

THE SPIRAL STRUCTURE OF GALAXIES

E. ATHANASSOULA

*Observatoire de Besancon, 41bis Av. de l'Observatoire, 25000 Besancon,
France*

and

*Observatoire de Marseille, 2 place Le Verrier, 13248 Marseille Cedex 04,
France*



NORTH-HOLLAND PHYSICS PUBLISHING-AMSTERDAM

THE SPIRAL STRUCTURE OF GALAXIES

E. ATHANASSOULA

Observatoire de Besançon, 41bis Av. de l'Observatoire, 25000 Besançon, France

and

Observatoire de Marseille, 2 place Le Verrier, 13248 Marseille Cedex 04, France

Received May 1984

Contents:

1. Introduction	321	9. Swing amplification	353
2. Morphology of spiral galaxies	322	10. Feedback cycles	357
2.1. Classification schemes	322	10.1. WASER	357
2.2. Grand design and flocculent spirals	323	10.2. Swing amplifier	358
2.3. Further properties of arms	325	11. Driving by companions	361
2.4. Bars	327	12. Driving by bars	363
2.5. Rings	329	13. Driving by asymmetries	365
2.6. Lenses	332	14. Modes	366
2.7. Dust lanes	333	15. <i>N</i> -body simulations	370
15.1. Stability mechanisms	371	15.2. Stability criteria	375
3. Basic properties of spiral galaxies	335	15.3. Non-linear properties of bars	377
4. Introduction to density-wave theory	338	15.4. Spiral structure	378
4.1. Dynamical background	338	16. Orbit and velocity fields in barred galaxies	379
4.2. Quasi-stationary density waves	338	17. Rings	385
4.3. The spiral	339	18. Stochastic self-propagating star formation	386
4.4. Orbits and resonances	340	19. Observational constraints	388
4.5. Kinematic quasi-stationary spirals	343	19.1. Rotation curves and mass distributions	388
5. Tightly-wrapped approximation	344	19.2. Location of resonances	389
5.1. The integral equation	344	19.3. Distribution of spiral tracers	389
5.2. The Lin-Shu-Kalnajs dispersion relation	344	19.4. Surface photometry	390
5.3. Axisymmetric stability	345	19.5. Velocity fields	391
5.4. Properties of the solution	346	19.6. Driving by companions	395
5.5. Velocity perturbations	347	20. In way of a conclusion	396
6. Shocks	348	References	398
7. Radial wave transport	349		
8. Shearing sheet	350		

Single orders for this issue

PHYSICS REPORTS (Review Section of Physics Letters) 114, Nos. 5 & 6 (1984) 319–403.

Copies of this issue may be obtained at the price given below. All orders should be sent directly to the Publisher. Orders must be accompanied by check.

Single issue price Dfl. 52.00, postage included.

1. Introduction

Spiral galaxies are often spectacular and pretty looking objects, and as such capture the imagination of many astronomers. The problems of the occurrence, formation and maintenance of spiral structure continue to be popular, and a vast literature exists on the subject.

The shapes and stellar content of the spiral arms vary widely from one galaxy to the other and the early work by Hubble brought out already the systematic variation of the form of the arms with other physical properties of the parent galaxy. Yet the question of how exactly the arms come about ran swiftly into a serious dilemma as more quantitative data on the detailed structure of spiral galaxies, including our own, became available [e.g. 206]. In the majority of cases the regions of galaxies containing the arms rotate differentially, so that if the arms are always composed of the same stars and gas they will wind up tighter and tighter in a timescale much shorter than the presumed lifetime of the galaxy. It is then difficult to reconcile the observed differential rotation with the large number of relatively open spirals.

If the spiral arms are not long lived material structures, then what are they? Several workers, starting with B. Lindblad, proposed that they are in fact the manifestation of density waves in the disk. In a long series of papers [170, and references therein] Lindblad exposed his theory, based on the idea that gravity can account for the formation of the spirals and that the observed arms are the crest of a non-evolving, uniformly rotating pattern. His ideas were not readily accepted by the astronomers of his time, one reason being his emphasis on leading spirals (i.e. with arms unwinding in the sense of rotation). Indeed, on the basis of rotation data and the occurrence of primary absorption lanes in galaxies seen nearly edge-on it was shown [113, 60] that spiral arms were trailing (i.e. they have arms that wind up in the sense of the rotation). On the contrary, the next attempt along this line, by C.C. Lin and his collaborators, received a very hearty welcome. They “regard 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 speed” (from [168]).

This hypothesis, coupled with several other simplifying assumptions, paved the way towards the construction of stellar density-wave models and to calculations of the gas flow in such models which in turn could be compared with the observations. Yet there is no convincing proof (nor for that matter disproof) that spirals are indeed long-lived. And when it was realized that tightly wound wave packets propagate radially and are damped in timescales of 10^9 yrs [269], this approach was questioned and other attempts to understand spiral structure, of a varied nature, sprang up.

Now, in parallel to the still continuing search for long-lived modes or to the consideration of spirals driven e.g. by stationary bars, come theories for grand design yet short-lived spirals triggered by companions and for more raggedy spirals as responses to large clumps of matter. All these attempts still assume gravity to be the dominant force. Yet even this assumption has sometimes been abandoned, when spiral structure has been attributed to the collective effects of star formation on galactic scales. On the observational side, more quantitative data became available as whole research programmes were devoted to the topic.

Thus the quest for unravelling the origin of spiral structure became as tangled as some of the spiral patterns themselves, with many individual contributions often a step forwards, but sometimes a step backwards. We will attempt here to review the situation as it stands now both in the theoretical and the observational fields. As A. Toomre wrote [272]: “Happily this remains a subject where it makes sense to start almost at the beginning”, and we propose to do just that. Before elaborating the theory, however, we will first present an account of the morphological structures we seek to explain, and the basic astrophysical data which may be used as input.

2. Morphology of spiral galaxies

In this section we summarize the main morphological characteristics of spiral galaxies, before turning to the various explanations theory can offer.

2.1. Classification schemes

There is a large variety of forms and shapes of spiral arms, and several classifications have been based on their characteristic appearances. Figure 1 shows Hubble's now classical tuning fork diagram, with the ellipticals, S0's, and the two families, ordinary (A) and barred (B) spiral galaxies, and three types (a, b, c). As explained in detail in the Hubble Atlas of Galaxies [227], the division according to type is based on the following criteria:

- the winding of the spiral arms,
- the degree of resolution of the arms into stars and HII-regions (knottiness or patchiness),
- the relative size of the unresolved central region.

For most galaxies these independent criteria are in good agreement with each other, but in some cases a conflict arises and then priority is set to the arm characteristics. The Hubble classification has been used very extensively for thousands of galaxies [e.g. 66, 232], and this fact demonstrates the merit of qualitative classification schemes. If one tries to quantify each characteristic by itself, however, one finds a relatively large intrinsic scatter within each Hubble type. As shown in [143], the mean pitch angle increases from $6^\circ \pm 3^\circ$ for Sa's to $18^\circ \pm 5^\circ$ for Sc's. And only a loose correlation exists between pitch angle and Yerkes type [143], which measures the degree of central light concentration and thus the relative size of the unresolved central region.

A second classification of disk galaxies, based on the degree of development of the spiral arms, has been proposed in [289]. Galaxies with long, thin, well developed arms of relatively high surface brightness can be distinguished from those with undeveloped, stubby arm patches. This seems to be in general correlated with the luminosity of galaxies, hence the name luminosity classification. Here also

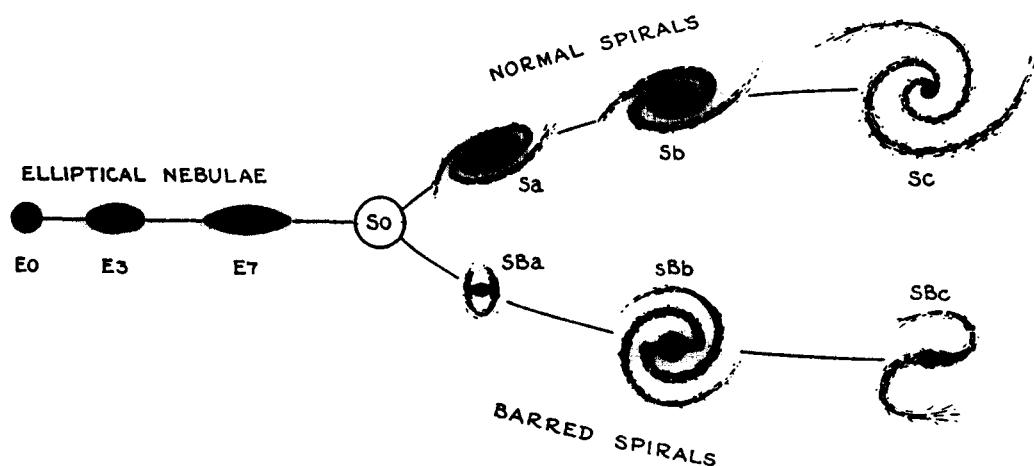


Fig. 1. The Hubble classification (from Hubble, *The Realm of the Nebulae*, Dover Publications, Inc., 1958, p. 45).

the replacement of qualitative impressions by quantitative data shows that an intrinsic scatter is present [e.g. 232]. Measurements of arm lengths, as given in [57, 144] show only a weak dependence on luminosity. A very wide range in arm lengths is present, and no bimodal distribution exists.

2.2. Grand design and flocculent spirals

The question of the coherence length of the spiral structure is often alluded to by making a distinction between on the one hand grand design spirals, where the spiral structure is dominated by a few, usually two, major arms, and on the other hand flocculent spirals, having only armlike filaments with a winding length of $\sim 30^\circ$ or less. These can be described as "a swirling hotch-potch of pieces of spiral arms" [91]. The contrast between NGC 7096 and NGC 2841 clearly illustrates this distinction (fig. 2).

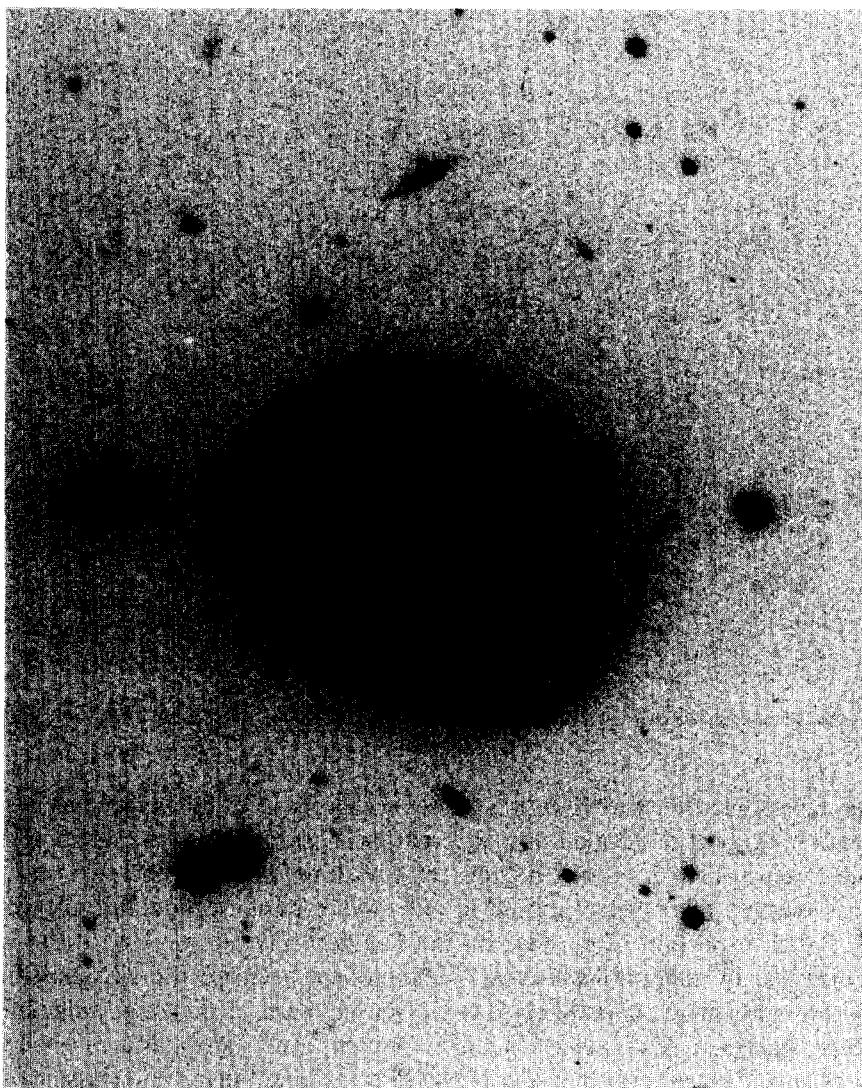


Fig. 2a. NGC 7096, a nice example of a smooth global spiral structure. Photograph courtesy of A. Sandage.

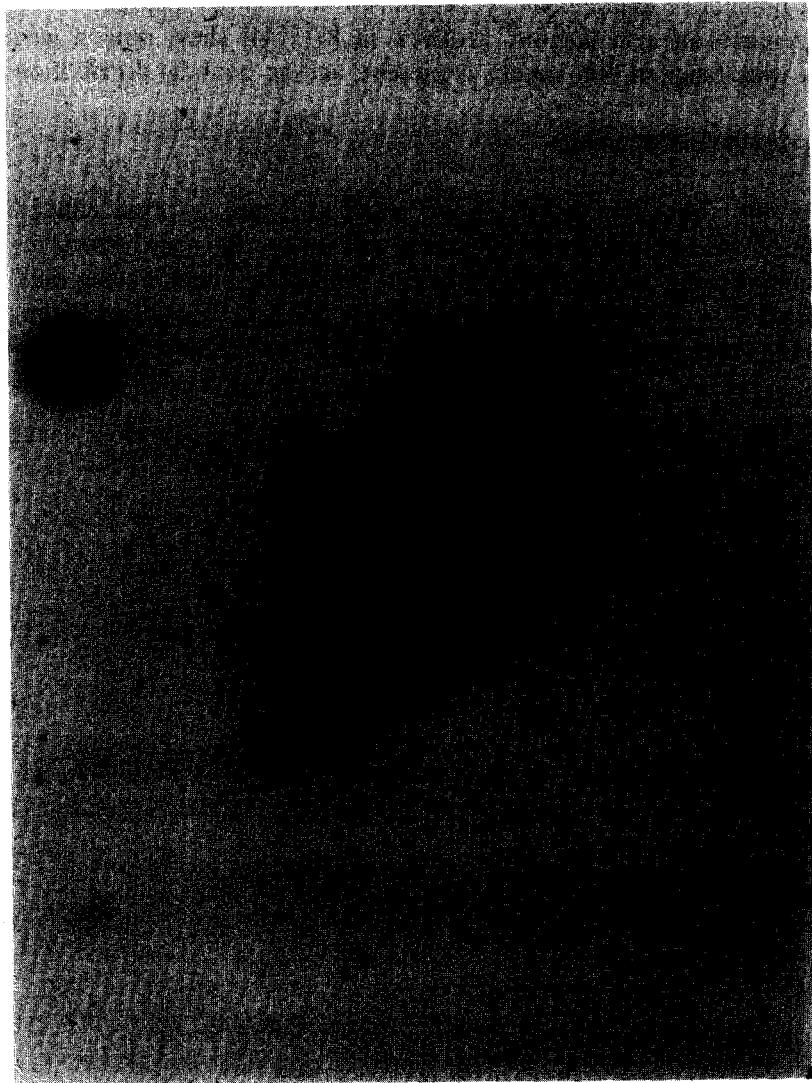


Fig. 2b. NGC 2841, the archetype of flocculent galaxies. Notice the very different spiral structures of figs. 2a and 2b. Photograph courtesy of A. Sandage.

Along this line a classification scheme with 5 classes on the basis of dust lanes has been attempted in [180]. Galaxies with long, well developed dust lane patterns can obviously be distinguished from those with scattered dust patches, and several degrees of transition between the two can be recognized. Similarly, a 12-class system has been proposed in [78] on the basis of blue and infrared photographs, supplemented by visual inspection of galaxies on the Palomar Sky Survey. Again the main variable is the degree of coherence in the spiral pattern. A comparison between the two classification schemes shows only moderate agreement, most likely due to the influence of personal judgment in attributing classes to particular galaxies. Also, since no further quantitative information is being used, several phenomena may be mixed together, especially in the intermediate bins of the classification.

These morphological studies provide also information on another aspect of the spiral structure problem, namely that spiral structure often does not come by itself. Indeed, as for example in M51 (fig. 3), the conspicuous presence of a companion fuels the suspicion that it played a prominent role in the creation of the spiral. Likewise, in galaxies with bars, e.g. NGC 1300 (fig. 4), one suspects that it is now the bar that has the major role in creating not only the spiral, but also the dust lanes.

From a large sample of galaxies, the following quantitative conclusions can be drawn on the frequency of occurrence of grand design spirals [78]: Only 30% of the isolated SA galaxies show a grand design, as opposed to 70% of SA's in binaries and groups. Barred spirals and intermediate types (SB and SAB) show a grand design in almost all cases, except for the very messy barred irregulars. Indeed examination of the plate material [230] of the Revised Shapley Ames Catalogue shows that regular barred galaxies with flocculent spiral structure are extremely rare. Two examples, NGC 1169 and NGC 7371, are given in fig. 5, although neither is as convincing as NGC 2841 itself, or as strongly barred as, say, NGC 1300.

2.3. Further properties of arms

In the vast majority of cases, the sense of winding of the arms with respect to the direction of rotation has been found to be trailing, i.e. the outer sections of an arm lag behind the inner parts in the same sense as the rotational velocities [113, 60, 210]. Very few examples of leading spirals have been suggested so far [210, 136]. Furthermore inspection of plates gives one the impression that bisymmetry is largely predominant, though a lot of multi-armed galaxies, and even a few one-armed ones [155, 136] are known. It would be very useful to confirm these impressions with some good statistics.

The spiral arms are outlined by several tracers. On broad band photographs of grand design spirals like M51 one usually sees the luminous arms made up by HII-regions and OB-associations, and dark lanes of absorbing material displaced very often towards the concave side. $H\alpha$ or UV-photographs bring out the arm structure as outlined by HII-regions. Modern 21-cm line studies are now sensitive enough and have sufficiently high angular resolution to show the spiral structure in the neutral gas, while recent work on the CO lines in several nearby spirals has demonstrated a close correspondence of the molecular gas with arms.

To trace spiral arms in galaxies with a grand design less obvious than M51 is not always straightforward. It is possible, however, to disentangle complicated patterns by making a Fourier analysis of the galaxy image. In brief, this involves decomposing the observed distribution into components with different angular periodicities. Then each component is analyzed separately into a suitably chosen basis to obtain its form, orientation and amplitude. For reasons mainly of mathematical convenience, logarithmic spirals have always been chosen as basis vectors, although in principle other sets, e.g. Legendre polynomials or Bessel functions, would do equally well. This method was first introduced and elaborated for the distribution of HII-regions in M31 [136]. A one-armed leading component was found to dominate the spiral up to $70'$. Inclusion of more data corroborated this result and led to the recognition of two more sets of two-armed spirals in the outer parts [48].

Several other galaxies have been studied in this manner [48, 124, 125, 126]. An interesting discrepancy arises for some galaxies when the optimum signal-to-noise ratios for two-armed components give orientation parameters for the disk which do not agree with those derived from surface photometry or kinematics. In particular such results have been found and discussed for the nearby spirals M33 and M51 [231, 48, 30].

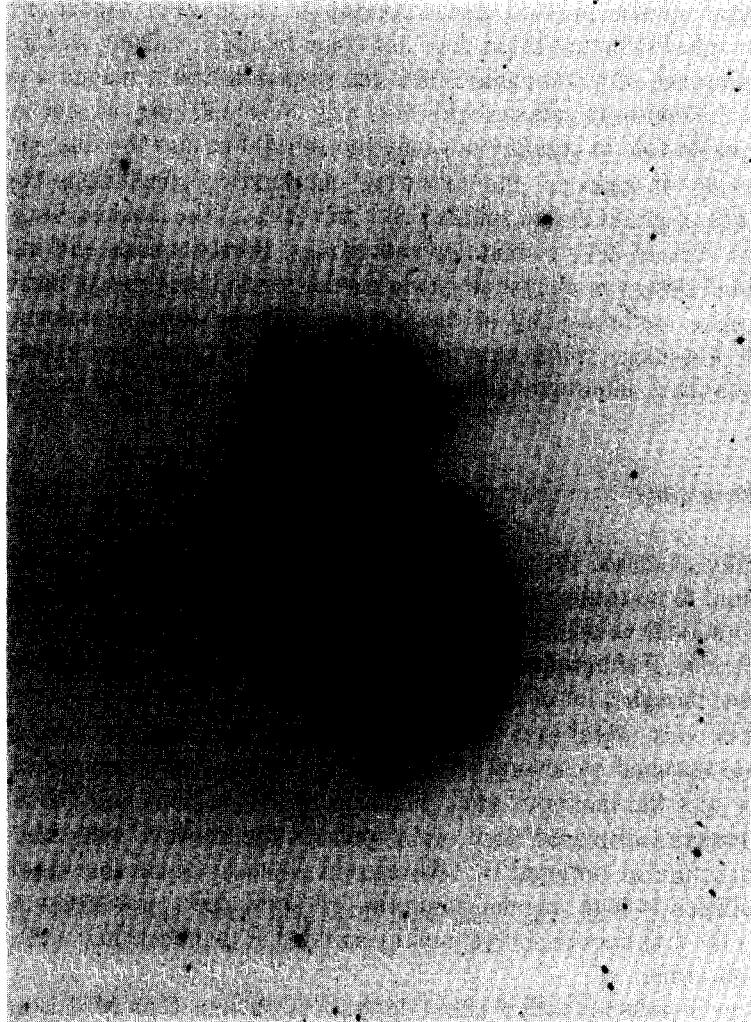


Fig. 3. M51, one of the best known global spiral structures, and its close companion NGC 5195. Theoretical calculations link directly the presence of the companion with the existence of the grand design spiral. Photograph courtesy of A. Sandage.

Some of the nicest and best developed spirals can be found in SB galaxies. These are often composed of two parts, an inner, high-surface-brightness part, and an outer part of low surface brightness. The latter parts of the arms turn back towards the more central part of the galaxy (as if they were trying to link with the side of the bar opposite to that from which they emanated), so that the complete spiral looks closed, like a pseudo-ring (cf. NGC 1300). Two such examples can be seen in figs. 4 and 6. In NGC 1365 the second, low surface brightness part is composed of small bits and pieces while in NGC 1300 the arms are much more continuous and in fact form two nearly complete intersecting ellipses.

In addition to the grand design spiral there are less important structures to be seen on photographs, usually called feathers, spurs or branches. Dust feathers often occur as little branches off the main dust

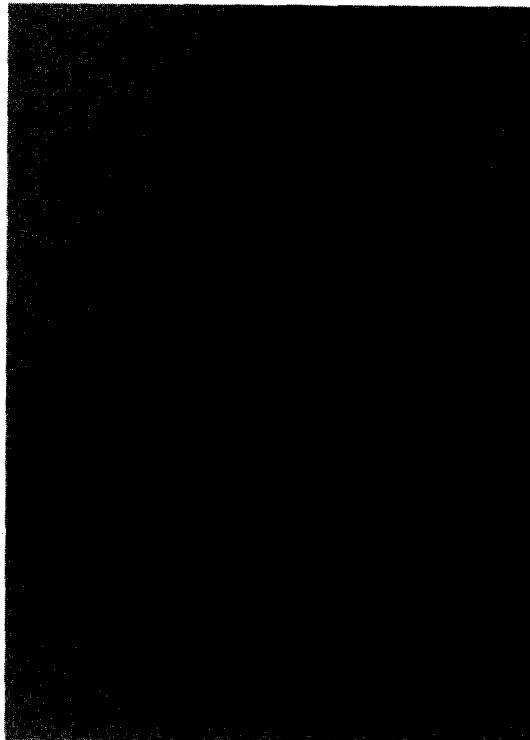


Fig. 4. NGC 1300, a southern barred galaxy. Notice the shape and position of the dust lanes along the bar and the low surface brightness parts of the arms, forming together with the inner arms two nearly complete intersecting ellipses. Photograph courtesy of A. Bosma.

lane. They are always found at the outer edges [177] and have typical pitch angles of $53^\circ \pm 12^\circ$. Luminous spurs can also be seen quite often. They too are found at the outer edges and have typical pitch angles of $63^\circ \pm 12^\circ$. The symmetrized version of M51 [140] shows that its spurs, like its two-armed grand design, are bisymmetric.

2.4. Bars

The differences between bars in various Hubble types are very distinct. In early type galaxies they are smooth, with no dust patches or lanes and no sign of resolution into lumps or knots. In many cases one finds that the luminosity of the bar does not decrease monotonically along its major axis, but rather that the bar can be described as consisting of a central part and two smooth luminous regions of varied size, called ansae [227], diametrically opposed from the center and equidistant from it. Good examples are NGC 2859, NGC 4262, NGC 5101 and NGC 2950, some of which are illustrated on page 42 of [227]. In the most extreme cases the three regions look completely detached.

If ansae are to be found only in types earlier than or up to SB_a, dust lanes become a major feature for types SB_b and later. Ansae and dust lanes along the bar seem to be mutually exclusive. In the not too early types one often finds instead of ansae two bright regions of intense star formation near the

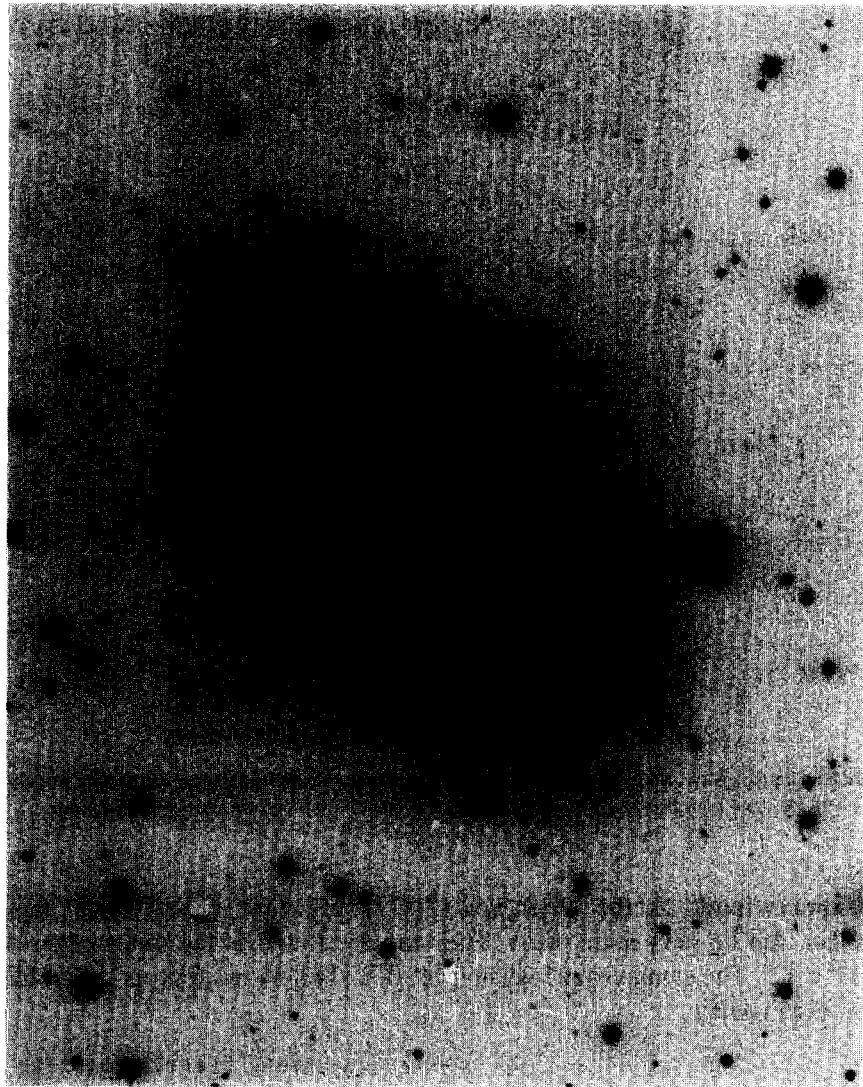


Fig. 5a. NGC 1169. See text for a description. Photograph courtesy of A. Sandage.

ends of the bar. Only in SBc and later types is the bar clearly resolved in luminous knots or stars. It thus seems that bars in early type galaxies are mainly composed of stars, while further along the Hubble sequence dust and gas are introduced in a gradually increased quantity.

A correlation has been found [13] between bulge size and bar length, both measured with respect to the disk size. The mean ratio of bar length to bulge diameter is 2.6 ± 0.7 , but there is some subjectivity in measuring bulge diameters. Moreover, the correction for inclination is not trivial, and the sample is biased towards galaxies with rings.

For a given Hubble type, both bar strength and axial ratio are used to assign the distinction SB, SAB

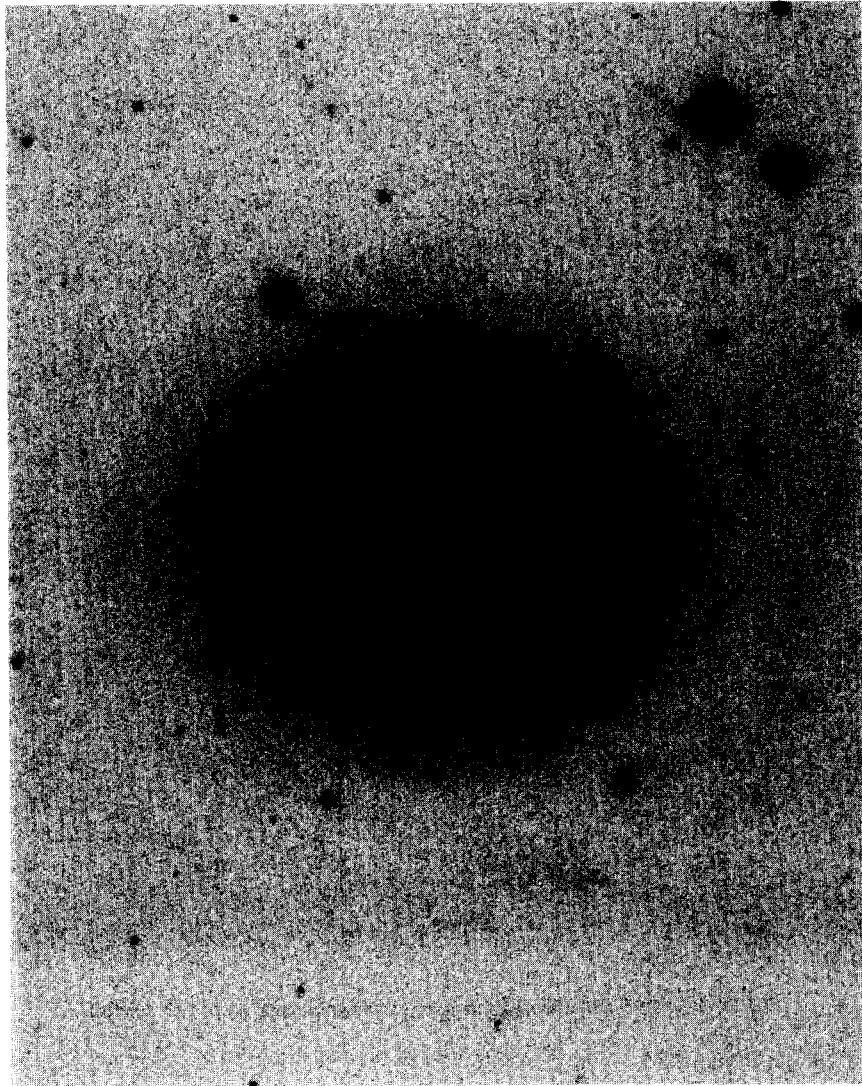


Fig. 5b. NGC 7371. See text for a description. Photograph courtesy of A. Sandage.

and SA [61]. Axial ratios vary roughly from 0.25 for strong bars to 0.9 for weak oval distortions. Thus there exists a wide variety of bar perturbations on the axisymmetric background. In many galaxies the bar turns at a certain radius rather abruptly into a spiral. Contrary to spirals, bars of symmetry $m > 2$ have never been observed. They would, had they existed, look like wheelspikes.

2.5. Rings

Rings frequently occur in both ordinary and barred spirals, inner rings being more frequent than

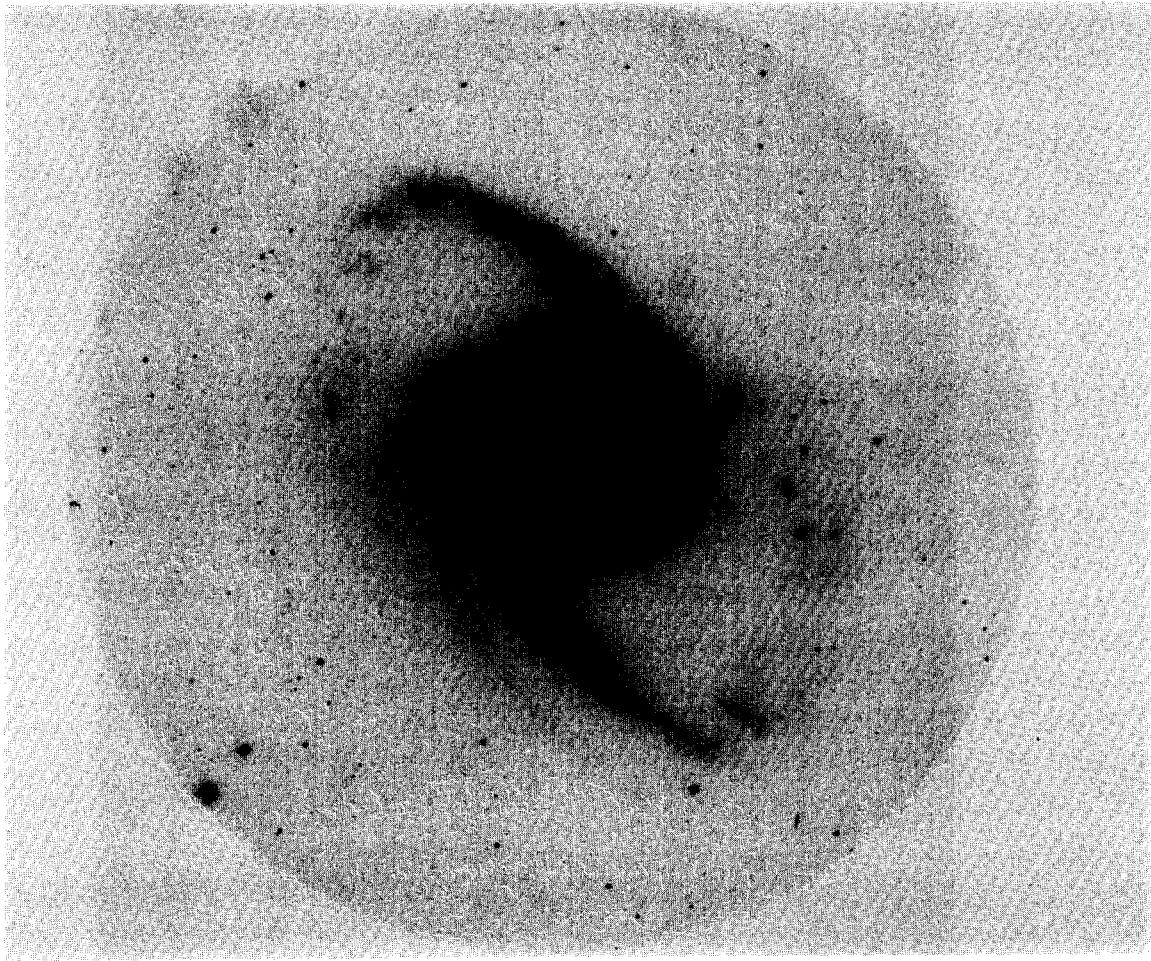


Fig. 6. NGC 1365, a southern barred galaxy. Photograph courtesy of P.O. Lindblad.

outer ones. About 20% of the SA's and 30% of the SB's are of the (r)-variety [61], and there is a large fraction of the (rs)-variety as well. Outer rings occur only in 4% of the 2300 galaxies classified in [65]. Statistics of the frequency of rings with Hubble type [61, 62] show a preponderance of rings among S0's and early type spirals, and their virtual absence in late type systems.

Inner rings in barred spirals surround the bar and more or less touch its extremities, while outer rings have a mean radius of order 2.2 times that of the bar major axis or inner ring size [147, 11]. Outer rings in ordinary spirals occur outside the main spiral structure and inner ones closer to the center. As the ratio of inner ring to disk size was found to vary regularly with Hubble type it has been used as a quaternary distance indicator [64].

Good examples of outer rings can be seen in NGC 1291, NGC 2217 and NGC 2859. Two extreme types can be distinguished: broad stellar and truly closed outer rings, which have a very smooth light distribution and red colors, and gaseous (near) rings consisting of two or more narrow spiral arms very

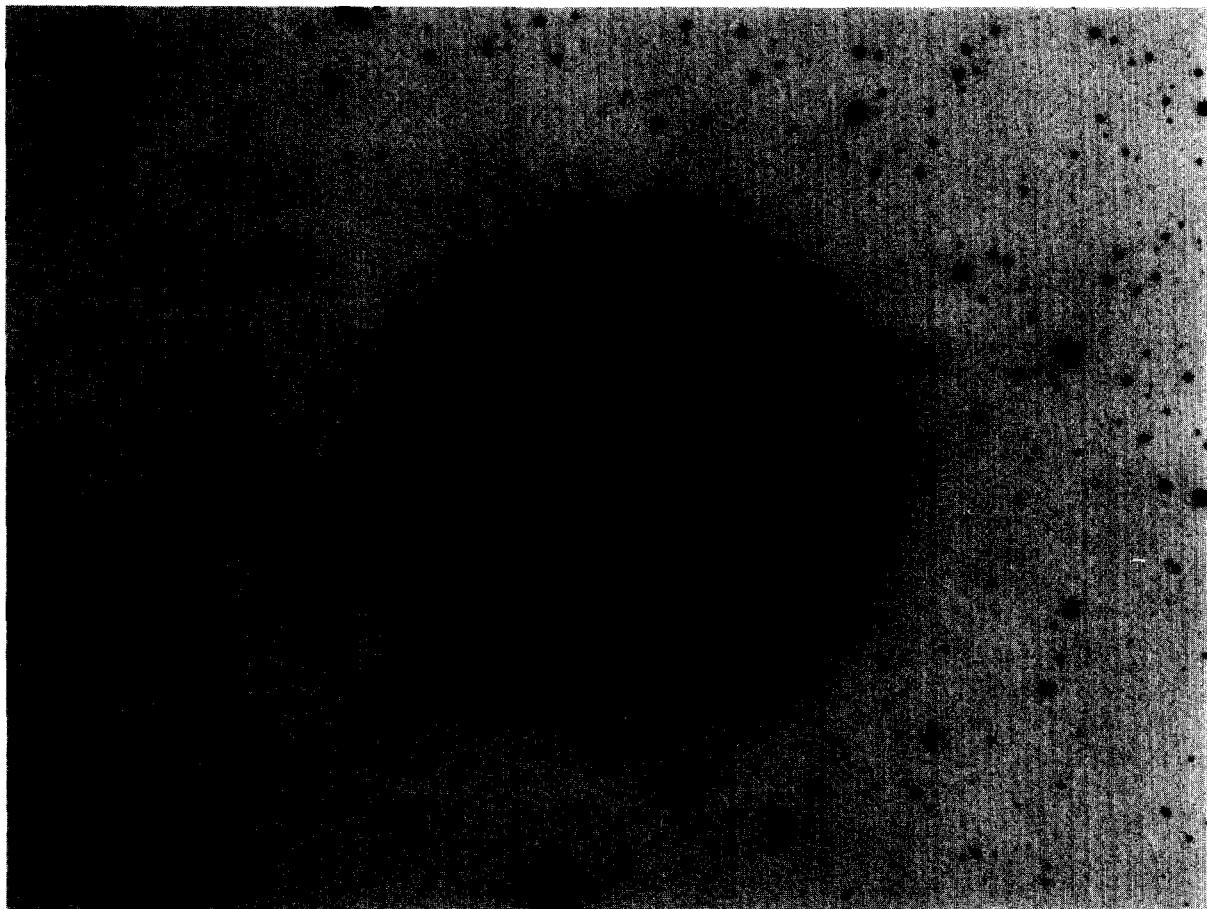


Fig. 7. NGC 2217, a barred galaxy with a prominent outer ring and an inner (pseudo) ring.

tightly wound. On low resolution plates gaseous rings look continuous, but on better plates regions of star formation can sometimes be easily recognized. However many rings show a mixture of both, i.e., patches of star formation are seen against a continuous stellar background (e.g. NGC 1543 has a stellar outer ring with two small arcs with ongoing star formation, at diagonally opposite sides of the nucleus). From the statistics of axial ratios of outer rings [11] it follows that their axial ratios lie somewhere between 0.8 and 1.0, and a sizeable fraction of rings with true axial ratios less than 0.7 can be excluded. Thus their axial ratio should range somewhere between 1.0 and 0.8. Although there is evidence that outer rings in barred spirals could be round, or simply rounder than in ordinary spirals, the data are too noisy to make a definite statement [11]. Furthermore in [147] and [242] it was found that outer rings lie preferentially perpendicular to the bar. In both cases, however, the samples were rather small, containing only 13 and 23 galaxies respectively.

Examples of well developed inner rings can be seen in NGC 1398, NGC 3351 and NGC 4392. They also have stellar and gaseous characteristics, and purely stellar inner rings are only found in very early types. On large scale plates some of the rings resolve into two tightly wound spiral arms with winding

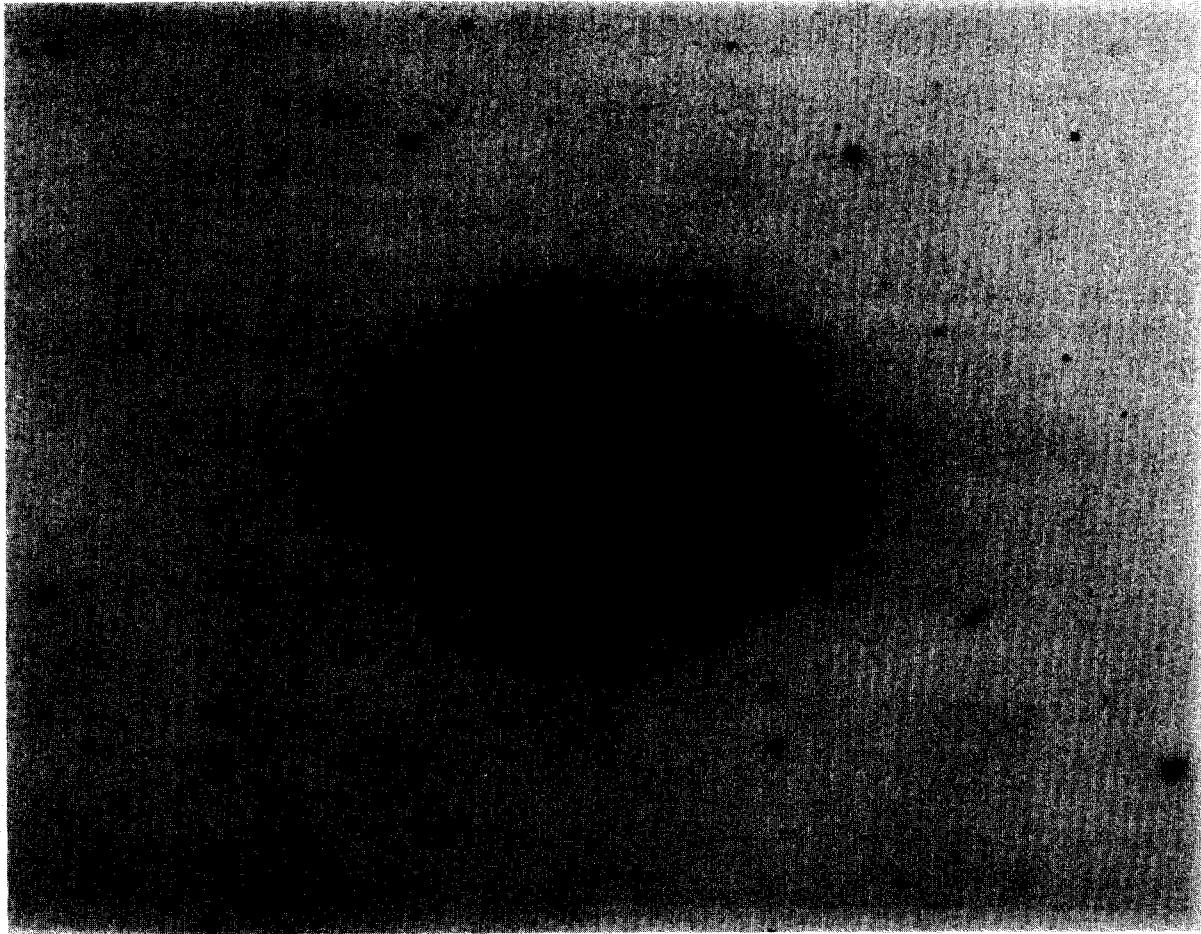


Fig. 8. NGC 1398, a barred galaxy with an inner ring and an outer (pseudo) ring. The area within the inner ring is filled by a lens. Photograph courtesy of A. Bosma.

length of $\sim 180^\circ$. Axial ratios of inner rings are between 0.8 and 1.0 [11] and there is some evidence that they are rounder in barred than in ordinary spirals. Furthermore they are elongated along the bar [242].

2.6. Lenses

A lens can be defined as a shelf in the luminosity distribution of a disk galaxy occurring in between the bulge and the disk, with a shallow brightness gradient and a steep outer edge. A well defined example is NGC 1553 [86]. Further examples can be found in [227, 232, 147]. Lenses are flat components, like the disk itself, as shown by the edge-on example NGC 4762 [280, 39]. At present there is still a great deal of confusion about the nature of lenses. According to [148] objects fitting the above description should be further subdivided. Thus one would distinguish between those located in early type galaxies (S0 to Sa), which would be called lenses, and those located in later types (Sb to Sm) which would be called ovals. In fact further work is necessary to show whether such a distinction is real and necessary. There is also a disagreement on the definition and size measurements of lenses between [147]

and [63] (see [11] for a comparison). Confusing the issue still further is the occurrence of lenses in some of the galaxies with low surface brightness disks [31].

A statistical analysis [11] showed that the lenses as recognized in [63] have intrinsic axial ratio mainly between 0.7 and 1.0. Hence an important fraction of them are eccentric enough to produce sizeable non-axisymmetric forces, thus qualitatively resembling bars. The occurrence of lenses in barred galaxies has been studied in [147] where it is argued that they are predominantly present in types SB0 to SBab. Inner rings take predominance at types SBb-SBbc and spiral arms at types SBc and later. However, there is a very significant fraction of galaxies without either lens or ring. It is also stressed that bars in early type disk galaxies fill the lens in one dimension. The diameter of bars, lenses and inner and outer rings correlates well with the total absolute magnitude of the galaxy.

2.7. Dust lanes

Dust lanes along the bar are often prominent in galaxies of type SBb or later. They can be found in two shapes:

i) Straight. The best example is NGC 1300 (fig. 4). Although it is usually believed that they are parallel to the bar major axis, careful inspection of large scale plates [230] gives the impression that at least in several cases they form an angle with it while staying parallel to each other, as is shown schematically in fig. 9a. Careful photometry would be needed to confirm this. In the inner parts they sometimes curl around the bulge (fig. 9b).

ii) Curved (fig. 9c), with their concave side towards the major axis. Good examples are NGC 6782 and NGC 1433 (fig. 11).

Quite a different category of dust lanes and arcs are seen as well, often occurring coextensively with the shapes discussed above. The dust pattern consists then of arcs running across rather than along the bar major axis, and often these are found on both sides of the bar. Frequently there are also peculiar or irregular shapes, which defy a systematic description.

Dust is not found only in the bar, but also in the spiral arms and in the inner and outer rings. In some grand design spirals, like M51, the primary dust lanes are continuous and located at the inside edge of the luminous spiral arms. However for the majority of galaxies the situation is more complicated and the dust and HII-regions do not form a nice spatial sequence. Quite often small dust arcs are found nearly everywhere, as a detailed perusal of the sketches and photographs in [179] shows. It is interesting

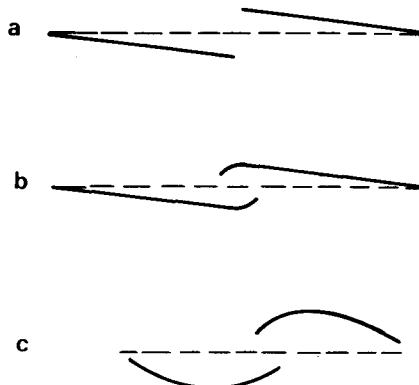
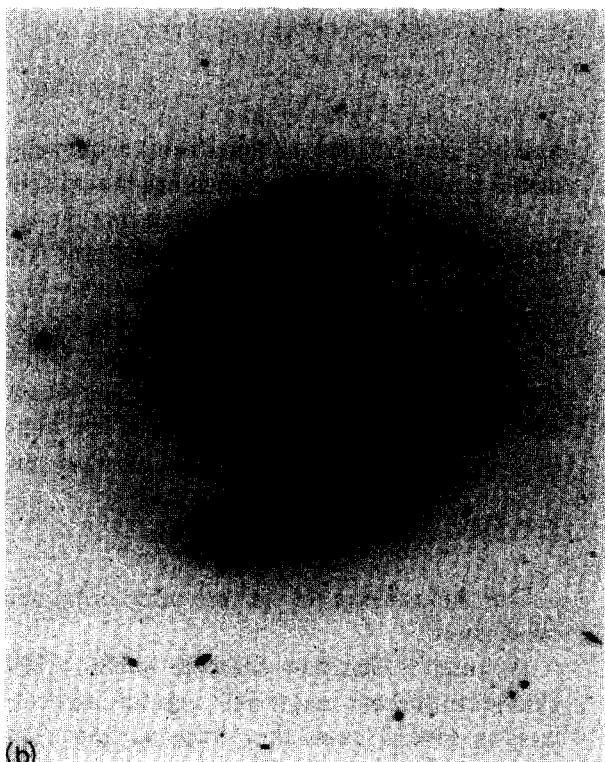


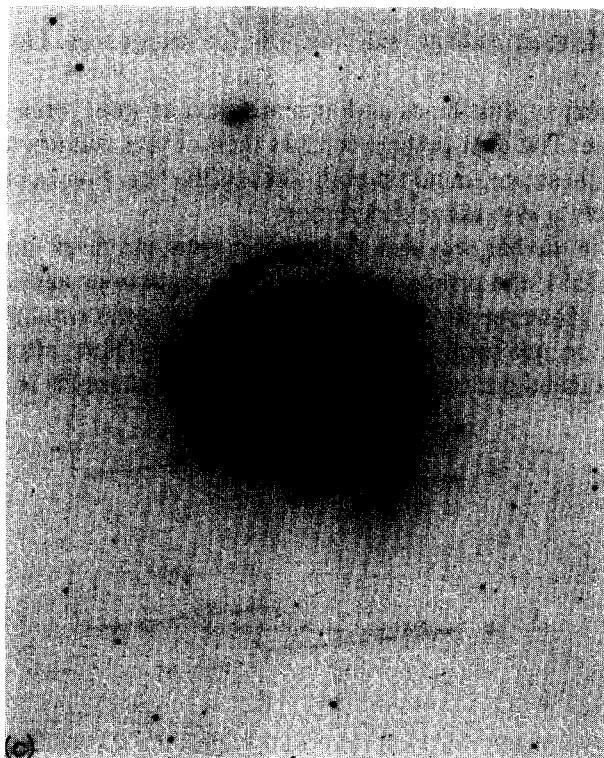
Fig. 9. Shapes of dust lanes in bars. The bar major axis is given by a dashed line.



(a)



(b)



(c)

Fig. 10. Shapes of dust lanes. (a) shows the dust lanes along the bar of NGC 5728. (b) shows the more complex structure of the dust in NGC 986, which starts as straight parallel dust lanes, bends sharply at the end of the bar, encircles, with several gaps, for roughly 180° the inner disk till it joins the opposite arm and then continues along it. (c) shows the dust lanes in NGC 4321, arced in the central region or following the spiral arms. Photographs courtesy of A. Sandage.

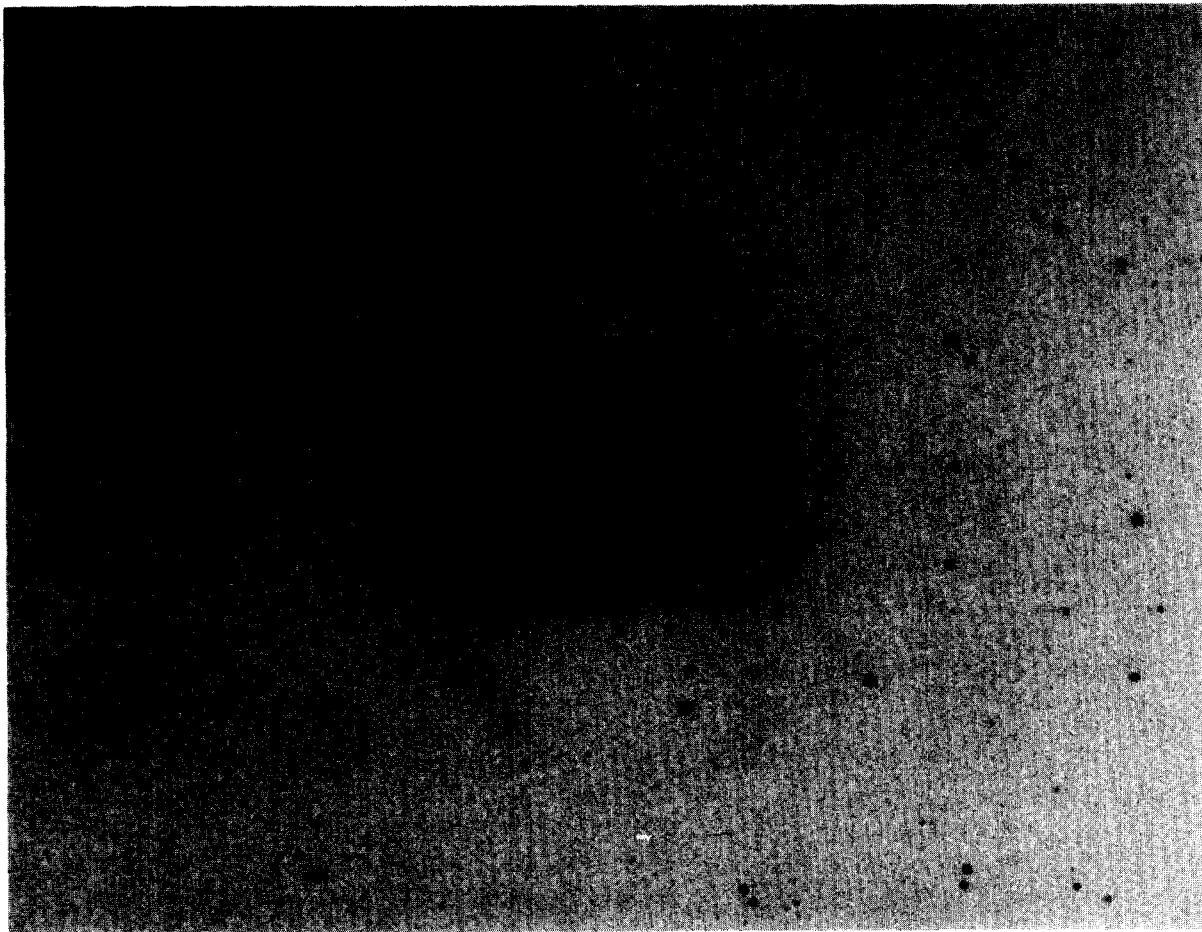


Fig. 11. NGC 1433, a barred galaxy with an inner ring and two outer arms in the shape of a (pseudo) ring. Notice the arced dust lanes in the bar. Photograph courtesy of A. Bosma.

to note that even these small dust arcs are very often distributed bisymmetrically with respect to the center of the galaxy. In the case of inner rings small dust arcs often cross the ring either perpendicular to the bar major axis or just at the ends of the bar.

A very interesting large scale dust pattern can sometimes be found along the outer boundaries of bars and ovals. Typical examples are NGC 986 (fig. 10) or NGC 613, though the latter has a somewhat complicated structure because of the multiarm structure in the outer parts. Schematically in the simpler two-armed case the dust starts out from the nucleus, crosses the bar or oval as straight dust lanes and then wraps around its outer boundary for about 180° to continue along the arms.

3. Basic properties of spiral galaxies

In this section we review a number of basic quantitative facts about spiral galaxies relevant for the theory of spiral structure. We have been selective here, since many stray facts are not yet integrated into a coherent theoretical framework. It seems very useful, however, to confront especially the theoretically minded reader with certain basic observations, which cannot be ignored with impunity.

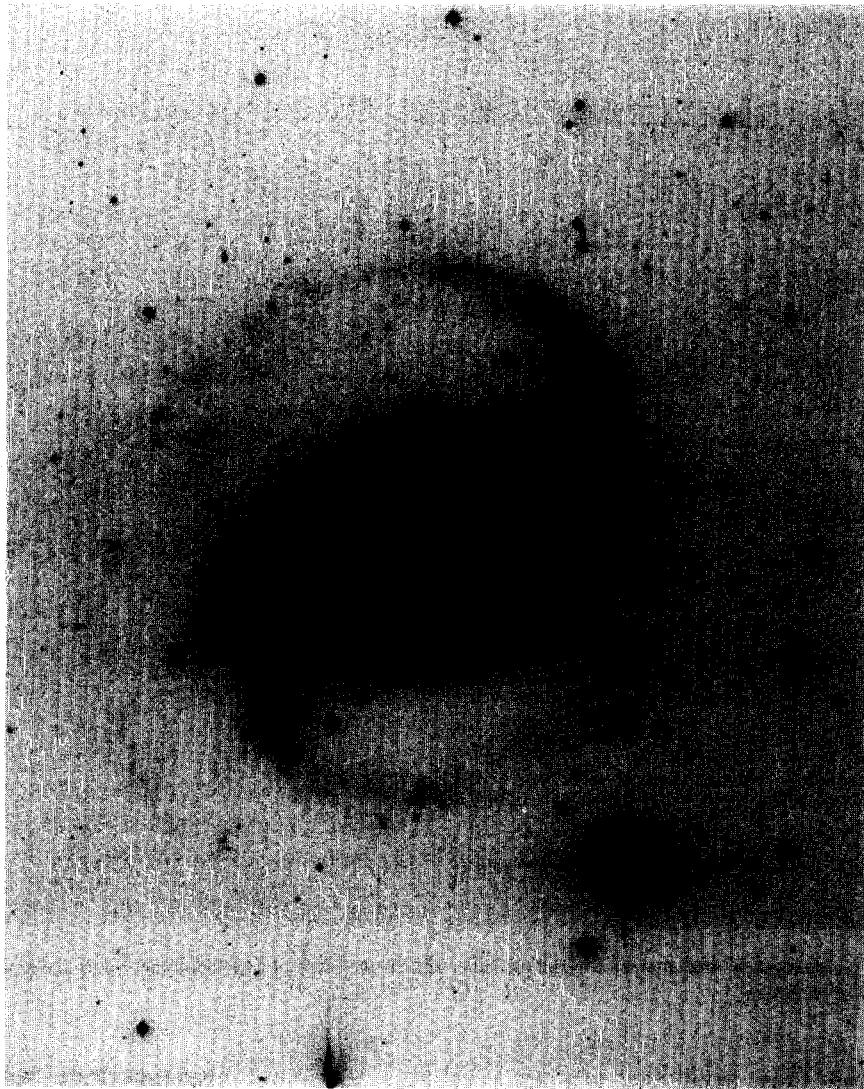


Fig. 12. NGC 1566, a galaxy with an interesting structure. The inner part has a smooth two-armed spiral structure and the outer a less well defined low surface brightness spiral. In the H infrared band this galaxy shows a prominent bar with an azimuthal brightness variation of 25%. Photograph courtesy of A. Sandage.

Already Hubble noted the large spread in total luminosities of spiral galaxies. The brightest spirals have an absolute magnitude in the blue spectral region of $M_B \sim -23$, and the faintest have $M_B \sim -17$. The brightest Sa's are intrinsically brighter than the brightest Sc's, and generally have also a higher mean surface brightness [e.g. 229]. The faintest Sc's overlap in luminosity with the brightest irregular galaxies. A good correlation, known as the Tully-Fisher relation [281], exists between the width of the integrated 21-cm line profile, corrected for inclination, and the total luminosity in the blue. Recently this relation has been extensively studied using also infrared magnitudes, and its dependence on Hubble type has been hotly debated [e.g. 1].

The corrected 21-cm line width gives only a global measure of the maximum rotational velocity of a

galaxy. In the theoretical modelling of a galaxy much more kinematical information is required. An important role is played by the rotation curve, describing the run of circular velocities of a test particle with radius. Since the gas in axisymmetric galaxies is on nearly circular orbits, and its random motions are small compared to the rotation, its kinematics can be used to derive rotation curves. The stellar motions are much more complicated, and the velocity dispersions are not necessarily negligible compared to the rotational velocities. Thus 'rotation curves' derived from absorption lines, which refer to stellar motions, should not be confused with rotation curves derived from gaseous emission lines.

Modern data [224, 30, and references therein] have given us a good idea of the shape of HII- and HI-rotation curves as function of Hubble type and luminosity. Schematically, most rotation curves can be characterized by a rising part in the central regions, with constant angular velocity Ω_0 , followed by a flat part further out, with constant circular velocity V_0 . For galaxies of equal luminosity those of type Sa have a higher V_0 than those of type Sc, the Sb's being intermediate. Within each type, a lower luminosity galaxy has a lower V_0 . For galaxies of different type with equal V_0 the Sa's have slightly higher Ω_0 than the Sc's. For all types the radius where the rotation curve turns over from a rising to a flat part is smaller when the luminosity is higher. This holds only in relative terms, e.g. as compared to the Holmberg radius of the galaxy. In absolute terms, less luminous galaxies are also much smaller in size.

If a galaxy has a flat rotation curve out to infinity, its mass surface density will vary as r^{-1} with radius. The observed radial distribution of luminosity, however, can be characterized roughly by an exponential law for the disk and a $r^{1/4}$ for the bulge [e.g. 85]. This implies that the local ratio of mass-to-luminosity will increase as function of radius, and reach values well over 100 in the outermost parts [e.g. 35]. Since the mass-to-light ratio of the stellar population is of order 1 to 10, a large amount of dark material must exist. Arguments have been presented that this material is distributed in a halo, but at present it is hard to obtain definite quantitative data on such halos [e.g. 278 and references therein].

The problems presented by the possible existence of halos are of direct interest to the theory of spiral structure, as we shall see. It is at present not possible to obtain an estimate of the mass surface density of the disk in which the spiral structure is observed, other than assuming some hypothesis about the mass-to-light ratio in conjunction with the distribution of light. Close inspection of the available data on radial luminosity profiles [e.g. 85, 39, 28, 291] show that there are a large number of exceptions to the simple exponential distribution. The presence of rings and lenses can usually be recognized, and sometimes effects like Type II profiles [85] and cutoffs in the disks [291] are present. Furthermore the stellar populations in the disk, and therefore its mass-to-light ratio, may change with radius or not be the same in say the lens as in the disk. For all these reasons it is very difficult to estimate the halo-to-disk mass and to choose a correct axisymmetric mass model.

The stellar and gaseous content of spiral galaxies also varies with Hubble type. The broad band integrated colors of Sc's are in general much bluer than those of Sa's, due to the larger fraction of young stars. Also the ratio of hydrogen mass to luminosity is higher in Sc's than in Sa's. Although a large intrinsic scatter exists, the trend is clearly that gas and young stars are much more important in Sc's than in Sa's. This is obviously reflected in the second criterion of the Hubble classification scheme, i.e. the degree of resolution into stars and HII-regions.

A few quantitative results are now known about the form of the spiral arms as function of Hubble type and luminosity. The variation of pitch angle with Hubble type has already been discussed in section 2. The intrinsic scatter may be partly ascribed to the correlation between pitch angle and luminosity (or alternatively the maximum rotational velocity) [145]. Measurements of arm widths on the basis of H photographs were discussed in [144]. The mean absolute armwidth increases with the luminosity of the

galaxy, but the ratio armwidth-to-disk size decreases with luminosity. Thus for bright galaxies the arms appear thinner, although their absolute widths are larger. The armwidth based on HII-regions does not increase with radius in a given galaxy [144], contrary to the result for dust lanes [177], which appear narrower in the central parts. However, these partly subjective measurements should be properly quantified.

4. Introduction to density-wave theory

4.1. Dynamical background

In order to understand a dynamical theory of spiral structure we need to introduce a number of general concepts like orbits and resonances, and some basic equations. We will introduce them here in the framework of the first theoretical attempt to understand spiral structure, i.e. the quasi-stationary density-wave theory, although their applicability is obviously more general.

Describing something as complex as a spiral galaxy by mathematical equations is obviously no easy matter. We will be compelled, all along the line, to make a long list of approximations and hypotheses. Let us start by neglecting the gaseous component (with the excuse that we will discuss it at length later on), the thickness of the galaxy and assume that gravity is the only force and that all stars have equal masses which we will take as unity. Still, keeping track of 10^{11} individual stars is impossible, so we generally describe the dynamics in terms of a continuous distribution of matter. Since encounters can be neglected we can use the collisionless Boltzmann equation

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} - \frac{\partial \Phi}{\partial \mathbf{x}} \frac{\partial f}{\partial \mathbf{v}} = 0$$

where $f = f(\mathbf{x}, \mathbf{v}, t)$ is the distribution function, describing the density in phase space, and Φ the potential. If we integrate f over all velocities we obtain the density

$$\mu = \int f d\mathbf{v}.$$

This equation is supplemented by Poisson's equation

$$\nabla^2 \Phi = 4\pi G \mu$$

relating the density with the potential that created it.

The final goal is to solve these two equations for the two unknowns (Φ and μ , or equivalently Φ and f). Such a solution is called selfconsistent as opposed to a case where the potential is assumed at the start and only the response to this forcing is calculated. We calculate both the density necessary to maintain a given gravitational field and the density response of the disk to this field. In selfconsistent cases these two densities must be equal.

4.2. Quasi-stationary density waves

The basic assumption of the density wave theory is that the spiral arms are not always composed of

the same stars and gas clouds but instead are the manifestation of the excess matter associated with the crest of a rotating wave pattern. Two further assumptions, that of linearity and quasi-stationarity were introduced from the onset. The assumption that in real galaxies the amplitude of the spiral or bar perturbations are small compared to the axisymmetric background is not trivial, but it is necessary to allow us to linearize the Boltzmann equation. The second hypothesis, that of quasi-stationarity, is even more questionable. It amounts to saying that the spiral will remain stationary in a frame of reference rotating with an appropriate angular velocity. Surprisingly little discussion or justification of this hypothesis has been presented so far [164], other than it enables us to carry out the analysis of the spiral structure problem relatively far. It is only a few years ago that the first doubts arose. For the remaining of this section and for the few following ones we will adopt it and see where it will lead us to, before stepping back to evaluate what we have thus accomplished.

These assumptions allow us to write any perturbation of the axisymmetric background in the form

$$\begin{Bmatrix} \Phi(r, \vartheta, t) \\ \mu(r, \vartheta, t) \\ f(r, \vartheta, v, t) \end{Bmatrix} = \sum_m \begin{Bmatrix} \Phi_m(r) \\ \mu_m(r) \\ f_m(r, v) \end{Bmatrix} \times \exp[i(\omega_m t - m\vartheta)]$$

where Φ , μ and f are the potential, the density, and the distribution function of the perturbation.

The summation index m indicates the symmetry of the component or, in the case of spirals, the number of arms. $m = 0$ corresponds to the axisymmetric background; $m = 2$, the most frequent perturbation, has a symmetry of π (i.e. the system remains unchanged after a rotation by a multiple of π), while the general m -term corresponds to a symmetry of $2\pi/m$. A word of caution is necessary here: Any m -armed spiral contributes to the m th term of the above summation but the opposite is not necessarily true. For example two symmetrical lumps of matter will give rise to even terms in the expansion and an asymmetry to odd ones. The summation reminds us that several components and symmetries are simultaneously present in each galaxy.

The constant ω_m can be complex: $\omega_m = \omega_{m,R} + i\omega_{m,I}$. The imaginary part corresponds to a growth or decay of the components in time and the real part to a rotation with constant angular velocity

$$\Omega_p = d\vartheta/dt = \omega_{m,R}/m$$

called the pattern speed. Since $\omega_{m,R}$ is a constant, independent of time or radius, each component will remain identical with time.

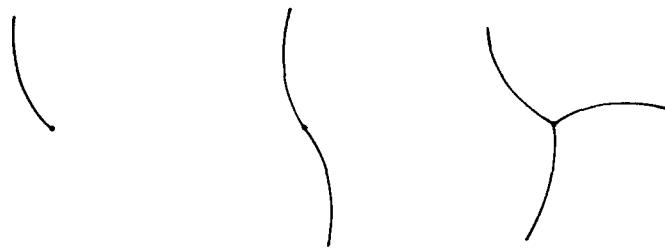
4.3. The spiral

If we consider only one component of the perturbation which we analyzed above we can write it in the usual wave notation:

$$A(r) \exp\{i[\varphi(r) + \omega t - m\vartheta]\}$$

where $A(r)$ is the amplitude of the wave and where we have dropped the subscript m for simplicity. By phase we usually denote either the whole expression in brackets in the exponent or its first and last term only. For every monotonic function $\varphi(r)$ this represents a spiral. Its wave number is defined as

$$k = d\varphi/dr.$$

Fig. 13. $m = 1, 2$ and 3 components of a spiral.

The winding of the spiral is measured by the angle i between the tangent to the spiral and a circle centered on the center of the galaxy. It is inversely proportional to the wave number,

$$\tan i = \frac{m}{r\varphi'(r)} = \frac{m}{kr}.$$

For tightly wound spirals i is small and k large. The inverse is true for open ones, the limiting case being a bar with $i = \pi/2$ and $k = 0$. A change of the sign of k corresponds to changing the sense of winding of the spiral. For positive values of k the spiral unwinds with growing ϑ and the spiral is called leading. For negative k (trailing spiral), the inverse is true.

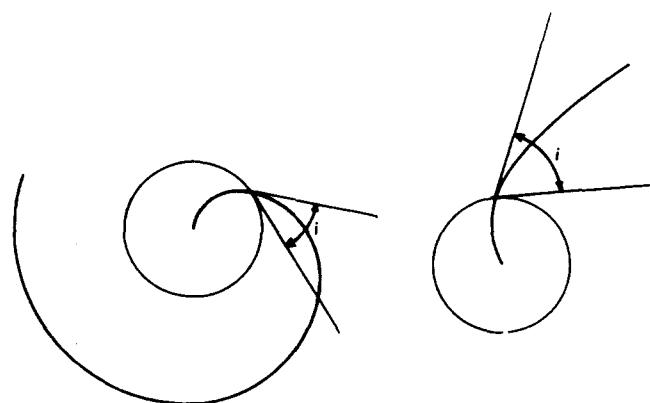
4.4. Orbits and resonances

For all analyses it is essential to have simple expressions for the orbits in the unperturbed galaxy. The zeroth order approximation is simply a circle on which the particle moves with angular velocity

$$\Omega = V/r$$

and rotational velocity

$$V(r) = (r d\Phi/dr)^{1/2}.$$

Fig. 14. Definition of pitch angle i .

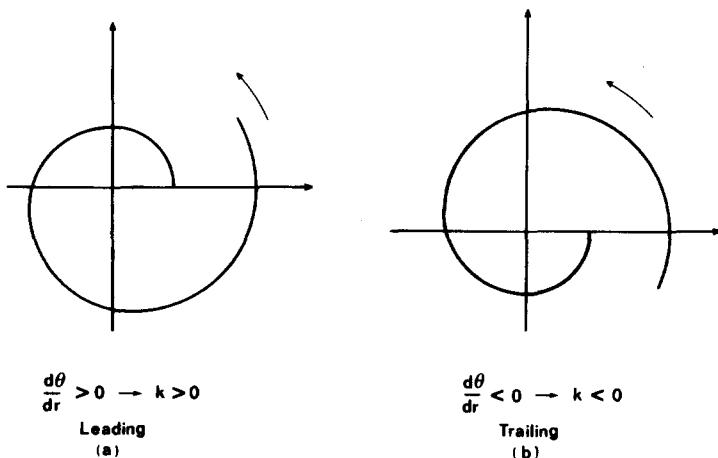


Fig. 15. Leading and trailing spiral patterns.

The curve $V = V(r)$ is called the rotation curve and may be obtained from observations. The first-order epicyclic theory [170] superposes on this rotation harmonic oscillations both in the radial and tangential directions with a characteristic frequency called the epicyclic frequency, given by

$$\kappa = 2\Omega [1 + \frac{1}{2} d \ln \Omega / d \ln r]^{1/2}.$$

Thus in the proper rotating frame the particle will move in a retrograde sense around a small ellipse with axial ratio $\kappa/2\Omega$, called an epicycle (fig. 16a). The resulting motion in the inertial frame is a rosette orbit, generally not closed (fig. 16b).

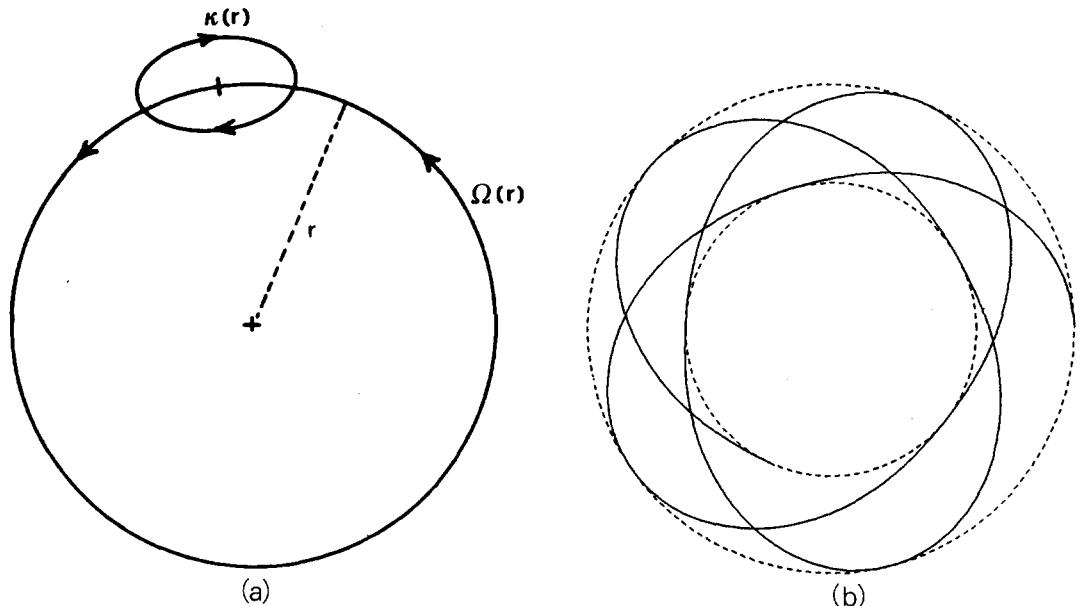


Fig. 16. (a) Epicyclic orbit in a rotating frame. (b) Rosette orbit in an inertial frame.

To a first approximation the motion of a star is thus governed by two frequencies. In a frame corotating with the spiral these are $\Omega(r) - \Omega_p$ and $\kappa(r)$. When they are commensurable, i.e. when the relative frequency

$$\nu = (\omega - m\Omega)/\kappa = m(\Omega_p - \Omega)/\kappa$$

is equal to a rational number, we have a resonance. The three main resonances in a galaxy are

- The inner Lindblad resonance (ILR) where $\nu = -1$ and $\Omega(r) - \kappa(r)/2 = \Omega_p$.
- The corotation resonance where $\nu = 0$ and $\Omega(r) = \Omega_p$.
- The outer Lindblad resonance (OLR) where $\nu = +1$ and $\Omega(r) + \kappa(r)/2 = \Omega_p$.

The location and even the existence of these resonances depends on the rotation curve, that is the mass distribution, and the pattern speed. Two examples, an isochrone [99], with

$$V = \frac{r}{\{(1 + (1 + r^2)^{1/2})^2 (1 + r^2)^{1/2}\}^{1/2}}$$

and a constant circular velocity rotation curve

$$V = \text{constant}$$

are given in figs. 17a and 17b. Note that the former may have none, one or two ILRs, depending on the value of the pattern speed, while the latter will always have one. Note also that the higher the m value the closer in radius the $\Omega_p = \Omega \pm \kappa/m$ resonances are located.

At resonances the orbits viewed from a reference frame rotating with angular frequency Ω_p are closed. At the Lindblad resonances the star goes twice in and out for once around while at corotation the orbit of the star is the little ellipse (see fig. 18).

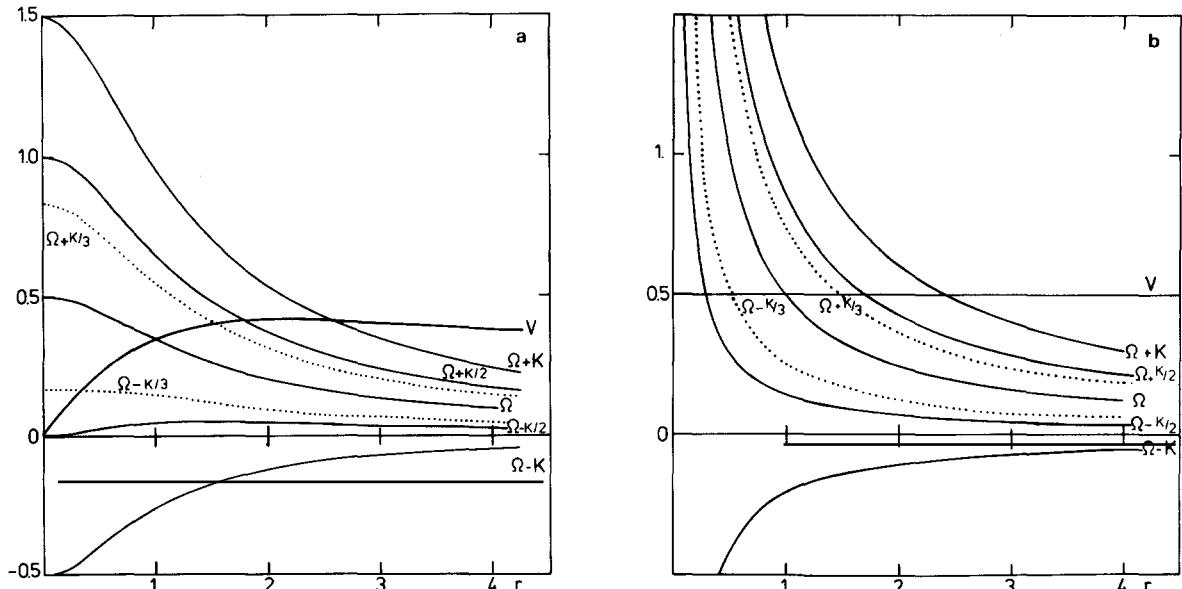


Fig. 17. Ω , $\Omega \pm \kappa$, $\Omega \pm \kappa/2$, $\Omega \pm \kappa/3$ (dotted line) as a function of radius for two different types of rotation curves (V), the isochrone (a) and the constant circular velocity model (b).

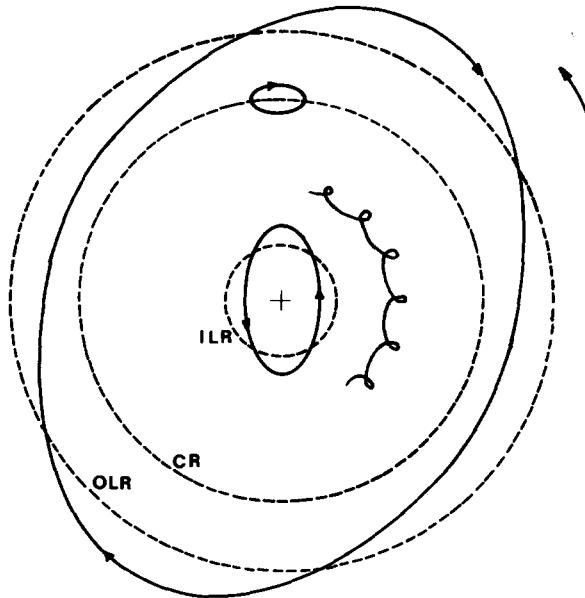


Fig. 18. Orbits at the three principal resonances, viewed in a frame of reference rotating with the spiral wave.

4.5. Kinematic quasi-stationary spirals

The basic concept of density waves can be illustrated using fig. 19 (following [135]). Let us populate with stars a number of orbits which, viewed from a suitable rotating frame, are concentric ellipses with major axes aligned. By turning each orbit by an angle

$$\vartheta = -a \ln R + b$$

where a and b are constants and R is the length of its major axis, we form two spiral arms (fig. 19b). Each star, orbiting on its ellipse, will cross the arms; in other words we have formed a density wave and not material arms. The density at a given point is inversely proportional to the distance between the two

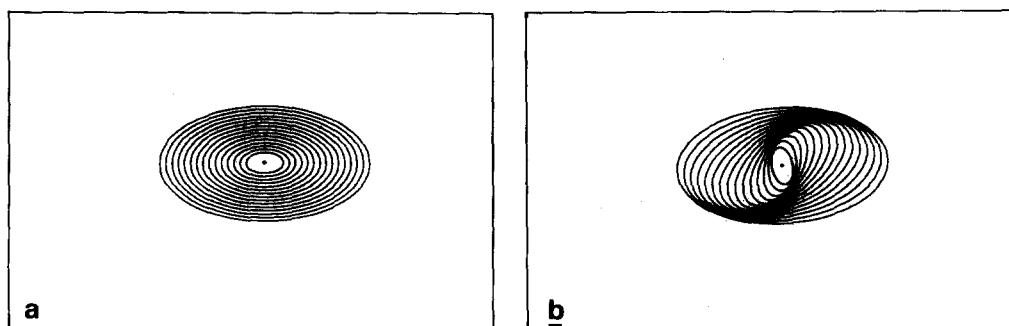


Fig. 19. Kinematic density waves.

neighboring orbits. It is thus the location of the ellipses which causes the maxima to have the shape of spiral arms.

Yet putting the orbits in a proper configuration, as in fig. 19b, is not the complete story. One should worry about how to keep them there. The shortcut taken by adopting the QSSS-hypothesis is simply to require that the spiral will remain stationary in time in the rotating frame. This leads to a mathematical formulation which is tractable, although there is no compelling reason to believe that observed spirals remain stationary.

5. Tightly-wrapped approximation

5.1. *The integral equation*

Applying the approximations discussed in the previous section to the Poisson and Boltzmann equations results in an integral equation [132, 137, 253], forming the backbone of all analyses of modes. Under the chosen set of hypotheses, essentially all the information on modes (if they exist) is contained in this integral equation and we rely on its solution to provide us with the radial distribution of the density of the perturbation as well as its pattern speed and growth rate. Yet we have been able so far to interrogate this equation only for discrete, exponentially growing modes. Furthermore, in order for a given mode to be adopted as a realistic solution, liable to have some application for real galaxies, its growth rate must be relatively low, since otherwise after a few galactic rotations the linear hypothesis would not be valid any more. This integral equation is, except in some simple cases, quite cumbersome. It can be reduced to a more manageable algebraic one, if one introduces further approximations which we will discuss next.

5.2. *The Lin–Shu–Kalnajs dispersion relation*

The mathematical formulation can be considerably simplified under two further assumptions [166, 161, 167].

- The epicyclic approximation is an acceptable approximation of the orbit equations, as is at least true for disks where the velocity dispersions are small compared to the rotational velocities.
- The wave is very tightly wound, i.e.

$$|k|r \gg 1. \quad (5.1)$$

This implies that the azimuthal variations of some quantities in the equations are small compared to the radial ones, and can therefore be neglected. It is known as the WKBJ⁽¹⁾ or asymptotic approximation. It has been heavily disputed, mainly in [132] where it was argued that spiral arms are too open to allow such a description and that even in cases where the observed spiral, i.e., that formed of young stars and gas clouds, is tightly wound the old population forms a rather open spiral [131, 132]. The above approximation is also known as the local approximation in the sense that the galaxy is supposed to be homogeneous over length scales of the order of a wavelength.

Under these assumptions the integral equation is easy to handle and simplifies to an algebraic

⁽¹⁾ From the initials of Wenzel, Kramer, Brillouin and Jeffreys who initially introduced this approximation in a different subject matter.

dispersion relation, known as the Lin–Shu–Kalnajs dispersion relation, which relates the wave number k to the pattern speed Ω_p (or to ν) [161, 168],

$$\kappa^2 - (\omega - m\Omega)^2 = 2\pi G\mu |k| F_\nu(x) \quad (5.2)$$

where $F_\nu(x)$ is a reduction factor. For the stellar WKBJ case it is equal to

$$F_\nu(x) = \frac{1 - \nu^2}{x} \left\{ 1 - \frac{\nu\pi}{2\pi \sin \nu\pi} \int_{-\pi}^{+\pi} \exp[-x(1 + \cos \gamma)] \cos \nu\gamma \, d\gamma \right\} \quad (5.3)$$

where $x = k^2 \sigma_u^2 / \kappa^2$. If instead of a stellar we consider a gaseous disk, then

$$F_\nu(x) = (1 + x/(1 - \nu^2))^{-1} \quad (5.4)$$

where now $x = k^2 c^2 / \kappa^2$ and c is the sound speed of the gas.

This dispersion relation was derived under the assumption that the disk has an infinitesimal thickness. For small but finite thickness the reduction factor changes somewhat [293, 168]. It is also not valid near the principal resonances, where a different approach must be followed. A WKBJ dispersion relation valid in the region near the ILR has been given in [182, 184]. The pattern speed Ω_p can not be derived from the dispersion relation, in contrast with the integral equation. So one must rely on observations to obtain it. Finally, if k is a solution, so is $-k$, i.e. this dispersion relation can not distinguish between leading and trailing solutions.

5.3. Axisymmetric stability

Before studying the spiral perturbations of a given galactic disk we must be reasonably confident of its axisymmetric stability. The analysis can be carried out with the help of eq. (5.2), though in time it preceded it. It can thus be shown [268] that in order for the disk to be safe from Jeans instabilities of wavelength $\lambda (= 2\pi/|k|)$ smaller than

$$\lambda_{\text{crit}} = 2\pi/k_{\text{crit}} = 4\pi^2 G\mu/\kappa^2$$

it must have a velocity dispersion σ_u , larger than

$$\sigma_{u,\text{min}} = 3.36 G\mu/\kappa$$

where G is the gravitational constant and σ_u is the radial velocity dispersion. The ratio of the actual radial velocity dispersion to this required minimum is known as Toomre's Q parameter

$$Q = \sigma_u/\sigma_{u,\text{min}}$$

and is used as a thermometer of galactic disks. Hot disks have large dispersions and high Q values, and the opposite is true for cool or cold disks. The solar neighborhood is the only place where some reasonably accurate values of Q can be derived. Even so the uncertainties in the various quantities

entering this formula are quite large and consequently Q is not well determined. The best estimate is around 1.6, but values in the range of 1.2 to 2.0 are allowed [271 and references therein]. Very recently observations of stellar velocity dispersions have become possible in external galaxies. Unfortunately, the uncertainties in the mass distribution are large and it is not straightforward to derive the corresponding Q values. With this proviso let us note that the observations of the Sb galaxy NGC 488, which has a weak 2-armed spiral, indicate Q values of the order of 2 [152].

If the epicyclic approximation holds there is a simple relation between the velocity dispersions in the radial direction, σ_u , and in the tangential direction, σ_v :

$$\sigma_v/\sigma_u = \kappa/2\Omega.$$

5.4. Properties of the solution

The relative wavelength, $\lambda/\lambda_{\text{crit}}$, derived from the Lin-Shu-Kalnajs dispersion relation (5.2) is plotted as a function of ν in fig. 20 for four values of Q . There are several points worth mentioning:

1) We find two solutions, one tightly wound and the other open, known as the short and long branches. The existence of the long one is questionable, since it is not in agreement with the asymptotic hypothesis from which it emanated, unless a very large fraction of the disk mass is frozen in a non-responsive component (e.g. a halo) [189]. On the other hand the short branch gives interarm distances of correct size and can, after some fiddling with free parameters, be fitted to the observations.

2) A spiral can extend only between the ILR and the OLR. As can be seen, e.g., from fig. 17, quite extensive two-armed spiral structures are thus possible, though for higher m values the region where

$$\Omega - \kappa/m < \Omega_p < \Omega + \kappa/m$$

is much more restricted. This is not true for the gaseous dispersion relation, where the Lindblad resonances are not boundaries of the spiral pattern.

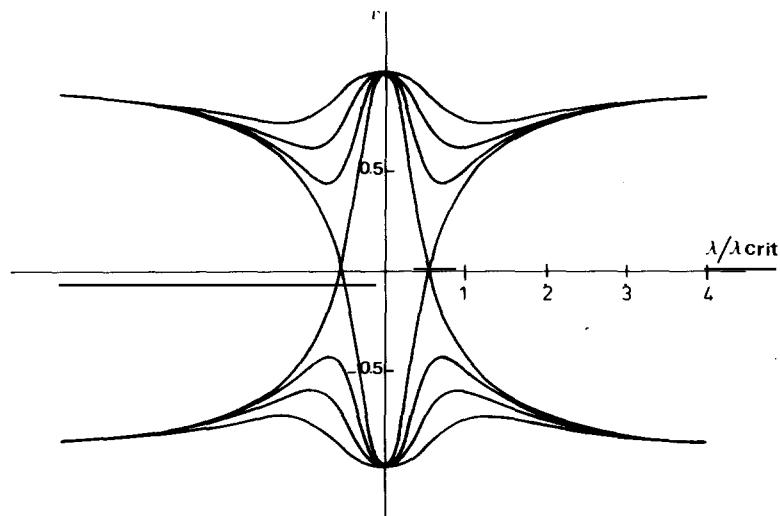


Fig. 20. Solutions of the dispersion relation for $Q = 1.0, 1.2, 1.5$ and 2.0 .

3) The most important limitation to the radial extent of the asymptotic spiral comes from the value of the stability parameter Q . For $Q = 1$ the whole extent between the two Lindblad resonances is permissible. For larger Q values, however, the spiral is excluded from an annulus surrounding corotation. The width of this annulus increases with Q and by $Q = 1.6$ the WKBJ spiral is nearly squashed out. This is clear in fig. 21, which shows spiral $m = 2$ solutions of the WKBJ dispersion relation for different values of Q . Note also that for larger Q values the spirals are more open.

These constraints on the extent of the asymptotic spirals for all but the lowest Q values present a major problem for the Lin-Shu-Kalnajs dispersion relation. It has to be assumed that corotation lies outside the spiral structure, unless the values of Q remain conveniently below 1.2 or 1.3. Either possibility runs into observational difficulties. The first alternative is problematic for galaxies such as UGC 2885 [225], which has a grand design over a very large radial interval. On the other hand low Q values do not seem to be in agreement with the values derived so far from the observations. Nevertheless the large uncertainties entering into the derivation of Q could allow some breathing space. More serious constraints come from numerical simulation results backed by arguments about the heating of disks [104, 107, etc.]. Indeed, if a disk is initially relatively cool, large and small scale instabilities form and heat it up. This heating has indeed been observed in numerical simulations [42, 251].

5.5. Velocity perturbations

Like all mass enhancements, WKBJ spirals will induce velocity perturbations, which can be qualitatively understood already from fig. 19. Because of angular momentum conservation the particles will have a tangential velocity lower than the mean near their apogalacticum and larger near the perigalacticum. Thus from the arrangement of the orbits it can be seen that the tangential velocity will in the mean be larger outside the arm compared to the inside. Similarly for the case of trailing spirals, one can see that the radial motion is toward the center if the arm is inside corotation and outwards if it is outside. The opposite is true for leading waves.

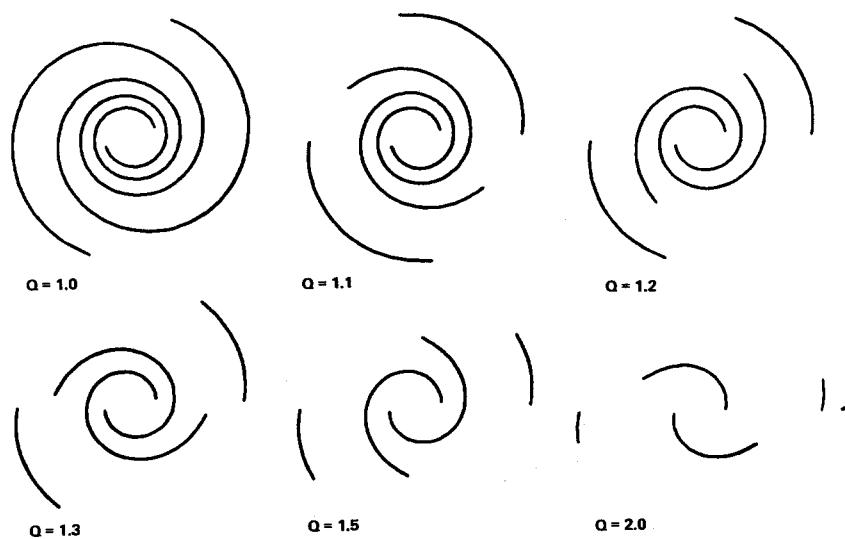


Fig. 21. Extent of the spiral solutions of the Lin-Shu-Kalnajs dispersion relation for six different values of Q .

The magnitude of these perturbations can be easily calculated in the framework of the WKBJ approximation. Unfortunately the comparison with observations did not prove as straightforward (see section 19).

This first version of the WKBJ theory is certainly not a final solution to all spiral structure problems, as had been initially hoped by some of its most ardent supporters, but it has still proved useful in many ways. When not pushed beyond its limits it remains a reasonable first-order solution or starting point, while the simplicity of the equations involved permitted fast estimates and evolution. Its comparisons with observations of the nearby big spirals and our own Galaxy were not “badly afoul” [272] and, furthermore, it paved the way to the gas and shock calculations which we will discuss in the next section.

6. Shocks

The linear WKBJ formulas show that the arm density contrast is inversely proportional to the square of the velocity dispersion of the responding population. Since there is a general correlation of stellar age with velocity dispersion [301], a small density contrast should be expected for old populations and a large one for young ones and for the gas. It was shown first in [88] and later in [219, 254], that gas responds non-linearly to all but the weakest forcings. A tightly wound spiral, assumed to be present in the old stellar disk, is used to drive the galactic gas, assumed to be isothermal with sound speed of about 10 km/s. The response of the gas along a typical streamline (fig. 22a) shows that shocks occur. It was thus proposed in [219] that the gas will undergo at the shock a very strong compression, which could trigger gravitational collapse of those gas clouds which are ripe for it, thus resulting in star formation. The density falls quite rapidly after the shock, so that star formation is confined to a rather narrow region. The dust will also experience a similar compression. If the gas rotates faster than the spiral pattern, i.e., if it is located within corotation the shock should be at the inner (concave) side of the arm (fig. 22b), while if it rotates slower, at its convex side. Thus a neat and clear picture was presented [219]. A sharp HI-peak and a narrow dust lane should be located at the concave (or convex) side of the arm, followed by a somewhat larger band of young stars and HII-regions. Indeed it has been known for a long time that these bright components delineate the spiral arms [e.g. 16]. The argument that compression occurs in the arms has been nicely corroborated at radio wavelengths, first in M51 where the radio continuum emission outlines a spiral coincident with the dust lanes [191], and later in M101 and M81, where the HI emission is likewise concentrated in the arms [2, 223].

Two important assumptions underlie the above gas dynamical results. The first is that the driving potential wave is tightly wound and the second that the interstellar medium can be adequately described as an isothermal gas. It is a great pity that these were introduced in the initial stages, since the results in many minds have been tightly associated with them. However, most of the results still stand if these assumptions are relaxed, e.g. it has been shown a number of times that a very open or even barred forcing can drive a similar response. This will be discussed in some detail in section 12.

Our picture of the interstellar medium has also evolved since the late sixties [193 and references therein]. We now picture it as quite violent and lumpy with large variations in temperature and density. On the cold end we find quite dense molecular clouds and at the very hot end the coronal gas ($T \sim 10^6$ K), heated by supernovae, and interconnected around the cold clouds like swiss cheese around its holes. This brew would be indeed very badly approximated by an isothermal gas with effective sound speed of 8–10 km/s. A new description of the gas was thus called for, and several attempts have been

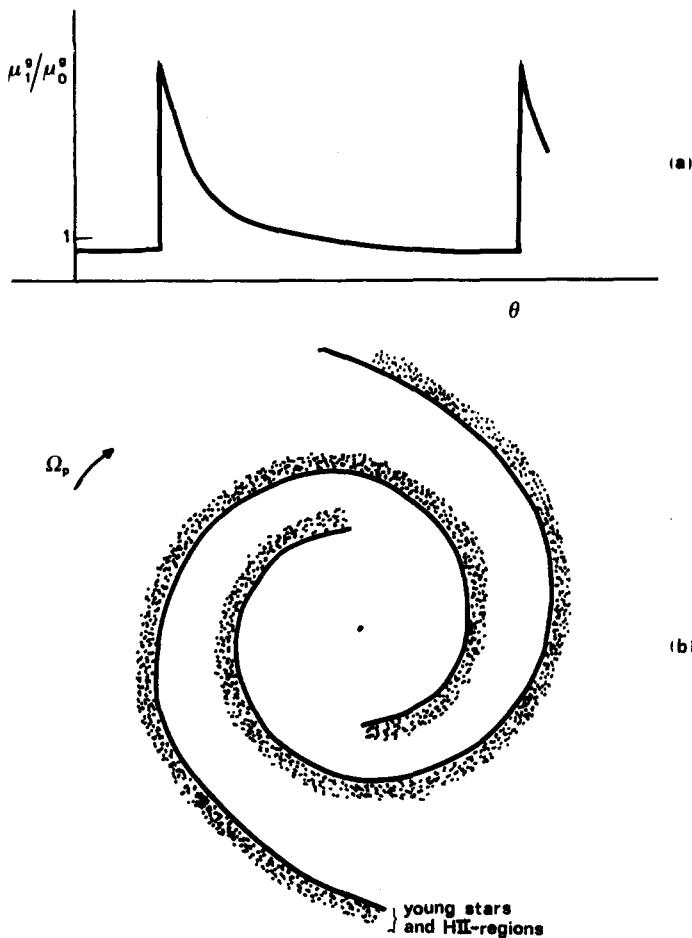


Fig. 22. Response of the gas (a) along a typical streamline and (b) in the plane of the galaxy. In (b) the solid line delineates the shock and the dotted area the region of young stars and HII-regions.

made by modelling it as a collection of clouds colliding inelastically [240, 241, 192, 46]. Since the observed velocity dispersion of interstellar clouds is about 8–10 km/s, the results using the cloud/particle picture of the gas are not drastically different from those using the isothermal gas picture. As shown in [160], the major characteristics of the response, and in particular the position of the shock, now defined as the location of collisions, remain basically the same. The shock region now loses its sharpness and becomes more ragged. Although this lumpy gas picture too will predict age gradients across an arm, these can now (because of the raggedness and the non-stationarity) be of significance only in a statistical sense and not individually.

7. Radial wave transport

The WKBJ analysis, discussed in section 5, and its logical sequitur into the non-linear analysis of gas response, sketched in section 6, generated a lot of interest in the early seventies. The discussion of the spiral structure of our own Galaxy [169, 163], of other galaxies [255, 221, 299], and the seemingly

convincing corroborations in the early data from high resolution radiotelescopes like the Westerbork array all contributed to this.

Yet the theory did not stand on a solid foundation. A discussion of the group velocity, which can be easily worked out for tightly wound wave packets, provided a major embarrassment [269]. Using little beyond the WKBJ dispersion relation and the definition of group velocity, it can be shown [269] that a packet of WKBJ waves propagates radially. Short trailing waves propagate from the region around corotation to the Lindblad resonances (this direction is reversed for long waves, and all directions reverse if the waves are leading instead of trailing). At the Lindblad resonances, provided they exist within the disk, the waves will be damped [182, 184]. This will happen in an uncomfortably small time scale of a few galactic revolutions. If no ILR is present the wave will reflect at the center of the galaxy, changing from trailing to leading (fig. 30).

The implications of propagation of WKBJ density-wave packets were “shatteringly destructive” [176] for the asymptotic theory, but very constructive for the theory of spiral structure itself. Now the idea of quasi-stationarity (i.e. the QSSS hypothesis) can only be maintained if the waves are somehow replenished. Driving mechanisms have to be found, or sources of energy to be tapped, if spiral waves are to be saved from extinction. This led to considerations of, e.g., local instabilities in the disk, tidal forcing from outside, or driving by large scale distortions of the disk like bars or ovals. We will discuss all these replenishment mechanisms in the next sections.

This is not the only damping experienced by density waves. As we have seen in section 6, gas can respond non-linearly to a given forcing and can shock. Contrary to stars it can dissipate energy at the expense of the wave, which will survive for only a few times 10^8 years [133, 222]. One more reason to search for energy reservoirs!

The propagation of angular momentum is derived and pictorially and very clearly explained in [176]. Resonant stars are the sources and sinks of angular momentum and energy. At the ILR they give to the wave and take away from it at corotation and OLR. Trailing waves can carry angular momentum outwards and in doing so decrease the rotational energy of the galaxy in favor of its random motions. This is in no way in conflict with the direction of propagation of a WKBJ wave packet [269]. Indeed such waves have negative energy and angular momentum inside corotation (and positive outside). Thus, as the wave propagates inwards, positive angular momentum is transferred outwards. To consider this mechanism as a possible way of maintaining spirals, however, is to place undue trust on a small percentage of stars. These can not emit or absorb energy and angular momentum without changing their orbits and becoming gradually trapped into quasi-periodic orbits. After this relatively short trapping time (of the order of the epicyclic period) they can no more act as sources or sinks.

8. Shearing sheet

Many of the results previously mentioned, as well as those to be discussed in the next few sections, can be nicely understood in the framework of a very simple model [91, 130, 269] where its advantages and limitations have been discussed, [274]. Let us consider a thin strip of a cold galaxy whose unperturbed velocity field is a shear flow and assume its distance from the center, r_0 , is large enough for its curvature to be neglected. Let us furthermore choose Cartesian coordinates (x, y) such that the x -axis points outwards, and the y -axis forwards in the direction of rotation. The equations of motion of

a particle undergoing small oscillations ($\Delta r \ll r_0$ and $\Delta\vartheta \ll 1$) can be written as

$$\begin{aligned}\ddot{x} - 2\Omega_0\dot{y} - 4\Omega_0 A_0 x &= f_x \\ \ddot{y} + 2\Omega_0\dot{x} &= f_y\end{aligned}\tag{8.1}$$

where

$$x = r - r_0, \quad y = r_0(\vartheta - \Omega_0 t),$$

Ω_0 is the angular velocity at radius r_0 , $A_0 = -\frac{1}{2}r_0(d\Omega/dr)_0$ is Oort's constant [205], and f_x and f_y are the radial and tangential components of a force which is either external or due to self-gravity. Let us now perturb this system with an instantaneous impulse, lying in the x -direction and varying as $\cos kx$. Strips parallel to the y -axis and of alternating positive and negative density excess will be formed. Using eqs. (8.1) and calculating the force due to self-gravity, one can determine the time dependence, $\tilde{x}(t)$, of the x -displacement as

$$\ddot{\tilde{x}}(t) + \kappa^2 \tilde{x}(t) = 2\pi G\mu_0 k \tilde{x}(t)\tag{8.2}$$

from which one gets the dispersion relation

$$\omega^2 = \kappa^2 - 2\pi G\mu_0 k.\tag{8.3}$$

Thus a particle would gyrate around its mean position with a frequency which, because of self-gravity, is smaller than the epicyclic frequency.

The above results hold for the extreme case of cold particles, i.e., particles with no velocity dispersion. Similar dispersion relations can be obtained in somewhat more realistic, yet still easy to handle, cases. For example hot gas can be described by adding a pressure term on the right-hand side of (8.2), so that the corresponding asymptotic dispersion relation becomes:

$$\omega^2 = \kappa^2 - 2\pi G\mu_0 k + c^2 k^2\tag{8.4}$$

where c is the sound speed. Similarly an asymptotic description of hot stars can be obtained with the help of the reduction factor $F_\nu(x)$ introduced in section 5,

$$\omega^2 = \kappa^2 - 2\pi G\mu_0 k F_\nu(x).\tag{8.5}$$

This is the same as the Lin–Shu–Kalnajs dispersion relation for $m = 0$.

An alternative compromise is to use “soft gravity” [194, 80], thereby simulating the effects of velocity dispersion by reducing the gravitational force between two stars. This is accomplished by replacing the Newtonian potential $-GM/R$ by $-GM/(R^2 + a^2)^{1/2}$, where a is the “softening length” or “vertical offset”. This last name comes from the fact that the softened gravity force is equal to the true Newtonian one that would be felt if one of the two particles was displaced from $z = 0$ to $z = a$, while

keeping their horizontal distance R fixed. The corresponding dispersion relationship is

$$\omega^2 = \kappa^2 - 2\pi G \mu_0 k e^{-ak}. \quad (8.6)$$

The three dispersion relations are compared in fig. 23. As could be expected they have fairly similar behavior. For large k values soft gravity, compared to hot stars, gives too small oscillation frequencies while hot gas gives too big ones. Notice also that in the latter case ω can become larger than κ , or equivalently, the wave can extend beyond the Lindblad resonances.

Equivalents of Toomre's stability criterion [268] can also be obtained for the other cases. Stability is ensured if $\omega^2 > 0$ for all values of k . For the cold disk this can only be true for

$$k < k_{\text{crit}} = \kappa^2 / 2\pi \mu_0 G. \quad (8.7)$$

For a gaseous disk a minimum sound speed of

$$c_{\text{min}} = \pi \mu_0 G / \kappa \quad (8.8)$$

and for soft gravity a minimum softening length of

$$a_{\text{min}} = 2\pi G \mu_0 / \kappa^2 e \quad (8.9)$$

are needed for stability.

In the above we applied the initial perturbation along the x -axis. Let us now consider the more general case of a perturbation at an angle γ with the y -axis [274]. Negative values of γ correspond to leading spirals, positive ones to trailing spirals. The unperturbed flow will change γ at the rate $d \tan \gamma / dt = 2A_0$ and the perturbation will shear. Thus a single free wave in such a shearing sheet can not be in a steady state. The equations of motion will now be written as

$$\begin{aligned} \ddot{x} - 2\Omega_0 \dot{y} - 4\Omega_0 A_0 x &= f_x \sin \gamma \\ \ddot{y} + 2\Omega_0 \dot{x} &= f_y \cos \gamma \end{aligned} \quad (8.10)$$

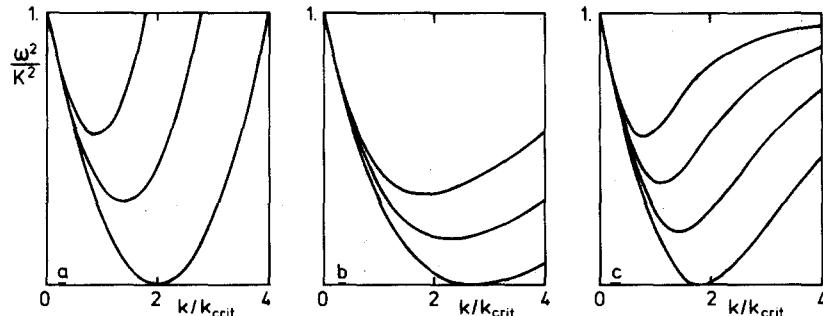


Fig. 23. Dispersion relation for (a) hot gas, with $c = c_{\text{min}}, 1.2c_{\text{min}}, 1.5c_{\text{min}}$, (b) soft gravity, with $a = a_{\text{min}}, 1.2a_{\text{min}}, 1.5a_{\text{min}}$ and (c) hot stars, with $Q = 1.0, 1.2, 1.5$ and 2.0 .

where f_\perp is the force perpendicular to the wave crests. Introducing the normal displacement

$$\xi = x \sin \gamma + y \cos \gamma \quad (8.11)$$

as a new variable, we can combine the two equations in one

$$\ddot{\xi} + S(\gamma)\xi = 0 \quad (8.12)$$

where

$$S(\gamma) = \kappa^2 - 8\Omega_0 A_0 \cos^2 \gamma + 12A_0^2 \cos^4 \gamma - 2\pi G\mu_0 k F \quad (8.13)$$

and the value of F depends on the case considered

$$F = 0 \quad \text{No self-gravity}$$

$$F = 1 \quad \text{Cold disk}$$

$$F = 1 - c^2 k / 2\pi G\mu_0 \quad \text{Hot gas}$$

$$F = e^{-ak} \quad \text{Soft gravity}$$

$$F = F(\gamma, Q, \dots) \quad \text{Hot stars.}$$

9. Swing amplification

Let us now, following [274], discuss the solutions of eq. (8.12) in various cases. These can be parametrized with the help of the dimensionless shear rate

$$\Gamma = \frac{2A_0}{\Omega} = -\frac{r}{\Omega} \frac{d\Omega}{dr}$$

and the X parameter

$$X = \lambda_y / \lambda_{\text{crit}} = k_{\text{crit}} / k_y = \kappa^2 / 2\pi G\mu k_y.$$

The value of Γ depends on the rotation curve. For a power law $V = V_0 r^\alpha$ we find $\Gamma = 1 - \alpha$. Specific cases are rigid rotation for which $\Gamma = 0$, flat rotation curves for which $\Gamma = 1$ and Keplerian ones for which $\Gamma = 1.5$. X can be roughly understood as a local measure of the ratio of disk mass (i.e., mass partaking in the oscillations) to total mass (i.e., disk + halo + bulge + ...).

Shear, as can be seen from fig. 24, lowers the effective spring rate (8.13). The opposite effect is obtained by raising Q or X , i.e. by lowering self-gravity forces with the help of either halos or velocity dispersions. Very high values of Q or X are necessary to keep the spring rate positive for all values of γ .

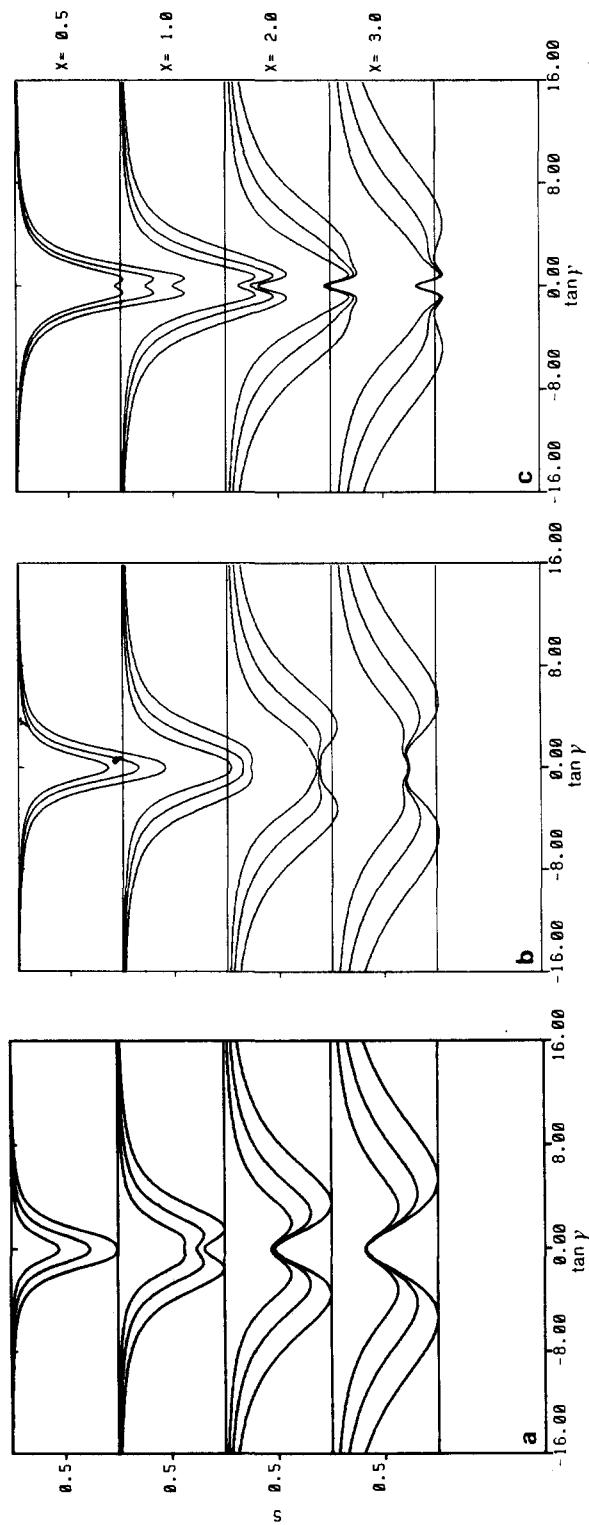


Fig. 24. Spring rates [as in 274, 275] in units of κ^2 . Values are $\Gamma = 0.0$ (a), $\Gamma = 0.5$ (b), $\Gamma = 1.0$ (c); $X = 0.5$ (top panels), $X = 1.0$ (second from top), $X = 2.0$ (third from top), $X = 3.0$ (bottom panels) and $Q = 1.0, 1.2$ and 1.5 (bottom, middle and top curves respectively).

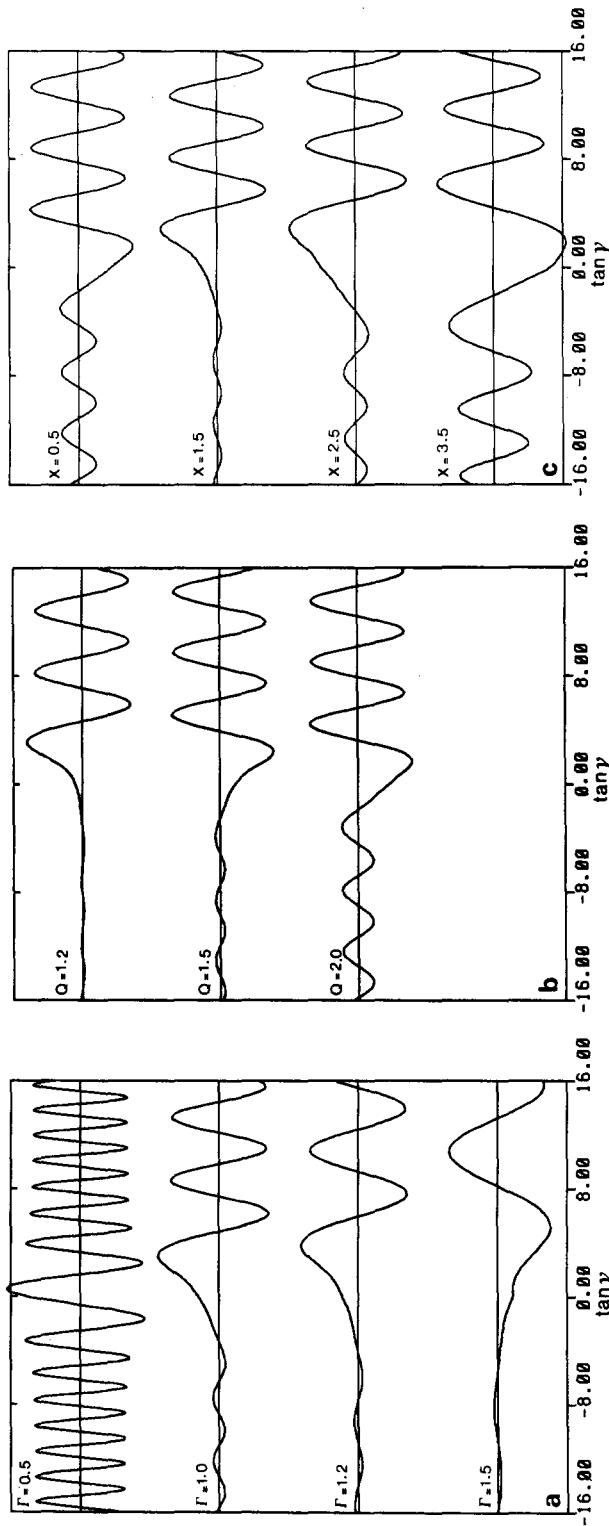


Fig. 25. Amplification factor from a full swing for optimal arrival phase. The scale is arbitrary and the amplification can be estimated from the ratio of the trailing and leading amplitudes. (a) shows the effect of shear and is calculated for $X = 2.0, Q = 1.5$ and $\Gamma = 0.5, 1.0, 1.2$ and 1.5 ; (b) shows the effect of Q and is calculated for $\Gamma = 1.0, X = 1.0$ and $Q = 1.0, 1.5$ and 2.0 ; (c) shows the effect of X and is calculated for $\Gamma = 1.0, Q = 1.0$ and $X = 0.5, 1.5, 2.5$ and 3.5 .

Some of the corresponding solutions of eq. (8.12) are given in fig. 25. One is impressed by the high amplifications thus obtained. Since these occur when the wave is swinging from leading to trailing they have been termed *swing amplifications*. They are most important when the shear is largest and/or when self-gravity is highest.

Although most informative, these plots are not the most compact way of conveying information. This can be better achieved by reading off every such graph the amplification that can be expected for given X , Γ and Q values provided the arrival phase is optimum. Figure 26 contains such information. Several remarks can be made:

- Very large amplifications, reaching several times 10, are possible for reasonable values of the parameters.

- Significant amplifications are found even for relatively hot disks.

- For cool disks the amplification is high even for a dimensionless shear rate as low as 0.5. For hot disks only large shear can be efficient.

- For X larger than a limiting value, X_{lim} , there is hardly any amplification. Furthermore this value depends on the shear rate but not on Q . Thus for $\Gamma = 0.5$, $X_{\text{lim}} = 1.5$; for $\Gamma = 1$, $X_{\text{lim}} = 3$; while for $\Gamma = 1.5$, $X_{\text{lim}} = 6$. In other words the amount of halo necessary to stop swing amplification depends on the form of the rotation curve. Since rotation curves are rising, often slowly, in the inner parts to become flat further out, one would expect a relatively small halo (or bulge) percentage to be able to kill swing amplification in the inner parts while a much larger percentage would be necessary in the outer parts.

In the above we have, for reasons of continuity and clarity, followed the presentation of [274], although swing amplification was first discovered and advocated in [91] for the case of a gaseous shearing sheet. It was concluded that "even when all purely radial disturbances are stable, there are still some sheared waves whose amplitudes grow by factors of more than 100. Moreover, this growth occurs

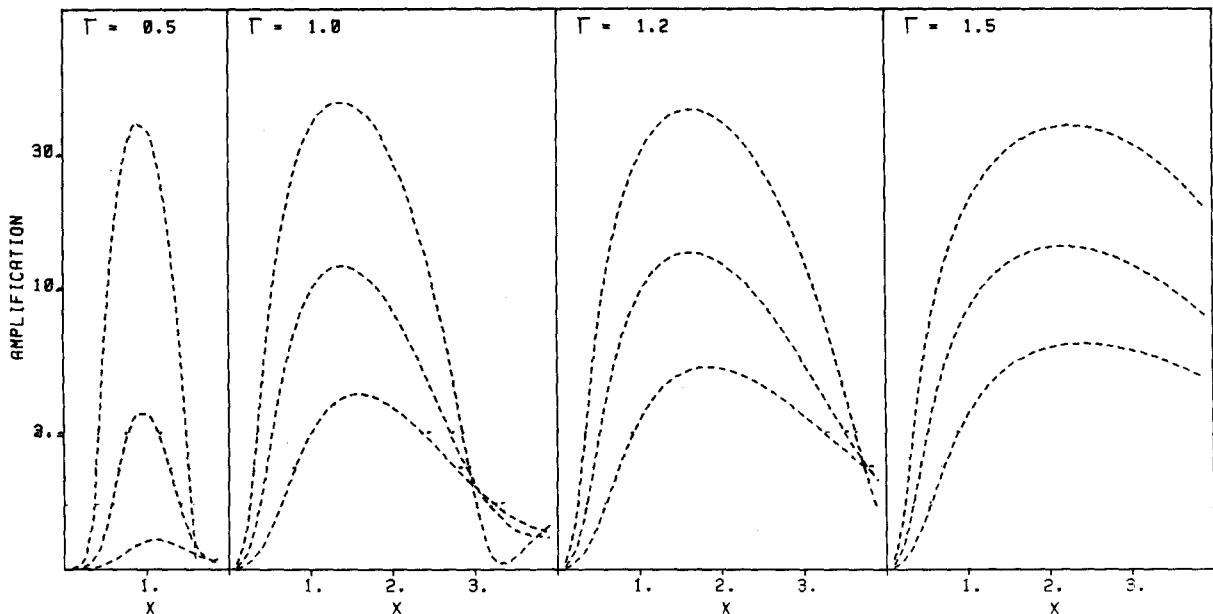


Fig. 26. Maximum Growth Factor from a full swing amplification as a function of X for $Q = 1.2, 1.5$ and 2.0 and $\Gamma = 0.5, 1.0, 1.2$ and 1.5 .

for waves of a well-defined wavelength and begins as the lines of constant density are sheared past the radial direction." Thus a theory of spiral structure based on the idea that spiral arms are short-lived, constantly forming and dying, was put forward. Barely a year after that a non-asymptotic analysis of a now stellar shearing sheet including random motions [130] showed equally large amplification factors. Although the mathematics are obviously more complex, the results are quite similar. The principal difference between the gaseous and stellar cases is the strong phase damping of the latter. Thus after reaching an often spectacular maximum, the amplitude of the disturbance drops considerably. This was further discussed and nicely illustrated in [274], where fig. 8 shows an impressive amplification of an initially leading wave, followed by a no less important decay due to group transport and to the subsequent damping at the Lindblad resonances. It is thus clear that swing amplification will not provide any stationary spirals, unless some input in the form of leading spirals is continuously provided. It can however provide spectacular global patterns, albeit short-lived.

10. Feedback cycles

In order to have a complete feedback cycle one needs to piece together propagation of groups of waves, an amplification mechanism, and a reflection in the central parts. Indeed if the amplification provided exceeds that lost by dissipation, a growing solution is possible. Two such cycles have been so far proposed.

10.1. WASER

The WASER, or Wave Amplification via Stimulated Emission of Radiation, was, ironically, first presented – though not named – in [269] and fully worked out in the framework of the WKBJ theory, discussed and baptized in [183, 185, 186, 187, 188]. It is schematically pictured in fig. 27. All three waves are trailing, one long and two short. At corotation the outgoing long trailing wave stimulates the emission of two short trailing waves. One propagates outwards and will be dissipated at the OLR while the other propagates inwards. The amplitude of these waves, and therefore the amplification factor (defined as the ratio of the amplitude of the inward short trailing wave to the amplitude of the incident long wave) depends critically on Q , as is shown in fig. 28. If $Q < 1$ the WASER mechanism is quite efficient; this case is only of academic interest, however, since axisymmetric instabilities will prevail. For $Q = 1$ it gives an amplification factor of ~ 1.5 while for larger values of Q it becomes rapidly useless.

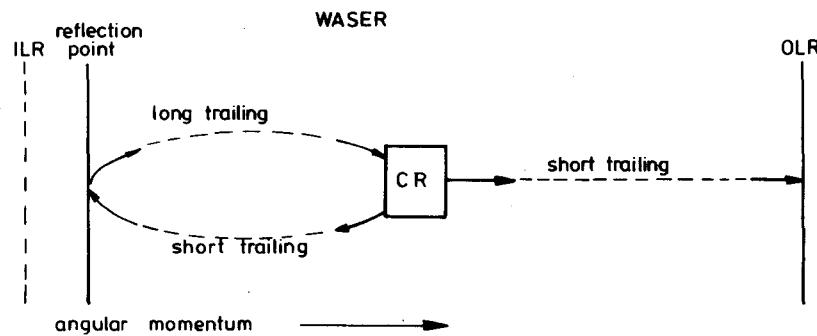


Fig. 27. The WASER feedback cycle.

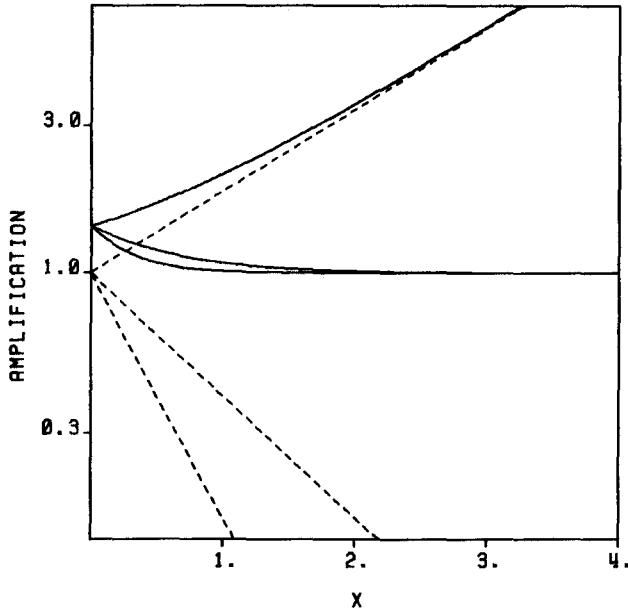


Fig. 28. Amplification by the WASER feedback cycle, from the approximations given in [186]. The amplification is given for the reverse (full line) and forward (dashed line) emitted wave at corotation, as a function of X for a flat rotation curve and $Q = 0.9, 1.2$ and 1.5 .

Significantly larger values of the amplification factor are obtained if the extra boost from shear is included [91, 159].

The cycle is closed by the reflection of the inward travelling short wave into a long one. This is possible if Q rises inwards fast enough (but on the other hand not too fast to invalidate the WKBJ approximation). Then there is a minimum radius past which the wave cannot propagate, while the long and short branches of the Lin–Shu–Kalnajs dispersion relation have equal wavelengths there. This sharp inwards rise of Q is by no means unphysical and simulates a hot center or bulge. Thus a fresh long wave is present to start the cycle anew. Obviously, if this mechanism is to work, the ILR must not lie on the path of the wave so as not to absorb it. This may mean either that there is no ILR or that it is situated at a smaller radius than the reflection point.

The WASER gives much smaller amplification factors than the swing amplification discussed below. Only if all help from shear [91, 159] and interaction with halo stars [187] is included it may yet prove in certain cases to be adequate to keep spiral structure alive.

10.2. Swing amplifier

This feedback cycle was introduced and nicely presented in [274]. What is astonishing is not that most of its ingredients were already well known for a long time [91, 130, 269], but that the piecing together, with hindsight so obvious, took so long to be made and get established. The feedback cycle (fig. 29) starts from a leading wave propagating outwards towards corotation. At the same time it unwinds to an open structure and then starts winding in the trailing sense. During this phase it is heavily amplified. We then have two trailing waves, one propagating from corotation inwards and the other propagating outwards towards OLR where it is dissipated. When the inwards trailing one reaches the

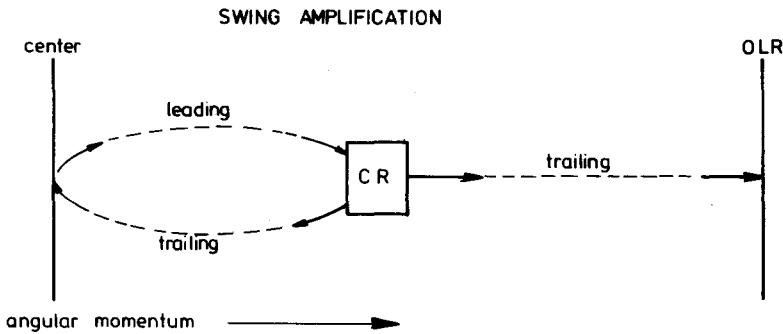


Fig. 29. The swing amplifier feedback cycle.

center it is reflected into a leading wave, propagating outwards. Thus the cycle is closed. A handwaving explanation of the latter reflection is given in fig. 30. Of course if an ILR is present the wave is absorbed before being reflected and the cycle is cut. Individual waves will in this case still be amplified, but will eventually die out.

The amplification factor of this feedback cycle is much larger than that of the previous one. It depends drastically on both the length ratio

$$X = \lambda_y / \lambda_{\text{crit}}$$

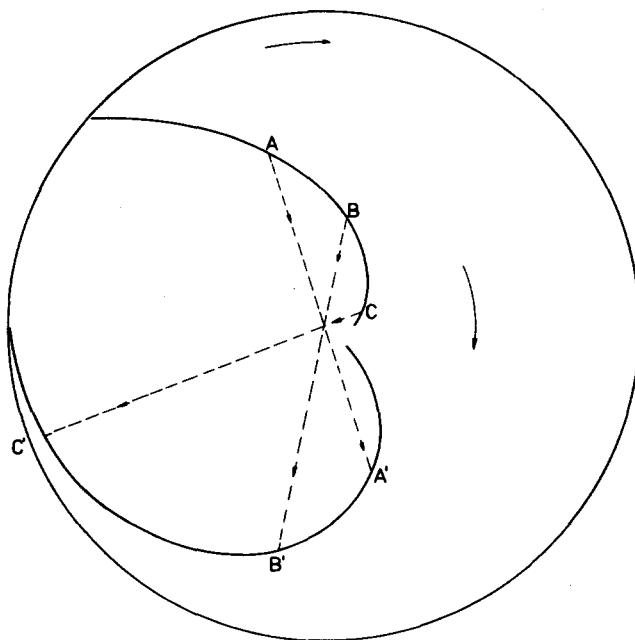


Fig. 30. Handwaving explanation of the reflection of a trailing wave at the center of a galaxy. If at a time $t = t_1$ the points A, B, C, ... are on a trailing spiral and we assume the inwards radial velocity to be independent of the distance from the galactic center then at a later time $t = t_2$ the points A', B', C', ... will be on a leading one, since the distance AA' should be equal to BB' and CC'.

and Q , as can be seen in fig. 26. For $\Gamma = 1$ maximum amplification is obtained for $X \sim 1.5$, while for $X > 3$ it is powerless.

A leading wave should thus be present in all galaxies in which this mechanism is operating. Its amplitude will be much smaller than that of the trailing wave, particularly in cases with strong amplification, and thus it will be very difficult to discern. The best chances would be in galaxies with small amplification factors, i.e., relatively large halos and/or hot disks, so that the amplitude of the trailing and leading waves are of the same order. Furthermore there should be no shocks so that the small amplification factor could be sufficient to maintain the wave. Then the superposition of the trailing and leading waves should give an interference pattern (see fig. 31) and Fourier analysis should reveal the existence of the two components.

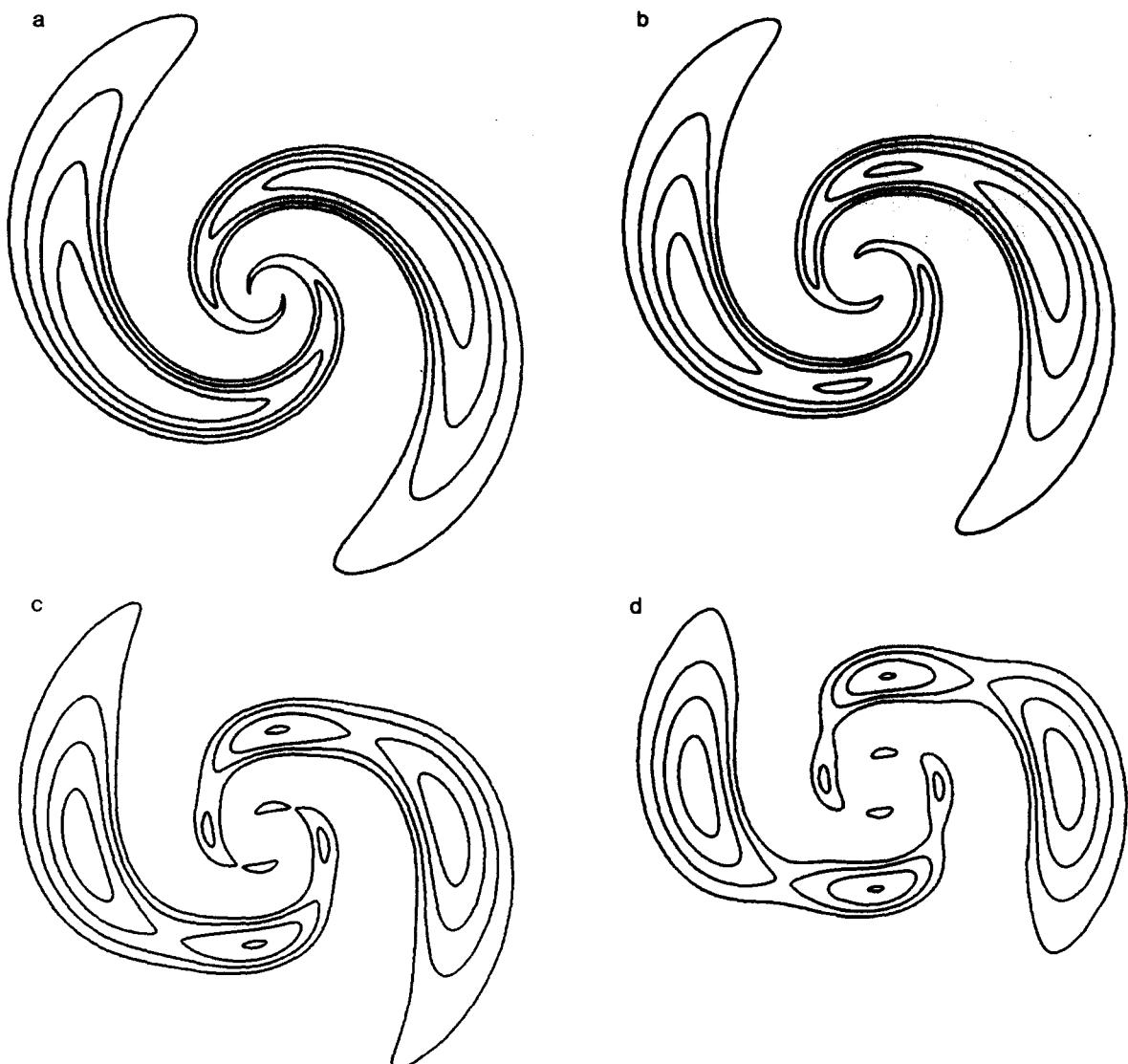


Fig. 31. Superposition of leading and trailing spirals. (a) 100% trailing; (b) 90% trailing 10% leading; (c) 80% trailing; (d) 70% trailing.

The parameter X depends on the number of arms m . Setting $\lambda_y = 2\pi r/m$ we get

$$X = rk^2/2\pi Gm\mu.$$

Thus it follows that the number of arms that will prevail will depend on properties of the axisymmetric background of the galaxy like its rotation curve and the percentage of its mass present in the disk. For example for the Schmidt model of our Galaxy [239] if all the mass is in the disk, the maximum growth factor is reached for $m = 2$ and $m = 3$. If on the other hand only half the mass is in the disk then maximum amplification will occur for $m = 4$ or 5 , i.e. a multi-armed structure. Thus the number of arms gives some clues about the halo mass. Higher m values are more liable to have an ILR, however, thus disarming the feedback cycle.

11. Driving by companions

Tidal forcing from an external galaxy was one of the first cures proposed [269] for the ailing WKBJ theory discussed in section 5. The first hints of its relevance were already presented and discussed [171, 213, 214] before it needed to come to rescue as replenisher of decaying spiral waves. From that early work and from subsequent more complete discussions [277, 270 and references therein] it was established that direct, relatively slow and close passages form nice trailing shapes, which subsequently disperse. In these calculations both galaxies are represented as point masses surrounded by test particles. This rather harsh approximation eases computations considerably, while its influence was underestimated since it was thought that the effects of encounters would be significant mainly in the outer parts. Thus the spiral could not and did not penetrate in the inner parts of the simulated disk, creating some unwarranted pessimism [272, 274]. Furthermore to include self-gravity in the calculations requires a stable disk as a starting configuration, something which had not yet been achieved at the time.

Self-gravity was included about five years later in analytical stationary treatments, where the companion is asked to stir the spiral continuously, i.e., it is a satellite on a circular [6, 93] or near-circular [92] orbit. In this case the angular velocity of the companion defines the pattern speed and the positions of resonances, and angular momentum is exchanged between it and the wave. This approach sets stringent constraints on the position and velocity of the companion, but can work with smaller forcing amplitudes than in the case of close encounters. Furthermore the spiral is guaranteed to last. If the companion is direct the $m = 2$ component of its forcing will conspire with the $m = 2$ ILR and excite a two-armed trailing wave. This can not be the case for a retrograde companion [6]. Here we must rely on the $m = 1$ component of the forcing (non-zero in the case of non-spherical or inter-penetrating mass distributions or close tides) to conspire with an $m = 1$ type ILR, since the $m = 2, 3, \dots$ inherently rotate in the forward direction. In such cases a one-armed leading spiral will be excited. These are undeniably special circumstances leading to a non-standard type of spiral, although it has been suggested [136, 5, 256] that one of our nearest neighbors, M31, is experiencing precisely such a forcing from its close companion M32 and has as a result developed a one-armed leading spiral. This spiral is very tightly wound, a sort of a pseudo-ring, an effect which is further enhanced by the large inclination angle of the galaxy. Angular momentum is exchanged between the spiral wave and the companion. The wave gains negative angular momentum while the companion loses it and, since it is on a retrograde orbit, it spirals inwards [5].

Although very suggestive the above results apply only to a rather restricted number of cases since flybys can not be thus analyzed. The only selfconsistent treatment in which the constraint of quasi-stationarity has been ignored can be found in [274]. A stable and yet still responsive axisymmetric disk is necessary, so as not to have to disentangle the response from the products of the instability. This is possible only if the swing amplifier cycle is stopped by an ILR, since a massive halo, while ensuring stability, will at the same time damp all responses. A constant circular velocity disk is guaranteed to have ILR's for all $m > 1$ (see also section 14 for a discussion of the modes of this disk). Fortunately, as we have seen in section 10, the X parameter is only half as big for $m = 2$ as it is for $m = 1$. It was thus possible, by freezing half the mass in a rigid halo and using a Q of 1.5, to curb the $m = 1$ instabilities while leaving the $m = 2$ waves in full swing. An $m = 2$ forcing was applied to this disk. Its form resembles the $m = 2$ component of a tidal force while its amplitude reaches at the most 1% or 2% of the central force. The resulting $m = 2$ spiral is quite spectacular, though transient since the feedback cycle cannot be closed. A couple of such examples are given in fig. 1 of [274], which is reproduced in fig. 32. Provided the switching on and off of the forcing was sudden enough the maximum perturbation can initially be found near the center and only in later times do the outer parts get severely distorted. The position of several particles has been marked on the figure by open or filled circles and can be followed throughout the evolution. One can thus see that the perturbations start off as material arms. The particles stay within the arm for quite long periods of time, roughly three quarters of a revolution or so, and only after that, when the spiral has become very tightly wound do they leave it and start travelling towards the other arm. Of course the nearer the particle to the center the sooner this will happen, since clocks in the central parts run faster. Thus at time 4 or 5 the inner parts of the spiral can

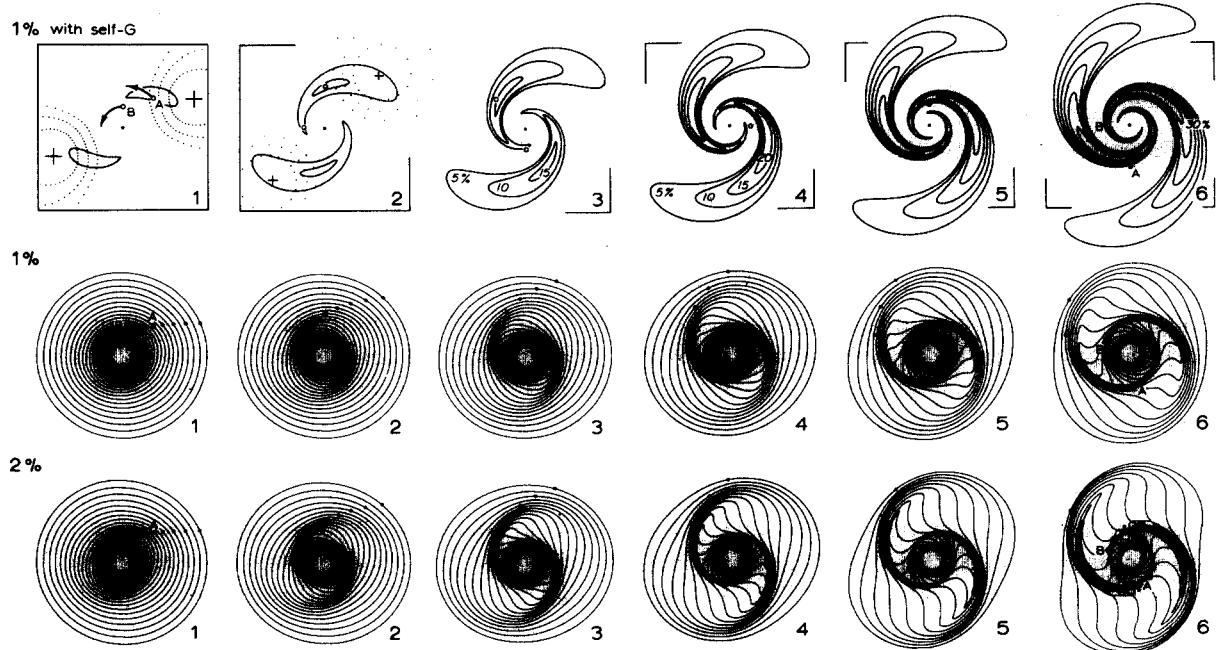


Fig. 32. From [274]. Responses of a $V = \text{constant}$ disk of stars to a transient gravity forcing from the imposed masses sketched at the upper left. The top row shows the excess densities among average stars. Other rows report the fates of various "cold" circles of test particles. The top rows presume that the maximum tidal force on particle A reached 1% of its central force; the third row supposes 2%.

be called density waves and the outer ones material arms. So ironically material arms and density waves collaborate and coexist nicely in one and the same frame.

The spiral cannot always reach the inner parts of the galaxy. If the passage of the companion is too slow its action is inefficient (compare the motion of a pendulum under a sudden or gradual forcing). The innermost particles will be less perturbed since their angular frequency is largest. Hence a not too fast passage may still excite an important spiral in the outer parts while leaving the inner parts more or less untouched.

It is thus clearly shown in [274] that galaxies which are stable because an ILR prevents the closure of the swing amplifier cycle can still respond quite vigorously to tidal forcing, provided they are not immersed in too massive a halo. This results in a global, though transient, spiral structure. If they are, on the contrary, immersed in a very massive halo they are unresponsive.

12. Driving by bars

It is questionable whether bars should be considered separately as a driving agent for spiral structure, since as long as their amplitudes remain small they may well be on equal footing with spirals, both resulting from the same instability. Nevertheless since the complete selfconsistent problem is very complicated, in particular if the effects of gas are included, it makes sense to separate the problem of the response of a gaseous and/or stellar disk to an imposed bar potential from that of bar formation.

The idea that a rotating bar or oval is an efficient way of driving spirals was put forward more than a decade ago [172, 269] and was tried out in numerous subsequent analytical and numerical calculations. These showed clearly that the result is indeed a spiral provided either the potential perturbation is growing (more precisely is non-steady) or the medium is dissipative. Put differently, the density maxima of the response need not coincide with the potential minima of the forcing. Thus gas flow calculations could proceed without the necessity of assuming a tightly wound spiral forcing (as was done in [e.g. 219]). This confirmed nicely the picture suggested in [131, 132] namely that a very open or bar-like perturbation in the stellar distribution drives a tightly wound response in the gas and young stars.

The first attempts were analytical, and discussed quasi-stationary tightly wound bar-driven gaseous spirals in an axisymmetrically unstable [83] or marginally stable disk [165].

The selfconsistent quasi-stationary response of a galaxy composed of both gas and stars to a growing and rotating bar was discussed in [7]. In many cases the response of both components was bar-like up to corotation, where it turned abruptly into a trailing two-armed spiral ending roughly at the OLR. In other cases, however, the spirals started well inside corotation. In no case did the bar-like response extend beyond the corotation radius. It was found that the potential of the stellar spiral provides a considerable fraction of the total forcing, and can therefore not be neglected in the calculations since the shape of the gaseous spiral is greatly influenced by it.

The form of the stellar spirals was found to be invariably more open than the gaseous ones, and their along-the-arm maximum occurs much closer to the center. Within the first quarter of a turn the thin gaseous arms are embedded in the broader stellar ones and leave them only at larger radii, where the stellar spiral is much weaker. Furthermore, a lower relative gas surface density leads to a larger radial extent and a longer winding length of the gaseous spiral. Also, the pitch angle of the arms increases outwards as we approach the OLR. Finally the arms are more open for faster growing forcings in good agreement with epicyclic orbit calculations [172, 7].

Transfer of positive angular momentum occurs from the bar to the spiral. If the bar is short it will

have negative energy and angular momentum and will thus, by losing positive amounts of each, be amplified. The opposite will be true for a sufficiently long bar. Thus a limit to the extent of the bar will be set [7].

The response of a distribution of test particles has been followed in [264, 265] to study the spiral structure generated in the stellar component by a growing oval distortion. For a sufficiently high growth rate of the bar a two-armed spiral develops beyond corotation. If the pattern speed is adjusted such that two ILR's occur relatively far apart, the response in the region in between these two resonances is greatly diminished.

The non-linear time-dependent response of the gas to an imposed bar or open stellar spiral has been studied using hydrodynamic codes. A great variety of axisymmetric models, perturbing potentials, and numerical schemes have been employed [218]. We will discuss here only the spiral density response in the gas, and leave the discussion of the flow in the bar region itself (which is much more dependent on the code used) to section 16.

The beam scheme [216, 235] and a quadrupole forcing have been used in [234, 119]. When a steady state is reached the gas forms a two-armed trailing spiral extending over the entire disk. Shocks develop along the inner boundaries of the arms, roughly between the ILR and corotation. Self-gravity in the gas has been included in [26, 121], but, as could have been expected, the results are not much affected by it.

The beam scheme and a model bar consisting of a homogeneous triaxial spheroid have been used in [237] to compare the steady state gas density with the morphology of SB_b galaxies. Broad open spiral enhancements in the gas are formed from weak perturbations, in good agreement with [234]. For strong perturbations, however, with the bar-axial ratio at least 1:3 and the corotation radius within 30% of the end of the bar, the gas responds by forming a conspicuous bar with strong linear offset shocks within it and weak trailing arms beyond it (see fig. 33).

A different picture of the interstellar medium has been adopted in [233]. The time-dependent response of test particles to a rotating oval distortion is found to be bar-like and to reinforce the bar. As

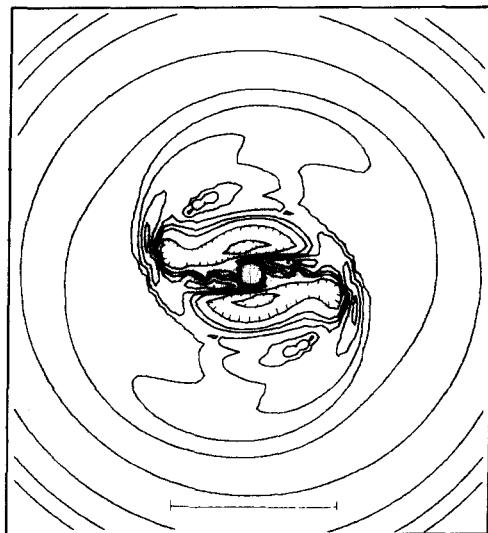


Fig. 33. Steady-state response of gas to a bar forcing [237]. The orientation and size of the bar are indicated by a horizontal line at the bottom. (Reprinted courtesy of R.H. Sanders and A.D. Tubbs, and The Astrophysical Journal, published by the University of Chicago Press; © 1980 The American Astronomical Society.)

a next step, the mass of each particle is distributed over a sufficiently large volume [174]. Thus the particles become fluid elements in the Lagrangian sense, and an artificial viscosity can be defined. The dissipation introduced in this way changes the response. A material bar is now formed within the ILR, and trailing density waves outside it. Although these calculations have a coarse resolution, they clearly show the effects dissipation can have on the response.

The most extensive study of the effects of bar forcing on a lumpy interstellar medium can be found in [240, 241]. Several forms of the axisymmetric background and of the perturbation potentials have been considered. The gas is modelled by “clouds” colliding inelastically, thereby losing a significant fraction of their relative motions. If collisions, and thereby dissipation, are neglected, the response of the stellar distribution shows spirality only when the bar is growing. If collisions are included, a two-armed trailing spiral is formed, extending roughly from corotation to the OLR. Collisions, and in particular those in which a large amount of energy is dissipated, occur predominantly in the arms. These are quite narrow, indicating the presence of shocks, with a pitch angle of about 15° . They become more tightly wound near the OLR, thereby giving the impression of a pseudo-ring, as is often observed in external galaxies. The total winding length of the arms shows little dependence on the bar strength, and is in general greater than the 180° predicted [234] on the basis of 90° per resonance. Due to the torque exerted by the bar, the spiral evolves. The region between corotation and the OLR becomes depleted and a ring is formed at the OLR. These later stages of evolution will be discussed in section 17. Correspondingly the gas exerts a torque on the bar so as to amplify a (negative) energy mode.

Somewhat different ways of modelling a lumpy interstellar medium have been used in [192, 46]. As in previous cases a two-armed trailing spiral response has been found, but the fine details are not the same.

The subject of bar-driven spiral waves has, as we have seen, generated an impressive number of publications. The main point to be remembered is that a spiral can be formed if the bar is growing and/or if the driven population is dissipative. The spiral may evolve in timescales of 10^9 years or so depending on what has been assumed for the gas (processes like gas recycling, or infall, may be important here).

13. Driving by asymmetries

A penchant for order and symmetry, as well as simplicity, has led theoreticians to consider one center in each galaxy, the same for all components. Thus in all calculations discussed in the previous section the center of the perturbing bar coincides with that of the disk. Furthermore, the shape of the bar was chosen bisymmetric, so that only $m = 2, 4, 6, \dots$ components need be considered in the calculations. Unfortunately nature is not always as tidy as theoreticians would like it to be. Late type galaxies in particular show more often than not important asymmetries, and the center of the various components does not necessarily coincide. The most spectacular case is that of offset bars, i.e. bars whose center does not coincide with that of the disk. The displacements between bar and disk centers are often large, sometimes as large as the bar semimajor axis length.

Since such displacements are observed so frequently we should try to understand and explain them. A number of scenarios have been put forward, but none of them backed up by serious calculations. In [67] it is argued that there is no particular reason why primordial gas, before forming a galaxy, should have spherical or even circular symmetry; the amount of asymmetry may thus determine the type of the particular galaxy formed. A second possibility is that displacements are, like bars, a type of instability

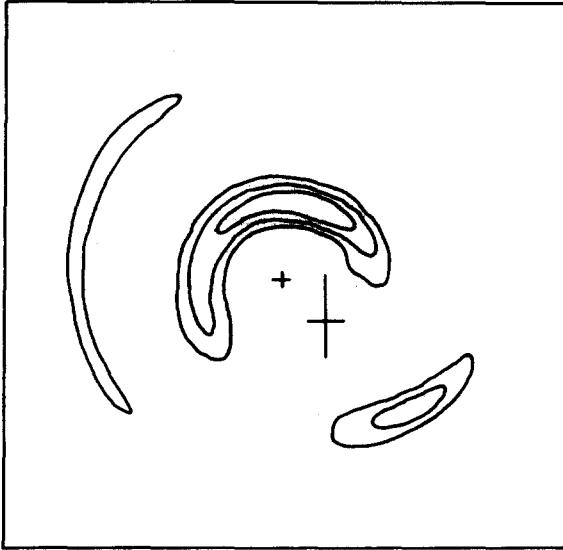


Fig. 34. One example of the response of a gaseous disk to an offcentered bar [45]. The center of the (non-responsive) stellar disk is marked with a small cross. Also marked are the major and minor axes of the bar.

dominant in certain types of disks. A third possibility is that displacements are due to asymmetric accretion of extragalactic material, be it large gas clouds or dwarf galaxies. And a fourth one is that the asymmetry is not real in mass, but only apparent in luminosity, and that the ubiquitous stochastic self-propagating star formation is responsible for it [82]. This last scenario, however, leaves the observed velocity asymmetries unexplained.

Whatever the origin of displacements, if they involve mass, we can decouple the problem of their formation from that of the response of a stellar or gaseous disk to their forcing, just as we did in the previous section for bars. In the asymmetric cases odd components, $m = 1, 3, 5, \dots$, will now be present in the perturbations, together with the even ones found in symmetric bars. Depending on a very large number of parameters, like the mass and eccentricity of the bar, its displacement with respect to the center of the disk, etc., odd or even components will dominate. This obviously reflects on the driven spiral [45], which can be anything between a slightly asymmetric two-armed spiral, or a system with two very unequal arms (fig. 34), or a predominantly one-armed spiral with only a small hint of a second (or third) arm. Indeed all these variations have been observed.

14. Modes

We have seen that spirals may be due to an external forcing. A possible alternative is that they are natural oscillations of galactic disks. In order to investigate this possibility a series of stability analyses were undertaken whose main aim was to find stable or very slowly growing modes in realistic disks. Modes, in contrast to wave packets, do not propagate radially. Thus the problems discussed in section 7 do not affect them. Furthermore global modes, if calculated properly, can offer interesting comparisons with asymptotic theory estimates and, more attractively, promises of global stability criteria.

For all these reasons a large number of mode calculations have been attempted. The results have

been viewed with pessimism [139], however, due to the large growth rates found. These can be reduced by assuming either that a large fraction of the mass is in a rigid non-responding halo or that the disk is very hot, i.e. has large velocity dispersions. Before discussing whether this pessimism is justified or not we will briefly review the results of the various calculations.

A proper mode calculation is not an easy matter. It involves solving the system of the linearized Boltzmann and Poisson's equations with proper boundary conditions. So far only exponentially growing modes in various axisymmetric disks have been investigated. For a fixed number of arms m , a sequence of eigenmodes, each characterized by a radial mode number n , is usually found. In view of the large difficulties involved, a number of more or less justifiable shortcuts and approximations have been made.

The most drastic one is to consider a cold system, i.e. one in which only circular motions are allowed. The problem then becomes fairly tractable and hence several axisymmetric disks have been analyzed in this way [116, 117, 118, 305, 200]. Axisymmetric, as well as non-axisymmetric instabilities have been found, the latter both trailing and leading. The important axisymmetric instabilities make it difficult to sort out the non-axisymmetric behavior in this case.

A more justifiable approximation is to consider a gaseous instead of a stellar disk. The former is easier to work with since less dynamical variables describe the response of the disk. The two approaches are related in the limit of small epicyclic orbital excursions. Hydrostatic pressure is in this case the analogue of velocity dispersions. However as can be seen, e.g. by comparing the corresponding dispersion relations (see fig. 23), gas responds in certain ways quite differently from stars. A most important difference is that gaseous density waves can easily cross the Lindblad resonances and extend throughout the entire disk. This is in sharp contradiction with the stellar case where the existence of an ILR is enough to stop the wave and put out of action the feedback mechanisms discussed previously. In spite of these shortcomings the analysis of various models of polytropic gaseous disks has given very interesting results. In [19] three types of Q profiles have been considered. One in which Q grows with r , one in which it is larger in the center than halfway out in the disk and an intermediate one in which Q is independent of r . The modes are bar-like in their central part and spiral further out. In the case of the first profile they are dominated by the central bar in which the density perturbation is maximum, so that they should be termed barred rather than spiral modes. On the contrary for the second Q profile modes are spiral. In the third case the density perturbation has comparable amplitudes in the bar and spiral. These results are consistent with modes which are barred in the center and spiral further out, with maximum amplitudes in regions of low Q . However the most interesting result is that the growth rates of the most unstable modes can be severely reduced if the disk is embedded in a massive rigid halo or if it has large velocity dispersions. A disk as cool as the solar neighborhood can be stable only if most of the mass is in an extended halo or central bulge. Furthermore it is possible to reach stability, at least for a Q profile such that Q grows with radius, by raising the mean temperature of the disk.

This stability analysis of a gaseous disk was followed by several others [3, 260, 261, 123], confirming its results. The stability of a Toomre disk [267] was studied in [4]. Rather open modes were found while in certain cases their amplitude undulated significantly along the spiral arms. The most rapidly growing modes of the first few azimuthal components have roughly the same pattern speed and growth times comparable with the period of rotation. Furthermore the amount of random motion was found to influence the shape of the mode in the sense that hotter models have more loosely wound and inconspicuous spiral arms, while the amplitude of the density undulates less. Also their pattern speed is lower. Higher order modes are more tightly wound and more undiluted.

Another category of calculations is based on the asymptotic approximation [e.g. 27, 158]. The axisymmetric models chosen have a hot center, i.e., a Q profile such that $Q \sim 1$ at intermediate and

large radii and much higher in the center. This configuration favors the WASER feedback mechanism. The extra amplification due to shear and self-gravity was discussed in [159]. As could have been predicted from [91, 130, 157] it was found that self-gravitating disks are much more responsive than disks in which only part of the mass is active, and shearing disks more than rigidly rotating ones. The combined effect can raise the growth rate significantly (factors of 5 or so). A non-asymptotic treatment of the Poisson equation was introduced in [209]. Comparison of its results with those predicted by the WASER amplification mechanism showed very good agreement. However the crucial question is not whether there exists a (or several) mode(s) in good agreement with the WASER, but rather whether the most unstable mode is correctly predicted by it. It is thus essential in all modes calculations to search for the fastest growing mode and then examine by which mechanism this mode is produced. For instance, several considerably faster-growing modes, due to swing amplification, have been found [275] in the disk of [209]. Very interesting in this respect are the preliminary results mentioned in [95]. Two $m = 2$ modes with nearly identical pattern speeds and growth rates are found. This degeneracy pointed to a mode due to the WASER. Indeed this amplification mechanism predicts the pattern speed and growth rate to within 5% and 15%. However, a faster growing $m = 3$ mode due to and explainable by swing amplification was also found!

Another shortcut used is that of soft gravity. This was applied to the Gaussian disk [267] in [80]. By raising the values of the characteristic length a of the gravity softening it is possible to stabilize the disk against axisymmetric instabilities. Non-axisymmetric instabilities are not discouraged that easily. The growth rate is of course lowered if higher values of a are considered but even for extreme values of the softening length (comparable to the scale length of the disk) non-axisymmetric stability was not achieved. This is in direct contradiction with the results of modes in gaseous disks, described above, and is probably due to the different approximations involved in the two cases. Although the gaseous treatment has been criticized as too stiff [274], it is not at all sure that soft gravity is that much better an approximation. It would thus be particularly interesting to attack this question using neither of the two approximations.

Softened gravity was again applied to a Gaussian disk in [274]. A series of modes for $m = 2$ was found. Two indications argue that these modes are swing amplified. First, as the percentage of halo mass is increased, interference patterns become obvious. As was said in section 10 this was to be expected since for large halo masses the amplification factor is small and the leading wave has a comparable amplitude to the trailing one, particularly near the center. Furthermore the successive density maxima in the modes (see fig. 12 of [274]) are roughly at 90° spacing, which shows equal wavelengths of the trailing and leading components. Secondly, the growth rates are consistent, to within 60% or 80%, with rough estimates from swing amplification.

In the subject of modes it seems to be an unwritten law that at any time the most interesting calculations have not reached publication yet, except under preliminary form. A notable counterexample is [134], which has been termed in [272] as “tour de force”. Indeed in that paper an undoubtedly simply model called the omega model has been treated completely analytically without any approximations whatsoever. The normal modes of a uniformly rotating self-gravitating disk have been separated and their characteristic frequencies calculated. It has been shown beyond doubt that non-axisymmetric modes can still be unstable after all axisymmetric instabilities have been suppressed. The most troublesome mode appears to be the bar-like elliptical deformation. Furthermore, it was shown that it is possible to stabilize this disk, either by increasing the percentage of mass in the halo, or by raising the velocity dispersion (lowering the mean rotation rate).

Three criticisms can be made about this work. The first concerns the distribution function used.

Indeed this is non-standard and has an (integrable) singularity in velocity space. Partly in order to test this distribution function assumption, the $m = 2, n = 2$ modes of these disk were calculated in [142] under the hypothesis that the mean as well as the circular velocity of the stars is of one uniform rotation, but without any further assumptions about the distribution function. The results are in very good agreement with [134]. The two other criticisms are not so easily disposed of. Indeed this disk

- a) does not have shear,
- b) has a rather sharp cutoff at its outer edge.

The importance of these two factors has been long underscored. For quite a long time these mode calculations have been considered, because of their quality, a reference point for all bar instabilities while it now turns out that these modes are due to a completely different mechanism, namely the sharp cutoff at the edge of the disk. This becomes clear, e.g., from the softened gravity mode calculations in [274] where, amongst the other modes calculated, one turned quite different. It was called D mode. It could have been confused with the other, swing amplified, modes of that disk, because for zero halo mass it has a pattern speed and growth rate falling in their midst. However, contrary to the other modes, for increasing halo mass it remains relatively hefty and does not show interference patterns. It is also not affected by a freezing of the inner parts, thus demonstrating that it does not rely on any reflection at the center. So it very ironically turns out that omega models, which have been long considered as the prime analytical example of a bar instability, do not belong to this category at all, but rely on completely different causes for their instabilities.

Another global calculation of modes has been given in [307, see also discussion in 272]. The model, which has a constant circular velocity at all radii and a surface density inversely proportional to radius, is self-similar, a property which was nicely exploited during the mode calculations in order to avoid unnecessary approximations. Plausible cutoffs were used both at the center and at large radii. Axisymmetric stability was obtained for similar values of the velocity dispersion as those predicted by local theory [268]. This axisymmetrically just stable disk was found to be also stable against $m = 2$ instabilities unless the central cutoff was too sharp! This can be nicely understood with the help of swing amplification. Indeed a constant circular velocity disk will always have an ILR which will absorb the incoming waves and prevent them from reflecting at the center. Thus the feedback cycle is broken and no mode can grow. If the central cutoff is sharp enough, however, reflection can occur at its edge, instead of at the center, provided this is outside the ILR. Thus the cycle will be closed and modes can grow. $m = 3$ and $m = 4$ modes were found to be more stable than $m = 2$ modes. Again this can be understood since the Lindblad resonances for $m > 2$ are nearer corotation than for $m = 2$ (see fig. 18) and thus have more chances of being outside the edge of the central cutoff and of stopping the wave. Finally $m = 1$ modes were found to be unstable even for quite high values of Q . Although these calculations are less reliable than those with $m > 1$, since the perturbations displace the center of mass, the instabilities found are most probably not an artifact. They can again be easily understood since ILR's for $m = 1$ do not exist for direct waves so that the feedback cycle cannot be tampered with.

To avoid problems with sharp cutoffs at the center or the edge of the disk, the isochrone disk [99] has been used in [141]. Since this model lacks self-similarity, it was only by relying heavily on numerical calculations that approximations, and in particular the epicyclic one, were avoided. Unfortunately only few preliminary results have so far reached print [138].

To summarize, this long list of modes calculations can be divided into two different groups

- a) Those due to the sharp edge of the disk, called edge modes. Examples are the omega models [134] and mode D in [274]. The obvious cure here is to smooth the edges of the system, something which real galaxies must have had ample time to do.

b) Those due to shear, wave propagation and reflection, in other words to one of the two feedback cycles described in section 10. Which of the two? In several disks considered so far which have central regions such that ILR's can be avoided, the dominant modes have been always found to be due to swing amplification. Although this and its high amplification rate make of swing amplification a most favorable candidate, one should not offhand exclude WASERs. Namely in some early type spiral galaxies it seems pretty difficult to avoid ILRs. If these disks have Q profiles rising sharply towards the center and shielding the wave from the ILR, then modes indebted to WASER action could perhaps survive.

The other important question is that of the growth rate. The initial aim in this quest for spiral modes was to find a stable or very mildly unstable mode. However, the analytical calculations show that self-gravitating disks with $Q \sim 1$ have been found to have quite healthily growing instabilities with growth rates comparable to pattern speeds. Two ways have been suggested for reducing this growth rate: either to allow only part of the density to respond (thus simulating the presence of a halo) or to consider sufficiently hot disks. Since halos are invisible there is no direct observational upper limit to their mass. The other alternative is to consider relatively hot disks. We know that Q in the solar neighborhood can not be larger than roughly two, but we have precious little knowledge of how it varies with radius. One can consider either a very hot center, thereby hoping to exclude swing amplification and allow only the tamer WASER, or a very hot corotation region where the wave would be amplified very little when swinging.

However we should be careful not to overdo it. To find a stable wave is in itself an interesting and challenging theoretical problem but it might well be that real galaxies need a reasonably hefty growth rate to overcome the damping from the shocks in the gas discussed in section 6. The modes calculations do not include, not even schematically, the energy dissipated in the shock since they describe only the stellar component. A quasi-stationary state could be reached through a balance between dissipation and the tendency for the wave to grow. Under this proviso, it could well be that unstable stellar modes driving a gaseous component are, at least in some galaxies, the answer to spiral structure.

15. N -body simulations

In the analytical approaches discussed so far it has been implicitly assumed that the number of stars in a galactic disk is large enough to allow us to treat it as a continuous system for which the Boltzmann equation applies, and that its graininess can be blissfully ignored. The opposite viewpoint is adopted in N -body computer simulations. Here a relatively small number of particles, typically 10^4 to 10^5 , stands for the 10^{11} stars in a galaxy. Each particle can be considered as a superstar of 10^6 to $10^7 M_\odot$, thus exaggerating the undoubtedly existing graininess.

The position and velocities of all particles are stored in the computer and their potential and force field is calculated. Direct summation of the forces between all particle pairs is prohibitively expensive for so large a number. Thus forces or potentials are calculated, using various techniques [102, 111, 127, 196, etc.], only at a restricted set of points forming a grid (cartesian or polar) and interpolated for all other positions. To sidestep problems of close encounters and binaries, forces are either truncated at small distances or substituted by approximations like soft gravity. Once the potential and forces are calculated, the particles are moved according to a given scheme [e.g. 102] and a new cycle starts.

Most of the simulations tried so far were two-dimensional and neglected z-motions and thickness of the disk (although soft gravity somewhat imitates the latter). Some fully three-dimensional calculations have been attempted [110, 101, 199, etc.]. The price to pay for the inclusion of the third dimension is a

worrisomely coarse grid, the thickness of the disk hardly exceeding a couple of mesh spaces. Furthermore the third dimension did not lead to appreciably different results, except for the case of unreasonably thick disks, and then only marginally [128, 47].

We shall not discuss here any of the technical aspects of the codes, except for two numerical hiccups which can, if not properly treated, severely influence the results and their astronomical implications. The first one is the necessarily incorrect description of Newtonian gravity at short distances. For example softened gravity, as was discussed in section 8, acts very much like velocity dispersions. The influence of softening on numerical simulations was emphasized in [249] where it was shown that it may, if too large [e.g. 25], lead to a spurious stability. The second problem comes from the graininess or noise due to the comparatively small number of particles. To curb this special starting conditions known as quiet starts are necessary. Comparisons of pattern speeds and growth rates from quiet start runs with theoretical predictions show an accuracy better than 10%, compared with 50% or worse of noisy starts [250].

In spite of their shortcomings, numerical simulations have contributed very significantly to our understanding of disk dynamics and in particular of their stability and spiral structure.

15.1. Stability mechanisms

A number of the early simulations [103, 104, 195, 107, etc.] verified that disks with $Q = 1$ were axisymmetrically stable. Just as in modes calculations, it soon became clear that such disks were violently unstable to bar instabilities. A typical example [from 104] is shown in fig. 35. A bar forms in the early stages of the evolution and, with relatively small changes of amplitude, shape and pattern speed, survives to the end of the calculation. Unlike analytical mode calculations, simulations are not restricted to the initial small amplitude perturbations, but can follow the evolution of the bar throughout. Thus numerical simulations can rightly claim to explain the origin and maintenance of bars and ovals in SB and SAB galaxies. More detailed comparisons of their results to observations have also been attempted [151, 199]. Not all galaxies have large bars, however, whereas this seemed the inevitable ultimate fate of all disks unless special care was taken to prevent it. Three ways of doing so have so far been suggested: halos, velocity dispersions, and the existence of an ILR.

15.1.1. Halos

A massive halo has been the most popular and most straightforward way to stability. All one has to do is to freeze part of the mass of the galaxy in a rigid non-responsive component, the halo. As is found also in analytical mode calculations, this reduces the growth rate of the instability. Several experiments have been run to determine the halo mass beyond which the instability cannot be detected. Obviously this limit depends on the code used and in particular on its softening as well as on the specific disk and halo model adopted. For reasonable values of the softening it is found that cool disks are unstable unless they are immersed in a rigid halo whose mass interior to the disk is at least comparable to the disk mass [208, 107, 101, 76, etc.]. This remains true for a non-rotating but non-rigid halo [109, 248].

The stabilizing effect of the halo can be easily understood in terms of swing amplification. Indeed by freezing part of the mass, one lowers the responsive density and raises the X parameter, so that the resulting amplification factors are diminished.

Can this be the process by which SA galaxies are stabilized or, equivalently, do such non-responsive massive components exist? Large bulges would be ideal candidates (though in most cases they can provide the necessary mass only for rather extreme values of their M/L ratio) but for the fact that they

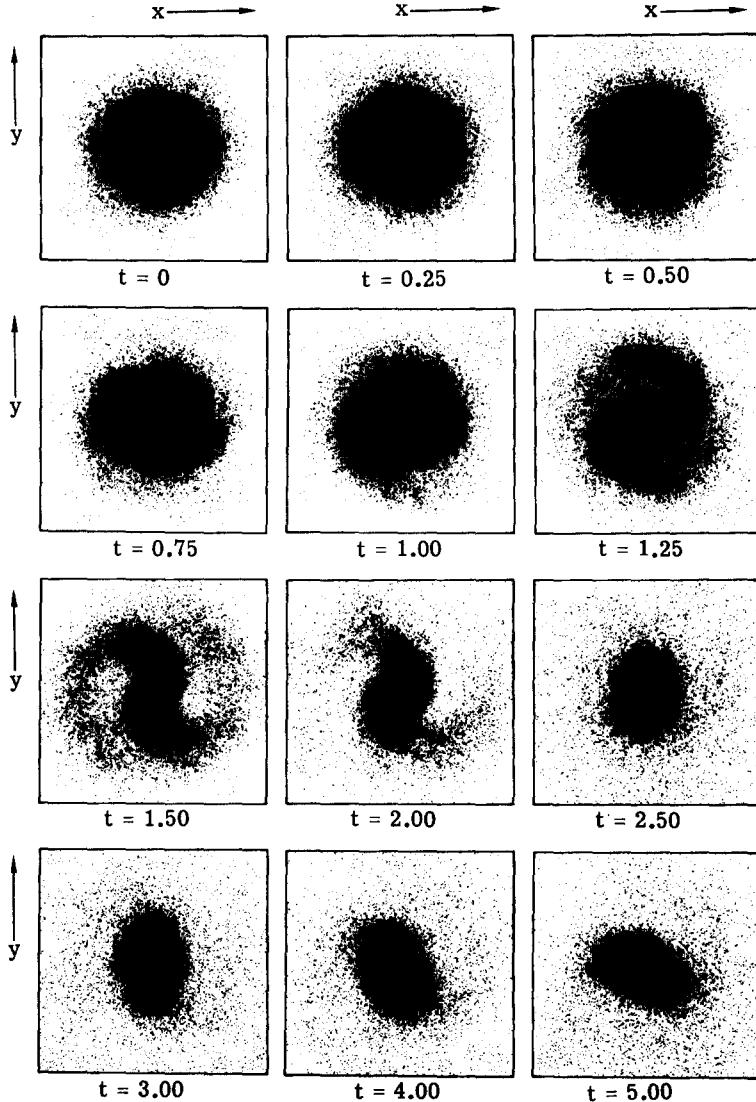


Fig. 35. Evolution of a disk of stars with an initially exponential mass distribution (from [104]).

are present only in early type galaxies. If they were the main stabilizing agent one would expect a predominance of barred galaxies in later Hubble types, which does not hold. No other visible component can contribute the necessary rigid mass. One must thus turn to a dark or invisible component, the halo. Its invisibility [87, 98, 257, 153, 58, 112, 36] makes it very hard to prove or disprove its existence, let alone measure its mass. Several hypotheses have been put forward as to what it is made of, ranging from neutrinos or very light stars ("Jupiters") to massive black holes.

The difficulties in obtaining an estimate of the halo to disk mass ratios of real galaxies are further aggravated by the fact that we are not interested in global estimates, but rather in the ratio of masses in the relatively inner parts, i.e. well within the optical disk. According to [249] it is only the innermost part of the halo, up to one or two times the radius within which the rotation curve is rapidly rising, that

will contribute to the disk stability. For early type galaxies this distance is often small, and radio techniques are not suitable for observing it, because of beamsmearing or lack of HI-emission [29]. A careful analysis of photometry and optical velocity data can at the best only give limits on M/L ratios [76, 12]. The only possible constraint is whether these are reasonable and consistent with stellar population models. One can also check the more detailed models built for several external galaxies [17, 292, 44] to see whether the mass placed in the halo is sufficient for stability. Very often the constraints on these models come from regions much further out from the galactic center and thus their extrapolation to the relatively inner parts may not be meaningful for our purposes. A similar comment could be made about our Galaxy, for which anyway some models [207] place enough mass in the halo, while others do not [18, 41].

An independent, but not easier or more reliable way, of obtaining some information on halo masses is via the spiral structure existing or not in a galaxy [12]. Indeed from the $X > X_{\text{lim}}$ criterion of swing amplification (section 9) one can see that too large halos would stifle all spiral structure, while in more moderate ones only multi-armed spirals can survive. This approach is complicated by observational inadequacies, however, and to a lesser extent by the fact that X is a local quantity, so that for galaxies where it varies noticeably with radius a global approach would be preferable.

Finally let us note that most of the theories on warps [276] rely on halos, although this provides neither proof of their existence nor a measure of their mass in the inner parts of galaxies.

15.1.2. Velocity dispersions

All numerical simulations and their determinations of the percentage of halo mass necessary for stability, discussed so far, started with an initially cool disk, barely axisymmetrically stable. From analytical mode analyses, however, it is known that introducing large velocity dispersions lowers significantly the growth rate of instabilities. Furthermore all numerical simulations of cool disks either self-gravitating or with a small halo to disk mass ratio, end up with high velocity dispersions. In one such case [105] the instability was allowed to run its course, and then the position angles of all the particles were randomized, and the evolution further continued. This disk was found to be stable thus suggesting that sufficiently large velocity dispersions could indeed put off the bar instability.

Until recently pressure support has not been as popular amongst numerical experimenters as halos. One of the reasons may be that it is not as straightforward to apply. The simple approach of arbitrarily choosing the radial velocity dispersion and then using the epicyclic approximation to determine the tangential component, breaks down whenever the velocity dispersion exceeds a small fraction of the circular velocity. One needs a distribution function which is a function of the integrals of motion and which integrates to the desired surface density. Thus the models will start off in equilibrium regardless of their degree of random motion [106, 197, 308, 14].

The stability of hot disks was particularly considered in [14]. It was shown that introducing velocity dispersions can reduce very substantially the growth rate of instabilities. In this way, models with arbitrarily small growth rates of the bar instability were constructed in the complete absence of any halo. Thus in order to ascertain whether a given mass model is stable, one needs to know both its halo mass and the random motions present in the disk. Inversely, the fact that a given galaxy is stable gives, on its own, no information on the amount of halo mass present. Several distribution functions, with Q profiles flat, increasing or decreasing with radius, were tried. High Q near the center of the disk proved much more effective at reducing the growth rates than increased random motions further out. Thus growth rates correlate better with central Q than with mean Q . The tightest correlation found was with

the fraction of mass in the model on nearly circular orbits, defined by the condition $(a - p)/(a + p) < 0.5$ where a and p are the apo and perigalactica of the unperturbed orbit of a star.

Retrograde stars are an alternative way of introducing velocity dispersions [137]. At least the edge instabilities of a Maclaurin disk can be subdued if the mean rotational velocity dropped below half the circular velocity [142]. A star on a retrograde orbit is more efficient in that respect than a non-rotating one. Retrograde stars as a cure to the bar instability were first introduced in the analytical mode calculations of the isochrone disk [141] and its numerical verification [308]. The agreement between the two approaches was very good. It was found that retrograde stars considerably lower the growth rate of the instability and also create fatter bars. The price to pay for that [308] was an $m = 1$ instability. This, however, may have been simply due to the way in which the retrograde stars were introduced. Indeed the distribution in phase space used in [308] definitely conveyed a two-stream impression, and this, as is well known, e.g. from plasma physics, may introduce instabilities. Thus the possibility of a two-stream instability severely limits the amount of retrograde stars that can be impunitively introduced in the case of a cold distribution function. On the other hand a hot distribution function needs at least a few retrograde stars to avoid sharp discontinuities in phase space, which could otherwise introduce yet other instabilities. Thus we should expect retrograde orbits to be found in the central parts of galaxies where the distributions are hottest. In fact numerical simulations [308, 14] show that if retrograde stars are not introduced from the onset, they appear of their own accord during the evolution.

Until recently the only place where relatively reliable estimates of velocity dispersions could be obtained was the solar neighborhood. Since this is far from the center of the Galaxy one would not expect to find much pressure support and in fact the radial velocity dispersion has been found to be roughly one fifth of the mean orbital velocity [301]. Dynamical considerations indicate that more pressure support should be expected near the centers of galaxies. Indeed velocity dispersions of the order of the mean orbital motion were reported by several observations of external galaxies [148, 149, 152, 33]. Although at the limit of technical feasibility, these results show clearly that galactic disks are not everywhere as cool as the solar neighbourhood. Unfortunately the z -component of the dispersion mixes in all those cases with the planar one, thus hindering precise comparisons with models. Somewhat less pressure support was reported for NGC 3115 [122], but the orientation of this galaxy is such that tangential velocity dispersions were mainly measured, whereas the radial components should be larger by roughly 50%.

15.1.3. Existence of an ILR

It was pointed out in [274] that in order to stabilize disks one needs only to break the swing amplification feedback cycle by introducing an ILR. There can be little doubt that this would in principle be effective, but whether this mechanism prevents an important fraction of SA galaxies from forming a bar is not clear. Indeed it has been shown in [249], for a class of models with gently rising rotation curves, that the bar forming mode adopts a pattern speed sufficiently high to avoid any ILRs. Very steeply rising rotation curves would therefore be required, as those observed in early type galaxies with large bulges. Whether an inner Lindblad resonance placed within such a spheroid would still cut the feedback cycle is an open question. Finally it is worth noting that high central angular velocities are not a particularity of SA galaxies, but are observed in SB galaxies as well [e.g. 129].

It has also been suggested [274] that galactic disks may well during their evolution redistribute their mass so as to increase their central concentration and therefore their $(\Omega - \kappa/2)_{\max}$ and in this way introduce an ILR. Spirals indeed help galaxies transfer their angular momentum outwards [176] thus helping the centers contract as desired, but there may not be enough mass involved to create the

necessary redistribution in the disk. A bar might prove more efficient, but even assuming it did manage to redistribute the mass as desired, one would still have to worry how, after that, to get rid of it, since so far no bar has vanished from a numerical simulation. One would have to rely on a violent event, e.g. a neighboring small galaxy falling in its center and thus perhaps destroying the bar and restoring the galaxy to its axisymmetry. The elaboration of such a scenario is an interesting problem for N -body codes.

15.2. Stability criteria

Toomre's local stability criterion [268] has proved so useful that it has prompted research for an analogous global non-axisymmetric one. Heedless of the difficulties involved, several attempts have been made.

15.2.1. The Ostriker and Peebles criterion [208]

According to this criterion a galaxy can not be stable unless the ratio of its organized kinetic energy (T_{mean}) to gravitational potential energy (W) is less than 0.14:

$$t = T_{\text{mean}}/W < 0.14. \quad (15.1)$$

Using the virial theorem

$$2T_{\text{mean}} + 2T_{\text{rand}} + W = 0$$

one can also write this criterion in the form

$$T_{\text{rand}} > 2.6T_{\text{mean}}$$

where T_{rand} is the kinetic energy in random motions. It is thus clear that only a small percentage of the total energy of a galaxy can be in rotation or ordered motions, least its stability be jeopardized. This can be achieved by introducing either a massive halo, i.e. increasing W , or sufficiently high velocity dispersions in the disk. This second possibility was not really explored in [208]. Note also that this criterion is necessary but not sufficient and that only the halo mass interior to the disk is effective, since a heavy spherical shell surrounding the galaxy would neither exert any forces on it nor contribute to its stability.

This criterion was corroborated by several 150 to 500-body calculations [208], as well as by a compilation of the up to then available numerical and analytical results for both bar and edge modes. It was severely criticized in [274], largely because it does not include the possibility of stabilizing by introducing an ILR. Several counterexamples found so far [4, 71, 307, 25, 19] are either due to this or to the fact that gaseous disks obey a different dispersion relation, which makes bar instabilities easier to suppress [274].

A somewhat revised version of this criterion has been put forward in [76]. Since it uses forces instead of potential energies it can deal correctly with cases such as a spherically symmetric shell surrounding the disk, but it does not answer the criticisms of [274].

15.2.2. The virial theorem [295, 296]

A second trial, based this time on analytical calculations, has been given in [295, 296]. It uses the linearization of the tensor virial equation of the second order and its main approximation consists in

using “trial functions” for the Lagrangian displacement (x, v) of a particle in phase space. Like [208] it quickly comes to the conclusion that too much rotation can be most destructive for the stability of galactic disks, but prefers to use the total angular momentum of the system, instead of its rotational kinetic energy, as a measure of rotation. Two quantitative tests of this criterion are available. First it reproduces exactly the point of marginal stability for the omega models [134]. Second it reproduces to within 3% an estimate, given in [297], of the point of marginal stability of the stellar analog of a uniformly rotating polytrope of index 0.5 [294]. However, if the density distribution is strongly non-homogeneous, e.g. it has a large central mass concentration, or if the system has strong differential rotation, the accuracy of the “trial functions” is debatable. Thus the application of this criterion to real galaxies may not be very enlightening.

15.2.3. Exponential disks and flat rotation curves [76]

This attempt [76] was much less ambitious, since it restricted itself to a particular model (namely an exponential disk embedded in a halo such that the total rotation curve be quite flat) and was aiming at a criterion easy to use for comparisons with observations. N -body experiments showed that such disks would be unstable unless

$$Y = V_m/(aM_D G)^{1/2} > 1.1 \quad (15.2)$$

where V_m is the maximum rotational velocity, a^{-1} is the scale length of the exponential disk and M_D its mass. This criterion was applied to observations and it was concluded that the M/L ratio of the disks of Sc galaxies could not exceed $1.5h$ (where h the value of the Hubble constant in units of $100 \text{ km s}^{-1} \text{ kpc}^{-1}$). This is indeed low but it was claimed to be consistent with population models [156, 266]. Furthermore since Y does not vary considerably between barred and non-barred galaxies (with the mass of the disk calculated in all cases from its luminosity) [81], it was conjectured that most galaxies lie on the borderline of stability.

This criterion is precisely what one would have obtained by assuming that the rotation curve was exactly flat ($V = \text{constant}$) and insisting that the X parameter (section 9) be larger than three everywhere in these models since

$$X = Y^2 e^{ar}/ar$$

and $X_{\lim} = e Y^2$. Furthermore it completely neglects the effects of velocity dispersions. Thus no galaxy with large pressure support and a small halo could be stable by this criterion although stability could well be thus achieved [14]. Finally its application (as for that matter the application of any criterion) to real galaxies may not be as straightforward as it looks. For example it relies heavily on the fact that the distribution of mass in the disks is exponential all the way to their centers. If, as claimed in [146], some exponential disks have central holes, this would severely enhance their stability by raising their X parameter.

15.2.4. Swing amplification [91, 130, 274]

Although not meant initially as a stability criterion, the requirement $X > X_{\lim}$ or diagrams like those of fig. 26 could perhaps be used as such. However this kills not only the bar instability but all spiral structure as well and it is local while Γ and Q can vary greatly with distance from the center.

We may thus conclude this section by saying that no simple stability criterion has been found so far which can substitute a global model analysis or a numerical experiment.

15.3. Non-linear properties of bars

A typical case of a bar-forming disk is shown in fig. 35. Such an evolution is always accompanied by severe heating of the disk. All bars ever formed in this way are direct, i.e. rotate in the same sense as the bulk of the stars in the disk. A detailed analysis of the orbits of the individual particles has not yet been made, but whatever is known about them [249, 199] is consistent with the general picture to be given in section 16. In particular there is a large number of elongated 2-lobes, 1-turn orbits.

The development of a bar can be considered as a two-stage process [249]. A relatively weak and short bar emerges from the initial bar instability. This intermittently forces a transient two-armed spiral, whose pattern speed equals that of the bar and places OLR near its outer end but still within the disk. The main effect of the spiral is to transport angular momentum radially outwards from the bar region to the stars near the OLR, as predicted in [176]. This results in a narrower, longer and less rapidly rotating bar, in good agreement with what has been predicted for individual orbits in [175]. The stars near the OLR heat up and after a few such episodes of transient spirals they are unable to absorb any more angular momentum so that the whole process stops, leaving a bar rotating in a hot featureless disk.

The initial length of the bar is determined by the axisymmetric rotation curve [249]. Indeed it is proportional to r_{\max} , the radius at which the rotation curve is maximum, the constant of proportionality being, for the model discussed in [249], roughly 1.4. Thus the bar instability is shown to be associated with the central parts of galaxies. If these were in one way or another prevented from forming a bar the whole disk would be safe. The final lengths of bars are found to be proportional to the modified Ostriker and Peebles parameter [76]. This agrees well with the previous remarks, since initially cooler or more massive disks will allow more spiral episodes before they heat up and therefore will end up with longer bars.

An observational verification of these results is prohibitively difficult. We can obviously measure the present and not the initial length of bars. Only by assuming that the final lengths are strongly dependent on the initial ones, which is not necessarily true [76], can some observational corroboration be sought. Moreover the orbits in the bars are highly elongated and the “rotation curves” obtained from measurements parallel and perpendicular to the bar are very different (see fig. 42 and corresponding discussion in section 19). A correlation of the bar length to the position of the maximum in the “rotation curve” along the bar should indeed show a tight correlation, were it not for the large uncertainties and inadequacies of the existing observational sample. However this would be more of a proof of the large eccentricities of the orbits within the bar than of the dependence of the initial bar length on the axisymmetric rotation curve. Since observed rotation curves seem to be inadequate for our purposes a more roundabout way is necessary. A tight correlation between bar length and bulge diameter, both measured in disk scale lengths, was found in [13] for galaxies of type SB_b and earlier. If indeed bulges dominate the inner parts of rotation curves of early type galaxies and determine their r_{\max} , then this correlation may lend some observational support to the numerical results.

The shape of the bar or oval formed in numerical simulations was discussed in [8]. It was found that the final eccentricity depends on how hot the initial distribution function was. Cool systems make thin bars and hot ones make fat ovals. The initial distribution function was used as a reference since after the bars have formed all quantities depend on the angle as well as the radius, thus making comparisons more difficult and less meaningful. Furthermore the differences between the velocity dispersions of the models are largely reduced as the runs evolve, since the initially cooler models heat up more than the initially hotter ones.

Three-dimensional calculations provide some information on the vertical structure and motions in a

bar [47]. When viewing the disk edge on one sees a box or peanut shape for the bar, depending on whether one is looking along its major or minor axes. Higher-order resonances exist between the bar motion and the z motions of the particles. In particular the fourth order resonance $\Omega_p = \Omega - \nu_z/4$, where ν_z is the frequency of the vertical oscillations, always occurs at the radius of the highest thickness of the peanut shape.

Bars slow down appreciably during the evolution [248, 249, 47, 76, etc.]. Most bars avoid ILRs during their formation [249, 248, 14]. This indeed would be expected if they were due to a feedback cycle as those discussed in section 10. The position of the ILR can be found from the rotation curve, as long as the amplitude of the bar is relatively low. Later on in the evolution it seems very unsafe to do so. A possibly accurate but much more cumbersome way would be via the orientation of the 2-lobe 1-turn orbits. Indeed it is known from periodic orbit calculations (see section 16) that these are elongated along the bar, except between the center and the ILR, if there is only one, or between the two ILRs, where they are perpendicular to the bar. According to this criterion ILRs in disks with strong bars may not exist even in cases where the crude method based on the rotation curve would say they do [288]. Thus a lot of caution is necessary. Three possible examples of bar instabilities with an ILR have been reported so far. Disks with constant rotational velocity throughout [307] have been numerically simulated in [197] and, in spite of the ILRs present for all m values except $m = 1$, were found unstable. This is obviously in disagreement with the theory. Nevertheless, the models were evolved for a particularly short time and the measured growth rates could perhaps be simply noise amplified by shear. The second case was found in a numerical simulation [308] of the Schmidt–Miyamoto models [201], but it is not reported whether this refers to the initial stages of the evolution. The same comment holds for the third case [76].

As will be discussed in more detail in section 16, it has been suggested on the basis of periodic orbit calculations that bars end at corotation. This seems to be in good agreement with the results of numerical simulations [249] since the orbital period of particles inside the bar is smaller than that of the bar, while the opposite is true for particles outside it. Furthermore since the pattern speed decreases during the evolution, corotation should move outwards and the bar should become longer, as is indeed the case. As the bar grows in length those orbits which join it change from retrograde to direct in the frame rotating with the (slowing down) pattern speed.

A study of the long term evolution of the bar is obviously very expensive, and few relevant results exist. In [47] the evolution was followed for many, typically twenty, revolutions of the bar and it was found that its amplitude, after reaching a maximum, declined with time. The physical relevance of this result was contested in [249] and it was attributed to the cutoff of that disk which would prevent angular momentum transport from the bar to the OLR, by eliminating the sink, and thus stop the secondary growth of the bar. The evolution of the runs in [76] was also followed for a similarly large number of revolutions and no sign of bar decay was found. Finally this question was specifically addressed in [43] using a very small number of particles, typically 400 to 4000, and it was proposed that the decay seen in [47] was due to relaxation effects.

15.4. Spiral structure

The initial goal of N -body simulations had been the understanding of spiral structure [198, 107 and references therein]. The formidable difficulties associated with the bar instability soon sidetracked this.

Transient spiral structures appear in the initial stages of each run, but, unless the bar instability has been suppressed, it will heat the disk to temperatures suffocating the spirals. Thus even if such a disk is triggered by a companion the driven spiral is weak, visible mainly in the cooler populations [107, 108].

Centrally concentrated halos stabilize the inner parts of a galaxy while allowing for some amplification due to shear in the outer parts. Thus in [24] a three-armed spiral amplified near corotation, then split in two parts, one propagating inwards towards ILR, the other outwards towards OLR. The transport of angular momentum by such spirals was found [252] to be in good agreement with theoretical predictions [176]. Furthermore its redistribution was closely correlated with the resonances of the pattern.

Since large velocity dispersions stifle spirals, the obvious way of continuously regenerating them is to cool the system or at least a subsystem of it. This was already realized fifteen years ago [198]. Two populations were present in these pioneering experiments, "stars" which were collisionless, and "gas" which was cooled at each step. "Stars" were created out of "gas" according to a simple recipe, in which the number created was proportional to the square of the number of "gas" particles in a given bin. It was shown that the stellar populations heated up, till they became largely pressure supported, so that it could not and did not show spiral structure. This should rather be sought in the cooler populations, in the same way as in galaxies the spirals are delineated better by gas and young stars. Indeed in the simulations the gaseous population showed very clear spiral waves and so did reasonably cool stars. The patterns were gradually lost for populations with higher dispersion.

Numerically much more reliable experiments including processes designed to mimic accretion and star formation are discussed in [251]. The disk showed short-lived spiral structures, continuously growing and decaying. Individual episodes of growth were shown to be well explained by swing amplification. This continuous spiral activity heated the disk to values of Q around two and led to an age–velocity relation consistent with the observations of the solar neighborhood [42]. Thus if disks were formed from accretion in massive pre-existing halos [300, 94] their stability may never have been threatened. In their infancy they could have been safe because of the large halos, while as the disk became more massive the random velocities build up and become the dominant stabilizing agent.

16. Orbits and velocity fields in barred galaxies

The behavior of individual stellar orbits underlies any study, be it the response of a disk to a given forcing or a selfconsistent calculation. It is thus essential in explaining the morphology and characteristics of barred galaxies, including those 'observed' in N -body models, to start by understanding the individual orbits in a simple bar model, composed of an axisymmetric disk and a rotating bar. Let us, for simplicity, work in a rotating frame of reference in which the bar is at rest and let the bar be along the y -axis. The motion of a particle in it is governed by the effective potential

$$\Phi_{\text{eff}} = \Phi(x, y) - 0.5 \Omega_p^2(x^2 + y^2) \quad (16.1)$$

where Ω_p is the pattern speed of the bar. The isopotentials of a typical case (given schematically in fig. 36) are not circular but rather elongated in shape, along the bar in the inner parts and perpendicular to it outside. The intermediate region is more complicated with two maxima (called Lagrangian points L4 and L5) perpendicular to the bar and two saddle points along it (L1 and L2). These four points lie in an annulus. In the limit of a very weak non-axisymmetric perturbation the four points lie on a circle, the corotation circle. For strong bars, whose maximum amplitude lies within the corotation radius, L4 and L5 are at smaller radii than L1 and L2.

Periodic orbits are the backbone of all orbital structure. Indeed if stable they are followed by

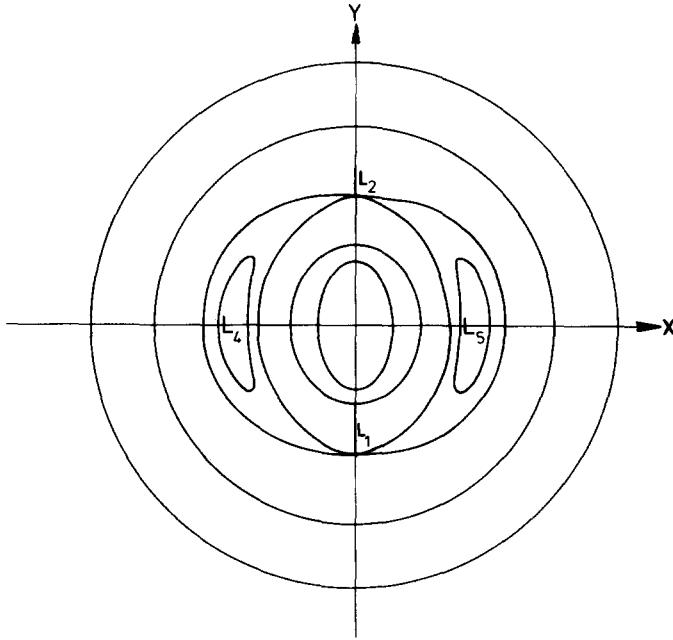


Fig. 36. Contours of the effective potential in a typical barred galaxy model.

quasi-periodic orbits, while if unstable they trigger ergodicity. Thus they crucially influence the general aspect of the density distribution. The main families of periodic orbits intersect perpendicular either the bar-axis, or the axis perpendicular to the bar, or both. They are given schematically in fig. 37 [compiled from 49, 53, 54, 51, 9]. Every orbit is represented by a point (H, x) , where x is the radius where it crosses the $y = 0$ axis, and

$$H = 0.5(\dot{x}^2 + \dot{y}^2) + \Phi(x, y) - 0.5\Omega_p^2(x^2 + y^2) \quad (16.2)$$

the Hamiltonian. The characteristics are given by solid lines if the corresponding periodic orbits are stable and by dashed lines if unstable. Obviously not every (H, x) point is possible, since for a given value of the radius the Hamiltonian cannot be less than

$$H = \Phi(x, 0) - 0.5\Omega_p^2x^2. \quad (16.3)$$

This curve, called curve of zero velocity (CZV), is given by a dotted line in fig. 37.

The main family inside corotation is x_1 , elongated along the bar and stable for most of its extent. In the limit of zero bar amplitude it is continuous and corresponds to the circular orbits, but for strong bars it becomes hopelessly broken up.

Also of dynamical importance are the so-called resonant families, each corresponding to a resonance n/m , where n is the number of radial oscillations in m turns. Starting at small radii the resonances are present in order of ascending n . The 1/1 characteristic is only present in the case of very strong bars (very massive or very eccentric) [211, 190]. The next resonance, 2/1, is well known as the ILR. As we have seen in section 4, there can be zero, one or two such resonances present. In the latter case there

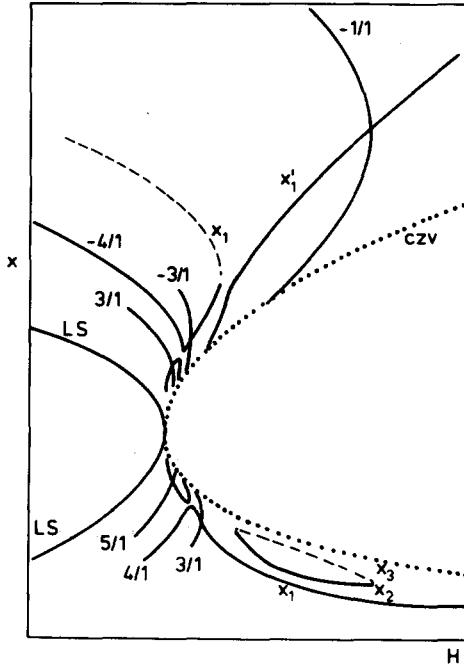


Fig. 37. Main families of periodic orbits intersecting perpendicularly the minor axis of the bar (schematically). The characteristics are given by solid lines if the corresponding periodic orbits are stable and by dashed lines if they are unstable. The curve of zero velocity (CZV) is given by a dotted line.

are two new families called for historical reasons x_2 and x_3 , whose characteristics are contained roughly between the two ILRs. They are elongated perpendicular to the bar. Thus their number should be relatively limited if the bar is to be selfconsistent. This can be achieved either if the two ILRs are relatively near each other, or if the bar is very strong [54, 288]. If there is no ILR these two families are absent, while if there is only one the perpendicular family extends all the way to the center [53, 9]. There follows a series of resonant orbits $n/1$ for n alternating odd and even. The characteristics of the odd resonances (i.e. $3/1$, $5/1$, etc.) bifurcate from x_1 while the even resonances (i.e. $4/1$, $6/1$ etc.) break x_1 in two and join smoothly onto the two branches. These $n/1$ branches are more and more densely packed as we approach corotation. For higher bar amplitudes the gaps become larger and the various resonances interact and introduce ergodicity. There are also an infinite number of families of type n/m with $m > 1$, but these are dynamically less important.

Closer still to corotation we have long period orbits librating around the Lagrangian points L4 or L5 having the shape of bananas. They decrease in size as the Hamiltonian increases, until they shrink to the corresponding Lagrangian point. One example is given in fig. 38. The transition between the type $n/1$ orbits circulating around the center and the librating ones occurs at the unstable Lagrangian points L1 and L2, where the period of both types of orbits tends to infinity [51]. For still higher values of the Hamiltonian we encounter the short period Lagrangian orbits, encircling L4 or L5, whose size increases with the Hamiltonian.

There is also inside corotation one family of purely retrograde orbits. Their shape is not far from circular, though somewhat elongated perpendicular to the bar.

Outside corotation we find a similar set of characteristics; Lagrangian short and long period orbits

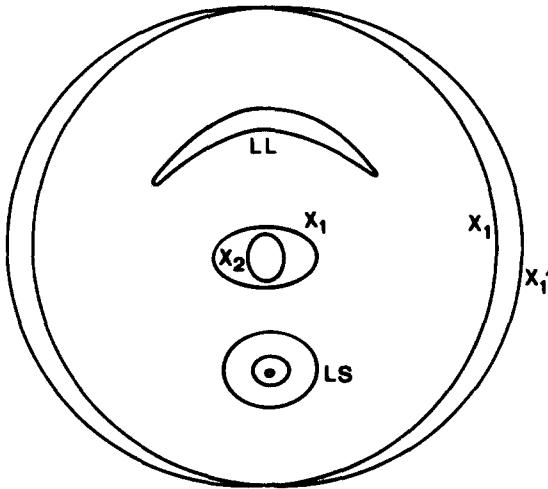


Fig. 38. Examples of periodic orbits belonging to x_1 , x'_1 , x_2 and the Lagrangian long and short periods (LL and LS correspondingly).

and resonant orbits denoted now by $-n/m$, the minus sign showing that they are retrograde in the rotating frame of reference. Mainly for historical reasons the nomenclature of the orbits of the OLR differs from the usual $n/1$. The characteristic of the x_1 is unstable for the case of strong bars.

The (H, x) diagram presented above contains information only about orbits which are simple periodic (i.e. close after one revolution) in a frame of reference corotating with the bar. A more general view of the orbital properties of a given system is obtained from the so-called surfaces of section originally introduced in [215, see also 100]. By considering the intersections of an orbit with the plane $y = 0$ and in particular those with $\dot{y} > 0$, we obtain a succession of points (x, \dot{x}) , called consequents. An m -periodic orbit (closing after m revolutions) is represented by m points. The quasi-periodic orbits trapped around a periodic orbit are represented by a sequence of points lying on a curve (called invariant curve). In other cases the consequents behave in a chaotic way. It is clearly impossible to draw a regular curve through them and they seem to be distributed at random in the area left free by the invariant curves. This situation corresponds to ergodic motion. Examples of both are given in fig. 39.

The amount of ergodicity has been found to depend on the strength of the bar, in particular on its eccentricity and mass [9, 263]. This may have important consequences for building selfconsistent bars. For example more massive bars will create more ergodicity and reduce the percentage of each surface of section covered by quasi-periodic orbits. This points to a selfregulating process which will determine the bar mass. Thus the latter can not grow to too large a percentage of the disk mass since in doing so it would kill its most ardent supporters and diminish itself. This agrees with observational evidence since the fraction of light in the bar has never been found to exceed roughly one third of the total light. Similar arguments can be made for the axial ratio of the bar.

Two reasons for ergodicity have been discussed at length in [52].:

a) As the amplitude of the bar grows the gaps produced by the even resonant orbits and the unstable regions produced by the odd ones become larger and the corresponding resonances interact.

b) The family x_1 may lose its stability at the bifurcation of some resonant family and does not recover it for larger values of the Hamiltonian. This bifurcation is followed by a cascade of infinite bifurcations leading to ergodicity (Feigenbaum effect).

The behavior of the orbits has also been used to set a limit to the length of the bar. The orientation

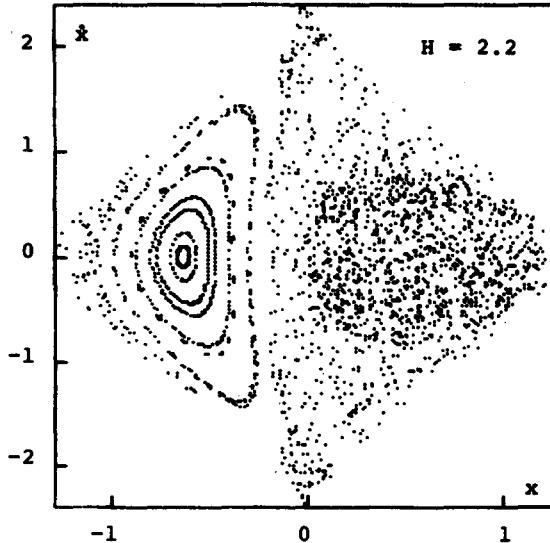


Fig. 39. Example of a surface of section (from [9]) having both well-behaved closed invariants curves (to the left, in the part corresponding to retrograde orbits and $x < 0$) and a large ergodic region (to the right, the part corresponding to direct orbits and $x > 0$).

and crowding of the orbits were found in [50] to be such as to enhance the bar between corotation and the outer ILR, and inside the inner ILR, if two such resonances are present. If there is only one ILR they enhance the bar between it and corotation. If there is no ILR, the bar is enhanced all the way from the center to corotation. No support is ever found outside corotation. These results hold when the bar amplitude decreases with radius or at least does not increase too rapidly. A different line of argument was followed in [175] where the trapping of the major axes of the orbital ellipses by the potential minimum was found possible only as long as $\Omega - \kappa/2$ increases with angular momentum, with the fast action variable J_f (i.e. the adiabatic invariant due to the rapid circulation around the orbit) being kept fixed. This will be the case in the inner parts of galaxies where the rotation curves rise substantially. No comparison of the results of the two theories has so far been attempted. Although they at first sight look terribly discordant, they could well be less so if the inner, relatively sharply rising portion of the rotation curve is contained roughly within the corotation radius.

There is a definite connection between the trapping of matter found around stable periodic orbits and morphological features of galaxies. For example, there might be a relation between the HI asymmetries often present in the outer parts of galaxies and the stability of the $-1/1$ asymmetric orbits. The connection of orbits to rings has been widely discussed in the literature. The instability of the main periodic orbits between corotation and OLR was linked in [240, 54] with the existence of outer rings, since it will lead to a depletion of the region between the two resonances while the stable orbits around the OLR will trap quasi-periodic orbits around them. This only happens if the perturbing force is strong between corotation and OLR. Thus depending on how mass is distributed in the axisymmetric background and the bar, the main families of periodic orbits between corotation and OLR may be stable or unstable, and large amounts of ergodicity may or may not be present. It is therefore crucial to find out what the relevant mass distributions in real galaxies are, and whether they are different in galaxies with or without outer rings [10]. Fewer attempts to link inner rings with families of periodic orbits have been made. In [68] the x_2 family was tried. These are stable and not far from round: somewhat elongated perpendicular to the bar. They can be responsible for inner rings only if there are

two ILRs and the bar terminates at the inner one, since otherwise the ring radius would be smaller than the major axis of the bar, i.e. the ring would intersect the bar whose ends would protrude outside it. For the case of longer bars extending all the way to corotation one might have to solicit the stable near circular periodic orbits $\pm n/1$ and/or the Lagrangian short period orbits. None of these possibilities have been properly investigated and it should be stressed that only the knowledge of the distribution function, i.e. the number of particles following a given orbit can definitely settle the question.

Finally the periodic orbits are directly linked to the gas flow in a barred galaxy, since under the zero pressure approximation a gas streamline will coincide with a closed periodic orbit. The converse is obviously not true since two given periodic orbits may well intersect whereas streamlines can not. To come back to our bar potential, a set of non-intersecting nested elongated orbits of the family x_1 can well approximate the streamlines in weak bars with no ILRs. For stronger bars it may not be possible to do this since the individual orbits will acquire cusps or intersect. The gas will not be able to follow them and will form shocks along the leading edges of the bar. Indeed at this location we observe the long, very often straight dust lanes. This logical link has already been made some twenty years ago [217]. On the other hand if there are two ILRs, set wide apart the response of the gas will be perpendicular to the imposed stellar bar [234]! The most realistic shocks have been found in cases where the bar is massive and centrally condensed, so that there are ILRs but near the center [234, 288].

By including the effects of dissipation, it can be shown [236] that periodic orbits will act as attractors [e.g. 279] in the phase space of the Hamiltonian. In other words viscous dissipation drives particles towards the simple periodic orbits.

Since in strong bars the simple and convenient periodic orbits picture does not work for the gas, we must return to a full hydrodynamical treatment, including pressure forces and viscosity, and the problem becomes much more complicated. In fact in a general case it can only be solved numerically. Several attempts, using various methods and models, have been made. We have already discussed the spiral density response in section 12, and will address here the problem of the gas flow in the bar region itself.

The beam scheme, devised in [216], is probably the most widely used up to now. It replaces the continuous Maxwellian distribution of gas particle velocities by few discrete beams of particles (usually three in one dimension and five in two), i.e. the distribution function is written as a sum of few, easy to handle, delta functions. A more complete description of this method can be found in [235, 119]. It is a first-order scheme, reasonably fast and easy to program, for which assets one pays with a relatively high viscosity. It has often been used for theoretical studies [234, 237, 238] as well as for modelling of observed barred galaxies [75].

This scheme was followed by many others, always aiming for lower viscosity and higher spatial resolution. In [220] low viscosity was traded for quasi-stationarity and an essentially one-dimensional scheme. The use of Godunov's method in [286] allows a return to two dimensions and time evolution while keeping the viscosity relatively low. The comparison of their results to those of the beam scheme is rather encouraging, since the gas streamlines computed by the two methods are not very different. A comparison of five codes in a one-dimensional problem has been given in [287], while the code which proved in this work to best approximate a perfect gas (FS2-[298]) was used in [285].

A word of caution is needed here: the least viscous code, although undoubtedly an interesting challenge, is not necessarily the best description of the lumpy interstellar medium. Indeed a second school has, as we have seen, modelled the interstellar medium as clouds colliding inelastically [241, 192, 46]. Shocks are recognized in this scheme as regions of large number of collisions.

The crucial question to ask is how the properties of the various codes affect the results, or in other words, how much the various features of the flow depend on the codes.

Intuitively one can understand viscosity as something that smooths or even blurs out sharp features. This will particularly apply to shocks but also, e.g., to how sharply the arms turn from the bar. Unfortunately the same kind of effect can also be achieved by using in the calculations less strong bars. So a relatively strong bar with a very viscous fluid could show similar flow behavior as a less strong one in a much less viscous fluid. The following analog, although very simplistic and extreme, is relevant: Suppose one foggy night you are shining your torch on some object. To achieve equal illumination in heavy fog you should use a very bright torch. If the only information you have is the amount of light incident on the object, you can not tell both how strong the torch is and how foggy the night.

Viscosity also affects the position of the shocks. In more viscous flows they are offset to the leading edge of the bar while for less viscous ones they approach the potential minimum [286]. Unfortunately again this is the effect you get with stronger bars. Another effect of a viscous code is to enhance the rate of gas infall to the nucleus of the galaxy. Again the same effect can be achieved by changing the strength of the bar (now particularly in the central part by keeping the mass of their inner disk constant while changing its scale length [237]).

All the above differences are quantitative rather than qualitative. One qualitative difference concerns the direction of flow after the shock. In viscous flows only inwards postshock motions can be found whereas in less viscous ones a small region of outwards postshock velocity can be found [220]. Any observational verification is extremely difficult, however, because the region is probably quite narrow and because one has to look along the bar.

Although one should certainly be aware of the various effects of viscosity one should not on the other hand err too much on the cautious side and disbelieve all results. Comparisons of, e.g., the beam and Godunov's scheme [286] show that there is room for optimism. Certainly the overall features of the flow are to be generally believed. Applications to particular galaxies are also meaningful and instructive, although one should realize that the choice of the code can influence the values of the parameters found for the best fit. It would indeed be very interesting to fit the same set of data with different codes in order to determine how consistent predictions of parameters like bar mass or pattern speed are.

17. Rings

Rings have been seen or hinted at in a large number of numerical simulations. For example in *N*-body stellar work outer rings have often been found, which, like their spiral analogues, are always transient. An understanding of the formation of gaseous rings has been achieved in [240, 241], where the response of a gaseous disk to a rotating bar potential was followed for several bar revolutions. Gas was modelled by "clouds" colliding inelastically, thereby losing an important fraction of their relative motions. In the first stages of the evolution a spiral was formed, extending from corotation to somewhat outside the OLR. Due to the torque exerted on the particles by the bar, the particles are pushed outwards, the region between the two resonances is slowly depleted and the spiral evolves, through a pseudoring, to a ring at the OLR. We have seen in the last section that at the OLR there are two main families of periodic orbits, x_1 and x'_1 , perpendicular and parallel correspondingly to the bar. If the bar is relatively weak and the gas is initially confined within the OLR radius, the gaseous particles will follow the family x_1 and the ring will be perpendicular to the bar. In this case one can say that this ring is inside OLR, in the sense that the angular momentum of the particles in it is less than that of the circular orbit at the OLR. On the other hand if the gas extends initially beyond the OLR, or the bar is stronger, the ring is formed around the x'_1 orbit and will be parallel to the bar. The angular momentum of the ring particles is in this case larger than that of circular orbits at the OLR. It was also argued in [240] that

perpendicular rings are not as robust as the parallel ones, since if there is even a little gas outside the OLR they can be disrupted by the collisions of the particles inside the OLR with those outside. However this does not seem to agree with observational results [147, 242] which show that the outer rings lie preferentially perpendicular to the bar.

If one now considers the case of a strong bar then there are several quantitative changes. The arm-to-arm density contrast as well as the pitch angle and width of the induced spiral are increased. During the evolution the pseudoring is perpendicular to the bar, i.e. the particles follow orbits of the family x_1 . In the final stages of the evolution, however, they are pushed past the OLR on orbits of the x'_1 which are near circular in shape. It may thus be expected that outer rings in very strongly barred galaxies be rounder than those in less barred or oval ones. Finally the rate of formation of the ring is influenced in the sense that rings appear faster when the bar is stronger.

Very similar results are found if the bar forcing is replaced by a spiral one. In particular the time scales involved are of the same order. Thus according to [240, 241] outer rings will be the end product of the evolution of all spirals, barred or not. However, the small percentage of galaxies having outer rings (see section 2) shows clearly that this is not the case. Three ways out were proposed in [240, 241]:

- a. The mass distribution in the models differs from that in real galaxies.
- b. The cloud collisions have not been correctly modelled.
- c. The gas between corotation and OLR is replenished by mass loss from stars or by infall.

Of these the third was favored in [240] where it was shown, by a numerical simulation including a replenishment roughly equivalent to $2 M_\odot \text{yr}^{-1}$, that the spiral structure may indeed persist for at least 50 bar rotations. However, the importance of the bar strength should not be underestimated since the rate of ring formation depends heavily on it and so a faster replenishment rate could be necessary. Thus one can conclude that a strong bar and/or a small replenishment rate will favor ring formation.

Different forms of potentials and a different collision law, conserving angular momentum, were used in [192]. The similarity of the results found enhances confidence. Furthermore a film [192] showing the evolution of some of the runs is very instructive. An evolutionary scenario leading from gaseous to stellar rings was briefly discussed following review [8], but should be backed by solid calculations.

The problem of inner rings has been discussed in much less breadth and depth [72, 73, 240, 46]. Inner rings have been found at the ends of the bar and elongated along it, in good agreement with observations. It is interesting to note that in these three approaches three different collision laws were used. For large values of the pattern speed no inner rings have ever been found. As the pattern speed is lowered, however, we first notice a doubling of the arms and then, for yet lower values of the pattern speed, an inner ring as a final product of the evolution. The ring does not lie at corotation but rather at the so-called ultraharmonic resonance, where $\Omega_p = \Omega - \kappa/4$. In a few cases and in particular for a very centrally condensed bar or a spiral forcing, a ring was formed at the ILR.

18. Stochastic self-propagating star formation

The backbone of all theories discussed so far is dynamics. Newton's law of gravity and the equations of hydrodynamics for the gas are the corner stones of all these approaches. The theory of stochastic self-propagating star formation (SSPSF) deviates from this. In this theory, star formation is considered as a percolation process, spreading through a galaxy like a fire through the forest, i.e. jumping from tree to tree. The proponents of this theory stress the importance of this process in explaining the morphological appearance of spiral and irregular galaxies. The idea of SSPSF was first introduced in

[202], and further advocated and developed in [89, 246, 90], where its implications for galaxy morphology and spiral structure were studied. An elaborate review of this subject has been given in [247], and we will briefly outline their theory before discussing its merits.

In its simplest form, the theory assumes that a given site of active star formation will stimulate its immediate neighborhood to form stars some time, t_p , later, with a stimulated probability P_{st} . The then burned-out site is refilled slowly with fresh gas until after some time τ it is ready to be lit up again. The basic galactic model has a rotation law which is taken solid-body out to one-tenth the radius of the model and flat at a level V_m thereafter. It is easy to put this scheme on a computer and produce an output picture with the regions of current star formation activity. Suitably chosen hard copies of the pictures convey a similar impression as photographs of some galaxies. The results depend on the total number of cells N considered (each cell is a site of star formation), the stimulated probability P_{st} , the time step t_p , the ‘refractory time’ τ , and the rotational velocity V_m . For each of these parameters an a posteriori justification can be given when the models are compared to the observations.

The stimulated probability P_{st} is thought to be locally determined by the process of star formation. The time step t_p is taken to be the lifetime of typical OB-associations increased with the propagation time of possible shocks (which supposedly induce the star formation) to the next cell. The refractory time τ has to be chosen with some care: if τ is large a large number of neighboring cells may still recover from recent star formation activity, no new cells will be lit up and the fire burns itself out. Thus τ and P_{st} have to be adjusted such that star formation is kept alive. In fact, if P_{st} is too small, even a negligible refractory time is not enough to keep the fireworks going. This condition, i.e. that P_{st} is larger than a given value P_c , is referred to as a phase transition. P_c depends on the dimension of the system: a 3D system has a lower P_c than a 2D-system. As argued in [247], this may qualitatively explain the difference between spheroids and disks.

A cluster formation rate (CFR) can be defined as the fraction of cells which are just lit up in a given time step. For a given refractory time τ the CFR increases as the rotational velocity increases. This would imply that galaxies with a high V_m have processed more gas into stars than galaxies with a low V_m [247]. If the display procedure is altered to introduce a luminosity – age relation for open clusters, and if some color data on evolving star clusters are introduced, models can be constructed in several color bands. The spiral structure in some of the models is preserved even in the red and infrared [247]. Spiral patterns for galaxies with high V_m are more tightly wound than the ones in galaxies with low V_m , due to the effect of shear. Global patterns for galaxies with the same V_m can be generated for larger values of the refractory time.

More complicated models including one or more gaseous components can be introduced. The stars and gas can now be followed separately, with gas turning into stars according to the process of stochastic self-propagating star formation. When a cluster is now created in a cell, the gas in that cell is distributed among neighboring cells to simulate the effects of OB stars on their surroundings. The inactive gas is allowed to return to the active state with the time constant τ . This last point comprises an efficient feedback for the cluster formation rate: if too many stars are created at once a lot of the gas becomes inactive and star formation would decrease.

The size effect inherent to the percolation process has consequences when small galaxies are modelled. Models with solid-body rotation but small rotational velocities begin to show large fluctuations in the cluster formation rate as their radius is decreased to about 8–12 cell sizes. The smaller the model, the larger the amplitude of what are now bursts of star formation, and the greater the time in between bursts. This could well explain the variety of properties of dwarfs irregulars, whose colors are sometimes far too blue to be explained with models assuming continuous star formation [244].

Are we, for the spiral structure problem, allowed to ignore gravity forces? All the previous sections in this review indicate very strongly that this is not possible. Furthermore, the ‘arms’ found in the SSPSF pictures are all composed of population I objects; hence the arms in the red and infrared pictures should be dominated by red supergiants. Thus bars, which are by definition straight, i.e. not shearing, and are found in the old stellar populations, can never be obtained by the SSPSF-process. Neither can the grand design smooth armed spirals as, e.g., shown in [229], for which it is very hard to believe that they are made up of young red stars. Also for normal grand design spirals, with both stars and gas, the surface photometry data indicates broad spiral arms in the old stellar populations. Thus SSPSF cannot be relevant for those. This leaves us with the very ragged-looking spirals and the truly flocculent ones. Here it can be argued that SSPSF may be of some relevance. On the other hand, disks do respond gravitationally to large clumps of matter, and can produce the more ragged structures in this way.

The idea of SSPSF may well be of interest for the theory of dwarf galaxies in which a large fraction of the ordinary mass is in gaseous form. Analytical formulations have been attempted for a 2-fluid model of stars and gas, without self-gravity, taking star formation, mass loss, and heating and cooling of the gas into account. A linear stability analysis of a general formulation of such a model showed indeed the possibility of self-propagating star formation [see 247]. A completely different formulation, tailored for dwarf galaxies, brought out clearly the oscillatory behavior with bursts of star formation [173]. Yet dwarf galaxies are besides the main issue considered here: spiral structure. And on this issue SSPSF provides a very incomplete formulation.

19. Observational constraints

A large wealth of observational results with a bearing on spiral structure theory is now available, but there are several pitfalls on the road to a meaningful confrontation of theory with observation. Idealized spiral patterns are not observed in large abundance, and the most decisive aspects of the theory do not lend themselves to direct observational tests. As will be clear at the end of this section, a lot more work is needed in this respect.

19.1. Rotation curves and mass distributions

Modern data [224, 30, and references therein] have given us a good idea of the shape of HII- and HI-rotation curves as function of Hubble type and luminosity of spiral galaxies. Although a good deal of systematics exists, the presence of intrinsic scatter prohibits making the unique correspondence between the observed kinematics and the morphology of a given galaxy. The continuing debate regarding the existence and predominance of a massive halo component [e.g. 139, 278], makes matters even more complicated.

In all dynamical models a rotation curve, or better an axisymmetric mass distribution based on it, is a necessary prerequisite. Specifically in the feedback cycles the form of the rotation curve and the mass distribution of the responding disk is very important (fig. 26). If the ratio of halo-to-disk mass is large, little amplification takes place and if the rotation curve and pattern speed are such that an ILR is present the swing amplification cycle is cut. The influence of the ILR has been discussed from the observational side in [155]. Indeed evidence has been presented that, for galaxies with no companions or bars, an ILR and a global spiral pattern are mutually exclusive. It has also been stressed, however, that grand design spirals, again for galaxies with no companions or bars, exist only in the inner parts of

galaxies, where the rotation curve indicates nearly solid-body rotation [155]. Yet several counterexamples are now known, e.g. UGC 2885 [225] and NGC 488. Moreover it must be remembered that in the regime of little differential rotation the swing amplifier is not very effective compared to the regime where the rotation curve is flat.

The study of stellar radial velocities and velocity dispersions has only recently yielded the first results for disk galaxies [148, 149, 152, 33]. In general, bulges have large velocity dispersions, and their observed flattening is consistent with the amount of stellar rotation [154]. Tri-axial bulges, often found in barred galaxies, rotate faster than the ordinary ones [150]. Only for galaxies of high surface brightness is information available beyond the bulge radius. The velocity dispersions are high in the central parts of bars and lenses, the latter in barred as well as non-barred galaxies, and decrease with radius to the instrumental measuring limit of ~ 60 km/s at the edge. The central parts of disks are appreciably hotter than the outer parts. The actual calculations of parameters like Q as function of radius, however, are not straightforward because of the unknown three-dimensional distribution of matter [152]. The uncertainties associated with the determination of the mass-to-light ratio makes the use of data on the luminosity distribution for this purpose rather hazardous. In late type spirals the contamination of the light of old disk stars with population I stars makes the determination of dispersions very difficult.

19.2. Location of resonances

As shown in section 17, several numerical simulations show that rings can form near resonances. A quantitative study of sizes of rings [11] indicates that indeed outer rings in barred spirals are most likely at the OLR, and inner rings near corotation or ultraharmonic resonance. No simple picture suggests itself for the location of rings in unbarred galaxies, although continuity arguments suggest that they should also be associated with resonances (in some cases now also with the ILR). If these associations of rings with resonances are correct, rotation data can be used to determine the pattern speed of the spiral or bar perturbation.

Other than the rings, no direct indicators exist to pinpoint the resonances and thus estimate pattern speeds. It was initially hoped [162, 255] that the outermost HII-region most conveniently pointed out the position of corotation. Later [221] the end of the prominent spiral and the edge of the “easily visible disk” were added. As a result corotation radii shrunk very considerably, without noticeably changing the quality of the fits. This shows clearly the dangers inherent in this type of criterion. One of the predictions of the WKBJ theory is that radial velocity perturbations inside and outside corotation have opposite sign (cf. section 5), and are thus an indirect indication of the position of the corotation radius, but no clearcut case has been found so far.

Theoretical arguments have been presented to show that bars cannot extend beyond corotation [50]. The good agreement between the diameters of bars and of inner rings [147] suggests very strongly that corotation is not far off the end of the bar.

19.3. Distribution of spiral tracers

The early results of calculations of the response to a two-armed spiral density-wave perturbation predicted the existence of large scale shocks [88, 219]. It was thus thought that due to the high compression star formation would result and, because of the advance of the wavefront with time, a definite sequence of spiral tracers would be observable. The regions of maximum compression were

identified with dust lanes, the compact HII-regions marked the onset of star formation, and the OB-associations, first without and later with red supergiants, should outline the luminous arms behind the shock. Thus a definite age sequence was suggested and looked for.

Observational support for this picture came from a number of studies. Maps of the radio continuum emission show a clear enhancement coinciding with the dust lanes in M51 [191], M81 [245], and in the barred spiral NGC 1097 [204], providing evidence for large scale compression. Maps of the neutral hydrogen emission in several galaxies show good coincidence of HI emission and the arms in a number of cases, foremost M81 [223], M101 [2], M31 [283, 38], and several others, including barred spirals and galaxies with oval distortions [32, 284, 290]. A comparison of the distribution of dust and HII-regions [177, 178, 179] suggests that in some grand design spirals, e.g. M51 and M83, the dust lanes are on the inside of the arms with respect to the HII-regions, but in general the situation is quite complex.

The distribution of the OB-associations, and the occurrence of a stellar ageing effect across the arms, are more difficult to establish. Since M31 is too edge-on to delineate the arm(s) on a blue photograph, it has not been used so far in this respect. The recent UV [59] and infrared maps [96] offer some new hope that the spiral structure of this galaxy might be disentangled. The only other galaxy close enough to study the stellar content in relation to the spiral structure is M33. Studies of the southern arm in this galaxy [55, 69, 70] indicate that at ~ 1.5 kpc radius the most compact HII-regions delineate the inner edge of the arm, while OB-associations of increasing diameter lie farther ahead of the arm in the direction of rotation. This is especially true of the old OB-associations containing red supergiants. However, at 3.4 kpc from the center the spatial segregation of the blue and red supergiants seems to be in the opposite sense [69], from which it has been inferred that corotation is at ~ 2.8 kpc radius. An extensive study of the stellar content of M33 [115] shows the difficulty in establishing these conclusions firmly. Indeed, the sequence dust lane, compact HII-regions, OB-associations is found in the major southern arm as is expected for a density wave, but the distribution of red supergiants is severely contaminated by foreground dwarfs. Moreover the major northern arm does not show the expected sequence. Thus this galaxy did not prove to be an ideal candidate for observational tests of spiral structure theories.

19.4. Surface photometry

The presumed age-effect across a density-wave spiral arm can also be studied with the help of color data. In an early attempt [22] it is claimed that the outside edges of the arms in M101 and NGC 628 are redder than the inside edges, in agreement with the naive density-wave shock picture. But a thorough study of six galaxies with pronounced spiral structure [243] does not confirm this picture. Instead, most of the arms are found to be broad and symmetrical, and no color gradient across the arms exists, except in a few cases with asymmetrical arm profiles. In those latter cases the asymmetry is in the sense expected from the simple density-wave theory. The broad symmetrical arms, which do not increase in width as function of radius, have a composite stellar content, indicating the presence of young stars superposed on a spiral enhancement in the disk itself. These stellar spirals increase in amplitude roughly linearly with radius to approximately 20% to 30% of the local mean axisymmetric background in the outer parts.

Similar results were found for M83 [262], where no asymmetry as that expected from an ageing effect was found. Star formation terminates equally abruptly at the inside and outside edges of the arms and there does not seem to be any transition zone on either side. Thus color maps show a two-zone structure, either zones of strong star formation, located mainly in the arms, or smooth red quiet zones, with little in the way of transition zones.

Indeed two effects can alter the naively expected picture and produce more symmetric profiles. On the one hand the newly born stars will not move in circular orbits after their formation. In [302] two cases were considered: a “pre-shock” and a “post-shock” case, with the initial velocities of the stars reflecting the gas velocity before and after the shock correspondingly. In both cases stars have a substantial radial velocity and initially move essentially along the spiral shock for a rather long time. The drift and broadening are very complicated, neither linear nor monotonic with time. The second effect, also helping to smear out asymmetries across the arm, is the fact that stars do not form instantaneously when the cloud crosses the shock but rather continuously for some $(10-20) \times 10^6$ years [304]. In that case the distinction between red and blue supergiants will also be blurred since the red supergiant phase occurs $\sim 5 \times 10^6$ years after formation. The theoretical profiles calculated with these two effects taken into account [20, 306] show no disagreement with the observational results of [243] and [262].

Further quantitative measures of the amplitudes of observed spirals has been discussed in [79] on the basis of blue and near infrared plates of several dozen nearby galaxies [77]. Arm/interarm contrast ratios were defined at both wavelengths, and the ratio of blue contrast to infrared contrast plotted against infrared contrast. An interesting result is found. For some galaxies there is a more or less constant blue-to-infrared contrast ratio with a widespread variation in infrared contrast, suggesting the dominance of a spiral enhancement in the stellar disk. For other galaxies no such behavior is found, the infrared contrast is roughly constant and some spread in the blue-to-infrared contrast ratio is apparent. There are also in between cases in which there is both spread in the blue-to-infrared contrast ratio and in the infrared contrast. A dependence of this behavior on spiral arm morphology was found. Extreme flocculent galaxies show very little structure in their old stellar population disks, while grand design spirals have smooth variations in the underlying old stellar disks, with amplitudes as high as $\sim 60\%$ of the local mean background. Yet the distinction grand design vs. flocculent spirals is not that unambiguous in the color contrast data. Several grand design galaxies, like NGC 628 and NGC 3344, have a similar behavior as flocculent galaxies, like NGC 7793, in the diagram of blue-to-infrared contrast ratio vs. infrared contrast.

Evidence for spiral structure in the relatively old underlying disk has been also found in the so-called smooth arm galaxies [258, 259, 303, 229], which have clear spiral structure but no sign of population I tracers.

A most interesting observation [97] of NGC 1566 shows clearly how different the responses of various stellar populations can be. This galaxy has two bright smooth and red inner arms but shows in the H infrared band a prominent bar. This is invisible in the ultraviolet and barely distinguishable in the blue, yet has an azimuthal brightness variation of roughly 25% in H!

Surface photometry can also be used to estimate parameters for barred spirals. Scarce data for 13 galaxies [84, 68, 56, 23, 203, 75, 74] suggests that the ratio of bar light to total light varies from 3% to 30% with a mean of 15%. There may well be a correlation of bar strength and axial ratio with other morphological features in barred galaxies, but there is as yet little quantitative information available on this.

19.5. Velocity fields

Velocity measurements are a more direct way of finding out how the mass is distributed in a galaxy than studying the light distribution. Every perturbation involving some mass will have detectable velocity perturbations if suitable tracers are present. Fortunately the neutral and ionized gas are pervasive enough to provide us with a wealth of information in this respect. However, in particular in

studies of HI and CO, the spatial and frequency resolution is often too poor to study many galaxies in great detail.

Any spiral will introduce, if sufficiently massive, regular deviations from circular motion which are observationally detectable. Indeed a perusal of the available data shows that “waves” on rotation curves and “wiggles” on velocity fields have often been observed. Illustrative examples of the former are NGC 2998 and UGC 2885 (fig. 4 in [225]), while conspicuous “wiggles” are present in the velocity fields of M81 [223, 299] and M31 [38]. For numerous other galaxies similar features have been found, but less convincing in most cases (see e.g. [29]).

For a long time hopes were high that the radial velocity measurements would reveal us the pattern speed of spirals with a grand design, since they change sign across the corotation radius. The most suitable candidate galaxy for this is M81, and observations at 21 cm have adequate resolution and sensitivity for this purpose [223, 299]. The attempt to obtain properties of the spiral wave [299] proved to be more difficult than expected. Initially a fit was made using WKBJ approximations for stars and gas, stellar wave amplitudes from photometry [243] and a pattern speed of $18 \text{ km s}^{-1} \text{ kpc}^{-1}$, placing corotation at 11.3 kpc and the inner Lindblad resonance at 2.5 kpc (distance 3.25 Mpc). This gave a good fit to the velocity perturbations for the eastern arm. Later the cloud/particle model was applied using the gas streamlines of the earlier model as trajectories for the clouds and simple rules for the star formation [21]. A reasonable resemblance to the real galaxy is found, but the regions of dark clouds and luminous HII-regions can only be separated in velocity space, which has yet to be done observationally. No model in which the grand design is due to the driving of the nearby companions has so far been attempted, although this seems a very promising hypothesis. Other candidates for this kind of work are for the moment less suitable. M31 is so edge-on and has a so much more complicated spiral structure that a major effort is required to disentangle it, even though beautiful datasets exist [283, 38]. In M33 the perturbations are much smaller and confined to the inner arms. M101 is probably too face-on, and no regular velocity perturbations have been found [34]. However, with the advent of scanning Fabry–Perot interferometers [15, 37] a new impact in this field can be expected, since suitable galaxies farther away can be chosen for observation.

The theoretical work on gas flow in bar potentials has been discussed in section 16. A typical example of a theoretical velocity field [120], projected so as to allow comparison with the observations of NGC 5383 [212, 226, 75] is given in fig. 40. Several large scale features are obvious, and have been repeatedly found in many other calculations. The lines of constant radial velocity tilt within the bar so that their orientation approaches that of the bar major axis. Thus the kinematical major axis in the inner parts of a barred spiral can differ greatly from the true major axis. The isovelocity contours also show some crowding and kinks as they leave the bar, so that the zero velocity contour has a Z-shape. Beam smearing, as occurs in HI data because of the coarse angular resolution, will smooth the contours. Velocity fields with these characteristics have been observed for several galaxies [29, 30 for a review].

The effect of orientation is shown in fig. 41, where the same velocity field is observed from different angles. The background stellar density distribution used in these calculations was taken from [249]. The effects of the bar are very obvious when the angle between its major axis and the true major axis is approximately 90° and best hidden when this angle is around 0° . This point should be kept in mind in selecting suitable candidate galaxies for observation.

For velocity fields so far from axisymmetry as those shown above, the term rotation curve loses its meaning. In fig. 42 (from [75]) we plot the tangential velocity as a function of radius for position angles along the bar (thin line) and across it (thick line), and compare it with a mean, i.e. azimuthally averaged in annuli, rotation curve (dashed line). The results are as one would expect for gas moving along

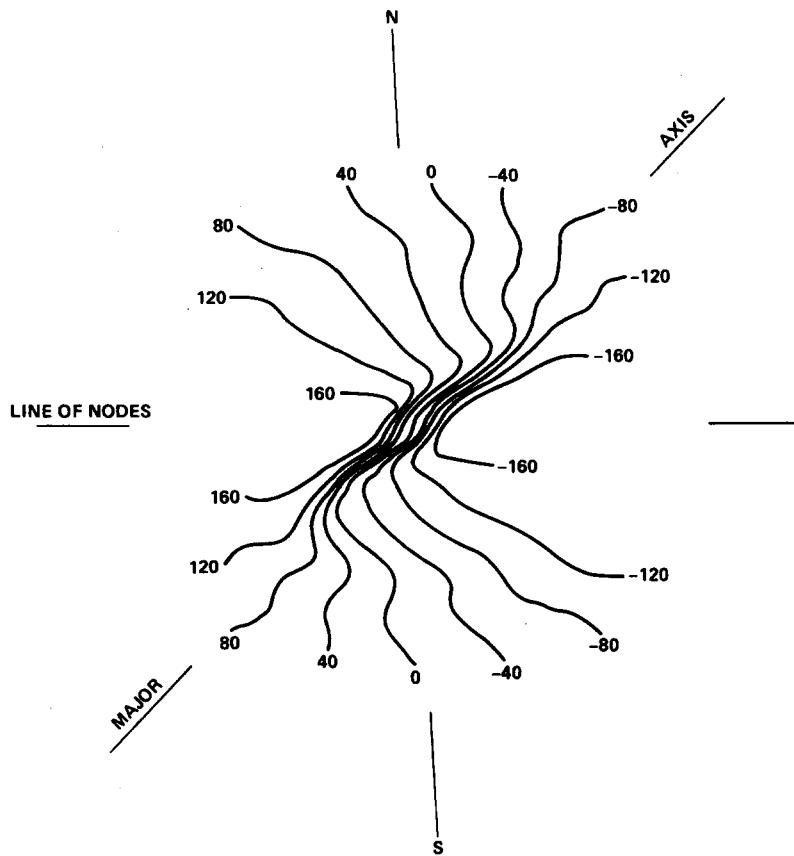


Fig. 40. Calculated line-of-sight velocities, in km/s, in a barred galaxy model [120]. The projection parameters and the length and velocity scaling have been chosen to match NGC 5383. (Reprinted courtesy of J.M. Huntley and The Astrophysical Journal, published by the University of Chicago Press; © 1978 The American Astronomical Society.)

elliptical streamlines. Across the bar we have a high maximum near the center while along it the maximum is lower and further out. Any trace of a maximum has disappeared on the mean curve and the result looks like the flat rotation curves so often observed.

In principle, a lot of useful parameters can be derived from comparison of gas flow models with observations, in particular if surface photometry is used to estimate the background stellar density distribution. In that case constraints on the bar strength and pattern speed, and on the mass-to-light ratio of the various components can be obtained. Such a study has been attempted for NGC 5383 [75], using HI data [226] and optical velocities [212, 75]. The beam smearing problem makes the HI data unsuitable for determination of critical bar parameters like bar strength and pattern speed, but in the outer parts the HI data are indispensable for deriving the orientation parameters and the axisymmetric background. The optical velocities put severe constraints on the allowable gas flow models, and from the surface photometry and the best fitting model mass-to-light ratios can be estimated. For NGC 5383 a well-defined sequence was found, with the bulge having a M/L of ~ 3 , the intermediate bar and lens region having a M/L of ~ 10 , and the outer disk and halo a M/L of $\sim 30 M_\odot/L_\odot$ (for a distance of 23.5 Mpc). Similar studies are now underway for a number of other galaxies, so that progress in this field will be made rapidly in the coming years.

So much for symmetric bars. If we now consider bars displaced with respect to the center of the disk

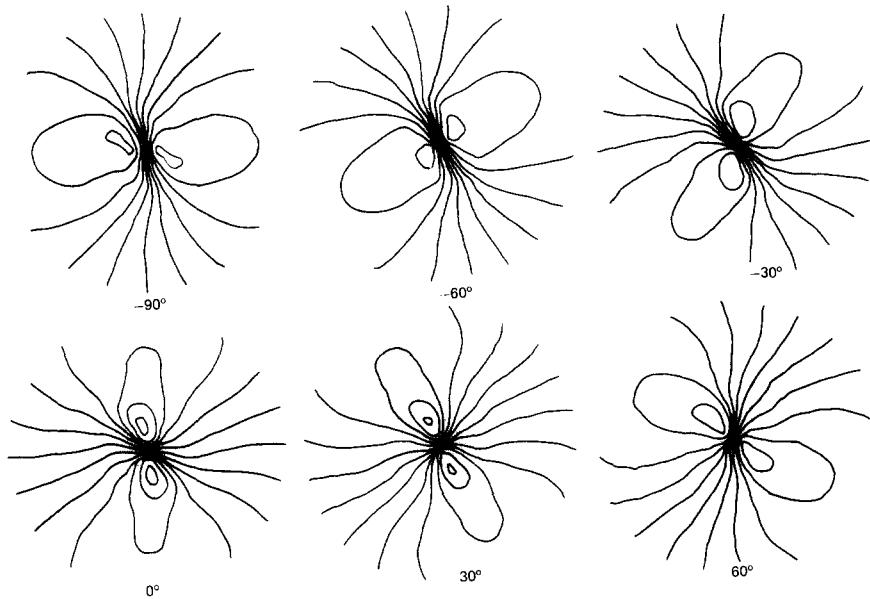


Fig. 41. Velocity field of a barred galaxy viewed from six different position angles. The density distribution is taken from [249] and the bar major axis is at roughly -60° . It is viewed successively from -90° , -60° , -30° , 0° , 30° and 60° . Notice that the signature of the bar is least pronounced when the bar is viewed along its minor axis, i.e. when the kinematical major axis is along the major axis of the bar.

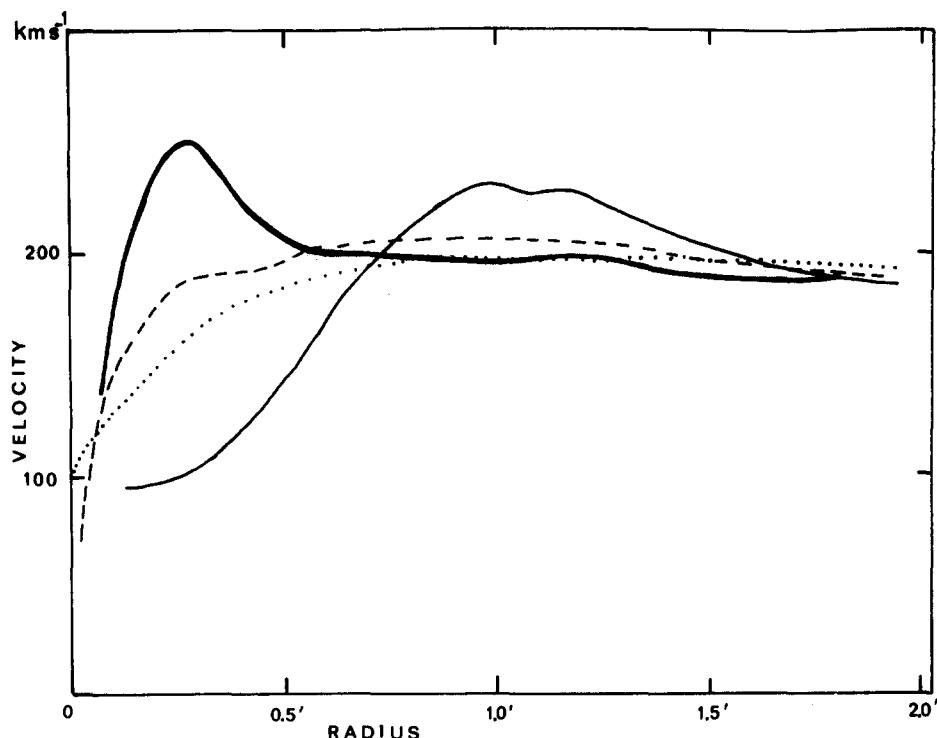


Fig. 42. Tangential velocity as a function of radius for position angles along the bar (thin line), across it (thick line) and azimuthally averaged in annuli (dashed line).

the problem of the gas flow is obviously not simplified. Observations show “a puzzling asymmetry” [67, see also 68], of the rotation curve and of the velocity field with respect to the center of the bar. A good example is NGC 1313 [181]. By calculating the gas flow in a displaced bar model, as prescribed by the photometry, one can reproduce the observations. Thus the asymmetry in the mass distribution is, at least in NGC 1313, enough to explain the important asymmetry in the observed velocity field [181].

19.6. Driving by companions

We have seen that a relatively close passage of a companion can force spiral structure in a disk, if this is not embedded in too massive a halo. Indeed some of the nicest observed spirals like M51 or M81 have obvious close companions. A statistical analysis [155] using a sample of 54 galaxies with measured rotation curves shows that all galaxies in the sample having close companions, a total of 7, have also a grand design spiral extending well beyond the inner parts of the galaxy. This would mean that at least these galaxies do not have too large halos. Statistics with a large sample [78] (305 galaxies out of which 52 are in binary systems and 174 in groups) show that among the isolated SA galaxies, $32 \pm 10\%$ have a well-developed spiral structure. Binary or group membership increases this probability to $67 \pm 6\%$. So, in contrast to [155], [78] deduce that, although companions can undoubtedly trigger spiral structure, they are neither perfectly effective nor always necessary.

In such statistical analyses, a certain care is needed as to what one names a companion. As can be seen from the calculations in [274] a 1% forcing is sufficient for provoking spirals. This can be roughly translated to magnitude or distance limits. Let us denote by D the distance between the center of a galaxy G of mass M_G and its companion C, of mass M_c , and let us compare the forces exerted on a star S at a distance r , from G. Let ϑ be the angle between D and r , and let $r \ll D$ so that perturbations expansions may be used. To first order the $m = 2$ tangential and radial forces from the companion will be

$$F_{r,2} = \frac{3GM_c}{2D^3} r \cos 2\vartheta, \quad F_{t,2} = \frac{3GM_c}{2D^3} r \sin 2\vartheta.$$

Let us also approximate the force from the parent galaxy G on the star S by

$$F_{\text{cent}} = GM_G/r^2.$$

Then for a 1% forcing

$$(F_{r,2}^2 + F_{t,2}^2)^{1/2} = 0.01 GM_G/r^2 \quad \text{or} \quad M_c/M_G = \frac{2}{3} \times 10^{-2} D^3/r^3.$$

(A correction for projection should be applied if the companion is not the plane of the galaxy.)

Thus	for $M_c = M_G/10$	$D = 2.5r$
	for $M_c = M_G$	$D = 5.3r$
	while for $M_c = 10M_G$	$D = 11.5r$

This indicates that the spirals in M51 and M81 could well have a tidal origin, while the Magellanic clouds can not have triggered any grand design spiral in our Galaxy. Furthermore it shows clearly that relatively far away companions have to be very massive in order to force a spiral, while small but nearby

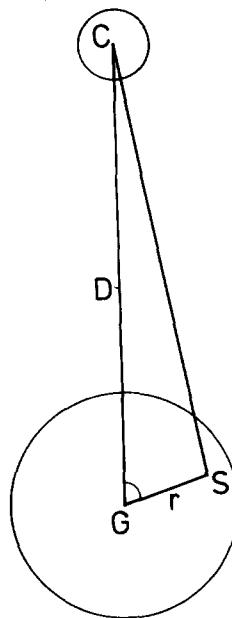


Fig. 43.

companions can be very efficient. The existence of such companions, however, is difficult to establish, unless large scale plates and redshifts are available. Proceeding from magnitude limited catalogues as in [282, 114] can be fallacious for all but the nearest spirals. Finally, although a given companion could by this formula be considered effective, it could well be hindered in its action by a large halo or velocity dispersion in the disk.

Some important clues on the velocity of the companion with respect to the galaxy it perturbs may be obtained from the minimum radius the spiral can reach. The word 'may' is necessary since the presence of a very substantial bulge component or, equivalently, the absence of a sizable disk fraction can also stop the spiral from reaching the center. Translating the time dependence of the forces in [274] into velocities and distances shows that a flyby has to be very rapid to excite a spiral in the central parts. However it is not a straightforward matter to apply this test to real galaxies since one needs in principle to know the geometry as well as the total relative velocity and not only its projection on the line of sight. One example where this is possible is the case of NGC 5194 and 5195. The geometry proposed in [273] is compatible with a sufficiently rapid passage, particularly if one remembers that the results in [274] assume $V = \text{constant}$ throughout the galaxy, while the rotation curve of the actual M51 rises not too abruptly with radius.

20. In way of a conclusion

Explaining the formation and evolution of the spiral patterns observed in many disk galaxies has not proved an easy task. A very large number of analytical calculations, numerical experiments and observations has led to some understanding of certain aspects of the problem and to a large number of speculations, stimulating yet more calculations, numerical experiments and observations.

Undeniable progress has been made on the theoretical side. It is now established that disks strive to

transfer their angular momentum outwards and also that shear and self-gravity can go a long way towards amplifying spirals. Thus we have a basis that helps us understand both modes and forced transient spirals, and probably also a good deal of the more ragged spiral structures.

On the numerical side also there has been an important progress. Numerical codes are now capable of calculating normal modes with a considerable accuracy. Thus there is good reason to believe that their description of the non-linear evolution will also be accurate.

However, all this evades our main question: Can we now explain the observed spiral structures? Have we gone far enough to be able to say something about that?

An important thing to note is that some of the nicest examples of global spiral structure have close companions. Nearly all reviews of this subject start by pointing out M51, M81, or NGC 5364. As theoretical calculations show clearly that companions can indeed force spectacular spirals (if the galactic disk is not immersed in too massive a halo and is not too hot) this is one possible solution of the problem.

Of course not all galaxies with a global spiral structure have close companions nor is this the only known way of creating spirals. A second distinct possibility is that such structures are modes. Under certain conditions (e.g. not too much halo) the bar mode is the most unstable one. Then the evolution of the galaxy should closely follow that of numerical experiments and the result would be an SB or SAB galaxy. If the disk was purely stellar then it could sustain only transient spirals and even these would stop after the disk had heated up. The result would then look like an SBO. If gas was present, however, the bar could drive relatively long-lived spirals. Infrared photometry may prove that this can be a solution for more galaxies than one would believe at first sight.

In other cases the most unstable mode is not bar-like but spiral. If this had a reasonable growth rate and shape it would be a viable solution. No modes calculated so far can claim to fulfill completely these two criteria. All examples have either too large growth rates or very lumpy shapes due to interference patterns. However, no reasonable calculations of modes in hot disks have been fully published so far. Second and most important, no dissipating gas has yet been included in any mode studies. Thus unstable modes with reasonably hefty growth rates, capable of compensating for the energy dissipated in gaseous shocks, may be a possible answer to spiral structure.

Finally the spiral pattern is often broken up, sometimes to the point where it is difficult to trace it. Multi-armed and flocculent spirals are no rarities and as such they deserve as much attention as the grand design two-armed spirals which capture the eye. At least three ways of accounting for these have been put forward so far. Non-gravitational stochastic theories propose such shapes. The superposition of a sufficiently large number of modes, provided of course these exist and have reasonable growth rates and comparable amplitudes, can give a very similar impression. Last but not least, large spurs, feathers or trailing arm segments can be understood as the responses of a sufficiently massive shearing disk triggered by relatively large clumps of matter.

Acknowledgements

While I was preparing this review I had a large number of discussions on the subject with many colleagues. The list of individual names would be too long to include but I thank them all heartily. I have particularly benefited from a series of seminars given by A. Toomre at the Stockholm Observatory in January 1982 and enjoyed some interesting discussions afterwards. Moreover, his extensive comments on a draft of this manuscript greatly helped in improving the presentation. I am also most

indebted to A. Sandage for allowing me to look for several days at his plate collection, from which a large part of section 2 was inspired. Figures 2, 3, 5, 10 and 12 are reproductions of his plates. Figures 4, 8 and 11 are reproductions of plates of A. Bosma and fig. 6 of a plate of P.O. Lindblad. A. Bosma read and commented on various preliminary versions. His imprint on the observational sections is particularly large. P. Vandervoort, J. Lequeux, J. Sellwood and G. Monnet read the manuscript and made several interesting suggestions. D. Schramm showed a lot of patience when I failed to meet deadlines. D. Chabod helped with the typing, S. Papaioannou with figs. 17, 19, 20, 28 and 31 and J. Colin with fig. 34.

References

- [1] M. Aaronson, in: *Highlights of Astronomy*, vol. 6, ed. R.M. West (Reidel Pub., 1983) p. 269.
- [2] R.J. Allen, W.M. Goss and H. van Woerden, *Astron. Astrophys.* 29 (1973) 447.
- [3] A. Ambastha and R.K. Varma, *J. Astrophys. Astr.* 3 (1982) 125.
- [4] S. Aoki, M. Noguchi and M. Iye, *Publ. Astron. Soc. J.* 31 (1979) 737.
- [5] E. Athanassoula, in: *Structure and Properties of Nearby Galaxies*, eds. E.M. Berkhuijsen and R. Wielebinski (Reidel Pub., 1978) p. 163.
- [6] E. Athanassoula, *Astron. Astrophys.* 69 (1978) 395.
- [7] E. Athanassoula, *Astron. Astrophys.* 88 (1980) 184.
- [8] E. Athanassoula, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 243.
- [9] E. Athanassoula, O. Bienayme, L. Martinet and D. Pfenninger, *Astron. Astrophys.* 127 (1983) 349.
- [10] E. Athanassoula and A. Bosma, in preparation (1984).
- [11] E. Athanassoula, A. Bosma, M. Creze and M.P. Schwarz, *Astron. Astrophys.* 107 (1982) 101 and 131 (1984) 175.
- [12] E. Athanassoula, A. Bosma and S. Papaioannou, in preparation (1984).
- [13] E. Athanassoula and L. Martinet, *Astron. Astrophys.* 87 (1980) L10.
- [14] E. Athanassoula and J.A. Sellwood, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 203 and in preparation.
- [15] P.D. Atherton and K. Taylor, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 31.
- [16] W. Baade, *Evolution of Stars and Galaxies* (MIT Press, 1963).
- [17] J.N. Bahcall, *Astrophys. J.* 267 (1983) 52.
- [18] J.N. Bahcall, M. Schmidt and R.M. Soneira, *Astrophys. J.* 265 (1983) 730.
- [19] J.M. Bardeen, in: *Dynamics of Stellar Systems*, ed. A. Hayli (Reidel Pub., 1975) p. 297.
- [20] F.N. Bash, *Astrophys. J.* 233 (1979) 524.
- [21] F.N. Bash and H.C.D. Visser, *Astrophys. J.* 247 (1981) 488.
- [22] W.A. Baum, in: *Spectral Classification and Multicolour Photometry*, eds. K. Loden, L.O. Loden and U. Sinnerstad (Academic Press, 1966) p. 288.
- [23] G.F. Benedict, *Astron. J.* 81 (1976) 799.
- [24] R.H. Berman, D.R.K. Brownrigg and R.W. Hockney, *Mon. Not. R. Astron. Soc.* 185 (1978) 861.
- [25] R.H. Berman and J.W-K. Mark, *Astron. Astrophys.* 77 (1979) 31.
- [26] R.H. Berman, D.J. Pollard and R.W. Hockney, *Astron. Astrophys.* 78 (1979) 133.
- [27] G. Bertin, Y.Y. Lau, C.C. Lin, J.W-K. Mark and L. Sugiyama, *Proc. Nat. Acad. Sci. USA* 74 (1977) 4726.
- [28] T. Boroson, *Astrophys. J. Suppl.* 46 (1981) 177.
- [29] A. Bosma, *Astron. J.* 86 (1981) 1791 and 1825.
- [30] A. Bosma, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 11.
- [31] A. Bosma, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 253.
- [32] A. Bosma, R.D. Ekers and J. Lequeux, *Astron. Astrophys.* 57 (1977) 97.
- [33] A. Bosma and K.C. Freeman, unpublished (1981).
- [34] A. Bosma, W.M. Goss and R.J. Allen, *Astron. Astrophys.* 93 (1981) 106.
- [35] A. Bosma and P.C. van der Kruit, *Astron. Astrophys.* 79 (1979) 281.
- [36] S.P. Boughn, P.R. Saulson and M. Seldner, *Astrophys. J.* 250 (1981) L15.
- [37] J. Boulesteix, Y.P. Georgelin, M. Marcellin and G. Monnet, in: *Instrumentation in Astronomy*, V (S.P.I.E., 1984).
- [38] E. Brinks, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 23.
- [39] D. Burstein, *Astrophys. J. Suppl.* 41 (1979) 435.
- [40] D. Burstein, *Astrophys. J.* 234 (1979) 435.

- [41] J.A.R. Caldwell and J.P. Ostriker, *Astrophys. J.* 251 (1981) 61.
- [42] R.G. Carlberg and J.A. Sellwood, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 127 and in preparation.
- [43] P. Carnevali, *Astrophys. J.* 265 (1983) 701.
- [44] S. Casertano, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 63.
- [45] J. Colin and E. Athanassoula, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 239 and in preparation.
- [46] F. Combes and M. Gerin, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 225 and in preparation.
- [47] F. Combes and R.H. Sanders, *Astron. Astrophys.* 96 (1981) 164.
- [48] S. Considere and E. Athanassoula, *Astron. Astrophys.* 111 (1982) 28.
- [49] G. Contopoulos, in: *Photometry, Kinematics and Dynamics of Galaxies*, ed. D.S. Evans (University of Texas Press, 1979) p. 425.
- [50] G. Contopoulos, *Astron. Astrophys.* 81 (1980) 198.
- [51] G. Contopoulos, *Astron. Astrophys.* 102 (1981) 265.
- [52] G. Contopoulos, *Astron. Astrophys.* 117 (1983) 89.
- [53] G. Contopoulos and C. Mertzannides, *Astron. Astrophys.* 61 (1977) 477.
- [54] G. Contopoulos and T. Papayannopoulos, *Astron. Astrophys.* 92 (1980) 33.
- [55] G. Courtes and R. Dubout-Crillon, *Astron. Astrophys.* 11 (1971) 468.
- [56] P. Crane, *Astrophys. J.* 197 (1975) 317.
- [57] C.-G. Danver, *Lund. Ann. No.* 10 (1942).
- [58] M. Davis, E. Feigelson and D.W. Latham, *Astr. J.* 85 (1980) 131.
- [59] J.M. Deharveng, P. Jakobsen, B. Milliard and M. Laget, *Astron. Astrophys.* 88 (1980) 52.
- [60] G. de Vaucouleurs, *Astrophys. J.* 127 (1958) 487.
- [61] G. de Vaucouleurs, *Astrophys. J. Suppl.* 8 (1963) 31.
- [62] G. de Vaucouleurs, *Astrophys. J. Suppl.* 29 (1975) 193.
- [63] G. de Vaucouleurs and R. Buta, *Astron. J.* 85 (1980) 637.
- [64] G. de Vaucouleurs and R. Buta, *Astrophys. J. Suppl.* 44 (1980) 451 and 48 (1982) 219.
- [65] G. de Vaucouleurs and A. de Vaucouleurs, *Reference Catalogue of Bright Galaxies* (The University of Texas Press, 1964).
- [66] G. de Vaucouleurs, A. de Vaucouleurs and H.G. Corwin, *Second Reference Catalogue of Bright Galaxies* (The University of Texas Press, 1976).
- [67] G. de Vaucouleurs and K.C. Freeman, in: *The Spiral Structure of our Galaxy*, eds. W. Becker and G. Contopoulos (Reidel Pub., 1970) p. 356.
- [68] G. de Vaucouleurs and K.C. Freeman, *Vistas in Astronomy* 14 (1972) 163.
- [69] M.E. Dixon, *Astrophys. J.* 164 (1971) 411.
- [70] R. Dubout-Crillon, *Astron. Astrophys.* 56 (1977) 293.
- [71] R.H. Durisen, *Astrophys. J.* 224 (1978) 826.
- [72] A. Duus, Ph.D. Thesis, Australian National University (1975).
- [73] A. Duus and K.C. Freeman, in: *La dynamique des galaxies spirales*, ed. L. Weiachew (CNRS Pub., 1975) p. 419.
- [74] M.F. Duval, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 237.
- [75] M.F. Duval and E. Athanassoula, *Astron. Astrophys.* 121 (1983) 297.
- [76] G. Efstathiou, G. Lake and J. Negroponte, *Mon. Not. R. Astron. Soc.* 199 (1982) 1069.
- [77] D.M. Elmegreen, *Astrophys. J. Suppl.* 47 (1981) 229.
- [78] D.M. Elmegreen and B.G. Elmegreen, *Mon. Not. R. Astron. Soc.* 201 (1982) 1021.
- [79] D.M. Elmegreen and B.G. Elmegreen, *Astrophys. J. Suppl.* 54 (1984) 127.
- [80] S.A. Erickson, Ph.D. Thesis, MIT (1974).
- [81] S.M. Fall and G. Efstathiou, *Mon. Not. R. Astron. Soc.* 193 (1980) 189.
- [82] J.V. Feitzinger, A.E. Glassgold, H. Gerola and P.E. Seiden, *Astron. Astrophys.* 98 (1981) 371.
- [83] S.I. Feldman and C.C. Lin, *Stud. Appl. Math.* 52 (1973) 1.
- [84] K.C. Freeman, in: *The Spiral Structure of our Galaxy*, eds. W. Becker and G. Contopoulos (Reidel Pub., 1970) p. 351.
- [85] K.C. Freeman, *Astrophys. J.* 160 (1970) 811.
- [86] K.C. Freeman, in: *Dynamics of Stellar Systems*, ed. A. Hayli (Reidel Pub., 1975) p. 367.
- [87] K.C. Freeman, D.W. Carrick and J.L. Craft, *Astrophys. J.* 198 (1975) L93.
- [88] M. Fujimoto, in: *Non-stable Phenomena in Galaxies* (The Publishing House of the Academy of Sciences of Armenian SSR, 1966) p. 453.
- [89] H. Gerola and P.E. Seiden, *Astrophys. J.* 223 (1978) 129.
- [90] H. Gerola, P.E. Seiden and L.S. Schulman, *Astrophys. J.* 242 (1980) 517.
- [91] P. Goldreich and D. Lynden-Bell, *Mon. Not. R. Astron. Soc.* 130 (1965) 125.
- [92] P. Goldreich and S. Tremaine, *Icarus* 34 (1978) 240.
- [93] P. Goldreich and S. Tremaine, *Astrophys. J.* 233 (1979) 857.
- [94] J.E. Gunn, in: *Astrophysical Cosmology*, eds H.A. Bruck, G.V. Coyne and M.S. Longair (Pontificia Academia Scientiarum, Vatican, 1982) p. 233.

- [95] J. Haass, in: Internal Kinematics and Dynamics of Galaxies, ed. E. Athanassoula (Reidel Pub., 1983) p. 121.
- [96] H.J. Habing, in: The Milky Way Galaxy, eds R.J. Allen, W.B. Burton and H. van Woerden (Reidel Pub., 1984) see also *Nature* 303 (1984) 287.
- [97] J.A. Hackwell and F. Schweizer, *Astrophys. J.* 265 (1983) 643.
- [98] D.J. Hegyi and G.L. Gerber, *Astrophys. J.* 218 (1977) L7.
- [99] M. Henon, *Ann. Astrophys.* 22 (1959) 126.
- [100] M. Henon and C. Heiles, *Astron. J.* 69 (1964) 73.
- [101] R.W. Hockney and D.R.K. Brownrigg, *Mon. Not. R. Astron. Soc.* 167 (1974) 351.
- [102] R.W. Hockney and J.W. Eastwood, *Computer Simulations Using Particles* (McGraw-Hill, 1981).
- [103] R.W. Hockney and F. Hohl, *Astron. J.* 74 (1969) 1102.
- [104] F. Hohl, *Astrophys. Sp. Sc.* 14 (1971) 91.
- [105] F. Hohl, *Astrophys. J.* 168 (1971) 343.
- [106] F. Hohl, *J. Comp. Phys.* 9 (1972) 10.
- [107] F. Hohl, in: *Dynamics of Stellar Systems*, ed. A. Hayli (Reidel Pub., 1975) p. 349.
- [108] F. Hohl, in: *La dynamique des galaxies spirales*, ed. L. Weliachew (CNRS Pub., 1975) p. 237.
- [109] F. Hohl, *Astron. J.* 81 (1976) 30.
- [110] F. Hohl, *Astron. J.* 83 (1978) 768.
- [111] F. Hohl and R.W. Hockney, *J. Comp. Phys.* 4 (1969) 306.
- [112] R.G. Hohlfeld and N. Krumm, *Astrophys. J.* 244 (1981) 476.
- [113] E. Hubble, *Astrophys. J.* 97 (1943) 112.
- [114] J. Huchra and T.X. Thuan, *Astrophys. J.* 216 (1977) 694.
- [115] R.M. Humphreys and A. Sandage, *Astrophys. J. Suppl.* 44 (1980) 319.
- [116] C. Hunter, *Mon. Not. R. Astron. Soc.* 126 (1963) 299.
- [117] C. Hunter, *St. Ap. Math. Soc. J.* 48 (1969) 55.
- [118] C. Hunter, *Astrophys. J.* 157 (1969) 183.
- [119] J.M. Huntley, Ph.D. Thesis, University of Virginia (1977).
- [120] J.M. Huntley, *Astrophys. J.* 225 (1978) L101.
- [121] J.M. Huntley, *Astrophys. J.* 238 (1980) 524.
- [122] G. Illingworth and P.L. Schechter, *Astrophys. J.* 256 (1982) 481.
- [123] M. Iye, *Publ. Astron. Soc. J.* 30 (1978) 223.
- [124] M. Iye, in: Internal Kinematics and Dynamics of Galaxies, ed. E. Athanassoula (Reidel Pub., 1983) p. 153.
- [125] M. Iye, M. Hamabe, M. Watanabe and S. Okamura, in: *Photometry, Kinematics and Dynamics of Galaxies*, ed. D.S. Evans (University of Texas Press, 1979) p. 407.
- [126] M. Iye, S. Okamura, M. Hamabe and M. Watanabe, *Astrophys. J.* 256 (1982) 103.
- [127] R.A. James, *J. Comp. Phys.* 25 (1977) 71.
- [128] R.A. James and J.A. Sellwood, *Mon. Not. R. Astron. Soc.* 182 (1978) 331.
- [129] S. Jorsater, in: *Photometry, Kinematics and Dynamics of Galaxies*, ed. D.S. Evans (University of Texas Press, 1979) p. 197.
- [130] W.H. Julian and A. Toomre, *Astrophys. J.* 146 (1966) 810.
- [131] A.J. Kalnajs, in: *The Spiral Structure of our Galaxy*, eds. W. Becker and G. Contopoulos (Reidel Pub., 1970) p. 318.
- [132] A.J. Kalnajs, *Astrophys. J.* 166 (1971) 275.
- [133] A.J. Kalnajs, *Astrophys. Lett.* 11 (1972) 41.
- [134] A.J. Kalnajs, *Astrophys. J.* 175 (1972) 63.
- [135] A.J. Kalnajs, *Proc. Astron. Soc. Aust.* 2 (1973) 174.
- [136] A.J. Kalnajs, in: *La dynamique des galaxies spirales*, ed. L. Weliachew (Paris, CNRS Pub., 1975) p. 103.
- [137] A.J. Kalnajs, *Astrophys. J.* 212 (1977) 637.
- [138] A.J. Kalnajs, in: *Structure and Properties of Nearby Galaxies*, eds. E.M. Berkhuijsen and R. Wielebinski (Reidel Pub., 1978) p. 113.
- [139] A.J. Kalnajs, in: Internal Kinematics and Dynamics of Galaxies, ed. E. Athanassoula (Reidel Pub., 1983) p. 87.
- [140] A.J. Kalnajs, in: Internal Kinematics and Dynamics of Galaxies, ed. E. Athanassoula (Reidel Pub., 1983) p. 109.
- [141] A.J. Kalnajs, in preparation (1984).
- [142] A.J. Kalnajs and E. Athanassoula-Georgala, *Mon. Not. R. Astron. Soc.* 168 (1974) 287.
- [143] R.C. Kennicutt, *Astron. J.* 86 (1981) 1847.
- [144] R.C. Kennicutt, *Astron. J.* 87 (1982) 255.
- [145] R.C. Kennicutt and P. Hodge, *Astrophys. J.* 253 (1982) 101.
- [146] J. Kormendy, *Astrophys. J.* 217 (1977) 406.
- [147] J. Kormendy, *Astrophys. J.* 227 (1979) 714.
- [148] J. Kormendy, in: *The Structure and Evolution of Normal Galaxies*, eds. S.M. Fall and D. Lynden-Bell (Cambridge University Press, 1981) p. 85.
- [149] J. Kormendy, in: *Morphology and Dynamics of Galaxies*, eds. L. Martinet and M. Mayor (Geneva Observatory, 1982) p. 113.
- [150] J. Kormendy, *Astrophys. J.* 257 (1982) 75.
- [151] J. Kormendy, *Astrophys. J.* 275 (1984) 527.

- [152] J. Kormendy, in: *The Milky Way Galaxy*, eds. R.J. Allen, W.B. Burton and H. van Woerden (Reidel Pub., 1984).
- [153] J. Kormendy and A.G. Bruzual, *Astrophys. J.* 223 (1978) L63.
- [154] J. Kormendy and G. Illingworth, *Astrophys. J.* 256 (1982) 460.
- [155] J. Kormendy and C.A. Norman, *Astrophys. J.* 233 (1979) 539.
- [156] R.B. Larson and B.M. Tinsley, *Astrophys. J.* 219 (1978) 46.
- [157] Y.Y. Lau and G. Bertin, *Astrophys. J.* 226 (1978) 508.
- [158] Y.Y. Lau, C.C. Lin and J.W-K. Mark, *Proc. Natl. Acad. Sci. USA* 73 (1976) 1379.
- [159] Y.Y. Lau and J.W-K. Mark, *Proc. Natl. Acad. Sci. USA* 73 (1976) 3785.
- [160] F.H. Levinson and W.W. Roberts, *Astrophys. J.* 245 (1981) 465.
- [161] C.C. Lin, *J. SIAM Appl. Math.* 14 (1966) 876.
- [162] C.C. Lin, in: *The Spiral Structure of our Galaxy*, eds. W. Becker and G. Contopoulos (Reidel Pub., 1970) p. 377.
- [163] C.C. Lin, in: *Highlights of Astronomy*, ed. C. de Jager (Reidel Pub., 1971) p. 88.
- [164] C.C. Lin, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 117.
- [165] C.C. Lin and Y.Y. Lau, *SIAM J. Appl. Math.* 29 (1975) 352.
- [166] C.C. Lin and F.H. Shu, *Astrophys. J.* 140 (1964) 646.
- [167] C.C. Lin and F.H. Shu, *Proc. Natl. Acad. Sci. USA* 55 (1966) 229.
- [168] C.C. Lin and F.H. Shu, in: *Astrophysics and General Relativity*, Vol. 2, eds. M. Chretien, S. Deser and J. Goldstein (Gordon and Breach Science Pub., 1971) p. 235.
- [169] C.C. Lin, C. Yuan and F.H. Shu, *Astrophys. J.* 155 (1969) 721.
- [170] B. Lindblad, *Handbuch der Physik* 53 (1959) 21.
- [171] P.O. Lindblad, *Stockholms Obs. Ann.* 21, No. 3 (1960).
- [172] P.O. Lindblad, *Stockholms Obs. Ann.* 21, No. 4 (1960).
- [173] H.H. Loose and K.J. Fricke, in: *The most Massive Stars*, eds. S. D'Odorico, D. Baade and K. Kjar (ESO Pub. 1981) p. 269.
- [174] L.B. Lucy, *Astron. J.* 82 (1977) 1013.
- [175] D. Lynden-Bell, *Mon. Not. R. Astron. Soc.* 187 (1979) 101.
- [176] D. Lynden-Bell and A.J. Kalnajs, *Mon. Not. R. Astron. Soc.* 157 (1972) 1.
- [177] B.T. Lynds, in: *The Spiral Structure of our Galaxy*, eds. W. Becker and G. Contopoulos (Reidel Pub., 1970) p. 26.
- [178] B.T. Lynds, in: *External Galaxies and Quasi-Stellar Objects*, ed. D.S. Evans (Reidel Pub., 1972) p. 56.
- [179] B.T. Lynds, *Astrophys. J. Suppl.* 28 (1974) 391.
- [180] B.T. Lynds, *Astrophys. J.* 238 (1980) 17.
- [181] M. Marcellin and E. Athanassoula, *Astron. Astrophys.* 105 (1982) 76.
- [182] J.W-K. Mark, *Proc. Natl. Acad. Sc. USA* 68 (1971) 2095.
- [183] J.W-K. Mark, in: *The Formation and Dynamics of Galaxies*, ed. J. Shakeshaft (Reidel Pub., 1974) p. 417.
- [184] J.W-K. Mark, *Astrophys. J.* 193 (1974) 539.
- [185] J.W-K. Mark, *Astrophys. J.* 203 (1976) 81.
- [186] J.W-K. Mark, *Astrophys. J.* 205 (1976) 363.
- [187] J.W-K. Mark, *Astrophys. J.* 206 (1976) 418.
- [188] J.W-K. Mark, *Astrophys. J.* 212 (1977) 645.
- [189] L.S. Marochnik, Y.N. Mishurov and A.A. Suchkov, *Astrophys. Space Sci.* 19 (1972) 285.
- [190] L. Martinet, *Astron. Astrophys.* 132 (1984) 381.
- [191] D.S. Mathewson, P.C. van der Kruit and W.N. Brouw, *Astron. Astroph.* 17 (1972) 468.
- [192] T. Matsuda and H. Isaka, *Prog. Theor. Phys.* 64 (1980) 1265, and private communication.
- [193] R. McCray and T.P. Snow, *Ann. Rev. Astr. Ap.* 17 (1979) 213.
- [194] R.H. Miller, *Astrophys. Space Sci.* 14 (1971) 73.
- [195] R.H. Miller, *Astrophys. J.* 190 (1974) 539.
- [196] R.H. Miller, *J. Comp. Phys.* 21 (1976) 400.
- [197] R.H. Miller, *Astrophys. J.* 223 (1978) 811.
- [198] R.H. Miller, K.H. Prendergast and W.J. Quirk, *Astrophys. J.* 161 (1970) 903.
- [199] R.H. Miller and B.F. Smith, *Astrophys. J.* 227 (1979) 785.
- [200] M. Miyamoto, *Publ. Astron. Soc. J.* 21 (1969) 319.
- [201] M. Miyamoto, *Astron. Astrophys.* 30 (1974) 441.
- [202] M.W. Mueller and W.D. Arnett, *Astrophys. J.* 210 (1976) 670.
- [203] S. Okamura, *Publ. Astron. Soc. Japan* 30 (1978) 91.
- [204] M. Ondrechen and J.M. van der Hulst, *Astrophys. J.* 269 (1983) L47.
- [205] J.H. Oort, *Bull. Astron. Inst. Netherlands* 3 (1927) 275.
- [206] J.H. Oort, in: *Interstellar Matter in Galaxies*, ed. L. Woltjer (W.A. Benjamin Press, 1962) p. 234.
- [207] J.P. Ostriker and J.A.R. Caldwell, in: *The Structure, Kinematics and Dynamics of the Milky Way*, ed. W.L.H. Shuter (Reidel Pub., 1983) p. 249.
- [208] J.P. Ostriker and P.J.E. Peebles, *Astrophys. J.* 186 (1973) 467.

- [209] R.F. Pannatoni, Ph.D. Thesis, MIT (1979).
- [210] I.I. Pasha and M.A. Smirnov, *Astrophys. Sp. Sc.* 86 (1982) 215.
- [211] T. Papayannopoulos and M. Petrou, *Astron. Astrophys.* 119 (1983) 21.
- [212] C.J. Peterson, V.C. Rubin, W.K. Ford and N. Thonnard, *Astrophys. J.* 219 (1978) 31.
- [213] J. Pfleiderer, *Z. Astrophys.* 58 (1963) 12.
- [214] J. Pfleiderer and H. Siedentopf, *Z. Astrophys.* 51 (1961) 201.
- [215] H. Poincaré, *Les Methodes nouvelles de la Mecanique Celeste*, III (Gauthier-Villars Ed., 1899).
- [216] K.H. Prendergast, unpublished (1970).
- [217] K.H. Prendergast, in: *Interstellar Matter in Galaxies*, ed. L. Woltjer (W.A. Benjamin, 1962) p. 217 and unpublished.
- [218] K.H. Prendergast, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 215.
- [219] W.W. Roberts, *Astrophys. J.* 158 (1969) 123.
- [220] W.W. Roberts, J.M. Huntley and G.D. van Albada, *Astrophys. J.* 233 (1979) 67.
- [221] W.W. Roberts, M.S. Roberts and F.H. Shu, *Astrophys. J.* 196 (1975) 381.
- [222] W.W. Roberts and F.H. Shu, *Astrophys. Lett.* 12 (1972) 49.
- [223] A.H. Rots and W.W. Shane, *Astron. Astrophys.* 45 (1975) 25.
- [224] V.C. Rubin, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 3.
- [225] V.C. Rubin, W.K. Ford and N. Thonnard, *Astrophys. J.* 238 (1980) 471.
- [226] R. Sancisi, R.J. Allen and W.T. Sullivan, *Astron. Astrophys.* 78 (1979) 217.
- [227] A. Sandage, *The Hubble Atlas of Galaxies* (Carnegie Institution of Washington, 1961).
- [228] A. Sandage, *Galaxies and the Universe*, in: *Stars and Stellar Systems*, vol. IX, eds. A. Sandage, M. Sandage and J. Kristian (University of Chicago Press, 1975) p. 1.
- [229] A. Sandage, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 367.
- [230] A. Sandage, unpublished (1984).
- [231] A. Sandage and R.M. Humphries, *Astrophys. J.* 236 (1980) L1.
- [232] A. Sandage and G. Tammann, *The Revised Shapley-Ames Catalogue* (Carnegie Institution of Washington, 1981).
- [233] R.H. Sanders, *Astrophys. J.* 217 (1977) 916.
- [234] R.H. Sanders and J.M. Huntley, *Astrophys. J.* 209 (1976) 53.
- [235] R.H. Sanders and K.H. Prendergast, *Astrophys. J.* 188 (1974) 489.
- [236] R.H. Sanders, P.J. Teuben and G.D. van Albada, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 221.
- [237] R.H. Sanders and A.D. Tubbs, *Astrophys. J.* 235 (1980) 803.
- [238] W.V. Schempp, *Astrophys. J.* 258 (1982) 96.
- [239] M. Schmidt, in: *Stars and Stellar Systems*, vol. 5, eds. A. Blaauw and M. Schmidt (University of Chicago Press, 1965).
- [240] M.P. Schwarz, Ph.D. Thesis, Australian National University (1979).
- [241] M.P. Schwarz, *Astrophys. J.* 247 (1981) 77.
- [242] M.P. Schwarz, *Astron. Astrophys.* 133 (1984) 222.
- [243] F. Schweizer, *Astrophys. J. Suppl.* 31 (1976) 313.
- [244] L. Searle, W.L.W. Sargent and W.G. Baguñolo, *Astrophys. J.* 179 (1973) 427.
- [245] A. Segalovitz, *Astron. Astrophys.* 55 (1977) 203.
- [246] P.E. Seiden and H. Gerola, *Astrophys. J.* 233 (1979) 56.
- [247] P.E. Seiden and H. Gerola, *Fund. Cosmic Physics* 7 (1982) 241.
- [248] J.A. Sellwood, *Astron. Astrophys.* 89 (1980) 296.
- [249] J.A. Sellwood, *Astron. Astrophys.* 99 (1981) 362.
- [250] J.A. Sellwood, *J. Comp. Phys.* 50 (1983) 337.
- [251] J.A. Sellwood and R.G. Carlberg, in preparation (1984).
- [252] J.A. Sellwood and R.A. James, *Mon. Not. R. Astron. Soc.* 187 (1979) 483.
- [253] F.H. Shu, *Astrophys. J.* 160 (1970) 89 and 99.
- [254] F.H. Shu, V. Milione and W.W. Roberts, *Astrophys. J.* 183 (1973) 819.
- [255] F.H. Shu, R.V. Stachnik and J.C. Yost, *Astrophys. J.* 166 (1971) 465.
- [256] F. Simien, E. Athanassoula, A. Pellet, G. Monnet, A. Maucherat and G. Courtes, *Astron. Astrophys.* 67 (1978) 73.
- [257] H. Spinrad, J.P. Ostriker, R.P.S. Stone, L.G. Chiu and A.G. Bruzual, *Astrophys. J.* 225 (1978) 56.
- [258] S.E. Strom, E.B. Jensen and K.M. Strom, *Astrophys. J.* 206 (1976) L11.
- [259] S.E. Strom and K.M. Strom, in: *Structure and Properties of nearby Galaxies*, eds. E.M. Berkhuijsen and R. Wielebinski (Reidel Pub., 1978) p. 69.
- [260] F. Takahara, *Prog. Theor. Phys.* 56 (1976) 1665.
- [261] F. Takahara, *Publ. Astron. Soc. J.* 30 (1978) 253.
- [262] R.J. Talbot, E.B. Jensen and R.J. Dufour, *Astrophys. J.* 229 (1979) 91.
- [263] P.J. Teuben and R.H. Sanders, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 211, and in preparation.

- [264] K.O. Thielheim and H. Wolff, *Astrophys. J.* 245 (1981) 39.
- [265] K.O. Thielheim and H. Wolff, *Mon. Not. R. Astron. Soc.* 199 (1982) 151.
- [266] B.M. Tinsley, *Mon. Not. R. Astron. Soc.* 194 (1981) 63.
- [267] A. Toomre, *Astrophys. J.* 138 (1963) 385.
- [268] A. Toomre, *Astrophys. J.* 139 (1964) 1217.
- [269] A. Toomre, *Astrophys. J.* 158 (1969) 899.
- [270] A. Toomre, in: *The Formation and Dynamics of Galaxies*, ed. J. Shakeshaft (Reidel Pub., 1974) p. 347.
- [271] A. Toomre, in: *Highlights of Astronomy*, ed. G. Contopoulos (Reidel Pub., 1974) p. 457.
- [272] A. Toomre, *Ann. Rev. Astron. Astrophys.* 15 (1977) 437.
- [273] A. Toomre, in: *The Large Scale Structure of the Universe*, eds. M.S. Longair and J. Einasto (Reidel Pub., 1978) p. 109.
- [274] A. Toomre, in: *The Structure and Evolution of Normal Galaxies*, eds. S.M. Fall and D. Lynden-Bell (Cambridge Univ. Press, 1981) p. 111.
- [275] A. Toomre, private communication (1982).
- [276] A. Toomre, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 177.
- [277] A. Toomre and J. Toomre, *Astrophys. J.* 178 (1972) 623.
- [278] S. Tremaine, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 411.
- [279] Y.M. Treve, in: *Topics in Nonlinear Dynamics. A Tribute to Sir Edward Bullard*, ed. S. Jorna (Amer. Inst. of Ph. Pub., 1978) p. 147.
- [280] V. Tsikoudi, *Astrophys. J. Suppl.* 43 (1980) 365.
- [281] R.B. Tully and J.R. Fisher, *Astron. Astrophys.* 54 (1977) 661.
- [282] E.L. Turner and J.R. Gott, *Astrophys. J. Suppl.* 32 (1976) 409.
- [283] S.C. Unwin, *Mon. Not. R. Astron. Soc.* 192 (1980) 243.
- [284] G.D. van Albada, *Astron. Astrophys.* 90 (1980) 123.
- [285] G.D. van Albada, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 227.
- [286] G.D. van Albada and W.W. Roberts, *Astrophys. J.* 246 (1981) 740.
- [287] G.D. van Albada, B. van Leer and W.W. Roberts, *Astron. Astrophys.* 108 (1982) 76.
- [288] T.S. van Albada and R.H. Sanders, *Mon. Not. R. Astron. Soc.* 201 (1982) 303.
- [289] S. van der Bergh, *Astrophys. J.* 131 (1960) 215 and 558.
- [290] J.M. van der Hulst, M.P. Ondrechen, J.H. Van Gorkom and E. Hummel, in: *Internal Kinematics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 233.
- [291] P.C. van der Kruit and L. Searle, *Astron. Astrophys.* 110 (1982) 61.
- [292] P.C. van der Kruit and G.S. Shostak, in: *Internal Kinematics and Dynamics of Galaxies*, ed. E. Athanassoula (Reidel Pub., 1983) p. 69.
- [293] P.O. Vandervoort, *Astrophys. J.* 161 (1970) 67.
- [294] P.O. Vandervoort, *Astrophys. J.* 240 (1980) 478.
- [295] P.O. Vandervoort, *Astrophys. J.* 256 (1982) L41.
- [296] P.O. Vandervoort, *Astrophys. J.* 273 (1983) 511.
- [297] P.O. Vandervoort and D.E. Welty, *Astrophys. J.* 263 (1982) 654.
- [298] B. van Leer, ICASE Report No. 81-11 (1981).
- [299] H.C.D. Visser, *Astron. Astrophys.* 88 (1980) 149 and 159.
- [300] S.D.M. White and M.J. Rees, *Mon. Not. R. Astron. Soc.* 183 (1978) 341.
- [301] R. Wielen, in: *Highlights of Astronomy*, ed. G. Contopoulos (Reidel Pub., 1974) p. 395.
- [302] R. Wielen, in: *Structure and Properties of Nearby Galaxies*, eds. E.M. Berkhuijsen and R. Wielebinski (Reidel Pub., 1978) p. 93.
- [303] M.S. Wilkerson, S.E. Strom and K.M. Strom, *Bull. Am. Astron. Soc.* 19 (1980) 649; *Astrophys. J.* 240 (1980) L115.
- [304] P.R. Woodward, *Astrophys. J.* 207 (1976) 484.
- [305] S. Yabushita, *Mon. Not. R. Astron. Soc.* 143 (1969) 231.
- [306] C. Yuan and P. Grosbol, *Astrophys. J.* 243 (1981) 432.
- [307] T.A. Zang, Ph.D. Thesis, MIT (1976).
- [308] T.A. Zang and F. Hohl, *Astrophys. J.* 226 (1978) 521.