On the dynamical basis of the classification of normal galaxies  
(density waves/galactic dynamics/spiral arms/spiral galaxies)

J. Haass, G. Bertin*, and C. C. Lin

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

Contributed by C. C. Lin, March 8, 1982

ABSTRACT 

Some realistic galaxy models have been found to support discrete unstable spiral modes. Here, through the study of the relevant physical mechanisms and an extensive numerical investigation of the properties of the dominant modes in a wide class of galactic equilibria, we show how spiral structures are excited with different morphological features, depending on the properties of the equilibrium model. We identify the basic dynamical parameters and mechanisms and compare the resulting morphology of spiral modes with the actual classification of galaxies. The present study suggests a dynamical basis for the transition among various types and subclasses of normal and barred spiral galaxies.

The starting point and some of the important final objectives of any theory of spiral structure in galaxies can be summarized by the questions posed by the morphological classification of Hubble, Sandage, and de Vaucouleurs (1) and by the luminosity classification of van den Bergh (2). The former scheme emphasizes the geometry of spiral structure, the properties of which correlate with a number of apparently independent physical parameters (such as gas content and bulge size). The frequent existence of a grand design poses the winding dilemma—i.e., the problem of how material objects (such as the young blue stars) can maintain their spiral arrangement in the presence of sizable differential rotation. The relevant questions are essentially related to the overall dynamics of self-gravitating disks. The latter classification scheme focuses on the distinction between broad and filamentary arms in different galaxies and on the related correlation with galaxy luminosity (thin-armed galaxies are brighter). The issues raised clearly bear on the mechanisms of shock generation and star formation in the interstellar medium in galaxy disks.

In the last two decades the density wave theory of spiral structure has provided a unifying working program for a number of theoretical and observational investigations on the subject. Specifically, the explanation of the existence of a grand design may be based on the idea that spiral arms are quasi-stationary wave patterns. The crucial basic assumption is that the long-range gravitational interaction among stars is behind the scenes and generates coherent collective phenomena. Even before a full dynamical understanding of the processes that lead to the excitation and maintenance of spiral structure was reached, the density wave theory had been supported by many successful predictions and comparisons with the observations (3, 4). In order to carry out the formidable task of investigating the dynamics of collisionless self-gravitating disks, three principal approaches have been followed: (i) N-body simulations (see refs. 5–7), (ii) stellar dynamical calculations (see refs. 8–14), and (iii) fluid models (see refs. 15–20). From the very beginning (5–7), it has been clear that axisymmetric disk models can be subject to instability with respect to regular global modes and thus tend to reach an evolving nonaxisymmetric spiral-like structure.

The hypothesis of quasi-stationary spiral structure stressed from the outset the importance of coexistence and superposition of different spiral features when interpreting the presence of a grand design of spiral structure in spite of some degree of irregularity. But now that unstable discrete spiral modes have been found to be supported in some realistic galaxy models (8, 15) and an evaluation of their implications has been made (21), the coexistence of spiral modes in galactic disks has assumed more importance in comparing theory with observations. In the present paper, we discuss and illustrate an attempt at a dynamical approach to the classification of spiral galaxies via the modal approach, which has yielded a particularly rich set of interesting results.

Modes, coexistence, and grand design of spiral structure

Not all spirals present a neat grand design. An explanation is that the more regular the observed spiral structure is, the smaller the number of spiral modes that should be involved in its dynamics. Certain observed galaxies should be essentially represented by a single mode, and in some cases a nonaxisymmetric spiral grand design (such as the galaxy D32) rules out any tidal interpretation. One remarkable case in which a single wave pattern study has produced a detailed qualitative and quantitative test (23) is that of M81. Recently, H. Visser and J. Haass* have been able to provide dynamical support to the successfull observational test by calculating a self-excited spiral mode for a model of M81 in agreement with the observed pattern and with the theoretical pattern used by Visser (23).

However, the observed structure is very often less regular, even if retaining a global two-armed appearance (as is the case of the galaxy D100† that is characterized by a spur coexisting with an otherwise fairly regular two-armed grand design). Haass’s example (Fig. 1) on the linear superposition of three global spiral modes readily shows that the effects of coexistence are likely to play a major role in the comparison of the theory with observations. On the other hand, it would be wrong to conclude that the presence of a number of modes greater than three already implies the fuzzy structures of the NCC2941 type. In fact, a linear superposition of three regular two-armed modes always leads to a coherent (though in general evolving in time) two-armed grand design (see ref. 21 for additional discussion).

Discrete unstable spiral modes: Comparison with morphology of galaxies

In this section, we describe the properties of the dominant two-armed modes in different galaxy models and discuss the possibility of a dynamical approach to the morphological classifi-

* Permanent address: Scuola Normale Superiore, 56100 Pisa, Italy.
† We are grateful to V. Rubin for bringing these galaxies to our attention. Indeed, she pointed out the resemblance between D100 and Fig. 1.
‡ Another interesting example of a nonaxisymmetric grand design spiral is NGC309 (see ref. 22).
§ The spiral pattern was published in ref. 21.
The modes discussed above do not exhibit inner Lindblad resonance. For models in which the modes may include such resonance, a proper treatment of the dynamics of stars should be included. Preliminary results for a sequence of models with variation of $\Theta(r)$ exhibit a gradual approach to the resonance condition. As this happens, the dominant modes show a substantial variation in pitch angle within the disk, the inclination being smaller in the central regions where the resonance condition is approached.

Comparison with the Morphological Classification of Galaxies. From the above results (and other material that is not presented in this article) we can draw many interesting conclusions and organize a dynamical approach to the classification of spiral galaxies.

The examples of Fig. 2 and Fig. 3 are indicative of two morphological transitions that are easily recognized in the Hubble diagram, from tightly wound (Sa) to more open (Sc) spirals and from normal (SA) to barred (SB) spiral galaxies through the intermediate SAB class. However, there are complications when we attempt to compare our dynamical framework to actual observed objects. As an example, we may recall that the two well-known galaxies M33 and NGC5364 are both classified as Sc, whereas their contrasting appearances imply that different dynamical phenomena are involved.

Even in our simple systematic scheme to model galaxies, it is easy to realize that, in order to adjust to the proper observational criteria that are commonly used in the morphological classification of galaxies, it would have been necessary to vary all three of the controlling parameters simultaneously, rather than change one parameter at a time. In fact, to reproduce the Hubble sequence, the shape of the rotation curve and the related bulge-halo size and structure should define the optimal paths to be followed in the three-dimensional parameter space. With this perspective we have chosen to take the view of defining a (somewhat independent) dynamical approach to the classification of galaxies and then comparing it with the existing observational classification schemes.

To be sure, galaxies with massive bars are not covered in the present linear modal approach. Still some barred (SB) and especially transition (SAB) galaxies may well turn out to be within our reach, as is partly hinted at by recent photometry of the older disk population (22).

In addition, our approach provides some other attractive suggestions, which touch on issues more subtle than those at the basis of the original Hubble program. For instance, we do see a distinction between normal spiral modes that do not reach the center (such as the one applicable to M81) and normal spiral modes that almost reach the center with patterns that resemble that of M51. Another interesting case has been found by exploring the role of the rotation curve. The results obtained tend to support the idea that rapidly rising rotation curves may be compatible with a situation of dual spiral structure (21), even though much more work has to be done in this direction. Finally, we may ask whether the transition from spiral galaxies with a grand design (like M81) to fuzzy spirals (like NGC2841) is within our dynamical program. This may be the case, but other explanations of fuzzy spirals should not be ruled out.

Dynamical mechanisms and further discussions

Much theoretical work (see refs. 5–21 and 24–30) has been directed towards the understanding of the mechanisms for the excitation and maintenance of large-scale spiral structures. This effort has benefited substantially from the flexible asymptotic approach (see refs. 8 and 15), which has identified the properties of elementary spiral waves and their role in certain global modes (see ref. 21). To extend our ability to calculate discrete global
spiral modes beyond the asymptotic theory, a numerical code (19) has been devised. To understand the dynamical mechanisms that sustain these modes, we adopt a spectral analysis (see ref. 21), with results such as those shown in Fig. 2 and Fig. 3. Such an analysis supports the view that the global structure of a mode is maintained by a proper superposition of four different kinds of density waves: leading and trailing waves, and, in each case, short and long waves.

Our present results support the following conclusions (additional discussion is given in ref. 21): (i) Normal spiral modes are essentially composed of two trailing spiral waves with opposite propagation properties inside of the corotation circle. (The inward-propagating short trailing wave is refracted back into an outward-propagating long trailing wave from the central regions in the general neighborhood where the nuclear bulge first appears.) The growth of these wave patterns is provided by a wave amplification by stimulated emission of "radiation" (WASER) mechanism (11), which is found to be aided by the presence of sizable tangential forces (14, 17). (ii) Unlike normal spiral modes, barred spiral modes include short leading waves in addition to trailing waves of both types. The incoming short trailing wave can reach the center and be reflected into a short leading wave most effectively if the mass distribution does not have a deep "hole" in the middle. The amplification of barred modes can still be interpreted in terms of a WASER mechanism—i.e., super-reflection induced by "radiation" of energy towards the outer regions of the disk. We note that, in the bar structures obtained, there are present not only a superposition of trailing and leading waves (in a way reminiscent of the "anti-spiral" theorem), but also a sizable component of very long "spoke-like" waves $k = 0$. (iii) The transition from normal to barred spiral modes is associated with different weights of short leading and long trailing components in the modal composition; both types contain short trailing waves that give galaxies the overall impression of trailing spirals.

We would like to reiterate that the transient approach (9, 16, 18, 25) does not address physical phenomena different from those discussed in the modal approach (8, 15). Only the emphasis is different and the regimes actually investigated do not always overlap in the parameter space. The relationship between these two approaches has been discussed in some detail in ref. 21. It is pointed out that eventually both approaches, as studied by local or asymptotic methods, need some extrapolation when used to describe truly open two-armed spiral features. The main reason for this limitation is the inherent long-range nature of the gravitational forces that dictate a global behavior of inhomogeneous collisionless stellar systems. But decisive progress has been made towards the understanding of open structures despite such restrictions.

Even though great progress has been made in developing a theory of spiral structure, many are the intrinsic limitations of the simplified models that are generally used in the calculation of spiral modes. However, qualitative understanding of specific issues is already available. Among these, we should recall the effects of finite thickness (see refs. 26–28) and the role of stellar dynamical effects on global modes. These factors, besides mod-
Fig. 3. Transition from normal to barred spiral modes. The same spiral mode is followed in a one-parameter sequence of equilibrium models, characterized by different properties of the nuclear bulge. The top frames illustrate density contours, the bottom frames the corresponding power spectra. The distribution of the stability parameter \( Q(r;q) \) is given by Eq. A3 in the Appendix, with \( q = 0.35, 0.15, \) and \( 0.05 \), respectively, for the three figures from left to right. The parameters in Eq. A1 and A2 are \( v = 1.75, \mu = 0 \).

ifying quantitative predictions on spiral modes, may play the crucial role of removing the singularities (and hence the related continuous spectra) that are present in more idealized models (21, 29).

One limitation of our approach is that we are relying on results based on a linear theory. The role of nonlinearities in spiral structure has long been discussed (e.g., see ref. 30 and comments and references in refs. 8 and 15). Some support for the relevance of linear theories derives from the fact that observed spiral features often have relatively low amplitude. Still, nonlinear effects at resonances have been claimed to play an important role (e.g., see ref. 30). Here we may remark that the phenomenon of trapping and the stability of trapped orbits (8) are not so obvious. In fact, the very coexistence of linear modes with small or large differences in pattern speeds may eliminate nonlinear trapping altogether, thus inhibiting the so-called nonlinear saturation at resonances.

Appendix

The family of galaxy models used in the present study essentially originates from the model C3 of ref. 13 and is characterized by three parameters: \( \mu, q, v \). These appear in the definition of the model in the following way. The angular rotation speed \( \Omega(r;v) \) in the disk required to balance the gravitational pull at radius \( r \) is given by

\[
\Omega^2(r;v) = \frac{16}{105} \left[ \frac{M_\odot}{a_1^2} \right] \frac{M_\odot}{v}\left[ 1 + \left( \frac{r}{a_1} \right)^2 \right]^{-3/2}, \quad [A1]
\]

in which

\[
H(x_1) = \frac{59.0625}{x_1^{11}} + \frac{26.25}{x_1^{10}} + \frac{16.6875}{x_1^9} + \frac{11.25}{x_1^8} + \frac{6.5625}{x_1^7}.
\]

and \( x_1 = \{1 + (r/a_1)^2\}^{1/2} \);

\[ a_1 = 12, M_1 = 2 \times 10^{10}, M_0 = 5 \times 10^9. \]

The active disk mass density \( \sigma(r;\mu) \) is:

\[
\sigma(r;\mu) = \frac{1}{a_2^2 x_1^{-11}} \frac{\mu}{20a_2^3} \left[ 1 + \left( \frac{r}{a_2} \right)^3 \right]^{-11/2} \frac{4.5}{\pi \mu} M_1, \quad [A2]
\]

with \( a_2 = 8 \). Finally, the stability parameter distribution \( Q(r;q) \) in the disk is specified by:

\[
Q(r;q) = 1 + q \exp[-r^2/a_3^2], \quad [A3]
\]

with \( a_3 = 1.45 \). The reference model \( S^0 \) is characterized by \( v = 1.75, \mu = 0, q = 1 \). In the above definitions \( G \) is the gravitational constant. In order to have numbers applicable to some realistic galaxies one may consider the (dimensional) lengths \( a_1, a_2, a_3 \), and \( v \) as measured in kiloparsecs, the masses \( M_1 \) and \( M_0 \) as measured in solar masses. We recall that the stability of a galactic disk is essentially determined by three functions: the rotation curve \( \Theta(r) = rK(r) \), the disk mass density \( \sigma(r) \), and the dispersion speed profile \( c(r) \). The rotation curve measures the total mass of the galaxy and provides the local values of the epicyclic frequency \( \kappa(r) \) and the shear rate \( d \ln \Omega/d \ln r \). The disk density \( \sigma(r) \), for a given \( \Theta(r) \), determines the global disk to halo mass ratio and the local typical scalelength of the disk—i.e., the dimensionless quantity \( e_\delta(r) = \pi G\sigma(r)R^2 \). For a spiral distur-
ance with m arms, the above functions completely determine the profile of the parameter \( j = m \varepsilon_0 (4 \Omega/\kappa) d \ln \Omega/d \ln r^{1/2} \), which plays a crucial role in the local stability and propagation properties of nonaxisymmetric waves (17). For a given \( \Theta(r) \) and \( \sigma(r) \) the "pressure" profile \( c(r) \) fixes the behavior of the local stability parameter \( Q = c \kappa/\pi G \sigma \). In fact, on dynamical grounds, we chose to specify directly \( Q \), rather than \( c \). Note that the parameters \( \mu \), \( q \), and \( v \) can be varied independently, changing one dynamical function at a time.

This work was partly supported by the National Science Foundation and partly by the Consiglio Nazionale delle Ricerche, under the United States–Italy Cooperative Science Program.