} d\chi , \quad \zeta = \frac{4\varpi a}{(\varpi + a)^2} \leq 1 . \tag{30c}$$

The asymptotic forms taken by $h_m(\zeta)$ have the appearance

$$h_m(\zeta) = \begin{cases} \dfrac{(2m)!}{2^{4m}(m!)^2}\zeta^m & \text{for} \quad \zeta \ll 1 , \quad \text{i.e.,} \quad \text{for} \quad a \ll \varpi \quad \text{or} \quad a \gg \varpi ; \\ -\frac{1}{2}\log(1 - \zeta^2) & \text{for} \quad \zeta \simeq 1 , \quad \text{i.e.,} \quad \text{for} \quad a \simeq \varpi . \end{cases} \tag{31}$$

Thus, the existence of the integral on the right-hand side of equation (30a) presupposes regular boundary conditions imposed on $S(\varpi)$ at $\varpi = 0$ and $\varpi = \infty$.

For self-sustained oscillations, the imposed surface density $S(\varpi)$ must be equated with the surface-density response $S_*(\varpi)$. This condition, combined with equations (28) and (30), leads to the following integral equation to be satisfied by $S(\varpi)$:

$$\varpi S(\varpi) = \int_0^\infty K_{m\omega}(\varpi, a)aS(a)da , \tag{32}$$

where the kernel $K_{m\omega}(\varpi, a)$ is given by successive integrations of $H_m(\varpi, a)$:

$$\begin{aligned} K_{m\omega}(\varpi, a) = & -4\pi m_* G \iint_{\Pi_0 > 0} \frac{dE_0 dJ}{\Pi_0(\varpi, E_0, J)} \\ & \times \left\{ \frac{\partial F_0}{\partial E_0} H_m(\varpi, a) - \frac{\omega \partial F_0/\partial E_0 + m\partial F_0/\partial J}{2\sin(\omega\tau_{12} - m\theta_{12})} \int_{-\tau_{12}}^{\tau_{12}} H_m(\varpi_*(\tau), a) \right. \\ & \left. \times \cos[\omega\tau - m\theta_*(\tau)]d\tau \right\} . \end{aligned} \tag{33}$$

Equation (32) is a linear homogeneous (singular) integral equation to be solved for the characteristic value ω and the eigenfunction $\varpi S(\varpi)$. Together with the kernel defined by equation (33), it constitutes the complete formulation of the problem of normal modes in the whole disk of stars with velocity dispersion.[2] The kernel in equation (33) has certain mathematical properties which are important for the physical problem.

[2] The final formulation in terms of the integral equation (32) is similar to that given by Kalnajs (1965). The major difference is that we have not adopted at the outset the epicyclic approximation for stellar orbits.

c) *The Possibility for Modes of Spiral Form*

Equation (31) implies that $K_{m\omega}(\varpi, a)$ is logarithmically singular at $a = \varpi$. While integrable, this singularity accentuates the gravitational influence of stars in the local neighborhood and admits the possibility of short-scale wavy disturbances which are self-sustained *locally*. Second, for real ω, the argument against waves of a spiral form with a preferred sense of twist requires the absence of resonant stars.

Suppose that ω is real and that there are no resonant stars; then $K_{m\omega}(\varpi, a)$ as given by equation (33) is purely real. Under these conditions, an "anti-spiral theorem" similar to that given by Lynden-Bell and Ostriker (1967) applies. By taking the complex conjugate of equation (32), we easily show that, if $S(\varpi)$ is a solution, then so is its complex conjugate $S^*(\varpi)$. If ω is not degenerate, $S^*(\varpi)$ can differ from $S(\varpi)$ only by a constant complex multiplicative factor of modulus unity, i.e.,

$$S^*(\varpi) = e^{-2i\chi}S(\varpi) \,,$$

where χ is some real constant. Equating the arguments of both sides results in the relation

$$\arg \{S(\varpi)\} = \chi \,.$$

Since the phase of $S(\varpi)$ is a strict constant, this normal mode shows no spiral structure; it is of the "wagon wheel" type (Prendergast 1967).

If ω is degenerate and $S(\varpi)$ and $S^*(\varpi)$ are linearly independent, they must correspond to spiral patterns of opposite twist; one leads wherever the other trails. Further, we may adopt

$$\tfrac{1}{2}[S(\varpi) + S^*(\varpi)] \quad \text{and} \quad \frac{1}{2i}[S(\varpi) - S^*(\varpi)]$$

rather than $S(\varpi)$ and $S^*(\varpi)$ themselves as the linearly independent solutions. Both these solutions are purely real and consequently do not exhibit spiral structure.[3]

If there are resonant stars, the simple form of the normal mode analysis given here breaks down, and the arguments leading to the "anti-spiral theorem" do not necessarily apply. It is not surprising that this is so. Basically, the "anti-spiral theorem" is a reflection of the time reversibility of the equations of motion. For neutrally stable oscillations, there is no "arrow of time" and no preferred sense of twist for spirals (when referred to the motion of the matter). That is, reversing time reverses the velocities of all stars but not the density pattern they form. When resonance exists, it becomes necessary to specify that the secular effects on the resonant stars began in the past; this specifies an "arrow of time." Similarly, as has been demonstrated explicitly for a "cold disk" by Hunter (1969), overstability may also serve to distinguish leading and trailing spirals.

IV. DISCUSSION

In summary, an "anti-spiral theorem," analogous to that given by Lynden-Bell and Ostriker (1967) for a gaseous disk, applies in the linearized theory to all neutral modes for which resonance does not occur and for which regular boundary conditions are applicable. Meanwhile, the study of the radial propagation of spiral waves made by Toomre (1969) has made especially urgent the need to understand even the long-term persistence of an existing spiral structure.

While such results pose obvious difficulties for the density-wave theory, they also

[3] When the eigenvalue spectrum is continuous, a superposition of two nonspiral modes whose eigenfrequencies are infinitesimally close can lead to a pattern which is spiral. Of course, individual modes lose much of their significance in this case, since an arbitrary disturbance will not generally excite a single mode.

suggest the permissibility of spirals to depend on the prevalence of one or a combination of the following conditions: (*a*) the existence of strong stellar resonances, (*b*) the effects of finite amplitude, (*c*) the existence of overstabilities, (*d*) the driving of the disk by other agencies.

It is difficult to accept the hypothesis that resonances can *by themselves* generate spiral patterns. Indeed, in the linear theory, the analogy with plasma physics would lead one to suspect that, except for special circumstances, resonances act to absorb density waves (by mechanisms similar to Landau damping). Of course, the conversion of wave energy into energy of particle motions cannot proceed indefinitely, and nonlinear effects—to which resonant stars are most susceptible—may eventually alter these conclusions.

Any expectation for waves to develop spontaneously is related to the possibility for instability. It has not been possible to provide a general statement regarding the stability of planar disturbances. However, an asymptotically valid result can be derived, in the WKBJ approximation, for a stellar disk whose distribution function for stars with low peculiar velocities is given locally by Schwarzschild's distribution. When stellar resonances are again ignored, this investigation (Paper II) reveals no short-scale instabilities in addition to the Jeans instability discussed for this geometry by Toomre (1964). The effects of "gradients," reported by Lin and Shu (1966) as a mechanism leading to overstability, provide growths of wave energy only for an observer moving with a wave group (Toomre 1969).

These results lead us to consider seriously that spiral structure may result from forcing by a yet unspecified agent. To account for the well-known correlations between spiral-arm spacing and the basic mass distribution, it is likely that this mechanism is internal to the galaxy. We leave for future investigations the further discussions of these points.

This work forms an extension of results obtained in the doctoral dissertation which I prepared at the Harvard College Observatory. I am grateful to Professor Max Krook and Professor C. C. Lin for their supervision of my thesis. I would also like to express my thanks to Professor G. Contopoulos, Professor Alar Toomre, and Dr. A. J. Kalnajs for various helpful discussions. This investigation was supported in part by the National Science Foundation.

REFERENCES

Hunter, C. 1969, *Ap. J.*, **157**, 183.
Kalnajs, A. J. 1965, unpublished Ph.D. thesis, Harvard University.
Lin, C. C. 1966, *SIAM, J. Appl. Math.*, **14**, 876.
Lin, C. C., and Shu, F. H. 1964, *Ap. J.*, **140**, 646.
———. 1966, *Proc. Nat. Acad. Sci.*, **55**, 229.
———. 1967, in *Proc. I.A.U. Symp. No. 31, Noordwijk, 1966*, p. 313.
Lin, C. C., Yuan, C., and Shu, F. H. 1969, *Ap. J.*, **155**, 721.
Lindblad, B. 1959, *Hdb. d. Phys.*, **53**, 21.
———. 1963, *Stockholm Obs. Ann.*, **22**, 3.
Lynden-Bell, D. 1962, *M.N.R.A.S.*, **124**, 279.
Lynden-Bell, D., and Ostriker, J. D. 1967, *M.N.R.A.S.*, **136**, 273.
Prendergast, K. H. 1967, in *Proc. I.A.U. Symp. No. 31, Noordwijk, 1966*, p. 303.
Shu, F. H. 1968, unpublished Ph.D. thesis, Harvard University.
———. 1970, *Ap. J.* **160**, 99 (Paper II).
Stix, T. H. 1962, *The Theory of Plasma Waves* (New York: McGraw-Hill Book Co.).
Toomre, A. 1964, *Ap. J.*, **139**, 1217.
———. 1969, *ibid.*, **158**, 899.
Vandervoort, P. O. 1967, *Ap. J.*, **147**, 384.