ON THE SPIRAL STRUCTURE OF DISK GALAXIES

C. C. Lin and Frank H. Shu
Department of Mathematics, Massachusetts Institute of Technology
Received March 20, 1964

ABSTRACT

It is shown that gravitational instability is a plausible basis for the formation of the spiral pattern in disk galaxies An explicit asymptotic formula is obtained for the form of the spiral. It gives reasonable numerical results for the galaxy, and qualitatively satisfactory trends for normal spirals of various types.

I. INTRODUCTION

The mechanism for the formation of the spiral patterns observed in most disk-shaped galaxies has not yet been fully understood. There is little doubt, from the observational data available, that these magnificent manifestations are associated with the interstellar gas and the brilliant young stars born in them. But could the old stars also play an important role in the formation of the spiral structure?

To construct a theory of the spiral structure, one must bear in mind the following important components of a galaxy:

a) *The stars*—with their gravitational forces, circular velocity, and velocity dispersion
b) *The interstellar gas*—with *its* gravitational field and pressure
c) *The magnetic field*—which exerts its influence through the highly conducting interstellar gas.

A complete theory should take all these components and forces into account, and put their relative importance into perspective. Such a theory is not yet available.

There are at least two possible types of spiral theories. The first alternative is to associate every spiral arm with a *given body of matter;* e.g., such an arm might essentially be a tube of gas primarily constrained by the interstellar magnetic field. The difficulty with the disrupting influence of differential rotation in such a theory is well known. The various issues associated with this point of view have been thoroughly discussed recently by Oort (1962). The second alternative is to regard the spiral structure as a *wave pattern*, which either remains stationary, or at least quasi-stationary, in a frame of reference rotating around the center of the galaxy at a proper angular speed (possibly zero).

Three years ago, through discussions with Professors B. Strömgren and L. Woltjer, one of us (C. C. L.) became interested in the possibility of spiral structure from the point of view of gravitational instabilities of the galactic disk, and several of us, including A. Toomre and C. Hunter, began to examine the problem in some detail. Our emphasis and conjectures are not yet in complete agreement. Since A. Toomre's (1964) point of view has been published, it seems desirable to publish our point of view even though the work is not yet as complete as the present writers would wish to have it. Furthermore, we feel that, in any case, the role of gravitational force, which is by far the largest in a galaxy, deserves to be carefully examined for a complete understanding of the problem, whether or not it eventually turns out to be the predominant mechanism.

Toomre tends to favor the first of the possibilities described above. In his point of view, the material clumping is periodically destroyed by differential rotation and regenerated by gravitational instability. It is somewhat difficult to see how this mechanism alone can account for the *relatively regular spiral pattern over the whole disk* in most of the flat galaxies. The present authors favor the second point of view, i.e., that the matter in the galaxy (stars and gas) can maintain a density wave through gravitational interaction in the presence of the differential rotation of the various parts of the disk. This density wave provides a spiral gravitational field which *underlies* the observable concentration of young stars and the gas. In this way, an observable spiral pattern can be maintained over the whole disk. Indeed, an explicit formula, equation (14), for its approximate description has been found.[1]

It is almost certain that the basic mechanism to be proposed here is related to the idea of "density waves" discussed by P. O. Lindblad (1960) and B. Lindblad (1961). Our analysis will, on the whole, tend to support their general conclusions. It is, however, somewhat difficult to make a detailed comparison between their ideas and ours, since the methods used are very different from each other. There is considerably more emphasis on individual stellar orbits in the Lindblads' theory, while the interaction of the stars over the disk as a whole is directly considered in our work through the use of the distribution function. This approach will enable us to exhibit certain spiral patterns very easily from general dynamical considerations (Sec. II). Some attempts at comparison will be made in Section IV. An investigation that further clarifies the relationship between these two approaches would no doubt enhance our understanding of the whole problem.

II. THE STABILITY PROBLEM OF THE GALACTIC DISK

It is natural to ask whether distribution of mass in the form of a thin rotating disk, such as that in our own Galaxy, is in a state of *stable* equilibrium. Instability can take the form of a warping of the disk or of motions in its plane. We shall concentrate on the latter mode as being more relevant to the problem at hand. In our idealized model, the mass is concentrated in an infinitesimally thin disk, with a surface density that is approximately equal to the projected density in the galactic plane. The gravitational forces are in balance with appropriate circular and random velocities. Such a representation necessarily involves a high surface density in the central region, where a typical galaxy has a bulge due to the large three-dimensional velocity dispersion of the stars. Thus, it will often be found convenient to have a singularity of the density distribution at the center, provided the total mass is finite.

To investigate the stability of such a disk, we may start with the basic equations of stellar dynamics expressed in terms of the distribution function in phase space (Chandrasekhar 1960). But it is sometimes convenient to use the infinite set of equations for the (tensorial) moments of the velocity components, and to terminate them by imposing a suitable approximation. As a first step, we shall neglect the velocity dispersion altogether.[2] One then has the following set of equations:

$$\mu_t + r^{-1}[(r\mu u)_r + (\mu v)_\theta] = 0\,, \tag{1a}$$

$$u_t + uu_r + (v/r)u_\theta - v^2/r = \phi_r\,, \tag{1b}$$

$$v_t + uv_r + (v/r)v_\theta + uv/r = \phi_\theta/r\,, \tag{1c}$$

$$\phi_{rr} + \phi_r/r + \phi_{\theta\theta}/r^2 + \phi_{zz} = -4\pi G\mu(r, \theta)\delta(z)\,, \tag{1d}$$

where the cylindrical system of coordinates (r, θ, z) is used, μ is the surface density, and ϕ is the negative of the gravitational potential. The first three equations are restricted to the plane $z = 0$, while equation (1d) is for the *three*-dimensional space. For the initial state of equilibrium, we consider a solution of the form $\mu = \mu_0(r)$, $u = 0$, $v = V(r) = r\Omega(r) > 0$. We then consider a disturbance from the equilibrium state; e.g., $\mu = \mu_0(r) + \mu'(r, \theta, t)$. As in all stability problems, we begin by considering the case of small dis-

[1] Preliminary results obtained by detailed numerical calculations carried out by Mr. Ronald Rehm confirm the validity of the formula and the general conclusions.

[2] In this approximation, the gaseous component of the galaxy is naturally included. We simply regard the density, μ, below as the *total* density of matter including both the stars and the gas.

turbances. The equations governing the quantities $\mu'(r, \theta, t)$, etc., can then be linearized.[3] We obtain

$$\mu'_t + \Omega\mu'_\theta + \frac{1}{r}(r\mu_0 u')_r + (\mu_0/r)v'_\theta = 0\,, \tag{2a}$$

$$u'_t + \Omega u'_\theta - 2\Omega v' = \phi'_r\,, \tag{2b}$$

$$v'_t + \Omega v'_\theta + (\kappa^2/2\Omega)u' = \phi'_\theta/r\,, \tag{2c}$$

$$\phi'_{rr} + \phi'_r/r + \phi'_{\theta\theta}/r^2 + \phi'_{zz} = -4\pi G\mu'\,\delta(z)\,. \tag{2d}$$

In the above equations, κ is the epicyclic frequency given by

$$\kappa^2 = 4\Omega^2[1 + (r/2\Omega)(d\Omega/dr)]\,. \tag{3}$$

The set of linearized equations (2a–d) admits solutions of the type

$$\mu' = Re\{\mu^{(1)}(r)\exp[i(\omega t - n\theta)]\}\,, \qquad \omega = \omega_r + i\omega_i\,, \tag{4}$$

where n is an integer (which may be taken to be positive without loss of generality), and ω_r and ω_i are the real and imaginary parts of the parameter ω. Unstable modes are given by $\omega_i < 0$.

Under such a general formulation, ω takes on characteristic values for the solution of a set of linear integro-differential equations in the single independent variable r, subjected to suitable boundary conditions at $r = 0$ and as $r \to \infty$. However, since we would admit a singularity at $r = 0$ in the solution, if we adopt the disk model right to the center of the galaxy, and the range of r is infinite, the set of characteristic values may be expected to have a *continuous spectrum* (in addition to any discrete part). Physically, this means that we are looking for a representation of the situation in the disklike part of the galaxy, and leaving the conditions near the central bulge, where the random velocities are large and three-dimensional, to adjust themselves to almost any requirement of the disk part. We also expect that all the disturbances would die off at infinity in a reasonable manner.

The solution (4) is clearly of the nature of a *density wave*. Indeed, it generally has a *spiral form*. To see this, let us write (as we can always do),

$$\mu^{(1)}(r) = S(r)\exp[i\Phi(r)]\,, \tag{5}$$

where $S(r)$ and $\Phi(r)$ are *real*. Then equations (4) give

$$\mu'(r, \theta, t) = S(r)e^{-\omega_i t}\cos[\omega_r t - n\theta + \Phi(r)]\,. \tag{6}$$

If $S(r)$ varies *slowly* with the radial distance r while $\Phi(r)$ varies *quickly*, then equation (6) gives a spiral impression in the density distribution at any instant of time, the form of the spiral being

$$\theta = \frac{1}{n}[\Phi(r) + \text{const.}]\,. \tag{7}$$

There are n arms in the spiral. These are *trailing spiral arms* if $\Phi'(r) < 0$, and leading ones if $\Phi'(r) > 0$. Note that we have taken $\Omega(r)$ and n to be positive by convention. By comparing equation (7) with observed two-armed galactic spirals, it is easy to see that Φ should change by an order of 4π over a typical radial distance in such cases as the whirlpool nebula.

[3] It is suggested that this is justifiable for our present purposes, even though the random velocities neglected might be comparable to the plausible amplitude of the disturbance, since we are dealing with the *co-operative* effect of grouping of stars, which acts in a manner quite different from the effect of velocity dispersion (see also Sec. IV).

III. AN ASYMPTOTIC SOLUTION

The stability problem formulated above is very similar to those in ordinary hydrodynamics.[4] It has, however, the additional mathematical difficulty that, after the variable z is eliminated, $\phi'(r, \theta, 0)$ and $\mu'(r, \theta)$ are connected by an integral relationship. Fortunately, this difficulty disappears in an asymptotic approximation based on the rapid change of the phase factor $\Phi(r)$ mentioned at the end of the last section. In this case, we may attempt to integrate the Poisson equation

$$\phi_{rr}{}^{(1)} + \phi_r{}^{(1)}/r - n^2\phi^{(1)}/r^2 + \phi^{(1)}{}_{zz} = -4\pi G\mu^{(1)}(r)\delta(z) \, , \tag{8}$$

by an asymptotic process.[5] We write

$$\mu^{(1)}(r) = S(r) \exp [i\lambda f(r)], \text{ and} \tag{9}$$

$$\phi^{(1)}(r, z) = \psi(r, z, \lambda) \exp [i\lambda h(r, z)] \, , \tag{10}$$

where all the functions involved may be complex, λ is a large real parameter introduced for the convenience of the formal asymptotic procedure, and $\psi(r, z, \lambda)$ is assumed to have an expansion in inverse powers of λ. In effect, we are looking for asymptotic solutions of the Laplace equation for $z > 0$ and $z < 0$, bounded at infinity (in a suitable sense for λ large), continuous at $z = 0$, and fulfilling the jump condition $[\phi_z] = -4\pi G\mu^{(1)}(r)$ at $z = 0$ that can be easily derived from equation (8). Such a solution leads to the simple result that, to a first approximation,

$$\phi_r{}^{(1)} = 2\pi i G\epsilon\mu^{(1)} \, , \qquad \phi_\theta{}^{(1)} = O(\mu^{(1)}/\lambda) \, , \tag{11}$$

where $\epsilon = \pm 1$ according to the sign of the real part of $f'(r)$.

Within this approximation, it is very easy to find solutions of the type (5) for the system of equations (2a–d). Our first aim is to determine the phase factor, which turns out to be

$$\lambda f'(r)\epsilon = [\kappa^2 - (\omega - n\Omega)^2]/2\pi G\mu_0 \, . \tag{12}$$

Thus, a solution of the type considered is possible only if the real part of $\kappa^2 - (\omega - n\Omega)^2$ is positive, i.e., if

$$\kappa^2 + \omega_i{}^2 - (\omega_r - n\Omega)^2 > 0 \, . \tag{13}$$

The sign of $\Phi'(r) = Re[\lambda f'(r)]$ can, however, be positive or negative. Thus, *both leading and trailing arms are permitted.* The choice between the two might be resolved only after the solution is found (e.g., by numerical integration) in the ranges where the inequality (13) is violated. In those ranges, the solution is expected to have a behavior different from that indicated above.[6] In particular, the condition at the center of the galaxy deserves special attention.

Another effect that might distinguish between leading and trailing arms is differential rotation. Although density waves are propagated primarily by gravitational forces, they would be modified by differential rotation, when the non-linear terms, omitted from equations (2a–d), are included. The effect might be analogous to that of fluid motion

[4] The similarity might hold not only for problem of initial instability, but also for the subsequent process of development into a quasi-stationary final state, with the random velocities of the stars and of the interstellar-gas clouds supplying the smoothing effects that limit the growth of the waves. Differential rotation corresponds to the shear flow which can supply the energy to the oscillations and random motions (cf. Sec. IV).

[5] For details see Appendix.

[6] Toomre (1964) carried out numerical calculations for axisymmetrical disturbances over the whole range of disk radii and found that the solution drops off roughly in an exponential manner at infinity.

that leads to the distortion of acoustic waves. In that case, a density decrease in the direction of wave propagation tends to be accentuated into a compression shock, whereas a density decrease would tend to be smoothed out by the motion of the fluid. Thus, it is conceivable that only trailing waves are stable in the presence of non-linear effects.

In the case of axisymmetrical disturbances, $n = 0$, the formula (12) reduces to one already obtained by Toomre (1964) for a local wavenumber in terms of an assumed frequency. The present asymptotic approach can also yield the slowly varying amplitude distribution over the disk, which is not available from the local theory. The importance of this approximate analysis over the whole disk will be seen in Section IV.

The present study shows that non-axisymmetrical disturbances can propagate around the disk without change of shape even *in the presence of differential rotation.* Indeed, for the range $r_1 < r < r_2$ in which expression (13) applies, the geometrical form of the spiral pattern is found from equations (7) and (12) to be given by

$$n(\theta - \theta_0) = -\int_{r_0}^{r} (2\pi G\mu_0)^{-1}[\kappa^2 + \omega_i^2 + (\omega_r - n\Omega)^2]\,dr\,. \tag{14}$$

The end points of this range, $r = r_1$ and r_2, correspond to the points of *local gravitational resonance* in the neutral case (cf. Sec. IV). As a typical example, if we refer to the density data given by Schmidt (1956) and the velocity data given by P. O. Lindblad (1960, Table 2) on the basis of Schmidt's model, and take $\omega_r = 20$ km/sec kpc and $\omega_i = 50$ km/sec kpc (which is somewhat less than the value of κ at 5 kpc from galactic center), we would have a wavelength of the order of 2–3 kpc in our neighborhood. The lower limit r_1 would be at about 2 kpc from the galactic center, while the upper limit r_2 would disappear altogether. Such a mode amplifies rather rapidly in the absence of velocity dispersion. Its expected behavior in the presence of velocity dispersion will be discussed in Section IV.

The contrast between the spiral patterns of Sa, Sb, and Sc galaxies can also be brought out analytically by equation (14). If there is a comparatively greater concentration of mass in the center, the density μ_0 of the disk part is relatively smaller. Equation (14) then predicts tighter spirals, as indeed are observed in Sa galaxies. More even distribution of matter corresponds to loosely wound spirals, as observed in Sc galaxies.

IV. HYPOTHESIS OF QUASI-STATIONARY SPIRAL STRUCTURE IN THE SPATIAL DISTRIBUTION OF STARS

The above analysis suggests that, in the absence of velocity dispersion of the stars, there are many possible spiral patterns of density modifications in a basically axisymmetrical disk of gravitating matter. Those modes which are strongly unstable may be expected to bring themselves into prominence. On the other hand, there is considerable velocity dispersion among the stars, and it is to be expected that all gravitational instabilities would tend to be smoothed out by their effects. Two possibilities then exist. The first is that all the modes of gravitational instability are suppressed. The second possibility is that the two tendencies might balance each other for certain modes, or a *group* of modes, in such a manner that a stable or nearly stable pattern might be maintained over long periods of time. Modes which are sufficiently different from this select group would, however, be suppressed.

It is not easy to examine the various factors that would enter into the decision between these two possibilities. A first attempt could be made by examining whether the velocity dispersion present is sufficient to suppress all the waves according to a *local* theory. Such a calculation has been carried out by Toomre (1964) for the solar neighborhood, and he reached the conclusion that an rms velocity dispersion σ_u of the order of 20–30 km/sec would be needed for the suppression of all the waves. This amount of velocity dispersion is actually present, and he concluded that the solar neighborhood is locally stable. However, if one would adopt the same method of analysis for a location at 4–5 kpc from the galactic center, the value of σ_u required would be about 70–90 km/sec (if we adopt Schmidt's model [1956] for our Galaxy). If such a high-velocity dispersion were actually present, it would imply that a considerable number of stars with high radial velocities would reach our neighborhood from the interior part of the Galaxy, contrary to observational evidence. Thus, that part of the disk is locally unstable.

This picture of a galactic disk, which is in part stable and in part unstable according to a *local* theory, would suggest the possibility of a balance resulting in a neutral density wave extending over the *whole* disk and having a scale of the order of (but smaller than) the distance between the stable and the unstable regions.[7]

The question may be raised as to why the velocity dispersion might cease to grow to the extent needed for complete stabilization. One possible reason is that large-scale waves are ineffective in producing local velocity dispersion. Furthermore, it may be conjectured that these large-scale waves are stabilized at small but finite amplitudes[8] by the joint effects of velocity dispersion and by the reaction of the gas together with its associated magnetic field (Chandrasekhar and Fermi 1953).

We now venture to suggest that there are indeed such large-scale neutral (or nearly neutral) waves for most of the disk galaxies, and formulate our ideas in the form of the following hypothesis:

The total stellar population, which has various degrees of velocity dispersion, forms a *quasi-stationary spiral structure* in space of the general nature discussed above. This is primarily due to the effect of gravitational instability as limited by velocity dispersion (and secondarily to the influence of the gas and the magnetic field). The extent of density variation in the spiral pattern may be only a small fraction of the symmetrical mean density distribution.

In advancing this hypothesis, we are not ignoring the fact that transient local patterns of a general spiral form are relatively easy to produce in a system in rotation. However, it seems rather difficult to account for a *coherent pattern over the whole disk* without bringing into account the co-operative effects of long-range gravitational forces.

If we accept the above hypothesis and follow the line of reasoning that led to it, the following inferences and conjectures may be made.

1. Due to the prevailing spiral gravitational field, all components of the galaxy, including the gas and the young stars, should form similar spiral patterns on the scale of the radius of the disk, whatever other secondary processes there might be. Different stellar components, with different extent of velocity dispersion, should lie on somewhat different spirals (cf. Zwicky 1957) and exhibit different degrees of unevenness. As an extreme case, one may even attempt to analyze the behavior of the gas, which has very little pressure, as a separate system (including possibly an appropriate magnetic field), which is subjected to the action of a spiral gravitational field produced by the rest of the galactic population. Its density contrast may therefore be expected to be far larger than that in the stellar components, as is indeed known from observations. These discussions should be supplemented by a consideration of the reaction of the gas on the stellar population, as mentioned above.

2. It is known from observations that galaxies devoid of gas do not show prominent spiral patterns. This is consistent with the present proposal. The velocity dispersion in such galaxies is usually thought to be comparatively large, and the mechanism of instability discussed here may indeed be suppressed completely. Even if this were not so,

[7] Although such waves cannot be analyzed by a strictly local theory, it is still feasible to apply the asymptotic method described in this paper; for the method does allow for the interrelationship between the various regions in the higher approximations. An asymptotic analysis that includes the effect of velocity dispersion would be extremely desirable.

[8] Cf. Landau's theory of stable oscillations at small but finite amplitudes in the problem of hydrodynamic stability (in Landau and Lifschitz 1959).

and the old stars did form a spiral structure, the lack of gas and the young stars stemming therefrom might so impair the manifestation of the spiral structure that it would hardly be observable.

3. One simple method to account for the effect of velocity dispersion and the reaction of the gas is to adopt the Lindblads' concept of a fixed smooth disk of stars belonging to the disk population and Population II and a mobile disk of Population I (gas and young stars). Then the basic equations (1a–d) should be modified by the addition of a force $F(r)$ in equation (1b) to represent the fixed field, and the surface density μ would refer to that of Population I alone. Except for this change of the interpretation of μ, the above theory will remain unchanged.[9] The spiral pattern is still given by equation (14), but the wavelength is much shorter, since μ is reduced. However, there is no sharp demarcation between the fixed and the mobile components. One could include in the mobile part also some older stars with relatively low-velocity dispersion. Thus, if we choose $\omega_r = 20$ km/sec kpc, $\omega_i = 0$, the mobile component in our neighborhood should be set at about 25 per cent of the total mass[10] in order to get a wavelength of approximately 2 kpc. In contrast to Section III, we consider neutral waves here, since the stabilizing influence of velocity dispersion has largely been taken into account.

4. Both the Lindblads emphasized the approximate constancy of $\Omega - \kappa/2$ over a large part of the Galaxy. Our theory does not seem to be dependent on it. Instead, we have a certain kind of local resonance only at two distinct radial locations in the galaxy where $\kappa^2 - (\omega - n\Omega)^2 = 0$ in the case of neutral waves (and also approximately for slightly amplifying and damped waves). This condition clearly means that the traveling spiral gravitational field is in step with the local epicyclic motion. Indeed, near these points, one would expect the gas, even more than the stars, to react very strongly and possibly to acquire very large radial velocities. These radial velocities should be outward and inward over alternate sectors of extent π/n each. It would be very interesting to pursue this point further both theoretically and in relation to observational facts (e.g., the 3-kpc arm).[11]

We are indebted to Professors B. Strömgren and L. Woltjer for introducing this subject to one of us (C. C. L.) at a conference in Princeton, N.J., in 1961 and for many helpful discussions during the course of this study. We had the pleasure of collaborating with Dr. A. Toomre. Frequent discussions with him have been to our mutual benefit, although we have placed somewhat different emphasis in our work and have arrived at somewhat different conjectures. We also had the pleasure of discussing the problem with Professors M. Krook and D. Layzer and Mr. A. J. Kalnajs.

This work is partially supported by a grant from the National Science Foundation.

APPENDIX

ASYMPTOTIC SOLUTION OF POISSON'S EQUATION

We now consider, in some detail, the solution of the equation for gravitational potential corresponding to a surface distribution of density in the plane $z = 0$. Especially we wish to demonstrate an asymptotic method used in the text.

[9] Dr. Kevin Prendergast first mentioned to me (C. C. L.) the idea that the present analysis might be made in the presence of the gravitational field of a fixed disk during a discussion at the early stages of this study.

[10] This is not unreasonable, especially since we expect this region to be locally stable and forced to oscillate by the inner regions. It might even be too low an estimate in view of the possible existence of some molecular hydrogen as yet undetected.

[11] See Oort (1962), pp. 3–22, and esp. his reference to B Lindblad on p. 12.

Since we are working with a cylindrical coordinate system, the convenient form for the basic equation is

$$\frac{\partial^2\phi}{\partial r^2}+\frac{1}{r}\frac{\partial\phi}{\partial r}+\frac{\partial^2\phi}{\partial z^2}+\frac{1}{r^2}\frac{\partial^2\phi}{\partial\theta^2}=-4\pi G\sigma(r,\theta)\delta(z). \tag{A1}$$

In other words, we look for solutions of Poisson's equation for $z > 0$ and for $z < 0$ such that, at $z = 0$, there is a discontinuity in ϕ_z given by

$$\left[\frac{\partial\phi}{\partial z}\right]=\phi_z(r,\theta,0+)-\phi_z(r,\theta,0-)=-4\pi G\sigma(r,\theta). \tag{A2}$$

Since the problem is linear, we may consider the individual harmonic components in

$$\sigma=\sigma_n(r)\cos n\theta\,,\qquad \phi=F_n(r,z)\cos n\theta \tag{A3}$$

and obtain the general solution by superposition. We thus have the partial differential equation

$$\frac{\partial^2F_n}{\partial r^2}+\frac{1}{r}\frac{\partial F_n}{\partial r}+\frac{\partial^2F_n}{\partial z^2}-\frac{n^2F_n}{r^2}=-4\pi G\sigma_n(r)\delta(z) \tag{A4}$$

to be solved under the condition

$$\left[\frac{\partial F_n}{\partial z}\right]=-4\pi G\sigma_n\,. \tag{A5}$$

We consider the case where the variation of $\sigma_n(r)$ is sufficiently rapid in some sense. That is, we consider a representation

$$\sigma_n(r)=Re\left[\sigma^{(1)}(r)e^{i\lambda f(r)}\right], \tag{A6}$$

where λ is large, while both $\sigma^{(1)}(r)$ and $f(r)$ vary but slowly in r. We may then expect a solution of the form

$$F_n(r,z)=Re\left[A(r,z,\lambda)e^{i\lambda\Phi(r,z)}\right], \tag{A7}$$

where $A(r, z, \lambda)$ depends on λ in the following asymptotic manner:

$$A(r,z,\lambda)=A^{(0)}(r,z)+\lambda^{-1}A^{(1)}(r,z)+\lambda^{-2}A^{(2)}(r,z)+\ldots \tag{A8}$$

The functions $\Phi(r, z)$, $A^{(0)}(r, z)$, . . . , etc., are slowly varying functions of r and z. The condition (A5) becomes the pair of relations

$$\left[\frac{\partial A}{\partial z}\right]+i\lambda A(r,0)\left[\frac{\partial\Phi}{\partial z}\right]=-4\pi G\sigma^{(1)}(r),\qquad \Phi(r,0)=f(r). \tag{A9}$$

We have imposed the conditions that $\Phi(r, z)$ and $A(r, z, \lambda)$ are continuous at $z = 0$.

Differential equations.—We now substitute

$$F=A(r,z,\lambda)e^{i\lambda\Phi(r,z)} \tag{A10}$$

into the equation (A4); we get

$$\begin{aligned}&\frac{\partial^2A}{\partial r^2}+\frac{\partial^2A}{\partial z^2}+\frac{1}{r}\frac{\partial A}{\partial r}-\frac{n^2A}{r^2}\\&+i\lambda\left[A\left(\frac{\partial^2\Phi}{\partial r^2}+\frac{\partial^2\Phi}{\partial z^2}+\frac{1}{r}\frac{\partial\Phi}{\partial r}\right)+2\left(\frac{\partial A}{\partial r}\frac{\partial\Phi}{\partial r}+\frac{\partial A}{\partial z}\frac{\partial\Phi}{\partial z}\right)\right]\\&-\lambda^2A\left[\left(\frac{\partial\Phi}{\partial r}\right)^2+\left(\frac{\partial\Phi}{\partial z}\right)^2\right]=0\,.\end{aligned} \tag{A11}$$

This leads immediately to the condition

$$\left(\frac{\partial\Phi}{\partial r}\right)^2+\left(\frac{\partial\Phi}{\partial z}\right)^2=0\,, \tag{A12}$$

which is satisfied by

$$\Phi=f_1(r+iz) \quad \text{or} \quad f_2(r-iz)\,, \tag{A13}$$

but not a combination of both. The second of the boundary conditions (A9) imposes the requirement that

$$f_1(r)=f_2(r)=f(r)\,. \tag{A14}$$

The choice of solution for $z>0$ and $z<0$ must be such that the solution decreases with $|z|$ increasing. Thus, we should have

$$\Phi=f(r+i\epsilon z) \quad \text{for} \quad z>0\,, \quad \text{and} \quad \Phi=f(r-i\epsilon z) \tag{A15}$$

for $z<0$, where $\epsilon=\pm 1$ accordingly as $f'(r)\gtrless 0$. The first condition (A9) becomes

$$\left[\frac{\partial A}{\partial z}\right]+i\lambda A\,(\epsilon 2i)\,f'(r)=-4\pi G\sigma^{(1)}(r)$$

or

$$\left[\frac{\partial A}{\partial z}\right]=-4\pi G\sigma^{(1)}(r)+2\lambda A\,|f'(r)|\,. \tag{A16}$$

Equation (A16) suggests that $A(r,z,\lambda)$ should be of the form

$$A(r,z,\lambda)=\lambda^{-1}[A^{(1)}(r,z)+\lambda^{-1}A^{(2)}(r,z)+\ldots]\,, \tag{A17}$$

and that $A^{(1)}(r,z)$ satisfies the condition

$$0=-4\pi G\sigma^{(1)}(r)+2A^{(1)}(r,0)\,|f'(r)|\,. \tag{A18}$$

With $\Phi(r,z)$ given in the form of expression (A13), equation (A11) can be somewhat simplified. Indeed, we get

$$\left(\frac{\partial^2 A}{\partial r^2}+\frac{\partial^2 A}{\partial z^2}+\frac{1}{r}\frac{\partial A}{\partial r}-\frac{n^2 A}{r^2}\right)+i\lambda f'(r\pm i\epsilon z)\left[\frac{A}{r}+2\left(\frac{\partial A}{\partial r}\pm i\frac{\partial A}{\partial z}\right)\right]=0\,. \quad (z\gtrless 0) \tag{A19}$$

Using the form (A17), the equation for $A^{(1)}(r,z)$ becomes

$$\frac{A_{\pm}^{(1)}}{2r}+\left[\frac{\partial A_{\pm}^{(1)}}{\partial r}\pm i\epsilon\frac{\partial A_{\pm}^{(1)}}{\partial z}\right]=0$$

or

$$\frac{\partial}{\partial r}[r^{1/2}A_{\pm}^{(1)}]\pm i\epsilon\frac{\partial}{\partial z}[r^{1/2}A_{\pm}^{(1)}]=0\,.$$

The solution of this is immediately seen to be

$$[r^{1/2}A_{\pm}^{(1)}]=g(r\pm i\epsilon z)\,, \quad (z\gtrless 0) \tag{A20}$$

where g is to be determined from the boundary condition (A18).

We have

$$r^{-1/2}g(r)=2\pi G\sigma^{(1)}(r)/|f'(r)|\,. \tag{A21}$$

Thus, to a first approximation, the surface distribution $\sigma^{(1)}(r)$ exp $[i\lambda f(r) - in\,\theta]$ gives rise to the potential

$$\phi(r,z) = A_{\pm}(r,z)\exp[i\lambda f(r \pm iz\epsilon) - in\theta], \qquad (z \gtrless 0) \quad \text{(A22)}$$

where $\epsilon = sgn\, f'(r)$,

$$A_{\pm}(r,z) = \lambda^{-1}r^{-1/2}g(r \pm iz\epsilon) \quad \text{(A23)}$$

and

$$r^{-1/2}g(r) = 2\pi G\sigma^{(1)}(r)/|f'(r)|. \quad \text{(A24)}$$

The force per unit mass in the plane $z = 0$ is given by

$$\phi_r(r,\,0) = 2\pi i G\sigma\epsilon, \qquad [\epsilon = \pm 1 \text{ according as } f'(r) \gtrless 0] \quad \text{(A25)}$$

where

$$\sigma = \sigma^{(1)}(r)\exp[i\lambda f(r) - in\theta].$$

Notice that n does not appear in the initial approximation for the amplitude functions. It appears only in the higher approximations.

It is understood in the above reasoning that $\sigma^{(1)}(r)$ and $f(r)$ are analytic functions. Otherwise, a more elaborate process is needed to justify the same results. See Shu (1963).

The higher approximations can be easily found by using the asymptotic representation (A17) for $A(r, z)$. It appears from the form of the solution that it holds only for a finite wedge ($|z/r| <$ const.) in the (z, r)-plane with the further restrictions that r should be finite, and that $f'(r)$ does not vanish over the range in question.

REFERENCES

Chandrasekhar, S. 1960, *Principles of Stellar Dynamics* (New York: Dover Publications), p. 184.

Chandrasekhar, S., and Fermi, E. 1953, *Ap. J.*, **118,** 113, 116.

Landau, L. D., and Lifschitz, E. M 1959, *Fluid Mechanics* (Reading, Mass.: Addison Wesley), p. 104.

Lindblad, B. 1961, *Stockholm Obs. Ann* , **21,** 8, esp. Sec. 5, "Spiral Structure."

Lindblad, P. O. 1960, *Stockholm Obs Ann.*, **21,** 4.

———. 1962, *Interstellar Matter in Galaxies*, ed. L. Woltjer (New York: W. A. Benjamin, Inc.), pp. 222–233.

Oort, J. H. 1962, *Interstellar Matter in Galaxies*, ed. L. Woltjer (New York: W. A. Benjamin, Inc), pp. 234–244.

Schmidt, M. 1956, *B.A.N* , **13,** No. 468, 15.

Shu, Frank, H. 1964, "Gravitational Instability and Spiral Structure in Disk Galaxies," thesis for B.Sc. in Physics, Massachusetts Institute of Technology.

Toomre, A. 1964, *Ap. J.*, **139,** 1217.

Zwicky, F. 1957, *Morphological Astronomy* (Berlin: Springer-Verlag), pp. 198–201.