The Astrophysical Journal, Vol. 155, March 1969

ON THE SPIRAL STRUCTURE OF DISK GALAXIES
III. COMPARISON WITH OBSERVATIONS*

C. C. Lin and C. Yuan
Massachusetts Institute of Technology
and
Frank H. Shu
Harvard College Observatory

Received June 21, 1968; revised August 16, 1968

ABSTRACT

The density-wave theory of galactic spirals is developed in a form slightly more general than that outlined by Lin and Shu in an earlier short communication. Only self-sustained waves are studied in this paper, and the problem of the origin of the spiral structure is barely discussed. The implications of the theory are examined, both in general terms and in detail. The conclusions are compared with observations. Specifically, we consider (1) the distribution of atomic hydrogen in the Galaxy, (2) the systematic motion of the gas, (3) the distribution of young stars and other optical objects, and (4) the migration of moderately young stars. Good agreements are obtained in all cases if we adopt a pattern speed of about 13 km $\sec^{-1}$ kpc^{-1}, and a spiral gravitational field equal to about 5 per cent of the symmetrical field. General discussions are also given on (*a*) the structure of the magnetic field and its role on the systematic motion of the gas, (*b*) the role of the density wave in the process of star formation, and (*c*) the distribution of H II regions as revealed by the 109α radio observations.

I. INTRODUCTION

The density-wave theory of galactic spirals was proposed by the late B. Lindblad a long time ago. An evaluation of the early phase of the theory was made by Chandrasekhar (1942). Many subsequent papers on this subject published by Lindblad and his co-workers may be traced through his last paper in the *Stockholm Observatory Annalen* (Lindblad 1963). Feelings of reservation were expressed by astronomers and astrophysicists about Lindblad's ideas, largely because the method adopted by him, which leans heavily on the properties of individual stellar orbits, cannot yield quantitative predictions for comparison with observations. Without making use of the statistical theory of stellar dynamics, he could hardly treat the behavior of stellar collective modes in a quantitative manner. Yet these modes indeed constitute the essence of the density waves. His discussions were therefore largely qualitative and did not convey sufficient conviction to the readers.

P. O. Lindblad (1960, 1962) attempted to study these collective modes by an extensive calculation of the orbits of a number of stars with the help of large computing machines. His model consists of a smooth background representing Population II stars, together with 192 stars representing Population I. Density waves of a spiral form were indeed found from these calculations, but they were rather transient and were not quasi-stationary, as suggested by B. Lindblad. In view of the limitations inherent in direct computational methods, the results obtained are perhaps to be described as semiquantitative, when comparison with observations is attempted.[1]

The explicit adoption of stellar collective modes as a basis for a density-wave theory

* A condensed version of this paper was reported at a meeting of Commission 33 at the General Assembly of the International Astronomical Union in Prague, August 22–31, 1967 (not to be published).

[1] More recently, Miller and Prendergast (1968) carried out extensive calculations for a disk model consisting of 120000 stars Others have dealt with cases involving cylindrical symmetry.

of galactic spirals was first published[2] by two of the present writers (Lin and Shu 1964, 1966, hereinafter referred to as Papers I and II, respectively). A description of the history and the evolution of ideas may be found in a survey article by Lin (1967*a*) on the dynamics of disk-shaped galaxies. In Paper I, an elementary theory was presented in which the dispersion of stellar velocities and the pressure (or turbulent motion) of the gas were neglected. Only speculative discussions were given concerning their effects. These speculations were confirmed by subsequent investigations, and the results were summarized in Paper II, which also contains an outline of the methods used for dealing with a galactic system composed of stars with velocity dispersion and of gas with pressure simulating the effect of turbulent motion. In the present paper, we shall also include brief discussions (in §§ I, V, and VIII) of the magnetic field and its influence on the gaseous component. The details of these investigations will be presented elsewhere (Yuan 1969).

As emphasized before (Oort 1962; Lin and Shu 1967; Lin 1968), the primary purpose of our theory is to attempt to explain the *grand design* over the whole disk, and in particular to explain its persistence. As one can see upon a little reflection, the problem of the origin of the spiral structure is mathematically more difficult, since we have to deal with a combined boundary-value and initial-value problem. These studies remain a challenge for future investigations.[3] In contrast, to explain persistence, one need only produce solutions representing self-sustained waves. The method outlined in Paper II is especially formulated for this purpose. Even here, there are some subtle points on the basic mechanism which should be clarified by future investigations. We shall bring out these points in the subsequent discussions.

a) Some General Implications of the Density-Wave Theory

The winding dilemma of material spiral arms implies also a difficulty in the associated magnetic field. As the winding process goes on, a material arm is stretched, and the magnetic field, being frozen into the gas, also continues to increase in its strength. These difficulties are all avoided if we associate the spiral pattern with a *density wave* that has a *quasi-stationary spiral structure.*

Consider the ideal case of a single wave. To an observer rotating with the pattern speed, the gas flow is steady and in *closed* stream lines[4] (which deviate only slightly from circular shape). Thus, one visualizes a situation in which the magnetic field always runs *approximately along* the spiral arm in the over-all pattern. There may also exist material arms which are being distorted by differential rotation. In such an arm, there might be even a fair amount of deviation between the field lines and the spiral arm, but it would still be a good first approximation to think of the field as primarily oriented along an arm. Such general conclusions appear to be in agreement with observations.

b) Objectives of the Present Paper

In the absence of a complete theory for the mechanism of density waves, the need for observational support is urgent. Accordingly, the main purpose of the present paper is to describe the applications of the theory to several observable features in our own Galaxy, in order to enable the reader to form an over-all picture of the extent to which the theory may yield results that agree with observations.[5] Some further details of the

[2] Kalnajs (1965) described a more ambitious scheme in attacking the same problem along similar lines. He envisaged the importance of collective modes of a very large scale. These are mathematically more difficult to treat, and his results have not yet been published. See also n. 3.

[3] See Shu (1968) for a fuller discussion of the complete mathematical formulation of the problem and of the possibility of tracing the origin of spiral structure to "gradient instability," as suggested in Paper II.

[4] Such flow patterns, including shock formation, have been recently calculated by Roberts (1969).

[5] An article by Lynds (1967*a*, *b*) gave a general account of these comparisons. The article also contains discussions of other data not discussed in detail in this paper.

analyses used and the comparisons made in each case will be published separately in later communications. Specifically, we shall consider (1) the distribution of atomic hydrogen, (2) the systematic motion of the gas, (3) the distribution of young stars and their place of formation, (4) the migration of moderately young stars and their place of formation, and, finally, (5) the role of the magnetic field in item (2), its alignment with the spiral arm having been mentioned above. We shall also examine the problems related to the local structure in the solar vicinity, in the light of these studies. Some comments will be made toward the end of this paper on the general status of the theory and on some general implications, especially on the process of star formation and on the evolution of the galaxies.

A secondary purpose of this paper is to give an orderly presentation of the theory in a form somewhat more general than that presented before. The principal results will be described in § 2 to pave the way for the discussion of the comparison of the theory with observations. The mathematical details of these developments may be found in the Appendices. The exposition of the present paper will be organized essentially as a continuation of Papers I and II cited above; only occasional references will be made to the other papers (e.g., Lin and Shu 1967; Lin 1966, 1967*a*, *b;* the reader is also referred to Prendergast 1967).

It should be mentioned from the outset that the detailed calculations in this work are based on an asymptotic analysis of the WKBJ type, valid for reasonably tightly wound spirals (pitch angle less than 15°). For larger inclinations, the results can be only qualitatively correct; accurate results can only be obtained from the type of approach formulated by Kalnajs (1965) and Shu (1968).

Our future plans include further applications of our theory to the structure of the Milky Way, and the study of other galaxies. A comprehensive discussion of the existing observable features that bear on the spiral structure of galaxies, especially our own, was given by Bok (1967) and by Bok and Contopoulos (1967). The reader should refer to them for further details, and for future observations recommended.

II. THEORY OF DENSITY WAVES

The density-wave theory was developed especially to resolve the problems of (i) the persistence of the spiral structure in the face of differential rotation and (ii) the large-scale nature of the phenomenon, i.e., the existence of a grand design over the whole galactic disk. We therefore start out with the QSSS hypothesis (hypothesis of the existence of density waves with a quasi-stationary spiral structure), rather than attempt to study the problem of the origin of the spiral structure.[6] As outlined in Paper II, the theory is based on the calculation of the responses of an infinitesimally thin disk of stars and gas to a *resultant* gravitational field of a spiral form. The effect of finite thickness has been recently examined by Shu (1968), but we shall not go into the details in this paper.

a) *Response of the Stellar Disk*

The stellar disk is described by a distribution function $\Psi(\varpi,\theta,c_\varpi,c_\theta,t)$, where t is the time, (ϖ,θ) are polar coordinates in the plane of the disk, and (c_ϖ,c_θ) are the peculiar velocities relative to a (differentially) rotating observer moving at the circular speed $\varpi\Omega(\varpi)$. In the absence of spiral modes, Ψ is given by a function $\Psi_0(\varpi,c_\varpi,c_\theta)$ which is time-independent and axisymmetric in the configurational space.

We now make some brief observations on the purely kinematical aspects of a smooth but tightly wound spiral pattern in the plane of a disk galaxy. First of all, we notice that these patterns may be described by functions of the form

$$F(\varpi,\theta,t) = \mathfrak{A}(\varpi) \exp \{i[\omega t - m\theta + \Phi(\varpi)]\} , \tag{2.1}$$

[6] The reader is referred to Lin (1967*b*) for a discussion of the "kinematical tests" of the feasibility of the QSSS hypothesis as the basis for a theory.

where ϖ,θ,t are defined as before; $\mathfrak{A}(\varpi)$ is a slowly varying function of ϖ; $\Phi(\varpi)$ is a slowly varying (monotonic) function multiplied by a large parameter; ω is a constant (real for neutral waves); and m is an integer. As usual, the real part of the above expression is to be taken. The lines of constant $\mathrm{Re}(F)$ are approximately given by the equation

$$m(\theta - \theta_0) = \Phi(\varpi) - \Phi(\varpi_0) , \tag{2.2}$$

which represents a spiral pattern with m arms.

The pattern given by equation (2.1) rotates with an angular velocity

$$\Omega_p = \omega_r/m , \tag{2.3}$$

where ω_r is the real part of ω. The radial wavenumber of the spiral pattern is given by $|k(\varpi)|$, where

$$k(\varpi) = \Phi'(\varpi) . \tag{2.4}$$

If the motion of the stars is in the direction of increasing θ, trailing waves correspond to $k(\varpi) < 0$, and leading waves, to $k(\varpi) > 0$.

In the presence of a small spiral gravitational potential,

$$\mathcal{V}_1 = A(\varpi) \exp \{i[\omega t - m\theta + \Phi(\varpi)]\} , \tag{2.5}$$

of the form of equation (2.1), the response $\Psi_1 = \Psi - \Psi_0$ in the distribution function should also have a spiral distribution in the configurational space. When the dispersion velocities of the stars are not large, this response Ψ_1 can be easily worked out. In this case, the basic distribution function Ψ_0 is of the form

$$\Psi_0 = \Psi_0(\xi^2 + \eta^2,\varpi) , \tag{2.6}$$

where ξ and η are defined by

$$c_\varpi = \xi(2\Omega/\kappa)(\varpi\Omega) = \xi V_1(\varpi) , \tag{2.7a}$$

and

$$c_\theta = \eta(\varpi\Omega) , \tag{2.7b}$$

in which κ is the epicyclic frequency defined by

$$\kappa^2 = (2\Omega)^2\left(1 + \frac{\varpi}{2\Omega}\frac{d\Omega}{d\varpi}\right) . \tag{2.8}$$

Within the approximation of small dispersion speeds is a parameter related to the integral for angular momentum, and (for given ϖ) $\xi^2 = \eta^2$ is the integral of epicyclic energy. The response Ψ_1 to the resultant potential (2.5) is found to be

$$\Psi_1 = \Psi_0'(\xi^2 + \eta^2)[1 - q(\alpha\xi,\alpha\eta,\nu)] \frac{\mathcal{V}_1}{V_1^2} , \tag{2.9}$$

in the case of *tightly wound spirals*, where Ψ_0' denotes the derivative of Ψ_0 with respect to its argument,[7] V_1 is defined in equation (2.7a), and the parameters α and ν are defined by

$$\alpha = (k\varpi)(2\Omega^2/\kappa^2) \tag{2.10}$$

and

$$\nu = (\omega - m\Omega)/\kappa . \tag{2.11}$$

[7] Here, as in other quantities, the dependence on the galactocentric distance ϖ is understood but not explicitly exhibited.

The function q is defined by the integral

$$q = \frac{\nu\pi}{\sin \nu\pi} \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp \{i[\nu s - a\xi \sin s + a\eta(1 + \cos s)]\} ds . \tag{2.12}$$

The derivation of the result (2.9) may be found in Appendix A. In all previous papers, we have presented results only for the case of the Schwarzschild distribution (for which Ψ_0 is an exponential function of $\xi^2 + \eta^2$). Clearly, this is an important case, and we shall continue to adopt it in following sections of this paper. But the generalization given here enables one to study, separately, the roles played by stars with different epicyclic energies; this may be done simply by taking Ψ_0 to be of the form of the Dirac δ-function. Such studies will be carried out in a separate investigation.

In the case of the Schwarzschild distribution, the perturbations in density and in the components of the mean velocity of the stars can be very readily calculated to exhibit their dependence on the dispersion velocity of the stars (see Appendix B). One obtains

$$\frac{\mathfrak{d}}{\sigma_*} = -\frac{k^2 A}{\kappa^2} \cdot \frac{1}{1 - \nu^2} \cdot \mathfrak{F}_\nu(x) , \tag{2.13a}$$

$$\frac{v_\varpi}{\varpi\Omega} = \frac{k^2 A}{\kappa\Omega} \cdot \frac{(k\varpi)^{-1}\nu}{1 - \nu^2} \cdot \mathfrak{F}_\nu^{(1)}(x) , \tag{2.13b}$$

$$\frac{v_\theta}{\varpi\Omega} = \frac{i}{2} \frac{k^2 A}{\Omega^2} \cdot \frac{(k\varpi)^{-1}}{1 - \nu^2} \cdot \mathfrak{F}_\nu^{(2)}(x) , \tag{2.13c}$$

where

$$x = k^2 \langle c_\varpi^2 \rangle / \kappa^2 \tag{2.14}$$

and the functions $\mathfrak{F}_\nu(x)$, $\mathfrak{F}_\nu^{(1)}(x)$, and $\mathfrak{F}_\nu^{(2)}(x)$ are "reduction factors" defined by equations (B9), (B14), and (B17) in Appendix B. They all take on the value unity for $x = 0$, i.e., in the case of zero dispersion. The function $\mathfrak{F}_\nu(x)$ is most important to our subsequent discussions. Its values have already been tabulated and presented in graphical form (Lin and Shu 1966; Lin 1967*b*). One form of the graph is reproduced in Figure 8 for convenience of reference.

We note that, since all the above results are given for tightly wound spirals, they are subject to an uncertainty of a factor of $1 + O(1/k\varpi)$.

b) *Response of the Gaseous Disk*

The response of the gaseous disk can be much more readily obtained if the turbulent motion of the gas is simulated by pressure. We then introduce the square of the equivalent acoustic velocity a, and define the quantity

$$x_g = k^2 a^2 / \kappa^2 , \tag{2.15}$$

analogous to x defined by equation (2.14). Then all the three *reduction factors* are found to be of the same form, i.e.,

$$\mathfrak{F}_\nu^{(g)}(x_g) = \frac{1}{1 + x_g/(1 - \nu^2)} . \tag{2.16}$$

For convenience of reference, we shall record the explicit formulae as follows:

$$\frac{\mathfrak{d}_g}{\sigma_0} = -\frac{k^2 A}{\kappa^2} \cdot \frac{1}{1 - \nu^2} \cdot \mathfrak{F}_\nu^{(g)}(x_g) , \tag{2.17a}$$

$$\frac{u_\varpi}{\varpi\Omega} = \frac{k^2 A}{\kappa\Omega} \cdot \frac{(k\varpi)^{-1}\nu}{1 - \nu^2} \mathfrak{F}_\nu^{(g)}(x_g) , \tag{2.17b}$$

$$\frac{u_\theta}{\varpi\Omega} = \frac{i}{2}\cdot\frac{k^2 A}{\Omega^2}\cdot\frac{(k\varpi)^{-1}}{1-\nu^2}\,\mathfrak{F}_\nu^{(g)}(x_g)\,, \tag{2.17c}$$

where u_ϖ and u_θ are the amplitude functions for the components of the gaseous velocity. Let us also note that, irrespective of pressure, the following ratios hold:

$$\frac{\delta_g}{\sigma_0}:\frac{u_\varpi}{\varpi\Omega}:\frac{u_\theta}{\varpi\Omega} = -k\varpi : m\left(\frac{\Omega_p}{\Omega}-1\right) : i\,\frac{\kappa^2}{2\Omega^2}\,, \tag{2.18}$$

where Ω_p is the pattern speed. This follows from the fact that the pressure gradient plays no role either in the equation of continuity or in the θ-component of the equation of motion (in the case of tightly wound spirals).

III. THE DISPERSION RELATION FOR DENSITY WAVES

The central relationship for any type of wave motion is that between wavelength and frequency. In the theory of self-sustained density waves, one must equate the density required to sustain the gravitational potential (eq. [2.1]) according to Poisson's equation, i.e.,

$$\delta_p = -|k|A/2\pi G\,, \tag{3.1}$$

with the *sum* of the responses in density variation in the stellar and gaseous disks. Thus, by using equations (2.13a) and (2.17a), we get the *dispersion relationship*,

$$\frac{k_*}{|k|}(1-\nu^2) = \mathfrak{F}_\nu(x) + \frac{\sigma_0}{\sigma_*}\mathfrak{F}_\nu^{(g)}(x_g)\,, \tag{3.2}$$

which connects $|k|$ and ν for a given galactic model. In the above equation, we define k_* by

$$k_* = \kappa^2/2\pi G\sigma_*\,, \tag{3.3}$$

and we must have

$$1-\nu^2 > 0\,. \tag{3.4}$$

The limiting cases $\nu = \pm 1$ correspond to Lindblad resonance $\Omega_p = \Omega \pm \kappa/m$; they deserve special treatment (see item *b* in § VIII).

The waves described by the dispersion relationship (eq. [3.2]) have the following properties:

a) They extend essentially over a range of the galactic disk for which the conditions

$$\Omega - \frac{\kappa}{m} < \Omega_p < \Omega + \frac{\kappa}{m} \tag{3.5}$$

are satisfied, where $\Omega(\varpi)$ is the angular speed of the stars, and $\kappa(\varpi)$ is the epicyclic frequency, as defined before. We shall refer to this range as the principal part of the spiral pattern. In Figure 1, we show the graphs of $\Omega(\varpi)$, $\kappa(\varpi)$, and $\Omega \pm \kappa/2$ on the basis of the model given by Schmidt (1965); the relevant numerical values may be found in Table 1. According to this diagram, for $m = 2$, the pattern would extend from $\varpi = 4$ kpc outward beyond 20 kpc, if $\Omega_p = 11$ km sec^{-1} kpc^{-1}. For other values of $m > 2$ this particular part would be quite limited in extent, whatever the value of Ω_p may be. This is the reason why *all spirals have two arms* (or only one arm), for the rotation curves are qualitatively the same. Note that the argument no longer holds for the outer part of a galaxy, leaving the possibility open for the existence of "multiple-armed" galaxies such as M101, which has still a well-defined two-armed structure in the interior parts. The possibility that one-armed galaxies exist cannot be ruled out by the present discussion, but it requires a more complicated study which will be presented later.

b) We notice that the dispersion relation for such waves is expressed in terms of the wavelength $\lambda = 2\pi/|k|$ and the frequency $\nu = m(\Omega - \Omega_p)/\kappa$ at which the stars see the gravitational field. This relationship includes two additional parameters, x and x_g, which measure the velocity dispersion of the stars and the turbulence in the gas. As a first attempt to appreciate the significance of this dispersion relationship, it is convenient to work out explicitly the results in the case of a pure stellar disk. In Figure 2, we show the result for the case where the velocity dispersion is barely enough to stabilize the disk against gravitational collapse (Toomre 1964). The margin of instability is obtained by setting $\nu^2 = 0$ in equation (3.2); conditions with $\nu^2 < 0$ correspond to instability, while

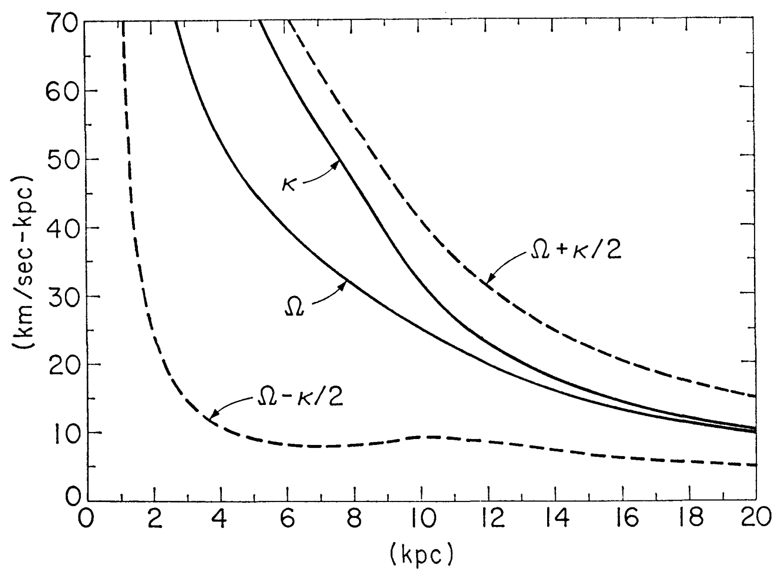


FIG. 1.—Rotation curve, etc., of our own Galaxy according to Schmidt model (symbols defined in Fig 2).

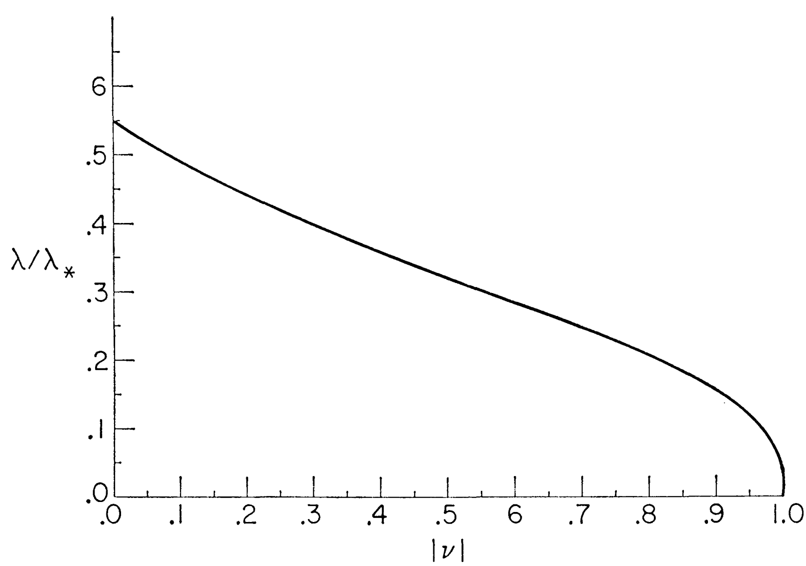


FIG. 2.—Dispersion relationship. Symbols are defined as follows: $\nu = m(\Omega - \Omega_p)/\kappa$, where $m = 2$ is the number of arms, $\Omega(\varpi)$ is the circular velocity in angular measure at galactocentric distance ϖ, $\kappa(\varpi)$ is the epicyclic frequency, λ is the spacing between two neighboring arms, and λ_* is a typical length scale defined by $\lambda_* = 4\pi^2 G\sigma_*/\kappa^2$, where σ_* is the projected surface density of the stars.

TABLE 1

1965 SCHMIDT MODEL OF THE GALAXY

A. ROTATION PARAMETERS

ϖ (kpc)	$\varpi\Omega$ (km sec^{-1})	Ω (km sec^{-1} kpc^{-1})	κ (km sec^{-1} kpc^{-1})	$\Omega-\kappa/2$ (km sec^{-1} kpc^{-1})	$\Omega+\kappa/2$ (km sec^{-1} kpc^{-1})
1	200	200 0	245 0	77 6	323 0
2	187	93 0	136 0	25 4	161 0
3	198	66 0	103 0	14 5	118 0
4	213	53 2	85 0	10 7	95 7
5	227	45 3	72 6	9 0	81 6
6	238	39 7	62 8	8 3	71 1
7	247	35 2	54 4	8 0	62 4
8	252	31 4	46 7	8 1	54 8
9	253	28 1	39 2	8 5	47 7
10	250	25 0	31 6	9 2	40 8
11	244	22 2	26 5	8 9	35 4
12	238	19 8	22 8	8 4	31 2
13	231	17 8	20 0	7 8	27 7
14	224	16 0	17.7	7 2	24 9
15	218	14 6	15 9	6 6	22 5
16	213	13 3	14 3	6 1	20 4
17	207	12 2	13 0	5 7	18 7
18	202	11 2	11 9	5 3	17 2
19	197	10 4	10 9	4 9	15 9
20	193	9 6	10 1	4 6	14 7

B. STRUCTURE PARAMETERS

ϖ (kpc)	$\sigma(\varpi)$ ($M_\odot$ pc^{-2})	$M(\varpi)$ ($10^9\ M_\odot$)	$\lambda_{tot}=4\pi^2(\sigma_*+\sigma_0)/\kappa^2$ (kpc)
1	1097 0	11	3 1
2	817 0	20	7 5
3	646 0	31	10 4
4	521 0	44	12 3
5	421 0	57	13 6
6	338 0	70	14 5
7	267 0	82	15 3
8	206 0	93	16 0
9	155 0	103	17 1
10	114 0	111	19 3
11	86 0	117	20 7
12	65 9	123	21 5
13	51 8	127	22 1
14	41 5	131	22 5
15	33 7	135	22 8
16	27 8	138	23 0
17	23 2	140	23 2
18	19 5	143	23 4
19	16 6	145	23 5
20	14 2	147	23 7

those with $\nu^2 > 0$ correspond to neutral waves. The condition $\nu^2 = 0$ leads to a minimum dispersion velocity $\langle c_\varpi^2 \rangle$ given by (see Toomre 1964)[8]

$$\frac{k_*^2 \langle c_\varpi^2 \rangle}{\kappa^2} = 0.2857 \simeq \tfrac{2}{7}, \tag{3.6}$$

below which there would be gravitational collapse. However, the corresponding wavelengths of the unstable waves, if $\langle c_\varpi^2 \rangle$ is slightly lowered from these critical values, are usually so large that their significance in the spiral theory is not clear. Thus, one might be inclined to think that even when $\langle c_\varpi^2 \rangle$ is not as large as that given by equation (3.6), the dispersion speeds may not increase any further.

As noted above, these calculations are based on an infinitesimally thin stellar sheet. The effect of a fair amount of gas and the modification due to finite thickness have been studied in detail by Shu (1968) and found to be not critical for the particular relationship in Figure 2. Some further discussion of these points will be given when our results are used in the following sections.

c) To the approximation discussed here, there are no indications that either trailing waves or leading waves are preferred, one to the other. However, such indications (though not conclusive) are found in the next approximation: Trailing patterns are preferred over leading patterns, if the velocity dispersion of stars increases toward the center. Other effects, especially non-linear terms representing convection of matter, would also tend to prefer trailing patterns. We are now in a position to compare our theoretical results with observed features in our own Galaxy.

IV. DISTRIBUTION OF ATOMIC HYDROGEN IN THE GALAXY

On the basis of the basic model shown in Figure 1, and the dispersion relationship shown in Figure 2, we can calculate a spiral pattern as soon as the pattern speed Ω_p is decided upon. A cursory examination of the numerical values obtained will show that Ω_p must be in the range of 11–13 km sec^{-1} kpc^{-1} to arrive at spacings between spiral arms that are reasonable compared with observations. One must also keep in mind that there is no reason why there should be only a single pattern associated with a single angular frequency Ω_p. A group of waves should in general be present. For simplicity, however, we shall try our best to compare the observational data with the theoretical results for a *single* pattern. It is therefore too much to expect very close agreement; yet we shall see that the results are surprisingly good.

In Figure 3, we reproduce the spiral pattern presented by Lin and Shu (1967), based on a pattern speed

$$\Omega_p = 11 \text{ km sec}^{-1} \text{ kpc}^{-1} .$$

The related basic data used are given in Table 1. When this pattern is compared with the observed distribution of atomic hydrogen, it appears to be in good agreement, although an even better agreement can probably be obtained by taking Ω_p slightly higher. As it is, the spiral pattern is perhaps a little too tight, but it does show the correct number of arms from the solar vicinity inward, up to the "3-kpc arm." The latter is tentatively associated with the strong motions that might be expected at inner Lindblad resonance, where

$$\Omega_p = \Omega - \kappa/2 .$$

According to the theory, the pattern ends here as a ring; there is no prominent pattern inside this ring. Indeed, no strong indications of a spiral structure have been obtained from observations. A detailed study of the conditions near Lindblad resonance requires

[8] Toomre considered exponentially damped or growing rings, not neutral spirals; but the mathematical theory is related to that in the present paper (see Lin and Shu 1966).

an extensive reformulation of the theory for dealing with a stellar sheet. This has been carried out by Shu (1968).

The Perseus arm and the Sagittarius arm can be fitted into an over-all spiral pattern with a skeleton shown by Figure 3, if each arm is assigned a width equal to about one-half the spacing. The distribution of the young optical objects in these arms will be discussed in § VI; no serious problems arise. The inner arm (the one inside the Sagittarius arm) also fits well into the picture. The Orion arm, however, presents somewhat of a problem. It is known to be inclined at about 30°, from observations of neutral hydrogen, optical objects, and magnetic field. If an arm with such an inclination were to be fitted into a

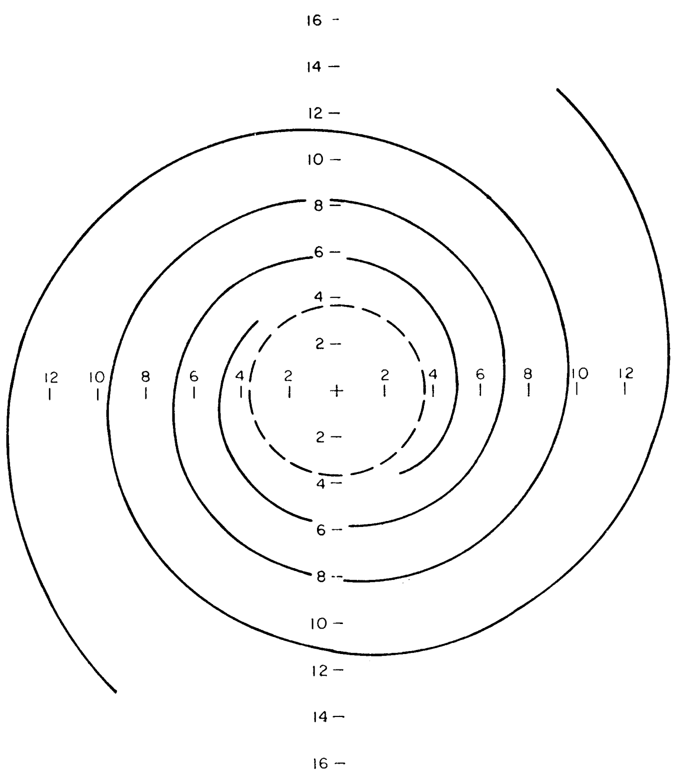


FIG. 3.—Spiral patterns at $\Omega_p = 11$ km sec^{-1} kpc^{-1}

two-armed pattern in our vicinity, the interarm spacing would be about 18 kpc, since the pitch angle i is related to this spacing λ by the simple relationship $\lambda = \pi\varpi_0 \tan i$, where ϖ_0 is the distance to the center of the Galaxy ($\varpi_0 = 10$ kpc in the present case). Such a spacing is clearly inadmissible, and we are therefore forced to conclude that the Orion arm (in which the Sun is located) is an interarm branch. Such branches, of course, occur quite frequently in the outer parts of disk-shaped spiral galaxies, and therefore we need not be unduly alarmed.

We shall leave the details of the choice of the pattern speed Ω_p and the initial point (ϖ_0,θ_0) in equation (2.2) to a separate discussion (Yuan 1969). Suffice it to mention here that the choices are limited to very narrow ranges.

Dispersion Velocities of Stars, Finite Thickness of the Galactic Disk

One subtle but rather crucial check of the general concepts is to examine the dispersion velocities of the stars predicted according to equation (3.6). In the above calculations, the disk is taken to be a pure stellar sheet of infinitesimal thickness. This leads to an rms value $\langle c_{\varpi}^2 \rangle^{1/2}$ of the radial dispersion velocity of about 52 km sec^{-1} for the solar vicinity, which is decidedly too high by at least 25 per cent. If the effect of the gas had been included, the discrepancy would have been larger. However, Shu (1968) has shown that all the discrepancies disappear when we take into consideration the effect of the finite thickness of the disks of stars and of gas. It is found that the pattern speed should also be increased slightly, to about 13.5 km sec^{-1} kpc^{-1}. (This value will be used in § VII.) Shu's calculations show that the rms value of the radial dispersion velocity should be about 37 km sec^{-1} in our vicinity (but presumably somewhat lower), if it is barely sufficient to sustain gravitational collapse. This value is in reasonable agreement with observations.

In the same study, Shu found the relative importance of gas and of stars to be about comparable. Here it is important to keep in mind that the gaseous disk is much thinner than the stellar disk. Consequently, the projected mass density of gas is not as large as indicated by the local volume density.

V. SYSTEMATIC MOTION[9] OF THE GAS

The possible existence of systematic motions was suspected quite early. Kerr (1962) suggested it as a possible source of difference between northern and southern observations. More recently, Burton (1966) considered the possible existence of a systematic motion of the gas near the outer edge of the Sagittarius arm, and concluded that "the observations are consistent with a situation in which the high-velocity stream is hydrogen flowing along the spiral arm, parallel to its axis." But he also suggested other possibilities. Shane and Bieger-Smith (1966) also discussed the role of systematic motions as part of their extensive effort in constructing alternative galactic models.

None of these authors, however, stressed the dynamical mechanism by which these systematical motions are produced. To be sure, whenever there is a circular (or nearly circular) arm of concentrated matter, its associated gravitational field would produce a systematic motion of the gas along its edges. It is easy to see this from considerations of angular momentum; indeed, the motion must be *with* the general rotation on the *outside* edge, and *against* the general rotation on the inside edge. But the density-wave theory goes deeper in the following two respects:

a) Since the pattern speed must in general differ from the speed of the material objects, which are in differential rotation, there *must* be systematic motion of the gas to maintain *conservation of matter.*

b) The existence of the gaseous arm clearly *indicates* the presence of a minimum in gravitational potential, but the gravitational potential is expected to be *jointly* maintained by the gas and by the stars with dispersion speeds comparable to the turbulent speeds in the gas. In many situations, *the stellar component would be dominant in supplying the gravitational field.* Thus, there is no need to be concerned about having sufficient contrast in gaseous density in order to *maintain* the field.

The systematic motion of the gas is estimated to have velocity components of the order of 10 km sec^{-1}. The radial component[10] cannot be easily singled out from observational data, but its magnitude is certainly compatible with the range of theoretically predicted values. The component in the circular direction would cause variations in the observed rotation curve. These variations have long been observed, but they were thought to be possibly the consequence of missing gas over interarm regions. A detailed study by Yuan (1969) has conclusively shown that the latter effect does not give significant contributions to the variation in velocity.

[9] By systematic motion, we mean the bulk motion of the gas *in addition to* the circular motion of the basic galactic model.

[10] For convenience of comparison with observational features, let it be noted that this component is directed toward the galactic center when the spiral pattern is trailing and traveling at a pattern speed lower than the local circular velocity. Cf. Miller (1968) for relevant data in the Perseus arm.

A comparison of theory with observation (Kerr 1964) is shown in Figures 4, 5, *a*, and 5, *b*. In Figure 4, we show the location of the gaseous arms as deduced from the location of the maxima and minima in the rotation curves shown in Figure 5. The figure is in very good agreement with the location of atomic hydrogen determined from direct observations. In Figure 5, we show the theoretical curve (dashed), assuming a certain "true rotation curve" on each side. There is good agreement in major features. The sudden dip in the northern curve at about 53° in galactic longitude l^{II} may be traced to the intersection of the Orion arm with the Sagittarius arm. The northern and southern rotation curves are still different, suggesting an oval distortion of the Galaxy as a whole. Such a distortion was suggested by Kerr. For a uniformly rotating disk, Hunter's (1963) theoretical work also indicates that it is an important mode of instability; more recently, Kalnajs (private communication) showed that this mode was particularly difficult to

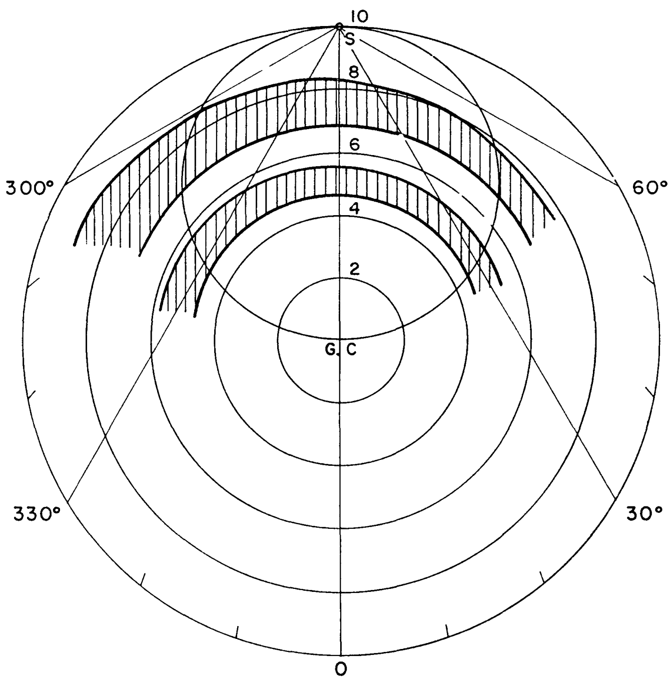


FIG. 4.—Locating gaseous arms from variations in the rotation curves

stabilize by stellar dispersion velocities. Work remains to be done to correlate the observational data with these general theoretical conclusions.

In the quantitative study of systematic motion, it is necessary to keep the role of the hydromagnetic forces in the picture. Yuan (1969) has discussed this point in great detail. The amplitudes of the velocity variation shown in Figures 4 and 5 are all about 8 km sec^{-1}. This could be caused by a spiral gravitational field equal to about 5 per cent of the axisymmetrical field (cf. § VII) acting on a gas with turbulent velocity equal to 7 km sec^{-1} (rms value of one component) and with a magnetic field of about 5×10^{-6} gauss. The amplitude of the total variation in mass is about 10 per cent of the mean, with gas and stars each contributing about half of the amount.

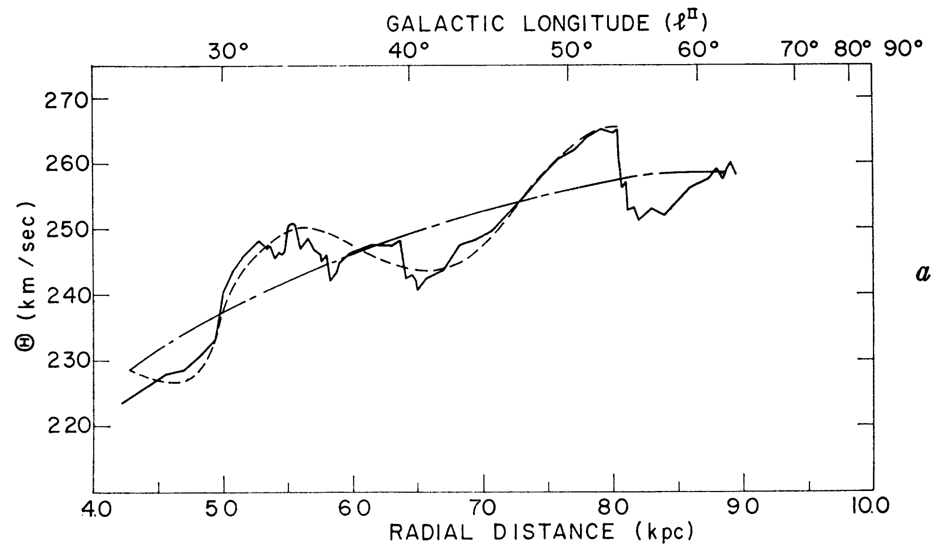
VI. THE DISTRIBUTION OF YOUNG STARS AND OTHER OPTICAL OBJECTS

It is well known that young O and B stars are indicators of the spiral arm. It is also generally agreed that they are formed at places of the highest gaseous concentration.[11]

[11] Recent work by Roberts (1968) shows that the process of star formation is intimately related to the formation of shocks in the gaseous flow, but the discussions in this section are not altered (see § VIII).

In the density-wave theory, this has a very important implication: The newly formed stars must eventually migrate out of the arms, since the former revolve at the angular velocity of the material (say 25 km sec^{-1} kpc^{-1}), while the latter move at the speed of the pattern (say 13 km sec^{-1} kpc^{-1}). Over a period of 10 million years, there would be a separation of the stars from the gaseous arm by something like 1.2 kpc in the solar vicinity. However, since the inclination of the spiral arms is small, the *radial* separation is only about one-tenth of the above amount, and becomes difficult to detect. (For a more detailed discussion, see Lin 1967*b*.)

With the spiral pattern discussed in § IV, it is found that the young optical objects indeed lie within the spiral arms, with the notable exception of the Orion arm (already discussed in § IV). Although the optical objects in the Perseus arm seem to indicate a fairly large local inclination, their radial spread is not so large as to cause any difficulty in locating them within an arm of small over-all inclination. There seems to be, however, some difficulty with the optical objects in the Vela-Carina direction. There may very



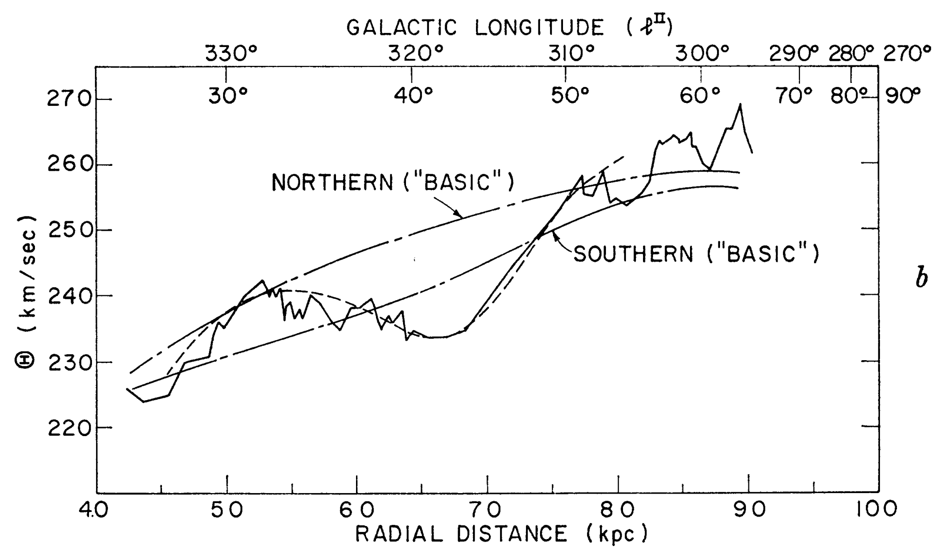


FIG. 5.—*a*, Northern Hemisphere. *b*, Southern Hemisphere. Rotation curves. Observation (Kerr 1964) ———; theory ----; mean motion, theory —-—-—-.

well be some local irregularities special to this region. Further theoretical and observational studies are certainly much to be desired.

At this point, it should be emphasized again that the theory primarily deals with the over-all pattern, the grand design, rather than with the detailed structure. It should, however, be possible to introduce local effects, such as the radial expansion of stars from an O-association, to explain local irregularities.

A closer examination of the structure of the spiral arm should then reveal the following three "lanes": (1) a dust lane indicating the place of highest gas concentration, (2) a lane of brilliant newly formed young stars with large H II regions—the string of beads noted by Morgan and his collaborators (see Sharpless 1965)—and (3) a lane of dying O stars with much smaller H II regions.[12]

Since the dust is suspended in the gas, it is plausible that it should be the thickest where the gas density is the highest. This is known to be true on a large scale. As the dust particles pass through the spiral arm, they will evaporate under the effect of the radiation of the brilliant young stars. (On a smaller scale, there must also be very little dust in the immediate vicinity of bright stars.) One might therefore conceive of a process of dust formation after the gas (moving in a nearly circular orbit) has left the bright and hot part of the spiral arm. This process continues to produce more dust, so that the highest concentration of dust is reached just before the particles are again evaporated at the next arm. If it is assumed that the pattern speed is about one-half the circular speed of the material (as is expected to be approximately true in our vicinity), the time between destruction and recongregation into a dust lane is about 1 galactic year. This should be compared with a typical time scale for dust formation, and a figure of 2×10^8 years is by no means unreasonable.

VII. THE MIGRATION OF MODERATELY YOUNG STARS

Ever since the method for the determination of stellar ages was greatly improved by Strömgren and his collaborators,[13] it has become meaningful to look for the birthplaces of the moderately young stars by following the history of their migration. It is expected that these stars would all be found to be formed inside spiral arms. For these calculations, one has, of course, to know the present location of the stars and to obtain their velocities from observations of their proper motions. All these are now obtainable with sufficient accuracy. Indeed, Contopoulos and Strömgren (1965) constructed tables for the easy calculation of such orbits, and had some success with the program. Some preliminary results were reported by Strömgren (1967), and it was concluded that these are generally compatible with the density-wave theory.

An important factor not included in this earlier analysis was the effect of the spiral gravitational field. Preliminary explorations made by Lin with the help of S. C. Wang, showed that even a small spiral field (equal to 5 per cent of the symmetrical field, for example) could be quite significant. It is thus necessary first to determine the spiral gravitational field for our Galaxy by using the results of § V. By experimenting with different choices of pattern frequencies and field strengths, Yuan (1969) found that a good choice would be a pattern speed of about 13.5 km sec^{-1} kpc^{-1} and a field strength of about 5 per cent of the symmetrical field. The results of his analysis are shown in Figures 6 and 7. In Figure 6, the places of formation of twenty-five stars are shown by using the ages determined by Strömgren *et al.* (private communication). Clearly, they appear to lie on one spiral arm which is *leading* and does not fit into any known struc-

[12] Specifically, A. Sandage pointed out to Lin the possibility of making these detailed observations and analyses in connection with M33. He also mentioned the plans of B. Lynds in this work.

[13] See Crawford (1963, 1966); Strömgren (1963*a*, *b*, 1964, 1965, 1966*a*, *b*); Strömgren and Perry (1965); Kelsall and Strömgren (1966); Crawford and Barnes (1966); Crawford and Perry (1966); and Crawford and Strömgren (1966).

ture from radio observations. In Figure 7, the places of formation of the same twenty-five stars are shown when the analysis includes the field described above. It is seen that all of them now lie in spiral arms.

The age determination suffers from an uncertainty of 10–15 per cent. We have taken the ages recommended by Strömgren *et al.*, with a few exceptions, but all well within the error limits given. In these few cases, the age was chosen to minimize the components of the dispersion velocity "at birth." In all cases, each component of the dispersion velocity is within 25 km sec^{-1}. This is deliberately taken to be higher than the initial turbulent velocity of the gas in order to allow for the increase during the formation process and for the inaccuracies of the parameters involved in our calculations.

VIII. CONCLUDING REMARKS

a) The density-wave theory has many attractive features. At the present time, there is as yet no definite criterion for the precise determination of the pattern speed Ω_p. The observational check of the theory therefore lies in the comparison with a multitude of data. One set of data may be regarded as a basis for determining the pattern speed, while the others are then used as checks. Indeed, the value of Ω_p turns out to be *rather sharply*

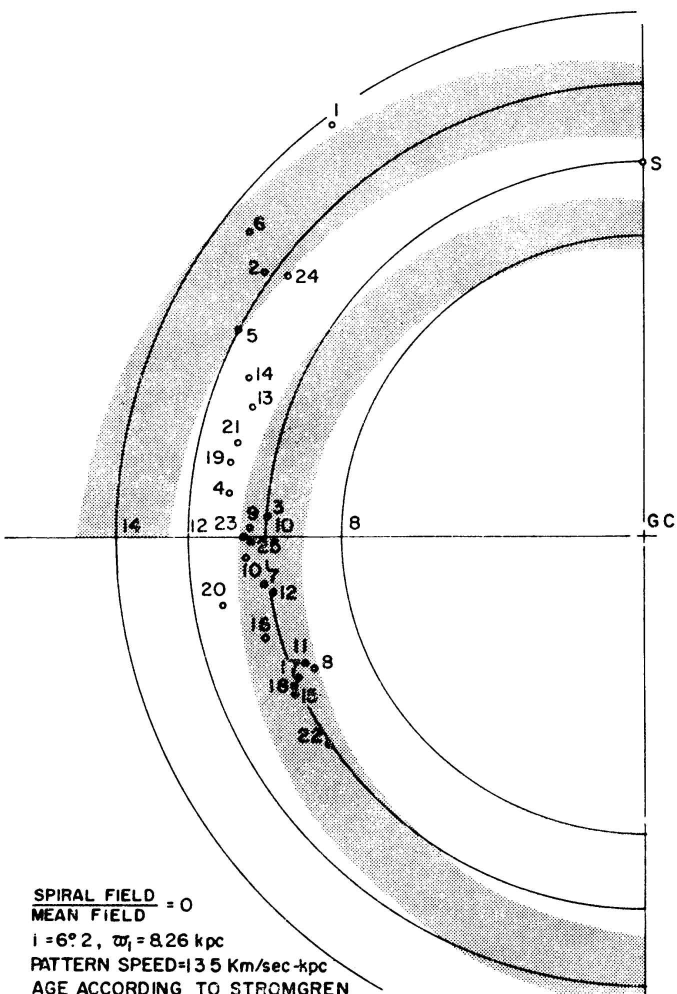


FIG. 6.—Places of formation of stars as determined without a spiral field

defined by the observational data. Other comparisons between theory and observations are also extremely satisfactory. For example, in the problem of the migration of moderately young stars, there are a great many factors that could cause the results to go astray, and yet we have obtained impressive agreement. Further continuation and extension of the observational program would be extremely desirable.

In general, it might be noted that the numerical values determined in this paper are subject to changes with the observational data and the theoretical models used, but we can have reasonable confidence in them, perhaps to within 20 per cent.

b) The range of values 11–13 km sec^{-1} kpc^{-1} obtained for the pattern speed Ω_p is also very satisfactory from the theoretical point of view. With reference to Figure 1, it is clear that Lindblad resonance is very nearly realized with Ω_p in this range, but *there is a definite deviation.* In the theory of Lindblad, co-operative effects were neglected; thus, he insisted on exact resonance, which can happen only if an extremely special rotation curve $\Omega(\varpi)$ holds—i.e., one for which $\Omega - \kappa/2 =$ const. We include the co-operative effects, and *the deviation is a measure of the scale of the spiral spacing* in the theory of the infinitesimally thin disk. A finite scale must be allowed; otherwise, the peculiar motions of the stars, both in the galactic plane and perpendicular to it, would

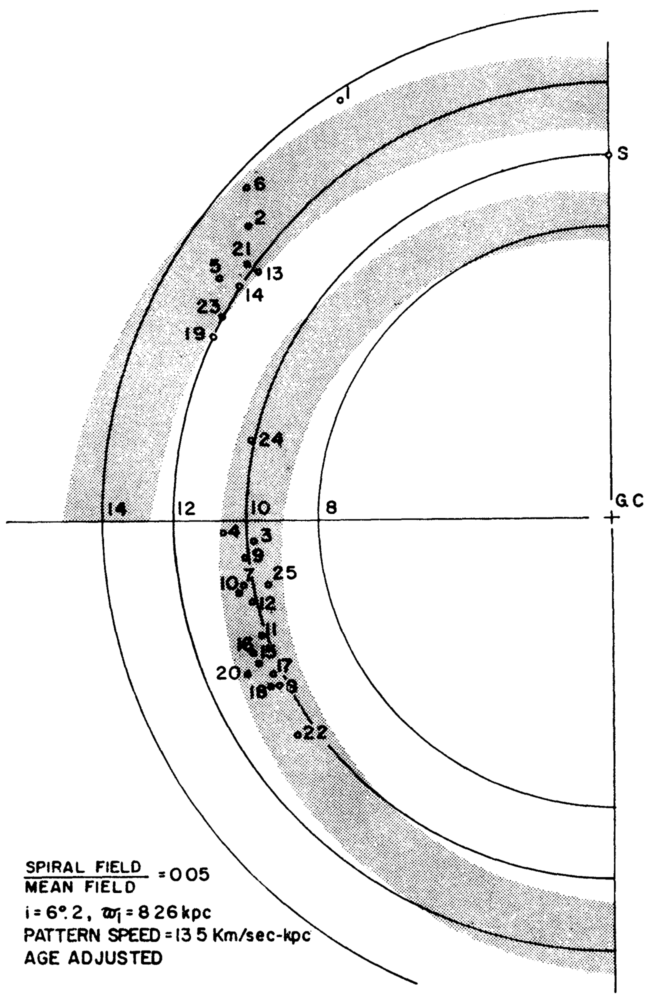


Fig. 7.—Places of formation of stars as determined with the inclusion of a spiral gravitational field travelling at a pattern speed of 13.5 km sec^{-1} kpc^{-1} and having an intensity equal to 5 per cent of that of the symmetrical gravitational field.

suppress the waves. Thus, by taking the pattern speed Ω_p in the range indicated by present investigations, we are "as close to resonance as possible," without giving up density waves altogether. Such is indeed the condition expected of any mechanical system.

c) The strength of the spiral field has now been estimated from the study of the migration of moderately young stars. It would be interesting to see how much contrast in gas density would ensue as a consequence. Calculations of this type have been made by Fujimoto (1966) and by Roberts (1969). Since the parameters chosen by Fujimoto do not correspond to those for a two-armed pattern discussed in this paper, only the results of Roberts are useful as a basis for checking the linear theory. A comparative study shows that estimates made on the basis of an extrapolation of the linear theory give generally satisfactory results.

d) The calculations of Fujimoto and of Roberts indicate the formation of a shock. This might indeed be an essential element in the process of star formation. One might imagine that the gas in turbulent motion has "clouds" before the shock, which are on the verge of gravitational collapse; the sudden compression would then trigger off the collapse of the clouds, which would lead to star formation. After the gas left the shock region, it would again be decompressed, and the process of star formation would cease. Superficially, such a process appears to be in such excellent agreement with our general concepts that it deserves further careful study in the immediate future.

An implication of this discussion is that the concentration of gas, though relatively high, need not be sufficient to imply extensive formation of new stars. The other prerequisite[14] of a shock must be fulfilled in many cases. This conclusion appears to be in general agreement with the abundance distribution of neutral hydrogen found by Roberts (1967) in certain galaxies. The highest abundance of hydrogen in certain galaxies is found in a ring *outside* the spiral structure. This is readily explained on the basis of the present theory, if it is supposed that shocks have not been able to form in these outer parts, because the conditions there (with too few stars and too much gas) are not favorable to the propagation of density waves.

Other corroborative evidence is the absence of H II regions within 4 kpc of the galactic center. Our theory suggests that this should be the case, since there cannot be any density waves inside the radius of inner Lindblad resonance (cf. Fig. 3). Indeed, the 109α radio observations reveal no H II regions in such localities (outside the dense nucleus).[15]

e) Once the process of star formation is clarified, one can proceed to study the evolution of a galaxy as a whole. Presumably, larger galaxies evolve faster, and the gas content in them would be more nearly depleted *by now* than in a smaller galaxy, if we assume that all the galaxies were formed around the same epoch. Thus, in a future epoch, a smaller galaxy might take on some of the present appearances of a larger galaxy, even though the two objects are *never* the same at any time. These speculations must be confirmed by careful analysis and substantiated by the observed stellar and gas contents in galaxies of various types.

f) The points of Lindblad resonance play important roles in the present theory. In the first place, they mark the extent of the spiral pattern in a theory involving stars. The scale of the wavelength reduces to zero at these points when the theory is developed on the basis of an infinitesimally thin disk. Thus, the spiral pattern is substantially "insulated" from the boundary conditions at the center of the galaxy. If the conditions are such that both the resonance points are absent, the existence of a *spiral* pattern becomes more doubtful. (This remark is presumably relevant to the anti-spiral theorem of Lynden-Bell and Ostriker (1967) in the case of a purely gaseous disk with pressure.

[14] Still another favorable condition is a high degree of turbulence, as will be discussed elsewhere.

[15] See Burke (1968) and the older work of Westerhout on background radiation, as quoted on page 196 in Kerr and Westerhout (1965).

In the gaseous case, the resonance condition described above does not exist; there is no case in which the wavelength reduces to zero.)

The behavior of stars near resonance has been studied thoroughly by Shu (1968). The behavior of gas has been studied by Roberts (1969), who found that there are "free modes" near resonance, i.e., there are solutions to equations of gas dynamics which correspond to non-circular motion even when the spiral component of the gravitational field is absent.

Clearly, the initiation of a spiral structure by mild instability of the type described in Paper II (caused by the gradient of dispersion velocities) is intimately related to the existence of a neutral spiral mode satisfying boundary conditions. Thus, the total absence of resonance points makes it unlikely for this instability mechanism to become operative. Consequently, conclusions obtained from the studies of a uniformly rotating disk need not apply to real, differentially rotating, disk-shaped galaxies.

During the course of this work, various discussions were held with numerous friends, too many to be acknowledged individually. We wish to register our gratitude collectively for their contributions to the theoretical ideas and to the proper interpretation of the observational data. We should, however, mention especially Professors B. Strömgren, L. Woltjer, and K. Prendergast for their invaluable help during the early and crucial phases of this work. Discussions with Professors J. H. Oort, Bart J. Bok, Bernard F. Burke, Frank Kerr, Morton S. Roberts, Allan R. Sandage, and Maarten Schmidt have given us valuable information on various observational facts. We have also benefited from discussions of the theoretical ideas with Professor D. Layzer at the early stages and from extensive discussions with Professor A. Toomre and Dr. A. Kalnajs over a period of the past few years.

This work is supported in part by grants from the National Science Foundation and from the National Aeronautics and Space Administration. Computations were carried out at the Computation Center of the Massachusetts Institute of Technology and at the Institute for Space Studies NASA in New York. We wish to express our thanks to Dr. Robert Jastrow for making the facilities available and to Dr. Kenneth Grossman for carrying out certain calculations related to the problem of star migration.

APPENDIX A

RESPONSE OF A STELLAR DISK TO A SPIRAL GRAVITATIONAL FIELD

We give here an outline of the calculation of the response of a stellar disk to a spiral gravitational field.

Let t be the time, and (Π, Θ, Z) be the components of total stellar velocity in the cylindrical coordinates (ϖ,θ,z). We shall restrict ourselves to the two-dimensional case and introduce the peculiar velocity components (c_ϖ,c_θ) by the relations

$$c_\varpi = \Pi\,, \qquad c_\theta = \Theta - \varpi\Omega(\varpi)\,, \tag{A1}$$

where $\Omega(\varpi)$ is the angular velocity of circular motion defined by balancing the centrifugal acceleration $\varpi\Omega^2$ with the symmetrical gravitational field. The distribution function $\Psi(\varpi,\theta,c_\varpi,c_\theta t)$ then satisfies the differential equation

$$\frac{\partial\Psi}{\partial t} + c_\varpi\frac{\partial\Psi}{\partial\varpi} + \left(\Omega + \frac{c_\theta}{\varpi}\right)\frac{\partial\Psi}{\partial\theta} + \left(a_\varpi + \Omega_\varpi{}^2 + 2\Omega c_\theta + \frac{c_\theta{}^2}{\varpi}\right)\frac{\partial\Psi}{\partial c_\varpi} + \left(a_\theta - \frac{\kappa^2}{2\Omega}c_\varpi - \frac{c_\varpi c_\theta}{\varpi}\right)\frac{\partial\Psi}{\partial c_\theta} = 0\,, \tag{A2}$$

where $\kappa(\varpi)$ is the epicyclic frequency defined by

$$\kappa^2 = (2\Omega)^2\left(1 + \frac{\varpi}{2\Omega}\frac{d\Omega}{d\varpi}\right) , \tag{A3}$$

and (a_ϖ, a_θ) are the components of acceleration of the prevailing gravitational field. In conformity with our scheme (Lin and Shu 1966), we do not impose a condition of self-consistency at this point. Instead, we assume that the resultant field is of the form

$$(a_\varpi, a_\theta) = (\varpi\Omega^2, 0) + (a_{\varpi 1}, a_{\theta 1}) , \tag{A4}$$

where $(a_{\varpi 1}, a_{\theta 1})$ are derivable from a spiral gravitational potential,

$$\mathcal{V}_1 = A(\varpi) \exp\{i[\omega t - m\theta + \Phi(\varpi)]\} ; \tag{A5}$$

that is,

$$a_{\varpi 1} = -\frac{\partial \mathcal{V}_1}{\partial \varpi} , \quad a_{\theta 1} = -\frac{1}{\varpi}\frac{\partial \mathcal{V}_1}{\partial \theta} . \tag{A6}$$

In equation (A5), both $A(\varpi)$ and $\Phi(\varpi)$ are real, and ω is a complex parameter (cf. eq. [2.1] in the text).

If Ψ_0 is the symmetrical distribution function corresponding to the case $A = 0$, and if we write

$$\Psi = \Psi_0(1 + \psi) = e^{-Q_0}(1 + \psi) , \tag{A7}$$

then Ψ satisfies the linearized equation,

$$\begin{aligned} &\frac{\partial \psi}{\partial t} + c_\varpi \frac{\partial \psi}{\partial \varpi} + \left(\Omega + \frac{c_\theta}{\varpi}\right)\frac{\partial \psi}{\partial \theta} + \left(2\Omega + \frac{c_\theta}{\varpi}\right)c_\theta \frac{\partial \psi}{\partial c_\varpi} \\ &- \left(\frac{\kappa^2}{2\Omega} + \frac{c_\theta}{\varpi}\right)c_\varpi \frac{\partial \psi}{\partial c_\theta} = a_{\varpi 1}\frac{\partial Q_0}{\partial c_\varpi} + a_{\theta 1}\frac{\partial Q_0}{\partial c_\theta} . \end{aligned} \tag{A8}$$

With the field components (eq. [A6]), Ψ is naturally expected to be of the form

$$\psi = \phi(\varpi, c_\varpi, c_\theta) \exp\{i[\omega t - m\theta + \Phi(\varpi)]\} . \tag{A9}$$

The amplitude function ϕ satisfies the equation

$$\begin{aligned} &c_\varpi \frac{\partial \phi}{\partial \varpi} + \left(2\Omega + \frac{c_\theta}{\varpi}\right)c_\theta \frac{\partial \phi}{\partial c_\varpi} - \left(\frac{\kappa^2}{2\Omega} + \frac{c_\theta}{\varpi}\right)c_\varpi \frac{\partial \phi}{\partial c_\theta} \\ &+ i\left[k c_\varpi - m\frac{c_\theta}{\varpi} + (\omega - m\Omega)\right]\phi = g_\varpi \frac{\partial Q_0}{\partial c_\varpi} + g_\theta \frac{\partial Q_0}{\partial c_\theta} , \end{aligned} \tag{A10}$$

where

$$k(\varpi) = \Phi'(\varpi) \tag{A11}$$

and

$$(a_{\varpi 1}, a_{\theta 1}) = (g_\varpi, g_\theta) \exp\{i[\omega t - m\theta + \Phi(\varpi)]\} . \tag{A12}$$

Asymptotic solution of equation (A10).—If we are dealing with the case of tightly wound spirals, we may attempt an asymptotic series of the form

$$\varphi = \varphi^{(0)} + \epsilon\varphi^{(1)} + \ldots , \tag{A13}$$

where ϵ is a parameter of the order of $(k\varpi)^{-1}$. Consistent with this approximation, we shall first introduce dimensionless velocities referred to local values; thus, we shall write

$$c_\varpi = \xi V_1 = \xi(2\Omega/\kappa)(\varpi\Omega) , \tag{A14a}$$

$$c_\theta = \eta(\varpi\Omega) . \tag{A14b}$$

(The scales are suggested by our experience with the theory of epicycles.) It is then not difficult to verify that the equation for $\varphi^{(0)}$ is

$$\eta \frac{\partial \varphi^{(0)}}{\partial \xi} - \xi \frac{\partial \varphi^{(0)}}{\partial \eta} + i(\nu + a\xi)\varphi^{(0)} = -\frac{iaA}{V_1^2}\frac{\partial Q_0}{\partial \xi}, \tag{A15}$$

where a and ν are defined by

$$a = (k\varpi)(2\Omega^2/\kappa^2), \tag{A15a}$$

and

$$\nu = (\omega - m\Omega)/\kappa. \tag{A15b}$$

Notice that a characteristic integral of equation (A15) is

$$\xi^2 + \eta^2 = C(\varpi), \tag{A16}$$

thanks to the choice of scales adopted. For our present purposes, we shall also take the basic distribution to be a function of $\xi^2 + \eta^2$ and ϖ; i.e.,

$$Q_0 = Q_0(\xi^2 + \eta^2, \varpi). \tag{A17}$$

Such a distribution function holds, when only stars of small dispersion speeds are considered. Since only a small fraction of stars have large dispersion speeds, and since they are not expected to be greatly influenced by the spiral gravitational field, equation (A17) is a good approximation for our present purposes. It is also compatible with the usually adopted Schwarzschild distribution. In that case, Q_0 is indeed of the form

$$Q_0 = \tfrac{1}{2}\left[\frac{c_\varpi^2}{\langle c_\varpi^2\rangle} + \frac{c_\theta^2}{\langle c_\theta^2\rangle}\right] + Q_{00}(\varpi), \tag{A18}$$

where $\langle c_\varpi^2\rangle$ and $\langle c_\theta^2\rangle$ are the mean-square values of the components of dispersion velocities for the location ϖ. For the moment, we shall consider general distribution functions of the form (A17). Note that ϖ is a parameter in all these calculations.

To integrate equation (A15), we introduce the characteristic integral (A16) as one of the variables; i.e., we introduce the polar coordinate system (τ,s) in the (ξ,η)-plane:

$$\xi = \tau \cos s, \quad \eta = \tau \sin s. \tag{A19}$$

Then equation (A15) becomes

$$-\frac{d\varphi^{(0)}}{ds} + i(\nu + a\tau \cos s)\varphi^{(0)} = -\frac{2iaA}{V_1^2} Q'(\tau^2)\tau \cos s, \tag{A20}$$

where $Q_0'(\tau^2)$ denotes the derivative of Q_0 with respect to τ^2, the dependence on ϖ being understood. Both τ^2 and ϖ are, of course, to be treated as parameters for the purposes of the differential equation (A20). As we shall demonstrate below, the solution of this differential equation is

$$\varphi^{(0)} = -\frac{2A}{V_1^2} Q_0'(\tau^2)[1 - q(a\xi, a\eta, \nu)], \tag{A21}$$

where

$$q = \frac{\nu\pi}{\sin(\nu\pi)} \cdot \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp\{i[\nu s - a\xi \sin s + a\eta(1 + \cos s)]\}\, ds. \tag{A22}$$

It is readily verified by direct substitution that equation (A22) is indeed a solution of the original equation (A15).

The derivation of equation (A21) from equation (A19) is also not difficult. The general solution of equation (A21), which involves one "constant" of integration depending both on τ (to be exhibited) and on ϖ (not to be exhibited), is

$$\frac{V_1^2}{2AQ_0'(\tau^2)}\varphi^{(0)} = e^{iF(s,\tau)}\left[C(\tau) + ia\int_0^s e^{-iF(s',\tau)}\tau \cos s' ds'\right], \tag{A23}$$

where

$$F(s,\tau) = \nu s + a\tau \sin s . \tag{A24}$$

To determine $C(\tau)$, we apply the condition of periodicity, equating $\varphi_0(s)$ with $\varphi_0(s + 2\pi)$. We then consider the integral in $\varphi_0(s + 2\pi)$ in its alternative form given by the right-hand side of the identity,

$$\int_0^{s+2\pi} = \int_0^{2\pi} + \int_{2\pi}^{2\pi+s} . \tag{A25}$$

The change of variable $s' = 2\pi + s''$ in the last integral of equation (A25) reduces it to the form

$$\int_0^s e^{-iF(s''+2\pi,\tau)}\tau \cos s'' ds'' = e^{2\pi i\nu}\int_0^s e^{-F(s'',\tau)}\tau \cos s'' ds'' . \tag{A26}$$

In the form on the right-hand side of equation (A26), the integral is the same as that occurring in $\varphi_0(s)$. Some obvious calculations would then yield the desired determination of $C(\tau)$, and thence the result (A21).

Some remarks on the solution (A21).—The solution (A21) may also be written in the following form. If we recall the definitions of ψ and φ (eqs. [A7] and [A9]), we see that the "response," in terms of the distribution function itself, is

$$\Psi_1 = \Psi_0\psi = \frac{2\mathcal{V}_1}{V_1^2}\Psi_0'(\tau^2)(1 - q) , \tag{A27}$$

where $q(a\xi,a\eta,\nu)$ is defined by equation (A22), ξ, η are defined by equation (A14), and $\tau^2 = \xi^2 + \eta^2$. The parameters a and ν are defined by equations (A15a) and (A15b). For the Schwarzschild distribution, we have

$$\frac{-2\mathcal{V}_1}{V_1^2}\Psi_0'(\tau^2) = \frac{\mathcal{V}_1}{\langle c_\varpi^2\rangle}\Psi_0(\tau^2) . \tag{A28}$$

It is important to notice that the density response, obtained by integrating equation (A27) with respect to c_ϖ and c_θ, is always *real for real* ω. To see this, we first note that $\Psi_0'(\tau^2)$ is even in η. The imaginary part of q appears in terms of the two integrals,

$$\mathfrak{J}_1 = \int_{-\pi}^{\pi} \cos(\nu s - a\xi \sin s) \sin[a\eta(1 + \cos s)]ds$$

and

$$\mathfrak{J}_2 = \int_{-\pi}^{\pi} \sin(\nu s - a\xi \sin s) \cos[a\eta(1 + \cos s)]ds .$$

Now, $\mathfrak{J}_2$ is clearly zero, since the integrand is odd in s. The integral $\mathfrak{J}_1$ does not vanish, but it is clearly odd in η. Thus, when the integration in η (i.e., in c_θ) is performed, the net contribution vanishes. Indeed, all even moments of c_θ, weighted with an arbitrary power of c_ϖ, vanish for the perturbation (A27).

APPENDIX B

CALCULATION OF RESPONSES IN DENSITY AND IN COMPONENTS OF MEAN STELLAR VELOCITY

We consider the case where the basic distribution function is the Schwarzschild distribution,

$$\Psi_0 = P_0(\varpi) \exp\left[-\frac{\mu_0}{2}(\xi^2+\eta^2)\right], \tag{B1}$$

where μ_0^{-1} is a measure of the mean-square value of the dispersion velocities

$$\mu_0^{-1} = \langle c_\varpi^2 \rangle / V_1^2 , \qquad V_1 = (2\Omega/\kappa)(\varpi\Omega) , \tag{B1a}$$

and P_0 is the normalization factor, so that

$$\sigma_{*0} = m_* \int\int \Psi_0 dc_\varpi dc_\theta . \tag{B2}$$

The response of distribution function to the spiral gravitational field is now of the form

$$\Psi_1 = \frac{-\mathcal{V}_1}{\langle c_\varpi^2 \rangle} \Psi_0(1-q) , \tag{B3}$$

where q is defined by equation (A21).

For the state $\Psi = \Psi_0 + \Psi_1$ we have

$$\sigma_*(\varpi,\theta,t) = m_* \int\int \Psi dc_\varpi dc_\theta , \tag{B4a}$$

and also the following relations for the mean stellar velocity components v_ϖ and v_θ:

$$\sigma_* v_\varpi = m_* \int\int \Psi c_\varpi dc_\varpi dc_\theta , \tag{B4b}$$

$$\sigma_* v_\theta = m_* \int\int \Psi c_\theta dc_\varpi dc_\theta . \tag{B4c}$$

It is easy to verify that, within the approximation of the linear theory,

$$\frac{\sigma_{*1}}{\sigma_{*0}} = \frac{\int\int \Psi_1 dc_\varpi dc_\theta}{\int\int \Psi_0 dc_\varpi dc_\theta} = \frac{-\mathcal{V}_1}{\langle c_\varpi^2 \rangle} \langle (1-q) \rangle , \tag{B5a}$$

$$\frac{v_\varpi}{V_1} = \frac{-\mathcal{V}_1}{\langle c_\varpi^2 \rangle} \langle \xi(1-q) \rangle , \tag{B5b}$$

$$\frac{v_\theta}{\varpi\Omega} = \frac{-\mathcal{V}_1}{\langle c_\varpi^2 \rangle} \langle \eta(1-q) \rangle , \tag{B5c}$$

where $\langle\ \rangle$ denotes the weighted average with respect to Ψ_0 or the function $\exp[-\frac{1}{2}\mu_0(\xi^2+\eta^2)]$.

The calculation of these averages depends on well-known relationships of the type

$$\langle e^{i\lambda\xi} \rangle = e^{-\lambda^2/2\mu_0} , \tag{B6a}$$

$$\langle \xi e^{i\lambda_1\xi} \rangle = \frac{i\lambda_1}{\mu_0} e^{-\lambda_1^2/2\mu_0} , \tag{B6b}$$

$$\langle \eta e^{i\lambda_2\eta} \rangle = \frac{i\lambda_2}{\mu_0} e^{-\lambda_2^2/2\mu_0} , \tag{B6c}$$

for Gaussian distribution functions. It is then easy to verify, by using equation (B5a), that

$$\langle (1-q) \rangle = 1 - \frac{\nu\pi}{\sin(\nu\pi)} \mathfrak{S}_\nu(x) , \tag{B7}$$

where $x = k^2\langle c_\varpi^2\rangle/\kappa^2$, and

$$\mathfrak{G}_\nu(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-x(1+\cos s)} \cos \nu s ds \,. \tag{B8}$$

The reduction factor $\mathfrak{F}_\nu(x)$ that occurs in equation (2.13a) in the text may then be introduced by defining it with the relation

$$\mathfrak{F}_\nu(x) = \frac{1-\nu^2}{x}\left[1 - \frac{\nu\pi}{\sin(\nu\pi)}\mathfrak{G}_\nu(x)\right], \tag{B9}$$

so that equation (2.13a) takes on a form directly comparable with the case of zero dispersion speeds.

TABLE 1A

NUMERICAL VALUES OF THE REDUCTION FACTOR $\mathfrak{F}(\nu,x)$

ν	x								
	0	1	2	3	4	5	6	7	8
0 0.	1	0 5342	0 3457	0 2523	0 1982	0 1633	0 1389	0 1209	0 1068
0 1.	1	.5333	.3447	.2514	1974	1625	.1381	.1202	1066
0 2.	1	.5305	.3416	.2484	1946	.1600	.1359	1182	.1046
0 3	1	.5258	3363	.2434	.1901	.1558	.1321	.1147	.1013
0.4	1	.5191	3286	.2362	.1835	.1499	.1267	.1097	0966
0 5	1	.5100	.3185	.2266	.1748	.1420	1195	.1032	.0907
0 6	1	.4984	.3056	.2145	.1638	.1321	.1105	.0950	.0833
0 7.	1	.4838	.2894	.1994	1503	.1199	.0995	.0849	0739
0 8.	1	.4658	.2695	.1810	.1337	.1050	.0861	.0726	0629
0.9.	1	.4435	2452	1586	.1136	.0871	.0700	.0581	0494
1 0..	1	0 4158	0 2153	0 1312	0 0894	0.0656	0 0507	0.0408	0.0341

The evaluation of the integral in equation (B5b) can be done with the help of equation (B6b), where $\lambda_1 = \alpha \sin s$. We find that an integral of the form

$$\mathfrak{H}_\nu(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-x(1+\cos s)} \sin \nu s \sin s ds \tag{B10}$$

is involved. By partial integration, this can be transformed into

$$\mathfrak{H}_\nu(x) = \frac{\sin(\nu\pi)}{\pi x}\left[1 - \frac{\nu\pi}{\sin(\nu\pi)}\mathfrak{G}_\nu(x)\right]. \tag{B11}$$

Thus, the result can be best summarized by writing

$$\langle\xi(1-q)\rangle : \langle(1-q)\rangle = -\frac{\nu}{\alpha} \tag{B12}$$

or

$$\frac{v_\varpi}{\varpi\Omega} : \frac{\sigma_{*1}}{\sigma_{*0}} = -\frac{\omega - m\Omega}{\Omega(k\varpi)} \,. \tag{B13}$$

It is worth noting that equation (B13) may be obtained directly by using the equation of continuity in the continuum form. The reduction factor $\mathfrak{F}_\nu^{(1)}(x)$ in equation (2.13b) is thus identical with $\mathfrak{F}_\nu(x)$:

$$\mathfrak{F}_\nu^{(1)}(x) = \mathfrak{F}_\nu(x) \,. \tag{B14}$$

The numerical values $\mathfrak{F}_\nu(x)$ are shown in Table 1A and Figure 8.

To evaluate equation (B4c), we follow a similar method by using equation (B6c) with $\lambda_2 = a(1 + \cos s)$. The result can be best summarized by writing

$$\langle \eta(1-q) \rangle : \langle \xi(1-q) \rangle = -i\mathfrak{G}_\nu'(x) : \mathfrak{H}_\nu(x) \,, \tag{B15}$$

or

$$\frac{\langle c_\theta \rangle}{\varpi\Omega} : \frac{\langle c_\varpi \rangle}{(\varpi\Omega)(2\Omega/\kappa)} = -i\mathfrak{G}_\nu'(x) : \mathfrak{H}_\nu(x) \,. \tag{B16}$$

In the form used in equation (2.13c) in text, we then have

$$\mathfrak{F}_\nu^{(2)}(x) = \frac{(1-\nu^2)(\nu\pi)}{\sin(\nu\pi)} [-\mathfrak{G}_\nu'(x)] \,, \tag{B17}$$

where $\mathfrak{G}_\nu'(x)$ is the derivative, with respect to its own argument x, of the function $\mathfrak{G}_\nu(x)$ defined by equation (B8).

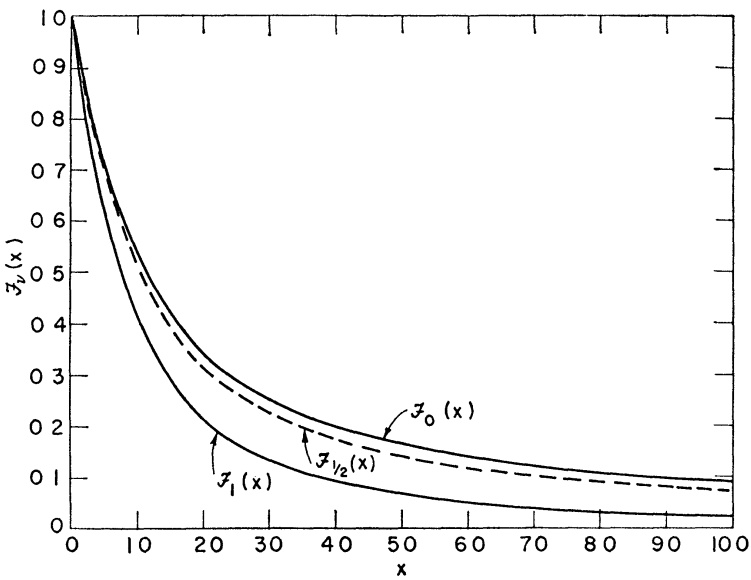


FIG. 8.—The reduction factor. Abscissa is a measure of the mean-square value of the radial component of the peculiar velocity of the stars.

APPENDIX C

AN ALTERNATIVE FORM OF THE DISPERSION RELATION

The dispersion relationship (3.2), in the case of a purely stellar disk, will now be presented in a form that reminds us of the similarity between our problem and that of waves in electromagnetic plasma. In this case, the dispersion relation (3.2) may also be written as

$$\frac{k_*}{|k|} = \frac{1}{x}\left[1 - \frac{\nu\pi}{\sin(\nu\pi)}\mathfrak{G}_\nu(x)\right] \,, \tag{C1}$$

where $\mathfrak{G}_\nu(x)$ is defined by equation (B8). Clearly, the function on the right-hand side has singularities (poles) at the points $\nu = \pm n$, where n is a positive integer. Otherwise, the function is analytic in the complex variable ν. Such meromorphic functions can be expanded into an infinite series of the principal parts at the poles, according to a well-known theorem of Mittag-Leffer. Clearly, the principal parts at $\nu = +n, -n$ are related, and it is therefore convenient to consider the series in the form

$$\frac{\nu\pi}{\sin(\nu\pi)}\mathfrak{G}_\nu(x) = \sum_{n=1}^{\infty}\frac{B_n(x)}{\nu^2 - n^2}\,. \tag{C2}$$

By taking the limit $\nu \rightarrow n$, we find that

$$B_n(x) = (-)^n n^2 I_n(x)\,, \tag{C3}$$

where $I_n(x)$ is the modified Bessel function of order n. Thus, equation (C1) may be written as

$$\frac{k_*}{k} = \frac{1}{x}\left[1 - \sum_{n=1}^{\infty}\frac{(-)^n I_n(x)}{(\nu/n)^2 - 1}\right]. \tag{C4}$$

The difference between this and the familiar formula for plasma physics (Bernstein 1958) lies essentially in (i) the difference in geometry and (ii) a reversal of sign caused by the difference between gravitational and electromagnetic forces.

The present form of the theory, as embodied in equation (3.2), is, however, much more perspicuous for the spiral problem, especially when there is also gas present. The concept of the reduction factor has been found to be particularly useful.

REFERENCES

Becker, W. 1964, *Zs. f. Ap.*, **58**, 202.
Bernstein, I. B. 1958, *Phys. Rev.*, **109**, 10.
Bok, B. J. 1967, *Am. Scientist*, **55**, 375.
Bok, B. J., and Contopoulos, G. 1967, *Rept. I.A.U. Commission*, No. 33.
Burke, B. F. 1968, presented at the Charlottesville Conference on H II regions.
Burton, W. B. 1966, *B.A.N.*, **18**, 247–255.
Chandrasekhar, S. 1942, *Principles of Stellar Dynamics* (rev. ed. 1960; New York: Dover Publications).
Contopoulos, G., and Strömgren, B. 1965, *Tables of Plane Galactic Orbits* (New York: Institute for Space Studies).
Crawford, D. L. 1963, *Ap. J.*, **137**, 523.
———. 1966, in *Proc. I.A.U. Symp. No. 24*, p. 170 (*Contributions from the Kitt Peak National Observatory*, No. 110).
Crawford, D. L., and Barnes, J. V. 1966, *A.J.*, **71**, 610 (*Contributions from the Kitt Peak National Observatory*, No. 192).
Crawford, D. L., and Perry, C. L. 1966, *A.J.*, **71**, 206 (*Contributions from the Kitt Peak National Observatory*, No. 147).
Crawford, D. L., and Strömgren, B. 1966, in *Vistas in Astronomy*, ed. A. Beer, **8** (New York: Pergamon Press), 149–157 (*Contributions from the Kitt Peak National Observatory*, No. 82).
Fujimoto, M. 1966, in *Proc. I.A.U. Symp. No. 29* (to be published).
Hunter, C. 1963, *M.N.R.A.S.*, **126**, 299.
———. 1965, *ibid.*, **129**, 321.
Kalnajs, A. J. 1965, unpublished Ph D. dissertation, Harvard University.
Kelsall, T., and Strömgren, B. 1966, in *Vistas in Astronomy*, ed. A. Beer, **8** (New York: Pergamon Press), 159–178.
Kerr, F. J. 1962, *M.N.R.A.S.*, **123**, 327.
———. 1964, *I.A.U. Symp. No. 20*, p. 81.
Kerr, F. J., and Westerhout, G. 1965, in *Galactic Structure*, ed. A. Blaauw and M. Schmidt (Chicago: University of Chicago Press), p. 167.
Lin, C. C. 1966, SIAM *J. Appl. Math.*, **14**, 876.
———. 1967*a*, *Ann. Rev. Astr. and Ap.*, **5**, 453.
———. 1967*b*, *Lectures in Applied Mathematics (1965)*, **2**: *Relativity Theory and Astrophysics* (New York: American Mathematical Society), p. 66.
———. 1968, in *Galaxies and the Universe*, ed. L. Woltjer (New York: Columbia University Press), p. 33.
Lin, C. C., and Shu, F. H. 1964, *Ap. J.*, **140**, 646.
———. 1966, in *Proc. Nat. Acad. Sci.*, **55**, 229.
———. 1967, in *Proc. I.A.U. Symp. No. 31, Noordwijk, 1966*, p. 313.
Lindblad, B. 1963, *Stockholm Obs. Ann.*, **22**, 3.
Lindblad, P. O. 1960, *Stockholm Obs. Ann.*, **21**, 3.

Lindblad, P. O., 1962, in *Interstellar Matter in Galaxies*, ed. L. Woltjer (New York: W. A. Benjamin), p. 222.
Lynden-Bell, D., and Ostriker, J. D. 1967, *M.N.R.A.S.*, **136**, 293.
Lynds, B. 1967*a*, *Sky and Tel.*, **33**, 343.
———. 1967*b*, *ibid.*, **34**, 18.
Miller, Joseph S. 1968, *Ap. J.*, **151**, 473.
Miller, R. H., and Prendergast, K. H. 1968, *Ap. J.*, **151**, 699.
Oort, J. H. 1962, in *Interstellar Matter in Galaxies*, ed. L. Woltjer (New York: W. A. Benjamin), p. 234.
Prendergast, K. H. 1967, in *Proc. I.A.U. Symp. No. 31, Noordwijk, 1966*, p. 303.
Roberts, M. S. 1967, presented at the Prague meeting of the I.A.U. (Commission 33).
Roberts, W. W. 1969, unpublished Ph.D. thesis, Massachusetts Institute of Technology.
Schmidt, M. 1965, in *Galactic Structure*, ed. A. Blaauw and M. Schmidt (Chicago: University of Chicago Press), p. 513.
Shane, W. W., and Bieger-Smith, G. P. 1966, *B.A.N.*, **18**, 263.
Sharpless, S. 1965, in *Galactic Structure*, ed. A. Blaauw and M. Schmidt (Chicago: University of Chicago Press), p. 131.
Shu, F. H. 1968, unpublished Ph.D. thesis, Harvard University.
Strömgren, B. 1963*a*, *Quart. J.R.A.S.*, **4**, 8.
———. 1963*b*, in *Basic Astronomical Data*, **3**, ed. K. Aa. Strand (Chicago: University of Chicago Press), p. 123.
———. 1964, *Astrophysica Norvegica*, **9**, 333 (*Contributions from the Kitt Peak National Observatory*, No. 59).
———. 1966*a*, in *Stellar Evolution*, ed. R. F. Stein and A. G. W. Cameron (New York: Plenum Press), p. 391.
———. 1966*b*, *Ann. Rev. Astro. and Ap.*, **4**, 433.
———. 1967, in *Proc. I.A.U. Symp. No. 31, Noordwijk, 1966.*
Strömgren, B., and Perry, C. 1965, *Photoelectric uvby Photometry for 1217 Stars Brighter than* $V = 6^{m}.5$, *Mostly of Spectral Classes A, F, and G* (2d version).
Toomre, Alar. 1964, *Ap. J.*, **139**, 1217.
Yuan, Chi. 1969 (to be published).